



Casa abierta al tiempo

UNIVERSIDAD AUTÓNOMA METROPOLITANA

Unidad Iztapalapa

División de Ciencias Básicas e Ingeniería

Representaciones Tiempo-Frecuencia para el Análisis de Sonidos Torácicos

Tesis para obtener el título de
Maestro en Ciencias (Ingeniería Biomédica)

Presenta:

Bersaín Alexander Reyes

Asesores:

Dra. Sonia Charleston Villalobos

Dr. Ramón González Camarena

Sinodales:

Dr. Rogelio Alcántara Silva

Dra. Sonia Charleston Villalobos

Presidente

Secretaria

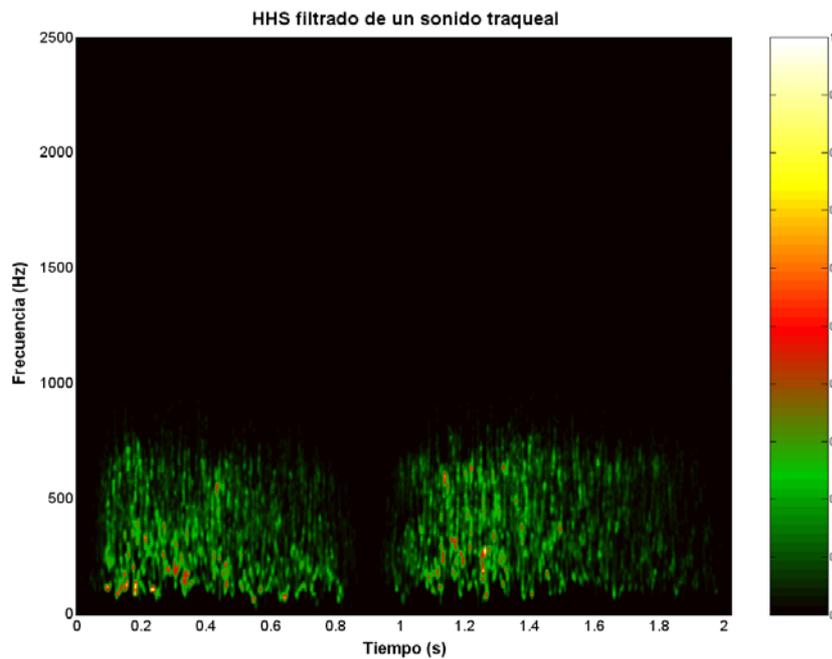
Vocal

DR. ANGEL TOMAS MUJAMA CORRALES

México, D.F. a 20 de Septiembre de 2010



Representaciones Tiempo-Frecuencia para el Análisis de Sonidos Torácicos



Bersáin Alexander Reyes

Maestría en Ciencias (Ingeniería Biomédica)

A

*Mi familia:
Angélica, Francisco,
David, Paula y Ricardo.*

*Mis hermanos de tesis:
Alejandra, Lidia y Lupita.*

*Mis mentores:
Sonia, Ramón y Tomás.*

*Mis compañeros y profesores de posgrado,
en especial a Joaquín, Raquel y Jatziri.*

Mi Casa Abierta al Tiempo.

Mis alumnos.

Mis amigos.

Gracias.

Este trabajo fue realizado con apoyo de la beca de posgrado otorgada por CONACyT.

*« Ninium ne crede colori »
Virgilio.*

ÍNDICE

RESUMEN	9
I. INTRODUCCIÓN.	13
I.1. Descripción de la señal en el dominio del tiempo y de la frecuencia.	13
I.1.1. Clasificación de las señales.	14
I.2. Importancia del análisis tiempo-frecuencia.	15
I.2.1. Algunos ejemplos analíticos.	15
I.3. Ideas fundamentales de las distribuciones tiempo-frecuencia.	17
I.4. Comentarios finales y estructura de la tesis.	18
Referencias.	20
II. ANTECEDENTES.	21
II.1. Sonidos respiratorios.	21
II.1.1. Sonidos pulmonares normales.	25
II.1.2. Sonidos traqueales normales.	30
II.1.3. Sonidos adventicios.	32
II.1.4. Estertores crepitantes, sonidos respiratorios adventicios discontinuos.	33
II.1.5. Estertores sibilantes, sonidos respiratorios adventicios continuos.	39
II.2. Sonidos cardiacos.	43
II.2.1. Primer sonido cardiaco, S1.	45
II.2.2. Segundo sonido cardiaco, S2.	48
Referencias.	53
III. PLANTEAMIENTO DEL PROBLEMA.	59
III.1. Introducción.	59
III.2. Hipótesis.	61
III.3. Objetivos.	61
III.3.1. Objetivo general.	61
III.3.2. Objetivos específicos.	61

IV.	REPRESENTACIONES TIEMPO-FRECUENCIA (TFRs).	63
IV.1.	Introducción.	63
IV.2.	La clase general de distribuciones.	63
IV.2.1.	La Distribución Wigner-Ville.	68
IV.2.1.1.	Formulación de la distribución.	68
IV.2.1.2.	Propiedades	70
IV.2.1.3.	Conceptos adicionales.	80
IV.2.1.4.	Comentario final.	81
IV.2.2.	El Espectrograma.	82
IV.2.2.1.	Formulación de la Transformada de Fourier de tiempo corto.	84
IV.2.2.2.	El Espectrograma.	85
IV.2.2.3.	Relación con la clase general de distribuciones.	86
IV.2.2.4.	Propiedades.	86
IV.2.2.5.	Ventanas de análisis comúnmente empleadas.	92
IV.2.2.6.	Relación y comparación con la distribución de Wigner.	94
IV.2.3.	La Distribución Choi-Williams.	96
IV.2.3.1.	Formulación de la distribución.	97
IV.2.3.2.	Propiedades.	99
IV.3.	El Espectrograma Reasignado.	103
IV.3.1.	Formulación de la distribución.	103
IV.3.2.	Propiedades.	106
IV.4.	Representación tiempo-frecuencia basado en el modelo autorregresivo variante en el tiempo de una señal estocástica.	109
IV.4.1.	Métodos paramétricos para la estimación del espectro de potencia.	111
IV.4.2.	Estimación del espectro de potencia mediante el modelado autorregresivo (AR).	113
IV.4.3.	El espectro de potencia instantánea y el modelo AR variante en el tiempo (TVAR).	115
IV.4.4.	Estimación de los parámetros del modelo TVAR.	117
IV.4.4.1.	El método de Burg.	117
IV.4.4.1.1.	Derivación del algoritmo.	117

IV.4.4.1.2.	Características del algoritmo.	120
IV.4.4.1.3.	Obtención de la representación tiempo-frecuencia.	123
IV.4.4.2.	El algoritmo de mínimos cuadrados recursivo (RLS).	124
IV.4.4.2.1.	Formulación del algoritmo.	125
IV.4.4.2.2.	Características.	131
IV.4.4.2.3.	Obtención de la representación tiempo-frecuencia.	135
IV.4.4.2.4.	Comparación con el método de Burg.	136
IV.4.4.3.	El algoritmo de mínimos cuadrados recursivo con factor de olvido variable (RLS-VFF).	137
IV.4.4.3.1.	Comparación con el algoritmo RLS estándar.	139
IV.4.4.3.2.	Obtención de la representación tiempo-frecuencia.	142
IV.4.4.4.	El filtro de Kalman.	144
IV.4.4.4.1.	Ecuaciones de estado del sistema y el proceso de innovación.	145
IV.4.4.4.2.	Formulación del filtro.	148
IV.4.4.4.3.	Obtención de la representación tiempo-frecuencia.	154
IV.5.	El Escalograma.	156
IV.5.1.	Conceptos teóricos de la representación wavelet.	156
IV.5.2.	Propiedades de la transformada wavelet continua.	161
IV.5.3.	El escalograma.	161
IV.5.4.	La wavelet Morlet compleja.	163
IV.6.	El espectro de Hilbert-Huang.	165
IV.6.1.	La transformada de Hilbert y el concepto de frecuencia instantánea.	168
IV.6.2.	Los modos de oscilación intrínsecos.	170
IV.6.3.	La descomposición empírica de modos.	171
IV.6.4.	El espectro de Hilbert-Huang.	174
	Referencias.	180

V.	METODOLOGÍA.	185
V.1.	Metodología general.	185
V.2.	Simulación computacional de las señales y sus representaciones tiempo-frecuencia ideales.	187
V.2.1.	Señales no biológicas.	187
	V.2.1.1. Señal no biológica concatenando sinusoidales puras.	188
	V.2.1.2. Señal no biológica mediante el modelo TVAR.	190
V.2.2.	Señales de sonidos cardiacos.	192
	V.2.2.1. Primer sonido cardiaco.	192
	V.2.2.2. Segundo sonido cardiaco.	195
V.2.3.	Señales de sonidos respiratorios adventicios.	197
	V.2.3.1. Sonidos adventicios discontinuos.	198
	V.2.3.2. Sonidos adventicios continuos.	200
V.3.	Índices de desempeño.	201
V.3.1.	La correlación bidimensional central.	202
	V.3.1.1. Enfoque global.	202
	V.3.1.2. Enfoque local.	202
V.3.2.	Estimación de la frecuencia instantánea mediante el primer momento temporal de la TFR.	204
	V.3.2.1. Error relativo.	205
	V.3.2.2. Coeficiente de correlación cruzada.	205
V.3.3.	Resolución tiempo-frecuencia.	206
V.4.	Parámetros de las representaciones tiempo-frecuencia.	206
V.5.	Adecuación de las representaciones tiempo-frecuencia.	208
V.6.	Señales de sonidos cardiacos y respiratorios adquiridas.	210
V.6.1.	Sonidos respiratorios.	210
	V.6.1.1. Sonidos respiratorios adventicios.	211
	V.6.1.2. Sonidos respiratorios normales.	215
V.6.2.	Sonidos cardiacos.	216
	Referencias.	218

VI. RESULTADOS.	221
VI.1. Introducción.	221
VI.2. Señal no biológica generada mediante el modelo TVAR.	221
VI.3. Señal no biológica concatenando sinusoidales puras.	249
VI.4. Primer sonido cardiaco y sus componentes.	254
VI.4.1. Componente muscular simulado (s_m) de S1.	254
VI.4.2. Componente valvular simulado (s_v) de S1.	259
VI.4.3. Primer sonido cardiaco, S1.	263
VI.4.4. Señales de S1 reales adquiridas.	267
VI.5. Segundo sonido cardiaco y sus componentes.	275
VI.5.1. Componente aórtico simulado (A2) de S2.	275
VI.5.2. Componente pulmonar simulado (P2) de S2.	280
VI.5.3. Segundo sonido cardiaco simulado, S2.	285
VI.5.4. Señales de S2 reales adquiridas.	290
VI.6. Estertor sibilante simulado.	303
VI.6.1. Estertores sibilantes reales adquiridos.	308
VI.7. Estertores crepitantes.	323
VI.7.1. Estertor crepitante fino simulado.	323
VI.7.2. Estertor crepitante grueso simulado.	328
VI.7.3. Estertores crepitantes reales adquiridos.	333
VI.8. Sonidos respiratorios normales adquiridos.	347
VI.8.1. Sonido pulmonar.	347
VI.8.2. Sonido traqueal.	351
VII. DISCUSIÓN.	355
VII.1. Introducción.	355
VII.2. Señal no biológica generada mediante el modelo TVAR.	355
VII.3. Señal no biológica concatenando dos tonos puros.	358
VII.4. Primer sonido cardiaco simulado y sus componentes.	360
VII.4.1. Componente muscular (s_m).	360
VII.4.2. Componente valvular simulado (s_v).	363
VII.4.3. Primer sonido cardiaco, S1.	366
VII.4.4. Señales de S1 reales adquiridas.	368

VII.5.	Segundo sonido cardiaco simulado y sus componentes.	369
VII.5.1.	Componente aórtico (A2).	369
VII.5.2.	Componente pulmonar simulado (P2).	370
VII.5.3.	Segundo sonido cardiaco simulado (S2).	371
VII.5.4.	Señales de S2 reales adquiridas.	373
VII.6.	Estertor sibilante simulado.	374
VII.6.1.	Estertores sibilantes reales adquiridos.	375
VII.7.	Estertores crepitantes.	376
VII.7.1.	Estertor crepitante fino simulado.	377
VII.7.2.	Estertor crepitante grueso simulado.	378
VII.7.3.	Estertores crepitantes reales adquiridos.	379
VII.8.	Sonidos respiratorios normales adquiridos.	380
VII.8.1.	Sonido pulmonar.	381
VII.8.2.	Sonido traqueal.	381
VIII.	CONCLUSIONES.	383
	ANEXOS.	384

RESUMEN.

La importancia del estudio de los sonidos generados en el tórax ha sido tal que, hace un par de siglos, Laennec mismo clasificó los eventos acústicos generados por la ventilación de los pulmones que escuchaba e intentó relacionarlos con los hallazgos anatómicos y patológicos que encontraba tras la autopsia. Aún en nuestros días, el médico usualmente describe estas señales acústicas mediante adjetivos relacionados con sonidos que le resulten cotidianos. Así, sólo la experiencia y el acuerdo mutuo han permitido cierto grado de uniformidad en la descripción de estos sonidos.

Las enfermedades del sistema cardiovascular y del sistema respiratorio han incrementado su incidencia alrededor del mundo, a tal punto que actualmente éstas se encuentran entre las 10 principales causas de morbi-mortalidad a nivel mundial. Es por ello que la auscultación del tórax con el estetoscopio constituye una técnica de aplicación sencilla desde el primer contacto con el paciente, no invasiva y de bajo costo, que posibilita la detección oportuna de algunos padecimientos cardio-respiratorios. Más aún, el campo de estudio de los sonidos cardiacos y de los sonidos respiratorios ha sido favorecido por el rápido crecimiento tecnológico y de procesamiento digital de señales, a tal grado que muchas de las limitaciones presentes en la auscultación torácica mediante el estetoscopio han sido superadas con la ayuda de métodos computarizados para la adquisición, procesamiento, análisis y visualización de los sonidos torácicos, entendidos como todas aquellas variaciones de presión acústicas que caen dentro del espectro audible y que se relacionan con eventos mecánicos fundamentales tales como la contracción cardiaca, el flujo sanguíneo, la mecánica respiratoria y el flujo en vías aéreas. Estos eventos funcionales generan sonidos que por la parte cardiovascular incluyen a los sonidos cardiacos (S1-S4), los murmullos y los soplos; en tanto que por la parte respiratoria corresponden a los sonidos ventilatorios normales, los sonidos adventicios y los sonidos producidos por los músculos respiratorios.

La complejidad de las estructuras biológicas de la cavidad torácica se ve reflejada en la complejidad de los sonidos que estas emiten. Estos sonidos, entre los que se cuentan los respiratorios y los cardiacos, son no estacionarios, lo que nos indica que poseen diversas estructuras e información cambiantes en el tiempo. Esta característica ha complicado la determinación de los mecanismos que dan génesis a algunos de ellos, aunque se han realizado desarrollos importantes, tanto en la fonocardiografía como en la fononeumografía, gracias a la creciente aplicación de técnicas de procesamiento digital de señales, tales como el análisis espectral, el reconocimiento de patrones y el análisis tiempo-frecuencia.

La principal motivación del análisis tiempo-frecuencia radica en que los espectros variantes en el tiempo existen en la naturaleza y por lo tanto debe entenderse su descripción y sus propiedades. La necesidad del análisis tiempo-frecuencia tiene sus semillas en el análisis clásico de las señales en los dominios temporal y espectral, y en particular en la comprensión de que el análisis aislado en estos dominios no es suficiente para describir completamente la naturaleza de estas señales, y que para ello se requiere obtener una distribución que represente la energía o la intensidad de la señal simultáneamente en el tiempo y en la frecuencia. En otras palabras se requiere una técnica de procesamiento capaz de tomar la información unidimensional de la señal y proporcionar simultáneamente el contenido en frecuencia, duración e intensidad de cada uno de sus componentes bien delimitados en una región del plano tiempo-frecuencia.

En el campo de los sonidos respiratorios el análisis tiempo-frecuencia ha sido una herramienta valiosa, y se han realizado diversos esfuerzos por encontrar una representación tiempo-frecuencia adecuada para analizarlos, e incluso para propósitos de reconocimiento automático de patrones. Desafortunadamente, en el campo de los sonidos respiratorios la herramienta clásica de análisis tiempo-frecuencia continúa siendo el espectrograma, cuya principal desventaja es su inherente compromiso entre su resolución temporal y su resolución espectral, el cual puede derivar en resultados pobres en registros de corta duración. Y a pesar de algunos esfuerzos por aplicar otras técnicas de representación tiempo-frecuencia para analizar los sonidos respiratorios, aún no se ha realizado una comparación cuantitativa entre la técnica clásica y otras de más reciente desarrollo. Así, el principal objetivo de esta tesis es realizar una comparación cuantitativa de diversas representaciones tiempo-frecuencia para el análisis de sonidos respiratorios, objetivo motivado por los prometedores resultados obtenidos en el campo de los sonidos cardíacos.

Las representaciones tiempo-frecuencia empleadas en esta tesis provienen de una gran variedad de técnicas de estimación, técnicas que podrían clasificarse de la forma siguiente: a) la clase general de distribuciones bilineales invariantes a corrimientos (el espectrograma, la distribución Wigner-Ville, y la distribución Choi-Williams), b) método de la reasignación (espectrograma reasignado), c) modelado autorregresivo variante en el tiempo (algoritmo de Burg, algoritmo de mínimos cuadrados recursivo, algoritmo de mínimos cuadrados recursivo con factor de olvido variable, y el filtro de Kalman), d) la clase afín de distribuciones (el escalograma), y e) descomposición de la señal en funciones base (el espectro de Hilbert-Huang).

En general, el cómputo de una representación tiempo-frecuencia mediante cada técnica depende de diferentes parámetros tales como el tipo y longitud de la ventana, o el orden del modelo. Como consecuencia, el desempeño de cada técnica tiempo-frecuencia dependerá de la adecuada selección de dichos parámetros. Así, tanto para la selección de parámetros como para la comparación de las diversas representaciones tiempo-frecuencia se emplearon señales de sonidos simulados, dado que su representación tiempo-frecuencia ideal se conoce de antemano, permitiendo la obtención de medidas de desempeño. Posteriormente, la técnica de análisis tiempo-frecuencia resultante fue aplicada a señales de sonidos torácicos reales adquiridos.

Las señales de sonido torácicos consideradas en esta tesis incluyen a los dos primeros sonidos cardiacos (S1 y S2), sonidos respiratorios adventicios discontinuos (estertores crepitantes) y discontinuos (estertores sibilantes), así como sonidos pulmonares y traqueales normales. Las representaciones tiempo-frecuencia ideales ofrecen una referencia contra la cual comparar el desempeño de los diversos métodos tiempo-frecuencia; sin embargo, para algunas señales no se contó con dicha referencia, por lo que el conocimiento obtenido de las señales previas resultó valioso.

Los índices empleados para evaluar el desempeño de las representaciones son: la correlación bidimensional central, aplicada de forma local (ρ_{promedio}) y global (ρ), el error cuadrático medio normalizado (NRMSE) y la correlación (ρ_{FI}) de la frecuencia instantánea estimada mediante el primer momento en el tiempo del centroide de la representación, y la resolución tiempo-frecuencia (res_{TF}).

Previo a la selección de las representaciones tiempo-frecuencia, cada una de estas fue adecuada tanto en dimensiones como en valores de intensidad, y su contenido visualizado adecuadamente mediante el diseño de una paleta de colores.

Las señales de sonidos respiratorios normales y adventicios se obtuvieron de dos bases de datos conformadas por los laboratorios de Fisiología Humana e Investigación en Procesamiento de Señales e Imágenes Biomédicas (LIPSIB) de la Universidad Autónoma Metropolitana Unidad Iztapalapa (UAM-I). La primera de ellas conformada por registros de pacientes con neumopatías intersticiales difusas (NID) y cuya adquisición se realizó en el Instituto Nacional de Enfermedades Respiratorias (INER); y la segunda conformada por registros de sujetos sanos y no fumadores adquirida en una cámara subamortiguada en la UAM-I. Además, se adquirieron registros cardiacos de un sujeto sano sin antecedentes de cardiopatías en el INER.

Ejemplos de los resultados obtenidos se presentan a continuación. Para el primer sonido cardiaco simulado: espectrograma reasignado ($\rho=0.6545$, $\rho_{\text{promedio}}=0.7750$, $\text{res}_{\text{TF}}=0.0002$), espectro de Hilbert-Huang ($\rho=0.5321$, $\rho_{\text{promedio}}=0.7668$, $\text{res}_{\text{TF}}=5.75 \times 10^{-7}$), y espectrograma clásico ($\rho=0.4303$, $\rho_{\text{promedio}}=0.7450$, $\text{res}_{\text{TF}}=0.0037$). Para el segundo cardiaco simulado: espectro de Hilbert-Huang ($\rho=0.6226$, $\rho_{\text{promedio}}=0.7258$, $\text{NRMSE}=0.1225$, $\rho_{\text{FI}}=0.9925$, $\text{res}_{\text{TF}}=8.39 \times 10^{-7}$), espectrograma reasignado ($\rho=0.6825$, $\rho_{\text{promedio}}=0.6914$, $\text{NRMSE}=0.1771$, $\rho_{\text{FI}}=0.9842$, $\text{res}_{\text{TF}}=2.11 \times 10^{-5}$), y espectrograma clásico ($\rho=0.4743$, $\rho_{\text{promedio}}=0.6239$, $\text{NRMSE}=0.1742$, $\rho_{\text{FI}}=0.9851$, $\text{res}_{\text{TF}}=0.0074$). Para la sibilancia simulada: espectro de Hilbert-Huang ($\rho=0.9247$, $\rho_{\text{promedio}}=0.9521$, $\text{NRMSE}=0.0601$, $\rho_{\text{FI}}=0.9982$, $\text{res}_{\text{TF}}=2.47 \times 10^{-6}$), espectrograma reasignado ($\rho=0.8821$, $\rho_{\text{promedio}}=0.9163$, $\text{NRMSE}=0.1400$, $\rho_{\text{FI}}=0.9903$, $\text{res}_{\text{TF}}=5.83 \times 10^{-5}$), y espectrograma clásico ($\rho=0.5300$, $\rho_{\text{promedio}}=0.8851$, $\text{NRMSE}=0.0074$, $\rho_{\text{FI}}=1$, $\text{res}_{\text{TF}}=0.0029$).

Además de estos valores cuantitativos, los resultados de esta tesis corroboraron algunas propiedades y características de las técnicas tiempo-frecuencia empleadas, y mostraron además las dificultades que sufre cada una de ellas para la estimación del espectro variante en el tiempo. De acuerdo a los índices de desempeño empleados, para diferentes señales resultaron diferentes representaciones tiempo-frecuencia como las más adecuadas, y además cada índice de desempeño señaló diferentes representaciones tiempo-frecuencia para una misma señal. Es decir, debido a las diferentes estructuras de las señales torácicas analizadas se encontró que diferentes representaciones tiempo-frecuencia resultaban más adecuadas para cada una señal particular, aunque en general destacan el Espectro de Hilbert-Huang, el espectrograma reasignado y la representación paramétrica obtenida por el método de Burg. Sin embargo, el espectro de Hilbert-Huang resulta, debido a las características de su método de descomposición, y al margen de sus limitantes, la técnica de análisis tiempo-frecuencia más adecuada y prometedora para el análisis de los sonidos respiratorios normales y adventicios de entre aquellas comparadas en este estudio.

I. INTRODUCCIÓN.

1. Descripción de la señal en el dominio del tiempo y de la frecuencia.

La necesidad de un análisis tiempo-frecuencia conjunto tiene sus orígenes y motivaciones en el análisis clásico de las señales en el dominio del tiempo y/o en el dominio de la frecuencia. El análisis de señales consiste en el estudio y caracterización de sus propiedades básicas, e históricamente, el análisis se desarrolló de manera simultánea al descubrimiento de las señales fundamentales en la naturaleza como el campo eléctrico, la corriente eléctrica y la onda de sonido. Sin embargo, actualmente el análisis de señales se ha extendido a diversos tipos de datos, incluyendo económicos y sociológicos [1].

El análisis de la variación de una señal con respecto del tiempo resulta fundamental debido a que el tiempo es una variable esencial para el estudio de diversos fenómenos. Sin embargo, si se desea entender de mejor forma el comportamiento del fenómeno, generalmente es ventajoso estudiar la señal asociada al fenómeno en una representación diferente, mediante su expansión en un conjunto completo de funciones. Desde un punto de vista matemático existe un número infinito de formas para definir la expansión. Así, además del tiempo, la representación más importante es la frecuencia [1, 2].

Los métodos de estimación del espectro* de potencia de una señal poseen una historia relativamente larga. Sin embargo, en un esfuerzo por entender la solución de la ecuación de onda, Jean Baptiste Joseph Fourier estableció en 1807 el concepto de las series de Fourier [3]. La idea de Fourier, que se basa en que una función discontinua puede representarse como la suma de funciones continuas – una idea que los grandes científicos de esa época, incluyendo a Laplace y Lagrange no dudaron en llamar absurda – se convirtió en una de las grandes innovaciones de las matemáticas y la ciencia.

La historia moderna de la estimación espectral comienza con el trabajo de J. W. Tukey, en 1949, el cual es la contraparte estadística del trabajo de Fourier realizado 142 años antes. Este resultado hizo posible un desarrollo activo del análisis espectral empírico por investigadores de todas las disciplinas científicas. Sin embargo, el análisis espectral resultaba computacionalmente costoso. Un gran avance computacional ocurrió con el descubrimiento en 1965 del algoritmo de la transformada rápida de Fourier realizado independientemente por J. Cooley y J. Tukey, y por Gordon Sande. El método Cooley-Tukey [4] hizo práctico el procesamiento de señales sobre formas de onda tanto en el dominio del tiempo como en el

* El término científico espectro fue acuñado por Sir Isaac Newton empleando la palabra latina para una imagen.

dominio de la frecuencia. La transformada de Fourier se convirtió no solamente en una descripción teórica sino también en una herramienta. Con el desarrollo de la transformada rápida de Fourier el campo del análisis espectral empírico saltó de la oscuridad a la luz, y actualmente es una disciplina central. Contribuciones posteriores importantes incluyen la introducción del análisis espectral de máxima entropía por John Parker Burg en 1967, el desarrollo de las ventanas espectrales por Emmanuel Parzen y otros investigadores en la década de 1950, la estimación espectral autorregresiva, y en la década de 1960, el criterio para la determinación del orden del modelo por Parzen e Hirotugu Akaike [3].

El análisis en la frecuencia o análisis espectral es fundamental debido a [5]: primero, es posible entender de mejor forma la fuente de la señal analizando espectralmente su forma de onda; segundo, la propagación de ondas a través de un medio generalmente depende de la frecuencia; y finalmente, la descomposición espectral simplifica nuestro entendimiento de la forma de onda. Aunque las sinusoidales simples son comunes en la naturaleza, es posible obtener la suma de sinusoides si la ecuación que gobierna el fenómeno físico es lineal. Así, aunque a menudo las señales lucen aleatorias, generalmente se trata de la simple superposición de ondas sinusoidales, lo cual es más simple de caracterizar. De acuerdo a lo anterior, surge la necesidad de descomponer una señal en sus frecuencias individuales, tarea que realiza el análisis de Fourier.

1.1. Clasificación de las señales.

Debido a la variedad de señales encontradas en la naturaleza, ha surgido una gran terminología para caracterizarlas [1]. Si una señal no cambia en algún sentido se dice que es estacionaria; en caso contrario es no-estacionaria. Si una señal dura un tiempo relativamente corto es generalmente llamada transiente. Si la señal se conoce explícitamente, se dice que es una señal determinista. Muy a menudo, debido a que el proceso físico que produce la señal es gobernado por eventos aleatorios, se dispone de muchas señales posibles, en cuyo caso se habla de una colección o ensamble de señales, una señal aleatoria, o una señal estocástica, donde una señal particular del ensamble es denominada una realización. Además, el contenido espectral se emplea en ocasiones para clasificar señales. Señales cuyo espectro está concentrado en una pequeña banda relativa a la frecuencia media son llamadas señales de banda estrecha, y en caso contrario son denominadas de banda ancha. Sin embargo estas clasificaciones son generales, puesto que para señales cuyo espectro cambia con el tiempo, no brindan una buena sensación de lo que acontece.

2. Importancia del análisis tiempo-frecuencia.

En el análisis de fenómenos fisiológicos es frecuente encontrar que los espectros de las señales asociadas son variantes con el tiempo. Sin embargo, el análisis en el tiempo y/o el análisis en la frecuencia no son suficientes para describir completamente la naturaleza de estas señales. En particular, el espectro indica que frecuencias están presentes en la señal pero no indica a que tiempo existen dichas frecuencias [2, 5]. Por lo tanto, surge la necesidad de describir cómo el contenido espectral de la señal cambia con el tiempo. Es decir, se requiere obtener una distribución que represente la energía o la intensidad de una señal simultáneamente en el tiempo y en la frecuencia. Mientras el espectro permite determinar que frecuencias existen en la señal, un análisis tiempo-frecuencia combinado indica la concentración de intensidad o energía en un tiempo y frecuencia particulares, i.e. permite determinar qué frecuencias existen en un tiempo particular [1].

Dos mecanismos físicos explican gran cantidad de situaciones donde el espectro de las señales varía con el tiempo [1] [5]. El primer mecanismo consiste en que la producción de frecuencias particulares depende de parámetros físicos que pueden cambiar con el tiempo. Íntimamente relacionada con esta cuestión se encuentran los posibles cambios del medio físico inmediato. Como ejemplo pueden considerarse las ondas sonoras producidas al final de un tubo, donde si la longitud del tubo es constante entonces el espectro de frecuencia de la señal de salida será constante a lo largo del tiempo, pero si la forma y longitud del tubo cambian con el tiempo, entonces la salida también variará con el tiempo. El segundo mecanismo consiste en que la propagación de las ondas en un medio, generalmente depende de la frecuencia.

2.1. Algunos ejemplos analíticos.

Ondas sinusoidales de duración finita.

En la figura I.1 se presentan dos señales compuestas por sinusoidales de duración finita a tres diferentes frecuencias. Ambas señales lucen diferentes en el dominio del tiempo (figura I.1.a), a pesar de que las dos señales poseen básicamente el mismo espectro (figura I.1.b), lo cual es razonable debido a que en ambos casos se tienen tres frecuencias. La diferencia radica en el tiempo en el cual dichas frecuencias existen, información que nos brinda su representación tiempo-frecuencia (figura I.1.c).

Chip con una amplitud Gaussiana.

En la figura I.2 se presenta una señal en la cual se observa que la frecuencia de sus oscilaciones se incrementa a lo largo del tiempo, sin embargo de su representación temporal (figura I.2.a) resulta difícil establecer el tipo de modulación en la frecuencia. Por su parte, la representación espectral (figura I.2.b) sólo indica el rango de frecuencias que contiene la señal, así como la intensidad de cada componente. La representación tiempo-frecuencia (figura I.2.c) indica el tiempo de ocurrencia e intensidad de dichos componentes.

Señales multicomponentes.

El análisis tiempo-frecuencia ofrece la ventaja de determinar si una señal es multicomponente o no, entendiendo por multicomponente a una señal que posee regiones bien delineadas en el plano tiempo-frecuencia [6]. En la figura I.3 se presentan dos ejemplos de señales multicomponentes.

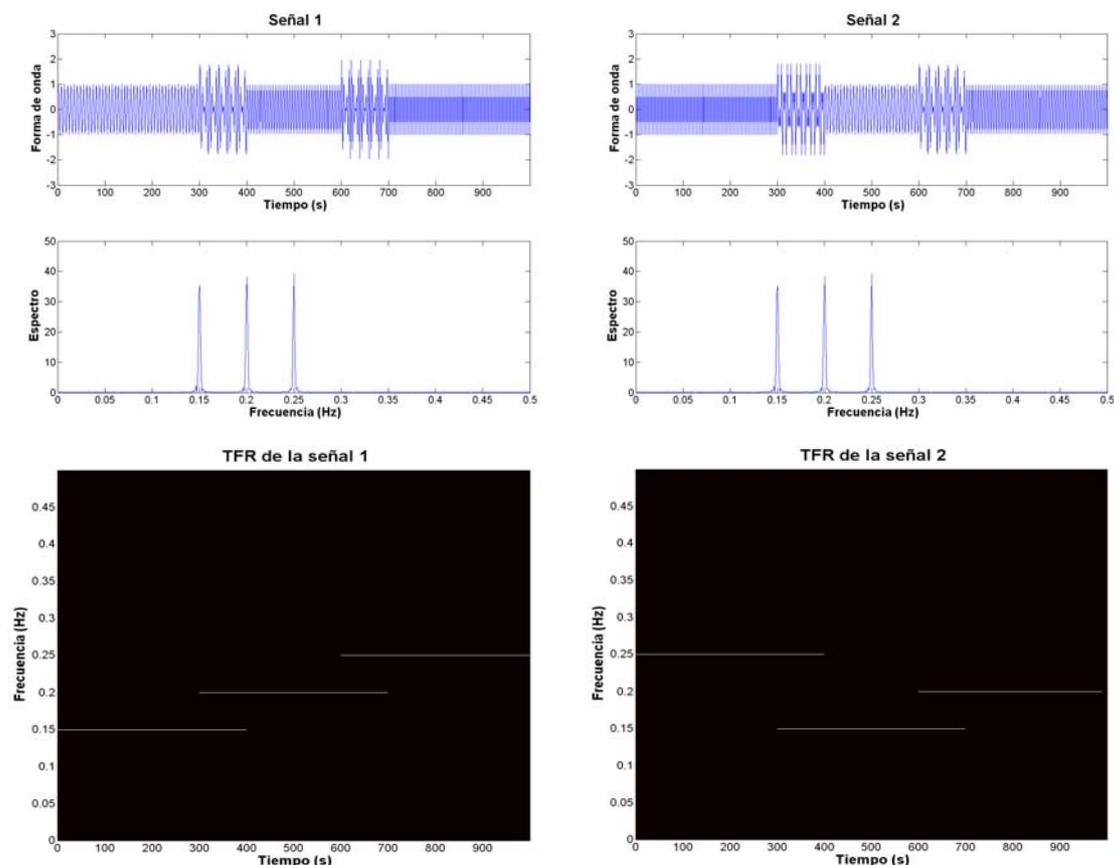


Fig. I.1. Dos señales compuestas por tres sinusoidales de longitud finita. a) Formas de ondas. b) Espectros. c) Representaciones tiempo-frecuencia.

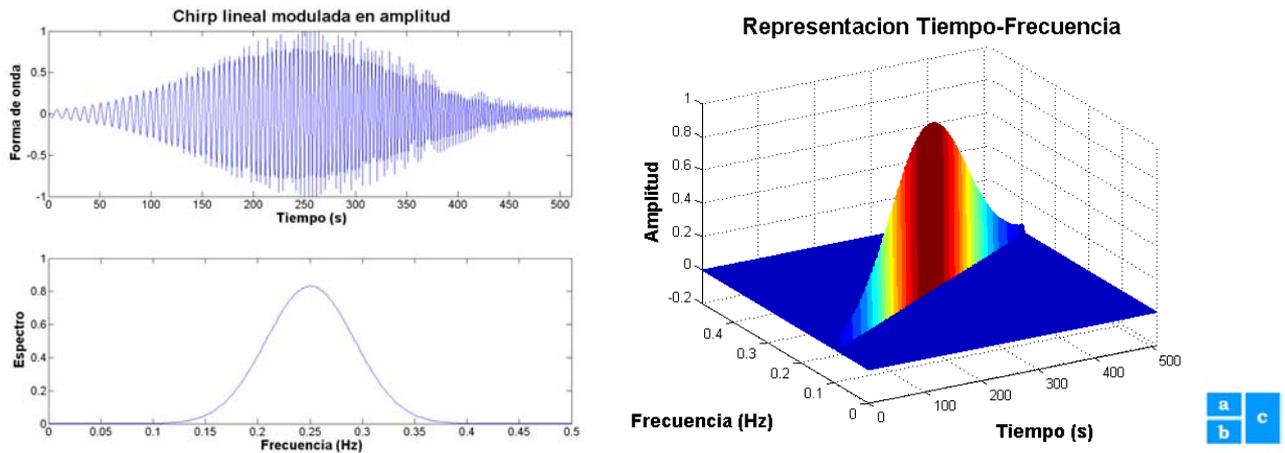


Fig. 1.2. Señal chirp lineal modulada en amplitud por una Gaussiana. a) Formas de onda. b) Espectro. c) Representación tiempo-frecuencia.

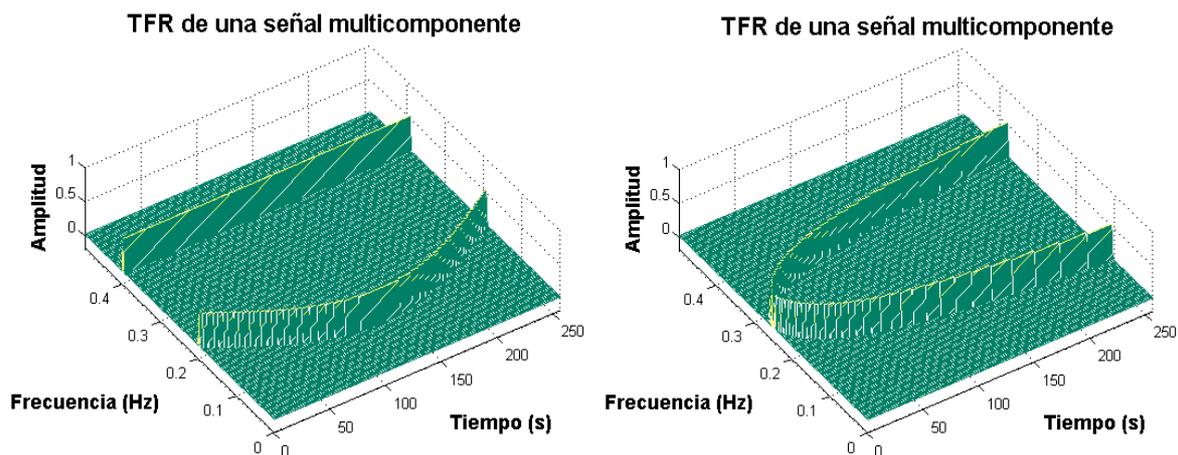


Fig. 1.3. Representaciones tiempo-frecuencia de dos señales multicomponentes.

3. Ideas fundamentales de las distribuciones tiempo-frecuencia.

El objetivo básico del análisis tiempo-frecuencia es generar una función que describa la densidad de energía de una señal simultáneamente en el tiempo y la frecuencia. Es decir, producir una función $P(t, \omega)$ igual a la intensidad en el tiempo t y la frecuencia ω ; donde $P(t, \omega)\Delta t\Delta \omega$ corresponde a la energía fraccional en la celda tiempo-frecuencia $\Delta t\Delta \omega$ en el punto (t, ω) [7]. Adicionalmente, conceptos fundamentales en las representaciones tiempo-frecuencia son los asociados con la señal analítica y la frecuencia instantánea [8]. A pesar de

que las señales en la naturaleza son reales, resulta ventajoso definir una señal compleja asociada con la señal real. El motivo de obtener la señal analítica es que presenta una fase a partir de la cual se puede obtener la frecuencia instantánea [1]. La frecuencia instantánea surgió con la llegada de la modulación en frecuencia para las transmisiones de radio en la década de 1920, pero fue hasta 1946 que la introducción de la señal analítica por Dennis Gabor, cristalizó los trabajos fundamentales [9]. A pesar de que la frecuencia instantánea es uno de los conceptos más primitivo e intuitivo, la descripción matemática exacta y la comprensión del concepto de frecuencias cambiantes esta lejos de ser obvio e incluso no es una cuestión resuelta [1]. Últimamente, ha existido la tendencia de definir la frecuencia instantánea como la derivada de la fase de la señal analítica. Una razón del porque la definición de la frecuencia instantánea es la derivada de la fase, radica en que su promedio temporal con la densidad de energía brinda la frecuencia media. Sin embargo, este resultado es válido para cualquier señal compleja [1]. Existen algunas dificultades respecto a la noción de frecuencia instantánea definida como la derivada de la fase de la señal analítica [10, 11]. Por ejemplo, la frecuencia instantánea puede no ser una de las frecuencias del espectro; la frecuencia instantánea puede ser negativa a pesar de que el espectro de la señal analítica es nulo para frecuencias negativas; la frecuencia instantánea puede estar fuera de la banda para una señal de banda limitada; la frecuencia instantánea puede ser continua y poseer un rango sobre un número infinito de valores para espectros que consisten únicamente de unos cuantos picos de frecuencia. Además, para obtener la señal analítica al tiempo t , debe conocerse la señal en todo momento. Entonces, debe mantenerse la mente abierta respecto a la definición apropiada de la señal compleja, i.e. la forma apropiada para definir la fase, amplitud y frecuencia instantánea. Probablemente, la última palabra sobre este tema aún no se ha dicho [1]. Finalmente, a pesar de la importancia de la señal analítica y la frecuencia instantánea su aplicación directa queda limitada a señales monocomponente, y esta es otra motivación para el análisis tiempo-frecuencia.

4. Comentarios finales y estructura de la tesis.

Los espectros variantes en el tiempo existen en la naturaleza y por lo tanto, su descripción y propiedades deben describirse y entenderse. A pesar de que en su estado actual, el análisis tiempo-frecuencia no es una teoría completa, las ideas y técnicas que se han desarrollado resultan útiles, es decir, brindan un conocimiento adicional con relación a la naturaleza de los fenómenos bajo estudio y en gran parte, los resultados hacen sentido con la intuición [1] [12].

El presente trabajo de tesis se enfoca a determinar una representación tiempo-frecuencia adecuada para el análisis de las señales acústicas producidas por la respiración humana en sujetos sanos y enfermos. La organización del escrito se describe a continuación.

El capítulo 2 aborda las características de las señales acústicas registradas sobre el área torácica a trabajar en esta tesis, los sonidos respiratorios, tanto normales como adventicios, y los dos primeros sonidos cardíacos. Para cada señal se brinda una descripción de su génesis y los mecanismos asociados, pero se hace énfasis en su análisis tiempo-frecuencia.

El capítulo 3 corresponde al planteamiento del problema de esta tesis. A partir de los antecedentes del capítulo 2, se establece lo valioso que resultaría un análisis tiempo-frecuencia conjunto de los sonidos respiratorios. En el mismo capítulo se establecen la hipótesis, y los objetivos, general y particulares, de esta tesis.

Una vez establecidas las señales a considerar y el objetivo de la tesis, en el capítulo 4 se introducen las diversas técnicas de análisis tiempo-frecuencia a evaluar: la distribución Wigner-Ville, el espectrograma, la distribución-Choi-Williams, el espectrograma reasignado, las representaciones resultantes de la estimación de los coeficientes del modelo autorregresivo variante en el tiempo por diversas técnicas (el método de Burg, el algoritmo de mínimos cuadrados estándar, el algoritmo de mínimos cuadrados con factor de olvido variable y el filtro de Kalman), el escalograma y el espectro de Hilbert-Huang. Para cada técnica se discuten sus principales características, ventajas y desventajas.

En el capítulo 5 se establece la metodología abordada para las diversas señales de sonidos torácicos. En particular se establece la simulación computacional de cada sonido, así como también la creación de su representación tiempo-frecuencia teórica a emplear como referencia en la evaluación de las técnicas mencionadas en el capítulo 4. Además, se establecen los índices de desempeños empleados con el propósito de cuantificar tal evaluación. Finalmente, se exponen las señales reales adquiridas para los diversos sonidos torácicos.

Con todas las herramientas anteriores, en el capítulo 6 se presentan los resultados obtenidos al procesar las señales acústicas con las técnicas de análisis tiempo-frecuencia. En este capítulo se presentan los valores de los índices de desempeño alcanzados con cada técnica para las señales simuladas, así como los resultados obtenidos al aplicar las técnicas seleccionadas a las señales reales adquiridas. De forma complementaria a los resultados mostrados en este capítulo, se incluye un CD con resultados adicionales.

Finalmente, el capítulo 7 se enfoca en la discusión de los resultados obtenidos, mientras que el capítulo 8 corresponde a las conclusiones y perspectivas de esta tesis.

Referencias.

- [1] L. Cohen, *Time-Frequency Analysis*. Englewood Cliffs, NJ: Prentice-Hall, 1995.
- [2] S. Qian, D. Chen, "Understanding the nature of signals whose power spectra change with time. Joint analysis," *IEEE Signal Processing Magazine*, pp. 52-67, March, 1999.
- [3] E. Robinson, "A historical perspective of spectrum estimation," *Proc. IEEE*, vol. 70, no.9, pp. 885-907, 1982.
- [4] J. Cooley, J. Tukey, "An algorithm for the machine calculation of complex Fourier Series," *Mathematics of Computation*, vol. 19, no. 90, pp. 297-301, 1965.
- [5] L. Cohen, "A primer on time-frequency distributions," en *Time-Frequency Signal Analysis*, editado por B. Boashash, pp. 3-42, Longman Cheshire, 1992.
- [6] B. Boashash, G. Jones, "Instantaneous frequency and time-frequency distributions," en *Time-Frequency Signal Analysis*, editado por B. Boashash, pp. 43-73, Longman Cheshire, 1992.
- [7] L. Cohen, "Time-Frequency Distributions – A Review," *Proc. of the IEEE*, vol. 77, pp. 941-981, 1989.
- [8] B. Boashash, "Estimating and interpreting the instantaneous frequency of a signal. Part 1: Fundamentals," *Proc. of the IEEE*, vol. 80, no. 4, pp. 520-538, 1992.
- [9] D. Gabor, "Theory of communication", *Journal of the IEE*, vol. 93, pp. 429-457, 1946.
- [10] J. Shekel, "Instantaneous frequency," *Proc. of IRE*, vol. 41, pp. 548, 1953.
- [11] L. Mandel, "Interpretation of instantaneous frequency," *Amer. J. Phys.*, vol. 42, pp. 840-846, 1974.
- [12] H. Hlawatsch, G. F. Boudreaux-Bartels, "Linear and quadratic time-frequency signal representations," *IEEE Signal Processing Magazine*, vol. 9, pp. 21-67, 1992.

II. ANTECEDENTES

En el capítulo anterior, se mencionó lo fundamental que resulta la variación en el tiempo de una señal y además, la importancia de su representación en el dominio de la frecuencia. En este capítulo se describen las características de las señales acústicas a tratar en esta tesis, los sonidos registrados sobre el área torácica, que incluyen a los sonidos respiratorios y a los sonidos cardiacos, entre otros. El tratamiento dado a cada sonido en este capítulo abarca una descripción de su génesis y los mecanismos asociados, además se enfatizan los esfuerzos en su análisis tiempo-frecuencia.

II.1 Sonidos respiratorios.

A principios del siglo XIX, tres años después de la invención del estetoscopio, Laennec, figura II.1, escribió en la reseña de su tratado clásico [1]:

“Fui consultado por una mujer joven con síntomas de padecimientos del corazón, en cuyo caso la percusión y aplicación de la mano fueron de poca ayuda en vista del alto grado de obesidad. Siendo el otro método recién mencionado [auscultación directa] inadmisibles debido a la edad y sexo de la paciente, recordé un simple y bien conocido hecho en acústica, ... la gran definición con la cual oímos el rasguño de una aguja en un extremo de una pieza de madera al aplicar nuestro oído sobre el otro extremo. Inmediatamente, sobre este indicio, enrolle un trozo de papel en forma de cilindro y aplique uno de sus extremos a la región del corazón y el otro a mi oído, y no fue una sorpresa pequeña encontrar que de ese modo pude percibir la acción del corazón en una forma mucho más clara y distinta de lo que yo había sido capaz con la aplicación inmediata de mi oído.”



Fig. II.1. René Laennec.

Aunque la auscultación directa puede rastrearse hasta Hipócrates, quien aconsejó acerca de la aplicación del oído sobre el tórax del paciente para escuchar sonidos transmitidos a través de la caja torácica, en los tiempos de Laennec esta técnica prácticamente había caído en desuso e incluso la idea resultaba desagradable. Así, Laennec experimentó con diversos cilindros y barras de madera, encontrando para su satisfacción que una barra sólida perforada con un calibre estrecho colocada entre su oído y el tórax del paciente mejoraba significativamente la transmisión del sonido. Hasta entonces, el examen clínico de la función pulmonar estaba dividida en 4 etapas: historia clínica, inspección, palpación y percusión [2], pero a partir de Laennec se popularizó la etapa de auscultación [3]. Laennec estaba más interesado en la auscultación pulmonar mas que en la cardiaca [1] *“La auscultación de los sonidos respiratorios con un cilindro (estetoscopio) produce señales audibles fácilmente interpretables y capaces de indicar la presencia y alcance de muchos desordenes en los órganos dentro de la caja torácica.”*

Una década después de que Laennec describiera su estetoscopio, diversos modelos fueron elaborados por otros médicos. En América, Nicholas Comins diseñó el estetoscopio biaural, en el cual el sonido era canalizado a través de tuberías a ambos oídos desde una sola membrana montada en una copa que era colocada sobre el tórax del paciente [2]. Este dispositivo permitió al médico no tener que asumir posiciones incómodas durante el examen. En palabras de Comins [4]: *“Es sorprendente que el descubrimiento de la auscultación no haya sugerido un instrumento flexible. Pero Laennec, como el inmortal Arquímedes, lidió con grandes ideas en regiones inexploradas del conocimiento... Se le ocurrió al escritor que ambos oídos podrían ser empleados simultánea y ventajosamente en la examinación mediante el estetoscopio... El instrumento adaptado para este propósito admite una fácil adaptación simultánea al paciente y a ambos oídos.”* Versiones más actuales del estetoscopio incluyen al estetoscopio electrónico, el cual incluso ha sido empleado para la evaluación a distancia de los sonidos respiratorios vía telefónica [5].

En nuestros días, los padecimientos pulmonares son una causa mayor de enfermedad alrededor del mundo. Las infecciones pulmonares tales como la neumonía y la bronquitis son comunes, y las enfermedades pulmonares intersticiales han incrementado su incidencia [6]. Aunque el diagnóstico de las enfermedades pulmonares se facilita con la auscultación pulmonar mediante el estetoscopio y la técnica es ampliamente empleada en la práctica clínica debido a su bajo costo, rapidez, fácil realización y sus características no invasivas e inocuas [2], este procedimiento presenta muchas limitaciones [6]. Entre las limitaciones del procedimiento de auscultación se encuentran: 1) es un proceso subjetivo que depende de la

audición, habilidad y experiencia del médico para diferenciar entre diferentes patrones de sonido, 2) no resulta sencillo obtener mediciones cuantitativas, 3) no es posible el almacenamiento de la información de sonidos respiratorios para su posterior análisis, seguimiento y correlación con otras señales fisiológicas y 4) cada estetoscopio posee una respuesta en frecuencia diferente, e incluso la campana y el diafragma de un mismo estetoscopio poseen diferentes características de transmisión, especialmente a bajas frecuencias [7].

Afortunadamente, en las últimas cuatro décadas, los métodos computarizados para el registro y análisis de sonidos respiratorios han superado muchas de las limitaciones de la auscultación simple, provocando que la actividad científica en el campo de la acústica respiratoria se incremente considerablemente. Actualmente, el análisis acústico respiratorio puede cuantificar cambios en los sonidos respiratorios, realizar registros permanentes de las mediciones hechas, y producir representaciones que ayuden al diagnóstico y tratamiento de pacientes que sufren de enfermedades pulmonares [6].

El estudio computarizado de los sonidos respiratorios ha estado inmerso en el rápido cambio tanto tecnológico como del procesamiento digital de señales. Gran parte del conocimiento ganado en el campo ha resultado del uso de una amplia variedad de técnicas de adquisición, procesamiento, análisis y visualización [8]. A la técnica de registro de los sonidos respiratorios se le conoce como fononeumografía. Los sistemas de registro de sonidos respiratorios usualmente consisten de un pneumotacógrafo para la medición de flujo de aire o volumen, y micrófonos para la captura de los sonidos. En la fononeumografía la intensidad del sonido se despliega conjuntamente con el flujo de aire. A continuación se presenta un panorama general de las técnicas empleadas en la adquisición, procesamiento, análisis y visualización, dejando al final y enfatizando las correspondientes al procesamiento de señales, dado que es el enfoque de esta tesis.

En lo referente a los métodos de adquisición, los sonidos registrados del sistema respiratorio son capturados por micrófonos electret acoplados por aire o sensores de contacto (acelerómetros) situados sobre el tórax, la tráquea u otro lugar. Junto con estos sonidos, es común adquirir simultáneamente otras señales fisiológicas tales como el flujo respiratorio, cambios en el volumen pulmonar o en la presión intratorácica y la saturación de oxígeno. Aunque típicamente sólo se emplea un canal de adquisición, cada vez resulta más común el empleo de múltiples canales [9-13]. Es frecuente aplicar un filtro analógico del tipo pasa altas o pasabanda e inclusive pasa bajas, a los sonidos capturados, donde las frecuencias de corte de estos filtros dependen de la aplicación particular. Los registros de sonidos respiratorios son

posteriormente digitalizados por un convertidor analógico-digital y almacenados en una computadora. Las frecuencias de muestreo empleadas durante el registro varían ampliamente.

Con respecto a los sonidos respiratorios resulta conveniente establecer algunos conceptos en cuanto a su generación así como la nomenclatura y su clasificación. La clasificación y nomenclatura de los sonidos respiratorios ha sido algo caótica desde los días de Laennec [14]. El propio Laennec clasificó los eventos acústicos generados por la ventilación de los pulmones que escuchaba y los relacionó con los hallazgos anatómicos y patológicos que encontraba tras realizar una autopsia [2]. La posterior traducción de su trabajo, y los inevitables esfuerzos por mejorarlo, han llevado en algunos casos a la confusión en la nomenclatura. Sin embargo, conforme los sonidos son más fácilmente registrados y clasificados en términos objetivos y físicos, los esfuerzos por uniformizar la terminología y clasificación están mostrando resultados útiles [3].

Los sonidos respiratorios corresponden a todos aquellos sonidos relacionados a la respiración, incluyendo a los sonidos ventilatorios, sonidos adventicios y a los sonidos producidos por los músculos respiratorios. Los sonidos de voz producidos durante la respiración no se consideran sonidos respiratorios. Por su parte, los sonidos ventilatorios son aquellos sonidos producidos al respirar, excluyendo a los sonidos adventicios, y escuchados o registrados sobre la pared torácica, la tráquea o en la boca de un sujeto sano. La generación de los sonidos ventilatorios se relaciona con el flujo de aire en el tracto respiratorio. Acústicamente se caracterizan como un ruido con un espectro de banda ancha cuyo rango de frecuencia depende del lugar de su recolección sobre el tórax. Mientras que los sonidos adicionales y añadidos a los sonidos ventilatorios se denominan sonidos adventicios, los cuales se clasifican de acuerdo a su duración en continuos y discontinuos. La presencia de sonidos adventicios usualmente indica un desorden pulmonar. Finalmente, tanto los sonidos ventilatorios como los adventicios escuchados o detectados sobre la pared torácica o dentro del tórax se denominan sonidos pulmonares para distinguirlos de aquellos escuchados sobre la tráquea o en la boca. En la figura II.2 se presenta un diagrama con las relaciones entre diversos términos asociados a los sonidos respiratorios [15].

En las siguientes secciones se describen algunas características de los sonidos pulmonares, haciendo énfasis en los esfuerzos por analizar esta señal acústica en los dominios del tiempo, la frecuencia y tiempo-frecuencia.

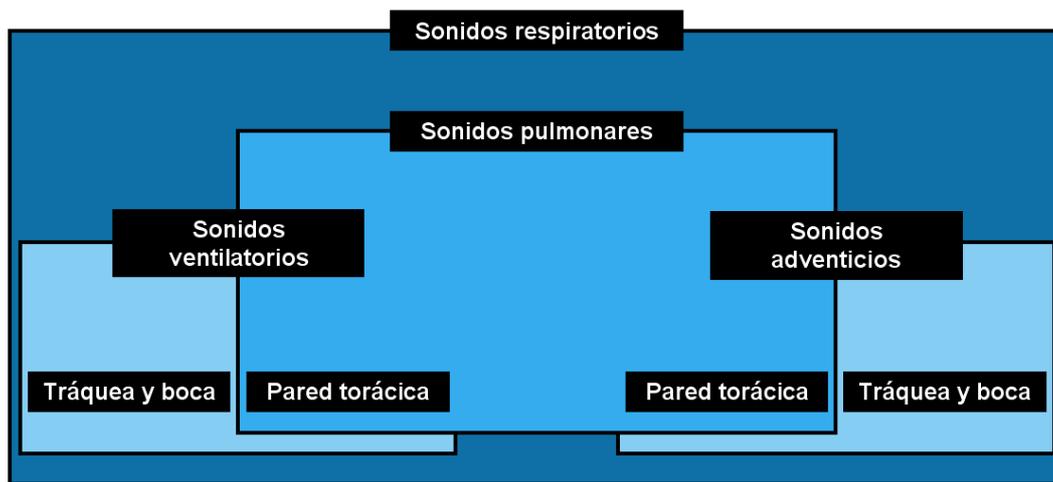


Fig. II.2. Relaciones entre los términos sonidos respiratorios, sonidos pulmonares, sonidos ventilatorios y sonidos adventicios.

II.1.1. Sonidos pulmonares normales.

Los sonidos generados en pulmones y vías aéreas saludables por la respiración normal espontánea, difieren de acuerdo a la localización de su registro y varían a lo largo del ciclo ventilatorio [2]. El origen de los sonidos producidos por la ventilación no está completamente claro y probablemente estén involucrados diversos mecanismos. El pulmón por sí mismo no puede generar sonido si no existe un flujo de aire; las diferencias de presión entre estructuras dentro del tórax o diferentes niveles de volumen pulmonar no pueden inducir sonidos por sí mismos en ausencia de un flujo de aire. Probablemente, los sonidos pulmonares sean inducidos por la turbulencia de aire a nivel de los bronquios lobares o segmentales. En los bronquios más pequeños, la velocidad del gas decrece por debajo del valor crítico necesario y no alcanza el número de Reynolds necesario para inducir turbulencias [16]. Por lo tanto, el flujo de aire en las vías aéreas más pequeñas se considera laminar y silente [17].

Desde un principio, el origen de los sonidos pulmonares normales escuchados sobre la superficie torácica ha creado controversia. En 1955, motivados por la aplicación del espectrógrafo de sonido al estudio de los sonidos cardiacos, se realizó la primera representación tiempo-frecuencia de los sonidos pulmonares normales y adventicios [18]. Empleando micrófonos de condensador y registrando las señales eléctricas resultantes sobre un disco magnético para diferentes frecuencias con un filtro variable, se representó la amplitud acústica mediante valores de intensidad de luz sobre un gráfica tiempo contra frecuencia. Durante el análisis se registraron amplitudes significativas sobre un rango de frecuencia de 0 a 400 Hz durante la inspiración y sólo una débil respuesta durante la espiración, compuesta principalmente de sonidos cardiacos. En este estudio se concluyó que

los sonidos “vesiculares” inspiratorios normales resultan de la turbulencia creada cuando corrientes de aire se expanden dentro de la multitud de sacos alveolares, no siendo necesario el movimiento de aire dentro de las vías aéreas más grandes para su producción. En contraparte, se concluyó que los sonidos espiratorios tienen su origen en las bifurcaciones de las vías aéreas más grandes. Finalmente, se señaló que el carácter de los sonidos pulmonares detectados sobre la superficie del tórax está determinada en gran medida por las propiedades de resonancia y conducción del sonido del tejido pulmonar entre las vías aéreas más grandes y el punto de detección, que la frecuencia natural del tórax en general varía inversamente con su volumen, y que procesos degenerativos en el pulmón y la pleura pueden ocasionar que el tórax resuene diferente de lo normal. Actualmente, a la representación tiempo-frecuencia de los sonidos respiratorios se le conoce como respirosonograma o sonograma.

Recientemente, en 1995 se encontró que los sonidos pulmonares normales son al menos producidos en parte de forma local como lo indica la evidencia encontrada en sujetos con la glotis cerrada durante una maniobra de Valsalva, en experimentos in vitro con la tráquea cerrada [2], y en pacientes a los cuales se les ha extraído la laringe [19]. Con respecto a los hallazgos encontrados al someter a los sujetos a una respiración intensa se ha planteado que los sonidos pueden deberse a una redistribución del gas dentro del pulmón desde el ápice hacia la base, donde la presión negativa ejercida por el diafragma es mayor, o tal vez a la presencia de vibraciones producidas por el estiramiento del tejido pulmonar [2].

Los avances técnicos en el registro y análisis de los sonidos han permitido realizar mediciones y comparaciones más precisas de su intensidad y contenido en frecuencia. Empleando la FFT para analizar los sonidos pulmonares de sujetos sanos en el dominio de la frecuencia, se encontró que la potencia de la señal decrece exponencialmente conforme se incrementa la frecuencia [20]. Las máximas frecuencias de los sonidos pulmonares inspiratorios fueron de 446 ± 143 Hz sobre la base del pulmón derecho, de 475 ± 115 Hz sobre la base del pulmón izquierdo, de 434 ± 130 Hz sobre la región inter-escapular y de 604 ± 302 Hz sobre el tórax anterior derecho. Mientras que para la fase espiratoria las máximas frecuencias fueron menores, a decir, 286 ± 53 Hz sobre la base del pulmón derecho, 284 ± 47 Hz sobre la base del pulmón izquierdo, de 338 ± 77 Hz sobre la región inter-escapular, y de 406 ± 205 sobre el tórax anterior derecho. En diferentes investigaciones se corroboró que el espectro de frecuencia de los sonidos pulmonares normales presenta una caída exponencial en sus amplitud que inicia entre 50-75 Hz y que se extiende hasta los 350 Hz [21], los 500 Hz [20] ó 1000 Hz [22], rangos que dependen de la investigación en particular. Además, se ha encontrado que ciertos espectros de frecuencia contienen numerosos picos o frecuencias

dominantes [21, 22]. Aunque se han observado frecuencias pico mayores en los sonidos registrados sobre el ápex del pulmón que sobre su base [23]; frecuencias que se han asociado a diferencias en las características de filtración y producción del sonido.

Ha sido reconocido que los sonidos pulmonares en los niños poseen una calidad distinta que generalmente se atribuye a la transmisión acústica a través de pulmones más pequeños con paredes torácicas más delgadas, por lo que el sonido parece ser más cercanamente producido a los oídos del médico durante la auscultación [24]. Siguiendo esta idea, la FFT fue empleada para investigar el efecto de la edad y el tamaño corporal sobre los sonidos pulmonares normales. En un primer estudio se analizaron los espectros de frecuencia en recién nacidos y se compararon los resultados obtenidos con los reportados para niños de mayor edad [25]. No se encontraron diferencias en las frecuencias pico medias ni en el rango de frecuencias entre los recién nacidos y los adultos. Argumentando que lo anterior resultaba contrario a la experiencia clínica, se realizó otro estudio para comparar el espectro de los sonidos pulmonares adquiridos bajo condiciones normales de auscultación empleando registros de los lóbulos superior e inferior [26]. Se concluyó que los sonidos pulmonares normales de los recién nacidos contienen componentes de frecuencia más altos que en adultos, y se postuló que esta diferencia resulta del menor filtrado de los sonidos pulmonares por los recién nacidos y por lo tanto, estas diferencias deben considerarse al valorar la presencia de anormalidades.

El análisis en la frecuencia también ha sido empleado para determinar si un componente de sonido proveniente de la contracción del músculo esquelético puede contribuir al sonido pulmonar inspiratorio [27]. Los datos mostraron que a frecuencias por debajo de los 200 Hz el componente del músculo esquelético se incrementa mientras que el componente de sonido pulmonar decrece, y se concluyó que el ruido generado por el músculo esquelético contamina seriamente al sonido pulmonar.

Además, se han realizado estudios acerca de la generación y transmisión de los sonidos pulmonares empleando la función de transferencia y la función de coherencia entre los sonidos registrados sobre la tráquea y sobre la superficie torácica de un pulmón canino seco e inflado [28]. Las observaciones sugieren que durante la espiración los sonidos registrados sobre la superficie se derivan de los sonidos traqueales, mientras que durante la inspiración los sonidos superficiales son generados de manera más próxima a la superficie pulmonar. Es decir, los sonidos espiratorios provienen de una fuente más central que los sonidos inspiratorios. Estas observaciones se han corroborado utilizando un modelo flexible de las vías aéreas caninas [29]. Los hallazgos sugieren que las fuentes predominantes de

generación de los sonidos pulmonares son las grandes vías aéreas para la fase inspiratoria y las vías aéreas de tamaño medio para la fase espiratoria.

En general, se tiene que el espectro de potencia de los sonidos pulmonares normales carecen de picos discretos y su carácter no es musical [30]. Y parece bien establecido que su componente inspiratorio es generado principalmente dentro de las vías aéreas lobares y segmentales, mientras que su componente espiratorio proviene de localidades más proximales [31-33].

El origen de los sonidos pulmonares también se ha investigado empleando el registro multicanal utilizando dos micrófonos ubicados en diferentes posiciones a lo largo de una línea horizontal del tórax [34-36]. Los autores plantean la hipótesis que las ondas sonoras registradas en dos sitios separados por unos cuantos centímetros sobre la pared torácica tenderán a ser relativamente similares si son producidas por una fuente distante común, mientras que se parecen en menor medida si son producidas localmente por fuentes separadas cercanas a cada micrófono. Se derivó un índice de cancelación o índice de substracción de intensidad el cual consiste en el cociente entre la substracción y la adición de la amplitud de ambas señales. Los hallazgos fueron consistentes con una fuente intra-pulmonar y probablemente intra-lobar, para la fase inspiratoria y una fuente ubicada en las vías áreas superiores, por lo menos parcialmente, para los sonidos de la fase espiratoria [35]. Además, se concluyó que es común encontrar una diferencia significativa en la amplitud de los sonidos pulmonares registrados sobre áreas homólogas opuestas del tórax, y que los sonidos pulmonares registrados sobre los ápices pulmonares son posiblemente producidos más centralmente que aquellos registrados sobre las bases pero que la tráquea no es la fuente de estos sonidos [34]. Experimentos similares fueron realizados empleando una función de coherencia para medir el grado de correlación entre el par de micrófonos [2]. Se obtuvo una buena correlación entre los 50 y 600 Hz con micrófonos separados por 6 cm, únicamente entre los 100 y 150 Hz para micrófonos separados entre 12 y 16 cm, y ninguna correlación entre micrófonos situados en lados opuestos de la caja torácica. Se indicó que un sitio intra-lobar o incluso intra-lobular era el responsable de la mayor parte del sonido pulmonar normal, y que quizás fuera una fuente de baja frecuencia situada centralmente. La aplicación de la función de coherencia a sonidos traqueales y sonidos registrados en la región posterior derecha llevó a concluir que parece improbable que los sonidos traqueales contribuyan a los sonidos escuchados en la base del pulmón, mientras que los sonidos traqueales constituyen un elemento significativo de los sonidos escuchados sobre el ápice pulmonar.

Otras investigaciones han examinado la relación entre los diferentes niveles de flujo de aire en la boca versus la amplitud y contenido espectral de los sonidos pulmonares. Se ha encontrado que la relación entre la amplitud del sonido pulmonar normal y el flujo de aire parece ser substancialmente lineal [34]. En contraste, se ha encontrado que la amplitud del sonido pulmonar durante la inspiración varía directamente con el cuadrado del flujo medido en la boca, mientras que cambios en la posición del sujeto y en el patrón respiratorio solo alteran la ganancia pero no su relación funcional [37].

La amplitud de los sonidos pulmonares varía en gran medida de un sujeto a otro incluso ventilando a flujos de aire similares, además incluso la corrección por altura, área de la superficie corporal y peso de los sujetos no resulta exitosa para abolir esta variabilidad [2]. Además, para un mismo sujeto estos sonidos exhiben una notable variación de amplitud a lo largo del tórax [30, 38-41]. En general, se concluye que la intensidad del sonido pulmonar es una buena medida de la ventilación regional si la intensidad es compensada por diferencias locales en la transmisión del sonido. Empleando el mapeo de la intensidad del sonido pulmonar sobre la superficie torácica se estudió en mejor medida su distribución espacial inhomogénea [13, 42, 43]. Se encontró que la amplitud se incrementa hacia la base posterior, decrece hacia la base anterior y permanece aproximadamente estable en un plano horizontal. Además se encontraron diferencias significativas entre regiones homólogas de los hemitórax izquierdo y derecho, y una marcada variación inter-sujeto [42]. Sin embargo, aunque se ha demostrado en detalle estas variaciones en amplitud, ningún estudio ha explicado las causas de este fenómeno [30].

Como ya se mencionó, el análisis de los sonidos pulmonares en el dominio de la frecuencia se ha realizado mediante el algoritmo de la transformada rápida de Fourier. Sin embargo, ha habido esfuerzos por utilizar otras alternativas para el análisis espectral, en especial estimadores paramétricos basados en el modelado autorregresivo. A mediados de la década de 1980, se empleó por primera vez el modelado AR para estimar las características de los sonidos pulmonares con propósitos de clasificación automática [44]. Posteriormente, se propuso un esquema que emplea el modelado AR para estimar la fuente y las características de transmisión de los sonidos pulmonares [45]. Además, se han realizado estudios encaminados a optimizar el uso del modelado AR para el análisis de los sonidos pulmonares normales, en particular la selección de un orden adecuado [46, 47]. Empleando el algoritmo de Burg para la determinación de los coeficientes AR se encontró que para los sonidos pulmonares se requieren órdenes en el intervalo de 6-8 mientras que para los sonidos traqueales se requieren al menos órdenes en el intervalo de 12-16. Adicionalmente, se realizó

la representación polar de las raíces AR calculadas a partir de los coeficientes AR, y se mostró que los sonidos pulmonares normales de un grupo de individuos están caracterizados por una baja variabilidad, lo cual sugirió que el modelado AR puede proporcionar una representación alternativa de estos sonidos [46]. El empleo del modelado AR ha sugerido que los efectos del volumen pulmonar sobre los sonidos pulmonares normales pueden ser separados de los efectos más prominentes del flujo de aire [48].

En el dominio tiempo-frecuencia la técnica clásica de análisis continúa basándose en el espectrograma. Pasaron más de dos décadas para producir una representación tiempo-frecuencia distinta de las imágenes en escala de grises a diferentes frecuencias producidas por el espectrógrafo de sonido [18]. Empleando equipos más modernos, en 1978 se presentaron mapas tridimensionales de los sonidos pulmonares que consistían en un arreglo en cascada de espectros de potencia [49]. Sin embargo, esta representación presenta la desventaja de ocultar información debido al acomodo sucesivo de curvas espectrales particularmente al desplegar diversos ciclos respiratorios, además de que las gráficas tridimensionales requieren un considerable tiempo computacional. Una década después, tomando ventaja de la FFT se produjeron los denominados respirosogramas digitales, en los cuales se presenta el tiempo y contenido en frecuencia de los sonidos respiratorios junto con su intensidad acústica en una escala de colores [30]. La FFT se obtuvo sobre segmentos sucesivos de 100 ms de duración, con un traslape entre segmentos sucesivos del 50%. Previamente al cálculo de la FFT, cada segmento fue delimitado en duración mediante una ventana de Hann. De manera similar al trabajo pionero de 1955, se presentaron respirosogramas de sonidos pulmonares normales y adventicios. Se encontró que los sonidos inspiratorios se extendían hasta los 800 Hz con intensidades máximas por debajo de 200 Hz, mientras que el contenido en frecuencia de los sonidos espiratorios resultó menor. Además, se encontró un ruido persistente de baja frecuencia y alta intensidad que posteriormente, fue identificado como sonido cardiaco, así como sonidos de baja frecuencia, por debajo de los 80 Hz, *en instantes de tiempo donde el sujeto mantenía la respiración*. Estos sonidos resultaban más notorios durante la respiración y se relacionaron con la actividad muscular.

II.1.2. Sonidos traqueales normales.

Los sonidos respiratorios detectados sobre la tráquea y las vías aéreas superiores se caracterizan por un carácter tubular o hueco, el cual de acuerdo a Laennec demostraba claramente que el aire está pasando a través de un canal más amplio que el de los alvéolos [1]. Estos sonidos poseen aproximadamente intensidades similares durante las fases inspiratoria y

espiratoria [15]. Los sonidos traqueales escuchados sobre la horquilla esternal o sobre la parte lateral del cuello, actualmente son un tópico de gran interés. En la década de 1920 surgió la preocupación de distinguir entre el contenido en frecuencia de los sonidos traqueales y de los sonidos pulmonares debido a la importancia de estos sonidos en el diagnóstico de la consolidación pulmonar. Los sonidos traqueales escuchados sobre pulmones consolidados muy probablemente representen una transmisión mejorada hacia la periferia del sonido producido en las grandes vías aéreas. El pulmón consolidado actúa como un medio conductor acústico continuo que no atenúa la transmisión del sonido, como lo hace un pulmón inflado. Probablemente el hallazgo de sonidos traqueales sobre un pulmón consolidado represente el reemplazo del sonido vesicular producido localmente (abolido por la consolidación y una consecuente ausencia de ventilación local) por sonido transmitido de un generador más central [19]. La generación de los sonidos traqueales está principalmente relacionada con el flujo turbulento de aire en las vías aéreas superiores, incluyendo la faringe, la glotis y regiones sub-glóticas. El flujo turbulento en glotis causa fluctuaciones de presión dentro de las vías aéreas. Las ondas de presión acústica en el interior del gas y el movimiento de las paredes de las vías aéreas también parecen contribuir a las vibraciones que alcanzan la superficie del cuello y que son registradas como sonidos traqueales. Debido a la corta distancia entre las varias fuentes de sonido en las vías aéreas superiores y los sensores en el cuello, y sin interposición de tejido pulmonar, los sonidos traqueales generalmente son interpretados como sonidos pulmonares más puros o menos filtrados [30]. El análisis en la frecuencia también ha mostrado la compatibilidad de resultados con la generación del sonido durante la inspiración mediante flujo turbulento en una red con bifurcaciones [50].

Empleando el espectrógrafo de sonido, se registraron sonidos traqueales con frecuencia de 60 a 600 Hz durante la inspiración, y hasta 700 Hz durante la espiración, concluyendo que el tono de los sonidos durante la inspiración y la espiración es a groso modo el mismo, pero que el sonido espiratorio es apreciablemente más intenso [18]. Además, se señaló una variación en el nivel de frecuencia central de los sonidos traqueales y que éstas parecen depender de las dimensiones de la tráquea, tanto de la longitud como del calibre. La representación tiempo-frecuencia también sirvió para mostrar que la neumonía produce cambios en las propiedades de conducción del sonido del tejido pulmonar, y por lo tanto, la espiración que ordinariamente es relativamente silente se vuelve audible como resultado de la transmisión de los sonidos producidos en el árbol traqueo-bronquial durante la consolidación. Finalmente, se mostró la presencia de sonidos cardiacos en el rango de frecuencias por debajo de los 240 Hz.

Se ha encontrado que el espectro de los sonidos traqueales exhibe picos y depresiones que se relacionan con las dimensiones de las vías aéreas y que dependen de la densidad del gas, lo cual indica su origen de resonancias dentro de las vías aéreas [30]. Las características del flujo turbulento están influenciadas por las dimensiones de los conductos, y las dimensiones traqueales son una función de la altura corporal [2]. Analizando sonidos traqueales a flujos estándar en niños y adultos se encontró que una fuerte relación entre las frecuencias de corte y la altura corporal; los niños con tráqueas más cortas tuvieron mayores frecuencias de corte que los adultos [51].

En resumen, la señal de sonido traqueal es intensa, cubriendo un rango más amplio de frecuencias que los sonidos pulmonares, con fases respiratorias bien definidas y cercanamente relacionadas con el flujo de aire [31, 52]. Se ha encontrado que una ley de potencia, con un exponente menor a 2, se ajustaba bien para la relación entre la potencia de los sonidos y el flujo [50]. En otro estudio con la FFT se encontró que las características espectrales de los sonidos traqueales dependen fuertemente del flujo de aire, un incremento en el flujo de aire eleva los componentes de alta frecuencia de su espectro de potencia [20, 53].

II.1.3. Sonidos adventicios.

Los sonidos adventicios son aquellos sonidos agregados a los sonidos pulmonares normales y se clasifican de acuerdo a su duración en continuos (sibilancias) y discontinuos (crepitancias). La presencia de los sonidos adventicios usualmente indica un desorden pulmonar [15]. Los mecanismos y sitios de producción de los sonidos adventicios no han sido del todo esclarecidos. Se piensa que los sonidos adventicios resultan de movimientos anormales de las vías aéreas o de materiales dentro de éstas durante la respiración [54]. Además la nomenclatura resulta mucho más confusa que la empleada para los sonidos respiratorios normales [14], aunque se han realizado diversos esfuerzos para coincidir sobre los nombres de las categorías más importantes. En la tabla II.I se muestra una clasificación de los sonidos adventicios, comparando la nomenclatura en uso actual con algunos sinónimos comunes y los términos originales de Laennec [19, 54].

Los sonidos respiratorios adventicios han sido objeto de una gran cantidad de investigaciones debido a que su aparición resulta común en diversas enfermedades cardiorespiratorias. A continuación se presentan las características de los sonidos adventicios continuos y discontinuos, así como algunos de los esfuerzos realizados por estudiarlos tanto en el dominio del tiempo como la frecuencia.

TABLA II.1. CLASIFICACIÓN DE LOS SONIDOS ADVENTICIOS.

Características acústicas	Nomenclatura recomendada por la ATS	Término en algunos libros de texto	Término británico	Término original de Laennec
Sonidos discontinuos, explosivos e interrumpidos				
Intensos, bajos en tono; valores promedio de IDW=1.25 ms, 2CD=9.32 ms	Crepitancia gruesa (coarse crackle)	Rale grueso	Crackle	Rale muquex ou gargouillement
Menos intensos, de menor duración y tono más alto que los anteriores; valores promedio de IDW=0.92 ms; 2CD=6.02 ms	Crepitancia fina (fine crackle)	Rale fino	Crackle	Rale humide ou crepitation
Sonidos continuos, duración mayor a 250 ms.				
Tono alto; frecuencia dominante > 400 Hz; sonido tipo siseo	Sibilancia (wheeze)	Rhonchus sibilante	Wheeze de tono alto	Rale sibilant sec ou sifflement
Tono alto; frecuencia dominante < 200 Hz; sonido tipo ronquido	Rhonchus	Rhonchus sonoro	Wheeze de tono bajo	Rale sec sonore ou ronflement
Definición de las abreviaturas: ATS = American Thoracic Society; IDW = Initial Deflection Width; 2CD = Two Cycle Duration.				

II.1.3.1. Estertores crepitantes, sonidos respiratorios adventicios discontinuos.

Los estertores crepitantes son sonidos adventicios discontinuos y explosivos que ocurren usualmente durante la inspiración durante desordenes pulmonares o problemas cardiacos. Se ha encontrado que la ocurrencia de crepitancias en los sonidos pulmonares usualmente refleja un proceso patológico en el tejido pulmonar o las vías aéreas [55]. Las crepitancias pueden ser caracterizadas con base en su forma de onda y duración, así como de acuerdo al tiempo de aparición en el ciclo respiratorio [15].

El término estertor crepitante fino se utiliza para referirse a sonidos con componentes elevados de frecuencia y corta duración y el término estertor crepitante grueso para referirse a sonidos con componentes de frecuencia más bajos y mayor duración, tal y como lo recomienda la American Thoracic Society [56]. En la tabla II.2 se muestra la clasificación de los estertores crepitantes de acuerdo a la ATS, aunque se reconoce que el procedimiento de registro y filtrado afecta las características de la forma de onda así como la duración de sus parámetros.

Por otra parte, el tiempo de aparición de los estertores crepitantes durante el ciclo respiratorio también ha sido estudiado; los estertores crepitantes pueden aparecer al inicio, en medio, o al final de las fases inspiratoria y espiratoria [55]. Se ha encontrado que el tiempo de aparición de las crepitancias durante la inspiración es de importancia clínica [57]. Los estertores crepitantes tempranos han sido asociados con severos problemas de obstrucción mientras que las crepitancias tardías han sido asociados con desordenes pulmonares restrictivos.

TABLA II.2. CLASIFICACIÓN DE LOS ESTERTORES CREPITANTES POR LA ATS.

Tipo	Intensidad	Tono	2CD	IDW
Crepitancia fina	Alta	Bajo	~ 10 ms	~ 1.5 ms
Crepitancia gruesa	Menos alta	Alto	< 5 ms	0.7 ms

Definición de las abreviaturas: ATS = American Thoracic Society; IDW = Initial Deflection Width; 2CD = Two Cycle Duration.

Debido a su corta duración y baja intensidad, la discriminación y caracterización de las crepitancias mediante la auscultación normal resulta difícil. Más aún, aunado al problema de las diferentes respuestas en frecuencia de los estetoscopios, existen diferencias en las propiedades auditivas de los médicos. La sensibilidad del oído a diferentes frecuencias depende de la intensidad relativa del sonido [58], la habilidad del oído para reconocer un tono de muy corta duración es pobre [59], el mínimo intervalo audible de separación entre dos sonidos también depende de sus componentes en frecuencia e intensidad [55]. Se ha encontrado una gran variabilidad entre médicos al describir los hallazgos de estertores crepitantes [60, 61]. En un estudio se probó la audición humana a estertores crepitantes simulados y añadidos al sonido pulmonar normal, encontrando que la falla en la detección de las crepitancias era más común cuando [62]: 1) el ruido de fondo era más intenso, 2) las crepitancias eran finas, y 3) la amplitud de las crepitancias era pequeña. Estos estudios han mostrado las grandes dificultades y variaciones en la descripción de las crepitancias.

A finales de la década de 1960 se empezó a utilizar una escala de tiempo expandida para desplegar los estertores crepitantes. Mediante la observación del intervalo de tiempo entre dos crepitancias consecutivas así como de su intensidad relativa, se mostró la naturaleza repetitiva de las crepitancias en ciclos respiratorios consecutivos [63]. Una década después, en 1977, la escala de tiempo fue expandida en mayor medida para mostrar la forma de onda de los sonidos pulmonares y traqueales normales, estertores crepitantes y sibilantes [64], originando el método de TEWA. En este estudio pionero, se mostró que las crepitancias aparecen como deflexiones de 10 ms o menos y que poseen una amplitud más grande que el

sonido pulmonar de base. Asimismo, se encontró que las crepitancias poseen una gran diversidad de patrones, y que las variaciones en el periodo de las deflexiones, i.e. la longitud en tiempo de un ciclo individual, y el número de total de ciclos parecieran ser sus características más notables. Finalmente, se encontró que las crepitancias finas mostraban periodos y duración más cortos que las crepitancias gruesas. Actualmente, el análisis con la forma de onda extendida, o TEWA, ha revelado que las crepitancias inician con una corta deflexión inicial seguida por deflexiones de mayor amplitud. Se considera que una crepitancia posee una amplitud de deflexión cuando menos dos veces mayor que la amplitud del sonido de fondo. Unos años más tarde, a inicios de la década de 1980, el método de expansión temporal para el análisis de la forma de onda fue retomado, pero esta vez con la diferencia de que cada estertor crepitante fue aislado para su análisis [65]. En este estudio se observó que las crepitancias además de caracterizarse por formas de onda transientes, poseen otras características: 1) *los intervalos de cada deflexión fueron inicialmente cortos y se hicieron progresivamente más grandes*, 2) las formas de onda estaban compuestas de un segmento inicial y segmentos con magnitud decreciente, 3) el segmento inicial presentaba características de una onda de choque y 4) la forma de onda se asemeja a una sinusoidal amortiguada con atenuación exponencial.

Basados en TEWA, se han empleado índices para clasificar a las crepitancias como finas o gruesas. Los dos primeros índices propuestos corresponden a la duración del ancho de la deflexión inicial (IDW) y a la duración de dos ciclos (2CD) determinada como el tiempo desde el inicio de IDW hasta el punto donde la crepitancia ha completado dos ciclos [66]. Un tercer índice que se ha propuesto corresponde al ancho de la deflexión más grande (LDW) definido como la duración de la deflexión con la amplitud más grande de la forma de onda [67]. Una ventaja del índice LDW sobre IDW y 2CD es que no presenta la dificultad de determinar el inicio preciso de un estertor crepitante, e incluso pareciera ser mejor parámetro que estos últimos para propósitos de clasificación. Posteriormente, se propusieron otros índices tales como la duración de la cuarta parte del primer ciclo (T_b) y la duración de los primeros $\frac{3}{4}$ (T_f) de la forma de onda de la crepitancia [68]. Incluso la polaridad de la forma de onda de las crepitancias también ha sido propuesta como un índice para su clasificación [69]. Dos nuevos índices se han sumado recientemente a la lista, ambos relacionados con la deflexión más elevada, o pico más alto: el medio periodo hacia la izquierda del pico más alto (T_1) y el medio periodo hacia la derecha del pico más alto (T_2) [70]. En este último estudio también se retomó como índice la polaridad de las crepitancias, entendiéndola como positiva

si el pico más alto está por encima de la línea de base y negativa en caso contrario. En la figura II.3 se esquematiza la forma de onda de un estertor crepitante, en una escala de tiempo extendida, para ilustrar los diversos índices propuestos para su clasificación a lo largo de los años. Finalmente, es importante notar que el filtrado, típicamente empleado para suprimir sonidos musculares de baja frecuencia, puede tener un efecto decisivo sobre la apariencia de la forma de onda de las crepitancias [71].

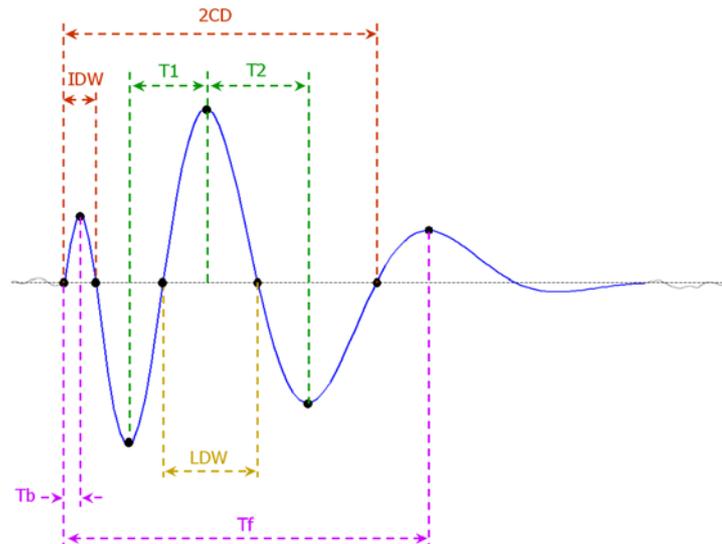


Fig. II.3. Parámetros propuestos para caracterizar la forma de onda de la crepitancia.

Antes de continuar con el análisis de las crepitancias en los dominios de la frecuencia y tiempo-frecuencia resulta interesante considerar las teorías acerca de la generación de los sonidos discontinuos. Inicialmente se creía que los estertores crepitantes eran generados por la ruptura de burbujas de aire en las vías aéreas o del movimiento de secreciones intraluminales [54], sin embargo la naturaleza repetitiva de estos sonidos en inspiraciones consecutivas contradice esta teoría. Actualmente, se considera que las crepitancias son causadas por la apertura súbita de vías aéreas anormalmente cerradas [63]. De acuerdo a esta teoría, durante la inspiración, un gradiente de presión se desarrolla a través de las vías aéreas colapsadas durante la espiración. Un sonido crepitante se genera cuando una vía aérea cerrada se abre abruptamente durante la inspiración o se cierra durante la espiración, induciendo una rápida equalización de la presión y una onda de presión. Cada una de las crepitancias podría representar una apertura o cierre abrupto de una sola vía aérea [55]. Esta teoría ha sido respaldada por el hallazgo de patrones recurrentes en las crepitancias durante ciclos respiratorios consecutivos, el cual sugiere que la secuencia repetitiva de sonidos implica que

las crepitancias se relacionan con la ocurrencia de cambios de presión y volumen dentro del pulmón, i.e. las crepitancias individuales ocurren a volúmenes inspirados y presiones transpulmonares similares durante inspiraciones sucesivas [72]. En contraste a la hipótesis de eualización explosiva de presión, se ha sugerido una hipótesis más simple para la generación de las crepitancias, la teoría *stress-relaxation quadrupoles* [73]. Esta teoría concuerda con la apertura súbita de vías aéreas cerradas, sin embargo asume que cuando una vía aérea se abre ocurren cambios bruscos en la tensión elástica. Entonces, eventos dinámicos en las paredes de las vías aéreas y no diferencias en la presión podrían crear una crepitancia, la cual podría dispersarse a partir de su sitio de generación. Este modelo predice cambios en la morfología y polaridad de la forma de onda dependiendo de la posición de observación con relación a la fuente, así como un *sucesivo incremento en los cruces por cero con respecto del tiempo*. Esta transición gradual de una respuesta más rápida a una más lenta hablaría de una dispersión de la onda. Los últimos resultados del análisis con TEWA acerca de la polaridad de la forma de onda de las crepitancias son consistentes con la teoría *stress-relaxation quadrupoles*, y muestran que las crepitancias escuchadas durante la espiración son causadas por un cierre abrupto de las vías aéreas y que un mecanismo similar pero de signo opuesto genera las crepitancias durante la inspiración, i.e. una reapertura abrupta [70]. El carácter de los sonidos crepitantes posiblemente dependa del diámetro de las vías aéreas que se abran y cierren, y esto a su vez depende de la patofisiología del tejido circundante. Se ha observado experimentalmente que las características de las crepitancias producidas en modelos de las vías aéreas varían con su tamaño; las vías aéreas más pequeñas producen crepitancias de menor duración (carácter fino) que las vías aéreas más grandes [73].

Empleando un arreglo de tres micrófonos sobre las bases pulmonares, se encontró que cada sensor sólo registraba crepitancias que ocurrían cercanos a éste [9], lo que sugiere un fenómeno local. Dado que el cierre de las pequeñas vías aéreas es un fenómeno que depende de la fuerza de gravedad, las crepitancias tienden a ocurrir primero en las áreas basales de los pulmones y cuando la enfermedad progresa también en zonas superiores [17]. El número de crepitancias por ciclo respiratorios está asociado con la severidad de la enfermedad en pacientes con desordenes pulmonares intersticiales [74]. La forma de onda y el tiempo de aparición de las crepitancias puede tener importancia clínica en el diagnóstico diferencial de enfermedades cardiorrespiratorias. Aplicaciones clínicas de los estertores crepitantes se encuentran en [55].

En el dominio de la frecuencia, desde la década de 1970 se han empleado los sistemas computacionales para analizar los estertores crepitantes mediante la FFT en pacientes con enfermedades respiratorias [75]. El empleo de la FFT mostró que la mayoría de los componentes en frecuencia de las crepitancias están limitados al intervalo de 100 a 1000 Hz [65]. El análisis espectral junto con TEWA mostró que los estertores crepitantes en diferentes patologías poseen características distintivas que pueden ser útiles en el diagnóstico médico [76]. Por ejemplo, se encontró que el contenido en frecuencia de las crepitancias inspiratorias es mayor en pacientes con fibrosis alveolitis que en aquellos con enfermedad pulmonar obstructiva crónica. La FFT también ha sido empleada para observar variaciones en las características de las crepitancias a lo largo de una patología [77], las crepitancias varían marcadamente durante el curso de la neumonía, e.g. el contenido en frecuencias se vuelve mayor durante la enfermedad. Además, la FFT también ha sido empleada para diferenciar a los estertores crepitantes [68], las frecuencias pico y máximas de las crepitancias finas son significativamente mayores que las frecuencias de las crepitancias gruesas. A pesar de estos esfuerzos con la FFT, la cual ha sido ampliamente utilizada para analizar sonidos, esta técnica no se ajusta bien a sonidos de corta duración como las crepitancias y por lo tanto, los índices más comúnmente empleados son IDW y 2CD [30].

En la década de 1950, empleando el espectrógrafo de sonido se realizaron las primeras representaciones tiempo-frecuencia de las crepitancias conjuntamente con los sonidos pulmonares normales [18], mostrando que las crepitancias son sonidos de duración limitada pero con un considerable contenido en frecuencia. Treinta años después, empleando la superposición de espectros de potencia obtenidos mediante la FFT de segmentos ventaneados con una función Hanning de 25 ms de ancho se obtuvo la representación tiempo-frecuencia de las crepitancias a lo largo de un ciclo respiratorio [78]. Se encontró que la energía de las crepitancias principalmente se encuentra en el rango de 100 a 2000 Hz; *éste es quizás el primer empleo de la STFT para el análisis de estertores crepitantes*. Basándose en la idea de estacionaridad espectral, se desarrolló un método para detectar crepitancias empleando representaciones tiempo-frecuencia [79], donde las crepitancias podrían distinguirse del ruido de fondo como picos. En particular, se empleó la STFT con ventanas rectangulares cortas para detectar crepitancias simples con una buena resolución temporal a expensas de la resolución en la frecuencia. Bajo el mismo esquema, se agregó la posibilidad de obtener un espectro basado en el modelado AR para la detección de crepitancias [80]. Otras técnicas que se han aplicado a la detección de estertores crepitantes inmersos en sonidos pulmonares en sujetos enfermos incluyen la descomposición wavelet [81] y técnicas de filtrado adaptable [82]. Con

respecto al análisis de las crepitancias en el dominio tiempo-frecuencia, en la década de los 90s se aplicaron tres distribuciones de la clase general, a decir, el espectrograma, la distribución de Wigner y la distribución Zhao-Atlas-Marks a un segmento de registro que contenía dos crepitancias [83], el estudio evidenció el compromiso en la resolución temporal y espectral del espectrograma, la presencia de términos cruzados en la distribución de Wigner y una mejoría en la resolución cuando se elige adecuadamente la función kernel. Sin embargo, la comparación entre las representaciones sólo fue cualitativa.

Recientemente se propuso la aplicación de la técnica de descomposición empírica de modos (EMD) para el análisis de los estertores crepitantes [84], sin embargo su aplicación se limitó a la descomposición de la señal en sus escalas de tiempo características sin llegar a obtener un espectro variante en el tiempo. Este estudio sugiere que la aplicación de EMD a sonidos no lineales y no estacionarios como las crepitancias resulta prometedora.

II.1.3.2. Estertores sibilantes, sonidos respiratorios adventicios continuos.

A diferencia de las crepitancias, los estertores sibilantes son sonidos adventicios continuos y de carácter más o menos musical. Acústicamente, una sibilancia se caracteriza por una forma de onda periódica con una frecuencia dominante usualmente por arriba de los 100 Hz y con una duración > 100 ms. Las sibilancias usualmente son asociadas con vías aéreas obstruidas debido a varias causas [15] y pueden ser caracterizadas por su localización, intensidad, tono, duración en el ciclo respiratorio y relación con la fase respiratoria [85].

La definición de una sibilancia ha variado con el tiempo. De acuerdo a las definiciones tempranas de la ATS [56], la palabra “continuo” implica que la duración de una sibilancia es mayor a 250 ms. La ATS además define a las sibilancias como sonidos continuos con tonos altos y califica a los sonidos continuos con tonos bajos como ronquidos. La nomenclatura de la ATS especifica que una sibilancia contiene una frecuencia dominante de 400 Hz o más, mientras que los ronquidos se caracterizan por una frecuencia dominante alrededor de 200 Hz o menos. Sin embargo, algunas investigaciones no concuerdan con estas características [86] [87]. De acuerdo a las recomendaciones de la *Computerized Respiratory Sound Analysis* (CORSA), la frecuencia dominante de una sibilancia es usualmente mayor a 100 Hz con una duración mayor a 100 ms [15]. Entonces, las sibilancias no tienen que extenderse necesariamente más de 250 ms como se propuso inicialmente, dado que típicamente duran más de 80 a 100 ms y su rango de frecuencia se extiende desde menos de 100 Hz hasta más de 1 kHz [30]. Las sibilancias son más intensas que los sonidos pulmonares de fondo y generalmente, son audibles en la boca abierta del paciente. Si la sibilancia contiene

esencialmente una sola frecuencia se denomina monofónica mientras que si se perciben múltiples frecuencias se denomina polifónica [15].

Los mecanismos fisiopatológicos que generan las sibilancias también no son del todo claros, pero parecen involucrar la interacción entre las paredes de las vías aéreas y el gas que se mueve a través de éstas. Se ha encontrado que la longitud de las vías aéreas que generan la sibilancia no muestra relación con el tono del sonido y que dicho tono no cambia al cambiar la mezcla de gas [85]. También se ha reportado que la masa y las propiedades elásticas de los bronquios son factores secundarios en la determinación del tono, mientras que la velocidad linear del flujo de aire juega un papel dominante [63]. En la misma investigación se indica que los sonidos musicales de tonos altos se producen cuando el calibre de las vías aéreas se reduce a tal grado que sus paredes opuestas están casi en contacto, dado que la aceleración del flujo de gas a través de las vías reducidas induce oscilación de sus paredes. Por lo tanto, las sibilancias resultan ser sonidos no estacionarios, donde las frecuencias altas no son necesariamente producidas por las vías aéreas periféricas y las frecuencias bajas por vías aéreas más centrales como sugirió inicialmente Laennec.

Se ha propuesto otra teoría basada en el análisis matemático de la estabilidad del flujo de aire a través de un tubo suave que puede colapsarse [88-90]. Analizando las características temporales y espectrales de los sonidos respiratorios en pacientes sanos y enfermos con sibilancias, y comparándolas con las predicciones realizadas por cinco teorías de producción de sibilancias, se encontró que la teoría *fluid dynamic flutter* se ajustaba mejor que los otros mecanismos sugeridos [86]. De acuerdo a este modelo, las sibilancias son producidas por la vibración de las paredes de las vías aéreas y el flujo de gas, donde la generación de la oscilación ocurre cuando la velocidad del gas en las vías aéreas alcanza un valor crítico determinado por las dimensiones y propiedades mecánicas de las paredes de las vías aéreas, e.g. grosor, densidad, tensión longitudinal y las propiedades mecánicas del gas, i.e. densidad y viscosidad. Dados estos parámetros, el modelo predice la velocidad de flujo crítica requerida para inducir las oscilaciones de las vías aéreas, así como la frecuencia y longitud de onda de las oscilaciones. La frecuencia de oscilación que corresponde al tono musical de la sibilancia se incrementaría con paredes más delgadas, una tensión longitudinal elevada, un canal más angosto y una elastancia elevada. Estos cambios en las características mecánicas del sistema pulmonar explicarían la presencia de sibilancias durante algunos desordenes pulmonares como el asma y enfermedades pulmonares obstructivas crónicas [85]. Además, los parámetros mecánicos pueden variar demasiado de una generación de vías aéreas a otra, siendo las vías aéreas más pequeñas menos estables. Sin embargo, las velocidades de aire a través de estas

vías se encuentran por debajo del nivel crítico, y por lo tanto las sibilancias son probablemente generadas en las primeras cinco o siete generaciones donde las velocidades de flujo son suficientemente altas [86]. Estudios sobre el efecto de la densidad del gas sobre el contenido de sibilancias espiratorias soportan la hipótesis de que el movimiento vibratorio de las vías aéreas es la fuente de energía acústica [91]. Finalmente, este modelo muestra que las sibilancias siempre están acompañadas por una limitación de flujo, pero que una limitación de flujo no está necesariamente acompañada por sibilancias [85].

El empleo de TEWA ha mostrado que una sibilancia produce deflexiones sinusoidales, u ondulaciones continuas [64], en la figura II.4 se presenta la forma de onda de una sibilancia real.

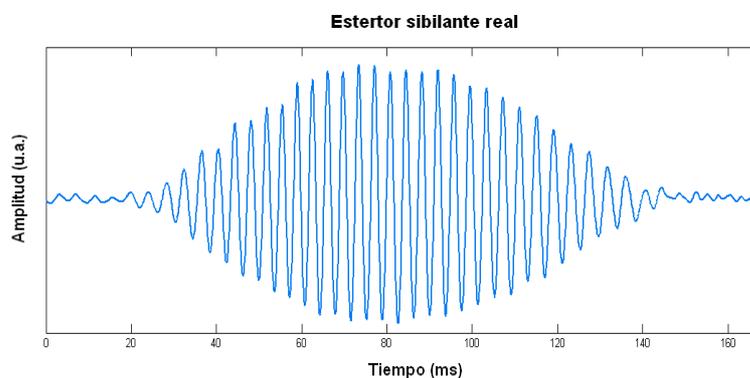


Fig. II.4. Forma de onda de una sibilancia monofónica.

La descripción en el dominio de la frecuencia de estas deflexiones sinusoidales generalmente se realiza calculando el espectro de potencia mediante la FFT. Las sibilancias generalmente producen picos bien definidos en el espectro de potencia, con un rango de frecuencias ampliamente variable de acuerdo a diversos investigadores, e.g. de 80 a 1600 Hz [86] ó de 350 a 950 Hz [87]. La transmisión de las sibilancias a través de las vías aéreas es mejor que la transmisión a través del pulmón hacia la superficie de la pared torácica. Los sonidos de mayor frecuencia son detectados con mayor claridad sobre la tráquea que sobre el pulmón [92], donde los componentes de alta frecuencia son principalmente absorbidos por el parénquima pulmonar [93]. Mediante el registro simultáneo de sonidos respiratorios sobre la tráquea y la pared torácica, se encontró que en ambas localidades aparecían picos en 480 y 650 Hz, sin embargo picos a 870 y 940 Hz observados sobre la tráquea estaban casi ausentes sobre la pared torácica [92]. Estas observaciones enfatizan el valor clínico de la auscultación traqueal en pacientes asmáticos [94].

La gran variabilidad en la frecuencia predominante de las sibilancias, así como la influencia del sitio de auscultación ilustran algunas de las dificultades encontradas en su análisis y cuantificación automática. Usualmente este análisis y cuantificación de las sibilancias se basa en su contenido espectral. Además, se han propuesto diversos criterios de amplitud para reconocer a una sibilancia, e.g. una sibilancia ha sido identificada por la presencia de un pico en el espectro de potencia con una frecuencia por encima de los 200 Hz y una amplitud mayor a quince veces la amplitud promedio del sonido de fondo en el rango de 110 a 1200 Hz [87], por la presencia de un pico significativo en el espectro cuando la frecuencia es mayor a 200 Hz y su amplitud es tres veces mayor que la señal de base [95], o como picos en el espectro por arriba de los 150 Hz y una amplitud tres veces mayor que la señal de base. Adicionalmente a la FFT, se ha empleado el modelado AR para estimar la máxima frecuencia de las sibilancias durante la respiración de pacientes asmáticos, y a partir de dicha frecuencia observar el comportamiento de las sibilancias antes y después de aplicar un broncodilatador [96].

Al igual que con las crepitancias, la primera representación tiempo-frecuencia de las sibilancias fue realizada con el espectrógrafo de sonido [18]. Las sibilancias de sujetos asmáticos aparecieron como armónicos que se extendieron a lo largo de la espiración. Posteriormente, se graficó los espectros de potencia obtenidos sucesivamente en el tiempo vía la FFT [97], mostrando que este método permite medir con más exactitud cambios, e.g. tanto del contenido en frecuencia como del tiempo de aparición y duración en el ciclo respiratorio producidos por un tratamiento médico.

Las representaciones tiempo-frecuencia de las sibilancias han sido empleadas para su análisis automático. Se han obtenido sonogramas mediante análisis armónico con base en la FFT y ventanas Hanning, así como con el modelado AR aplicado a segmentos del sonido adquirido [80]. Otra técnica empleada para generar la representación tiempo-frecuencia, aparte del clásico espectrograma [98-101], es la reportada recientemente en [102] y que utiliza la transformada wavelet continua.

Finalmente, un gran número de circunstancias son capaces de producir sonidos discontinuos adventicios. Así, las sibilancias no solamente ocurren durante el asma. Una lista detallada de las enfermedades que pueden asociarse a las sibilancias, así como la importancia clínica de estos sonidos se presenta en [85].

A manera de resumen de los tipos de sonidos respiratorios, en la Tabla II.3 se presentan las principales de las características que los distinguen y sus posibles orígenes.

TABLA II.3. CATEGORÍAS DE LOS SONIDOS RESPIRATORIOS.

Sonido respiratorio	Mecanismos	Origen	Acústica	Relevancia
Sonidos básicos				
Sonido pulmonar normal	Vórtices de flujo turbulento, otros mecanismos	Vías aéreas centrales (expiración) vías aéreas lobares a segmentales (inspiración)	Ruido filtrado con un pasa-bajas (rango <100 Hz hasta >1000 Hz)	Ventilación regional, calibre de las vías aéreas
Sonido traqueal normal	Flujo turbulento, colisión del flujo sobre las paredes de las vías aéreas	Faringe, laringe, tráquea, grandes vías aéreas	Ruido con resonancias (rango <100 Hz hasta > 3000 Hz)	Configuración de las vías aéreas superiores
Sonidos adventicios				
Sibilancia	Vibración de las paredes de las vías aéreas, difusión de los vórtices	Vías aéreas centrales y más pequeñas	Sinusoide (rango ~ 100 Hz hasta > 1000 Hz; duración típica > 80 ms).	Obstrucción de las vías aéreas, limitación de flujo
Crepitancia	Estrés-relajación de las paredes de las vías aéreas	Vías aéreas centrales y más pequeñas	Onda de deflexión fuertemente amortiguada; duración típica < 20 ms)	Cierre de las vías aéreas, secreciones

II.2. Sonidos cardiacos.

El primero en atribuir a la auscultación cardiaca un potencial diagnóstico fue Robert Hook, figura II.5, hasta quien puede rastrearse la invención del análisis de señales: *“He sido capaz de escuchar de manera clara los latidos del corazón de un hombre... ¿Quién sabe?, digo, pero podría ser posible descubrir los movimientos de las partes internas de los cuerpos... por el sonido que estos producen, esto podría descubrir los trabajos realizados en las diversas oficinas y tiendas del cuerpo humano, y consecuentemente descubrir que instrumentos o maquinaria no está funcionando.”* Con el advenimiento del estetoscopio de Laennec en 1816, la auscultación cardiaca se convirtió en una herramienta clínica fundamental que prevalece hasta nuestros días [103].



Fig II.5. Robert Hooke.

La auscultación es una de las más antiguas y útiles técnicas para el diagnóstico de enfermedades cardiacas. Actualmente, el valor de la auscultación cardiaca para el diagnóstico ha sido confirmado por un gran número de complejas técnicas como la angiografía y la ecocardiografía Doppler, bidimensional y en modo M [104]. Los eventos que toman lugar en el corazón producen variaciones de presión acústicas las cuales caen dentro del espectro clasificado como sonido. El registro gráfico de los sonidos cardiacos es denominado fonocardiografía, y la señal resultante fonocardiograma (PCG) [105].

El primer fonocardiografo fue desarrollado por Einthoven en 1894 con el propósito de proveer al médico de una herramienta complementaria para registrar los sonidos cardiacos escuchados durante la auscultación, para obtener mediciones específicas de sus parámetros temporales, y ganar un mejor entendimiento de los mecanismos básicos responsables de su génesis [104]. A pesar de que la fonocardiografía es una técnica no invasiva y de gran potencial para la detección de varias enfermedades cardiacas, e.g. enfermedades valvulares, hipertensión pulmonar, disfunción ventricular y cardiomiopatías, su evolución ha sido lenta en comparación con otras técnicas debido, entre otros aspectos, a que [103, 104]: 1) el PCG es una señal compleja difícil de analizar visualmente, 2) aún existe controversia respecto a la génesis de los sonidos cardiacos y 3) el advenimiento de la imagenología por ultrasonido ha hecho de la fonocardiografía y la auscultación cardiaca herramientas clínicas secundarias. Sin embargo, se han realizado desarrollos importantes en muchos aspectos de la fonocardiografía debido a la creciente aplicación de técnicas de procesamiento digital de señales, tales como el análisis espectral, reconocimiento de patrones y el análisis tiempo-frecuencia. Tales desarrollos han conducido a lo que actualmente se conoce como fonocardiografía digital [104].

II.2.1. Primer sonido cardiaco, S1.

Los sonidos cardiacos resultan de los eventos dinámicos asociados con la contracción y relajación de la aurícula y los ventrículos, los movimientos valvulares y el flujo sanguíneo. El primer sonido cardiaco (S1) ocurre con la contracción ventricular durante el periodo de contracción isovolumétrica, cuando la presión del ventrículo izquierdo crece rápidamente, como se muestra en la figura II.6, indicando el inicio de la sístole ventricular. Cuatro eventos valvulares determinan a S1: el cierre de la válvula mitral, el cierre de la válvula tricúspide, la apertura de la válvula pulmonar y la apertura de la válvula aórtica, generalmente en esa secuencia [103, 104, 106]. La duración total de S1 es en promedio de 100 a 120 ms en un adulto normal [107].

Aunque las relaciones temporales entre los eventos eléctricos, mecánicos y acústicos en el corazón están bien establecidas, los mecanismos involucrados en la génesis de los sonidos cardiacos todavía están sujetos a controversia [104, 106, 108]. De acuerdo a la teoría cardiorhémica, se asume que las cavidades cardiacas, las válvulas y la sangre representan un sistema interdependiente que vibra como un ente [104]. Para este modelo, el cierre de las válvulas mitral y tricúspide inicia vibraciones de forma indirecta al evitar el contraflujo de los ventrículos a las aurículas, con la subsecuente excitación de todas las estructuras cardiacas [106]. En esta teoría, S1 se inicia con un componente de baja frecuencia sincronizado con la primera contracción miocárdica, después del incremento en la presión ventricular. El segundo componente, de mayor frecuencia, es causado por la tensión de las estructuras ventriculares izquierdas, la contracción del miocardio, y la desaceleración de la sangre. El tercer componente, que ocurre en el instante de la apertura de la válvula aórtica, está relacionado con la aceleración súbita de la sangre dentro de la aorta e involucra la oscilación de la sangre entre la aorta y las paredes ventriculares. El cuarto componente es debido a la turbulencia en el flujo sanguíneo a través de la aorta ascendente. Finalmente, el corazón derecho no contribuye significativamente en la génesis de S1 [104]. A pesar de que la teoría cardiorhémica es generalmente aceptada, aún existe controversia respecto a la importancia relativa de las válvulas y el miocardio en la determinación de las propiedades en la frecuencia de S1 [103]. Se ha argumentado por un modelo de resonancia valvular, que la frecuencia de S1 está determinada principalmente por la rigidez y el área de la válvula mitral y por la masa sanguínea del ventrículo izquierdo [109-111]. Alternativamente, se ha propuesto que el cierre de la válvula mitral inicia vibraciones resonantes a través del miocardio, con una frecuencia natural determinada por la tensión, rigidez y geometría ventricular [112]. Ambos modelos, el valvular y el miocárdico predicen que las frecuencias de S1 son proporcionales a la presión

ventricular izquierda y dado que la presión ventricular izquierda se incrementa rápidamente a lo largo de la generación de S1, ambos modelos predicen una frecuencia que se incrementa con el tiempo [103].

A finales de la década de los 60s e inicios de los 70s se sugirió que podría obtenerse información significativa del análisis en la frecuencia de los sonidos cardiacos [112]. A partir de entonces, el análisis espectral de los sonidos cardiacos ha sido empleado con diferentes propósitos. A mediados de la década de los 70s, el análisis espectral de S1 registrado sobre el área apical de sujetos sanos mostró que su espectro de frecuencia contiene un pico en un rango de baja frecuencia entre 10 y 50 Hz, y otro pico entre 50 y 140 Hz [113]. El análisis espectral también ha sido empleado para estudiar S1 en pacientes con isquemia miocárdica [114], para la detección automática del primer y segundo ruidos cardiacos [115], para valorar la rigidez de válvulas mitrales bioprostéticas [116], para comparar métodos de reconocimiento de patrones en la clasificación de sonidos cardiacos en pacientes con válvulas bioprostéticas [117], para la detección de oclusiones coronarias [118], y para la comparación del espectro del primer y segundo ruidos cardiacos antes y después de la implantación de una válvula mecánica [119]. Entre las diversas técnicas espectrales empleadas se incluyen la FFT y el espectro de potencia obtenido mediante modelado autorregresivo (AR), mediante el modelado autorregresivo y de promedio móvil (ARMA), y mediante el modelado con sinusoidales amortiguadas basado en la aproximación de Prony.

Como ya se mencionó, tanto el modelo valvular como el miocárdico predicen un incremento monótono en la frecuencia de resonancia debido a un incremento en la presión ventricular izquierda. Adicionalmente a esta dinámica en la frecuencia, la apertura y cierre de las válvulas cardiacas pueden producir perturbaciones súbitas en las vibraciones cardiacas. Este comportamiento no estacionario, tiene como consecuencia que las técnicas de análisis en la frecuencia clásicas que asumen estacionariedad de la señal, tales como la transformada de Fourier o la estimación espectral paramétrica, resulten inadecuadas debido a que consideran que la frecuencia de señal permanece constante durante el intervalo de la transformación. Consecuentemente, se requieren técnicas dinámicas de análisis en la frecuencia para relacionar las estructuras cardiacas con las vibraciones que emiten, y con ello revelar los complejos mecanismos de la generación de S1 [103, 120]. La no estacionariedad de los sonidos cardiacos ha sido reconocida desde al menos la mitad del siglo XX [103], y resulta importante el aplicar técnicas de procesamiento más adecuadas como las transformaciones tiempo-frecuencia [120]. A mediados de la década de los 80s, se propuso una teoría que integra las observaciones de la teoría valvular y la cardiorhemica, bajo la hipótesis de que el

corazón es un sistema viscoelástico con grados de libertad múltiples en el cual cada estructura vibrante, además de la turbulencia del flujo sanguíneo, puede contribuir al PCG, el cual es una señal multicomponente. En esta nueva teoría, los eventos productores de sonido que dan lugar al PCG (las variaciones súbitas de presiones intracardiacas) resultan de fluctuaciones en la tensión de las estructuras del corazón (componentes valvulares, cuerdas tendinosas y miocardio) y de la turbulencia en el flujo sanguíneo. Dichas variaciones de presión son producidas durante el cierre y apertura de las cuatro válvulas cardiacas, la contracción y relajación del miocardio, y la resultante aceleración y desaceleración de la masa sanguínea [104, 121].

Los primeros análisis tiempo-frecuencia fueron realizados en la década de los 50s empleando un dispositivo denominado espectrógrafo de sonido [122-124]. El espectrógrafo fue empleado para el diagnóstico no invasivo del mal funcionamiento de válvulas prostéticas mediante el uso de la frecuencia pico instantánea [125]. A partir de la posibilidad de obtener una mayor resolución mediante el empleo de la FFT que mediante filtros pasabanda analógicos [113], los métodos de tiempo corto tales como el espectrograma y el modelado espectral autorregresivo empezaron a emplearse con mayor frecuencia en el análisis de S1. El espectrograma fue empleado para caracterizar la “firma” espectral de válvulas aórticas prostéticas [126] y por su parte el análisis paramétrico fue empleado para extraer el patrón espectral de señales fonocardiográficas [127].

Un estudio enfocado en definir la duración óptima de la ventana de tiempo del espectrograma para el análisis de S1 mostró que el rango de duración se encuentra entre 16 y 32 ms, al producir un suavizado mínimo en el tiempo y la frecuencia [128]. Estos datos sugieren que se requieren ventanas de tiempo más cortas que las empleadas en estudios previos cuyas duraciones igualan al periodo de contracción isovolumétrica (60-90 ms).

A inicios de la década de los 90s se aplicó la transformada wavelet al fonocardiograma [129]. En la investigación realizada por J. Bulgrin et al., al comparar el desempeño de la STFT, la transformada wavelet y una técnica basada en el filtrado estadístico se demostró que la STFT y la transformada wavelet proporcionan una resolución tiempo-frecuencia comparable al analizar los eventos acústicos cardiacos, mientras que la técnica restante caracterizó en menos medida la naturaleza multicomponente de los sonidos cardiacos [130]. Además, se encontró una alta variabilidad inter-paciente en las representaciones tiempo-frecuencia de los PCG intracardiacos y una baja variabilidad latido a latido intra-sujeto. La similitud entre el espectrograma y la transformada wavelet no resulta sorprendente si se considera que el espectro de los sonidos cardiacos posee un ancho de banda relativamente

angosto y por consiguiente el filtro o wavelet de análisis varía poco y produce un espectro tiempo-frecuencia comparable con un espectrograma con una ventana fija [103]. El estudio comparativo entre diversas técnicas tiempo-frecuencia sugirió que una mejoría en la localización de eventos acústicos durante el ciclo cardiaco puede resultar útil en el desarrollo de dispositivos de auscultación automáticos [130].

A mediados de la década de los 90s se realizaron estudios para comparar diversas representaciones tiempo-frecuencia con el propósito de seleccionar la técnica más adecuada para el análisis de S1. Mediante la construcción de una señal S1 sintética consistente de dos sinusoidales amortiguadas, o componentes valvulares, y una señal con una modulación en la frecuencia de tipo creciente, o componente miocárdico, se probaron cinco representaciones tiempo-frecuencia [108, 131, 132]. Las distribuciones evaluadas fueron el espectrograma, el modelado autorregresivo aplicado a segmentos de la señal, una distribución con interferencia reducida, la distribución Zhao-Atlas-Marks [133] y la distribución Bessel [134]. Los resultados mostraron que el espectrograma produce la resolución tiempo-frecuencia más baja y que para la señal simulada, ninguna de las técnicas caracterizó adecuadamente el componente muscular. Esto ilustra que diferentes transformaciones favorecen diferentes aspectos de una señal variante en el tiempo y por lo tanto, la técnica óptima dependerá de la característica de interés de la señal [106]. Otro estudio comparó tres distribuciones tiempo-frecuencia pertenecientes a la clase de Cohen: una distribución con interferencia reducida, la distribución Zhao-Atlas-Marks y la distribución positiva [135]. Los resultados sugieren que tales distribuciones pueden resultar útiles en la caracterización de la naturaleza de los sonidos cardiacos, pero que la elección de una distribución es una decisión que debiera ser validada rigurosamente, en particular cuando la señal de interés esta pobremente comprendida.

II.2.2. Segundo sonido cardiaco, S2.

Mientras que S1 indica el inicio de la sístole ventricular, el segundo sonido cardiaco S2, marca el fin de la sístole ventricular y el inicio de la relajación ventricular, seguida por el cierre de las válvulas sigmoideas, i.e. las válvulas aórtica y pulmonar [104]. Los sonidos cardiacos se encuentran separados aproximadamente por 300 a 400 ms [107], en la figura II.3 se ilustran los eventos anteriores. Los mecanismos asociados con la génesis de los componentes de S2 son más ampliamente aceptados [136]. Mediante la auscultación, los componentes de S2 se han identificado fácilmente, por lo que la ecocardiografía sólo ha añadido poco a su entendimiento excepto por su habilidad para identificar el cierre pulmonar en algunos casos extremadamente difíciles [137].

De acuerdo a la teoría valvular, S2 está compuesto de dos ondas transientes de alta frecuencia, A2 y P2, relacionadas con el cierre de las válvulas aórtica y pulmonar, respectivamente [104, 138-140]. El componente aórtico, A2, generalmente es producido durante el cierre y vibración de la válvula aórtica y su tejido circundante. Además, A2 es seguido por el componente pulmonar, P2, que a su vez resulta del cierre y vibración de la válvula pulmonar y tejidos que la rodean [136]. Por su parte, la teoría cardiorresonante señala que S2 es producido por la desaceleración y el regreso de flujo hacia la aorta y arteria pulmonar, por lo que los dos componentes transientes ocurren principalmente dentro de la aorta y la arteria pulmonar así como en los tractos ventriculares izquierdo y derecho [104, 136]. Usualmente, la duración de cada componente es menor a 50 ms [107, 142]. Normalmente, A2 está compuesto de componentes de alta frecuencia y posee una amplitud mayor que P2 que se caracteriza por componentes de menor frecuencia [141].

Diferentes factores influyen en el tiempo de retardo entre A2 y P2; la respiración del sujeto es la más común [141]. Durante la espiración, los dos componentes normalmente están separados por un intervalo menor a 30 ms, o incluso menor a 15 ms. Durante la inspiración, el intervalo de separación se amplía, principalmente debido al retardo de P2; el tiempo de retardo puede estar entre 30 y 80 ms [142]. Finalmente, la representación temporal de S2 cuando A2 y P2 se encuentran cercanos entre sí puede variar significativamente dependiendo de la diferencia entre sus fases instantáneas en el momento del inicio de P2 [143].

A mediados de la década de los 70s, se empleó la FFT para analizar S2 [144]. En el estudio se analizaron sonidos de S2 registrados sobre las áreas aórtica y pulmonar, encontrando que los sujetos presentaban al menos dos picos en el rango de frecuencias de 80-400 Hz. Los resultados anteriores sugirieron que cuando menos ocurren dos eventos mecánicos resonantes en la producción del segundo sonido cardíaco. Posteriormente, se emplearon técnicas de análisis espectral, como el método de máxima entropía, que permiten una mayor resolución en la frecuencia que la obtenida por los métodos paramétricos [145].

En [146, 147] se demostró que la frecuencia de resonancia de P2 es proporcional a la presión arterial pulmonar y consecuentemente, la frecuencia instantánea de P2 puede asumirse proporcional a dicha presión. Una hipótesis similar aplica entre la frecuencia instantánea de A2 y la presión aórtica [136]. Empleando el espectrograma y técnicas de reconocimiento de patrones se mostró que S2 puede emplearse para estimar la presión aórtica con un error comparable al obtenido con esfigmomanómetros automáticos [148]. La estimación de la presión sanguínea mediante el análisis espectral de S2 se basa en la ley de Laplace, la cual asume que la tensión de la pared aórtica, o pulmonar, es proporcional a la presión sanguínea

dentro de la arteria correspondiente. Como consecuencia, se espera que la frecuencia de resonancia de la columna de sangre dentro de la aorta, o la arteria pulmonar, sea proporcional a la tensión en la pared arterial y por lo tanto, a la presión arterial [104].

Se han empleado diversas técnicas de modelado matemático para el análisis y síntesis de los dos primeros sonidos cardíacos. Un primer modelo consiste en emplear sinusoidales amortiguadas el cual ofrece las ventajas de parametrizar por completo a la señal y brindar un significado físico claro. Este modelo asume que dado que los sonidos producidos por el cierre y vibración de las válvulas cardíacas están compuestas de señales transientes de corta duración, se pueden modelarse mediante la suma finita de sinusoidales amortiguadas con diferentes frecuencias y constantes de decaimiento [141, 149-153]. Este modelo tiene las desventajas de: (1) no permitir la separación de A2 y P2 dentro de S2, (2) no representar de manera correcta componentes chirp transientes y (3) para caracterizar a S2 se requiere un gran número de componentes. Un segundo modelo consiste en emplear el método de *matching pursuit* con el cual los sonidos cardíacos son modelados mediante la suma de sinusoidales con una envolvente, por lo general de tipo Gaussiana que origina el llamado diccionario Gabor, que se encuentran localizados con gran resolución en el plano tiempo-frecuencia [154]. El método también requiere de un gran número de componentes. Una tercera aproximación consiste en modelar S1 o S2 mediante una señal con frecuencia modulada (chirp) lineal de amplitud variable [155], sin embargo el modelo tiene la desventaja de no permitir una correcta representación de los componentes individuales. Una cuarta aproximación consiste en el empleo de las representaciones tiempo-frecuencia para modelar S2. Argumentando que la frecuencia instantánea de S2, así como de sus componentes A2 y P2, generalmente no posee una relación lineal con el tiempo, se propuso que las representaciones de A2 y P2 deben estar compuestas de señales chirp transientes de corta duración que poseen una frecuencia instantánea proporcional a la presión decreciente aórtica o pulmonar, respectivamente [136]. Empleando la distribución de Wigner-Ville, bajo el argumento de que ofrece una buena resolución tiempo-frecuencia, se mostró que es posible modelar cada componente mediante una señal chirp no lineal de corta duración cuya distribución de energía se concentra alrededor de su frecuencia instantánea, que presenta un rápido decremento con respecto del tiempo. Posteriormente, el modelo fue empleado para extraer y aislar ambos componentes de S2 mediante la estimación y reconstrucción de sus funciones de amplitud y fase instantáneas [143]. En este estudio se encontró que el error entre la señal reconstruida y la original es proporcional al intervalo de traslape entre ambos componentes.

En forma similar al análisis de S1, se han comparado diversas representaciones tiempo-frecuencia para el análisis de S2. Entre las distribuciones empleadas se encuentran el espectrograma, la distribución de Wigner y la transformada wavelet [156, 157]. En estos estudios se encontró que el espectrograma no detecta adecuadamente los dos componentes del segundo sonido cardiaco, la distribución de Wigner puede proporcionar características tiempo-frecuencia del PCG, mientras que la transformada wavelet es capaz de detectar A2 y P2 de un PCG normal. Entonces, la transformada wavelet resulta de ayuda para obtener medidas cualitativas y cuantitativas de las características tiempo-frecuencia de S2. Una aplicación de la transformada wavelet al análisis de S2 consistió en la medición del retardo entre ambos componentes, que es importante en el diagnóstico de diversas patologías como estenosis mitral o pulmonar [158]. Recientemente surgió una técnica de análisis no lineal y no estacionaria denominada descomposición empírica de modos (EMD). Esta técnica de análisis tiempo-frecuencia ha sido aplicada a los dos primeros sonidos cardiacos tanto en escenarios simulados, mediante modelos matemáticos para S1 y S2, como en sonidos cardiacos adquiridos sobre la superficie del tórax [159]. Al ser una descomposición adaptable que extrae la información directamente de los sonidos cardiacos sin el empleo de funciones kernel o wavelets prototipo, su aplicación resulta prometedora. Las señales de sonido cardiaco representan señales multicomponentes desafiantes, al ser señales no estacionarias que provienen de sistemas biológicos complejos. El problema de procesamiento de señales dominante consiste en la descomposición de los sonidos en sus componentes individuales y su subsiguiente clasificación como normal o patológico. Por lo tanto, la aplicación de técnicas recientes de procesamiento de señales no estacionarias podría reflejarse en un mejor entendimiento de la generación de los sonidos cardiacos y desempeño clínico [106].

En el presente capítulo se ha descrito la posible génesis de los sonidos respiratorios, pulmonares y adventicios, y de los sonidos cardiacos. Además, se ha expuesto la forma de análisis que se ha realizado en un gran número de investigaciones acerca de los sonidos respiratorios y cardiacos. Finalmente, es de esperarse que nuevas formas de análisis, como las representaciones tiempo-frecuencia, motivo de la presente tesis, abran nuevas avenidas para las investigaciones de mecanismos subyacentes en algún fenómeno fisiológico y en particular, en el análisis de los sonidos respiratorios.

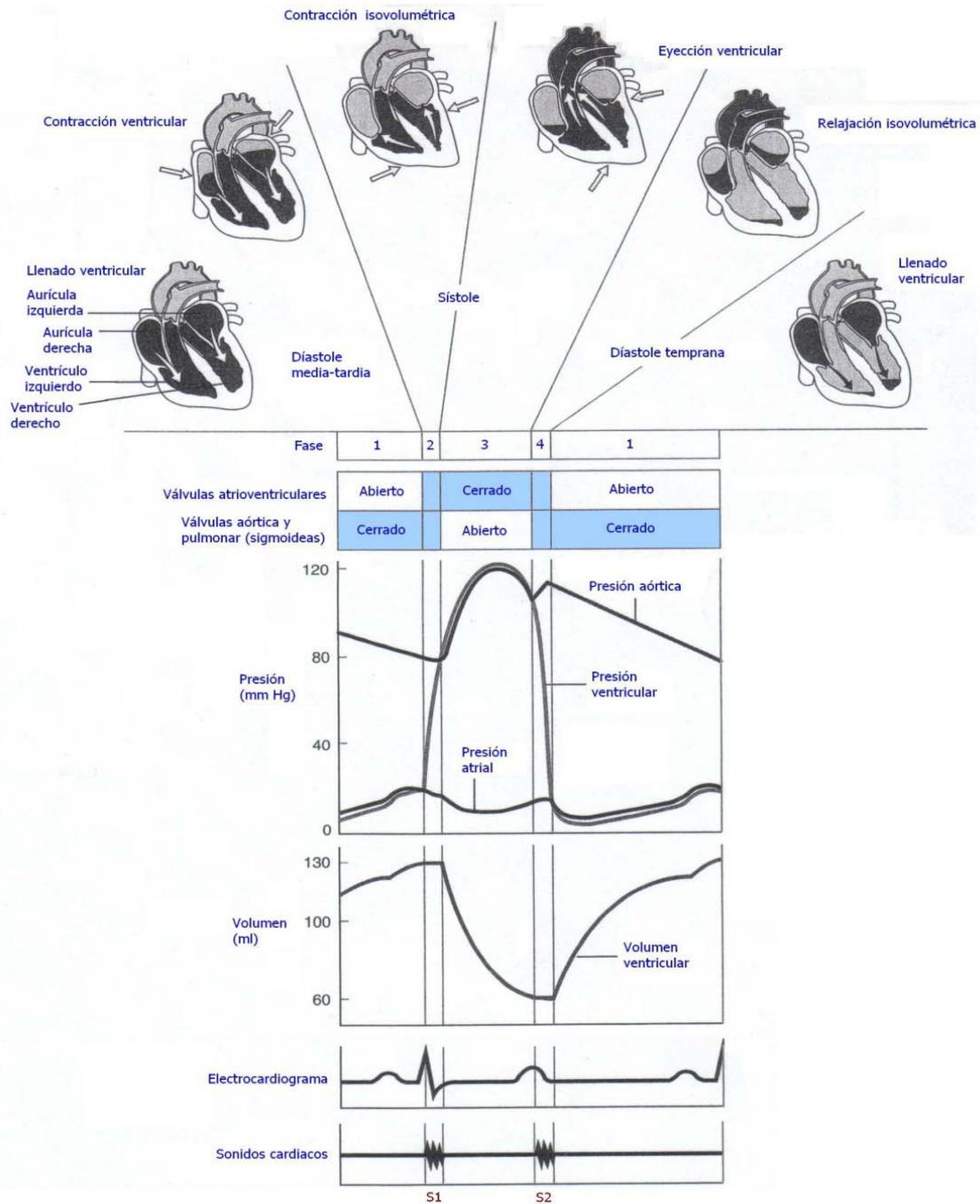


Fig. II.6. Diagrama Wiggers del ciclo cardíaco. La diástole corresponde a las fases 1 y 4, mientras que la sístole a las fases 2 y 3. Existe correlación entre los gradientes de presión a través de las válvulas, el cierre y apertura de éstas, así como el cambio en el volumen ventricular. La señal de ECG se correlaciona con los eventos mecánicos del corazón, la onda P precede la contracción auricular, el complejo QRS precede la contracción ventricular y la onda T precede la relajación ventricular. Los sonidos cardíacos se correlacionan con el cierre de las válvulas cardíacas. La figura fue tomada y modificada de [16].

Referencias.

- [1] R. Laennec, *Treatise' is De l'Auscultation Médiante ou Traité du Diagnostic des Maladies des Poumons et du Coeur*, Brosson & Chaudé, Paris, 1819. Traducción al inglés por J. Forbes, Underwood, London, 1821.
- [2] F. Dalmay, M. Antonini, P. Marquet, R. Menier, "Acoustic properties of the normal chest," *Eur. Respir. J.*, vol. 8, pp. 1761-1769, 1995.
- [3] R. Loudon, "The Lung exam," *Clinics in Chest Medicine*, vol. 8, no. 2, pp. 265-272, 1987.
- [4] N. Comins, "New Stethoscope," Extracto de *The London Medical Gazette*, 12 de Agosto de 1829.
- [5] B. Krishna, P. Scott, E. Howard, P. Boggs, "Evaluation of lungs by telephone," *Annals of Allergy*, vol. 54, pp. 109-111, 1985.
- [6] A. Sovijärvi, J. Vanderschoot, J. Earis, "Standardization of computerized respiratory sound analysis," *Eur. Respir. Rev.*, vol. 10, no. 77, pp. 585-590, 2000.
- [7] M. Abella, J. Formolo, D. Penny, "Comparison of acoustic properties of six popular stethoscopes," *J. Acoust. Soc. Am.*, vol. 91, pp. 2224-2228, 1992.
- [8] J. Earis, B. Cheetham, "Current methods used for computerized respiratory sound analysis," *Eur. Respir. Rev.*, vol. 10, no. 77, pp. 586-590, 2000.
- [9] G. Benedetto, F. Dalmaso, R. Spagnolo, "Surface distribution of crackling sounds," *IEEE Trans. Biomed. Eng.*, vol. 35, pp. 406-412, 1988.
- [10] H. Pasterkamp, R. Consunji, Y. Oh, J. Holbrow, "Chest surface mapping of lung sounds during methacholine challenge," *Pediatr. Pulmonol.*, vol. 23, pp. 21-30, 1997.
- [11] T. Bergstresser, D. Ofengeim, A. Vyshedskiy, J. Shane, R. Murphy, "Sound transmission in the lung as a function of lung volume," *J. Appl. Physiol.*, vol. 93, pp. 667-674, 2002.
- [12] M. Nissan, N. Gavriely, "A computer based lung sound analysis," *Comput. Methods Programs Biomed.*, vol. 40, pp. 7-13, 1993.
- [13] S. Charleston-Villalobos, S. Cortes-Rubiano, R. Gonzalez-Camarena, G. Chi-Lem, T. Aljama-Corrales, "Respiratory acoustic thoracic imaging (RATHI): assessing deterministic interpolation techniques," *Med. Biol. Eng. Comput.*, vol. 42, no. 5, pp. 618-626, 2004.
- [14] J. Robertson, R. Coope, "Rales, rhonchi, and Laennec," *The Lancet*, vol. 31, pp. 417-423, 1957.
- [15] A. Sovijärvi, F. Dalmaso, J. Vaderschoot, L. Malmberg, G. Righini, S. Stoneman, "Definition of terms for applications of respiratory sounds," *Eur. Respir. Rev.*, vol. 10, no. 77, pp. 597-610, 2000.
- [16] T. Pedley, R. Schroter, M. Sudlow, "The prediction of pressure drop and variation of resistance within the human bronchial airways," *Resp. Physiol.*, vol. 9, pp. 387, 1970.
- [17] A. Sovijärvi, L. Malmberg, G. Charbonneau, J. Vanderschoot, F. Dalmaso, C. Sacco, M. Rossi, J. Earis, "Characteristics of breath sounds and adventitious respiratory sounds," *Eur. Respir. Rev.*, vol. 10, no. 77, pp. 591-596, 2000.
- [18] V. McKusick, J. Jenkins, G. Webb, "The acoustic basis of the chest examination: studies by means of sound spectrography," *Am. Rev. Tuberc.*, vol. 72, pp. 12-34, 1955.
- [19] R. Loudon, R. Murphy, "Lung sounds," *Am. Rev. Respir. Dis.*, vol. 130, pp. 663-673, 1984.
- [20] N. Gavriely, Y. Palti, G. Alroy, "Spectral characteristics of normal breath sounds," *J. Appl. Physiol.*, vol. 50, pp. 307-314, 1981.
- [21] R. Hallgren, S. Huang, S. McMahon, T. Shockey, "Breath sounds: development of a system for measurement and analysis," *J. Clin. Eng.*, vol. 7, pp. 135-141, 1982.
- [22] S. Chowdhury, A. Majumder, "Digital spectrum analysis of respiratory sound," *IEEE Trans. Bomed. Eng.*, vol. 28, pp. 784-788, 1981.
- [23] J. Schreiber, W. Anderson, M. Wegman, W. Waring, "Frequency analysis of breath sounds by phonopneumography," *Med. Instrum.*, vol. 15, pp. 331-334, 1981.
- [24] E. Scarpelli, *Examination of the lung. Pulmonary disease of the fetus, newborn and child*, Lea & Febiger, Philadelphia, 1978.
- [25] H. Pasterkamp, R. Fenton, F. Leahy, V. Chernick, "Spectral analysis of breath sounds in normal newborn infants," *Med. Instrum.*, vol. 17, pp. 355-357, 1983.
- [26] J. Kanga, S. Kraman, "Comparison of the lung sound frequency spectra of infants and adults," *Pediatr. Pulmonol.*, vol. 2, pp. 292-295, 1986.
- [27] S. Kraman, "Does vesicular lung sound come only from the lungs?," *Am. Rev. Respir. Dis.*, vol. 128, pp. 622-626, 1983.

-
- [28] N. Gavriely, "Generation and transmission of lung sounds. A study in dried, inflated, dog lung model," *Am. Rev. Respir. Dis.*, vol. Suppl. 127, pp. 258, 1983.
- [29] S. Kraman, P. Wang, "Airflow-generated sound in a hollow canine airway cast," *Chest*, vol. 97, pp. 461-466, 1990.
- [30] H. Pasterkamp, S. Kraman, G. Wodicka, "Respiratory sounds. Advances beyond the stethoscope," *Am. J. Respir. Crit. Care Med.*, vol. 156, pp. 974-987, 1997.
- [31] P. Forgacs, "The functional basis of pulmonary sounds," *Chest*, vol. 73, no. 3, pp. 399-405, 1978.
- [32] S. Kraman, "Does laryngeal noise contribute to the vesicular lung sound?," *Am. Rev. Respir. Dis.*, vol. 124, pp. 92-94, 1981.
- [33] A. Bohadana, J. Kanga, S. Kraman, "Does airway closure affect lung sound generation?," *Clin. Physiol.*, vol. 8, pp. 341-349, 1988.
- [34] S. Kraman, "The relationship between airflow and lung sound amplitude in normal subjects," *Chest*, vol. 86, pp. 225-229, 1984.
- [35] S. Kraman, "Determination of the site of production of respiratory sounds by subtraction phonopneumography," *Am. Rev. Respir. Dis.*, vol. 122, pp. 303-309, 1980.
- [36] S. Kraman, O. Austrheim, "Comparison of lung sound and transmitted sound amplitude in normal men," *Am. Rev. Respir. Dis.*, vol. 128, pp. 451-454, 1983.
- [37] B. Shyoff, Y. Ploysongsang, H. Chang, "Airflow and normal lung sounds," *Am. Rev. Respir. Dis.*, vol. 137, pp. 872-876, 1988.
- [38] J. Nairn, M. Turner-Warwick, "Breath sounds in emphysema," *Br. J. Dis. Chest*, vol. 63, pp. 28-37, 1969.
- [39] P. LeBlanc, P. Macklem, W. Ross, "Breath sounds and distribution of pulmonary ventilation," *Am. Rev. Respir. Dis.*, vol. 102, pp. 10-16, 1970.
- [40] Y. Ploysongsang, R. Martin, W. Ross, R. Loudon, P. Macklem, "Breath sounds and regional ventilation," *Am. Rev. Respir. Dis.*, vol. 116, pp. 187-199, 1977.
- [41] Y. Ploysongsang, P. Macklem, W. Ross, "Distribution of regional ventilation measured by breath sounds," *Am. Rev. Respir. Dis.*, vol. 117, pp. 657-664, 1978.
- [42] D. O'Donnell, S. Kraman, "Vesicular lung sound amplitude mapping by automated flow-gated phonopneumography," *J. Appl. Physiol.*, vol. 53, pp. 603-609, 1982.
- [43] R. Dosani, S. Kraman, "Lung sound intensity variability in normal men. A contour phonopneumographic study," *Chest*, vol. 83, pp. 628-631, 1983.
- [44] A. Cohen, D. Landsberg, "Analysis and automatic classification of breath sounds," *IEEE Trans. Biomed. Eng.*, vol. BME-31, pp. 585-590, 1984.
- [45] V. Iyer, P. Ramamoorthy, Y. Ploysongsang, "Autoregressive modeling of lung sounds: characterization of source and transmission," *IEEE Trans. Biomed. Eng.*, vol. 36, pp. 1133-1137, 1989.
- [46] N. Gabriely, M. Herzberg, "Parametric representation of normal breath sounds," *J. Appl. Physiol.*, vol. 73, no. 5, pp. 1776-1784, 1992.
- [47] J. Vanderschoot, N. Kappeyne, H. Schreur, "AR model orders of lung sounds," *Proc. Annual Int. Conf. IEEE-EMBS*, pp. 2531-2532, 1992.
- [48] J. Vanderschoot, H. Schreur, "Flow and volume related AR-modelling of lung sounds," *Proc. Annual Int. Conf. IEEE-EMBS*, pp. 385-386, 1991.
- [49] F. Wooten, W. Waring, M. Wegmann, W. Anderson, J. Conley, "Method for respiratory sound analysis," *Med. Instrum.*, vol. 12, pp. 254-257, 1978.
- [50] N. Gavriely, D. Cugell, "Airflow effects on amplitude and spectral content of normal breath sounds," *J. Appl. Physiol.*, vol. 80, pp. 5-13, 1996.
- [51] I. Sanchez, H. Pasterkamp, "Tracheal sound spectra depend on body height," *Am. Rev. Respir. Dis.*, vol. 148, pp. 1083-1087, 1993.
- [52] P. Forgacs, "Breath sounds," *Thorax*, vol. 33, pp. 681-683, 1978.
- [53] C. Lessard, W. Wong, "Correlation of constant flow rate with frequency spectrum of respiratory sounds when measured at the trachea," *IEEE Trans. Biomed. Eng.*, vol. 33, pp. 461-463, 1986.
- [54] R. Murphy, S. Holford, "Lung sounds," *ATS News*, vol. 8, no. 4, pp. 24-29, 1980.
- [55] P. Piirilä, A. Sovijärvi, "Crackles: recording, analysis and clinical significance," *Eur. Respir. J.*, vol. 8, pp. 2139-2148, 1995.
- [56] American Thoracic Society, "Updated nomenclature for membership reaction. Reports of the ATS-ACCP ad hoc committee," *ATS news*, vol. 3, pp. 5-6, 1977.
- [57] A. Nath, L. Capel, "Inspiratory crackles- early and late," *Thorax*, vol. 29, pp. 223-227, 1974.
-

- [58] M. Rappaport, H. Sprague, "Physiologic and physical laws that govern auscultation, and their clinical application," *Am. Heart J.*, vol. 21, pp. 257-318, 1941.
- [59] G. Stewart, "Problems suggested by and uncertainty principle in acoustics," *J. Acoust. Soc. Am.*, vol. 3, pp. 325-329, 1931.
- [60] H. Smyllie, L. Blendis, P. Armitage, "Observer disagreement in physical signs of the respiratory system," *Lancet*, vol. 2, pp. 412-413, 1965.
- [61] L. Hudson, R. Conn, R. Matsubara, A. Pribble, "Rales: diagnostic uselessness of qualitative adjectives," *Am. Rev. Respir. Dis.*, vol. 113, pp. 187, 1976.
- [62] H. Kiyokawa, M. Greenberg, K. Shirota, H. Pasterkamp, "Auditory detection of simulated crackles in breath sounds," *Chest*, vol. 119, no. 6, pp. 1887-1892, 2001.
- [63] P. Forgacs, "Crackles and wheezes," *Lancet*, pp. 203-205, July 1967.
- [64] R. Murphy, S. Holford, W. Knowler, "Visual lung-sound characterization by time-expanded wave-form analysis," *N. Eng. J. Med.*, vol. 296, pp. 968-971, 1977.
- [65] M. Mori, K. Kinoshita, H. Morinari, T. Shiraishi, S. Koike, S. Murao, "Waveform and spectral analysis of crackles," *Thorax*, vol. 35, pp. 843-850, 1980.
- [66] S. Holford, "Discontinuous adventitious lung sounds: measurement, classification and modeling," *Doctoral Thesis*, Massachusetts Institute of Technology, 1981.
- [67] J. Hoevers, R. Loudon, "Measuring crackles," *Chest*, vol. 98, no. 5, pp. 1240-1243, 1990.
- [68] M. Munakata, H. Ukita, I. Doi, Y. Ohtsuka, Y. Masaki, Y. Homma, Y. Kawakami, "Spectral and waveform characteristics of fine and coarse crackles," *Thorax*, vol. 46, pp. 651-657, 1991.
- [69] M. Matsuzaki, "Polarity of crackle waveforms: a new index for crackle differentiation," *Hokkaido Igaku Zasshi*, vol. 60, pp. 104-113, 1985.
- [70] A. Vyshedskiy, R. Alhashem, R. Paciej, M. Ebril, I. Rudman, J. Fredberg, R. Murphy, "Mechanism of inspiratory and expiratory crackles," *Chest*, vol. 135, pp. 156-164, 2009
- [71] R. Katila, P. Piirilä, K. Kallio, E. Paajanen, T. Rosqvist, A. Sovijärvi, "Original waveform of lung sound crackles: a case study of the effect of high-pass filtration," *J. Appl. Physiol.*, vol. 71, pp. 2173-2177, 1991.
- [72] A. Nath, L. Capel, "Inspiratory crackles and mechanical events of breathing," *Thorax*, vol. 29, pp. 695-698, 1974.
- [73] J. Fredberg, S. Holford, "Discrete lung sounds: crackles (rales) as stress-relaxation quadrupoles," *J. Acoust. Soc. Am.*, vol. 73, no. 3, 1036-1046, 1983.
- [74] G. Epler, C. Carrington, E. Gaensler, "Crackles (rales) in the interstitial pulmonary diseases," *Chest*, vol. 73, pp. 333-339, 1978.
- [75] R. Murphy, K. Sorensen, "Chest auscultation in the diagnosis of pulmonary asbestosis," *J. Occup. Med.*, vol. 15, pp. 272-276, 1973.
- [76] P. Piirilä, A. Sovijärvi, T. Kaisla, H. Rajala, T. Katila, "Crackles in patients with fibrosing alveolitis, bronchiectasis, COPD and heart failure," *Chest*, vol. 99, pp. 1076-1083, 1991.
- [77] P. Piirilä, "Changes in crackle characteristics during the clinical course of pneumonia," *Chest*, vol. 102, pp. 176-183, 1992.
- [78] F. Dalmaso, M. Guarene, R. Spagnolo, G. Benedetto, G. Righini, "A computer system for timing and acoustical analysis of crackles: a study in criptogenic fibrosing alveolitis," *Bull. Eur. Physiopathol. Respir.*, vol. 20, 139-144, 1984.
- [79] T. Kaisla, A. Sovijärvi, P. Piirilä, H. Rajala, S. Haltsonen, T. Rosqvist, "Validated method for automatic detection of lung sound crackles," *Med. Biol. Eng. Comput.*, vol. 29, pp. 517-521, 1991.
- [80] T. Rosqvist, E. Paajanen, K. Kallio, H. Rajala, T. Katila, P. Piirilä, P. Malmberg, A. Sovijärvi, "Toolkit for lung sound analysis," *Med. Biol. Eng. Comput.*, vol. 33, pp. 190-195, 1995.
- [81] B. Sankur, E. Guler, Y. Kahya, "Multiresolution biological transient extraction applied to respiratory crackles," *Comput. Biol. Med.*, vol. 26, pp. 25-39, 1996.
- [82] M. Ono, K. Arakawa, M. Mori, T. Sugimoto, H. Harashima, "Separation of fine crackles from vesicular sounds by a nonlinear digital filter," *IEEE Trans. Biomed. Eng.*, vol. 36, pp. 286-291, 1989.
- [83] H. Pasika, D. Pengelly, "Lung sound crackle analysis using generalised time-frequency representations," *Med. Biol. Eng. Comput.*, vol. 32, pp. 688-690, 1994.
- [84] S. Charleston-Villalobos, R. González-Camarena, G. Chi-Lem, T. Aljama-Corrales, "Crackle sounds analysis by empirical mode decomposition," *IEEE Eng. Med. Biol. Mag.*, pp. 40-47, January/February 2007.
- [85] N. Meslier, G. Charbonneau, J. Racineux, "Wheezes," *Eur. Respir. J.*, vol. 8, pp. 1942-1948, 1995.
- [86] N. Gavriely, Y. Palti, G. Alroy, J. Grotberg, "Measurement and theory of wheezing breath sounds," *J. Appl. Physiol.*, vol. 57, no. 2, pp. 481-492, 1984.

- [87] H. Pasterkamp, T. Asher, F. Leahy, R. Fenton, V. Chernik, "The effect of anticholinergic treatment on postexertional wheezing in asthma studied by phonopneumography and spirometry," *Am. Rev. Respir. Dis.*, vol. 132, pp. 16-21, 1985.
- [88] J. Grotberg, S. Davis, "Fluid dynamic flapping of a collapsible channel: sound generation and flow limitation," *J. Biomech.*, vol. 13, pp. 219-230, 1980.
- [89] J. Grotberg, E. Reiss, "A subsonic flutter anomaly," *J. Sound Vib.*, vol. 80, pp. 444-446, 1982.
- [90] J. Grotberg, E. Reiss, "Subsonic flapping flutter," *J. Sound Vib.*, vol. 92, pp. 349-361, 1984.
- [91] Y. Shabtai-Musih, J. Grotberg, N. Gavriely, "Spectral content of forced expiratory wheezes during air, He, and SF₆ breathing in normal humans," *J. Appl. Physiol.*, vol. 72, no. 2, pp. 629-635, 1992.
- [92] R. Fenton, H. Pasterkamp, A. Tal, V. Chernik, "Automated spectral characterization of wheezing in asthmatic children," *IEEE Trans. Biomed. Eng.*, vol. 32, pp. 50-55, 1985.
- [93] G. Wodicka, K. Stevens, H. Golub, E. Cravalho, D. Shannon, "A model of acoustic transmission in the respiratory system," *IEEE Trans. Biomed. Eng.*, vol. 36, pp. 925-934, 1989.
- [94] H. Pasterkamp, R. Fenton, A. Tal, V. Chernik, "Tracheal vs lung sounds in acute asthma," *Am. Rev. Respir. Dis.*, vol. 129, pp. 256A, 1984.
- [95] R. Baughman, R. Loudon, "Stridor: differentiation from asthma or upper airway noise," *Am. Rev. Respir. Dis.*, vol. 139, pp. 1407-1409, 1989.
- [96] S. Cortes, R. Jané, J. Fiz, J. Morera, "Monitoring of wheeze duration during spontaneous respiration in asthmatic patients," *Proc. 27th Annual Int. Conf. IEEE-EMBS*, pp. 6141-6144, 2005.
- [97] R. Baughman, R. Loudon, "Quantitation of wheezing in acute asthma," *Chest*, vol. 86, pp. 718-722, 1984.
- [98] S. Taplidou, L. Hadjileontiadis, T. Penzel, V. Gross, S. Panas, "WED: An efficient wheezing-episode detector based on breath sounds spectrogram analysis," *Proc. 25th Annual Int. Conf. IEEE-EMBS*, pp. 2531-2534, 2003.
- [99] A. Homs-Corbera, J. Fiz, J. Morera, R. Jané, "Time-frequency detection and analysis of wheezes during forced exhalation," *IEEE Trans. Biomed. Eng.*, vol. 51, pp. 182-186, 2004.
- [100] R. Riella, P. Nohama, R. Borges, A. Stelle, "Automatic wheezing recognition in recorded lung sounds," *Proc. 25th Annual Int. Conf. IEEE-EMBS*, pp. 2535-2538, 2003.
- [101] H. Meng-Lun, C. Jen-Chien, C. Feng-Chia, W. Huey-Dong, C. Fok-Ching, "Respiratory wheeze detection system," *Proc. 27th Annual Int. Conf. IEEE-EMBS*, pp. 7553-7559, 2005.
- [102] S. Taplidou, L. Hadjileontiadis, I. Kitsas, K. Panoulas, T. Penzel, V. Gross, S. Panas, "On applying Continuous Wavelet Transform in wheeze analysis," *Proc. 26th Annual Int. Conf. IEEE-EMBS*, pp. 3832-3835, 2004.
- [103] J. Wood, T. Barry, "Time-frequency analysis of the first heart sound," *IEEE Eng. Med. Biol. Mag.*, pp. 144-151, vol. 14, 1995.
- [104] L. Durand, P. Pibarot, "Digital signal processing of the phonocardiogram: review of the most recent advancements," *CRC Critical Reviews Biomed. Eng.*, vol. 23, pp. 163-219, 1995.
- [105] D. Winer, L. Perry, C. Caceres, "Heart sound analysis: a three dimensional approach," *Am. J. Cardiol.*, vol. 16, pp. 547-551, 1965.
- [106] J. Wood, D. Barry, "Time-frequency analysis of skeletal muscle and cardiac vibrations," *Proc. of the IEEE*, vol. 84, no. 9, pp. 1281-1294, 1996.
- [107] T. Joo, J. McClellan, R. Foale, G. Myers, R. Lees, "Pole-zero modeling and classification of phonocardiograms," *IEEE Trans. Biomed. Eng.*, vol. BME-30, no. 2, pp. 110-118, 1983.
- [108] D. Chen, L. Durand, H. Lee, "Time-frequency analysis of the first heart sound. Part 1: simulation and analysis," *Med. Biol. Eng. Comput.*, vol. 35, pp. 306-310, 1997.
- [109] P. Stein, H. Sabbah, J. Lakier, D. Magilligan, D. Goldstein, "Frequency of the first heart sound in the assessment of stiffening of mitral bioprosthetic valves," *Circulation*, vol. 63, pp. 200-203, 1981.
- [110] L. Durand, J. De Guise, G. Cloutier, R. Guardo, M. Brais, "Evaluation of the FFT-based and modern parametric methods for spectral analysis of bioprosthetic valve sounds," *IEEE Trans. Biomed. Eng.*, vol. BME-33, no. 6, pp. 572-578, 1986.
- [111] T. Hearn, J. Mazumdar, R. Hubbard, G. Eyster, "Temporal and heart-size in first-heart-sound spectra," *Med. Biol. Eng. Comput.*, vol. 17, pp. 563-568, 1979.
- [112] R. Adolph, J. Stephens, K. Tanaka, "The clinical value of frequency analysis of the first heart sound in myocardial infarction," *Circulation*, vol. 41, pp. 1003-1014, 1970.
- [113] A. Yoganathan, R. Gupta, F. Udawadia, J. Miller, W. Corcoran, R. Sarma, J. Jonson, R. Bing, "Use of the fast Fourier transform for frequency analysis of the first heart sound in normal man," *Med. Biol. Eng.*, vol. 14, pp. 69-73, 1976.

- [114] W. Clarke, S. Austin, P. Shah, P. Griffen, J. Dove, J. McCullough, B. Schreiner, "Spectral energy of the first heart sound in acute myocardial ischemia," *Circulation*, vol. 57, pp. 593-598, 1978.
- [115] A. Iwata, N. Ishii, N. Suzumura, K. Ikegaya, "Algorithm for detecting the first and second heart sounds by spectral tracking," *Med. Biol. Eng. Comput.*, vol. 18, pp. 19-26, 1980.
- [116] P. Stein, H. Sabbah, J. Lakier, S. Kemp, D. Magilligan, "Fundamental basis and clinical validation of frequency analysis of the valve closure sound as an indicator of stiffening of bioprosthetic valves," *J. Cardiology*, vol. 14, pp. 57-58, 1984.
- [117] L. Durand, M. Blanchard, H. Sabbah, M. Hamid, S. Kemp, P. Stein, "A Bayes model for automatic detection and quantification of bioprosthetic valve degeneration," *Math. Comput. Modell.*, vol. 11, pp. 158-163, 1988.
- [118] M. Akay, J. Semmlow, Welkowitz, M. Bauer, J. Kostis, "Detection of coronary occlusions using autoregressive modeling of diastolic heart sounds," *IEEE Trans. Biomed. Eng.*, vol. 37, no. 4, pp. 366-373, 1990.
- [119] H. Sava, J. McDonnell, K. Fox, "Spectral analysis of first and second heart sounds before and after mechanical heart valve implantation," *Proc. 16th Annual Int. Conf. IEEE Eng. Med. & Biol. Soc.*, pp. 1280-1281, 1994.
- [120] J. Wood, A. Buda, D. Barry, "Time-frequency transforms: a new approach to first heart sound frequency dynamics," *IEEE Trans. Biomed. Eng.*, vol. 39, no. 7, pp. 730-740, 1992.
- [121] L. Durand, J. Genest, R. Guardo, "Modeling of the transfer function of the heart-thorax acoustic system in dogs," *IEEE Trans. Biomed. Eng.*, vol. BME-32, no. 8, pp. 592-601, 1985.
- [122] R. Koenig, H. K. Dunn, L. Y. Lacy, "The sound spectrograph," *J. Acoust. Soc. Amer.*, vol. 18, pp. 19-49, 1946.
- [123] G. Geckeler, W. Mason, R. Riesz, C. Wirth, "Cardiospectrograms," *Am. Heart J.*, vol. 48, pp. 190-196, 1954.
- [124] V. McKusick, G. Webb, J. Humphries, J. Reid, "On cardiovascular sound: further observations by means of spectral phonocardiography," *Circulation*, vol. 11, pp. 849-870, 1955.
- [125] J. Hylen, F. Kloster, R. Herr, A. Starr, H. Griswold, "Sound spectrographic diagnosis of aortic ball variance," *Circulation*, vol. 34, pp. 849-858, 1969.
- [126] D. Suobanck, A. Yoganathan, E. Harrison, W. Corcoran, "A quantitative method for the in vitro study of sounds produced by prosthetic aortic heart valves. Parts 2-3," *Med. Biol. Eng. Comput.*, vol. 22, pp. 40-54, 1984.
- [127] A. Iwata, N. Suzumura, K. Ikegaya, "Pattern classification of the phonocardiogram using linear prediction analysis," *Med. Biol. Eng. Comput.*, vol. 15, pp. 407-412, 1977.
- [128] G. Jamous, L. Durand, Y. Langlois, T. Lanthier, P. Pibarot, S. Cario, "Optimal time-window duration for computing time/frequency representations of normal phonocardiograms in dogs," *Med. Biol. Eng. Comput.*, vol. 30, pp. 503-508, 1992.
- [129] L. Khadra, M. Matalgah, B. El-Asir, S. Mawagdeh, "The wavelet transform and its applications to phonocardiograms signal analysis," *Med. Inform.*, vol. 16, pp. 271-277, 1991.
- [130] J. Bulgrin, B. Rubal, C. Thompson, J. Moody, "Comparison of short-time Fourier, wavelet and time-domain analyses of intracardiac sounds," *Biomed. Sci. Instrum.*, vol. 29, pp. 465-472, 1993.
- [131] D. Chen, L. Durand, H. Lee, "Selection of a time-frequency representation for analysis of the first heart sound signal," *Proc. 16th Annual Int. Conf. IEEE-EMBS*, pp. 1276-1277, 1994.
- [132] D. Chen, L. Durand, Z. Guo, H. Lee, "Time-frequency analysis of the first heart sound. Part 2: an appropriate time-frequency representation technique," *Med. Biol. Eng. Comput.*, vol. 35, pp. 311-323, 1997.
- [133] Y. Zhao, L. Atlas, R. Marks, "The use of Cone-shaped kernels for generalized time-frequency representations of nonstationary signals," *IEEE Trans. Signal Process.*, vol. 38, no. 7, pp. 1084-1091, 1990.
- [134] Z. Guo, L. Durand, H. Lee, "The time-frequency distributions of nonstationary signals based on a Bessel kernel," *IEEE Trans. Signal Process.*, vol. 42, no. 7, pp. 1700-1707, 1994.
- [135] J. Bulgrin, B. Rubal, T. Posch, J. Moody, "Comparison of Binomial, ZAM and minimum cross-entropy time-frequency distributions of intracardiac heart sounds," *Record of the 28th Asilomar Conf. on Signals, Systems and Computers*, pp. 383-387, 1994.
- [136] J. Xu, L. Durand, P. Pibarot, "Nonlinear transient chirp signal modeling of the aortic and pulmonary components of the second heart sound," *IEEE Trans. Biomed. Eng.*, vol. 47, pp. 1328-1335, 2000.
- [137] A. Leatham, "Auscultation of the heart since Laënnec," *Thorax*, vol. 36, pp. 95-98, 1981.
- [138] C. Wooley, "Intracardiac phonocardiography. Intracardiac sound and pressure in man," *Circulation*, vol. 57, no. 6, pp. 1039-1054, 1978.

- [139] J. Abrams, "Current concepts of the genesis of heart sounds I. First and second heart sounds," *JAMA*, vol. 239, pp. 2787-2789, 1978.
- [140] S. Hirschfeld, J. Liebman, C. Bormuth, "Intracardiac pressure-sound correlates of echographic aortic valve closure," *Circulation*, vol. 55, pp. 602-604, 1977.
- [141] G. Cloutier, M. Grenier, R. Guardo, L. Durand, "Spectral analysis of closing sounds produced by Ionescu-Shiley bioprosthetic aortic heart valves. Part 2," *Med. Biol. Eng. Comput.*, vol. 25, pp. 492-496, 1987.
- [142] J. Shaver, R. Salerni, P. Reddy, "Normal and abnormal heart sounds in cardiac diagnosis. Part 1: systolic sounds," *Curr. Prob. Cardiol.*, vol. 10, pp. 2-68, 1985.
- [143] J. Xu, L. Durand, P. Pibarot, "Extraction of the aortic and pulmonary components of the second heart sound using a nonlinear transient chirp signal model," *IEEE Trans. Biomed. Eng.*, vol. 48, no. 3, pp. 277-283, 2001.
- [144] A. Yoganathan, R. Gupta, W. Corcoran, F. Udawadia, R. Sarma, R. Bing, "Use of the fast Fourier transform in the frequency analysis of the second heart sound in normal man," *Med. Biol. Eng.*, vol. 14, pp. 455-460, 1976.
- [145] R. Foale, T. Joo, J. McClellan, R. Metzinger, G. Grant, G. Myers, R. Lees, "Detection of aortic porcine valve dysfunction by maximum entropy spectral analysis," *Circulation*, vol. 68, pp. 42-49, 1983.
- [146] S. Aggio, E. Baracca, C. Longhini, C. Brunazzi, L. Longhini, G. Musacci, C. Fersini, "Noninvasive estimation of the pulmonary systolic pressure from the spectral analysis of the second heart sound," *Acta Cardiol.*, vol. 45, pp. 199-202, 1990.
- [147] C. Longhini, E. Baracca, C. Brunazzi, M. Vaccari, L. Longhini, F. Barbaresi, "A new noninvasive method for estimation of pulmonary arterial pressure in mitral stenosis," *Am. J. Cardiol.*, vol. 68, pp. 398-401, 1991.
- [148] A. Bartels, D. Harder, "Non-invasive determination of systolic blood pressure by heart sound pattern analysis," *Clin. Phys. Physiol. Meas.*, vol. 13, pp. 249-256, 1992.
- [149] H. Köymen, B. Altay, Y. Ziya, "A study of prosthetic heart valve sounds," *IEEE Trans. Biomed. Eng.*, vol. BME-34, no. 11, pp. 853-863, 1987.
- [150] Y. Tang, C. Danmin, L. Durand, "The synthesis of the aortic valve closure sound of the dog by the mean filter of forward and backward predictor," *IEEE Trans. Biomed. Eng.*, vol. 39, pp. 1-8, 1992.
- [151] H. Sava, T. McDonnell, "New analysis-synthesis of first heart sounds using forward-backward overdetermined Prony's method," *Proc. 16th Annual Int. Conf. IEEE Eng. Med. & Biol. Soc.*, pp. 1274-1275, 1994.
- [152] H. Sava, T. McDonnell, "Differences in spectral composition between monostrut Bjork-Shiley and Carbomedics valves implanted in the aortic position," *Med. Biol. Eng. Comput.*, vol. 33, pp. 689-694, 1995.
- [153] A. Baykal, Y. Ziya, H. Köymen, "Distribution of aortic mechanical prosthetic valve closure sound model parameters on the surface of the chest," *IEEE Trans. Biomed. Eng.*, vol. 42, no.4, 358-370, 1995.
- [154] X. Zhang, L. Durand, L. Senhadji, H. Lee, J. Coatrieux, "Analysis-synthesis of the phonocardiogram based on the matching pursuit method," *IEEE Trans. Biomed. Eng.*, vol. 45, no. 8, pp. 962-971, 1998.
- [155] T. Tran, N. Jones, J. Fothergill, "Heart sound simulator," *Med. Biol. Eng. Comput.*, vol. 33, pp. 357-359, 1995.
- [156] M. Obaidat, "Phonocardiogram signal analysis: techniques and performance comparison," *J. Med. Eng. Technol.*, vol. 17, no. 6, pp. 221-227, 1993.
- [157] S. Debbal, F. Bereksi, "Time-frequency analysis of the second cardiac sound in phonocardiogram signals," *Med. Phys.*, vol. 32, no. 9, pp. 2911-2917, 2005.
- [158] S. Debbal, F. Bereksi, "Automatic measure of the split in the second cardiac by using the wavelet transform technique," *Comput. Biol. Med.*, vol. 37, pp. 269-276, 2007.
- [159] S. Charleston-Villalobos, T. Aljama-Corrales, R. Gonzalez-Camarena, "Analysis of simulated heart sounds by intrinsic mode functions," *Proc. 28th Annual Int. Conf. IEEE-EMBS*, pp. 2848-2851, 2006.

III. PLANTEAMIENTO DEL PROBLEMA.

III.1. Introducción

Aunque la auscultación de los sonidos producidos por los órganos internos de la cavidad torácica quizá se remonta hasta el siglo V a. C., no fue sino hasta el siglo XVIII que a la técnica de auscultación le fue atribuida un potencial diagnóstico. Un siglo después, con el posterior advenimiento del estetoscopio, la auscultación se popularizó a tal grado que actualmente permanece como una herramienta valiosa debido a sus características no invasivas, entre otras.

Durante la auscultación, el médico está interesado en una variedad de sonidos tanto normales como anormales producidos dentro del tórax con el propósito final de determinar la naturaleza de estos sonidos mediante su estudio cuidadoso. Sin embargo, la auscultación clínica posee ciertas desventajas, siendo las más importantes la respuesta en frecuencia del estetoscopio, la ausencia de almacenamiento de la información para su posterior procesamiento y la descripción cualitativa de los sonidos escuchados por el médico. Al igual que Laennec lo hiciera dos siglos atrás, el médico usualmente describe las señales acústicas mediante adjetivos relacionados con sonidos cotidianos para él. Por ejemplo, las siguientes palabras se emplean para describir a los sonidos torácicos: musical, suave, rugoso, moderadamente intenso, intenso y de tono bajo, tono moderadamente alto, mojado, fino, grueso, etc. La mayoría de los términos anteriores son puramente relativos, y sólo mediante la experiencia y el acuerdo mutuo ha existido un grado de uniformidad razonable en la fraseología empleada para describir estos sonidos.

Los sonidos provenientes de la cavidad torácica son diversos y con una estructura que se modifica rápidamente en el tiempo, reflejando la complejidad de los sistemas biológicos que los emiten. Tanto los sonidos respiratorios como los cardiacos comparten la característica de no estacionaridad. Como resultado, los mecanismos que dan génesis a algunos de los sonidos torácicos aún suscitan controversia y otros, ni siquiera han sido establecidos. Quizá ninguno de estos sonidos aparezca como simples tonos musicales, caracterizados por su frecuencia central, siendo más probable que consistan de una mezcla de componentes con diversas duraciones en una o más regiones en el dominio de la frecuencia. Además, es posible que la importancia de los sonidos dependa de las intensidades relativas de sus diversos componentes espectrales.

Para contender con estas características de no estacionaridad y de componentes múltiples de la señal, es necesario una técnica de procesamiento capaz de tomar la información unidimensional de la señal y proporcionar simultáneamente el contenido en frecuencia, duración e intensidad de cada uno de sus mono-componentes, i.e. una señal bien delimitada en una región del plano tiempo-frecuencia. Este enfoque lleva naturalmente al empleo de representaciones tiempo-frecuencia, las cuales se espera que proporcionen los elementos requeridos para investigar los vínculos entre las características acústicas de las señales y los procesos físicos y fisiológicos que las generan.

Las técnicas de procesamiento de señales han ayudado a esclarecer el origen de los diversos sonidos torácicos. En particular, el análisis tiempo-frecuencia ha sido una herramienta valiosa en el campo de los sonidos cardiacos, donde se han realizado diversos esfuerzos por encontrar una representación tiempo-frecuencia adecuada para su análisis y su posterior aplicación en el reconocimiento de patrones. Sin embargo, en el campo de los sonidos respiratorios la herramienta clásica de análisis tiempo-frecuencia continúa siendo el espectrograma, en este caso denominado respirosonograma. La principal desventaja del espectrograma, aplicado a una señal ventaneada, consiste en su inherente compromiso entre su resolución temporal y su resolución espectral, como dicta el principio de incertidumbre de Heisenberg-Gabor, y por consiguiente su aplicación en registros de corta duración puede derivar en resultados pobres. A pesar de los esfuerzos por aplicar otras técnicas de representación tiempo-frecuencia para analizar a los sonidos respiratorios, aún no se ha realizado una comparación cuantitativa entre las técnicas clásicas y las más recientes.

Considerando los resultados obtenidos al emplear y comparar representaciones tiempo-frecuencia para analizar los dos primeros sonidos cardiacos, la aplicación de técnicas de tiempo-frecuencia recientes resulta prometedora en el campo de los sonidos respiratorios. Además, dado que las señales de sonido cardiaco se han estudiado y modelado en mayor medida que las señales de sonido respiratorio, la experiencia en el análisis de los sonidos cardiacos resulta un buen punto de partida. De forma complementaria, el análisis tiempo-frecuencia de los sonidos cardiacos, a la par de los sonidos respiratorios, también resulta necesario si se considera que una de las tareas de procesamiento más básicas consiste en la eliminación de señales de interferencia; una de las líneas de investigación en sonidos respiratorios es minimizar la presencia de los sonidos cardiacos. Además, algunas de las características presentes en los sonidos cardiacos resultan similares a las encontradas en los

sonidos respiratorios. En particular, el comportamiento del contenido en frecuencia del segundo sonido cardiaco es similar al esperado para un sonido respiratorio adventicio discontinuo. Así, aunque el enfoque de esta tesis es el análisis tiempo-frecuencia de los sonidos respiratorios, el empleo de los sonidos cardiacos está justificado dado que el conocimiento que pueda alcanzarse con ellos resultará útil al confrontarse con los sonidos producidos durante la respiración.

III.2. Hipótesis

Debido a la no estacionaridad y posiblemente no linealidad de los sonidos respiratorios, las técnicas tiempo-frecuencia más recientes proporcionan una mejor representación en el espacio tiempo-frecuencia de los sonidos pulmonares y adventicios.

III.3. Objetivos

III.3.1. Objetivo general.

Evaluar representaciones tiempo-frecuencia desarrolladas recientemente, que permitan aumentar la resolución en el análisis temporal y espectral para los sonidos respiratorios, contrastando los resultados obtenidos con las representaciones clásicas tiempo-frecuencia.

III.3.2. Objetivos específicos.

- Proponer y emplear diversas técnicas de análisis tiempo-frecuencia obtenidas recientemente para analizar a los dos primeros sonidos cardiacos y los sonidos respiratorios normales y adventicios.
- Programar las representaciones tiempo-frecuencia de tal forma que los resultados obtenidos con cada una de ellas resulten adecuados para su posterior comparación.
- Obtener una adecuada visualización de los resultados obtenidos con las técnicas de análisis tiempo-frecuencia.
- Emplear señales de sonidos torácicos simulados para verificar la programación de las diversas técnicas.
- Para cada una de las técnicas, variar sus parámetros y seleccionar la representación tiempo-frecuencia más adecuada.

- Establecer un esquema para comparar cuantitativamente las diversas técnicas tiempo-frecuencia para cada señal de sonido.
- Aplicar a señales reales adquiridas la técnica de análisis tiempo-frecuencia que resulte más adecuada para analizar cada tipo de sonido torácico y contrastar los resultados con los obtenidos por la técnica clásica.

IV. REPRESENTACIONES TIEMPO-FRECUENCIA.

IV.1. Introducción.

Una distribución tiempo-frecuencia es dependiente de la señal, y por lo tanto representa a la señal en el dominio del tiempo y de la frecuencia, lo cual brinda lugar a la frase *representación tiempo-frecuencia* [1]; la idea de la representación es similar a la empleada en una partitura musical, la cual muestra las frecuencias a ejecutar a lo largo del tiempo. Recientemente, la representación tiempo-frecuencia ha recibido una considerable atención como herramienta para el análisis de una gran variedad de señales y sistemas. En particular, si el contenido en frecuencia de la señal varía con el tiempo, como es el caso de las señales no-estacionarias, entonces la aproximación resulta sumamente atractiva. Aunque tanto la descripción en el tiempo o en la frecuencia ofrecen una información completa acerca de la señal, asumiendo que se trata de señales deterministas, ninguna de ellas revela explícitamente el espectro de frecuencia a un tiempo particular o el tiempo al cual ocurre un componente específico de frecuencia. En contraparte, mediante el mapeo de una función unidimensional de tiempo o frecuencia hacia una función bidimensional de tiempo y frecuencia, se obtiene la representación tiempo-frecuencia que localiza la energía en ambas direcciones [2].

Para que una distribución particular corresponda a una distribución de la energía de la señal en el tiempo y la frecuencia, la distribución debe poseer ciertas propiedades [3]. Por ejemplo, la distribución tiempo-frecuencia debe ser una distribución evaluada en los reales tal que el corrimiento de una señal, en el tiempo o en la frecuencia, resulte en el correspondiente corrimiento de la distribución. Además, la proyección de la distribución sobre el dominio del tiempo, o sobre el dominio de la frecuencia, debe ser igual a la potencia instantánea, o la densidad espectral de la señal. Más aún, el centroide de la distribución en cada instante de tiempo, o en cada frecuencia, debe ser igual a la frecuencia instantánea, o al retardo de grupo de la señal [4].

IV.2. La clase general de distribuciones.

Los dos trabajos originales que establecieron el concepto de la distribución tiempo-frecuencia se deben a Gabor [5] y a Wigner [6]. La distribución de Wigner fue el primer ejemplo de una representación tiempo-frecuencia cualitativamente diferente del espectrograma. La principal motivación de Wigner radicaba en calcular la corrección cuántica de un coeficiente que indica como un gas se desvía de la ley de los gases ideales;

clásicamente, para calcular dicho coeficiente se necesita una distribución conjunta de la posición y el momento. En 1932, Wigner ideó una representación basada en las distribuciones marginales de la información mecánica cuántica de la posición y el momento [6]. La aplicación de la distribución de Wigner al análisis de señales fue abordada inicialmente por Ville [7]. Es importante mencionar, desde el punto de vista cronológico, que la idea del espectrograma fue cristalizada a mediados de la década de 1940s [5, 8] y que independientemente a este desarrollo existió una considerable actividad, en las décadas comprendidas entre 1940s a 1960s, con el propósito de idear distribuciones que fueran similares a la distribución de Wigner en el sentido de que satisficieran las marginales, la condición de frecuencia instantánea y otras propiedades deseables. Entre las distribuciones propuestas, tanto en el análisis de señales como en mecánica cuántica, se encontraban la Rihaczek [9], Page [10] y Margenau-Hill [11].

En el año de 1966 fue concebido un método que podía generar de una manera simple un número infinito de nuevas distribuciones [12], caracterizada cada una de ellas mediante una función auxiliar, la función kernel. Esta clase general de distribuciones ha sido denominada clase de distribuciones de Cohen, en honor al trabajo pionero de Leon Cohen. Todas las distribuciones tiempo-frecuencia de la clase general, denominadas en el presente texto como $C(t, \omega)$, se pueden obtener mediante:

$$C(t, \omega) = \frac{1}{4\pi^2} \iiint s^*(u - \frac{1}{2}\tau) s(u + \frac{1}{2}\tau) \phi(\theta, \tau) e^{-j\theta t - j\tau\omega + j\theta u} du d\tau d\theta \quad (1)$$

donde $s(u)$ es la señal bajo análisis en el dominio del tiempo, $s^*(u)$ es su complejo conjugado, y $\phi(\theta, \tau)$ es una función bidimensional denominada por Claasen y Mecklenbräuker como el kernel de la distribución [13], un término que se ha acuñado debido a las importantes contribuciones de estos autores al entendimiento de la clase general, particularmente en el contexto del análisis de señales. La clase general de distribuciones también puede expresarse en términos de la transformada de Fourier de la señal bajo análisis de la forma siguiente:

$$C(t, \omega) = \frac{1}{4\pi^2} \iiint S^*(u + \frac{1}{2}\theta) S(u - \frac{1}{2}\theta) \phi(\theta, \tau) e^{-j\theta t - j\tau\omega + j\tau u} d\theta d\tau du \quad (2)$$

Existen otras formas de escribir la clase general de distribuciones tiempo-frecuencia que resultan convenientes para diseñar un kernel o implementar una distribución tiempo-frecuencia [2, 12]. Dos de estas alternativas corresponden a las formulaciones mediante la función característica de la distribución y la función de autocorrelación dependiente del tiempo de la señal. En el primer caso, la función característica es la doble transformada de Fourier de la distribución tiempo-frecuencia y por lo tanto, utilizando la transformada inversa de Fourier, es posible escribir la clase general de distribución como [3, 12, 14]:

$$C(t, \omega) = \frac{1}{4\pi^2} \iint M(\theta, \tau) e^{-j\theta t - j\tau \omega} d\theta d\tau \quad (3)$$

donde, es posible identificar a partir de la ecuación (1) que:

$$\begin{aligned} M(\theta, \tau) &= \phi(\theta, \tau) \int s^*(u - \frac{1}{2}\tau) s(u + \frac{1}{2}\tau) e^{j\theta u} du \\ M(\theta, \tau) &= \phi(\theta, \tau) A_s(\theta, \tau) \end{aligned} \quad (4)$$

Dado que a la función bidimensional $A_s(\theta, \tau)$ se le conoce como la función de ambigüedad simétrica de la señal, entonces la función característica $M(\theta, \tau)$ puede ser apropiadamente denominada la función de ambigüedad generalizada.

La segunda formulación corresponde a una generalización de las relaciones de Wiener-Khinchin [15], es decir, pensando a la distribución tiempo-frecuencia como la transformada de Fourier de una función de autocorrelación local, $R_l(\tau)$:

$$C(t, \omega) = \frac{1}{2\pi} \int R_l(\tau) e^{-j\tau \omega} d\tau \quad (5)$$

donde, comparando con la ecuación de la clase general, se tiene que:

$$R_l(\tau) = \frac{1}{2\pi} \iint s^*(u - \frac{1}{2}\tau) s(u + \frac{1}{2}\tau) \phi(\theta, \tau) e^{j\theta(u-t)} d\theta du \quad (6)$$

y $R_l(\tau)$ se denomina la función de autocorrelación local determinista generalizada.

Como se observa, la función kernel $\phi(\theta, \tau)$ juega un papel primordial en la definición y caracterización de las distribuciones tiempo-frecuencia, y ofrece varias ventajas: a) se puede obtener y estudiar las propiedades de las distribuciones aplicando ciertas restricciones al kernel; lo anterior permite escoger aquellas funciones kernel que produzcan distribuciones con propiedades deseables y prescritas [1]; b) las propiedades de la distribución pueden ser fácilmente establecidas mediante un análisis del kernel, e.g. en [2, 4, 14] se reportan las

propiedades que satisfacen algunas distribuciones de la clase general y los requerimientos de las funciones kernel asociadas; c) dado un kernel, es sencillo generar una distribución tiempo-frecuencia, i.e. *cualquier miembro de la clase de distribuciones de Cohen puede encontrarse aplicando la doble transformada de Fourier al resultado del producto de la función de ambigüedad simétrica por la función kernel* [16]. Una característica importante de una distribución tiempo-frecuencia de Cohen es su invarianza a traslaciones en el tiempo y en la frecuencia [17].

Tres representaciones tiempo-frecuencia importantes en el análisis de señales que pertenecen a la clase general de Cohen son: la distribución de Wigner, el Espectrograma y la distribución Choi-Williams. En particular, la distribución de Wigner ha recibido una gran atención como una herramienta conveniente para el análisis de señales monocomponentes, entendidas como aquellas señales cuya frecuencia instantánea y retardo de grupo corresponden a la misma curva en el plano tiempo-frecuencia [16, 18], ejemplos de estas señales monocomponentes incluyen a las señales sinusoidales y señales chirp. En el caso de señales que pueden ser expandidas como la suma de sus monocomponentes, la estructura bilineal inherente a las distribuciones de la clase de Cohen genera términos cruzados de interferencia que son indeseables [19]. Estos términos cruzados pueden resultar problemáticos cuando se trata de utilizar las características tiempo-frecuencia con propósitos de reconocimiento de patrones y clasificación. En general, pueden ocurrir diferentes combinaciones de componentes que pertenecen a la representación tiempo frecuencia correcta de la señal (auto-términos) y términos cruzados, e.g. los términos cruzados pueden ocurrir en las posiciones de los auto términos y oscurecerlos. La cantidad y forma de esta interferencia están directamente relacionadas con la función kernel empleada para obtener la distribución particular. *Por lo tanto, un kernel efectivo para la representación tiempo-frecuencia debería disminuir los efectos de los términos cruzados conservando las propiedades deseables que sostienen la validez de la distribución tiempo-frecuencia* [3].

En la tabla IV.I se presentan las funciones kernel correspondientes a la representación tiempo-frecuencia de Wigner, Espectrograma y Choi-Williams; y en la figura IV.1 se muestra el comportamiento de las funciones kernel mencionadas en el dominio de la función de ambigüedad, donde θ es el eje doppler y τ es el eje de retardo. En la tabla IV.II se presenta una comparación de las propiedades que satisfacen estas tres distribuciones. Una comparación de las resoluciones de estas y otras distribuciones de la clase general se encuentra en [19].

En las siguientes secciones se detallan, de acuerdo a su aparición cronológica, las tres distribuciones de la clase de Cohen empleadas en esta tesis: la distribución de Wigner, el espectrograma y la distribución Choi-Williams.

TABLA IV.I. ALGUNAS DISTRIBUCIONES DE LA CLASE DE COHEN Y SUS FUNCIONES KERNEL.

Nombre	Kernel: $\phi(\theta, \tau)$	Distribución: $C(t, \omega)$
Clase general	$\phi(\theta, \tau)$	$\frac{1}{4\pi^2} \iiint s^*(u - \frac{1}{2}\tau) s(u + \frac{1}{2}\tau) \phi(\theta, \tau) e^{-j\theta t - j\tau\omega + j\theta u} du d\tau d\theta$
Wigner	1	$\frac{1}{2\pi} \int s^*(t - \frac{1}{2}\tau) s(t + \frac{1}{2}\tau) e^{-j\tau\omega} d\tau$
Espectrograma	$\frac{1}{2\pi} \int h^*(u - \frac{1}{2}\tau) h(u + \frac{1}{2}\tau) e^{-j\theta u} du$	$\left \frac{1}{\sqrt{2\pi}} \int s(\tau) h(\tau - t) e^{-j\tau\omega} d\tau \right ^2$
Choi-Williams	$e^{-\frac{\theta^2 \tau^2}{\sigma}}$	$\iint \sqrt{\frac{\sigma}{4\pi\tau^2}} e^{-\frac{\sigma(u-t)^2}{4\tau^2} - j\tau\omega} s^*(u - \frac{1}{2}\tau) s(u + \frac{1}{2}\tau) du d\tau$

* Aunque para la distribución de Wigner su kernel es la función unitaria; es posible escribir la forma general de tal manera que el kernel de cualquier distribución sea uno, en cuyo caso el kernel de la distribución de Wigner sería otra función. Para un estudio de la transformación entre distribuciones véase [1].

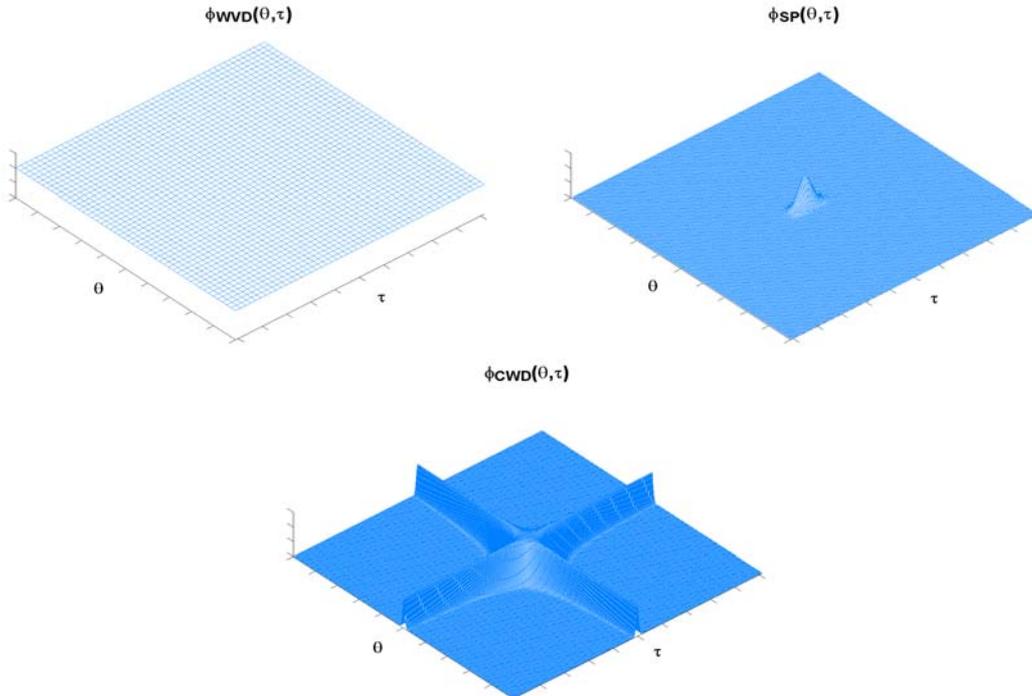


Fig. IV.1. Funciones kernel de tres miembros de la clase de Cohen. a) La distribución de Wigner-Ville, $\phi_{WVD}(\theta, \tau)$. b) El espectrograma, $\phi_{SP}(\theta, \tau)$. c) La distribución Choi-Williams, $\phi_{CWD}(\theta, \tau)$.

TABLA IV.II. PROPIEDADES DE ALGUNAS DISTRIBUCIONES DE LA CLASE DE COHEN.

Propiedad	Wigner	Espectrograma	Choi-Williams
No-negatividad		+	
Evaluada en los reales	+	+	+
Corrimiento en el tiempo*	+	+	+
Corrimiento en la frecuencia*	+	+	+
Marginal en el tiempo	+		+
Marginal en la frecuencia	+		+
Frecuencia instantánea	+		+
Retardo de grupo	+		+
Soporte en el tiempo	+		
Soporte en la frecuencia	+		
Interferencias reducidas		+	+

* Los miembros de la clase de distribuciones de Cohen son invariantes a corrimientos en el tiempo y en la frecuencia.

IV.2.1. La Distribución Wigner-Ville.

La distribución de Wigner fue introducida en el análisis de señales por Ville [7] aproximadamente 15 años después del trabajo de Eugene Wigner [6] y en consecuencia, actualmente, es común denominarla como la distribución Wigner-Ville. Además, Ville derivó la distribución de Wigner empleando un método basado en las funciones características y en 1948, estableció la definición de la distribución empleando el concepto de señal analítica, la cual puede definirse en el dominio del tiempo mediante la transformada de Hilbert de la señal real o en el dominio de la frecuencia mediante la supresión de la parte negativa del espectro de la señal real. Así, suele referirse a la distribución de Wigner-Ville, WVD, cuando se emplea la señal analítica, y reservar el término distribución de Wigner, WD, cuando se emplea la señal real [20].

IV.2.1.1. Formulación de la distribución.

La distribución Wigner-Ville de la señal $s(t)$ se define como [1, 3, 17, 21]:

$$WVD(t, \omega) = \frac{1}{2\pi} \int s^* \left(t - \frac{1}{2}\tau\right) s \left(t + \frac{1}{2}\tau\right) e^{-j\tau\omega} d\tau \quad (7)$$

y en términos de su espectro $S(\omega)$ como:

$$WVD(t, \omega) = \frac{1}{2\pi} \int S^*(\omega + \frac{1}{2}\theta) S(\omega - \frac{1}{2}\theta) e^{-j\theta} d\theta \quad (8)$$

Debido a que la señal o su espectro intervienen dos veces en el cálculo, la distribución de Wigner se considera del tipo bilineal.

A partir de la definición, se observa que para obtener el valor de la distribución para un tiempo particular es necesario considerar el producto de la señal a un tiempo pasado y a un tiempo futuro, con el mismo corrimiento hacia ambos lados. Por lo tanto, para determinar las propiedades de la distribución de Wigner al tiempo t se pliega, en forma similar al procedimiento gráfico para determinar la convolución, la parte izquierda de la señal sobre su parte derecha para ver si existe un traslape. Si existe traslape, entonces las propiedades al tiempo pasado y futuro estarán presentes al tiempo actual de interés, denotado por t . Es decir, la distribución de Wigner es altamente no local dado que pondera los tiempos lejanos igual a los tiempos cercanos [1]. Por otra parte, debido a que la forma de distribución de Wigner es básicamente idéntica en los dominios temporal y espectral, lo válido para un dominio se mantiene válido para el otro. Si estas dos ideas se mantienen en mente, muchos resultados obtenidos para la distribución de Wigner son más fáciles de interpretar.

Para establecer la relación de la distribución de Wigner-Ville con la formulación de la clase general de distribuciones consideremos la función característica de la distribución mediante la doble transformada de Fourier [1, 17]:

$$\begin{aligned} M_{WVD}(\theta, \tau) &= \iint WVD(t, \omega) e^{j\theta t + j\tau \omega} dt d\omega \\ M_{WVD}(\theta, \tau) &= \frac{1}{2\pi} \iiint e^{j\theta t + j\tau \omega} s^*(t - \frac{1}{2}\tau') s(t + \frac{1}{2}\tau') e^{-j\tau' \omega} d\tau' dt d\omega \\ M_{WVD}(\theta, \tau) &= \iint e^{j\theta t} \delta(\tau - \tau') s^*(t - \frac{1}{2}\tau') s(t + \frac{1}{2}\tau') d\tau' dt \\ M_{WVD}(\theta, \tau) &= \int s^*(t - \frac{1}{2}\tau) s(t + \frac{1}{2}\tau) e^{j\theta t} dt \\ M_{WVD}(\theta, \tau) &= A_s(\theta, \tau) \end{aligned} \quad (9)$$

Como se puede observar, la función característica corresponde a la función de ambigüedad simétrica de la señal. Entonces, la distribución de Wigner y la función de ambigüedad son duales en el sentido de que conforman un par transformado de Fourier [3]. En términos del espectro de la señal, la función característica de la distribución de Wigner es [1]:

$$M_{WVD}(\theta, \tau) = \int S^*(\omega + \frac{1}{2}\theta) S(\omega - \frac{1}{2}\theta) e^{j\tau \omega} d\omega \quad (10)$$

que nuevamente corresponde a $A_s(\theta, \tau)$ [3].

Desde su derivación inicial por Ville y Moyal [22] la función de ambigüedad $A_s(\theta, \tau)$ ha jugado un papel importante en el análisis de señales. La función de ambigüedad puede ser interpretada como una función de correlación tiempo-frecuencia conjunta. Específicamente, $A_s(\theta, \tau)$ satisface las propiedades marginales de correlación al evaluarla a lo largo de sus variables independientes, i.e. $A_s(0, \tau)$ implica la correlación en el dominio del tiempo y $A_s(\theta, 0)$ corresponde a la correlación en el dominio de la frecuencia. Además, el valor máximo de la función de ambigüedad ocurre en el origen de sus coordenadas y corresponde a la energía de la señal, i.e. $|A_s(\theta, \tau)| \leq A_s(0, 0) = \int |s(t)|^2 dt$ [3]. A la magnitud al cuadrado de la función de ambigüedad, $|A_s(\theta, \tau)|^2$, se le denomina la superficie de ambigüedad. Sin embargo, es común que a su complejo conjugado $A_s^*(\theta, \tau)$, o su magnitud $|A_s(\theta, \tau)|$, o inclusive a $|A_s(\theta, \tau)|^2$ se les asigne el nombre de función de ambigüedad [17].

Finalmente, mediante la comparación directa entre la función característica de la distribución de Wigner $M_{WVD}(\theta, \tau)$ y la función característica de la clase general, o función de ambigüedad generalizada, $M(\theta, \tau) = \phi(\theta, \tau)A_s(\theta, \tau)$, se observa que la función kernel de la distribución de Wigner-Ville es [1]:

$$\phi_{WVD}(\theta, \tau) = 1 \quad (11)$$

A continuación se establecen y analizan algunas de las propiedades de la distribución de Wigner-Ville que son importantes en el desarrollo del presente trabajo.

IV.2.1.2. Propiedades

De entre las distribuciones tiempo-frecuencia, la distribución de Wigner-Ville resulta una de las más interesantes debido a que el entendimiento de sus propiedades ha significado un gran avance en el desarrollo del campo y a que satisface una gran cantidad de propiedades deseables en una representación tiempo-frecuencia.

1. Evaluada en lo reales.

La distribución de Wigner siempre es real, incluso si la señal es compleja. Es decir, la distribución de Wigner es igual a su complejo conjugado [1]:

$$\begin{aligned}
WVD^*(t, \omega) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} s(t - \frac{1}{2}\tau) s^*(t + \frac{1}{2}\tau) e^{j\tau\omega} d\tau \\
WVD^*(t, \omega) &= -\frac{1}{2\pi} \int_{\infty}^{-\infty} s(t + \frac{1}{2}\tau) s^*(t - \frac{1}{2}\tau) e^{-j\tau\omega} d\tau \\
WVD^*(t, \omega) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} s(t + \frac{1}{2}\tau) s^*(t - \frac{1}{2}\tau) e^{-j\tau\omega} d\tau \\
WVD^*(t, \omega) &= WVD(t, \omega)
\end{aligned} \tag{12}$$

El hecho de que la distribución de Wigner sea real para cualquier señal, también puede deducirse a partir de la función característica. En este caso, la condición que se debe cumplir para que una distribución sea real es [1]:

$$M^*(-\theta, -\tau) = M(\theta, \tau) \tag{13}$$

donde la función característica para la distribución de Wigner es la función de ambigüedad de la señal $A_s(\theta, \tau)$, la cual satisface esta propiedad.

De forma equivalente, una condición suficiente y necesaria para que la distribución sea real consiste en que su función kernel satisfaga [17]:

$$\phi^*(-\theta, -\tau) = \phi(\theta, \tau) \tag{14}$$

que resulta válido para la función kernel constante de la distribución de Wigner.

2. Marginales.

De forma directa puede verificarse que la marginal en el tiempo se cumple integrando la distribución de Wigner sobre la frecuencia:

$$\begin{aligned}
P_{WVD}(t) &= \int WVD(t, \omega) d\omega = \frac{1}{2\pi} \iint s^*(t - \frac{1}{2}\tau) s(t + \frac{1}{2}\tau) e^{-j\tau\omega} d\tau d\omega \\
P_{WVD}(t) &= \int s^*(t - \frac{1}{2}\tau) s(t + \frac{1}{2}\tau) \delta(\tau) d\tau \\
P_{WVD}(t) &= |s(t)|^2
\end{aligned} \tag{15}$$

De forma similar, se puede verificar que la marginal en la frecuencia corresponde a:

$$P_{WVD}(\omega) = |S(\omega)|^2 \tag{16}$$

Por otra parte, ambas ecuaciones pueden verificarse considerando las marginales de la función característica de la distribución de Wigner. Para la marginal en el tiempo:

$$M_{WVD}(\theta, 0) = \int |s(t)|^2 e^{j\theta t} dt \tag{17}$$

y para la marginal en la frecuencia:

$$M_{WVD}(0, \tau) = \int |S(\omega)|^2 e^{j\tau\omega} d\omega \quad (18)$$

y por lo tanto, la propiedad asociada con las marginales, es decir:

$$\int WVD(t, \omega) d\omega = |s(t)|^2$$

$$\int WVD(t, \omega) dt = |S(\omega)|^2$$

se satisface [1].

En términos de la función kernel se debe cumplir [17]:

$$\phi(\theta, 0) = 1 \quad (19)$$

$$\phi(0, \theta) = 1 \quad (20)$$

para la marginal en el tiempo y en la frecuencia, respectivamente.

Nuevamente, para el caso de la función kernel constante ambos requerimientos se cumplen.

3. Conservación de la energía.

Debido a que los marginales de las ecuaciones (15) y (16) se satisfacen, la condición de energía total se satisface automáticamente [1]:

$$E_s = \iint WVD_s(t, \omega) d\omega dt = \int |s(t)|^2 d\tau = 1 \quad (21)$$

Esta condición equivale a tener un kernel que cumpla [17]:

$$\phi(0, 0) = 1 \quad (22)$$

que corresponde a la condición de normalización. Esta condición es más débil que las condiciones marginales del kernel, dado que puede existir una distribución conjunta cuya energía total sea la energía de la señal, pero que no satisfaga las marginales, e.g. el espectrograma.

4. No positividad.

Una función que satisface los marginales no es positiva a lo largo de todo el plano tiempo-frecuencia; ésta será negativa en algún lugar. En la figura IV.2, se muestra un ejemplo de las negatividades que pueden existir en la distribución WVD. Entonces, dado que la distribución de Wigner satisface las marginales, es de esperarse que siempre posea regiones de valores negativos para cualquier señal. Este será siempre el caso excepto para señales chirp con modulación Gaussiana en amplitud de la forma

$s(t) = \left(\frac{\alpha}{\pi}\right)^{\frac{1}{4}} e^{-\frac{\alpha}{2}t^2 + j(\frac{1}{2}\beta t + \omega_0)t}$ [23]. La razón es que la distribución de Wigner para esta señal no es realmente bilineal [1], figura IV.3.

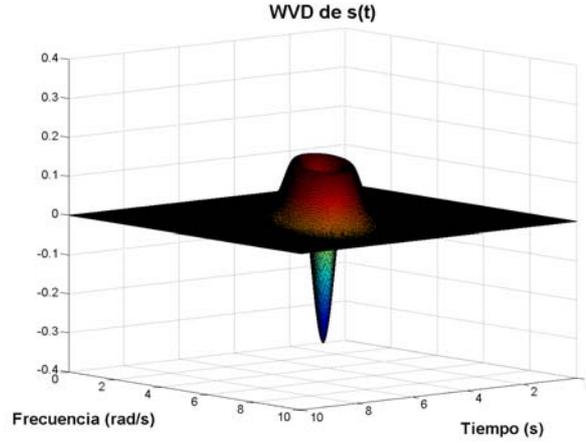


Fig. IV.2. En general, la distribución de Wigner de una señal es negativa en algún lugar del plano tiempo-frecuencia para cada señal. En este caso la señal corresponde a

$$s(t) = \left(\frac{4\alpha^3}{\pi}\right)^{\frac{1}{4}} (t-t_0) e^{-\frac{\alpha}{2}(t-t_0)^2 + j(\frac{1}{2}\beta(t-t_0) + \omega_0)(t-t_0)}, \quad \text{donde } \omega_0 = 5 \text{ rad/s}, \quad \alpha = 1, \\ \beta = 1 \text{ rad/s}^2, \text{ y } t_0 = 5 \text{ s}.$$

5. Perfecta localización de señales chirp lineales puras.

Para el caso de una chirp lineal sin modulación en amplitud se tiene que la distribución de Wigner está concentrada alrededor de su frecuencia instantánea $\omega_i(t) = \beta t + \omega_0$, es decir [1]:

$$\begin{aligned} \text{Si} \quad & s(t) = e^{j(\frac{1}{2}\beta t + \omega_0)t} \\ \text{Entonces} \quad & WVD_s(t, \omega) = \delta(\omega - \beta - \omega_0) \end{aligned} \quad (23)$$

6. Simetría.

Para señales cuyos espectros son simétricos, i.e. señales reales, la distribución de Wigner es simétrica en el dominio de la frecuencia [1], es decir:

$$\begin{aligned} \text{Si} \quad & S(\omega) = S(-\omega) \\ \text{entonces} \quad & WVD_s(t, \omega) = WVD_s(t, -\omega) \end{aligned} \quad (24)$$

De forma similar, para señales simétricas en el dominio del tiempo, i.e. espectros reales, la distribución de Wigner es simétrica en el dominio del tiempo:

$$\begin{aligned}
 \text{Si} & \quad s(t) = s(-t) \\
 \text{entonces} & \quad WVD_s(t, \omega) = WVD_s(-t, \omega)
 \end{aligned} \tag{25}$$

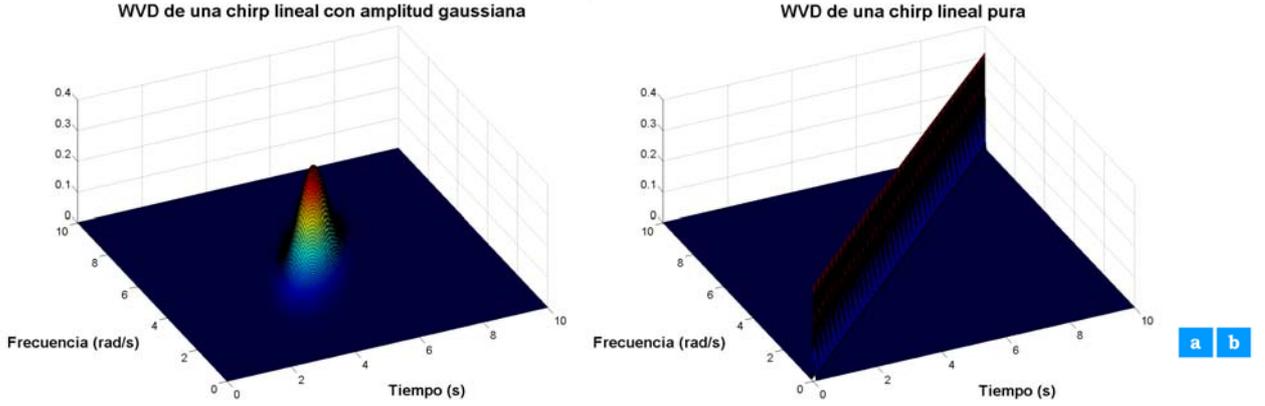


Fig. IV.3. Distribución de Wigner de señales chirp lineales con modulaciones Gaussianas. La señal $s(t) = \left(\frac{\alpha}{\pi}\right)^{\frac{1}{4}} e^{-\frac{\alpha}{2}(t-t_0)^2 + j(\frac{1}{2}\beta(t-t_0) + \omega_0)(t-t_0)}$ es la más general para la cual la distribución de Wigner es positiva a lo largo de todo el plano tiempo-frecuencia. Los parámetros de las señales en ambos casos son $\omega_0 = 5 \text{ rad/s}$, $\beta = 1 \text{ rad/s}^2$, y $t_0 = 5 \text{ s}$. a) $\alpha = 1$. b) $\alpha = 0.001$. Conforme la envolvente se vuelve más constante la WVD se concentra más alrededor de la frecuencia instantánea de la señal, $\omega_i(t) = \beta(t - t_0) + \omega_0$.

7. Corrimientos en el tiempo y en la frecuencia.

Si se tiene una señal $x(t)$ desplazada en el tiempo por t_0 y su espectro se desplaza por ω_0 , entonces la distribución de Wigner preservara dichos corrimientos [1]:

$$\text{Si} \quad y(t) = e^{j\omega_0 t} x(t - t_0) \tag{26}$$

entonces

$$\begin{aligned}
 WVD_y(t, \omega) &= \frac{1}{2\pi} \int e^{-j\omega_0(t-\frac{\tau}{2})} x^*(t - t_0 - \frac{1}{2}\tau) e^{j\omega_0(t+\frac{\tau}{2})} x(t - t_0 + \frac{1}{2}\tau) e^{-j\tau\omega} d\tau \\
 WVD_y(t, \omega) &= \frac{1}{2\pi} \int x^*(t - t_0 - \frac{1}{2}\tau) x(t - t_0 + \frac{1}{2}\tau) e^{-j\tau(\omega - \omega_0)} d\tau \\
 WVD_y(t, \omega) &= WVD_x(t - t_0, \omega - \omega_0)
 \end{aligned} \tag{27}$$

Es decir, la distribución de Wigner es invariante a corrimientos en el tiempo y en la frecuencia [3].

8. Dilataciones.

La distribución de Wigner también preserva dilataciones [3]

$$\begin{aligned} \text{Si} \quad & y(t) = \sqrt{k}x(kt) \quad , \quad k > 0 \\ \text{entonces} \quad & WVD_y(t, \omega) = WVD_x\left(kt, \frac{\omega}{k}\right) \end{aligned} \quad (28)$$

9. Convolución en el tiempo y en la frecuencia.

Esta propiedad indica que la distribución de Wigner es compatible con filtrados y modulaciones [3]. En el primer caso, si la señal $y(t)$ corresponde a la salida de un filtro con una respuesta al impulso $h(t)$ cuya entrada es $x(t)$, entonces la WVD de $y(t)$ corresponde a la convolución temporal entre la WVD de $h(t)$ y la WVD de $x(t)$ [17] de la forma:

$$\begin{aligned} \text{Si} \quad & y(t) = h(t) * x(t) = \int h(t-t')x(t')dt' \\ \text{entonces} \quad & WVD_y(t, \omega) = \int WVD_h(t-t', \omega)WVD_x(t', \omega)dt' \end{aligned} \quad (29)$$

En el segundo caso, si la señal $y(t)$ corresponde a la modulación de una señal $x(t)$ por una función $m(t)$, i.e. a la convolución en la frecuencia, entonces la WVD de $y(t)$ corresponde a la convolución en la frecuencia entre la WVD de $m(t)$ y la WVD de $x(t)$ [17]:

$$\begin{aligned} \text{Si} \quad & y(t) = m(t)x(t) \\ \text{entonces} \quad & WVD_y(t, \omega) = \int WVD_m(t, \omega - \omega')WVD_x(t, \omega')d\omega' \end{aligned} \quad (30)$$

10. La frecuencia instantánea y el retardo de grupo.

Dado que la distribución de Wigner satisface las marginales, el primer momento en la frecuencia (centro de gravedad) es:

$$\langle \omega \rangle_t = \frac{1}{|s(t)|^2} \int \omega WVD(t, \omega) d\omega = \varphi'(t) \quad (31)$$

mientras que el primer momento en el tiempo corresponde a:

$$\langle t \rangle_\omega = \frac{1}{|S(\omega)|^2} \int t WVD(t, \omega) dt = -\psi'(\omega) \quad (32)$$

Entonces, la frecuencia instantánea $\varphi'(t)$ y el retardo de grupo $\psi'(\omega)$ de la señal analítica asociada a $s(t)$ pueden obtenerse a partir de la distribución de Wigner [1,17]. Considerando que $\langle \omega \rangle_t$ es un promedio, resulta interesante preguntarse por la dispersión alrededor de esa media:

$$\sigma_{\omega|t}^2 = \langle \omega^2 \rangle_t - \langle \omega \rangle_t^2 \quad (33)$$

donde

$$\langle \omega^2 \rangle_t = \frac{1}{|s(t)|^2} \int \omega^2 WVD(t, \omega) d\omega \quad (34)$$

Sin embargo, $\sigma_{\omega|t}^2$ puede ser negativa, resultando en una desviación estándar condicional $\sigma_{\omega|t}$ imaginaria, difícil de interpretar. Por lo tanto, mientras la distribución de Wigner provee un excelente resultado para la frecuencia promedio, ofrece un resultado muy pobre para la dispersión de esa frecuencia [1].

11. Soporte finito en el sentido amplio o débil.

Para una señal de duración infinita, la distribución de Wigner es diferente de cero para todo tiempo debido a que el pliegue de la parte derecha con la parte izquierda de la señal genera valores diferentes de cero, sin importar el tiempo que se elija. En contraste, para una señal de duración finita en el intervalo de t_1 a t_2 , la distribución de Wigner será cero antes de que la señal inicie y después de que termine, es decir, el traslape existe solamente en el intervalo (t_1, t_2) , es decir:

$$\begin{aligned} \text{Si} & \quad s(t) \text{ es cero fuera de } (t_1, t_2) \\ \text{entonces} & \quad WVD_s(t, \omega) = 0 \quad \text{para } t \text{ fuera de } (t_1, t_2) \end{aligned} \quad (35)$$

Comentarios similares se pueden realizar para una señal de banda limitada, ver figura IV.4.a. En consecuencia, la distribución de Wigner satisface las propiedades de soporte finito en el tiempo y en la frecuencia [1, 17].

Para el caso de señales intermitentes, la distribución de Wigner generalmente no es cero en un tiempo para el cual la señal es cero, debido a que al plegar las partes izquierda y derecha de la señal se obtiene una respuesta distinta de cero, ver figura IV.4.b. Por lo tanto, la distribución de Wigner no es necesariamente cero para tiempos donde la señal es cero y no es necesariamente cero para frecuencias que no existen en el espectro, originando dificultades en la interpretación de la distribución de Wigner.

Manifestaciones de este fenómeno han sido llamadas interferencias o términos cruzados y han sido erróneamente atribuidas a la estructura bilineal en la señal de la distribución de Wigner [1]. Los términos cruzados pueden ser minimizados mediante un suavizado, pero esta operación destruye algunas otras propiedades deseables de la distribución de Wigner [17].

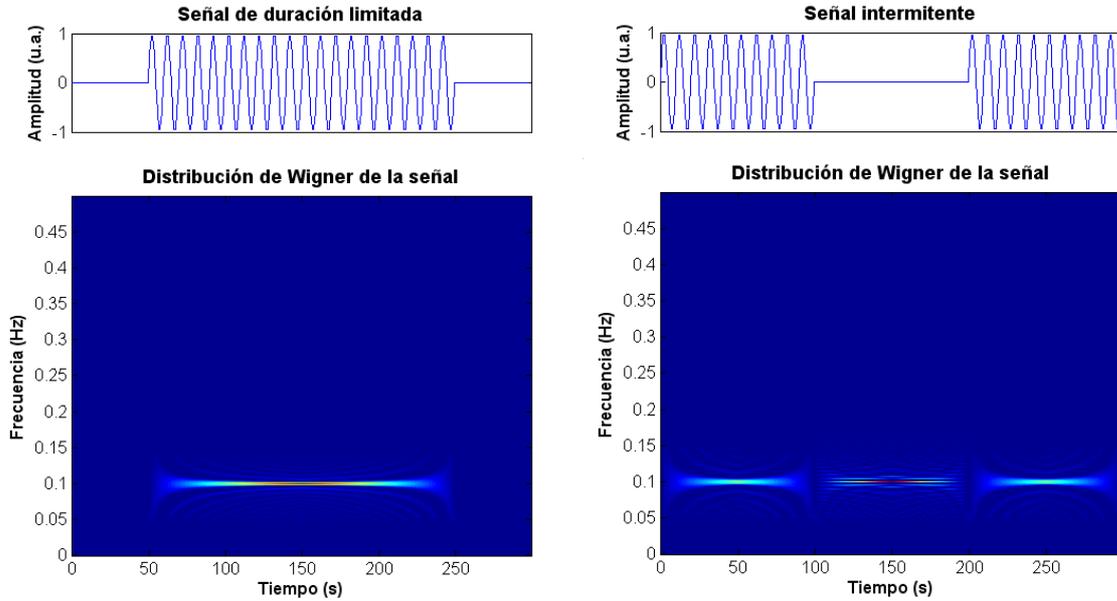


Fig. IV.4. Distribución de Wigner-Ville para: a) señal de duración finita, y b) señal intermitente que origina términos cruzados.

12. Términos cruzados.

Una propiedad importante de la distribución de Wigner es que la suma de dos señales no es la suma de las distribuciones de Wigner de cada señal [1]:

$$\begin{aligned}
 \text{Si} \quad & y(t) = x_1(t) + x_2(t) \\
 \text{Entonces} \quad & WVD_y(t, \omega) = WVD_{x_1}(t, \omega) + WVD_{x_2}(t, \omega) \\
 & \quad \quad \quad + WVD_{x_1, x_2}(t, \omega) + WVD_{x_2, x_1}(t, \omega)
 \end{aligned} \tag{36}$$

donde

$$WVD_{x_1, x_2}(t, \omega) = \frac{1}{2\pi} \int x_1^*(t - \frac{1}{2}\tau) x_2(t + \frac{1}{2}\tau) e^{-j\tau\omega} d\tau \tag{37}$$

es denominada la distribución de Wigner cruzada de $x_1(t)$ y $x_2(t)$, y en términos de los espectros de las señales corresponde a [1]:

$$WVD_{x_1, x_2}(t, \omega) = \frac{1}{2\pi} \int S_1^*(\omega + \frac{1}{2}\theta) S_2(\omega - \frac{1}{2}\theta) e^{-j\theta t} d\theta \tag{38}$$

La distribución de Wigner cruzada es compleja y por lo tanto se cumple:

$$WVD_{x_1, x_2}(t, \omega) = WVD_{x_2, x_1}^*(t, \omega) \quad (39)$$

y en consecuencia:

$$WVD_y(t, \omega) = WVD_{x_1}(t, \omega) + WVD_{x_2}(t, \omega) + 2\Re\{WVD_{x_1, x_2}(t, \omega)\} \quad (40)$$

Los dos primeros términos se denominan auto-componentes, mientras que el término restante es generalmente denominado término cruzado [1]. Dado que las señales reales, generalmente se consideran constituidas por la suma de varias señales individuales, es importante establecer el comportamiento de la distribución de Wigner para este caso. De forma general, si la señal $y(t)$ esta compuesta por la suma de N componentes:

$$y(t) = \sum_{i=1}^N x_i(t) \quad (41)$$

entonces, su distribución de Wigner poseerá N auto-componentes y $\frac{1}{2}N(N-1)$ términos cruzados, i.e., un término cruzado para cada par de auto-términos, ver figura IV.5, y estará dada por [24]:

$$WVD_y(t, \omega) = \sum_{i=1}^N WVD_{x_i}(t, \omega) + \sum_{k=1}^{N-1} \sum_{l=k+1}^N 2\Re\{WVD_{x_k, x_l}(t, \omega)\} \quad (42)$$

Algunas de las propiedades conocidas de los términos cruzados son [24]:

- 1) El término cruzado $WVD_{x_k, x_l}(t, \omega)$ ocurre en la distancia media entre dos auto-componentes en el espacio tiempo-frecuencia, i.e., si las coordenadas tiempo-frecuencia centrales de los auto-términos $WVD_{x_k}(t, \omega)$ y $WVD_{x_l}(t, \omega)$ son (t_{x_k}, ω_{x_k}) y (t_{x_l}, ω_{x_l}) , respectivamente, el término cruzado $WVD_{x_k, x_l}(t, \omega)$ ocurre en $\left(\frac{t_{x_k} + t_{x_l}}{2}, \frac{\omega_{x_k} + \omega_{x_l}}{2}\right)$.
- 2) Los términos cruzados son de naturaleza oscilatoria, y oscilan en la dirección ortogonal a la línea que conecta a los auto-términos $WVD_{x_k}(t, \omega)$ y $WVD_{x_l}(t, \omega)$ con una frecuencia de oscilación proporcional a la distancia en el plano tiempo-frecuencia entre los dos auto-componentes.

- 3) Los términos cruzados pueden poseer una amplitud igual al doble del producto de las magnitudes de $WVD_{x_k}(t, \omega)$ y $WVD_{x_l}(t, \omega)$. Por lo tanto, la presencia de los términos cruzados dificulta la interpretación de la distribución de Wigner.

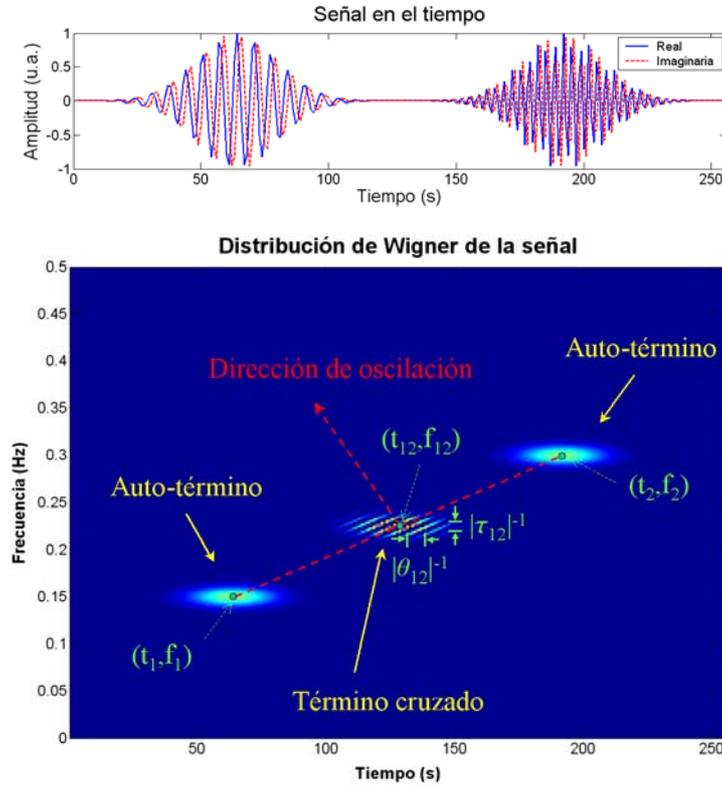


Fig. IV.5. Geometría de las interferencias de la distribución Wigner-Ville. a) Señal analítica compuesta por la suma de dos señales Gaussianas centradas en el plano tiempo-frecuencia alrededor de los puntos (t_1, f_1) y (t_2, f_2) . b) WVD de la señal compuesta por los auto-términos ubicados alrededor de (t_1, f_1) y (t_2, f_2) y el término cruzado localizado alrededor del punto medio entre ambos auto-términos, i.e., $(t_{12} = \frac{1}{2}(t_1 + t_2), f_{12} = \frac{1}{2}(f_1 + f_2))$. El término cruzado oscila perpendicularmente a la línea que une los auto-términos con un periodo de oscilación en el tiempo de $|\theta_{12}|^{-1} = |f_1 - f_2|^{-1}$ y con un periodo de oscilación en la frecuencia de $|\tau_{12}|^{-1} = |t_1 - t_2|^{-1}$.

En el plano de la función de ambigüedad, los auto-términos están principalmente localizados alrededor del origen mientras que los términos cruzados se encuentran alejados de éste, en una relación directa con la distancia entre los auto-componentes involucrados [25], ver figura IV.6.a. En consecuencia, dado que la función de ambigüedad generalizada de la distribución de Wigner es la función de ambigüedad en

sí, i.e. $\phi_{WVD}(\theta, \tau) = 1$, la distribución de Wigner pondera por igual los términos cruzados que los auto-términos y por lo tanto, esta distribución presenta la mayor cantidad de términos cruzados de entre las distintas distribuciones tiempo-frecuencia, ver figura IV.6.b.

Aunque los términos cruzados pueden suprimirse mediante la elección adecuada de un kernel que opere como un filtro pasa-bajas en el plano de la función de ambigüedad [26], si se desea interpretar una representación tiempo-frecuencia cuadrática como una distribución tiempo-frecuencia de la energía de la señal, donde sus marginales respectivos son igual a las densidades de energía unidimensionales, entonces dicha representación tiempo-frecuencia debe contener los términos cruzados [3].

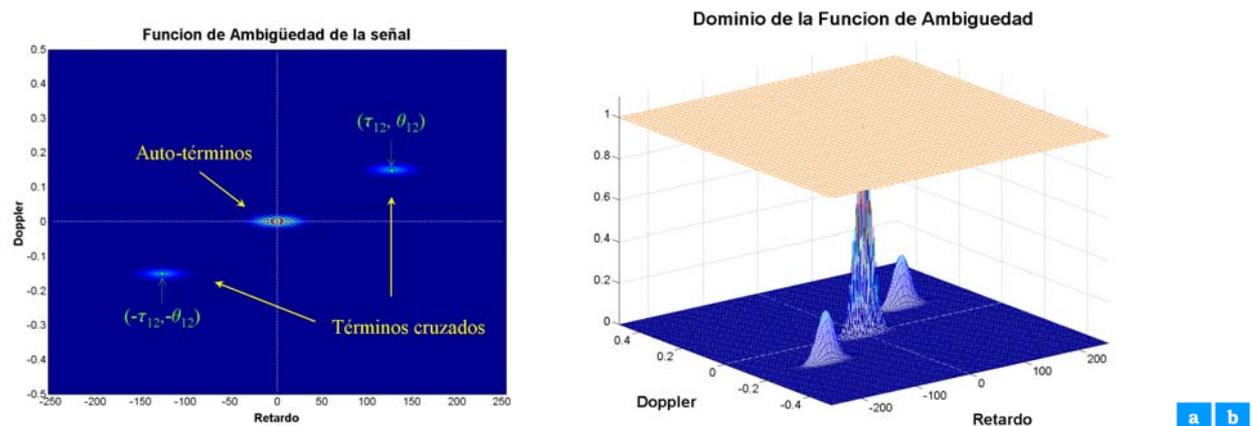


Fig. IV.6. Geometría de las interferencias en la función de ambigüedad, donde la señal considerada corresponde a la señal en la figura IV.4.a . a) Los auto-términos se localizan alrededor del origen del plano de la función de ambigüedad mientras que el término cruzado consiste de dos sub-términos localizados alrededor de los puntos $(-\tau_{12}, -\theta_{12})$ y (τ_{12}, θ_{12}) , respectivamente, donde τ_{12} y θ_{12} corresponden a la diferencia entre los centros en la frecuencia y en el tiempo de los términos de la señal, i.e., $\tau_{12} = |t_1 - t_2|$ y $\theta_{12} = |f_1 - f_2|$, y b) Ponderación uniforme de los auto-términos y los términos cruzados debido al kernel constante en el dominio de la función de ambigüedad.

IV.2.1.3. Conceptos adicionales.

Señal monocomponente

En el dominio tiempo-frecuencia, un concepto fundamental es el asociado con señales monocomponente y multicomponente. Si una señal está bien delineada en una región

tiempo-frecuencia, se le denomina monocomponente y si la señal posee más de una región bien definida se le denomina multicomponente. En otras palabras, la frase “señal multicomponente” alude a una señal bien delineada en regiones del plano tiempo-frecuencia y no simplemente a una señal expresada como la arbitraria suma de partes.

Precisamente, uno de los logros del análisis tiempo-frecuencia consiste en que la distribución de Wigner, u otra, indica si la señal es o no multicomponente [1].

Señal analítica.

Existen tres argumentos para emplear la señal analítica y no la señal real en la obtención de la distribución de Wigner, ver figura IV.7:

1. Si se emplea la señal real, que posee un espectro tanto en la parte positiva como negativa del eje de frecuencias, existirán términos cruzados entre ambos semiplanos tiempo-frecuencia. Por lo tanto, dado que la señal analítica suprime las frecuencias negativas, su distribución de Wigner será cero para dichas frecuencias y esto elimina los términos cruzados entre ambos semiplanos tiempo-frecuencia [20, 21]. Sin embargo, sin importar si se emplea la señal real o su contraparte analítica, existirán términos cruzados entre las frecuencias positivas [17].
2. Generalmente se desea que el primer momento condicional de la distribución corresponda a la derivada de la fase, i.e. la frecuencia instantánea. Por lo tanto, debe emplearse la señal analítica dado que la frecuencia instantánea se define en términos de ésta [17].
3. La señal analítica elimina la necesidad de emplear una frecuencia de muestreo igual al doble de la frecuencia de muestro de Nyquist [20]. Es decir, al emplear la señal analítica la ausencia de un espectro en las frecuencias negativas elimina el problema de *aliasing* que pudiera ocurrir en la distribución al muestrear la señal real a la frecuencia de Nyquist [27].

IV.2.1.4. Comentario final.

Históricamente el principal motivo para modificar la distribución de Wigner fue el intento de lograr una distribución positiva. Una forma de aproximarse al problema consiste en intentar suavizar la distribución de Wigner mediante su convolución con una función de suavizado para obtener una distribución de Wigner suavizada [1]. Una distribución que ha sido ampliamente estudiada es la distribución Choi-Williams, la cual puede obtenerse

mediante la convolución de la distribución de Wigner con cierta función. Sin embargo, sería erróneo pensar en la distribución Choi-Williams como una distribución de Wigner suavizada, dado que también es válido que la distribución de Wigner puede ser obtenida de la Choi-Williams mediante una convolución y se podría argumentar que la distribución de Wigner es una versión suavizada de la distribución Choi-Williams [1]. Otra distribución interesante tiempo-frecuencia que puede obtenerse mediante la convolución de la distribución de Wigner con otra función es el espectrograma.

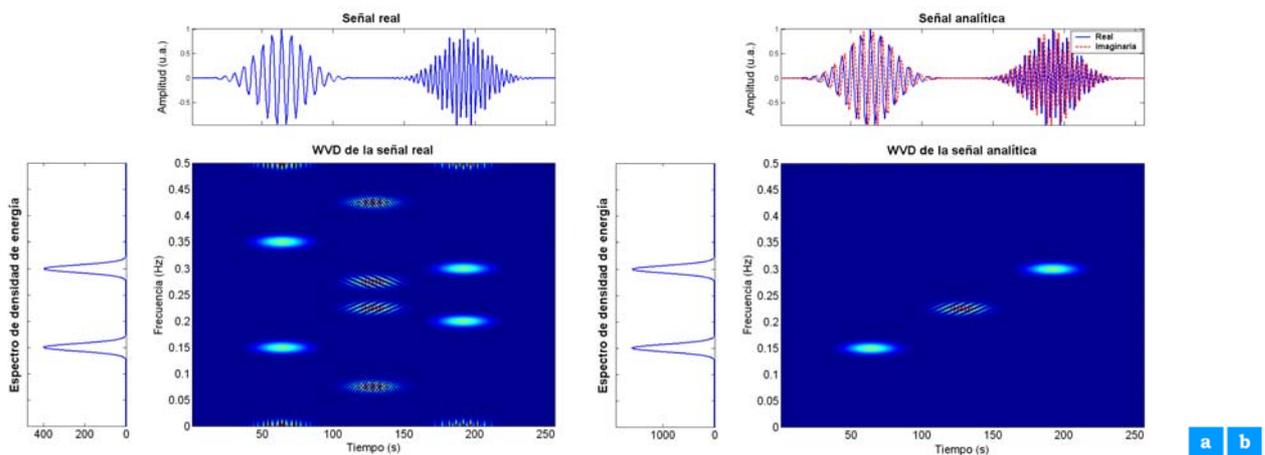


Fig. IV.7. Distribución de Wigner-Ville de una señal compuesta por dos señales Gaussianas elementales, cada una centrada en el plano tiempo-frecuencia en (64 seg, 0.15 Hz) y (192 seg, 0.30 Hz). a) WVD empleando la señal real donde aparece aliasing espectral así como términos cruzados adicionales, y b) WVD empleando la señal analítica. Como se observa, no importa cual señal se emplee existen términos cruzados.

IV.2.2. El Espectrograma.

En 1946 Dennis Gabor [5] sentó las bases para el análisis tiempo-frecuencia que subsecuentemente, en la década de 1980s, impactaron en el interés por el análisis basado en funciones wavelets. El punto inicial consistió en notar que la transformada de Fourier es solamente una abstracción teórica, la cual implica formas de onda de duración infinita. En la nueva representación de Gabor se considera que la señal es una función tanto del tiempo como de la frecuencia y se emplea solamente frecuencias positivas mediante la introducción de una señal compleja cuya parte real es la señal física bajo análisis; por los mismos años Ville [7] se enfocó a la teoría de señales complejas o analíticas. En forma general, Gabor [28] estableció que es posible representar a una señal de infinitas maneras, mediante una combinación lineal de ciertas funciones elementales asociadas con celdas individuales. Gabor

dividió el medio plano tiempo-frecuencia en celdas de área unitaria, $T\Omega = 1$, donde T es la duración efectiva de la celda y Ω es su ancho de banda, ver figura IV.8. Además, limitó las funciones elementales a señales armónicas moduladas con una señal Gaussiana, las cuales además de poseer la misma morfología en el dominio de Fourier su producto entre la duración efectiva T y el ancho de banda efectivo Ω es el mínimo posible; otro resultado clásico debido a Gabor [29].

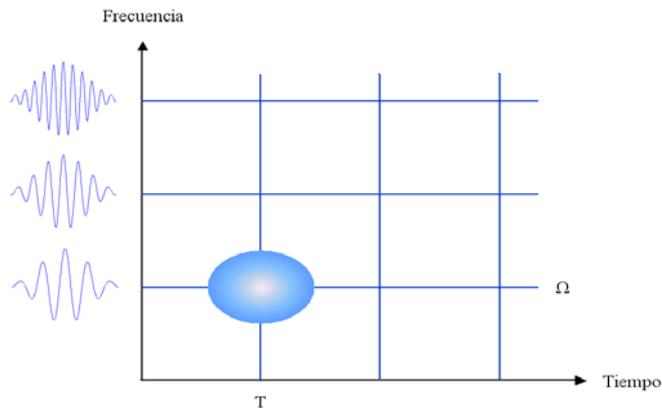


Fig. IV.8. Representación tiempo-frecuencia de Gabor. El medio plano tiempo-frecuencia es dividido en celdas de área unitaria, a cada una de las cuales se le asocia una señal elemental.

Alrededor del tiempo de introducción de la representación o expansión de Gabor apareció otra aproximación: la transformada de Fourier de tiempo corto, cuyas ideas fueron cristalizadas al término de la segunda guerra mundial con el desarrollo del espectrógrafo [8]. Actualmente, la transformada de Fourier de tiempo corto es el método más ampliamente usado para estudiar señales no estacionarias. El concepto es simple pero útil, suponga que escucha una pieza musical con una hora de duración y que al inicio hay violines y al final tambores. Al aplicar el análisis de Fourier sobre toda la hora, el espectro de energía mostrará picos en las frecuencias correspondientes a los violines y tambores pero no dará ninguna indicación de cuando fueron tocados estos instrumentos. Lo más sencillo que puede hacerse es dividir la hora en segmentos y aplicar el análisis de Fourier sobre cada intervalo. En el espectro de cada segmento es posible observar en que intervalos de tiempo se tocaron los violines y los tambores; evidentemente, la interpretación depende de la longitud de cada segmento. La transformada de Fourier de tiempo corto permite observar la variación del espectro con respecto del tiempo [1]. Sin embargo, surge una pregunta interesante: ¿es posible considerar intervalos de tiempo tan cortos como se desee?, en otras palabras ¿puede continuarse con el proceso de ruptura de la señal hasta lograr una localización temporal

siempre más fina? La respuesta es no; después de ciertas reducciones el espectro carecerá de sentido y no mostrará relación alguna con el espectro de la señal original. Lo anterior no debe atribuirse a alguna limitación fundamental, sino más a una limitación de la técnica que produce señales de corta duración con el propósito de estimar el espectro. El problema anterior se asocia con el principio de incertidumbre que se analiza más adelante. También debe tenerse en cuenta que la transformada de Fourier de tiempo corto mezcla las propiedades de la señal con las propiedades de la función ventana que se utiliza para segmentarla.

A pesar de las dificultades anteriores, el método de la transformada de Fourier de tiempo corto es ideal en muchos aspectos. La transformada está bien definida, basada sobre principios físicos razonables, y para muchas señales y situaciones brinda una excelente estructura tiempo-frecuencia consistente con la intuición. Sin embargo, para ciertas situaciones no es el mejor método disponible en el sentido de que no siempre otorga una imagen lo más clara posible de lo que está ocurriendo [1].

IV.2.2.1. Formulación de la Transformada de Fourier de tiempo corto.

Dado que la duración de cada bloque es corto en comparación con la señal completa a analizar, el proceso se denomina transformada de Fourier de tiempo corto, STFT [21]. Para estudiar las propiedades de la señal al tiempo t , se enfatiza la señal en ese tiempo y se suprime la señal original en otros tiempos mediante la multiplicación por una función ventana, $h(t)$, centrada en t , para producir la señal modificada [1, 21]:

$$s_t(\tau) = s(\tau)h(\tau - t) \quad (43)$$

donde la señal modificada es una función de dos tiempos, el tiempo fijo en el cual estamos interesados, t , y el tiempo τ . La función ventana, ver figura IV.9, permite que la señal modificada cumpla con lo siguiente:

$$s_t(\tau) \approx \begin{cases} s(\tau) & \text{para } \tau \text{ cercanos a } t \\ 0 & \text{para } \tau \text{ lejanos a } t \end{cases} \quad (44)$$

Finalmente, debido a que la señal modificada enfatiza la señal original alrededor del tiempo t , su transformada de Fourier reflejará la distribución de la frecuencia alrededor de este tiempo [1, 21, 30]:

$$\begin{aligned} S_t(\omega) &= \frac{1}{\sqrt{2\pi}} \int e^{-j\omega\tau} s_t(\tau) d\tau \\ S_t(\omega) &= \frac{1}{\sqrt{2\pi}} \int e^{-j\omega\tau} s(\tau)h(\tau - t) d\tau \end{aligned} \quad (45)$$

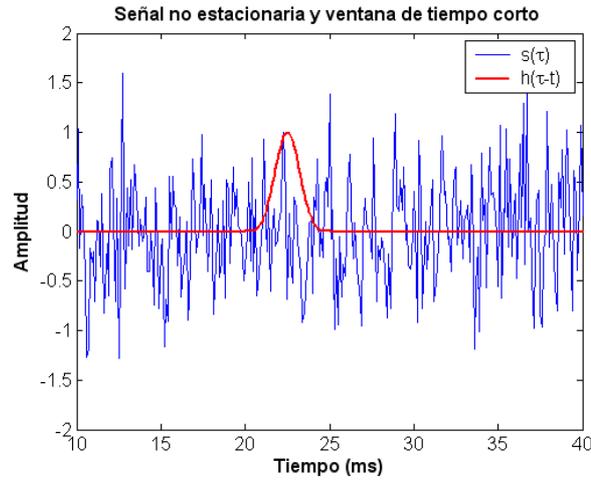


Fig. IV.9. Ejemplo de señal no estacionaria $s(\tau)$ y una ventana de análisis de tiempo corto $h(\tau-t)$, centrada al tiempo t .

IV.2.2.2. El Espectrograma.

El espectrograma se define como la magnitud al cuadrado de la representación lineal correspondiente a la transformada de Fourier de tiempo corto. Es decir, el correspondiente espectro de la densidad de energía al tiempo t es [1, 3, 21]:

$$SP(t, \omega) = |S_t(\omega)|^2 = \left| \frac{1}{\sqrt{2\pi}} \int e^{-j\omega\tau} s(\tau) h(\tau-t) d\tau \right|^2 \quad (46)$$

donde para cada instante de tiempo se obtiene un espectro diferente y la totalidad de ese espectro es la distribución tiempo-frecuencia, $SP(t, \omega)$. A dicha distribución se le han otorgado diversos nombres dependiendo del campo de aplicación, e.g. sonograma, siendo el más común el de espectrograma. Dependiendo de la duración de la ventana de tiempo $h(t)$, el espectrograma resultante se denomina como espectrograma de banda ancha si $h(t)$ es de corta duración, i.e. $H(\omega)$ es ancha, o espectrograma de banda angosta si $h(t)$ es de larga duración, i.e. $H(\omega)$ es angosta.

Hasta ahora, se ha discutido una primera interpretación de la STFT mediante la transformada de Fourier de la señal ventaneada. Una segunda interpretación consiste en pensar a la STFT como un análisis de banco de filtros [3, 30]. Sin embargo por ahora nos conformaremos con la primera.

IV.2.2.3. Relación con la clase general de distribuciones.

Recordando que la función característica de una distribución corresponde a su doble transformada de Fourier, la función característica del espectrograma es [1]:

$$\begin{aligned} M_{SP}(\theta, \tau) &= \iint |S_t(\omega)|^2 e^{j\theta\omega + j\tau\omega} dt d\omega \\ M_{SP}(\theta, \tau) &= A_s(\theta, \tau) A_h(-\theta, \tau) \end{aligned} \quad (47)$$

donde

$$A_s(\theta, \tau) = \int s^*(t - \frac{1}{2}\tau) s(t + \frac{1}{2}\tau) e^{j\theta t} dt \quad (48)$$

es la función de ambigüedad de la señal, y $A_h(\theta, \tau)$ es la función de ambigüedad de la ventana definida de manera idéntica para $h(t)$ en lugar de $s(t)$. Mediante la comparación directa entre la función característica del espectrograma, $M_{SP}(\theta, \tau)$, y la función característica de la clase general, $M(\theta, \tau) = \phi(\theta, \tau) A_s(\theta, \tau)$, se observa que la función kernel del espectrograma es [17]:

$$\phi_{SP}(\theta, \tau) = A_h(-\theta, \tau) = \int h^*(t - \frac{1}{2}\tau) h(t + \frac{1}{2}\tau) e^{-j\theta t} dt \quad (49)$$

que corresponde a la función de ambigüedad de la ventana con θ reemplazado por $-\theta$. El kernel también puede expresarse en términos de la transformada de Fourier de la ventana [17], de la forma:

$$\phi_{SP}(\theta, \tau) = \int H^*(\omega - \frac{1}{2}\theta) H(\omega + \frac{1}{2}\theta) e^{j\tau\omega} d\omega \quad (50)$$

IV.2.2.4. Propiedades.

Las propiedades básicas del espectrograma se presentaron en la tabla IV.II de este capítulo y a continuación se detallan algunas de ellas. En ocasiones, se utilizará el hecho de que muchas de las propiedades del espectrograma se heredan de la STFT, dada la relación entre ambas.

1. Positividad.

Esta propiedad se cumple a partir de la definición del espectrograma de la señal como la magnitud al cuadrado de la transformada de Fourier de tiempo corto o como la magnitud al cuadrado de la transformada temporal de frecuencia corta, es decir:

$$SP_s(t, \omega) = |S_t(\omega)|^2 = |s_\omega(t)|^2 \geq 0 \quad , \quad \text{para todo } (t, \omega) \quad (51)$$

2. Corrimientos en el tiempo y en la frecuencia.

La STFT preserva corrimientos en la frecuencia y corrimientos en el tiempo, por lo tanto el espectrograma también preservará corrimientos en el tiempo y en la frecuencia, esto es:

$$\begin{aligned} \text{Si} \quad & y(t) = x(t) \cdot e^{j\omega_0 t} \\ \text{entonces} \quad & STFT_y(t, \omega) = STFT_x(t, \omega - \omega_0) \\ \text{por lo tanto} \quad & SP_y(t, \omega) = SP_x(t, \omega - \omega_0) \end{aligned} \quad (52)$$

$$\begin{aligned} \text{Si} \quad & y(t) = x(t - t_0) \\ \text{entonces} \quad & STFT_y(t, \omega) = STFT_x(t - t_0, \omega) \cdot e^{j\omega t_0} \\ \text{por lo tanto} \quad & SP_y(t, \omega) = SP_x(t - t_0, \omega) \end{aligned} \quad (53)$$

Además, el espectrograma es un miembro de la clase de distribuciones de Cohen y en consecuencia es una distribución invariante al corrimiento [3].

3. Energía total.

La energía total de una distribución se obtiene integrándola sobre todo tiempo y frecuencia, o de forma equivalente evaluando su función característica en el origen del plano de la función de ambigüedad [1], es decir:

$$\begin{aligned} E_{SP} &= \iint SP_s(t, \omega) dt d\omega = M_{SP}(0,0) = A_s(0,0)A_h(0,0) \\ E_{SP} &= \iint SP_s(t, \omega) dt d\omega = \int |s(t)|^2 dt \cdot \int |h(t)|^2 dt \end{aligned} \quad (54)$$

Por lo tanto, si la ventana es de energía unitaria, la energía del espectrograma corresponde a la energía de la señal. En términos del kernel, la energía se preserva si

$$\phi_{SP}(0,0) = \int |h(t)|^2 dt = 1 \quad (55)$$

que se logra al normalizar la ventana de análisis [17].

4. Las marginales.

La marginal de tiempo es obtenida integrando la distribución sobre la frecuencia, de la forma:

$$\begin{aligned}
P(t) &= \int |S_i(\omega)|^2 d\omega \\
P(t) &= \frac{1}{2\pi} \iiint s(\tau)h(\tau-t)s^*(\tau')h^*(\tau'-t)e^{-j\omega(\tau-\tau')}d\tau d\tau' d\omega \\
P(t) &= \iint s(\tau)h(\tau-t)s^*(\tau')h^*(\tau'-t)\delta(\tau-\tau')d\tau d\tau' \\
P(t) &= \int s(\tau)h(\tau-t)s^*(\tau)h^*(\tau-t)d\tau = \int |s(\tau)|^2 |h(\tau-t)|^2 d\tau \neq |s(t)|^2
\end{aligned} \tag{56}$$

En forma similar, la marginal de frecuencia es:

$$P(\omega) = \int |S(\omega')|^2 |H(\omega-\omega')|^2 d\omega' \neq |S(\omega)|^2 \tag{57}$$

En conclusión, el espectrograma generalmente no satisface los marginales correctos, i.e. $|s(t)|^2$ y $|S(\omega)|^2$. La razón es que el espectrograma mezcla la distribución de energía de la ventana con la distribución de energía de la señal, lo que introduce efectos no relacionados con las propiedades de la señal original [1].

En términos del kernel, para satisfacer la marginal en el tiempo debe cumplirse que

$$\phi_{SP}(\theta, 0) = \int |h(t)|^2 e^{-j\theta t} dt = 1 \tag{58}$$

y para la marginal en la frecuencia

$$\phi_{SP}(0, \tau) = \int |H(\omega)|^2 e^{j\tau\omega} d\omega = 1 \tag{59}$$

Para satisfacer ambas marginales las dos cantidades anteriores deben ser iguales a uno. La única forma en que $\phi_{SP}(\theta, 0) = 1$ consiste en elegir una ventana cuyo valor cuadrático se aproxime a una función delta. Entre más cercana sea la elección a la función delta, más cercana estará la marginal en el tiempo del espectrograma a la energía instantánea. Sin embargo, para una función angosta, su transformada de Fourier será muy ancha y la marginal en la frecuencia del espectrograma estará muy alejada del espectro de la densidad de energía [17].

5. Soporte finito.

La propiedad de soporte finito en el tiempo indica que para señales de duración finita se espera que la distribución sea cero antes de que la señal inicie y después de que esta termine, ver figura IV.10. Para el caso del espectrograma esta propiedad generalmente no se cumple. Por ejemplo, si se considera un instante de tiempo t anterior al inicio de la señal $s(t)$, el espectrograma puede no ser cero en ese instante, debido a que la

señal modificada $s(\tau)h(\tau - t)$ puede no serlo ya que la ventana puede conservar parte de $s(t)$ [1].

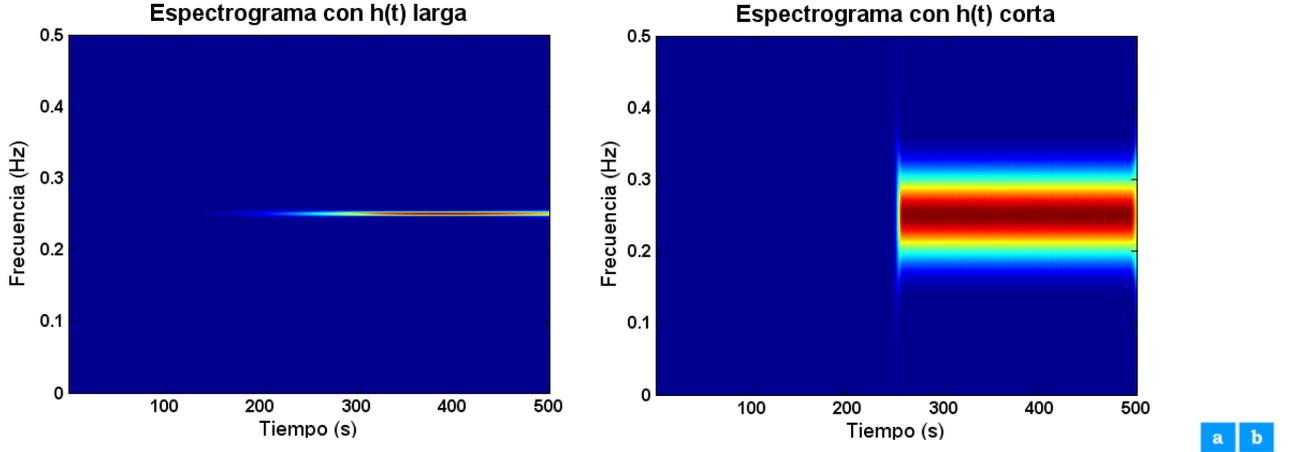


Fig. IV.10. Propiedad de soporte finito. Espectrogramas de la señal $s(t) = \sin(2\pi f_0(t - t_0))$ con $f_0 = 0.25 \text{ Hz}$ y $t_0 = 250 \text{ s}$ y utilizando una ventana rectangular con una duración t_h . a) Espectrograma de banda angosta, $t_h = 455 \text{ s}$, y b) Espectrograma de banda ancha, $t_h = 15 \text{ s}$.

6. Resolución en el tiempo y en la frecuencia.

La resolución en el tiempo de la STFT se obtiene al considerar una señal igual a un impulso de Dirac,

$$\begin{aligned} \text{Si} \quad & x(t) = \delta(t - t_0) \\ \text{entonces} \quad & STFT_x(t, \omega; h) = e^{-j\omega t_0} h(t - t_0) \end{aligned} \quad (60)$$

De manera similar, la resolución en la frecuencia de la STFT se obtiene al considerar un impulso de Dirac en el dominio de la frecuencia, i.e. una sinusoidal compleja,

$$\begin{aligned} \text{Si} \quad & x(t) = e^{j\omega_0 t} \\ \text{entonces} \quad & STFT_x(t, \omega; h) = e^{-j\omega_0 t} H(\omega - \omega_0) \end{aligned} \quad (61)$$

Se puede observar que la resolución en el tiempo del espectrograma, i.e., $|STFT|^2$, es proporcional a la duración de la ventana de análisis $h(t)$, mientras que su resolución en la frecuencia es proporcional al ancho de banda de $h(t)$. Una buena resolución en el tiempo involucra una $h(t)$ de corta duración, pero una buena resolución en la frecuencia requiere un filtro $H(\omega)$ con una banda angosta, i.e. una $h(t)$ de larga duración. Desafortunadamente, $h(t)$ y $H(\omega)$ no pueden ser arbitrariamente angostas

como consecuencia del principio de incertidumbre, o desigualdad de Heisenberg-Gabor. Así, la principal limitación del espectrograma es que existe un compromiso entre la resolución en el tiempo y la resolución en la frecuencia [1], ver figura IV.11.

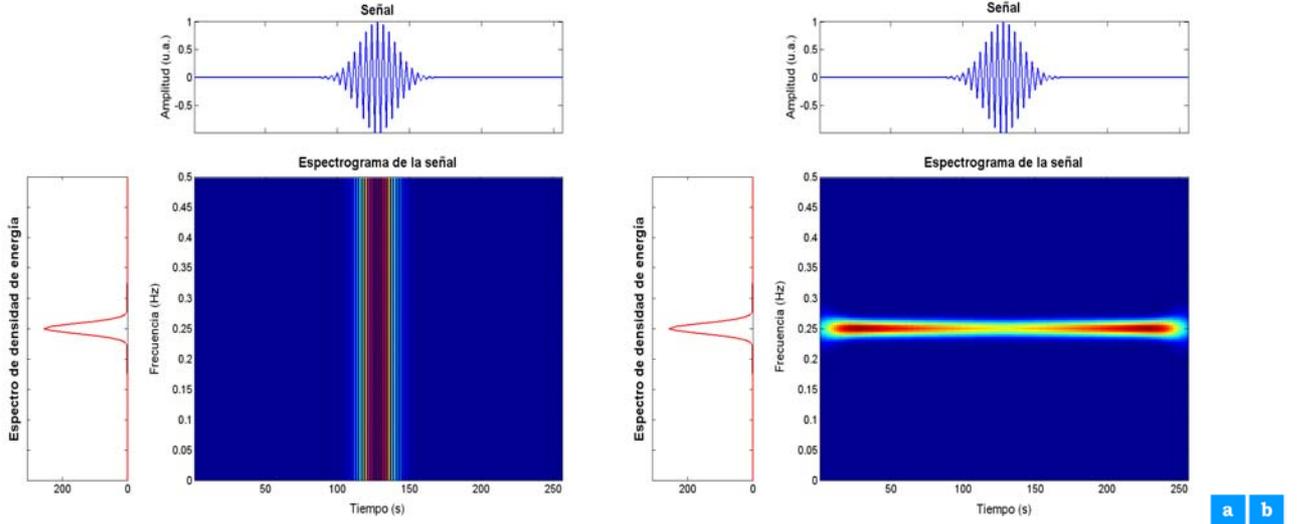


Fig. IV.11. Resolución en el tiempo y la frecuencia del espectrograma. a) Al considerar una $h(t)$ igual a un impulso de Dirac se obtiene una perfecta resolución en el tiempo pero una resolución nula en la frecuencia, y b) al considerar una $h(t)$ constante se obtiene una perfecta resolución en la frecuencia pero una resolución nula en el tiempo.

7. Estructura de las interferencias.

El espectrograma de la suma de dos señales no es la suma de los espectrogramas, es decir:

$$\begin{aligned} \text{Si} \quad & y(t) = x_1(t) + x_2(t) \\ \text{entonces} \quad & SP_y(t, \omega) = SP_{x_1}(t, \omega) + SP_{x_2}(t, \omega) + 2\Re\{SP_{x_1, x_2}(t, \omega)\} \end{aligned} \quad (62)$$

Por lo tanto, como cualquier distribución cuadrática, el espectrograma presenta términos de interferencia dados por $SP_{x_1, x_2}(t, \omega) = STFT_{x_1}(t, \omega) \cdot STFT_{x_2}^*(t, \omega)$. De forma general, si la señal $y(t)$ está compuesta por la suma de N componentes de la forma:

$$y(t) = \sum_{i=1}^n x_i(t) \quad (63)$$

entonces, el espectrograma de la señal corresponde a [24]:

$$\begin{aligned}
SP_y(t, \omega) &= \sum_{i=1}^N SP_{x_i}(t, \omega) \\
&+ 2 \sum_{k=1}^{N-1} \sum_{l=k+1}^N |STFT_{x_k}(t, \omega)| |STFT_{x_l}(t, \omega)| \cdot \cos(\varphi_{x_k}(t, \omega) - \varphi_{x_l}(t, \omega))
\end{aligned} \tag{64}$$

donde, los primeros N términos $SP_{x_i}(t, \omega)$ son los auto-términos y el segundo conjunto corresponde a los términos cruzados que están modulados por un coseno cuyo argumento es una función de la diferencia de las funciones de fase de las transformadas de Fourier de tiempo corto $STFT_{x_k}(t, \omega)$ y $STFT_{x_l}(t, \omega)$. La importancia de la ecuación anterior es la siguiente:

- 1) El espectrograma presentará términos de interferencia en aquellas regiones donde un par de auto-términos se traslapen. Por lo tanto, dependiendo de la cantidad de traslape entre cada par de auto-términos, la distribución de energía del espectrograma de la señal $y(t)$ poseerá un mínimo de cero y un máximo de $\frac{1}{2}N(N-1)$ términos cruzados o de interferencia.
- 2) Los términos de interferencia del espectrograma son oscilatorios y pueden presentar una magnitud máxima igual al producto de la magnitud de los auto-términos. Teniendo esto en mente, si los componentes de la señal se encuentran muy cercanamente espaciados se generaran problemas con la extracción de parámetros y más aún, el número de términos de interferencia puede crecer de forma cuadrática con el número de componentes de la señal, lo cual dificulta el análisis.

8. Frecuencia instantánea y retardo de grupo.

Desafortunadamente los estimados obtenidos con el espectrograma dependen de la ventana elegida. En particular, se espera que la frecuencia promedio del espectrograma definida como

$$\langle \omega \rangle_t = \frac{\int \omega |STFT(t, \omega)|^2 d\omega}{\int |STFT(t, \omega)|^2 d\omega} \tag{65}$$

corresponda a la frecuencia instantánea definida como la derivada de la fase de la señal analítica. Si se considera una ventana real, $h(t)$, que se estrecha de tal forma que $|h(t)|^2 \rightarrow \delta(t)$, $h(t) \in \mathfrak{R}$, entonces la frecuencia promedio es:

$$\langle \omega \rangle_t \rightarrow \varphi'(t) \tag{66}$$

sin embargo [17]:

$$\sigma_{\omega|t} \rightarrow \infty \quad (67)$$

Es decir, conforme se reduce la duración de la ventana se incrementa la desviación estándar y es más difícil de discernir la frecuencia instantánea, como resultado del teorema del producto tiempo-ancho de banda [1]. Finalmente, todas las puntualizaciones mencionadas con respecto del estimado de la frecuencia instantánea se aplican al retardo de grupo, excepto que en este último caso se estrecha la ventana en el dominio de la frecuencia [17].

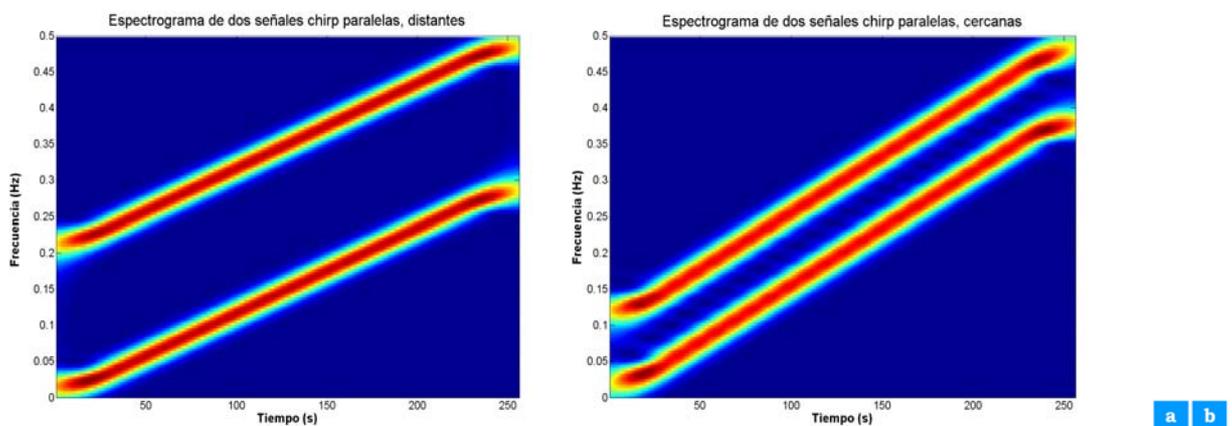


Fig. IV.12. Términos cruzados en el espectrograma. Señal bajo análisis compuesta por dos señales chirp con una modulación lineal en la frecuencia. a) Componentes distantes en el plano tiempo-frecuencia, y b) componentes cercanos en el plano tiempo-frecuencia.

IV.2.2.5. Ventanas de análisis comúnmente empleadas.

Finalmente, aún queda la cuestión de cuál ventana es la más adecuada. Retomando el ejemplo de la estimación de la frecuencia instantánea, se encontró que se debe minimizar la desviación estándar local del estimado, lo cual resulta en elegir la ventana de tal forma que $\sigma_{\omega|t}$ sea mínima. Debido a que esta desviación estándar depende de la señal, la ventana óptima también será dependiente de la señal [1].

Las ventanas han sido empleadas clásicamente en el análisis armónico para reducir algunos efectos no deseables relacionados a la fuga espectral. Entre otras se pueden mencionar la rectangular, triangular, Gaussiana, Hanning, Hamming, Blackman, Blackman-Harris, Kaiser, Tukey, y la lista continua. Cada una de estas ventanas tiene una justificación así como parámetros significativos. A continuación se presentan las tres ventanas empleadas en esta tesis para la obtención del espectrograma, ver figura IV.13 [31]:

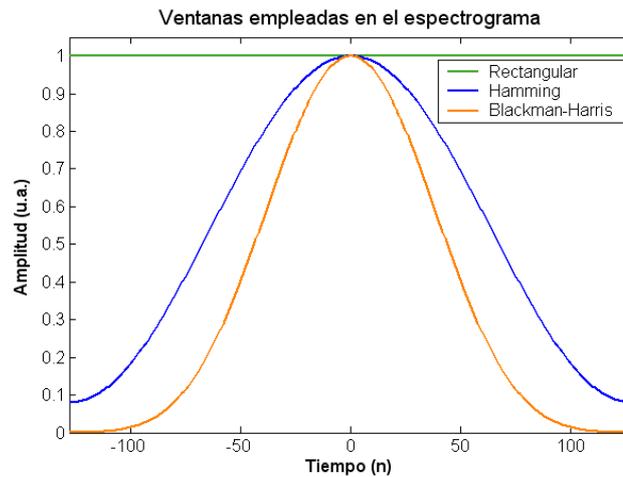


Fig. IV.13. Funciones ventana empleadas para la obtención del espectrograma.

1. Ventana rectangular o Dirichlet.

Definición:

$$h(n) = \begin{cases} 1 & , \quad n = -\frac{N}{2}, \dots, -1, 0, 1, \dots, \frac{N}{2} \\ 0 & , \quad \text{en otro caso} \end{cases} \quad (68)$$

La transformada de Fourier de esta ventana corresponde al kernel Dirichlet que exhibe un lóbulo principal. Debido a su morfología, esta ventana sufre en gran medida del fenómeno de Gibbs en las discontinuidades, y es precisamente este comportamiento oscilatorio el que se trata de controlar empleando otras ventanas.

2. Ventana Hamming.

Definición:

$$h(n) = \begin{cases} 0.54 + 0.46 \cos \left[\frac{2\pi}{N} n \right] & , \quad n = -\frac{N}{2}, \dots, -1, 0, 1, \dots, \frac{N}{2} \\ 0 & , \quad \text{en otro caso} \end{cases} \quad (69)$$

Esta ventana puede pensarse como una ventana de Hann modificada. La ventana espectral de Hann se construye mediante la suma de tres kernel Dirichlet con diferentes fases de tal forma que se cancele la estructura de los lóbulos laterales. Cuando se alcanza la perfecta cancelación del primer lóbulo, la ventana resultante es la ventana Hamming.

3. Ventana Blackman-Harris.

Es posible construir ventanas mediante la suma de $(2K - 1)$ kernel Dirichlet. Blackman examinó estas ventanas para una $K = 3$. Sin embargo, empleando la técnica de búsqueda por gradiente, Harris encontró la familia de ventanas Blackman-Harris de 3 y 4 términos distintos de cero, las cuales alcanzan un nivel del lóbulo lateral mínimo. La ventana Blackman-Harris mínima de 4 términos esta definida por:

$$h(n) = \begin{cases} a_0 + \sum_{k=1}^3 a_k \cos\left[\frac{2\pi}{N} kn\right] & , \quad n = -\frac{N}{2}, \dots, -1, 0, 1, \dots, \frac{N}{2} \\ 0 & , \quad \text{en otro caso} \end{cases} \quad (70)$$

donde

$$a_0 = 0.35875$$

$$a_1 = -0.48829$$

$$a_2 = 0.14128$$

$$a_3 = -0.01168$$

IV.2.2.6. Relación y comparación con la distribución de Wigner.

La relación entre la distribución de Wigner y el espectrograma fue inicialmente mostrada por Mark [32]. Al realizar la convolución bidimensional entre la distribución de Wigner de la señal y la distribución de Wigner de una ventana, en el dominio tiempo-frecuencia, se obtiene el espectrograma, es decir:

$$SP(t, \omega) = \frac{1}{2\pi} \iint WVD_s(u, \Omega) WVD_h(t-u, \omega-\Omega) du d\Omega \quad (71)$$

$$SP(t, \omega) = \left| \frac{1}{2\pi} \int s(\tau) h(\tau-t) e^{-j\omega\tau} d\tau \right|^2$$

Entonces, el espectrograma puede pensarse como la distribución de Wigner de la señal suavizada con la distribución de Wigner de la ventana. Aunque esta relación es interesante, no debe conferirse un significado particular ya que se sabe que esta relación es tan sólo un caso especial de las relaciones que conectan dos distribuciones tiempo-frecuencia bilineales diferentes [17].

Comúnmente se dice que una de las ventajas de la distribución de Wigner sobre el espectrograma es que no hay que lidiar con la elección de la ventana. Sin embargo, el espectrograma no es una distribución, es una clase infinita de distribuciones. El asunto es que la distribución de Wigner resulta mejor que cualquier espectrograma en algunos aspectos, ya que incluso después de lidiar con la búsqueda de la ventana quizá no sea posible hallar una que produzca un espectrograma que sea mejor que la distribución de Wigner. En particular, la distribución de Wigner brinda una idea clara de la frecuencia instantánea y el retardo de grupo, ya que sus promedios condicionales corresponden a dichas cantidades. En contraste, lo anterior nunca ocurre con el espectrograma, aunque algunas veces se logra una buena aproximación [1].

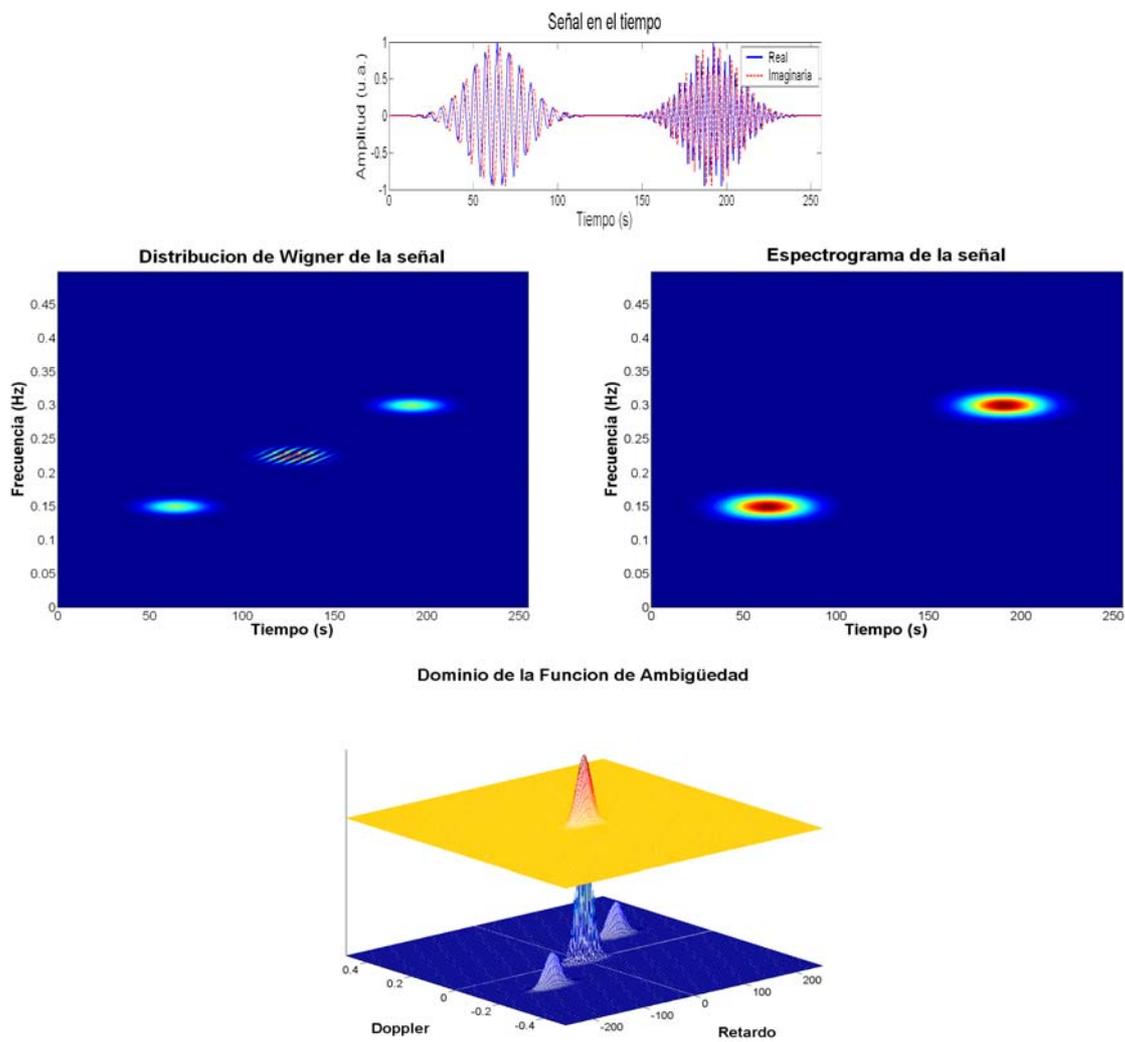


Fig. IV.14. Comparación entre la distribución de Wigner-Ville y el espectrograma de una señal compleja que consiste en la suma de dos señales Gaussianas. a) Señal en el tiempo, b) Distribución de Wigner-Ville de la señal, c) Espectrograma de la señal, y d) Acción pasa-bajas del kernel del espectrograma en el dominio de la función de ambigüedad.



La reducción de los términos cruzados en el espectrograma se debe a que su función kernel, correspondiente a la función de ambigüedad de la ventana de análisis, actúa como un filtro pasa-bajas en el dominio de la función de ambigüedad, ponderando en mayor medida los términos de la señal que se encuentran alrededor del origen y atenuando los términos cruzados que se encuentran alejados del origen, ver figura IV.14. Sin embargo, esta limpieza de los términos cruzados se realiza a expensas de una disminución en la resolución tiempo-frecuencia del espectrograma.

Por otra parte, una de las ventajas del espectrograma es que es una distribución apropiada en el sentido de que es positiva y los resultados obtenidos pueden ser interpretados, aunque estos pueden ser pobres o erróneos.

IV.2.3. La Distribución Choi-Williams.

Tradicionalmente, el espectrograma ha desempeñado un papel importante en la visualización del contenido en frecuencia variante en el tiempo para varios tipos señales, tales como la voz humana y los sonidos respiratorios. Sin embargo, como ya se mencionó, la desventaja inherente de esta técnica radica en el compromiso inevitable entre las resoluciones temporal y espectral. Las limitaciones de la distribución de Wigner-Ville y del espectrograma motivaron el diseño de nuevas representaciones tiempo-frecuencia. Choi, Jeong y Williams desarrollaron la teoría de las distribuciones con interferencia reducida [2, 4, 14], así como las ideas que permiten diseñar funciones kernel para alcanzar tal objetivo. Su primer ejemplo fue la distribución introducida por Choi y Williams, una distribución con un kernel de tipo exponencial que nombraron Distribución Exponencial, ED [4]. Sin embargo, actualmente los investigadores la denominan Distribución Choi-Williams, CWD. La CWD parcialmente mejora las desventajas del espectrograma y la WVD, proporcionando una alta resolución en el tiempo y en la frecuencia y suprimiendo los términos cruzados.

Previo al trabajo de Choi y Williams, diversos estudios reportaron intentos por reducir, o remover, los términos cruzados mediante operaciones sobre la distribución [33, 34]. Sin embargo, ninguna de estas aproximaciones logró conservar las propiedades deseables en la distribución tales como las marginales de energía y los primeros momentos locales. Choi y Williams observaron que mediante una elección adecuada de la función kernel $\phi(\theta, \tau)$ es posible minimizar los términos cruzados y aún mantener las propiedades deseables de los

auto-términos. Finalmente, encontraron que un kernel de tipo exponencial era una buena elección.

IV.2.3.1. Formulación de la distribución.

Como ya se mencionó, la distribución Choi-Williams (CWD) es un miembro de la clase general de Cohen que posee la función kernel [4]:

$$\phi_{CWD}(\theta, \tau) = e^{-\frac{\theta^2 \tau^2}{\sigma}} \quad (72)$$

donde $\sigma > 0$ es un factor de escalamiento. Al sustituir este kernel en la expresión de la clase general se obtiene en el dominio del tiempo:

$$CWD(t, \omega) = \iint \sqrt{\frac{\sigma}{4\pi\tau^2}} e^{-\frac{\sigma(u-t)^2}{4\tau^2} - j\tau\omega} s^*(u - \frac{1}{2}\tau) s(u + \frac{1}{2}\tau) dud\tau \quad (73)$$

y en el dominio de la frecuencia:

$$CWD(t, \omega) = \frac{1}{2\pi} \iint \sqrt{\frac{\sigma}{4\pi\theta^2}} e^{-\frac{\sigma(u-\omega)^2}{4\theta^2} - j\theta t} S(u + \frac{1}{2}\theta) S^*(u - \frac{1}{2}\theta) dud\theta \quad (74)$$

Observe que si considera un valor de σ demasiado grande, entonces la distribución Choi-Williams se aproxima a la distribución de Wigner-Ville, debido a que su kernel se aproxima a la unidad [1], es decir:

$$\lim_{\sigma \rightarrow \infty} CWD(t, \omega, \sigma) \rightarrow WVD(t, \omega) \quad (75)$$

Una primera interpretación de la CWD puede hacerse desde el punto de vista de la estimación de la densidad espectral. Recordando que una distribución tiempo-frecuencia de la clase de Cohen, $C(t, \omega)$, puede ser considerada como la transformada de Fourier de la función de autocorrelación local $R_t(\tau)$ la cual es estimada para un tiempo dado, t , es decir:

$$C(t, \omega) = \frac{1}{2\pi} \int R_t(\tau) e^{-j\tau\omega} d\tau \quad (76)$$

donde

$$R_t(\tau) = \iint \phi(\theta, \tau) s(u + \frac{1}{2}\tau) s^*(u - \frac{1}{2}\tau) e^{j\theta(u-t)} dud\theta \quad (77)$$

entonces, la función kernel $\phi(\theta, \tau)$ desempeña un papel importante en la determinación de las características de la función de autocorrelación local, estimada mediante promedios temporales. La cuestión ahora consiste en determinar que propiedades debe poseer el kernel para la estimación de la función de autocorrelación. La función kernel necesita ponderar por un gran peso a $s(u + \frac{1}{2}\tau) s^*(u - \frac{1}{2}\tau)$ cuando u esta cercana a t y ponderar dicha cantidad en

menor medida en el caso contrario [4]. Es decir, dado que el interés principal de la distribución es estudiar los fenómenos que ocurren localmente, se debe de ponderar en mayor medida $s(u + \frac{1}{2}\tau)s^*(u - \frac{1}{2}\tau)$ cuando u esta cercana a t .

En la distribución Choi-Williams, la autocorrelación local es estimada mediante:

$$R_t(\tau)_{CWD} = \int \sqrt{\frac{\sigma}{4\pi\tau^2}} e^{-\frac{\sigma(u-t)^2}{4\tau^2}} s(u + \frac{1}{2}\tau)s^*(u - \frac{1}{2}\tau) du \quad (78)$$

Entonces, cada valor de la función de autocorrelación local es obtenido de un conjunto de muestras vecinas con pesos selectivos, el tamaño del conjunto vecino esta controlado por la variable τ [4].

La segunda interpretación de la CWD está dada en términos de la función de ambigüedad generalizada [12] y resulta muy útil para poder visualizar la efectividad del kernel exponencial de la CWD en la reducción de los términos cruzados. Como se ha mencionado, la función de ambigüedad generalizada posee la siguiente estructura bilineal

$$M(\theta, \tau; \phi) = \phi(\theta, \tau) \int e^{j(\theta u)} s(u + \frac{1}{2}\tau)s^*(u - \frac{1}{2}\tau) du \quad (79)$$

Por lo tanto, para una señal multicomponente, representada por la suma de sus componentes, la función de ambigüedad puede ser descompuesta en auto-términos y términos cruzados [4], de la forma:

$$M(\theta, \tau; \phi) = \sum_{i=1}^n M_{i,i}(\theta, \tau; \phi) + \sum_{n \neq m} M_{n,m}(\theta, \tau; \phi) \quad (80)$$

donde

$$M_{i,i}(\theta, \tau; \phi) = \phi(\theta, \tau) \int e^{j(\theta u)} s_i(u + \frac{1}{2}\tau)s_i^*(u - \frac{1}{2}\tau) du \quad (81)$$

es la contraparte de un auto-término, y

$$M_{n,m}(\theta, \tau; \phi) = \phi(\theta, \tau) \int e^{j(\theta u)} s_n(u + \frac{1}{2}\tau)s_m^*(u - \frac{1}{2}\tau) du \quad (82)$$

es la contraparte de un término cruzado. Por lo tanto, para poder enfatizar los auto-términos de una distribución tiempo-frecuencia y disminuir los términos cruzados, resulta casi natural asignar un gran factor de peso cuando θ y τ están cercanos al origen del plano de ambigüedad y un factor pequeño cuando θ y τ se encuentran lejanos del origen [4, 35].

La importancia del trabajo de Choi y Williams es la formulación y efectiva implementación de una forma de elegir distribuciones que minimicen los valores espurios causados por los términos cruzados. Así como el hecho de haber conectado las propiedades de una distribución con la función de autocorrelación local y la función característica. Aunque el kernel exponencial de la CWD conforma una familia con un parámetro único, el método de

Choi y Williams puede ser empleado para encontrar muchas otras funciones kernel que posean las propiedades deseables [17].

IV.2.3.2. Propiedades.

A continuación se discuten algunas propiedades de la distribución Choi-Williams.

1. Evaluada en los reales.

La CWD es una distribución evaluada en los reales, debido a que su kernel exponencial cumple que [4]:

$$\phi_{CWD}^*(\theta, \tau) = \phi_{CWD}(-\theta, -\tau) \quad (83)$$

2. Corrimientos en el tiempo y en la frecuencia.

Debido a que la distribución Choi-Williams es una distribución de la clase general, satisface la propiedad de ser invariante al corrimiento en ambos dominios. En términos del kernel, la condición que se debe cumplir para que la distribución preserve los corrimientos de la señal en el tiempo [1], es:

$$\phi(\theta, \tau) \text{ es independiente de } t \quad (84)$$

y en la frecuencia es:

$$\phi(\theta, \tau) \text{ es independiente de } \omega \quad (85)$$

Dado que ambas condiciones se cumplen para $\phi_{CWD}(\theta, \tau)$, la CWD es invariante a corrimientos en el tiempo y en la frecuencia.

3. Marginales.

La distribución Choi-Williams satisface las marginales en el tiempo y en la frecuencia. Para el caso de la marginal en el tiempo, la integral de la CWD sobre el tiempo para cada frecuencia es igual a la densidad espectral de la señal en esa frecuencia. En términos del kernel la CWD cumple la condición:

$$\phi_{CWD}(\theta, 0) = 1 \quad \text{para toda } \theta \quad (86)$$

En forma similar, para la marginal en la frecuencia la integral de la CWD sobre todas las frecuencias para cada instante de tiempo corresponde a la potencia instantánea en ese tiempo, debido a que su kernel satisface:

$$\phi_{CWD}(0, \tau) = 1 \quad \text{para toda } \tau \quad (87)$$

4. Energía total.

La integral de la CWD sobre todo el plano tiempo-frecuencia es igual a la energía total de la señal. Esta propiedad es el resultado de que la CWD satisface ambas marginales. En términos del kernel, la CWD cumple que

$$\phi_{CWD}(0,0) = 1 \quad (88)$$

5. Frecuencia instantánea y retardo de grupo.

El centroide en frecuencia de la CWD para cada instante de tiempo es igual a la frecuencia instantánea de la señal, dado que su kernel cumple [4]:

$$\begin{aligned} \phi_{CWD}(\theta,0) &= 1 && \text{para todo } \theta \\ \left. \frac{d}{d\tau} \phi_{CWD}(\theta,\tau) \right|_{\tau=0} &= 0 && \text{para toda } \theta \end{aligned} \quad (89)$$

El centroide temporal de la CWD para cada frecuencia es igual al retardo de grupo. En términos del kernel la CWD cumple que [4]:

$$\begin{aligned} \phi_{CWD}(0,\tau) &= 1 && \text{para toda } \tau \\ \left. \frac{d}{d\theta} \phi_{CWD}(\theta,\tau) \right|_{\theta=0} &= 0 && \text{para toda } \tau \end{aligned} \quad (90)$$

6. No positividad.

Aunque la CWD cumple con las propiedades anteriores, no es una distribución no-negativa debido a que su función kernel no corresponde a la forma de una combinación lineal de funciones de ambigüedad [4].

7. Reducción de interferencias.

El kernel de la CWD actúa como un filtro pasa-bajas en el dominio de la función de ambigüedad. El parámetro σ de la función kernel exponencial de la CWD permite ajustar el compromiso entre la resolución de los auto-términos y la supresión de los términos cruzados [4]. En la figura IV.15 se presenta el efecto del kernel CWD.

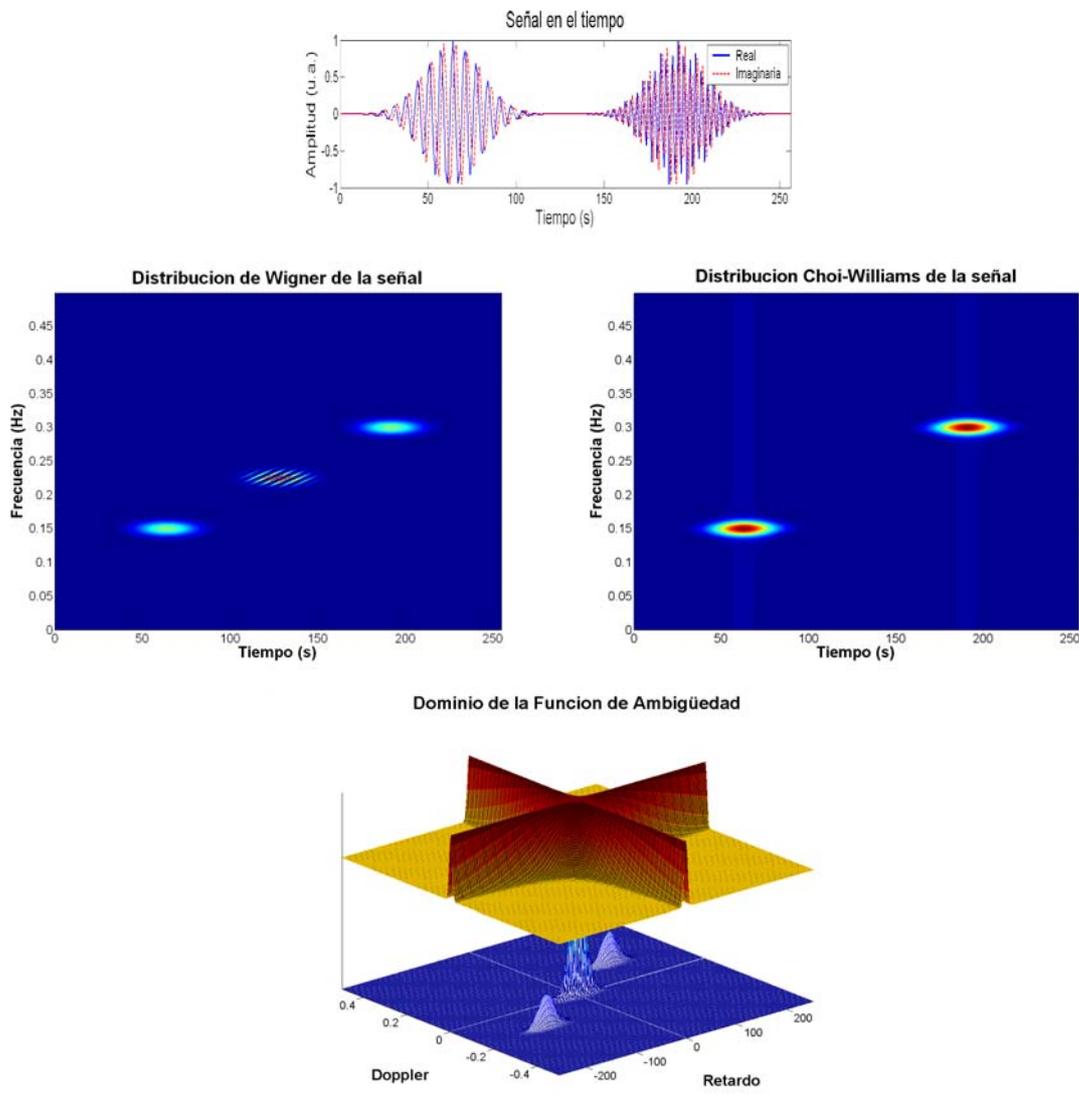


Fig. IV.15. Comparación entre la distribución de Wigner-Ville y la distribución Choi-Williams de una señal compleja que consiste en la suma de dos señales Gaussianas. a) Señal en el tiempo, b) Distribución de Wigner, c) Distribución Choi-Williams con un factor de escalamiento $\sigma = 300$, y d) Acción pasa-bajas del kernel de la CWD en el dominio de la función de ambigüedad. Observe que el $\phi_{CWD}(\theta, \tau)$ es unitario a lo largo de los ejes y decae rápidamente lejos del origen.



En la Figura IV. 16 se muestra el efecto del parámetro σ del kernel exponencial de la distribución Choi-Williams en el dominio de la función de ambigüedad y el dominio tiempo-frecuencia. En la figura se utilizaron valores de $\sigma = 300$, $\sigma = 1000$ y $\sigma = 10000$. Mediante la variación del parámetro σ se controla la caída relativa del kernel (curvas del mismo nivel de energía). Como se puede observar, conforme σ se incrementa los términos cruzados aumentan y la distribución Choi-Williams se aproxima la distribución Wigner-Ville.

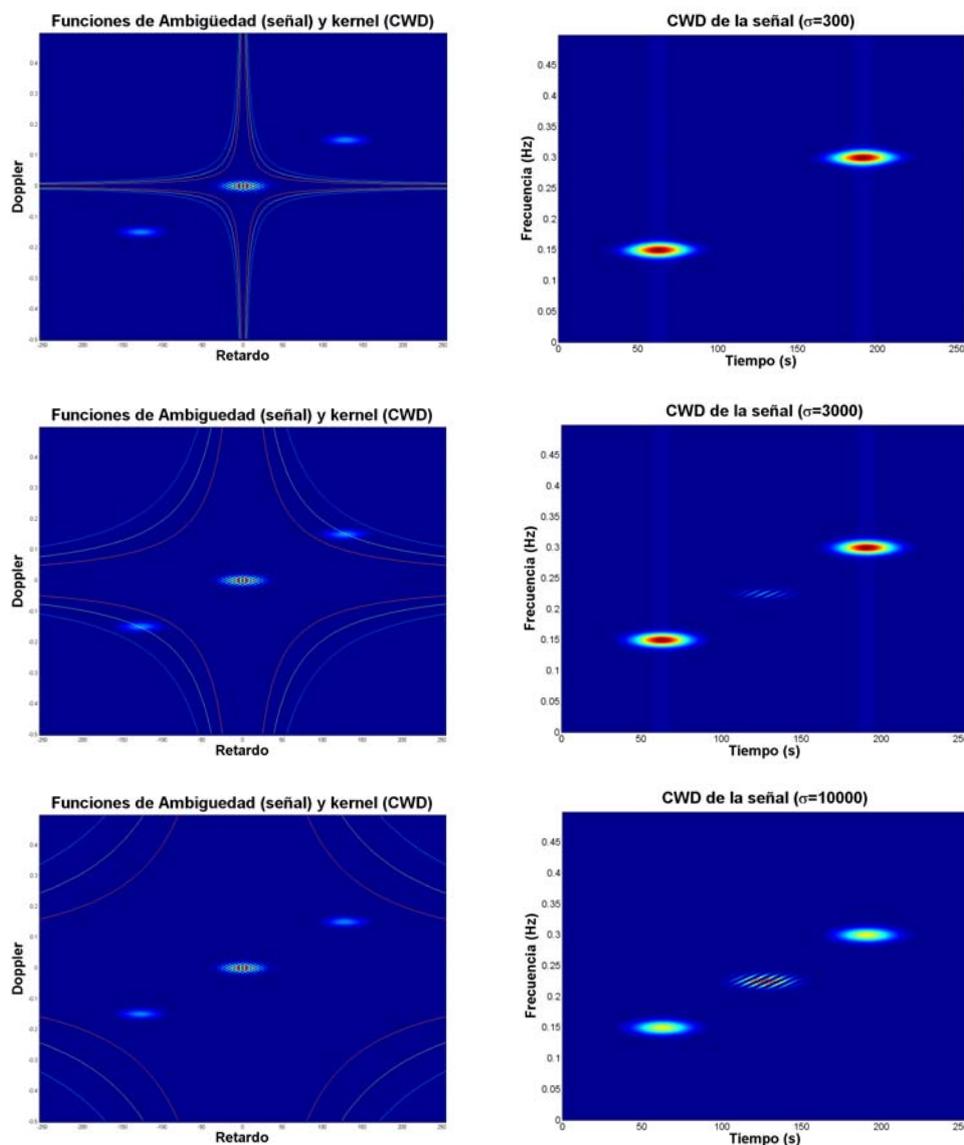


Fig. IV.16. Efecto del parámetro σ en la distribución Choi-Williams para la señal de la figura IV.15.a.

IV.3. El Espectrograma Reasignado.

Hasta ahora se han presentado tres distribuciones pertenecientes a la clase general de distribuciones, en particular se han mencionado algunas de las limitantes de estas técnicas como por ejemplo el compromiso existente entre la cantidad de términos cruzados y la resolución tiempo-frecuencia. Se han propuesto diversas técnicas de procesamiento para mejorar la legibilidad de la representación de la señal., e. g. la descomposición de la señal en componentes elementales y la utilización de la suma de las representaciones de estos componentes como la representación de la señal original [36], o el intento por reconocer los términos de interferencia debido a su geometría particular y su estructura oscilatoria para su remoción mediante el procesamiento de imágenes [37]. Otro tipo de procesamiento de la representación de la señal consiste en realizar un incremento de la concentración de componentes de la señal. El *método de la reasignación* [38, 39] pertenece a este último tipo de esfuerzos. A continuación se detalla la aplicación de dicho método al espectrograma, aplicación que produce el espectrograma modificado o reasignado.

IV.3.1. Formulación de la distribución.

Como ya se mencionó, el espectrograma, i.e. la magnitud al cuadrado de la transformada de Fourier de tiempo corto, puede expresarse como la convolución bidimensional de la distribución Wigner-Ville de la señal $x(t)$ con la distribución Wigner-Ville de alguna ventana de análisis $h(t)$ [32]

$$SP_h(x; t, \omega) = \iint WVD(x; u, \Omega) WVD(h; t - u, \omega - \Omega) du \frac{d\Omega}{2\pi} \quad (91)$$

por lo tanto, para el caso del espectrograma la función kernel en el dominio del tiempo es [39]

$$\phi_{SP}(u, \Omega) = WVD(h; u, \Omega) \quad (92)$$

La idea de la reasignación fue introducida como un intento por mejorar el espectrograma. Koderá et al. [38] sugirieron modificar el espectrograma moviendo cada valor calculado en cualquier punto (t, ω) a otro punto (t', ω') que sea más representativo de la localización de la energía de la señal en la vecindad de (t, ω) . El punto (t', ω') fue simplemente escogido como el centro de gravedad de los valores de energía de la señal localizados alrededor de (t, ω) . Esta operación, condujo al espectrograma modificado, o reasignado [38].

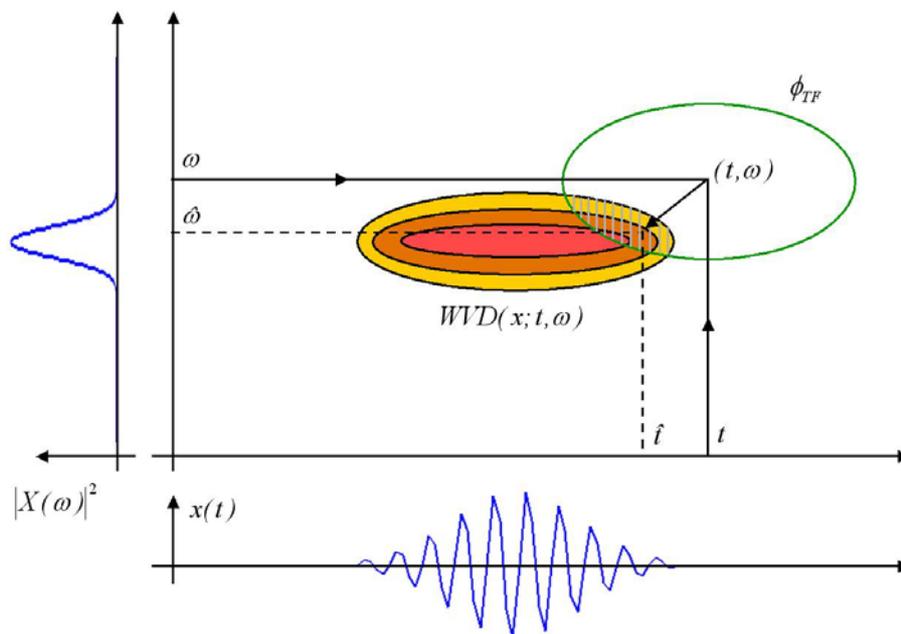


Fig. IV.17. Principio del método de reasignación.

El punto clave del método de la reasignación radica en que no hay razón para que los valores de la distribución estén simétricamente distribuidos alrededor del punto (t, ω) , el cual es el centro geométrico de dicho dominio, y por lo tanto, su promedio no debería ser asignado a este punto, sino al centro de gravedad de dicho dominio, el cual es mucho más representativo de la distribución energética local de la señal. La reasignación del espectrograma permite compensar su pobre concentración tiempo-frecuencia. Puede mostrarse que los operadores de reasignación del espectrograma corresponden a las coordenadas del centro de gravedad de la energía de la señal localizada en un dominio cerrado y centrado en (t, ω) . Desafortunadamente las expresiones obtenidas por Kodera et al. no conducen a una implementación eficiente para señales de tiempo discreto. Auger et al., encontraron que dichas coordenadas pueden ser fácilmente calculadas empleando solamente transformadas de Fourier de tiempo corto [39].

El operador de reasignación para el tiempo está dado por [39]

$$\begin{aligned}
\hat{t}(x; t, \omega) &= t - \frac{\iint u \cdot WVD(h; u, \Omega) WVD(x; t - u, \omega - \Omega) du \frac{d\Omega}{2\pi}}{\iint WVD(h; u, \Omega) WVD(x; t - u, \omega - \Omega) du \frac{d\Omega}{2\pi}} \\
\hat{t}(x; t, \omega) &= t - \Re e \left\{ \frac{\iint u \cdot Ri^*(h; u, \Omega) Ri(x; t - u, \omega - \Omega) du \frac{d\Omega}{2\pi}}{\iint Ri^*(h; u, \Omega) Ri(x; t - u, \omega - \Omega) du \frac{d\Omega}{2\pi}} \right\} \\
\hat{t}(x; t, \omega) &= t - \Re \left\{ \frac{STFT_{T_h}(x; t, \omega) \cdot STFT_h^*(x; t, \omega)}{|STFT_h(x; t, \omega)|^2} \right\}
\end{aligned} \tag{93}$$

y su contraparte en la frecuencia por [39]

$$\begin{aligned}
\hat{\omega}(x; t, \omega) &= \omega - \frac{\iint \Omega \cdot WVD(h; u, \Omega) WVD(x; t - u, \omega - \Omega) du \frac{d\Omega}{2\pi}}{\iint WVD(h; u, \Omega) WVD(x; t - u, \omega - \Omega) du \frac{d\Omega}{2\pi}} \\
\hat{\omega}(x; t, \omega) &= \omega - \Re \left\{ \frac{\iint \Omega \cdot Ri^*(h; u, \Omega) Ri(x; t - u, \omega - \Omega) du \frac{d\Omega}{2\pi}}{\iint Ri^*(h; u, \Omega) Ri(x; t - u, \omega - \Omega) du \frac{d\Omega}{2\pi}} \right\} \\
\hat{\omega}(x; t, \omega) &= \omega + \Im m \left\{ \frac{STFT_{D_h}(x; t, \omega) \cdot STFT_h^*(x; t, \omega)}{|STFT_h(x; t, \omega)|^2} \right\}
\end{aligned} \tag{94}$$

donde

$$Ri^*(x; t, \omega) = x(t) \cdot X^*(\omega) e^{-j\omega t} \tag{95}$$

y

$$T_h = t \cdot h(t) \tag{96}$$

$$D_h = Dh(t) = h'(t) = \frac{dh(t)}{dt} \tag{97}$$

corresponden a los operadores de multiplicación y derivación, respectivamente. Respecto a las expresiones correspondientes a los operadores de reasignación, solamente las intermedias fueron utilizadas usadas por Kodera, Gendrin y de Villemary para presentar la reasignación del espectrograma, mientras que las restantes fueron introducidas por Auger y Flandrin.

Por lo tanto, la versión modificada, o reasignada, del espectrograma en cada punto (t', ω') se define como [39]:

$$SPR_h(x; t', \omega') = \iint SP_h(x; t, \omega) \delta(t' - \hat{t}(x; t, \omega)) \delta(\omega' - \hat{\omega}(x; t, \omega)) dt \frac{d\omega}{2\pi} \tag{98}$$

IV.3.2. Propiedades.

Una comparación de las propiedades básicas del espectrograma y su versión reasignada puede verse en la tabla IV.III. Al igual que el espectrograma, su versión reasignada es no-negativa, y retiene todas las propiedades del espectrograma excepto la bilinealidad, satisface la invarianza a corrimientos en el tiempo y la frecuencia y la propiedad de conservación de la energía si se cumple que la ventana de análisis posee una energía unitaria. A pesar de perder la bilinealidad, cualquier espectrograma reasignado gana el beneficio de localizar perfectamente señales chirp e impulsos, es decir [39]:

$$\text{Si } x(t) = A \cdot e^{j(\omega_1 t + \frac{1}{2} \alpha t^2)} \quad (99)$$

$$\text{entonces } \hat{\omega}(x; t, \omega) = \omega_1 + \alpha t(x; t, \omega)$$

$$\text{Si } x(t) = A \cdot \delta(t - t_1)$$

$$\text{entonces } \hat{t}(x; t, \omega) = t_1 \quad (100)$$

TABLA IV.III. ALGUNAS PROPIEDADES DEL ESPECTROGRAMA Y SU VERSIÓN REASIGNADA.

Propiedad	Espectrograma	Espectrograma reasignado
Bilinealidad	+	
Corrimiento en el tiempo	+	+
Corrimiento en la frecuencia	+	+
Conservación de la energía	+	+
Positividad	+	+
Potencia instantánea		
Valores nulos de la señal		
Localización perfecta de chirps		+
Localización perfecta de sinusoidales		+
Localización perfecta de impulsos		+

Kodera, Gendrin y de Villemary mostraron que el método del espectrograma reasignado origina una representación tiempo-frecuencia que emplea, a diferencia del espectrograma, tanto el modulo al cuadrado como la fase de la transformada de Fourier de

tiempo-corto, debido a que los operadores de reasignación son iguales a la frecuencia instantánea y al retardo de grupo de la señal filtrada $y(t) = STFT_h(x; t, \omega)$

$$\hat{t}(x; t, \omega) = -\frac{\partial}{\partial \omega} [\psi_h(x; t, \omega)] \quad (101)$$

$$\hat{\omega}(x; t, \omega) = \omega + \frac{\partial}{\partial t} [\psi_h(x; t, \omega)] \quad (102)$$

donde $\psi_h(x; t, \omega)$ es la fase de la transformada de Fourier de tiempo corto de la señal $x(t)$, es decir:

$$STFT_h(x; t, \omega) = |STFT_h(x; t, \omega)| e^{j\psi_h(x; t, \omega)} \quad (103)$$

Aunque las expresiones para los operadores de reasignación en términos de la fase son físicamente significativas, no llevan a una implementación eficiente. Esto probablemente explica porque el método de la ventana móvil modificada no había sido ampliamente utilizado. Por otra parte, la introducción de las expresiones para los operadores de reasignación derivada por Auger y Flandrin permite una implementación realizable y sin un incremento drástico en la complejidad computacional.

Finalmente, el espectrograma reasignado previene de forma efectiva la dispersión de los componentes de la señal, que usualmente se observa en el espectrograma, lo cual reduce la dependencia del análisis en la selección de una ventana arbitraria [40]:

$$\begin{aligned} \text{Si} \quad & x(t) = |x(t)| e^{j\varphi_x(t)} \\ \text{entonces} \quad & \lim_{h(t) \rightarrow \delta(t)} SP_h(x; t, \omega) = |x(t)|^2 \end{aligned} \quad (104)$$

$$\text{y} \quad \lim_{h(t) \rightarrow \delta(t)} SPR_h(x; t, \omega) = |x(t)|^2 \delta(\omega - \varphi'_x(t))$$

$$\begin{aligned} \text{Si} \quad & X(\omega) = \int x(t) e^{-j\omega t} dt = |X(\omega)| e^{j\psi_x(\omega)} \\ \text{entonces} \quad & \lim_{h(t) \rightarrow 1} SP_h(x; t, \omega) = |X(\omega)|^2 \end{aligned} \quad (105)$$

$$\text{y} \quad \lim_{h(t) \rightarrow 1} SPR_h(x; t, \omega) = |X(\omega)|^2 \delta(t + \psi'_x(\omega))$$

En palabras, conforme la dispersión de la ventana de análisis tiende a cero en el dominio del tiempo (a cero en el dominio de la frecuencia) el espectrograma tiende hacia la potencia instantánea (el espectro de la densidad de energía), con la respectiva pérdida de una dimensión en la representación, y por su parte el espectrograma reasignado tiende hacia la representación tiempo-frecuencia instantánea (representación retardo de grupo-frecuencia),

produciendo una adecuada localización de la señal en el plano tiempo-frecuencia [40]. En la Fig. IV.18. se compara el espectrograma reasignado con el espectrograma y la distribución de Wigner-Ville para una señal compuesta de la suma de un componente sinusoidal y un componente con amplitud constante y frecuencia instantánea descrita por medio periodo de una senoide. Como se observa la WVD presenta términos de interferencia entre los auto-términos, el espectrograma disminuye la presencia de los términos cruzados con la desventaja de aumentar la dispersión de la distribución. Finalmente se puede observar que el espectrograma reasignado incrementa la legibilidad de la información en el dominio tiempo-frecuencia que se aproxima a la representación ideal mostrada en la figura IV.18.a.

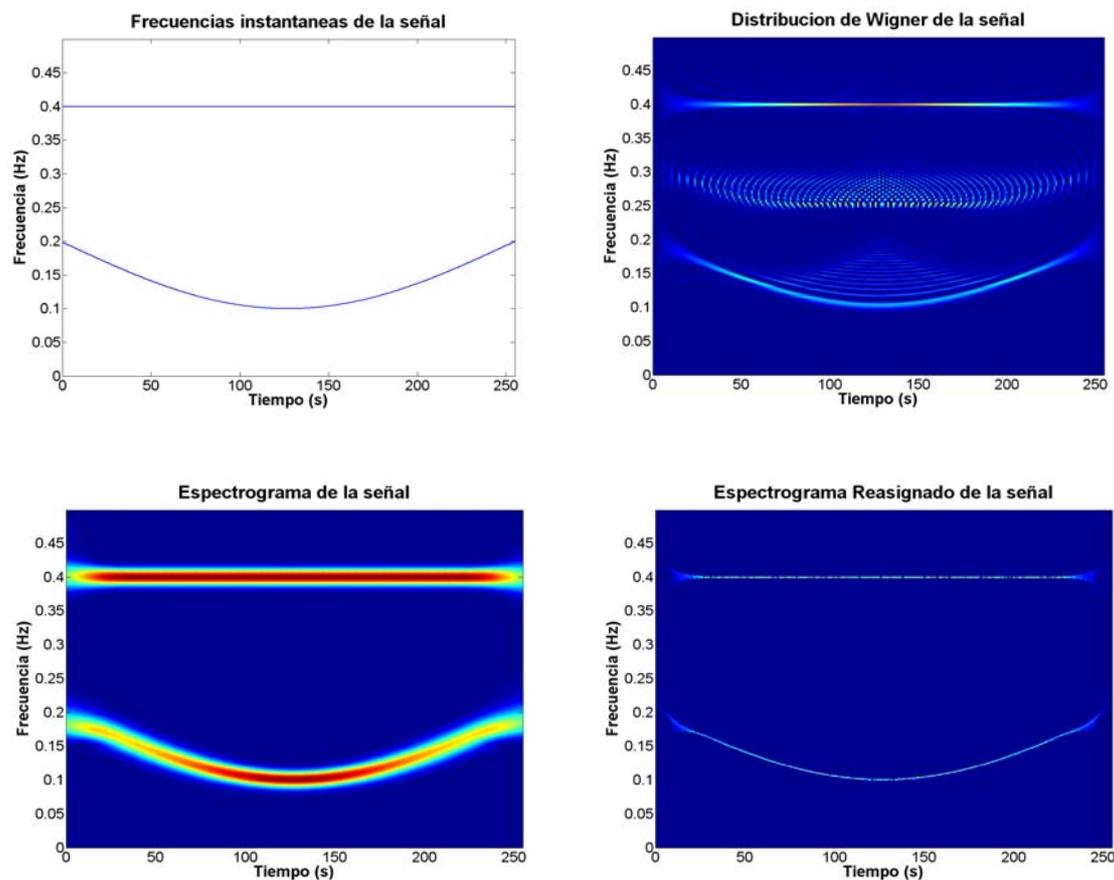


Fig. IV.18. Comparación del Espectrograma reasignado con otras distribuciones. a) Representación tiempo-frecuencia de la señal simulada en términos de la frecuencia instantánea de cada componente, b) Distribución de Wigner-Ville, c) Espectrograma, y d) Espectrograma reasignado.

IV.4. Representación tiempo-frecuencia basado en el modelo autorregresivo variante en el tiempo de una señal estocástica.

Hasta ahora se han presentado las distribuciones tiempo-frecuencia de la clase general y se han discutido sus principales características y limitaciones, por ejemplo la aparición de términos cruzados y la introducción de una función kernel para atenuarlos a expensas de una reducción en la resolución tiempo-frecuencia. Otra aproximación interesante consiste en obtener una representación tiempo-frecuencia de la señal mediante el enfoque de procesos estocásticos. En esta tesis, se emplearán técnicas paramétricas basadas en el modelado autorregresivo variante en el tiempo para considerar la no estacionaridad inherente de los sistemas fisiológicos.

Como punto de partida en la formulación de la representación tiempo-frecuencia de una señal estocástica se establecerán algunos conceptos que son importantes en la obtención del espectro de potencia de una señal discreta [15, 41-44].

El espectro de potencia de una señal estocástica $x(n)$ se puede obtener mediante dos métodos, ver figura IV.19. El primero, *método directo*, involucra el cálculo de la transformada de Fourier de la serie de tiempo $\{x(n)\}$, es decir:

$$S_{xx}(f) = |X(f)|^2 = \left| \sum_{n=-\infty}^{\infty} x(n) e^{-j2\pi fn} \right|^2 \quad (105)$$

La segunda aproximación, *método indirecto*, requiere de dos pasos: a) obtener la función de autocorrelación $r_{xx}(k)$ a partir de $x(n)$, y b) obtener el espectro de potencia mediante su transformada de Fourier, es decir:

$$S_{xx}(f) = \sum_{k=-\infty}^{\infty} r_{xx}(k) e^{-j2\pi kf} \quad (107)$$

En la práctica, sólo se dispone de la secuencia de duración finita $x(n)$, $0 \leq n \leq N-1$, para calcular el espectro de la señal y en consecuencia al estimar el espectro de potencia se debe considerar el efecto de la ventana rectangular, debido a que se genera una dispersión del espectro de potencia de la señal bajo análisis. La dispersión del espectro debido al ventaneo empleado resulta un problema cuando deseamos resolver entre señales cuyos componentes en frecuencia se encuentran muy cercanos entre ellos. Además, debe ser claro que el espectro de potencia de la secuencia ventaneada $\tilde{x}(n)$ es una aproximación del espectro deseado de la secuencia original $x(n)$.

Los problemas de estimación del espectro de potencia de una serie de tiempo son las bases de la motivación de nuevos métodos. Algunos de estos métodos se basan en la estimación no-paramétrica del espectro de potencia, otros en la estimación paramétrica, y otros más en la estimación espectral de varianza mínima.

Los métodos no-paramétricos clásicos no asumen ningún modelo acerca de la generación de la serie de tiempo. Su énfasis radica en obtener un estimado consistente del espectro de potencia a través del empleo de alguna operación de promediación o suavizado realizado directamente sobre el periodograma o sobre la función de autocorrelación. Sin embargo, el efecto de estas operaciones reduce aún más la resolución en la frecuencia, a expensas de reducir la varianza del estimador. Estos métodos clásicos fueron desarrollados por Bartlett en 1948 [45], Blackman y Tukey en 1958 [46], y Welch en 1967 [47].

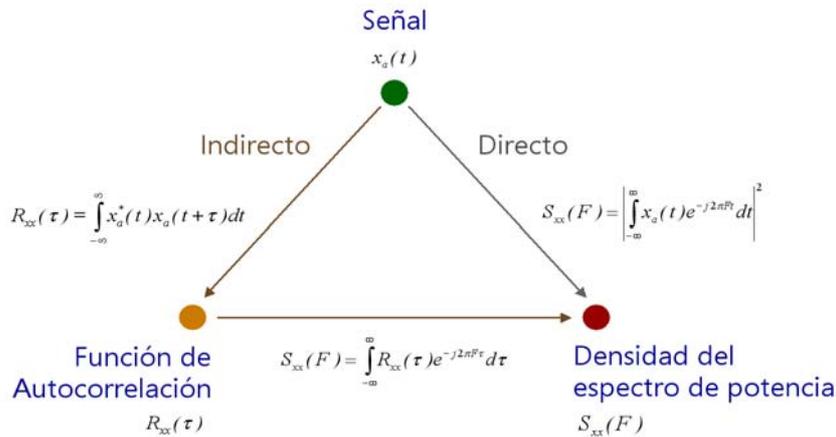


Fig. IV.19. Definiciones equivalentes del espectro de potencia.

Los métodos de estimación del espectro de potencia no-paramétricos son relativamente simples y fáciles de calcular empleando el algoritmo de la transformada rápida de Fourier. Sin embargo, estos métodos requieren de series de tiempo de una longitud considerable para poder obtener la resolución en frecuencia adecuada en muchas aplicaciones. La limitación básica de los métodos no-paramétricos radica en asumir que el estimado de la función de autocorrelación $r_{xx}(m)$ es cero para $m \geq N$, que se ve reflejado en la estimación del espectro de potencia, es decir:

$$S_{xx}(f) = \sum_{m=-(N-1)}^{N-1} r_{xx}(m) e^{-j2\pi f m} \quad (108)$$

que impacta considerablemente la resolución en la frecuencia y la calidad del espectro de potencia estimado, además, la estimación del periodograma asume que los datos son periódicos con periodo N . Sin embargo, ninguna de estas consideraciones es realista.

IV.4.1. Métodos paramétricos para la estimación del espectro de potencia.

En contraste, los métodos paramétricos de estimación espectral se basan en algún modelo acerca de la generación de los datos. En general, los métodos basados en modelos, que se han desarrollado en las últimas tres décadas, proporcionan una más alta resolución que los métodos clásicos.

Los métodos paramétricos de estimación del espectro de potencia *extrapolan* los valores de la autocorrelación para valores $m \geq N$; la extrapolación es posible si se posee información *a priori* acerca de la generación de los datos. En tal caso, se puede construir un modelo a partir de los datos observados. Finalmente, el espectro de densidad de potencia se puede calcular utilizando los parámetros estimados del modelo.

La aproximación mediante el modelado elimina la necesidad de emplear funciones ventana, así como de asumir que la secuencia de autocorrelación es cero para $|m| \geq N$. Como consecuencia, los métodos de estimación del espectro de potencia paramétricos, o basados en modelos, evitan el problema de fuga del espectro y proporcionan una mejor resolución en la frecuencia que los métodos no-paramétricos. Esto es de especial importancia en aplicaciones donde se dispone de registros cortos de datos debido a que el fenómeno es transiente o variante en el tiempo, e.g. análisis de sonidos pulmonares adventicios discontinuos.

Como punto de partida, consideremos aquellos métodos paramétricos que modelan a la secuencia de datos $x(n)$ como la salida de un sistema lineal caracterizado por una función de transferencia racional de la forma:

$$H(z) = \frac{B(z)}{A(z)} = \frac{\sum_{k=0}^q b_k z^{-k}}{1 + \sum_{k=1}^p a_k z^{-k}} \quad (109)$$

cuya ecuación de diferencias correspondiente es:

$$x(n) = -\sum_{k=1}^p a_k x(n-k) + \sum_{k=0}^q b_k w(n-k) \quad (110)$$

donde $w(n)$ es la secuencia de entrada al sistema y los datos observados $x(n)$ representan la secuencia de salida. En la estimación del espectro de potencia, la secuencia de entrada no es observable. En consecuencia, y recurriendo a la teoría de sistemas lineales con entradas aleatorias estacionarias, el espectro de densidad de potencia se puede obtener mediante:

$$S_{xx}(f) = |H(f)|^2 S_{ww}(f) \quad (111)$$

donde $S_{ww}(f)$ es el espectro de densidad de potencia de la secuencia de entrada y $H(f)$ es la respuesta en frecuencia del modelo. Para estimar el espectro de densidad de potencia $S_{xx}(f)$, se asume que la secuencia de entrada $w(n)$ es un proceso de ruido blanco con media cero y con una secuencia de autocorrelación dada por $\gamma_{ww}(m) = \sigma_w^2 \delta(m)$, donde σ_w^2 es la varianza de la secuencia de ruido blanco, i.e. $\sigma_w^2 = E[|w(n)|^2]$, con $E[\cdot]$ representando la operación de valor esperado. Por lo tanto, el espectro de densidad de potencia de los datos observados corresponde simplemente a:

$$S_{xx}(f) = |H(f)|^2 \sigma_w^2 = \sigma_w^2 \frac{|B(f)|^2}{|A(f)|^2} \quad (112)$$

En resumen, el procedimiento de estimación del espectro de una señal $x(n)$ consiste en dos etapas. Primero, dada la secuencia de datos $x(n)$, $0 \leq n \leq N-1$, se estiman los parámetros $\{a_k\}_{k=1,\dots,p}$ y $\{b_k\}_{k=0,\dots,q}$ del modelo. Segundo, a partir de estos estimados, se calcula el estimado del espectro de densidad de potencia. Existe una relación básica entre la secuencia de autocorrelación $\{\gamma_{xx}(m)\}$ y los parámetros $\{a_k\}$ y $\{b_k\}$ del sistema $H(z)$ que genera el proceso mediante el filtrado de la secuencia de ruido blanco $w(n)$. Esta relación puede obtenerse multiplicando la ecuación de diferencias del modelo por el término $x^*(n-m)$ y tomando el valor esperado en ambos lados de la ecuación resultante. Es decir:

$$E[x(n)x^*(n-m)] = -\sum_{k=1}^p a_k E[x(n-k)x^*(n-m)] + \sum_{k=0}^q b_k E[w(n-k)x^*(n-m)] \quad (113)$$

por lo tanto:

$$\gamma_{xx}(m) = -\sum_{k=1}^p a_k \gamma_{xx}(m-k) + \sum_{k=0}^q b_k \gamma_{wx}(m-k) \quad (114)$$

donde $\gamma_{wx}(m)$ es la secuencia de correlación cruzada entre $w(n)$ y $x(n)$. La correlación cruzada $\gamma_{wx}(m)$ se relaciona con la respuesta al impulso del sistema de la siguiente manera:

$$\begin{aligned}
\gamma_{wx}(m) &= E[x^*(n)w(n+m)] \\
\gamma_{wx}(m) &= E\left[\sum_{k=0}^{\infty} h(k)w^*(n-k)w(n+m)\right] \\
\gamma_{wx}(m) &= \sigma_w^2 h(-m)
\end{aligned} \tag{115}$$

donde, en el último paso se utiliza el hecho de que la secuencia $w(n)$ corresponde a un ruido blanco y en consecuencia:

$$\gamma_{wx}(m) = \begin{cases} 0 & , m > 0 \\ \sigma_w^2 h(-m) & , m \leq 0 \end{cases} \tag{116}$$

Combinando el resultado de la ecuación anterior con el obtenido en la ecuación (114) se obtiene la siguiente relación:

$$\gamma_{xx}(m) = \begin{cases} -\sum_{k=1}^p a_k \gamma_{xx}(m-k) & , m > q \\ -\sum_{k=1}^p a_k \gamma_{xx}(m-k) + \sigma_w^2 \sum_{k=0}^{q-m} h(k)b_{k+m} & , 0 \leq m \leq q \\ \gamma_{xx}^*(-m) & , m < 0 \end{cases} \tag{117}$$

La ecuación anterior representa la relación no lineal entre $\gamma_{xx}(m)$ y los parámetros $\{a_k\}$ y $\{b_k\}$ para el proceso ARMA.

IV.4.2. Estimación del espectro de potencia mediante el modelado autorregresivo.

A partir de la ecuación (109) para una función sistema racional, se pueden distinguir tres casos específicos: El modelo autorregresivo de promedio móvil (ARMA), el modelo autorregresivo (AR), y el modelo de promedio móvil (MA). De estos tres modelos lineales, el modelo AR es por mucho el más empleado. Las razones son básicamente cuatro [48]. Primera, muchas señales tales como la voz humana, la música o posiblemente las señales de sonidos pulmonares, poseen una estructura autorregresiva subyacente. Segunda, la estructura de únicamente polos del modelo AR permite una buena representación de espectros de potencia con picos agudos (resonancias), el cual lo hace un candidato adecuado para situaciones como las que posiblemente se presentan en sonidos pulmonares. Tercera, la estimación de los parámetros del modelo AR involucra la solución de un sistema lineal de ecuaciones que puede ser resuelta de manera eficiente. Cuarta, en general, cualquier señal, no

necesariamente de naturaleza AR, puede ser modelada mediante un proceso AR si se selecciona un orden de modelo suficientemente grande.

El teorema de descomposición de Wold, establece que cualquier proceso ARMA o MA puede ser representado de manera unívoca por un modelo AR. En vista de este teorema, la cuestión de adoptar un modelo se reduce a seleccionar aquel que requiera el menor número de parámetros y que además sean fáciles de calcular. Como regla general, el modelo MA requiere muchos más coeficientes para representar espectros con picos agudos y consecuentemente, es raramente utilizado para la estimación espectral. Por lo tanto, en la práctica y en la presente tesis utiliza el modelo AR.

El modelo autorregresivo (AR) incluye los coeficientes a_k , $1 \leq k \leq p$, es decir, la función sistema corresponde a un sistema lineal con únicamente polos (*all-pole*) dado por:

$$H(z) = \frac{1}{A(z)} = \frac{1}{1 + \sum_{k=1}^p a_k z^{-k}} \quad (118)$$

cuya ecuación de diferencias para la relación entrada-salida es:

$$x(n) + \sum_{k=1}^p a_k x(n-k) = w(n) \quad (119)$$

A partir de la ecuación (117) la relación entre los parámetros AR y la secuencia de autocorrelación se obtiene efectuando $q = 0$, es decir:

$$\gamma_{xx}(m) = \begin{cases} -\sum_{k=1}^p a_k \gamma_{xx}(m-k) & , \quad m > 0 \\ -\sum_{k=1}^p a_k \gamma_{xx}(m-k) + \sigma_w^2 & , \quad m = 0 \\ \gamma_{xx}^*(-m) & , \quad m < 0 \end{cases} \quad (120)$$

que corresponde a una relación lineal entre $\gamma_{xx}(m)$ y los parámetros $\{a_k\}$. En este caso, los parámetros AR pueden obtenerse de la solución de las ecuaciones normales o de Yule-Walker, es decir:

$$\begin{bmatrix} \gamma_{xx}(0) & \gamma_{xx}(-1) & \dots & \gamma_{xx}(-p+1) \\ \gamma_{xx}(1) & \gamma_{xx}(0) & \dots & \gamma_{xx}(-p+2) \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{xx}(p-1) & \gamma_{xx}(p-2) & \dots & \gamma_{xx}(0) \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{bmatrix} = - \begin{bmatrix} \gamma_{xx}(1) \\ \gamma_{xx}(2) \\ \vdots \\ \gamma_{xx}(p) \end{bmatrix} \quad (121)$$

y la varianza del proceso de ruido blanco puede calcularse a partir de la ecuación:

$$\sigma_w^2 = \gamma_{xx}(0) + \sum_{k=1}^p a_k \gamma_{xx}(-k) \quad (122)$$

Estas dos últimas ecuaciones se combinan en una sola ecuación matricial para obtener las ecuaciones normales aumentadas, de la forma:

$$\begin{bmatrix} \gamma_{xx}(0) & \gamma_{xx}(-1) & \gamma_{xx}(-2) & \cdots & \gamma_{xx}(-p) \\ \gamma_{xx}(1) & \gamma_{xx}(0) & \gamma_{xx}(-1) & \cdots & \gamma_{xx}(-p+1) \\ \gamma_{xx}(2) & \gamma_{xx}(1) & \cdot & & \gamma_{xx}(-p+2) \\ \vdots & \vdots & & & \vdots \\ \gamma_{xx}(p) & \gamma_{xx}(p-1) & \gamma_{xx}(p-2) & \cdots & \gamma_{xx}(0) \end{bmatrix} \begin{bmatrix} 1 \\ a_1 \\ a_2 \\ \vdots \\ a_p \end{bmatrix} = \begin{bmatrix} \sigma_w^2 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad (123)$$

Debido a que la matriz de autocorrelación en las ecuaciones de Yule-Walker es *Toeplitz*, i.e. $\Gamma_p(i, j) = \Gamma_p(i - j)$, y además *Hermitian*, i.e. $\Gamma_p(i, j) = \Gamma_p^*(i, j)$, los parámetros del modelo se pueden calcular de manera eficiente mediante el empleo del algoritmo Levinson-Durbin. El método de solución de Levinson-Durbin procede de manera recursiva, iniciando con un solo coeficiente, i.e. primer orden, y entonces incrementa el orden del modelo recursivamente para obtener la solución de órdenes más altos mediante el empleo de las soluciones de órdenes más bajos, previamente obtenidos.

En resumen, los parámetros del modelo $AR(p)$ se determinan a partir del conocimiento de la secuencia de autocorrelación $\gamma_{xx}(m)$ para $0 \leq m \leq p$, y con los parámetros se estima el espectro de potencia de la serie de tiempo bajo análisis.

Para el caso del modelado autorregresivo, el espectro de potencia adopta la siguiente forma:

$$S_{xx}(f) = \sigma_w^2 \frac{1}{|A(f)|^2} \quad (124)$$

$$S_{xx}(f) = \frac{\sigma_w^2}{\left| 1 + \sum_{k=1}^p a_k e^{-j2\pi f k} \right|^2}$$

IV.4.3. El espectro de potencia instantánea y el modelo AR variante en el tiempo.

El hecho de que es posible obtener una alta resolución espectral utilizando modelos autorregresivos, en comparación con las técnicas no paramétricas, puede emplearse para construir una representación tiempo-frecuencia paramétrica con una alta resolución tiempo-

frecuencia. Para alcanzar dicho objetivo, se puede extender el modelo autorregresivo para considerar la posible no estacionaridad de la señal, permitiendo que los coeficientes AR varíen en el tiempo. En este caso, se produce un modelo autorregresivo variante en el tiempo, TVAR, el cual para un proceso estocástico $x(n)$ corresponde a:

$$x(n) = -\sum_{k=1}^p a_k(n)x(n-k) + w(n) \quad (125)$$

donde p es el orden del modelo AR variante en el tiempo, el conjunto $\{a_k(n)\}_{k=1}^p$ corresponde a los coeficientes TVAR al tiempo n , el conjunto $\{x(n-k)\}_{k=1}^p$ representa las muestras pasadas de $x(n)$, y $w(n)$ es un proceso de ruido blanco. Consecuentemente, el espectro de potencia de $x(n)$ al tiempo n , esta dado por

$$S_{xx,TVAR}(n, f) = \frac{\sigma_w^2(n)}{\left|1 + \sum_{k=1}^p a_k(n)e^{-j2\pi fk}\right|^2} \quad (126)$$

donde $\sigma_w^2(n)$ representa la varianza del error del modelo TVAR de orden p al tiempo n . El espectro de potencia instantáneo de ecuación (118) constituye la representación tiempo-frecuencia basado de la serie de tiempo $x(n)$, con base en el modelo TVAR.

Debido a las características de cada técnica de estimación de los coeficientes AR, la representación tiempo-frecuencia (TFR), basada en el modelado TVAR, puede obtenerse mediante tres aproximaciones.

La primera aproximación consiste en adoptar un enfoque similar al espectrograma. Es decir, dado que las técnicas paramétricas permiten obtener una alta resolución espectral empleando registros de corta duración, entonces, la TFR puede ser obtenida deslizando una función ventana a lo largo de la señal, estimando los coeficientes AR para cada segmento de la señal, y finalmente calculando el espectro para segmento. Este procedimiento asume estacionaridad en la señal dentro de la duración de la ventana y por ello resulta más adecuado para analizar señales con una baja no-estacionaridad.

Una segunda aproximación consiste en emplear algoritmos recursivos para la estimación de los parámetros AR, de esta manera, los p coeficientes del modelo AR se obtienen mediante la actualización de sus valores en instantes previos, y al final del procedimiento se contará con $N - p + 1$ coeficientes que abarcarán la duración de la señal. Estos métodos ofrecen la ventaja de poder ser aplicados en tiempo real y con ello un mayor seguimiento de la no estacionaridad de la señal.

Una tercera aproximación, que no se explora en esta tesis, consiste en modelar los coeficientes AR mediante una combinación lineal de funciones base deterministas, las cuales, desde luego, son una función del tiempo. Entonces, para cada instante de tiempo, es posible obtener un estimado del espectro de potencia para producir la TFR paramétrica [49, 50]. Esta mejoría hace posible un rápido seguimiento de espectros variantes, sin embargo introduce la dificultad de seleccionar *a priori* las funciones base más adecuadas.

IV.4.4. Estimación de los parámetros del modelo TVAR.

En esta tesis se consideraron cuatro métodos para la obtención de los parámetros del modelo TVAR, y a partir de ellos se calculó la representación tiempo-frecuencia de los sonidos torácicos. Los métodos de estimación considerados para este problema son:

- El método de Burg.
- El algoritmo de mínimos cuadrados recursivo (RLS).
- El algoritmo de mínimos cuadrados recursivo con factor de olvido variable (RLS-VFF).
- El filtro de Kalman.

IV.4.4.1. El método de Burg.

Circa 1967 fue introducido uno de los algoritmos más tempranos y mejor conocidos para la estimación de los parámetros del modelo autorregresivo, a partir de N muestras, por John Burg [51]. El método de Burg se basa en la minimización de los errores hacia delante y hacia atrás de los filtros de predicción lineales, sujeta a la restricción de que los parámetros AR satisfagan la recursión de Levinson-Durbin.

IV.4.4.1.1. Derivación del algoritmo.

Para derivar el estimador, se parte de un conjunto de datos $x(n)$, $n = 0, 1, \dots, N - 1$, y se asume un proceso estocástico estacionario de únicamente polos. Entonces, si se consideran los estimados de orden m para sus filtros de predicción lineal hacia delante y hacia atrás, los errores de predicción hacia delante, $f_m(n)$, y hacia atrás, $b_m(n)$, quedan definidos mediante [44, 52]:

$$\begin{aligned}
f_m(n) &= x(n) + \sum_{k=1}^m a_m(k)x(n-k) \\
b_m(n) &= x(n-m) + \sum_{k=1}^m a_m^*(k)x(n+k-m)
\end{aligned}
\tag{127}$$

donde los $a_m(k)$, $0 \leq k \leq m-1$, $m = 1, 2, \dots, p$, son los coeficientes de predicción, o coeficientes AR. Además, los coeficientes del filtro de predicción hacia atrás son simplemente los conjugados de los coeficientes del filtro de predicción hacia delante [53]. Para obtener los estimados de los parámetros AR debe minimizarse la suma de las energías de los errores de predicción hacia delante y hacia atrás, i.e. el error cuadrático medio definido como [44, 53]:

$$\varepsilon_m = \sum_{n=m}^{N-1} \left[|f_m(n)|^2 + |b_m(n)|^2 \right]
\tag{128}$$

La minimización se realiza mediante la selección de los coeficientes de predicción, sujetos a la restricción de satisfacer la recursión de Levinson-Durbin dada por [44]:

$$a_m(k) = a_{m-1}(k) + K_m a_{m-1}^*(m-k), \quad 1 \leq k \leq m-1, \quad 1 \leq m \leq p
\tag{129}$$

donde $K_m = a_m(m)$ es el m -ésimo coeficiente de reflexión en la realización lattice del filtro de predicción. En la figura IV.20.a, se muestran las diferentes representaciones y el enlace matemático de un modelo autorregresivo, mientras que en la figura IV.20.b se muestra la estructura de un filtro lattice de predicción. La energía del error de predicción ε_m al orden m es una función de $a_m(m)$, el cual es desconocido, y del resto de los parámetros AR previos, $a_m(1)$ hasta $a_m(m-1)$, los cuales siempre pueden ser conocidos a partir del orden $m-1$ de la recursión [53]. Por lo tanto, ε_m es una función de un simple parámetro $a_m(m)$, denominado coeficiente de reflexión debido a su interpretación como parámetro físico en un modelo acústico de la voz mediante un tubo [54] o en modelos sísmicos.

Al emplear la restricción de Levinson-Durbin en las ecuaciones para los errores de predicción $f_m(n)$ y $b_m(n)$, se obtienen las siguientes recursiones [44, 53]:

$$\begin{aligned}
f_m(n) &= f_{m-1}(n) + K_m b_{m-1}(n-1), \quad m = 1, 2, \dots, p \\
g_m(n) &= K_m^* f_{m-1}(n) + b_{m-1}(n-1), \quad m = 1, 2, \dots, p
\end{aligned}
\tag{130}$$

donde la inicialización corresponde a $f_0(n) = b_0(n) = x(n)$. Al sustituir ambas ecuaciones recursivas para los errores de predicción en la definición del error cuadrático medio ε_m y realizar la minimización de ε_m con respecto a los coeficientes de reflexión K_m , se tiene [44, 52, 53]:

$$\hat{K}_m = \frac{-\sum_{n=m}^{N-1} f_{m-1}(n)b_{m-1}^*(n-1)}{\frac{1}{2}\sum_{n=m}^{N-1} [f_{m-1}(n)|^2 + |b_{m-1}(n-1)|^2]}, \quad m = 1, 2, \dots, p \quad (131)$$

donde el término del numerador corresponde a un estimado de la correlación cruzada entre los errores de predicción hacia delante y hacia atrás.

Debido a los factores de normalización del numerador se tiene que $|K_m| = |a_m(m)| < 1$, para toda m de tal manera que el modelo de únicamente polos obtenidos de los datos es estable [52], como consecuencia de la localización de los coeficientes del modelo en el interior del círculo unitario del plano z , condición necesaria para producir un filtro de predicción lineal de fase mínima [53, 55]. Un filtro estable significa que se generaran valores de autocorrelación finitos y acotados, como se requiere para un proceso estacionario [42].

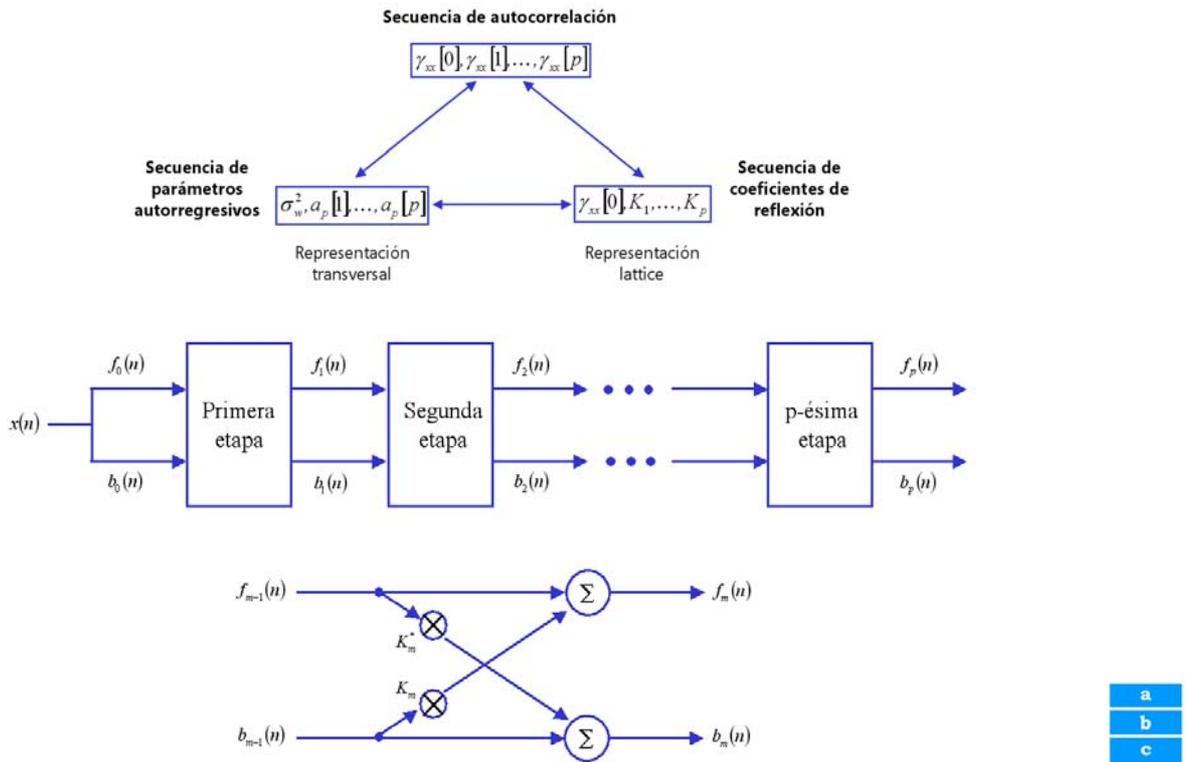


Fig. IV.20. a) Representaciones equivalentes de un proceso autorregresivo mediante la estructura transversal y de lattice, y su relación matemática con base en el algoritmo Levinson-Durbin. b) Estructura de un filtro lattice de predicción con p etapas. c) Una etapa del filtro lattice.

Al expandir el denominador de la ecuación (131), la ecuación de los coeficientes de reflexión adquiere la siguiente forma [44]:

$$\hat{K}_m = \frac{-\sum_{n=m}^{N-1} f_{m-1}(n)b_{m-1}^*(n-1)}{\frac{1}{2}\sum_{n=m}^{N-1}|f_{m-1}(n)|^2 + \frac{1}{2}\sum_{n=m}^{N-1}|b_{m-1}(n)|^2}, \quad m = 1, 2, \dots, p \quad (132)$$

y por lo tanto, el denominador corresponde al estimado de mínimos cuadrados de los errores hacia delante y hacia atrás, E_{m-1}^f y E_{m-1}^b , respectivamente, donde [44]:

$$\begin{aligned} E_{m-1}^f &= E\left[|f_{m-1}(n-1)|^2\right] \\ E_{m-1}^b &= E\left[|b_{m-1}(n-1)|^2\right] \end{aligned} \quad (133)$$

Finalmente, los coeficientes de reflexión pueden ser expresados como:

$$\hat{K}_m = \frac{-\sum_{n=m}^{N-1} f_{m-1}(n)b_{m-1}^*(n-1)}{\frac{1}{2}[\hat{E}_{m-1}^f + \hat{E}_{m-1}^b]}, \quad m = 1, 2, \dots, p \quad (134)$$

donde $\hat{E}_{m-1}^f + \hat{E}_{m-1}^b$ es un estimado del error cuadrático total E_m .

En 1978, Andersen mostró que el denominador también puede ser calculado de forma recursiva mediante la relación [56]:

$$\hat{E}_m = (1 - |\hat{K}_m|^2)\hat{E}_{m-1} - |f_{m-1}(m-1)|^2 - |b_{m-1}(m-2)|^2 \quad (135)$$

donde $\hat{E}_m \equiv \hat{E}_m^f + \hat{E}_m^b$ corresponde al error cuadrático mínimo total. Como consecuencia de dicha recursión, se reduce el esfuerzo computacional para encontrar los coeficientes de reflexión K_m .

IV.4.4.1.2. Características del algoritmo.

El algoritmo de Burg requiere $3Np - p^2 - 2N - p$ adiciones complejas, $3Np - p^2 - N + 3p$ multiplicaciones complejas, p divisiones reales, y $3N + p + 2$ almacenamientos de números complejos [53].

Las principales ventajas del método de Burg para la estimación de los parámetros del modelo AR son básicamente tres: 1) produce un modelo AR estable, 2) es computacionalmente eficiente, y 3) resulta en un estimado del espectro de potencia con alta resolución en la frecuencia [44]. Sin embargo, el método de Burg es conocido por sufrir dos problemas: apertura de la línea espectral y un corrimiento en la posición de los picos espectrales. Por apertura de la línea espectral debe entenderse que el espectro de la señal

puede poseer un único pico agudo, pero el método de Burg resulta en dos o más picos cercanamente espaciados, ver figura IV.21.a. Este problema ocurre con mayor frecuencia cuando: a) la relación señal a ruido es alta, b) la fase inicial de los componentes sinusoidales es algún múltiplo impar de 45° , c) la duración de la secuencia de datos es tal que los componentes sinusoidales poseen un número impar de cuartos de ciclo, y d) el número de parámetros AR estimados corresponde a un gran porcentaje del número de datos empleados para la estimación. Además, para modelos de alto orden, el método introduce picos espurios [57]. Finalmente, para señales sinusoidales inmersas en ruido, el método de Burg es sensible a la fase inicial de una señal sinusoidal, especialmente en registros de corta duración, ver figura IV.21.b. Esta sensibilidad, se manifiesta como un sesgo en la posición de los picos espectrales estimados con respecto a las localizaciones verdaderas de dichos picos [58].

En resumen, para un modelo autorregresivo de orden p , el algoritmo de Burg obtiene los coeficientes de reflexión, mediante el empleo de las ecuaciones:

$$\hat{K}_m = \frac{-\sum_{n=m}^{N-1} f_{m-1}(n)b_{m-1}^*(n-1)}{\frac{1}{2}[\hat{E}_{m-1}^f + \hat{E}_{m-1}^b]}, \quad m = 1, 2, \dots, p$$

$$\hat{E}_m = (1 - |\hat{K}_m|^2)\hat{E}_{m-1} - |f_{m-1}(m-1)|^2 - |b_{m-1}(m-2)|^2$$

y emplea el algoritmo de Levinson-Durbin para obtener los parámetros del modelo autorregresivo. A partir de los estimados de los parámetros AR se puede determinar el estimado del espectro de potencia correspondiente al método de Burg [44], de la forma:

$$PSD_{Burg}(f) = \frac{\hat{E}_p}{\left|1 + \sum_{k=1}^p \hat{a}_p(k)e^{-j2\pi fk}\right|^2} \quad (136)$$

El estimado del espectro de potencia autorregresivo ha sido frecuentemente denominado método de máximo entropía de análisis espectral [55, 59]. Esta interpretación fue introducida en el trabajo doctoral de Burg [60]: “*El método digital usual para obtener un estimado del espectro de potencia a partir de mediciones de la función de autocorrelación considera que la función de correlación es cero para todos los retardos en los cuales no se dispone de un estimado y emplea algún procedimiento para reducir el efecto del truncamiento de la función de autocorrelación. El método discutido en este trabajo, por otra parte, conserva todas las estimaciones en los retardos sin modificación y emplea estimados no nulos para aquellos no estimados directamente. El principio de estimación empleado es*

que el estimado espectral debe ser el más aleatorio o poseer la máxima entropía de cualquier espectro de potencia que sea consistente con los datos medidos. Esta nueva técnica de análisis proporciona un estimado espectral con una mayor resolución que el obtenido por las técnicas convencionales y con muy poco incremento en el tiempo computacional”.

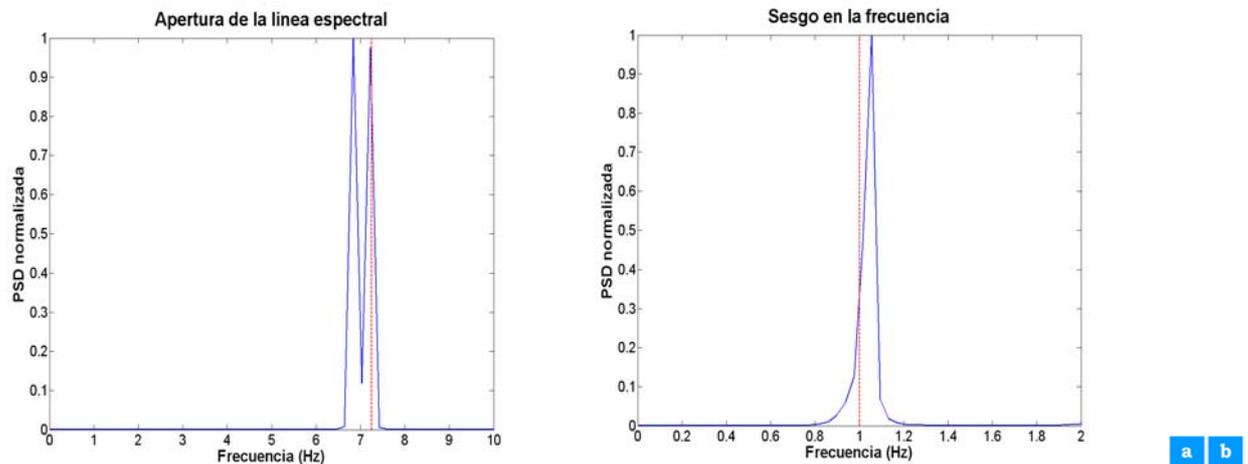


Fig. IV.21. Método de Burg. a) Apertura de la línea espectral para una sinusoidal de amplitud unitaria de longitud de 101 muestras, fase inicial de 45° y frecuencia de 7.25 Hz (línea punteada), inmersa en ruido blanco con relación señal a ruido de 50 dB con una frecuencia de muestreo de 100 Hz. b) Sesgo del estimador espectral cuando la señal consiste de una sinusoidal de 31 muestras inmersa en ruido blanco con una relación señal a ruido de 10 dB, amplitud unitaria, frecuencia de 1 Hz (línea punteada), fase inicial de 70° y una frecuencia de muestreo de 200 Hz.

La estimación espectral autorregresiva mediante el algoritmo de Burg presenta las siguientes características:

- No aplica un ventaneo a los datos.
- Minimiza los errores de predicción hacia delante y hacia atrás en el sentido cuadrático medio, donde los coeficientes del modelo AR están sujetos a la recursión de Levinson-Durbin.
- Presenta una alta resolución en la frecuencia para registros cortos.
- Siempre produce un modelo estable.
- La localización de los picos depende en gran medida de la fase inicial.
- Puede sufrir de división de líneas espectrales para sinusoidales inmersas en ruido, o cuando el orden del modelo empleado es demasiado grande.
- Presenta un corrimiento en la frecuencia al estimar sinusoidales en ruido.

IV.4.4.1.3. Obtención de la representación tiempo-frecuencia.

Una vez que los coeficientes se han estimado de forma óptima mediante el algoritmo de Burg, la TFR se construye mediante la graficación del espectro de potencia en cada instante de tiempo, que corresponde al centro temporal de la función ventana [61], es decir:

$$P_{TVAR}(tw, f) = \frac{\sigma_w^2(tw)}{\left| 1 + \sum_{k=1}^p \hat{a}_k(tw) e^{-j2\pi f k} \right|^2} \quad (137)$$

La función ventana entonces se desliza a lo largo de toda la señal, con un incremento temporal predeterminado para producir la TFR paramétrica. En esta tesis se emplearon dos funciones ventana para segmentar la información de los sonidos torácicos, la ventana de Hamming y la Blackman-Harris.

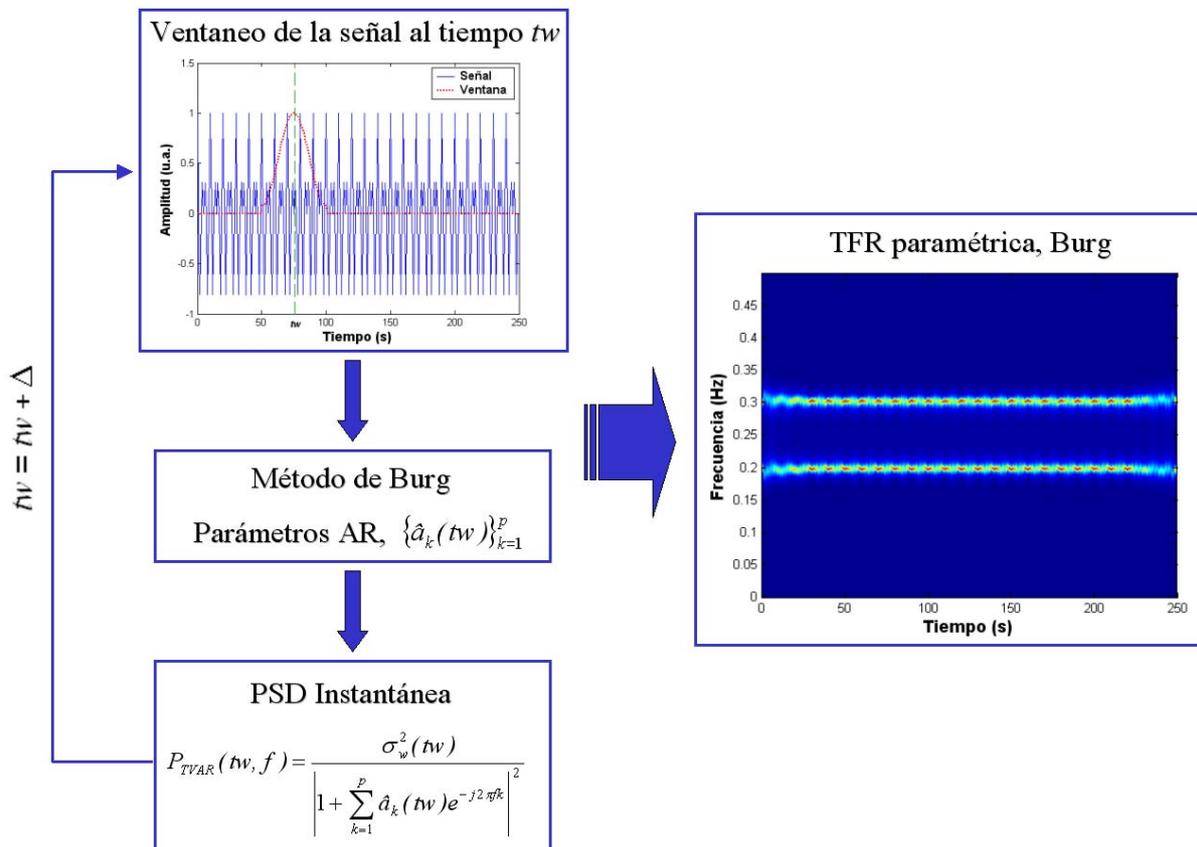


Fig. IV.22. Obtención de la TFR paramétrica utilizando el algoritmo de Burg. A la señal se le aplica una función ventana al tiempo tw y posteriormente se obtienen los p parámetros AR mediante Burg y finalmente, se obtiene el estimado de la potencia espectral al instante de tiempo tw . Para conformar la TFR paramétrica el centro la ventana se desliza hasta abarcar la longitud de la señal.

Una de las ventajas de la representación tiempo-frecuencia paramétrica es que no presenta términos de interferencia inherentes a las distribuciones de la clase general. Además, el empleo de ventanas temporales cortas no implica una degradación en la resolución espectral. Sin embargo, presenta el problema de seleccionar una ventana lo suficientemente larga de tal forma que la señal ventaneada aún posea suficiente información para una buena estimación espectral. Y en este último aspecto, no existen reglas para orientar la decisión. En la figura IV.22, se esquematiza la metodología a seguir para determinar la representación tiempo-frecuencia de una serie de tiempo mediante la determinación de los coeficientes del modelo AR utilizando el algoritmo de Burg.

IV.4.4.2. El algoritmo de mínimos cuadrados recursivo.

El algoritmo de Burg para la estimación espectral paramétrica presentado con anterioridad pertenece al denominado conjunto de métodos de estimación directa en bloques, dado que los datos disponibles se dividen en bloques y los parámetros AR son estimados para cada bloque sin el cálculo previo de la función de autocorrelación. Por su parte, las técnicas de estimación adaptables, secuenciales o recursivas, representan una alternativa donde los coeficientes estimados al tiempo n dependen de sus estimados previos al tiempo $n - 1$ [52].

En esta sección se plantea al algoritmo recursivo de mínimos cuadrados para la obtención de una representación tiempo-frecuencia paramétrica. Se inicia con el desarrollo del algoritmo RLS revisando las ideas básicas referentes al método de mínimos cuadrados. Posteriormente, se explota una relación del álgebra matricial conocida como la identidad de Woodbury para obtener el algoritmo recursivo de interés. Adicionalmente, se presentan algunos resultados acerca de la convergencia y capacidad de seguimiento del algoritmo. Finalmente, se plantea el procedimiento para la obtención de la TFR basada en el modelado TVAR.

Sin duda, uno puede trazar los orígenes de la teoría de la estimación hasta tiempos muy antiguos, sin embargo a Karl Friederich Gauss se le conoce como su progenitor. El trabajo realizado por Norbert Wiener, Rudolph Kalman, y muchos otros, ha expandido el legado de Gauss, y ha brindado una rica colección de métodos y algoritmos de estimación a elegir. Actualmente, la teoría de la estimación es ampliamente utilizada en gran variedad de ramas de la ciencia y la tecnología [62].

La teoría de la estimación es un producto de la necesidad y la tecnología. Por ejemplo, Gauss en 1795 necesitaba predecir el movimiento de los planetas y cometas a partir de

mediciones telescópicas. Los movimientos pueden caracterizarse completamente mediante seis parámetros, por lo tanto, el problema de estimación considerado por Gauss consistió en inferir los valores de tales parámetros a partir de los datos de medición. Esta necesidad propició el *método de mínimos cuadrados* que constituye la piedra angular para gran parte de la teoría de la estimación, tanto clásica como moderna [62].

La estimación de mínimos cuadrados puede estudiarse desde dos puntos de vista, el procesamiento clásico en bloques, en el cual todas las mediciones son procesadas al mismo tiempo y un procesamiento recursivo más moderno, en el cual sólo unas cuantas o incluso una medición es procesada. Es posible extender el uso del método de mínimos cuadrados para desarrollar un algoritmo recursivo para el diseño de filtros transversales adaptables tales que, dado el estimado de mínimos cuadrados del vector de pesos del filtro al tiempo $n - 1$, se pueda calcular el estimado actualizado de este vector al tiempo n conforme van llegando los nuevos datos. El algoritmo resultante se conoce como el *algoritmo de mínimos cuadrados recursivo* (Recursive Least-Squares, RLS) [52, 63].

IV.4.4.2.1. Formulación del algoritmo.

En las implementaciones recursivas del método de mínimos cuadrados, el cálculo se inicia con condiciones iniciales conocidas y se emplea la información contenida en los nuevos datos para actualizar los estimados anteriores. Como consecuencia, la función de costo a minimizar, $E(n)$, es una función de la longitud variable de los datos observables, n . Por otra parte, es común introducir un factor de ponderación en la definición de $E(n)$ [63], es decir:

$$E(n) = \sum_{i=1}^n \beta(n, i) |e(i)|^2 \quad (138)$$

donde $e(i)$ es la diferencia entre la respuesta deseada $d(i)$ y la salida $y(i)$ producida por un filtro transversal, ver figura IV.23. Las entradas al tiempo i son iguales a $u(i), u(i-1), \dots, u(i-M+1)$. Es decir, $e(i)$ está definida como:

$$\begin{aligned} e(i) &= d(i) - y(i) \\ e(i) &= d(i) - \bar{w}^H(n) \bar{u}(i) \end{aligned} \quad (139)$$

donde $\bar{u}(i)$ es el vector de entrada al tiempo i , y corresponde a $\bar{u}(i) = [u(i), u(i-1), \dots, u(i-M+1)]^T$, mientras que $\bar{w}(n)$ es el vector de pesos del filtro al tiempo n de la forma $\bar{w}^H(n) = [w_0(n), w_1(n), \dots, w_{M-1}(n)]^T$.

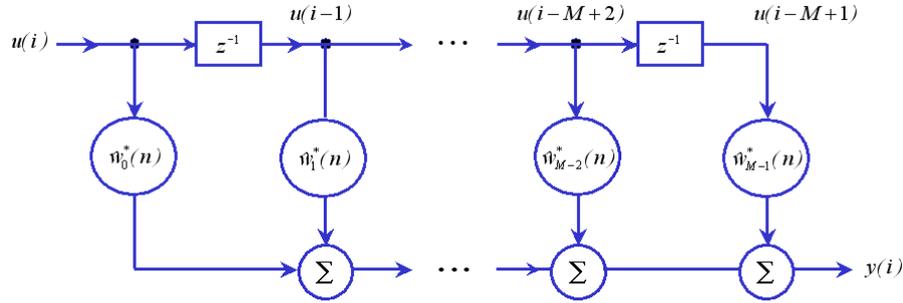


Fig. IV.23. Filtro transversal. El vector de pesos del filtro permanece fijo durante el intervalo de observación $1 \leq i \leq n$ para el cual está definida la función de costo $E(n)$.

El factor de ponderación, $\beta(n, i)$, posee la propiedad de estar definido en el rango $0 < \beta(n, i) \leq 1$, $i = 1, 2, \dots, n$, [42, 63]. Una forma especial de ponderación que comúnmente se emplea es el *factor de ponderación exponencial* o *factor de olvido* definido como $\beta(n, i) = \lambda^{n-i}$, $i = 1, 2, \dots, n$, donde λ es una constante positiva cercana a, pero menor que, la unidad. Hablando a groso modo, $(1 - \lambda)^{-1}$ es una medida de la memoria del algoritmo, en el caso particular de $\lambda = 1$ se obtiene el método ordinario de mínimos cuadrados, que corresponde a una memoria infinita. Entonces, en el método de mínimos cuadrados con ponderación exponencial, se minimiza la función de costo correspondiente al error cuadrático ponderado exponencialmente [63], es decir:

$$E(n) = \sum_{i=1}^n \lambda^{n-i} |e(i)|^2 \quad (140)$$

donde se proporciona una ponderación mayor a los errores actuales y una menor a los errores más antiguos, lo que permite un mejor seguimiento de la señal [42].

El valor óptimo del vector de pesos, $\hat{w}(n)$, para el cual la función de costo $E(n)$ alcanza su valor mínimo, esta definido por las siguientes *ecuaciones normales* escritas en forma matricial [63]:

$$\Phi(n)\hat{w}(n) = \bar{\theta}(n) \quad (141)$$

donde la matriz de correlación, de tamaño $M \times M$, está definida por:

$$\Phi(n) = \sum_{i=1}^n \lambda^{n-i} \bar{u}(i)\bar{u}^H(i) \quad (142)$$

y el vector de correlación cruzada $\bar{\theta}(n)$, de tamaño $M \times 1$, entre la entrada al filtro transversal y la respuesta deseada esta definido como:

$$\bar{\theta}(n) = \sum_{i=1}^n \lambda^{n-i} \bar{u}(i) d^*(i) \quad (143)$$

donde el asterisco denota la operación de conjugación compleja y $(\cdot)^H$ corresponde al operador Hermitian, i.e. complejo conjugado traspuesto. Por su parte, el empleo de $i=1$ como el limite inferior de la sumatoria, indica que se asume la técnica de pre-ventaneo, de acuerdo con la cual los datos de entrada anteriores al tiempo $i=1$ son iguales a cero, $u(i) = 0$, $i \leq 0$.

La matriz de correlación $\Phi(n)$, también se puede expresar como:

$$\Phi(n) = \lambda \left[\sum_{i=1}^{n-1} \lambda^{n-1-i} \bar{u}(i) \bar{u}^H(i) \right] + \bar{u}(n) \bar{u}^H(n) \quad (144)$$

mediante el aislamiento del término correspondiente a $i=n$ del resto de la sumatoria. Pero por definición, la expresión dentro del paréntesis corresponde a la matriz de correlación al tiempo $n-1$. Por lo tanto, la recursión para actualizar el valor de la matriz de correlación de los datos de entrada es:

$$\Phi(n) = \lambda \Phi(n-1) + \bar{u}(n) \bar{u}^H(n) \quad (145)$$

donde $\Phi(n-1)$ es el valor anterior de la matriz de correlación y el producto matricial $\bar{u}(n) \bar{u}^H(n)$ desempeña un término de corrección en la operación de actualización. En forma similar, el vector de correlación cruzada entre las entradas y la respuesta deseada puede escribirse recursivamente como:

$$\begin{aligned} \bar{\theta}(n) &= \lambda \left[\sum_{i=1}^{n-1} \lambda^{n-1-i} \bar{u}(i) d^*(i) \right] + \bar{u}(n) d^*(n) \\ \bar{\theta}(n) &= \lambda \bar{\theta}(n-1) + \bar{u}(n) d^*(n) \end{aligned} \quad (146)$$

De acuerdo con las ecuaciones normales, para calcular el estimado de mínimos cuadrados $\hat{w}(n)$ para el vector de pesos, es necesario determinar la inversa de la matriz de correlación i.e., $\hat{w}(n) = \Phi^{-1}(n) \bar{\theta}(n)$. Sin embargo, en la práctica, se evita tal operación debido a que puede consumir demasiado tiempo, particularmente si el número de pesos del filtro es elevado. Por otra parte, el interés radica en obtener un algoritmo recursivo y en consecuencia, calcular $\hat{w}(n)$ de manera recursiva para $n = 1, 2, \dots, \infty$. Afortunadamente, asumiendo que las condiciones iniciales han sido elegidas para asegurar la no singularidad de

$\Phi(n)$, es posible alcanzar ambos objetivos empleando un resultado básico del álgebra matricial conocido como el *lema de inversión de matrices* [41, 62-64]. Así, de acuerdo al lema de inversión de matrices, la inversa de la matriz de correlación está dada por:

$$\Phi^{-1}(n) = \lambda^{-1}\Phi^{-1}(n-1) - \frac{\lambda^{-2}\Phi^{-1}(n-1)\bar{u}(n)\bar{u}^H(n)\Phi^{-1}(n-1)}{1 + \lambda^{-1}\bar{u}^H(n)\Phi^{-1}(n-1)\bar{u}(n)} \quad (147)$$

Si se define a $P(n) = \Phi^{-1}(n)$ y además:

$$\bar{k}(n) = \frac{\lambda^{-1}P(n-1)\bar{u}(n)}{1 + \lambda^{-1}\bar{u}^H(n)P(n-1)\bar{u}(n)} \quad (148)$$

entonces, la ecuación recursiva (147) se puede escribir como:

$$P(n) = \lambda^{-1}P(n-1) - \lambda^{-1}\bar{k}(n)\bar{u}^H(n)P(n-1) \quad (149)$$

que corresponde a la ecuación de Riccati para el algoritmo RLS. A la matriz $P(n)$, de dimensiones $M \times M$, se le conoce como la matriz de correlación inversa, mientras que al vector $\bar{k}(n)$, de dimensiones $M \times 1$, se le refiere como el vector de ganancia.

Es posible reordenar la ecuación para el vector de ganancia de la siguiente manera:

$$\begin{aligned} \bar{k}(n) &= \frac{\lambda^{-1}P(n-1)\bar{u}(n)}{1 + \lambda^{-1}\bar{u}^H(n)P(n-1)\bar{u}(n)} \\ \bar{k}(n) &= \lambda^{-1}P(n-1)\bar{u}(n) - \lambda^{-1}\bar{k}(n)\bar{u}^H(n)P(n-1)\bar{u}(n) \\ \bar{k}(n) &= \left[\lambda^{-1}P(n-1) - \lambda^{-1}\bar{k}(n)\bar{u}^H(n)P(n-1) \right] \bar{u}(n) \end{aligned} \quad (150)$$

donde la expresión dentro del paréntesis es igual a $P(n)$. Por lo tanto, la expresión del vector de ganancia se simplifica de la forma:

$$\begin{aligned} \bar{k}(n) &= P(n)\bar{u}(n) \\ \bar{k}(n) &= \Phi^{-1}(n)\bar{u}(n) \end{aligned} \quad (151)$$

que corresponde al vector de entrada $\bar{u}(n)$ transformado por la inversa de la matriz de correlación $\Phi(n)$.

A partir de las ecuaciones normales, de la definición de $P(n)$, y de las recursiones para $\theta(n)$ y $P(n)$, el valor óptimo vector de pesos $\hat{w}(n)$ al tiempo n esta dado por:

$$\begin{aligned} \hat{w}(n) &= \Phi^{-1}(n)\bar{\theta}(n) \\ \hat{w}(n) &= P(n)\bar{\theta}(n) \\ \hat{w}(n) &= P(n)\lambda\bar{\theta}(n-1) + P(n)\bar{u}(n)d^*(n) \\ \hat{w}(n) &= \lambda \left[\lambda^{-1}P(n-1) - \lambda^{-1}\bar{k}(n)\bar{u}^H(n)P(n-1) \right] \bar{\theta}(n-1) + P(n)\bar{u}(n)d^*(n) \\ \hat{w}(n) &= P(n-1)\bar{\theta}(n-1) - \bar{k}(n)\bar{u}^H(n)P(n-1)\bar{\theta}(n-1) + P(n)\bar{u}(n)d^*(n) \end{aligned} \quad (152)$$

donde $P(n-1)\bar{\theta}(n-1) = \hat{w}(n-1)$ y $\bar{k}(n) = P(n)\bar{u}(n)$, por lo que:

$$\begin{aligned}\hat{w}(n) &= \hat{w}(n-1) - \bar{k}(n)\bar{u}^H(n)\hat{w}(n-1) + \bar{k}(n)d^*(n) \\ \hat{w}(n) &= \hat{w}(n-1) + \bar{k}(n)[d^*(n) - \bar{u}^H(n)\hat{w}(n-1)]\end{aligned}\quad (153)$$

Finalmente, la ecuación recursiva para la actualización del vector de pesos del filtro es:

$$\hat{w}(n) = \hat{w}(n-1) + \bar{k}(n)\alpha^*(n) \quad (154)$$

donde $\alpha(n)$, denominada la secuencia de *innovación*, está definida por:

$$\alpha(n) = d(n) - \hat{w}^H(n-1)\bar{u}(n) \quad (155)$$

que también se conoce como *error de estimación a priori* debido a que el producto interno $\hat{w}^H(n-1)\bar{u}(n)$ representa un estimado de la respuesta deseada $d(n)$, basado en el estimado del vector de pesos del filtro realizado al tiempo $n-1$. En general, la innovación $\alpha(n)$ es diferente del *error de estimación a posteriori* $e(n) = d(n) - \hat{w}^H(n)\bar{u}(n)$ cuyo cálculo involucra el estimado *actual* del vector de pesos del filtro disponible al tiempo n . Por lo tanto, $\alpha(n)$ puede verse como un valor tentativo de $e(n)$ antes de actualizar el vector de pesos del filtro.

La ecuación recursiva para la matriz de correlación inversa, $P(n) = \lambda^{-1}P(n-1) - \lambda^{-1}\bar{k}(n)\bar{u}^H(n)P(n-1)$ requiere inicializarse con un valor $P(0)$ que asegure la no singularidad de la matriz de correlación $\Phi(n)$. Para ello, puede evaluarse la matriz inversa:

$$\left[\sum_{i=-n_0}^0 \lambda^{-i} \bar{u}(i)\bar{u}^H(i) \right]^{-1} \quad (156)$$

donde el vector de entrada $\bar{u}(i)$ se obtiene de un bloque inicial de datos para $-n_0 \leq i \leq 0$. Sin embargo, resulta más sencillo modificar la expresión de $\Phi(n)$ manteniendo intactas las recursiones del algoritmo RLS, de la forma:

$$\Phi(n) = \sum_{i=1}^n \lambda^{n-i} \bar{u}(i)\bar{u}^H(i) + \delta\lambda^n I \quad (157)$$

donde I es la matriz identidad, de tamaño $M \times M$ y δ es una constante positiva pequeña. Para $n=0$, la matriz de correlación modificada resultante es $\Phi(0) = \delta I$ y el valor inicial de $P(n) = \Phi^{-1}(n)$ es $P(0) = \delta^{-1}I$. La inicialización anterior equivale a forzar el dato

desconocido $u(-M+1)$ al valor de $\lambda^{\frac{(-M+1)}{2}} \delta^{\frac{1}{2}}$ en lugar de considerarlo cero, modificando el método de pre-ventaneo en:

$$u(n) = \begin{cases} \lambda^{\frac{(-M+1)}{2}} \delta^{\frac{1}{2}}, & n = -M+1 \\ 0, & n < 0, \quad n \neq -M+1 \end{cases} \quad (158)$$

Para un filtro transversal con M pesos, el índice $n = -M+1$ se refiere al último peso del filtro. Cuando la primera muestra de datos diferente de cero, $u(i)$, ingresa al filtro la entrada de inicialización $u(-M+1)$ abandona el filtro y de ahí en adelante el algoritmo RLS inicia su funcionamiento.

Por otra parte, el valor inicial para el vector de pesos del filtro se escoge como $\bar{w}(0) = \bar{0}$, donde $\bar{0}$ es un vector nulo de tamaño $M \times 1$. El procedimiento de inicialización que incorpora las ecuaciones $P(0) = \delta^{-1}I$ y $\bar{w}(0) = \bar{0}$ se conoce como una inicialización con restricciones suaves donde únicamente se requiere la constante positiva δ [62, 63]. La elección recomendada de δ es un valor pequeño es decir $0.01\sigma_u^2$, donde σ_u^2 es la varianza de la muestra de datos $\bar{u}(n)$. La elección se basa en la experiencia práctica con el algoritmo RLS y está soportada por un análisis estadístico del algoritmo con una inicialización con restricciones suaves [65, 66]. Sin embargo, cuando se emplean datos de longitud considerable, el valor exacto de la constante de inicialización δ tiene un efecto insignificante [63].

En resumen, el algoritmo de mínimos cuadrados recursivo queda constituido por las ecuaciones mostradas en la Tabla IV.IV. En particular, la ecuación correspondiente al error de estimación a priori, $\alpha(n)$, describe la operación de filtrado del algoritmo. Por su parte, la ecuación correspondiente al estimado del vector de pesos del filtro, $\hat{w}(n)$, describe la operación adaptable del algoritmo, donde $\hat{w}(n)$ es actualizado mediante un incremento en su valor anterior por una cantidad igual al complejo conjugado del error de estimación a priori multiplicado por el vector de ganancia variante en el tiempo $\bar{k}(n)$. Las ecuaciones restantes permiten actualizar el valor del vector de ganancias.

TABLA IV.IV. EL ALGORITMO RLS.

Inicializar el algoritmo ajustando

$$P(0) = \delta^{-1}I$$

$$\hat{w}(0) = \vec{0}$$

Para cada instante de tiempo, $n = p, \dots, N-1$, calcular

$$\bar{k}(n) = \frac{\lambda^{-1}P(n-1)\bar{u}(n)}{1 + \lambda^{-1}\bar{u}^H(n)P(n-1)\bar{u}(n)}$$

$$\alpha(n) = d(n) - \hat{w}^H(n-1)\bar{u}(n)$$

$$\hat{w}(n) = \hat{w}(n-1) + \bar{k}(n)\alpha^*(n)$$

$$P(n) = \lambda^{-1}P(n-1) - \lambda^{-1}\bar{k}(n)\bar{u}^H(n)P(n-1)$$

IV.4.4.2.2. Características.

Una característica importante del algoritmo RLS descrito por este conjunto de ecuaciones radica en que la inversión de la matriz de correlación, $\Phi(n)$, es reemplazada por una simple división escalar en cada paso del algoritmo. Además, el algoritmo RLS presenta una rápida velocidad de convergencia comparado con otros algoritmos, a expensas de un alto costo computacional. Un análisis detallado del número de operaciones realizadas por el algoritmo para señales reales y complejas se encuentra en [67]. En la figura IV.24, se esquematiza el algoritmo RLS.

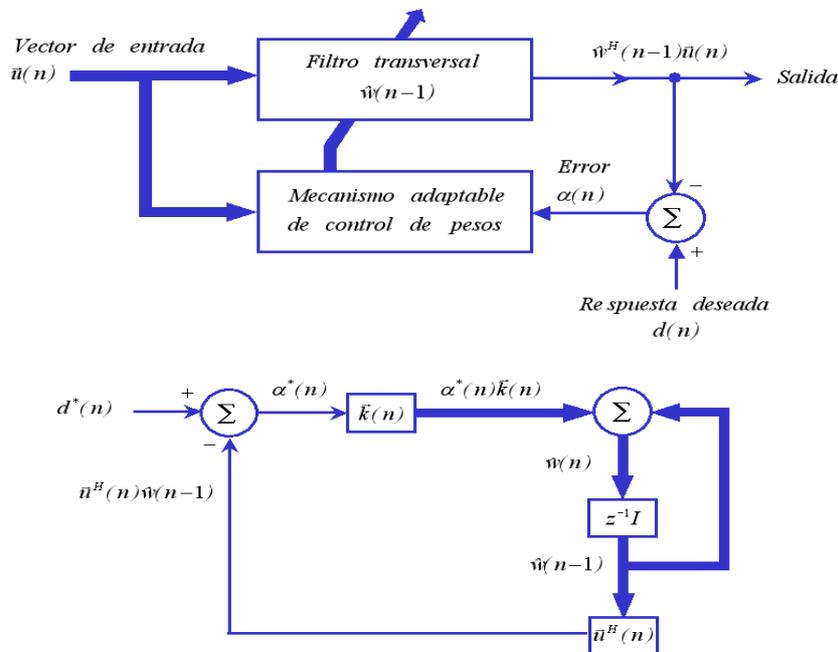


Fig. IV.24. Algoritmo recursivo de mínimos cuadrados, RLS. a) Diagrama de bloques. b) Diagrama de flujo de la señal.

Se ha estudiado la convergencia del algoritmo RLS en un ambiente estacionario y no estacionario. En un *ambiente estacionario*, la superficie del error de desempeño se encuentra fija y el objetivo consiste en encontrar, paso a paso, el punto mínimo de esa superficie y con ello asegurar el desempeño óptimo [52]. Se puede demostrar que el algoritmo converge a la solución de Wiener \bar{w}_o en el valor esperado del vector de pesos conforme se incrementa el valor de n , es decir:

$$E[\hat{w}(n)] = \bar{w}_o + \bar{b}(n) \quad (159)$$

donde el sesgo:

$$\bar{b}(n) \approx -\frac{\delta}{n} \Phi^{-1}(n) \bar{w}_o \quad (160)$$

se aproxima a cero para n grandes, y por lo tanto el algoritmo RLS produce un estimado asintóticamente sin sesgo del vector de parámetros de regresión \bar{w}_o . Además, también existe convergencia en el error cuadrático medio del vector de pesos. Se puede demostrar que:

$$\text{tr}\left\{E\left[\left(\hat{w}(n) - \bar{w}_o\right)\left(\hat{w}(n) - \bar{w}_o\right)^H\right]\right\} = \frac{\sigma^2}{n - M - 1} \sum_{i=1}^M \frac{1}{\lambda_i}, \quad n > M + 1 \quad (161)$$

donde $\sigma^2 = E[e_o(i)e_o^*(j)]$ para $i = j$, e_o corresponde al error de medición de los datos, y λ_i son los eigenvalores de la matriz R de correlación promedio del ensamble. En consecuencia, el error cuadrático medio del vector de pesos $\hat{w}(n)$ es amplificado por la inversa del eigenvalor más pequeño λ_{min} . Entonces, en una primera aproximación, la sensibilidad del algoritmo RLS a la dispersión de los eigenvalores esta inicialmente determinada en proporción a la inversa de λ_{min} . Por ello, los problemas de mínimos cuadrados mal condicionados propician malas propiedades de convergencia. Además, el error cuadrático medio del vector de pesos $\hat{w}(n)$ decae casi linealmente con el número de iteraciones, n . Por lo tanto, el estimado $\hat{w}(n)$ producido por el algoritmo RLS para el vector de pesos converge en la norma (i.e., la media cuadrática) al vector de parámetros \bar{w}_o del modelo de regresión lineal múltiple de forma casi lineal con el tiempo.

El desempeño del algoritmo RLS se puede establecer con base a la curva de aprendizaje del promedio de ensamble sobre el error de estimación a priori $\alpha(n)$, es decir:

$$J'(n) = E\left[|\alpha(n)|^2\right] \quad (162)$$

donde para el modelo de regresión lineal

$$\begin{aligned}\alpha(n) &= d(n) - \hat{w}^H(n-1)\bar{u}(n) \\ \alpha(n) &= e_o(n) - [\hat{w}(n-1) - \bar{w}_o]^H \bar{u}(n)\end{aligned}\quad (163)$$

conduce finalmente a:

$$J'(n) = \sigma^2 + \frac{M\sigma^2}{n-M-1}, \quad n > M+1 \quad (164)$$

El resultado anterior conduce a las siguientes conclusiones [52, 63]:

- La curva de aprendizaje del algoritmo RLS converge en alrededor de $2M$ iteraciones, donde M es el número de pesos del filtro transversal.
- Conforme el número de iteraciones, n , se aproxima a infinito, el error cuadrático medio $J'(n)$ se aproxima a un valor final igual a la varianza σ^2 del error de medición $e_o(n)$. Es decir, cuando el algoritmo RLS opera en un ambiente estacionario se obtiene la solución óptima de Wiener [68].

En *condiciones no estacionarias* el punto mínimo de la superficie de error de desempeño ya no se encuentra fijo, y como consecuencia el algoritmo de filtrado adaptable ahora tiene la tarea adicional de seguir el punto mínimo [69]. En otras palabras, el algoritmo requiere seguir continuamente las variaciones estadísticas de los datos de entrada, cuya ocurrencia se asume lo suficientemente lenta para ser posible. En el caso no estacionario, la señal de error $e(n)$ involucrada en el proceso adaptable, se define mediante:

$$\begin{aligned}e(n) &= d(n) - y(n) \\ e(n) &= \bar{w}_o^H(n)\bar{u}(n) + v(n) - \hat{w}^H(n)\bar{u}(n)\end{aligned}\quad (165)$$

donde $\hat{w}(n)$ es el vector de pesos del filtro transversal adaptable, y $v(n)$ es el ruido de medición considerado como un proceso de ruido blanco con media cero y varianza σ_v^2 . El vector de pesos $\bar{w}_o(n)$ representa el objetivo a seguir por el filtro adaptable. En consecuencia, el vector de error de pesos se establece como:

$$\bar{\epsilon}(n) = \hat{w}(n) - \bar{w}_o(n) \quad (166)$$

A partir del vector de error de pesos o de su matriz de correlación, $K(n) = E[\bar{\epsilon}(n)\bar{\epsilon}^H(n)]$, se establecen dos figuras de mérito para valorar la capacidad de seguimiento del algoritmo RLS: a) la desviación cuadrática media $D(n)$, y b) el desajuste $M(n)$. En general, $D(n)$ y $M(n)$ dependen del número de iteraciones y ambas resaltan diferentes aspectos del problema de seguimiento de una forma complementaria. Para el algoritmo RLS, y utilizando la teoría de

Markov de primer orden para describir los cambios del vector $\bar{w}_0(n)$, el vector de error de pesos se puede establecer como:

$$\bar{\epsilon}(n) \approx \lambda \bar{\epsilon}(n-1) + (1-\lambda)R^{-1}\bar{u}(n)v^*(n) - \bar{\omega}(n) \quad , \quad \text{para } n \text{ grande} \quad (167)$$

donde R^{-1} es la inversa de la matriz de correlación de promedio de ensamble del vector de entrada, $\bar{\omega}(n)$ es el vector de ruido del proceso con media cero y matriz de correlación Q , y λ es el factor de olvido, que se considera cercano a la unidad. En consecuencia, la matriz de correlación de $\bar{\epsilon}(n)$ esta dada por:

$$K(n) \approx \frac{1-\lambda}{2} \sigma_v^2 R^{-1} + \frac{1}{2(1-\lambda)} Q \quad , \quad \text{para } n \text{ grande} \quad (168)$$

Por lo tanto, la desviación cuadrática media, entre el vector de pesos actual \bar{w}_o del sistema dinámico desconocido y el vector de pesos ajustado $\hat{w}(n)$ del filtro adaptable, definida como

$$\begin{aligned} D(n) &= E\left[\|\bar{\epsilon}(n)\|^2\right] \\ D(n) &= \text{tr}[K(n)] \end{aligned} \quad (169)$$

y sustituyendo la ecuación (168), finalmente se obtiene:

$$D(n) \approx \frac{1-\lambda}{2} \sigma_v^2 \text{tr}[R^{-1}] + \frac{1}{2(1-\lambda)} \text{tr}[Q] \quad , \quad \text{para } n \text{ grande} \quad (170)$$

donde el primer término es la varianza de la estimación debida al ruido de medición $v(n)$ y el segundo término es la varianza debida al vector de ruido de medición $\bar{\omega}(n)$. Estas dos contribuciones varían en proporción a $(1-\lambda)$ y $(1-\lambda)^{-1}$, respectivamente. Por lo tanto, el valor óptimo del factor de olvido, λ_{opt} , ocurre cuando estas dos contribuciones son iguales, y corresponde a

$$\lambda_{opt} \approx 1 - \frac{1}{\sigma_v} \left(\frac{\text{tr}[Q]}{\text{tr}[R^{-1}]} \right)^{\frac{1}{2}} \quad (171)$$

Para obtener un buen desempeño se requiere que $D(n)$ sea pequeño. Sustituyendo la ecuación anterior obtenemos que la desviación cuadrática media mínima está dada por:

$$D_{min} \approx \sigma_v \left(\text{tr}[R^{-1}] \text{tr}[Q] \right)^{\frac{1}{2}} \quad (172)$$

En cuanto al desajuste se puede demostrar que el valor óptimo del factor de olvido variable, λ_{opt} , ocurre cuando:

$$\lambda_{opt} \approx 1 - \frac{1}{\sigma_v} \left(\frac{\text{tr}[RQ]}{M} \right)^{\frac{1}{2}} \quad (173)$$

y que conduce a un valor mínimo del desajuste correspondiente a

$$M_{\min} = \frac{(M \text{tr}[RQ])^{\frac{1}{2}}}{\sigma_v^2}, \quad \text{para } n \text{ grande} \quad (174)$$

Como ya se ha mencionado, el término $(1 - \lambda)^{-1}$ proporciona una medida burda de la memoria del algoritmo RLS. Por lo tanto, conforme el factor de olvido se reduce el desajuste se incrementa. Es decir, una rápida velocidad de adaptación del algoritmo RLS, en general, resulta en un proceso adaptable más ruidoso [63].

Finalmente, en ambientes estacionarios o no estacionarios, cuando se emplea un valor de $\lambda \neq 1$ el algoritmo RLS con ponderación exponencial, y en general los estimadores que emplean una ventana finita, causan ruido en la estimación del vector de pesos. Lo anterior provoca alejarse de la solución óptima o de la solución de Wiener [68].

IV.4.4.2.3. Obtención de la representación tiempo-frecuencia.

Considerando que se requiere la representación tiempo-frecuencia paramétrica del proceso estocástico $\{x(n)\}$, $0 \leq n \leq N-1$, con base a p coeficientes de un modelo autorregresivo, entonces, el vector de entrada al filtro de predicción lineal hacia delante con una estructura transversal, ver figura IV.25, al tiempo n corresponde al conjunto de muestras pasadas $\{x(n-k)\}_{k=1}^p$, i.e. $\bar{x}(n-1) = [x(n-1), x(n-2), \dots, x(n-p)]^T$, y la respuesta deseada corresponde a la muestra actual $d(n) = x(n)$. Por lo tanto, mediante la comparación de la definición del error de estimación a posteriori y de la ecuación recursiva, el algoritmo RLS determina los parámetros del filtro de predicción que se relacionan con los coeficientes del modelo del modelo TVAR de la forma:

$$\begin{aligned} w_0(n) &= a_0(n) = 1 \\ \hat{w}_k(n) &= -\hat{a}_k(n), \quad k = 1, \dots, p \end{aligned} \quad (175)$$

Lo anterior, permite obtener una representación tiempo-frecuencia mediante:

$$P_{RLS}(n, f) = \frac{\sigma_w^2(n)}{\left| 1 + \sum_{k=1}^p \hat{a}_k(n) e^{-j2\pi f k} \right|^2} \quad (176)$$

En el caso del algoritmo RLS trabajando como un filtro de predicción lineal de orden p , puede obtenerse una representación tiempo-frecuencia para instantes de tiempo en el intervalo $p \leq n \leq N-1$, donde el límite inferior refleja el primer instante de tiempo donde el algoritmo proporciona una salida.

Finalmente, la inicialización del algoritmo al tiempo $n=0$ se realiza empleando restricciones suaves, y el valor del factor de olvido se mantiene fijo a lo largo de todo el proceso de estimación. Sin embargo, se han desarrollado esquemas que permiten modificar, o actualizar, el factor de olvido del algoritmo RLS para contender con las variaciones estadísticas de la señal cuando esta cambia de un ambiente estacionario a uno no estacionario, y viceversa.

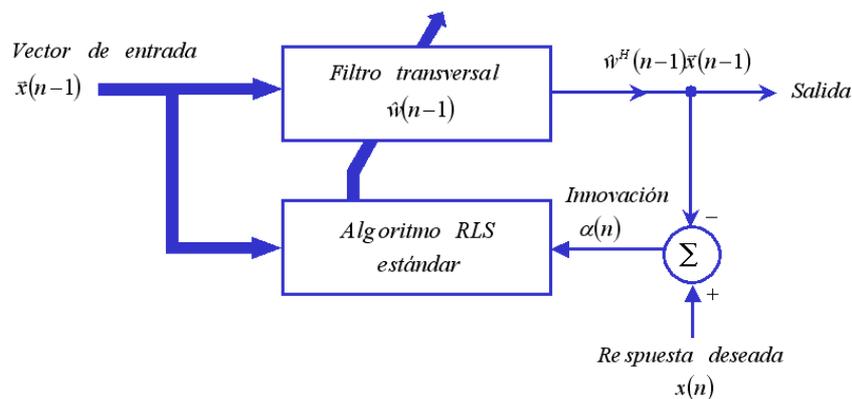


Fig. IV.25. Empleo del algoritmo RLS para la estimación de los parámetros de un modelo $TVAR(p)$ mediante el empleo de un filtro de predicción lineal hacia adelante con estructura transversal.

IV.4.4.2.4. Comparación con el método de Burg.

El método de mínimos cuadrados es particularmente interesante para determinar la representación tiempo-frecuencia de una serie de tiempo. Además, el método basado en mínimos cuadrados puede considerarse una alternativa al algoritmo de Burg para resolver problemas de modelado AR. Sin embargo, existen tres diferencias básicas entre ambos procedimientos:

1. El algoritmo de mínimos cuadrados estima los coeficientes de un modelo transversal equivalente para los datos de entrada, mientras que el algoritmo de Burg estima los coeficientes de reflexión de un modelo lattice.
2. En el método de mínimos cuadrados, no se hacen consideraciones acerca de la estadística de los datos de entrada mientras que el algoritmo de Burg, asume que los

datos de entrada constituyen un proceso débilmente estacionario. Además, el algoritmo basado en mínimos cuadrado no sufre del problema de separación de líneas espectrales.

3. El algoritmo de Burg produce una solución de fase mínima en el sentido de que los coeficientes de reflexión del predictor de lattice poseen una magnitud menor o igual a la unidad. En contraparte, el algoritmo de mínimos cuadrados no garantiza tal solución, sin embargo en la estimación espectral la falta de una solución de fase mínima no es de particular interés.

IV.4.4.3. El algoritmo de mínimos cuadrados recursivo con factor de olvido variable.

Es posible representar las señales no estacionarias mediante modelos TVAR donde los coeficientes $\{a_k(n)\}_{k=1}^p$ son representados mediante una combinación lineal de un pequeño número de funciones de tiempo conocidas de antemano, las denominadas *funciones base* [49]. Aunque esta aproximación ofrece la ventaja de guiar al mismo tipo de procedimientos de identificación que los modelos AR con parámetros constantes, o invariantes en el tiempo, aun se requiere mucho trabajo para desarrollar un procedimiento sistemático para diseñar las mejores funciones base para clases específicas de señales no estacionarias debido a que un conjunto base dado no significa que sea el mejor para todas las señales no estacionarias

Otra aproximación consiste en considerar representaciones de señales no estacionarias usando *factores de olvido variables*, VFF. El concepto de VFF fue introducido en la teoría de control para evitar un control inestable en ambientes que requieren un auto-ajuste [70-72]. A continuación se presenta la formulación de un algoritmo de mínimos cuadrados recursivo con factor de olvido variable, RLS-VFF [73, 74]. Sin embargo, se iniciará por recordar el algoritmo de mínimos cuadrados estándar para la estimación de los parámetros AR haciendo especial énfasis en la función que desempeña el factor de olvido.

El algoritmo RLS estándar posee tres posibles fuentes de error. El primero de ellos surge del ruido, e. g. el ruido de medición. El segundo error de estimación es debido al empleo de datos finitos; este error posee una varianza que decrece con la longitud de la ventana. El tercer error es causado por el empleo de datos no estacionarios y posee una varianza que se incrementa con la longitud de los datos; este error podría ser minimizado

excluyendo rápidamente los datos pasados y basando los estimados de manera predominante en los datos más recientes [68, 73, 75].

En el caso estacionario los parámetros pueden estimarse con un factor de olvido unitario, $\lambda = 1$ mientras que en el caso no estacionario los parámetros TVAR son calculados empleando el mismo valor $\lambda < 1$ para cada muestra del proceso y en consecuencia, la estimación espectral podría no ser estadísticamente óptima. Es decir, dado un valor de λ pequeño se puede estimar rápidamente la tendencia global de una señal no estacionaria pero a expensas de una mayor varianza debido a una menor cantidad de datos disponibles. Siguiendo la misma idea, con un valor de λ cercano a la unidad, al algoritmo RLS le tomará un tiempo relativamente largo estimar los parámetros correctos, pero eventualmente estos serán estimados con gran precisión cuando la señal experimente características estacionarias [73].

La velocidad de adaptación del algoritmo RLS está determinada por la longitud de la memoria asintótica que se define como:

$$N = \frac{1}{1 - \lambda} \quad (177)$$

En el caso estacionario, el espectro es estimado a partir de una longitud de memoria N de la señal. Si consideramos que una señal estacionaria esta compuesta por varias sub-señales con diferentes longitudes de memoria que varían entre una longitud de memoria mínima N_{min} y una longitud de memoria máxima N_{max} , entonces el espectro variante en el tiempo óptimo debería obtenerse asignando el factor de olvido adecuado con valor entre λ_{min} y λ_{max} , para cada sub-señal. Sin embargo, en la práctica no es realista conocer de antemano las longitudes de memoria ni los puntos de inicio de cada sub-señal. Por lo tanto, es necesario estimar el grado de no estacionaridad de la señal para calcular el siguiente valor del factor de olvido. Para alcanzar este objetivo es posible definir el error de predicción extendido como [73, 74]:

$$Q(n) = \frac{1}{H} \sum_{i=0}^{H-1} e^2(n-i) \quad (178)$$

Dado que el error proveniente del ruido es un proceso aleatorio, una promediación adecuada del error de predicción resulta útil para minimizar el efecto de un gran ruido aditivo espurio. El parámetro H debe ser pequeño en comparación con la memoria asintótica mínima, N_{min} , de tal manera que la promediación no oscurezca la no estacionaridad de la señal. Una estrategia para elegir el factor de olvido variable puede definirse mediante:

$$\lambda(n) = 1 - \frac{1}{N(n)} \quad (179)$$

donde

$$N(n) = \frac{\sigma_e^2 N_{max}}{Q(n)} \quad (180)$$

con σ_e^2 igual a la varianza del ruido esperado basada en el conocimiento real del proceso. Entonces, la longitud de memoria asintótica máxima N_{max} controlará la velocidad de adaptación. Para un proceso estacionario, el error de predicción extendido se aproximará a la varianza del proceso y por lo tanto la longitud de memoria asintótica $N(n)$ se aproximará a N_{max} . Dado que este esquema de adaptación del factor de olvido no garantiza que la $\lambda(n)$ producida no sea negativa, debe imponerse un límite inferior λ_{min} . Finalmente, empleando las ecuaciones anteriores, el factor de olvido variable $\lambda(n)$ se actualiza para cada iteración del algoritmo RLS mediante:

$$\lambda(n) = 1 - \frac{\frac{1}{H} \sum_{i=0}^{H-1} e^2(n-i)}{\sigma_e^2 N_{max}} \quad (181)$$

Observe que el esquema de adaptación del factor de olvido depende del conocimiento *a priori* de la varianza del ruido esperado σ_e^2 .

En la tabla IV.V se presenta el algoritmo RLS modificado que incluye el esquema de actualización del factor de olvido.

IV.4.4.3.1. Comparación con el algoritmo RLS estándar.

A continuación se presenta la aplicación del algoritmo RLS estándar y del algoritmo RLS con factor de olvido variable para una señal de prueba. La señal de prueba consiste en 600 muestras de un proceso estocástico generado mediante un proceso autorregresivo de segundo orden definido por la siguiente ecuación de diferencias recursiva:

$$x(n) + \sum_{k=1}^2 a_k(n)x(n-k) = w(n)$$

donde

$$a_1(n) = \begin{cases} -1.618 & , \quad 0 < n < 299 \\ 0.618 & , \quad 300 < n < 599 \end{cases}$$

$$a_2(n) = -1$$

y $w(n)$ es un proceso de ruido blanco con media cero y varianza unitaria. La señal fue generada empleando una frecuencia de muestreo de 1 Hz. Los coeficientes y la señal de salida del modelo TVAR se muestran en la Fig. IV.26. Observe que el parámetro $a_1(n)$, y por lo tanto la estadística de la señal, cambia abruptamente a la mitad de la duración de la señal en la muestra $n = 300$.

TABLA IV.V. EL ALGORITMO RLS-VFF.

Inicializar el algoritmo

$$P(0) = \delta^{-1} I$$

$$\bar{w}(0) = \bar{0}$$

$$\lambda(0) = \lambda_0$$

Establecer

$$\lambda_{min}, N_{max}, H, \sigma_e^2$$

Para cada instante de tiempo, $n = p, \dots, N-1$, calcular

$$\bar{k}(n) = \frac{\lambda^{-1} P(n-1) \bar{u}(n)}{1 + \lambda^{-1} \bar{u}^H(n) P(n-1) \bar{u}(n)}$$

$$\alpha(n) = d(n) - \hat{w}^H(n-1) \bar{u}(n)$$

$$\hat{w}(n) = \hat{w}(n-1) + \bar{k}(n) \alpha^*(n)$$

$$P(n) = \lambda^{-1} P(n-1) - \lambda^{-1} \bar{k}(n) \bar{u}^H(n) P(n-1)$$

Para $n = H, \dots, N-1$, actualizar el factor de olvido

$$\lambda(n) = 1 - \frac{1}{H} \sum_{i=0}^{H-1} e^2(n-i)$$

$$\lambda(n) = 1 - \frac{\sigma_e^2 N_{max}}{\sigma_e^2 N_{max}}$$

Si $\lambda(n) < \lambda_{min}$

$$\lambda(n) = \lambda_{min}$$

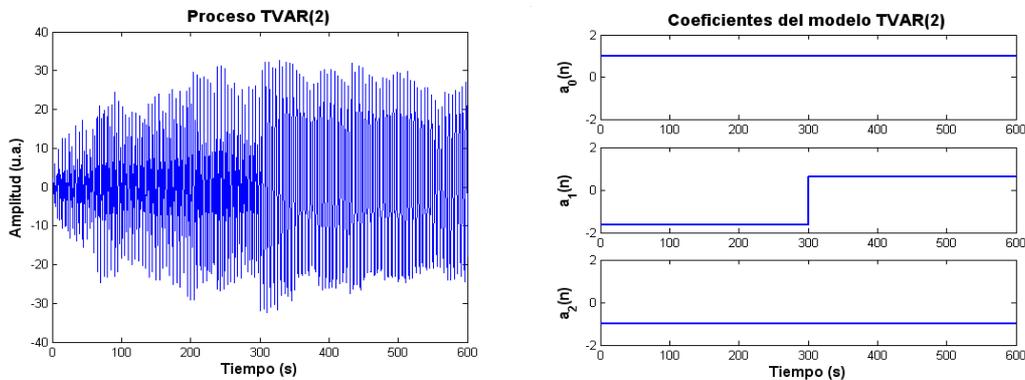


Fig. IV.26. Señal usada para la comparación entre los algoritmos RLS estándar y RLS con factor de olvido variable. a) Proceso estocástico $TVAR(2)$. b) Variación temporal de los coeficientes AR.

El algoritmo RLS estándar requiere de la selección del factor de olvido λ que se mantiene constante. Para propósitos ilustrativos se emplean los valores de $\lambda = 1$, que corresponde a una memoria infinita, y $\lambda = 0.75$. Los coeficientes estimados se presentan en la figura IV.27. Al emplear el algoritmo RLS estándar con un factor de olvido constante $\lambda = 1$, se observa que los parámetros TVAR son estimados con una gran precisión, sin embargo el algoritmo requiere un gran número de iteraciones para ajustarse al cambio abrupto de la estadística de la señal ocurrido a la mitad del registro, ver el comportamiento de la señal en color azul en la figura IV.27.a. En este caso, la velocidad de adaptación del algoritmo es demasiado lenta, tanto que no alcanza a converger al valor real los parámetros TVAR dentro de la duración del registro. En contraste, al emplear un factor de olvido pequeño $\lambda = 0.75$, la velocidad de adaptación del algoritmo es lo suficientemente alta para estimar la tendencia global debido al empleo de una menor cantidad de datos considerados por el factor de ponderación exponencial. Sin embargo, el incremento en la velocidad de adaptación se realiza a expensas de un incremento en la varianza de los estimados de ambos parámetros TVAR, observe las variaciones en la línea de base en color azul desde $n = 0$ hasta $n = 300$. En general, para el algoritmo RLS es difícil de determinar de antemano la elección del valor más adecuado del factor de olvido constante a lo largo de todo el registro, y por lo tanto para un valor de λ existirá un compromiso entre la velocidad de adaptación y la varianza de los estimados.

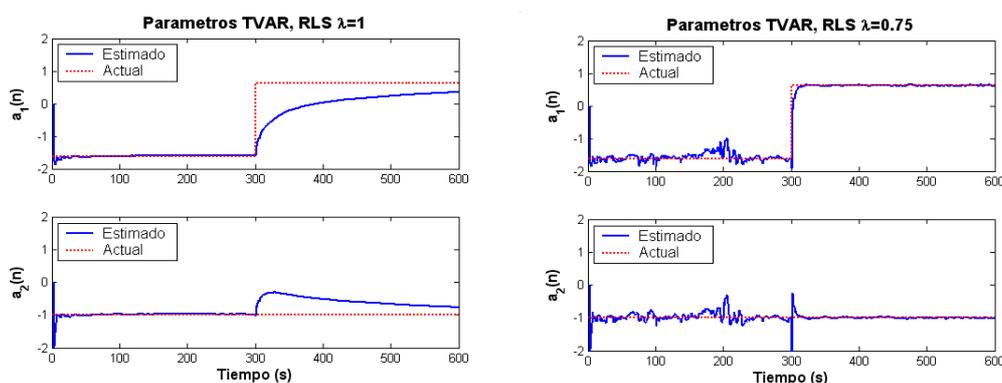


Fig. IV.27. Estimación de los parámetros del modelo $TVAR(2)$ mediante el algoritmo RLS estándar.

a) $\lambda = 1$, y b) $\lambda = 0.75$.

En la figura IV.28, se muestra como el error de predicción extendido detecta el cambio en la estacionaridad de la señal al tiempo $n = 300$, lo cual conduce a una disminución del factor de olvido con el propósito de emplear una menor cantidad de datos y ajustarse con una

mayor velocidad a los cambios en la estadística de la señal. Como se mencionó anteriormente, este esquema de actualización no asegura que el factor de olvido sea positivo, por lo tanto es necesario establecer un límite inferior, que se estableció en $\lambda_{min} = 0.75$. Observe que aunque el algoritmo RLS-VFF fue inicializado con un valor de $\lambda(0) = \lambda_{min}$, el algoritmo fue capaz de detectar la estacionaridad de la señal y ajustar el valor del factor de olvido a valores cercanos a la unidad y con ello obtener estimados de los parámetros TVAR con gran precisión. De manera similar, observe que al tiempo $n = 300$ el factor de olvido alcanza su límite inferior debido a la detección de una no estacionaridad significativa, logrando estimar rápidamente la tendencia global de los parámetros. Conforme los parámetros estimados se aproximan a sus valores reales, el factor de olvido se acerca nuevamente a la unidad permitiendo estimados con baja varianza. Es decir, para esta señal el algoritmo RLS-VFF presenta una gran velocidad de adaptación en una situación no estacionaria y coeficientes estimados con una baja varianza en una situación estacionaria.

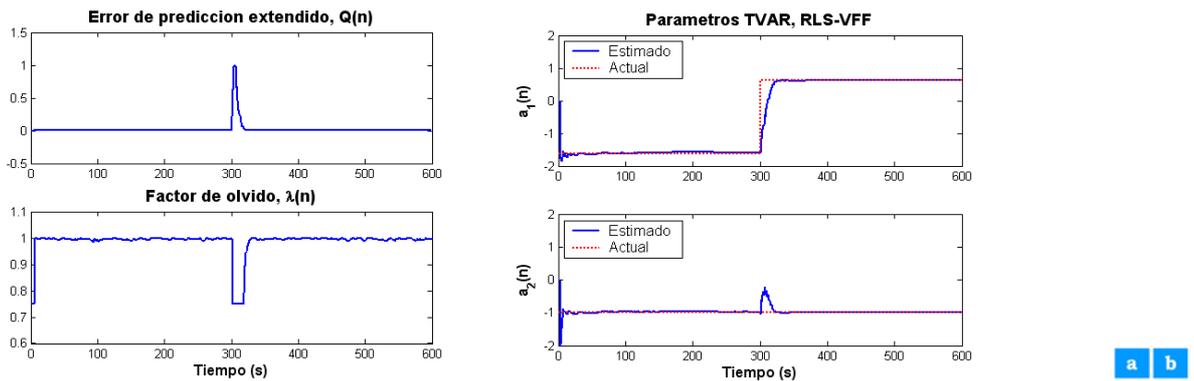


Fig. IV.28. Estimación de los parámetros del modelo $TVAR(2)$ mediante el algoritmo RLS-VFF con parámetros $H = 5$, $N_{max} = 1$ y $\sigma_e^2 = 1$. a) Detección de la estacionaridad de la señal mediante el error de predicción extendido y modificación del factor de olvido, y b) estimación de los parámetros AR.

IV.4.4.3.2. Obtención de la representación tiempo-frecuencia.

Al igual que en el caso del algoritmo RLS estándar, la disposición de los parámetros TVAR a lo largo de la duración del registro permite obtener una representación tiempo-frecuencia paramétrica mediante la aplicación de la fórmula de estimación espectral paramétrica AR con la frecuencia deseada. En particular, para un modelo TVAR de orden p , es posible estimar el espectro de potencia instantáneo para instantes de tiempo en el intervalo

$p \leq n \leq N-1$, siendo N la cantidad total de muestras del registro, para conformar la TFR mediante

$$P_{RLS-VFF}(n, f) = \frac{\sigma_w^2(n)}{\left| 1 + \sum_{k=1}^p \hat{a}_k(n) e^{-j2\pi f k} \right|^2} \quad (182)$$

Mediante el esquema de estimación en bloques del algoritmo de Burg, se obtienen estimados tanto de los parámetros AR como de la varianza del error de estimación $\sigma_w^2(n)$ para cada señal ventaneada que pueden ser aplicados directamente en la estimación espectral paramétrica. Sin embargo, también es común reemplazar el valor de $\sigma_w^2(n)$ por un vector unitario constante de longitud N , $\sigma_w^2(n) = 1$, $0 \leq n \leq N-1$, para producir la representación tiempo-frecuencia paramétrica basada en la estimación recursiva, o en bloques, de los parámetros AR [48, 49]

$$P_{RLS-VFF}(n, f) = \frac{1}{\left| 1 + \sum_{k=1}^p \hat{a}_k(n) e^{-j2\pi f k} \right|^2} = \frac{1}{|FFT[\bar{a}(n)]|^2} \quad (183)$$

donde el vector $\bar{a}(n)$ representa el vector de parámetros variantes en el tiempo del modelo $TVAR(p)$, y FFT denota la transformada rápida de Fourier.

Retomando el ejemplo comparativo entre los algoritmos RLS estándar y con factor de olvido variable, a continuación se presentan las representaciones tiempo-frecuencia obtenidas con cada algoritmo.

Mediante el empleo de identidades trigonométricas puede mostrarse que el coeficiente $TVAR$ $a_1(n)$ corresponde a:

$$a_1(n) = 2 \cos(2\pi f_i(n))$$

por lo tanto, la representación tiempo-frecuencia esperada debe presentar un decremento en la frecuencia instantánea al tiempo $n = 300$, dado por

$$f_i(n) = \begin{cases} 0.4 \text{ Hz} & , \quad 0 < n < 299 \\ 0.2 \text{ Hz} & , \quad 300 < n < 599 \end{cases}$$

En la figura IV.29, se muestra la representación tiempo-frecuencia de la señal de la figura IV.26, utilizando los algoritmos RLS estándar y con factor de olvido variable. Se observa que el algoritmo RLS-VFF para un ambiente estacionario su estimación es equiparable a la obtenida mediante el algoritmo RLS estándar con el factor de olvido fijo

$\lambda = 1$. Además, la transición entre las frecuencias instantáneas es equiparable a la obtenida mediante el algoritmo RLS estándar con el factor de olvido fijo $\lambda = 0.75$. Lo anterior es debido al empleo del error de predicción extendido que permite detectar cambios en la estacionaridad de la señal para modificar el factor de olvido o equivalentemente la memoria del algoritmo RLS. Sin embargo, en general la representación tiempo frecuencia obtenida por RLS-VFF es más cercana a la representación teórica esperada.

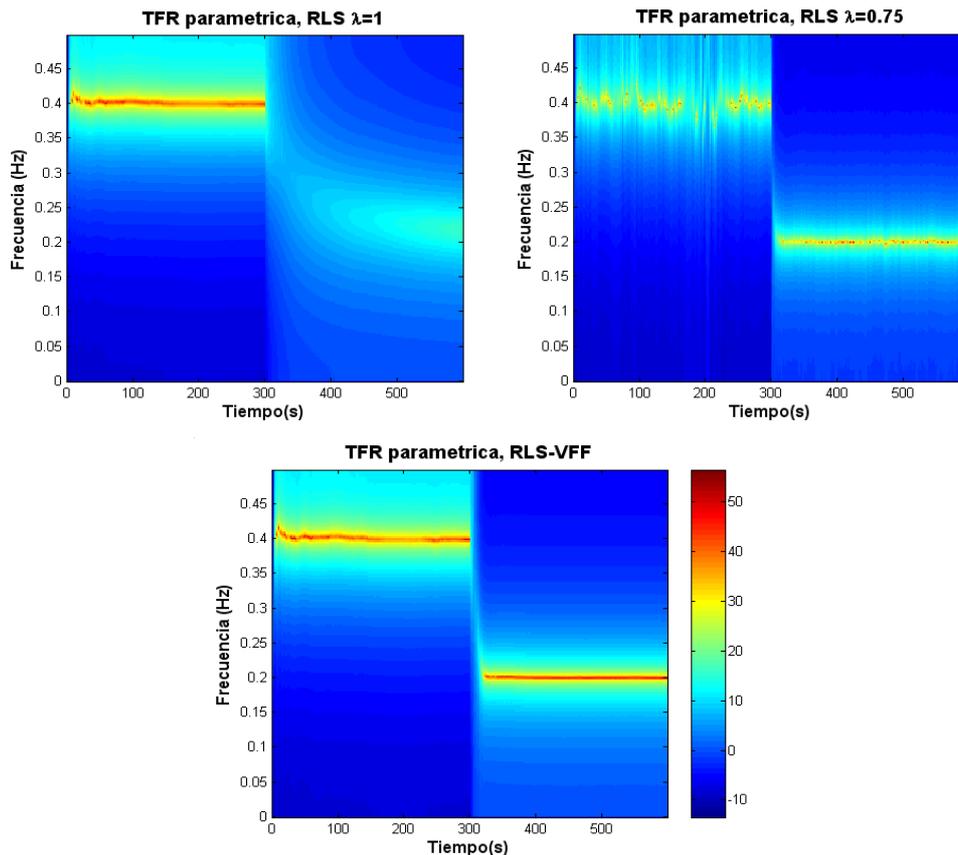


Fig. IV.29. Representaciones tiempo-frecuencia paramétricas mediante RLS y RLS-VFF. a) RLS estándar con $\lambda = 1$, b) RLS estándar con $\lambda = 0.75$, y c) RLS-VFF. En todos los casos se consideró una varianza del error de estimación unitaria, y la misma paleta de colores.

IV.4.4.4. El filtro de Kalman.

El último estimador de los coeficientes autorregresivos variantes en el tiempo empleado en esta tesis se basa en el filtro de Kalman, el cual consiste en un filtro recursivo de error cuadrático medio. El filtro fue desarrollado por Rudolph Kalman en 1959 [76], sin embargo alrededor del mismo tiempo fueron derivados de forma independiente resultados recursivos equivalentes al de Kalman por Peter Swerling [77] y Stratonovich [78].

Dos características distintivas de un filtro de Kalman son: 1) su formulación matemática se describe en términos de conceptos de ecuaciones de estado, y 2) su solución es calculada de forma recursiva, ya que el estimado actual del estado se calcula a partir de los estimados previos y de los nuevos datos de entrada. El procedimiento recursivo del filtro de Kalman puede obtenerse por varias vías, e. g. modelos estadísticos Bayesianos [79], conceptos de la teoría de estimación [62], o mediante el proceso estocástico de innovación [80].

Para la estimación de los parámetros TVAR mediante el filtro de Kalman se iniciará planteando el problema en su forma de ecuaciones de estado y posteriormente se derivará el filtro recursivo usando el proceso de innovación. Finalmente, se harán las conexiones del filtro resultante con el problema del modelo autorregresivo dinámico de nuestro interés.

IV.4.4.4.1. Ecuaciones de estado del sistema y el proceso de innovación.

Para poder emplear el filtro de Kalman discreto se debe poseer un modelo de la señal en su forma de ecuaciones de estado. La evolución del estado del sistema se establece mediante una ecuación de diferencias de primer orden y debe ser estimado a partir de observaciones ruidosas. Considerando un vector M -dimensional $\bar{x}(n)$ que denota el *estado* de un sistema dinámico, lineal y de tiempo discreto, y un vector N -dimensional $\bar{y}(n)$ que denota los *datos observados* del sistema, entonces, la forma general de las ecuaciones de estado que describen al sistema [63, 81, 82], se pueden expresar como:

$$\begin{aligned}\bar{x}(n+1) &= \Phi(n+1, n)\bar{x}(n) + \bar{v}_1(n) \\ \bar{y}(n) &= C(n)\bar{x}(n) + \bar{v}_2(n)\end{aligned}\tag{184}$$

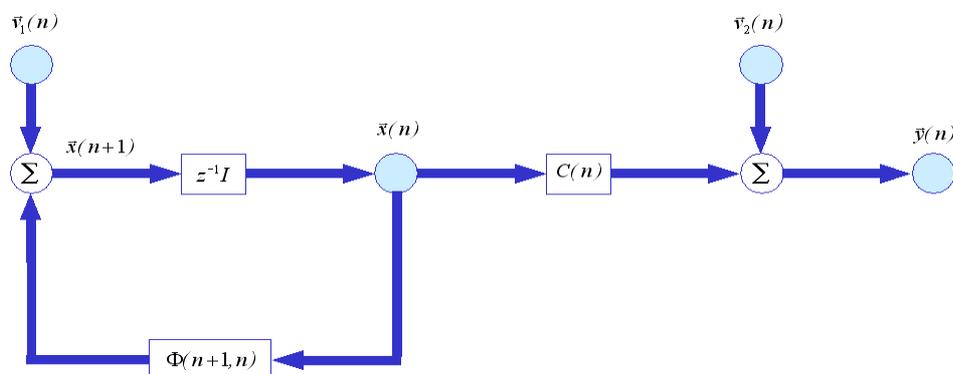


Fig. IV.30. Diagrama de flujo de la representación de un sistema dinámico, lineal de tiempo discreto mediante ecuaciones de estado.

Las ecuaciones anteriores representan la ecuación del proceso y la ecuación de medición, respectivamente. Una representación gráfica de estas ecuaciones se muestra en la figura IV.30, y a continuación se describe cada una de ellas:

1. *Ecuación del proceso.*

Esta ecuación corresponde a:

$$\bar{x}(n+1) = \Phi(n+1, n)\bar{x}(n) + \bar{v}_1(n) \quad (185)$$

donde $\bar{x}(n)$ es el estado del sistema al tiempo n , $\Phi(n+1, n)$ es una matriz de dimensiones $M \times M$, denominada matriz de transición o de flujo, que relaciona los estados del sistema al tiempo $n+1$ y n , y el vector M -dimensional $\bar{v}_1(n)$ representa el ruido del proceso.

2. *Ecuación de medición.*

Esta ecuación describe al vector de observación mediante:

$$\bar{y}(n) = C(n)\bar{x}(n) + \bar{v}_2(n) \quad (186)$$

donde $\bar{y}(n)$ es el vector de observaciones al tiempo n , $C(n)$ es la matriz de medición de dimensiones $N \times M$, y el vector N -dimensional $\bar{v}_2(n)$ es denominado el ruido de medición.

Los ruidos del proceso y de medición son modelados como procesos Gaussianos de ruido blanco con media cero y con matrices de correlación definidas mediante:

$$E[\bar{v}_1(n)\bar{v}_1^H(k)] = \begin{cases} Q_1(n) & , \quad n = k \\ 0 & , \quad n \neq k \end{cases} \quad (187)$$

$$y \quad E[\bar{v}_2(n)\bar{v}_2^H(k)] = \begin{cases} Q_2(n) & , \quad n = k \\ 0 & , \quad n \neq k \end{cases} \quad (188)$$

respectivamente. Además, se asume que ambos vectores de ruido son estadísticamente independientes, de tal forma que $E[\bar{v}_1(n)\bar{v}_2^H(k)] = 0$, para todo n y k .

Si el problema puede ser formulado mediante ecuaciones de estado, y se conocen las matrices de transición $\Phi(n+1, n)$, de medición $C(n)$, de correlación $Q_1(n)$ y $Q_2(n)$, entonces puede emplearse el filtro de Kalman discreto para estimar óptimamente el estado del sistema. Es decir, el problema consiste en emplear los datos observados, dados por los

vectores $\bar{y}(1), \bar{y}(2), \dots, \bar{y}(n)$, para encontrar los estimados de mínimos cuadrados de los componentes del estado $\bar{x}(i)$, para cada instante de tiempo $n \geq 1$. De acuerdo al valor de i el problema resulta en un problema de filtrado ($i = n$), de predicción ($i > n$) o de suavizado ($1 \leq i < n$) [63]. Dado que el problema de filtrado de Kalman se abordará desde el punto de vista del proceso de innovación, resulta conveniente describirlo.

Desde el desarrollo del método de mínimos cuadrados por Gauss, siempre ha existido el interés de evaluar el desempeño de un filtro, en el sentido de que los estimados deben satisfacer todas las observaciones de la manera más precisa posible. Una forma lógica consiste en considerar los residuos como una medida de desempeño del filtro. Kailath [80] señaló que la secuencia de salida del filtro de Kalman es una secuencia de ruido blanco, y dado que se puede obtener explícitamente, puede ser examinada en cada etapa para verificar que la secuencia de residuo o innovación posee las características estadísticas apropiadas. A continuación se utilizan las propiedades del proceso de innovación para derivar las fórmulas recursivas del filtro de Kalman.

El *proceso de innovación* asociado con $\bar{y}(n)$ corresponde a [80]:

$$\bar{\alpha}(n) = \bar{y}(n) - \hat{y}(n | y_{n-1}) \quad , \quad n = 1, 2, \dots \quad (189)$$

donde $\hat{y}(n | y_{n-1})$ representa el estimado del vector $\bar{y}(n)$ utilizando la información hasta $n-1$, $\bar{\alpha}(n)$ de dimensiones $M \times 1$ representa la nueva información en los datos observados $\bar{y}(n)$. La parte de la observación $\bar{y}(n)$ que es realmente nueva está contenida en $\bar{\alpha}(n)$, y de ahí su nombre de innovación. El proceso de innovación posee las siguientes propiedades [80]:

Propiedad 1. El proceso de innovación $\bar{\alpha}(n)$, asociado con los datos observados $\bar{y}(n)$ al tiempo n , es ortogonal a todas las observaciones pasadas $\bar{y}(1), \bar{y}(2), \dots, \bar{y}(n-1)$, es decir $E[\bar{\alpha}(n)\bar{y}^H(k)] = 0 \quad , \quad 1 \leq k \leq n-1$.

Propiedad 2. El proceso de innovación consiste de una secuencia de variables aleatorias vectoriales que son ortogonales entre sí, es decir $E[\bar{\alpha}(n)\bar{\alpha}^H(k)] = 0 \quad , \quad 1 \leq k \leq n-1$. Esta propiedad nos indica que la secuencia $\{\bar{\alpha}(n)\}$ corresponde a un ruido blanco.

Propiedad 3. Existe una correspondencia uno a uno entre la secuencia de variables aleatorias vectoriales $\{\bar{y}(1), \bar{y}(2), \dots, \bar{y}(n)\}$ y la secuencia de variables aleatorias vectoriales $\{\bar{\alpha}(1), \bar{\alpha}(2), \dots, \bar{\alpha}(n)\}$, donde una secuencia puede ser obtenida de la otra por medio de operadores lineales estables sin pérdida de información,

$$\{\bar{y}(1), \bar{y}(2), \dots, \bar{y}(n)\} \leftrightarrow \{\bar{\alpha}(1), \bar{\alpha}(2), \dots, \bar{\alpha}(n)\} \quad (190)$$

Es posible conformar la secuencia de variables aleatorias que definen al proceso de innovación empleando el procedimiento de ortogonalización de Gram-Schmidt [64].

IV.4.4.4.2. Formulación del filtro.

Para formular el filtro de Kalman, el primer reto consiste en expresar la matriz de correlación del proceso de innovación $\bar{\alpha}(n)$ en términos de los elementos de la ecuación de medición. Dicha matriz de correlación, $\Sigma(n)$, está definida como [63]:

$$\Sigma(n) = E[\bar{\alpha}(n)\bar{\alpha}^H(n)] \quad (191)$$

Empleando la ecuación de medición, y dado el espacio vectorial y_{n-1} , representado por las observaciones $\bar{y}(1), \bar{y}(2), \dots, \bar{y}(n-1)$, se tiene que el estimador cuadrático medio mínimo del valor actual del vector de observación $\bar{y}(n)$ corresponde a:

$$\hat{y}(n | y_{n-1}) = C(n)\hat{x}(n | y_{n-1}) + \hat{v}_2(n | y_{n-1}) \quad (192)$$

donde $\hat{x}(n | y_{n-1})$ es el vector de predicción de un paso hacia adelante del estado. A partir de la ecuación de medición se deduce que $\hat{v}_2(n)$ es ortogonal a las observaciones pasadas $\bar{y}(1), \bar{y}(2), \dots, \bar{y}(n-1)$, es decir $E[\bar{y}(k)\hat{v}_2^H(n)] = 0$, $0 \leq k \leq n-1$, y por lo tanto el estimado $\hat{v}_2(n | y_{n-1})$ del vector del ruido de medición es cero, en consecuencia:

$$\hat{y}(n | y_{n-1}) = C(n)\hat{x}(n | y_{n-1}) \quad (193)$$

Al sustituir la expresión anterior para $\hat{y}(n | y_{n-1})$, y la expresión para $\bar{y}(n)$ dada por la ecuación de medición en la definición del proceso de innovación, se obtiene:

$$\begin{aligned} \bar{\alpha}(n) &= \bar{y}(n) - \hat{y}(n | y_{n-1}) \\ \bar{\alpha}(n) &= C(n)\bar{x}(n) + \bar{v}_2(n) - C(n)\hat{x}(n | y_{n-1}) \\ \bar{\alpha}(n) &= C(n)\bar{\varepsilon}(n, n-1) + \bar{v}_2(n) \end{aligned} \quad (194)$$

donde $\bar{\varepsilon}(n, n-1)$ es el vector de error de predicción del estado al tiempo n empleando los datos hasta el tiempo $n-1$, y corresponde a la diferencia entre el vector de estado $\bar{x}(n)$ y el vector de predicción de un paso $\hat{x}(n | y_{n-1})$, es decir $\bar{\varepsilon}(n, n-1) = \bar{x}(n) - \hat{x}(n | y_{n-1})$,

ortogonal tanto al ruido del proceso $\bar{v}_1(n)$ como al ruido de medición $\bar{v}_2(n)$. Empleando los resultados anteriores, la matriz de correlación del proceso de innovación puede expresarse mediante:

$$\begin{aligned}\Sigma(n) &= E[\bar{\alpha}(n)\bar{\alpha}^H(n)] \\ \Sigma(n) &= E\left[\left(C(n)\bar{\varepsilon}(n,n-1) + \bar{v}_2\right)\left(\bar{\varepsilon}^H(n,n-1)C^H(n) + \bar{v}_2^H\right)\right] \\ \Sigma(n) &= C(n)K(n,n-1)C^H(n) + Q_2(n)\end{aligned}\quad (195)$$

donde $Q_2(n)$ es la matriz de correlación del ruido de medición $\bar{v}_2(n)$, y la matriz $K(n,n-1)$, de dimensiones $M \times M$, es la matriz de correlación del error de estado de predicción, definida como $K(n,n-1) = E[\bar{\varepsilon}(n,n-1)\bar{\varepsilon}^H(n,n-1)]$, y que se emplea como la descripción estadística del error en el estimado $\hat{x}(n | y_{n-1})$.

El siguiente reto consiste en derivar el estimado cuadrático medio mínimo del estado $\bar{x}(i)$ a partir del proceso de innovación [63]. Considerando la relación uno a uno entre las secuencias $\{\bar{y}(1), \bar{y}(2), \dots, \bar{y}(n)\}$ y $\{\bar{\alpha}(1), \bar{\alpha}(2), \dots, \bar{\alpha}(n)\}$ el estimado puede ser expresado como una combinación lineal de la secuencia del proceso de innovación como:

$$\hat{x}(i | y_n) = \sum_{k=1}^n E[\bar{x}(i)\bar{\alpha}^H(k)]\Sigma^{-1}(k)\bar{\alpha}(k) \quad (196)$$

separando la sumatoria en dos términos, se tiene:

$$\begin{aligned}\hat{x}(i | y_n) &= \sum_{k=1}^{n-1} E[\bar{x}(i)\bar{\alpha}^H(k)]\Sigma^{-1}(k)\bar{\alpha}(k) \\ &\quad + E[\bar{x}(n)\bar{\alpha}^H(n)]\Sigma^{-1}(n)\bar{\alpha}(n)\end{aligned}\quad (197)$$

para $i = n + 1$ se puede describir como:

$$\hat{x}(n+1 | y_n) = \sum_{k=1}^{n-1} E[\bar{x}(n+1)\bar{\alpha}^H(k)]\Sigma^{-1}(k)\bar{\alpha}(k) + E[\bar{x}(n)\bar{\alpha}^H(n)]\Sigma^{-1}(n)\bar{\alpha}(n) \quad (198)$$

La ecuación anterior relaciona el estado $\bar{x}(n+1)$ al tiempo $n+1$ con el estado $\bar{x}(n)$ al tiempo n . Post-multiplicando la ecuación (185) por $\bar{\alpha}^H(k)$ y tomando el valor esperado, se tiene que para $0 \leq k \leq n$:

$$E[\bar{x}(n+1)\bar{\alpha}^H(k)] = E\left[\left(\Phi(n+1,n)\bar{x}(n) + \bar{v}_1(n)\right)\bar{\alpha}^H(k)\right] \quad (199)$$

dado que $\bar{\alpha}(k)$ solamente depende de los datos observados $\bar{y}(1), \dots, \bar{y}(k)$, y además $\bar{v}_1(n)$ y $\bar{\alpha}(k)$ son ortogonales para $0 \leq k \leq n$, entonces:

$$E[\bar{x}(n+1)\bar{\alpha}^H(k)] = \Phi(n+1,n)E[\bar{x}(n)\bar{\alpha}^H(k)] \quad (200)$$

Por lo tanto, el primer término de la ecuación (198) corresponde a:

$$\sum_{k=1}^{n-1} E[\bar{x}(n+1)\bar{\alpha}^H(k)]\Sigma^{-1}(k)\bar{\alpha}(k) = \Phi(n+1,n)\sum_{k=1}^{n-1} E[\bar{x}(n)\bar{\alpha}^H(k)]\Sigma^{-1}(k)\bar{\alpha}(k) \quad (201)$$

donde, la sumatoria del lado derecho es la expansión de $\hat{x}(n|y_{n-1})$ en términos de la secuencia de procesos de innovación, por lo tanto:

$$\sum_{k=1}^{n-1} E[\bar{x}(n+1)\bar{\alpha}^H(k)]\Sigma^{-1}(k)\bar{\alpha}(k) = \Phi(n+1,n)\hat{x}(i|y_{n-1}) \quad (202)$$

Finalmente, el estimado cuadrático medio mínimo del estado $\hat{x}(n+1|y_n)$ queda expresado en términos del proceso de innovación $\bar{\alpha}(n)$ mediante:

$$\hat{x}(n+1|y_n) = \Phi(n+1,n)\hat{x}(i|y_{n-1}) + E[\bar{x}(n)\bar{\alpha}^H(n)]\Sigma^{-1}(n)\bar{\alpha}(n) \quad (203)$$

La matriz de dimensiones $M \times N$, definida como

$$G(n) = E[\bar{x}(n)\bar{\alpha}^H(n)]\Sigma^{-1}(n) \quad (204)$$

se denomina la ganancia de Kalman [63]. Empleando esta matriz de ganancia, el estimado $\hat{x}(n+1|y_n)$ corresponde a:

$$\hat{x}(n+1|y_n) = \Phi(n+1,n)\hat{x}(n|y_{n-1}) + G(n)\bar{\alpha}(n) \quad (205)$$

Esta ecuación es importante ya que indica que el estimado cuadrático medio mínimo $\hat{x}(n+1|y_n)$ del estado de un sistema dinámico lineal puede calcularse añadiendo un término de corrección al estimado previo $\hat{x}(n|y_{n-1})$ que se encuentra pre-multiplicado por la matriz de transición $\Phi(n+1,n)$. El término de corrección es igual al proceso de innovación $\bar{\alpha}(n)$ pre-multiplicado por la matriz $G(n)$, y de ahí su nombre de ganancia.

Ahora, es necesario hallar una expresión para $G(n)$ adecuada para su cálculo. Empleando la ecuación del proceso y la expresión obtenida para el proceso de innovación, el valor esperado de la ganancia de Kalman puede expresarse de la siguiente manera

$$\begin{aligned} G(n) &= E[\bar{x}(n)\bar{\alpha}^H(n)]\Sigma^{-1}(n) \\ G(n) &= E\left[(\Phi(n+1,n)\bar{x}(n) + \bar{v}_1(n))(C(n)\bar{\varepsilon}(n,n-1) + \bar{v}_2(n))^H\right]\Sigma^{-1}(n) \end{aligned} \quad (206)$$

pero se tiene que $\bar{x}(n)$ y $\bar{v}_2(n)$ no están correlacionados, $\bar{\varepsilon}(n,n-1)$ y $\bar{v}_1(n)$ son ortogonales, $\bar{v}_1(n)$ y $\bar{v}_2(n)$ son estadísticamente independientes, y $\Phi(n+1,n)$ se asume como conocida. Por lo tanto:

$$G(n) = \Phi(n+1, n)E[\bar{x}(n)\bar{\varepsilon}^H(n, n-1)]C^H(n)\Sigma^{-1}(n) \quad (207)$$

Utilizando la definición del vector de error del estado de predicción, y observando que $\bar{\varepsilon}(n, n-1)$ y $\hat{x}(n | y_{n-1})$ son ortogonales, se puede reducir la ecuación anterior a:

$$\begin{aligned} G(n) &= \Phi(n+1, n)E[(\bar{\varepsilon}(n, n-1) + \hat{x}(n | y_{n-1}))\bar{\varepsilon}^H(n, n-1)]C^H(n)\Sigma^{-1}(n) \\ G(n) &= \Phi(n+1, n)E[\bar{\varepsilon}(n, n-1)\bar{\varepsilon}^H(n, n-1)]C^H(n)\Sigma^{-1}(n) \end{aligned} \quad (208)$$

donde el valor esperado de la ecuación anterior es precisamente la matriz de correlación del error de estado de predicción, es decir

$$G(n) = \Phi(n+1, n)K(n, n-1)C^H(n)\Sigma^{-1}(n) \quad (209)$$

El último reto consiste en derivar una expresión recursiva para el cálculo de $K(n, n-1)$ [63]. Por definición, la matriz de correlación del vector de error $\bar{\varepsilon}(n+1, n)$ corresponde a

$$K(n+1, n) = E[\bar{\varepsilon}(n+1, n)\bar{\varepsilon}^H(n+1, n)] \quad (210)$$

donde

$$\bar{\varepsilon}(n+1, n) = \bar{x}(n+1) - \hat{x}(n+1 | y_n) \quad (211)$$

Sustituyendo la ecuación (184), ecuación (205) y la ecuación (194), la ecuación anterior se puede expresar como:

$$\bar{\varepsilon}(n+1, n) = [\Phi(n+1, n) - G(n)C(n)]\bar{\varepsilon}(n, n-1) - G(n)\bar{v}_2(n) + \bar{v}_1(n) \quad (212)$$

dado que $\bar{\varepsilon}(n, n-1)$, $\bar{v}_1(n)$ y $\bar{v}_2(n)$ son vectores mutuamente no correlacionados, $K(n+1, n)$ está dada por:

$$\begin{aligned} K(n+1, n) &= [\Phi(n+1, n) - G(n)C(n)]K(n, n-1)[\Phi(n+1, n) - G(n)C(n)]^H \\ &\quad + Q_1(n) + G(n)Q_2(n)G^H(n) \end{aligned} \quad (213)$$

que puede simplificarse, sustituyendo las expresiones para la ganancia de Kalman de las ecuaciones (210) y el hecho que $\bar{\varepsilon}(n, n-1) = \bar{x}(n) - \hat{x}(n | y_{n-1})$, en:

$$K(n+1, n) = \Phi(n+1, n)K(n)\Phi^H(n+1, n) + Q_1(n) \quad (214)$$

La ecuación anterior es denominada la ecuación de diferencias de Riccati, en honor a Jacopo Riccati [63].

Empleando la propiedad de la matriz de transición, es decir:

$$\Phi(n+1, n)\Phi(n, n+1) = I \quad (215)$$

se tiene la recursión para la matriz $K(n)$, de dimensiones $M \times M$, de la forma:

$$K(n) = K(n, n-1) - \Phi(n, n+1)G(n)C(n)K(n, n-1) \quad (216)$$

que corresponde a la matriz de correlación de $\bar{\varepsilon}(n)$ y se emplea como la descripción estadística del error en el estimado $\hat{x}(n | y_n)$.

Para utilizar el algoritmo del filtro de Kalman es necesario especificar las condiciones iniciales. A pesar de que el estado inicial de un filtro de Kalman no se conoce con precisión, usualmente se establece en términos de su media y matriz de correlación. Al tiempo $n = 0$, en ausencia de cualquier dato observado, el estimado inicial se puede seleccionar como $\hat{x}(0 | y_0) = E[\hat{x}(0)] = \bar{0}$ asumiendo que el vector de estado $\bar{x}(n)$ posee una media nula y que su matriz de correlación puede seleccionarse como $K(0) = E[\hat{x}(0)\hat{x}^H(0)] = P_0$. A partir de estas condiciones iniciales, los valores siguientes se calculan empleando la matriz de transición $\Phi(n+1, n)$ mediante:

$$\begin{aligned} \hat{x}(1 | y_0) &= \Phi(1,0)\hat{x}(0 | y_0) \\ K(1,0) &= \Phi(1,0)K(0)\Phi^H(1,0) + Q_1(0) \end{aligned} \quad (217)$$

A partir de este instante, el algoritmo puede proseguir su curso.

El diagrama a bloques del filtro de Kalman se muestra en la figura IV.31 y en la Tabla IV.VI se muestra el resumen del filtro de Kalman, incluyendo condiciones iniciales.

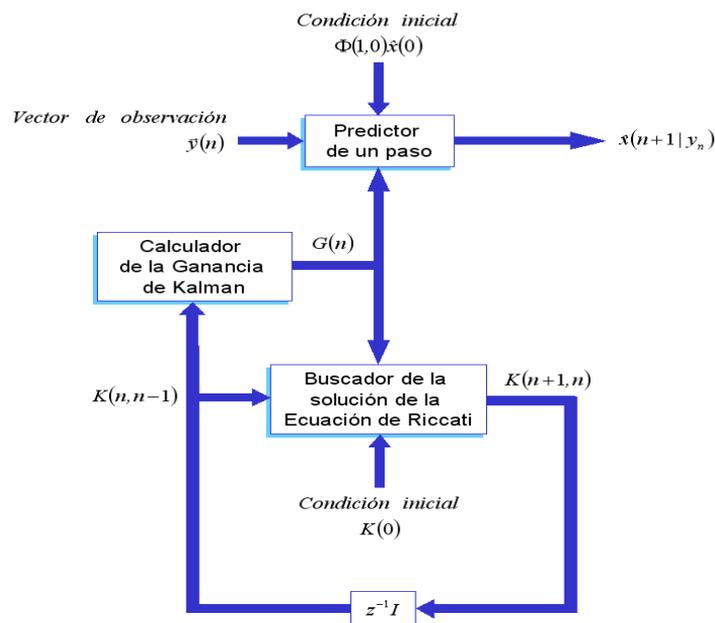


Fig. IV.31 Diagrama a bloques del filtro de Kalman basado en tres componentes: 1) La ganancia de Kalman, 2) Solución de la ecuación de Riccati, y 3) Predictor de un paso hacia adelante. El diagrama de cada componente se muestra en la figura IV.32.

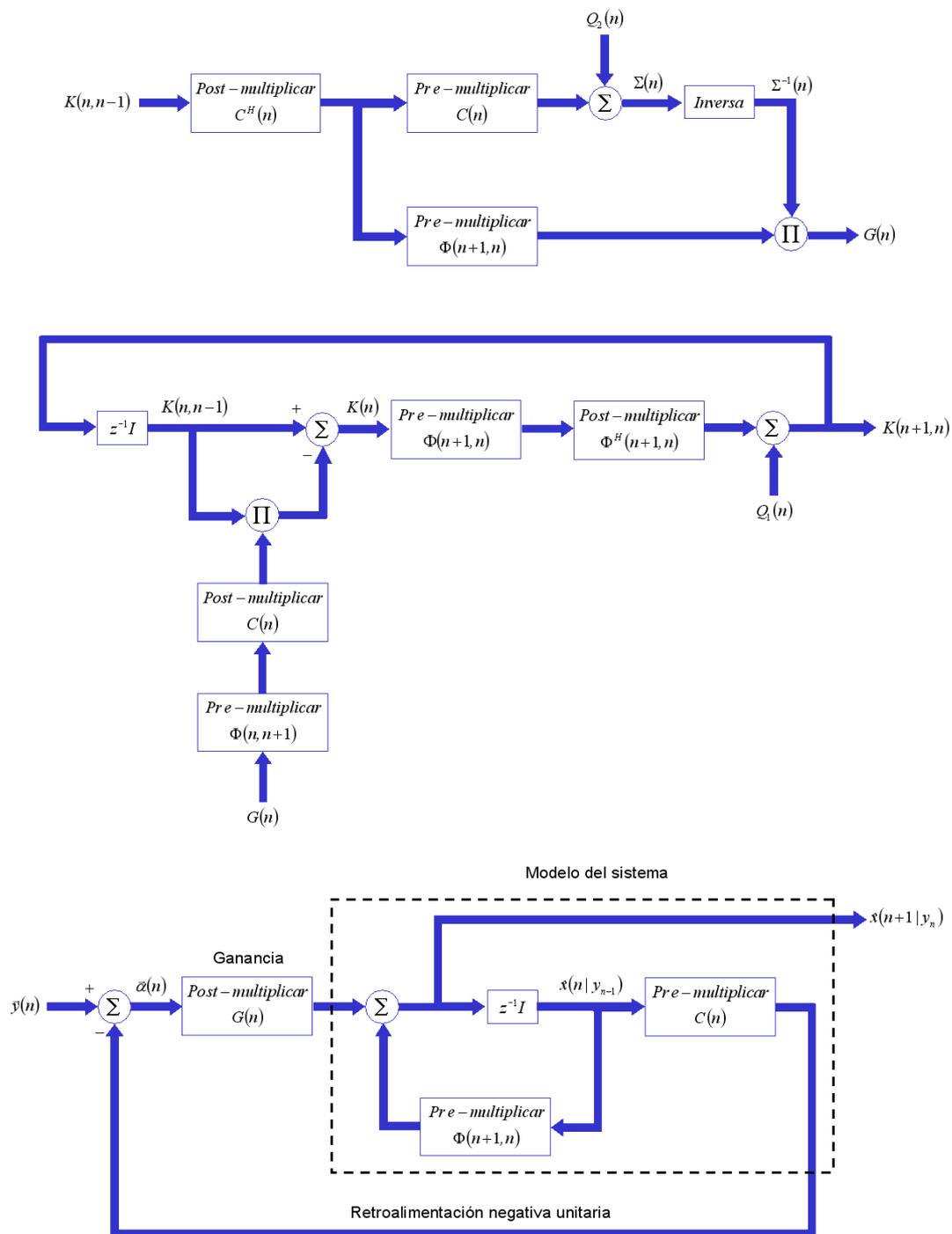


Fig. IV.32. Componentes del filtro de Kalman. a) Ganancia de Kalman. Una vez calculado $G(n)$ se puede actualizar el predictor de un paso para obtener el estimado $\hat{x}(n+1 | y_n)$. b) Solución de la ecuación de Riccati. Dado el valor de $K(n, n-1)$ calcula el valor actualizado $K(n+1, n)$. c) Predictor de un paso hacia adelante. Calcula el estimado del vector de estado $\hat{x}(n+1 | y_n)$ dado su valor previo $\hat{x}(n | y_{n-1})$.

TABLA IV.VI. EL FILTRO DE KALMAN.

Entrada	$Observaciones = \{\bar{y}(1), \bar{y}(2), \dots, \bar{y}(n)\}$
Salida	$Predicción del vector de estado = \hat{x}(n+1 y_n)$
Parámetros conocidos	$Matriz de transición del estado = \Phi(n+1, n)$ $Matriz de medición = C(n)$ $Matriz de correlación del ruido del proceso = Q_1(n)$ $Matriz de correlación del ruido de medición = Q_2(n)$
Inicializar el algoritmo ajustando	$\hat{x}(0 y_0) = E[\hat{x}(0)] = \bar{0}$ $K(0) = E[\hat{x}(0)\hat{x}^H(0)] = P_0$
Para cada instante de tiempo, $n = 1, 2, \dots$, calcular	$G(n) = \Phi(n+1, n)K(n, n-1)C^H(n)(C(n)K(n, n-1)C^H(n) + Q_2(n))^{-1}$ $\bar{\alpha}(n) = \bar{y}(n) - C(n)\hat{x}(n y_{n-1})$ $\hat{x}(n+1 y_n) = \Phi(n+1, n)\hat{x}(n y_{n-1}) + G(n)\bar{\alpha}(n)$ $K(n) = K(n, n-1) - \Phi(n, n+1)G(n)C(n)K(n, n-1)$ $K(n+1, n) = \Phi(n+1, n)K(n)\Phi^H(n+1, n) + Q_1(n)$

IV.4.4.4.3. Obtención de la representación tiempo-frecuencia.

Recordando que un proceso autorregresivo variante en el tiempo (TVAR) de orden p $\{x(n)\}$, esta descrito por la siguiente recursión

$$x(n) = \sum_{k=1}^p a_k(n)x(n-k) + w(n)$$

donde $x(n)$ es la señal bajo análisis al instante n , $\{a_k(n)\}_{k=1}^p$ son los parámetros del modelo TVAR, $\{x(n-k)\}_{k=1}^p$ son muestras retrasadas de la señal, y $\{w(n)\}$ un proceso de ruido estacionario con media cero y varianza σ^2 . Nuestro interés radica en estimar los parámetros del modelo TVAR mediante el filtro de Kalman, y a partir de ellos obtener un estimado del espectro de la densidad de potencia instantánea. Las ecuaciones de estado considerando a los parámetros del modelo TVAR pueden formularse de la forma siguiente [48, 83]

$$\begin{aligned} \bar{a}(n+1) &= \Phi(n+1, n)\bar{a}(n) + \bar{v}(n) \\ y(n) &= \bar{y}^T(n-1)\bar{a}(n) + w(n) \end{aligned} \quad (218)$$

que corresponden a la ecuación del proceso y la ecuación de medición, respectivamente, del proceso AR dinámico. En esta formulación los parámetros $\bar{a}(n)$ se convierten en las

variables de estado. La matriz de transición $\Phi(n+1,n)$ es una matriz diagonal, $\Phi(n+1,n) = \rho I$, y corresponde a la correlación entre los estados $\bar{a}(n+1)$ y $\bar{a}(n)$. El valor de ρ controla la velocidad de adaptación y la capacidad de seguimiento del algoritmo. Para señales en las cuales los parámetros TVAR cambian lentamente con el tiempo, el valor de ρ se selecciona muy cercano a la unidad. En el caso de $\rho = 1$ la matriz de transición es igual a $\Phi(n+1,n) = I$. En la tabla IV.VII se muestran las relaciones obtenidas al comparar las ecuaciones del modelo TVAR con las variables de estado del filtro de Kalman.

TABLA IV.VII. RELACIONES ENTRE EL FILTRO DE KALMAN Y EL MODELO TVAR.

Parámetro	Modelo de estado del filtro de Kalman	Modelo TVAR
Vector de estado	$\bar{x}(n)$	$\bar{a}(n)$
Matriz de transición	$\Phi(n+1,n)$	I
Vector de ruido del proceso	$\bar{v}_1(n)$	$\bar{v}(n)$
Matriz de correlación del vector de ruido del proceso	$Q_1(n)$	qI
Vector de observación	$\bar{y}(n)$	$y(n)$
Matriz de medición	$C(n)$	$\bar{y}^T(n-1)$
Vector de ruido de medición	$\bar{v}_2(n)$	$w(n)$
Matriz de correlación del vector de ruido de medición	$Q_2(n)$	σ^2

Utilizando estas relaciones, puede emplearse el filtro de Kalman para modelar un proceso no estacionario $\{y(n)\}$ mediante un proceso TVAR. Los parámetros TVAR pueden estimarse de manera óptima utilizando el siguiente algoritmo recursivo:

$$\begin{aligned}
 G(n) &= K(n,n-1)\bar{y}(n-1)\left[\bar{y}^T(n-1)K(n,n-1)\bar{y}(n-1) + \sigma^2\right]^{-1} \\
 \alpha(n) &= y(n) - \bar{y}^T(n-1)\hat{a}(n-1) \\
 \hat{a}(n) &= \hat{a}(n-1) + G(n)\alpha(n) \\
 K(n) &= K(n,n-1) - G(n)\bar{y}^T(n-1)K(n,n-1) \\
 K(n+1,n) &= K(n) + qI
 \end{aligned} \tag{219}$$

donde $\hat{a}(n)$ es el estimado del estado (los parámetros TVAR) al tiempo n . Observe que el empleo del algoritmo anterior asume el conocimiento previo de la varianza σ^2 del residuo $w(n)$, y la varianza q de cada elemento del vector de ruido del proceso $\bar{v}(n)$. Las condiciones iniciales de esta recursión están dadas por:

$$\begin{aligned}
 \hat{a}(0) &= \bar{0} \\
 K(0) &= cI \quad , \quad c = \text{constante positiva pequeña}
 \end{aligned} \tag{220}$$

Finalmente, la obtención de la representación tiempo-frecuencia basada considerando un modelo TVAR de orden p y determinado los parámetros del modelo mediante el algoritmo de Kalman se formular como:

$$P_{Kalman}(n, f) = \frac{1}{\left| 1 + \sum_{k=1}^p \hat{a}_{k, Kalman}(n) e^{-j2\pi f k} \right|^2} \quad (221)$$

IV.5. El Escalograma.

En esta sección se establecerá la distribución tiempo-frecuencia mediante el uso de la teoría de la transformada *wavelet* (WT). La transformada *wavelet* proporciona una alternativa a la transformada de Fourier de tiempo corto (STFT) de interés en el análisis de señales no estacionarias. Mientras que la STFT emplea una ventana de análisis única la WT emplea ventanas de tiempo cortas a altas frecuencias y ventanas de tiempo extensas a bajas frecuencias. La WT se visualizará como una descomposición de la señal en un conjunto de funciones base, las *wavelets*, que son obtenidas a partir de una *wavelet* prototipo o *wavelet* madre, mediante dilataciones y contracciones o escalamientos, así como por desplazamientos. Finalmente, se introducirá la noción de *escala* empleada en la WT como una alternativa a la frecuencia, originando la representación tiempo-escala, lo cual significa que la señal se mapea a un plano tiempo-escala, y no a un plano tiempo-frecuencia, existiendo sin embargo una relación entre ambas nociones.

IV.5.1. Conceptos teóricos de la representación *wavelet*.

Un gran número de técnicas que han sido desarrolladas de forma independiente para varias aplicaciones de procesamiento de señales, entre las que se incluyen el procesamiento multiresolución de señales y las expansiones en series *wavelet*, han sido reconocidas como diferentes facetas de una misma teoría, la *teoría wavelet*, que proporciona un marco unificado.

Aunque las ideas entorno a los *wavelets* surgieron desde inicios desde la década de 1940 con la publicación del trabajo de Gabor [5], solo recientemente la teoría *wavelet* se ha desarrollado de forma unificada. Mientras que la idea de observar a una señal en varias escalas y analizarla con varias resoluciones surgió independientemente en diferentes campos de las matemáticas, la física y la ingeniería, fue a mediados de la década de 1980 que Morlet,

Grossmann y Meyer [84] construyeron los fundamentos matemáticos alrededor de lo que actualmente denominamos *wavelets*. Morlet fue el primero en proponer la transformada *wavelet* como una herramienta para el análisis de señales, en particular de datos sísmicos. Por su parte, Daubechies [85] y Mallat [86] establecieron las conexiones con el procesamiento discreto de señales e imágenes, y desde entonces se ha realizado un gran número de contribuciones teóricas y prácticas en varios aspectos de la transformada *wavelet*.

Es conveniente recordar que la transformada de Fourier de tiempo-corto (STFT) se puede visualizar ya sea como una sucesión de transformadas de Fourier de un segmento ventaneado de la señal (barras verticales), ver figura IV.33.a, o como análisis mediante un banco de filtros modulados (barras horizontales). Sin embargo, la resolución en el tiempo y en la frecuencia no pueden ser arbitrariamente pequeñas, debido a que su producto, denominado producto tiempo-ancho de banda, se encuentra acotado inferiormente por el principio de incertidumbre, o la desigualdad de Heisenberg-Gabor [5]:

$$\text{Producto Tiempo - Ancho de banda} = \Delta t \Delta f \geq \frac{1}{4\pi} \quad (222)$$

que indica que existe un compromiso entre la resolución en el tiempo y la resolución en la frecuencia [1]. Además, una vez que una ventana es seleccionada en la STFT la resolución tiempo-frecuencia se fija a lo largo de todo el plano tiempo frecuencia, implicando funciones base que únicamente cambian en frecuencia y no en duración, ver figura IV.33.b. Por ejemplo, si la señal está compuesta por varios componentes, entonces cada componente puede ser analizado con una buena resolución en el tiempo o en la frecuencia, pero no con ambas simultáneamente [87].

Una forma de solventar la limitación en la resolución de la STFT consiste en imaginar que las resoluciones Δt y Δf pueden variar a lo largo del plano tiempo-frecuencia y con ello obtener un análisis multiresolución, ver figura IV.33.c, donde ahora las funciones base varían en duración y en frecuencia, ver figura IV.33.d. Cuando el análisis se considera como un banco de filtros, la resolución en el tiempo debe incrementarse con la frecuencia central de los filtros de análisis. Por lo tanto debe imponerse que Δf sea proporcional a f [87], es decir:

$$\frac{\Delta f}{f} = c \quad (223)$$

donde c es una constante. Entonces el banco de filtros de análisis estará compuesto de filtros pasa-banda con un ancho de banda relativo constante. En otras palabras, en lugar de que la respuesta en frecuencia del filtro de análisis se encuentre regularmente espaciada a lo largo

del eje de frecuencia, como ocurre en la STFT, éstas estarán regularmente espaciadas en una escala logarítmica, como ocurre con la audición humana [85]. En el caso del análisis con un ancho de banda relativo constante tanto Δf como Δt cambian en función de la frecuencia central del filtro de análisis, y aún se satisface el principio de incertidumbre pero la resolución en el tiempo es arbitrariamente buena en altas frecuencias, mientras que la resolución en la frecuencia es arbitrariamente buena a bajas frecuencias. Por lo tanto, este tipo de análisis resulta más útil para señales conformadas por componentes de altas frecuencias de corta duración y componentes de bajas frecuencias de larga duración [87].

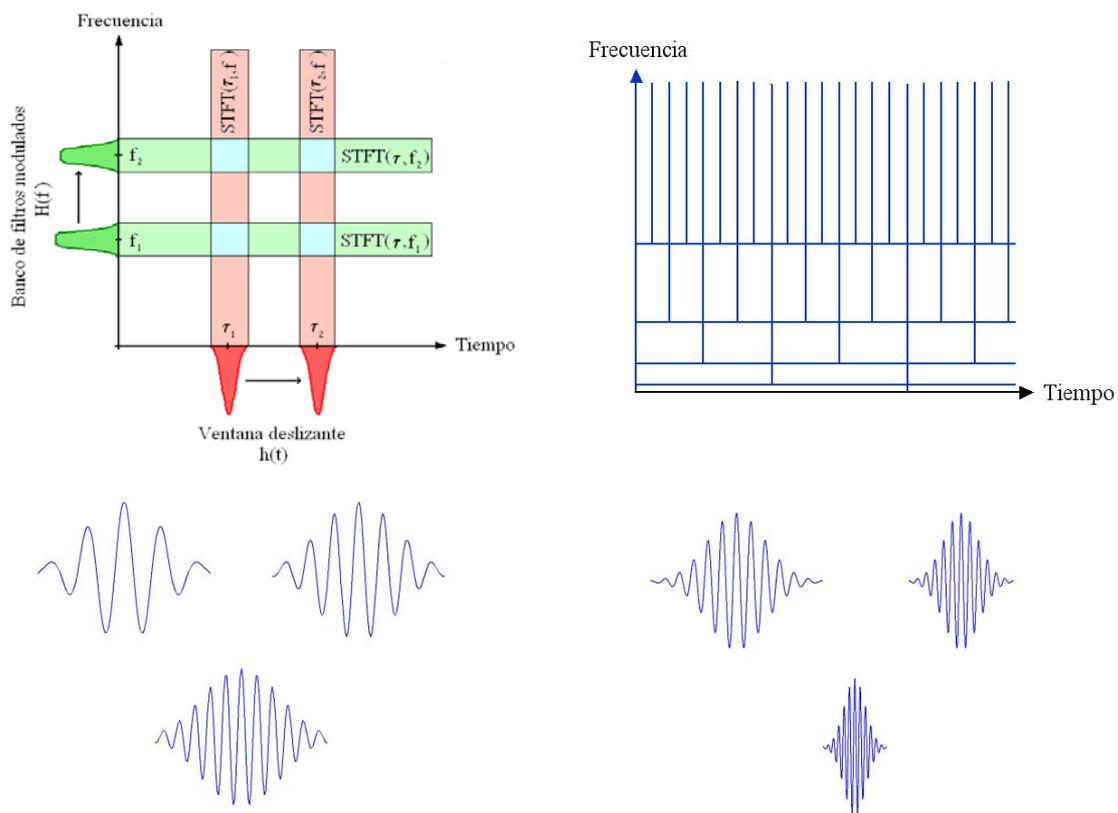


Fig. IV.33. Resolución tiempo-frecuencia y funciones base de la transformada de Fourier de tiempo corto (STFT) y la transformada *wavelet* (WT). a) Plano tiempo-frecuencia de la transformada de Fourier de tiempo-corto (STFT). b) Átomos tiempo-frecuencia de la STFT. c) Plano tiempo-frecuencia de la WT con ventanas de análisis variable. d) Átomos de la WT, *wavelets*.

Dado que en el presente trabajo, para el análisis de los sonidos torácicos en el dominio tiempo-frecuencia, se utilizó la transformada *wavelet* continua (CWT), a continuación se mencionan algunos de sus aspectos teóricos más importantes.

La CWT utiliza la idea del análisis con un ancho de banda relativo constante, sin embargo agrega una simplificación: todas las respuestas al impulso del banco de filtro son definidas como versiones escalas de una función prototipo $\psi(t)$ [87], esto es:

$$\psi_a(t) = \frac{1}{\sqrt{|a|}} \psi\left(\frac{t}{a}\right) \quad (224)$$

donde a es el factor de escala, y el término $1/\sqrt{|a|}$ es empleado para normalizar la energía de la función $\psi(t)$. En general, cuando una función $f(t)$ es escalada, ($f(t) \rightarrow f(at)$, $a > 0$), si $a > 1$ entonces $f(t)$ se contrae y si $a < 1$ entonces $f(t)$ se expande [85]. Lo anterior conduce a la siguiente definición de la CWT de una señal $x(t)$ [87]:

$$CWT_x(\tau, a) = \frac{1}{\sqrt{|a|}} \int x(t) \psi^*\left(\frac{t-\tau}{a}\right) dt \quad (225)$$

que puede escribirse de forma equivalente mediante un cambio de variable como:

$$CWT_x(\tau, a) = \sqrt{|a|} \int x(at) \psi^*\left(t - \frac{\tau}{a}\right) dt \quad (226)$$

Dado que la función prototipo $\psi(t)$, denominada la *wavelet* básica o *wavelet* madre, es empleada para todas las respuestas al impulso, ninguna escala es privilegiada y por lo tanto, el análisis *wavelet* es auto-similar para todas las escalas.

La interpretación de la CWT mediante la ecuación (225) es la siguiente: conforme la escala se incrementa, la respuesta al impulso del filtro $\psi((t-\tau) \cdot a^{-1})$ se dispersa más en el tiempo y sólo considera comportamientos temporales largos. De forma equivalente, la interpretación de la ecuación (226) indica que conforme la escala crece, una versión contraída de la señal es vista a través de un filtro con longitud constante. Por lo tanto, el factor de escala posee una interpretación similar a la empleada en los mapas geográficos, donde escalas grandes implican vistas globales y escalas pequeñas implican vistas detalladas [87].

Es posible entender el concepto de la transformada *wavelet* continua de una forma alternativa mediante la definición de las funciones base. A partir de la definición de la CWT de la ecuación (225), la transformación puede visualizarse como un producto interno de la forma [87, 88]:

$$CWT_x(\tau, a) = \frac{1}{\sqrt{|a|}} \int x(t) \psi_{a,\tau}^*(t) dt \quad (227)$$

que mide la similitud entre la señal bajo análisis $x(t)$ y las funciones base, denominadas *wavelets*, donde:

$$\psi_{a,\tau}(t) = \frac{1}{\sqrt{|a|}} \psi\left(\frac{t-\tau}{a}\right) \quad (228)$$

Las *wavelets* $\psi_{a,\tau}(t)$ consisten en versiones escaladas, por un factor a ($a > 0$) y trasladadas por τ , de una *wavelet* madre $\psi(t)$. El análisis *wavelet* resulta en un conjunto de coeficientes que indican que tan cercana se encuentra una señal de una función base particular. Entonces, una función pueda ser representada mediante una descomposición en *wavelets*, $\psi_{a,\tau}(t)$ ortogonales, en el sentido de que la función original pueda sintetizarse mediante la suma de átomos de morfología constante pero de diferente tamaño y amplitud [87].

Mientras el análisis se realiza mediante la ecuación (227), la síntesis consiste en sumar todas las proyecciones ortogonales, de la forma:

$$x(t) = c \iint_{a>0} CWT_x(\tau, a) \psi_{a,\tau}(t) \frac{dad\tau}{a^2} \quad (229)$$

donde la constante c depende únicamente de $\psi(t)$.

Las funciones base $\psi_{a,\tau}(t)$ no son ciertamente ortogonales debido a que son redundantes al estar definidas para los parámetros de variación continua a y τ . Sin embargo, la fórmula de síntesis es válida cuando $\psi(t)$ es de energía finita y pasa-banda, lo que implica que oscila en el tiempo como una onda corta, y de ahí su nombre de *wavelet*, implicando que la transformada de Fourier de la *wavelet* madre $\Psi(f)$ debe satisfacer la condición de admisibilidad dada por [89], es decir:

$$\int |\Psi(f)|^2 \frac{df}{|f|} = 1 \quad (230)$$

que resulta más restrictiva que en el caso de la STFT donde solamente se requiere que la energía de la ventana de análisis $h(t)$ sea de energía finita. De forma más precisa, si $\psi(t)$ se asume como suficientemente regular, y posee un decaimiento al infinito, entonces la condición de reconstrucción equivale a $\int \psi(t) dt = 0$, que significa que una señal puede ser reconstruida sólo con media cero [87].

IV.5.2. Propiedades de la transformada *wavelet* continua.

Algunas de las propiedades que posee la CWT son [3, 87]:

1. La CWT es invariante a traslaciones y escalamientos en el tiempo.

$$\text{Si } x(t) \Rightarrow CWT_x(t, a; \psi),$$

entonces:

$$y(t) = \sqrt{|a_0|} x(a_0(t - t_0)) \Rightarrow CWT_y(t, a, \psi) = CWT_x(a_0(t - t_0), \frac{a}{a_0}; \psi)$$

Este grupo de transformaciones es denominado grupo afin [85], sin embargo la CWT no preserva traslaciones en la frecuencia.

2. La señal $x(t)$ puede ser recuperada a partir de su CWT.

$$\text{Sea } CWT_x(\tau, a) = \frac{1}{\sqrt{|a|}} \int x(t) \psi^* \left(\frac{t - \tau}{a} \right) dt,$$

$$\text{entonces: } x(t) = c \iint_{a>0} CWT_x(\tau, a) \psi_{a,\tau}(t) \frac{dad\tau}{a^2}$$

si se cumple la condición de restricción para la *wavelet* madre, es decir: $\int \psi(t) dt = 0$.

3. La resolución en el tiempo y la frecuencia, al igual que en el caso de la STFT, están relacionadas mediante la desigualdad de Heisenberg-Gabor. Sin embargo, en la CWT, la resolución en la frecuencia se vuelve más pobre conforme la frecuencia de análisis se incrementa mientras que la resolución en el tiempo mejora.
4. El escalograma, i.e. la magnitud al cuadrado de la CWT, posee las propiedades de conservación de energía, ser una distribución evaluada en los reales y ser siempre positiva [87].

IV.5.3. El escalograma.

Para una señal se puede definir una distribución tiempo-frecuencia similar al espectrograma pero utilizando la transformada *wavelet*. Dado que la CWT es una descomposición en bases ortonormales, puede mostrarse que preserva la energía de la señal [87, 89], es decir:

$$\iint |CWT_x(\tau, a)|^2 \frac{d\tau da}{a^2} = \int |x(t)|^2 dt = E_x \quad (231)$$

donde E_x es la energía de la señal $x(t)$. Lo anterior permite definir la representación tiempo-escala denominada *escalograma*, como la magnitud al cuadrado de la CWT, de la forma:

$$SC_x(\tau, a; \psi) = \frac{1}{|a|} \left| \int x(t) \psi^* \left(\frac{t - \tau}{a} \right) dt \right|^2 \quad (232)$$

El escalograma es una distribución de la energía de la señal en el plano tiempo-escala asociada con la medida $\frac{d\tau da}{a^2}$, que es equivalente a $dtdf$, y por ello se expresa en unidades de potencia por unidad de frecuencia, como en el caso del espectrograma. En la figura IV.34, se muestran las diferencias entre un escalograma y un espectrograma para dos diferentes señales: un impulso de Dirac y una señal compuesta por tres sinusoidales de frecuencias $f_0, 2f_0, 4f_0$. Para el impulso de Dirac, en el caso de la CWT, la representación alrededor del punto $t = t_0$ está limitada a un cono en el plano tiempo-escala, estando muy localizada alrededor de t_0 a escalas pequeñas, ver figura IV.34.a. En contraste, para la STFT, la representación es tan larga como la extensión de la ventana de análisis, para todas las frecuencias, ver figura IV.34.b. En el caso de la señal compuesta por sinusoides, dado que el análisis tiempo-escala es logarítmico con la frecuencia, el ancho de la representación de las sinusoidales se incrementa con la frecuencia en el caso del escalograma, figura IV.34.c, mientras que permanece constante en el espectrograma, ver figura IV.34.d.

Tanto el espectrograma como el escalograma producen una representación bidimensional de la señal más o menos fácil de interpretar, donde cada patrón en el plano tiempo-escala o tiempo-frecuencia contribuye a la energía global de la señal. Sin embargo, tales representaciones de energía poseen desventajas. Por ejemplo, tanto el espectrograma como el escalograma son funciones bilineales de la señal analizada y por ello aparecen términos cruzados como términos de interferencia entre patrones en el plano tiempo-frecuencia o tiempo-escala. Los términos de interferencia del escalograma y del espectrograma, están restringidos a aquellas regiones del plano tiempo-frecuencia donde se traslapan los auto-escalogramas, los términos de la señal. Por ello, si dos componentes de la señal se encuentran lo suficientemente separados en el plano tiempo-frecuencia, sus escalogramas-cruzados serán esencialmente cero [24].

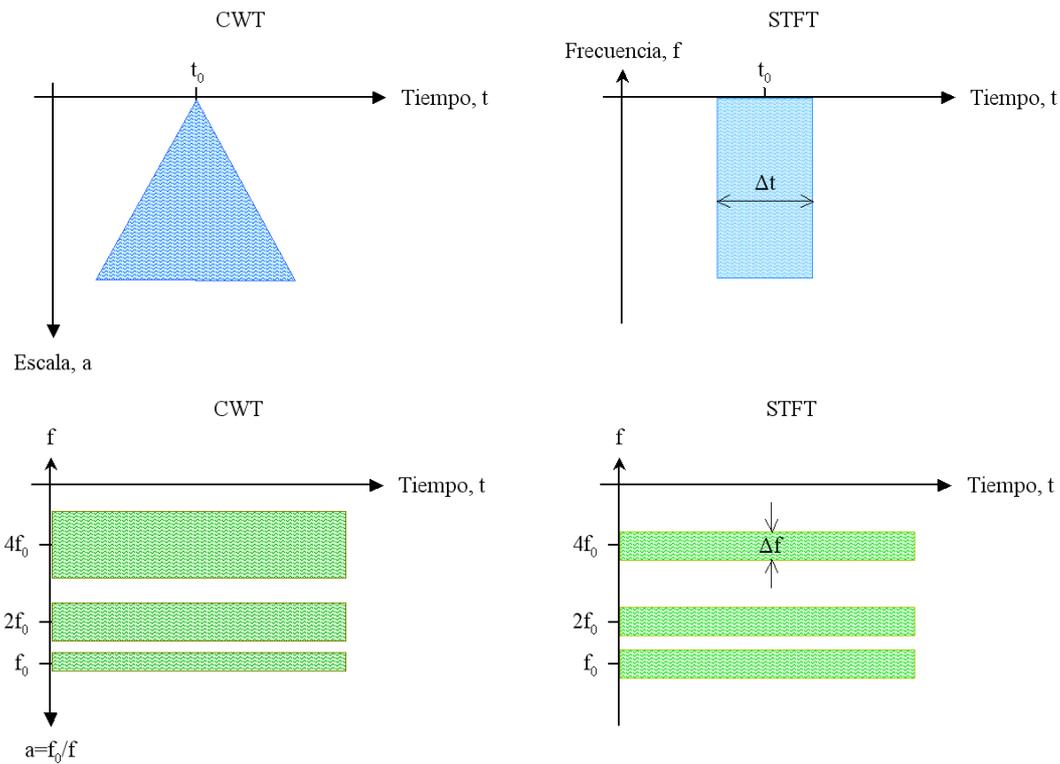


Fig. IV.34. Comparación entre escalograma y espectrograma. a) Escalograma de una delta de Dirac al tiempo $t = t_0$ y b) su Espectrograma correspondiente. c) Escalograma de tres sinusoidales de frecuencias f_0 , $2f_0$, $4f_0$, y d) su Espectrograma correspondiente.

IV.5.4 La *wavelet* Morlet compleja.

Debido a la flexibilidad en la elección de una *wavelet* prototipo, un gran número de funciones han sido empleadas como *wavelets*. En particular, nos centraremos en una *wavelet* madre especialmente útil en la CWT, la *wavelet* Morlet, denominada así en honor a Jean Morlet y formulada originalmente por Goupillaud, Grossmann y Morlet en 1984 [2]. La *wavelet* Morlet consiste en una *wavelet* compleja con una envolvente Gaussiana y se define como [90]:

$$\psi(t) = \pi^{-\frac{1}{4}} e^{j2\pi f_0 t} e^{-\frac{t^2}{2}} \quad (233)$$

La envolvente Gaussiana confina la forma de onda sinusoidal alrededor de una frecuencia central f_0 , la frecuencia característica de la *wavelet* Morlet. La *wavelet* Morlet, es una sinusoidal compleja ventaneada con una función Gaussiana, como se muestra en la figura IV.35.a, y por lo tanto su decaimiento exponencial resulta en una buena localización temporal

durante su aplicación en la transformada *wavelet* continua. En la figura IV.35.b se muestra la representación en la frecuencia de la ecuación (233). Además, debido a que la *wavelet* de Morlet es una función compleja, permite extraer información tanto de la magnitud como de la fase del proceso bajo análisis [90].

En la figura IV.36 se presenta nuevamente el escalograma para el impulso de Dirac y las señales sinusoidales utilizando la *wavelet* Morlet. En el caso del escalograma para el impulso de Dirac se observa que la resolución en el tiempo depende de la frecuencia de análisis; mientras que en las señales sinusoidales, la resolución en la frecuencia también se ve afectada por la frecuencia de análisis.

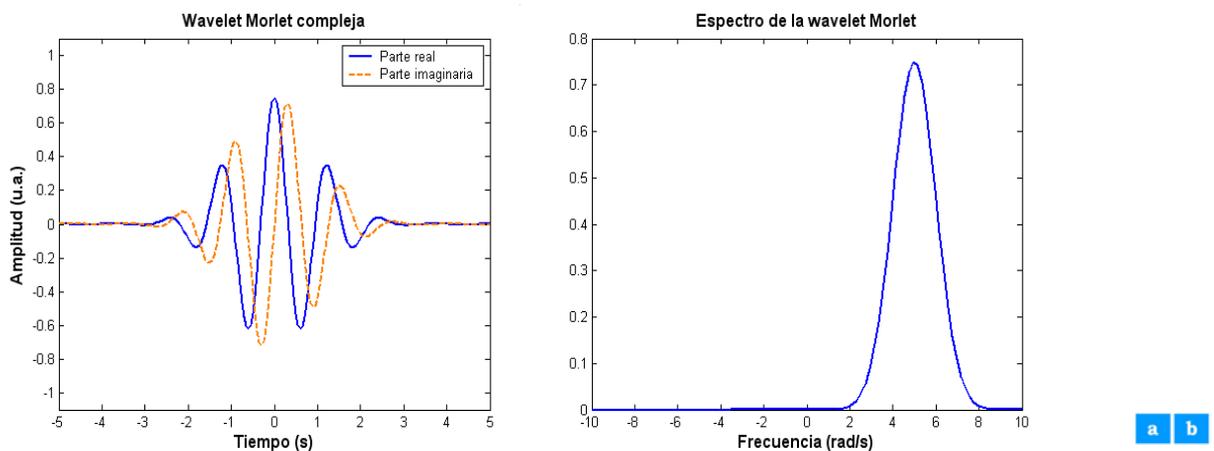


Fig. IV.35. *Wavelet* Morlet. a) Partes real (línea continua) e imaginaria (línea punteada) de la *wavelet*, para $\omega_0 = 2\pi f_0 = 5 \text{ rad} / \text{s}$. b) Espectro de potencia de la *wavelet*.

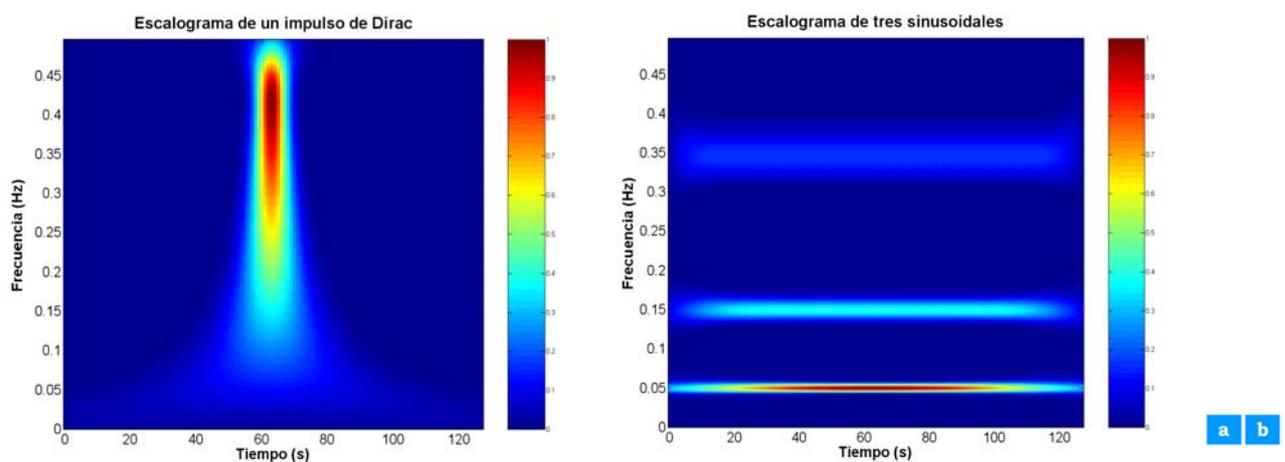


Fig. IV.36. Escalograma utilizando una *wavelet* Morlet de: a) un impulso de Dirac al tiempo $t = 64 \text{ s}$, y b) de tres señales sinusoidales con frecuencias $f_0, 2f_0, 4f_0$, con $f_0 = 0.25 \text{ Hz}$.

Finalmente, el escalograma se puede considerar como miembro de la clase afín de distribuciones, y se ha demostrado que se puede expresar como una versión suavizada de la distribución de Wigner-Ville [89], de la forma:

$$SC_x(\tau, a; \psi) = \iint WVD_x(u, \theta) WVD_\psi\left(\frac{u-t}{a}, a\theta\right) du d\theta \quad (234)$$

donde ψ es la *wavelet* madre. A partir de la expresión anterior, el escalograma puede considerarse como una correlación afín bidimensional entre la distribución de Wigner-Ville de la señal y la distribución de Wigner-Ville de la *wavelet* básica. Además, la expresión (234) señala la fuerte relación existente entre el escalograma y la distribución de Wigner-Ville. Ambas distribuciones pertenecen al denominada *clase afín de distribuciones* [91], la cual es la clase de distribuciones de energía tiempo-frecuencia covariante a traslaciones en el tiempo y dilataciones. Finalmente, la distribución de Wigner establece el vínculo entre el escalograma y el espectrograma.

IV.6. El espectro de Hilbert-Huang.

Hasta ahora se han discutido diversas representaciones tiempo-frecuencia, algunas basadas en una clase de transformaciones invariantes a corrimientos en el tiempo y en la frecuencia, así como sus versiones modificadas, otras dadas en términos del modelado de la señal bajo análisis mediante ecuaciones recursivas e inclusive, otras basadas en transformaciones invariantes a corrimientos y escalamientos en el tiempo. En la presente sección se introduce la última técnica considerada en este trabajo de tesis para la obtención de una representación tiempo-frecuencia, el *espectro de Hilbert-Huang* (HHS) [92].

Históricamente, el término *espectro* se ha convertido casi en sinónimo de la transformada de Fourier de los datos y se ha aplicado a todo tipo de señales, debido a su simplicidad. Sin embargo, existen dos restricciones cruciales del análisis de Fourier: el sistema debe ser lineal y los datos deben ser estrictamente periódicos o estacionarios. Desafortunadamente, al tratar con señales reales es común encontrarse con uno o más de los siguientes problemas: la duración total de los datos es corta, los datos no son estacionarios y representan procesos no lineales. En consecuencia, el análisis espectral de Fourier es de uso limitado, sin embargo todavía es empleado para procesar señales reales. Por las razones brevemente mencionadas, es importante el desarrollo de nuevas técnicas para el análisis de

señales en el dominio del tiempo y la frecuencia o en ambos dominios simultáneamente. En este sentido el espectro de Hilbert-Huang ha permitido representar a las señales en el espacio tiempo-frecuencia con base en nuevos conceptos teóricos, como la descomposición en modos de oscilación empíricos. Resulta muy interesante aplicar el espectro de Hilbert-Huang a sonidos torácicos y comparar sus resultados con aquellos obtenidos por representaciones que durante cierto tiempo han permanecido como las técnicas clásicas de representación en el espacio tiempo-frecuencia.

Con el propósito de formular el espectro de Hilbert-Huang es conveniente recordar algunos conceptos fundamentales. Los métodos de representación tiempo-frecuencia presentados en las secciones anteriores están diseñados para modificar la representación del análisis de Fourier pero fallan de una u otra forma. Sin embargo, establecen las condiciones necesarias que deben de cumplir las funciones base para representar a una serie de tiempo no lineal y no estacionaria, a saber: que sean completas, que sean ortogonales, que sean locales y adaptables. La condición de *completas* garantiza el grado de precisión de la expansión, la condición de *ortogonalidad* garantiza la positividad de la energía y evita fugas. Ambas constituyen los requerimientos estándar de todo método de expansión lineal. Las condiciones restantes son particulares para los datos no lineales y no estacionarios. La condición de *localidad* es la más crucial para los datos no estacionarios, donde no existe una escala de tiempo y por lo tanto, todos los eventos tienen que ser identificados por el tiempo de sus ocurrencias; se requiere que tanto la amplitud, o la energía, así como la frecuencia sean funciones del tiempo. La condición de *adaptabilidad* también es fundamental, dado que al adaptarse a las variaciones locales de los datos, la descomposición podrá describir completamente la física del proceso. La capacidad de adaptación del método es especialmente importante para fenómenos no lineales, y una manifestación de esta no linealidad es la distorsión armónica en el análisis de Fourier. Considerando los aspectos antes mencionados es difícil que una base predeterminada se ajuste a todos los fenómenos y en consecuencia, una alternativa para su generación es obtenerla directamente a partir de los datos.

En la derivación del espectro de Hilbert-Huang, el método de la descomposición empírica de modos (EMD) juega un papel crucial. El EMD permite que cualquier conjunto de datos pueda descomponerse en un número finito de funciones base, los denominados modos de oscilación intrínsecos (IMFs), donde cada uno de estos modos admite una transformada de Hilbert bien comportada. A partir de la transformada de Hilbert se obtienen la frecuencia y la amplitud instantánea de cada IMF como una función del tiempo, y cuya representación final resulta en el HHS. Un concepto importante es que las IMFs se obtienen directamente de los datos bajo análisis lo que significa que el EMD es un método adaptable. El EMD, al basarse

en las escalas de tiempo características y locales de los datos produce frecuencias instantáneas que están relacionadas con estructuras del proceso bajo análisis.

La explicación del método se realizará de la forma siguiente: se iniciará con una revisión de la transformada de Hilbert y las condiciones que debe cumplir una señal para que sea posible obtener una frecuencia instantánea físicamente significativa, posteriormente se enlazarán estas condiciones con los modos de oscilación intrínsecos que las satisfagan, y finalmente se presentará el proceso para obtener los modos a partir de los datos originales, mediante la denominada descomposición empírica de modos. En la figura IV.37, se esquematiza el procedimiento para generar el HHS de una señal, como se puede observar el procedimiento inicia con la extracción de los IMFs.

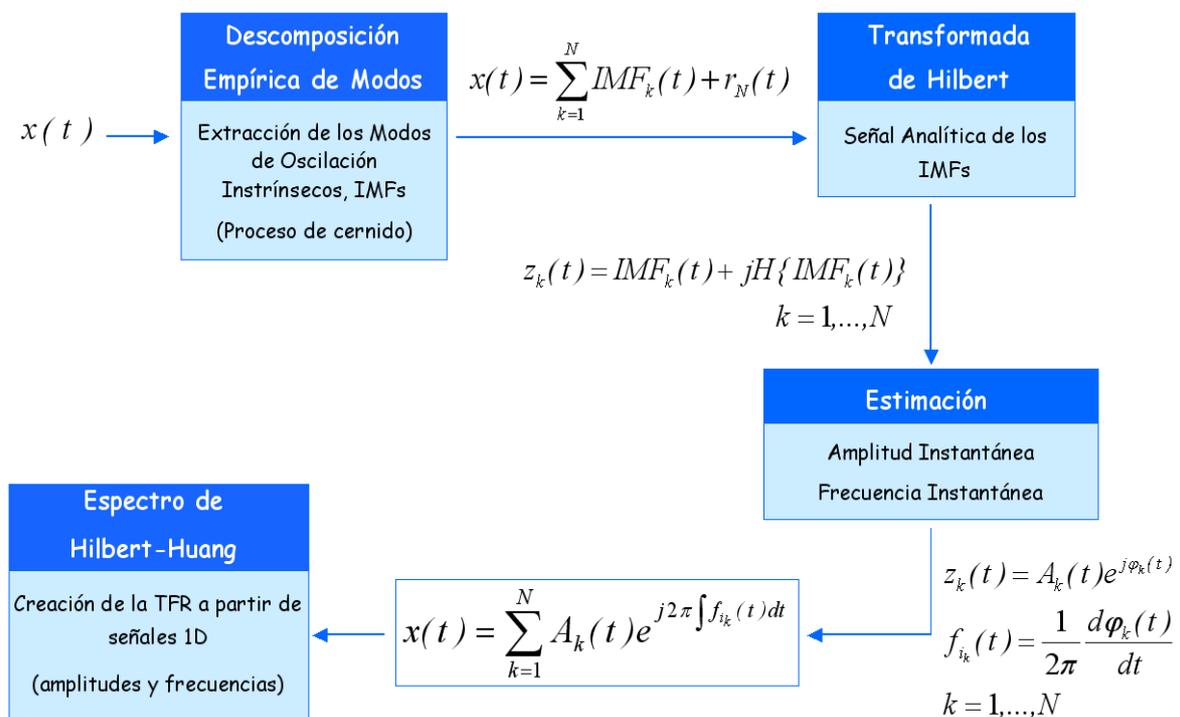


Fig. IV.37. Método para obtener la representación de datos no estacionarios y no lineales mediante el espectro de Hilbert-Huang (HHS). El método requiere de dos pasos: 1) el pre-procesamiento mediante la descomposición empírica de modos para representar la señal en términos de funciones base derivadas de los datos, y 2) la aplicación de la transformada de Hilbert a los modos de oscilación intrínsecos generados para construir el HHS.

IV.6.1. La transformada de Hilbert y el concepto de frecuencia instantánea.

Mientras que la noción de energía instantánea de la señal ha sido bien aceptada, la noción de frecuencia instantánea ha sido controversial debido principalmente a dos elementos: 1) la gran influencia del análisis espectral de Fourier, y 2) que no hay una forma única de definir a la frecuencia instantánea.

El análisis de Fourier define la frecuencia para las funciones seno y coseno con una extensión igual a la duración de los datos. En este sentido, la frecuencia instantánea debería estar relacionada con alguna de estas funciones sinusoidales y por lo tanto, se necesitaría al menos una oscilación completa para definir el valor de la frecuencia local. Sin embargo, esta definición no hace sentido para datos no estacionarios en los cuales la frecuencia varía temporalmente. La definición de la frecuencia instantánea no resulta complicada cuando se cuenta con una forma única para definir la señal analítica mediante la transformada de Hilbert [7, 16, 17].

Para una serie de tiempo arbitraria $x(t)$, se tiene que su transformada de Hilbert, $H\{x(t)\}$, corresponde a:

$$H\{x(t)\} = \frac{1}{\pi} P \int_{-\infty}^{\infty} \frac{x(\tau)}{t - \tau} d\tau \quad (235)$$

donde P indica el valor principal de Cauchy. A partir de la definición de las señales $x(t)$ y $H\{x(t)\}$, se define la señal compleja $z(t)$, correspondiente a la señal analítica de la serie de datos, como:

$$z(t) = x(t) + jH\{x(t)\} = a(t)e^{j\varphi(t)} \quad (236)$$

donde la amplitud y la fase instantánea se definen como:

$$\begin{aligned} a(t) &= \sqrt{(x(t))^2 + (H\{x(t)\})^2} \\ \varphi(t) &= \arctan\left(\frac{H\{x(t)\}}{x(t)}\right) \end{aligned} \quad (237)$$

La transformada de Hilbert puede considerarse como la convolución de $x(t)$ con $\frac{1}{t}$, lo cual enfatiza las propiedades locales de la serie de tiempo. Aún con la utilización de la transformada de Hilbert existen controversias en utilizar la definición de la frecuencia instantánea dada por Ville, es decir:

$$f_i(t) = \frac{\omega_i}{2\pi} = \frac{1}{2\pi} \frac{d\varphi(t)}{dt} \quad (238)$$

Por ejemplo, se necesitan ciertas condiciones de los datos para que la frecuencia instantánea $f_i(t)$ posea un valor único en cualquier instante de tiempo; es decir, que represente un solo componente. Lo anterior, obligó a la introducción del término de señal monocomponente por Leon Cohen [1], sin embargo no existe una definición clara que permita considerar una señal como monocomponente. Por consiguiente se ha adoptado el concepto de banda angosta como la limitación para que los datos posean una frecuencia instantánea significativa [16].

Sin embargo, las definiciones de banda angosta, basadas en momentos estadísticos del espectro de la señal [1], definen el ancho de banda de la señal de forma global y resultan demasiado restrictivas. Otra restricción que debe imponerse a los datos para que posean una frecuencia instantánea significativa consiste en que la parte real de su transformada de Fourier solamente exista para frecuencias positivas [5], sin embargo esta restricción también es global. Entonces, para el análisis de datos se requiere trasladar estas restricciones globales a restricciones locales que se puedan implementar mediante etapas de un método simple.

Con el propósito de entender el tipo de restricciones que se le imponen a una señal para que la frecuencia instantánea que se obtiene vía la transformada de Hilbert posea un significado físico, consideremos la señal $x(t) = \alpha + \text{sen}(t)$ con un valor inicial de α igual a cero. En este caso, la señal analítica $z(t)$ para la señal $x(t)$ es:

$$z(t) = \text{sen}(t) + jH\{\text{sen}(t)\} = \text{sen}(t) + j\cos(t)$$

en consecuencia, de acuerdo con las ecuaciones (237) y (238), la fase y la frecuencia instantánea de $z(t)$ son t y un valor constante, respectivamente. En la figura IV.38 se muestran las gráficas correspondientes al plano de fase, la fase y la frecuencia instantánea. Para el caso de α igual a cero, todas las gráficas hacen sentido, por ejemplo, la gráfica de frecuencia instantánea es constante, como corresponde a la teoría. Sin embargo, si α es menor o mayor a 1, las gráficas de frecuencia instantánea ya no tienen un significado físico dado que presentan valores negativos y una serie de oscilaciones.

Por lo tanto, una restricción que debe tener una función para poseer una frecuencia instantánea que tenga sentido físico es presentar simetría local con respecto del nivel cero. Huang *et al* demostraron que para una función, su frecuencia instantánea significativa solamente puede definirse si su forma de onda se restringe a ser localmente simétrica con respecto a un valor promedio cero, y poseer el mismo número de cruces por cero y de máximos o mínimos o a lo más diferir en una unidad. Además, sugirieron un método para

descomponer los datos en una clase de funciones que satisfacen dichas condiciones y que designaron como modos de oscilación intrínsecos para los cuales es posible definir su frecuencia instantánea en todo momento [92].

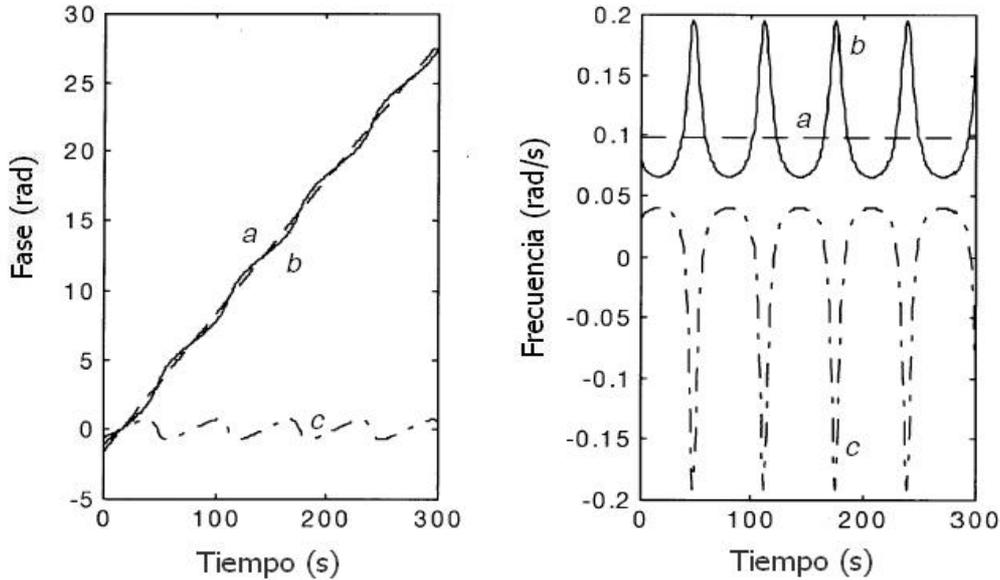


Fig. IV.38. Concepto de fase y frecuencia instantánea utilizando una señal senoidal con un valor de DC (α). a) Valor de $\alpha=0$, b) Valor de $\alpha<1$, y c) Valor de $\alpha>1$.

IV.6.2. Los modos de oscilación intrínsecos.

De manera formal, una función es un modo de oscilación intrínseco (IMF) si satisface dos condiciones: 1) a lo largo de todo el conjunto de datos, un IMF posee el mismo número de extremos que de cruces por cero o a lo más éstos difieren por uno, y 2) para cualquier instante de tiempo, el valor medio de su envolvente definida por el máximo local y su envolvente definida por el mínimo local es cero. Un IMF puede ser expresado en su forma analítica, $z(t)$, mediante su transformada de Hilbert, como:

$$z(t) = IMF(t) + jH\{IMF(t)\} = A(t)e^{j\varphi(t)} \quad (239)$$

Si se aplica la transformada de Fourier a $z(t)$ se obtiene

$$Z(\omega) = \int_{-\infty}^{\infty} A(t)e^{j\varphi(t)} e^{-j\omega t} dt = \int_{-\infty}^{\infty} A(t)e^{j(\varphi(t)-\omega t)} dt \quad (240)$$

Por lo tanto, la máxima contribución a $Z(\omega)$ se obtiene por la frecuencia $\tilde{\omega}$, que satisface la condición:

$$\frac{d}{dt}\{\varphi(t) - \omega t\} = 0 \quad (241)$$

que resulta en la definición de la frecuencia instantánea introducida por Ville [7], es decir:

$$\tilde{\omega} = \omega_i(t) = \frac{d\varphi(t)}{dt} \quad (242)$$

A partir de la forma polar de $z(t)$ y del resultado obtenido para $\omega_i(t)$, se observa que la definición de frecuencia anterior brinda el mejor ajuste sinusoidal local, y por lo tanto, no es necesario un periodo de oscilación completo para definir un valor de frecuencia instantánea dado que ésta puede ser definida para cada instante de tiempo. Ahora sólo resta representar los datos, $x(t)$, en un conjunto de componentes, los IMFs, de tal forma que pueda asignarse un valor de frecuencia instantánea a cada uno de ellos.

IV.6.3. La descomposición empírica de modos.

Dado que en la mayoría de los casos el conjunto de datos analizado no es un IMF, es necesario descomponerlos en sus modos de oscilación intrínsecos. Generalmente, dicha descomposición resulta necesaria ya que para cualquier instante de tiempo los datos pueden involucrar más de un modo de oscilación, y por lo tanto su simple transformada de Hilbert no proporciona una descripción completa de su contenido en frecuencia. La descomposición se realiza mediante el denominado *proceso de cernido* [92], que se basa en las características locales de los IMFs. Para descomponer la señal se debe cumplir: 1) la señal debe poseer al menos dos extremos, un máximo y un mínimo, 2) la escala de tiempo característica está definida por el lapso de tiempo entre los extremos, y 3) si la señal está desprovista de extremos pero contiene puntos de inflexión, puede derivarse una o más ocasiones para revelar los extremos. A continuación se describe el proceso de cernido para extraer sistemáticamente los IMFs de una serie de tiempo $x(t)$, $t = 1, 2, \dots, T$.

El primer paso consiste en determinar todos los extremos de los datos que corresponden a los máximos y mínimos locales de la señal $x(t)$. Una vez identificados todos los máximos locales se deben conectar mediante una interpolación spline cúbica para conformar la envolvente superior. Un procedimiento similar se efectúa con los mínimos locales para obtener la envolvente inferior. Posteriormente, se obtiene su valor medio, designado como $m_1(t)$. Ahora, se obtiene un primer posible IMF mediante la siguiente diferencia:

$$h_1(t) = x(t) - m_1(t) \quad (243)$$

Para que $h_1(t)$ sea un IMF debe cumplir la propiedad de simetría con respecto al valor promedio cero y debe de poseer el mismo número de cruces por cero que de máximos o mínimos o diferir al menos en una unidad. Sin embargo, en el primer paso del algoritmo puede ser que $h_1(t)$ no cumpla con estas condiciones y es necesario repetir el procedimiento de cernido. En un segundo paso, el componente $h_1(t)$ es considerado como los nuevos datos, al cual se le obtiene el valor medio de sus envolventes $m_{1,1}(t)$, para producir un nuevo componente mediante la diferencia:

$$h_{1,1}(t) = h_1(t) - m_{1,1}(t) \quad (244)$$

y el proceso se repite k veces hasta que el componente $h_{1,k}(t)$ generado sea efectivamente un IMF, es decir:

$$h_{1,k}(t) = h_{1,k-1}(t) - m_{1,k}(t) \quad (245)$$

donde $h_{1,k}(t)$ cumple con las propiedades mencionadas y es designado como el primer modo de oscilación intrínseco de los datos, de la forma:

$$IMF_1(t) = h_{1,k}(t) \quad (246)$$

Para determinar los restantes IMFs, el $IMF_1(t)$, componente con el periodo más corto de la señal o con la escala de tiempo más fina, se separa de los datos originales, es decir:

$$r_1(t) = x(t) - IMF_1(t) \quad (247)$$

Dado que el residuo, $r_1(t)$, todavía puede contener información de componentes con periodos más largos, debe ser tratado como los nuevos datos y procesado mediante el proceso de cernido descrito anteriormente. Este procedimiento puede repetirse sobre todos los residuos subsecuentes, y resultar en:

$$\begin{aligned} r_2(t) &= r_1(t) - IMF_2(t) \\ &\vdots \\ r_N(t) &= r_{N-1}(t) - IMF_N(t) \end{aligned} \quad (248)$$

Al final, la descomposición puede ser detenida mediante dos criterios: 1) cuando el modo intrínseco de oscilación $IMF_N(t)$, o el residuo $r_N(t)$, es demasiado pequeño como para tener una consecuencia sustancial, y 2) cuando el residuo $r_N(t)$ se convierte en una función monótona de la cual no pueden extraerse más IMFs. En la práctica, para garantizar que los IMFs resultantes retengan suficiente sentido físico, se debe imponer un criterio para detener el proceso de cernido. Esto puede llevarse a cabo limitando la desviación estándar

SD , de dos componentes resultantes de cernidos consecutivos, a un valor entre 0.2 y 0.3, es decir:

$$0.2 \prec SD = \sum_{t=1}^T \left[\frac{(h_{1,k-1}(t) - h_{1,k}(t))^2}{(h_{1,k-1}(t))^2} \right] \prec 0.3 \quad (249)$$

Mediante el EMD el número de extremos decrece conforme se avanza de un residuo al siguiente, y por ello, se garantiza que la descomposición proporcione un número finito de modos que permite representar a la señal original mediante la siguiente expansión en N funciones base:

$$x(t) = \sum_{k=1}^N IMF_k(t) + r_N(t) \quad (250)$$

En la tabla IV.VIII se resumen los pasos del EMD para obtener los IMFs de una señal $x(t)$, $t = 1, 2, \dots, T$, mediante el proceso de cernido.

Con respecto a la completitud de las funciones base, teóricamente se encuentra dada por la identidad en la expansión en IMFs, ecuación (250). De forma práctica, existe una diferencia entre la señal original y la señal reconstruida mediante la suma de todos los modos intrínsecos de oscilación y el residuo, sin embargo la amplitud máxima del error es demasiado pequeña y es debida al error de redondeo de la computadora.

En lo referente a la ortogonalidad, prácticamente se cumple pero no se encuentra garantizada teóricamente. Debido a la descomposición, las funciones base debieran de ser todas ortogonales entre sí, dado que cada elemento es obtenido mediante la diferencia entre la señal y su media local a través de las envolventes máximas y mínimas. Sin embargo lo anterior no es estrictamente cierto ya que la media que se obtiene mediante las envolventes no es la media verdadera. Además, cada IMF es sólo una parte de la media de la señal. Por ello, es inevitable que exista una pequeña fuga. Huang et. al. establecieron un índice de ortogonalidad, para el cual un valor cercano a cero indica una fuga pequeña. Aunque la medida pareciera ser global, el significado realmente es local, i.e. para algunos datos especiales, los IMFs vecinos pueden tener secciones de datos que porten las mismas frecuencias pero a diferentes instantes de tiempo. Entonces, localmente cualesquiera dos componentes deben ser ortogonales para todo propósito práctico.

TABLA IV.VIII. LA DESCOMPOSICIÓN EMPÍRICA DE MODOS.

- 1) Identificar todos los extremos, máximos y mínimos locales de $x(t)$.
- 2) Generar su envolvente superior $e_{sup}(t)$, e inferior $e_{inf}(t)$, mediante la interpolación spline cúbica de los extremos respectivos.
- 3) Promediar las dos envolventes para obtener la media local, $m_1(t) = (e_{sup}(t) + e_{inf}(t))/2$.
- 4) Extraer el primer componente, $h_1(t) = x(t) - m_1(t)$.
- 5) Verificar las propiedades de $h_1(t)$:
 - a. Si $h_1(t)$, cumple con las propiedades de un modo de oscilación intrínseco, entonces $IMF_1(t) = h_1(t)$ y se extrae de los datos para obtener el residuo, $r_1(t) = x(t) - IMF_1(t)$.
 - b. Si $h_1(t)$, no es un IMF, entonces éste reemplaza a los datos originales, $h_1(t) \rightarrow x(t)$.
- 6) Repetir el procedimiento de los pasos 1 al 5, hasta que se cumpla con alguno de los siguientes criterios de paro:
 - a. El N-ésimo modo de oscilación intrínseco $IMF_N(t)$ o el residuo final $r_N(t)$ se vuelve demasiado pequeño.
 - b. $r_N(t)$ es una función monótona de la cual no pueden extraerse más modos de oscilación.

IV.6.4. El espectro de Hilbert-Huang.

Siguiendo con la metodología indicada en la figura IV.37, se observa que una vez que se dispone de todos los IMFs ya no existe mayor dificultad en aplicar la transformada de Hilbert a cada uno de los N modos de oscilación resultantes para calcular su frecuencia instantánea, de acuerdo a la definición de Ville, es decir:

$$z_k(t) = IMF_k(t) + jH\{IMF_k(t)\} = A_k(t)e^{j\varphi_k(t)}, \quad k = 1, \dots, N \quad (251)$$

Por lo tanto, utilizando la ecuación de expansión en términos de los IMFS, es posible expresar a la señal $x(t)$ mediante:

$$x(t) = \sum_{k=1}^N A_k(t)e^{j2\pi \int f_{i,k}(t) dt} \quad (252)$$

donde en la ecuación anterior se omitió el residuo final $r_N(t)$ dado que es una función monótona o una constante y se utilizó el hecho de que la frecuencia instantánea corresponde a la derivada de la fase instantánea de la señal analítica, o $\varphi(t) = \int 2\pi f_i(t) dt$. Si la señal $x(t)$ es expandida mediante una representación de Fourier como:

$$x(t) = \sum_{k=1}^N A_k e^{j2\pi f_k t} \quad (253)$$

donde A_k y f_k son constantes, entonces comparando la ecuación anterior con la ecuación (252), la expansión obtenida mediante EMD representa una expansión de Fourier generalizada, donde tanto la amplitud y la frecuencia de cada componente son funciones del tiempo. La amplitud variable y la frecuencia instantánea no sólo mejoran la eficiencia de la expansión, sino que le permiten contender con datos no estacionarios. Además, la expansión en IMFs puede separar las modulaciones en amplitud y en frecuencia, y rompe con la restricción de la expansión de Fourier en funciones de amplitud y frecuencia constantes.

La expansión en términos de los modos intrínsecos de oscilación permite representar la amplitud y frecuencia instantáneas como funciones del tiempo en una gráfica tridimensional, en la cual la amplitud es el contorno sobre el plano tiempo-frecuencia. Esta distribución tiempo-frecuencia de la amplitud es denominada espectro de amplitud de Hilbert, o simplemente espectro de Hilbert, o espectro de Hilbert-Huang, $HHS(t, f)$ [92].

Existen diversas formas de representar el $HHS(t, f)$, e. g. mediante mapas con un código de colores, o mapas de contorno, ambas con o sin suavizado. En general, el espectro de Hilbert presenta una apariencia bastante distinta en comparación con la ofrecida por el espectrograma o el escalograma, debido a que el espectro de Hilbert solamente enfatiza las variaciones de frecuencia de cada IMF, mientras que las otras representaciones ofrecen un mapa suavizado que muestra una rica distribución de armónicos. Entonces, si se desea una forma más continua, pueden aplicarse métodos de suavizado espacial [93]. Un primer método consiste en emplear un filtro espacial o máscara de ponderación, con el cual se promedia directamente sobre un rango determinado de celdas tiempo-frecuencia. Una segunda opción consiste en utilizar un filtro pasa-bajas diseñado en el dominio de la frecuencia espacial y emplear el teorema de convolución para producir la imagen filtrada. Tanto en el dominio original de la imagen, como en el dominio de la frecuencia espacial, uno de los filtros más utilizados es el Gaussiano. Aunque el espectro suavizado ofrece una apariencia más similar a la obtenida con el espectrograma y el escalograma, no necesariamente brinda una mejor interpretación física que la representación original, dado que la información contenida ha sido degradada. Sin embargo, la versión suavizada permite identificar más fácilmente la densidad de energía y su tendencia de evolución como funciones del tiempo y la frecuencia, y por ello resulta útil echar un primer vistazo al espectro suavizado. En general, si se desean resultados

más cuantitativos el $HHS(t, f)$ original resulta más adecuado, mientras que si se desean resultados más cualitativos su versión suavizada será más conveniente.

A continuación se presenta un ejemplo para ilustrar la metodología para la obtención del $HHS(t, f)$ y familiarizarse con algunas ideas importantes. Considere una señal $x(t)$, creada a partir de la adición de tres componentes: dos señales, $x_1(t)$ y $x_2(t)$, de amplitud constante y modulación sinusoidal en frecuencia, y una señal, $x_3(t)$, correspondiente a un tono puro con una envolvente en amplitud de tipo Gaussiana. Las formas de onda de la señal y los componentes, así como la representación tiempo-frecuencia ideal, obtenida a partir de sus expresiones en amplitud y en frecuencia, se muestran en la figura IV.39.a y IV.39.b. Observe que, a diferencia de sus componentes, la señal $x(t)$ no posee un valor medio de cero, sin embargo este no es un requisito para este método. Los IMFs resultantes de la descomposición se muestran en la figura IV.39.c. Como resultado de aplicar sucesivamente el proceso de cernido, la señal $x(t)$ puede ser expandida con solamente nueve IMFs y una señal residual, lo cual representa una ventaja en comparación con la tradicional expansión de Fourier. Observe que los modos de oscilación $IMF_1(t) - IMF_3(t)$ presentan una gran correlación con los tres componentes individuales de la señal simulada. Los IMFs restantes contribuyen en menor medida a la energía de la señal reconstruida como se observa en las figura IV.39.d. Este ejemplo también ilustra que comúnmente el número de IMFs obtenidos no coincide con el número de componentes de la señal analizada.

Una prueba de la completitud de EMD se muestra en la figura IV.40, donde se observa que el error entre la señal reconstruida y la señal original es demasiado pequeño y puede atribuirse a la precisión de la computadora. Con respecto a la ortogonalidad de los IMFs, el índice de ortogonalidad es de 0.0183. A partir de la descomposición mostrada en la figura IV.39.c, se puede observar que los componentes tienen escalas de tiempo que no se traslapan localmente. Sin embargo, observe que los componentes $IMF_1(t)$ e $IMF_2(t)$ son intermitentes, ocasionando que contengan oscilaciones de la misma escala pero en diferentes localizaciones temporales. En otras palabras, los componentes son ortogonales localmente.

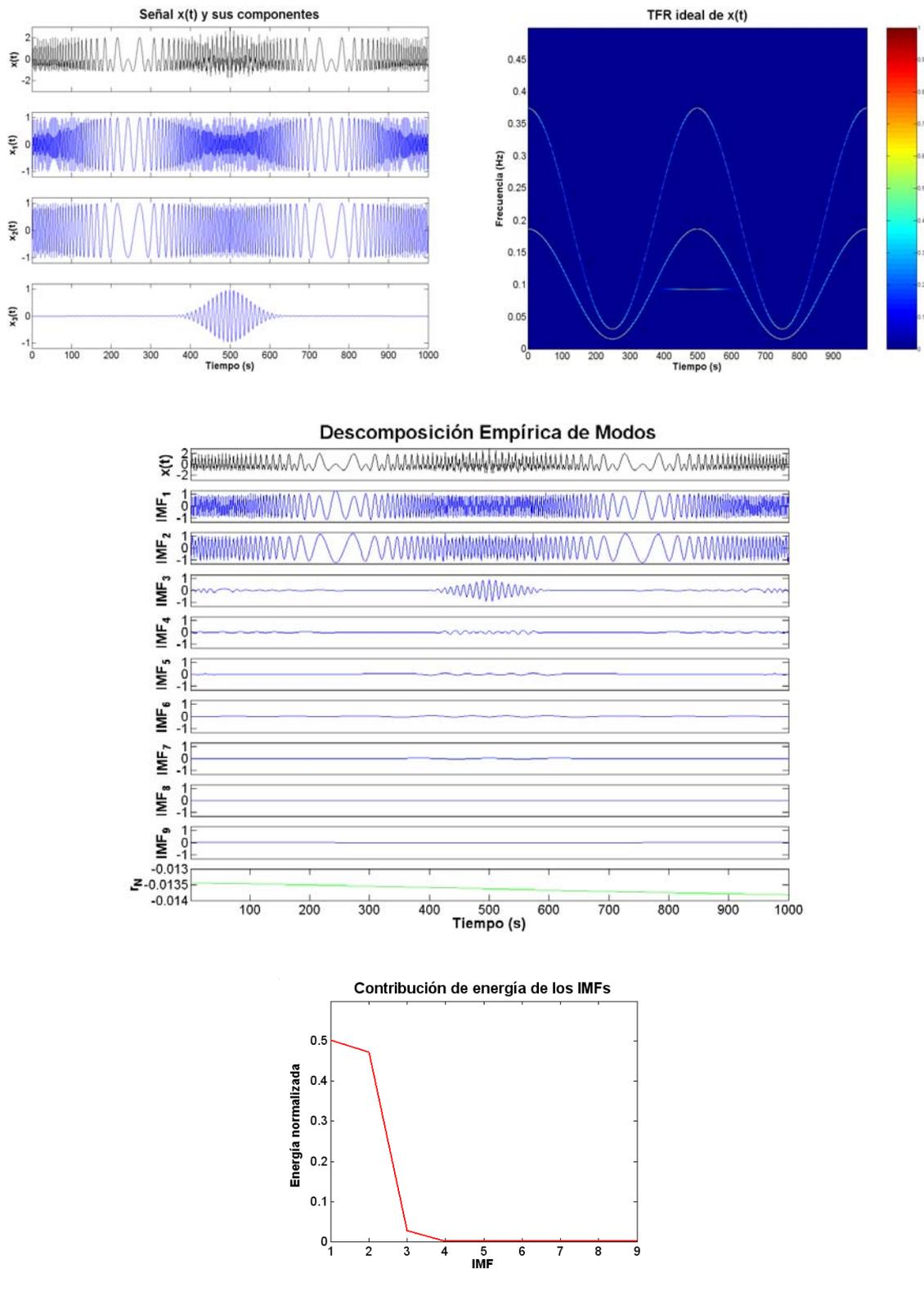


Fig. IV.39. a) Señal $x(t) = x_1(t) + x_2(t) + x_3(t)$, b) TFR ideal de la señal $x(t)$, c) Descomposición de la señal en $IMF_1(t) - IMF_9(t)$ y la señal de residuo $r_N(t)$, y d) Contribución a la energía total de la señal de cada modo de oscilación intrínseco (IMF).



El segundo paso para la obtención del $HHS(t, f)$ es la aplicación de la transformada de Hilbert a cada uno de los IMFs resultantes para obtener la expresión polar de su señal analítica correspondiente. Para este ejemplo, se determinaron las amplitudes y frecuencias instantáneas de los tres primeros modos de oscilación intrínsecos a través de la transformada de Hilbert. Los resultados se muestran en la figura IV.41. Observe que aunque el EMD extrae exitosamente las escalas de tiempo intrínsecas del proceso los $IMF_1(t)$ y $IMF_2(t)$ no presentan amplitudes exactamente constantes, como ocurre con sus contrapartes simuladas, lo cual puede deberse a la acumulación de errores de los ajustes de interpolación a lo largo del proceso de cernido. Por otra parte, observe que la frecuencia instantánea para el $IMF_3(t)$ presenta saltos relevantes, sin embargo ocurren en instantes de tiempo cuando su amplitud instantánea asociada es demasiado pequeña como para notarse en el espectro de Hilbert-Huang.

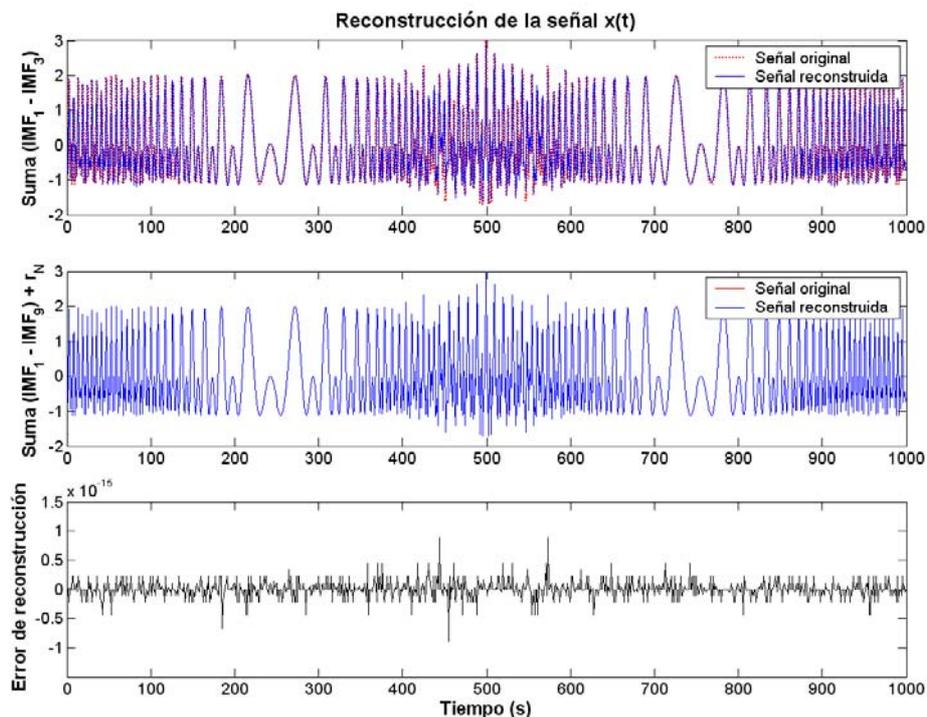


Fig. IV.40. Reconstrucción de la señal original $x(t)$ a partir de los modos de oscilación intrínsecos generados por el EMD. a) Señal $x(t)$ (línea roja discontinua) y la suma de las funciones base $IMF_1(t) - IMF_3(t)$ (línea azul continua), b) Señal $x(t)$ (línea roja discontinua) y la suma de las funciones base $IMF_1(t) - IMF_9(t)$ y el residuo $r_N(t)$ (línea azul continua), y c) Diferencia entre la señal original y la señal reconstruida.

a
b
c

Finalmente, a partir de las señales unidimensionales instantáneas, es posible conformar el espectro de Hilbert-Huang de la señal. En este caso la amplitud instantánea representa la intensidad de la imagen correspondiente a la representación tiempo-frecuencia. Las versiones original y suavizada del $HHS(t, f)$ se muestran en la figura IV.41.

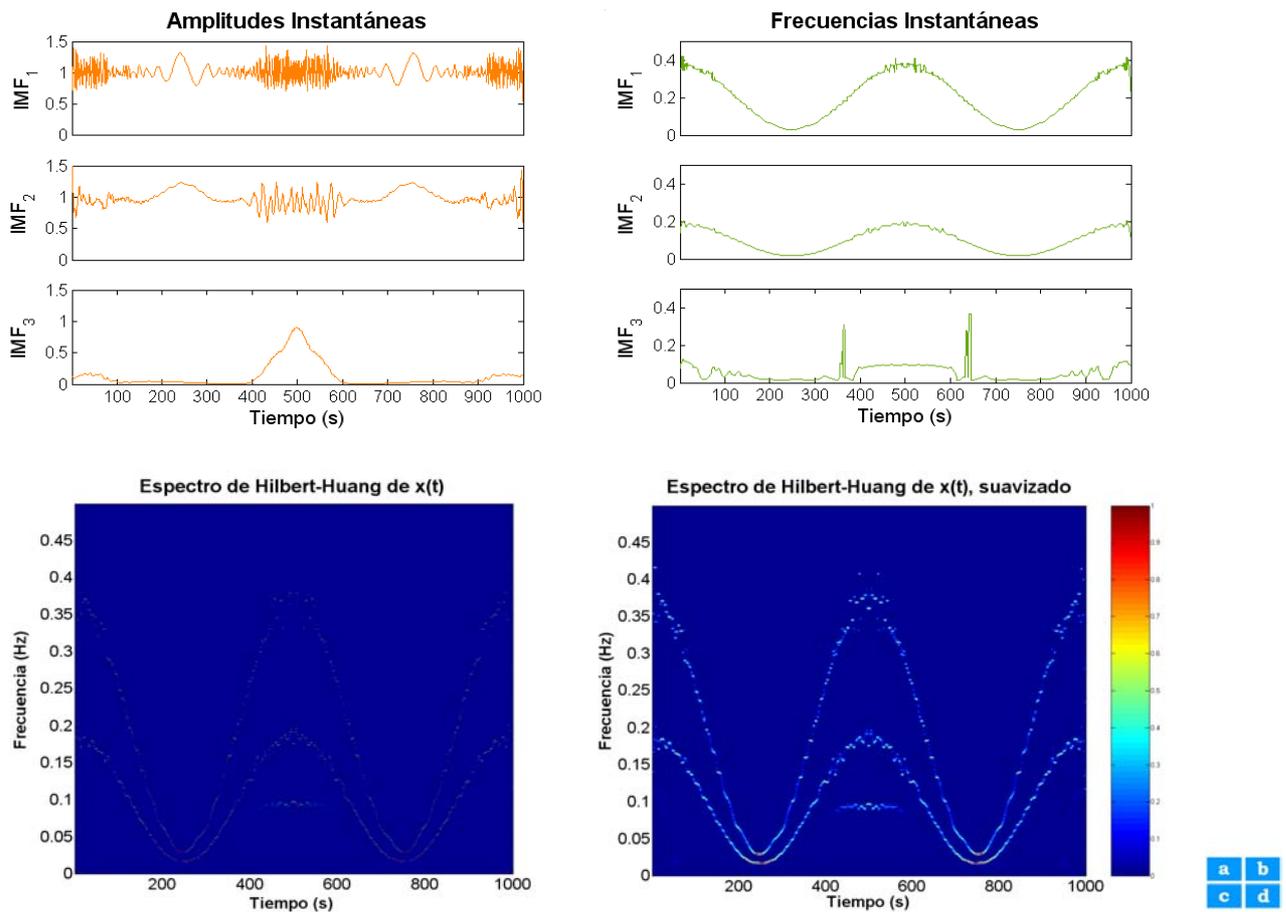


Fig. IV.41. a) Amplitudes instantáneas, y b) frecuencias instantáneas de los primeros tres modos de oscilación .
 c) Espectro de Hilbert-Huang, $HHS(t, f)$ original, y d) $HHS(t, f)$ suavizado con un filtro pasa-bajas Gaussiano.

Referencias.

- [1] L. Cohen, *Time-Frequency Analysis*. Englewood Cliffs, NJ: Prentice-Hall, 1995.
- [2] J. Jeong, W. Williams, "Kernel design for reduced interference distributions," *IEEE Trans. Sig. Proc.*, vol. 40, pp. 402-4012, 1992.
- [3] H. Hlawatsch, G. F. Boudreaux-Bartels, "Linear and quadratic time-frequency signal representations," *IEEE Signal Processing Magazine*, vol. 9, pp. 21-67, 1992.
- [4] H. I. Choi, W. J. Williams, "Improved time-frequency representation of multicomponent signals using exponential kernels," *IEEE Trans. on Acoust., Speech, Signal Processing*, vol. 37, pp. 862-871, 1989.
- [5] D. Gabor, "Theory of communication", *Journal of the IEE*, vol. 93, pp. 429-457, 1946.
- [6] E. P. Wigner, "On the quantum correction for thermodynamic equilibriumm," *Physical Review*, vol. 40, pp. 749-759, 1932.
- [7] J. Ville, "Theorie et applications de la notion de signal analytique, *Cables et Transmissions*, vol. 2A, pp. 61-74, 1948. Traducción al ingles por I. Sellin, "Theory and applications of the notion of complex signal," RAND Comporation Technical Report T-92, Santa Monica, CA, 1958.
- [8] R. Koenig, H. K. Dunn, L. Y. Lacy, "The sound spectrograph," *J. Acoust. Soc. Amer.*, vol. 18, pp-19-49, 1946.
- [9] A. W. Rihaczek, "Signal energy distribution in time and frequency," *IEEE Trans. Info. Theory*, vol 14, pp. 369-374, 1968.
- [10] C. H. Page, "Instantaneous power spectra," *Jour. Appl. Phys.*, vol. 23, pp. 103-106, 1952.
- [11] H. Margenau, R. N. Hill, "Correlation between measurements in quantum theory," *Prog. Theoret. Phys.*, vol. 26, pp. 722-738, 1961.
- [12] L. Cohen, "Generalized phase-state distribution functions," *J. Math. Phys.*, vol. 7, pp. 781-786, 1966.
- [13] T. Claasen, W. Mecklenbräuker, "The Wigner distribution- a tool for time-frequency signal analysis- Part III: Relations with other time-frequency signal transformations," *Philips Jour. Research.*, vol. 35, pp. 372-389, 1980.
- [14] W. Williams, J. Jeong, "Reduced interference time-frequency distributions," en *Time-Frequency Signal Analysis*, editado por B. Boashash, pp. 74-97, Longman Cheshire, 1992.
- [15] P. Z. Peebles, *Probability, random variables, and random signal principles*. McGraw-Hill, 3ra. Ed., 1993.
- [16] B. Boashash, G. Jones, "Instantaneous frequency and time-frequency distributions," en *Time-Frequency Signal Analysis*, editado por B. Boashash, pp. 43-73, Longman Cheshire, 1992.
- [17] L. Cohen, "Time-Frequency Distributions – A Review," *Proc. of the IEEE*, vol. 77, pp. 941-981, 1989.
- [18] K. Kodera, R. Gendrin, C. Villedary, "Analysis of time-varying signals with small BT values," *IEEE Trans. Acoust. Speech, Signal Processing*, vol. 26, pp. 64-76, 1978.
- [19] D. L. Jones, T. W. Parks, "A resolution comparison of several time-frequency representations," *IEEE Trans. on Signal Processing*, vol. 40, pp. 413-420, 1992.

- [20] B. Boashash, "Note on the use of the Wigner distribution for time-frequency signal analysis," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, pp. 1518-1521, 1988.
- [21] L. Cohen, "A primer on time-frequency distributions," en *Time-Frequency Signal Analysis*, editado por B. Boashash, pp. 3-42, Longman Cheshire, 1992.
- [22] J. Moyal, "Quantum mechanics as a statistical theory," *Proc. Camb. Phil. Soc.*, vol. 45, pp. 99-124, 1949.
- [23] R. Hudson, "When is the Wigner quasi-probability function non-negative?," *Rep. Math. Phys.*, vol. 6, pp. 249-252, 1974.
- [24] S. Kadambe, G. Boudreaux-Bartels, "A comparison of the existence of 'cross terms' in the Wigner distribution and the squared magnitude of the wavelet transform and the short time Fourier transform," *IEEE Trans. Sig. Proc.*, vol. 40, no. 10, 1992.
- [25] P. Flandrin, "Some features of time-frequency representations of multi-component signals," *IEEE Int. Conf. Acoust., Speech, Sig. Proc.*, pp. 41.B.4.1-4, San Diego, CA, 1984.
- [26] W. Williams, J. Jeong, "Reduced interference time-frequency distributions," en *Time-Frequency Signal Analysis*, editado por B. Boashash, pp. 74-97, Longman Cheshire, 1992.
- [27] B. Boashash, P. Black, "An efficient real-time implementation of the Wigner-Ville distribution," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-35, pp. 1611-1618, 1987.
- [28] D. Gabor, "Communication theory and physics," *Transactions of the IRE Professional Group on Information Theory*, vol. 1, no. 1, pp. 48-59, 1953.
- [29] P. Hill, "Dennis Gabor - Contributions to communication theory & signal processing," *EUROCON, The International Conference on "Computer as a Tool"*, pp. 2632-2637, Varsovia, 9-12 Sept, 2007.
- [30] J. Allen, L. Rabiner, "A unified approach to short-time Fourier transform analysis and synthesis," *Proc. IEEE*, vol. 65, no. 11, pp. 1558-1564, 1977.
- [31] F. Harris, "On the use of windows for harmonic analysis with the discrete Fourier transform," *Proc. IEEE*, vol. 66, no. 1, pp. 51-84, 1978.
- [32] W. Mark, "Spectral analysis of the convolution and filtering of non-stationary stochastic processes," *Jour. Sound Vib.*, vol. 11, pp. 16-63, 1970.
- [33] W. Martin, "Wigner-Ville spectral analysis of non-stationary processes," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-33, pp. 1461-1470, 1985.
- [34] N. Marinovic, G. Eichmann, "An expansion of Wigner distribution and its application," *Proc. IEEE Int. Conf. ASSP*, vol. 3, pp. 27.3.1-27.3.4, 1985.
- [35] P. Flandrin, "Some features of time-frequency representations of multi-component signals," *IEEE Int. Conf. Acoust., Speech, Sig. Proc.*, pp. 41.B.4.1-4, San Diego, CA, 1984.
- [36] S. Qian, J. Morris, "Wigner distribution decomposition and cross-terms deleted representation," *Signal Processing*, vol. 27, no. 2, pp. 125-144, 1992.
- [37] M. Sun, C. Li, L. Sekhar, R. Scabassi, "Elimination of the cross-components of the discrete pseudo Wigner-Ville distribution via image processing," *Proc. IEEE-ICASSP*, pp. 2230-2233, 1989.

-
- [38] K. Kodera, C. Villedary, R. Gendrin, "A new method for the numerical analysis of time-varying signals with small BT values," *Phys. Earth Planet. Interiors*, no. 12, pp. 142-150, 1976.
- [39] F. Auger, P. Flandrin, "Improving the readability of time-frequency and time-scale representations by the reassignment method," *IEEE Trans. on Signal Proc.*, vol. 43, no. 5, pp. 1068-1089, 1995.
- [40] F. Auger, P. Flandrin, "Generalization of the reassignment method to all bilinear time-frequency and time-scale representations," *Proc. IEEE-ICASSP*, pp. IV-317-320, 1994.
- [41] S. Kay, S. Marple, "Spectrum analysis-a modern perspective," *Proc. IEEE*, vol. 69, no. 11, 1380-1419, 1981.
- [42] S. Marple, *Digital Spectral Analysis with Applications*, Prentice-Hall, 1987.
- [43] S. Marple, "A tutorial overview of modern spectral estimation," *Proc. IEEE-ICASSP*, pp. 2152-2157, 1989.
- [44] J. Proakis, D. Manolakis, *Digital Signal Processing: Principles, Algorithms, and Applications*, 3era. Edición, Prentice-Hall, 1996.
- [45] M. Bartlett, "Smoothing periodograms from time series with continuous spectral," *Nature*, vol. 161, no. 4096, pp. 686-687, 1948.
- [46] R. Blackman, J. Tukey, *The measurement of power spectra from the point of view of communications Engineering*, Dover, 1959.
- [47] P. Welch, "The use of the fast Fourier transform for the estimation of power spectra: A method based on time averaging over short modified periodograms," *IEEE Trans. Audio Electroacoust.*, vol. AU-15, no. 2, 1967.
- [48] M. Aboy, O. Márquez, J. McNames, R. Hornero, T. Thong, B. Goldstein, "Adaptive modeling and spectral estimation of nonstationary biomedical signals based on Kalman filtering," *IEEE Trans. Biomed. Eng.*, vol. 52, no. 8, pp. 1485-1489, 2005.
- [49] R. Zou, H. Wang, K. Chon, "A robust time-varying identification algorithm using basis functions," *Ann. Biomed. Eng.*, vol. 31, pp. 840-853, 2003.
- [50] R. Zou, K. Chon, "Robust algorithm for estimation of time-varying transfer functions," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 2, pp. 219-227, 2004.
- [51] J. Burg, "Maximum entropy spectral analysis," *Proc. 37th Meeting of the Society of Exploration Geophysicists*, 1967.
- [52] S. Haykin, *Adaptive filter theory*, 3ra. Edición, Prentice-Hall, 1996.
- [53] L. Marple, "A new autoregressive spectrum analysis algorithm," *IEEE Trans. Acoust. Speech, Signal Processing*, vol. ASSP-28, no. 4, pp. 441-454, 1980.
- [54] J. Makhoul, "Linear prediction: a tutorial review," *Proc. IEEE*, vol. 63, no. 4, pp. 561-580, 1975.
- [55] T. Barnard, "The maximum entropy spectrum and the Burg technique," *Technical Report no. 1: Advanced Signal Processing, Texas Instruments Incorporated*, 1975.
- [56] N. Andersen, "Comments on the performance of maximum entropy algorithms," *Proc. IEEE*, vol. 66, no. 11, pp. 1581-1582, 1978.
- [57] P. Fougere, E. Zawalick, H. Radoski, "Spontaneous line splitting in maximum entropy power spectrum analysis," *Phys. Earth and Plan. Inter.*, vol. 12, pp. 201-207, 1976.
-

- [58] T. Ulrych, R. Clayton, "Time series modeling and maximum entropy," *Phys. Earth and Plan. Int.*, vol. 12, pp. 188-200, 1976.
- [59] E. Robinson, "A historical prespective of spectrum estimation," *Proc. IEEE*, vol. 70, no.9, pp. 885-907, 1982.
- [60] J. Burg, "Maximum entropy spectral analysis," *Tesis doctoral*, Universidad de Standford, 1975. Resumen.
- [61] L. Padovese, "Hybrid time-frequency methods for non-stationary mechanical sigal analysis," *Mechanical Systems and Signal Processing*, vol. 18, pp. 1047-1064, 2004.
- [62] J. Mendel, *Lessons in Digital Estimation Theory*, Prentice-Hall, 1986.
- [63] S. Haykin, *Adaptive filter theory*, Prentice-Hall, 1986.
- [64] L. Hogben, *Handbook of linear algebra*, Chapman & Hall/CRC, 2007.
- [65] N. Hubing, S. Alexander, "Statistical analysis of the soft constrained initialization of recursive least squares algorithms", *Proc. ICASSP-90*, vol. 3, pp. 1277-1280, 1990.
- [66] N. Hubing, S. Alexander, "Statistical analysis of initialization methods for RLS adaptive filters", *IEEE Trans. Signal Processing*, vol. 39, no. 8, pp. 1793-1804, 1991.
- [67] A. Sayed, "Adaptive filters", IEEE Press, Wiley-Interscience: IEEE Press, 2008.
- [68] E. Eleftheriou, D. Falconer, "Tracking properties and steady-state performance of RLS adaptive filter algorithms," *IEEE Trans. Acoust, Speech, Signal Processing*, vol. 34, no. 5, pp. 1097-1109, 1986.
- [69] B. Widrow, E. Walach, "On the statistical efficiency of the LSM algorithm with non-stationary inputs," *IEEE Trans. Information Theory*, vol. IT-30, no. 2, pp. 211-221, 1984.
- [70] T. Fortescue, L. Kershenbaum, B. Ydstie, "Implementation of self-tuning regulators with variable forgetting factors," *Automatica*, vol. 17, pp. 831-835, 1981.
- [71] A. Cordero, D. Mayne, "Deterministic convergence of a self-tuning regulator with variable forgetting factor," *IEE Proc.*, vol. 128, pp. 16-23, 1981.
- [72] P. Wellstead, S. Sanoff, "Extended self-tuning algorithm", *Int. J. Control*, vol. 34, pp. 433-455, 1981.
- [73] Y. Cho, S. Kim, E. Powers, "Time-frequency analysis using AR models with variable forgetting factors," *Proc. ICASSP'90*, pp. 2479-2482, 1990.
- [74] Y. Cho, S. Kim, E. Powers, "Time-varying spectral estimation using AR models with variable forgetting factors," *IEEE Trans. Signal Processing*, vol. 39, pp. 1422-1425, 1991.
- [75] B. Toplis, S. Pasupathy, "Tracking improvements in fast RLS algorithm using a variable forgetting factor," *IEEE. Trans. Acoust. Speech, Signal Processing*, vol. 36, pp. 206-227, 1988.
- [76] R. Kalman, "A new approach to linear filtering and prediction problems," *Transactions of the ASME-Journal of Basic Engineering*, vol. 82 D, pp. 35-45, 1960. Reimpresión.
- [77] P. Swerling, "First order error propagation in a stagewise smoothing procedure for satellite observations," *J. Astronautical Sci.*, vol. 6, pp. 46-52, 1959. Reimpresión.
- [78] R. Stratonovich, "Application of the theory of Markoff processes in optimal sigal detection," *Radio Eng. Electron. Phy.*, vol. 1, pp. 1-19, 1960.
- [79] Y. Ho, "A Bayesian approach to problems in stochastic estimation and control," *IEEE Trans. Automatic Control*, vol. AC-9, pp. 333-339, 1964.

-
- [80] T. Kailath, "An innovations approach to least-squares estimation-Part I: linear filtering in additive white noise," *IEEE Trans. Automatic Control*, vol. AC-13, pp. 646-655, 1968.
- [81] W. Penny, S. Roberts, "Dynamic linear models, recursive least squares and steepest-descent learning," Reporte técnico, Imperial College of Science, Technology and Medicine, Londres, pp. 1-11, 1998.
- [82] N. Sorenson, "Kalman filtering techniques," en *Advances in Control Systems Theory and Applications*, editado por C. Leondes, vol. 3, pp. 219-292, Academic Press, 1966.
- [83] M. Aboy, J. McNames, O. Márquez, R. Hornero, T. Thong, B. Goldstein, "Power spectral density estimation and tracking of nonstationary pressure signals based on Kalman filtering", *Proc. 26th Annual International Conference of the IEEE-EMBS*, pp. 156-159, 2004.
- [84] P. Goupillaud, A. Grossman, J. Morlet, "Cycle-octave and related transforms in seismic signal analysis, *Geoexploration*, vol. 23, pp. 85-102, 1984.
- [85] I. Dauechies, "The wavelet transform, time-frequency localization and signal analysis," *IEEE Trans. on Info. Theory*, vol. 36, no. 5, pp. 961-1005, 1990.
- [86] S. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," *IEEE Trans. on Pattern Analysis and Machine Intell*, vol. 11, no. 7, pp. 674-693, 1989.
- [87] O. Rioul, M. Vetterli, "Wavelets and signal processing," *IEEE Sig. Proc. Magazine*, pp. 14-38, Octubre, 1991.
- [88] K. Gram-Hansen, "A bandwidth concept for CPB time-frequency analysis," *IEEE Int. Conf. Acoust. Speech, Sig. Proc.*, pp. 2033-2036, 1991.
- [89] O. Rioul, P. Flandrin, "Time-scale energy distributions: a general class extending wavelet transforms," *IEEE Trans. Sig. Proc.*, vol. 40, no. 7, 1746-1757, 1992.
- [90] P. Kumar, E. Foufoula-Georgiou, "Wavelet analysis in geophysics: an introduction," en *Wavelets in Geophysics*, editado por E. Foufoula-Georgiou y P. Kumar, pp. 1-44, Academic Press, 1994.
- [91] P. Flandrin, O. Rioul, "Affine smoothing of the Wigner-Ville distribution," *IEEE Int. Conf. Acoust. Speech, Sig. Proc.*, pp. 2455-2458, 1990.
- [92] N. Huang, Z. Shen, S. Long, M. Wu, H. Shih, A. Zheng, N. Yen, C. Tung, H. Liu, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and nonstationary time series analysis," *Proc. Roy. Soc. Lond.*, vol. A 454, pp. 903-995, 1998.
- [93] R. González, R. Woods, S. Eddins, *Digital image processing using MATLAB*, Prentice-Hall, 2004.

V. METODOLOGÍA

V.1. Metodología general.

Para la comparación de las diversas representaciones tiempo-frecuencia se emplearon señales de sonidos simulados, dado que su representación tiempo-frecuencia ideal se conoce de antemano, permitiendo la obtención de medidas de desempeño. Posteriormente, la técnica de análisis tiempo-frecuencia resultante fue aplicada a señales de sonidos torácicos adquiridos. La metodología de este trabajo incluye: la simulación de las diversas señales de sonidos torácicos y su representación tiempo-frecuencia ideal, el computo de las representaciones tiempo-frecuencia con un conjunto de definido de parámetros, la selección de la representación tiempo-frecuencia representativa de cada técnica, la comparación final entre las técnicas tiempo-frecuencia, y la aplicación de la técnica tiempo-frecuencia seleccionada para el análisis de sonidos reales adquiridos.

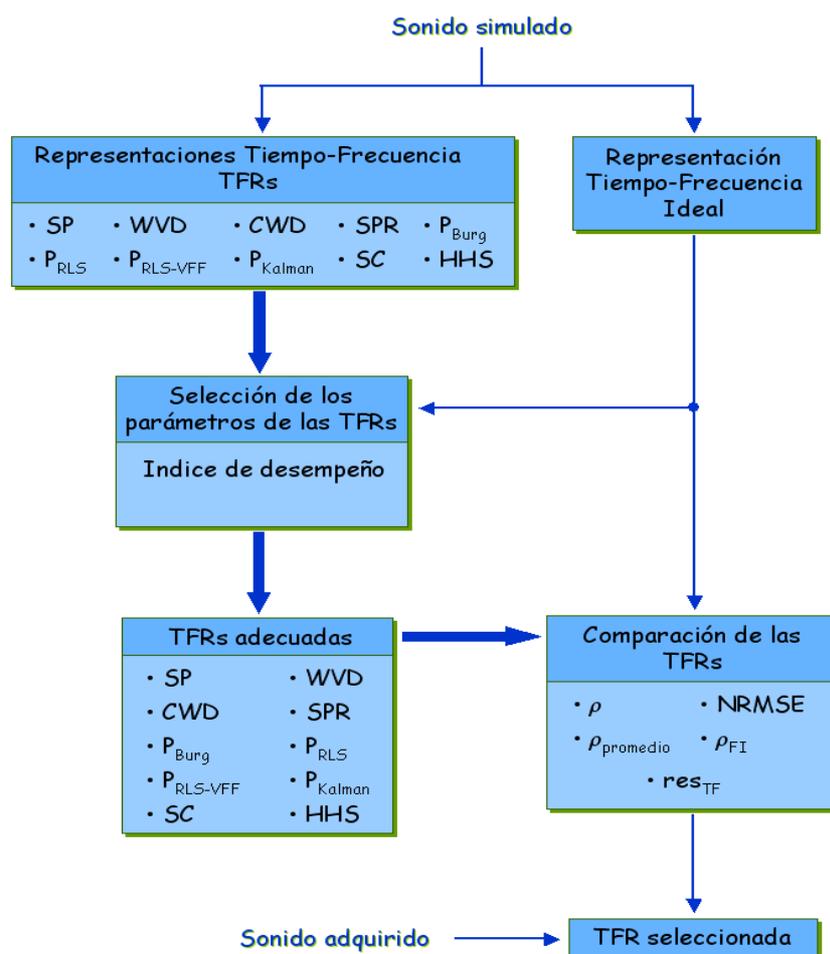


Fig. V.1. Diagrama de bloques de la metodología a seguir para las señales de las que se dispone su representación tiempo-frecuencia ideal.

La metodología anterior resulta válida para las señales de prueba no biológicas, los dos primeros sonidos cardiacos y los estertores sibilantes, ya que en todos estos casos se dispone de la señal simulada así como de su representación tiempo-frecuencia ideal, considerada la referencia. Sin embargo, para el caso de los estertores crepitantes se dispone de la señal simulada pero no de su representación tiempo-frecuencia ideal. Pero dado que esta señal comparte con el segundo cardiaco el hecho de que su forma de onda en el tiempo presenta un incremento sucesivo en los intervalos de cruce por cero, i.e. una caída en su contenido en frecuencia, se consideró como su representación tiempo-frecuencia ideal aquella obtenida con la técnica de análisis seleccionada para este sonido cardiaco.

En la figura V.1 se presenta el diagrama a bloques con la metodología empleada para la selección de la representación tiempo-frecuencia más adecuada para el análisis de los dos primeros sonidos cardiacos y sus componentes, y los estertores sibilantes. Esta misma metodología se empleó para las señales no biológicas simuladas, con excepción de la aplicación a señales adquiridas. Por su parte, en la figura V.2 se presenta la metodología seguida para el caso de los estertores crepitantes.

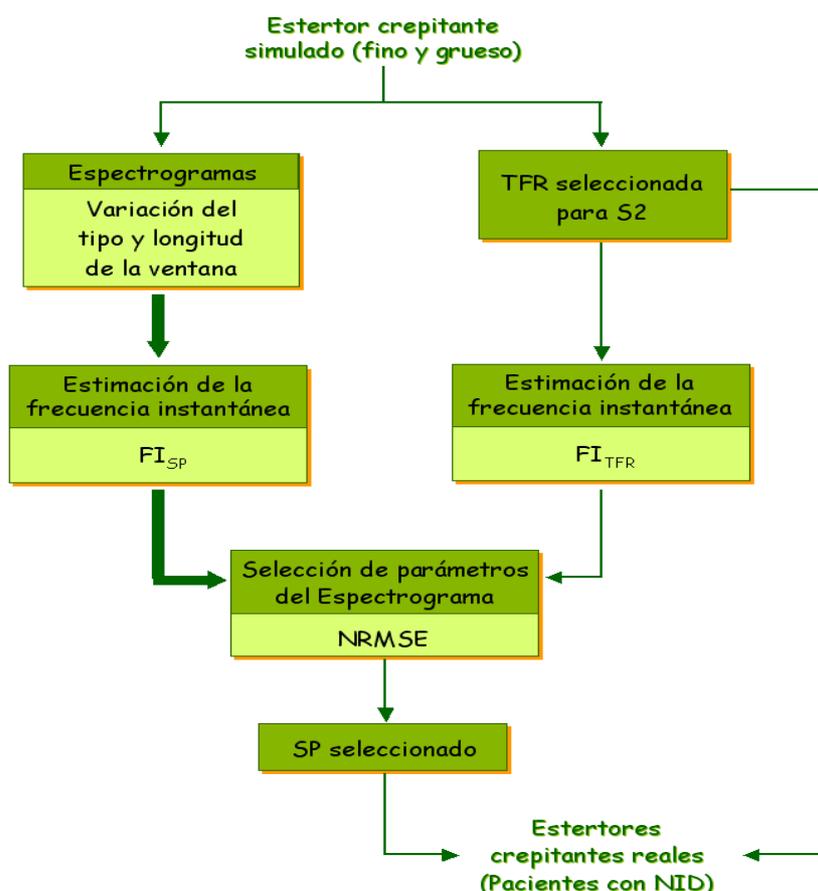


Fig. V.2. Diagrama de bloques de la metodología a seguir para los sonidos respiratorios adventicios discontinuos.

De ambos diagramas a bloques se observa que las etapas básicas de este trabajo incluyen: la creación de las señales simuladas, la formación de sus representaciones tiempo-frecuencia ideales, la definición de los índices de desempeño empleadas para evaluar las diversas técnicas, la definición del conjunto de parámetros de las representaciones tiempo-frecuencia, y la aplicación de las técnicas seleccionadas a sonidos torácicos adquiridos. A continuación se describen estas etapas.

V.2 Simulación computacional de las señales y sus representaciones tiempo-frecuencia ideales.

Los sonidos torácicos incluyen una variedad de sonidos entre los que se encuentran los sonidos respiratorios normales y anormales, el primer y segundo ruido cardíaco, los sonidos gástricos y los sonidos musculares, entre otros. Considerando que no existe un análisis tiempo-frecuencia exhaustivo de los sonidos respiratorios y con el propósito de evaluar las diferentes técnicas en la presente tesis, en esta sección se incluye la simulación de los sonidos torácicos más conocidos (ruidos cardíacos) así como la simulación matemática de los sonidos pulmonares adventicios (estertores crepitantes y sibilantes). Entre los sonidos torácicos, los ruidos cardíacos han sido analizados con mayor profundidad, y uno de los propósitos de interés se ha centrado en establecer una representación tiempo-frecuencia con mayor resolución que facilite comprender o establecer los mecanismos subyacentes en su generación. La simulación matemática de los ruidos cardíacos, con base en los posibles eventos fisiológicos que los constituyen, ha incluido funciones matemáticas que permiten reproducir la forma de onda del primer y segundo ruido cardíaco.

En esta sección se aborda la simulación computacional y la representación tiempo-frecuencia ideal de señales no biológicas, de los sonidos cardíacos S1 y S2, así como de los sonidos respiratorios adventicios continuos y discontinuos. Mientras que en la siguiente sección el interés se enfocará en definir índices de comparación que permitan determinar la mejor representación tiempo-frecuencia de los sonidos torácicos.

V.2.1. Señales no biológicas

La simulación computacional de señales no biológicas es común en la ingeniería puesto que sirven como señales de prueba para los diversos algoritmos de procesamiento y resultan de gran utilidad para ganar un entendimiento acerca de dichos algoritmos, pero más importante aún, permiten tener cierto grado de confiabilidad en la interpretación de los resultados obtenidos al aplicarlos a señales reales. Como primer paso, en esta tesis se simula

una señal no biológica con base en señales sinusoidales puras con un cambio abrupto en su frecuencia instantánea, mediante dos enfoques distintos. El primero de ellos mediante la concatenación de dos sinusoidales puras de duración finita con diferentes frecuencias y el segundo, mediante el empleo del modelado autorregresivo variante en el tiempo. Cabe mencionar que los escenarios de simulación indicados se han empleado frecuentemente para la evaluación de diversos algoritmos de representación tiempo-frecuencia, [1-8]. A continuación se definen y generan las señales no biológicas, mediante cada uno de los enfoques antes mencionados, así como sus representaciones tiempo-frecuencia ideales.

V.2.1.1. Señal no biológica concatenando sinusoidales puras.

Consideremos primero el caso de sinusoidales puras cuya frecuencia cambia abruptamente de un valor a otro [1-4]. Este ejemplo clásico se emplea para evaluar representaciones tiempo-frecuencia debido a que considera un cambio repentino en los parámetros de la señal y que las técnicas de procesamiento deben ser capaces de seguir con la mayor resolución tiempo-frecuencia posible. La señal simulada se genera mediante la concatenación de dos sinusoidales de duración finita siguiendo la fórmula:

$$s(t) = \begin{cases} \sin(2\pi(0.1F_s)t) & , \quad 0 \leq t < 250 \text{ s} \\ \sin(2\pi(0.2F_s)t) & , \quad 250 \leq t < 500 \text{ s} \end{cases} \quad (233)$$

donde F_s corresponde a la frecuencia de muestreo, que se estableció en $F_s = 1 \text{ Hz}$. Como indica la ecuación (233), la señal simula un cambio abrupto a la mitad del registro, en la muestra 250, mediante el cambio de una sinusoidal de baja frecuencia a una sinusoidal de alta frecuencia, ver figura V.3.

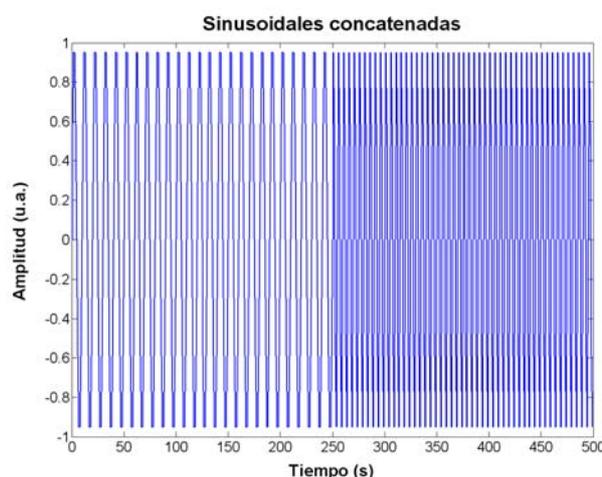


Fig. V.3. Señal simulada consistente en la concatenación de dos sinusoidales puras de duración finita.

Observe que la amplitud instantánea de la señal de la figura V.3 es la misma, mientras que su frecuencia instantánea es diferente, es decir:

$$A_i(t) = 1 \quad , \quad 0 \leq t < 500 \text{ s} \quad (234)$$

además:

$$f_i(t) = \begin{cases} 0.1 \text{ Hz} & , \quad 0 \leq t < 250 \text{ s} \\ 0.2 \text{ Hz} & , \quad 250 \text{ s} \leq t < 500 \text{ s} \end{cases} \quad (235)$$

A partir de estas dos últimas ecuaciones es posible construir la representación tiempo-frecuencia ideal y que se representa mediante una imagen; mapeando los valores de amplitud o intensidad de la señal a su celda correspondiente de frecuencia con base en su valor de frecuencia instantánea, al tiempo t . Para ello, es posible crear una matriz cuyas columnas correspondan a cada instante de tiempo de la señal y cuyos NF renglones correspondan a valores de frecuencia igualmente espaciados en el rango de $0-0.5F_s$. Por lo tanto, la representación tiempo-frecuencia ideal resultará en una matriz de dimensiones $NF \times N$, donde N corresponde a la longitud de la señal. Para el caso de esta primera señal simulada, el total de celdas en la frecuencia fue ajustado en $NF = 512$. En la figura V.4 se esquematiza el procedimiento de mapeo para la conformación de la representación tiempo-frecuencia ideal. Mientras que en la figura V.5 se presenta la representación tiempo-frecuencia ideal para la primer señal simulada.

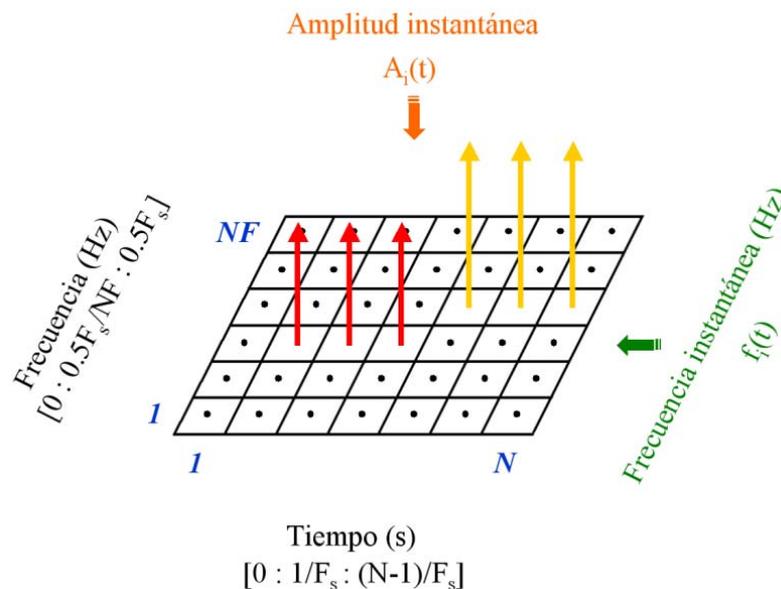


Fig. V.4. Esquematización del mapeo de la amplitud y frecuencia instantánea de una señal para obtener su representación tiempo-frecuencia ideal. La amplitud de la señal se define por el color en la paleta de colores que se adjunta a la imagen.

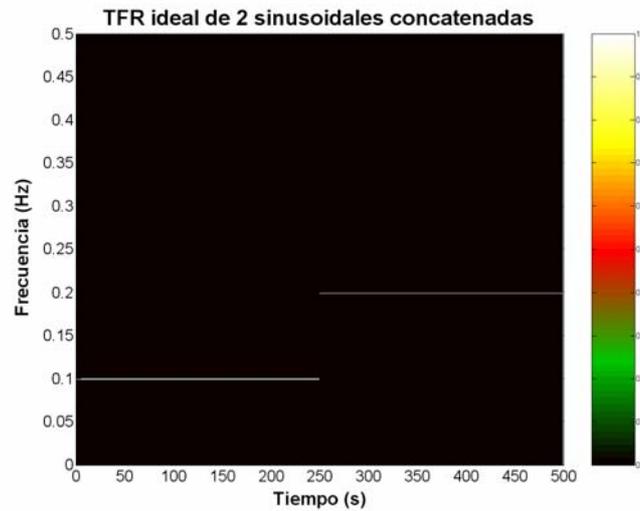


Fig. V.5. Representación tiempo-frecuencia ideal de la señal compuesta mediante la concatenación de dos sinusoidales.

V.2.1.2. Señal no biológica mediante el modelo TVAR.

El segundo enfoque para simular la señal no biológica consiste en el empleo del modelado autorregresivo variante en el tiempo de segundo orden. La señal consiste en 500 muestras de un proceso estocástico, generado mediante la siguiente ecuación de diferencias recursiva, es decir:

$$x(n) + \sum_{k=1}^2 a_k(n)x(n-k) = w(n) \quad (236)$$

donde los coeficientes del modelo presentan el siguiente comportamiento:

$$a_1(n) = \begin{cases} 1.618 & , \quad 0 \leq n \leq 249 \\ 0.618 & , \quad 250 \leq n \leq 499 \end{cases} \quad (237)$$

$$a_2(n) = -1$$

y $w(n)$ es un proceso de ruido blanco con media cero y varianza unitaria. La señal fue generada empleando una frecuencia de muestreo de 1 Hz. Los coeficientes y la señal de salida del modelo del modelo TVAR se muestran en la Fig. V.6. Observe que el parámetro $a_1(n)$, y por lo tanto la estadística de la señal, cambia abruptamente a la mitad de la duración de la señal, $t = 300$ s, cambiando su contenido de bajas a altas frecuencias.

Empleando identidades trigonométricas se puede demostrar que el coeficiente TVAR $a_1(n)$ corresponde a:

$$a_1(n) = 2 \cos(2\pi f_i(n)) \quad (238)$$

y que en consecuencia, la representación tiempo-frecuencia ideal debe presentar un incremento en la frecuencia instantánea al tiempo $t = 300$ s, que corresponde a:

$$f_i(t) = \begin{cases} 0.1 \text{ Hz} & , \quad 0 < t < 299 \text{ s} \\ 0.2 \text{ Hz} & , \quad 300 \text{ s} < t < 599 \text{ s} \end{cases} \quad (239)$$

la razón de esta simulación es que ha sido ampliamente empleada para la valoración de diversos algoritmos de representación tiempo-frecuencia mediante la estimación del contenido espectral de la señal con base en modelos autorregresivos [5-8]. En la figura V.6 se muestra la representación tiempo-frecuencia ideal para la señal definida por las ecuaciones anteriores, en contraste con la señal simulada con los parámetros indicados por las ecuaciones (234) y (235), la señal generada por el modelo TVAR contiene diferentes amplitudes instantáneas. La representación tiempo-frecuencia ideal fue generada empleando la fórmula teórica para la estimación espectral paramétrica basada en el modelado TVAR, empleando 512 puntos en la estimación de la FFT, es decir:

$$P_{TVAR}(n, f) = \frac{1}{\left| 1 + \sum_{k=1}^p a_k(n) e^{-j2\pi f k} \right|^2} = \frac{1}{|FFT[\bar{a}(n)]|^2} \quad (240)$$

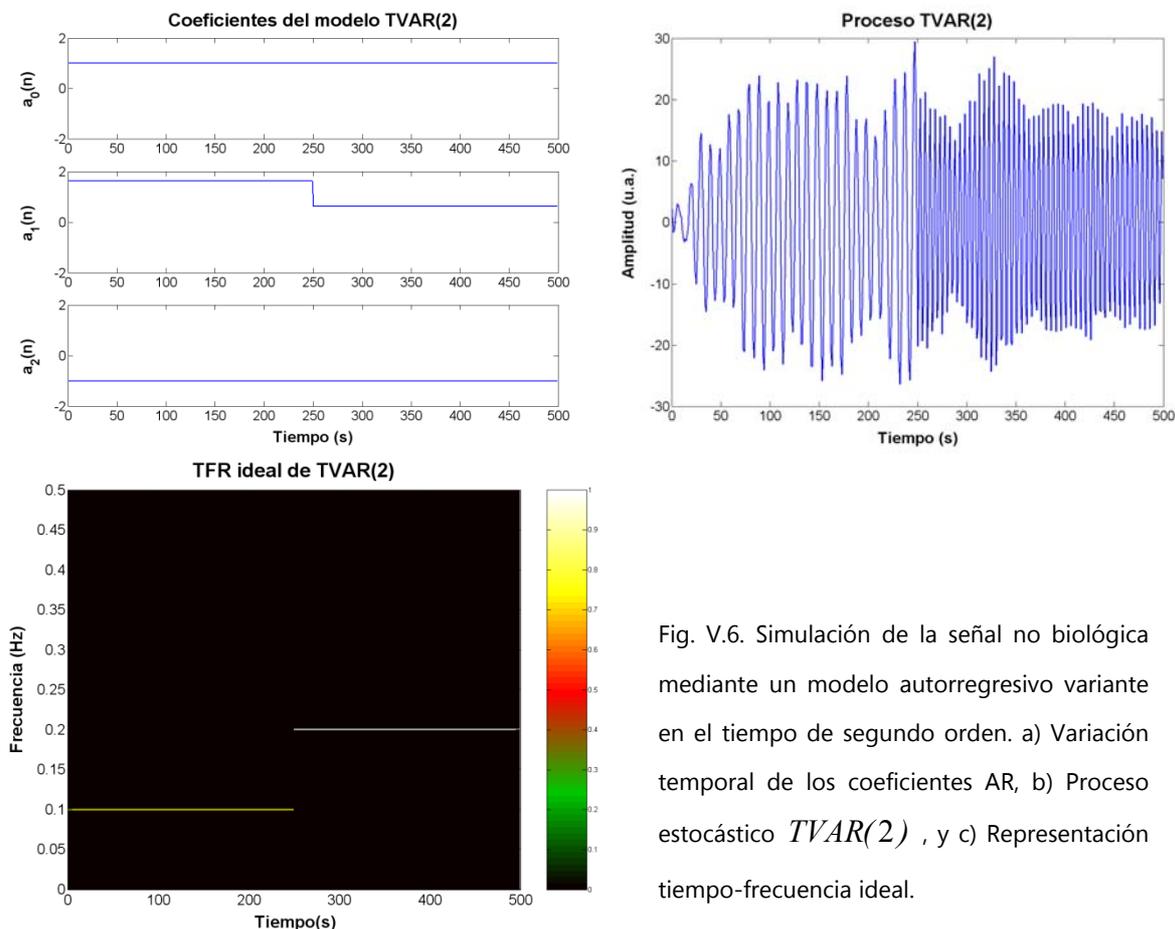


Fig. V.6. Simulación de la señal no biológica mediante un modelo autorregresivo variante en el tiempo de segundo orden. a) Variación temporal de los coeficientes AR, b) Proceso estocástico $TVAR(2)$, y c) Representación tiempo-frecuencia ideal.



V.2.2. Señales de sonidos cardiacos.

De acuerdo a las características de los componentes de los dos primeros sonidos cardiacos, se simuló cada uno de ellos así como su representación tiempo-frecuencia ideal correspondiente. Posteriormente, se generaron los sonidos cardiacos S1 y S2 así como su TFR ideal. A continuación se presenta el procedimiento de simulación para cada sonido cardiaco.

V.2.2.1. Primer sonido cardiaco.

El primer sonido cardiaco, S1, fue simulado de acuerdo a las expresiones propuestas por Chen et. al. [9], quienes generaron una señal S1 sintética basada en la hipótesis de que el sonido registrado sobre el tórax, en el área apical del corazón, está compuesta de vibraciones de frecuencia constante debidas a la válvula mitral y de una vibración de frecuencia modulada debida al miocardio. De acuerdo a la hipótesis anterior, el sonido S1 simulado consiste esencialmente de un componente valvular $s_v(t)$ y un componente miocárdico $s_m(t)$, es decir:

$$S1(t) = s_m(t) + \begin{cases} 0 & , \quad 0 \leq t < t_0 \\ s_v(t - t_0) & , \quad t \geq t_0 \end{cases} \quad (241)$$

donde t_0 representa un retardo entre el tiempo de aparición de ambos componentes, dado que la válvula mitral se cierra posteriormente a la contracción del miocardio. La duración de S1 fue ajustada a 100 ms y el retardo entre componentes a $t_0 = 10$ ms, mientras que la frecuencia de muestreo fue de 5 kHz.

El componente valvular simulado consiste de un sonido transiente de corta duración y amplitud decreciente producido por el cierre y vibración de las válvulas cardiacas. Chen et. al. modelaron $s_v(t)$ mediante un conjunto de sinusoidales con un decaimiento exponencial en amplitud, de acuerdo al modelo de sinusoidales amortiguadas [10-13], es decir:

$$s_v(t) = \sum_{i=1}^N A_i e^{-k_i t} \sin(2\pi f_i t + \phi_i) \quad (242)$$

donde N es el número de sinusoides, A_i es la amplitud, k_i es el factor de amortiguamiento, f_i es la frecuencia y ϕ_i es la fase de la i -ésimo componente armónico. En este modelo, la señal cardiaca valvular se considera como la respuesta del sistema acústico corazón-tórax a las vibraciones armónicas de las válvulas cardiacas, cuyo amortiguamiento es debido a la sangre y la transmisión acústica a través del tórax. En la simulación, los parámetros del componente valvular fueron ajustados en $N = 2$, $f_1 = 50$ Hz, $A_1 = 1$,

$\phi_1 = -\pi$, $k_1 = 60$, $f_2 = 150$ Hz, $A_2 = 0.5$, $\phi_2 = -\pi$, y $k_2 = 60$. En la figura V.7.a, se muestra la morfología del componente valvular resultante. En consecuencia, la TFR ideal de $s_v(t)$ fue creada a partir de las señales de amplitud instantánea, $A_i(t) = A_i e^{-k_i t}$ y frecuencia instantánea $f_i(t) = f_i$ de cada componente armónico, mediante su respectivo mapeo a una imagen. El número de celdas en la frecuencia fue ajustado en $NF = 1024$. La TFR ideal para el componente valvular se muestra en la figura V.7.b.

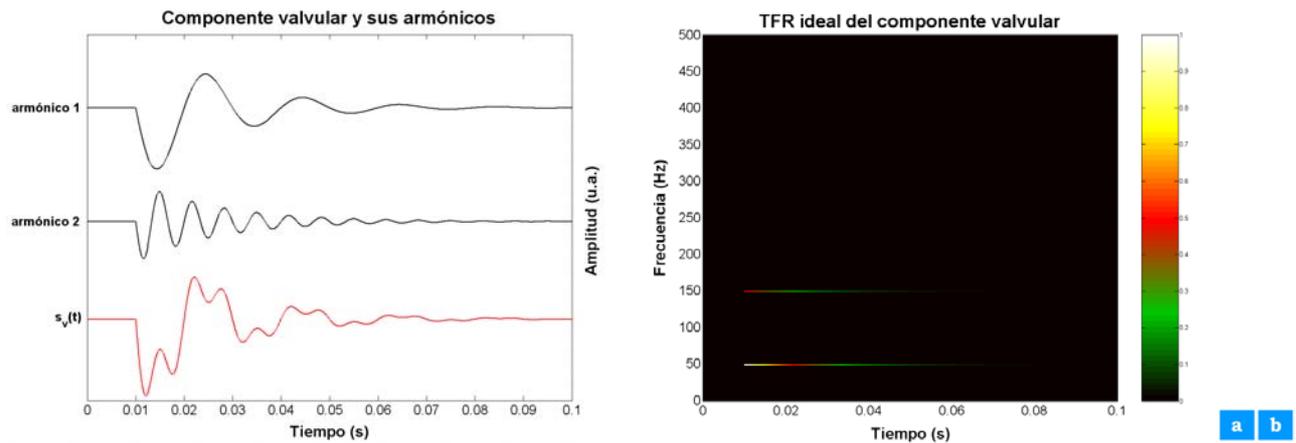


Fig. V.7. Simulación del componente valvular. a) Componentes armónicos y morfología del componente valvular de S1, y b) Representación tiempo-frecuencia ideal.

El componente miocárdico fue simulado mediante una señal determinista con modulación tanto en amplitud como en frecuencia. Chen et. al. simularon este componente mediante un sonido con frecuencia modulada asociado con la tensión generada por el miocardio durante la contracción ventricular izquierda, [14, 15]. Por lo tanto, el componente $s_m(t)$ se representa por la siguiente ecuación:

$$s_m(t) = A_m(t) \sin(2\pi(f_0 + f_m(t)t + \phi_m(t))) \quad (243)$$

donde $A_m(t)$ es la función de modulación en amplitud, f_0 es la frecuencia portadora, $f_m(t)$ es la función de modulación en frecuencia y $\phi_m(t)$ es la función de fase. Los parámetros de modulación en amplitud se establecieron de acuerdo a:

$$A_m(t) = \begin{cases} 0.275(1.1 - 0.9 \cos(83.4\pi t)) & , \quad 0 \leq t \leq 12 \text{ ms} \\ 0.55 & , \quad 12 \text{ ms} < t \leq 30 \text{ ms} \\ 0.275 & , \quad 30 \text{ ms} < t \leq 60 \text{ ms} \\ 0 & , \quad t > 60 \text{ ms} \end{cases} \quad (244)$$

y los parámetros de modulación en la frecuencia mediante:

$$f_m(t) = \begin{cases} -40 \cos(18.3\pi t) & , \quad 0 \leq t \leq 30 \text{ ms} \\ 0 & , \quad 30 \text{ ms} < t \leq 60 \text{ ms} \end{cases} \quad (245)$$

$$f_0 = \begin{cases} 60 \text{ Hz} & , \quad 0 \leq t \leq 30 \text{ ms} \\ 100 \text{ Hz} & , \quad 30 \text{ ms} < t \leq 60 \text{ ms} \end{cases}$$

$$\phi_m(t) = 0 \quad 0 \leq t \leq 60 \text{ ms}$$

En la figura V.8.a se muestra la morfología del componente $s_m(t)$ así como su representación tiempo-frecuencia ideal (ver figura V.8.b), resultante del mapeo de su amplitud y frecuencia instantánea. El número de celdas en la frecuencia nuevamente se ajustó en $NF = 1024$. Finalmente, a partir de las señales de los componentes $s_m(t)$ y $s_v(t)$ se simuló el primer sonido cardiaco de acuerdo a la ecuación (241). En la figura V.9.a se muestra la señal S1 simulada y su TFR ideal resultante del mapeo de las funciones de amplitud y frecuencia instantánea de cada componente. Observe que la frecuencia de la señal se incrementa durante la contracción y permanece constante después de alcanzar una meseta en la fuerza de contracción. Por su parte, la amplitud se incrementa rápidamente al inicio, permanece constante por un instante para finalmente decaer a cero.

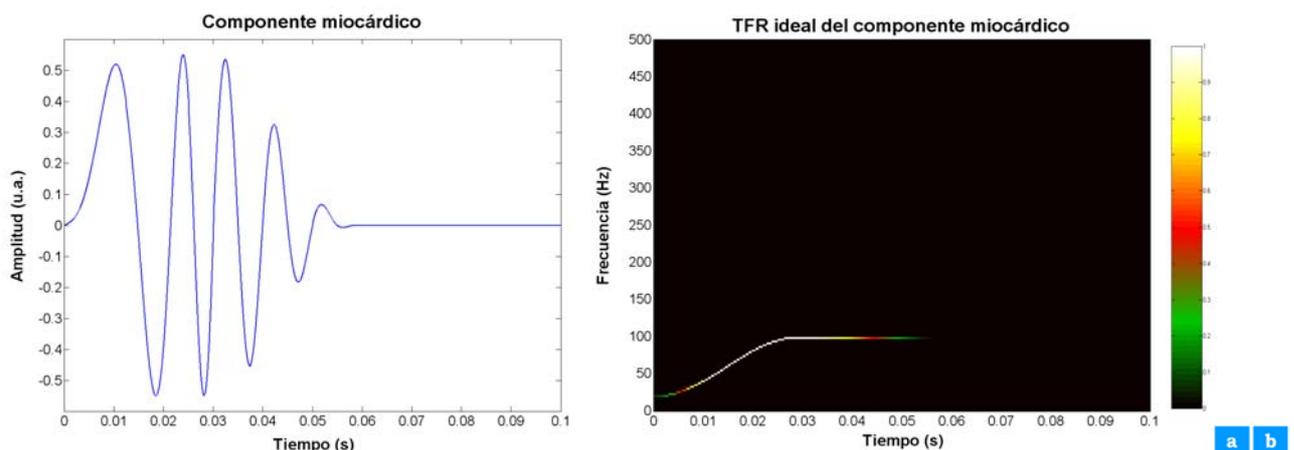


Fig. V.6. Simulación del componente miocárdico. a) Forma de onda de $s_m(t)$ y b) Representación tiempo-frecuencia ideal de $s_m(t)$ simulado.

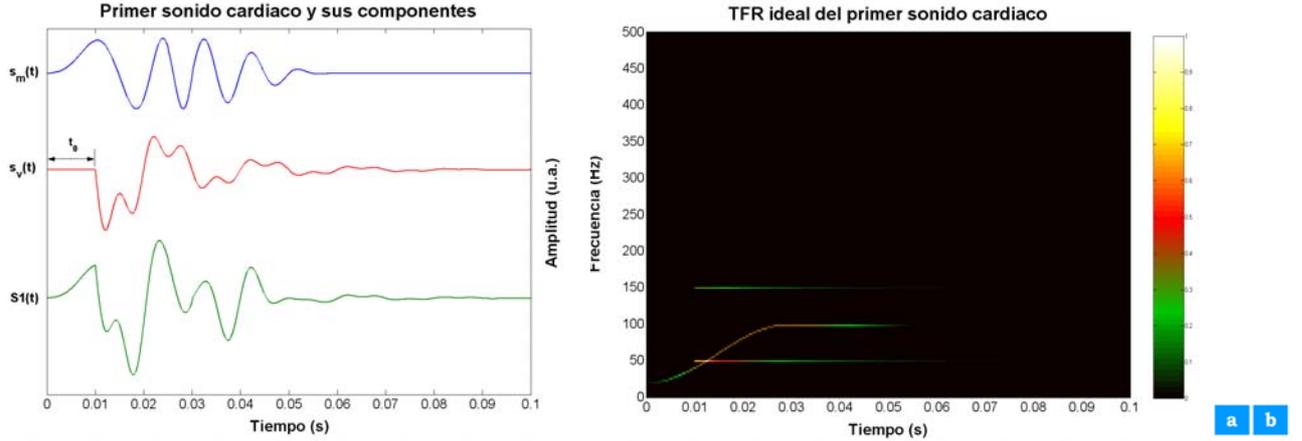


Fig. V.9. Señal simulada S1. a) Componentes y morfología de S1, y b) representación tiempo-frecuencia ideal.

V.2.2.2. Segundo sonido cardiaco.

En esta tesis, el segundo sonido cardiaco S2 fue simulado a través de sus componentes aórtico y pulmonar, A2 y P2, respectivamente, de acuerdo a las expresiones matemáticas propuestas por Xu et al. [16]. Cada uno de los componentes se simula mediante una señal chirp no lineal y se permite un traslape entre los componentes aórtico y pulmonar, es decir:

$$\begin{aligned} S_2(t) &= A2(t) + P2(t - t_0) \\ S_2(t) &= A_A(t) \sin(\varphi_A(t)) + A_P(t - t_0) \sin(\varphi_P(t - t_0)) \end{aligned} \quad (246)$$

donde $A(t)$ y $\varphi(t)$ son las funciones de amplitud y fase instantáneas del componente respectivo, denotado por el subíndice A para el aórtico y P para el pulmonar, mientras que t_0 define el tiempo de retardo entre la aparición de ambos componentes. En nuestra simulación, el tiempo de retardo se estableció en $t_0 = 35 \text{ ms}$ y la duración de cada componente se fijó en 60 ms , con la amplitud y frecuencia instantánea de cada uno de los componentes determinada por:

$$\begin{aligned} IF_A(t) &= 24.3 + 225.7(t+1)^{-\frac{1}{2}} \\ A_A(t) &= 1.5 \left(1 - e^{-\frac{t}{8}}\right) e^{-\frac{t}{16}} \sin\left(\frac{\pi t}{60}\right) \\ IF_P(t) &= 21.83 + 178.17(t+1)^{-\frac{1}{2}} \\ A_P(t) &= \left(1 - e^{-\frac{t}{8}}\right) e^{-\frac{t}{16}} \sin\left(\frac{\pi t}{60}\right) \end{aligned} \quad (247)$$

Observe que cada componente es una señal de corta duración que posee una frecuencia instantánea decreciente. Además, el primer factor de la amplitud instantánea de cada componente es una ponderación de la energía de la señal, el segundo y tercer factor proveen parámetros para controlar el decaimiento de la envolvente en amplitud del sonido simulado, mientras que el cuarto factor es empleado para asegurar que la duración finita de cada componente este limitada a 60 ms . Para poder simular A2 y P2, y por consiguiente S2, utilizando la ecuación (246) sólo resta obtener la fase instantánea de cada uno de ellos, la cual se obtiene a partir de la definición de la frecuencia instantánea dada por Ville mediante la integral de la frecuencia instantánea, es decir:

$$\varphi(t) = \int_{-\infty}^t IF(t) dt \quad (248)$$

En la figura V.10 se muestra la amplitud y frecuencia instantánea de cada componente, incluyendo el desplazamiento en el tiempo del componente pulmonar. En cada caso, la frecuencia instantánea de la señal sigue un comportamiento decreciente no lineal. También se muestra el segundo sonido cardiaco simulado con un traslape de 25 ms entre ambos componentes.

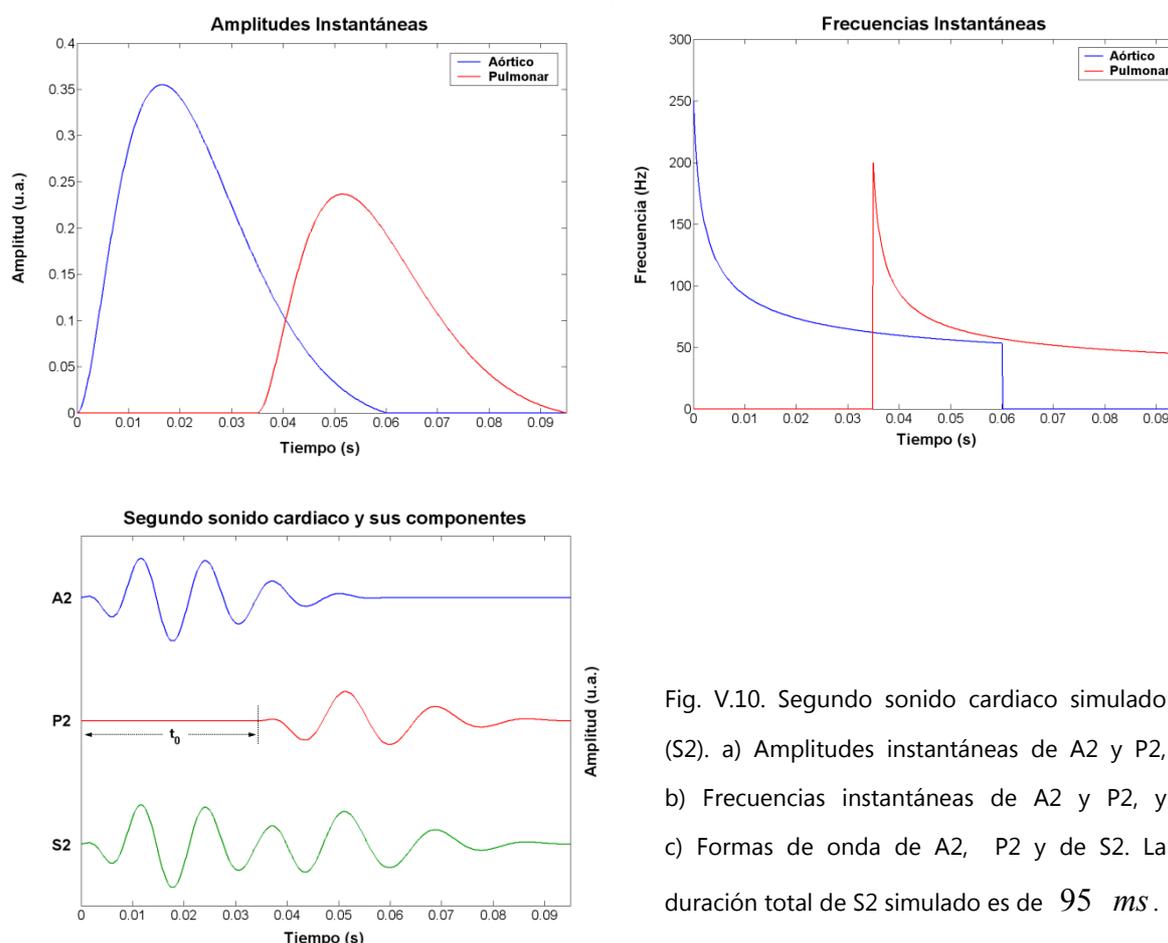


Fig. V.10. Segundo sonido cardiaco simulado (S2). a) Amplitudes instantáneas de A2 y P2, b) Frecuencias instantáneas de A2 y P2, y c) Formas de onda de A2, P2 y de S2. La duración total de S2 simulado es de 95 ms .



Finalmente, la representación tiempo-frecuencia ideal de S2 fue obtenida mediante el mapeo de las funciones de amplitud y frecuencia instantánea de sus dos componentes, empleando $NF = 1024$ celdas en la frecuencia. La frecuencia de muestreo para las simulaciones fue ajustada en 5 kHz. En la figura V.11.a y V.11.b se presenta la TFR ideal de los componentes aórtico y pulmonar de S2 simulado y en la figura V.11.c la TFR ideal para S2.

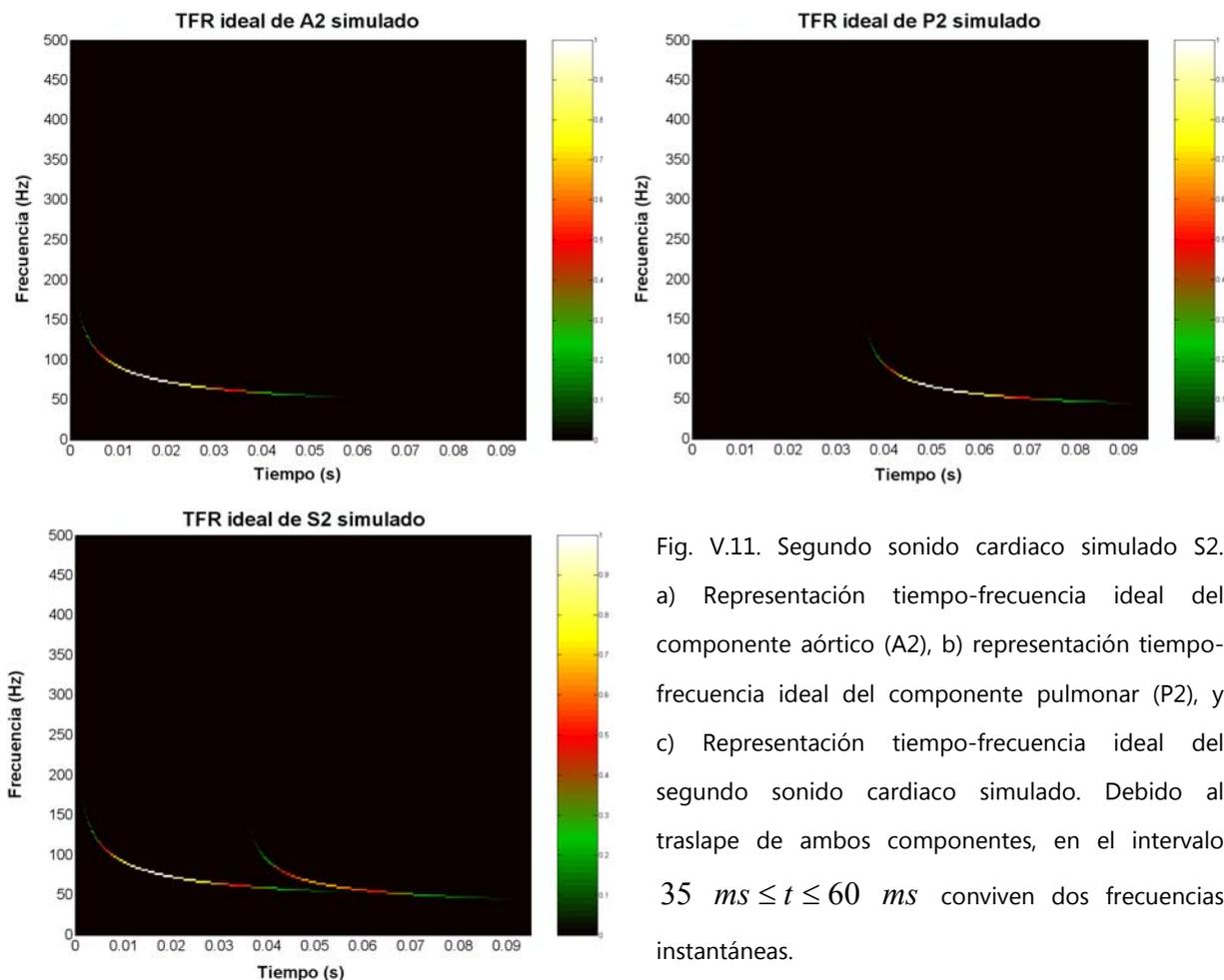


Fig. V.11. Segundo sonido cardiaco simulado S2. a) Representación tiempo-frecuencia ideal del componente aórtico (A2), b) representación tiempo-frecuencia ideal del componente pulmonar (P2), y c) Representación tiempo-frecuencia ideal del segundo sonido cardiaco simulado. Debido al traslape de ambos componentes, en el intervalo $35 \text{ ms} \leq t \leq 60 \text{ ms}$ conviven dos frecuencias instantáneas.



V.2.3. Señales de sonidos respiratorios adventicios.

Ya se ha mencionado que los sonidos respiratorios contienen información de importancia acerca de la fisiología y patologías de los pulmones y vías aéreas. En particular las características de los sonidos adventicios resultan diferentes en problemas pulmonares, además de que sus características acústicas indican no solo la presencia de anomalía en el sistema respiratorio sino también su severidad y localización. A continuación se simulan

señales que corresponden a los dos tipos de sonidos respiratorios adventicios de acuerdo a su duración: continuos y discontinuos.

V.2.3.1 Sonidos adventicios discontinuos.

El sonido respiratorio adventicio discontinuo por excelencia es el estertor crepitante. Los sonidos crepitantes fueron generados de acuerdo a las funciones matemáticas propuestas por Kiyokawa et. al. [17], quienes las diseñaron para producir formas de onda similares a aquellas reportadas en sonidos pulmonares anormales reales. Kiyokawa et. al. definieron la ecuación de una forma de onda que: 1) posee dos ciclos, 2) la localización de su primer intercepto positivo aparece explícitamente como un parámetro de la fórmula, 3) la mayor parte de su potencia está concentrada alrededor de su origen y 4) su morfología resultante tiene una muy buena aproximación con la morfología descrita por los expertos para un crepitante real.

De acuerdo a las características anteriores, los estertores crepitantes finos y gruesos fueron simulados mediante las siguientes ecuaciones:

$$\begin{aligned} y(t) &= m(t)y_0(t) \\ m(t) &= \sin^2(\pi\sqrt{t}) \quad , \quad 0 \leq t \leq 1 \\ y_0(t) &= \sin\left(4\pi \frac{\log(0.25)}{\log(t_0)} t\right) \end{aligned} \quad (249)$$

donde $m(t)$ es una función de modulación en amplitud que desplaza la mayor parte de la potencia de la función oscilatoria $y_0(t)$ al inicio de la forma de onda y donde $y_0(t)$ es una función que posee dos ciclos cuyo primer intercepto positivo se localiza en t_0 . Observe que en las ecuaciones anteriores la duración de la señal se encuentra normalizada entre 0 y 1.

Empleando terminología de sonidos respiratorios, la duración del estertor crepitante corresponde al parámetro de la duración de dos ciclos (2CD), mientras que el primer intercepto positivo corresponde al parámetro de ancho de deflexión inicial (IDW). Los parámetros de los estertores crepitantes simulados fueron ajustados como se muestra en la tabla V.I, de acuerdo a los criterios de Murphy et al. [18]. La frecuencia de muestreo fue ajustada en 10 kHz. En la figura V.12.a se presentan las funciones correspondientes a la modulación en amplitud $m(t)$ y a la función oscilatoria $y_0(t)$ empleadas para simular el estertor crepitante fino, en la figura V.12.b se presentan los crepitantes finos y gruesos, mientras que en la figura V.12.c se muestra su contenido espectral.

TABLA V.I. CARACTERÍSTICAS TEMPORALES DE LOS ESTERTORES CREPITANTES.

Tipo	IDW	2CD
Fino	0.5 ms	5 ms
Grueso	1.2 ms	9 ms

Para el caso de los estertores crepitantes, no se dispone de una representación tiempo-frecuencia ideal que sirva como referencia para la evaluación de las diversas técnicas de análisis. Sin embargo, mediante inspección visual de las señales simuladas se observa que las formas de onda tienden a extenderse conforme el tiempo evoluciona. Lo anterior se puede confirmar al estimar la frecuencia instantánea de $y_0(t)$ así como la del estertor crepitante $y(t) = m(t)y_0(t)$, mediante la derivada de la fase de la señal analítica. En la figura V.13 se muestra el comportamiento de la frecuencia instantánea para el estertor crepitante fino simulado (línea punteada) y su componente oscilatorio (línea continua). Observe que debido a la baja amplitud de las funciones hacia el final del estertor simulado, la estimación presenta gran cantidad de error.

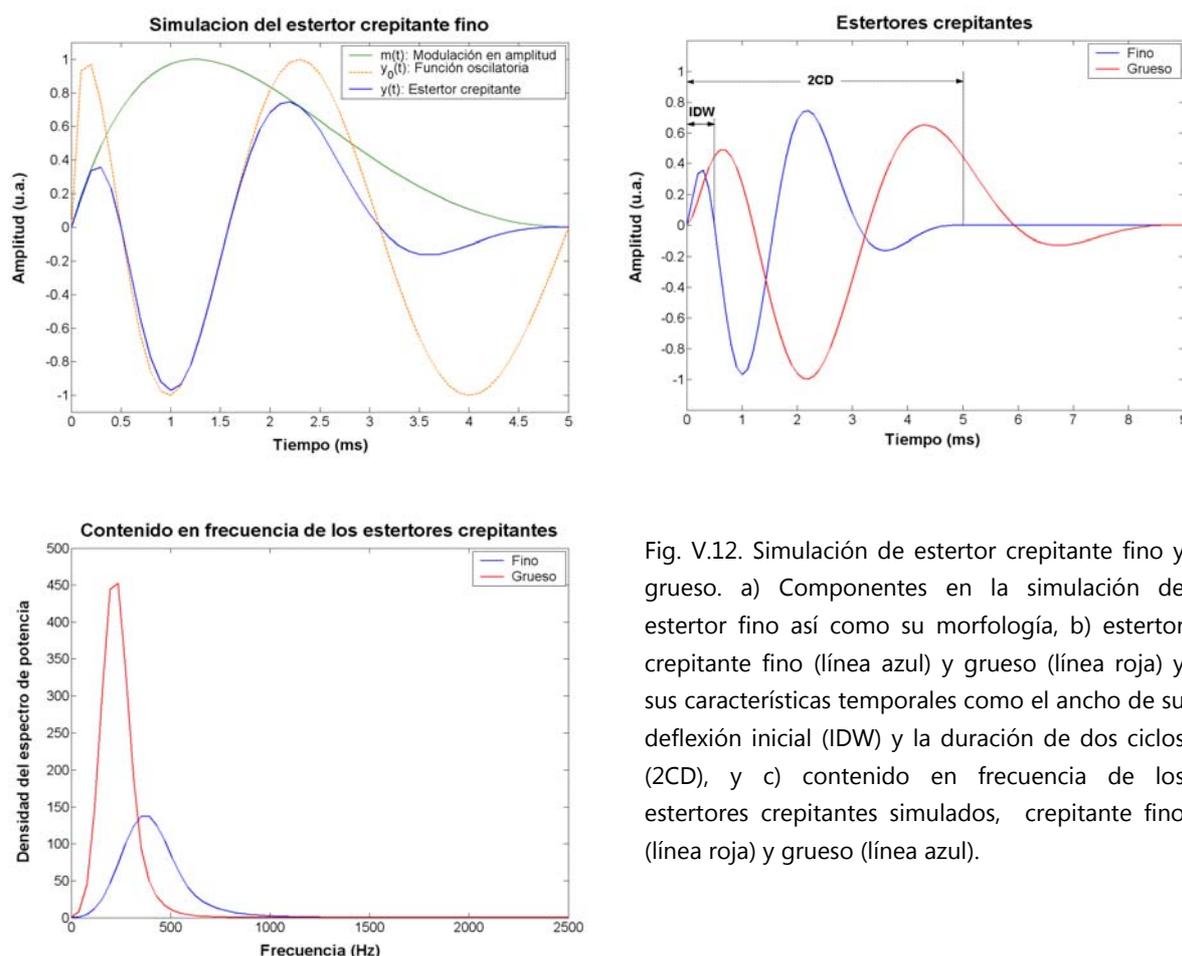


Fig. V.12. Simulación de estertor crepitante fino y grueso. a) Componentes en la simulación de estertor fino así como su morfología, b) estertor crepitante fino (línea azul) y grueso (línea roja) y sus características temporales como el ancho de su deflexión inicial (IDW) y la duración de dos ciclos (2CD), y c) contenido en frecuencia de los estertores crepitantes simulados, crepitante fino (línea roja) y grueso (línea azul).

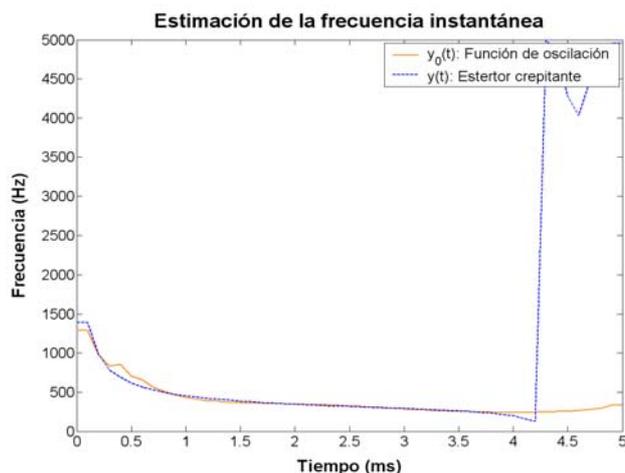


Fig. V.13. Estimación de la frecuencia instantánea mediante la derivada de la fase de la señal analítica del estertor fino simulado (línea punteada) y de su componente oscilatorio $y_0(t)$ (línea continua), observe que hacia el final del estertor simulado debido a la baja amplitud de las funciones la estimación presenta gran cantidad de error.

V.2.3.2. Sonidos adventicios continuos.

En esta tesis, el sonido adventicio continuo que se simuló corresponde a un estertor sibilante. De acuerdo a las definiciones presentadas por la European Respiratory Review Task Force en los lineamientos CORSA [19], la frecuencia dominante de un estertor sibilante es usualmente mayor a 100 Hz con una duración mayor a 100 ms, pudiendo ser monofónico cuando solamente un tono es audible o polifónico cuando múltiples frecuencias se perciben simultáneamente.

La simulación del estertor sibilante se realizó mediante la creación de una forma de onda semejante a las encontradas en los registros de sujetos enfermos que cumplieran con los criterios CORSA antes mencionados. En particular, se simuló un estertor sibilante monofónico con una frecuencia central de 275 Hz con una duración de 130 ms, mediante la modulación en amplitud de una sinusoidal de duración finita, es decir:

$$y(t) = A(t) \sin(2\pi f_0 t) \quad (250)$$

donde $A(t)$ es una función de modulación en amplitud y f_0 es la frecuencia del tono. La función de modulación empleada corresponde a una ventana Hamming. La frecuencia de muestreo fue ajustada en 5 kHz. La señal resultante se muestra en la figura V.14.a.

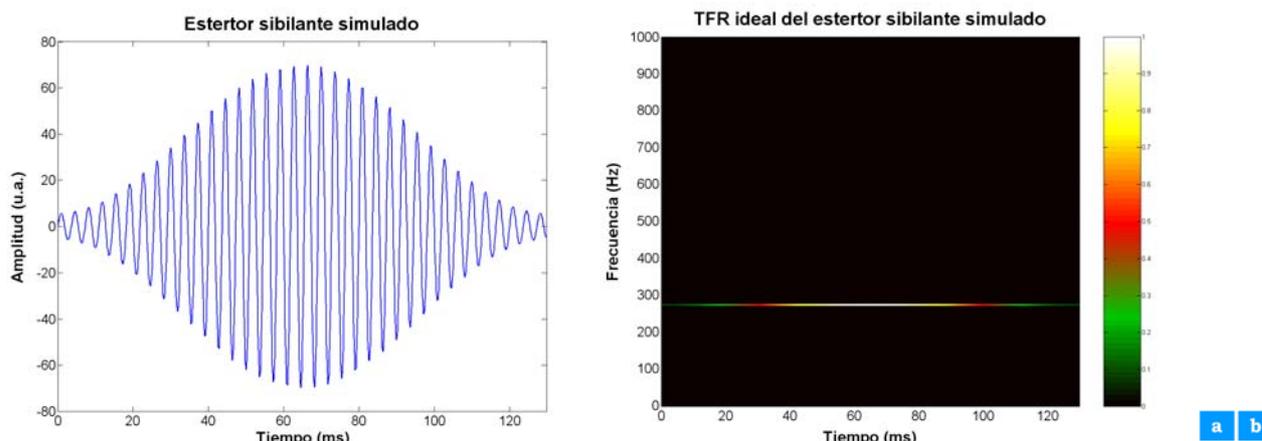


Fig. V.14. Sonido adventicio continuo. a) Estertor sibilante simulado mediante la modulación en amplitud con una ventana Hamming de un tono puro, y b) representación tiempo-frecuencia ideal.

Mediante el mapeo de la amplitud instantánea del estertor sibilante, i.e. magnitud de la ventana Hamming y su frecuencia instantánea, i.e. la frecuencia constante del tono, se generó su representación tiempo-frecuencia ideal. El total de celdas en la frecuencia fue ajustado a $NF = 512$. La TFR ideal resultante se presenta en la figura V.12.b.

V.3. Índices de desempeño.

En la sección anterior se abordó la simulación de señales con diferentes características tiempo-frecuencia y se construyeron sus TFRs ideales. Las representaciones ideales ofrecen una referencia contra la cual comparar el desempeño de los diversos métodos tiempo-frecuencia. Esta sección se enfoca a la definición de los índices de comparación, los cuales proporcionan medidas de desempeño de las técnicas tiempo-frecuencia que se utilizan para determinar la TFR más adecuada para la representación tiempo-frecuencia de los sonidos torácicos.

Los índices empleados para evaluar el desempeño de las TFR estimadas son: (a) la correlación bidimensional central, (b) el error cuadrático medio normalizado y (c) la correlación de la frecuencia instantánea estimada mediante el primer momento en el tiempo de la TFR. A continuación se describen cada uno de los índices mencionados.

V.3.1. La correlación bidimensional central.

El índice de desempeño con base en la correlación bidimensional central se define de la siguiente forma [20]:

$$\rho = \frac{\sum_l \sum_k TFR_d(l,k)TFR_e(l,k)}{\sqrt{\sum_l \sum_k (TFR_d(l,k))^2 \sum_l \sum_k (TFR_e(l,k))^2}} \quad (251)$$

donde $TFR_d(l,k)$ corresponde a la TFR ideal, $TFR_e(l,k)$ corresponde a la TFR estimada por cada uno de los métodos tiempo-frecuencia y las variables l y k representan los índices de tiempo y frecuencia discretos, respectivamente.

En la presente tesis, la aplicación del índice de desempeño de la ecuación (251) se utilizó de dos formas: 1) mediante un enfoque global, i.e. considerando las TFRs como imágenes completas y 2) mediante un enfoque local, i.e. dividiendo cada TFR en sub-imágenes.

V.3.1.1. Enfoque global.

El índice de correlación bidimensional central ha sido empleado para medir la diferencia, en términos de la distribución de energía entre la TFR estimada y su versión ideal. Para la correlación bidimensional central, una TFR con un índice de ρ cercano a la unidad significa que la TFR se aproxima en mejor medida a la TFR ideal; es decir, la técnica de estimación asociada con un valor de $\rho \approx 1$ presenta el mejor desempeño. El índice de desempeño se obtuvo mediante la comparación entre la representación tiempo-frecuencia estimada mediante cada método mencionado en el capítulo IV y la representación tiempo-frecuencia ideal de cada señal.

V.3.1.2. Enfoque local.

Adicionalmente a la aplicación de la correlación bidimensional central entre cada TFR estimada y su versión ideal, cada imagen TFR fue dividida en un arreglo de 5x5 sub-imágenes y el índice ρ fue calculado entre una sub-imagen ideal y la correspondiente sub-imagen estimada. En la figura V.15 se muestra un ejemplo del procedimiento del enfoque local, como puede observarse en las figuras V.15.a y V.15.b las imágenes TFR ideal y estimada se dividen en 25 sub-imágenes, mientras que en la figura V.15.c se muestran los valores del índice de correlación bidimensional para cada sub-imagen, para este ejemplo se observa que la estimación de la representación tiempo-frecuencia no es cercana a la ideal, dado los valores de correlación en las sub-imágenes donde existe información de la representación ideal.

Finalmente, para decidir la robustez de la técnica de estimación se obtuvo un valor promedio de ρ . La TFR con el valor de $\rho_{promedio}$ más cercano a uno representa la TFR con el mejor desempeño. La intención de emplear este enfoque consiste en valorar con mayor resolución, distribución local de la energía, el desempeño de las TFRs.

Una situación que es importante considerar al determinar el índice de correlación bidimensional central es que pueden existir pares de sub-imágenes que contengan únicamente valores nulos lo cual llevaría a una indeterminación de ρ . Para evitar este conflicto, antes de realizar la segmentación en sub-imágenes, se añade una constante positiva pequeña a ambas imágenes TFRs, i.e. a la ideal y a la estimada. Finalmente, debido a las elevadas frecuencias de muestreo empleadas en comparación con el contenido en frecuencia de las señales, el tamaño de las representaciones tiempo-frecuencia fue recortado en el eje de frecuencia, para enfocarse al área que contiene información relevante.

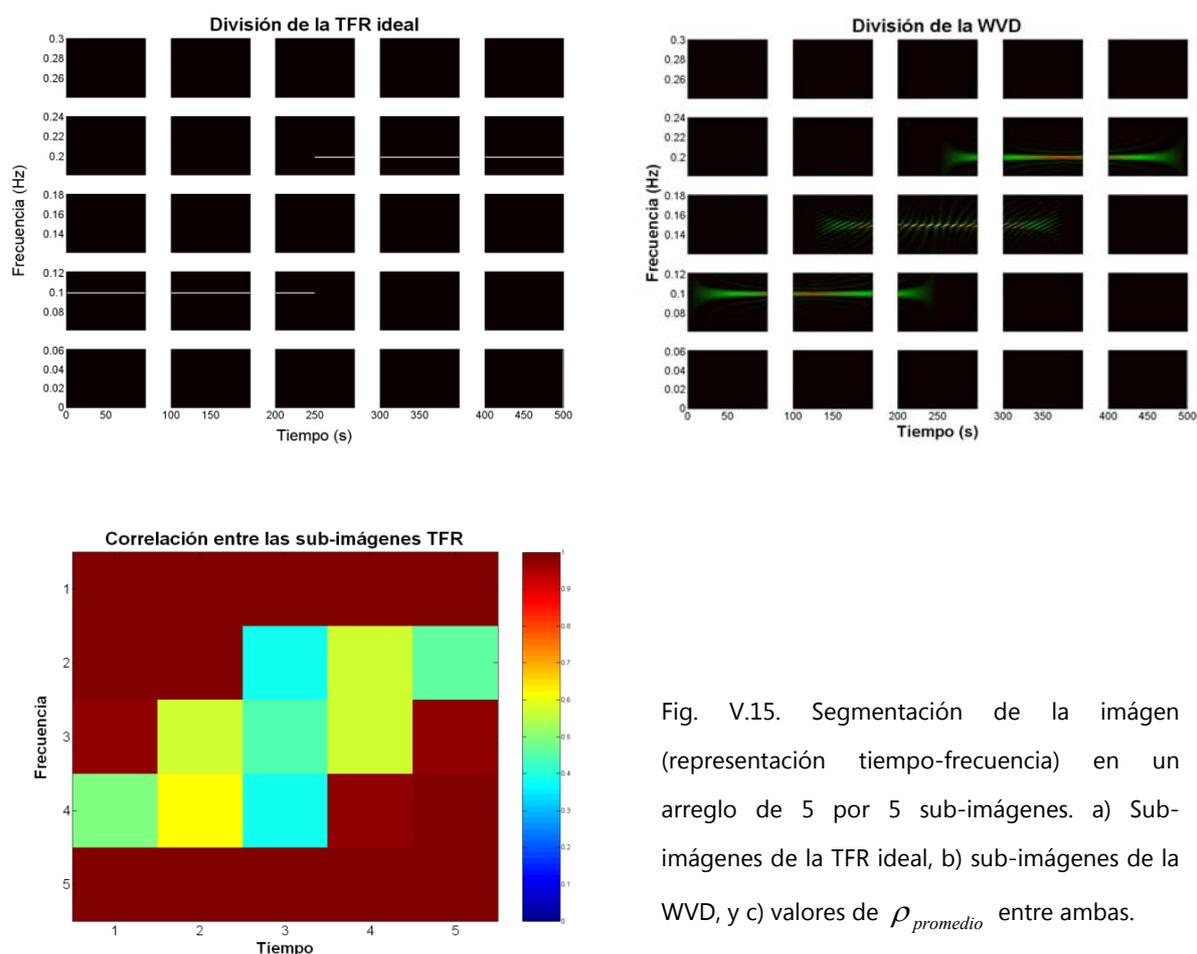


Fig. V.15. Segmentación de la imagen (representación tiempo-frecuencia) en un arreglo de 5 por 5 sub-imágenes. a) Sub-imágenes de la TFR ideal, b) sub-imágenes de la WVD, y c) valores de $\rho_{promedio}$ entre ambas.



V.3.2. Estimación de la frecuencia instantánea mediante el primer momento temporal de la TFR.

La función de frecuencia instantánea también fue empleada para valorar el desempeño de los diversos métodos de estimación de la representación tiempo-frecuencia. La razón de su empleo se debe en gran medida a que ha sido demostrado que la frecuencia instantánea proporciona información fisiológica relevante, por ejemplo, la frecuencia instantánea es proporcional a la contracción muscular en la generación del primer sonido cardiaco, además, es proporcional a la presión arterial generada por las arterias pulmonar y aórtica en la generación de los componentes del segundo sonido cardiaco. Considerando que para los sonidos respiratorios no existe un análisis relacionado con su representación en el dominio tiempo-frecuencia y dada la importancia teórica y práctica del comportamiento de la frecuencia instantánea, en la presente tesis se utiliza para generar dos índices para evaluar de mejor forma las técnicas de estimación de TFR, ya que la frecuencia instantánea pudiera caracterizar adecuadamente a los sonidos adventicios.

La estimación de las frecuencias instantáneas se realizó empleando los momentos de las TFRs [21]. Al considerar el primer momento de la TFR con respecto a la frecuencia, primer momento temporal, puede obtenerse un estimado de la frecuencia instantánea de la señal $x(t)$, de la siguiente forma:

$$FI(t) = \frac{\int_{-\infty}^{\infty} fTFR(t, f)df}{\text{var}(x(t))} \quad (252)$$

que en general resulta en una versión suavizada de la frecuencia instantánea teórica [22]. En la figura V.14 se muestran las frecuencias instantáneas para el segundo sonido cardiaco simulado S2 y estimadas a partir de la TFR ideal y de un espectrograma, utilizando la ecuación (252). Los dos índices de desempeño basados en la frecuencia instantánea estimada consisten en: (a) el error relativo obtenido mediante el error cuadrático medio normalizado y (b) el coeficiente de correlación cruzada; ambos índices se calculan entre las TFRs estimadas e ideales. Los índices se basan únicamente en la estimación de la frecuencia instantánea de las señales, a diferencia de los índices de correlación bidimensional central que además contemplan la distribución en tiempo y frecuencia de la amplitud alrededor de dicha frecuencia instantánea.

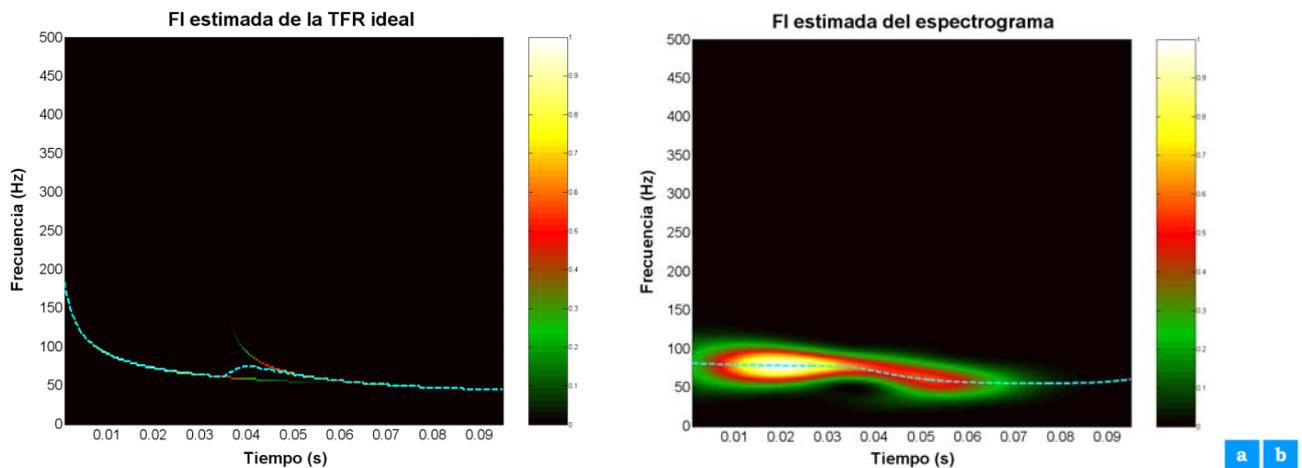


Fig. V.16. Estimación de la frecuencia instantánea (línea discontinua) mediante el primer momento temporal de las TFRs. a) Utilizando la representación tiempo-frecuencia ideal. b) Utilizando la representación tiempo-frecuencia generada por el Espectrograma. Observe que cuando existen dos frecuencias instantáneas al mismo tiempo, la estimación resulta en el centro de gravedad de ambas.

V.3.2.1. Error relativo.

Este índice se define mediante el error cuadrático medio normalizado [23], es decir:

$$NRMSE = \sqrt{\frac{\sum_l (FI_d(l) - FI_e(l))^2}{\sum_l (FI_d(l))^2}} \quad (253)$$

donde $FI_d(l)$ y $FI_e(l)$ corresponden a la frecuencia instantánea estimada de la TFR ideal y la TFR estimada, respectivamente. En la ecuación (254), l representa el índice de tiempo discreto. Como puede deducirse de la ecuación (254), la TFR con el menor valor de $NRMSE$ se asocia con el mejor desempeño.

V.3.2.2. Coeficiente de correlación cruzada.

Este índice está definido mediante [23]:

$$\rho_{FI} = \frac{\sum_l FI_d(l)FI_e(l)}{\sqrt{\sum_l (FI_d(l))^2 \sum_l (FI_e(l))^2}} \quad (254)$$

donde $FI_d(l)$ y $FI_e(l)$ corresponden a la frecuencia instantánea estimada de la TFR ideal y la TFR estimada, respectivamente y l representa el índice de tiempo discreto. Para este índice, la TFR con el valor de ρ_{FI} más cercano a uno se asocia con el mejor desempeño.

V.3.3. Resolución tiempo-frecuencia.

El último índice de desempeño utilizado en la presente tesis, es una medida de la resolución en el dominio tiempo-frecuencia. Este índice se emplea para medir la concentración de la energía o resolución de la representación tiempo-frecuencia estimada. El índice de resolución tiempo-frecuencia se define como [20]:

$$res_{TF} = \frac{\sum_l B(l)}{N} \cdot \frac{\sum_k T(k)}{N} \quad (255)$$

y

$$B(l) = \frac{\sum_k TFR_e(l,k)}{\max_k \{TFR_e(l,k)\}} \quad , \quad T(k) = \frac{\sum_l TFR_e(l,k)}{\max_l \{TFR_e(l,k)\}} \quad (256)$$

y $TFR_e(l,k)$ corresponde a la TFR estimada por cada uno de los métodos tiempo-frecuencia, N representa el número de muestra en el tiempo y la frecuencia, y l , k representan los índices de tiempo y frecuencia discretos, respectivamente. Para este caso, la TFR con el valor de res_{TF} más próximo a cero implica el mejor desempeño.

V.4. Parámetros de las representaciones tiempo-frecuencia.

El cómputo de una representación tiempo-frecuencia mediante cada técnica depende de diferentes parámetros tales como el tipo y longitud de la ventana, o el orden del modelo. Como consecuencia, el desempeño de cada técnica tiempo-frecuencia dependerá de la adecuada selección de dichos parámetros.

Considérese el ejemplo del espectrograma. Para esta técnica, un parámetro importante es la duración de la ventana empleada para aislar la porción de la señal en un instante de tiempo dado. Si la ventana es demasiado corta, la resolución temporal será precisa pero la resolución en la frecuencia resultará pobre. En contraparte, si la duración de la ventana es demasiado larga, la resolución en la frecuencia será alta pero la resolución en el tiempo será pobre. Entonces, debe encontrarse un balance entre la resolución en el tiempo y en la frecuencia mediante la determinación de su ventana óptima. Este compromiso no es trivial ya que requiere conocimiento *a priori* de la distribución tiempo-frecuencia teórica de la señal simulada. Sin embargo, en nuestro caso se dispone de la distribución tiempo-frecuencia teórica de los sonidos torácicos simulados. Así, la mejor representación tiempo-frecuencia de cada técnica se obtuvo mediante la variación de sus parámetros y la posterior selección de aquella con el mayor índice de correlación con respecto a la distribución tiempo-frecuencia teórica del sonido simulado.

A continuación se describen los conjuntos de parámetros empleados para calcular las representaciones tiempo-frecuencia mediante cada una de las técnicas mencionadas en el capítulo anterior.

- **El espectrograma, SP.**

Tipo de ventana: Rectangular, Hamming y Blackman-Harris.

Longitud de la ventana: 9 valores equidistantes en el intervalo $[N/10 - N/2]$, donde N es la duración de la señal^{***}. Las longitudes se redondearon al valor impar más cercano.

- **La distribución Choi-Williams, CWD.**

Ancho de la función kernel (σ): 27 valores equidistantes en el intervalo $[1 - \sigma_{\max}]$, donde σ_{\max} correspondiente a la CWD con una correlación ≥ 0.85 con la WVD.

- **El espectrograma reasignado, SPR.**

Tipo de ventana: Rectangular, Hamming y Blackman-Harris.

Longitud de la ventana: 9 valores equidistantes en el intervalo $[N/10, N/2]$, donde N es la duración de la señal^{***}. Las longitudes se redondearon al valor impar más cercano.

- **Modelado TVAR con la técnica de Burg, P_{Burg} .**

Orden del modelo TVAR: $[2 - 5]^*$.

Tipo de ventana: Hamming y Blackman-Harris.

Longitud de la ventana: 9 valores equidistantes en el intervalo $[N/10, N/2]$, donde N es la duración de la señal^{***}. Las longitudes se redondearon al valor impar más cercano.

- **Modelado TVAR con el algoritmo RLS, P_{RLS} .**

Orden del modelo TVAR: $[2 - 5]^*$.

Factor de olvido: 9 valores equidistantes en el intervalo $[0.75 - 1]^{**}$.

^{***} Para el caso de los sonidos cardíacos, se consideraron longitudes de ventana con 5 ms de diferencia.

^{*} Para el caso de las dos señales no biológicas el orden se estableció en 2.

^{**} Para el caso de las dos señales no biológicas se consideraron 18 valores en dicho intervalo.

- **Modelado TVAR con el algoritmo RLS-VFF, $P_{\text{RLS-VFF}}$.**

Orden del modelo TVAR: $[2 - 5]^*$.

Varianza del ruido esperado (σ_e^2): 9 valores logarítmicamente equidistantes en un intervalo dependiente de la señal.

- **Modelado TVAR con el filtro de Kalman, P_{Kalman} .**

Orden del modelo TVAR: $[2 - 5]^*$.

Varianza del ruido de medición (σ^2): Se estableció en 0.1.

Varianza del ruido del proceso (q): 9 valores logarítmicamente equidistantes en un intervalo dependiente de la señal.

- **El escalograma, SC.**

Wavelet madre: Mortlet compleja.

Longitud de la wavelet en la escala más gruesa: 9 valores en diversos intervalos dependiendo de la señal.

Las dos técnicas restantes, i.e. la **distribución Wigner-Ville (WVD)** y el **espectro de Hilbert-Huang (HHS)**, no requirieron del ajuste de parámetros. En particular el espectro de Hilbert-Huang es una técnica adaptable que no necesita la definición de cierta función kernel y que se basa en los datos. Sin embargo, en nuestro caso para la conformación de la representación tiempo-frecuencia se seleccionaron aquellos modos intrínsecos de oscilación que contuvieran la mayor correlación y cantidad de energía de la señal acústica original.

V.5. Adecuación de las representaciones tiempo-frecuencia.

Para la comparación entre las diversas representaciones tiempo-frecuencia, es necesario contar con imágenes de dimensiones adecuadas, i.e. de las mismas dimensiones de la representación tiempo-frecuencia teórica. Estas dimensiones corresponden al tamaño de la matriz empleada al comparar de manera global las representaciones. Sin embargo, posteriormente las dimensiones fueron reducidas antes de realizar la comparación por sub-imágenes debido a las elevadas frecuencias de muestreo empleadas y a la baja ubicación del

* Para el caso de las dos señales no biológicas el orden se estableció en 2.

contenido en frecuencia de las señales. En la tabla V.2 se presentan las dimensiones originales y recortadas de las distribuciones ideales.

TABLA V.2. DIMENSIONES DE LAS REPRESENTACIONES TIEMPO-FRECUENCIA IDEALES.

Señal de sonido	Tamaño original	Tamaño reducido
Señales no biológicas	512x500	308x500
Primer sonido cardiaco (S1)	1024x501	123x501
Componente muscular (sm)	1024x501	123x501
Componente valvular (sv)	1024x501	123x501
Segundo sonido cardiaco (S2)	1024x476	123x476
Componente aórtico (A2)	1024x476	123x476
Componente valvular (P2)	1024x476	123x476
Estertor sibilante	512x650	205x650
Estertor crepitante fino	512x51	-
Estertor crepitante grueso	512x91	-

Con el propósito de comparar las diversas representaciones tiempo-frecuencia, así como para la supresión de valores negativos que carecen de significado físico, e.g. en la distribución de Wigner, cada representación, incluyendo a la ideal, fue normalizada con respecto a sus valores máximo y mínimo para producir una imagen con una intensidad en el intervalo $[0 - 1]$, mediante:

$$RTF_{norm}(t, f) = \frac{RTF(t, f) - \min(RTF(t, f))}{\max(RTF(t, f)) - \min(RTF(t, f))} \quad (257)$$

Finalmente, con el propósito de visualizar adecuadamente el contenido de las distribuciones tiempo-frecuencia, i.e. las intensidades relativas de los componentes, se diseñó una paleta de colores adecuada. En la figura V.17 se presenta la paleta de colores diseñada. Por ejemplo, esta paleta de colores permitió visualizar en mejor medida los términos cruzados del espectrograma cuando estos estaban presentes y difícilmente se distinguían con otras paletas.

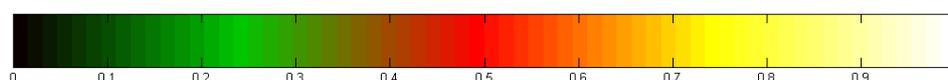


Fig. V.17. paleta de colores diseñada para la visualización de las representaciones tiempo-frecuencia.

V.6. Señales de sonidos cardiacos y respiratorios adquiridas.

La técnica de análisis tiempo-frecuencia seleccionada con ayuda de la señal simulada y su representación tiempo-frecuencia teórica se aplicó a sonidos reales adquiridos. A continuación se describe la etapa de adquisición de los diversos sonidos torácicos, así como la selección de los sonidos torácicos reales.

V.6.1. Sonidos respiratorios.

Las señales de sonidos respiratorios normales y adventicias se obtuvieron de dos bases de datos conformadas por el Laboratorio de Investigación en Procesamiento de Señales e Imágenes Biomédicas (LIPSIB) de la Universidad Autónoma Metropolitana Unidad Iztapalapa (UAM-I). Ambas bases de datos comparten el equipo y programas de adquisición empleados.

- **Equipo y programa de adquisición.**

Las señales fueron adquiridas y almacenadas empleando una computadora personal con un procesador AMD Duron @ 1GHz de velocidad y 250 MB en RAM. El convertidor A/D empleado fue el PCI-6071E de National Instruments con 12 bits de resolución. La frecuencia de muestreo fue ajustada en 10 kHz.

Se adquirieron 26 canales de sonidos respiratorios, de los cuales 25 corresponden a sonidos pulmonares y el restante es un sonido traqueal. Se emplearon los micrófonos subminiatura BT-1834 Knowles® acoplados por aire, de tipo electret y con una respuesta en frecuencia plana de 50 a 3000 Hz. Se adquirió un registro adicional de ruido de base a campo abierto para calibrar las señales de sonidos respiratorios.

Adicionalmente se registró simultáneamente la señal de flujo respiratorio mediante un pneumotacógrafo tipo Fleish acoplado a un transductor de presión Validyne®. La calibración se realizó mediante un rotámetro.

El programa que permitió la adquisición y visualización en tiempo real de las señales adquiridas fue desarrollado en el LIPSIB en la plataforma LabVIEW®. Este programa también permite la exportación de los archivos binarios a un formato de texto para su posterior procesamiento con el programa Matlab®.

- **Arreglo de micrófonos.**

Los micrófonos se distribuyeron en un arreglo matricial de 5 x 5. La primera fila de micrófonos fue colocada 6 cm por debajo de la apófisis espinosa de la séptima vértebra cervical. La nomenclatura de los micrófonos incluye tres letras: cara, i.e. anterior (A) o posterior (P); lado del tórax, i.e. izquierdo (L, left) o derecho (R, right); y línea vertical, i.e. media (M), clavicular (C) o axilar (X). Finalmente, la nomenclatura se completa con un número que indica la altura de la fila. En la figura V.18 se presenta el arreglo de micrófonos junto con su nomenclatura.

- **Maniobra.**

La maniobra realizada por cada sujeto tuvo una duración de 15 segundos, divididos en aproximadamente 5 segundos de apnea inspiratoria, 5 segundos de respiración normal a un flujo máximo establecido, y 5 segundos de apnea espiratoria. La maniobra fue dirigida por un médico con ayuda del despliegue de la señal de flujo en un osciloscopio digital Tektronix®.

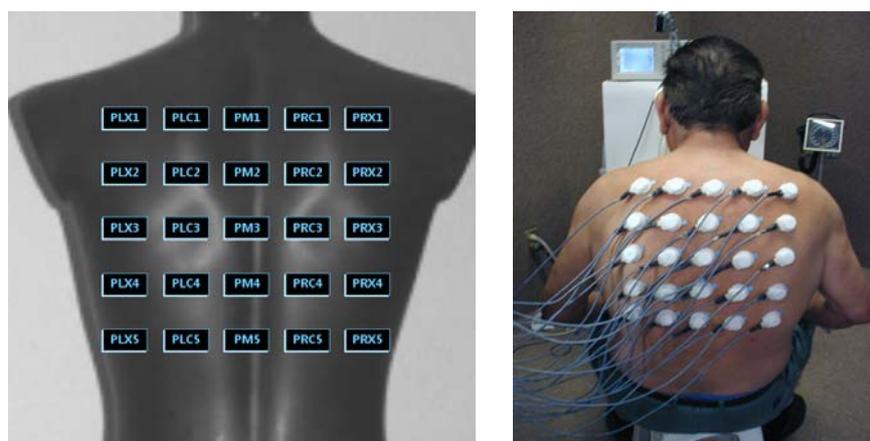


Fig. V.18. Ubicación y nomenclatura del arreglo de micrófonos sobre el tórax posterior.

V.6.1.1. Sonidos respiratorios adventicios.

Esta base de datos está conformada por registros de pacientes con neumoopatías intersticiales difusas (NID). La adquisición se llevó a cabo en el Instituto Nacional de Enfermedades Respiratorias (INER), donde neumólogos realizaron el diagnóstico de los pacientes en base a sus características clínicas y funcionales. Adicionalmente a las señales de sonidos respiratorios y el flujo, se adquirió la señal de ECG mediante una derivación bipolar.

La maniobra fue realizada con los sujetos sentados. El flujo inspiratorio máximo durante la maniobra fue de aproximadamente 1.5 l/s.

En la figura V.19 se muestra la adquisición de un registro de sonidos respiratorios de un paciente con sonidos respiratorios adventicios. Por su parte, en la figura V.20 se muestra el registro multicanal con los 25 sonidos pulmonares adquiridos y la señal de flujo respiratorio.



Fig. V.19. Adquisición multicanal de un sujeto con neumatías intersticiales difusas, donde se observa el arreglo de micrófonos sobre el tórax posterior, el pneumotacógrafo y el electrocardiografo.

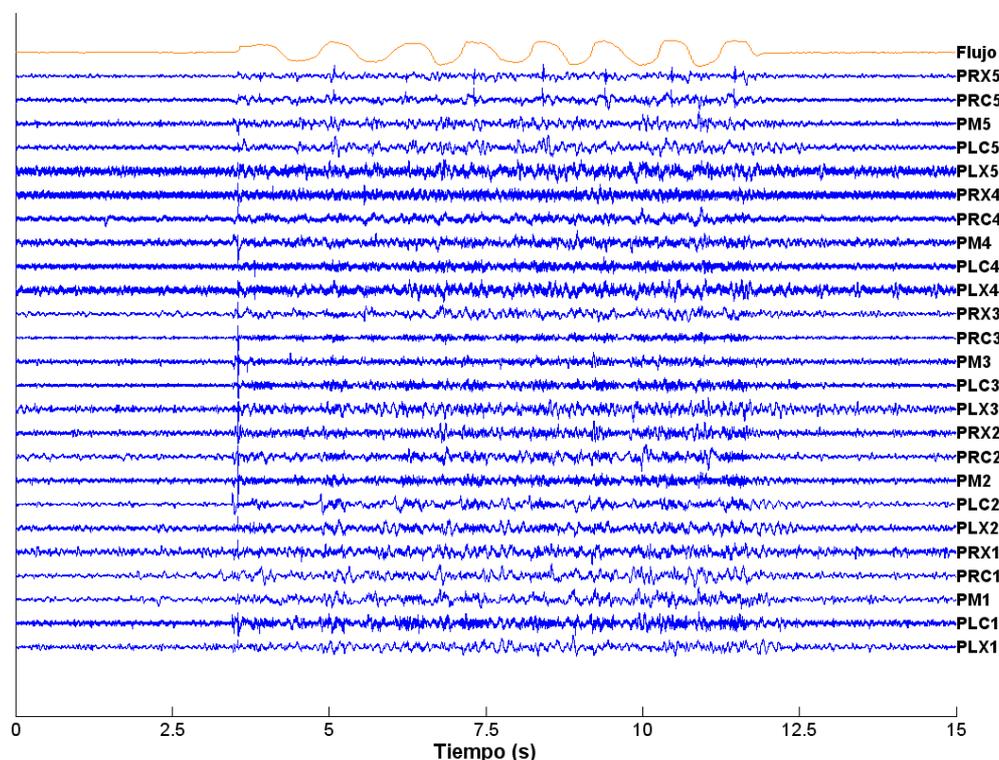


Fig. V.20. Registro multicanal de sonidos respiratorios de un paciente con neumatías intersticiales difusas. Se presentan las señales adquiridas sin filtrar de los 25 canales de sonidos pulmonares, así como la señal de flujo respiratorio.

Previo a la selección manual de las señales de sonidos respiratorios adventicios, las señales de sonido fueron filtradas con un filtro FIR pasabanda con un ancho de banda de 75 a 1000 Hz. El orden del filtro FIR fue de 201, y la ventana empleada para minimizar el fenómeno de Gibbs fue la Hamming. En la figura V.21 se muestra la magnitud de la respuesta en frecuencia del filtro FIR diseñado. En la figura V.22 se muestran los sonidos respiratorios filtrados.

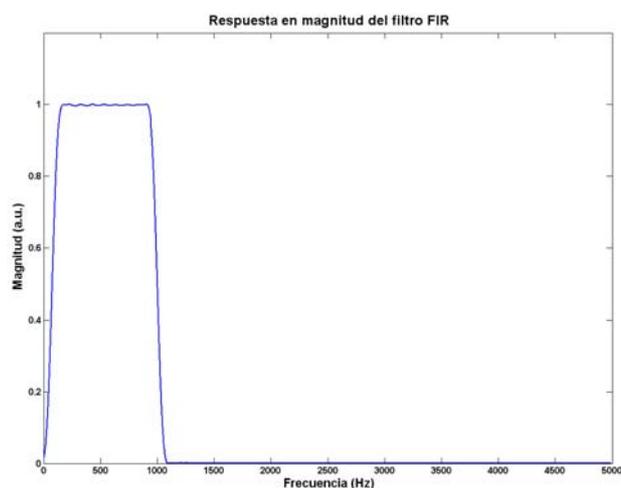


Fig. V.21. Filtro FIR pasabanda diseñado para filtrar los sonidos respiratorios adventicios.

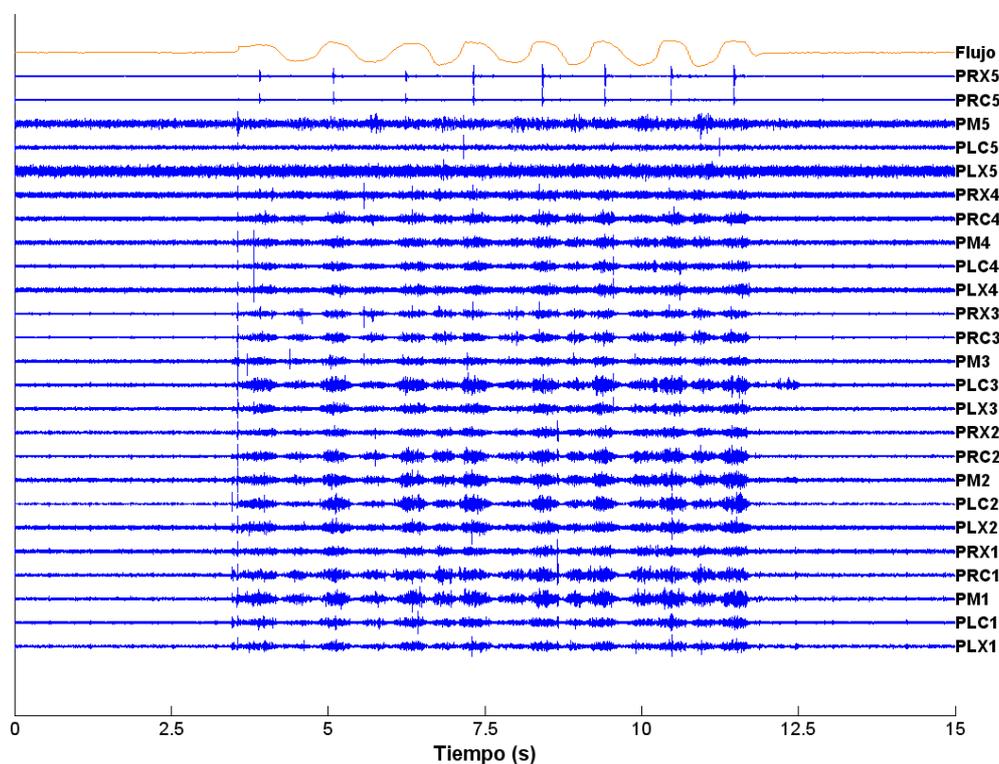


Fig. V.22. Registro multicanal de sonidos respiratorios filtrados de un paciente con neumoopatías intersticiales difusas.

A partir de los registros multicanal de sonidos pulmonares filtrados, se seleccionaron manualmente sonidos respiratorios adventicios tanto continuos como discontinuos. La selección se realizó analizando canal por canal de los diversos registros de sujetos con NID, y extrayendo segmentos que incluyeran formas de onda que concordaran con los criterios definidos en el capítulo de antecedentes. En la figura V.23 se presenta un segmento extraído de los registros de sonido, que incluye a un estertor crepitante real. Mientras que en la figura V.24 la forma de onda de un estertor sibilante real.

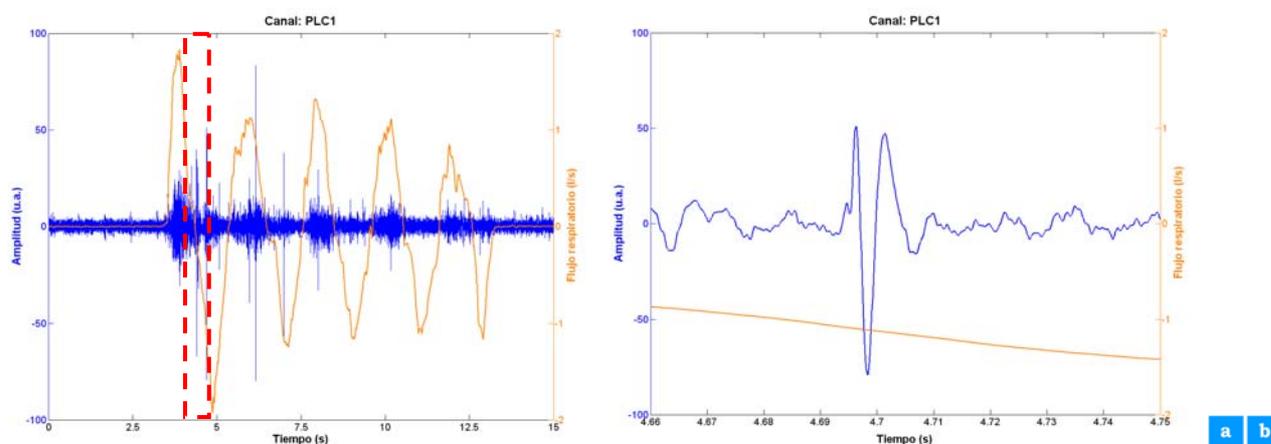


Fig. V.23. Sonido respiratorio adventicio discontinuo real. a) Canal de sonido respiratorio con sonidos agregados. El recuadro en líneas discontinuas indica el segmento extraído. b) Segmento de sonido respiratorio que incluye un sonido adventicio discontinuo.

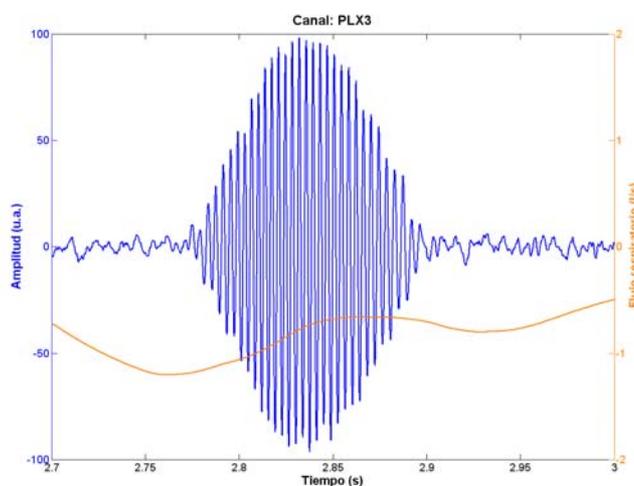


Fig. V.24. Segmento de sonido respiratorio que incluye un sonido adventicio continuo real.

V.6.1.2. Sonidos respiratorios normales.

Esta base de datos esta conformada por registros de sujetos sanos y no fumadores respirando a diversos flujos respiratorios. La adquisición se realizó en una cámara subamortiguada en la (UAM-I). La maniobra fue realizada con los sujetos de pie, con las manos en la nuca y los codos ligeramente hacia delante para minimizar la interferencia de las escápulas. Los registros se realizaron a tres diferentes flujos inspiratorios máximos de 1, 1.5, o 2 l/s cada uno.

Previo a su análisis tiempo-frecuencia, las señales de sonidos respiratorios fueron filtradas con un filtro FIR pasabanda de orden 201 con frecuencias de corte de 75 Hz y 2000 Hz y diseñado con una ventana Hamming. Además, las señales de sonidos respiratorios fueron calibradas empleando el canal de ruido de base. En la figura V.25 se presenta el registro multicanal de sonidos respiratorios filtrado y calibrado.

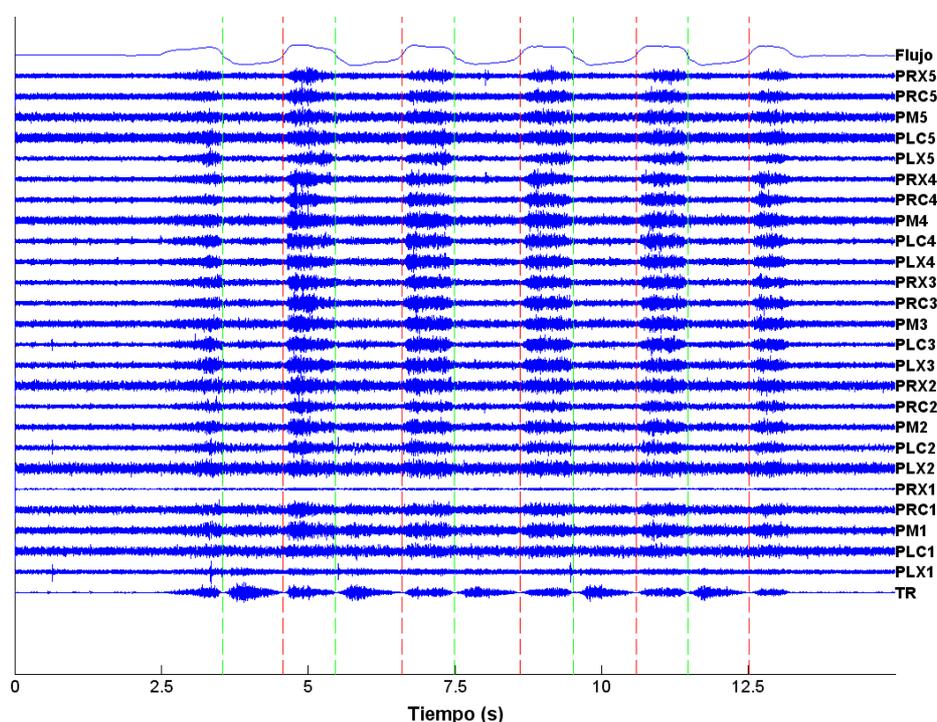


Fig. V.25. Registro multicanal de sonidos respiratorios normales. Señales filtradas y calibradas. Las líneas verticales delimitan las fases del ciclo respiratorio.

A partir del registro multicanal de sonidos respiratorios, se seleccionó un ciclo respiratorio del canal adquirido sobre la traquea y del canal PLC3. Esto permitirá la comparación de las características entre el sonido pulmonar y el traqueal, y las fases inspiratoria y espiratoria. En la figura V.26 se muestran las señales de sonidos respiratorios reales extraídas del registro para su análisis tiempo-frecuencia.

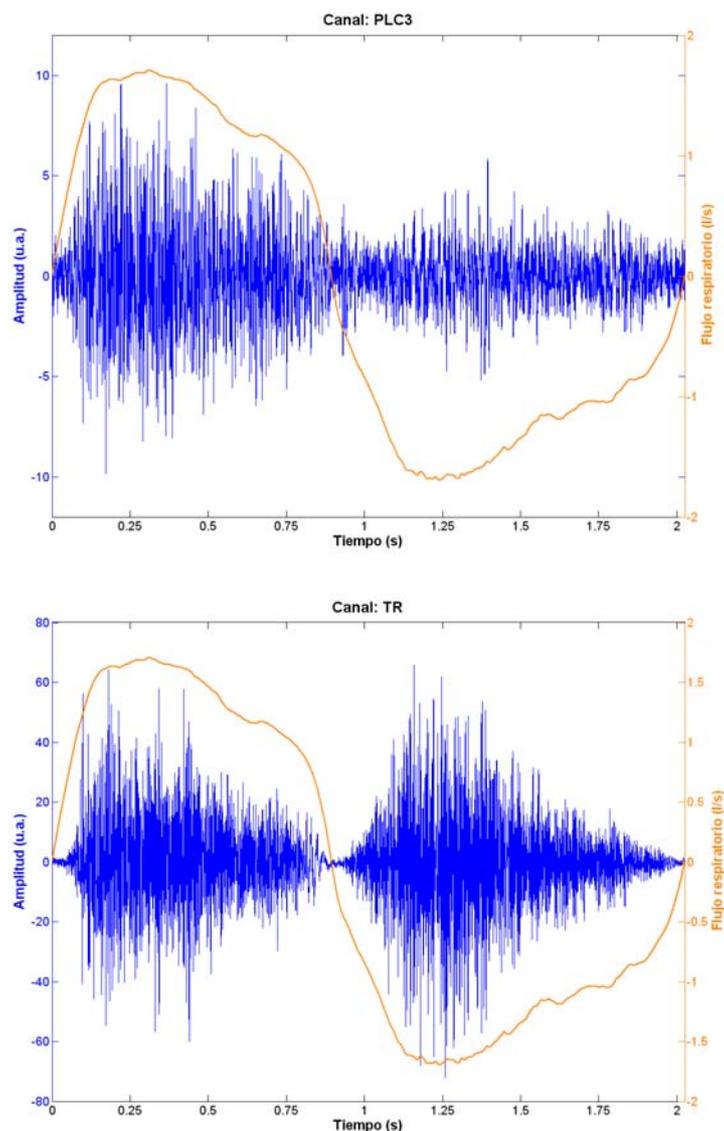


Fig. V.26. Señales de sonidos respiratorios reales adquiridas. a) Sonido pulmonar. b) Sonido traqueal.

V.6.2. Sonidos cardiacos.

Se adquirieron registros de un sujeto sano sin antecedentes de cardiopatías. La adquisición se realizó en el Instituto Nacional de Enfermedades respiratorias. El equipo de adquisición empleado corresponde al reportado en la sección de adquisición de sonidos respiratorios. Los micrófonos fueron ubicados sobre los focos pulmonar y aórtico de auscultación cardiaca, y la adquisición se realizó durante un periodo de inspiración profunda. La frecuencia de muestreo empleada fue de 10 kHz. En la figura V.27 se presenta un registro fonocardiográfico adquirido, y en las figuras V.28 y VI.29 algunas señales de sonidos cardiacos aisladas.

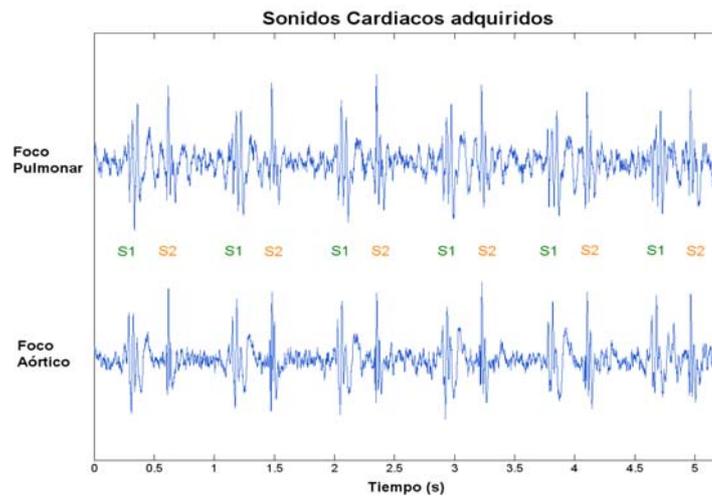


Fig. V.27. Registro de sonidos cardiacos adquiridos sobre los focos pulmonares y aórticos. Las señales fueron filtradas con un filtro FIR pasabajas de orden 201 con una frecuencia de corte de 1 kHz.

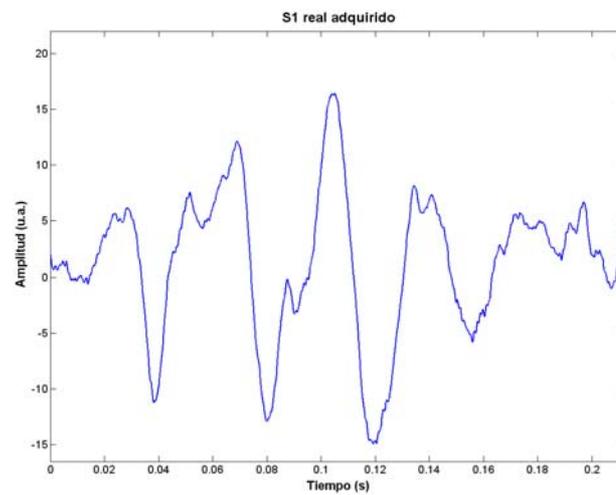


Fig. V.28. Señal de primer sonido cardiaco adquirida.

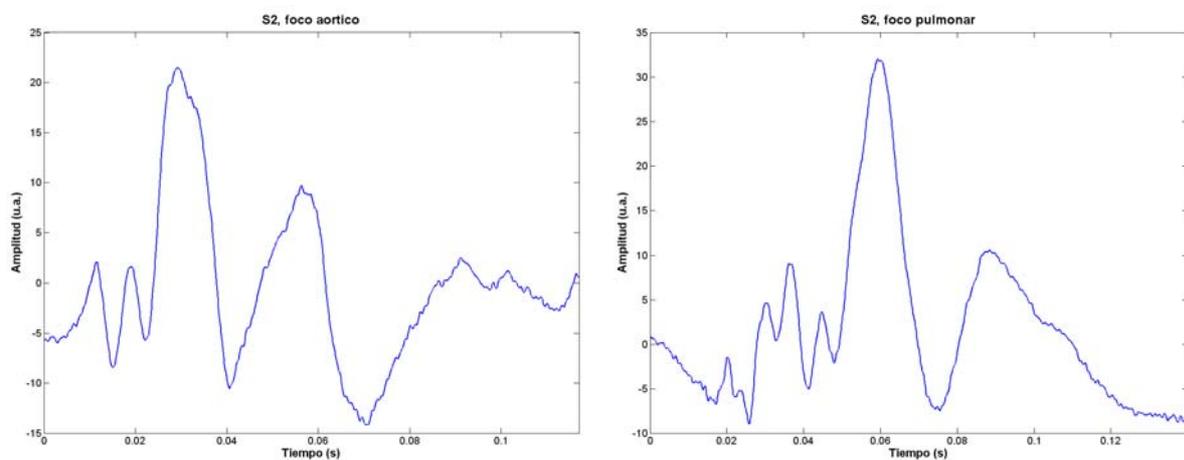


Fig. V.29. Señales adquiridas del segundo sonido cardiaco para su análisis tiempo-frecuencia. a) S2 adquirido sobre el foco aórtico. b) S2 adquirido sobre el foco pulmonar.

a b

Referencias.

- [1] N. Huang, Z. Shen, S. Long, M. Wu, H. Shih, A. Zheng, N. Yen, C. Tung, H. Liu, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and nonstationary time series analysis," *Proc. Roy. Soc. Lond.*, vol. A 454, pp. 903-995, 1998.
- [2] H. Hlawatsch, G. F. Boudreaux-Bartels, "Linear and quadratic time-frequency signal representations," *IEEE Signal Processing Magazine*, vol. 9, pp. 21-67, 1992.
- [3] L. Cohen, *Time-Frequency Analysis*. Englewood Cliffs, NJ: Prentice-Hall, 1995.
- [4] A. Bianchi, L. Mainardi, E. Petrucci, M. Signorini, M. Mainardi, S. Cerutti, "Time-variant power spectrum analysis for the detection of transient episodes in HRV signal," *IEEE Trans. Biomed. Eng.*, vol. 40, no. 2, pp. 136-144, 1993.
- [5] A. Bianchi, L. Mainardi, C. Meloni, S. Chierchia, S. Cerutti, "Continuous monitoring of the sympatho-vagal balance through spectral analysis," *IEEE Engineering in Medicine and Biology Magazine*, vol. 16, no. 5, pp. 64-73, 1997.
- [6] R. Zou, H. Wang, K. Chon, "A robust time-varying identification algorithm using basis functions," *Ann. Biomed. Eng.*, vol. 31, pp. 840-853, 2003.
- [7] Y. Cho, S. Kim, E. Powers, "Time-frequency analysis using AR models with variable forgetting factors," *Proc. ICASSP'90*, pp. 2479-2482, 1990.
- [8] Y. Cho, S. Kim, E. Powers, "Time-varying spectral estimation using AR models with variable forgetting factors," *IEEE Trans. Signal Processing*, vol. 39, pp. 1422-1425, 1991.
- [9] D. Chen, L. Durand, H. Lee, "Time-frequency analysis of the first heart sound. Part I: Simulation and analysis," *Med. Biol. Eng. Comput.*, vol. 35, pp. 306-310, 1997.
- [10] G. Cloutier, M. Grenier, R. Guardo, L. Durand, "Spectral analysis of closing sounds produced by Ionescu-Shiley bioprosthetic aortic heart valves. Part 2: Computer simulations of aortic closing sounds and estimation of their truncation level and signal-to-noise-ratio," *Med. Biol. Eng. Comput.*, vol. 25, pp. 492-496, 1987.
- [11] Y. Tang, C. Danmin, L. Durand, "The synthesis of the aortic valve closure sound of dog by the mean filter of forward and backward predictor," *IEEE Trans. Biomed. Eng.*, vol. 39, pp. 1-8, 1992.
- [12] R. Bedi, J. McDonnell, "Evaluation of parametric methods for the spectral analysis of Carpentier-Edwards heart valve sounds," *Proc. Computers in Cardiology*, pp. 675-678, 1993.
- [13] H. Sava, J. McDonnell, "New analysis-synthesis of first heart sounds using forward-backward overdetermined Prony's method," *Proc. 16th Annual International IEEE-EMBS Conf.*, pp. 1274-1275.
- [14] J. Wood, A. Buda, D. Barry, "Time-frequency transforms: A new approach to first heart sound frequency dynamics," *IEEE Trans. Biomed. Eng.*, vol. 39, no. 7, pp. 730-740, 1992.
- [15] J. Wood, D. Barry, "Quantification of first heart sound frequency dynamics across the human chest wall," *Med. Biol. Eng. Comput.*, vol. 32, pp. S71-S78, 1994.

-
- [16] J. Xu, L. Durand, P. Pibarot, "Extraction of the aortic and pulmonary components of the second heart sound using a nonlinear transient chirp signal model," *IEEE Trans. Biomed. Eng.*, vol. 47, no. 7, pp. 1328-1335, 2000.
- [17] H. Kiyokawa, M. Greenberg, K. Shirota, and H. Pasterkamp, "Auditory detection of simulated crackles in breath sounds," *Chest*, vol. 119, pp. 1886-1892, 2001.
- [18] R. Murphy, S. Holford, W. Knowler, "Visual lung sound characterization by time-expanded wave-form analysis," *N. Engl. J. Med.*, 296, pp. 968-971, 1977.
- [19] A. Sovijärvi, F. Dalmaso, J. Vanderschoot, L. Malmberg, G. Righini, S. Stoneman, "Definitions of terms for applications of respiratory sounds," *Eur. Respir. Rev.*, vol. 10, pp. 597-610, 2000.
- [20] L. Rankine, N. Stevenson, M. Mesbah, B. Boashash, "A Quantitative Comparison of Non-Parametric Time-Frequency Representations," *Proc. 13th European Signal Processing Conference (EUSIPCO2005)*, Turkia, 2005.
- [21] L. White, B. Boashash, "Estimating the instantaneous frequency of a Gaussian random process using the Wigner-Ville Distribution," *IEEE Trans. Acoust. Spech. Sig. Proc.*, vol. 36, no. 3, pp. 417-420, 1988.
- [22] B. Boashash, G. Jones, "Instantaneous frequency and time-frequency distributions," en *Time-Frequency Signal Analysis*, editado por B. Boashash, pp. 43-73, Longman Cheshire, 1992.
- [23] D. Chen, L. Durand, Z. Guo, H. Lee, "Time-frequency analysis of the first heart sound. Part II: An appropriate time-frequency representation technique," *Med. Biol. Eng. Comput.*, vol. 35, pp. 311-317, 1997.

VI. RESULTADOS.

VI.1. Introducción.

A continuación se presentan los resultados obtenidos al aplicar las diversas técnicas tiempo-frecuencia a las señales no biológicas y biológicas, simuladas y reales. Para cada técnica se ilustrará el proceso de selección de sus parámetros y posteriormente, se presentará la comparación entre las representaciones tiempo-frecuencia adecuadas para la selección final.

VI.2. Señal no biológica generada mediante el modelo TVAR.

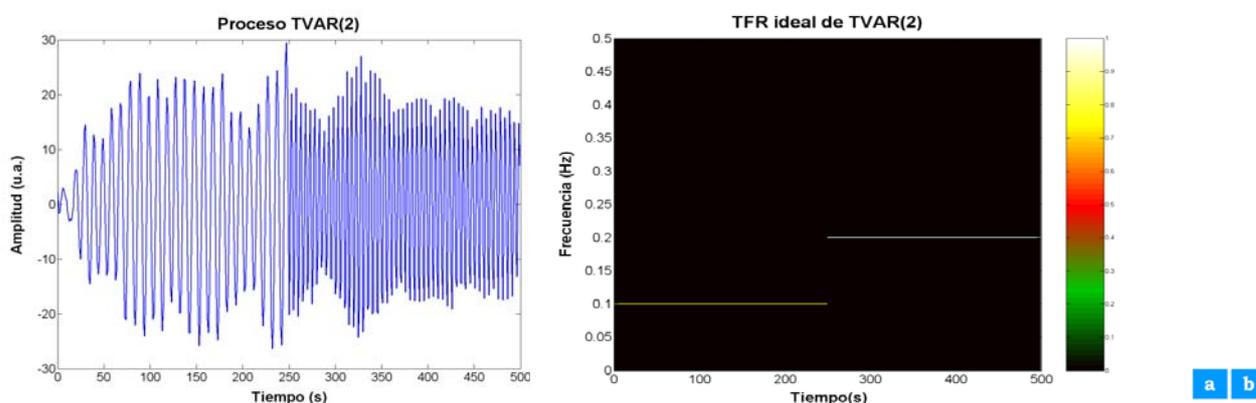


Fig. VI.1. Proceso TVAR de segundo orden. a) Forma de onda, y b) Representación tiempo-frecuencia teórica.

▪ Espectrograma.

En las figuras VI.2 – VI.4 se muestran los espectrogramas resultantes al emplear funciones ventana Hamming, Rectangular y Blackman-Harris con nueve valores de longitud. Los espectrogramas se muestran entre las frecuencias de 0 y 0.3 Hz. La paleta de colores de todas las representaciones tiempo-frecuencia es la mostrada en la figura VI.2. Esta paleta emplea un color negro para el valor mínimo de intensidad, un color rojo para valores medios y un color blanco para el valor máximo.

En las figuras VI.5 – VI.7 se muestran los valores de correlación obtenidos entre la imagen de la figura VI.1.b (representación tiempo-frecuencia ideal) y las imágenes de las figuras VI.2-VI.4. La correlación se efectuó considerando sub-imágenes en cada uno de los espectrogramas. La paleta de colores empleada para la comparación local de las representaciones tiempo-frecuencia utiliza el color azul para valores

mulos y el color rojo para una correlación unitaria, como se muestra en la figura VI.5.

En la figura VI.8 se presentan los valores de correlación promedio para las tres representaciones tiempo-frecuencia en función de la longitud de la ventana, mientras que en la figura VI.9 se presenta el espectrograma seleccionado de acuerdo al criterio de correlación local.

En el Anexo I se presenta la función de ambigüedad de la señal y las funciones kernel para cada espectrograma correspondiente de las figuras VI.2 – VI.4, además se presenta la tabla con los valores de correlación promedio correspondiente a la figura VI.8.

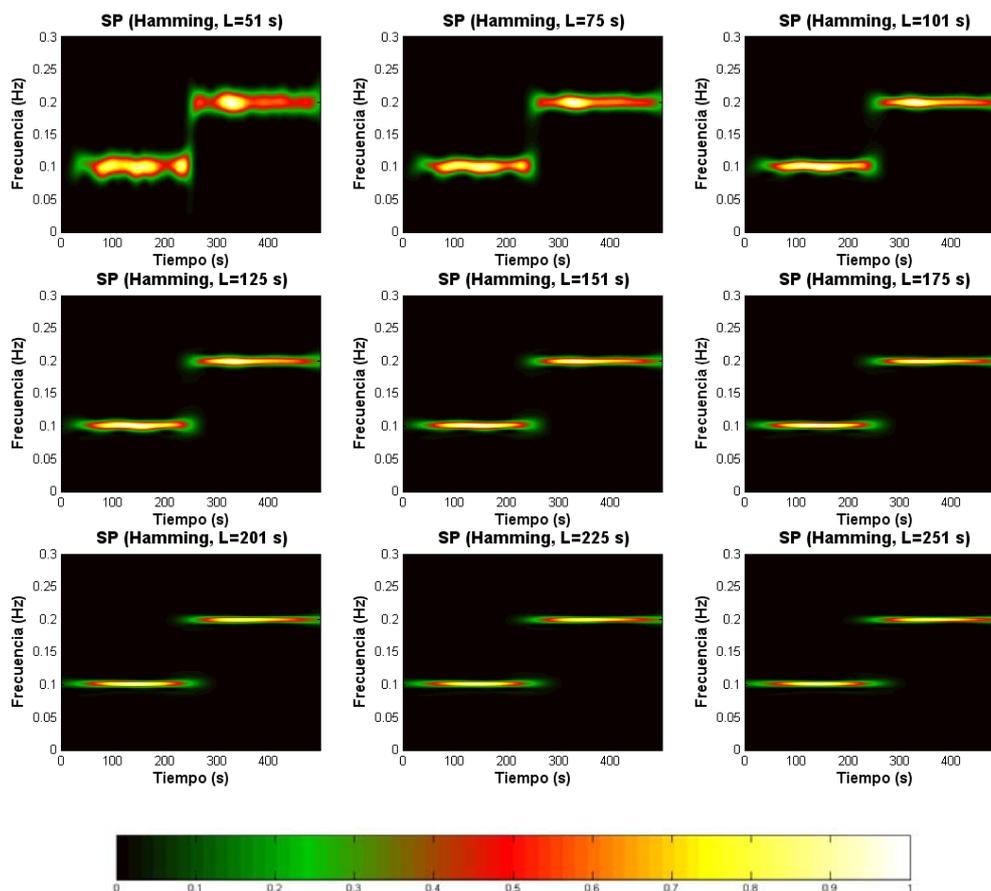


Fig. VI.2. Espectrogramas obtenidos con una ventana Hamming con nueve valores de longitud.

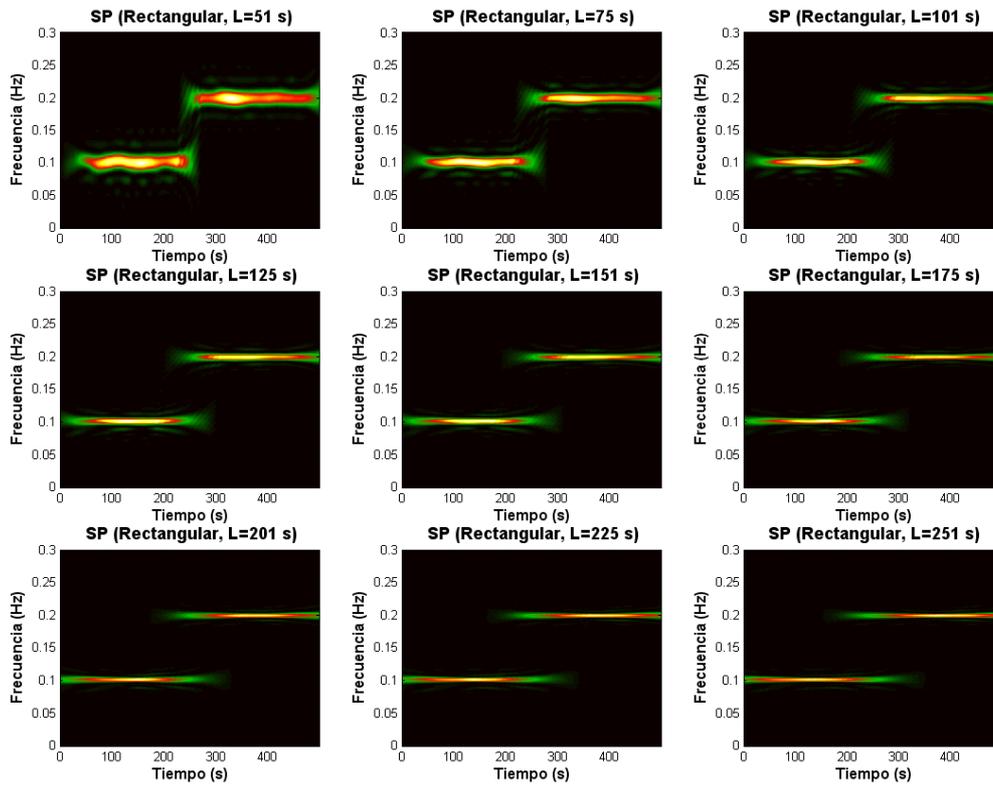


Fig. VI.3. Espectrogramas obtenidos con una ventana Rectangular con nueve valores de longitud.

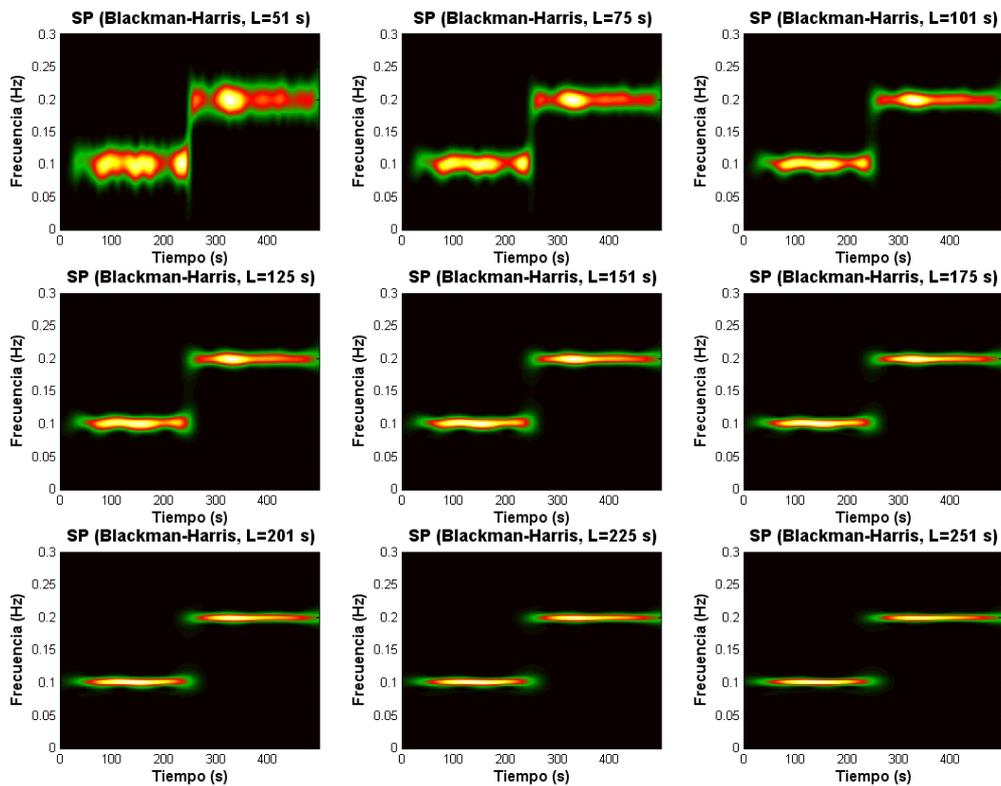
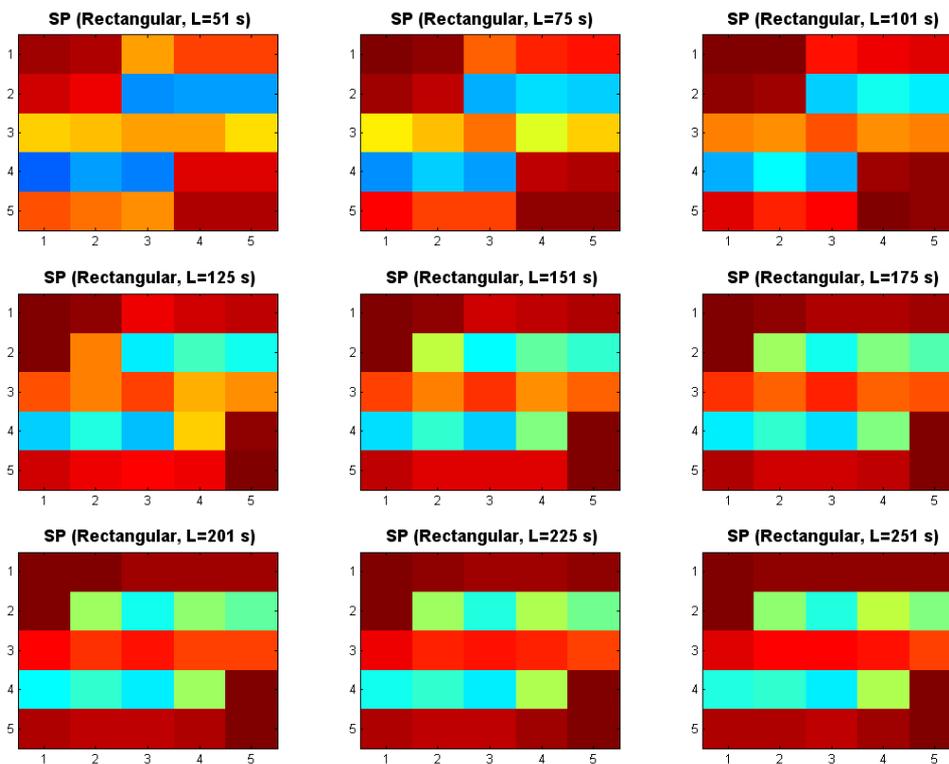
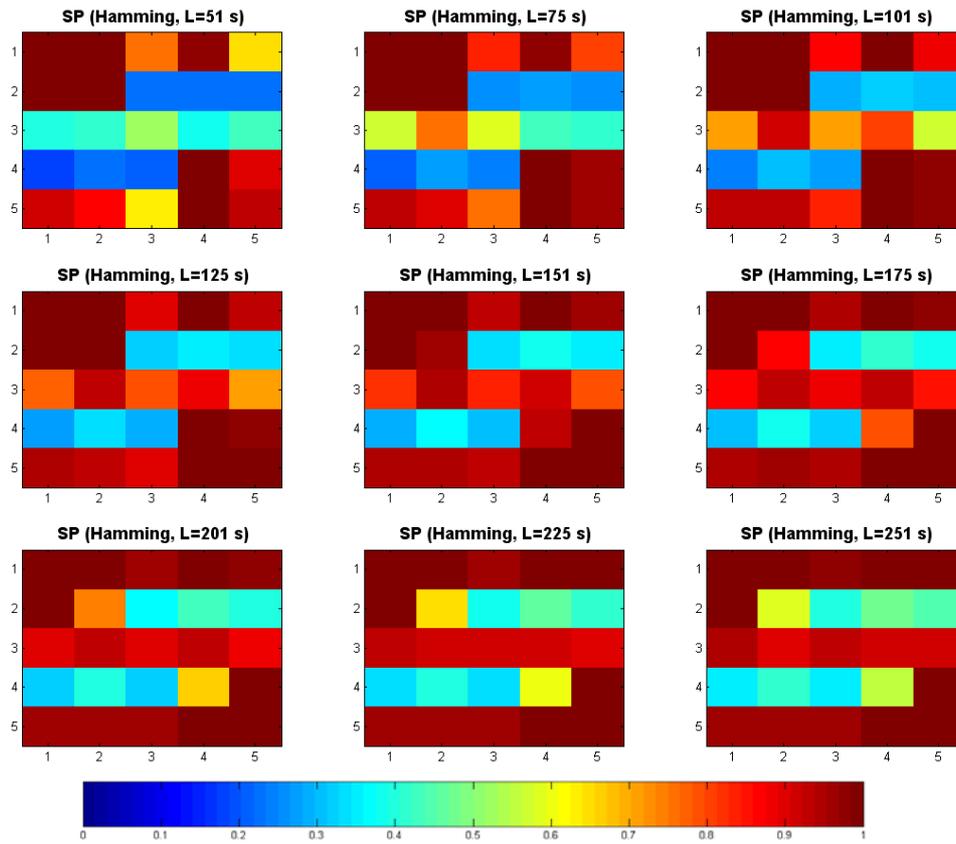


Fig. VI.4. Espectrogramas obtenidos con una ventana Blackman-Harris con nueve valores de longitud.



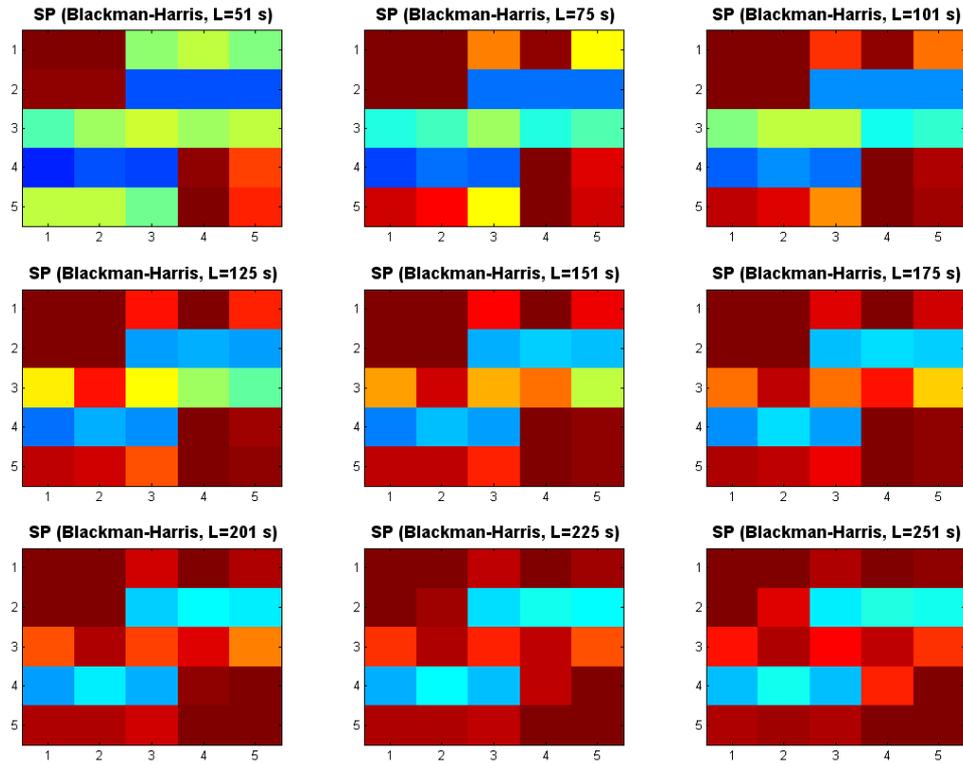


Fig. VI.7. Valores de correlación de las sub-imágenes de los espectrogramas con la ventana Blackman-Harris.

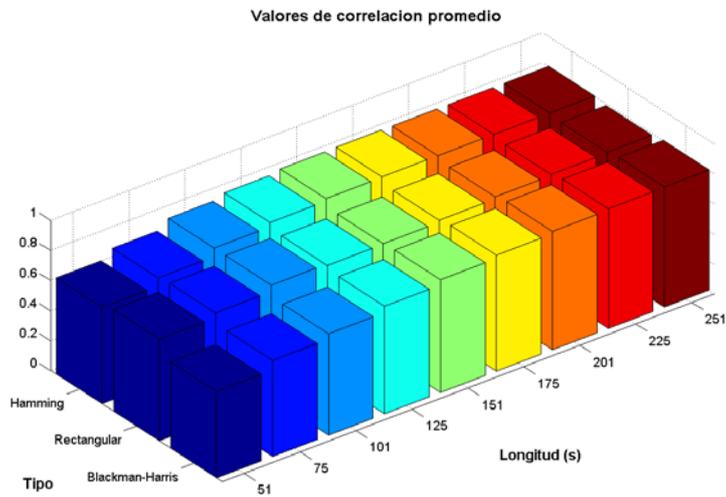


Fig. VI.8. Valores de $\rho_{promedio}$ entre cada espectrograma y la representación tiempo-frecuencia ideal.



- **Distribución Wigner-Ville.**

En la figura VI.10 se muestra la distribución de Wigner-Ville del proceso TVAR. En la figura VI.11 se muestra la comparación local entre la distribución de Wigner-Ville y la representación tiempo-frecuencia teórica del proceso TVAR. El valor de $\rho_{promedio}$ es igual a 0.6110.

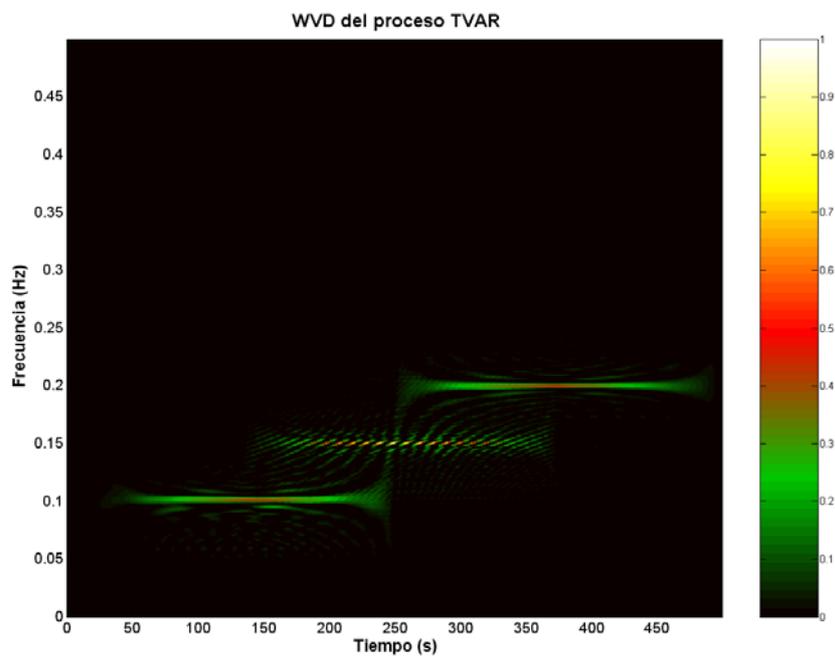
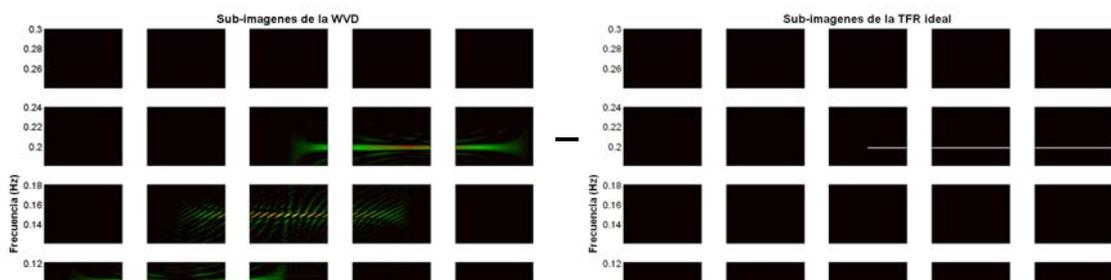


Fig. VI.10. Distribución Wigner-Ville del proceso TVAR.





- **Distribución Choi-Williams.**

En la figura VI.12 se presenta la selección del valor de σ_{max} , mientras que en las figuras VI.13 y VI.14 se presentan las distribuciones Choi-Williams para 27 valores del parámetro σ en el intervalo $[1 - \sigma_{max}]$. El cálculo de la correlación entre sub-imágenes de la distribución Choi-Williams con la representación ideal se realizó como en el caso del Espectrograma. Por motivos de espacio no se presentan las imágenes de correlación pero la información se reporta mediante la figura VI.15 donde se grafican los valores de $\rho_{promedio}$ en función de σ ; en el Anexo I se reportan los valores de $\rho_{promedio}$ empleados para seleccionar la CWD. En la figura VI.16 se reporta la CWD seleccionado con base en el criterio de máxima correlación promedio.

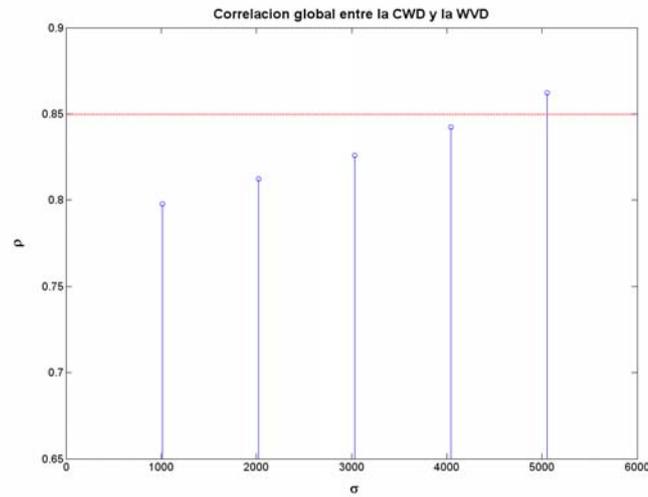


Fig. VI.12. Determinación del rango de valores de σ para el computo de las distribuciones Choi-Williams. La línea punteada indica una correlación global de 0.85 entre la WVD y la CWD.

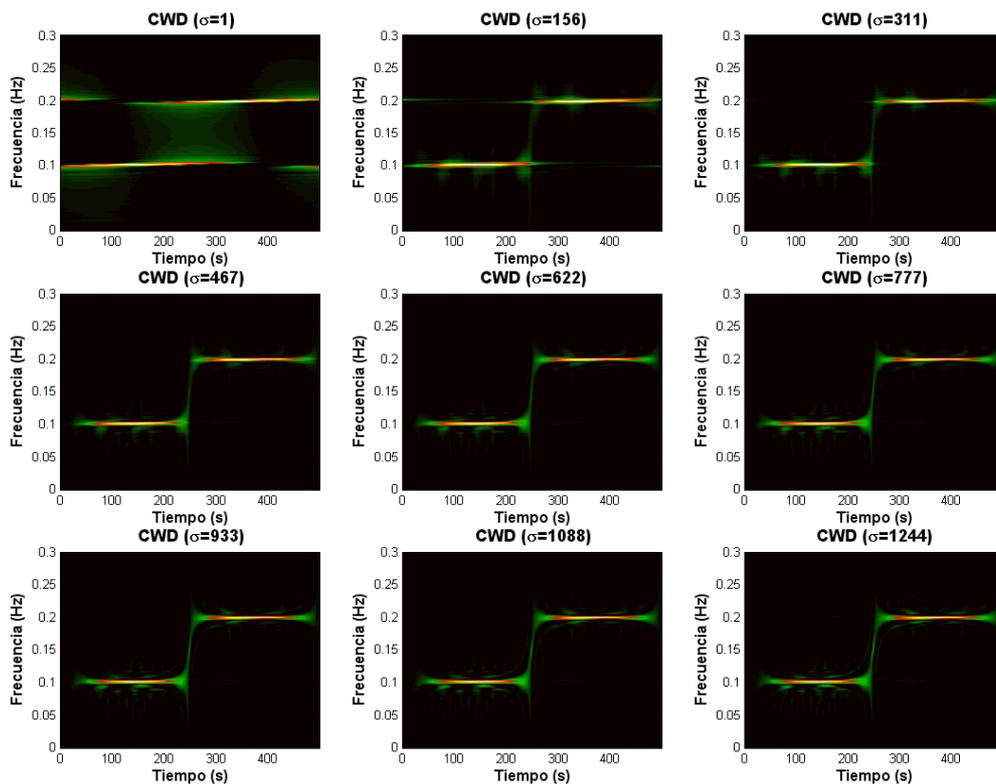
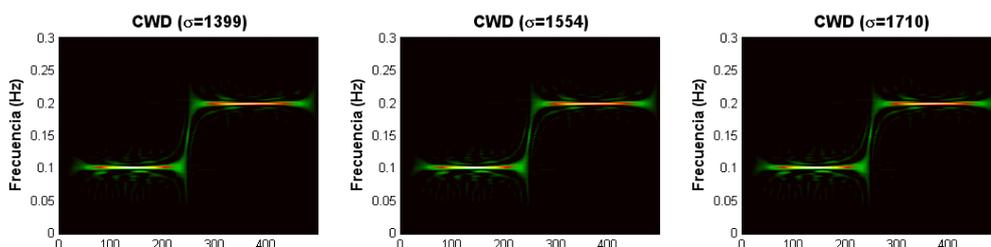
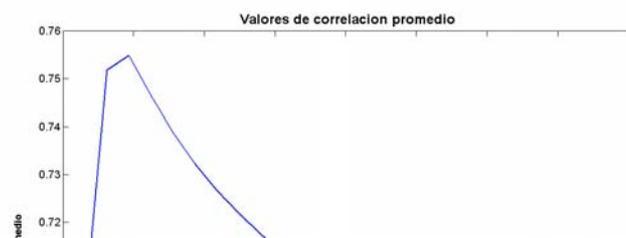


Fig. VI.13. Distribuciones Choi-Williams del proceso TVAR para distintos valores de σ .





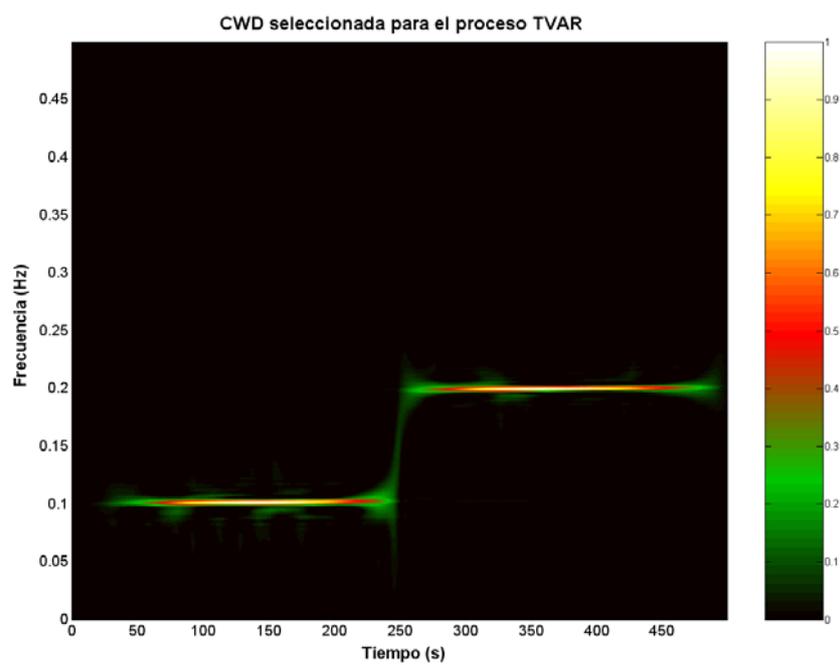


Fig. VI.16. Distribución Choi-Williams seleccionada para el proceso TVAR. $\sigma = 467$.

- **Espectrograma reasignado.**

En las figuras VI.17 – VI.19 se presenta la reasignación de los espectrogramas mostrados en las figuras VI.2 – VI.4, respectivamente. En el Anexo I se presentan los valores de correlación obtenidos mediante la comparación local entre los espectrogramas reasignados y la representación tiempo-frecuencia ideal de la señal. En la figura VI.20 se presentan los valores de correlación promedio para cada espectrograma reasignado, a partir de los cuales se seleccionó el espectrograma reasignado final, ver figura VI.21. En el Anexo I se presenta la tabla con los valores de correlación promedio correspondientes.

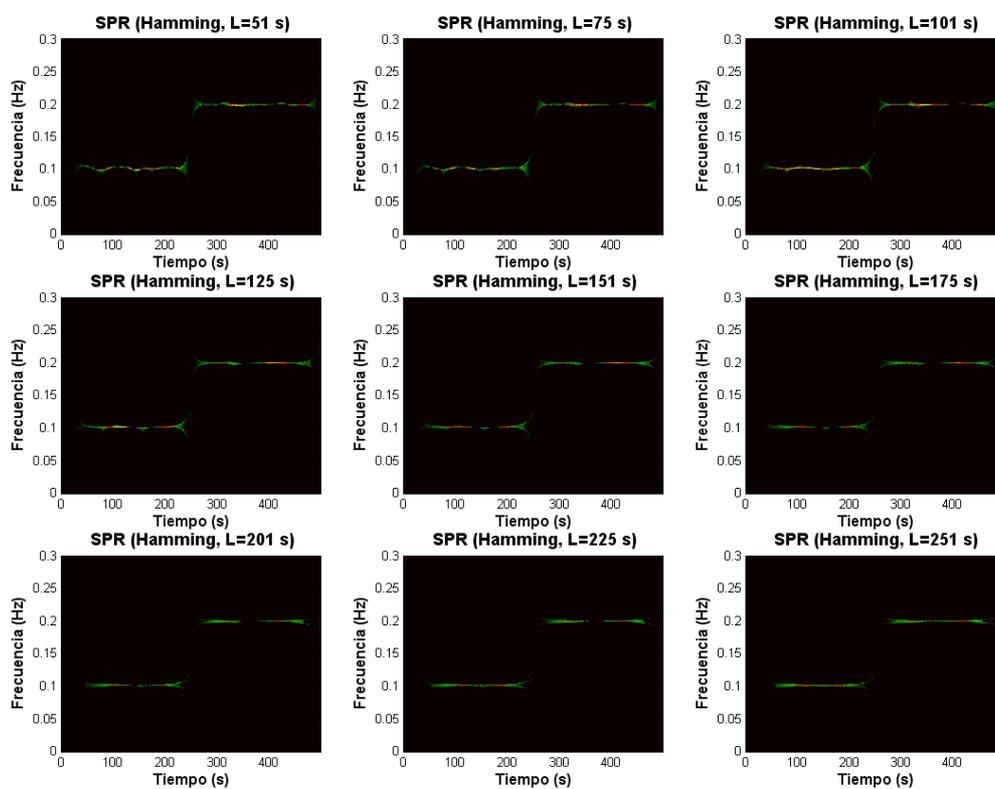
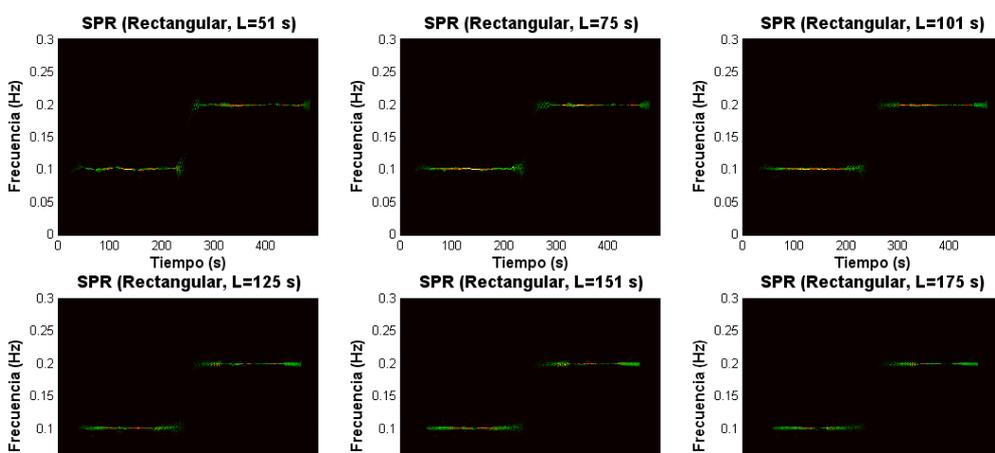


Fig. VI.17. Espectrogramas reasignados obtenidos con una ventana Hamming con nueve valores de longitud.



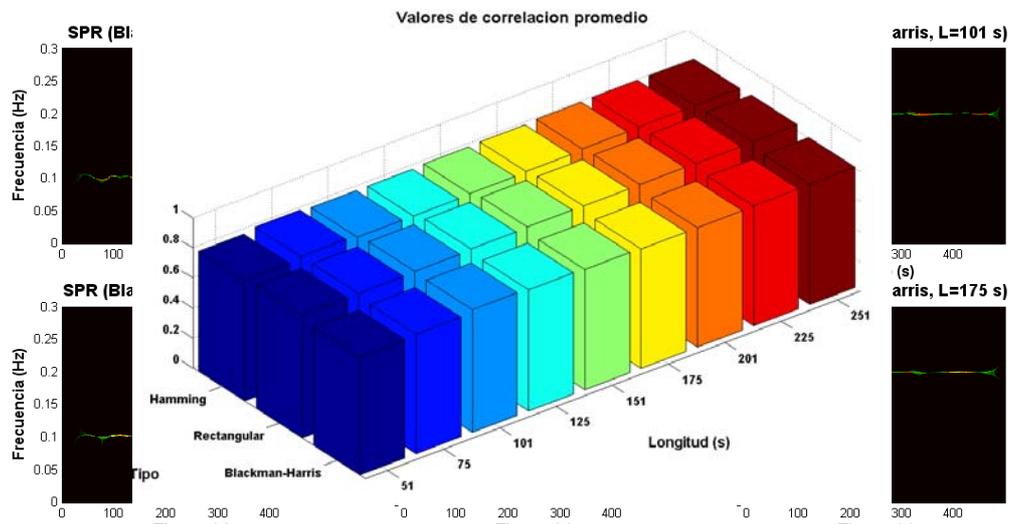


Fig. VI.20. Valores de $\rho_{promedio}$ entre cada espectrograma reasignado y la representación tiempo-frecuencia ideal para el proceso TVAR.

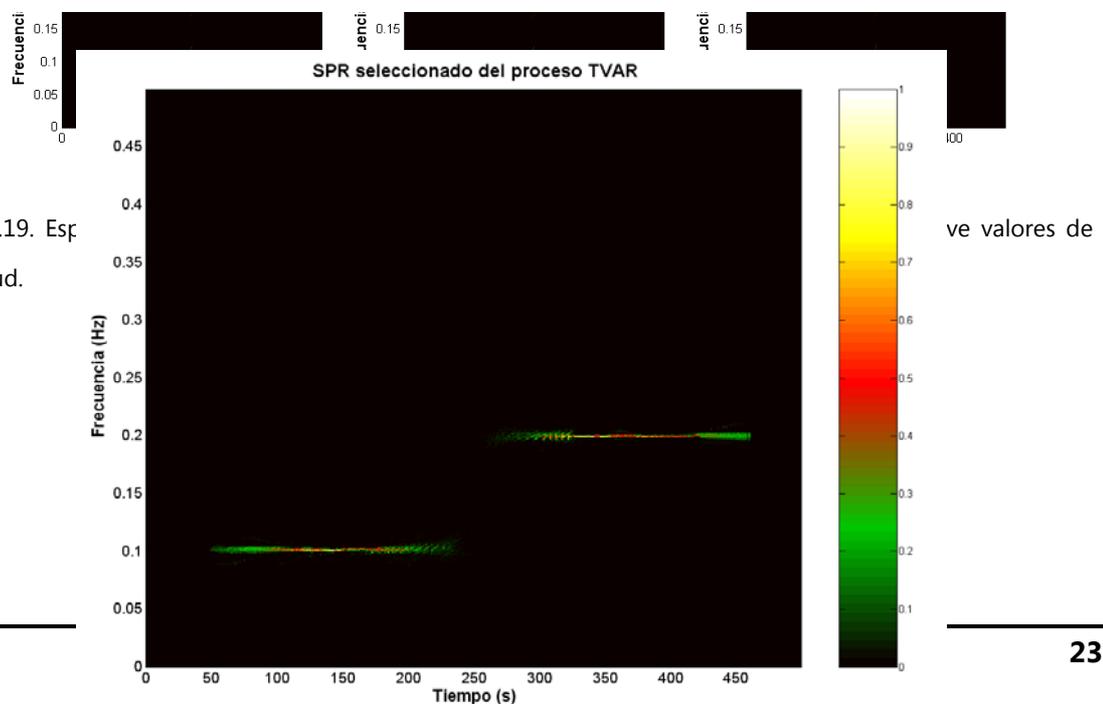


Fig. VI.19. Espectrograma reasignado seleccionado para el proceso TVAR. Ventana Rectangular de 151 s de longitud.

Fig. VI.21. Espectrograma reasignado seleccionado para el proceso TVAR. Ventana Rectangular de 151 s de longitud.

▪ **Representación tiempo-frecuencia mediante P_{BURG} .**

En la figura VI.22 y VI.23 se presentan las representaciones tiempo-frecuencia obtenidas vía el modelado autorregresivo por el método de Burg en función del orden del modelo, tipo de ventana y diversas longitudes de la ventana.

En el Anexo I se presentan los valores de correlación obtenidos mediante la comparación local entre las representaciones paramétricas por el método de Burg y la representación tiempo-frecuencia teórica de la señal. En la tabla VI.I se presentan los valores de correlación promedio para cada representación paramétrica de Burg. Finalmente, en la figura VI.24 se presenta la representación tiempo-frecuencia de Burg seleccionada para el proceso TVAR.

TABLA VI.I. VALORES DE $\rho_{promedio}$ PARA LAS REPRESENTACIONES TIEMPO-FRECUENCIA OBTENIDAS VIA EL MÉTODO DE BURG.

M=2	Hamming	Blackman-Harris
51 s	0.7454	0.6794
75 s	0.7692	0.7426
101 s	0.7774	0.7650
125 s	0.7886	0.7692
151 s	0.7961	0.7763
175 s	0.7972	0.7838
201 s	0.7925	0.7919
225 s	0.7867	0.7950
251 s	0.7793	0.7966

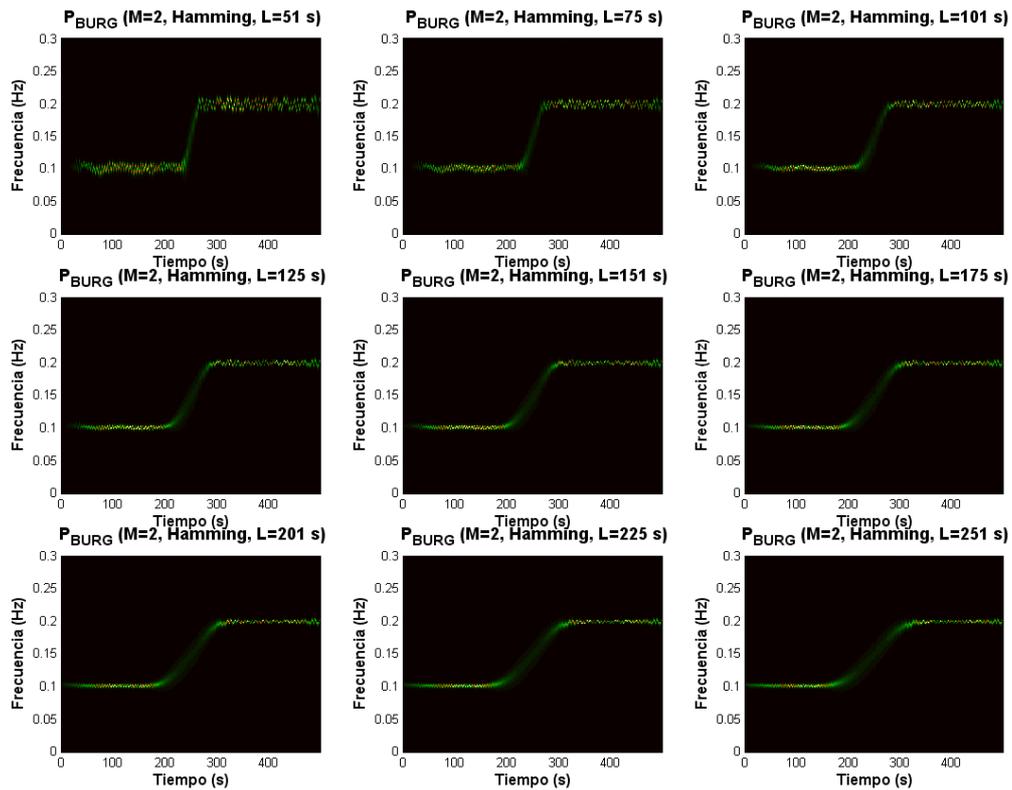


Fig. VI.22. Representaciones tiempo-frecuencia de Burg para el proceso TVAR obtenidas con un orden del modelo AR igual a 2, ventana Hamming y nueve valores de longitud.

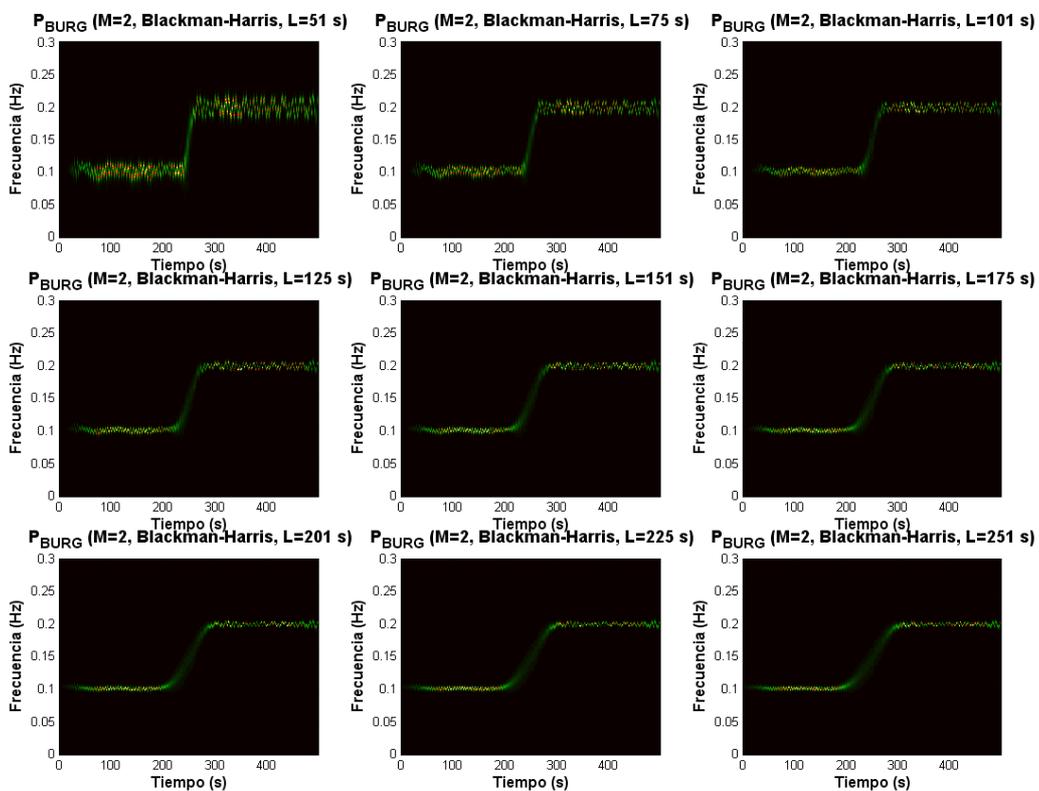


Fig. VI.23. Representaciones tiempo-frecuencia de Burg para el proceso TVAR, obtenidas con un orden del modelo AR igual a 2, ventana Blackman-Harris y nueve valores de longitud.

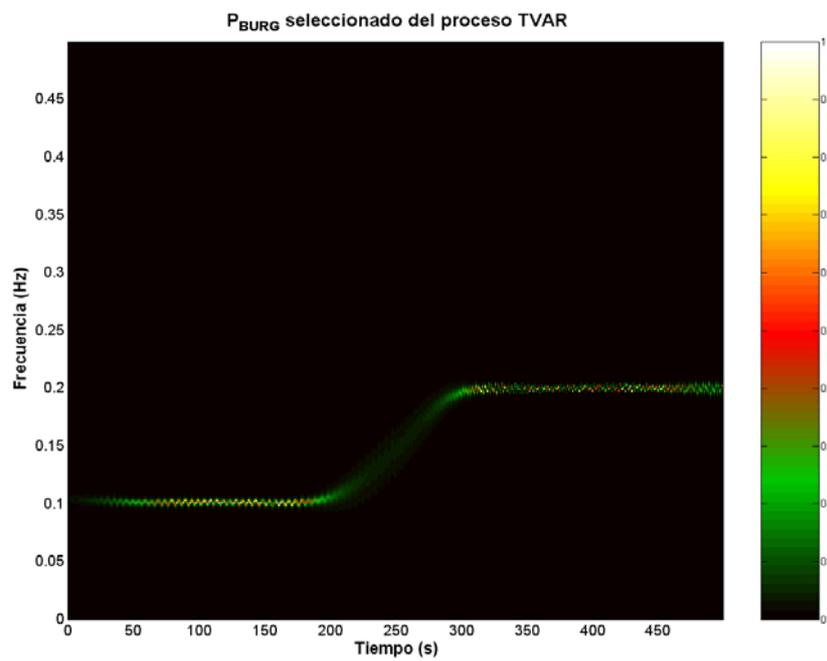


Fig. VI.24. Representación tiempo-frecuencia paramétrica vía el método de Burg seleccionada para el proceso TVAR. Orden del modelo igual a 2, ventana Hamming de 175 s de longitud.

- **Representación tiempo-frecuencia mediante P_{RLS}.**

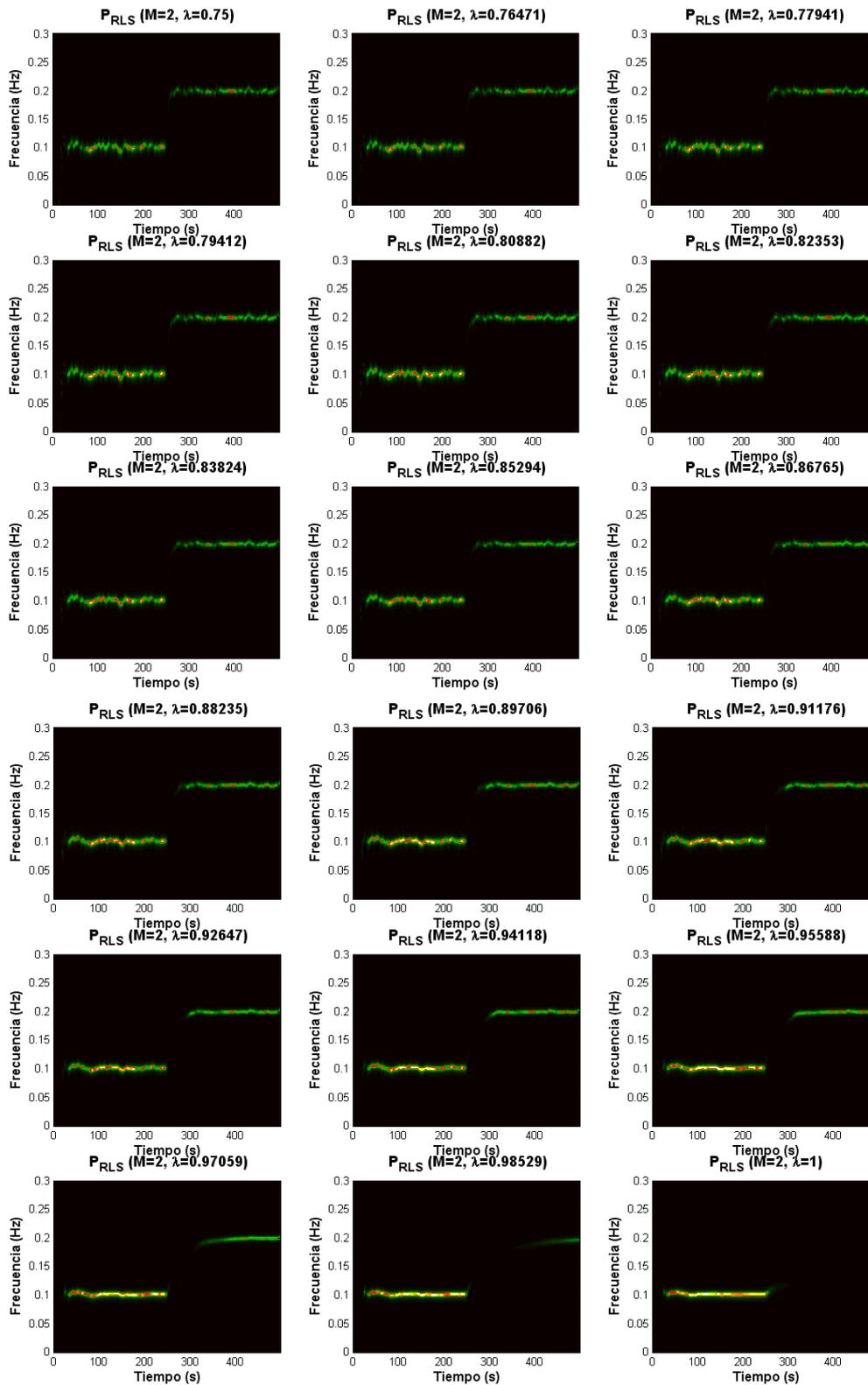


Fig. VI.25. Representaciones tiempo-frecuencia para el proceso TVAR vía el algoritmo RLS para un orden del modelo igual a 2 y diversos valores del factor de olvido.

En la figura VI.25 se muestran las representaciones tiempo-frecuencia paramétricas obtenidas vía el algoritmo RLS (P_{RLS}) para un modelo AR de segundo orden y

diversos valores del factor de olvido. En la figura VI.26 se presenta los valores de correlación promedio en función del factor de olvido, en el Anexo I se presenta la tabla con los valores de correlación promedio. La representación tiempo-frecuencia P_{RLS} seleccionada para esta señal se presenta en la figura VI.27.

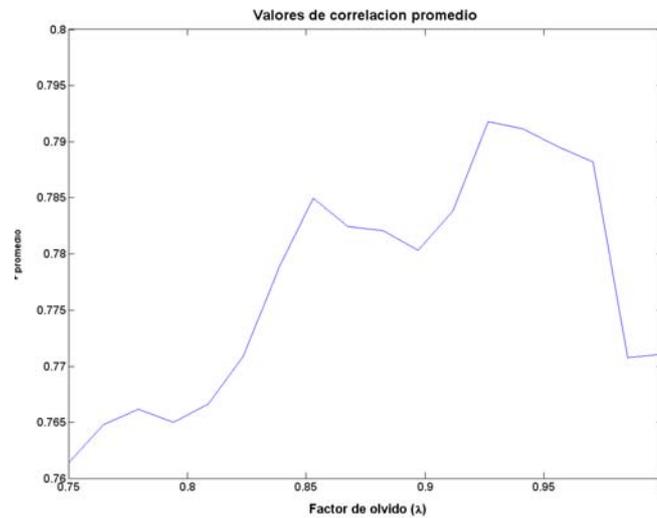


Fig. VI.26. Valores de $\rho_{promedio}$ entre cada representación tiempo-frecuencia paramétrica vía el algoritmo RLS y la representación tiempo-frecuencia ideal para el proceso TVAR.

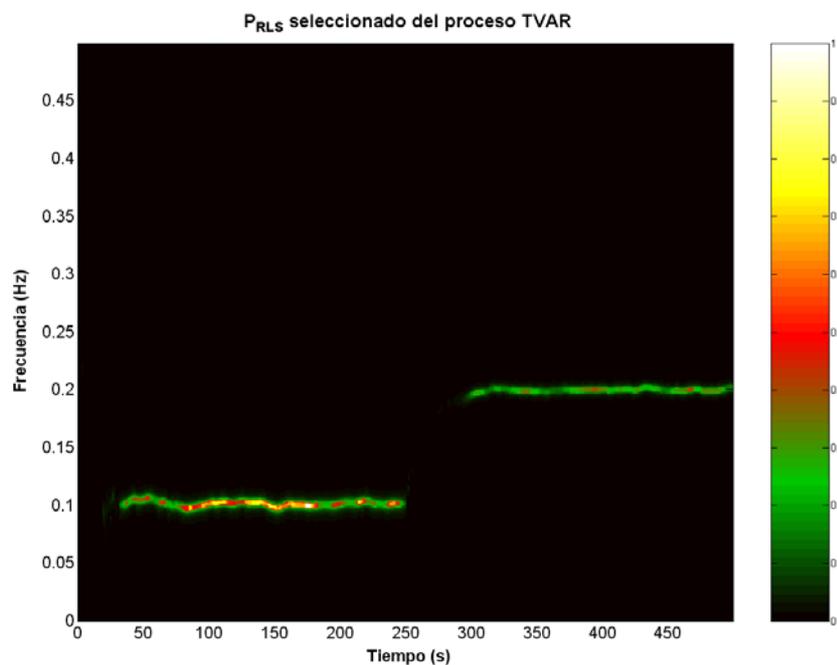


Fig. VI.27. Representación tiempo-frecuencia vía el algoritmo RLS seleccionada para el proceso TVAR. El factor de olvido es igual a 0.9265.

▪ **Representación tiempo-frecuencia mediante $P_{\text{RLS-VFF}}$.**

En la figura VI.28 se muestran las representaciones tiempo-frecuencia paramétricas obtenidas vía el algoritmo RLS-VFF para un modelo AR de segundo orden y diversos valores de la varianza del ruido esperado (σ_e^2). En la figura VI.29 se muestran los valores de correlación promedio obtenidos de la comparación local de cada representación $P_{\text{RLS-VFF}}$ y en el Anexo I se incluye una tabla con los valores de correlación promedio. En la figura VI.30 se muestra la representación $P_{\text{RLS-VFF}}$ seleccionada para esta señal.

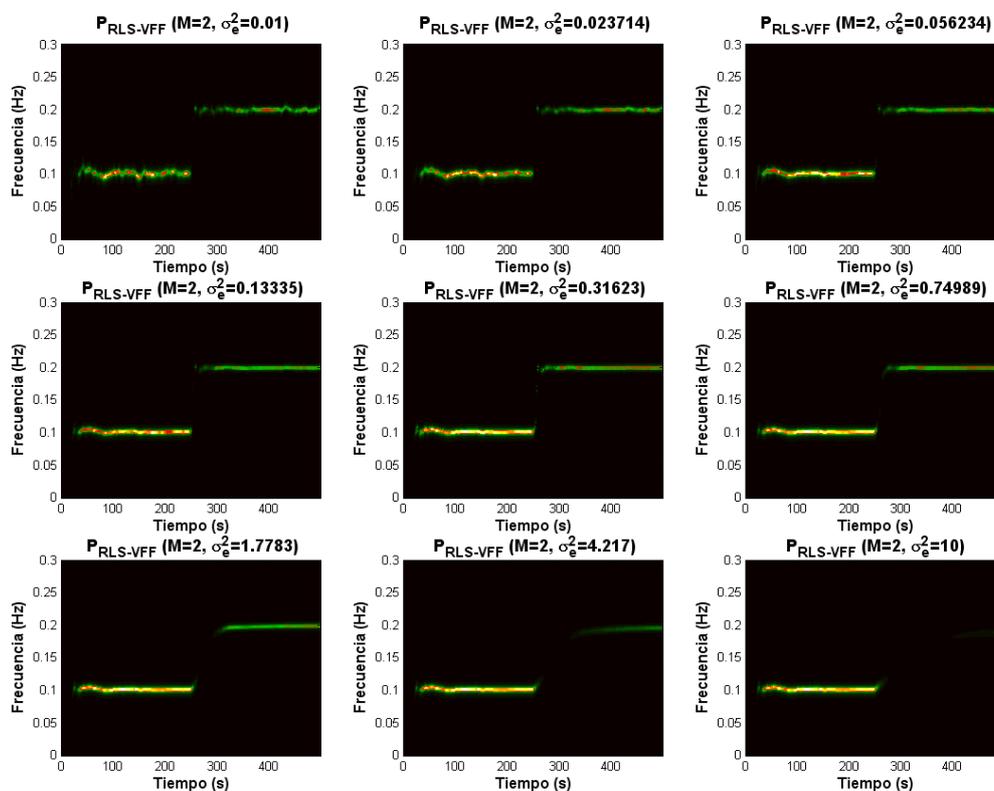


Fig. VI.28. Representaciones tiempo-frecuencia para el proceso TVAR vía el algoritmo RLS-VFF para un orden del modelo igual a 2 y diversos valores la varianza del ruido esperado (σ_e^2).

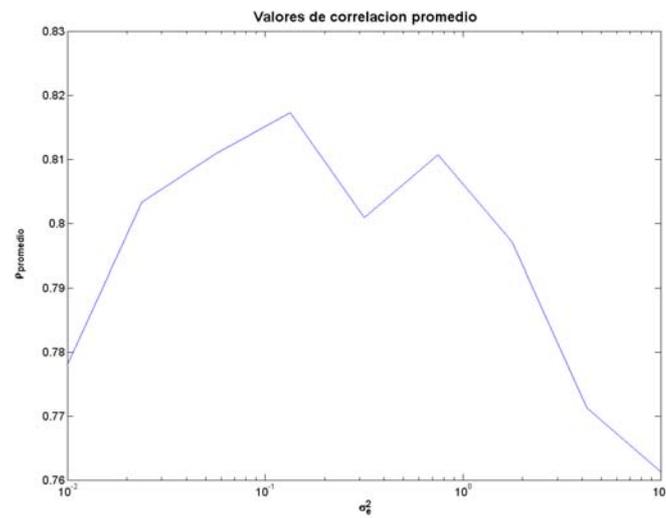


Fig. VI.29. Valores de $\rho_{promedio}$ entre cada representación tiempo-frecuencia paramétrica vía el algoritmo RLS-VFF y la representación tiempo-frecuencia ideal para el proceso TVAR.

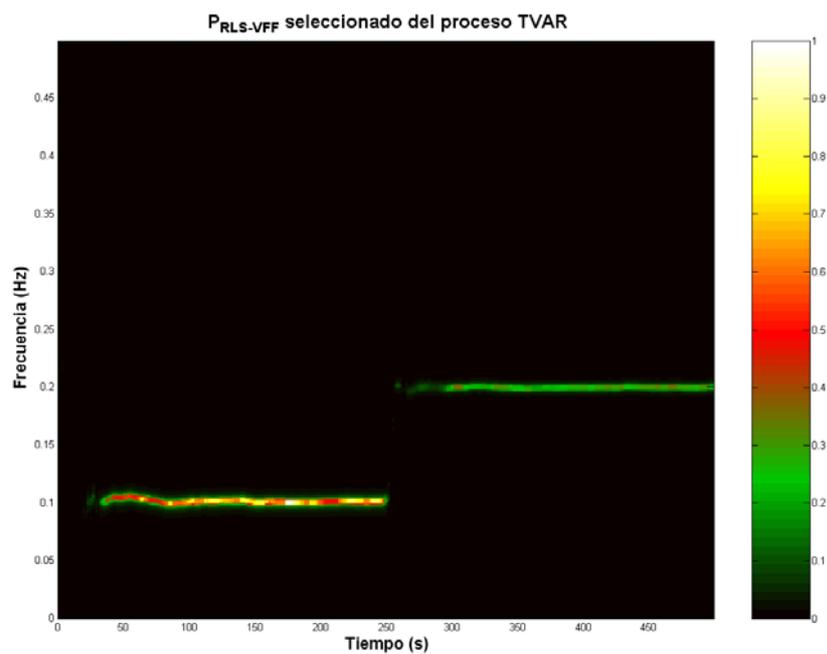


Fig. VI.30. Representación tiempo-frecuencia seleccionada vía el algoritmo RLS-VFF para el proceso TVAR. La varianza del ruido esperado es igual a 0.1334.

- Representación tiempo-frecuencia mediante P_{KALMAN} .

Las representaciones tiempo-frecuencia paramétricas del proceso TVAR obtenidas vía el filtro de Kalman se presentan en la figura VI.31 para un orden del modelo igual a 2, una varianza del ruido de medición (σ^2) igual a 0.1, y diversos valores de la varianza del ruido del proceso (q). Los resultados de $\rho_{promedio}$ para la selección de la representación tiempo-frecuencia vía el filtro de Kalman se presentan en la figura VI.32, y los valores obtenidos se reportan en el Anexo I. La representación P_{KALMAN} seleccionada para esta señal se presenta en la figura VI.33.

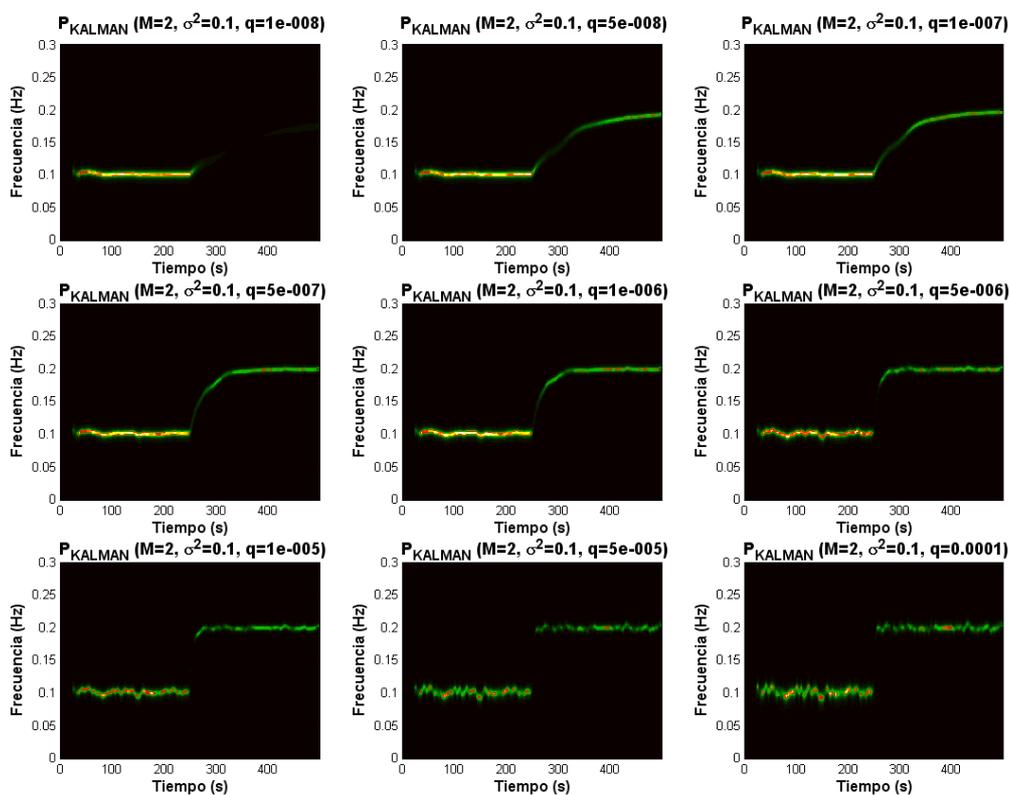


Fig. VI.31. Representaciones tiempo-frecuencia para el proceso TVAR vía el filtro de Kalman para un orden del modelo AR(2), varianza del ruido de medición igual a 0.1 y diversos valores la varianza del ruido proceso.



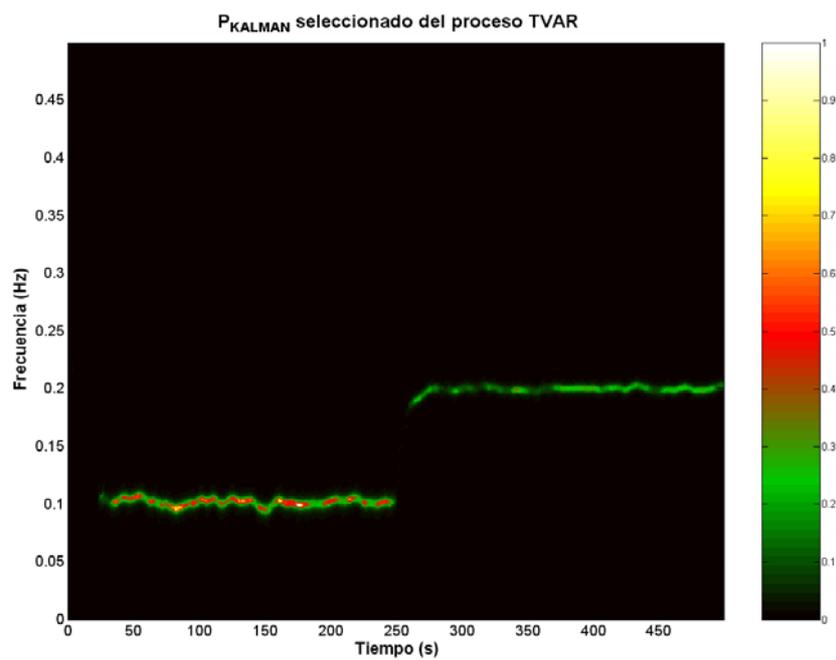


Fig. VI.33. Representación tiempo-frecuencia seleccionada vía el filtro de Kalman para el proceso TVAR. La varianza del ruido del proceso es igual a 1×10^{-5} .

Escalograma.

Los escalogramas Morlet obtenidos para el proceso TVAR con diferentes longitudes de la *wavelet* en la escala más gruesa se presentan en la figura VI.34. En la figura VI.35 se presentan los valores de correlación promedio obtenidos mediante la comparación local de cada escalograma con la representación tiempo-frecuencia ideal; en el anexo I se presenta la tabla con los valores correspondientes a la figura VI.35. En la figura VI.36 se presenta el escalograma seleccionado para esta señal.

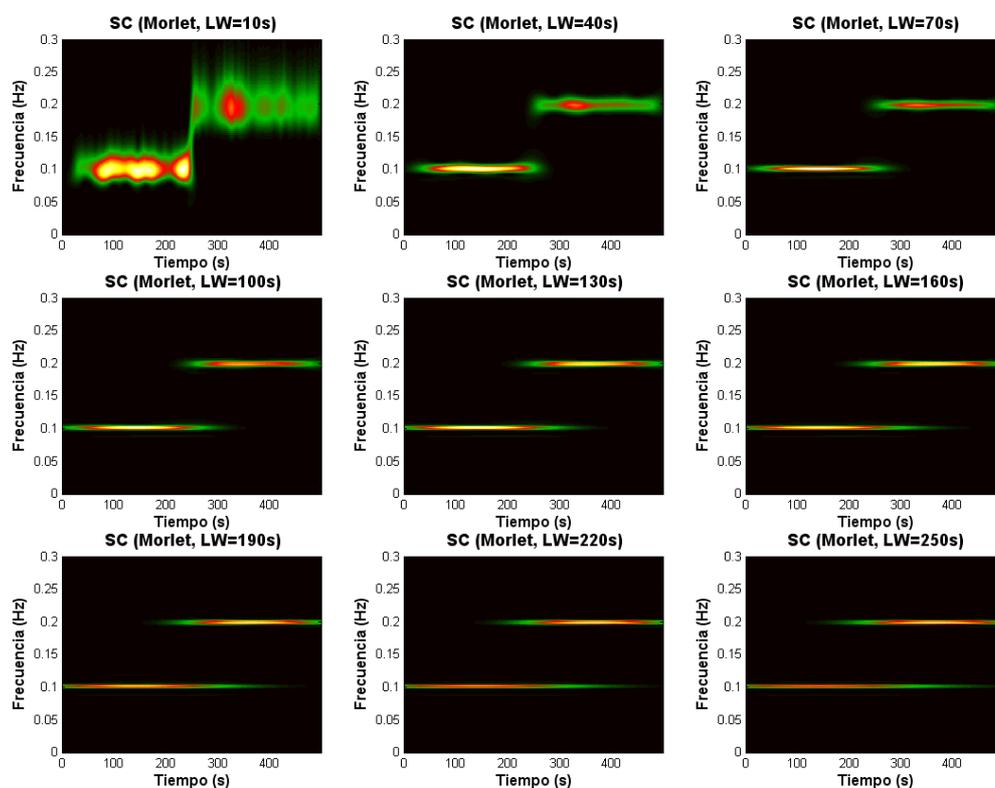
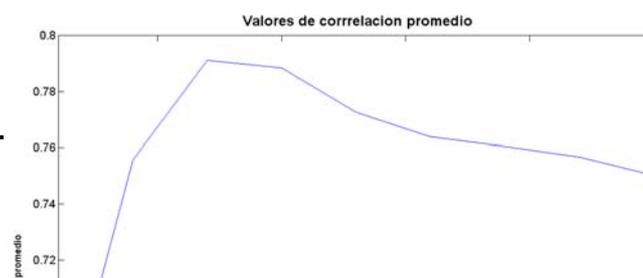


Fig. VI.34. Escalogramas del proceso TVAR obtenidos con una *wavelet* Morlet y diversas longitudes de la *wavelet* en la escala más gruesa.



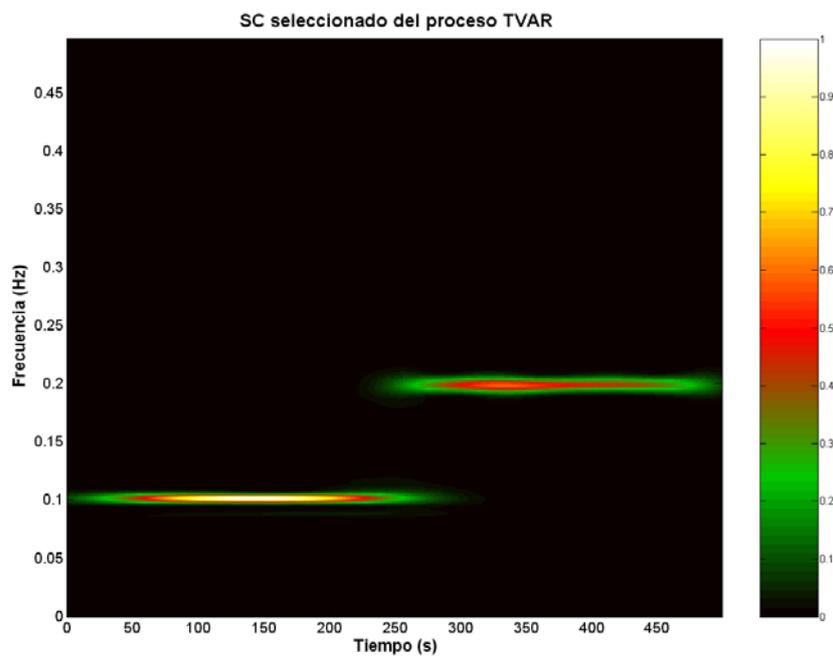


Fig. VI.36. Escalograma seleccionado para el proceso TVAR. Wavelet Morlet de 70 s de longitud en la escala más gruesa.

- **Espectro de Hilbert-Huang.**

La descomposición empírica de modos del proceso TVAR se presenta en la figura VI.37. En la figura VI.38 se presentan la amplitud y la frecuencia instantánea del primer modo de oscilación intrínseco, seleccionado para obtener el espectro de Hilbert-Huang conforme al criterio de energía, ver Anexo I. En la misma figura se incluye la frecuencia instantánea del IMF_1 convolucionada con un filtro de mediana de siete muestras de longitud. El espectro de Hilbert-Huang del proceso TVAR se presenta en la figura VI.39. Su comparación local con la representación tiempo-frecuencia teórica de la señal se presenta en el Anexo I y el valor de $\rho_{promedio}$ es igual a 0.7919.

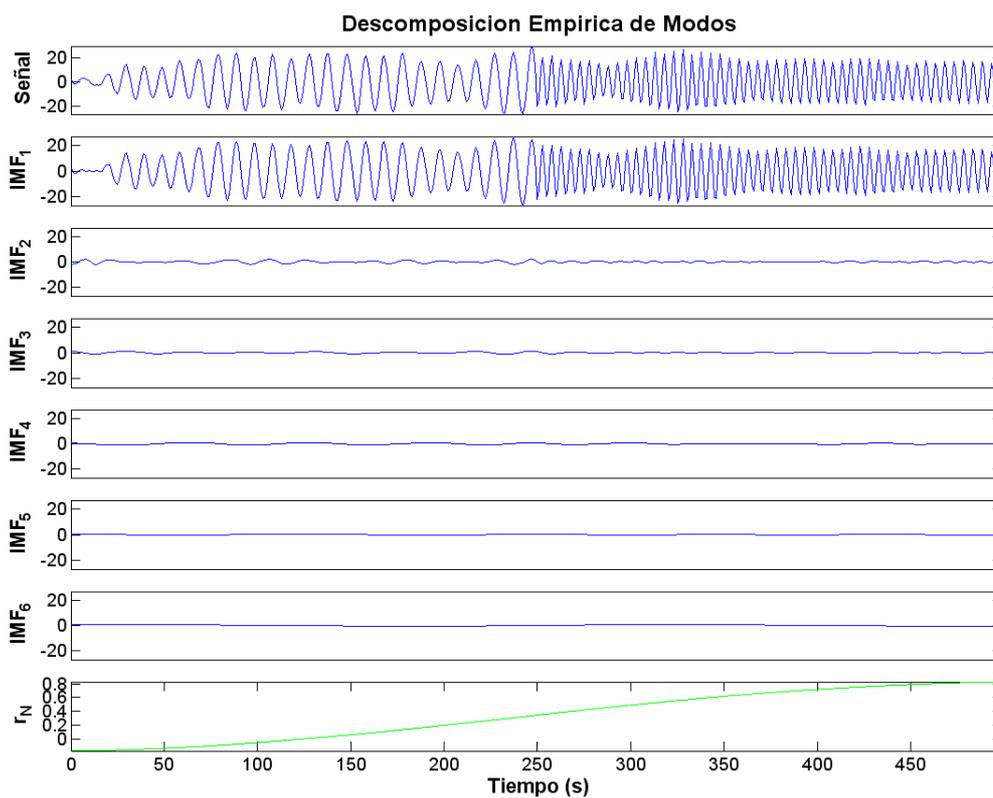


Fig. VI.37. Descomposición empírica de modos de la señal del proceso TVAR. Se presentan la señal original, los modos de oscilación intrínsecos (IMFs) y la señal de residuo (r_N) resultante.

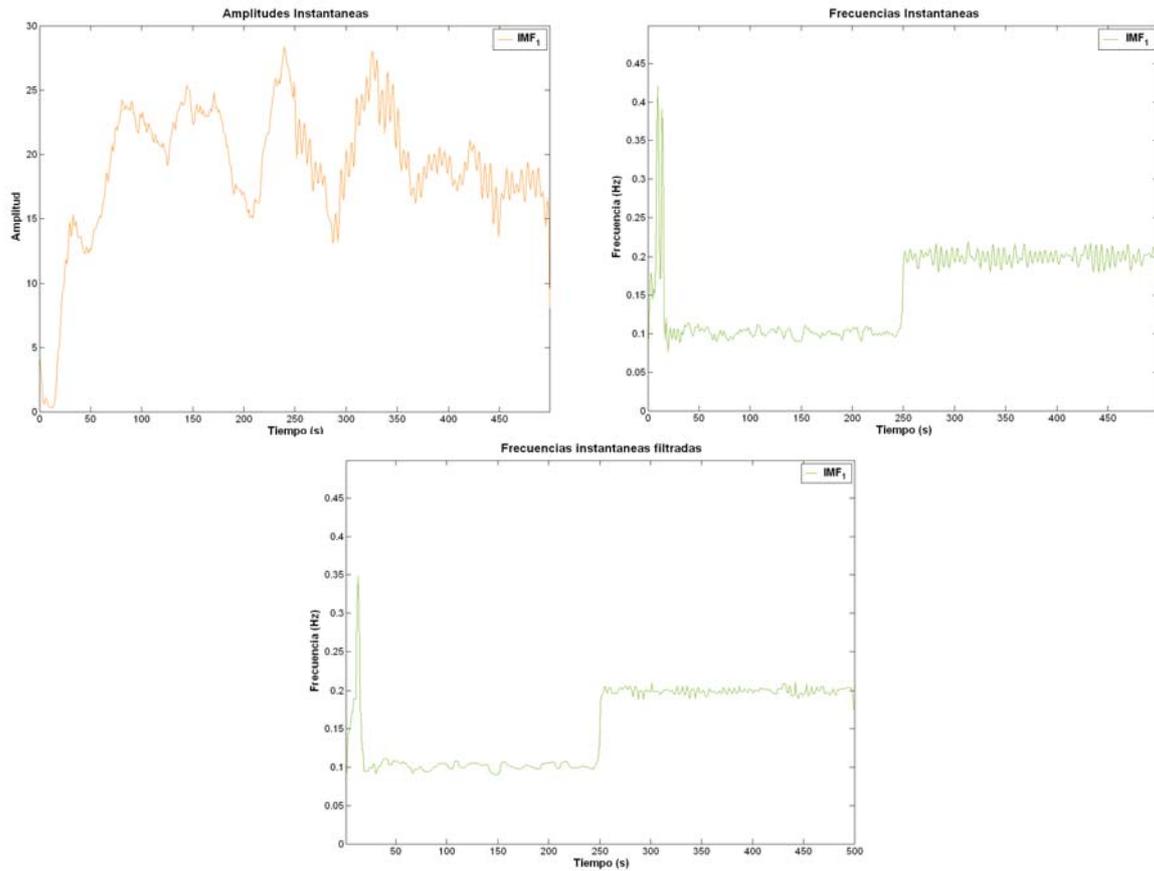


Fig. VI.38. Señal analítica del primer modo de oscilación intrínseco. a) Amplitud instantánea. b) Frecuencia instantánea. c) Frecuencia instantánea filtrada con un filtro de mediana.

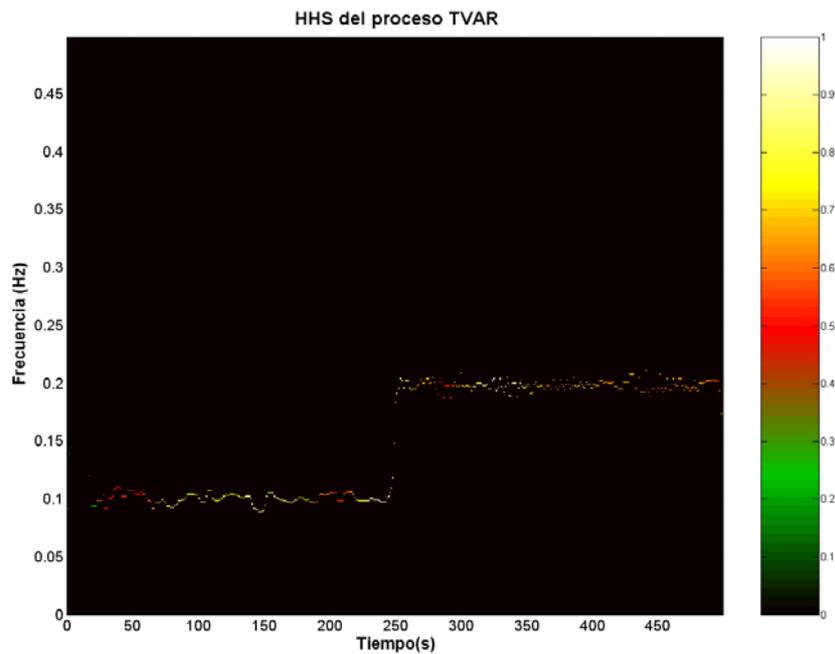


Fig. VI.39. Espectro de Hilbert-Huang para el proceso TVAR.

▪ **Comparación de las técnicas tiempo-frecuencia.**

En la figura VI.40 se muestran las representaciones seleccionadas para el proceso TVAR de acuerdo al criterio de correlación local mientras que en la figura VI.41 se presenta la estimación de la frecuencia instantánea de la señal a partir del centroide de cada representación.

En la tabla VI.II se presentan los valores de los índices de desempeño obtenidos para cada una de las representaciones tiempo-frecuencia mostrados en la figura VI.40.

TABLA VI.II. INDICES DE DESEMPEÑO DE LAS REPRESENTACIONES TIEMPO-FRECUENCIA SELECCIONADAS PARA EL PROCESO TVAR.

	ρ	ρ_{promedio}	NRMSE	ρ_{FI}	res _{TF}
SP	0.4975	0.8002	0.0781	0.9970	0.0057
WVD	0.5676	0.6110	0.2203	0.9755	0.0046
CWD	0.5887	0.7549	0.0782	0.9970	0.0059
SPR	0.6205	0.8387	0.0625	0.9982	0.0006
P_{BURG}	0.5372	0.7972	0.0802	0.9969	0.0026
P_{RLS}	0.5290	0.7918	0.1211	0.9929	0.0042
P_{RLS-VFF}	0.5416	0.8173	0.1023	0.9948	0.0043
P_{KALMAN}	0.5244	0.7909	0.1009	0.9949	0.0040
SC	0.4807	0.7913	0.0830	0.9967	0.0061
HHS	0.3966	0.7919	0.1294	0.9916	2.98x10 ⁻⁶

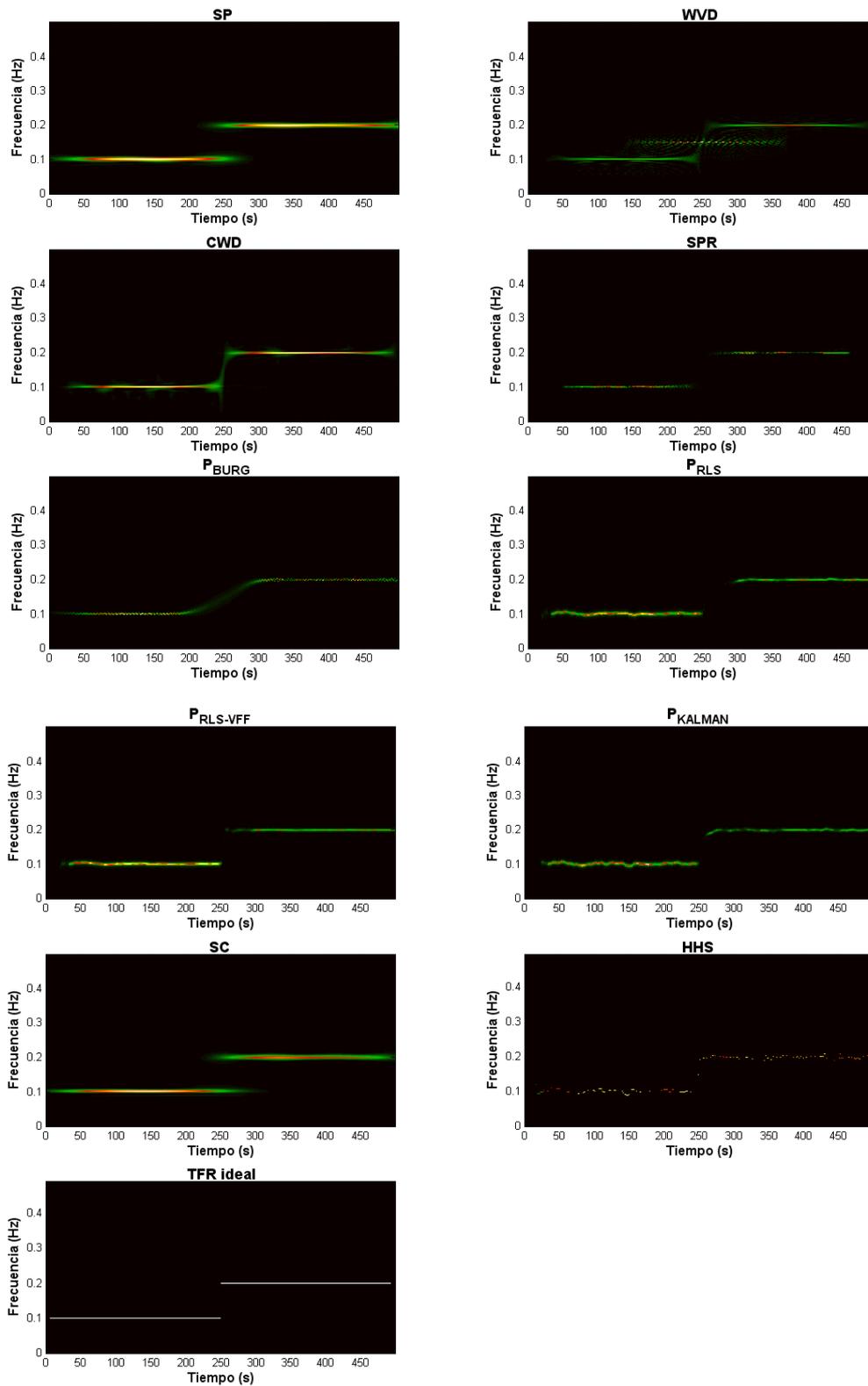


Fig. VI.40. Representaciones tiempo-frecuencia seleccionadas para cada técnica de análisis tiempo-frecuencia así como la representación tiempo-frecuencia teórica del proceso TVAR.

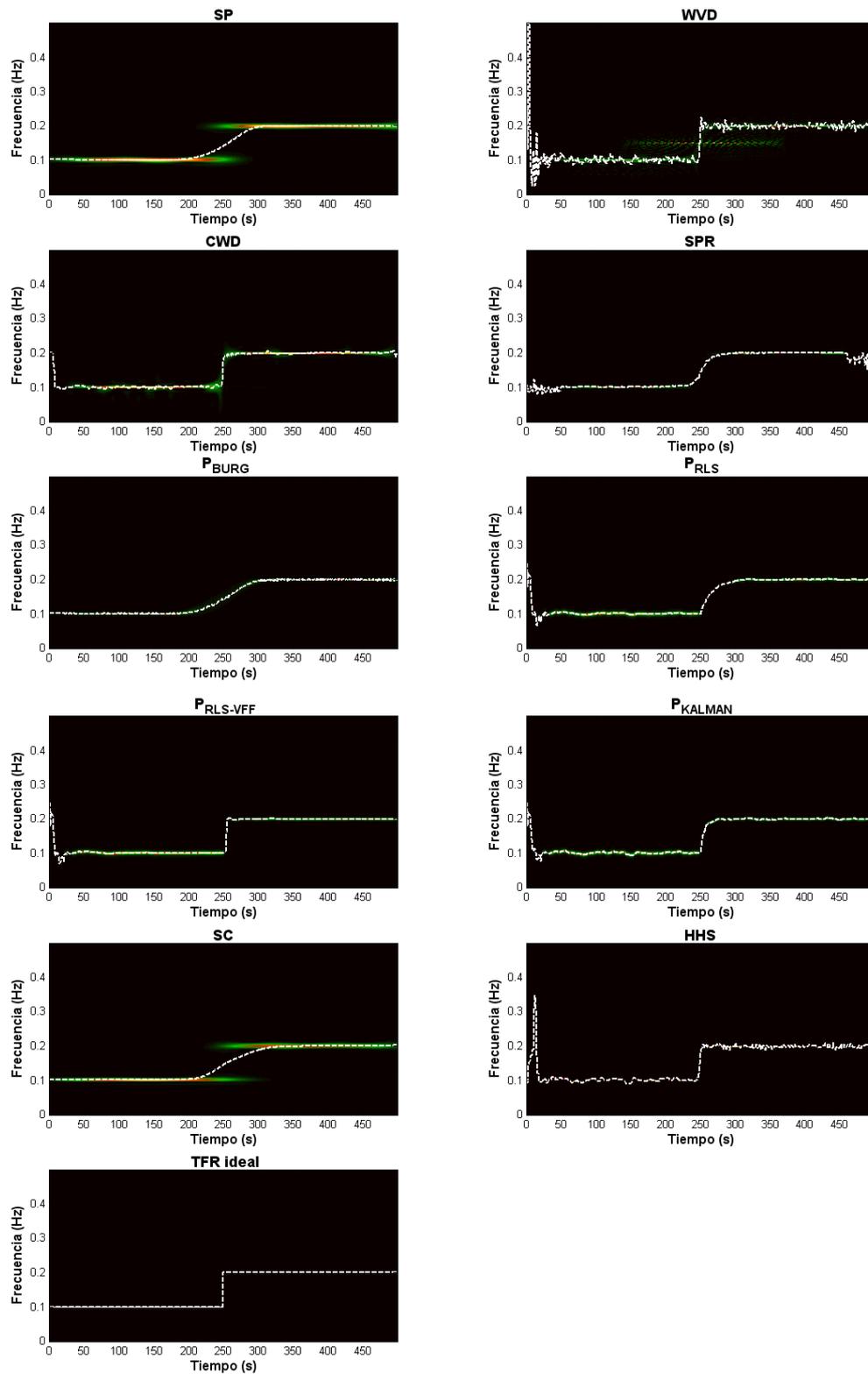


Fig. VI.41. Estimación de la frecuencia instantánea mediante el primer momento de cada una de las representaciones tiempo-frecuencia seleccionadas y teórica del proceso TVAR.

VI.3. Señal no biológica concatenando sinusoidales puras.

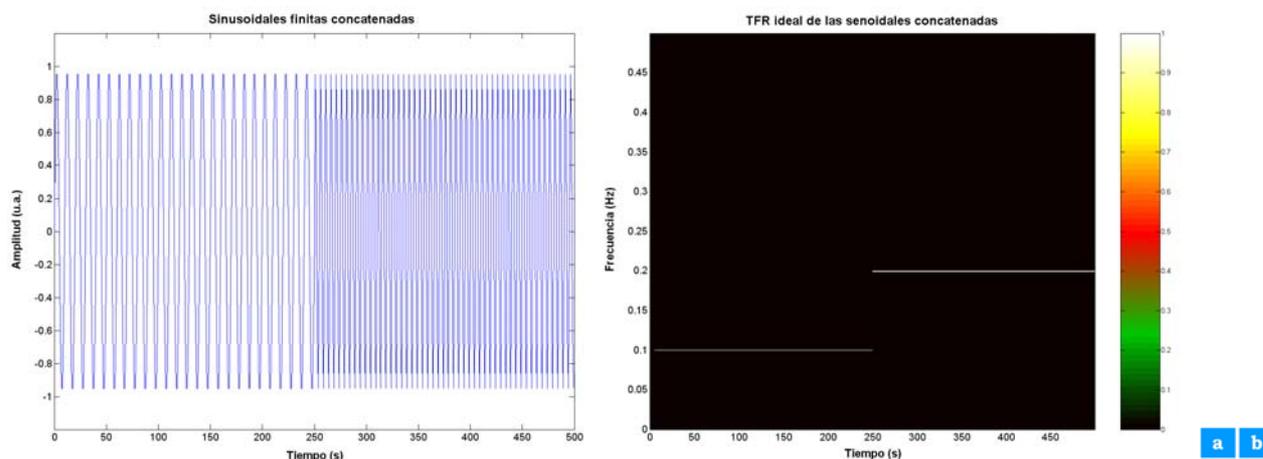


Fig. VI.42. Señal conformada por la concatenación de sinusoidales de duración finita. a) Forma de onda, b) Representación tiempo-frecuencia teórica.

En forma similar al caso de señal no biológica simulada mediante el modelo TVAR, se obtuvieron todas las representaciones tiempo-frecuencia utilizadas en la presente tesis, se obtuvieron las sub-imágenes de correlación local, se realizó el cálculo de la correlación promedio y se seleccionaron las mejores representaciones tiempo-frecuencia.

Para el caso del espectrograma, la mejor representación tiempo-frecuencia fue obtenida con una ventana Hamming de 251 s de longitud. En el Anexo I se presentan la función de ambigüedad de la señal, cada espectrograma calculado, la comparación local de cada espectrograma, el espectrograma seleccionado para la señal, y se incluyen los valores de correlación promedio en una tabla.

Para la distribución de Wigner-Ville el valor de $\rho_{promedio}$ obtenido es igual a 0.7673. La comparación local con la representación teórica se presenta en el Anexo I.

Para la distribución Choi-Williams la mejor representación tiempo-frecuencia se obtuvo con una $\sigma = 467$. En el Anexo I se presentan las distribuciones Choi-Williams para diferentes valores de σ , así como su comparación local con la representación tiempo-frecuencia ideal. Además, en el Anexo se presenta la tabla con los valores de $\rho_{promedio}$ de cada distribución, y la distribución Choi-Williams seleccionada para esta señal.

Para el espectrograma reasignado la mejor representación tiempo-frecuencia fue el obtenido con una ventana Blackman-Harris 51 s de longitud. En el Anexo I se presentan los espectrogramas reasignados obtenidos, los valores de correlación obtenidos mediante su

comparación local con la representación tiempo-frecuencia ideal de la señal, la tabla con los valores de correlación promedio, y el espectrograma reasignado seleccionado.

Para la representación tiempo-frecuencia por el método de Burg, la mejor representación con base en el criterio de la máxima correlación promedio se obtuvo con un modelo AR de orden 2 y una ventana Hamming de 201 s de longitud. En el Anexo I se presentan las representaciones tiempo-frecuencia obtenidas vía el método de Burg variando la longitud de las ventanas Hamming y Blackman-Harris. Además, se incluye la comparación local entre cada una de estas representaciones de Burg y la representación tiempo-frecuencia teórica, una tabla con los valores de $\rho_{promedio}$ obtenidos, y la representación tiempo-frecuencia de Burg seleccionada para esta señal.

En el caso de la representación tiempo-frecuencia vía el algoritmo RLS, la mejor representación correspondió al factor de olvido de 0.75, con un modelo AR de segundo orden. En el Anexo I se presenta cada representación tiempo-frecuencia P_{RLS} obtenida al variar el factor de olvido, así como su comparación local con la representación teórica de las sinusoidales concatenadas. Además, se presenta la tabla con los valores de correlación promedio correspondientes, y la representación tiempo-frecuencia P_{RLS} seleccionada.

Para el algoritmo RLS-VFF la mejor representación tiempo-frecuencia se obtuvo con $\sigma_e^2 = 0.0001$, para un modelo AR de segundo orden. En el Anexo I se presentan las representaciones tiempo-frecuencia calculadas con esta técnica, la comparación local entre dichas representaciones y la representación tiempo-frecuencia teórica, la tabla con los valores de correlación promedio correspondientes y la representación $P_{RLS-VFF}$ seleccionada.

En el caso de la representación tiempo-frecuencia mediante el filtro de Kalman, la mejor representación correspondió a la varianza del ruido del proceso igual a $q = 5 \times 10^{-1}$ para un modelo AR de segundo orden con varianza del ruido de medición (σ^2) igual a 0.1. En el Anexo I se presenta cada representación P_{Kalman} junto con su comparación local con la representación tiempo-frecuencia teórica. Además, se incluye la tabla con los valores de correlación correspondientes y la representación tiempo-frecuencia P_{Kalman} seleccionada para esta señal.

Para el escalograma, la Wavelet Morlet de 70 s de longitud en la escala más gruesa originó la mejor representación. En el Anexo I se presentan los escalogramas obtenidos al variar la longitud de la ventana de Morlet, su comparación local con la representación teórica, la tabla con los valores de correlación promedio, y el escalograma seleccionado.

Finalmente, para el espectro de Hilbert-Huang el valor de $\rho_{promedio}$ es igual a 0.8892. En el Anexo I se presenta la descomposición empírica de modos de la señal, la amplitud y frecuencia instantáneas de su modo de oscilación intrínseco. Además, se incluye el Espectro de Hilbert-Huang de las sinusoidales concatenadas, la comparación local, y una versión suavizada de dicho espectro.

▪ **Comparación entre las representación tiempo-frecuencia.**

Las representaciones tiempo-frecuencia seleccionadas para las sinusoidales concatenadas de acuerdo al criterio de correlación local se presentan en la figura VI.43. La estimación de la frecuencia instantánea de la señal a partir del centroide de cada una de estas representaciones se presenta en la figura VI.44, mientras que en la tabla VI.III se presentan los valores de los índices de desempeño obtenidos para cada una de las representaciones tiempo-frecuencia mostrados en la figura VI.43.

TABLA VI.III. INDICES DE DESEMPEÑO DE LAS REPRESENTACIONES TIEMPO-FRECUENCIA SELECCIONADAS PARA LAS SINUSOIDALES CONCATENADAS.

	ρ	$\rho_{promedio}$	NRMSE	ρ_{FI}	res_{TF}
SP	0.5592	0.8545	0.0859	0.9963	0.0064
WVD	0.6031	0.7673	0.0549	0.9985	0.0057
CWD	0.6377	0.8554	0.0257	0.9997	0.0060
SPR	0.6232	0.8685	0.0288	0.9996	3.89×10^{-5}
P_{BURG}	0.5738	0.8157	0.0769	0.9971	0.0021
P_{RLS}	0.5408	0.8685	0.0782	0.9969	0.0053
P_{RLS-VFF}	0.5414	0.8704	0.0886	0.9961	0.0040
P_{KALMAN}	0.5425	0.8700	0.0734	0.9973	0.0040
SC	0.5194	0.8417	0.0736	0.9973	0.0068
HHS	0.6597	0.8892	0.0180	0.9998	3.79×10^{-6}

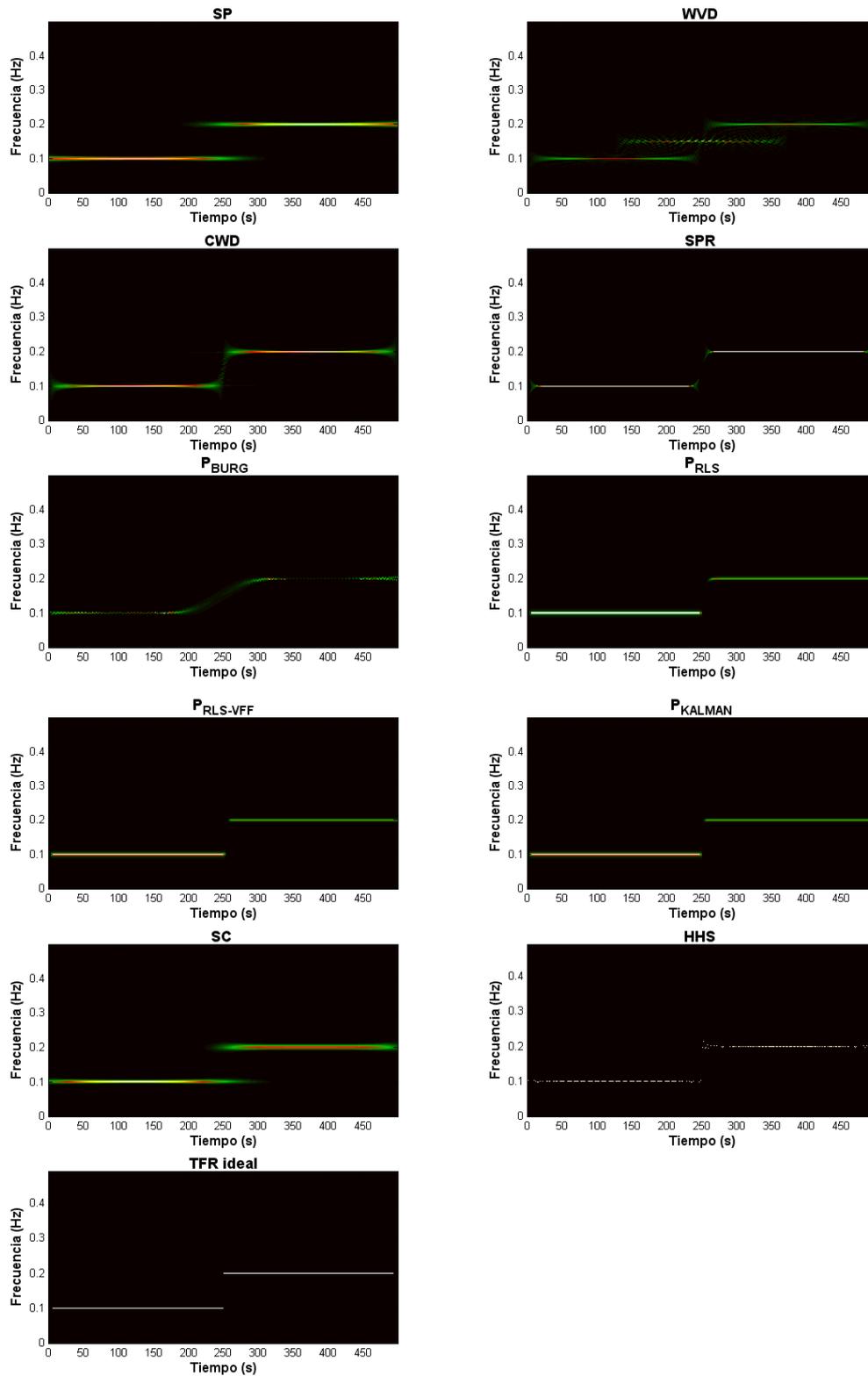


Fig. VI.43. Representaciones tiempo-frecuencia seleccionadas para cada técnica de análisis tiempo-frecuencia así como la representación tiempo-frecuencia teórica de las sinusoidales concatenadas.

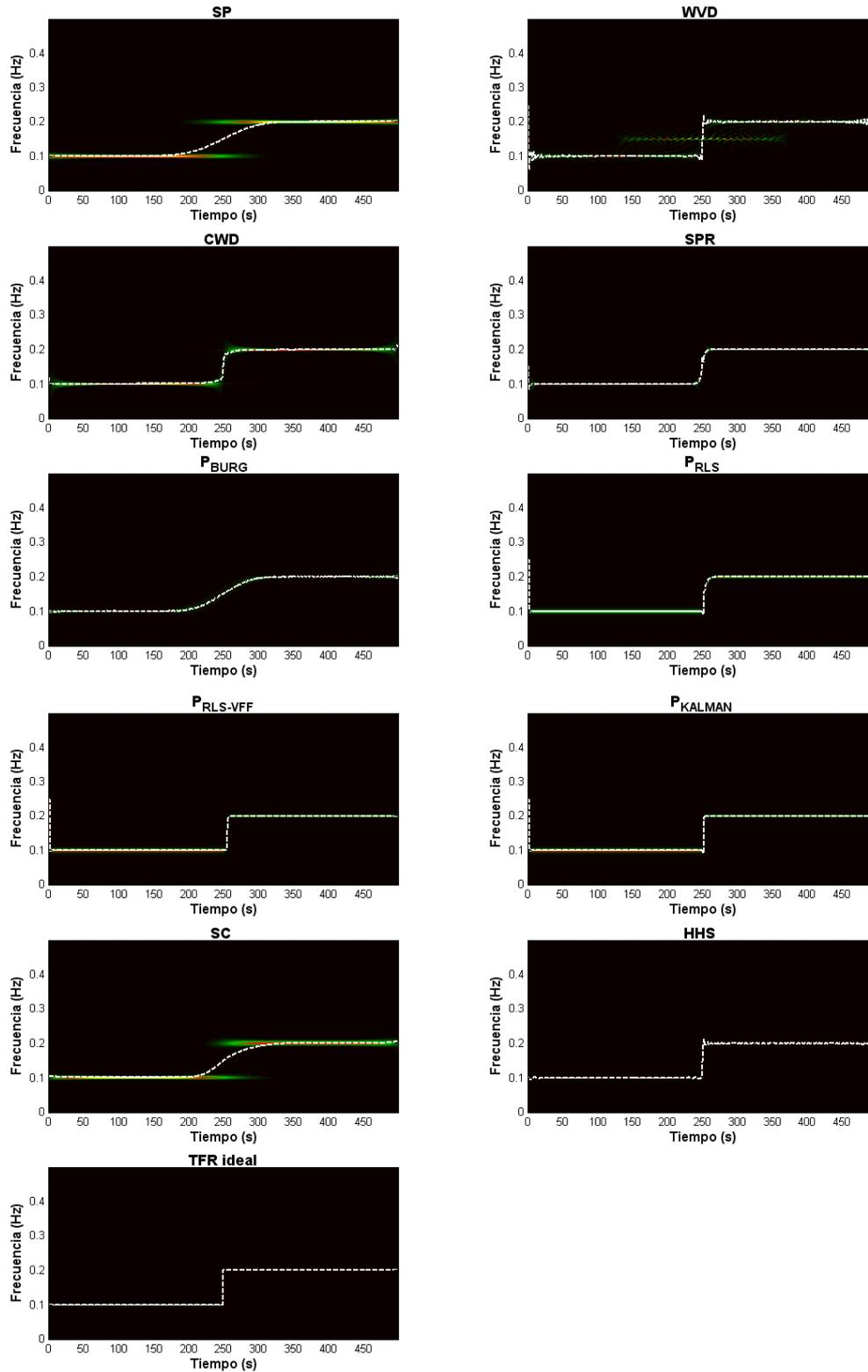


Fig. VI.44. Estimación de la frecuencia instantánea mediante el primer momento de cada representación tiempo-frecuencia seleccionada así como teórica de las sinusoidales concatenadas.

VI.4. Primer sonido cardiaco y sus componentes.

VI.4.1. Componente muscular simulado (s_m) de S1.

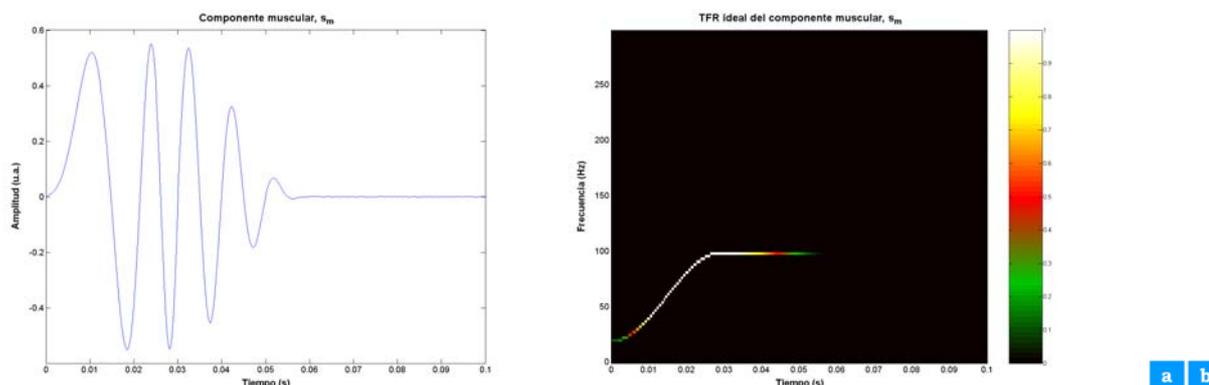


Fig. VI.45. Componente muscular de S1. a) Forma de onda. b) Representación tiempo-frecuencia teórica.

En forma similar al caso de las dos señales no biológicas simuladas, se obtuvieron todas las representaciones tiempo-frecuencia utilizadas en la presente tesis, se obtuvieron las subimágenes de correlación local, se realizó el cálculo de la correlación promedio y se seleccionaron las mejores representaciones tiempo-frecuencia.

Para el caso del espectrograma, la mejor representación tiempo-frecuencia fue obtenida con una ventana Blackman-Harris 25 ms de longitud. En el Anexo I se presentan los espectrogramas calculados, la tabla con los resultados de comparación local estos espectrogramas con la representación tiempo-frecuencia teórica, y el espectrograma seleccionado para el componente muscular simulado de S1.

Para la distribución Wigner-Ville de esta señal, el valor de $\rho_{promedio}$ obtenido es igual a 0.7745. La distribución Wigner-Ville del componente muscular simulado de S1 se presenta en el anexo I.

Para la distribución Choi-Williams del componente muscular simulado, la mejor distribución se obtuvo con una $\sigma = 428$. En el Anexo I se presentan las distribuciones Choi-Williams para diferentes valores de σ , los resultados de su comparación local con la representación tiempo-frecuencia ideal, y la distribución Choi-Williams seleccionada.

Para el espectrograma reasignado, la mejor representación tiempo-frecuencia del componente muscular simulado fue el obtenida con una ventana Hamming de 20.2 ms de longitud. En el Anexo I se presentan los espectrogramas reasignados obtenidos, la tabla con

los valores de correlación promedio resultantes de la comparación local con la representación tiempo-frecuencia ideal, y el espectrograma reasignado seleccionado para esa señal.

Para la representación tiempo-frecuencia obtenida vía el método de Burg, la mejor representación del componente muscular simulado se obtuvo con un modelo AR de orden 2 y una ventana Blackman-Harris de 45 ms de longitud. En el Anexo I se presentan las representaciones tiempo-frecuencia de Burg obtenidas para diversos ordenes del modelo AR y funciones ventana. Además, se incluye la tabla con los resultados de la comparación local de estas representaciones con la representación tiempo-frecuencia teórica, y la representación tiempo-frecuencia de Burg seleccionada para el componente muscular simulado.

En el caso de la representación tiempo-frecuencia vía el algoritmo RLS, la mejor representación correspondió a un modelo AR de orden 2 y un factor de olvido de 0.90625. En el Anexo I se presenta cada representación tiempo-frecuencia P_{RLS} obtenida al variar el orden del modelo AR y el factor de olvido, los resultados de su comparación local con la representación teórica de la señal, y la representación tiempo-frecuencia P_{RLS} seleccionada para el componente muscular simulado.

Para el algoritmo RLS-VFF la mejor representación tiempo-frecuencia se obtuvo con una varianza del ruido esperado $\sigma_e^2 = 1 \times 10^{-11}$, para un modelo AR de segundo orden. En el Anexo I se presentan las representaciones tiempo-frecuencia calculadas con esta técnica para esta señal simulada, la tabla con los valores de correlación promedio resultantes de su comparación local, y la representación $P_{RLS-VFF}$ seleccionada para el componente muscular simulado de S1.

En el caso de la representación tiempo-frecuencia mediante el filtro de Kalman, la mejor representación de esta señal correspondió a la varianza del ruido del proceso igual a $q = 0.17783$, para un modelo AR de orden 3 y una varianza del ruido de medición (σ^2) unitaria. En el Anexo I se presenta cada representación P_{Kalman} obtenida para esta señal, la tabla con los resultados de su comparación local con la representación tiempo-frecuencia teórica, y la representación tiempo-frecuencia P_{Kalman} seleccionada para el componente muscular simulado de S1.

Para el escalograma, la Wavelet Morlet de 1.2 ms de longitud en la escala más gruesa originó la mejor representación tiempo-frecuencia. En el Anexo I se presentan los escalogramas obtenidos al variar la longitud de la ventana de Morlet, la tabla con los resultados de su comparación local con la representación teórica, y el escalograma seleccionado.

Para el espectro de Hilbert-Huang, su comparación con la representación teórica de la señal produjo un valor de $\rho_{promedio}$ igual a 0.8459. En el Anexo I se presenta la descomposición empírica de modos de la señal, la amplitud y frecuencia instantáneas de su primer modo de oscilación intrínseco, el Espectro de Hilbert-Huang del componente muscular simulado de S1, y una versión suavizada de dicho espectro.

▪ **Comparación de las técnicas tiempo-frecuencia.**

En la figura VI.46 se muestran las representaciones tiempo-frecuencia seleccionadas para el componente muscular de S1 simulado, de acuerdo al criterio de correlación local. En la figura VI.47 se presenta la estimación de la frecuencia instantánea de la señal a partir del centroide de cada una de estas representaciones. En la tabla VI.IV se presentan los valores de los índices de desempeño obtenidos para cada una de las representaciones tiempo-frecuencia mostradas en la figura VI.46.

TABLA VI.IV. INDICES DE DESEMPEÑO DE LAS REPRESENTACIONES TIEMPO-FRECUENCIA SELECCIONADAS PARA s_m .

	ρ	$\rho_{promedio}$	NRMSE	ρ_{FI}	res_{TF}
SP	0.4264	0.8144	0.4264	0.9152	0.0178
WVD	0.5826	0.7745	0.5826	0.0036	0.0032
CWD	0.5254	0.8110	0.5254	0.8151	0.0039
SPR	0.6989	0.8454	0.6989	0.7247	5.22×10^{-5}
P_{BURG}	0.6864	0.8459	0.6864	0.9594	0.0015
P_{RLS}	0.5819	0.8371	0.5819	0.8607	0.0024
P_{RLS-VFF}	0.6357	0.8379	0.6649	0.9721	0.0022
P_{KALMAN}	0.6649	0.8197	0.7101	0.4239	0.0052
SC	0.4168	0.7702	0.4168	0.8449	0.0778
HHS	0.7101	0.8459	0.6357	0.9268	6.74×10^{-7}

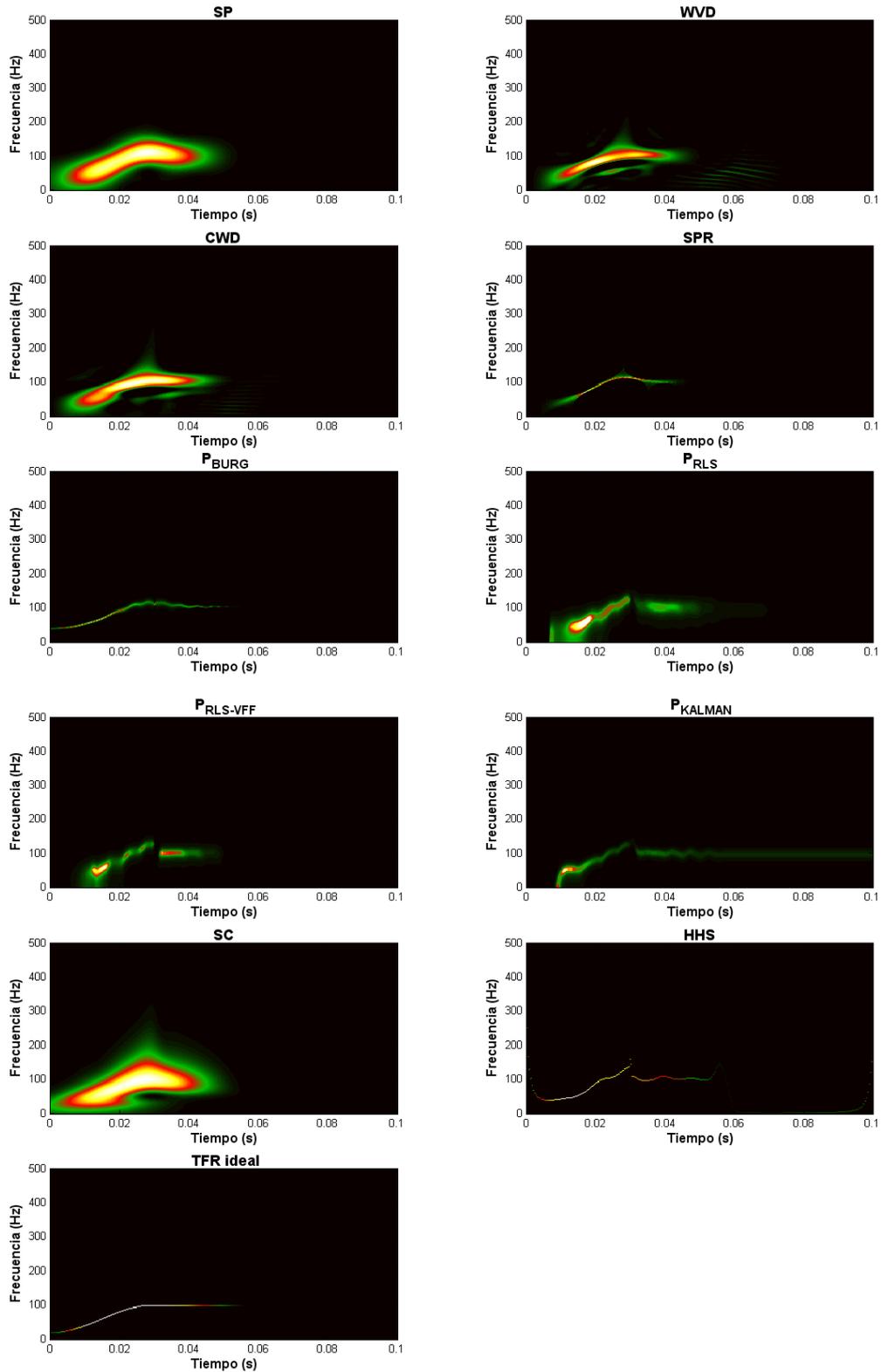


Fig. VI.46. Representaciones tiempo-frecuencia seleccionadas para cada técnica de análisis tiempo-frecuencia así como la representación tiempo-frecuencia teórica del componente muscular simulado de S1.

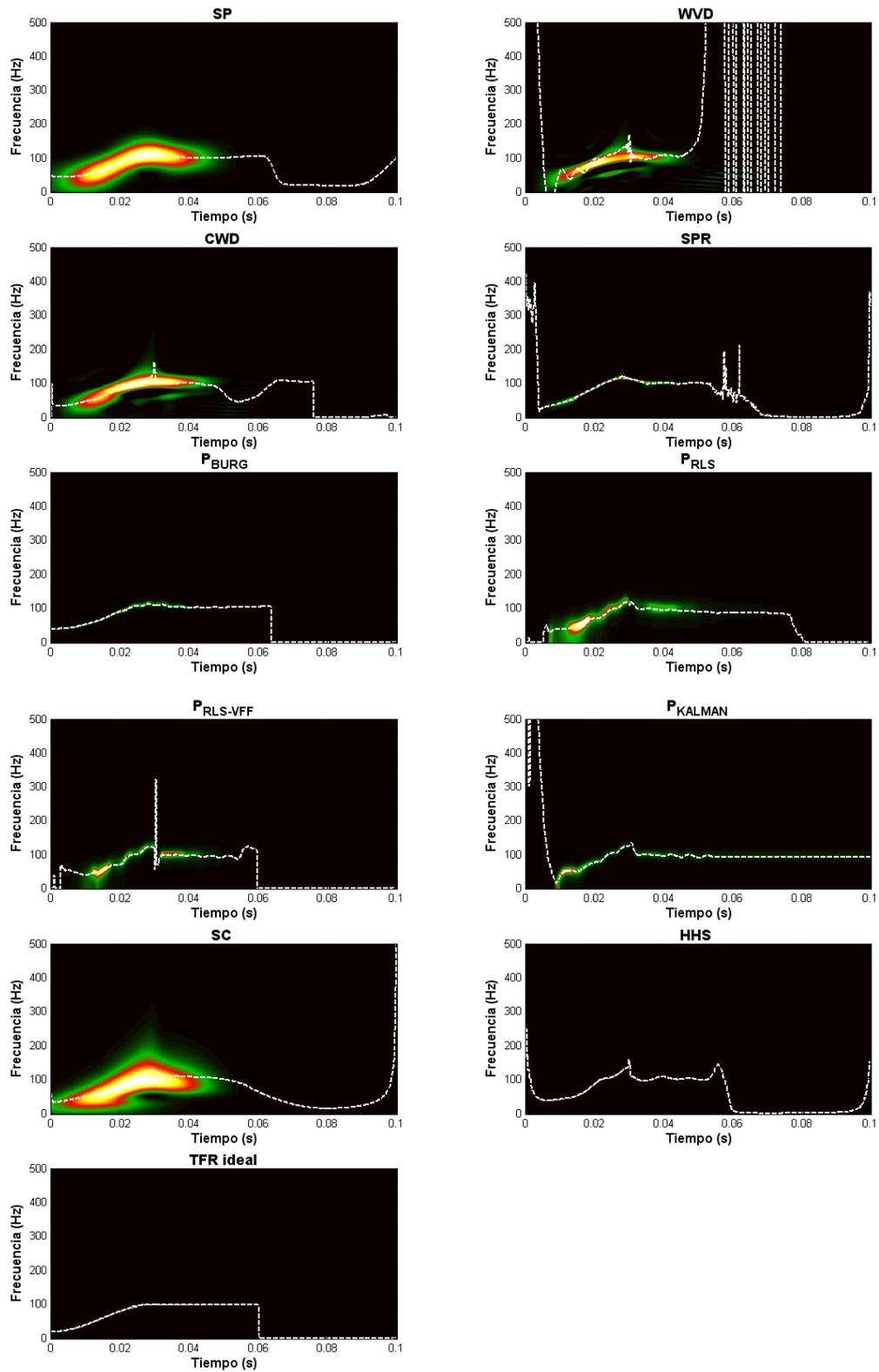


Fig. VI.47. Estimación de la frecuencia instantánea mediante el primer momento de cada representación tiempo-frecuencia seleccionada y la teórica del componente muscular simulado de S1.

VI.4.2. Componente valvular simulado (s_v) de S1.

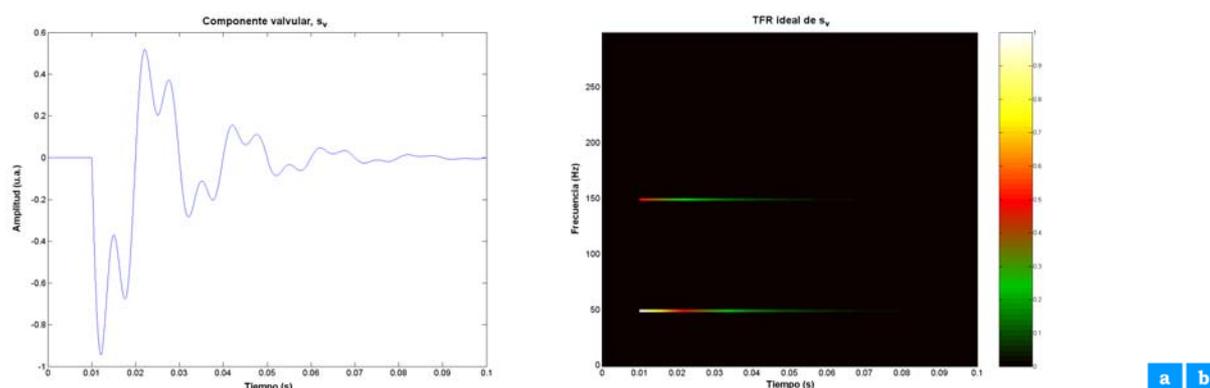


Fig. VI.48. Componente valvular de S1. a) Forma de onda. b) Representación tiempo-frecuencia teórica.

Para el componente valvular simulado del primer sonido cardiaco también se obtuvieron todas las representaciones tiempo-frecuencia utilizadas en la presente tesis, se obtuvieron las sub-imágenes de correlación local, se realizó el cálculo de la correlación promedio y se seleccionaron las mejores representaciones tiempo-frecuencia.

Para el caso del espectrograma, la mejor representación tiempo-frecuencia fue obtenida con una ventana Rectangular 35 ms de longitud. En el Anexo I se presentan los espectrogramas calculados, la tabla con los resultados de comparación local estos espectrogramas con la representación tiempo-frecuencia teórica, y el espectrograma seleccionado para el componente valvular simulado de S1.

Para la distribución Wigner-Ville de esta señal, el valor de $\rho_{promedio}$ obtenido es igual a 0.7177. La distribución Wigner-Ville del componente valvular simulado de S1 se presenta en el anexo I.

Para la distribución Choi-Williams, la mejor distribución se obtuvo con una $\sigma = 428$. En el Anexo I se presentan las distribuciones Choi-Williams obtenidas con diferentes valores de σ , los resultados de su comparación local con la representación tiempo-frecuencia ideal, y la distribución Choi-Williams seleccionada para el componente valvular simulado.

Para el espectrograma reasignado, la mejor representación tiempo-frecuencia fue el obtenida con una ventana Blackman-Harris de 45 ms de longitud. En el Anexo I se presentan los espectrogramas reasignados obtenidos, la tabla con los valores de correlación promedio resultantes de la comparación local con la representación tiempo-frecuencia ideal, y el espectrograma reasignado seleccionado para el del componente valvular simulado de S1.

Para la representación tiempo-frecuencia obtenida vía el método de Burg, la mejor representación del componente valvular simulado se obtuvo con un modelo AR de orden 5 y una ventana Blackman-Harris de 45 ms de longitud. En el Anexo I se presentan las representaciones tiempo-frecuencia de Burg obtenidas para diversos ordenes del modelo AR y funciones ventana. Además, se incluye la tabla con los resultados de la comparación local de estas representaciones tiempo-frecuencia, y la representación P_{Burg} seleccionada para el componente valvular simulado.

En el caso de la representación tiempo-frecuencia vía el algoritmo RLS, la mejor representación se obtuvo con un modelo AR de orden 5 y un factor de olvido igual a 0.84375. En el Anexo I se presenta cada representación tiempo-frecuencia P_{RLS} obtenida al variar el orden del modelo AR y el factor de olvido, los resultados de su comparación local con la representación teórica de la señal, y la representación tiempo-frecuencia P_{RLS} seleccionada para el componente valvular simulado de S1. Además, se incluyen representaciones tiempo-frecuencia P_{RLS} obtenidas para modelos AR de órdenes 6 al 15 y diversos factores de olvido.

Para el algoritmo RLS-VFF la mejor representación tiempo-frecuencia se obtuvo con una varianza del ruido esperado $\sigma_e^2 = 1 \times 10^{-10}$, para un modelo AR de orden 5. En el Anexo I se presentan las representaciones tiempo-frecuencia calculadas con esta técnica para el componente valvular, la tabla con los valores de correlación promedio resultantes de su comparación local, y la representación $P_{\text{RLS-VFF}}$ seleccionada para esta señal. Al igual que para la técnica RLS estándar, se incluyen representaciones $P_{\text{RLS-VFF}}$ obtenidas con modelos AR de órdenes 6 al 15 y diversas varianzas del ruido esperado.

En el caso de la representación tiempo-frecuencia mediante el filtro de Kalman, la mejor representación del componente valvular simulado correspondió a la varianza del ruido del proceso igual a $q = 1 \times 10^{-3}$, para un modelo AR de orden 5 y una varianza del ruido de medición (σ^2) unitaria. En el Anexo I se presenta cada representación P_{Kalman} obtenida para esta señal, la tabla con los resultados de su comparación local con la representación tiempo-frecuencia teórica, y la representación tiempo-frecuencia P_{Kalman} seleccionada para el componente valvular del primer sonido cardiaco simulado.

Para el escalograma, la Wavelet Morlet de 1.8 ms de longitud en la escala más gruesa originó la mejor representación tiempo-frecuencia. En el Anexo I se presentan los escalogramas obtenidos al variar la longitud de la ventana de Morlet, la tabla con los resultados de su comparación local con la representación teórica, y el escalograma seleccionado para el componente valvular simulado.

Para el espectro de Hilbert-Huang del componente valvular, la comparación local con la representación teórica de la señal produjo un valor de $\rho_{promedio}$ igual a 0.8415. En el Anexo I se presenta la descomposición empírica de modos del componente valvular de S1 simulado, la magnitud y la frecuencia instantáneas de los dos primeros modos de oscilación intrínsecos, el Espectro de Hilbert-Huang del componente muscular valvular de S1 simulado, y una versión suavizada de dicho espectro.

▪ **Comparación de las técnicas tiempo-frecuencia.**

En la figura VI.49 se muestran las representaciones tiempo-frecuencia seleccionadas para el componente valvular simulado de S1, de acuerdo al criterio de correlación local.

Los valores de los índices de desempeño obtenidos para cada una de las representaciones tiempo-frecuencia mostradas en la figura VI.49 se presentan en la tabla VI.V.

TABLA VI.V. INDICES DE DESEMPEÑO DE LAS REPRESENTACIONES TIEMPO-FRECUENCIA SELECCIONADAS PARA s_v .

	ρ	$\rho_{promedio}$	NRMSE	ρ_{FI}	res _{TF}
SP	0.5811	0.7776	-	-	0.0216
WVD	0.5827	0.7177	-	-	0.0022
CWD	0.6108	0.7662	-	-	0.0019
SPR	0.7880	0.8006	-	-	4.88x10 ⁻⁵
P_{BURG}	0.5662	0.7008	-	-	0.0012
P_{RLS}	0.4845	0.6989	-	-	0.0057
P_{RLS-VFF}	0.7192	0.7779	-	-	0.0271
P_{KALMAN}	0.4619	0.5970	-	-	0.1290
SC	0.6507	0.7649	-	-	0.0233
HHS	0.7828	0.8415	-	-	1.12x10 ⁻⁶

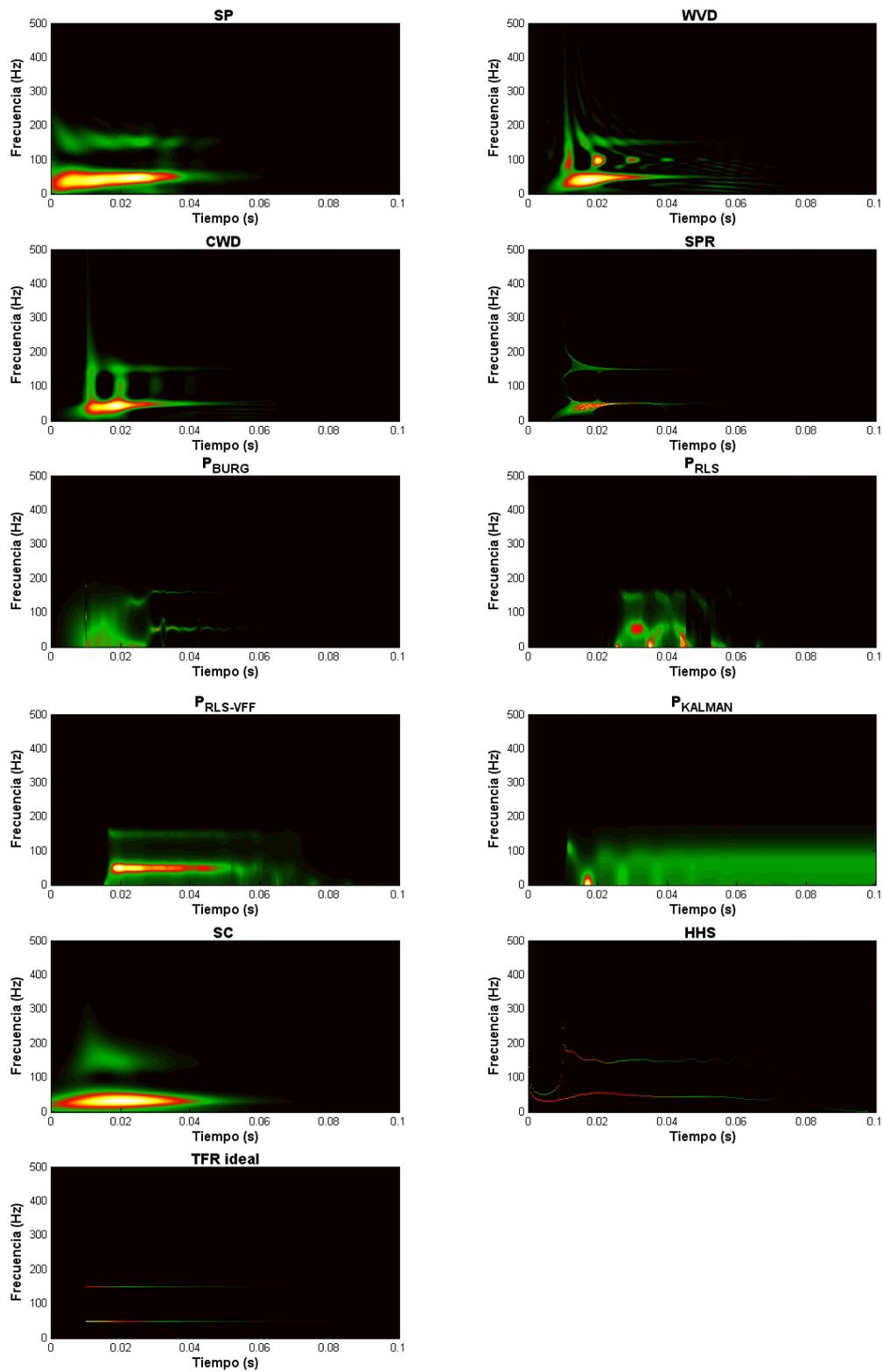


Fig. VI.49. Representaciones tiempo-frecuencia seleccionadas para cada técnica de análisis tiempo-frecuencia así como la representación tiempo-frecuencia teórica del componente valvular simulado de S1.

VI.4.3. Primer sonido cardiaco, S1.

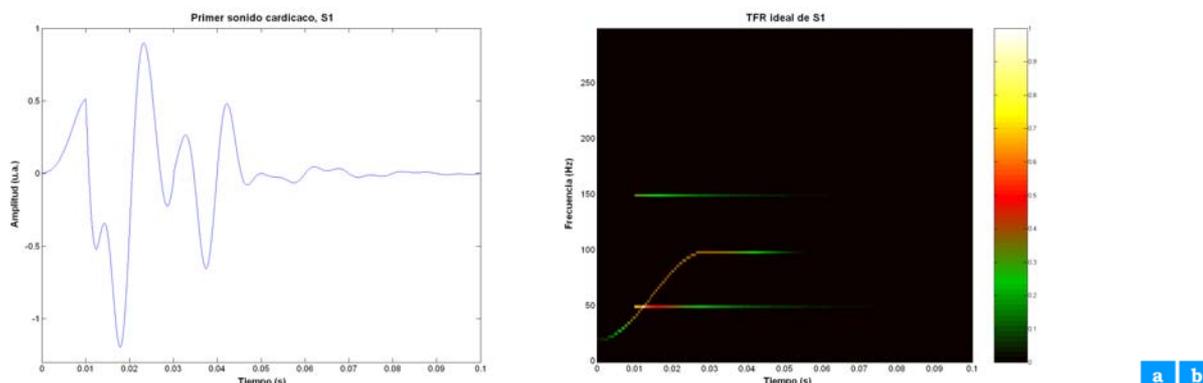


Fig. VI.50. Primer sonido cardiaco simulado (S1). a) Forma de onda. b) Representación tiempo-frecuencia teórica.

En forma similar al caso de los componentes valvular (s_v) y muscular (s_m) de S1, se obtuvieron todas las representaciones tiempo-frecuencia utilizadas en la presente tesis para el primer sonido S1 simulado, se generaron las sub-imágenes de correlación local, se realizó el cálculo de la correlación promedio y se seleccionaron las mejores representaciones tiempo-frecuencia con base en el criterio del máximo valor de correlación promedio.

Para el caso del espectrograma, la mejor representación tiempo-frecuencia fue la generada con una ventana Hamming de 50.2 ms de longitud. La función de ambigüedad de la señal, la tabla con los valores de correlación promedio correspondientes a la comparación local entre cada espectrograma generado y la representación tiempo-frecuencia teórica, y el espectrograma seleccionado para S1 simulado se presentan en el Anexo I.

Con relación a la distribución de Wigner-Ville el valor de $\rho_{promedio}$ obtenido fue igual a 0.7126. Esta representación tiempo-frecuencia se presenta en el Anexo I.

La mejor representación Choi-Williams se obtuvo con el parámetro $\sigma = 428$. Los resultados de la comparación local entre cada distribución Choi-Williams de S1 simulado y su representación tiempo-frecuencia ideal se presenta en el Anexo I, donde además se presenta la distribución Choi-Williams seleccionada para esta señal.

Para la técnica del espectrograma reasignado la mejor representación se obtuvo para la ventana Hamming de 50.2 ms de longitud. En el Anexo I se presentan los espectrogramas reasignados calculados, la tabla con sus valores de correlación promedio, y el espectrograma reasignado seleccionado para S1 simulado.

Para la representación tiempo-frecuencia obtenida mediante el modelado autorregresivo vía el algoritmo de Burg, la mejor representación se obtuvo para un modelo AR de orden 2 con una ventana Blackman-Harris de 50.2 ms de longitud. En el Anexo I se presentan las representaciones P_{Burg} de esta señal, obtenidas al variar el orden del modelo AR, el tipo y la longitud de la ventana de análisis. Además, se incluye una tabla con los resultados de la comparación local de estas representaciones tiempo-frecuencia, y la representación P_{Burg} seleccionada para S1 simulado.

Con relación a la representación tiempo-frecuencia vía el algoritmo RLS, la mejor representación se obtuvo con un modelo AR de segundo orden y un factor de olvido igual a 0.96875. En el Anexo I se presentan las representaciones P_{RLS} obtenidas para S1 simulado, la tabla con los valores de correlación promedio correspondientes a su comparación local, y la representación P_{RLS} seleccionada para esta señal. Adicionalmente, se incluyen representaciones tiempo-frecuencia P_{RLS} obtenidas con modelos AR de órdenes 6 al 15 y diversos factores de olvido.

Para el algoritmo RLS-VFF la mejor representación se obtuvo con un modelo AR de orden 5 y una varianza del ruido esperado igual a $\sigma_e^2 = 3.1623 \times 10^{-10}$. Las representaciones $P_{\text{RLS-VFF}}$ obtenidas para S1 simulado, la tabla con los valores de correlación promedio resultantes de la comparación local, y la representación tiempo-frecuencia $P_{\text{RLS-VFF}}$ seleccionada para esta señal, se presentan en el Anexo I. En el mismo anexo se incluyen representaciones $P_{\text{RLS-VFF}}$ obtenidas con modelos AR para órdenes 6 al 15 y diversas varianzas del ruido esperado.

Para la representación tiempo-frecuencia vía el filtro de Kalman, la mejor representación para S1 simulado se obtuvo con un modelo AR de orden 2, una varianza del ruido del proceso (q) igual a 1×10^{-3} , y una varianza del ruido de medición (σ^2) unitaria. En el Anexo I se presentan las representaciones P_{Kalman} de esta señal, la tabla con sus valores de correlación promedio, y la representación tiempo-frecuencia P_{Kalman} seleccionada para S2 simulado.

Con relación a la técnica del escalograma utilizando la wavelet de Morlet, la mejor representación se obtuvo para la wavelet Morlet de longitud 2.1 ms en la escala más gruesa. En el Anexo I se presenta la tabla con los valores de correlación promedio correspondientes a la comparación local entre cada representación obtenida por esta técnica y la representación tiempo-frecuencia teórica de la señal, así como el escalograma Morlet seleccionado para el primer sonido cardiaco simulado.

Para el espectro de Hilbert-Huang, se obtuvo el valor de $\rho_{promedio}$ igual a 0.7668 mediante la comparación local con la distribución teórica de la señal. En el Anexo I se presentan la descomposición empírica de modos de S1 simulado, las señales analíticas de los dos primeros modos de oscilación intrínsecos, y el espectro de Hilbert-Huang de S1 simulado que resulta de estos dos modos de oscilación. Además, se presenta una versión suavizada del espectro de Hilbert-Huang de S1 simulado.

▪ **Comparación de las técnicas tiempo-frecuencia.**

Las representaciones tiempo-frecuencia seleccionadas de acuerdo al criterio de correlación local para S1 simulado se presentan en la figura VI.51. Los valores de los índices de desempeño obtenidos para cada una de las representaciones tiempo-frecuencia mostradas en la figura VI.51 se presentan en la tabla VI.VI donde se observa que la mejor representación tiempo-frecuencia para el sonido S1 simulado, en base al criterio de compasión local, se obtiene mediante el espectro de Hilbert-Huang con una correlación promedio de 0.7668.

TABLA VI.VI. INDICES DE DESEMPEÑO DE LAS REPRESENTACIONES TIEMPO-FRECUENCIA SELECCIONADAS PARA S1.

	ρ	$\rho_{promedio}$	NRMSE	ρ_{FI}	res _{TF}
SP	0.4303	0.7450	-	-	0.0037
WVD	0.4270	0.7126	-	-	0.0020
CWD	0.4417	0.7505	-	-	0.0017
SPR	0.6545	0.7750	-	-	0.0002
P_{BURG}	0.2801	0.6799	-	-	0.0007
P_{RLS}	0.3203	0.6191	-	-	0.0108
P_{RLS-VFF}	0.3649	0.6377	-	-	0.0014
P_{KALMAN}	0.2503	0.5052	-	-	0.0110
SC	0.3999	0.7220	-	-	0.0063
HHS	0.5321	0.7668	-	-	5.75x10 ⁻⁷

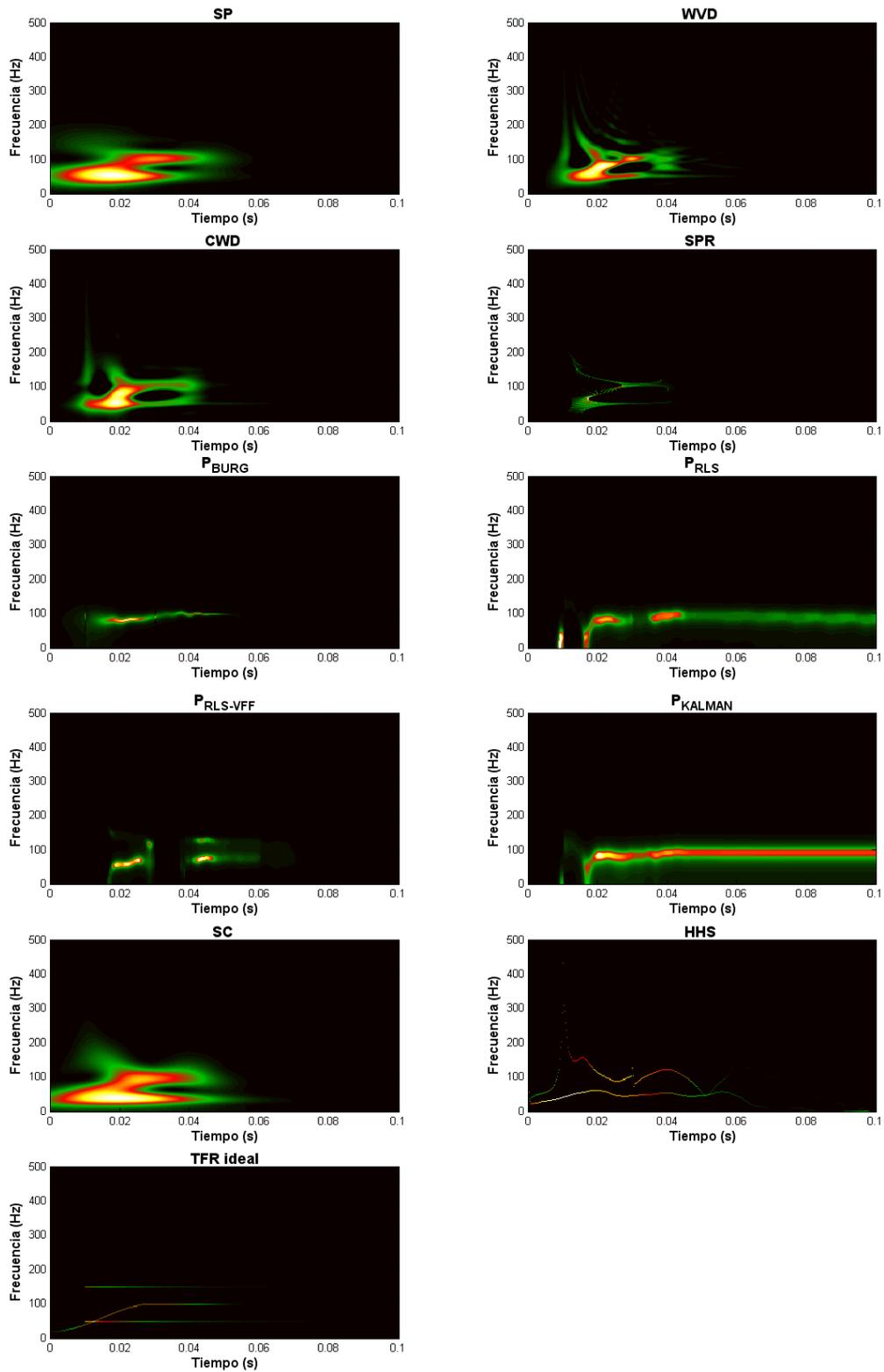


Fig. VI.51. Representaciones tiempo-frecuencia seleccionadas para cada técnica de análisis así como la representación tiempo-frecuencia teórica de S1 simulado.

VI.4.4. Señales de S1 reales adquiridas.

- **S1 adquirido 1.**

A continuación se contrasta los resultados obtenidos al aplicar representaciones tiempo-frecuencia basados en el espectrograma con la representación obtenida mediante el espectro de Hilbert-Huang para tres señales reales de S1. En la figura VI.52 se presenta una señal de primer sonido cardiaco adquirida de un sujeto sano. En la figura VI.53 se muestra el espectrograma de la señal del primer sonido cardiaco mientras que en la figura VI.54 se presenta la versión reasignada. Tanto para el espectrograma como para su versión reasignada se empleó una ventana Hamming de 60 ms de longitud. En la figura IV.55 se muestra la descomposición empírica de modos de la señal del sonido cardiaco mientras que en la figura VI.56 se presenta su Espectro de Hilbert-Huang.

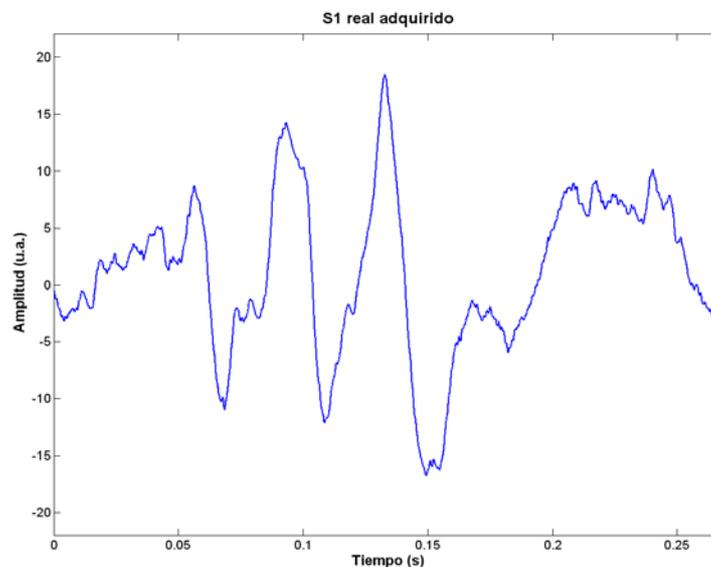


Fig. VI.52. Primer sonido cardiaco adquirido.

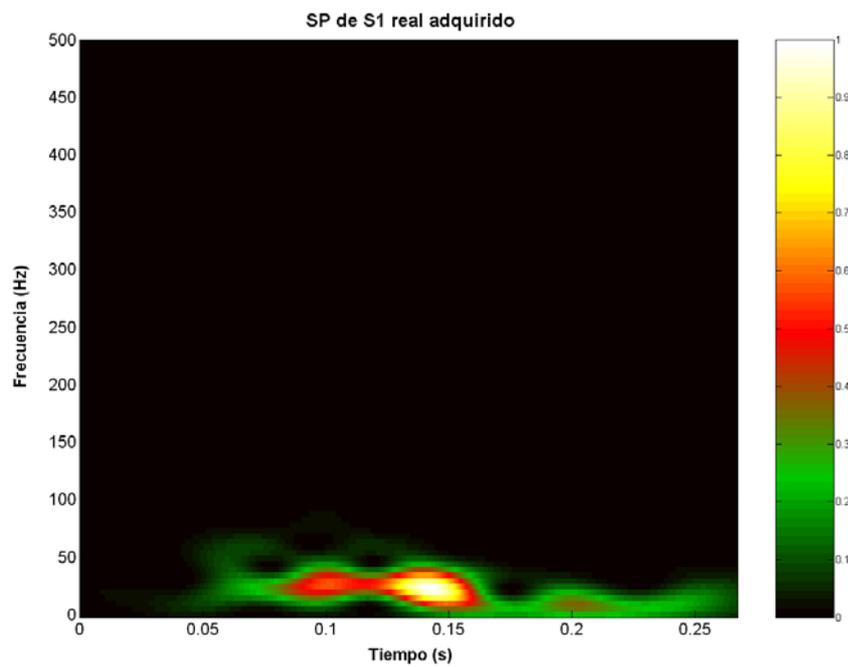


Fig. VI.53. Espectrograma de S1 adquirido mostrado en la VI.52, utilizando una ventana Hamming de 60 ms de longitud.

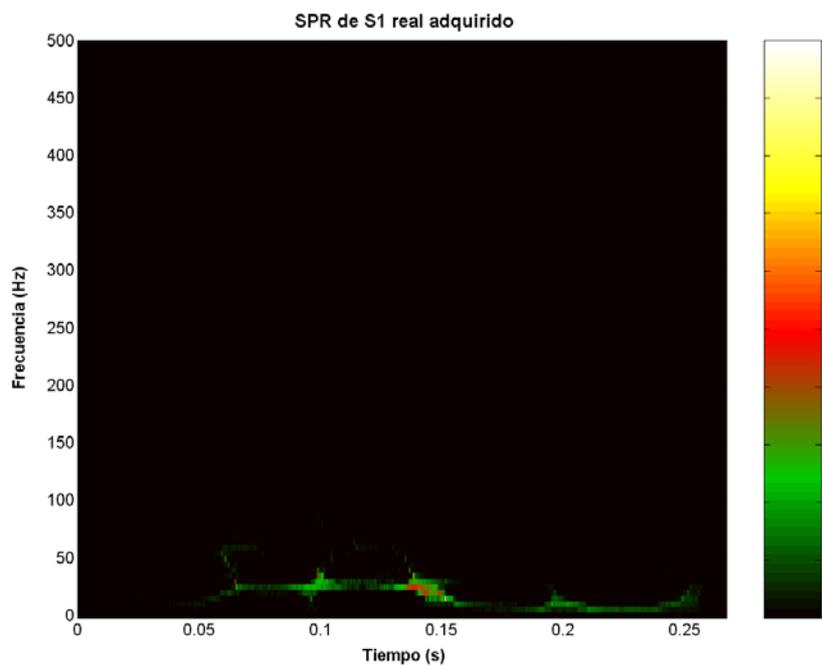


Fig. VI.54. Espectrograma reasignado de S1 adquirido.

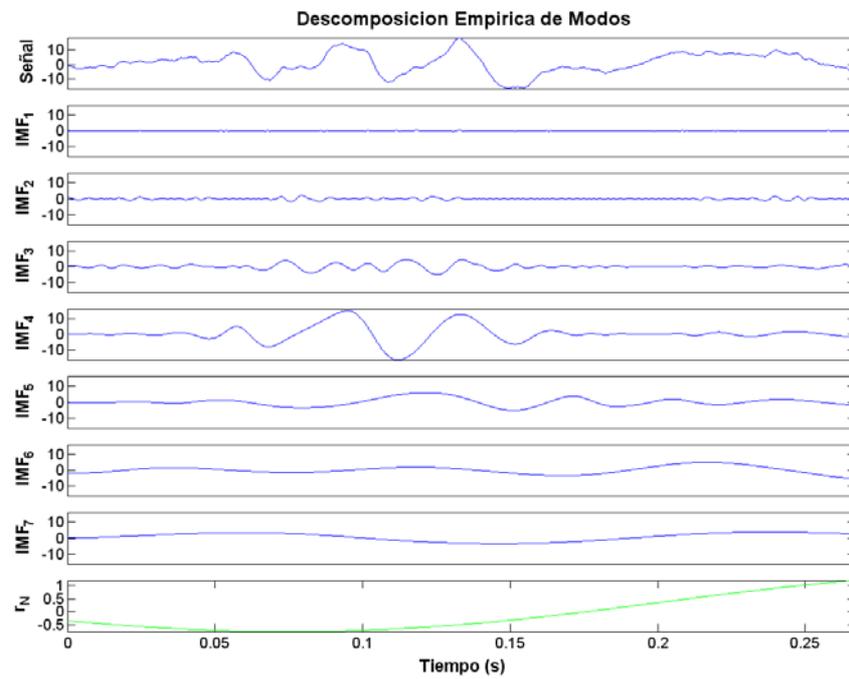


Fig. VI.55. Descomposición empírica de modos de la señal de S1 adquirida.

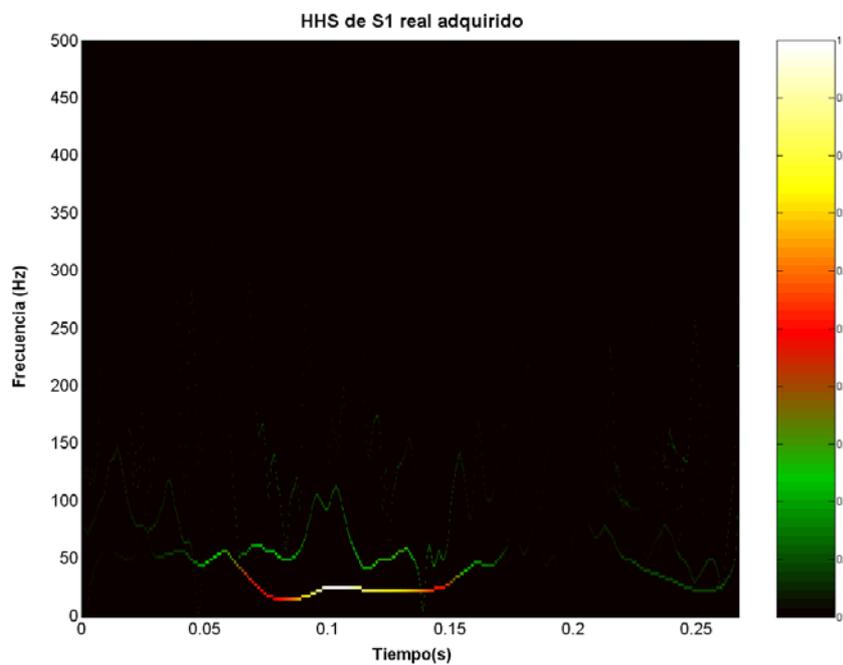


Fig. VI.56. Espectro de Hilbert-Huang del primer sonido cardiaco adquirido.

▪ S1 adquirido 2.

En la figura VI.57 se presenta una segunda señal del primer sonido cardiaco adquirido. El espectrograma de la señal se presenta en la figura VI.58 mientras que en la figura VI.59 se presenta el espectrograma reasignado. La descomposición empírica de modos de la señal se presenta en la figura VI.60 mientras que en la figura VI.61 se presenta su correspondiente Espectro de Hilbert-Huang.

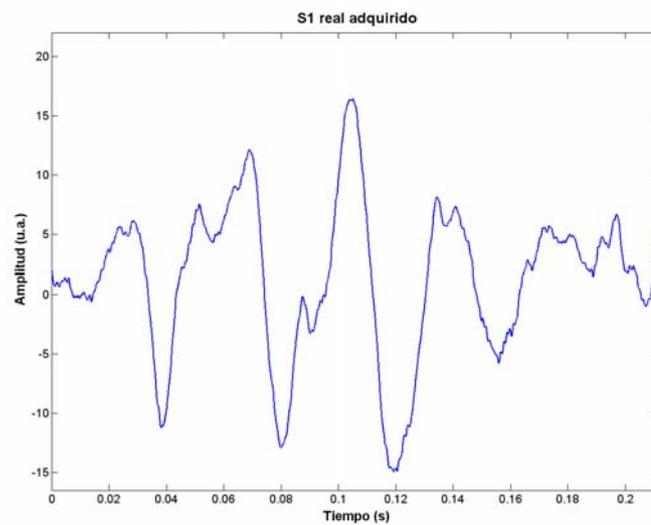


Fig. VI.57. Primer sonido cardiaco adquirido.

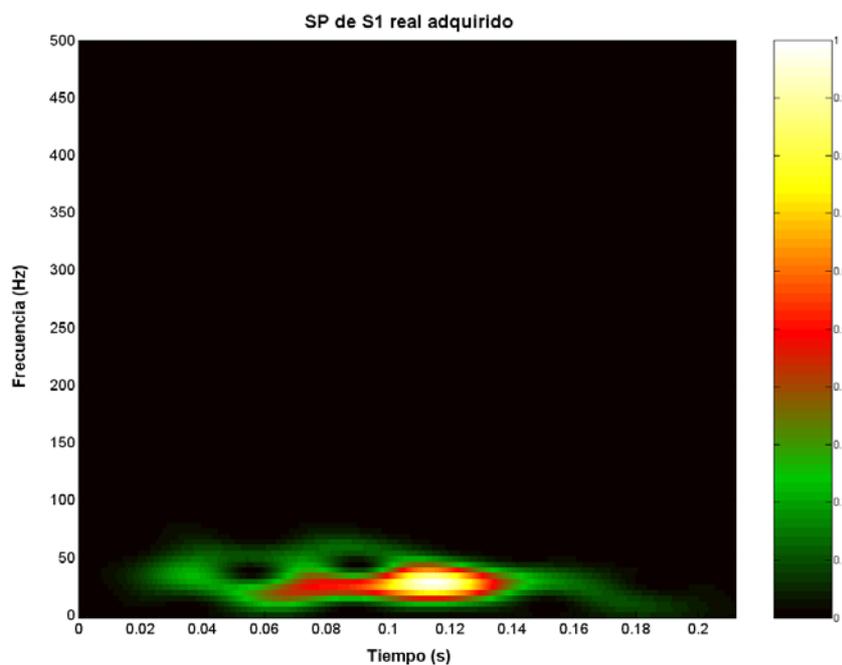


Fig. VI.58. Espectrograma de S1 adquirido de la figura VI.57, utilizando una ventana de Hamming de 60 ms de longitud.

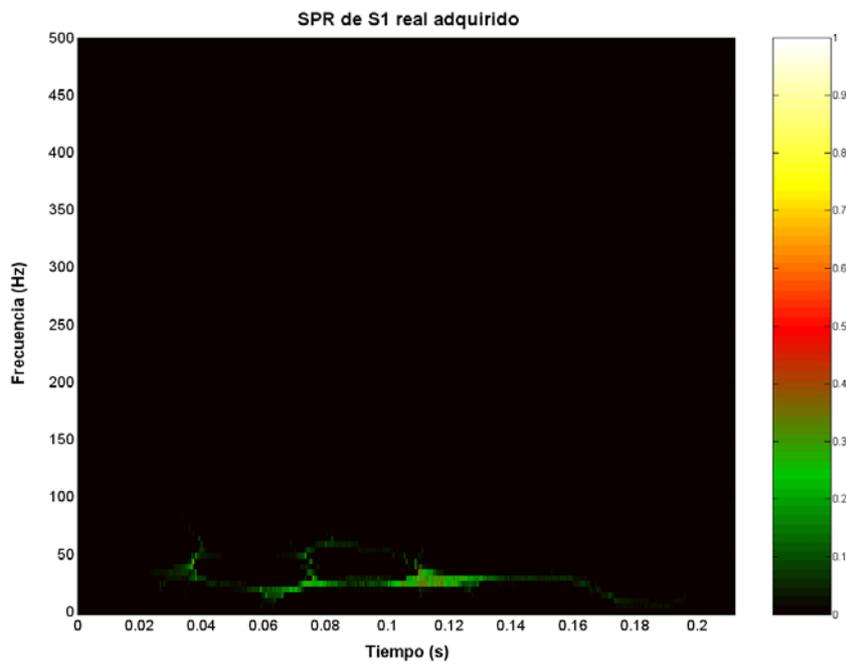


Fig. VI.59. Espectrograma reasignado de S1 adquirido de la figura VI.57.

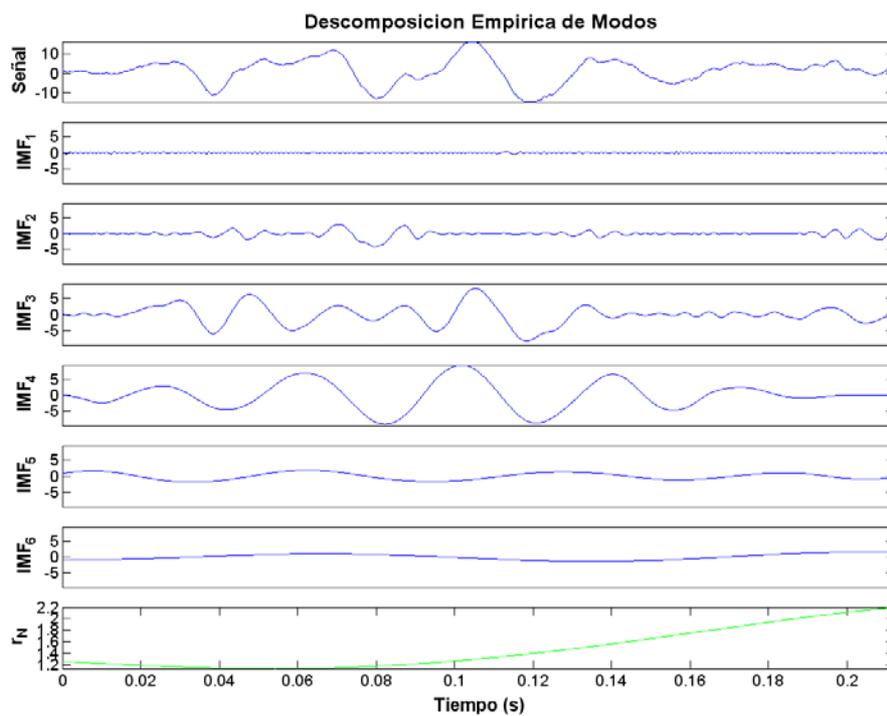


Fig. VI.60. Descomposición empírica de modos de la señal de S1 adquirida de la figura VI.57.

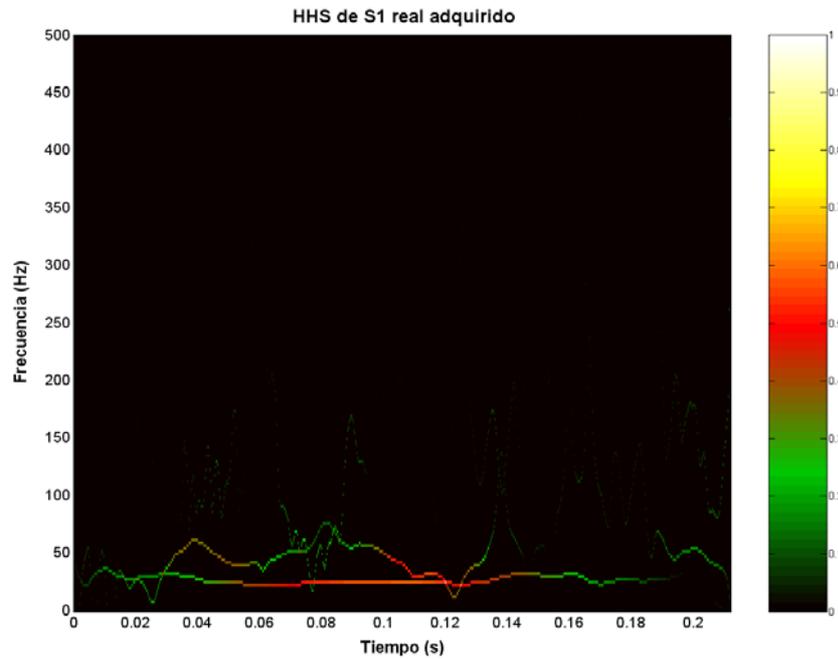


Fig. VI.61. Espectro de Hilbert-Huang del primer sonido cardíaco mostrado en la figura VI.57.

▪ S1 adquirido 3.

En la figura VI.62 se presenta una tercera señal de primer sonido cardíaco adquirida. El espectrograma de la señal se muestra en la figura VI.63 mientras que el espectrograma reasignado se muestra en la figura VI.64. La descomposición empírica de modos de la señal se presenta en la figura VI.65 y su correspondiente Espectro de Hilbert-Huang se incluye en la figura VI.66.

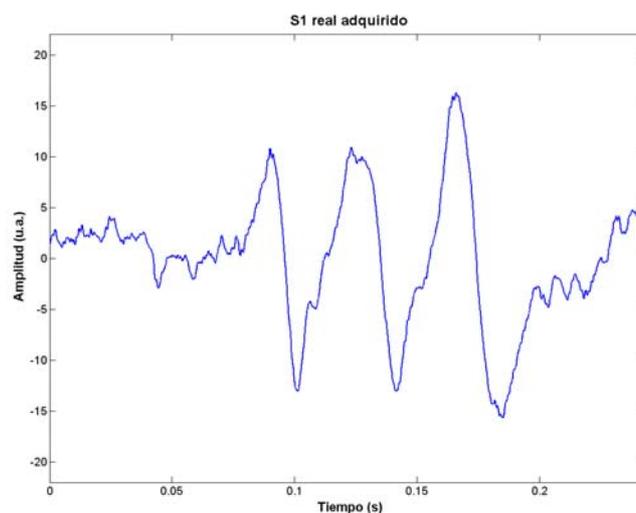


Fig. VI.62. Primer sonido cardíaco adquirido.

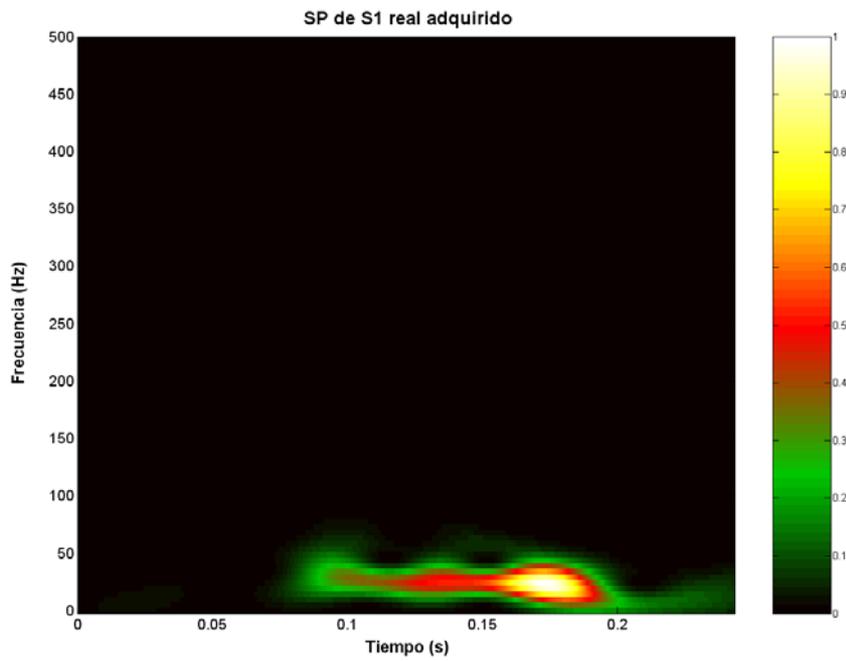


Fig. VI.63. Espectrograma de S1 adquirido mostrado en la figura VI.62, utilizando una ventana de Hamming de 60 ms de longitud.

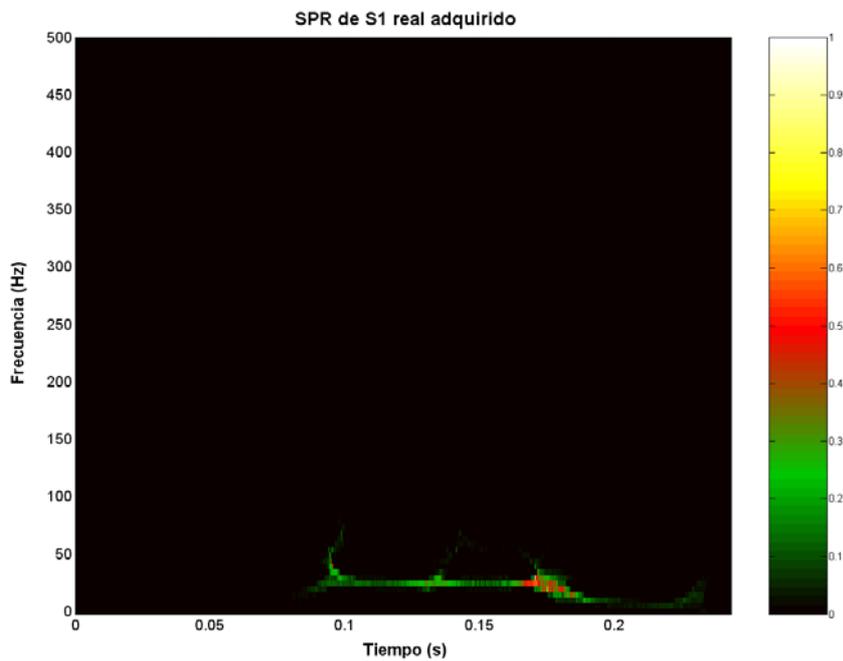


Fig. VI.64. Espectrograma reasignado de S1 adquirido mostrado en la figura VI.62.

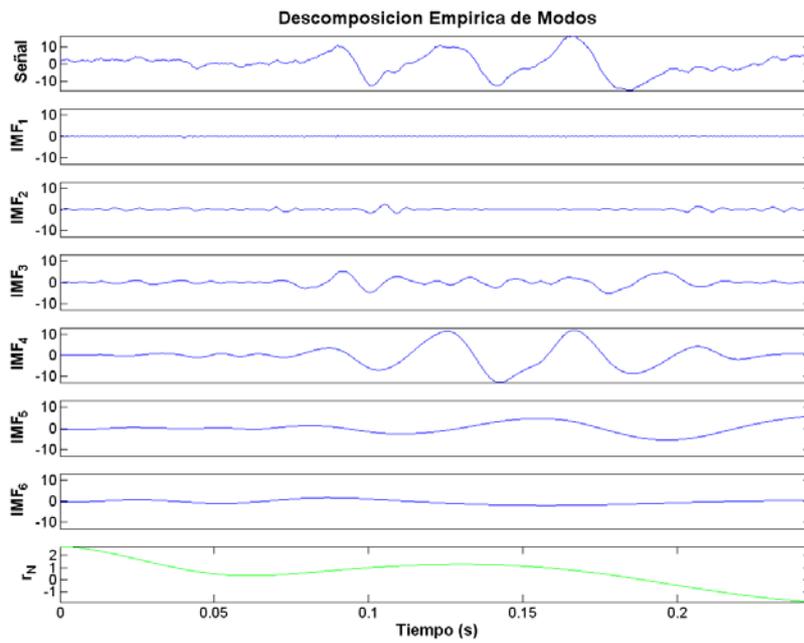


Fig. VI.65. Descomposición empírica de modos de la señal de S1 adquirida de la figura VI.62.

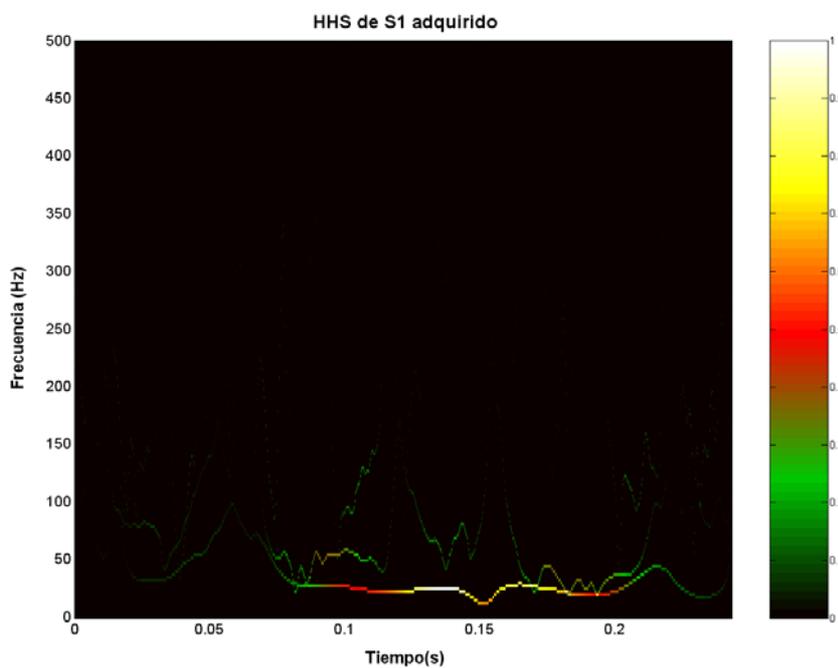


Fig. VI.66. Espectro de Hilbert-Huang del primer sonido cardíaco mostrado en la figura VI.62.

VI.5. Segundo sonido cardíaco y sus componentes.

VI.5.1. Componente aórtico simulado (A2) de S2.

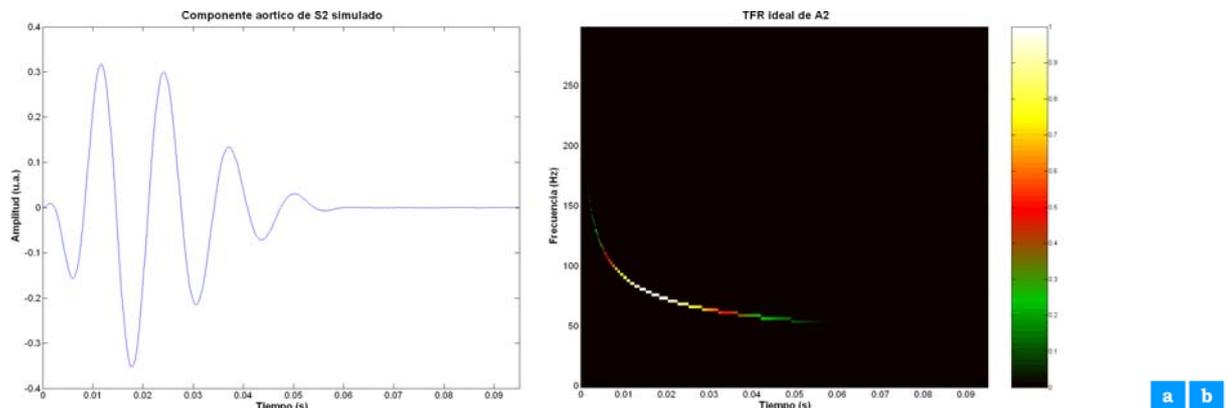


Fig. VI.67. Componente aórtico del segundo sonido cardiaco simulado (A2). a) Formas de onda.
b) Representación tiempo-frecuencia teórica.

En forma similar a S1 simulado y sus componentes, se obtuvieron todas las representaciones tiempo-frecuencia utilizadas en la presente tesis para el componente aórtico del segundo ruido cardiaco simulado, A2. Se generaron las sub-imágenes de correlación local, se realizó el cálculo de la correlación promedio y se seleccionaron las mejores representaciones tiempo-frecuencia con base en el criterio del máximo valor de correlación promedio.

Para el caso del espectrograma, la mejor representación tiempo-frecuencia para A2 fue la generada con una ventana Hamming de 44.6 ms de longitud. La función de ambigüedad de la señal, cada espectrograma generado al variar el tipo de ventana y su longitud, la tabla con los valores de correlación promedio resultantes de su comparación local, y el espectrograma seleccionado para A2 simulado se incluyen en el Anexo I.

Con respecto a la representación vía la Distribución de Wigner-Ville, la comparación local con la representación teórica mostrada en la figura VI.67.b generó un valor de $\rho_{promedio}$ igual a 0.7292. La distribución de Wigner de A2 simulado se incluyó en el Anexo I.

La mejor distribución de Choi-Williams para A2 simulado se obtuvo con un parámetro $\sigma = 661$. En el Anexo I se muestran las distribuciones Choi-Williams de la señal, obtenidas al variar el parámetro σ . Además, en el mismo anexo se presentan los resultados de la comparación local entre estas distribuciones y la representación tiempo-frecuencia ideal de la señal simulada, y la distribución Choi-Williams seleccionada para A2 simulado.

Para la técnica del espectrograma reasignado, la mejor representación para A2 se obtuvo con una ventana Blackman-Harris de 9.4 ms de longitud. En el Anexo I se presentan los espectrogramas reasignados generados para esta señal, la tabla con los valores de correlación promedio para cada uno de ellos, y el espectrograma reasignado seleccionado para A2 simulado.

Con relación a la representación tiempo-frecuencia mediante el modelo AR vía el algoritmo de Burg, la mejor representación se obtuvo para un modelo AR de orden 2 y una ventana Blackman-Harris de 29.4 ms de longitud. En el Anexo I se incluyen las representaciones tiempo-frecuencia de Burg obtenidas para A2 simulado, los resultados de su comparación local con la representación tiempo-frecuencia teórica, y la representación P_{Burg} seleccionada para el componente aórtico de S2 simulado.

Con respecto a la técnica por el algoritmo RLS, la mejor representación de A2 simulado se obtuvo con un modelo AR de segundo orden y un factor de olvido de 0.8125. Cada representación P_{RLS} generada al variar el orden del modelo y el factor de olvido, la tabla con sus valores de correlación promedio correspondientes, y la representación tiempo-frecuencia P_{RLS} seleccionada para A2 simulado, se incluyen en el anexo I.

Para la técnica de representación tiempo-frecuencia vía el algoritmo RLS-VFF, la mejor representación se obtuvo con un modelo AR de orden 2 y una varianza del ruido esperado igual a 5.6234×10^{-11} . En el Anexo I se incluyen las representaciones tiempo-frecuencia $P_{\text{RLS-VFF}}$ generadas para A2 simulado, la tabla con los valores de correlación promedio obtenidos para cada una de ellas, y la representación $P_{\text{RLS-VFF}}$ seleccionada para A2 simulado.

Al utilizar el algoritmo de Kalman para generar la mejor representación de A2 se obtuvo que los parámetros adecuados eran un modelo AR de orden 2, una varianza del ruido de medición unitaria y una varianza del ruido del proceso igual a 5.6234×10^{-4} . En el Anexo I se incluyen las representaciones P_{Kalman} generadas para A2 simulado, los resultados de su comparación local, y la representación P_{Kalman} seleccionada para esta señal.

Para el escalograma, la mejor representación de A2 se obtuvo con la *wavelet* Morlet de 2.1 ms de longitud en la escala más gruesa. En el Anexo I se incluye cada escalograma generado al variar la longitud de la *wavelet*. Además, se incluye la tabla con los valores de correlación promedio de estos escalogramas, así como el escalograma seleccionado para el componente aórtico de S2 simulado.

Con respecto al espectro de Hilbert-Huang de A2 simulado, se obtuvo el valor de $\rho_{promedio}$ igual a 0.7677 mediante su comparación con la representación teórica de la señal. La descomposición empírica de modos de A2 se presenta en el Anexo I, donde también se incluyen la contribución de energía de cada modo de oscilación, la señal analítica del primer modo de oscilación intrínseco, el espectro de Hilbert-Huang que resulta de este modo de oscilación, una versión suavizada de este espectro de Hilbert-Huang de A2 simulado.

▪ **Comparación de las técnicas tiempo-frecuencia.**

Las representaciones tiempo-frecuencia seleccionadas para A2 de acuerdo al criterio de correlación local se presentan en la figura VI.68. La estimación de la frecuencia instantánea de la señal a partir del centroide de cada una de estas representaciones se muestra en la figura VI.69. En la tabla VI.VII se presentan los valores de los índices de desempeño obtenidos para cada una de las representaciones tiempo-frecuencia mostrados en la figura VI.68.

TABLA VI.VII. INDICES DE DESEMPEÑO DE LAS REPRESENTACIONES TIEMPO-FRECUENCIA SELECCIONADAS PARA A2.

	ρ	$\rho_{promedio}$	NRMSE	ρ_{FI}	res_{TF}
SP	0.4973	0.7123	0.6642	0.7508	0.0471
WVD	0.5676	0.7292	1.0000	0.0015	0.0036
CWD	0.5727	0.7242	0.7728	0.6358	0.0036
SPR	0.7408	0.7617	0.7437	0.6715	2.46×10^{-5}
P_{BURG}	0.7169	0.7488	0.8706	0.4921	0.0023
P_{RLS}	0.6974	0.7392	0.9898	0.1516	0.0030
P_{RLS-VFF}	0.6395	0.7306	0.7869	0.6589	0.0030
P_{KALMAN}	0.4456	0.6751	0.7851	0.6797	0.0951
SC	0.4944	0.7147	0.7835	0.6374	0.0490
HHS	0.7520	0.7677	0.9958	0.0924	8.21×10^{-7}

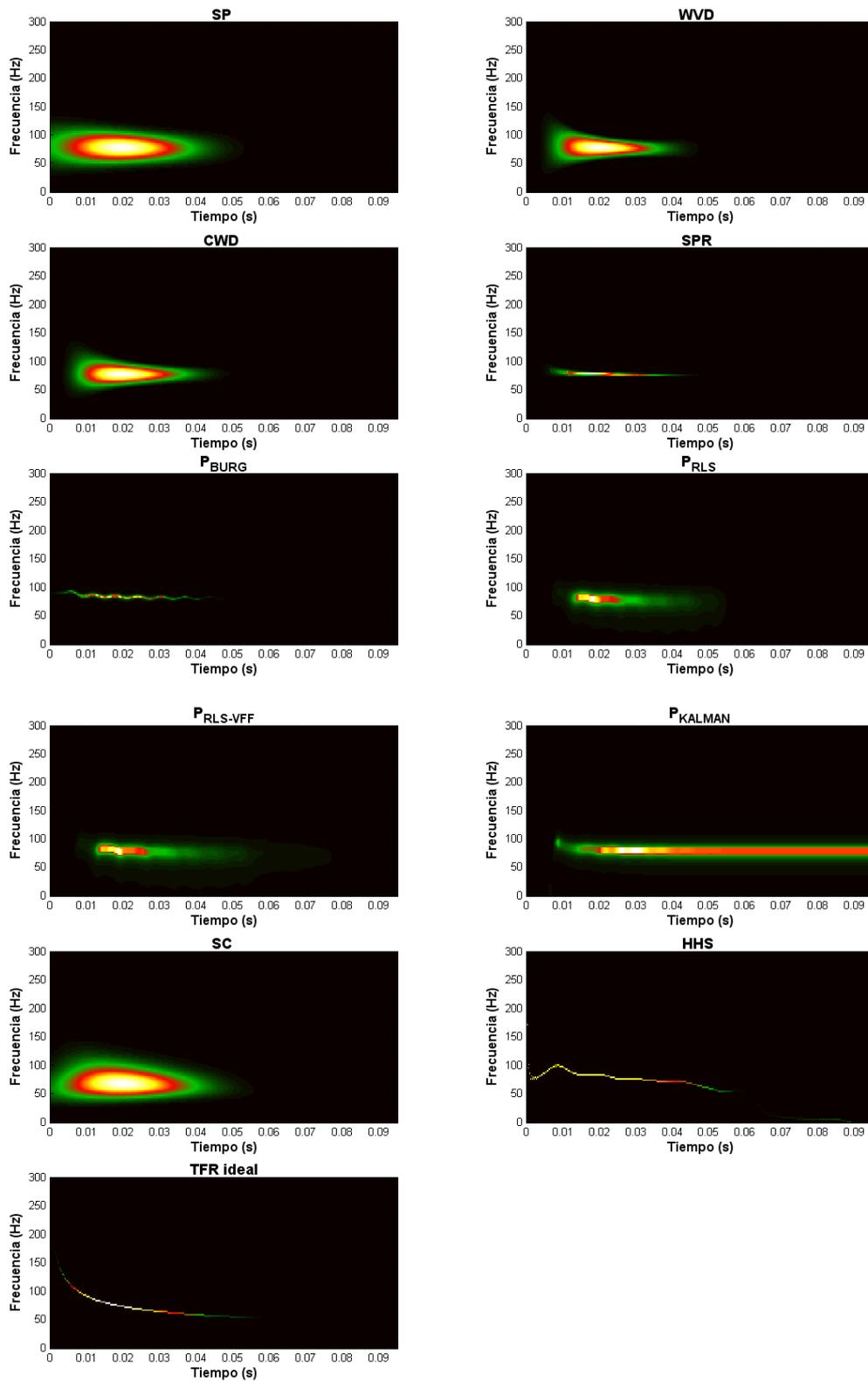


Fig. VI.68. Representaciones tiempo-frecuencia seleccionadas para cada técnica de análisis tiempo-frecuencia así como la representación tiempo-frecuencia teórica de A2 simulado.

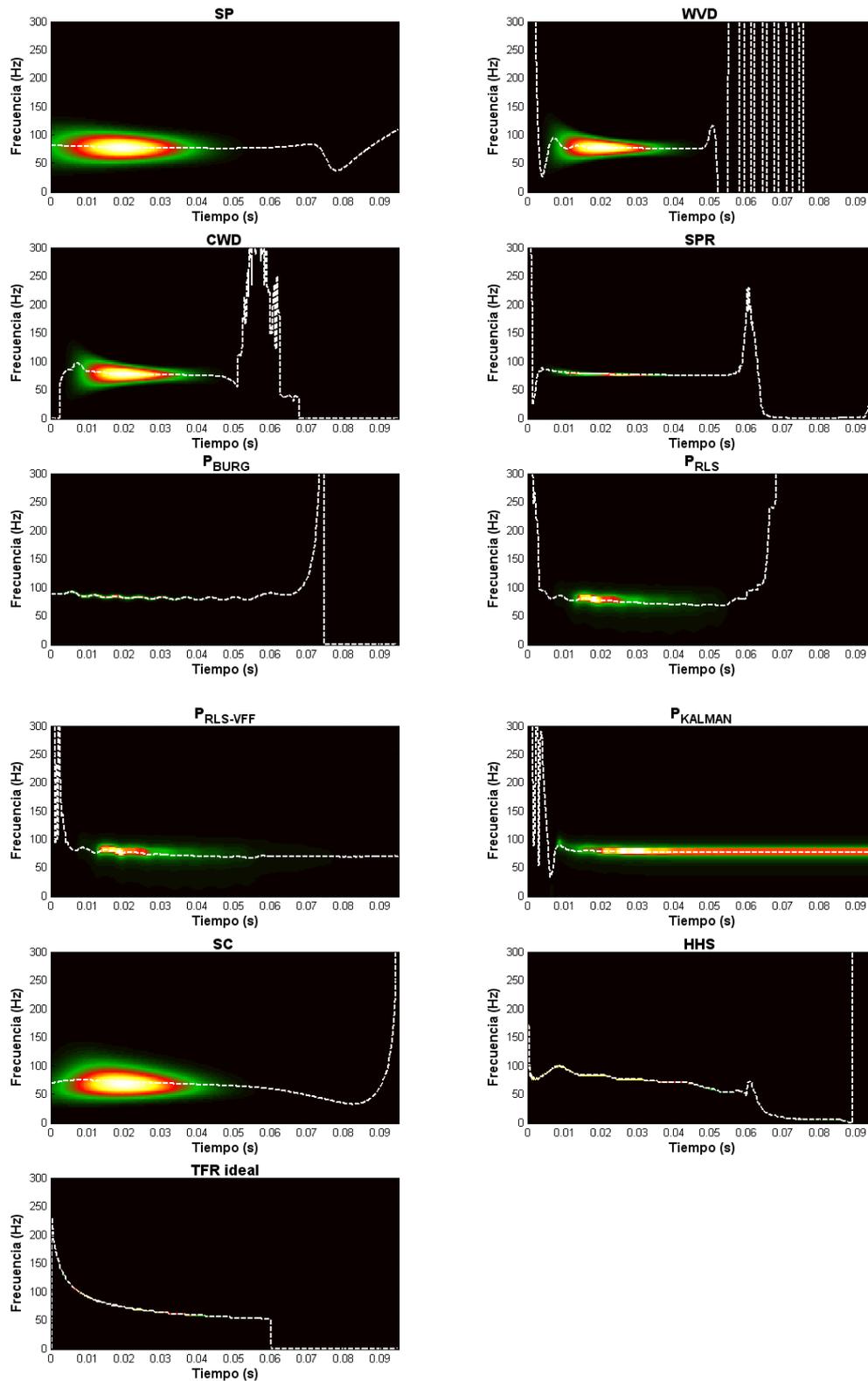


Fig. VI.69. Estimación de la frecuencia instantánea mediante el primer momento de cada representación tiempo-frecuencia seleccionada así como teórica de A2 simulado.

VI.5.2. Componente pulmonar simulado (P2) de S2.

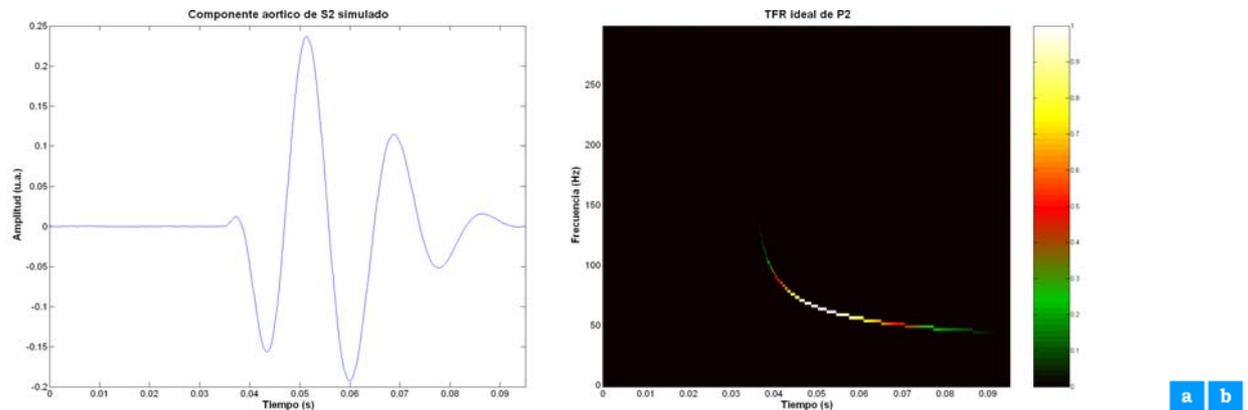


Fig. VI.70. Componente pulmonar del segundo sonido cardiaco (P2). a) Forma de onda. b) Representación tiempo-frecuencia teórica.

En forma similar al caso del componente aórtico del segundo ruido cardiaco (A2), se obtuvieron todas las representaciones tiempo-frecuencia utilizadas en la presente tesis para el componente pulmonar simulado del segundo ruido cardiaco, P2. Se generaron las subimágenes de correlación local, se realizó el cálculo de la correlación promedio y se seleccionaron las mejores representaciones tiempo-frecuencia con base en el criterio del máximo valor de correlación promedio.

Para el caso del espectrograma, la mejor representación tiempo-frecuencia para P2 simulado fue la generada con una ventana Hamming de 44.6 ms de longitud. De forma similar a las señales anteriores, la función de ambigüedad de la señal, cada espectrograma generado al variar el tipo de ventana y su longitud, los resultados de su comparación local con la representación tiempo-frecuencia teórica, así como el espectrograma seleccionado para P2 simulado se incluyen en el Anexo I.

La distribución Wigner-Ville de P2, presentada en el Anexo I, generó un valor de $\rho_{promedio}$ igual a 0.7300.

Para la Distribución Choi-Williams, la mejor representación se obtuvo para el valor de $\sigma = 1011$. En el Anexo I se incluye cada distribución generada Choi-Williams de P2 simulado. Además, en el mismo anexo se presentan los resultados de la comparación local de cada una de estas distribuciones, así como la distribución Choi-Williams seleccionada para el componente pulmonar de S2 simulado.

Para el espectrograma reasignado, la mejor representación se obtuvo con una ventana Blackman-Harris de 29.4 ms de longitud. Los espectrogramas reasignados obtenidos al variar σ , la tabla con los que contiene sus valores de correlación promedio correspondientes, y el espectrograma reasignado seleccionado para P2 simulado se incluyen en el Anexo I.

Cada representación tiempo-frecuencia de Burg que se generó para P2 simulado, así como la tabla con sus valores de correlación promedio correspondientes, se incluyen en el Anexo I. En este anexo también se presenta la representación P_{Burg} seleccionada para esta señal, la cual en este caso se obtuvo con un modelo AR de segundo orden y una ventana Blackman-Harris de 34.6 ms de longitud.

Para la técnica mediante el algoritmo de RLS, la mejor representación se obtuvo con un modelo AR de orden 2 y un factor de olvido igual a 0.84375. Cada representación tiempo-frecuencia generada con esta técnica para P2 simulado se presenta en el Anexo I, donde también se incluye una tabla con los valores de correlación promedio correspondientes, y la representación P_{RLS} seleccionada para este componente cardíaco simulado.

Para la técnica de representación tiempo-frecuencia mediante el algoritmo RLS-VFF, la mejor representación se obtuvo para un modelo AR orden 2 y una varianza del ruido esperado igual a 5.6234×10^{-11} . En el Anexo I se incluye el conjunto de representación $P_{\text{RLS-VFF}}$ calculadas para esta señal, así como su comparación local con la representación tiempo-frecuencia teórica. En el mismo anexo se presenta la representación $P_{\text{RLS-VFF}}$ seleccionada.

Con respecto a la representación mediante el filtro de Kalman, la mejor representación tiempo-frecuencia se obtuvo empleando un modelo AR de orden 3, una varianza del ruido del proceso igual a 4.217×10^{-2} , y una varianza del ruido de medición unitaria. En el Anexo I se presenta cada representación P_{Kalman} generada para P2 simulado, así como sus valores de correlación resultantes de su comparación local, y la representación P_{Kalman} seleccionada para el componente pulmonar simulado.

Para el escalograma Morlet, la mejor representación se obtuvo con una wavelet de 1.275 ms de longitud en la escala más gruesa. En el Anexo I se presentan los escalogramas Morlet obtenidos para P2 simulado. Además, se incluye la tabla con los resultados de la comparación local entre cada escalograma y la representación tiempo-frecuencia teórica de esta señal, así como el escalograma Morlet seleccionado para P2 simulado.

Con respecto al espectro de Hilbert-Huang del componente pulmonar simulado de S2, en el Anexo I se presenta la descomposición empírica de modos de esta señal simulada, la contribución de energía de cada modo de oscilación resultante. Además, en el mismo anexo se presenta la señal analítica del primer modo de oscilación intrínseco, y el espectro de Hilbert-

Huang de P2 simulado que resulta de este primer modo de oscilación intrínseco. La comparación local entre este espectro de Hilbert-Huang y la representación tiempo-frecuencia teórica de la señal produjo un valor de $\rho_{promedio}$ igual a 0.7754. Finalmente, también se anexa una versión suavizada del espectro de Hilbert-Huang del componente pulmonar de S2 simulado.

▪ **Comparación de las técnicas tiempo-frecuencia.**

En la figura VI.71 se presentan las representaciones tiempo-frecuencia seleccionadas de acuerdo al criterio de correlación local para P2 simulado. En la figura VI.72 se presenta la estimación de la frecuencia instantánea de la señal a partir del centroide de cada una de estas representaciones seleccionadas. En la tabla VI.VIII se presentan los valores de los índices de desempeño obtenidos para cada una de las representaciones tiempo-frecuencia mostrados en la figura VI.71.

TABLA VI.VIII. INDICES DE DESEMPEÑO DE LAS REPRESENTACIONES
TIEMPO-FRECUENCIA SELECCIONADAS PARA P2.

	ρ	$\rho_{promedio}$	NRMSE	ρ_{FI}	res _{TF}
SP	0.5127	0.7242	0.6384	0.7905	0.0947
WVD	0.5974	0.7300	0.9974	0.0729	0.0033
CWD	0.6027	0.7267	0.6772	0.8548	0.0047
SPR	0.7890	0.7940	0.7046	0.7481	2.29x10 ⁻⁵
P_{BURG}	0.7744	0.7535	0.4926	0.8776	0.0023
P_{RLS}	0.7793	0.7335	0.4803	0.9061	0.0028
P_{RLS-VFF}	0.7646	0.7366	0.4065	0.9301	0.0065
P_{KALMAN}	0.5564	0.6993	0.6626	0.8507	0.2424
SC	0.4523	0.7253	0.6220	0.8925	0.1879
HHS	0.6754	0.7754	0.6496	0.7766	5.72x10 ⁻⁷

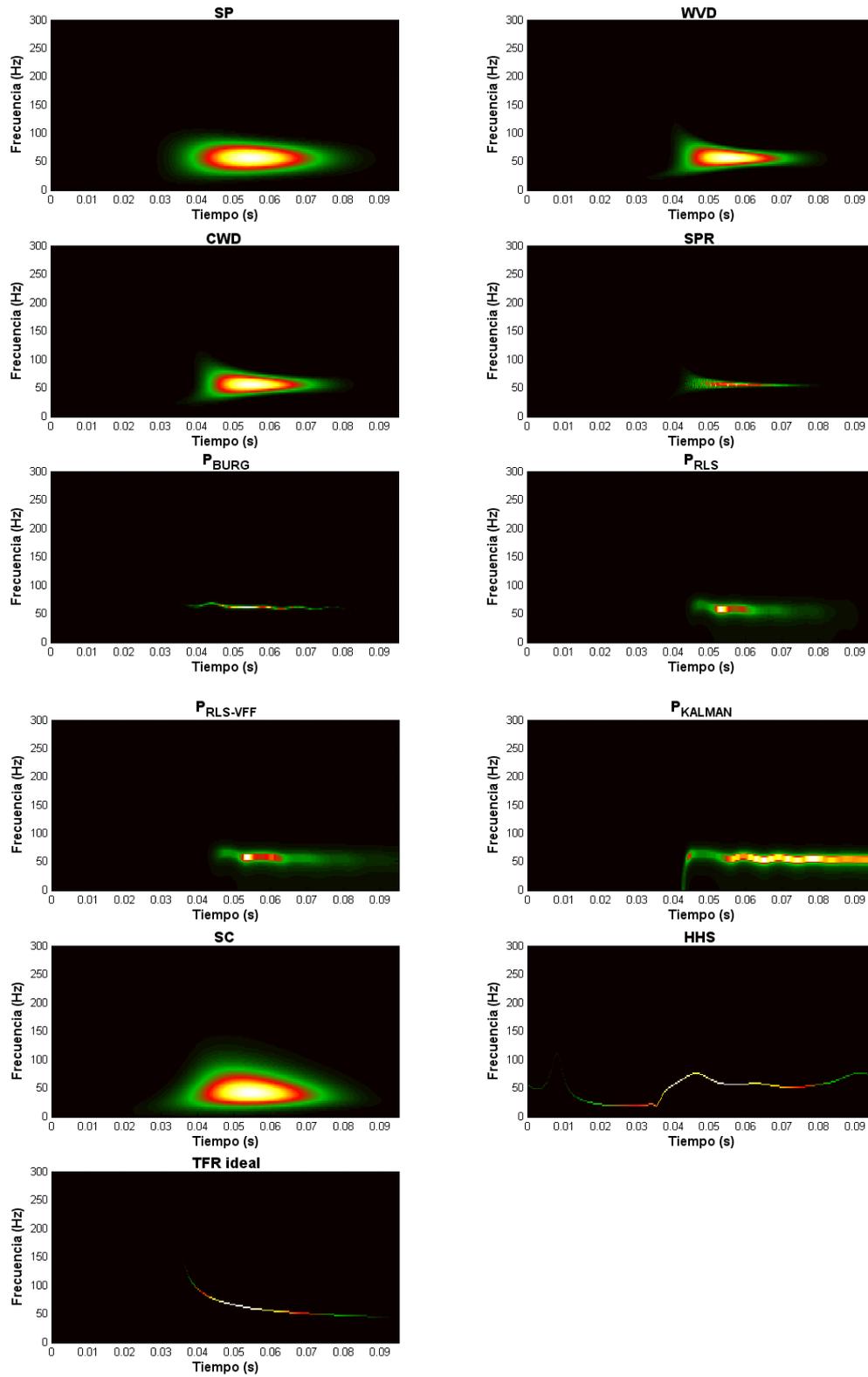


Fig. VI.71. Representaciones tiempo-frecuencia seleccionadas para cada técnica de análisis tiempo-frecuencia así como la representación tiempo-frecuencia teórica de P2 simulado.

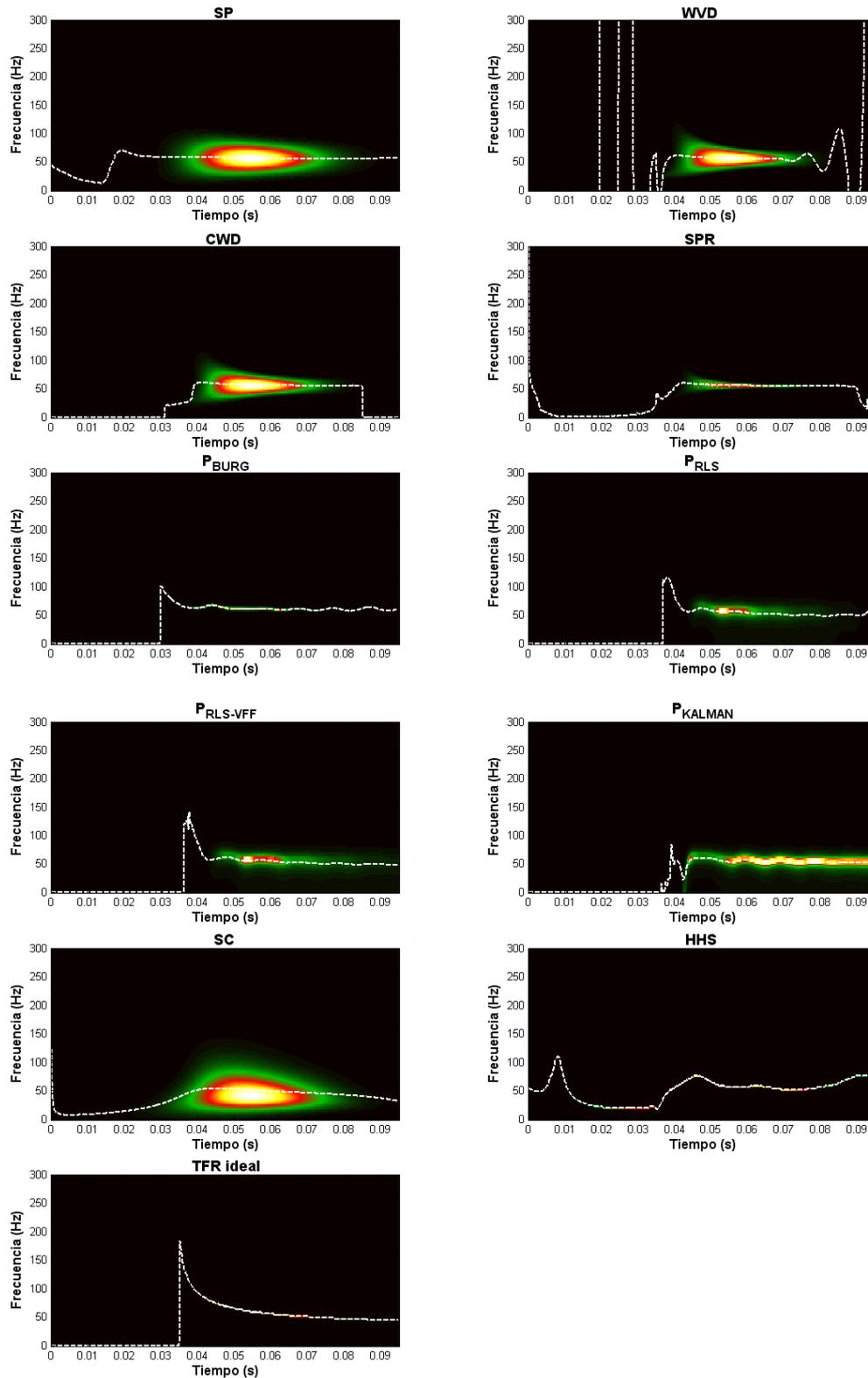


Fig. VI.72. Estimación de la frecuencia instantánea mediante el primer momento de cada representación tiempo-frecuencia seleccionada así como la teórica de P2 simulado.

VI.5.3. Segundo sonido cardiaco simulado (S2).

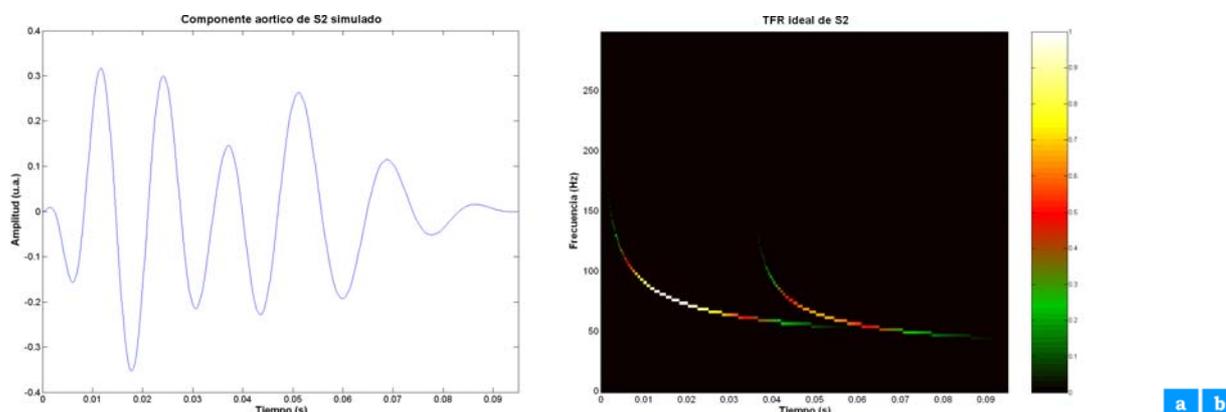


Fig. VI.73. Segundo sonido cardiaco simulado (S2). a) Forma de onda. b) Representación tiempo-frecuencia teórica.

Para el segundo sonido cardiaco simulado, también se obtuvieron todas las representaciones tiempo-frecuencia utilizadas en la presente tesis, se generaron las subimágenes de correlación local, se realizó el cálculo de la correlación promedio y se seleccionaron las mejores representaciones tiempo-frecuencia con base en el criterio del máximo valor de correlación promedio.

Para el caso del espectrograma, la mejor representación tiempo-frecuencia fue la generada con una ventana Hamming de 44.6 ms de longitud. La función de ambigüedad de la señal, la comparación local entre cada espectrograma generado para S2 y la representación tiempo-frecuencia teórica, la tabla con los valores de correlación promedio correspondientes, y el espectrograma seleccionado para esta señal se presentan en el Anexo I.

Para la Distribución Wigner-Ville el valor de $\rho_{promedio}$ que se obtuvo fue de 0.6329. La distribución Wigner-Ville de S2 simulado se presenta en el Anexo I.

Para la Distribución Choi-Williams, la mejor representación se obtuvo con el parámetro $\sigma=1011$. En el Anexo I se incluye cada distribución Choi-Williams calculada para esta señal con diferentes valores de σ . En el mismo anexo se presentan los resultados de la comparación local de estas distribuciones, así como la distribución Choi-Williams seleccionada para este sonido cardiaco.

En el caso del espectrograma reasignado la mejor representación se obtuvo con una ventana Blackman-Harris de 9.4 ms de longitud. Los espectrogramas reasignados generados, al variar el tipo de ventana y su longitud, se incluyen en el anexo I. En el mismo anexo se incluye la tabla con los valores de correlación promedio obtenidos para estas distribuciones, así como el espectrograma reasignado seleccionado para S2 simulado.

Con relación a la representación tiempo-frecuencia por el algoritmo de Burg, la mejor representación se obtuvo con un modelo AR de orden 2 y una ventana Blackman-Harris de 29.4 ms de longitud. Las representaciones tiempo-frecuencia P_{Burg} generadas para S2 simulado, los resultados de su comparación local con la representación tiempo-frecuencia teórica, y la representación tiempo-frecuencia de S2 seleccionada para esta técnica, se incluyen en el Anexo I.

En el caso de la representación obtenida mediante el algoritmo RLS, la mejor representación se obtuvo con un modelo AR de orden 2 y un factor de olvido igual a 0.84375. Cada representación paramétrica generada con este algoritmo, al variar el orden del modelo AR y el factor de olvido, se incluyen en el Anexo I junto con sus valores de correlación promedio y la representación P_{RLS} seleccionada para S2 simulado.

Para la versión del RLS con factor de olvido variable, RLS-VFF, la mejor representación de S2 simulado se obtuvo con un modelo AR orden 2 y una varianza del ruido esperado igual a 3.1623×10^{-11} . En el Anexo I se incluye en conjunto de representaciones tiempo-frecuencia $P_{\text{RLS-VFF}}$ obtenidas para esta señal, así como los resultados de su comparación local con la representación tiempo-frecuencia teórica, y la representación $P_{\text{RLS-VFF}}$ seleccionada para esta señal.

Con respecto a la representación obtenida vía el filtro de Kalman, la mejor representación se obtuvo con un modelo AR de orden 2, una varianza del ruido de medición (σ^2) unitaria y una varianza del ruido del proceso (q) igual a 0.1. Cada representación P_{Kalman} generada para S2 simulado, la tabla con los valores de correlación promedio resultantes, y la representación tiempo-frecuencia seleccionada para esta técnica se incluyen en el Anexo I.

El escalograma que proporcionó la mejor representación se obtuvo con la wavelet Morlet de 1.95 ms de longitud en la escala más gruesa. En el Anexo I se presenta el conjunto de escalogramas obtenidos para esta señal, así como los resultados de su comparación local con la representación tiempo-frecuencia teórica, y el escalograma Morlet seleccionado para S2 simulado.

Finalmente, el Espectro de Hilbert-Huang de S2 simulado produjo un valor de $\rho_{promedio}$ igual a 0.7258. En el Anexo I se presenta la descomposición empírica de modos del segundo sonido cardiaco simulado. En el mismo anexo también se incluye la contribución de energía de cada modo de oscilación resultante, la señal analítica del primer modo de oscilación intrínseco de S2 simulado, el espectro de Hilbert-Huang que resulta de dicho modo de oscilación intrínseco, y una versión suavizada del espectro de Hilbert-Huang de S2 simulado.

▪ **Comparación de las técnicas tiempo-frecuencia.**

En la figura VI.74 se presentan las representaciones tiempo-frecuencia seleccionadas de acuerdo al criterio de correlación local para S2 simulado. En la figura VI.75 se presenta la estimación de la frecuencia instantánea de la señal a partir del centroide de cada una de estas representaciones seleccionadas. Los valores de los índices de desempeño obtenidos para cada una de las representaciones tiempo-frecuencia mostrados en la figura VI.74 se presentan en la tabla VI.IX. En la tabla VI.X se presentan los índices de desempeño basados en la estimación frecuencia instantánea para instantes de tiempo en los cuales la amplitud de la señal es mayor al 5% de su valor máximo.

TABLA VI.IX. INDICES DE DESEMPEÑO DE LAS

REPRESENTACIONES TIEMPO-FRECUENCIA SELECCIONADAS PARA S2.

	ρ	$\rho_{promedio}$	NRMSE	ρ_{FI}	res_{TF}
SP	0.4743	0.6239	0.3026	0.9593	0.0074
WVD	0.5277	0.6329	0.9072	0.4895	0.0031
CWD	0.5086	0.6036	0.5434	0.8811	0.0039
SPR	0.6825	0.6914	0.6648	0.7479	2.11×10^{-5}
P_{BURG}	0.7184	0.6656	0.2804	0.9602	0.0006
P_{RLS}	0.6729	0.6312	0.7449	0.6916	0.0056
P_{RLS-VFF}	0.6730	0.6302	0.7494	0.6895	0.0069
P_{KALMAN}	0.5084	0.6247	0.7496	0.7011	0.0097
SC	0.4746	0.6187	0.3919	0.9637	0.0110
HHS	0.6226	0.7258	0.9733	0.2512	8.39×10^{-7}

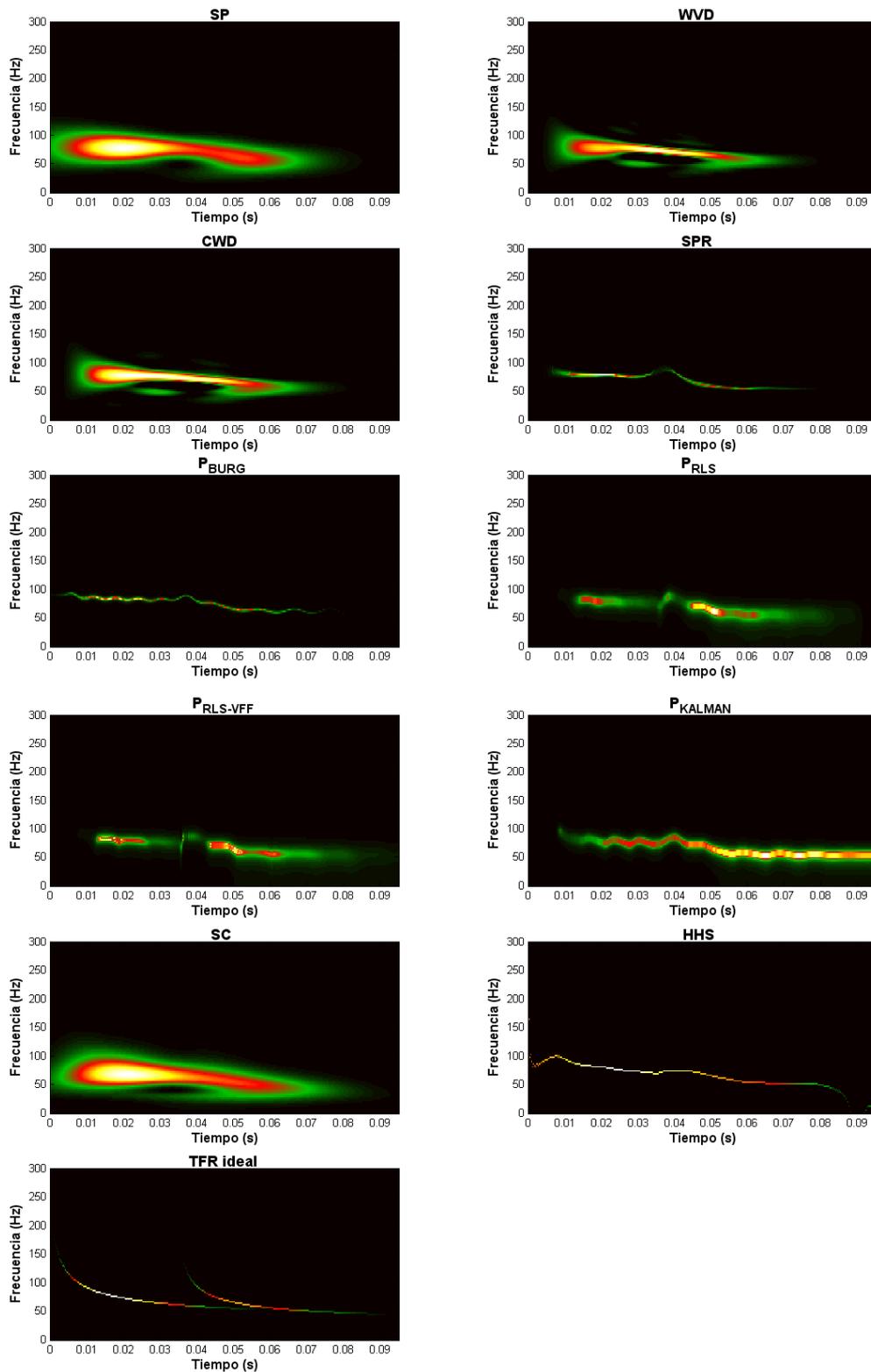


Fig. VI.74. Representaciones tiempo-frecuencia seleccionadas para cada técnica de análisis tiempo-frecuencia así como la representación tiempo-frecuencia teórica de S2 simulado.

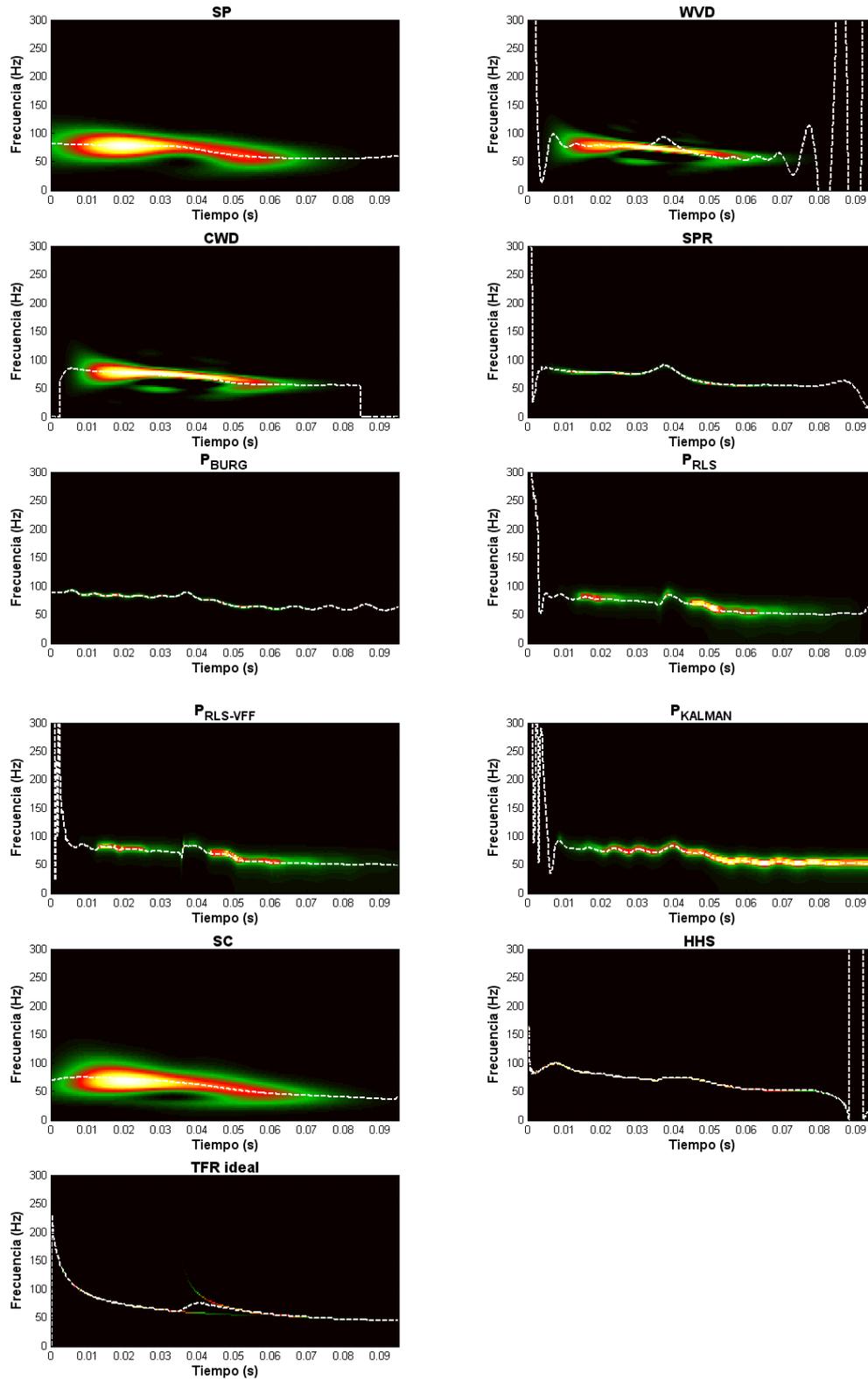


Fig. VI.75. Estimación de la frecuencia instantánea mediante el primer momento de cada representación tiempo-frecuencia seleccionada así como la teoría de S2 simulado.

TABLA VI.X. INDICES DE DESEMPEÑO DE LAS REPRESENTACIONES TIEMPO-FRECUENCIA SELECCIONADAS PARA S2, OBTENIDOS PARA VALORES DE AMPLITUD DE LA SEÑAL > 5% DEL MÁXIMO.

	NRMSE	ρ_{FI}
SP	0.1742	0.9851
WVD	0.7026	0.7122
CWD	0.2017	0.9807
SPR	0.1771	0.9842
P_{BURG}	0.1792	0.9854
P_{RLS}	0.1779	0.9849
P_{RLS-VFF}	0.1462	0.9902
P_{KALMAN}	0.2789	0.9621
SC	0.2379	0.9877
HHS	0.1225	0.9925

VI.5.4. Señales de S2 reales adquiridas.

- **S2 adquirido 1 (foco aórtico).**

En la figura VI.76 se presenta una señal de segundo sonido cardiaco adquirida de un sujeto sano sobre el foco aórtico de auscultación. La descomposición empírica de modos de la señal se presenta en la figura VI.77, mientras que en la figura VI.78 se presenta la señal analítica de los modos de oscilación intrínsecos seleccionados con base al criterio de energía. En las figuras VI.79 y VI.80 se presentan el Espectro de Hilbert-Huang del sonido cardiaco adquirido y una versión suavizada de dicho espectro, respectivamente.

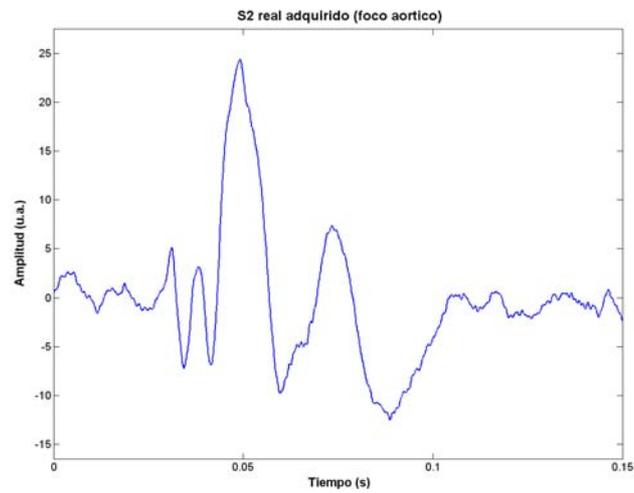


Fig. VI.76. Segundo sonido cardíaco adquirido sobre el foco aórtico.

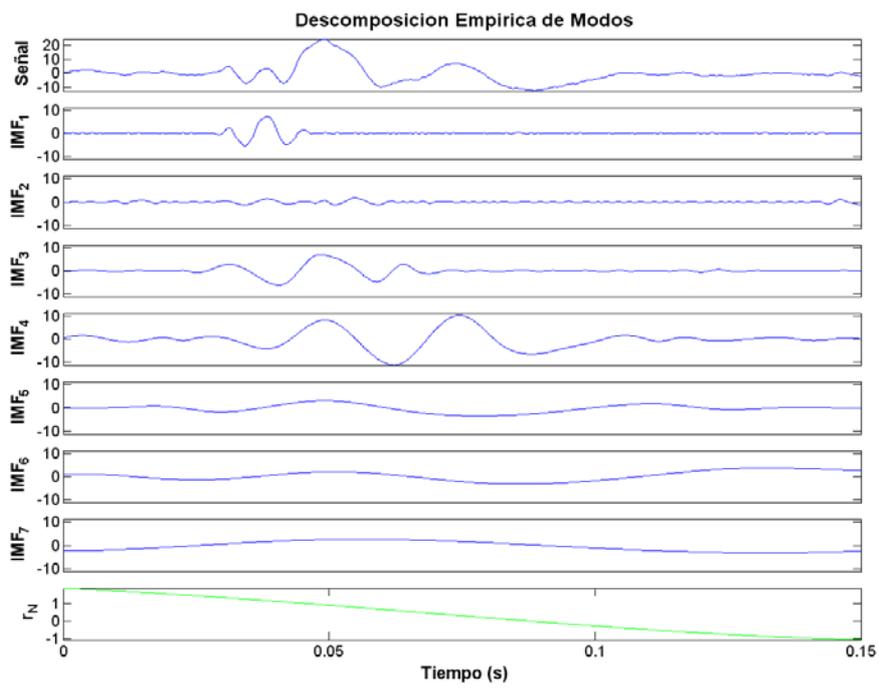


Fig. VI.77. Descomposición empírica de modos de la señal de S2 adquirida de la figura VI.76.

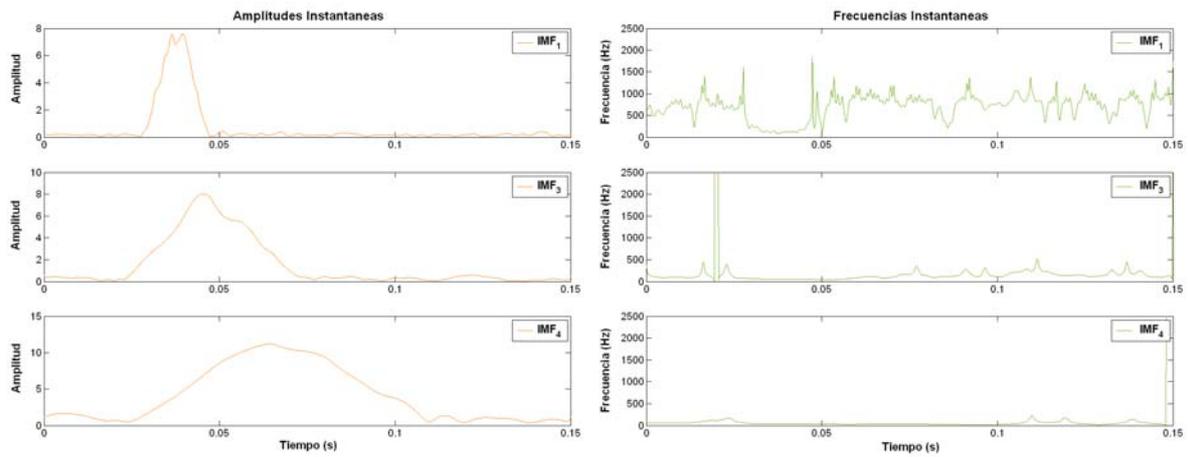


Fig. VI.78. Señales analíticas de los tres primeros modos de oscilación intrínsecos de la señal S2 adquirida. Amplitud instantánea (columna izquierda) y Frecuencia instantánea (columna derecha).

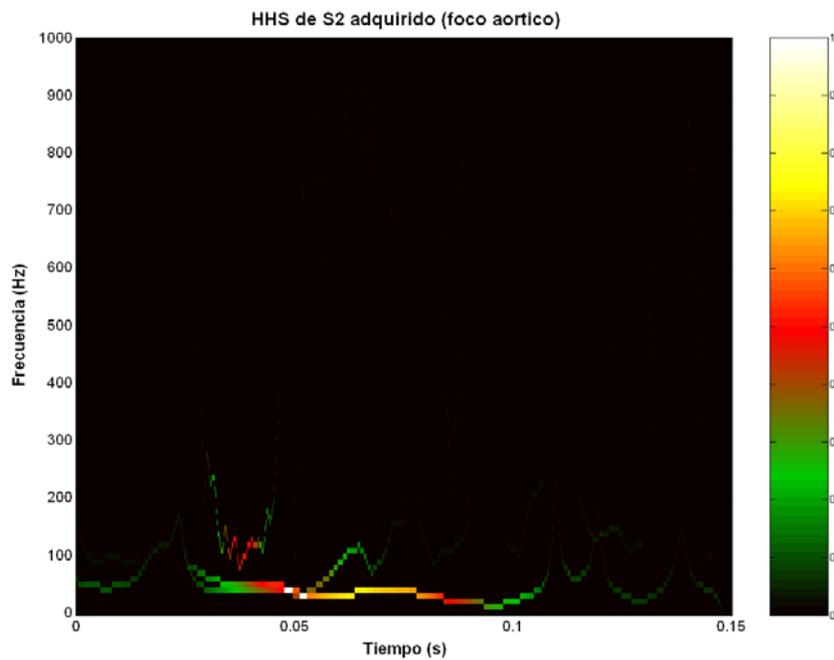


Fig. VI.79. Espectro de Hilbert-Huang del segundo sonido cardiaco mostrado en la figura VI.76.

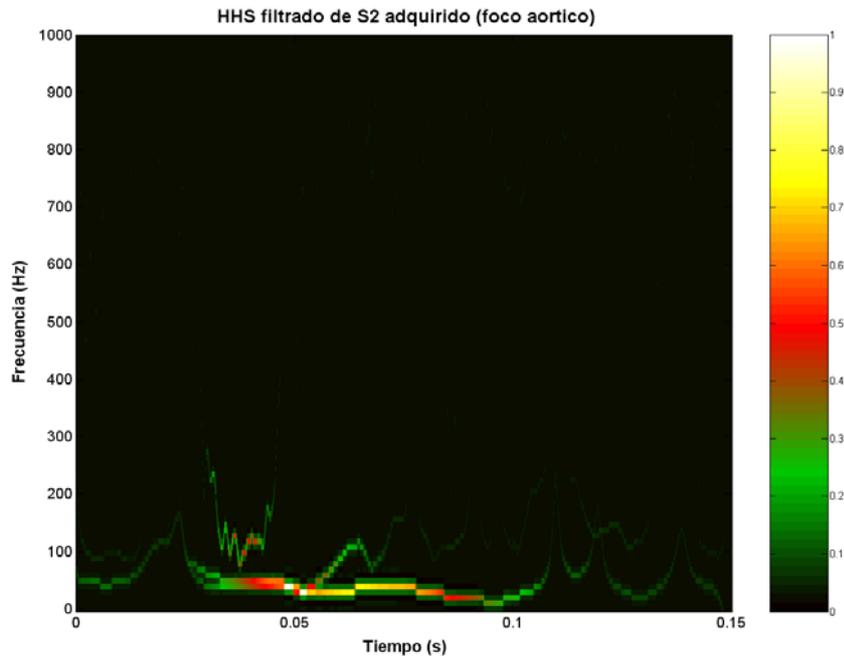


Fig. VI.80. Versión suavizada del Espectro de Hilbert-Huang mostrado en la figura VI.79.

- **S2 adquirido 1 (foco pulmonar).**

La señal de S2 adquirida de forma simultánea a la señal de la figura VI.76, pero sobre el foco pulmonar de auscultación se presenta en la figura VI.81. En la figura VI.82 se presenta su descomposición empírica de modos. La señal analítica de sus dos modos de oscilación intrínsecos seleccionados de acuerdo al criterio de energía se presentan en la figura VI.83.

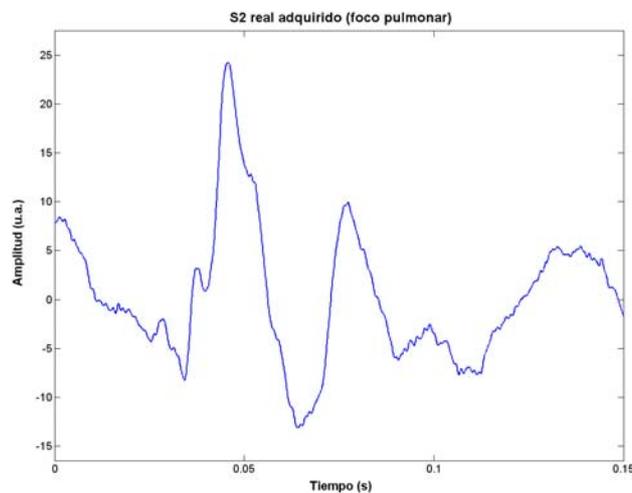


Fig. VI.81. Segundo sonido cardiaco adquirido sobre el foco pulmonar.

El espectro de Hilbert-Huang de la señal de S2 adquirida sobre el foco pulmonar se presenta en la figura VI.84 mientras que una versión suavizada se presenta en la figura VI.85.

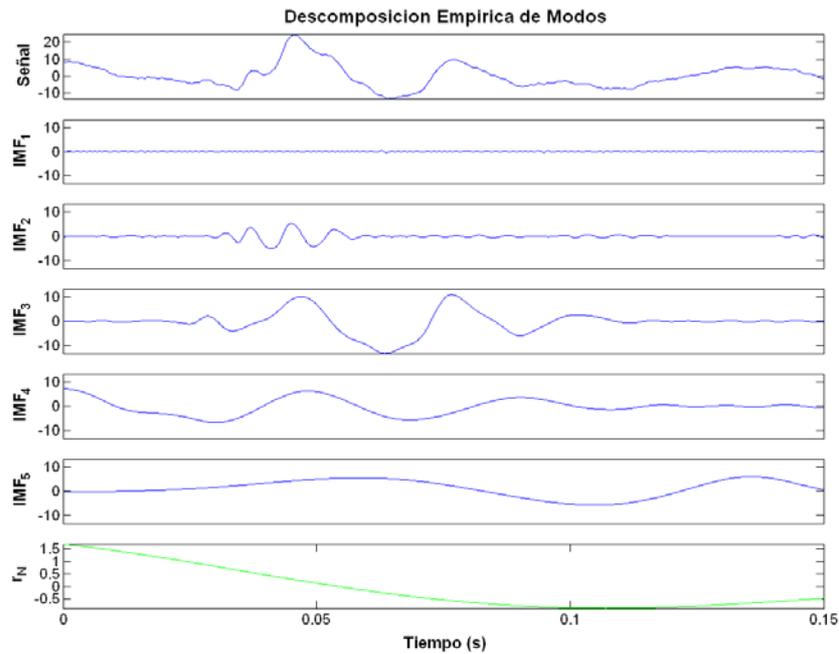


Fig. VI.82. Descomposición empírica de modos de la señal de S2 de la figura VI.81.

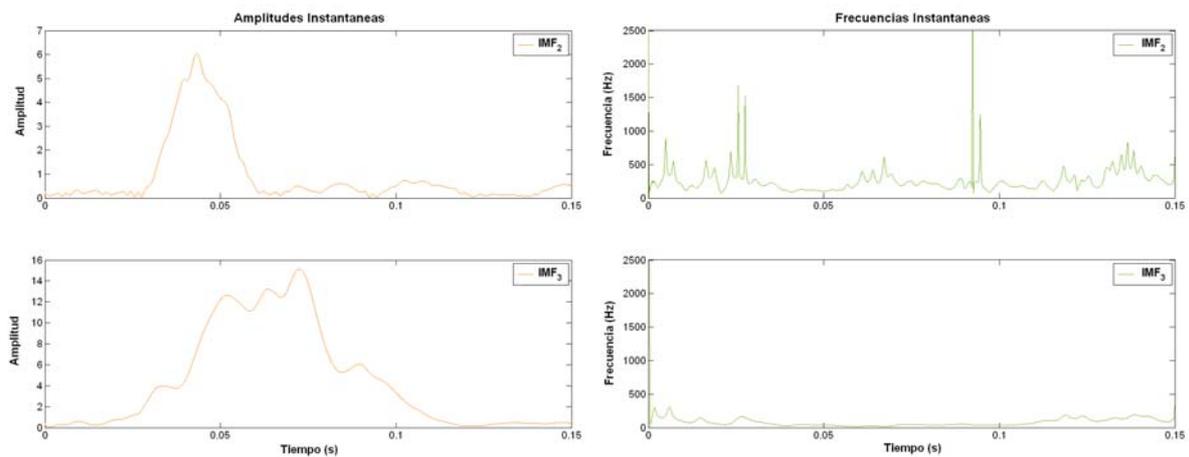


Fig. VI.83. Señales analíticas de los modos de oscilación intrínsecos seleccionados. Amplitud instantánea (columna izquierda) y Frecuencia instantánea (columna derecha).

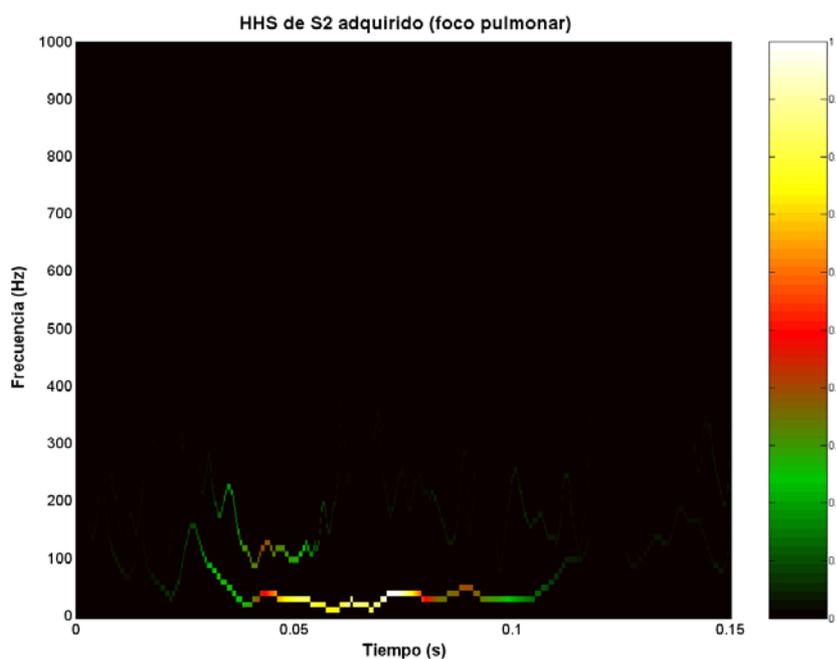


Fig. VI.84. Espectro de Hilbert-Huang del segundo sonido cardiaco mostrado en la figura VI.81.

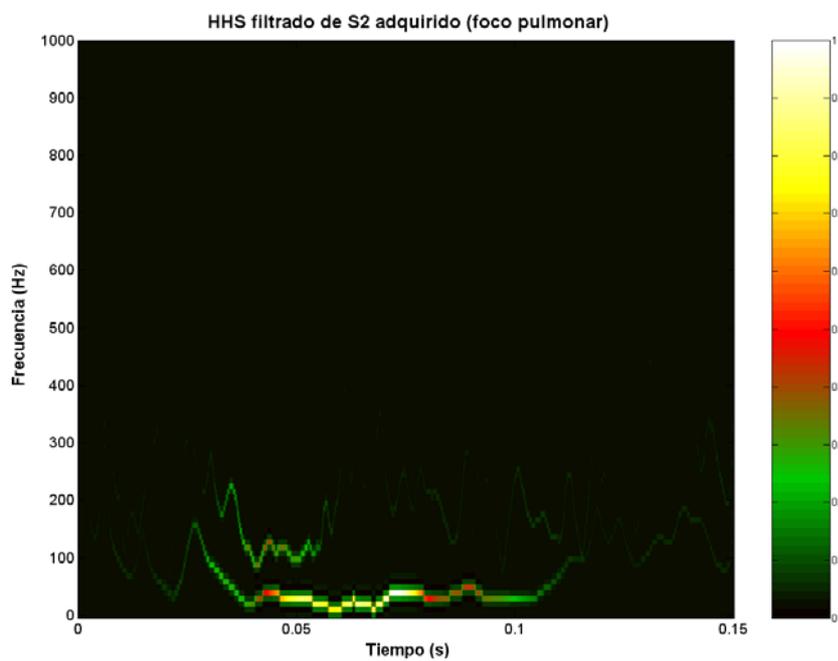


Fig. VI.85. Versión suavizada del Espectro de Hilbert-Huang mostrado en la figura VI.84.

- **S2 adquirido 2 (foco aórtico).**

Una segunda señal de segundo sonido cardiaco adquirida sobre el foco aórtico se presenta en la figura VI.86. Su descomposición empírica de modos se muestra en la figura VI.87. La señal analítica de sus modos de oscilación intrínsecos seleccionados se presenta en la figura VI.88. El Espectro de Hilbert-Huang de este segundo sonido cardiaco se presenta en la figura VI.89. De manera complementaria, en la figura VI.90 se presenta una versión suavizada del Espectro de Hilbert de esta señal adquirida.

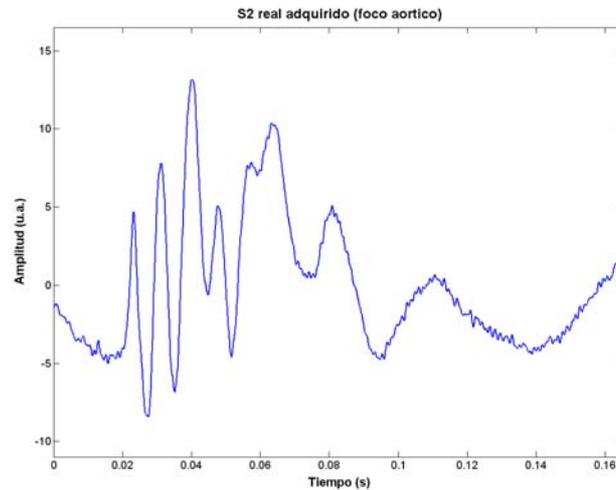


Fig. VI.86. Segundo sonido cardiaco adquirido sobre el foco aórtico.

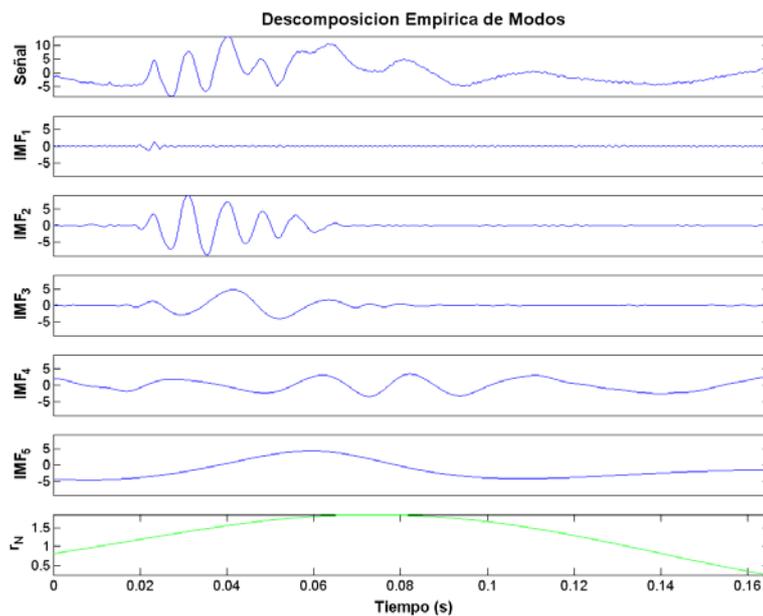


Fig. VI.87. Descomposición empírica de modos de la señal de S2 de la figura VI.86.

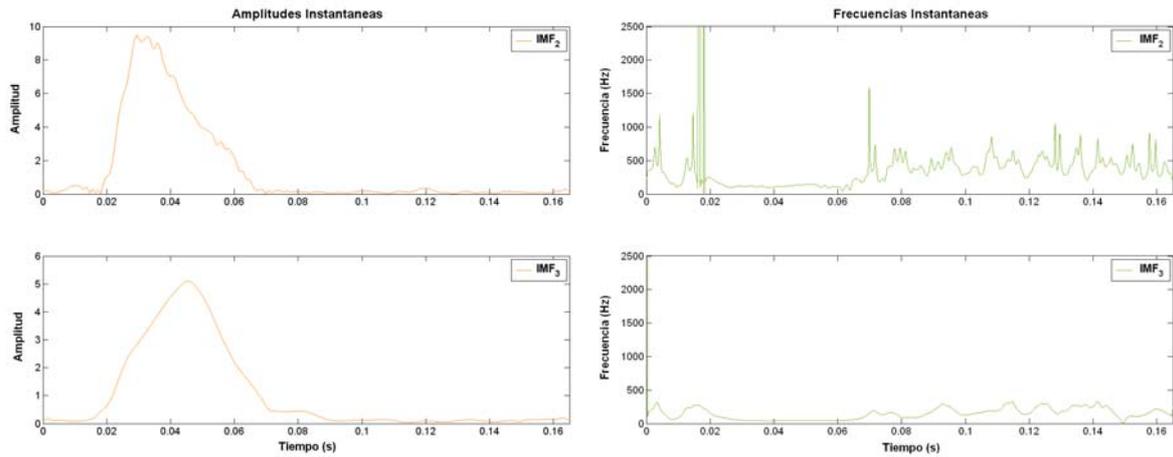


Fig. VI.88. Señales analíticas de los modos de oscilación intrínsecos seleccionados. Amplitud instantánea (columna izquierda) y Frecuencia instantánea (columna derecha).

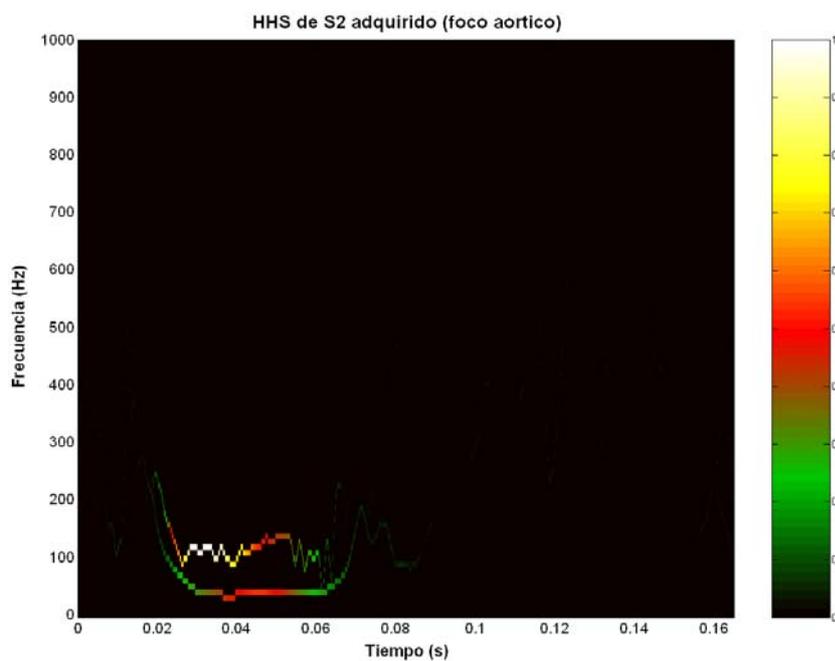
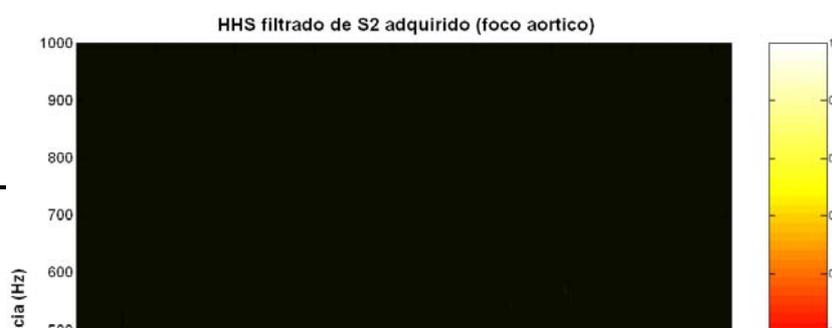


Fig. VI.89. Espectro de Hilbert-Huang del segundo sonido cardiaco mostrado en la figura VI.86.



- **S2 adquirido 2 (foco pulmonar).**

La señal de S2 adquirida sobre el foco pulmonar de auscultación que complementa a la señal de la figura VI.86, se presenta en la figura VI.91. La descomposición empírica de modos de S2 adquirida sobre el foco pulmonar se presenta en la figura VI.92. En la figura VI.93 se presenta la señal analítica de sus tres primeros modos de oscilación intrínsecos seleccionados para conformar el Espectro de Hilbert, mismo que se presenta en la figura VI.94. Una versión suavizada de este Espectro de Hilbert se presenta en la figura VI.95.

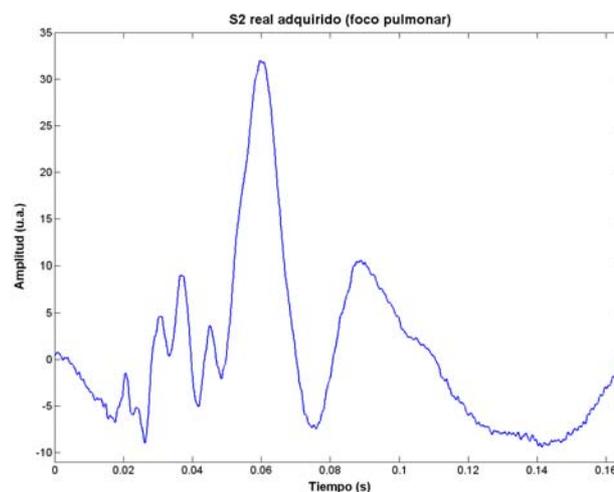


Fig. VI.91. Segundo sonido cardiaco adquirido sobre el foco pulmonar.

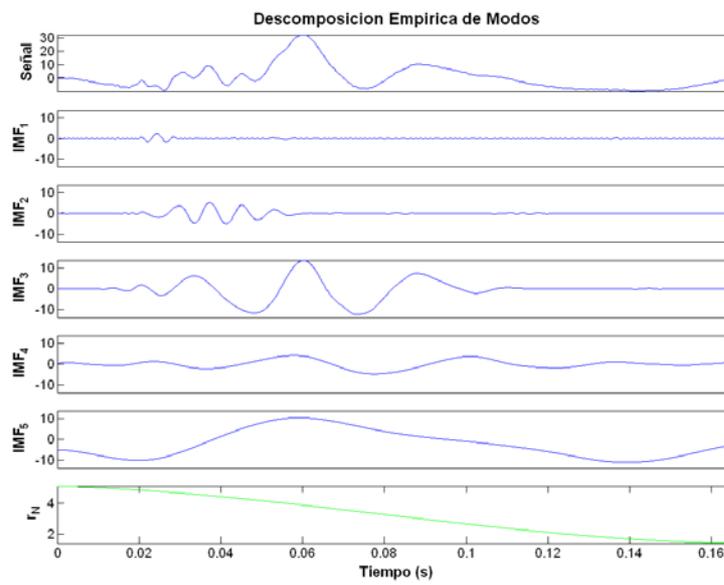


Fig. VI.92. Descomposición empírica de modos de la señal de S2 de la figura VI.91.

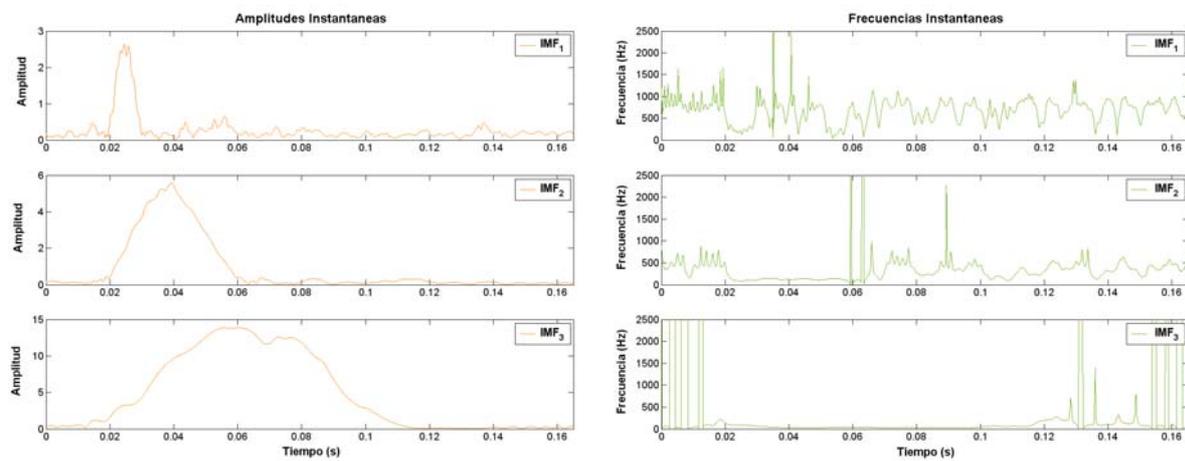


Fig. VI.93. Señales analíticas de los tres primeros modos de oscilación intrínsecos. Amplitud instantánea (columna izquierda) y Frecuencia instantánea (columna derecha).

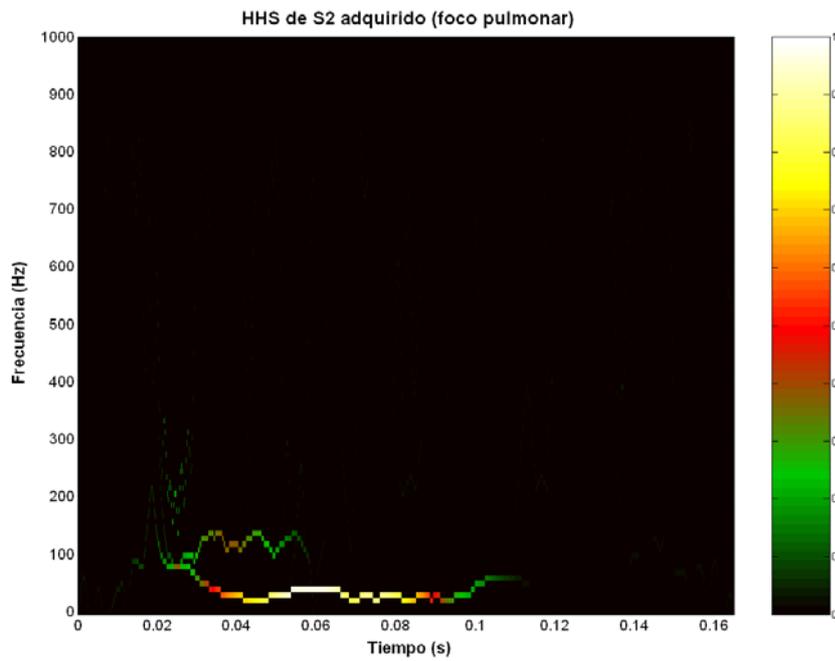


Fig. VI.94. Espectro de Hilbert-Huang del segundo sonido cardiaco mostrado en la figura VI.91.

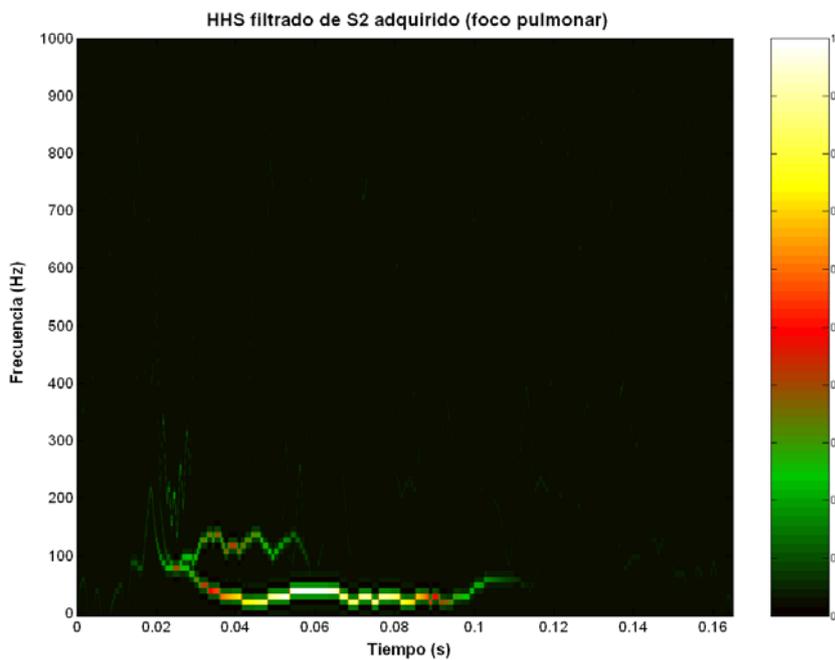


Fig. VI.95. Versión suavizada del Espectro de Hilbert-Huang mostrado en la figura VI.94.

- **S2 adquirido (sin filtrar).**

Con el propósito de observar los modos de oscilación de una señal con ruido de alta frecuencia, en la figura VI.96 se presenta la señal de S2 mostrada en la figura VI.91 sin pre-procesarla con el filtro FIR pasa-bajas con frecuencia de corte de 1 kHz.

La descomposición empírica de modos de S2 sin filtrar se presenta en la figura VI.97.

Su espectro de Hilbert-Huang se presenta en la figura VI.98.

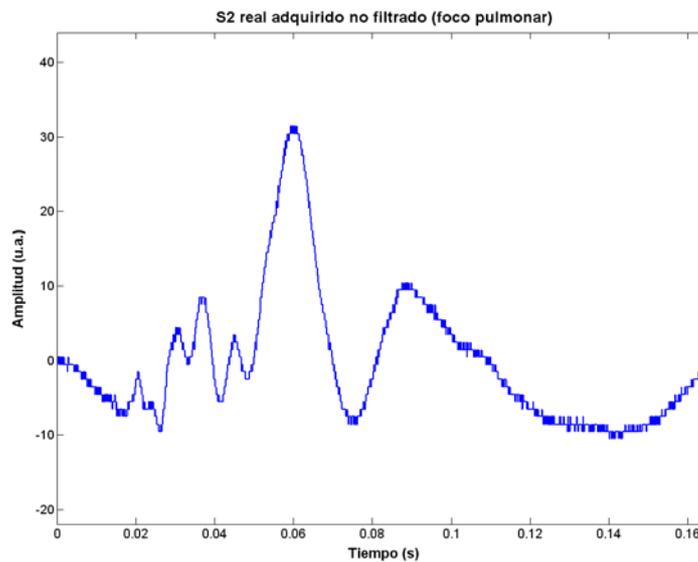


Fig. VI.96. Segundo sonido cardíaco adquirido sobre el foco pulmonar sin filtrar.

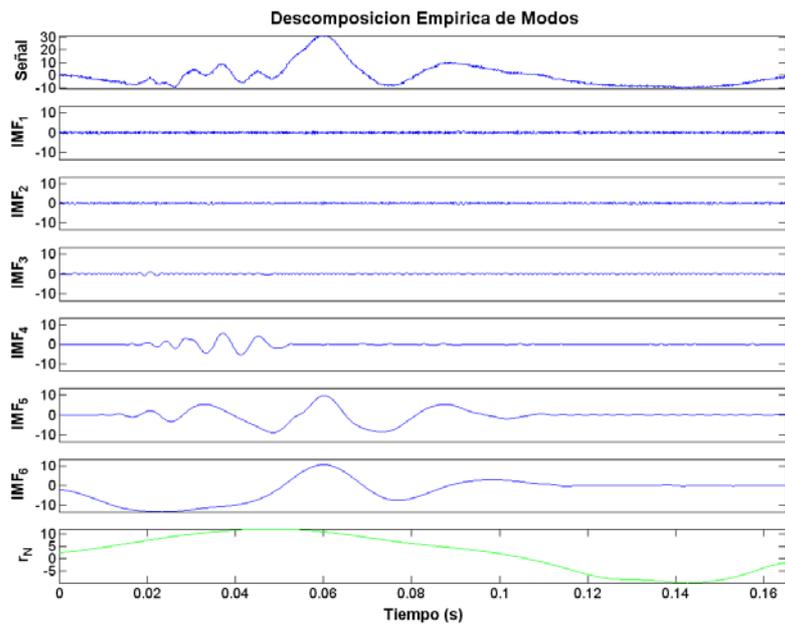


Fig. VI.97. Descomposición empírica de modos de la señal de S2 de la figura VI.96.

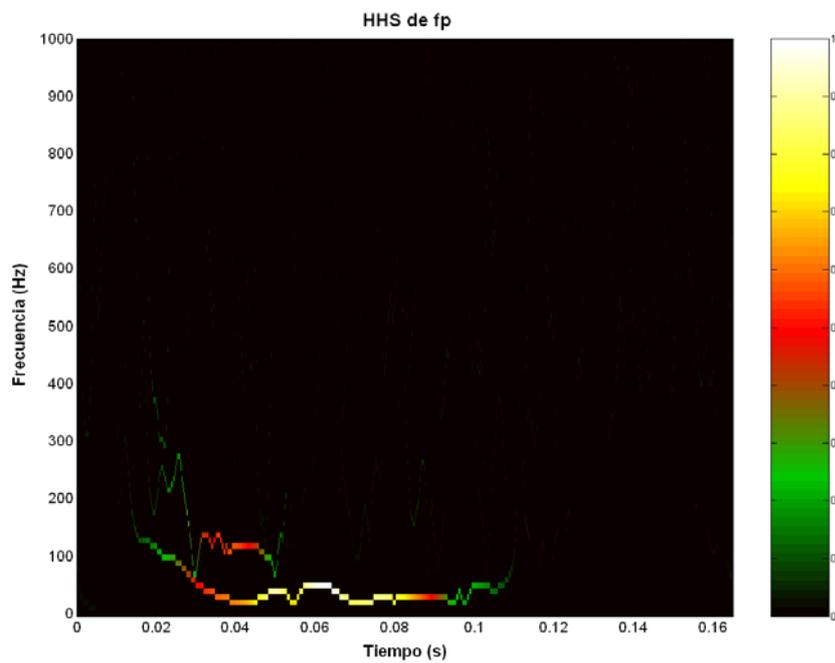


Fig. VI.98. Espectro de Hilbert-Huang del segundo sonido cardíaco mostrado en la figura VI.96.

VI.6. Estertor sibilante simulado.

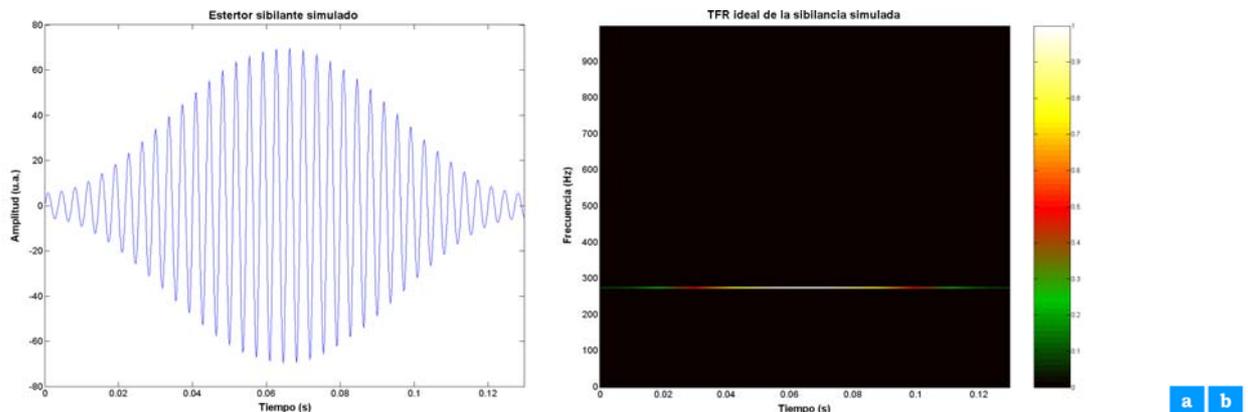


Fig. VI.99. Sibilancia monofónica simulada. a) Formas de onda. b) Representación tiempo-frecuencia teórica

En forma similar al caso de las señales simuladas no biológicas y de sonidos cardiacos, se obtuvieron todas las representaciones tiempo-frecuencia utilizadas en la presente tesis para un sonido respiratorio adventicio continuo simulado, denominado estertor sibilante (figura VI.99.a). Se generaron las sub-imágenes de correlación local, se realizó el cálculo de la correlación promedio y se seleccionaron las mejores representaciones tiempo-frecuencia con base en el criterio del máximo valor de correlación promedio.

Para el caso del espectrograma, la mejor representación tiempo-frecuencia fue la generada con una ventana Hamming de 65 ms de longitud. En el Anexo I se incluyen la función de ambigüedad de la sibilancia simulada, cada espectrograma generado al variar el tipo de ventana y su longitud, la tabla con los valores de su correlación promedio con la representación tiempo-frecuencia teórica de la señal, y el espectrograma seleccionado para esta sibilancia simulada.

La distribución Wigner-Ville de la sibilancia simulada generó un valor de correlación promedio $\rho_{promedio} = 0.8762$ con la representación teórica de la figura VI.99.b. La distribución Wigner-Ville de esta señal se presenta en el Anexo I.

Con relación a la Distribución Choi-Williams, la mejor representación se obtuvo con el parámetro $\sigma = 385$. En el Anexo I se presenta el conjunto de distribuciones Choi-Williams generadas para la sibilancia simulada al variar σ , la tabla con sus valores de correlación promedio correspondientes, y la distribución Choi-Williams seleccionada para esta señal.

La mejor representación tiempo-frecuencia vía el espectrograma reasignado se logró con una ventana Hamming de 13 ms de longitud. Los espectrogramas reasignados generados al variar el tipo de ventana y su longitud, los resultados de su comparación local con la representación tiempo-frecuencia ideal de la sibilancia simulada, y el espectrograma reasignado seleccionado, se presentan en el Anexo I.

Respecto de la representación obtenida vía el modelo AR con el método de Burg, la mejor representación tiempo-frecuencia se obtuvo con un modelo AR de orden 4 y una ventana Hamming de 65 ms de longitud. Los valores de correlación promedio de cada representación tiempo-frecuencia de Burg que se obtuvo al variar el orden del modelo AR, el tipo de ventana y su longitud, así como cada una de estas representaciones, se presentan en el Anexo I. En el mismo anexo también se incluye la representación P_{Burg} seleccionada para el estertor sibilante simulado.

Con relación a la representación tiempo-frecuencia generada vía el algoritmo RLS, la mejor se obtuvo con un modelo AR de orden 4 y un factor de olvido igual a 0.75. Cada representación P_{RLS} obtenida con diversos ordenes del modelo AR y valores del factor de olvido para la sibilancia simulada, se presenta en el Anexo I junto con la tabla de valores de correlación promedio correspondientes y la representación P_{RLS} seleccionada.

Para el algoritmo RLS-VFF la mejor representación tiempo-frecuencia se obtuvo con un modelo AR de orden 2 y una varianza del ruido esperado igual a 1×10^{-6} . En el Anexo I se presenta cada representación tiempo-frecuencia $P_{\text{RLS-VFF}}$ obtenida para diversos órdenes del modelo AR y valores de la varianza del ruido esperado, así como los resultados de su comparación local con la representación tiempo-frecuencia teórica. La representación $P_{\text{RLS-VFF}}$ seleccionada para esta sibilancia simulada también se presenta en dicho anexo.

Para la representación tiempo-frecuencia obtenida vía el filtro de Kalman, la mejor se obtuvo con los parámetros siguientes: modelo AR de orden $p = 2$, varianza del ruido de medición $\sigma^2 = 1$ y varianza del ruido del proceso $q = 0.001$. En el Anexo I se presentan las representaciones P_{Kalman} generadas para la sibilancia simulada, así como sus valores de correlación promedio y la representación tiempo-frecuencia seleccionada de esta técnica de análisis.

Para el escalograma, la mejor representación se obtuvo con la wavelet Morlet de 20 ms de longitud en la escala más gruesa. En el Anexo I se presentan los escalogramas Morlet de la sibilancia simulada, generado al variar la longitud de la wavelet. En el mismo anexo se incluyen los resultados de la comparación local de estos escalogramas con la representación

tiempo-frecuencia teórica de la señal simulada, así el escalograma Morlet seleccionado para el estertor sibilante simulado.

Con relación al espectro de Hilbert-Huang de la sibilancia simulada, se obtuvo un valor de $\rho_{promedio}$ igual a 0.9521 a través de su comparación local con la representación teórica. En el Anexo I se presenta la descomposición empírica de modos del estertor sibilante simulado, la señal analítica del único modo de oscilación intrínseco resultante, y el espectro de Hilbert-Huang que resulta del mapeo de este modo de oscilación. Finalmente, en el mismo anexo se presenta la comparación local del espectro de Hilbert-Huang con la representación tiempo-frecuencia teórica de la sibilancia simulada.

▪ **Comparación de las técnicas tiempo-frecuencia.**

En la figura VI.100 se presentan las representaciones tiempo-frecuencia seleccionadas de acuerdo al criterio de correlación local para la sibilancia simulada. La estimación de la frecuencia instantánea de la señal a partir del centroide de cada una de estas representaciones seleccionadas, presenta en la figura VI.101. En la tabla VI.XI se presentan los valores de los índices de desempeño obtenidos para cada una de las representaciones tiempo-frecuencia mostrados en la figura VI.100.

TABLA VI.XI. INDICES DE DESEMPEÑO DE LAS REPRESENTACIONES TIEMPO-FRECUENCIA SELECCIONADAS PARA LA SIBILANCIA SIMULADA.

	ρ	$\rho_{promedio}$	NRMSE	ρ_{FI}	res_{TF}
SP	0.5300	0.8851	0.0074	1.0000	0.0029
WVD	0.7325	0.8762	0.2139	0.9769	0.0009
CWD	0.6882	0.8965	0.0124	0.9999	0.0017
SPR	0.8821	0.9163	0.1400	0.9903	5.83×10^{-5}
P_{BURG}	0.4564	0.8798	0.0199	0.9999	0.0048
P_{RLS}	0.4804	0.8563	0.1996	0.9800	0.0026
P_{RLS-VFF}	0.4367	0.8560	0.1991	0.9800	0.0086
P_{KALMAN}	0.4350	0.8570	0.1991	0.9800	0.0086
SC	0.6214	0.9222	0.0114	1.0000	0.0033
HHS	0.9247	0.9521	0.0601	0.9982	2.47×10^{-6}

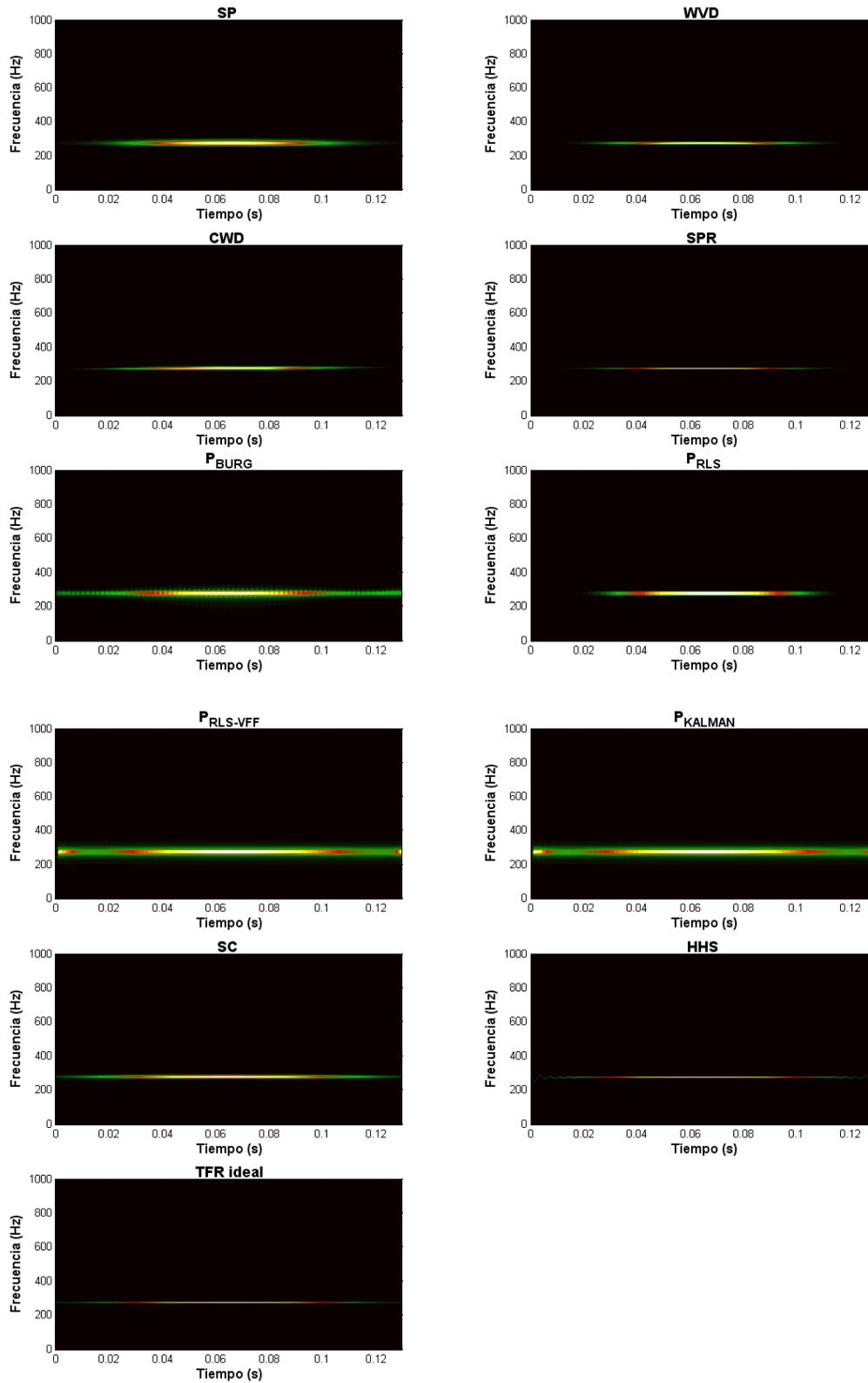


Fig. VI.100. Representaciones tiempo-frecuencia seleccionadas para cada técnica de análisis tiempo-frecuencia así como la representación tiempo-frecuencia teórica de la sibilancia simulada.

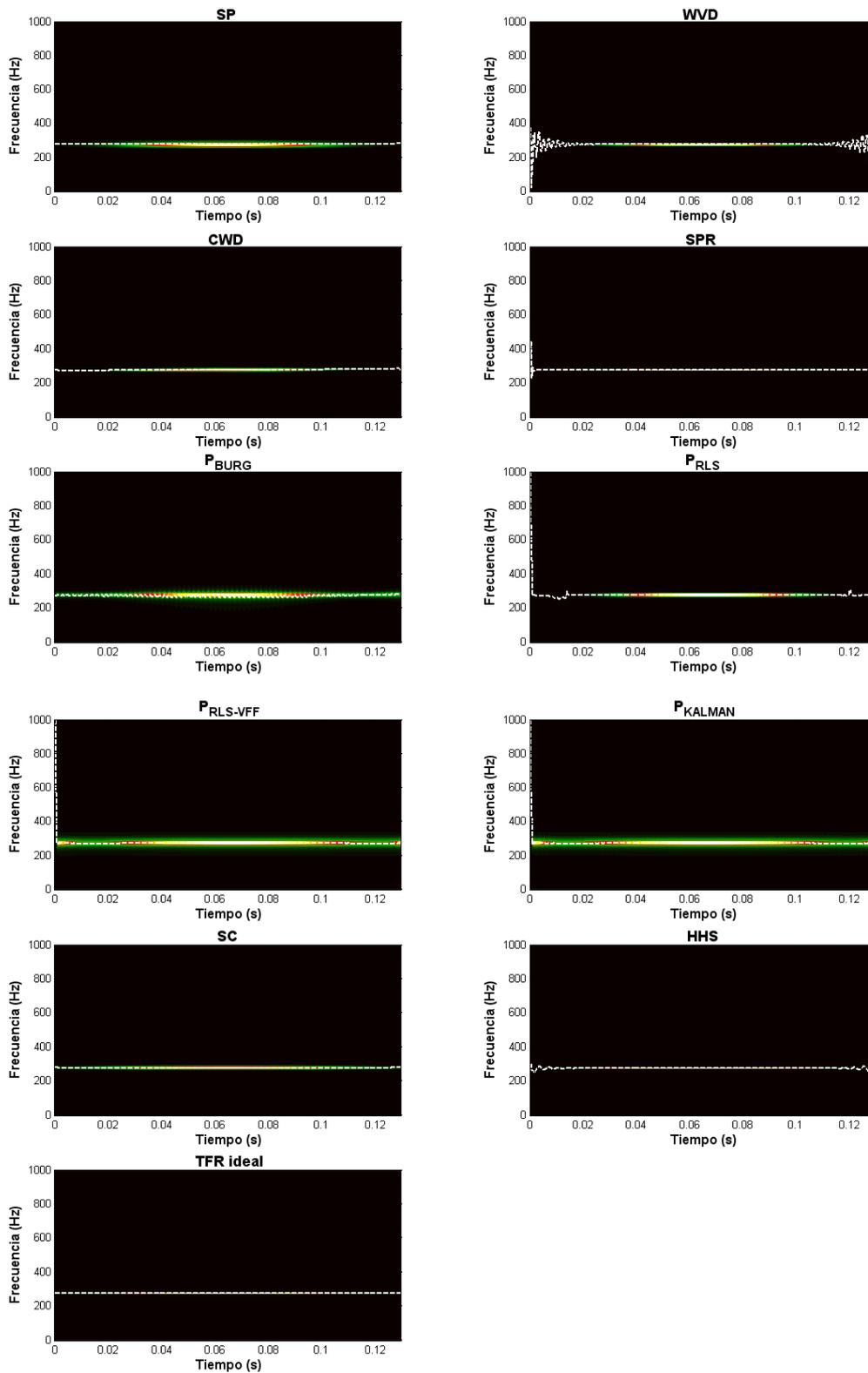


Fig. VI.101. Estimación de la frecuencia instantánea mediante el primer momento de cada representación tiempo-frecuencia seleccionada así como teórica del estertor sibilante simulado.

VI.6.1. Estertores sibilantes reales adquiridos.

▪ Sibilancia real 1.

En la figura VI.102 se presenta una sibilancia real adquirida de un sujeto con Neumopatía Intersticial Difusa (NID). La descomposición empírica de modos de la sibilancia se presenta en la figura VI.103. La señal analítica del primer modo de oscilación intrínseco se presenta en la figura VI.104. En las figura VI.105 y VI.106 se presentan el Espectro de Hilbert-Huang de la sibilancia adquirida y una versión suavizada, respectivamente. En la figura VI.107 se presenta el espectrograma de la señal adquirida, calculado con una ventana del tipo y longitud indicados por los resultados de la simulación.

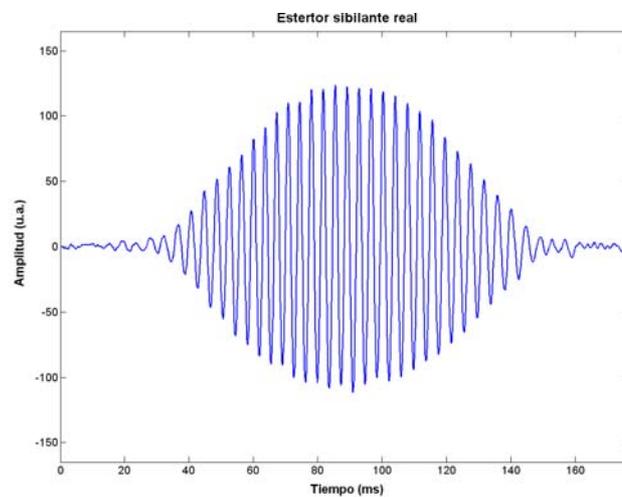


Fig. VI.102. Forma de onda de una sibilancia adquirida.

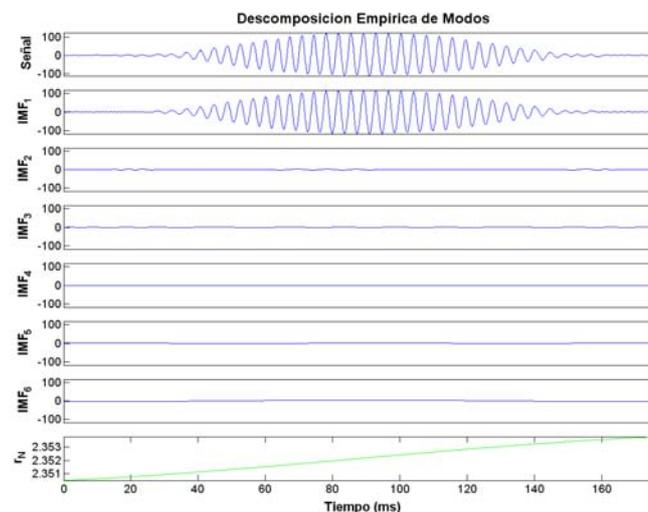


Fig. VI.103. Descomposición empírica de modos del estertor sibilante de la figura VI.102.

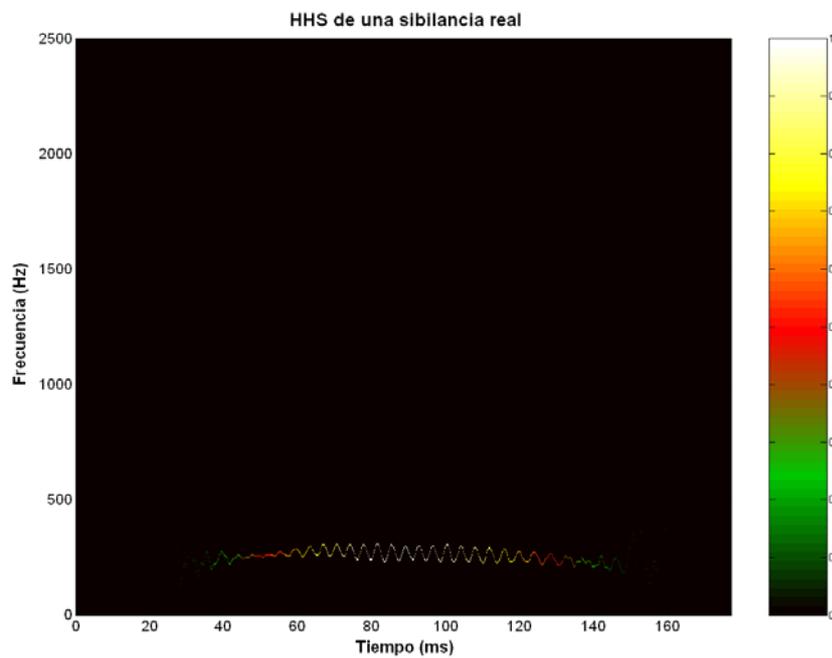
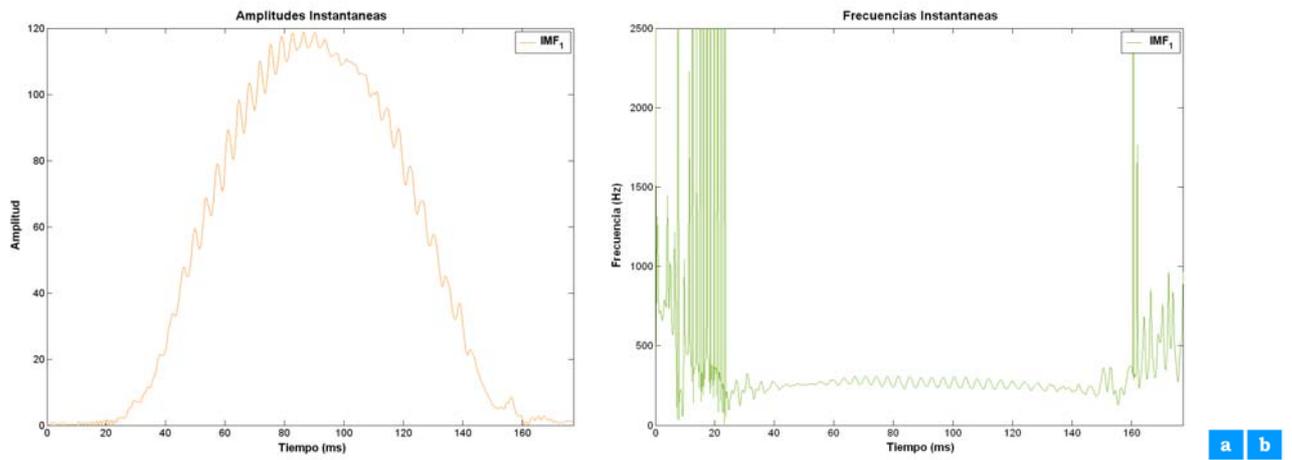


Fig. VI.105. Espectro de Hilbert-Huang de la sibilancia adquirida mostrada en la figura VI.102.

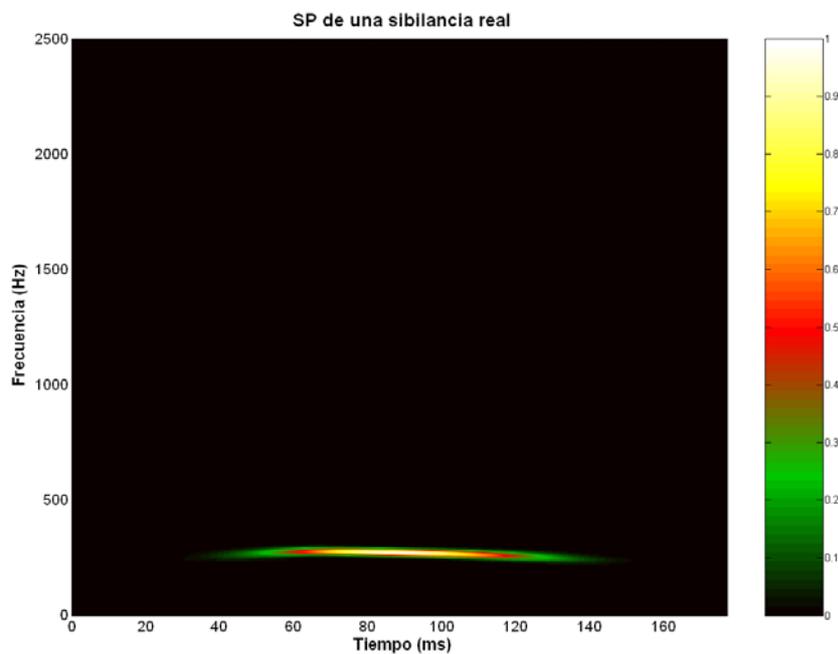


Fig. VI.107. Espectrograma de la sibilancia adquirida mostrada en la VI.102 al utilizar una ventana Hamming de 60 ms de longitud.

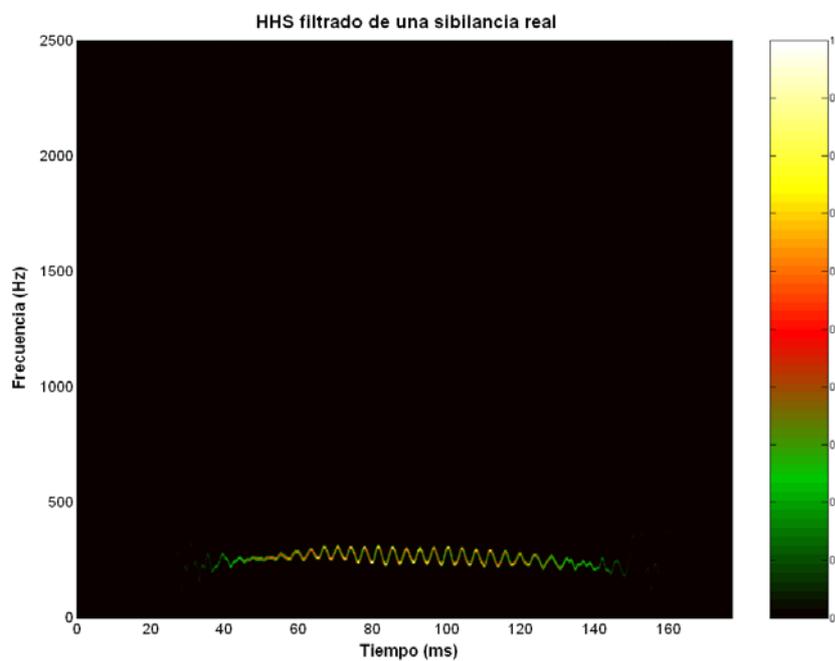


Fig. VI.106. Versión suavizada del Espectro de Hilbert-Huang del estertor sibilante adquirido mostrado en la figura VI.105.

▪ Sibilancia real 2.

En la figura VI.108 se presenta una segunda sibilancia real adquirida de un sujeto con NID mientras que en la figura VI.109 se presenta la descomposición empírica de modos de esta sibilancia. La señal analítica del primer modo de oscilación intrínseco resultante se presenta en la figura VI.110. En la figura VI.111 se presenta el Espectro de Hilbert-Huang de la sibilancia adquirida y una versión suavizada se presenta en la figura VI.112. El espectrograma obtenido para esta segunda sibilancia adquirida se presenta en la figura VI.113.

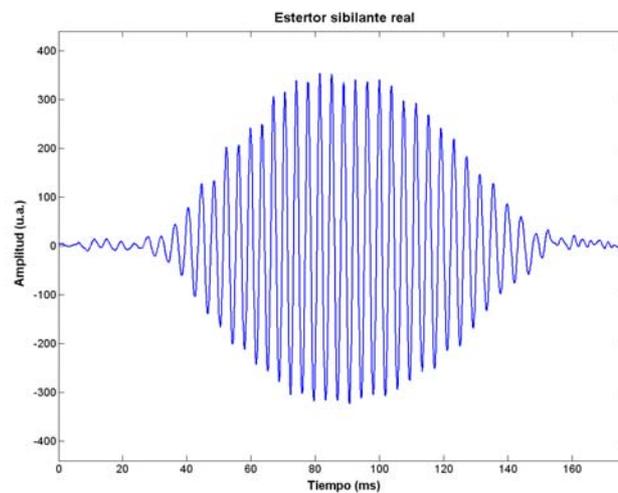


Fig. VI.108. Forma de onda de una segunda sibilancia adquirida.

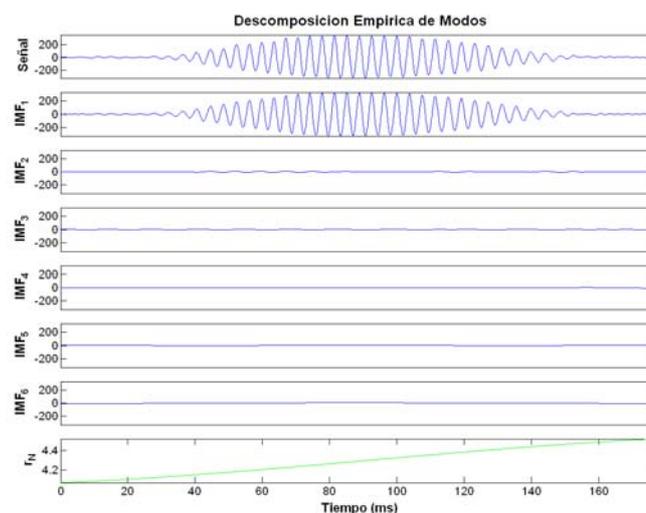


Fig. VI.109. Descomposición empírica de modos del estertor sibilante de la figura VI.108.

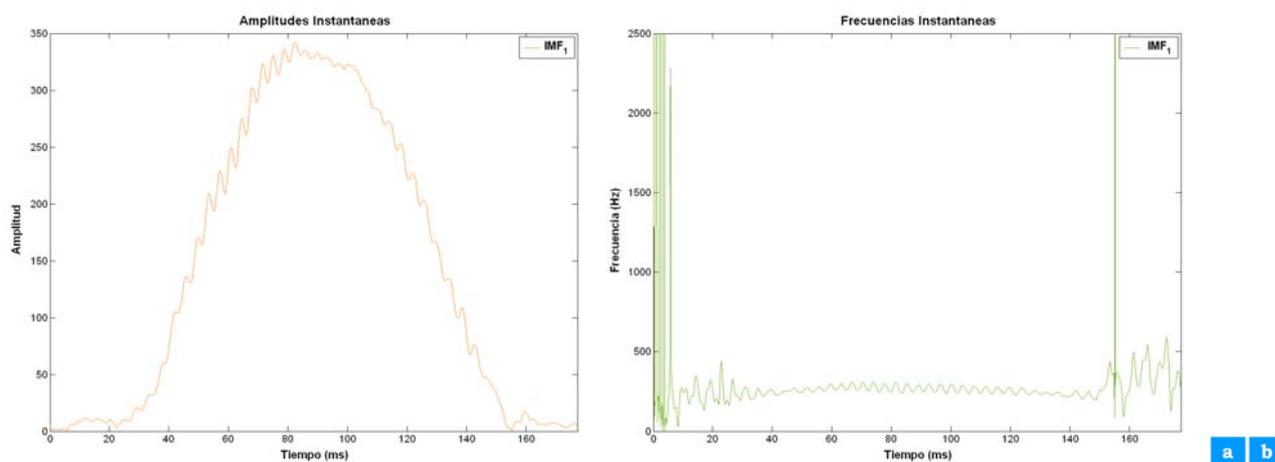


Fig. VI.110. Señal analítica del primer modo de oscilación intrínseco. a) Amplitud instantánea. b) Frecuencia instantánea.

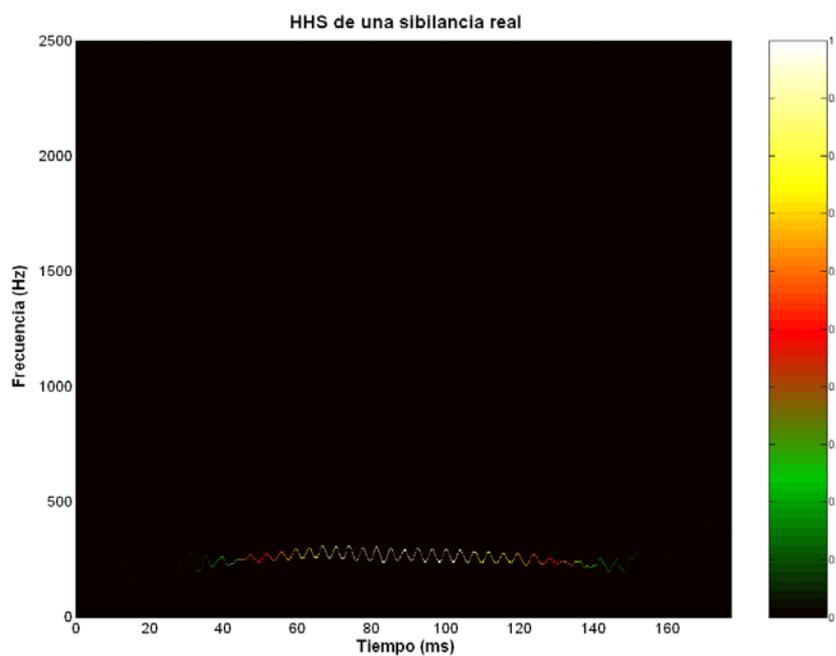


Fig. VI.111. Espectro de Hilbert-Huang de la sibilancia adquirida mostrada en la figura VI.108.

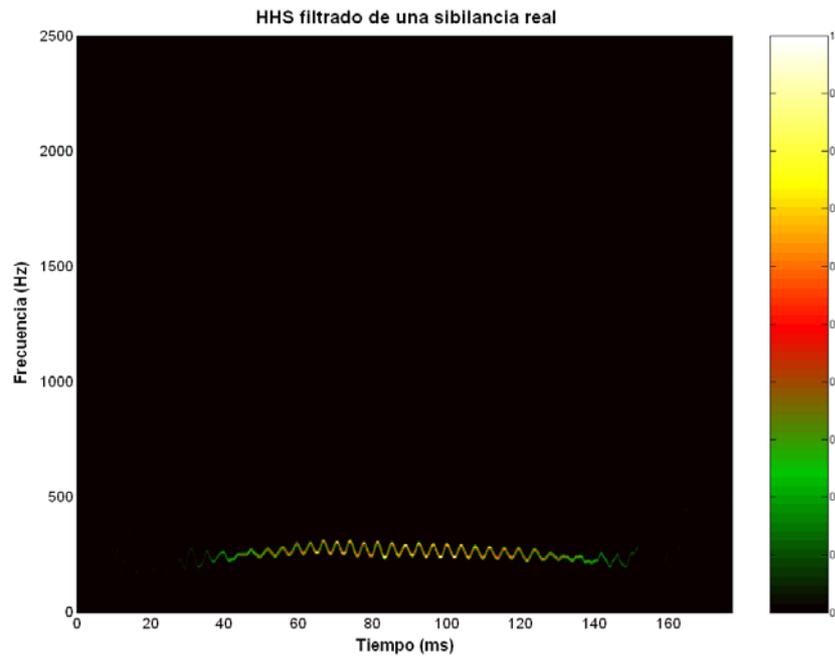


Fig. VI.112. Versión suavizada del Espectro de Hilbert-Huang del estertor sibilante adquirido mostrado en la figura VI.111.

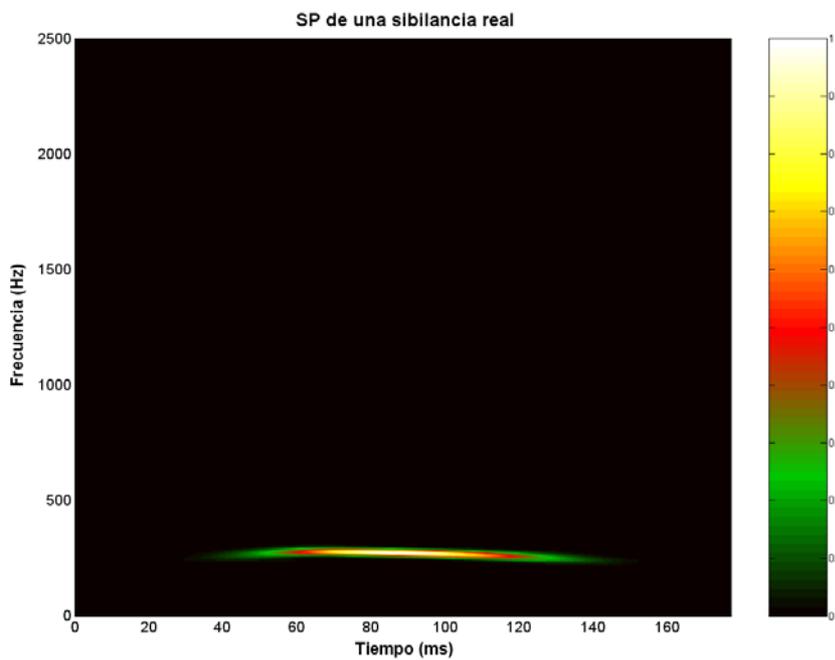


Fig. VI.113. Espectrograma de la sibilancia adquirida mostrada en la VI.108. Ventana Hamming de 60 ms de longitud.

▪ Sibilancia real 3.

Una tercera sibilancia adquirida de un sujeto con NID se presenta en la figura VI.114. La descomposición empírica de modos de esta sibilancia adquirida se presenta en la figura VI.115, mientras que en la figura VI.116 se presenta la señal analítica de su primer modo de oscilación intrínseco. El Espectro de Hilbert-Huang correspondiente se presenta en la figura VI.117. En la figura VI.118 se presenta una versión suavizada de dicho espectro. El espectrograma obtenido para el estertor sibilante adquirido se presenta en la figura VI.119.

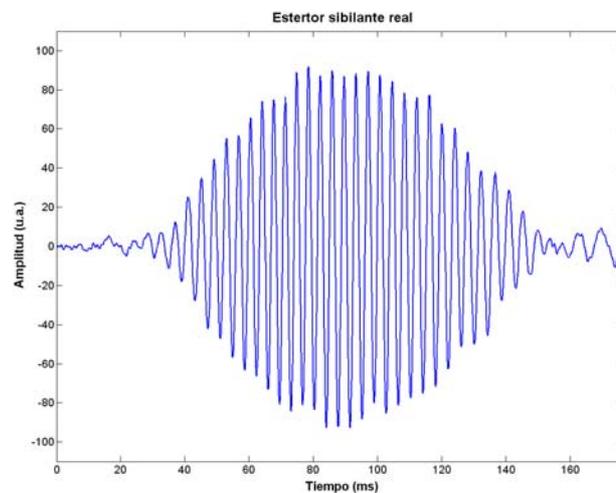


Fig. VI.114. Forma de onda de una sibilancia adquirida.

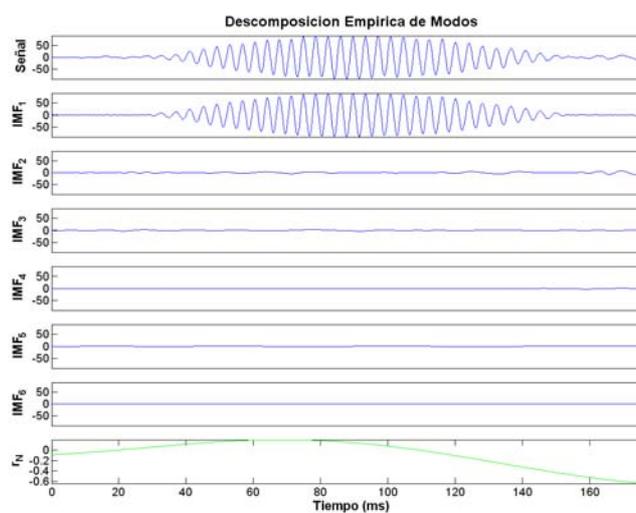


Fig. VI.115. Descomposición empírica de modos del estertor sibilante de la figura VI.114.

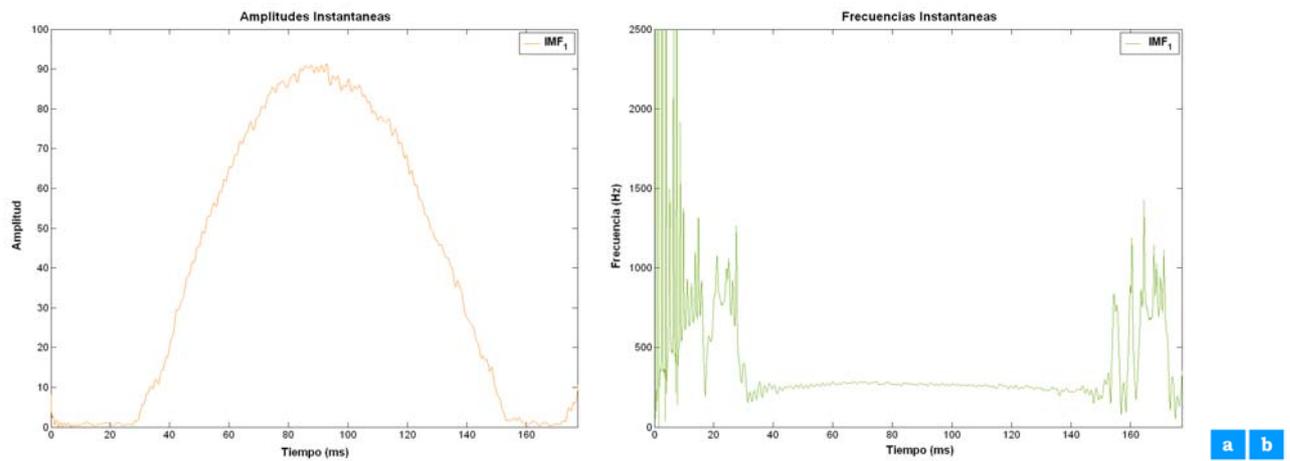


Fig. VI.116. Señal analítica del primer modo de oscilación intrínseco. a) Amplitud instantánea. b) Frecuencia instantánea.

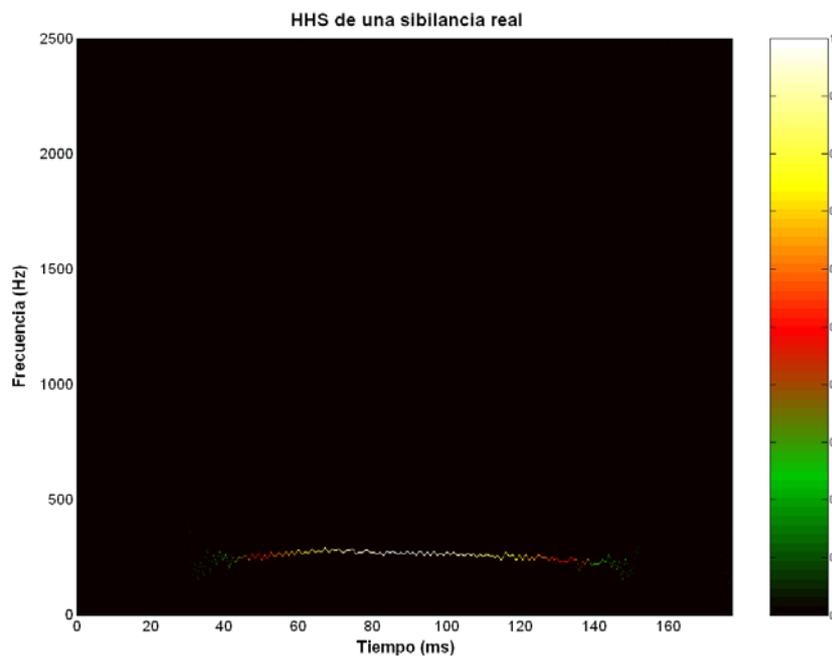


Fig. VI.117. Espectro de Hilbert-Huang de la sibilancia adquirida mostrada en la figura VI.114.

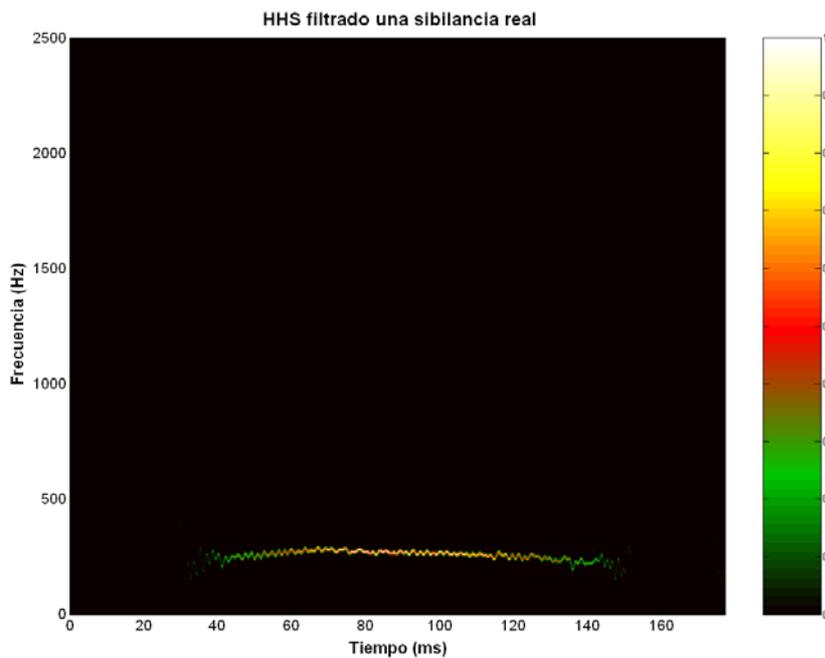


Fig. VI.118. Versión suavizada del Espectro de Hilbert-Huang del estertor sibilante adquirido mostrado en la figura VI.117.

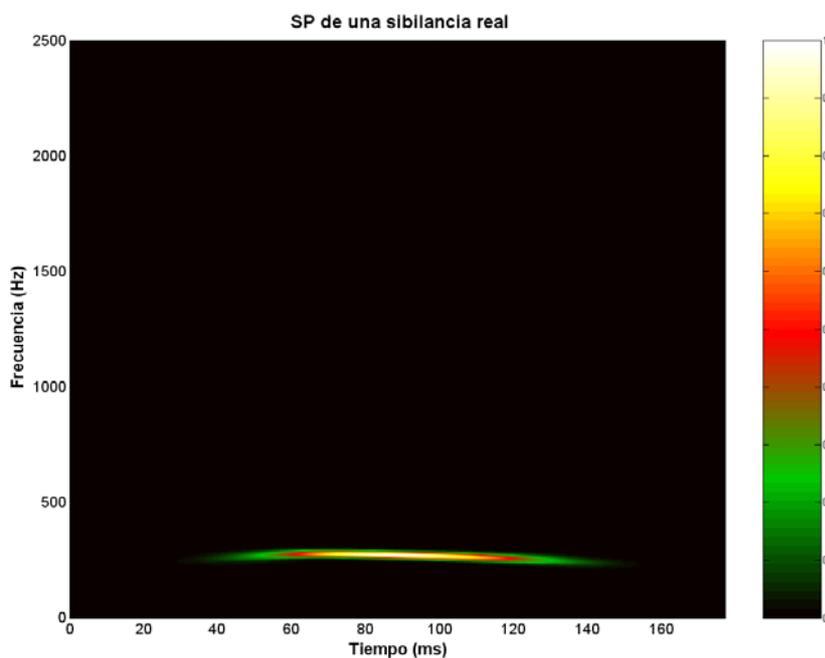


Fig. VI.119. Espectrograma de la sibilancia adquirida mostrada en la VI.114 con una ventana Hamming de 60 ms de longitud.

■ Sibilancia real 4.

En la figura VI.120 se presenta una cuarta sibilancia real adquirida de un sujeto con NID. Su descomposición empírica de modos se muestra en la figura VI.121 y en la figura VI.122 se presenta la señal analítica de su primer modo de oscilación intrínseco. En la figura VI.123 se presenta el Espectro de Hilbert-Huang y una versión suavizada se muestra en la figura VI.124. En la figura VI.125 se presenta el espectrograma obtenido para esta sibilancia adquirida.

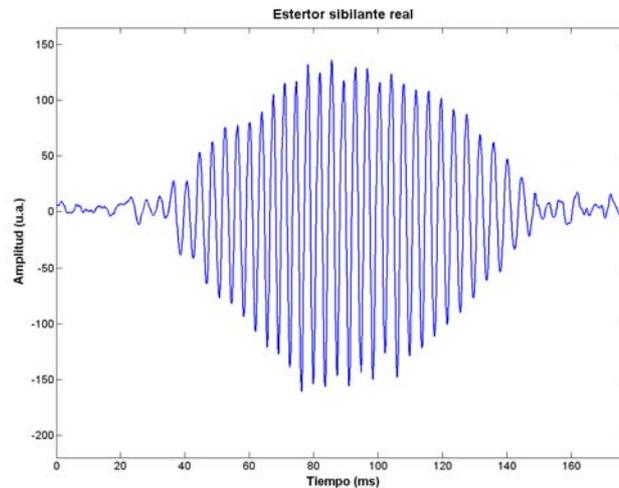


Fig. VI.120. Forma de onda de una sibilancia adquirida.

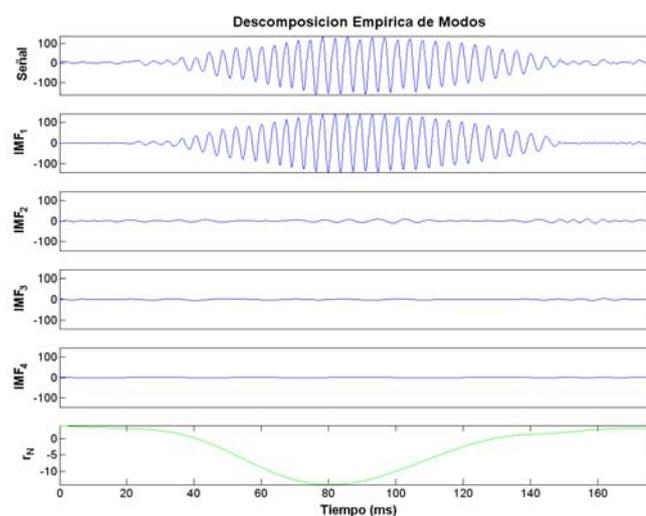


Fig. VI.121. Descomposición empírica de modos del estertor sibilante de la figura VI.120.

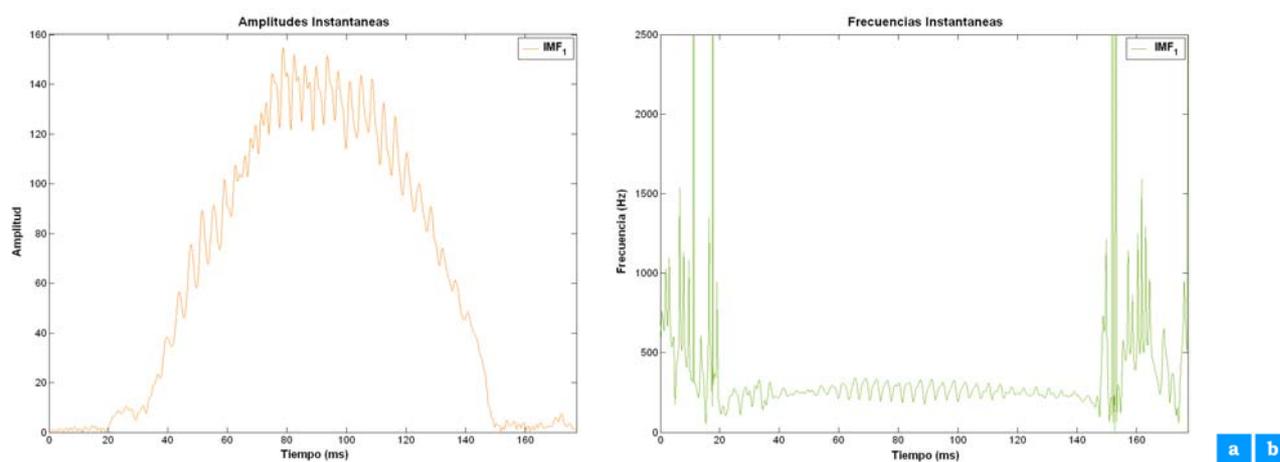


Fig. VI.122. Señal analítica del primer modo de oscilación intrínseco. a) Amplitud instantánea. b) Frecuencia instantánea.

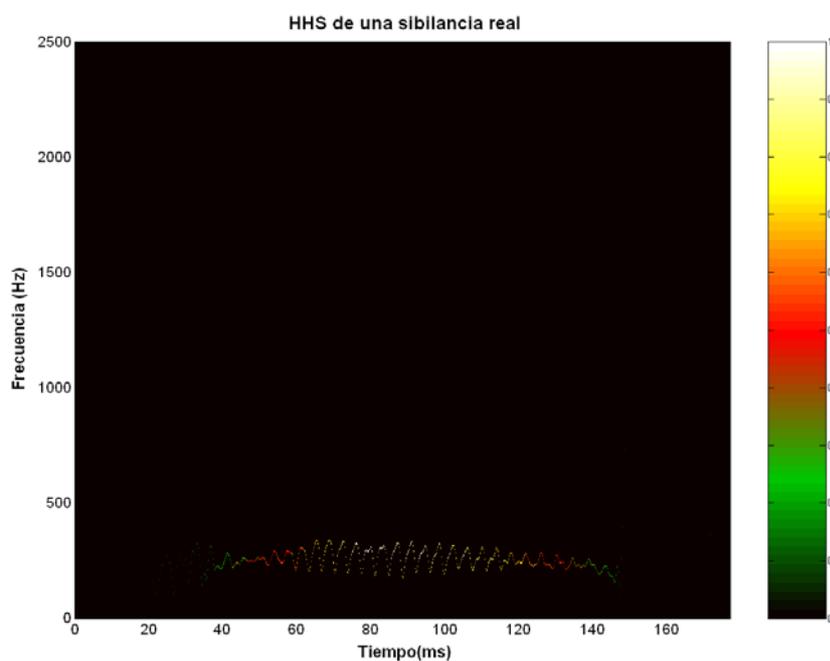


Fig. VI.123. Espectro de Hilbert-Huang de la sibilancia adquirida mostrada en la figura VI.120.

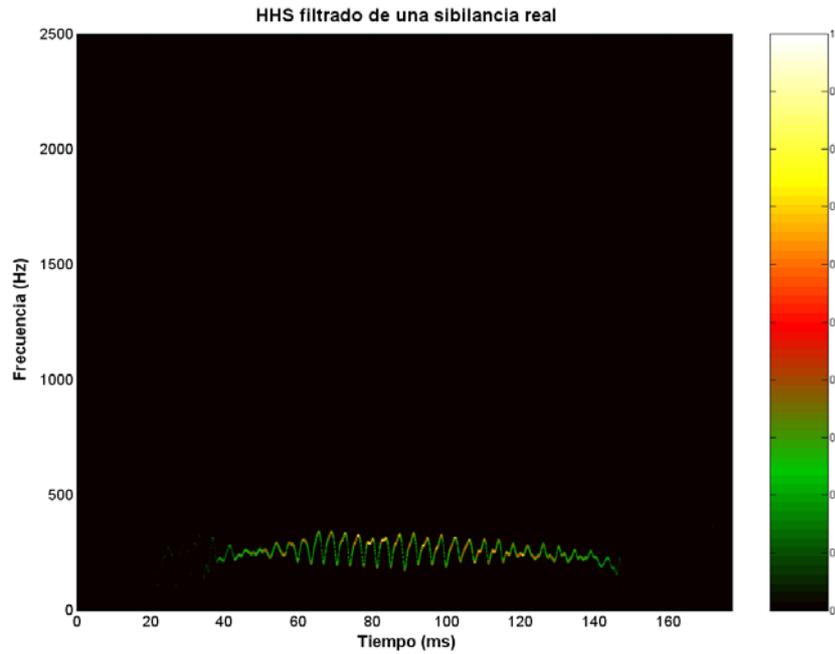


Fig. VI.124. Versión suavizada del Espectro de Hilbert-Huang del estertor sibilante adquirido mostrado en la figura VI.123.

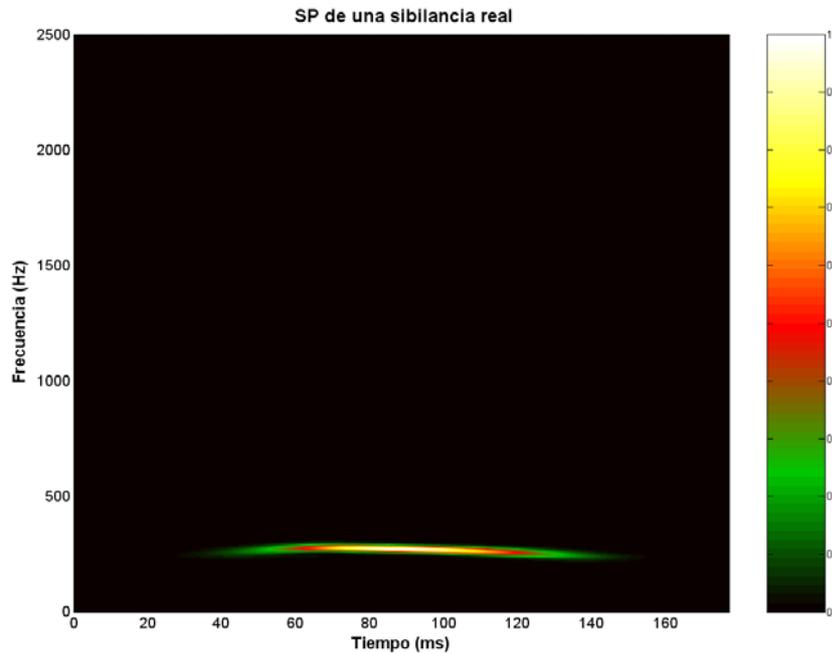


Fig. VI.125. Espectrograma de la sibilancia adquirida de la figura VI.120 con una ventana Hamming de 60 ms de longitud.

▪ Sibilancia real 5.

En la figura VI.126 se presenta un sonido respiratorio real adquirido de un paciente con NID a lo largo de un ciclo espiratorio. La señal adquirida contiene un sonido adventicio continuo así como un sonido adventicio discontinuo. La descomposición empírica de modos de este sonido respiratorio se muestra en la figura VI.127, mientras que en la figura VI.128 se presenta la señal analítica de sus dos primeros modos de oscilación intrínsecos.

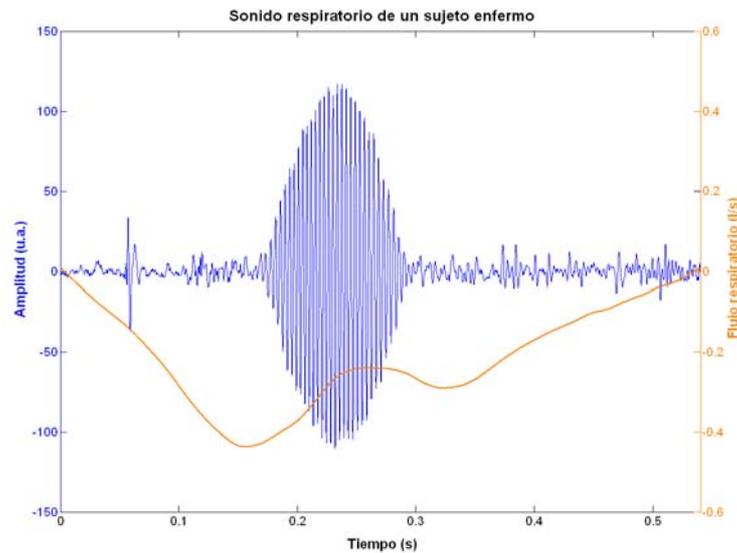


Fig. VI.126. Sonido respiratorio adquirido de un paciente con NID durante la espiración.

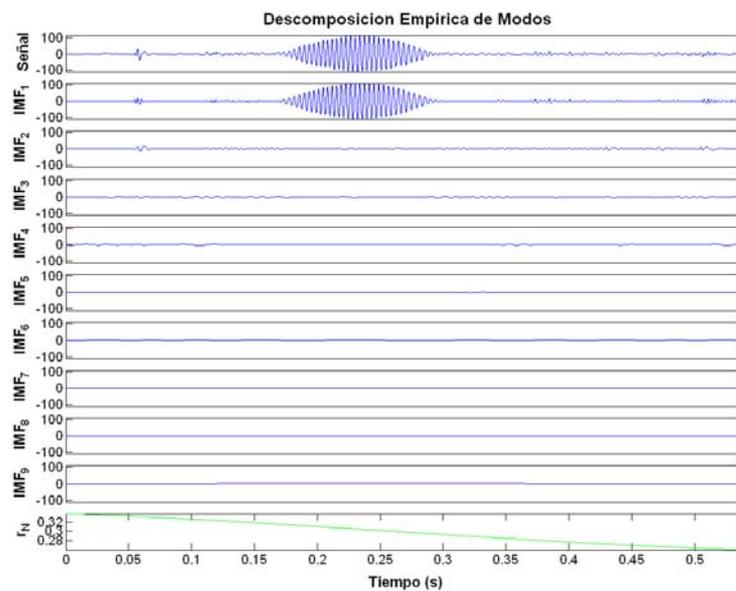


Fig. VI.127. Descomposición empírica de modos del sonido respiratorio de la figura VI.126.

En la figura VI.129 se presenta el Espectro de Hilbert-Huang de la señal y en la figura VI.130 se presenta una versión suavizada de dicho espectro. En la figura VI.131 se presenta el espectrograma obtenido para este sonido respiratorio con una ventana Hamming de 100 ms de longitud, de acuerdo a lo reportado en la bibliografía.

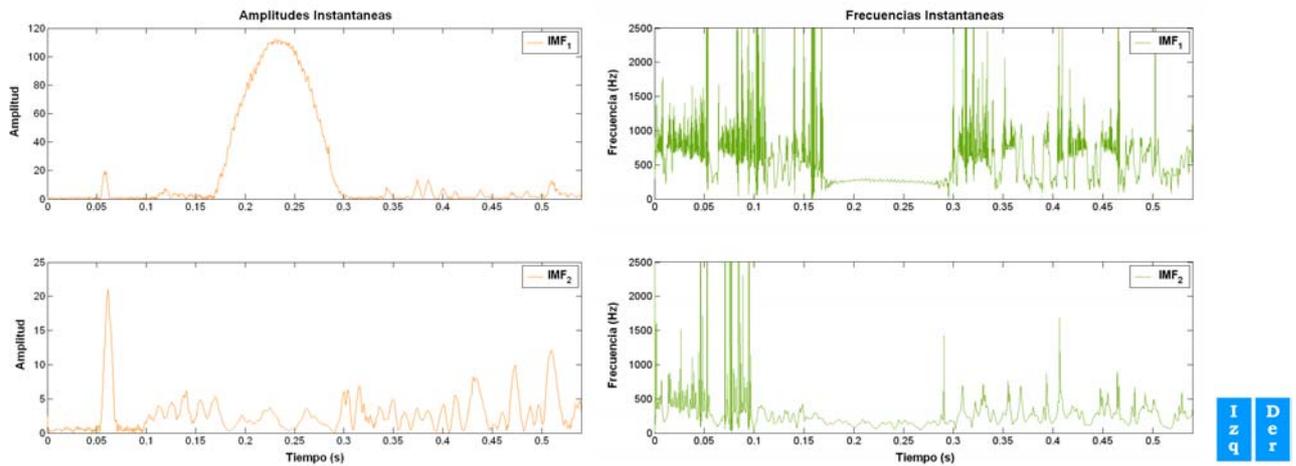


Fig. VI.128. Señal analítica de los dos primeros modos de oscilación intrínsecos. Amplitud instantánea (columna izquierda), y Frecuencia instantánea (columna derecha).

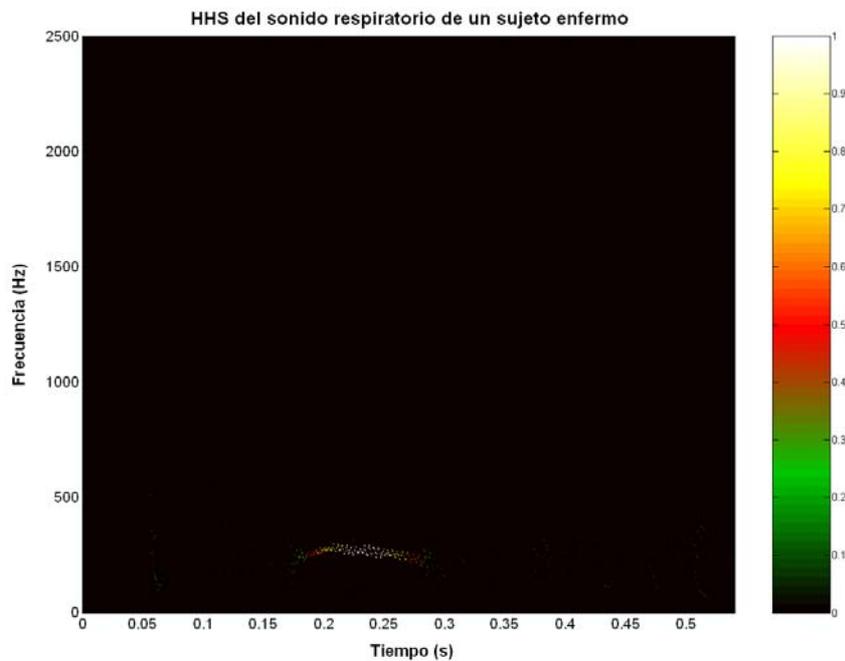


Fig. VI.129. Espectro de Hilbert-Huang del sonido respiratorio mostrado en la figura VI.126.

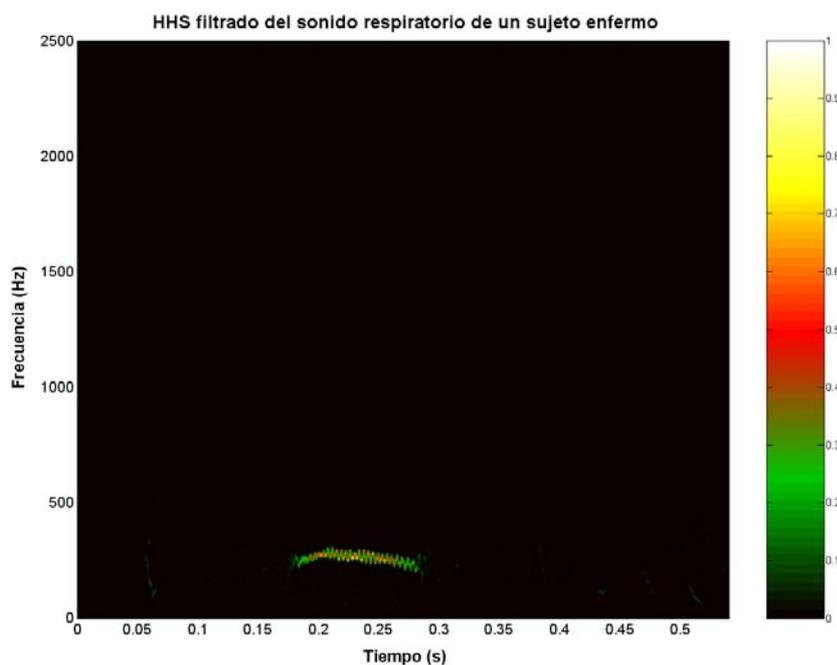


Fig. VI.130. Versión suavizada del Espectro de Hilbert-Huang de sonido respiratorio adquirido mostrado en la figura VI.129.

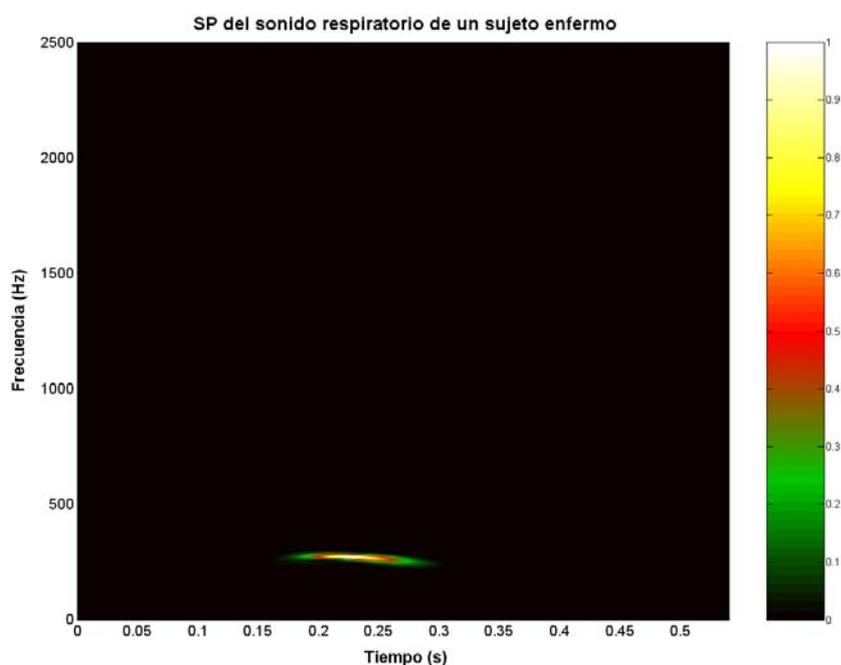


Fig. VI.131. Espectrograma del sonido respiratorio adquirido de la figura VI.126, obtenido con una ventana Hamming de 100 ms de longitud.

VI.7. Estertores crepitantes.

VI.7.1. Estertor crepitante fino simulado.

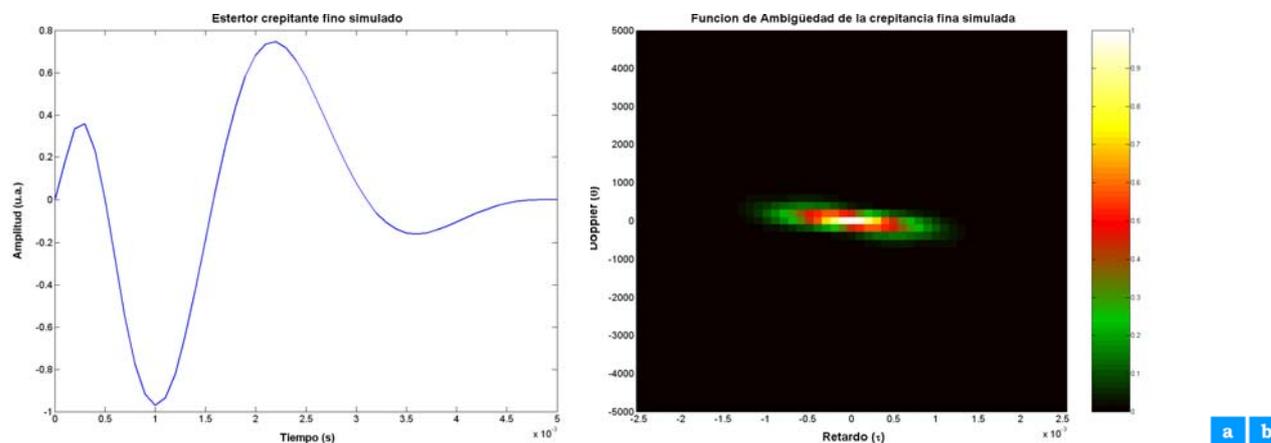


Fig. VI.132. Crepitancia fina simulada. a) Forma de onda. b) Función de ambigüedad.

En forma similar al caso de las señales simuladas no biológicas, de sonidos cardiacos y de sonidos sibilantes, se obtuvieron todas las representaciones tiempo-frecuencia utilizadas en la presente tesis para un sonido respiratorio adventicio discontinuo simulado, denominado estertor crepitante fino (figura VI.132.a). Debido a que no se cuenta con la expresión matemática de la frecuencia instantánea (representación teórica) correspondiente a este sonido, el espectro de Hilbert-Huang y la frecuencia instantánea estimada a partir de esta representación se tomaron como referencia para el resto de las representaciones. Por lo tanto se inicia con esta técnica para el análisis de las técnicas de representación restantes.

En la figura VI.133 se presenta la descomposición empírica de modos del estertor crepitante fino simulado, y la señal analítica de su único modo de oscilación intrínseco se presenta en la figura VI.134. En la figura VI.135 se presenta el espectro de Hilbert-Huang, mientras que en la figura VI.136 se presenta una versión suavizada de dicho espectro. El espectro de Hilbert-Huang mostrado en la figura VI.135 fue tomado como la representación tiempo-frecuencia de referencia.

Para el espectrograma, la mejor representación se obtuvo con una ventana Hamming de 2.5 ms de longitud. Los valores del error cuadrático medio obtenidos entre cada frecuencia instantánea estimada de los espectrogramas y la frecuencia instantánea estimada del espectro de Hilbert-Huang (referencia) del estertor crepitante se presentan en la tabla VI.XII.

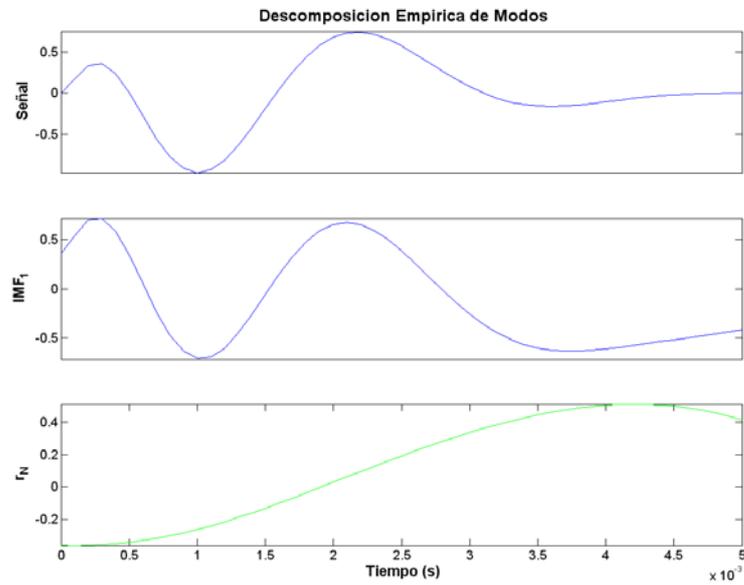


Fig. VI.133. Descomposición empírica de modos del estertor crepitante fino simulado.

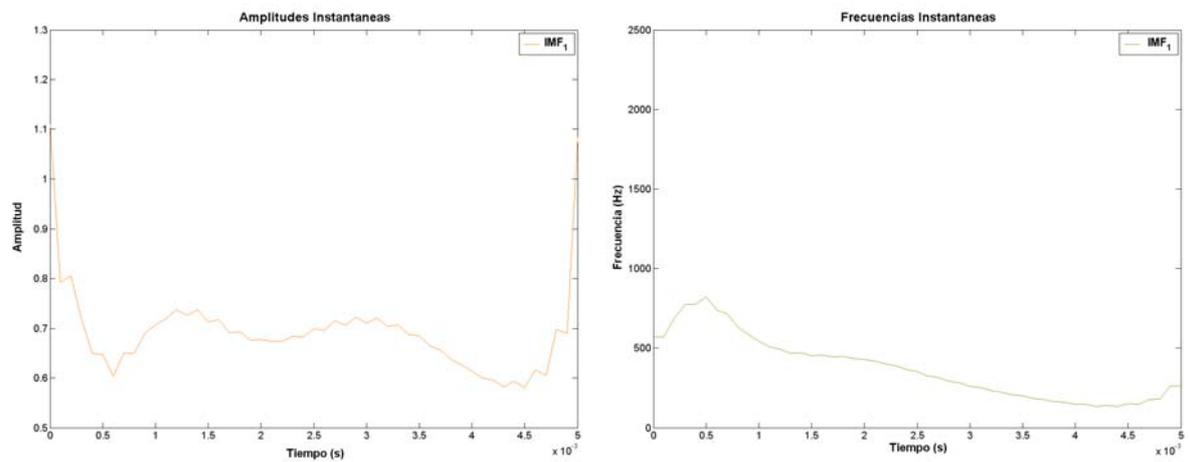


Fig. VI.124. Señal analítica del primer modo de oscilación intrínseco. a) Amplitud instantánea. b) Frecuencia instantánea.

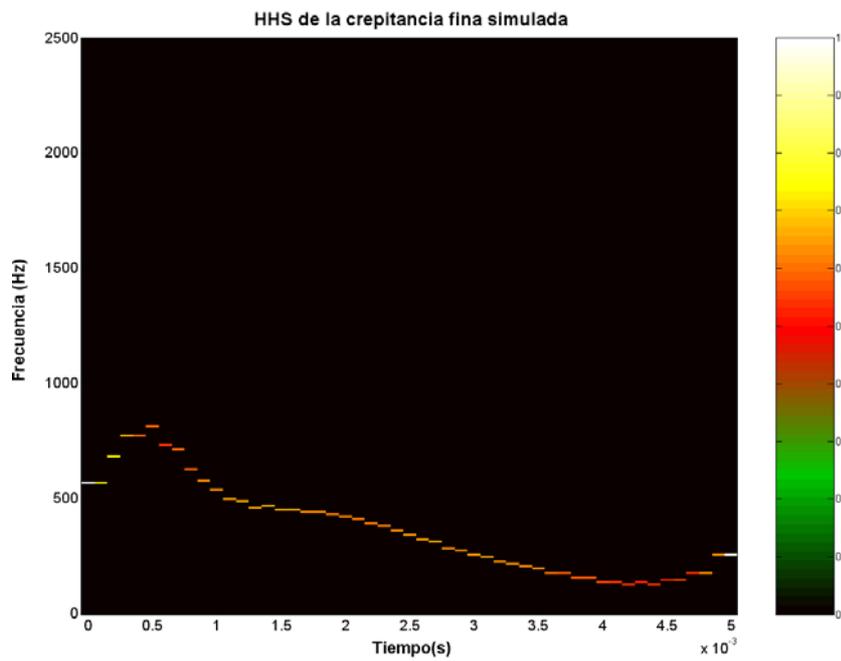


Fig. VI.135. Espectro de Hilbert-Huang de la crepitancia fina simulada.

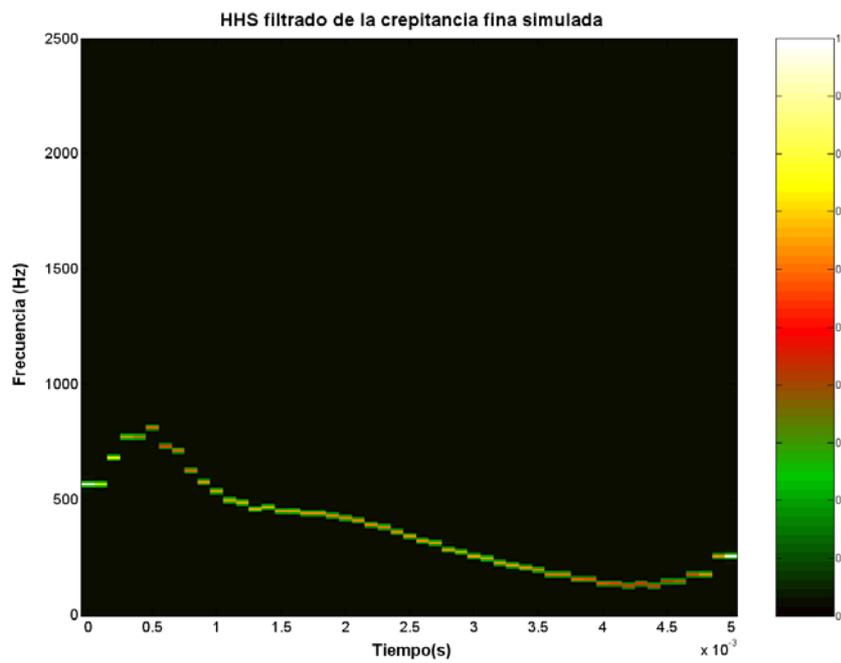


Fig. VI.136. Espectro de Hilbert-Huang suavizado de la crepitancia fina simulada.

TABLA VI.XII. VALORES DE NRMSE PARA LOS ESPECTROGRAMAS DE
LA CREPITANCIA FINA SIMULADA.

	Hamming	Rectangular	Blackman-Harris
0.5 ms	199.9500	166.1600	336.8000
0.7 ms	132.2400	137.7700	210.4300
1.1 ms	86.1560	117.5900	121.9100
1.3 ms	78.0660	112.0800	102.5200
1.5 ms	72.2650	97.8340	90.3480
1.7 ms	67.4190	87.7360	82.4740
2.1 ms	53.4550	72.0100	72.8550
2.3 ms	45.2880	67.1160	69.0120
2.5 ms	40.0880	63.5430	64.9310

El conjunto de espectrogramas de la crepitancia fina simulada, obtenidos al variar el tipo y longitud de la ventana de análisis se presentan en el Anexo I. En el mismo anexo también se incluyen las diversas representaciones tiempo-frecuencia obtenidas con cada técnica de análisis tiempo-frecuencia al variar sus parámetros correspondientes. La distribución de Wigner-Ville, y el resto de las representaciones tiempo-frecuencia seleccionadas para esta señal se presentan en la figura VI.137 y se detallan en el Anexo I. A continuación se presentan los parámetros correspondientes de cada una de estas representaciones tiempo-frecuencia seleccionadas.

Para la Distribución de Choi-Williams, la mejor representación se obtuvo con $\sigma = 95$. Para la técnica del espectrograma reasignado, la mejor representación se obtuvo con una ventana Blackman-Harris de 1.1 ms de longitud. Con respecto al cálculo de la representación tiempo-frecuencia vía el modelo TVAR y el algoritmo de Burg, la mejor representación se obtuvo con un modelo AR de orden 2 y una ventana Hamming de 2.5 ms de longitud. Para la técnica tiempo-frecuencia utilizando el algoritmo RLS, la mejor representación se obtuvo con un modelo AR de orden 2 y un factor de olvido igual a 0.78125. Con relación al algoritmo RLS-VFF, la mejor representación se obtuvo con un modelo AR orden 2 y una varianza del ruido esperado igual a 1×10^{-6} . Para la representación con base en el filtro de Kalman, la mejor se obtuvo con los parámetros siguientes: modelo AR de orden 2, varianza del ruido de medición igual a 0.1 y varianza del ruido del proceso igual a 0.017783. Con respecto al

escalograma, la mejor representación se obtuvo con una wavelet Morlet de 0.6 ms de longitud en la escala más gruesa.

▪ **Representaciones tiempo-frecuencia seleccionadas.**

En la figura VI.137 se presentan las representaciones tiempo-frecuencia seleccionadas para la crepitancia fina de acuerdo al criterio del error cuadrático medio de las frecuencias instantáneas estimadas mediante el centroide.

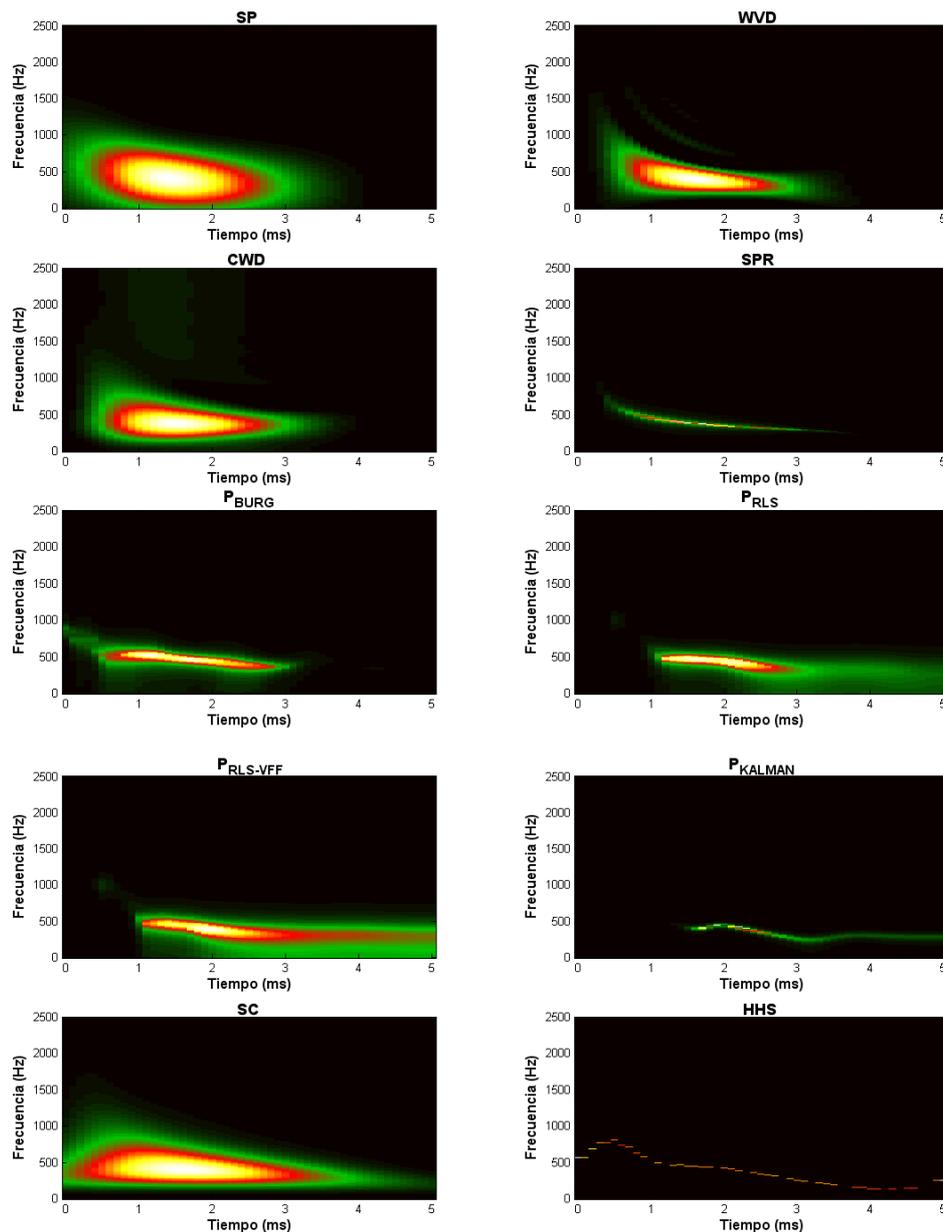


Fig. VI.137. Representaciones tiempo-frecuencia seleccionadas para cada técnica de análisis tiempo-frecuencia, así como el Espectro de Hilbert-Huang de la crepitancia fina simulada.

VI.7.2. Estertor crepitante grueso simulado.

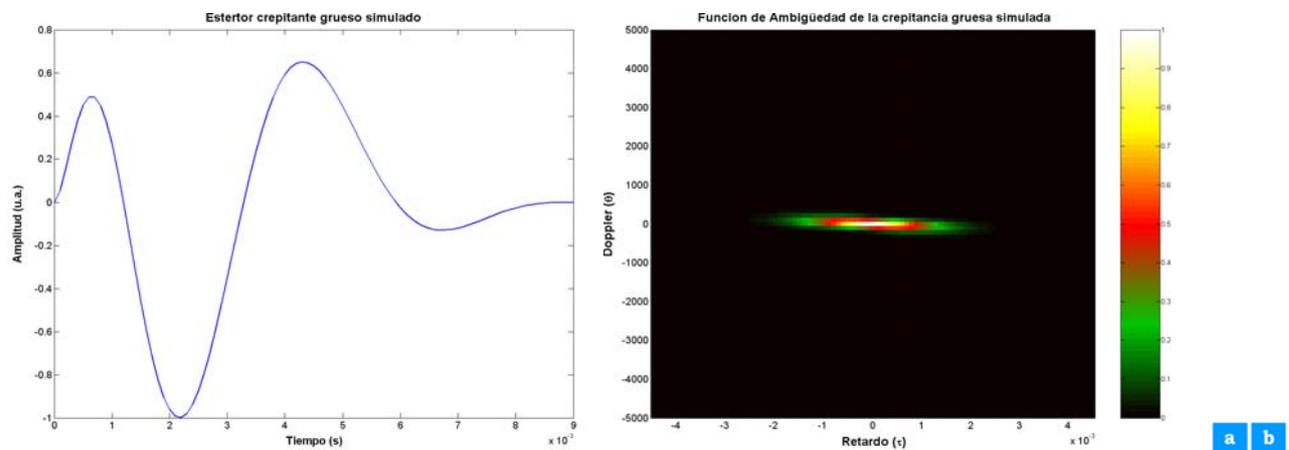


Fig. VI.138. Crepitancia gruesa simulada. a) Forma de onda. b) Función de ambigüedad.

En forma similar al caso de la crepitancia fina simulada, se obtuvieron todas las representaciones tiempo-frecuencia utilizadas en la presente tesis para un segundo sonido respiratorio adventicio discontinuo simulado, denominado estertor crepitante grueso (figura VI.138.a). El espectro de Hilbert-Huang se consideró nuevamente como la representación de referencia para esta señal. Por lo tanto, otra vez se inicia con esta técnica para el análisis de las técnicas de representación restantes. La descomposición empírica de modos del crepitante grueso se presenta en la figura VI.139 mientras que la señal analítica de su modo de oscilación intrínseco se muestra en la figura VI.140. El espectro de Hilbert-Huang correspondiente que resulta de este modo de oscilación se presenta en la figura VI.141 y una versión suavizada del espectro se presenta en la figura VI.142.

Con respecto del espectrograma, la mejor representación en términos del NRMSE entre las frecuencias instantáneas de cada espectrograma generado, al variar el tipo de ventana y su longitud, y de la representación tiempo-frecuencia ideal (espectrograma de Hilbert-Huang) fue la generada con una ventana Hamming de 4.5 ms de longitud. En la tabla VI.XIII se incluyen todos los valores de NRMSE para las diferentes ventanas y longitudes utilizadas en este caso.

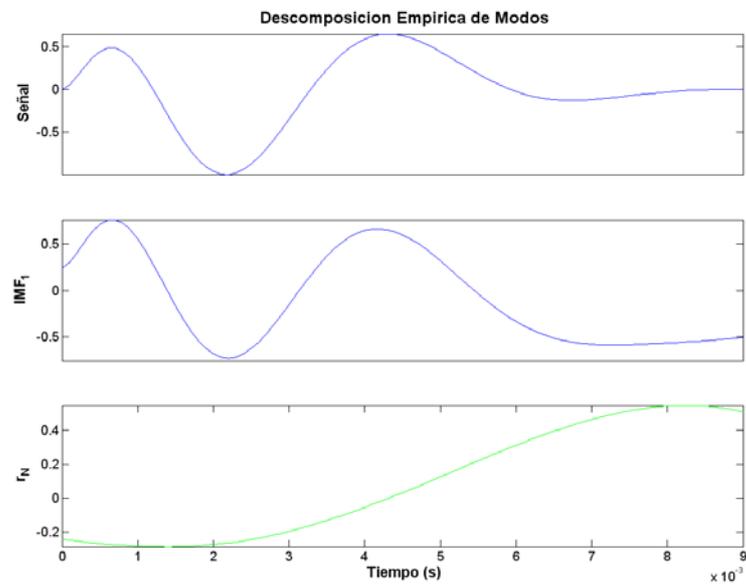


Fig. VI.139. Descomposición empírica de modos del estertor crepitante grueso simulado.

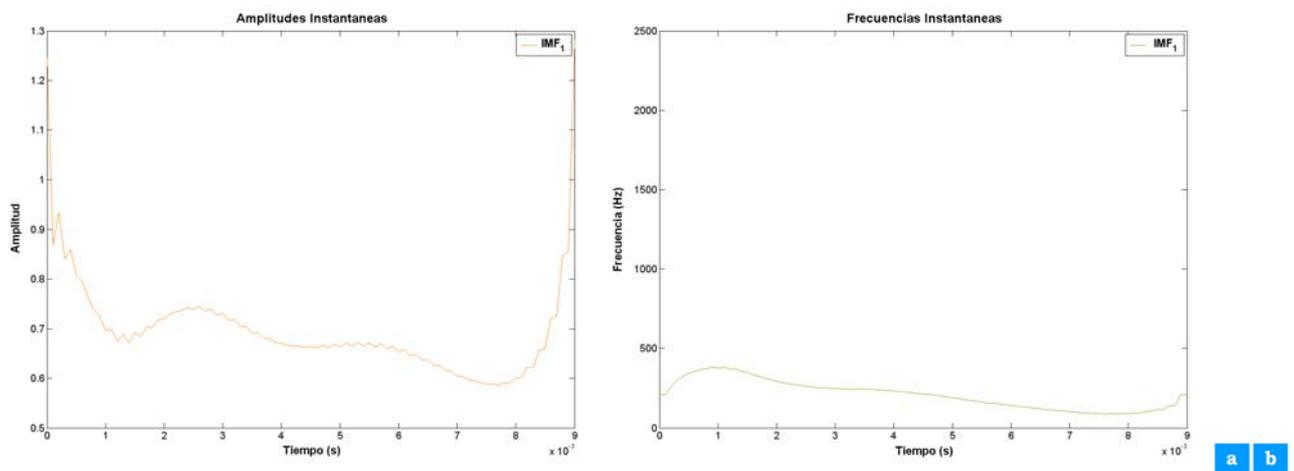


Fig. VI.140. Señal analítica del primer modo de oscilación intrínseco para el crepitante grueso.
a) Amplitud instantánea. b) Frecuencia instantánea.

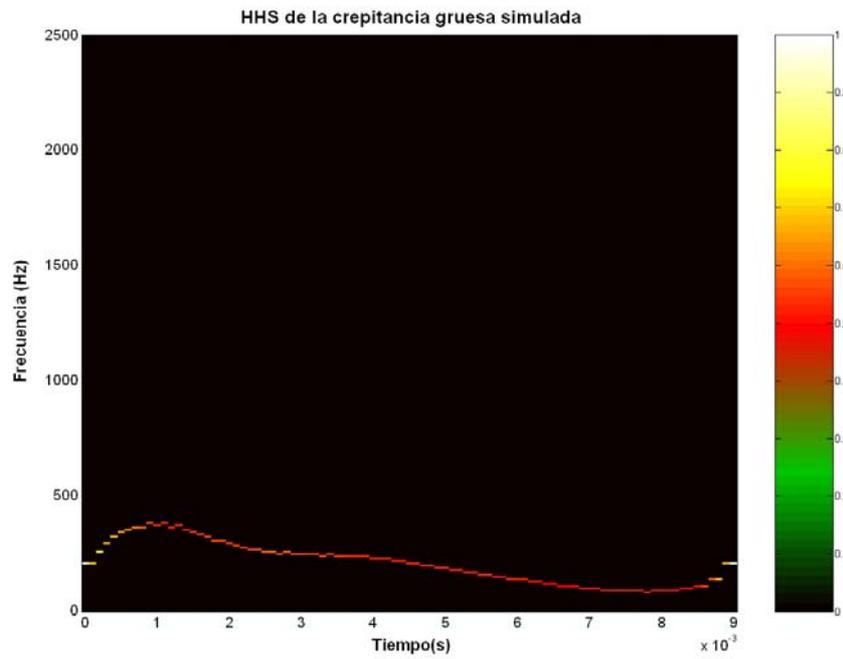


Fig. VI.141. Espectro de Hilbert-Huang de la crepitancia gruesa simulada.

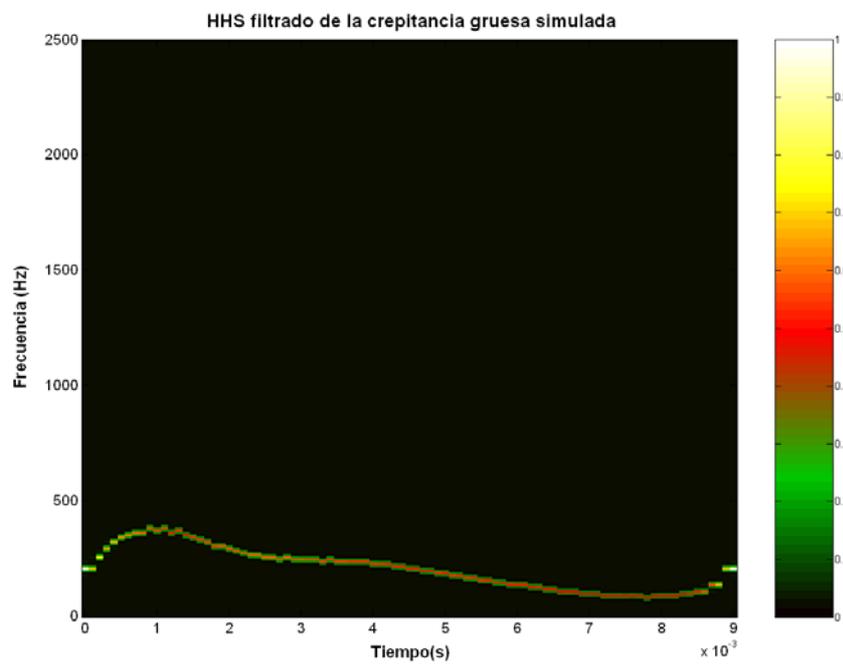


Fig. VI.142. Espectro de Hilbert-Huang suavizado de la crepitancia gruesa simulada.

TABLA VI.XIII. VALORES DE NRMSE PARA LOS ESPECTROGRAMAS DE LA CREPITANCIA GRUESA SIMULADA.

	Hamming	Rectangular	Blackman-Harris
0.9 ms	201.4400	196.3200	313.8500
1.3 ms	138.6800	154.7000	204.3000
1.9 ms	103.2600	121.5200	138.8000
2.3 ms	90.1140	106.9300	118.2200
2.7 ms	78.9560	95.8340	104.9400
3.1 ms	68.5900	86.9310	95.0840
3.7 ms	54.7040	77.2980	82.9210
4.1 ms	47.2220	72.2130	75.5110
4.5 ms	41.4110	68.0640	68.4070

De forma similar a la crepitancia fina simulada, en el Anexo I se presentan las diversas representaciones tiempo-frecuencia de la crepitancia gruesa simulada, obtenidas para técnica de análisis al variar sus parámetros correspondientes. Además, en el mismo anexo se incluye la representación tiempo-frecuencia seleccionada para cada técnica mediante el criterio del error cuadrático medio entre las frecuencias instantáneas estimadas a partir del centroide de las distribuciones.

En el caso de la Distribución de Choi-Williams, la mejor representación se obtuvo para el parámetro $\sigma = 60$. Con respecto a la técnica del espectrograma reasignado, la mejor representación se obtuvo para una ventana Blackman-Harris de 1.9 ms de longitud.

En el caso de la representación vía el algoritmo de Burg, la mejor representación se obtuvo para un modelo AR de orden 2 y una ventana de Hamming de 4.5 ms de longitud. Con relación a la representación vía el algoritmo RLS, la mejor representación para este caso se obtuvo con un modelo AR de orden 2 y un factor de olvido igual a 0.78125. Para la técnica con base en el algoritmo RLS-VFF, la mejor representación se obtuvo con un modelo AR orden 2 y una varianza del ruido esperado igual a 5.6234×10^{-9} . En el caso del filtro de Kalman, la mejor representación se obtuvo con un modelo AR de orden 2, una varianza del ruido de medición igual a 0.1 y una varianza del ruido del proceso igual a 3.1623×10^{-3} . Para el escalograma la mejor representación se obtuvo con una wavelet Morlet de 0.6 ms de longitud en la escala más gruesa.

- **Representaciones tiempo-frecuencia seleccionadas.**

Las representaciones tiempo-frecuencia seleccionadas es estertor crepitante grueso de acuerdo al criterio del error cuadrático medio de las frecuencias instantáneas estimadas, se presentan en la figura VI.143.

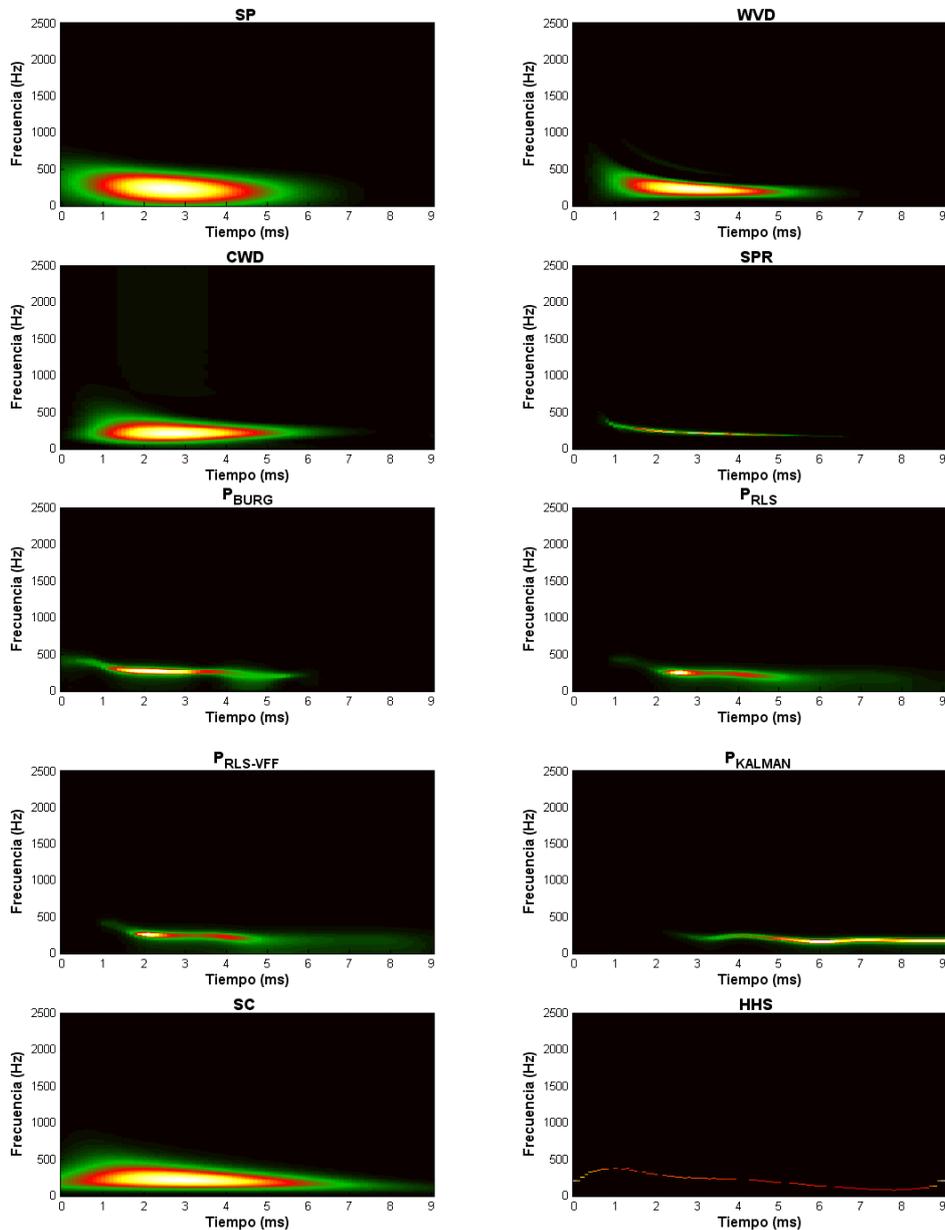


Fig. VI.143. Representaciones tiempo-frecuencia seleccionadas para cada técnica de análisis tiempo-frecuencia, así como el Espectro de Hilbert-Huang de la crepitancia gruesa simulada.

VI.7.3. Estertores crepitantes reales adquiridos.

▪ Crepitancia real 1.

En la figura VI.144 se presenta una crepitancia real adquirida de un sujeto con NID. Su descomposición empírica de modos se presenta en la figura VI.145, la contribución de energía de cada modo de oscilación resultante se presenta en la figura VI.146, mientras que en la figura VI.147 se presenta la señal analítica de sus dos primeros modos de oscilación intrínsecos. En la figura VI.148 se presenta el Espectro de Hilbert-Huang, y en la figura VI.149 se muestra una versión suavizada de dicho espectro. En la figura VI.150 se presenta el espectrograma obtenido para esta crepitancia adquirida con una ventana acorde a los resultados de la simulación.

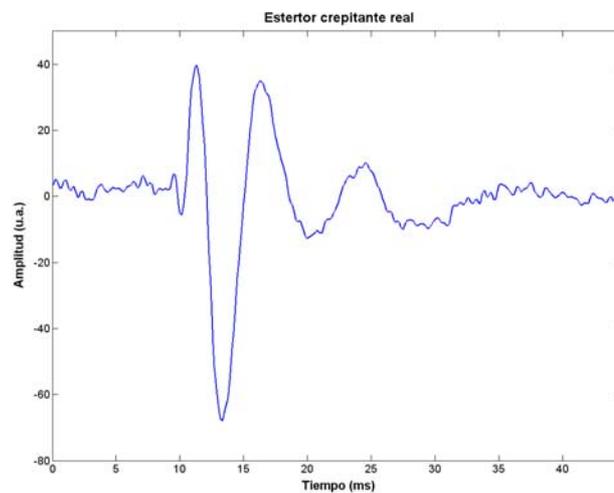


Fig. VI.144. Forma de onda de una crepitancia adquirida.

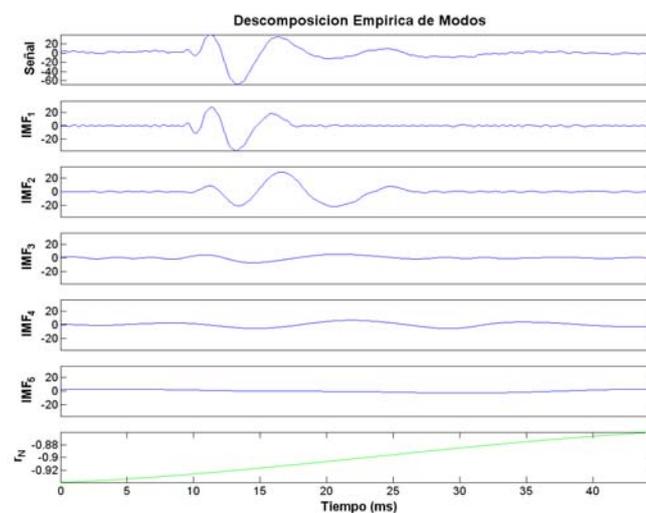


Fig. VI.145. Descomposición empírica de modos del estertor crepitante de la figura VI.144.

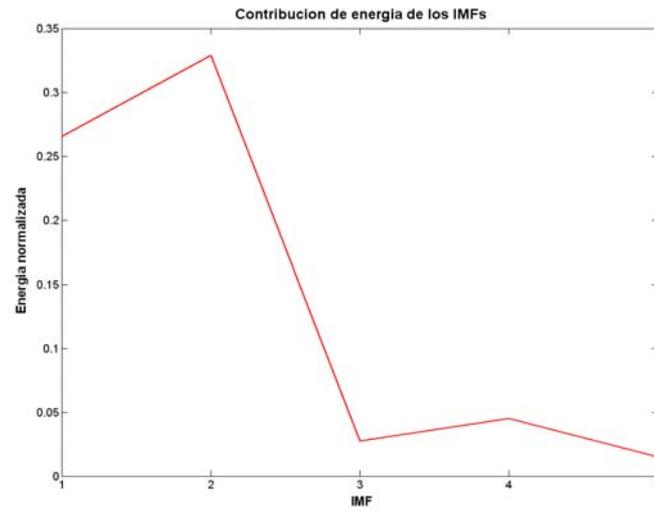


Fig. VI.146. Contribución de energía de cada modo de oscilación intrínseco de la figura VI.145.

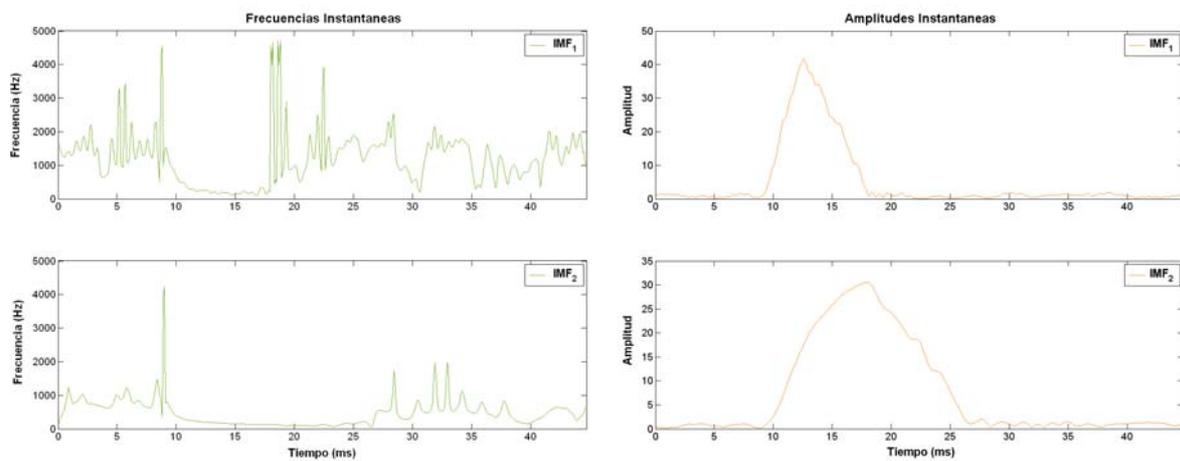


Fig. VI.147. Señales analíticas de los dos primeros modos de oscilación intrínsecos. Amplitud instantánea (columna izquierda) y Frecuencia instantánea (columna derecha).

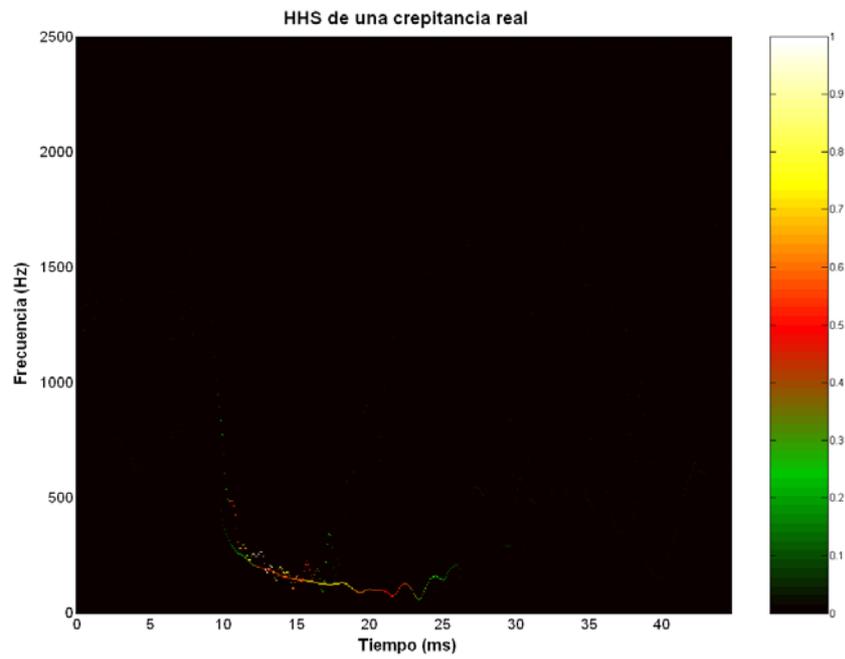


Fig. VI.148. Espectro de Hilbert-Huang de la crepitancia adquirida mostrada en la figura VI144.

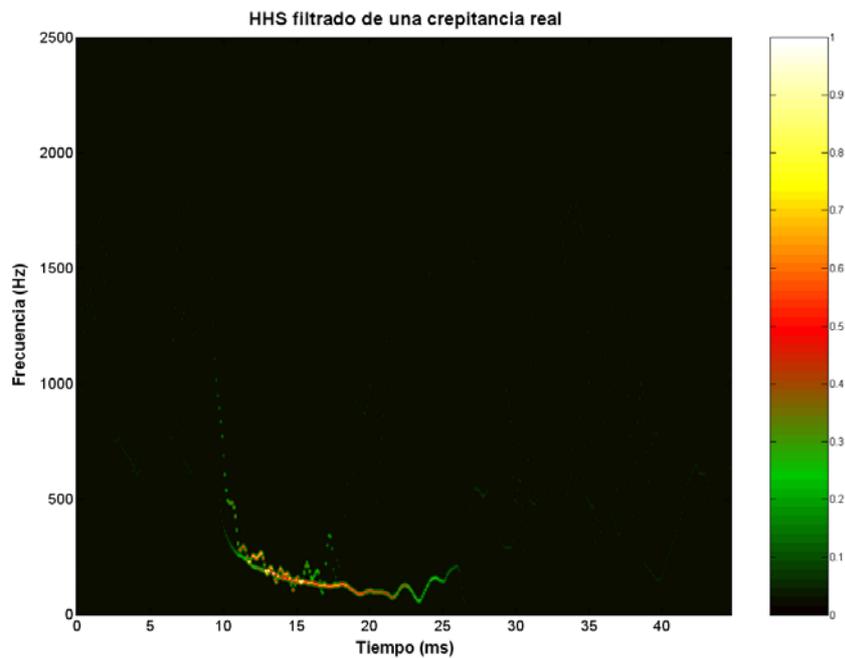


Fig. VI.149. Espectro de Hilbert-Huang suavizado de la crepitancia mostrada en la figura VI.144.



- **Crepitancia real 2.**

Una segunda crepitancia real adquirida se presenta en la figura VI.151. En la figura VI.152 se presenta la descomposición empírica de modos de esta crepitancia y la señal analítica correspondiente de sus tres primeros modos de oscilación intrínsecos se presenta en la figura VI.153. El Espectro de Hilbert-Huang de esta crepitancia se presenta en la figura VI.154 y su versión suavizada en la figura VI.155. El espectrograma de esta segunda crepitancia real se presenta en la figura VI.156.

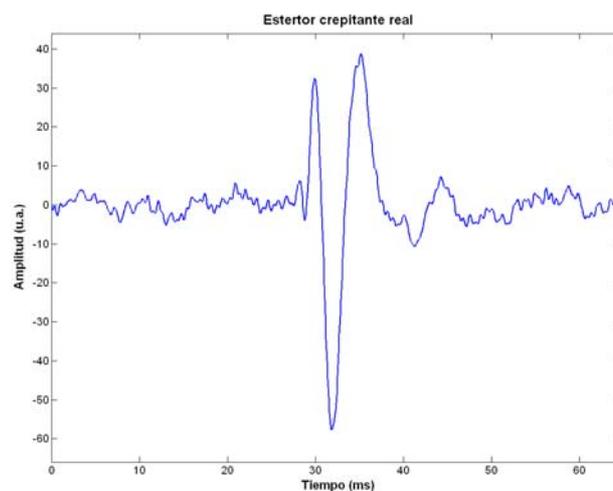


Fig. VI.151. Forma de onda de una segunda crepitancia adquirida.

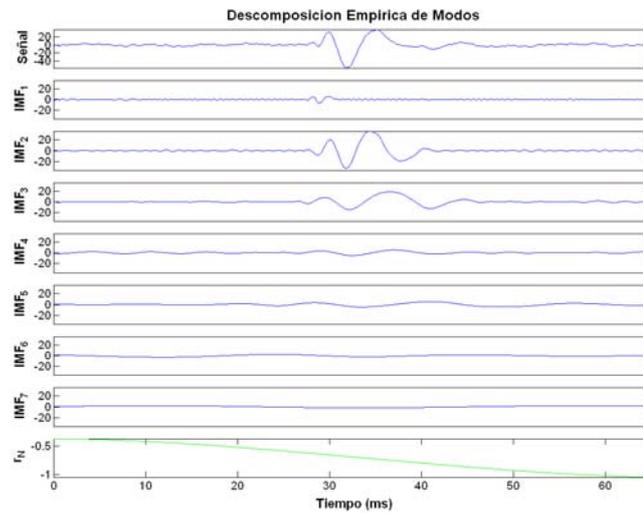


Fig. VI.152. Descomposición empírica de modos del estertor crepitante de la figura VI.151.

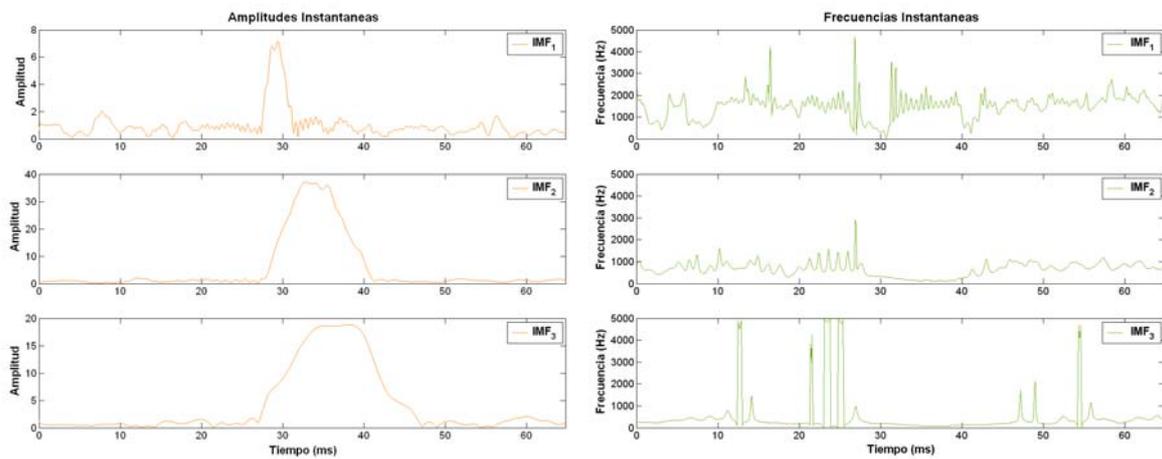


Fig. VI.153. Señales analíticas de los tres primeros modos de oscilación intrínsecos. Amplitud instantánea (columna izquierda) y Frecuencia instantánea (columna derecha).

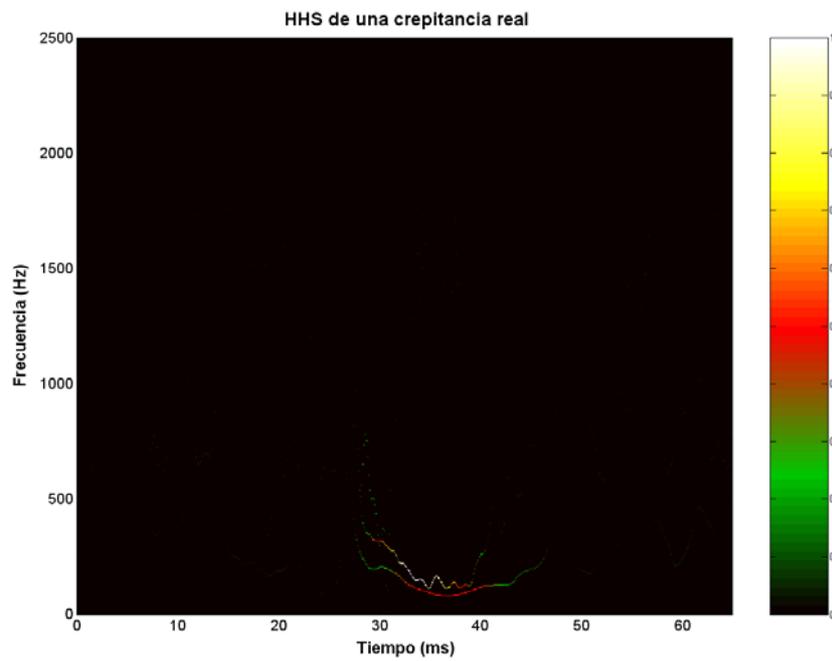


Fig. VI.154. Espectro de Hilbert-Huang de la crepitancia adquirida mostrada en la figura VI.151.

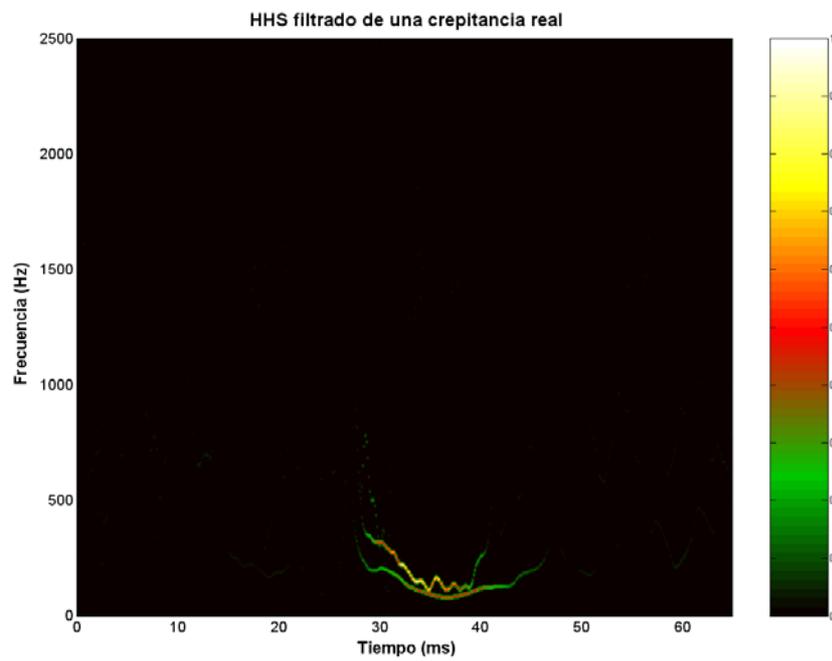


Fig. VI.155. Espectro de Hilbert-Huang suavizado de la crepitancia mostrada en la figura VI.151.

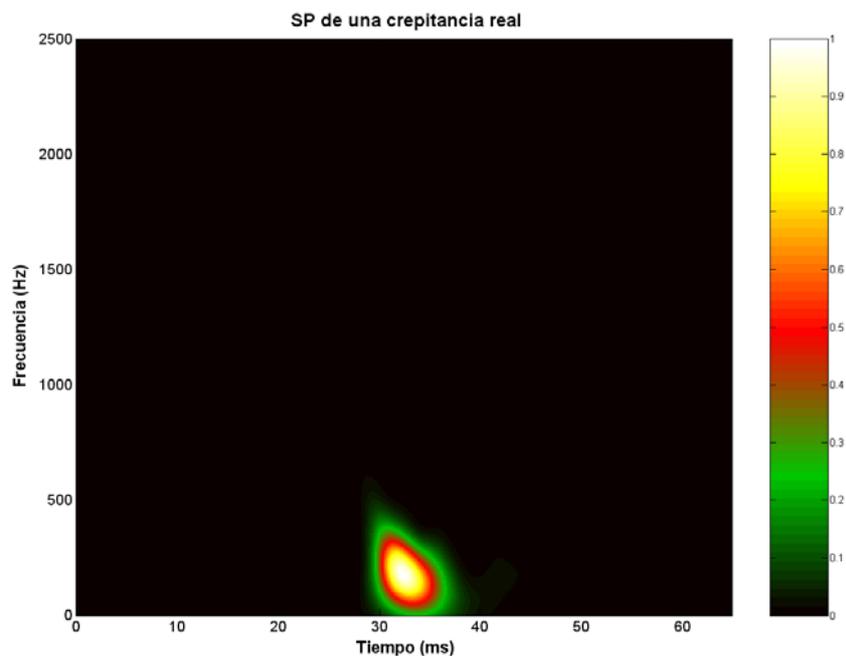


Fig. VI.156. Espectrograma de la crepitancia adquirida mostrada en la VI.151, obtenido con una ventana Hamming de 6 ms de longitud.

▪ Crepitancia real 3.

En la figura VI.157 se presenta una tercera crepitancia real adquirida de un sujeto con NID. En las figuras VI.158 se presenta la descomposición empírica de modos de esta crepitancia y en la figura VI.159 se presenta la señal analítica de los dos modos de oscilación intrínsecos seleccionados para esta señal.

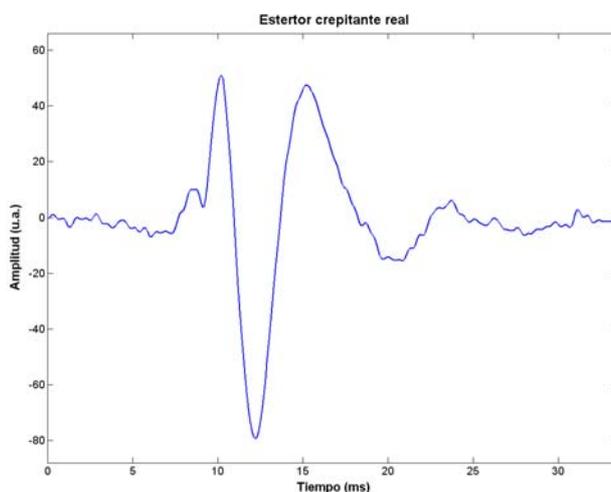


Fig. VI.157. Forma de onda de una crepitancia adquirida.

En la figura VI.160 se presenta el Espectro de Hilbert-Huang de esta crepitancia, mientras que el espectro suavizado se presenta en la figura VI.161. En la figura VI.162 se presenta el espectrograma de esta crepitancia real. El espectrograma de cada crepitancia real se obtuvo con una ventana Hamming de una longitud igual a la mitad de la duración de la crepitancia de acuerdo a los resultados de la simulación.

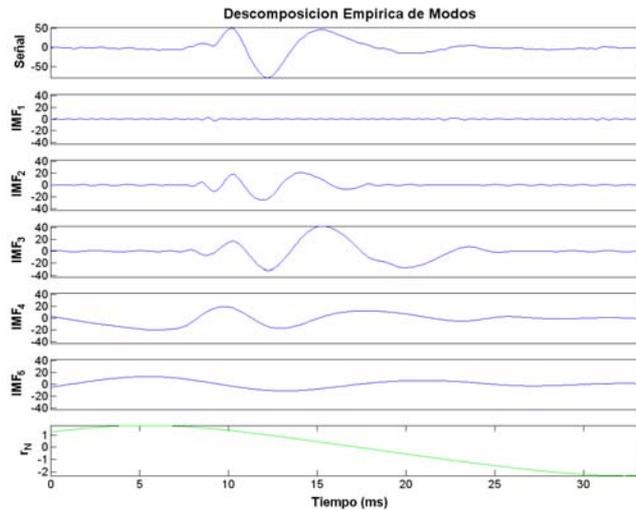


Fig. VI.158. Descomposición empírica de modos del estertor crepitante de la figura VI.157.

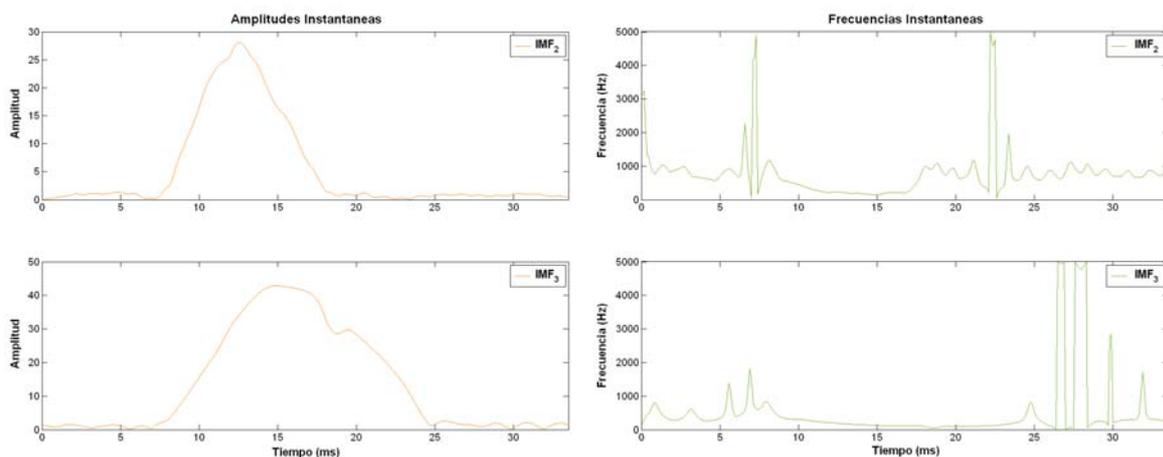


Fig. VI.159. Señales analíticas del segundo y tercer modos de oscilación intrínsecos. Amplitud instantánea (columna izquierda) y Frecuencia instantánea (columna derecha).

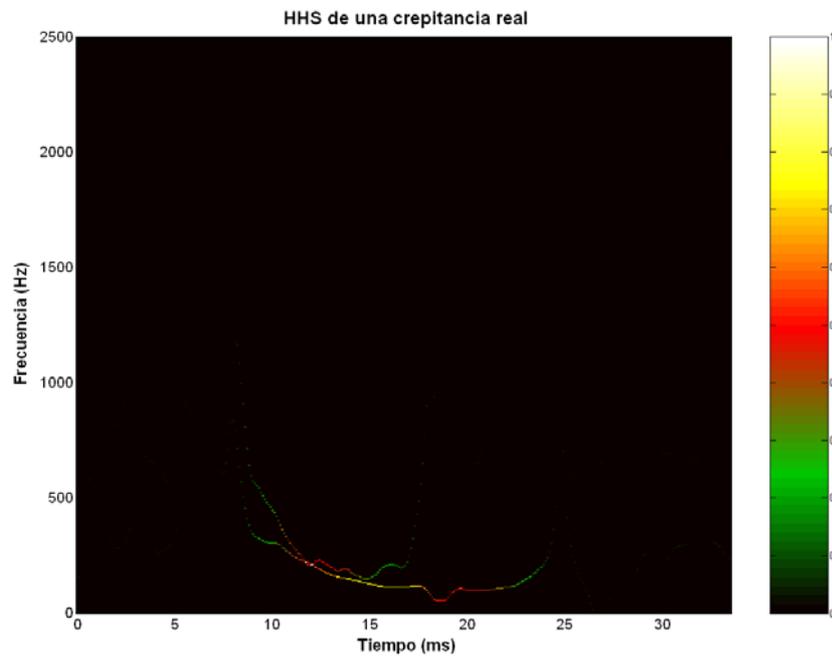


Fig. VI.160. Espectro de Hilbert-Huang de la crepitancia adquirida mostrada en la figura VI.157.

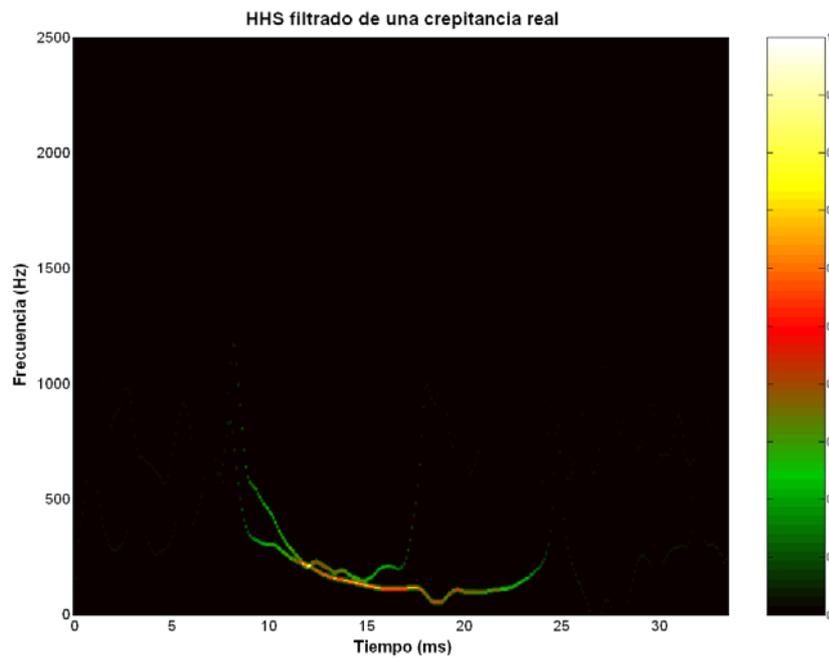


Fig. VI.161. Espectro de Hilbert-Huang suavizado de la crepitancia mostrada en la figura VI.157.



- **Crepitancia real 4.**

En la figura VI.163 se presenta un sonido respiratorio real adquirido de un paciente con NID a lo largo de un ciclo inspiratorio. La señal adquirida contiene dos sonidos adventicios discontinuos. La descomposición empírica de modos de este sonido respiratorio se muestra en la figura VI.164. En la figura VI.165 se presenta la contribución de cada modo intrínseco de oscilación de la descomposición a la energía de la señal, mientras que en la figura VI.166 se presenta la señal analítica de sus dos primeros modos de oscilación intrínsecos.

En la figura VI.167 se presenta el Espectro de Hilbert-Huang de la señal y en la figura VI.168 se presenta una versión suavizada de dicho espectro.

En la figura VI.169 se presenta el espectrograma obtenido para este sonido respiratorio con una ventana Hamming de 100 ms de longitud, de acuerdo a lo reportado en la bibliografía. Por otra parte, en la figura VI.170 se presenta el espectrograma obtenido para la misma señal con una ventana Hamming de 5 ms de longitud, de acuerdo a los resultados de la simulación.

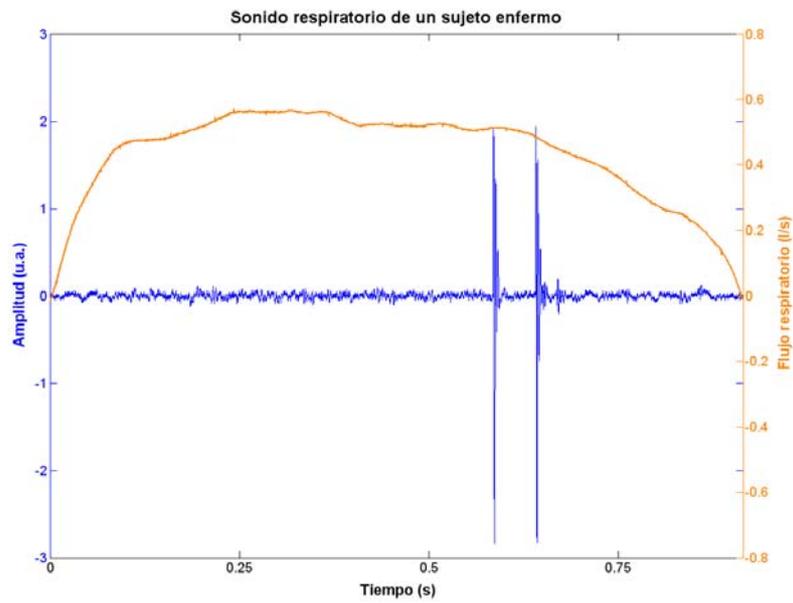


Fig. VI.163. Sonido respiratorio adquirido de un paciente con NID durante la inspiración.

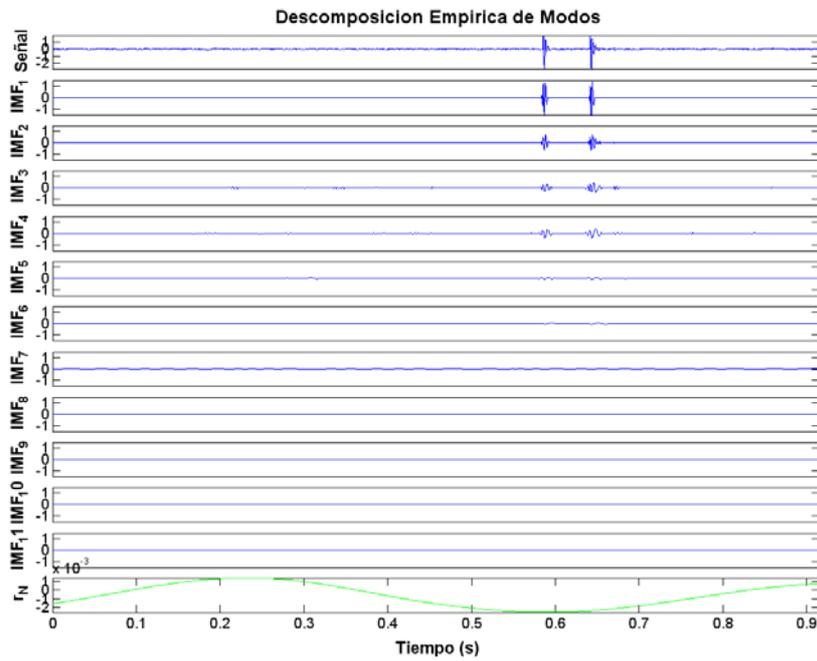


Fig. VI.164. Descomposición empírica de modos del sonido respiratorio de la figura VI.163.

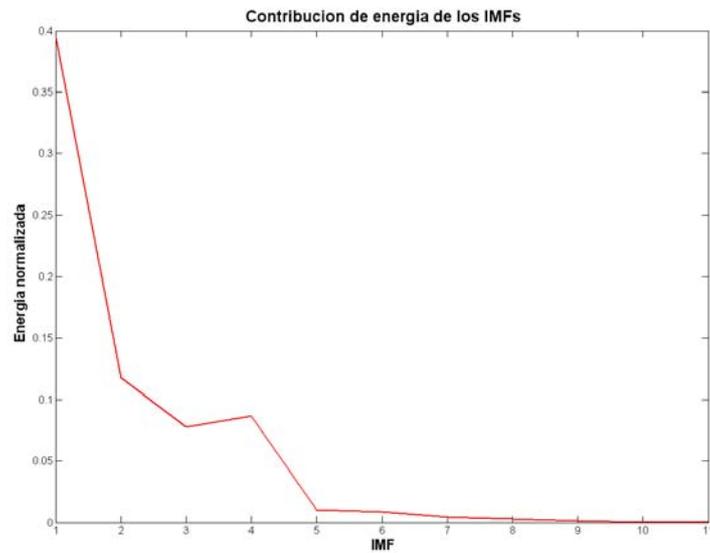


Fig. VI.165. Contribución de energía de cada modo de oscilación intrínseco del sonido respiratorio con crepitaciones de la figura VI.163.

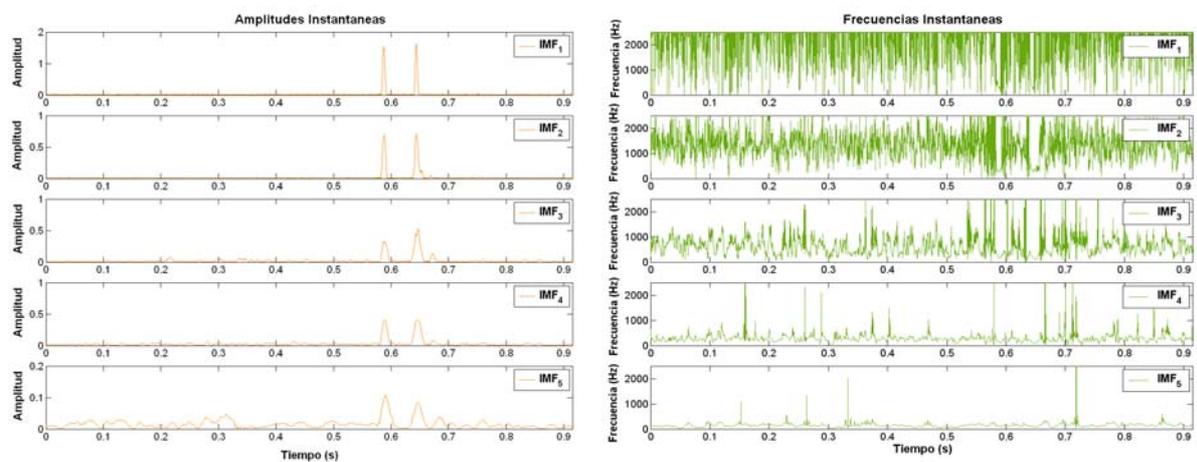


Fig. VI.166. Señal analítica de los cuatro primeros modos de oscilación intrínsecos. Amplitud instantánea (columna izquierda), y Frecuencia instantánea (columna derecha).

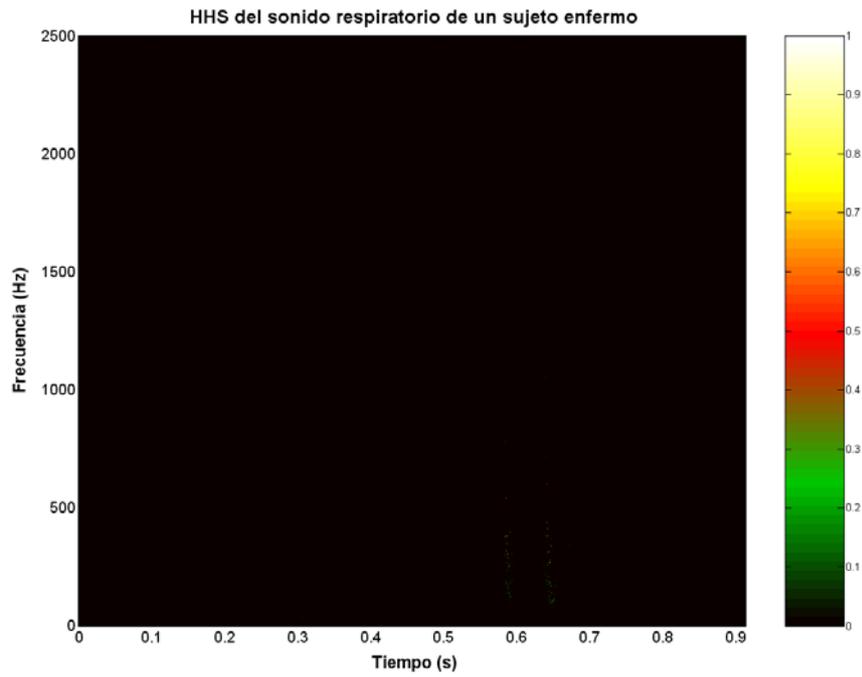


Fig. VI.167. Espectro de Hilbert-Huang del sonido respiratorio mostrado en la figura VI.163.

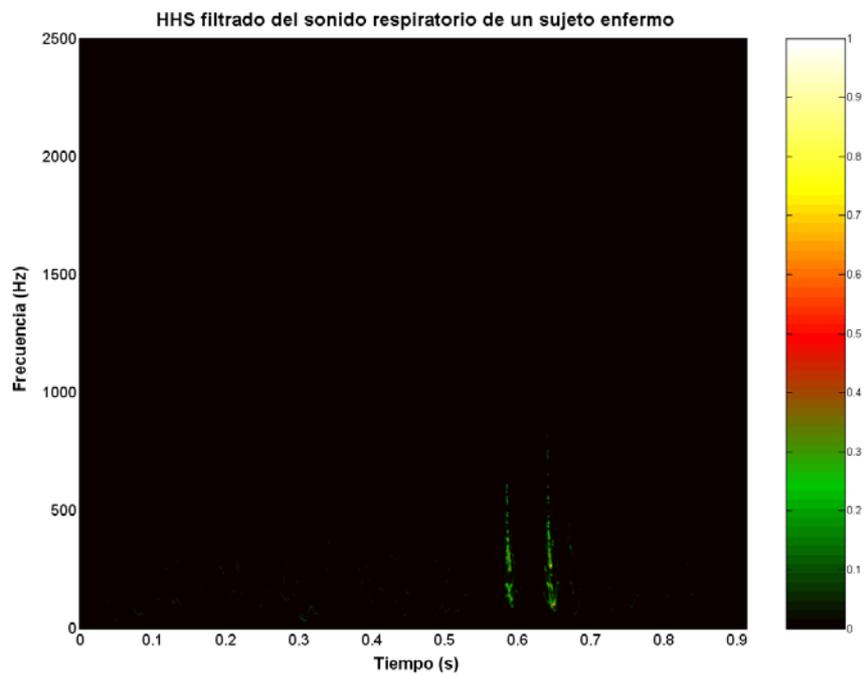


Fig. VI.168. Versión suavizada del Espectro de Hilbert-Huang de sonido respiratorio adquirido mostrado en la figura VI.163.

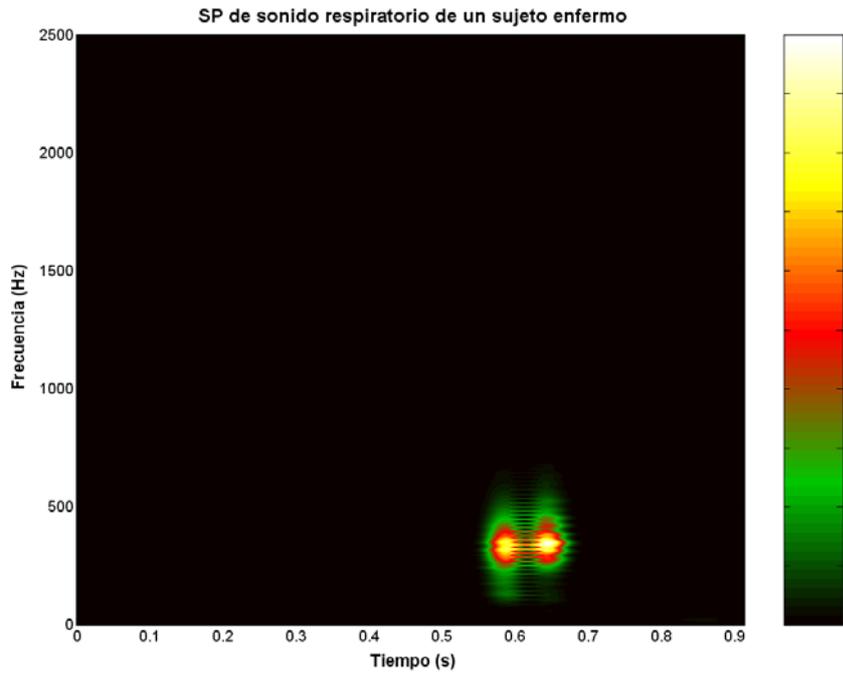


Fig. VI.169. Espectrograma del sonido respiratorio adquirido de la figura VI.163, obtenido con una ventana Hamming de 100 ms de longitud.

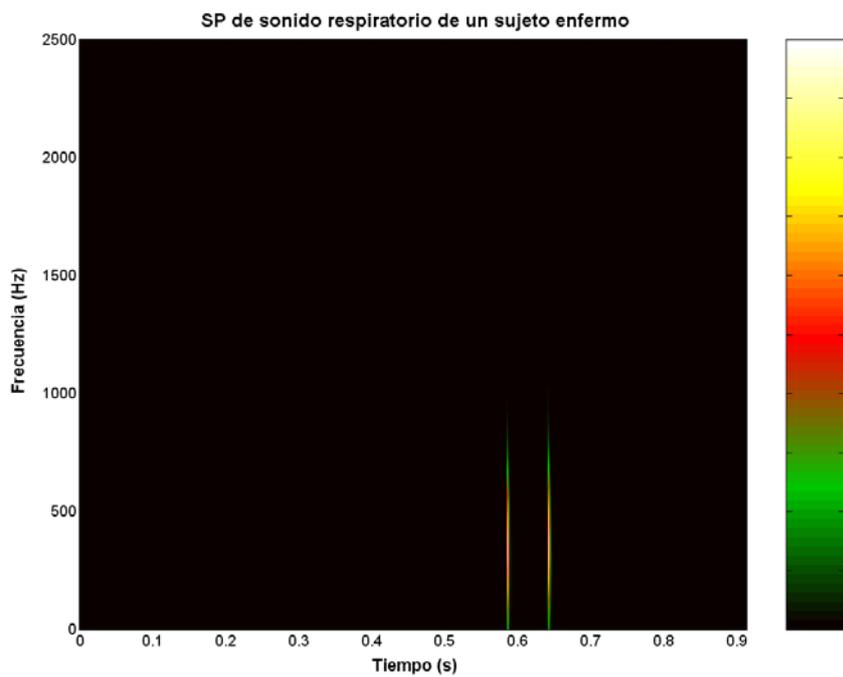


Fig. VI.170. Espectrograma del sonido respiratorio adquirido de la figura VI.163, obtenido con una ventana Hamming de 5 ms de longitud.

VI.8. Sonidos respiratorios normales adquiridos.

VI.8.1 Sonido pulmonar.

Una señal típica de sonido pulmonar normal adquirida sobre la superficie torácica de un sujeto sano, durante un ciclo respiratorio conteniendo las fases inspiratoria y espiratoria, se presenta en la figura VI.171. La descomposición empírica de modos de esta señal adquirida se presenta en la figura VI.172, la descomposición generó once modos de oscilación intrínsecos. En la figura VI.173 se presenta la contribución de energía de cada modo de oscilación intrínseco del sonido pulmonar normal. Las señales analíticas de los modos intrínsecos de oscilación seleccionados de acuerdo a su energía y su correlación con la amplitud del flujo respiratorio se presentan en la figura VI.174. En la figura VI.175 se presenta el Espectro de Hilbert-Huang del sonido pulmonar normal que se construyó a partir de las señales analíticas de la figura VI.174. Además, por cuestiones de visualización se presenta una versión suavizada de dicho espectro en la figura VI.176. El espectrograma correspondiente se presenta en la figura VI.177. El espectrograma de esta señal se obtuvo con una ventana Hamming de 100 ms de longitud, de acuerdo a lo reportado en la bibliografía.

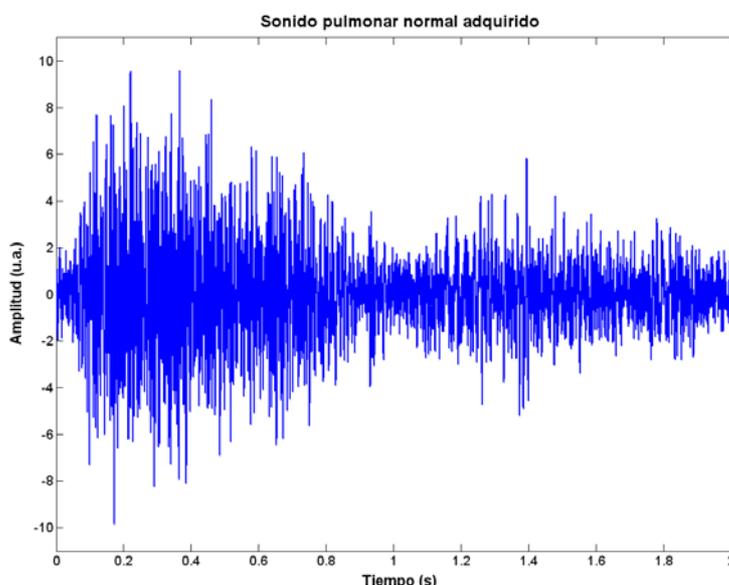


Fig. VI.171. Forma de onda del sonido pulmonar normal adquirido a lo largo de un ciclo respiratorio.

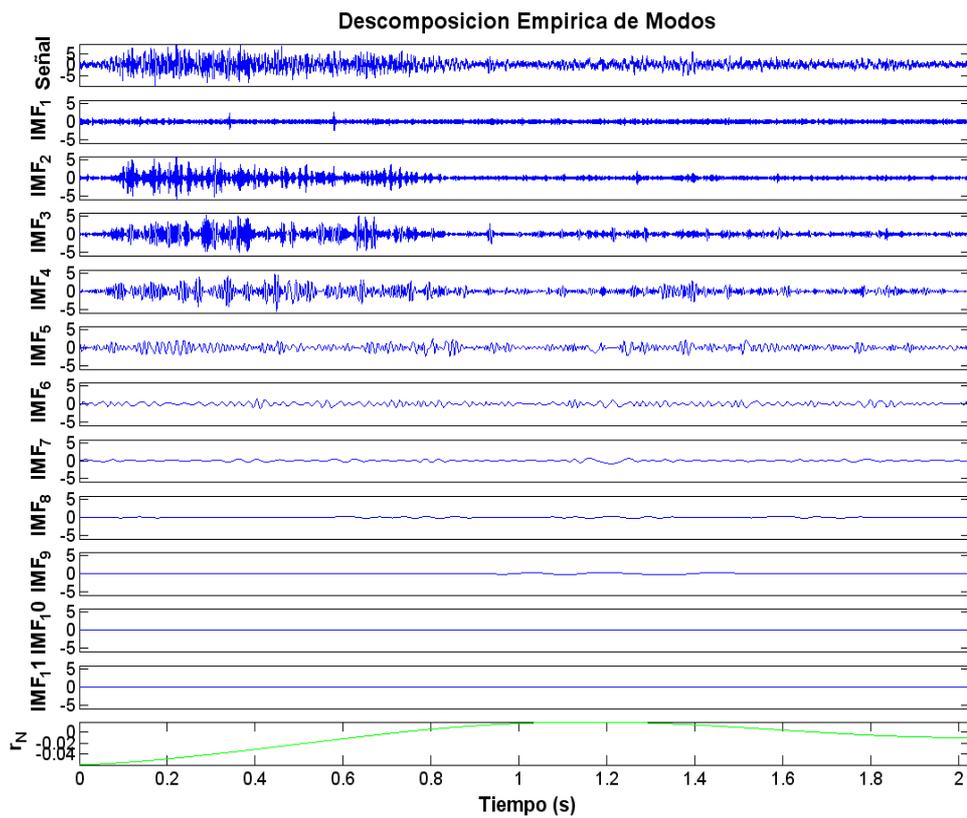


Fig. VI.172. Descomposición empírica de modos del sonido pulmonar normal adquirido.

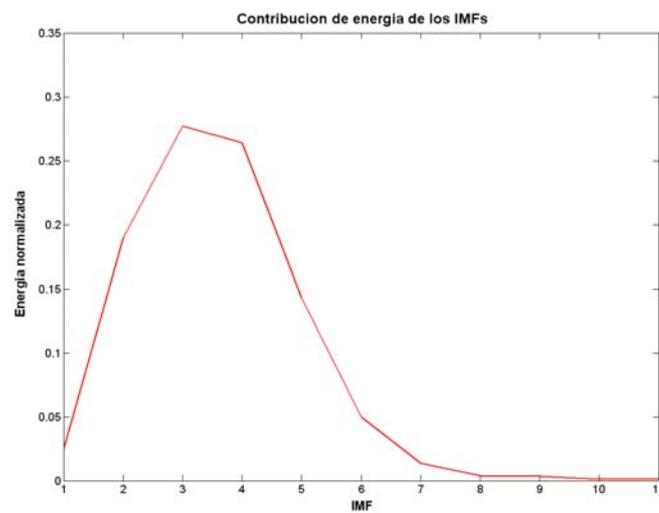


Fig. VI.173. Contribución de energía de cada modo de oscilación intrínseco del sonido pulmonar normal mostrado en la figura VI.171.

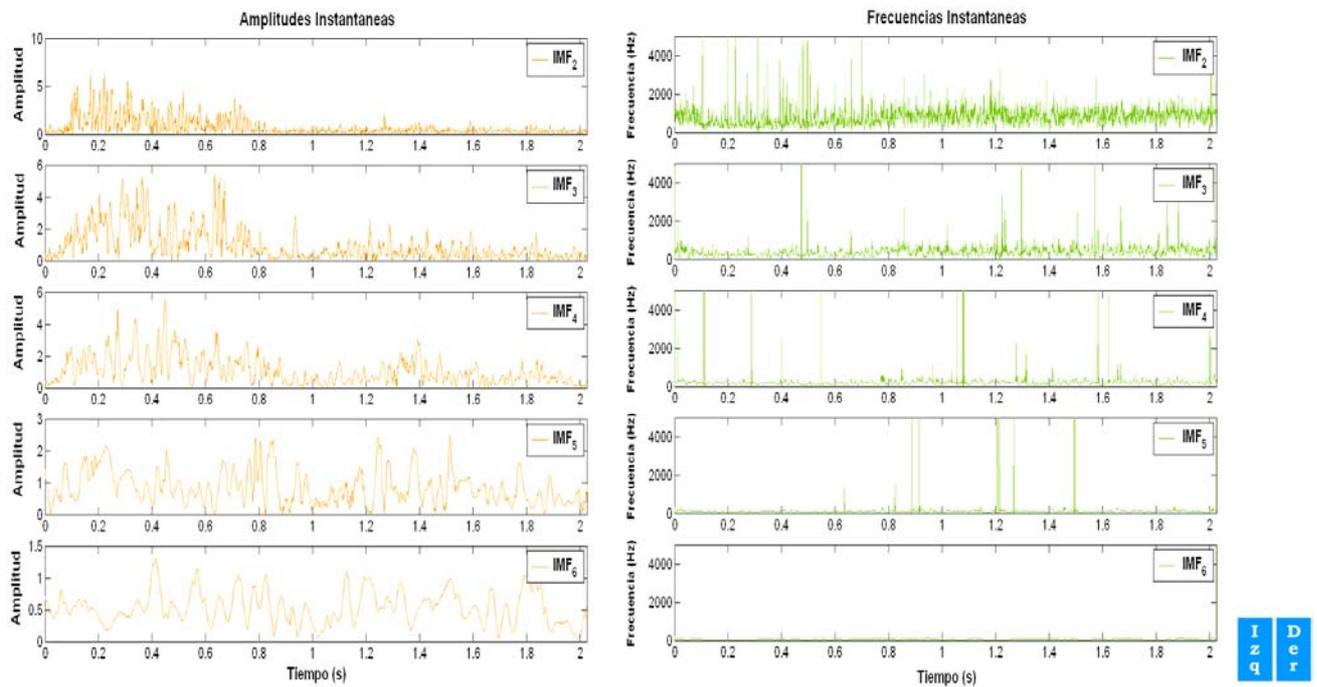


Fig. VI.174. Señales analíticas de los algunos modos de oscilación intrínsecos del sonido pulmonar normal adquirido. Amplitud instantánea (columna izquierda) y Frecuencia instantánea (columna derecha).

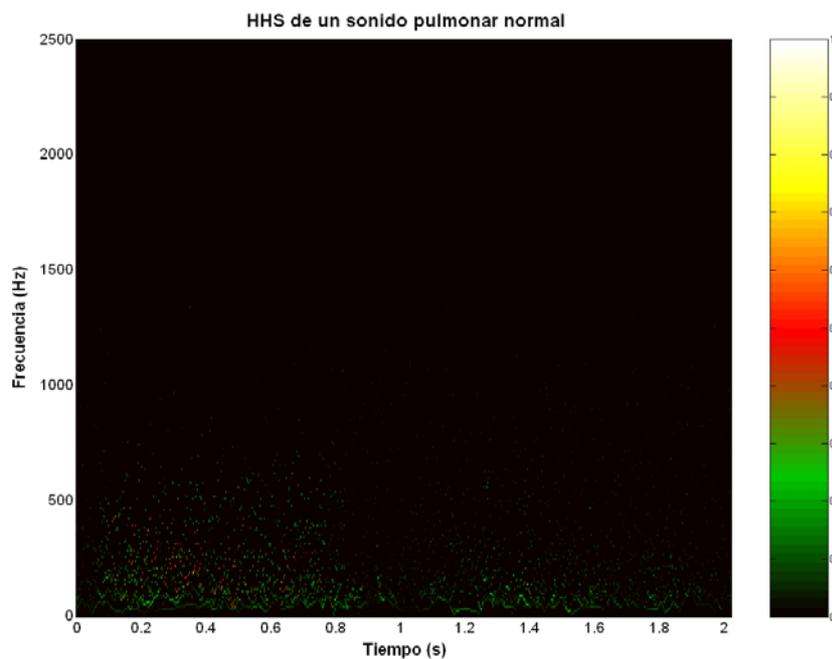


Fig. VI.175. Espectro de Hilbert-Huang del sonido pulmonar normal adquirido.

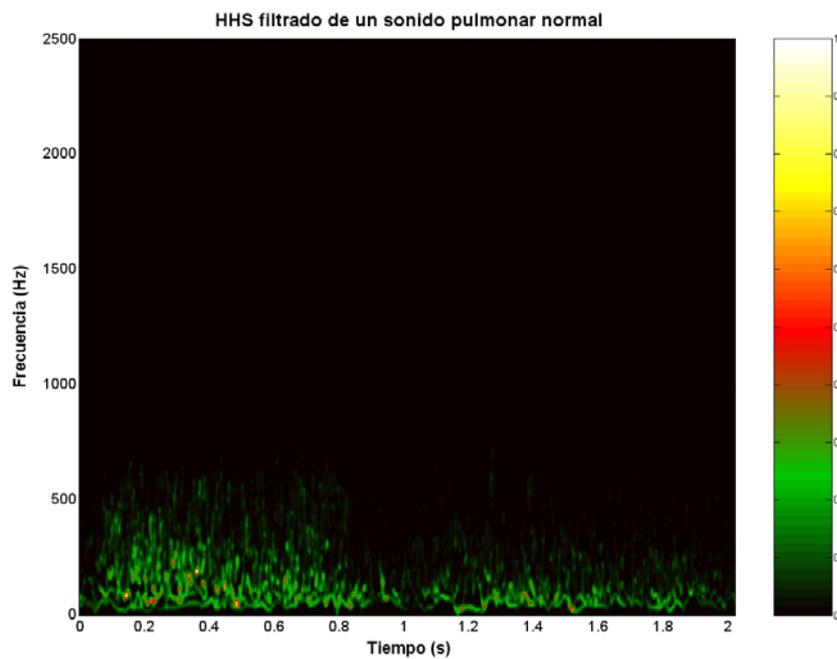


Fig. VI.176. Versión suavizada del Espectro de Hilbert-Huang del sonido pulmonar normal adquirido.

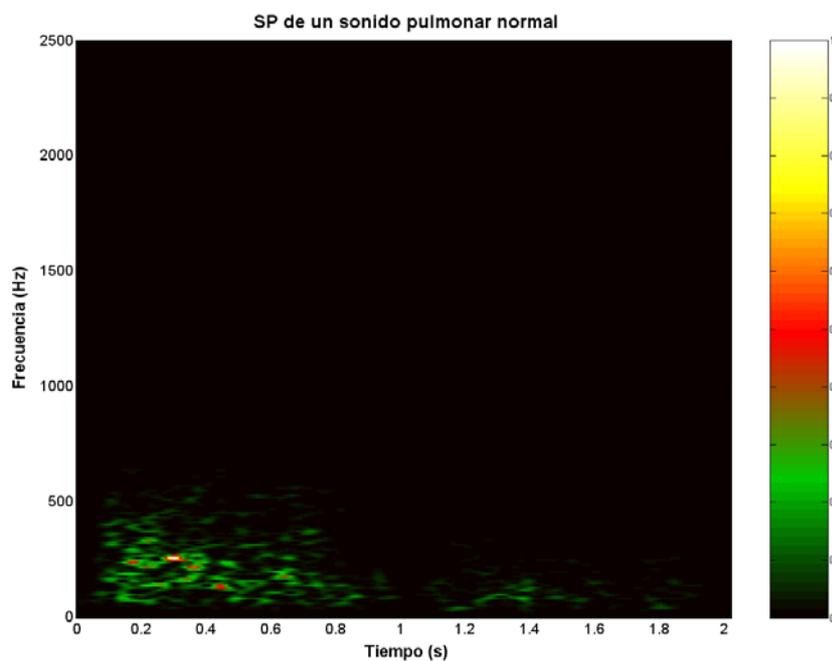


Fig. VI.177. Espectrograma del sonido pulmonar normal adquirido con una ventana Hamming de 100 ms de longitud.

VI.8.2 Sonido traqueal.

En la figura VI.178 se presenta una señal de sonido traqueal de un sujeto sano. La descomposición empírica de modos del sonido traqueal se presenta en la figura VI.179 mientras que la contribución de energía de cada modo de oscilación intrínseco se presenta en la figura VI.180. En la figura VI.181 se presentan las señales analíticas de los modos intrínsecos de oscilación seleccionados. El Espectro de Hilbert-Huang de la señal de sonido traqueal normal se presenta en la figura VI.182. De forma complementaria, se presenta una versión suavizada del Espectro de Hilbert de la señal en la figura VI.183. Finalmente, en la figura VI.184 se presenta el espectrograma del sonido traqueal adquirido, donde nuevamente se empleo una ventana Hamming de 100 ms de longitud para su obtención.

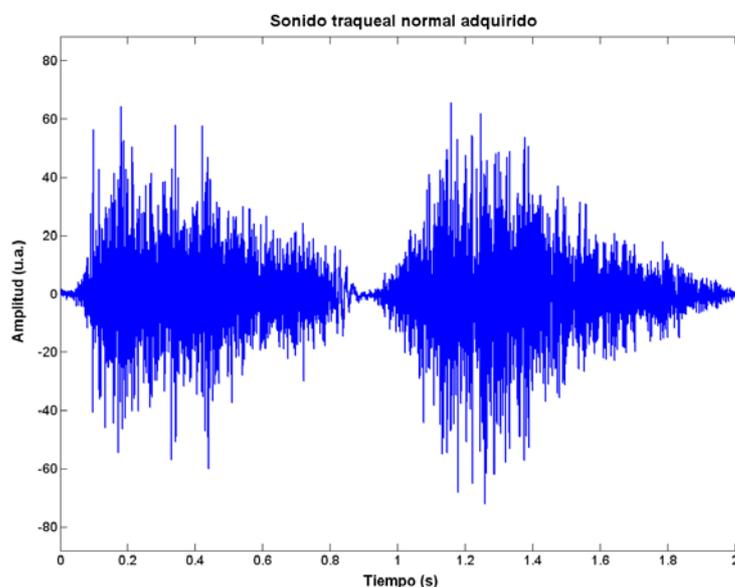


Fig. VI.178. Forma de onda del sonido traqueal normal adquirido a lo largo de un ciclo respiratorio.

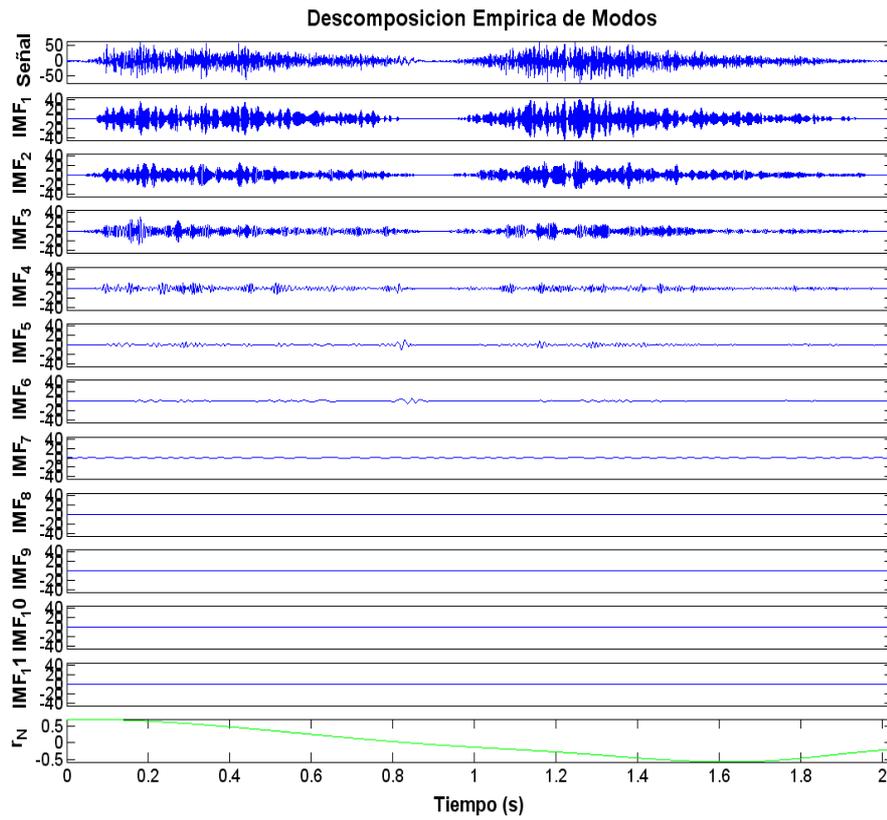


Fig. VI.179. Descomposición empírica de modos del sonido traqueal normal adquirido.

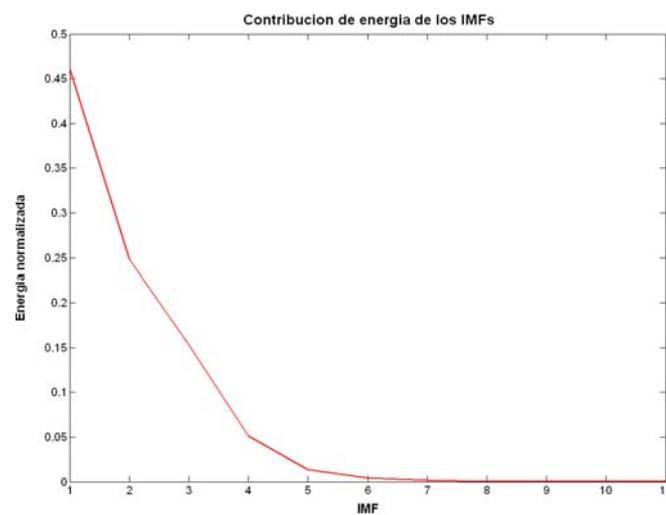


Fig. VI.180. Contribución de energía de cada modo de oscilación intrínseco del sonido traqueal normal adquirido.

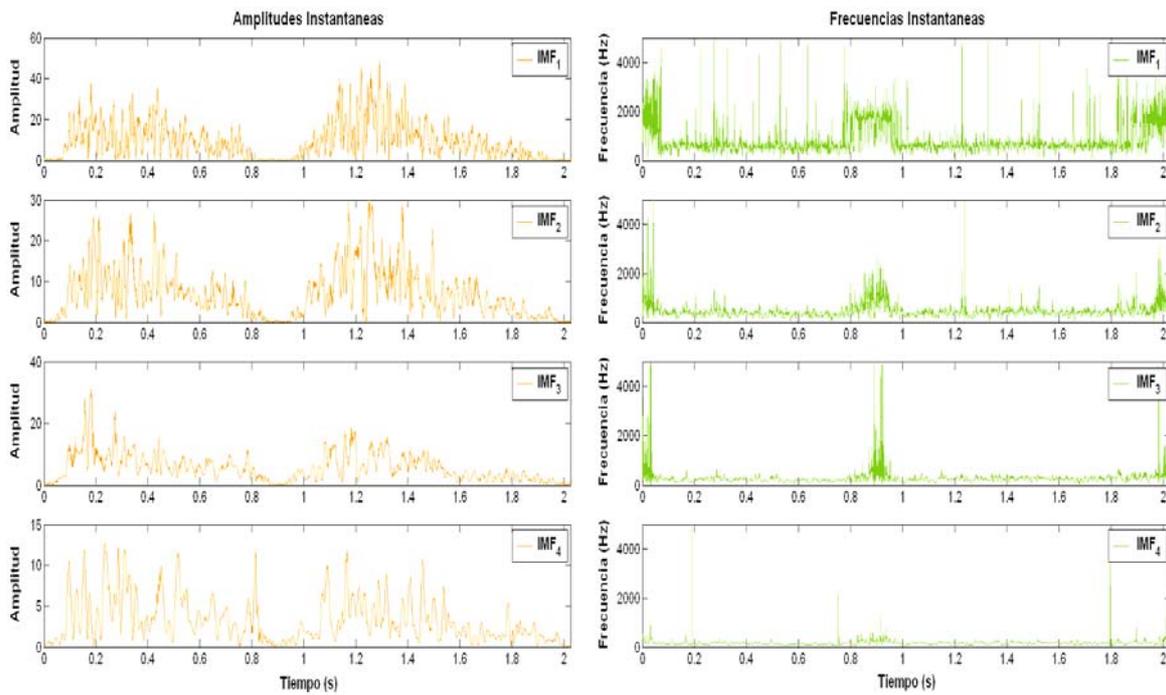


Fig. VI.181. Señales analíticas de los primeros cuatro modos de oscilación intrínsecos del sonido traqueal normal adquirido. Amplitud instantánea (columna izquierda) y Frecuencia instantánea (columna derecha).

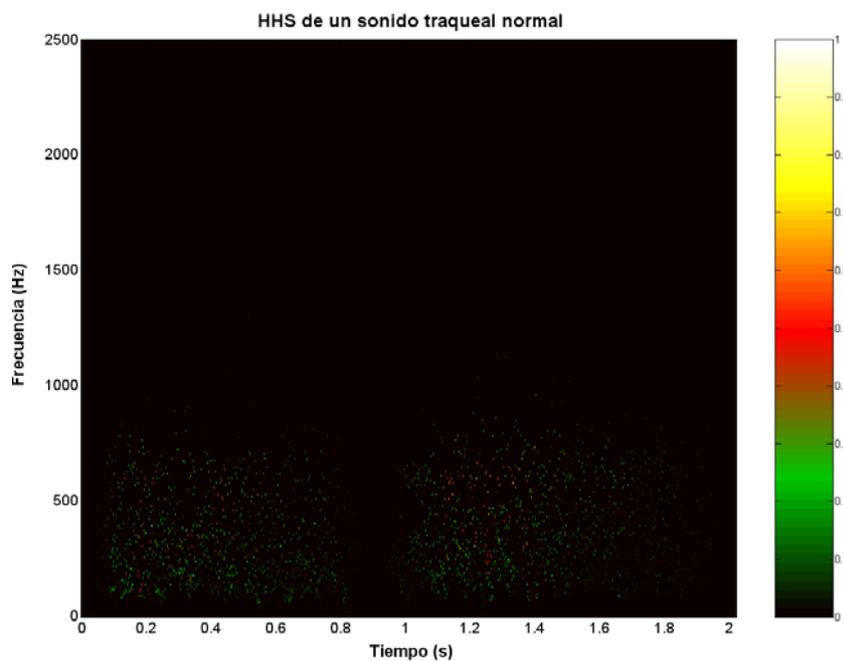


Fig. VI.182. Espectro de Hilbert-Huang del sonido traqueal normal adquirido.

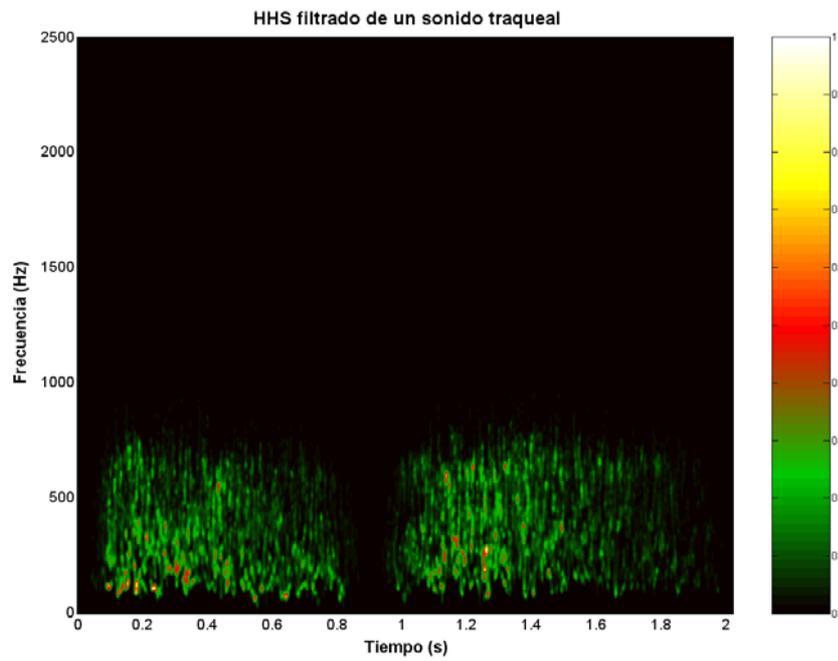


Fig. VI.183. Versión suavizada del Espectro de Hilbert-Huang del sonido traqueal normal adquirido.

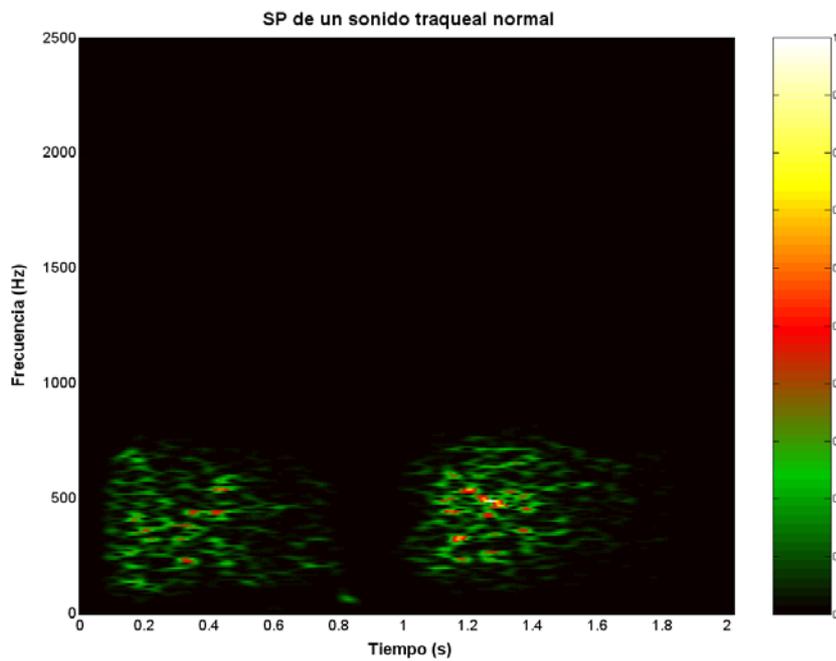


Fig. VI.184. Espectrograma del sonido traqueal normal adquirido.

VII. DISCUSIÓN

VII.1. Introducción.

A continuación se discuten los resultados obtenidos para las diversas representaciones tiempo-frecuencia de las señales biológicas y no biológicas, simuladas y reales presentados en el capítulo VI y complementados en el anexo I de esta tesis.

VII.2. Señal no biológica generada mediante el modelo TVAR.

Para el caso de los dos tonos generados mediante el modelo TVAR, se corroboró la dependencia del espectrograma con la longitud de la ventana de análisis. En las figuras VI.2 – VI.4 se observa que a medida que la longitud de la ventana aumenta, mejora la resolución en la frecuencia pero disminuye la resolución en el tiempo. Este comportamiento se ve reflejado en la comparación local de cada espectrograma con la representación tiempo-frecuencia teórica de la señal, donde en las figuras VI.5 – VI.7 se observan valores de correlación bajos en celdas superiores e inferiores a las subimágenes correspondientes a las de la señal, y gradualmente estos valores de correlación se desplazan hacia celdas horizontales adyacentes. La selección de los parámetros del espectrograma para esta señal sintética concuerda con lo esperado, es decir se seleccionó un espectrograma (figura VI.9) con una adecuada resolución de los componentes en frecuencia y con una buena transición en el tiempo de ocurrencia de ambos, además la selección puede corroborarse al observar las funciones kernel de cada espectrograma de forma sobrepuesta sobre la función de ambigüedad de la señal como se muestra en el anexo I.

Como se esperaba, la distribución Wigner-Ville de esta señal (figura VI.10) presenta términos de interferencia ubicados en la zona tiempo-frecuencia intermedia a los auto-términos de la señal. A pesar de la alta resolución tiempo-frecuencia de los auto-términos, la presencia de los términos de interferencia dificulta la legibilidad de la distribución dado que su intensidad resulta equiparable a los términos de la señal. Los valores bajos de correlación obtenidos durante la comparación local (figura VI.11) dan muestra de lo mencionado anteriormente.

Para el caso de la distribución Choi-Williams, a medida que el parámetro σ aumenta, el filtrado pasabajas de su función kernel permite el paso de una mayor cantidad de términos cruzados como se observa en las figuras VI.13 y VI.14. La distribución Choi-Williams seleccionada para esta señal (figura VI.16) mediante la comparación local resulta adecuada

después de observar el comportamiento de cada función kernel sobre la señal en el dominio de la función de ambigüedad (anexo I).

Para el caso del espectrograma modificado, la reasignación de los valores de los espectrogramas a sus centros de gravedad resulta en representaciones tiempo-frecuencia con una alta resolución en la frecuencia y con algunas dificultades en los extremos, lo cual habla de una baja dependencia en la ventana empleada para el análisis.

Con respecto a las representaciones tiempo-frecuencia obtenidas mediante la aplicación del método de Burg sobre segmentos ventaneados de la señal, se observa que de forma similar al espectrograma se tiene un compromiso en la resolución temporal y espectral a medida que aumenta la longitud de la ventana de análisis. Además, la técnica tiempo-frecuencia traslada las dificultades del método de Burg para la estimación espectral originando estimados con alta varianza. La representación tiempo-frecuencia P_{Burg} seleccionada presenta un equilibrio entre la resolución en la frecuencia y en el tiempo.

El empleo del algoritmo de mínimos cuadrados recursivo con factor de olvido fijo para producir la representación tiempo-frecuencia P_{RLS} , traslada las características de los estimados de los coeficientes del modelo autorregresivo. A decir, cuando se emplea un factor de olvido bajo la representación tiempo-frecuencia es capaz de seguir los cambios en los tonos de la señal pero a expensas de una gran varianza en los estimados y a medida que el factor de olvido se acerca a la unidad, la varianza disminuye pero a expensas de una lenta recuperación. En consecuencia, una P_{RLS} adecuada se obtuvo con un factor de olvido menor a la unidad que produjera un adecuado compromiso entre el tiempo de transición de los componentes en frecuencia y la varianza en la estimación de dichos componentes, tal como se observa en la representación P_{RLS} seleccionada (figura VI.27).

Por su parte, las representaciones tiempo-frecuencia obtenidas con un algoritmo de mínimos cuadrados con factor de olvido variable $P_{\text{RLS-VFF}}$, muestran que el parámetro de la varianza del ruido desempeña un papel similar al factor de olvido en el algoritmo RLS estándar, por lo que la adecuada selección de este parámetro resulta indispensable. La comparación local de cada representación $P_{\text{RLS-VFF}}$ (anexo I) muestra lo mencionado anteriormente. Finalmente, la representación $P_{\text{RLS-VFF}}$ seleccionada para esta señal resulta en una representación tiempo-frecuencia con una adecuada estimación tanto en varianza como en duración de los componentes de la señal.

Para el caso de las representaciones tiempo-frecuencia obtenidas al emplear el filtro de Kalman (P_{Kalman}), descrito en el capítulo IV, se observa (figura VI.31) que para un modelo autorregresivo y una varianza del ruido de medición dados, la varianza del ruido del proceso

se comporta de forma inversa al factor de olvido del algoritmo RLS, i.e. valores bajos de la varianza del ruido del proceso resultan en estimados espectrales con baja varianza pero con un tiempo de ajuste bastante largo, mientras que valores altos de la varianza del ruido del proceso resultan en estimados con mucha varianza pero con una adecuada transición entre los componentes. Nuevamente, la selección de la representación P_{Kalman} concuerda con la que se realizaría por inspección visual.

Con respecto al escalograma Morlet, se observa (especialmente para ventanas de corta duración) que la resolución espectral disminuye conforme aumenta la frecuencia, además se observa que, de forma similar al espectrograma, la resolución en el tiempo (resolución en la frecuencia) disminuye (aumenta) conforme la longitud de la *wavelet* se incrementa. Lo mencionado anteriormente respecto al compromiso en la resolución temporal y espectral, y la dependencia de esta última resolución con la frecuencia, queda evidenciado en la comparación local de cada escalograma con la representación teórica de la señal (anexo I). El escalograma seleccionado para esta señal presenta un adecuado compromiso en la resolución tiempo-frecuencia.

Respecto al espectro de Hilbert-Huang, se tiene que la descomposición de la señal resulta en varios modos de oscilación intrínsecos, sin embargo, el primer modo presenta la mayor correlación y contribución de energía con la señal original (figura VI.37). La estimación de la amplitud y la frecuencia instantánea de la señal analítica del primer modo intrínseco de oscilación (figura VI.38), a pesar de que proporciona una idea bastante adecuada, resulta en estimaciones ruidosas. Resulta interesante observar que para instantes de tiempo donde la amplitud instantánea estimada es pequeña, la frecuencia instantánea estimada presenta valores elevados. Es posible que la estimación de la frecuencia instantánea mediante el empleo de la derivada sea lo que ocasione dichos problemas. Para producir el espectro de Hilbert-Huang de la señal, se trató de compensar los cambios abruptos en la frecuencia instantánea mediante el empleo de un filtro de mediana y no mediante un filtro de media (con el propósito de preservar las transiciones entre los componentes en frecuencia de la señal). Debido a que el espectro de Hilbert de la señal (figura VI.39) sólo considera una frecuencia instantánea (señal 1D), su comparación local presenta valores máximos de correlación sobre una gran cantidad de subimágenes (anexo I) pero no en aquellas donde se ubican los componentes de la señal.

Finalmente, se observa que todas las técnicas brindan una adecuada representación tiempo-frecuencia de la señal simulada en el sentido de que indican la presencia de dos tonos puros de duración finita con diferentes tiempos de ocurrencia, con excepción de la

distribución Wigner-Ville cuya legibilidad se ve deteriorada por la presencia de términos cruzados. Con respecto a la estimación de la frecuencia instantánea se observa que las representaciones WVD, CWD, SPR, $P_{RLS-VFF}$, P_{Kalman} y HHS generan resultados adecuados al indicar una transición abrupta entre los componentes, sin embargo presentan dificultades en la estimación de la frecuencia instantánea al inicio de la señal. Por su parte, el SP, P_{Burg} y SC producen estimados de la frecuencia instantánea bastante suaves aún en los extremos, aunque con una transición bastante lenta entre ambos componentes, lo cual podría explicarse debido al empleo de una función ventana.

De acuerdo a los índices de desempeño mostrados en la tabla VI.II, el espectrograma reasignado resulta la técnica de análisis tiempo-frecuencia más adecuada para la señal del proceso TVAR. El valor de la resolución tiempo-frecuencia del espectrograma reasignado sólo se ve superado por el espectro de Hilbert-Huang. Sin embargo, es importante señalar que el espectro de Hilbert-Huang presenta la ventaja de ser una técnica de análisis adaptable, i.e. depende solamente de la señal y no de una función ventana u otros parámetros. Con respecto a los índices de desempeño, los valores de correlación promedio resultan mayores a los valores de correlación global. Por su parte, para esta señal los índices de desempeño basado en la estimación de la frecuencia instantánea de las representaciones pondera en mayor medida las estimaciones con poca varianza en comparación con la ponderación a grandes transiciones entre componentes, como muestran los valores obtenidos para el espectrograma, la distribución Choi-Williams, y el espectro de Hilbert-Huang y la distribución Wigner-Ville.

VII.3. Señal no biológica concatenando dos tonos puros.

Aunque para el caso de la señal no biológica obtenida mediante un proceso autorregresivo variante el tiempo, la señal generada presenta dos componentes de distintas frecuencias, su envolvente, o amplitud instantánea, dista de ser constante. Por esa razón se decidió probar las técnicas tiempo-frecuencia con una señal que presentará tales características.

Para el caso del espectrograma y la WVD existen comentarios similares a los presentados en el caso anterior. Para la distribución Choi-Williams, en el Anexo I, se observa que a medida que el parámetro σ aumenta, la *frecuencia de corte* de la función kernel exponencial se incrementa, lo que gradualmente incluye una mayor cantidad de términos cruzados, mejorando la resolución tiempo-frecuencia y tendiendo hacia la distribución Wigner-Ville de la señal. En particular, resulta interesante observar que un filtrado excesivo

de la WVD (parámetro σ pequeño) dispersa demasiado los auto-términos de la señal sobre el plano tiempo-frecuencia.

El espectrograma reasignado resulta en una gran concentración de auto-términos y en una supresión de los términos cruzados. A pesar de que los resultados son similares, como lo indican los valores de correlación promedio en el anexo I, algunos espectrogramas reasignados sobresalen debido a que la intensidad obtenida para los componentes es cercana a la unidad, a excepción de los valores extremos. Así, el espectrograma reasignado seleccionado para esta señal se asemeja demasiado a la representación tiempo-frecuencia teórica.

Comentarios similares al caso anterior se pueden establecer para el caso de la representación tiempo-frecuencia P_{Burg} .

Con respecto al empleo del algoritmo RLS estándar, se observa que para factores de olvido bajos (anexo I) las representaciones tiempo-frecuencia obtenidas son bastante parecidas entre sí, como lo indica el valor de correlación promedio, mostrando una adecuada estimación tanto en el tiempo como en la frecuencia de los componentes de la señal. Además, se observa que cuando el algoritmo trabaja en un ambiente estacionario (factor de olvido unitario), la representación P_{RLS} producida no logra estimar el segundo componente en frecuencia de la señal y por lo tanto para esta imagen se tiene el menor valor de correlación promedio.

Comentarios similares al caso anterior se pueden establecer para el caso de la representación tiempo-frecuencia mediante el algoritmo RLS-VFF (anexo I). La representación $P_{\text{RLS-VFF}}$ seleccionada para esta señal es bastante similar a la P_{RLS} seleccionada, con algunas diferencias alrededor de la transición entre ambos componentes.

Para el filtro de Kalman, escalograma y el espectro de Hilbert-Huang, presentaron comportamientos similares al caso de la señal modelada por un proceso TVAR (anexo I). Es importante señalar que el espectro de Hilbert-Huang de esta señal se asemeja bastante a la representación tiempo frecuencia ideal.

De acuerdo a los índices de desempeño de la tabla VI.III, la técnica de análisis tiempo-frecuencia más adecuada para las dos sinusoidales de duración finita es el espectro de Hilbert-Huang. A diferencia de lo obtenido para el proceso TVAR, el espectro de Hilbert-Huang de esta segunda señal presenta estimados constantes para los tonos puros, con oscilaciones en la transición que no sobrepasan el área delimitada por los lóbulos del espectrograma. Además, esta técnica ofrece una resolución tiempo-frecuencia equiparable a la resolución de la representación tiempo-frecuencia teórica de la señal.

VII.4. Primer sonido cardíaco simulado y sus componentes.

VII.4.1. Componente muscular (s_m).

A diferencia de las señales no biológicas anteriores, para el componente muscular del primer sonido cardíaco simulado su función de ambigüedad (anexo I) resulta en una morfología tal que la discriminación entre los auto-términos y los términos cruzados resulta complicada.

Para el caso del espectrograma, se observa que para la ventana rectangular aparecen términos de interferencia que disminuyen la legibilidad de la representación (anexo I). El espectrograma seleccionado para s_m simulado ofrece un adecuado compromiso entre las resoluciones tiempo y frecuencia, así como una baja presencia de términos cruzados.

La distribución Wigner-Ville de la señal da indicación de los términos cruzados de s_m . Se observa que existe una gran cantidad de términos de interferencia, donde los de mayor intensidad se ubican próximos a los auto-términos.

Con respecto a la distribución Choi-Williams, se observa que para valores de σ pequeños las distribuciones tiempo-frecuencia (anexo I) resultan bastante afectadas al grado de que incluso hacia el final de la señal, para instantes de tiempo en los cuales la intensidad de la señal es muy baja, aparecen términos de considerable energía que al parecer corresponden con los del inicio de la señal. Gradualmente, la concentración de los auto-términos va aumentando con la consecuente inclusión de los términos cruzados, y la CWD tiende a la WVD de la señal. Así, la CWD seleccionada resulta en una ligera limpieza de los términos cruzados a expensas de un deterioro en la resolución tiempo-frecuencia.

La reasignación de los espectrogramas de la señal (anexo I) resulta en una mayor concentración de los auto-términos de la señal alrededor de la frecuencia instantánea de la señal, lo cual se refleja en su comparación local con la representación teórica mediante una mayor cantidad de valores de correlación cercanos a la unidad con respecto a los obtenidos para el espectrograma. Para el caso de la ventana rectangular, la presencia de una gran cantidad de términos cruzados en los espectrogramas, se ve reflejada en sus versiones reasignadas. El espectrograma reasignado seleccionado para esta señal presenta una buena resolución espectral pero no da indicación de lo acontecido para instantes de tiempo tempranos de la señal, aunque tampoco ninguna de las distribuciones de la clase general mencionadas anteriormente otorgó dicha información.

Para el caso de las representaciones tiempo-frecuencia obtenidas mediante el método de Burg (anexo I) se observa que, para un orden dado, la resolución en la frecuencia se hace

difusa para ventanas de tiempo cortas, además de que presenta ondulaciones. Para el caso de la ventana Blackman-Harris se tiene que a medida que el orden del modelo AR se incrementa, aparecen otros componentes espurios en la frecuencia los cuales incluso vuelven ilegible la representación tiempo-frecuencia, alejándola de la representación teórica de la señal. El comportamiento discutido anteriormente se refleja en la comparación local. La representación tiempo-frecuencia P_{Burg} seleccionada para s_m genera una adecuada idea de la variación temporal del espectro de la señal, incluso al inicio de la señal aunque la estimación espectral presenta un corrimiento considerable.

Con respecto a la aplicación del algoritmo RLS para la estimación de la potencia instantánea de s_m , se encontró que únicamente para un modelo TVAR de orden 2 se produjeron representaciones tiempo-frecuencia que concuerdan con lo indicado por la representación teórica de la señal, y para los órdenes de modelo restantes se obtuvieron imágenes ilegibles (anexo I). Lo anterior puede deberse a que, a diferencia del método de Burg para la estimación de los coeficientes AR, el algoritmo RLS no asegura un modelo estable, i.e. los polos no se ubican necesariamente al interior del círculo unitario, lo cual origina valores de intensidad bastante dispares entre un instante de tiempo y otro. A pesar de que las disparidades en la magnitud de la intensidad trataron de compensarse mediante el empleo de un filtro de mediana sobre las representaciones tiempo-frecuencia, no se obtuvieron resultados adecuados. Por otra parte, a pesar de que la amplitud instantánea teórica de la señal es insignificante después de 60 ms, el empleo de un factor de olvido unitario provoca que los últimos estimados espectrales se extiendan hacia el final de la señal.

Las representaciones tiempo-frecuencia obtenidas con el algoritmo RLS con factor de olvido variable presentan las mismas dificultades que las discutidas para el caso del algoritmo RLS estándar. La representación P_{RLS} seleccionada se asemeja a la obtenida con el algoritmo RLS estándar.

Para la representación tiempo-frecuencia P_{Kalman} del componente muscular simulado (anexo I), se observa que para todos los órdenes del modelo TVAR, sin importar el valor de la varianza del ruido del proceso, la intensidad de la señal se extiende hacia instantes de tiempo mayores a los 60 ms de forma similar al caso del algoritmo RLS con un factor de olvido unitario. Este comportamiento resulta visible en la comparación local de las representaciones con la representación tiempo-frecuencia deseada. Así, la representación P_{Kalman} de s_m presenta dificultades tanto al inicio de la señal, como en el caso de P_{RLS} y $P_{\text{RLS-VFF}}$, como también para instantes de tiempo posteriores a la señal.

Con respecto a los escalogramas Morlet, se observa la división no uniforme del plano tiempo-frecuencia (anexo I). En particular, para la *wavelet* Morlet de menor duración se observa que a altas frecuencias la resolución espectral es pobre y para la *wavelet* de mayor duración se observa que a bajas frecuencias la resolución temporal es pobre dispersando demasiado el componente de la señal en el tiempo. El escalograma seleccionado para esta señal resulta en un compromiso entre ambas dificultades.

Aunque la descomposición empírica de modos de la señal (anexo I) produce cinco modos de oscilación intrínsecos, el primer IMF presenta una mayor correlación y contribución de energía a la señal de s_m simulada. La amplitud instantánea estimada para dicho modo de oscilación indica que la mayor parte de la energía de la señal se concentra en los primeros 60 ms y posteriormente decrece abruptamente. Por su parte, la estimación de la frecuencia instantánea indica que el espectro de la señal se incrementa durante los primeros 30 ms para posteriormente centrarse alrededor los 100 Hz durante los siguientes 30 ms. A pesar de que estos resultados concuerdan con la representación tiempo-frecuencia teórica de la señal, el HHS de la señal presenta dificultades en instantes iniciales y finales donde las frecuencias instantáneas estimadas crecen abruptamente.

De acuerdo a los índices de desempeño de la tabla VI.IV, que se basan en la correlación entre la imagen tiempo-frecuencia estimada y teórica, el espectro de Hilbert-Huang, el espectrograma reasignado y la representación P_{Burg} resultan las técnicas más adecuadas para describir la variación temporal del espectro del componente muscular simulado de S1. Estas tres técnicas de análisis producen la mejor concentración de energía alrededor de la frecuencia instantánea de la señal simulada como lo indica el índice de resolución tiempo-frecuencia. Por su parte, de acuerdo al índice de desempeño NRMSE, el espectrograma y el escalograma resultan las representaciones más adecuadas para esta señal, sin embargo, tales representaciones muestran los peores valores de resolución tiempo-frecuencia. Así, aunque la frecuencia instantánea estimada con el centroide del espectrograma y el escalograma presenten un error pequeño, su ancho de banda instantáneo estimado es muy alto indicando una gran dispersión, o varianza, alrededor de los estimados de frecuencia instantánea. Por otra parte, las representaciones $P_{\text{RLS-VFF}}$, P_{Burg} y el HHS resultan las técnicas de análisis que producen los valores de correlación más altos entre sus frecuencias instantáneas estimadas y la frecuencia instantánea teórica de la señal. Con lo anterior en mente, el espectro de Hilbert-Huang y la representación P_{Burg} resultan las técnicas de análisis tiempo-frecuencia más adecuadas para el componente muscular de S1 simulado. Sin embargo el desempeño de la representación P_{Burg} esta condicionado a la selección del orden del modelo

AR, el tipo y longitud de la ventana de análisis, como se ilustra en el apartado correspondiente a esta técnica.

VII.4.2. Componente valvular simulado (s_v).

Mientras que para el componente muscular simulado su frecuencia instantánea se incrementa hasta un valor tope para mantenerse ahí por un intervalo de tiempo, para el caso del componente valvular simulado debe tenerse en mente que se trata de una señal multicomponente, para la cual no podemos hablar de una sola frecuencia instantánea dado que para cada instante de tiempo conviven dos componentes en frecuencia. Nuevamente, la función de ambigüedad de esta señal se encuentra muy concentrada alrededor del origen (anexo I), aunque observando la señal en el tiempo se espera que existan términos cruzados en la región tiempo-frecuencia intermedia a los dos monocomponentes, que se traducirán en términos paralelos al eje doppler en dominio de la función de ambigüedad y con una distancia lejana al origen.

Con respecto al espectrograma de s_v (anexo I), se tiene que para ventanas de tiempo cortas la resolución en la frecuencia es baja pero la resolución en el tiempo es adecuada al grado que se obtiene información del tiempo de aparición de los componentes, aunque no de su duración. Conforme la duración de las ventanas de análisis se incrementa, mejora la resolución espectral a expensas de una baja resolución temporal, al grado de indicar que los componentes aparecen en instantes de tiempo anteriores a los simulados. En los espectrogramas obtenidos también resulta notoria la presencia de términos cruzados. El espectrograma seleccionado para esta señal, conforme a la comparación local, brinda una adecuada resolución en la frecuencia indicando los límites finales de los componentes, pero no su tiempo de aparición.

La distribución Wigner-Ville de la señal, presenta una resolución temporal que permite definir la extensión en el tiempo de sus dos componentes con una adecuada resolución en la frecuencia alrededor de las dos frecuencias instantáneas. Sin embargo, la WVD de la señal presenta términos cruzados con una intensidad equiparable a la del auto-término de menor frecuencia y que se extienden a lo largo de la duración de la señal.

Para el caso de la distribución Choi-Williams de s_v (anexo I), nuevamente se observa el efecto pasabajas de su kernel exponencial que gradualmente va mejorando la resolución tiempo-frecuencia de los auto-términos pero a expensas de una mayor inclusión de términos cruzados. La CWD seleccionada para esta señal, si bien presenta una mejoría en la resolución tiempo-frecuencia en comparación con el espectrograma, también presenta términos cruzados

como en el caso de la WVD. Sin embargo, dadas las características de los términos cruzados en el dominio de la función de ambigüedad no se esperaba suprimirlos por completo debido la morfología del kernel exponencial de la CWD.

Los espectrogramas reasignados del componente muscular simulado (anexo I) muestran un aumento de la resolución tiempo-frecuencia comparado con sus respectivos espectrogramas, alrededor del tiempo de aparición de la señal. Para ventanas de corta duración, debido a la baja resolución en la frecuencia de los espectrogramas sus versiones reasignadas conservan componentes que se extienden verticalmente de un término de la señal a otro. Además, en algunos casos los espectrogramas reasignados presentan ondulaciones alrededor del componente de baja frecuencia. El espectrograma reasignado seleccionado para s_v simulado proporciona una gran concentración de los componentes alrededor de sus regiones tiempo-frecuencia simuladas, siendo menor al inicio de los componentes.

Al aplicar el método de Burg sobre segmentos ventaneados de s_v (anexo I) para obtener representaciones tiempo-frecuencia, no se logró una representación adecuada de la señal con los órdenes del modelo AR y las diversas ventanas (tipos y longitudes) empleadas. Se tiene que para la ventana Hamming, para cualquier orden del modelo, las representaciones tiempo-frecuencia alcanzadas consisten en un componente centrado alrededor de la frecuencia media de ambos componentes (100 Hz) y que presenta un comportamiento ondulatorio, comportamiento que se disminuye conforme aumenta la longitud de la ventana. Para el caso de la ventana Blackman-Harris se tiene el mismo comportamiento, aunque a partir de un modelo de orden 3 aparecen indicios de los términos simulados para la señal pero corridos en la frecuencia. Si bien la representación P_{Burg} seleccionada para s_v señala la existencia de dos componentes en frecuencia para instantes de tiempo mayores a los 25 ms, para instantes previos la representación tiempo-frecuencia obtenida no es correcta.

Por su parte, la representación P_{RLS} también presenta dificultades (anexo I). Para modelos TVAR de orden 2 y 3, a partir de los 10 ms el espectro estimado presenta un comportamiento ondulatorio al decrecer desde una frecuencia alrededor de los 100 Hz hasta una frecuencia nula para posteriormente crecer y repetir el proceso. Para estos órdenes del modelo, con un factor de olvido cercano a la unidad las ondulaciones se estabilizan hasta convertirse en un componente de frecuencia constante con intensidad elevada. Para esta señal, se presentó de nuevo el problema de que las representaciones tiempo-frecuencia producidas no proporcionaron información relacionada con la señal. Para un modelo TVAR de orden 5, las representaciones empiezan a mostrar componentes alrededor de los simulados para algunos valores del factor de olvido. Haciendo un gran esfuerzo, se observa que la P_{RLS}

seleccionada para esta s_v indica la presencia, entre otras más, de dos componentes alrededor de 50 y 150 Hz, aunque estos aparecen para instantes de tiempo bastante alejados de los simulados. De forma adicional a los órdenes empleados para todas las representaciones generadas mediante el modelo TVAR, para esta técnica se obtuvieron P_{RLS} con órdenes del modelo en el intervalo 6-15, para los cuales, con factores de olvido bajos, gradualmente se va aumentando la concentración de los componentes alrededor de las regiones tiempo-frecuencia esperadas, aunque en ninguna de ellas los componentes inician en 10 ms.

Al igual que con los resultados obtenidos con las señales no biológicas y con el componente muscular simulado, la representación tiempo-frecuencia generada con el algoritmo RLS-VFF (anexo I) muestra que la varianza del ruido esperado desempeña un papel similar al del factor de olvido en el algoritmo RLS estándar. Las $P_{RLS-VFF}$ obtenidas presentan características similares a las P_{RLS} , sin embargo, para esta técnica se empezaron a obtener dos componentes que concuerdan con los simulados a partir de un modelo TVAR de orden 4. En comparación con la representación P_{RLS} seleccionada, la $P_{RLS-VFF}$ seleccionada presenta una mejor legibilidad y resolución de los componentes de s_v simulado.

Con respecto a la representación producida por el filtro de Kalman (anexo I), se observa nuevamente que para valores bajos de la varianza del proceso, las ondulaciones del componente estimado cercano a los 100 Hz se estabilizan rápidamente y conforme el valor de la varianza aumenta, también lo hacen las ondulaciones. Para esta técnica de análisis tiempo-frecuencia no fue posible producir una representación tiempo-frecuencia adecuada del componente valvular simulado. La P_{Kalman} seleccionada para esta señal consiste básicamente de una región constante en el plano tiempo frecuencia a partir aproximadamente 15 ms y hasta alrededor de 100 Hz.

Para el caso de los escalogramas Morlet de s_v (anexo I), se observa la estructura piramidal en la división del plano tiempo-frecuencia por la transformada *wavelet*. Además, se observa la presencia de términos cruzados en la región tiempo-frecuencia intermedia entre los dos monocomponentes, como lo señala la relación del escalograma con la distribución Wigner-Ville. El escalograma seleccionado para esta señal, presenta dos componentes en el plano tiempo-frecuencia alrededor de los simulados aunque con un gran suavizado.

La descomposición empírica de modos de s_v simulado (anexo I) produce cuatro modos intrínsecos de oscilación, de los cuales se seleccionaron los dos primeros para producir la representación tiempo-frecuencia de la señal. El primer modo de oscilación corresponde a componente de 150 Hz y el segundo a 50 Hz. A pesar de que la amplitud instantánea estimada para cada componente presenta un decaimiento con el tiempo, los valores de amplitud para

instantes de tiempo menores a los 10 ms resultan elevados debido a la interpolación empleada para producir cada modo intrínseco de oscilación, como se observa en la descomposición de la señal. Así, aunque la estimación de la frecuencia instantánea de cada modo de oscilación resulta adecuada, los valores de intensidad estimados degradan el espectro de Hilbert-Huang para instantes de tiempo previos a la aparición del componente valvular simulado.

Para el caso del componente valvular simulado de S1, la señal está compuesta por dos componentes de diferentes frecuencias que conviven en el tiempo. Por tal razón, no se obtuvieron estimados de la frecuencia instantánea mediante el centroide de las representaciones, ya que se tienen dos frecuencias instantáneas bien definidas. De acuerdo a los índices de desempeño basados en las correlaciones local y global (tabla VI.V), el espectro de Hilbert-Huang y el espectrograma reasignado resultan las técnicas de análisis más adecuadas para el componente valvular de S1 simulado. Además, estas dos técnicas de análisis producen los mejores valores de resolución tiempo-frecuencia. En particular, la reasignación mejora en gran medida la concentración de energía del espectrograma dejando ver el esqueleto de la distribución, sin embargo esta técnica requiere la selección de la longitud y tipo de la ventana de análisis. Por su parte, el espectro de Hilbert-Huang resulta visualmente similar a la representación tiempo-frecuencia teórica de la señal, pero también presenta dificultades debido al empleo del interpolador y de la derivada para obtener la señal analítica y la frecuencia instantánea, respectivamente.

VII.4.3. Primer sonido cardiaco, S1.

La señal de primer sonido cardiaco simulada con base en la suma de los componentes muscular y valvular anteriores, genera una forma de onda difícil de analizar. La inspección visual de la función de ambigüedad de la señal no revela gran información acerca de los auto-términos ni de los términos cruzados. La TFR ideal de S1 posee su mayor intensidad en el cruce entre la señal de baja frecuencia del componente valvular y el componente muscular, hecho que ocasiona que la intensidad de la señal de alta frecuencia del componente valvular se vea ensombrecida.

El espectrograma seleccionado para S1 simulado indica la presencia de un componente de frecuencia centrado alrededor de los 50 Hz, de un componente cuya frecuencia se incrementa hasta mantenerse alrededor de los 100 Hz y de un componente de corta duración y baja intensidad alrededor de los 150 Hz. Sin embargo, la resolución de la representación tiempo-frecuencia es bastante baja.

La resolución tiempo-frecuencia de los componentes de la señal mejora en la distribución Wigner-Ville, aunque se ve opacada por la presencia de términos de interferencia que la vuelven ruidosa. La WVD de S1 simulado no proporciona información acerca del comportamiento del componente muscular sobre el intervalo de 0 a 10 ms.

Por su parte, la distribución Choi-Williams seleccionada para esta señal simulada resulta en una ligera limpieza de la WVD. Al igual que la WVD, la CWD de S1 pareciera seguir inicialmente un componente de alta frecuencia sin embargo, posteriormente decrece abruptamente para seguir al componente centrado alrededor de los 100 Hz.

El espectrograma reasignado seleccionado, al igual que el espectrograma, indica el comportamiento de los tres componentes de la señal, aunque en forma similar a WVD y CWD sólo sigue al componente valvular de mayor frecuencia durante un corto periodo de tiempo. Hasta ahora, sólo el espectrograma ha producido un componente centrado alrededor de los 150 Hz de una duración relativamente adecuada, aunque esto se debe más bien a la pobre resolución tiempo-frecuencia obtenida con su kernel de tipo pasabajas.

Para el caso de las representaciones tiempo-frecuencia paramétricas obtenidas mediante el modelo TVAR, ninguna técnica de análisis produjo una representación que reflejara lo simulado en la señal. Al parecer, todas ellas, con excepción del algoritmo RLS-VFF, siguieron en mayor medida al componente muscular. En el caso del algoritmo RLS estándar y el filtro de Kalman no lograron seguir los cambios hacia el final de la señal de S1 simulada.

Con respecto al escalograma seleccionado, básicamente resulta en una versión similar al espectrograma, donde la diferencia más notoria es en la extensión del componente valvular de baja frecuencia, que era de esperarse debido al empleo de ventanas de tiempo largas a bajas frecuencias para producir el análisis multiresolución.

La descomposición empírica de modos de S1 simulado (anexo I) produce cinco modos de oscilación intrínsecos y la señal de residuo. De estos componentes, los dos primeros son los que contribuyen con mayor energía a la reconstrucción de la señal original y por lo tanto, fueron seleccionados para producir el espectro de Hilbert-Huang. La frecuencia instantánea estimada del primer modo intrínseco de oscilación crece abruptamente para después decrecer con una velocidad similar y establecerse momentáneamente alrededor de los 150 Hz (componente valvular de alta frecuencia), finalmente su valor decrece para oscilar alrededor de los 100 Hz (meseta en frecuencia del componente muscular). Por su parte, la frecuencia instantánea estimada del segundo modo intrínseco de oscilación parte alrededor de los 25 Hz y se incrementa alrededor de los 15 ms (parte inicial del componente muscular) para

finalmente, oscilar alrededor de los 50 Hz (componente valvular de baja frecuencia). Los cambios repentinos en las frecuencias instantáneas para seguir un componente de la señal en particular quizá se deban a la intensidad relativa de cada monocomponente.

Al igual que para el componente valvular del primer sonido cardiaco simulado, para S1 no se obtuvieron estimados de la frecuencia instantánea a partir de las imágenes tiempo-frecuencia. Los índices de desempeño de la tabla VI.VI indican que las mejores representaciones tiempo-frecuencia del primer sonido cardiaco simulado corresponden al espectrograma reasignado y al espectro de Hilbert-Huang. Adicionalmente, estas representaciones presentan los mejores índices de resolución tiempo-frecuencia para S1 simulado.

VII.4.4. Señales de S1 reales adquiridas.

Como se observó anteriormente, el análisis de la señal simulada del primer sonido cardiaco resulta complicado. De las técnicas de análisis tiempo-frecuencia anteriores, se obtuvo que el espectrograma reasignado y el espectro de Hilbert-Huang presentan la mayor concentración de energía alrededor de las frecuencias instantáneas simuladas. Por lo tanto, estas dos técnicas, junto con el espectrograma se aplicaron a señales de S1 reales adquiridas. En las tres señales procesadas, el espectrograma proporcionó la peor resolución tiempo-frecuencia en el análisis de los componentes, sin embargo la representación refleja la existencia de un solo componente alrededor de 25 Hz, con una mayor intensidad cerca de los 150 ms. El espectrograma reasignado por su parte, también refleja la presencia de un componente alrededor de los 25 Hz, con una intensidad constante para todos los casos. Con respecto al espectro de Hilbert-Huang, se obtuvieron varios modos de oscilación donde los primeros dos se pueden considerar ruido de fondo de la señal analizada. A partir del IMF3, la energía de los modos de oscilación se incrementa en todos los casos analizados. Del análisis visual de los espectros de Hilbert-Huang se puede concluir que los S1 analizados contienen preponderantemente una sola frecuencia constante de 25 Hz, que se asemeja al componente de baja frecuencia del componente valvular de S1. Sin embargo, para el segundo caso analizado (Fig. VI.66) se observa un componente que incrementa su frecuencia desde los 10 Hz, a los 20 ms, hasta llegar a los 50 Hz, a los 40 ms. Este comportamiento se asemeja al componente muscular de S1. Sin embargo, en los casos analizados no se obtuvo un comportamiento similar a los componentes de la representación teórica (Fig. VI.50.b), confirmando que el análisis del primer ruido cardiaco resulta complicado

VII.5. Segundo sonido cardiaco simulado y sus componentes.

VII.5.1. Componente aórtico (A2).

Dada la frecuencia instantánea de la señal su función de ambigüedad (anexo I) tiene la morfología de una elipse, cuyo eje mayor presenta una ligera rotación. Con esto en mente es posible anticipar el comportamiento de las representaciones tiempo-frecuencia de la clase general.

La distribución Choi-Williams de A2 simulado visualmente luce igual a la distribución Wigner-Ville. Además, debido a la morfología del kernel exponencial de la CWD a partir de un valor del parámetro σ los cambios entre una CWD y la siguiente son muy pequeños, debido a que se ha incluido la mayor parte de los auto-términos de la señal. Por su parte, el espectrograma tiene una menor resolución tiempo-frecuencia en comparación con las representaciones anteriores, lo cual concuerda con lo esperado debido a que las funciones de ambigüedad de su ventana se encuentran muy concentradas alrededor del origen.

El espectrograma reasignado seleccionado para A2 presenta un aumento en la resolución tiempo-frecuencia. Aunque dadas las características de los espectrogramas de la señal, el espectrograma reasignado no indica la caída exponencial en la frecuencia instantánea del componente aórtico simulado.

La representación tiempo-frecuencia obtenida mediante el algoritmo de Burg presenta una alta resolución tiempo-frecuencia al igual que para el espectrograma reasignado. Sin embargo, P_{Burg} presenta oscilaciones debidas al corrimiento en la estimación del espectro de potencia.

Las representaciones tiempo-frecuencia obtenidas mediante los algoritmos RLS estándar y RLS-VFF resultan bastante similares, con diferencias en el tiempo de duración de la señal. En ambos casos, la estimación no genera información de lo acontecido con el espectro de la señal para instantes de tiempo menores a los 100 ms.

Por su parte, el filtro de Kalman presenta las mismas dificultades en la estimación del espectro de la señal para instantes de tiempo menores a 10 ms. En un pequeño intervalo de tiempo la estimación presenta una caída, pero el algoritmo no sigue los cambios en la señal y la estimación prevalece incluso para instantes de tiempo donde la señal ya no existe.

El escalograma seleccionado resulta básicamente en un espectrograma con una ventana de tiempo larga a bajas frecuencias y una ventana corta a altas frecuencias.

El espectro de Hilbert-Huang muestra una caída en frecuencia más pronunciada que el resto de las representaciones tiempo-frecuencia. Además, los valores de intensidad estimados

resultan muy cercanos a los valores indicados por la representación teórica de la señal. Sin embargo, el HHS no estima adecuadamente la frecuencia instantánea de A2 al inicio de la señal y presenta dificultades en la estimación hacia el final de la representación.

De acuerdo a los valores de correlación local y global mostrados en la tabla VI.VII, el espectro de Hilbert-Huang, el espectrograma reasignado y la representación P_{Burg} son las técnicas de análisis tiempo-frecuencia más adecuadas para el componente aórtico. Sin embargo, ninguna de estas tres técnicas estima adecuadamente la frecuencia instantánea al inicio de la señal. Al margen de los índices de desempeño, el espectro de Hilbert-Huang produce valores de intensidad similares a los de la representación teórica de la señal, además en comparación con el espectrograma reasignado muestra de forma más clara el descenso en el espectro del componente aórtico simulado.

VII.5.2. Componente pulmonar simulado (P2).

El componente pulmonar de S2 simulado, básicamente es una versión trasladada en tiempo del componente A2 simulado de S2, diferenciándose por los valores extremos de sus frecuencias instantáneas.

Las representaciones tiempo-frecuencia de la clase general seleccionadas, i.e. el espectrograma, la distribución Wigner-Ville y la distribución Choi-Williams, presentan características similares a sus contrapartes seleccionada para el componente P2 simulado, como se esperaba dado que la clase general de representaciones es una clase invariante al corrimiento en el tiempo.

Así, el espectrograma reasignado seleccionado para P2 simulado nuevamente no indica la caída exponencial en la frecuencia instantánea de la señal debido a las características de su espectrograma correspondiente.

Por su parte, las representaciones tiempo-frecuencia obtenidas mediante el modelado TVAR presentan características similares a las obtenidas para A2 simulado. La P_{Burg} de la señal muestra una gran resolución tiempo-frecuencia en comparación con P_{RLS} , $P_{\text{RLS-VFF}}$ y P_{Kalman} , aunque presenta corrimientos en las frecuencias estimadas. Ninguna de estas representaciones paramétricas proporciona información del decaimiento exponencial de P2 simulado, sobre todo en el inicio de la señal.

El escalograma seleccionado para P2 simulado, también presenta las mismas características que el escalograma seleccionado para el componente A2 simulado, lo cual era de esperarse debido a la relación existente entre esta representación y la WVD.

La descomposición empírica del componente P2 simulado (anexo I) arroja diversos modos intrínsecos de oscilación, siendo el primero de ellos el que presenta mayor correlación con la señal original. La interpolación empleada para generar el primer modo de oscilación intrínseco presenta oscilaciones en el IMF1 inexistentes en P2 simulado, tales oscilaciones se ven reflejadas en una inadecuada oscilación de la amplitud y frecuencia instantáneas de la señal analítica. Por lo tanto, el HHS generado para P2 simulado dista de lo esperado al inicio de la señal. Nuevamente, se observan problemas en la estimación de la frecuencia instantánea al inicio y final de la señal simulada, aunque nuevamente los valores de intensidad resultan bastante similares a los simulados para la parte intermedia de la señal.

De los valores de la tabla VI.VIII, el espectrograma reasignado, P_{Burg} y el espectro de Hilbert-Huang resultan las mejores representaciones tiempo-frecuencia del componente pulmonar de S2 simulado de acuerdo a la correlación local. La mejor concentración de energía alrededor de la frecuencia instantánea la presentan el espectro de Hilbert-Huang, el espectrograma reasignado y la representación P_{Burg} . Así, los diversos índices de desempeño señalan diferentes representaciones tiempo-frecuencia como aquellas con el mejor desempeño. En general, el espectrograma reasignado, las representaciones paramétricas basadas en el modelado TVAR y el espectro de Hilbert-Huang ofrecen mejores resultados que los obtenidos con el espectrograma.

VII.5.3. Segundo sonido cardiaco simulado (S2).

La forma de onda del segundo sonido cardiaco simulado no da indicación acerca de la presencia de los componentes aórtico y pulmonar simulados.

El espectrograma seleccionado para esta señal muestra dos lóbulos de energía centrados alrededor de los instantes de tiempo donde la energía de cada componente simulado es más intensa. Sin embargo, su resolución tiempo-frecuencia es baja y no muestra un decaimiento en frecuencia correlacionado al componente aórtico.

La distribución Wigner-Ville mejora la resolución tiempo-frecuencia pero su lectura se dificulta por la presencia de gran cantidad de términos cruzados. Los términos cruzados, debido a su estructura en el dominio de la función de ambigüedad no son eliminados por la distribución Choi-Williams, por lo que ambas distribuciones resultan muy similares.

El espectrograma reasignado seleccionado para S2 simulado presenta una mejoría notable en la resolución tiempo-frecuencia con respecto a la obtenida con el espectrograma. A pesar de no mostrar un decaimiento en la frecuencia para A2, alrededor de los 35 ms, el espectro se incrementa ligeramente para nuevamente decrecer hasta un valor límite.

La representación P_{Burg} seleccionada presenta una alta resolución tiempo-frecuencia, pero presenta corrimientos en las estimaciones del espectro que se traducen en ondulaciones en la representación tiempo-frecuencia. Al igual que para el espectrograma reasignado, la P_{Burg} muestra un incremento en el espectro alrededor de los 35 ms y un decaimiento conforme transcurre el tiempo.

El resto de las representaciones obtenidas mediante el modelado TVAR no generan información para instantes de tiempo por debajo de los 10 ms, mientras que la P_{Kalman} no se adapta al comportamiento de la señal en su parte final. Sin embargo, las tres representaciones muestran un incremento en el espectro de la señal alrededor de los 35 ms, así como el posterior decaimiento del contenido en frecuencia.

El escalograma de S2 simulado nuevamente resulta muy similar al espectrograma, observándose el efecto de emplear ventanas muy largas a bajas frecuencias mediante una dispersión del término cruzado localizado por debajo de los auto-términos.

La descomposición empírica de modos de la señal produce tres modos intrínsecos de oscilación y un residuo de baja magnitud. El primero de estos modos de oscilación presenta una alta correlación con la señal original. La estimación de la amplitud instantánea del primer IMF indica que el componente A2 contribuye con mayor energía que el componente P2 a la señal de S2 simulada. Por su parte, la frecuencia instantánea estimada para el primer IMF muestra un decaimiento en la parte correspondiente al componente A2, un incremento en alrededor de los 35 ms, y un posterior decaimiento en la parte correspondiente al componente P2. Así, el HHS de S2 simulado resulta visualmente muy similar a la representación tiempo-frecuencia teórica de la señal, manteniéndose los problemas debido a la interpolación al inicio y final de la señal.

De acuerdo a los índices de desempeño mostrados en la tabla VI.IX, diferentes representaciones tiempo-frecuencia son adecuadas para el análisis del segundo sonido cardiaco simulado. La correlación local señala al espectro de Hilbert-Huang y al espectrograma reasignado; la correlación global a P_{Burg} y el espectrograma reasignado; el error entre las frecuencias instantáneas a P_{Burg} y el espectrograma, la correlación entre las frecuencias instantáneas al escalograma y P_{Burg} ; y la resolución tiempo-frecuencia al espectro de Hilbert-Huang, el espectrograma reasignado y P_{Burg} . Diversas representaciones tiempo-frecuencia presentan dificultades para estimar adecuadamente el espectro de S2 simulado al inicio y al final, y brindan estimados adecuados para el resto de la señal. Por esta razón, se obtuvieron nuevamente los índices de desempeño basados en la estimación del centroide omitiendo valores extremos de la señal. De acuerdo a estos índices, mostrados en la tabla

VI.X, el espectro de Hilbert-Huang y P_{Burg} resultan las dos representaciones con el mejor desempeño en la estimación de la frecuencia instantánea. El espectro de Hilbert-Huang ofrece la ventaja de ser una técnica de análisis que depende únicamente de la señal y de brindar una adecuada estimación de la frecuencia instantánea de la señal. Por lo tanto, el espectro de Hilbert-Huang fue seleccionada como la técnica más adecuada para el análisis tiempo-frecuencia de las señales de S2 reales. Además, considerando la similitud morfológica de S2 con los sonidos respiratorios adventicios discontinuos (estertores crepitantes), el espectro de Hilbert-Huang se seleccionó como la técnica más adecuada para su análisis.

VII.5.4. Señales de S2 reales adquiridas.

La técnica de análisis tiempo-frecuencia seleccionada para la señal de S2, el espectro de Hilbert-Huang, fue aplicada a sonidos de S2 reales adquiridos de sujetos sanos sobre los focos pulmonares y aórticos. En general, resulta complicado determinar la presencia de los componentes aórtico (A2) y pulmonar (P2) del segundo sonido cardiaco, dado su traslape temporal.

La descomposición empírica de modos produce en promedio cinco modos de oscilación intrínsecos para las señales de S2. De estos modos de oscilación, se seleccionaron dos o tres de ellos para producir el espectro de Hilbert-Huang. En general, se observó que la parte inicial de las señales de S2 reales está dominada por componentes de alta frecuencia (primeros IMFs) y que decrecen en amplitud rápidamente, mientras que los componentes de menor frecuencia (IMFs posteriores) tienen una mayor extensión. Este comportamiento resulta evidente en la estimación de las amplitudes instantáneas. Los espectros de Hilbert-Huang contruidos para las señales reales a partir de las amplitudes y frecuencias instantáneas, muestran que el contenido en frecuencia de los modos de oscilación decrece con el tiempo hasta mantenerse alrededor de un valor dado con un comportamiento de tipo exponencial.

Para el caso de una señal de S2 sin filtrar y aplicando Hilbert-Huang se obtuvo que la descomposición empírica de modos de la señal ubica el ruido de alta frecuencia de la señal en los primeros modos de oscilación, mientras que la información de segundo sonido cardiaco se manifiesta en modos de oscilación más bajos, resultando nuevamente en componentes con características similares a los obtenidos con la señal S2 filtrada. Así, el espectro de Hilbert-Huang nos ofrece la ventaja de eliminar un filtro para suprimir el ruido de alta frecuencia de la señal.

VII.6. Estertor sibilante simulado.

El estertor sibilante simulado consiste en un tono puro modulado en intensidad por una ventana Hamming. Así, su función de ambigüedad (anexo I) es una línea recta que corre a lo largo del eje de retardo con una intensidad que decrece con la distancia al origen.

De acuerdo la estructura de los términos cruzados de la sibilancia simulada, su distribución Wigner-Ville presenta una alta resolución tiempo-frecuencia, en comparación con la obtenida con el espectrograma debido al efecto pasabajas del empleo de una ventana de análisis. Sin embargo, aunque la ventana de tiempo largo disminuye la resolución espectral, en este caso representa una mejoría en la resolución temporal en comparación con la WVD de la señal. Por su parte, la distribución Choi-Williams de la señal presenta una resolución tiempo-frecuencia mayor a la mostrada por el espectrograma.

El espectrograma reasignado seleccionado aumenta notablemente la concentración del auto-término alrededor de su frecuencia instantánea.

La representación P_{Burg} presenta una resolución espectral similar a la del espectrograma, aunque en este caso el empleo de la función ventana ocasiona que la intensidad de la señal sea intermitente y que exista fuga en el espectro. Para esta representación la intensidad de la señal se extiende visiblemente a lo largo de toda la duración de la señal, quizá nuevamente debido al empleo de la ventana de análisis.

Aunque la representación P_{RLS} estima adecuadamente el espectro de la señal durante su parte intermedia, los valores de intensidad estimados en sus extremos resultan demasiado bajos y proporciona una lectura errónea de la extensión de la señal.

Por su parte, las representaciones $P_{\text{RLS-VFF}}$ y P_{Kalman} resultan bastante similares, ofreciendo una baja resolución tiempo-frecuencia de la sibilancia simulada. En particular la resolución espectral es bastante baja y ambas representaciones presentan una sobrestimación de la intensidad de la señal en su parte inicial y final.

El escalograma presenta una gran resolución tiempo-frecuencia de la sibilancia simulada. En este caso, el empleo de una ventana de tiempo larga a bajas frecuencias presenta el beneficio de extender la intensidad de la señal hacia otros instantes de tiempo, aunada a una elevada resolución espectral.

Por su parte, la descomposición empírica de modos de la sibilancia simulada únicamente arroja un modo de oscilación intrínseco y ninguna señal de residuo, lo cual indica que la señal ya presenta características adecuadas para obtener su señal analítica. La estimación de la amplitud instantánea de la señal recupera la función Hamming de

modulación en amplitud, o envolvente de la señal. La frecuencia instantánea estimada está centrada en la frecuencia del tono simulado, aunque nuevamente presenta dificultades en los extremos de la señal. Así, el espectro de Hilbert-Huang resultante para la sibilancia simulada luce casi idéntico a la representación tiempo-frecuencia teórica.

A partir de los índices de desempeño mostrados en la tabla VI.XI, el espectro de Hilbert-Huang resulta la técnica de análisis más adecuada para el análisis de la sibilancia simulada conforme a los criterios de correlación local y global. Además de que esta técnica de análisis produce la representación con la mayor resolución tiempo-frecuencia al mostrar una concentración de energía similar a la teórica. En consecuencia, el espectro de Hilbert-Huang fue seleccionado como la técnica más adecuada para el análisis tiempo-frecuencia de estertores sibilantes, debido a que produce una distribución de la energía de la señal bastante cercana a deseada sin la necesidad depender de parámetros tales como ventanas, ordenes del modelo, o varianzas del ruido.

VII.6.1. Estertores sibilantes reales adquiridos.

De los cinco estertores sibilantes reales adquiridos de pacientes con NID, las primeras cuatro señales corresponden a sibilancias aisladas, mientras que la restante corresponde a una señal de sonido respiratorio adquirida a lo largo de la fase espiratoria e incluye tanto un estertor sibilante como un estertor crepitante. A continuación se discuten los resultados obtenidos con el espectrograma y el espectro de Hilbert-Huang.

Para el caso de las sibilancias aisladas, el método de la descomposición empírica de modos produce diferente número de modos de oscilación intrínseco. Sin embargo, el primer modo de oscilación, o IMF, representa casi en su totalidad a la sibilancia. Así, el primer IMF de cada señal fue seleccionado para producir el espectro de Hilbert-Huang. Como en el caso de la señal simulada, la amplitud instantánea estimada recupera la función de modulación en amplitud, o envolvente, de la sibilancia, aunque a diferencia de la simulación la amplitud instantánea presenta oscilaciones. Por su parte, la estimación de la frecuencia instantánea, al igual que para las señales simuladas, presenta transiciones abruptas para instantes de la señal en los cuales la amplitud instantánea es baja. A diferencia de lo encontrado en las simulaciones, la frecuencia instantánea de la sibilancia real presenta una modulación en la frecuencia alrededor de una frecuencia central. La variación en el comportamiento del espectro de la señal resulta visible tanto en el espectro de Hilbert-Huang de la señal, como en su versión suavizada debido a la alta resolución tiempo-frecuencia producida por esta técnica. Por su parte, los espectrogramas se obtuvieron con una ventana Hamming 60 ms de longitud

de acuerdo a lo indicado por las simulaciones. En cada caso, el espectrograma de la señal estima el tono central de la sibilancia con una adecuada resolución temporal aunque con una baja resolución en la frecuencia. Quizá este suavizado a lo largo del eje de frecuencias no permita observar a mayor detalle el comportamiento instantáneo del espectro de las sibilancias.

Con respecto a la señal durante la fase espiratoria, se observa la presencia de una crepitancia de baja intensidad al inicio de la fase y de un sibilancia en la parte intermedia. Su descomposición empírica de modos produce diversos modos de oscilación intrínsecos, aunque la información de la sibilancia queda contenida en su mayoría en el primero de ellos y la información de la crepitancia en los dos primeros modos de oscilación, como lo reflejan sus estimados de amplitud instantánea. En particular, para el estertor crepitante el primer modo de oscilación es de menor duración que el segundo. Con respecto a las frecuencias instantáneas estimadas, se observa que con excepción de los instantes de tiempo donde ocurren ambos estertores el comportamiento resulta desordenado. Para el estertor crepitante, se observa un decremento en la frecuencia instantánea en ambos modos de oscilación y para el estertor sibilante se observa una frecuencia con oscilaciones alrededor de un valor medio. El espectro de Hilbert-Huang indica tanto la presencia de los estertores como el comportamiento de sus espectros e incluso para la crepitancia a pesar de su carácter transiente.

Para el caso de las señales de sonidos respiratorios, la ventana empleada en la bibliografía pionera del respirosónograma tiene una extensión de 100 ms. Esta misma longitud de ventana se empleó para producir el espectrograma de la señal de sonidos respiratorios adventicios. A diferencia del espectro de Hilbert-Huang, el espectrograma resultante únicamente muestra la presencia de la sibilancia pero no la presencia de la crepitancia debido a la gran extensión de la ventana de análisis que degrada la resolución temporal de la representación. Nuevamente, la resolución tiempo-frecuencia del espectrograma resulta mucho menor a la correspondiente del espectro de Hilbert-Huang.

VII.7. Estertores crepitantes.

Para los estertores crepitantes simulados, su espectro de Hilbert-Huang fue considerado como la representación tiempo-frecuencia de referencia, debido a que a partir de los resultados obtenidos con las señales simuladas anteriormente, esta técnica de análisis resulta en una adecuada estimación de la frecuencia instantánea. En particular, los estertores crepitantes simulados comparten con la señal de S2 simulada la característica de que su frecuencia instantánea decrece con el tiempo. Tanto para la crepitancia fina simulada como

para la gruesa, la descomposición empírica de modos produjo un único modo intrínseco de oscilación y una señal de residuo.

Uno de los objetivos de esta tesis consistió en comparar representaciones tiempo-frecuencia desarrolladas recientemente con la herramienta clásica en el campo de los sonidos respiratorios, el espectrograma. En este caso, el índice de desempeño empleado para seleccionar los parámetros del espectrograma (tipo y longitud de la ventana) fue el error cuadrático medio, NRMSE, entre las frecuencias instantáneas estimadas mediante el centroide. La elección de este índice se debe a que por el momento el interés se centra más en el comportamiento del contenido en frecuencia que en los valores de intensidad, con miras a un posible empleo de estas representaciones tiempo-frecuencia con propósitos de reconocimiento y clasificación de patrones.

Finalmente, además de seleccionar los parámetros del espectrograma para cada tipo de crepitancia simulada, también se aplicó el mismo esquema para las nueve técnicas de análisis tiempo-frecuencia restantes. A continuación se discuten los resultados encontrados.

VII.7.1. Estertor crepitante fino simulado.

Para el caso del espectrograma, para los tipos de ventana empleados, se tiene que el error en la estimación de la frecuencia instantánea de la crepitancia fina simulada disminuye al aumentar la longitud de la ventana de análisis. Observe que las longitudes de las ventanas probadas en esta tesis van de un décimo hasta un medio de la duración de la crepitancia simulada.

La distribución de Wigner-Ville presenta una mayor resolución tiempo-frecuencia que la obtenida con el espectrograma, aunque nuevamente su mayor dificultad es la presencia de términos cruzados que termina por degradar su legibilidad.

La distribución Choi-Williams de la señal presenta una pequeña mejoría en la resolución tiempo-frecuencia comparada con el espectrograma, aunque al parecer su kernel exponencial pasabajas presenta dificultades para suprimir los términos cruzados.

La representación tiempo-frecuencia lograda con el método de la reasignación resulta en una enorme mejoría en la resolución tiempo-frecuencia del espectrograma y también de las demás representaciones de la clase de Cohen anteriores. Así, el espectrograma reasignado nos indica un decremento en el espectro de la crepitancia a lo largo del tiempo. Sin embargo, no proporciona información de lo acontecido al inicio y al final de la señal, aunque tampoco lo hacen las representaciones anteriores.

Con respecto a las representaciones tiempo-frecuencia obtenidas mediante el modelado TVAR y la estimación paramétrica del espectro de potencia, en todos los casos un modelo AR de orden 2 resultó apropiado. La P_{Burg} de la crepitancia fina simulada, al igual que las representaciones anteriores sólo brinda una representación adecuada para la parte inicial de la señal, con una mejora en la resolución tiempo-frecuencia en comparación con el espectrograma, aunque no así de lo acontecido en su parte final. Las representaciones P_{RLS} y $P_{\text{RLS-VFF}}$ resultan bastante similares con excepción de los valores de intensidad, en especial en la parte final de la señal. Por su parte, la representación P_{Kalman} muestra una alta resolución, sin embargo durante el primer tercio de la señal la representación resulta ilegible.

El escalograma de la crepitancia fina simulada se extiende a lo largo de toda la señal, sin embargo es probable que se deba al empleo de ventanas de tiempo largas a bajas frecuencias más que una adecuada estimación espectral. La resolución del escalograma resulta muy similar a la del espectrograma, aunque en este caso se observa la división no uniforme del plano tiempo-frecuencia por la transformada *wavelet*.

Finalmente, aunque gran parte de las representaciones tiempo-frecuencia probadas anteriormente nos indican, en mejor o peor medida, un comportamiento decreciente en el espectro de la crepitancia fina simulada, cada una de ellas presenta diversas dificultades para proporcionar estimaciones adecuadas a lo largo de toda la extensión de la señal. A pesar de que el espectro de Hilbert-Huang al parecer brinda tales estimados, hay que considerar las dificultades encontradas con el resto de las señales simuladas para estimar la frecuencia instantánea al inicio y al final de la señal.

La alta resolución tiempo-frecuencia del espectro de Hilbert-Huang de la señal nos señala, al margen de lo acontecido en los extremos, que el contenido en frecuencias de la crepitancia fina simulada decrece desde un valor cercano a los 820 Hz hasta un valor de 140 Hz. A pesar de que el decaimiento es visualmente notorio, resulta complicado hablar de un comportamiento puramente lineal o exponencial o de otro tipo.

VII.7.2. Estertor crepitante grueso simulado.

Debido a que la morfología de los estertores crepitantes finos y gruesos simulados es la misma, diferenciándose en su duración y tiempo del primer cruce por cero, las representaciones tiempo-frecuencia obtenidas para la crepitancia gruesa simulada mantienen las mismas características discutidas para la crepitancia fina simulada. El contenido en frecuencia de la crepitancia gruesa simulada resulta menor al contenido en frecuencia de la crepitancia fina, dado que la señal es una versión comprimida de esta última.

Nuevamente, el espectro de Hilbert-Huang de la señal presenta la más alta resolución tiempo-frecuencia de entre las representaciones tiempo-frecuencia probadas. El espectro de la crepitancia gruesa decrece desde un valor cercano a los 200 Hz hasta un valor cercano a los 80 Hz. Como en el caso de la crepitancia fina simulada, no se observa un comportamiento bien definido en el descenso del espectro. Aunque, en comparación con la crepitancia fina, la crepitancia gruesa presenta un descenso más lento en su contenido en frecuencia.

VII.7.3 Estertores crepitantes reales adquiridos.

Para los tres primeros casos, la señal contiene únicamente una crepitancia real, adquirida de un paciente con NID, con un sonido respiratorio de baja intensidad. Los espectrogramas de los estertores crepitantes reales se calcularon con una ventana Hamming de 6 ms de longitud de acuerdo a los resultados obtenidos para las crepitancias simuladas. Los espectrogramas de estas crepitancias consisten de un lóbulo de energía con una baja resolución temporal y espectral. Así, el espectrograma señala una gran concentración de energía al inicio de la crepitancia y no brinda información de lo ocurrido con su espectro durante su parte final. La baja resolución espectral indica que el contenido en frecuencia de la señal se extiende desde frecuencias cercanas a 0 Hz hasta los 500 Hz, y al parecer el contenido en frecuencias desciende conforme el tiempo evoluciona.

Por su parte, la descomposición empírica de modos produce diferente número de modos de oscilación intrínsecos para las primeras tres señales reales. Además, de esta diferencia en el número de componentes, se tiene que los IMFs que contribuyen en mayor medida a la producción de la crepitancia y que están más correlacionados con ésta, aparecen en los primeros modos de oscilación intrínsecos, con excepción de la tercera señal donde el primer IMF al parecer consiste de ruido de alta frecuencia asociado al proceso de adquisición. Así, la información de las crepitancias reales parece propagarse más allá de una sola escala de tiempo. En consecuencia, el espectro de Hilbert-Huang de estas crepitancias reales consiste de dos monocomponentes, donde cada uno exhibe un descenso en su espectro. Para estas señales se tiene que un componente se extiende a lo largo de toda la duración de la señal, mientras que el otro es más corto y presenta una caída más rápida. En general, los espectros de Hilbert-Huang de las crepitancias aisladas muestran que el espectro decrece rápidamente en la parte inicial, desde aproximadamente 500 Hz hasta un valor cercano a los 80 Hz para posteriormente, incrementarse ligeramente hasta cerca de los 220 Hz.

A diferencia de las tres primeras señales, la cuarta señal consiste de una señal de sonido respiratorio adquirida a lo largo de una fase inspiratoria de un paciente con NID. Esta

última señal real presenta dos crepitancias reales durante la segunda mitad de la fase inspiratoria. La amplitud de estas crepitancias es notoriamente mayor que el sonido respiratorio de fondo, facilitado su ubicación visual. Al descomponer la señal de sonido respiratorio en sus modos de oscilación intrínseco, se producen once componentes. Al igual que en el caso de las crepitancias reales aisladas, se tiene que la información de las dos crepitancias contenidas en la señal se propaga hacia diversos IMFs, donde la contribución de energía y el contenido en frecuencia va decreciendo. Los estimados en la amplitud instantánea indican claramente el tiempo de ocurrencia de cada estertor crepitante. Por su parte, la estimación de las frecuencias instantáneas muestra que en particular para los tres primeros modos de oscilación el espectro de la señal presenta un comportamiento de ruido de color y muestra un descenso sostenido en los instantes de tiempo donde ocurren las crepitancias. Así, con estos estimados en las amplitudes y frecuencias instantáneas, el espectro de Hilbert-Huang de la señal proporciona información del descenso en el contenido espectral de cada estertor crepitante, con una gran resolución temporal y espectral. En cada caso, el espectro desciende desde aproximadamente los 600 Hz hasta los 100 Hz.

Para esta última señal, se obtuvieron dos espectrogramas con una ventana Hamming, el primero con una ventana de 100 ms de longitud de acuerdo a lo indicado en la bibliografía y el segundo con una ventana de 5 ms de longitud, de acuerdo a lo indicado por las simulaciones de esta tesis. Para el primer espectrograma, espectrograma de banda angosta, se observan dos lóbulos de información, aunque la extensión de la ventana de análisis resulta demasiado larga y produce una baja resolución temporal así como una notable presencia de términos de interferencia entre ambas crepitancias reales. Para el segundo espectrograma, espectrograma de banda ancha, la resolución temporal aumenta notoriamente y con ella la disminución de los términos cruzados, sin embargo la resolución espectral se ve deteriorada y extiende el contenido en frecuencias de la señal desde los 0 Hz hasta frecuencias cercanas a los 900 Hz.

VII.8. Sonidos respiratorios normales adquiridos.

De forma similar a los trabajos pioneros del análisis tiempo-frecuencia de los sonidos respiratorios, en la sección de resultados se presentan los espectrogramas de dos señales representativas de sonidos respiratorios normales, uno correspondiente a un sonido pulmonar y el otro correspondiente a un sonido traqueal. Además, dados los resultados obtenidos con el espectro de Hilbert-Huang en las señales simuladas y reales anteriores, se presenta su

representación tiempo-frecuencia con esta técnica. A continuación se discuten los resultados obtenidos.

VII.8.1 Sonido pulmonar.

De la forma de onda de la señal (figura VI.171) se observa que en promedio la intensidad del sonido pulmonar es mayor durante la fase inspiratoria que durante la espiratoria, y que además la intensidad instantánea de la señal varía de acuerdo al flujo, como se ha señalado en la bibliografía.

La descomposición empírica de la señal (figura VI.172) produce once modos de oscilación intrínsecos y una señal de residuo. Los primeros seis modos de oscilación contribuyen en mayor medida a la composición de la señal (figura VI.173), sin embargo la intensidad instantánea del primer modo de oscilación no muestra una relación con la señal de flujo, como los cinco modos posteriores y quizá represente ruido de alta frecuencia debido al muestreo durante el proceso de adquisición. Por lo tanto, el IMF1 fue descartado para la conformación del espectro de Hilbert-Huang de la señal. Las señales analíticas de los modos de oscilación seleccionados (figura VI.174) muestran que la intensidad instantánea resulta mayor para la fase inspiratoria que para la espiración para los tres primeros modos de oscilación intrínsecos (frecuencias más altas) de la señal. Realizado el mapeo de las frecuencias y amplitudes instantáneas para producir el espectro de Hilbert-Huang (figura VI.175), se observa el espectro de la señal no sigue un comportamiento bien definido y se asemeja al de un ruido de color. La suavización del espectro de Hilbert-Huang (figura VI.176) muestra de forma clara que, además de la intensidad, el ancho de banda de la fase inspiratoria es mayor al de la fase espiratoria.

Por su parte, el espectrograma de la señal (figura VI.177), obtenido con una ventana de tiempo de la misma duración que la empleada en la bibliografía para señales respiratorias, también muestra que la intensidad y contenido en frecuencia de la inspiración es mayor al de la espiración. Sin embargo, la baja resolución del espectrograma, en particular su resolución temporal, no indica el comportamiento del espectro a lo largo de toda la duración de la señal.

VII.8.2 Sonido traqueal.

Para el caso del sonido traqueal normal, la forma de onda (figura VI.178) muestra que ambas fases, inspiración y espiración, poseen valores de intensidad instantánea similares, y que además existe un silencio en la transición de la inspiración a la espiración. Estas

características concuerdan con lo indicado en la bibliografía para un sonido respiratorio adquirido sobre la tráquea.

Al descomponer la señal en sus modos de oscilación intrínsecos se obtienen nuevamente once componentes (figura VI.179), de los cuales los primeros cuatro representan la mayor energía del sonido traqueal (figura VI.180) y muestran una correlación con la señal de flujo. La estimación de la amplitud instantánea de los primeros cuatro modos de oscilación muestra que la intensidad promedio de ambas fases es similar, además de que durante el periodo de transición de la inspiración a la espiración, la amplitud instantánea de estos modos presenta valores muy bajos. Por otra parte, en los instantes de tiempo en los cuales la intensidad instantánea es baja, las frecuencias instantáneas estimadas muestran valores muy elevados, quizá relacionados con dificultades al emplear la derivada. El espectro de Hilbert-Huang del sonido traqueal normal (figura VI.182) muestra que tanto para la fase inspiratoria como para la espiratoria, el contenido en frecuencia de la señal corresponde a un ruido de color, donde ambos límites superiores resultan bastante similares. Además, se observa el periodo de silencio entre ambas fases. Estas características resultan más notorias en la versión suavizada del espectro de Hilbert-Huang (figura VI.183).

Finalmente, el espectrograma del sonido traqueal (VI.184) muestra contenidos en frecuencia similares para ambas fases respiratorias y el silencio entre ellas, aunque de nuevo con una baja resolución tiempo-frecuencia. Como se ha mostrado a lo largo de esta tesis, el empleo de una función ventana repercute en la información ofrecida por el espectrograma resultante. Así, el empleo de una ventana con una longitud mayor quizá mostraría el comportamiento del espectro para toda la señal, aunque la información espectral se vería degradada.

VIII. CONCLUSIONES

A lo largo de esta tesis, con la ayuda de diversas representaciones tiempo-frecuencia se analizaron señales que presentan distintos contenidos en frecuencia variantes en el tiempo. La complejidad y desconocimiento del comportamiento tiempo-frecuencia de las señales analizadas fue gradualmente en aumento. Este hecho permitió la corroboración de las características y propiedades de las técnicas de análisis tiempo-frecuencia empleadas, antes de su aplicación a las señales biológicas de nuestro interés principal, los sonidos respiratorios.

El espectrograma, considerada como la herramienta de análisis tiempo-frecuencia clásica y de referencia, proporcionó resultados adecuados en la mayoría de las aplicaciones debido al comportamiento de su función kernel en el dominio de la función de ambigüedad. Sin embargo, otras técnicas se desempeñaron mejor para esa misma señal. En particular la reasignación del espectrograma, la representación tiempo-frecuencia basada en el modelado autorregresivo variante en el tiempo vía el algoritmo de Burg, y el espectro de Hilbert vía la descomposición empírica de modos, resultaron en mejores herramientas tiempo-frecuencia alternativas de acuerdo a los índices de desempeño empleados. Aunque la búsqueda de una técnica tiempo-frecuencia que resulte de uso universal para todo tipo de señales continua, el espectro de Hilbert-Huang se destacó de las dos representaciones restantes debido a su dependencia única y en la señal a procesar, i.e. por su adaptabilidad. Además, en ocasiones permitió la omisión de una etapa de pre-acondicionamiento por un filtro pasabajas debido a la naturaleza de su descomposición. Por otra parte, la obtención del espectro de Hilbert-Huang como se empleó en esta tesis también presenta limitaciones debidas a la interpolación durante la conformación de los modos o a la derivada durante la estimación de la frecuencia instantánea.

El espectro de Hilbert-Huang resultó en una herramienta tiempo-frecuencia más adecuada para el análisis de los sonidos respiratorios adventicios y normales, así como para los dos primeros sonidos cardiacos, debido a su elevada concentración tiempo-frecuencia en comparación con el espectrograma clásico. Para el caso de los estertores crepitantes permitió observar el descenso en su contenido en frecuencia con el tiempo, comportamiento tiempo-frecuencia que pudiera emplearse para su clasificación automática por algoritmos de reconocimiento de patrones. Para el caso de los estertores sibilantes indicó un comportamiento tiempo-frecuencia que podría estar relacionado con su mecanismo de producción. Y finalmente, para el caso de los sonidos pulmonares y traqueales se corroboró la información reportada previamente en la literatura.

ANEXOS.

Los siguientes anexos se incluyen en un CD-ROM que acompaña a esta tesis.

ANEXO I. RESULTADOS COMPLEMENTARIOS.

En este capítulo se incluyen resultados que complementan el contenido del capítulo VI de esta tesis. Estos resultados complementarios fueron extraídos del cuerpo principal de la tesis con el propósito de agilizar la lectura la misma.

ANEXO II. PARTICIPACIONES EN CONGRESOS.

Parte del trabajo de esta tesis fue presentado en los siguientes marcos internacionales:

“Time-Frequency Representations for Second Heart Sound Analysis.”

30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society.

Vancouver, Canadá, Agosto de 2008.

“Analysis of Discontinuous Adventitious Lung Sounds by Hilbert-Huang Spectrum.”

30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society.

Vancouver, Canadá, Agosto de 2008.