



UNIVERSIDAD NACIONAL DEL LITORAL
Facultad de Ingeniería y Ciencias Hídricas
Instituto de Desarrollo Tecnológico para la Industria Química

**TÉCNICAS BASADAS EN
MEDIDAS DE COMPLEJIDAD PARA EL
ANÁLISIS DE SEÑALES BIOMÉDICAS**

Juan Felipe Restrepo Rincoar

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

2016

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



ACTA DE EVALUACIÓN DE TESIS DE DOCTORADO

En la sede de la Facultad de Ingeniería y Ciencias Hídricas de la Universidad Nacional del Litoral, a los veintidós días del mes de septiembre del año dos mil dieciséis, se reúnen los miembros del Jurado designado para la evaluación de la Tesis de Doctorado en Ingeniería titulada **“Técnicas Basadas en Medidas de Complejidad para el Análisis de Señales Biomédicas”**, desarrollada por el Ing. Juan Felipe RESTREPO RINCKOAR, DNI N° 19.035.724. Ellos son: Dr. Jorge Luis Moiola, Dr. Franco Martín Pessana y Dra. Luciana De Micco.

Luego de escuchar la Defensa Pública y de evaluar la Tesis, el Jurado resuelve:

En primer lugar la versión escrita de esta tesis es muy clara y con abundantes y muy pertinentes referencias bibliográficas. El tratamiento matemático es muy profundo con aportes originales y las aplicaciones validan los desarrollados en el plano teórico. El tesisista presentó en forma ordenada y con una exposición clara los aspectos más salientes de sus contribuciones. Por lo anteriormente expuesto, este Jurado resuelve en su aprobación con la máxima calificación (10, sobresaliente)

Sin más, se da por finalizado el Acto Académico con la firma de los miembros del Jurado al pie de la presente. -----

Dr. Jorge Luis Moiola

Dr. Franco Martín Pessana

Dr. JOSÉ LUIS MACOR
SECRETARIO DE POSGRADO
Facultad de Ingeniería y Cs. Hídricas

Dra. Luciana De Micco (*)

() La Drz. De Micco participó de la defensa de Tesis por video conferencia.*

Dr. JOSÉ LUIS MACOR
SECRETARIO DE POSGRADO
Facultad de Ingeniería y Cs. Hídricas

Universidad Nacional del Litoral
Facultad de Ingeniería y
Ciencias Hídricas

Secretaría de Posgrado

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



UNIVERSIDAD NACIONAL DEL LITORAL
Facultad de Ingeniería y Ciencias Hídricas

Santa Fe, 22 de septiembre de 2016.

Como miembros del Jurado Evaluador de la Tesis de Doctorado en Ingeniería titulada **“Técnicas Basadas en Medidas de Complejidad para el Análisis de Señales Biomédicas”**, desarrollada por el Ing. Juan Felipe RESTREPO RINCKOAR, en el marco de la Mención “Inteligencia Computacional, Señales y Sistemas”, certificamos que hemos evaluado la Tesis y recomendamos que sea aceptada como parte de los requisitos para la obtención del título de Doctor en Ingeniería.

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

Dr. Jorge Luis Moiola

Dr. Franco Martín Pessana

Dra. Luciana De Micco (se)

La Dra De Micco participó de la defensa de Tesis por videoconferencia

Dr. JOSÉ LUIS MACOR
SECRETARIO DE POSGRADO
Facultad de Ingeniería y Cs. Hídricas

Santa Fe, 22 de septiembre de 2016

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

Dr. Gastón Schlotthauer
Codirector de Tesis

Dra. María Eugenia Torres
Directora de Tesis

Universidad Nacional del Litoral
Facultad de Ingeniería y
Ciencias Hídricas

Secretaría de Posgrado

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

Resumen

Existe en la literatura evidencia que sugiere que la dinámica de los sistemas fisiológicos es compleja y no lineal. Dicha complejidad provee al organismo de un mecanismo de respuesta ante cambios abruptos del ambiente. En este sentido, el grado de complejidad de un sistema, está estrechamente relacionado con los estados normales y patológicos del organismo. Es importante contar con medios para medir cuantitativamente la complejidad de un sistema fisiológico, que sirva de ayuda tanto para el diagnóstico como para medir la respuesta de un individuo ante algún tratamiento. En las señales biomédicas se refleja de manera directa el estado del sistema fisiológico del cual provienen. Surge entonces la necesidad de desarrollar tecnologías que permitan medir de forma confiable el grado de complejidad de un sistema fisiológico a partir de ellas. Para cubrir esta necesidad, se han desarrollado conceptos teóricos y herramientas algorítmicas, que se agrupan principalmente en la teoría de dinámicas no lineales, la teoría del caos y la teoría de la información.

El objetivo principal de esta tesis doctoral es presentar dos desarrollos diferentes que permiten caracterizar la complejidad de un sistema. El primero está relacionado con la entropía aproximada (ApEn) de una serie temporal y el segundo con la dimensión de correlación (D) y la entropía de correlación (K_2) de un sistema dinámico.

La ApEn es un estadístico capaz de medir la regularidad de una serie temporal. Uno de sus mayores problemas es la alta dependencia de sus parámetros. Por este motivo, algunos autores propusieron la $ApEn_{max}$ como un estimador de regularidad más consistente. No obstante, el nivel de ruido (σ) presente en la serie temporal disminuye la capacidad de discriminación tanto de la ApEn como de la $ApEn_{max}$. En respuesta a esto propusimos el estimador h_{max} , que se define como el valor de escala h para el cual ApEn alcanza su valor máximo. Utilizando señales provenientes de sistemas de alta y baja dimensión, hemos estudiado el comportamiento de $ApEn_{max}$ y h_{max} ante diferentes niveles de ruido y longitudes de datos. Como conclusión, h_{max} provee valiosa información para diferenciar entre dinámicas de diferente regularidad, especialmente en casos con alto nivel de ruido. Además, para diferenciar entre dinámicas de distinta regularidad, es mejor utilizar conjuntamente

$ApEn_{max}$ y h_{max} que cada uno por separado. Hemos desarrollado dos aplicaciones: una para la detección de episodios ictales a partir de una señal de electroencefalografía y un detector de voces patológicas a partir de la señal de voz, superando los desempeños reportados en la literatura.

El segundo desarrollo de esta tesis doctoral guarda relación con la estimación de \mathbf{D} y \mathbf{K}_2 . Estos invariantes se calculan a partir de la integral de correlación, estimada mediante el algoritmo de correlación. Esta integral recibe varios nombres dependiendo de la función núcleo que se utilice para su cálculo. Por un lado está la integral de correlación estándar (Grassberger y Procaccia), y por otro, la integral de correlación de núcleo Gaussiano (ICG) propuesta por Diks. El problema de la integral de correlación estándar es que no permite estimar \mathbf{D} y \mathbf{K}_2 en presencia de ruido. La ICG mejora parcialmente este inconveniente. Sin embargo, el cálculo de \mathbf{K}_2 mediante la ICG presenta problemas de convergencia y necesita contar con un gran volumen de datos. Como primer paso en la solución de este inconveniente, proponemos una nueva integral de correlación, la cual denominamos integral de correlación asistida por ruido (ICAR). Para calcularla hemos desarrollado el algoritmo de correlación asistido por ruido. Mediante resultados analíticos y evidencia numérica, demostramos que la integral de correlación estándar y la ICG pueden considerarse casos particulares de la ICAR. Como segundo paso, proponemos la integral de correlación U (ICU). Esta integral de correlación también es un caso particular de la ICAR. A diferencia de la integral de correlación estándar y la ICG, la ICU permite incorporar información de la dimensión de inmersión en la función núcleo. La consecuencia directa de esto es una mejoría notable en la estimación de \mathbf{K}_2 . A partir de la ICU, desarrollamos una serie de estimadores que nos permiten calcular \mathbf{D} , \mathbf{K}_2 y σ a partir de una serie temporal. Estudiamos el comportamiento de estos estimadores ante diferentes condiciones de ruido y longitud de datos. Los resultados muestran que los estimadores propuestos brindan estimaciones confiables de los invariantes \mathbf{D} , \mathbf{K}_2 y σ . Además, diseñamos un algoritmo automático para la estimación de invariantes, que ha sido validado estadísticamente.

Abstract

There exist a lot of evidence suggesting that the dynamics of physiological systems is complex and nonlinear. Thanks to this complexity, the individual has a defense mechanism to respond to environmental changes. In this sense, the degree of complexity of the system is linked to normal or pathological states of the organism. Therefore, it is important to quantitatively measure the complexity of a physiological system, since this information can be useful to diagnose or to follow-up the treatment of a certain pathology. The condition of a physiological system is directly reflected in the biomedical signals that it produces. There have been proposed theoretical concepts and algorithms founded on nonlinear dynamics theory, chaos theory and information theory. However, it is still necessary to develop new methods and technologies able to provide reliable measurements of the degree of complexity of a system from the biomedical signals.

The objective of this thesis is to introduce two different approaches that allow us to characterize the complexity of a system, in particular in presence of noise. The first one is related with the approximate entropy (ApEn) of a temporal series and the second one with the correlation dimension (D) and the correlation entropy (K_2) of a dynamical system.

The ApEn is a statistic that measures the regularity of a signal. One of its major problems is related to the high dependency on the value of its parameters. For this reason some authors have proposed $ApEn_{max}$ as a more consistent regularity measure. However, the noise level (σ) in the temporal series undermine the discrimination capacity of both ApEn and $ApEn_{max}$. As solution we have proposed the estimator h_{max} , that is defined as the value of the scale h at which ApEn achieves its maximum value. Using time series from low- and high-dimensional systems, we have studied the behavior of $ApEn_{max}$ and h_{max} under different noise conditions and data lengths. We can conclude that h_{max} provides valuable information useful for classification purposes. Moreover, the use of h_{max} in conjunction with $ApEn_{max}$ as features, can significantly decrease the misclassification rate of a linear classifier, in comparison with their isolated use. Based on these results, we have developed an ictal episode detector from electroencephalographic signals and a pathological

voice detector that outperform the results reported in the literature.

The second result of this thesis is related with the estimations of D and K_2 . These invariants are estimated through the correlation integral, which is calculated using the correlation algorithm. There exist two types of correlation integrals: on one hand there is the classical correlation integral (Grassberger and Procaccia), and on the other one there is the Gaussian kernel correlation integral (GCI) proposed by Diks. The main problem of the classical correlation integral is that in presence of noise it is unable to estimate D and K_2 . The Gaussian kernel correlation integral (GCI) partially solves this issue, however it presents convergence problems when computing K_2 . As a first step in the solution of this problem we have proposed the noise assisted correlation integral (NCI) and for its estimation, we have created the noise assisted correlation algorithm. Through analytic deductions and numerical simulations we have shown that the classical correlation integral and the GCI are particular cases of the NCI. As a second step, we have developed the U correlation integral (UCI). This correlation integral is also a particular case of the NCI. Unlike the GCI and the standard correlation integral, the UCI is able to incorporate information about the embedding dimension in its kernel function. The direct consequence is a notably improvement in the estimation of K_2 . Based on the UCI we have deduced coarse-grained estimators for D , K_2 and σ . We have studied their behavior under different noise conditions and data lengths. The results show that these estimators lead to reliable estimations of D , K_2 and σ . Moreover, we have designed an automatic estimator of invariants which have been statistically validated.

Key words: nonlinear dynamics, complexity measures, biomedical signals, approximate entropy, correlation integral, correlation dimension, correlation entropy, U correlation integral, noise assisted correlation integral.

Índice general

Título	I
Acta de evaluación	III
Resumen	V
Abstract	VII
Índice de figuras	XIII
Índice de tablas	XVII
Índice de algoritmos	XIX
Lista de abreviaturas	XXI
Lista de Símbolos	XXV
1 Dinámicas no lineales, atractores y caos	1
1.1 Dinámicas no lineales y sistemas	2
1.2 Atractores	5
1.3 Reconstrucción del espacio de estados	6
1.3.1 Retardo de inmersión	9
1.3.2 Dimensión de inmersión	11
1.4 Sistemas fisiológicos y dinámicas no lineales	15
1.4.1 Sistema nervioso central	16
1.4.2 Sistema fonador	21
1.5 Conclusiones	25
2 Medidas de complejidad	27
2.1 Introducción	27
2.2 Dimensiones	30
2.2.1 Dimensión de conteo por cajas	31

Índice general

2.2.2	Dimensión de correlación	32
2.3	Entropía	33
2.3.1	Entropía de Kolgomorov-Sinai	34
2.3.2	Entropía de correlación	35
2.4	Divergencia entre trayectorias: Exponentes de Lyapunov	36
2.5	Integral de Correlación	38
2.5.1	Estimadores dependientes de la escala	42
2.5.2	Algoritmo de correlación	43
2.6	Entropía aproximada	49
2.7	Conclusiones	54
3	Entropía aproximada y cambios de regularidad	55
3.1	Introducción	56
3.2	$ApEn_{max}$, h_{max} y cambios de regularidad	58
3.2.1	Simulaciones con señales artificiales	58
3.2.2	Aplicaciones con señales reales	69
3.3	Conclusiones	79
4	Integral de correlación asistida por ruido	81
4.1	Introducción	81
4.2	Deducción analítica	83
4.2.1	Caso particular 1: Integral de correlación estándar	85
4.2.2	Caso particular 2: Integral de correlación de núcleo Gaussiano	85
4.3	Simulaciones	86
4.4	Conclusiones	90
5	Integral de correlación U	91
5.1	Introducción	91
5.2	Deducción analítica	92
5.3	Caso particular: $ICU^{\beta=2}$	93
5.4	Caso particular: $ICU^{\beta=m}$	94
5.5	Simulaciones	95
5.6	Conclusiones	99
6	Estimadores basados en la integral de correlación U	101
6.1	Introducción	102
6.2	Estimador dependiente de la escala para D	103
6.3	Estimador dependiente de la escala para K_2	107
6.4	Estimador dependiente de la escala para σ	111
6.5	Estimador automático de invariantes	113
6.6	Conclusiones	120

7 Conclusiones finales y trabajos futuros	121
7.1 Trabajos futuros	123
A Apéndice A	125
A.1 Funcional para D basado en $ICU^{\beta=m}$	125
A.2 Comportamiento de $d \ln U_m^{\beta=m}(h) / d \ln h$	127
B Apéndice B	129
B.1 Funcional para K_2 basado en $ICU^{\beta=m}$	129
B.2 Comportamiento de $\ln U_{m+2}^{\beta=m+2}(h) / U_m^{\beta=m}(h)$	130
C Apéndice C	133
C.1 Funcional para σ basado en $ICU^{\beta=m}$	133
Lista de Publicaciones	135
Bibliografía	137

Índice de figuras

1.1 Transformación logística. Soluciones con diferentes condiciones iniciales y atractor	4
1.2 Reconstrucción de un atractor mediante retardo de coordenadas	8
1.3 Reconstrucción del atractor del sistema de Rössler con diferentes retardos de inmersión	10
1.4 Falsos vecinos. Atractor de Hénon.	12
1.5 Anatomía y fisiología del cerebro	17
1.6 Señal de electroencefalograma. Segmentos de EEG normal y patológico (ictal)	20
1.7 Aparato fonador	22
1.8 Señal de voz. Segmentos de voz normal y patológica	23
2.1 Transformación triangular	28
2.2 Conjunto de Cantor	32
2.3 Partición del espacio-tiempo para una solución	35
2.4 Algoritmo de correlación	44
2.5 Gráfica log-log de la ICG para la transformación triangular sin ruido	46
2.6 Estimación de D utilizando la ICG para la transformación triangular sin ruido	47
2.7 Estimación de K_2 utilizando la ICG para la transformación triangular sin ruido	47
2.8 Gráfica log-log de la ICG para la transformación triangular con ruido ($\sigma = 0.05$)	48
2.9 Estimación de σ utilizando la ICG para la transformación triangular con ruido	48
2.10 Entropía aproximada en función de h para la transformación triangular	51
2.11 Entropía aproximada y $ApEn_{max}$ para señales de distinta complejidad	52
3.1 Entropía aproximada en función de h , $ApEn_{max}$ y h_{max} para señales de diferente complejidad bajo la influencia de ruido (RSR= 30 dB)	57

Índice de figuras

3.2	Intervalo de confianza del 95 % para $ApEn_{max}$ y h_{max} . Sistema de Mackey-Glass sin ruido y con ruido (RSR = 5dB)	60
3.3	Intervalo de confianza del 95 % para $ApEn_{max}$ y h_{max} . Sistema de Shilnikov sin ruido y con ruido (RSR = 5dB)	62
3.4	Gráfica de dispersión de $ApEn_{max}$ vs h_{max} para los sistemas de Mackey-Glass y Shilnikov con ruido (RSR = 5dB)	63
3.5	Validación cruzada de 10 particiones para la tasa de error de clasificación en función de m	66
3.6	Validación cruzada de una iteración para la tasa de error de clasificación en función de N	67
3.7	Señal de EEG con episodios pre-ictales e ictales	70
3.8	Comportamiento de $ApEn_{max}$ y h_{max} en función de m para dos episodios pre-ictales y dos ictales	71
3.9	Detección de episodios ictales utilizando $ApEn_{max}$ y h_{max} a partir de una señal de EEG con ruido (RSR = 5 dB)	73
3.10	Comportamiento de $ApEn_{max}$ y h_{max} para voces normales y patológicas	76
3.11	Clasificación de voces sanas y patológicas utilizando $ApEn_{max}$ y h_{max} . Sencibilidad, precisión y especificidad en función de m	77
3.12	Detector automático de voces patológicas. Gráfico de dispersión para voces patológicas y sanas	78
4.1	Algoritmo de correlación asistido por ruido	82
4.2	Gráfica log-log para la ICG y la ICAR con ruido de Rayleigh para la transformación de Hénon sin ruido	87
4.3	Estimación de D utilizando la ICG y la ICAR con ruido de Rayleigh para la transformación de Hénon sin ruido	88
4.4	Estimación de K_2 utilizando la ICG y la ICAR con ruido de Rayleigh para la transformación de Hénon sin ruido	89
4.5	Gráfica log-log para la ICG y la ICAR en presencia de ruido ($\sigma = 0.05$)	90
5.1	Gráfica log-log de la ICG y la $ICU^{\beta=m}$ para la transformación de Hénon sin ruido	96
5.2	Estimación de D utilizando la ICG y la $ICU^{\beta=m}$ para la transformación de Hénon sin ruido	97
5.3	Estimación de K_2 utilizando la ICG y la $ICU^{\beta=m}$ para la transformación de Hénon sin ruido	97
5.4	Gráfica log-log de la ICG y la $ICU^{\beta=m}$ para la transformación de Hénon con ruido ($\sigma = 0.05$)	98
5.5	Estimación de σ utilizando la ICG y la $ICU^{\beta=m}$ para la transformación de Hénon con ruido ($\sigma = 0.05$)	98

6.1	Transformación de Hénon. EDE para la dimensión de correlación $D_m^U(h)$	106
6.2	Transformación de Hénon. EDE para la entropía de correlación $K_m^U(h)$	110
6.3	Transformación de Hénon. EDE para el nivel de ruido $\sigma_m^U(h)$	112
6.4	Transformación de Hénon. Gráfico de cajas de la estimación automática de los invariantes σ , D y K_2 para diferente número de vectores de estado	116
6.5	Sistema de Mackey-Glass. Gráfico de cajas de la estimación automática de los invariantes σ , D y K_2 para diferente número de vectores de estado	119

Índice de tablas

3.1 Modelos y parámetros utilizados en las simulaciones con $ApEn_{max}$ y	
h_{max}	59
3.2 Resultados de Arias-Londoño y col. en la clasificación de voces pato-	
lógicas	75
3.3 Matriz de confusión para el detector automático de voces normales	
y patológicas	79

Índice de algoritmos

1	Algoritmo de conteo por cajas.	32
2	Algoritmo de correlación de Grassberger y Procaccia.	45
3	Algoritmo de correlación asistido por ruido.	86
4	Detector automático de invariantes utilizando los estimadores dependientes de la escala $\sigma_m^U(h)$, $D_m^U(h)$ y $K_m^U(h)$.	115

Lista de abreviaturas

D	dimensión de correlación.
K_2	entropía de correlación.
σ	nivel de ruido.
$ICU^{\beta=2}$	integral de correlación U con $\beta = 2$.
$ICU^{\beta=m}$	integral de correlación U con $\beta = m$.
ApEn	entropía aproximada.
CAD	convertor análogo-digital.
CUSUM	algoritmo de detección de cambios en valores medios.
EDE	estimador dependiente de la escala.
EEG	electroencefalograma.
FDA	función de distribución acumulada.
FDAC	función de distribución acumulada complementaria.
FDP	función de distribución de probabilidad.
GApEn	entropía aproximada de núcleo Gaussiano.
GSampEn	entropía muestral de núcleo Gaussiano.
ICAR	integral de correlación asistida por ruido.
ICG	integral de correlación de núcleo Gaussiano.
ICU	integral de correlación U.
iid	independiente e idénticamente distribuida.
MEL	máximo exponente de Liapunov.
MSV	máquinas de soporte vectorial.

Lista de abreviaturas

RSR	relación señal a ruido.
SampEn	entropía muestral.
TEC	tasa de error de clasificación.
VFC	variabilidad de frecuencia cardíaca.

Lista de Símbolos

$\ \cdot\ ^2$	Cuadrado de la distancia Euclídea.
z_ω	Cuadrado de la distancia Euclídea entre dos vectores de estado.
d_0	Dimensión de conteo por cajas.
m	Dimensión de inmersión.
β	Dimensión del ruido utilizado en el algoritmo de correlación asistido por ruido - Grados de libertad.
$\ \cdot\ $	Distancia Euclídea.
\tilde{z}_ω	Distancia Euclídea entre dos vectores de estado.
$\ \cdot\ _\infty$	Distancia Infinito.
χ_β	Distribución Chi.
χ_β^2	Distribución Chi cuadrada.
$\delta[n]$	Distribución delta de Kroenecker.
$\mathcal{N}(0, \sigma^2)$	Distribución normal de media cero y varianza σ^2 .
$\mathcal{N}(\mathbf{0}, \Sigma^2)$	Distribución normal multivariada de media cero y matriz de covarianza Σ^2 .
$\mathcal{U}(a, b)$	Distribución uniforme con soporte en (a, b) .
$ApEn(m, h, N)$	Entropía aproximada.
$ApEn_{max}$	Entropía aproximada máxima.
h	Escala o umbral.
h_{max}	Escala máxima.
$D_m^T(h)$	Estimador dependiente de la escala para la dimensión de correlación basado en la $T_m(h)$.
$D_m^U(h)$	Estimador dependiente de la escala para la dimensión de correlación basado en la $U_m^{\beta=m}(h)$.

Lista de Símbolos

$K_m^T(h)$	Estimador dependiente de la escala para la entropía de correlación basado en la $T_m(h)$.
$K_m^U(h)$	Estimador dependiente de la escala para la dimensión de correlación basado en la $U_m^{\beta=m}(h)$.
$\sigma_m^T(h)$	Estimador dependiente de la escala para el nivel de ruido basado en la $T_m(h)$.
$\sigma_m^U(h)$	Estimador dependiente de la escala para la dimensión de correlación basado en la $U_m^{\beta=m}(h)$.
$\Delta_m^U(h)$	Funcional para el nivel de ruido basado en $U_m^{\beta=m}(h)$.
f_t	Función de densidad de probabilidad de la variable aleatoria t .
\tilde{F}_t	Función de distribución acumulada complementaria de la variable aleatoria t .
F_t	Función de distribución acumulada de la variable aleatoria t .
δ_{x_i}	Función delta de Dirac $\delta(x - X_i)$.
$H(t)$	Función escalón.
${}_1F_1(a, b; t)$	Función Hipergeométrica Confluente.
${}_2F_1(a, b; c; t)$	Función Hipergeométrica de Gauss.
$\Gamma(a)$	Función Gamma.
$\Gamma(a, t)$	Función Gamma incompleta superior.
$S_m(h)$	Integral de correlación asistida por ruido.
$U_m^\beta(h)$	Integral de correlación U.
$U_m^{\beta=2}(h)$	Integral de correlación U con $\beta = 2$.
$U_m^{\beta=m}(h)$	Integral de correlación U con $\beta = m$.
$T_m(h)$	Integral de correlación de núcleo Gaussiano.
$C_m(h)$	Integral de correlación estándar.
$G_m(h)$	Integral de correlación generalizada.
N	Longitud de la serie temporal.
m_0	Mínima dimensión de inmersión.
Q	Número de comparadores que componen el algoritmo de correlación - Número de distancias entre pares de vectores de estado.

L	Número de vectores de estado.
μ_ω	Realización de ruido utilizada en el algoritmo de correlación asistido por ruido.
τ	Retardo de inmersión.
$\widehat{S}_m(h)$	Suma de correlación asistida por ruido.
$\widehat{C}_m(h)$	Suma de correlación estándar.
$\widehat{G}_m(h)$	Suma de correlación generalizada.
σ_s	Varianza de la serie temporal s .
σ_η	Varianza del ruido η presente en la serie temporal.

El caos es un orden por descifrar.

Saramago

1

Dinámicas no lineales, atractores y caos

El estudio de la dinámica de los sistemas físicos tiene su origen a mediados del siglo XVI, cuando Sir Isaac Newton desarrolló las ecuaciones diferenciales para resolver el problema de calcular la trayectoria de la tierra alrededor del sol, a partir de la ley de atracción gravitacional. Décadas más tarde, algunos físicos y matemáticos buscaron resolver este mismo problema pero considerando también la masa de la luna (problema de los tres cuerpos). Después de muchos intentos, se concluyó que era imposible encontrar una solución analítica. En el siglo XIX el Francés Henri Poincaré introdujo una metodología [6], basada en la geometría, que permite analizar de forma cualitativa el comportamiento a largo plazo de sistemas como éste. Gracias a ello, Poincaré describió la posibilidad de tener sistemas determinísticos cuya dinámica es impredecible, introduciendo de esta forma la primera noción de caos. Este importante descubrimiento quedó relegado en la literatura hasta mediados del siglo XX, cuando la invención de las computadoras permitió experimentar con la dinámica de sistemas no lineales. En 1963 el matemático y meteorólogo Edward Norton Lorenz, en su análisis de un modelo simplificado del movimiento por convección de la atmósfera terrestre, “redescubrió” las dinámicas caóticas [7]. Él observó que la solución a sus ecuaciones nunca convergía a puntos estables o a órbitas periódicas. En cambio, oscilaban de forma no periódica e irregular. También observó que si partía de dos condiciones iniciales muy próximas, las soluciones ob-

tenidas eran muy diferentes. Si analizamos este resultado en términos físicos, se podría traducir en la imposibilidad de ser exacto a la hora de hacer predicciones meteorológicas. A pesar de todo lo anterior, Lorenz demostró que existe estructura dentro del caos, ya que al graficar las trayectorias en tres dimensiones, el resultado fue un conjunto de puntos que asemeja la forma de una mariposa. En la década de los 70s surgieron muchos más desarrollos en esta materia. En 1971 Ruelle y Takens [8] utilizaron el término “atractor extraño” para referirse al tipo de estructura formada por la dinámica de un sistema en régimen caótico. Pocos años después, May [9] encontró un ejemplo de caos en la dinámica de un modelo de población biológica (transformación logística). De esta misma forma se han descubierto sistemas caóticos en campos como la química, la mecánica de fluidos, la electrónica, etc. El continuo avance de la tecnología ha hecho posible un estudio más profundo de sistemas con dinámicas no lineales. Hasta la fecha se conoce que un gran número de fenómenos naturales poseen características caóticas. Esto lleva a pensar que tal vez el caos es la forma más eficiente que tiene el universo para encontrar un orden.

1.1 Dinámicas no lineales y sistemas

Existen dos grandes tipos de sistemas dinámicos: las ecuaciones diferenciales y las aplicaciones iteradas (ecuaciones en recurrencias o los mapas iterados). La diferencia entre éstos radica en que el primero modela sistemas de tiempo continuo, como el sistema de Lorenz. El segundo modela sistemas de tiempo discreto, como la transformación logística (mapa logístico o aplicación logística) [10]. Se dice que un sistema continuo o discreto es no lineal si sus variables y/o sus derivadas se relacionan de forma no lineal, por ejemplo, un producto, una potencia u otro tipo de relaciones como las trigonométricas. La principal diferencia entre los sistemas lineales y los no lineales es que para encontrar la solución de un sistema lineal, este puede ser dividido en partes más simples. Luego, cada subsistema puede ser resuelto por separado y la solución final será la suma de cada solución. Es decir, un sistema lineal goza de la propiedad matemática de aditividad de sus soluciones, y físicamente se puede interpretar como la suma de sus partes. Por otro lado, aunque un sistema no lineal puede ser descompuesto en subsistemas, éstos no trabajan de forma independientes. Siempre que exista algún tipo de relación entre subsistemas (cooperación, competencia, etc) ésta será modelada de forma no lineal. Por ejemplo, el siguiente modelo:

$$x_{n+1} = \mu x_n - \mu x_n^2, \quad (1.1)$$

permite estudiar la evolución de la cantidad x_n de bacterias de una población, cuyo comportamiento dependerá del parámetro $\mu \in \mathbb{R}^+$. Es posible ver que el número de bacterias en la siguiente generación (x_{n+1}) depende de la diferencia entre dos términos, uno lineal y otro no lineal. Para $\mu > 1$, el término lineal describe el crecimiento de la población que sería incontrollable si los recursos fueran ilimitados. Sin embargo, al disponer recursos limitados existe una competencia entre los diferentes individuos, lo cual es modelado mediante el término cuadrático [11].

Las no linealidades en estos sistemas hacen que encontrar una solución analítica muchas veces sea una tarea imposible. Existe una forma alternativa de estudiar fenómenos cuyas expresiones analíticas son difíciles de resolver, esta radica en un análisis geométrico de la dinámica. Para ello necesitamos una representación geométrica de la dinámica en un espacio abstracto. Por ejemplo, tomemos el movimiento de un péndulo cuya ecuación está dada por:

$$\ddot{\theta} + \frac{g}{L} \operatorname{sen} \theta = 0,$$

donde θ es el ángulo del péndulo respecto a la vertical, $\ddot{\theta}$ su segunda derivada temporal, g es la aceleración de la gravedad y L es la longitud del péndulo. Esta ecuación diferencial puede ser transformada en el sistema no lineal:

$$\begin{cases} \dot{x} = y, \\ \dot{y} = -\frac{g}{L} \operatorname{sen} x, \end{cases}$$

donde x representa la posición del péndulo e y su velocidad. Ahora nuestra dinámica está caracterizada por dos magnitudes que evolucionan conjuntamente en el tiempo. Por ser un sistema determinístico, una vez conocida su posición inicial $x(0)$ y su velocidad inicial $y(0)$ es posible calcular la solución del sistema para cualquier tiempo t , en otras palabras, conocemos $x(t)$ y $y(t)$. Ahora imaginemos un espacio bidimensional abstracto en cual los puntos $(x(t), y(t))$, $\forall t$, forman una curva. Esta curva representa en el espacio bidimensional \mathbb{R}^2 , lo que se denomina trayectoria del sistema y el espacio es conocido como espacio de fase o espacio de estados [10]. La cuestión principal es que el comportamiento geométrico de la trayectoria en el espacio de estados (para $t \rightarrow \infty$) nos da mucha información acerca del comportamiento del sistema modelado por la ecuación diferencial. Es decir, nos describe la naturaleza de la dinámica del fenómeno sin necesidad de resolver analíticamente su ecuación¹ [12].

¹Observar que en el caso que nos ocupa, podríamos obtener una solución aproximada explícita, mediante la aproximación lineal del término no lineal $-\frac{g}{L} \sin \theta$.

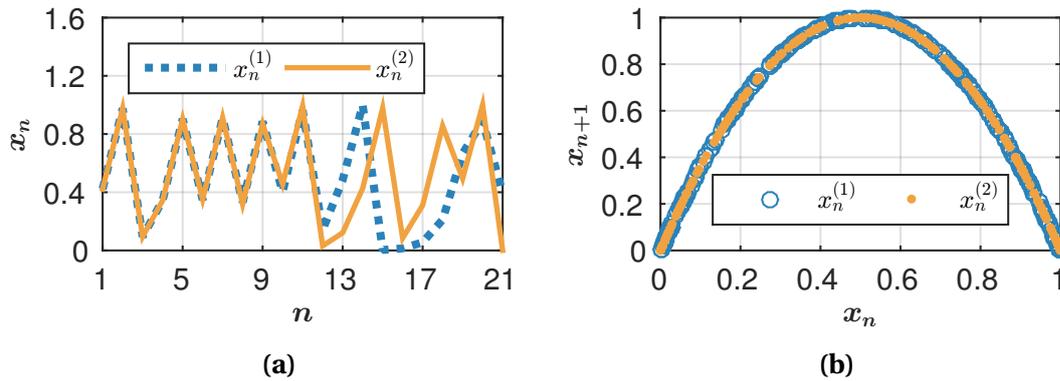


Figura 1.1. Transformación logística con $\mu = 4$. (a) Se muestran dos soluciones $x_n^{(1)}$ y $x_n^{(2)}$ cuya diferencia en la condición inicial es $x_0^{(1)} - x_0^{(2)} = 10^{-4}$. (b) En el espacio de estados bi-dimensional, se aprecia el atractor del sistema. Las trayectorias correspondientes a ambas condiciones iniciales se sumergen en la misma región geométrica.

Otro fenómeno que surge de la no linealidad de los sistemas caóticos es la sensibilidad a las condiciones iniciales. Errores del orden de milésimas en la medición de la condición inicial se magnifican al correr del tiempo. Es decir que, aún conociendo la ecuación de la dinámica, es imposible hacer predicciones acertadas en futuros distantes. Lo anterior puede verse en la figura 1.1a donde se presentan dos soluciones, $x_n^{(1)}$ y $x_n^{(2)}$, de la transformación logística (ecuación 1.1) con $\mu = 4$ (valor para el cual este sistema es caótico). La diferencia entre las condiciones iniciales de ambas soluciones es menor a 10^{-4} . Se observa que ambas soluciones son cercanas para las primeras iteraciones de la transformación, sin embargo, rápidamente sus comportamientos son muy diferentes. Esto nos indica que, dos soluciones pueden ser muy distintas a pesar de que provienen del mismo sistema y que sus condiciones iniciales sean muy próximas. Esto dificulta enormemente su análisis.

Este comportamiento impredecible es conocido como *caos*. Si bien el concepto de caos ya era conocido por el matemático Henri Poincaré en el siglo XIX, este fenómeno no tomó importancia hasta 1975, cuando Tien Y. Li y James A. Yorke [13] acuñaron el término “caos” en sus estudios sobre el transformación logística. El caos se define como una dinámica determinística, aperiódica, acotada y con sensibilidad a las condiciones iniciales [14]. En esta definición cada palabra tiene un significado específico. *Determinístico* significa que existe una regla de evolución definida sin elementos aleatorios. Es decir, existe un único valor para el estado \mathbf{x}_{t+1} determinado por la regla $f(\mathbf{x}_t)$. *Aperiodicidad* se refiere a que el mismo estado no se repite más de una vez. En la práctica, esto no se cumple estrictamente debido a que contamos con precisión finita en nuestras computadoras. Sin embargo, la presencia de largos ciclos de aperiodicidad en simulaciones computacionales es evidencia de posible

caos. *Acotado* se refiere a que en las sucesivas iteraciones de la dinámica, las soluciones permanecen en un rango finito de valores, en otras palabras, las trayectorias se mueven en una región acotada del espacio de fase. Por último, la *sensibilidad a las condiciones iniciales*, significa que dos trayectorias correspondientes a condiciones iniciales muy próximas, se separarán exponencialmente. Lo anterior tiene consecuencias profundas, ya que implica que nunca podremos predecir cual será el estado de un sistema caótico a largo plazo, debido a que es imposible conocer con certeza absoluta el valor de la condición inicial.

Una de las motivaciones para el estudio de las trayectorias es que, para sistemas disipativos, éstas permanecerán acotadas para casi toda condición inicial. Es decir, si tomamos un volumen inicial en el espacio de estados donde cada punto es una condición inicial y dejamos que el sistema evolucione según su dinámica, cada trayectoria, con el tiempo, será atrapada por un subconjunto acotado de puntos en el espacio de estados. Este conjunto es llamado atractor del sistema [12]. Este hecho tiene grandes consecuencias para el análisis de dinámicas no lineales ya que implica que un conjunto de puntos que está disperso por el espacio de estados, será contraído por acción de la dinámica hacia el atractor del sistema. Es decir, que cualquier trayectoria, sin importar su condición inicial, terminará en el atractor del sistema. En la figura 1.1b se muestra el atractor de la transformación logística en el espacio de estados. La reconstrucción fue hecha a partir de dos trayectorias con diferentes condiciones iniciales. Se observa que aunque las soluciones difieren temporalmente (ver figura 1.1a), las trayectorias formadas a partir de estas soluciones visitan todos los puntos que conforman el atractor. Esto reduce el estudio de la dinámica de un fenómeno al análisis de las propiedades de su atractor. Por tanto, las características morfológicas y dinámicas de este atractor proporcionan mucha información acerca del fenómeno físico de interés.

Finalmente, pero no menos importante, debemos destacar que a pesar de las dificultades que puede representar el modelar un fenómeno físico mediante sistemas de ecuaciones no lineales (o sistemas de ecuaciones en diferencias no lineales), existe mucha evidencia que indica que este tipo de dinámicas abundan en la naturaleza [10, 12, 15, 16].

1.2 Atractores

Para comenzar a hablar de conjuntos atractores es necesario conocer dos conceptos importantes. El primero es el concepto de conjunto invariante y el segundo es el concepto de cuenca de un conjunto invariante. Dada una topología en el espa-

cio de fase, un conjunto Ω en este espacio es invariante respecto a la dinámica f si $f(\Omega) = \Omega$. Es decir, si una trayectoria comienza dentro del conjunto Ω , permanecerá en él por siempre [12]. Un conjunto abierto U que contenga a Ω se llama cuenca de Ω si para todo subconjunto abierto V que contenga a Ω , $f^t(U) \subseteq V$ (f^t denota la composición de f con ella misma t veces), para todo t lo suficientemente grande [17]. En otras palabras, las trayectorias que comienzan en U se aproximan a Ω de forma rápida y uniforme. Entonces, llamamos atractor de un sistema a un conjunto invariante Ω que tiene una cuenca de atracción U [17].

Una definición alternativa para el atractor de un sistema está basada en la idea de que en experimentos físicos y simulaciones numéricas es inevitable la presencia de ruido dinámico. Por tanto, un atractor es observable físicamente si éste es robusto ante ruido dinámico de muy baja magnitud. Entonces, un atractor es un conjunto invariante para el cual, en presencia de ruido dinámico de muy baja magnitud, cualquier trayectoria que comience en él permanecerá en él. Además se pide que exista “comunicación” entre diferentes puntos del atractor, es decir, que el atractor no pueda ser dividido en atractores más simples [17]. Observar que esta definición ya no considera una cuenca de atracción.

La diferencia con la primera definición puede ser analizada mediante la transformación logística (ecuación [1.1]) con $\mu = 4$. Si una trayectoria comienza en $x_0 = 0$ permanece en el conjunto invariante $\{0\}$. Es sabido que para este valor de μ la transformación logística cuenta con dos conjuntos invariantes [14]. El $\{0\}$ y el intervalo $(0, 1)$, al cual converge cualquier trayectoria que comience en este intervalo. Esto nos deja con dos conjuntos invariantes para un mismo sistema, en este caso. No obstante, bajo ruido dinámico infinitesimal el conjunto $\{0\}$ ya no será un conjunto invariante debido a que cualquier trayectoria que comience en $x_0 = 0$ escapará rápidamente hacia el conjunto $(0, 1)$, a causa de la perturbación del ruido, revelando la verdadera dinámica del sistema [17].

1.3 Reconstrucción del espacio de estados

La información contenida en el atractor de un sistema es suficiente para caracterizar su dinámica. Para reconstruir el atractor cuando no se dispone de las ecuaciones del sistema, es necesaria la observación directa de sus variables. Por ejemplo, para reconstruir el atractor del péndulo, si se tuviera un péndulo físico y no se conocieran los parámetros de la ecuación [1.1], sería necesario medir directamente su posición y velocidad de manera conjunta por un periodo prolongado de tiempo. En la práctica muchas veces las mediciones directas son imposibles. En ciertos casos, se puede

1.3. Reconstrucción del espacio de estados

contar con mediciones indirectas del tipo $s_{in} = h_i(\mathbf{x}_n)$ donde h_i es una función continuamente diferenciable y de valor real. Por ejemplo, la medición directa de la distancia de una estrella a la tierra es una tarea imposible, en su lugar los astrónomos utilizan la intensidad de la luz irradiada por la estrella que se observa (mide) en los telescopios terrestres. El problema es cómo reconstruir el atractor de un sistema a partir de medidas indirectas de los estados. La solución fue dada por Mañé en [18]. Él demostró que si tenemos m mediciones de un sistema $\mathbf{y}_n = (s_{1n}, s_{2n}, \dots, s_{mn})$ y si $m > 2d_0$, donde d_0 es la dimensión de conteo por cajas del atractor Ω (ver subsección 2.2.1), entonces la función $G : \mathbf{x}_n \in \Omega \rightarrow \mathbf{y}_n \in \mathbb{R}^m$ tiene la propiedad genérica² de ser uno a uno. Lo anterior nos dice que, salvo en algunas excepciones, si $m > 2d_0$ entonces el atractor reconstruido con mediciones \mathbf{y}_n será topológicamente equivalente al verdadero atractor del sistema.

En otros escenarios, no se cuenta con mediciones de todas las variables del sistema, sólo con mediciones temporales de una sola de las variables. La pregunta ahora es cómo reconstruir un atractor a partir de una única serie temporal. Esta brecha entre la práctica y la teoría encontró una primera solución cuando Packard y col. [19] propusieron el método de “derivative coordinates” que consiste en calcular las derivadas de orden superior de la serie temporal y utilizarlas en lugar de las variables faltantes del sistema. Por ejemplo, para el sistema dinámico del péndulo puede ser difícil medir de forma simultánea la posición $x(t)$ y velocidad $y(t)$. Sin embargo si contamos con mediciones de la posición, podemos estimar la velocidad mediante la derivada temporal $\dot{x}(t)$ de la posición. Así la trayectoria del sistema puede ser reconstruida utilizando los puntos $(x(t), \dot{x}(t))$. El mayor problema de esta metodología es que el proceso de derivación tiende a producir ruido de alta frecuencia, que sumado al ruido proveniente de la serie temporal hacen que las reconstrucciones no sean muy buenas [20].

Como alternativa, Takens [21] propuso el método de retardo de coordenadas (“delay coordinates”). Sea M una variedad (“manifold”) d -dimensional suave ($M \in C^2$) que constituye el espacio de estados de un sistema y $\phi : M \rightarrow M$ su dinámica. Supongamos que sólo tenemos acceso a una función escalar $s(n) = h(\mathbf{x}_n)$ dada por la función $h : M \rightarrow \mathbb{R}$, donde $\mathbf{x}_n = \phi^n(\mathbf{x}_0)$ es un estado del sistema que evolucionó a partir de la condición inicial \mathbf{x}_0 y ϕ^n denota la composición de ϕ con ella misma n veces. Es posible construir una función de retardo de coordenadas F como [21, 22]:

$$F : M \rightarrow \mathbb{R}^m$$

$$\mathbf{x}_n \rightarrow \mathbf{y}_n = F(\mathbf{x}_n) = (s(n), s(n + \tau), \dots, s(n + (m - 1)\tau)),$$

²Se dice que una propiedad es genérica para una clase de funciones F si F tiene un conjunto abierto y denso para el cual la propiedad se cumple [17].

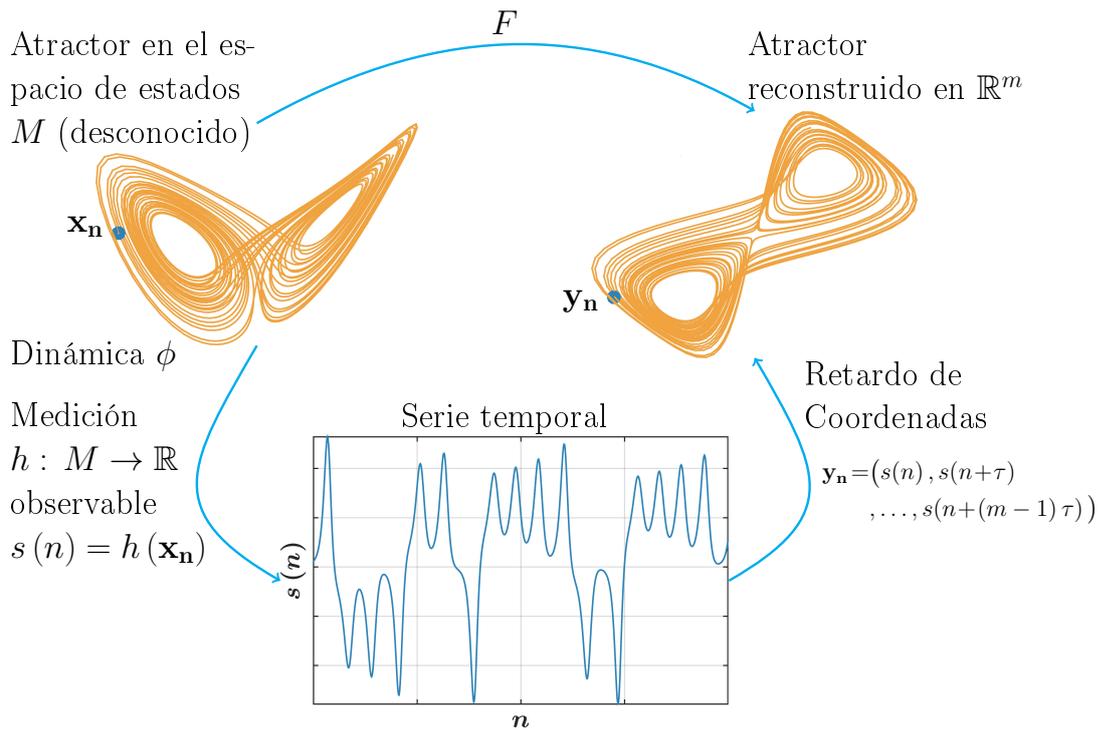


Figura 1.2. Reconstrucción de un atractor mediante retardo de coordenadas. Se presenta un diagrama de la reconstrucción del espacio de estados de una dinámica a partir de una serie temporal. Figura reproducida de [22].

donde m es la dimensión de inmersión y τ es el retardo de inmersión. Esta función lleva un estado \mathbf{x}_n del espacio original del sistema $M \in \mathbb{R}^d$ a un espacio reconstruido en \mathbb{R}^m . Takens demostró que si $m \geq 2d + 1$, entonces F tiene la *propiedad genérica* de ser una inmersión (“embedding”) de M en \mathbb{R}^m . Es decir F es una función uno a uno entre M y \mathbb{R}^m [22]. En este caso por propiedad genérica se entiende que el subconjunto de pares (h, τ) que conducen a una inmersión es abierto y denso dentro del conjunto de todos los pares (h, τ) [22]. En la figura 1.2 se muestra un esquema del teorema de reconstrucción.

Este teorema fue modificado por Sauer y col. de dos maneras [23, 24]. La primera reemplaza la condición $m \geq 2d + 1$ por $m \geq 2d_0$, donde d_0 es la dimensión de conteo por cajas del atractor $\Omega \in M$ (ver subsección 2.2.1). Esto representa una gran ventaja para sistemas físicos que posean un atractor de baja dimensión que viva en un espacio de estados de alta dimensión, como por ejemplo, el sistema de ecuaciones de Navier-Stokes [25]. En este sentido, para reconstruir un atractor con dimensión $d_0(\Omega) = 5$, inmerso en un espacio con dimensión $d = 100$, sería suficiente una dimensión de inmersión $m > 10$.

La segunda modificación del teorema es que F tiene la *propiedad prevalente* de ser una inmersión (no genérica como indica Takens), lo cual significa que “casi todo”³ par (h, τ) conducirá a una inmersión [22]. Esto quiere decir que el método de retardo de coordenadas puede reconstruir una gran diversidad de atractores, por este motivo es el método de reconstrucción más utilizado en la literatura referente al tema.

Es de resaltar que el método de retardo de coordenadas depende básicamente de dos parámetros que deben ser fijados antes de la reconstrucción. El primero es el retardo de inmersión (τ) y el segundo es la dimensión de inmersión (m). La correcta selección de sus valores es crucial para la adecuada reconstrucción del atractor y en consecuencia para la estimación de los invariantes del sistema.

1.3.1 Retardo de inmersión

El teorema de reconstrucción no impone restricciones sobre cómo elegir el retardo de inmersión, es decir, desde un punto de vista matemático τ podría tomar un valor arbitrario [22]. En otras palabras, diferentes valores de τ pueden conducir a atractores difeomórficamente equivalentes. Sin embargo, en la práctica solo se cuenta con un número de datos limitado y de precisión finita. Debido a esto, τ no puede ser fijado de manera arbitraria y su valor es muy importante para la adecuada reconstrucción del atractor de un sistema [12].

Podemos analizar como varía la reconstrucción de un sistema dinámico dependiendo del parámetro τ . Por ejemplo en el caso del sistema de Rössler:

$$\begin{cases} \dot{x} = -y - z, \\ \dot{y} = x + ay, \\ \dot{z} = b + z(x - c), \end{cases}$$

con $a = b = 0.2$ y $c = 5.7$. Cuando el valor de τ es muy pequeño las coordenadas del vector reconstruido son muy parecidas y por tanto la nube de puntos se agrupan alrededor de una recta en el espacio de estados (ver figura 1.3a). En este caso, es muy difícil reconocer cualquier tipo de estructura.

Por otro lado un valor muy grande de τ puede producir dobleces en el atractor reconstruido. Estos dobleces son un problema ya que hacen que puntos que no son

³En el correspondiente contexto se está trabajando en un espacio topológico medible por lo que existe una medida μ . “Casi todo” se refiere a que el conjunto de puntos donde no vale la propiedad tiene medida cero y entonces se dice que la propiedad vale “ μ -para casi todo punto”.

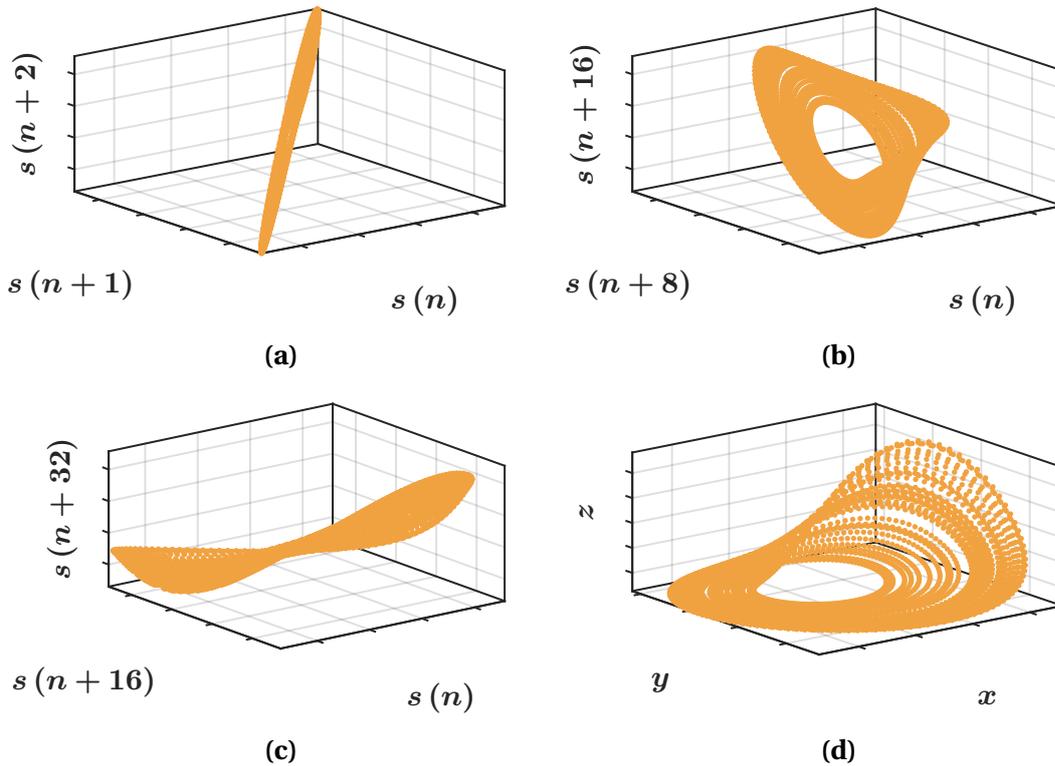


Figura 1.3. Reconstrucción del sistema de Rössler con diferentes retardos de inmersión. (a) $\tau = 1$, muy pequeño. (b) $\tau = 8$, adecuado. (c) $\tau = 16$, muy grande. (d) Atractor original del sistema.

vecinos en el atractor original lo sean en el reconstruido. Particularmente se debe tener cuidado en presencia de ruido donde los vectores de estado presentan una desviación de su lugar original. Además, un valor muy grande de τ puede ocasionar que los vectores de estado sean totalmente independientes, removiendo cualquier relación con la dinámica del sistema. Esto se evidencia como un plegado excesivo del atractor reconstruido, en el cual las trayectorias se entrecruzan unas con otras (ver figura 1.3c) y en consecuencia la reconstrucción deja de ser biyectiva por problemas de precisión numérica [22].

Un valor adecuado de τ garantiza que la reconstrucción preserve las características del atractor original (ver figuras 1.3b y 1.3d). Existen varios criterios para elegir el valor del retardo de inmersión. No obstante, los dos más utilizados están basados, uno en la función de autocorrelación y otro en la función de información mutua [15].

La función de autocorrelación da cuenta de las correlaciones lineales entre muestras desplazadas de la serie temporal. Por este motivo, se suele elegir como valor de τ al tiempo para el cual esta función alcanza su primer cruce por cero. Muchas veces

esta función siempre es positiva; para tales casos, se elige como retardo de inmersión al valor temporal para el cual la función de autocorrelación decae por debajo de e^{-1} [11, 15]. El principal problema de utilizar la función de autocorrelación es que ésta es un estadístico lineal, por tanto no tiene la capacidad para detectar correlaciones de tipo no lineal debidas a la dinámica del sistema. No tener en cuenta este tipo de correlaciones puede llevar a utilizar un valor de τ más grande del necesario.

Como alternativa se utiliza la función de información mutua [15]:

$$I(T) = \sum_{n=1}^N \Pr\{s(n), s(n+T)\} \log_2 \left[\frac{\Pr\{s(n), s(n+T)\}}{\Pr\{s(n)\} \Pr\{s(n+T)\}} \right], \quad (1.2)$$

donde $\Pr\{s(n)\}$ es la probabilidad de observar $s(n)$ y $\Pr\{s(n), s(n+T)\}$ es la probabilidad de observar conjuntamente $s(n)$ y $s(n+T)$. Esta función mide el grado de redundancia o correlación no lineal entre muestras de la serie temporal. En otras palabras, la función de información mutua nos indica cuánta información sobre la muestra $s(n+T)$ está contenida en la muestra $s(n)$ [12, 15]. En este sentido, τ es fijado como el valor de tiempo para el cual esta función encuentra su primer cruce por cero [11, 12, 15, 26].

1.3.2 Dimensión de inmersión

El teorema de reconstrucción nos indica que la dimensión de inmersión m debe ser mayor a dos veces la dimensión de conteo por cajas del atractor. En esencia, cualquier valor de $m > 2d_0$ sería suficiente para alcanzar una inmersión, sin embargo, se deben tener en cuenta dos cuestiones. La primera es que la mayoría de los invariantes de un sistema requieren para su cálculo la estimación de una medida de probabilidad inherente al atractor (ver capítulo 2). Por consiguiente, si la dimensión de inmersión crece, también lo hará el número de datos necesarios para alcanzar una buena estimación. Es decir, si en una dimensión uno son necesarios L vectores de estado para una buena estimación, en dos dimensiones serán necesarios L^2 vectores de estado. La segunda es que el incremento en el número de datos lleva consigo un incremento en el costo computacional de los algoritmos de estimación [12]. En base a esto y sabiendo que, por lo general, no se tiene conocimiento previo de la dimensión de conteo de cajas de un sistema, es necesario encontrar el valor de m más pequeño que cumpla con la condición impuesta por el teorema de reconstrucción. Este valor de m se lo llama mínima dimensión de inmersión m_0 .

Existen varios criterios para estimar la mínima dimensión de inmersión, sin embargo el más conocido es el método de “falsos vecinos” propuesto por Kennel y col.

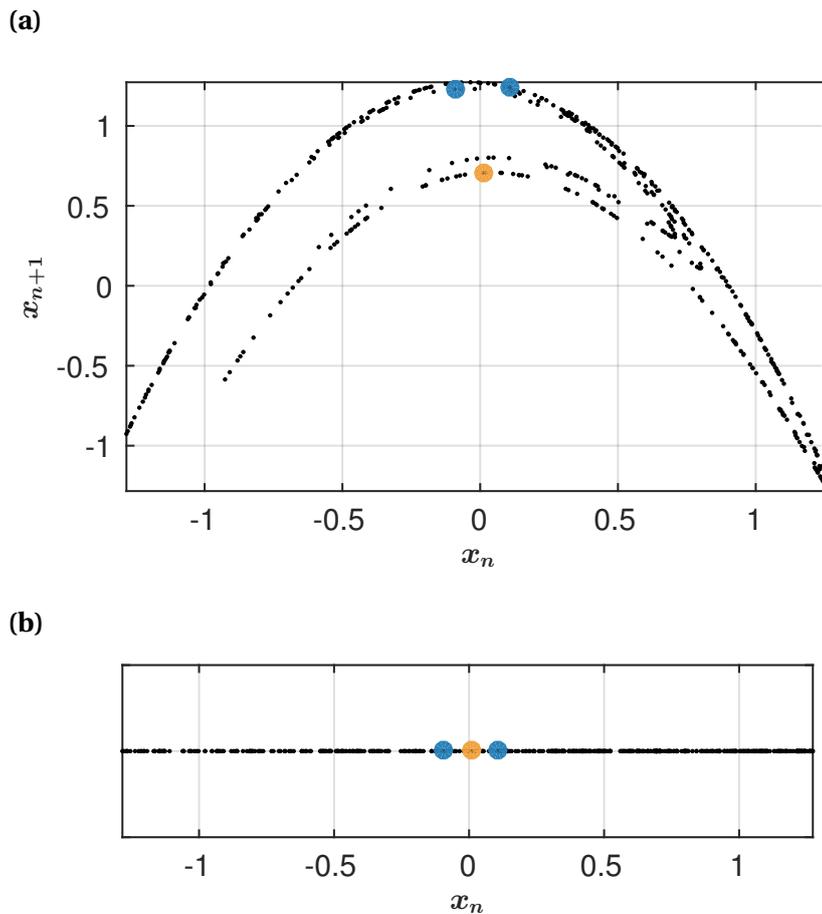


Figura 1.4. Falsos vecinos: se muestra sobre el atractor de Hénon tres puntos, dos azules y uno naranja. En una dimensión $m = 2$ el punto naranja no es vecino de los azules (a), pero cuando se proyectan en una dimensión menor $m = 1$ si lo son (b). Figura reproducida de [27].

[27]. Conceptualmente, los falsos vecinos son puntos que son cercanos debido al proceso de inmersión y no a la dinámica del sistema [15, 28]. Para explicar esta idea retornemos por un instante al los espacios de estados de sistemas dinámicos dados por ecuaciones diferenciales y consideraremos la hipótesis de que la dinámica de evolución de un sistema en el espacio de estados es una función suave (función continuamente diferenciable). En este sentido, si un par de estados son vecinos, las trayectorias que comiencen en ellos deberán permanecer cercanas en un corto periodo de tiempo, incluso si la dinámica es caótica [12]. Por tanto, los falsos vecinos son estados aparentemente cercanos (debido a que son proyecciones en dimensiones inferiores a la mínima dimensión de inmersión) cuya evolución temporal inmediata es muy diferente. Para ilustrar esto, ahora en relación con series de datos

1.3. Reconstrucción del espacio de estados

discretas, digamos que contamos con muestras $s(n)$ de un sistema cuya mínima dimensión de inmersión es $m_0 = 2$. Ahora, supongamos un espacio reconstruido con $m = 1$ donde los vectores de estado $\mathbf{x}_n = s(n)$ y $\mathbf{y}_{\hat{n}} = s(\hat{n})$, $n \neq \hat{n}$, son vecinos, pero su evolución luego de un tiempo τ (corto) $\mathbf{x}_{n+\tau} = s(n + \tau)$ y $\mathbf{y}_{\hat{n}+\tau} = s(\hat{n} + \tau)$ son muy diferentes. Si realizamos la inmersión con $m = 2$ obtendremos que los vectores de estado correspondientes en esta dimensión $\mathbf{x}_n = (s(n), s(n + \tau))$ y $\mathbf{y}_{\hat{n}} = (s(\hat{n}), s(\hat{n} + \tau))$ no son vecinos, ya que su segunda coordenada es muy diferente. Por ejemplo, en la figura 1.4 se presenta una reconstrucción, en una y dos dimensiones, de la transformación de Hénon:

$$x_{n+1} = 1 - a x_n^2 + b x_{n-1}, \quad (1.3)$$

donde $a = 4$ y $b = 0.3$. Este sistema dinámico discreto es uno de los más estudiados en el área de las dinámicas no lineales ya que está bien caracterizado. Podemos ver en la figura 1.4a claramente cómo el punto naranja no es vecino de los puntos azules. Sin embargo, en la figura 1.4b se observa que al proyectar el atractor en una dimensión menor, los puntos azules y el naranja se convierten en vecinos. Es decir, la segunda componente del punto naranja y aquella correspondiente a los puntos azules difieren en gran medida. En consecuencia, en un espacio de dimensión uno, el punto naranja es un “falso vecino” respecto a los puntos azules, mientras que estos últimos son verdaderos vecinos.

Debemos ahora pensar en un estadístico que nos permita estudiar este fenómeno. Si m es mayor o igual a la mínima dimensión de inmersión, el número de falsos vecinos debe ser igual a cero. En este sentido, un posible estadístico es el porcentaje de falsos vecinos presentes en una inmersión dada [27]. Consideremos entonces una serie temporal $\{s(i)\}_{i=1}^N$ y los vectores de estado reconstruidos:

$$\mathbf{x}_n = (s(n), s(n + \tau), \dots, s(n + (m - 1)\tau)) \quad \text{con} \quad n = 1, 2, \dots, L = N - (m - 1)\tau,$$

de dimensión m . Para cada vector de estado \mathbf{x}_n debemos encontrar su vecino más próximo $\mathbf{y}_{\hat{n}}$ utilizando alguna medida de distancia. Luego debemos incrementar la dimensión de inmersión en una unidad, obteniendo así un vector de estado $\bar{\mathbf{x}}_n$ en un espacio de dimensión $(m + 1)$ y su respectivo vecino más cercano $\bar{\mathbf{y}}_{\hat{n}}$ en esa dimensión:

$$\begin{aligned} \bar{\mathbf{x}}_n &= (s(n), s(n + \tau), \dots, s(n + (m - 1)\tau), s(n + m\tau)) \\ \bar{\mathbf{y}}_{\hat{n}} &= (s(\hat{n}), s(\hat{n} + \tau), \dots, s(\hat{n} + (m - 1)\tau), s(\hat{n} + m\tau)). \end{aligned}$$

Un criterio para saber si \mathbf{x}_n y $\mathbf{y}_{\hat{n}}$ son falsos vecinos consiste en comparar la diferencia entre sus distancias en un espacio m -dimensional y $(m + 1)$ -dimensional contra un

parámetro R_T [27]:

$$FNN = \begin{cases} 1, & \text{si } \frac{\|\bar{\mathbf{x}}_n - \bar{\mathbf{y}}_n\|^2 - \|\mathbf{x}_n - \mathbf{y}_n\|^2}{\|\mathbf{x}_n - \mathbf{y}_n\|^2} = \frac{(s(n+m\tau) - s(\hat{n}+m\tau))^2}{\|\mathbf{x}_n - \mathbf{y}_n\|^2} \geq R_T, \\ 0, & \text{en otro caso.} \end{cases}$$

En base a este criterio podemos obtener el porcentaje de falsos vecinos como:

$$\%FNN = 100 \sum_{n=1}^L \frac{FNN}{L},$$

donde L es el número total de vectores de estado. El parámetro R_T debe ser fijado según la distribución espacial de los puntos en el espacio de estados reconstruidos [15]. Si R_T es muy pequeño, verdaderos vecinos serán contados como falsos. Por el contrario, si este parámetro es muy grande, algunos falsos vecinos no serán tenidos en cuenta. Además, a medida que m crece el atractor se va desdoblado, por tanto R_T debe cambiar su valor [15]. Kennel y col. [27], basados en conocimientos empíricos sugieren que $10 \leq R_T \leq 30$.

El mayor problema de la metodología propuesta por Kennel y col. es que el valor del parámetro R_T es dependiente de la aplicación y de los datos. No existe un consenso general para elegir su valor. Como alternativa Cao [29] propuso un algoritmo diferente para encontrar la mínima dimensión de inmersión. Si bien este algoritmo está basado en el principio de contar “falsos vecinos”, su formulación elimina la necesidad de fijar un parámetro externo, como R_T , para contarlos. Cao propone calcular la cantidad:

$$a(n, m) = \frac{\|\bar{\mathbf{x}}_n - \bar{\mathbf{y}}_n\|}{\|\mathbf{x}_n - \mathbf{y}_n\|},$$

con $n = 1, 2, \dots, L = N - m\tau$ que describe la razón entre la distancia de dos puntos $(m+1)$ -dimensionales y la distancia entre sus correspondientes proyecciones en un espacio de dimensión m . Con estas cantidades se define el estadístico:

$$El(m) = \frac{E(m+1)}{E(m)},$$

donde

$$E(m) = \frac{1}{N - m\tau} \sum_{n=1}^{N-m\tau} a(n, m).$$

Cao encontró que si la dinámica proviene de un atractor, $E1(m)$ deja de cambiar cuando m es mayor que algún valor m_0 . Por tanto $m_0 + 1$ es la mínima dimensión de inmersión. El algoritmo de Cao puede estimar la mínima dimensión de inmersión de forma guiada por los datos. Por este importante motivo hemos decidido utilizarlo en nuestras simulaciones.

1.4 Sistemas fisiológicos y dinámicas no lineales

En medicina se define una crisis cómo un punto crucial, en la evolución de una patología, donde se define la vida o muerte del paciente [30]. Es una prioridad de la medicina poder detectar crisis, ya que a partir de ellas es posible realizar un diagnóstico que indique su causa. Desde tiempos ancestrales la medicina ha tratado de abordar esta laboriosa tarea. En palabras del mismo Hipócrates de Cos: “Detecta una crisis, descubre su causa y luego enseña cómo diagnosticarla” [30]. Con este fin, el médico debe recaudar información a partir de señales del cuerpo humano, con ella debe detectar anomalías y mediante un proceso de razonamiento lógico debe encontrar su causa. Sin embargo, hasta un médico experimentado es propenso a pasar por alto algún síntoma, lo cual puede repercutir en el posterior diagnóstico. Por este motivo, uno de los objetivos de los médicos e ingenieros es mejorar la sensibilidad y especificidad en la detección de crisis. Esto ha derivado en la creación e implementación de diversas herramientas electrónicas e informáticas que permitan sobrepasar la sensibilidad de los sentidos humanos.

Con el fin de automatizar el proceso de detección de patologías, es necesario primero registrar digitalmente las señales fisiológicas provenientes del cuerpo humano. En este punto dichas señales se transforman en información que puede ser guardada y procesada mediante algoritmos. La naturaleza de estas metodologías de procesamiento está estrechamente ligada a la naturaleza del sistema en estudio, más concretamente a las hipótesis que el investigador establezca sobre este sistema. En el procesamiento digital de señales existen dos corrientes básicas para el análisis de señales fisiológicas. La primera es la escuela clásica que trata a los sistemas fisiológicos como sistemas de naturaleza lineal. Es decir, cada sistema fisiológico puede ser subdividido en sistemas más pequeños e independientes entre sí. De esta forma cada subsistema puede ser estudiado en forma aislada. Los adeptos a esta escuela encuentran especial utilidad en técnicas lineales, como el análisis tiempo-frecuencia⁴ mediante la transformada de Fourier o la transformada Ondita [31].

⁴Los esquemas tiempo-frecuencia o tiempo-escala en realidad suponen que el sistema puede considerarse lineal pero durante un período lo suficientemente corto de tiempo [31].

Por otro lado, está la corriente que modela los sistemas fisiológicos como sistemas complejos cuya dinámica es no lineal. Al igual que los sistemas lineales, este tipo de sistemas también puede ser dividido en subsistemas, sin embargo, los mismos no son independientes entre sí. En otras palabras, los diferentes subsistemas comparten información. Esta interacción es la que causa que la dinámica sea no lineal. Existe amplia evidencia experimental que muestra que los sistemas fisiológicos son sistemas complejos cuya dinámica es no lineal [30, 32-39]. Está bien establecida la idea de que el grado de complejidad de un sistema fisiológico está estrechamente ligado con estados normales y patológicos [32]. En estado normal, el grado de complejidad de los sistemas de control fisiológicos es alto. Dicha complejidad provee al organismo de “grados de libertad” para poder reaccionar ante cambios intempestivos en el entorno. Por lo tanto, la pérdida de complejidad por parte los sistemas de control fisiológico a causa de factores como el envejecimiento, es el origen de algunas patologías [32].

Las señales biomédicas son el resultado de la dinámica de los sistemas fisiológicos de los cuales provienen. En este sentido, contienen información acerca del estado en el que dichos sistemas se encuentran. Debido a la naturaleza de la fisiología humana, tiene sentido estudiar esta señales por medio de técnicas basadas en la teoría de las dinámicas no lineales.

En las siguientes subsecciones describiremos dos sistemas fisiológicos de especial interés para el desarrollo de este trabajo. El sistema nervioso central y el sistema productor de la voz humana. Hemos decidido trabajar con estos sistemas debido a que sus señales fisiológicas pueden ser adquiridas de manera no invasiva y por la amplia experiencia en ellas que acredita el grupo de trabajo en el que se desarrolló esta tesis doctoral.

1.4.1 Sistema nervioso central

El sistema nervioso humano puede ser dividido en dos grandes subsistemas: el sistema nervioso central y el sistema nervioso periférico. El sistema nervioso central está, a su vez, compuesto por el encéfalo o cerebro y por la médula espinal. El sistema nervioso periférico está compuesto por nervios que conectan al sistema nervioso central con el resto de órganos del cuerpo [40]. El sistema nervioso central, es el centro de procesamiento de la información que proviene de las ramas aferentes del sistema nervioso periférico. En particular, el cerebro es quién determina si, ante un estímulo, es necesaria una respuesta. De serlo, ésta es enviada por la vía eferente hacia su órgano objetivo (principalmente músculos y glándulas) [41].

1.4. Sistemas fisiológicos y dinámicos no lineales

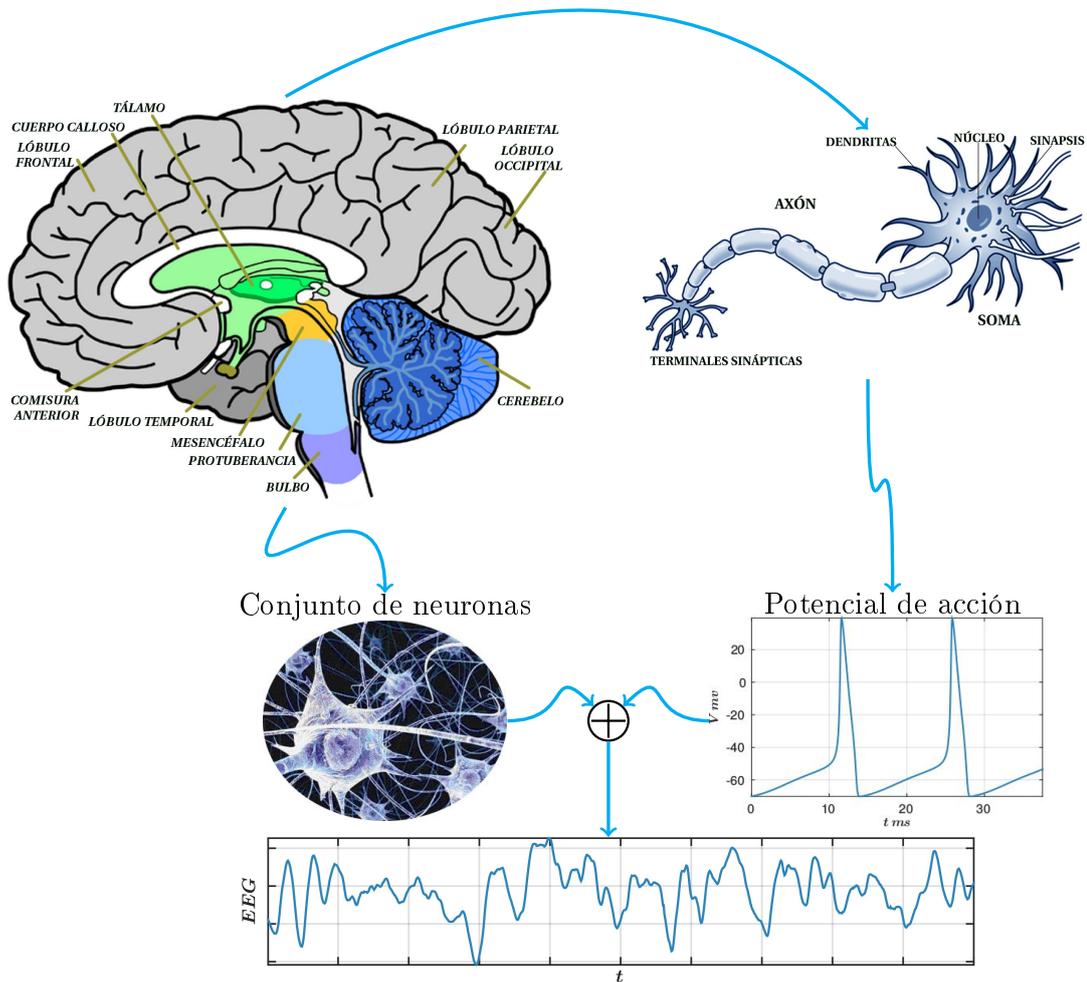


Figura 1.5. Se presenta la anatomía y fisiología del cerebro. Algunas imágenes fueron modificadas de [42].

El cerebro es la estructura más importante que compone el sistema nervioso central. Su alta complejidad y versatilidad ha maravillado a los investigadores por siglos. Este ha sido objeto de estudio constante, partiendo desde niveles básicos como su composición química, hasta el estudio de mecanismos avanzados como la memoria y el aprendizaje [40]. El encéfalo, que se presenta en la figura 1.5, está dividido en: los hemisferios cerebrales, el diencefalo, el mesencéfalo, el cerebelo, la protuberancia y el bulbo raquídeo [43]. La unidad básica funcional que compone el sistema nervioso es la neurona. Estas células se encargan de procesar y transmitir información de un lugar a otro de este sistema. Existen muchos tipos de neuronas, las cuales se diferencian por su tamaño, forma y función [40]. Generalmente, una neurona está conformada por el soma o cuerpo, del cual se desprenden las dendritas y el axón (ver figura 1.5). Las dendritas son pequeñas ramificaciones utilizadas

para recibir información procedente de neuronas vecinas [41], la cual es retransmitida a otras neuronas por medio del axón [43]. El mecanismo de transmisión de información entre neuronas es de tipo electro-químico. Este comienza en la sinapsis, que es el lugar donde la parte terminal del axón de una neurona, llamada neurona presináptica, se encuentra con las dendritas o el soma de otra neurona, denominada postsináptica. En la sinapsis, la información eléctrica que viaja por la neurona presináptica, llamada potencial de acción (ver figura 1.5), es transformada en información química (neurotransmisores) para ser comunicada a la neurona postsináptica [40]. Luego esta información es codificada de nuevo en impulsos eléctricos que viajan a través del axón de la neurona postsináptica. Esto se da gracias a un intercambio de iones entre la neurona y el medio circundante [41]. Al final, la neurona postsináptica se transforma en neurona presináptica para continuar con la transmisión de la información.

La transmisión sináptica es el paso más vulnerable en el proceso de comunicación entre neuronas. Esto se debe a que los receptores sinápticos están expuestos al fluido extracelular, lo que los hace accesibles a diferentes compuestos químicos como drogas y venenos [41]. Diferentes problemas en la transmisión sináptica han sido vinculados con algunas patologías como la esquizofrenia, la depresión y el Parkinson [41].

Una patología de especial interés en esta tesis es la epilepsia. Ésta se caracteriza por una actividad excesiva e incontrolada de las neuronas en todo el sistema nervioso o en partes de él [43]. Esta enfermedad se puede clasificar en tres tipos: grand mal, petit mal y focal. La epilepsia grand mal se distingue por descargas neuronales extremas en todas las áreas del cerebro. Estas son transmitidas por la médula espinal y los nervios a los músculos, provocando convulsiones tónicas generalizadas en la totalidad del cuerpo. Estos ataques pueden durar desde pocos segundos hasta cuatro minutos. La epilepsia petit mal se caracteriza por tres a treinta segundos de inconsciencia. Durante este periodo la persona puede sufrir de contracciones musculares leves en la boca o en los párpados. Este tipo de epilepsia se presenta, por lo general, al final de la infancia y eventualmente puede iniciar un ataque epiléptico de tipo grand mal. La epilepsia focal puede involucrar partes de la corteza o estructuras internas más profundas. Esta es el resultado de lesiones en el tejido nervioso o de anomalías funcionales localizadas, como un tumor. El tejido afectado por estas lesiones produce descargas rápidas que pueden expandirse al tejido circundante. Por ejemplo, cuando la lesión está cerca del área motora puede desencadenar movimientos involuntarios. En ocasiones, estas descargas focalizadas pueden ser tan fuertes que desencadenan un episodio epiléptico de gran mal [43].

1.4. Sistemas fisiológicos y dinámicas no lineales

Para desarrollar los diferentes procesos que permiten el normal desenvolvimiento de una persona, es necesario que un número grande de neuronas estén trabajando de forma colectiva. En una neurona, la transmisión de potenciales de acción modifica electroquímicamente el medio circundante. Este efecto es muy débil. Sin embargo, si sumamos los efectos que producen millones de neuronas trabajando juntas obtenemos una señal eléctrica que puede ser medida en el cuero cabelludo mediante electrodos. Esta señal eléctrica nos provee mucha información sobre el funcionamiento del cerebro y puede ser adquirida mediante una técnica llamada electroencefalografía (ver figura 1.5 [43]).

Señales de Electroencefalografía

Los diferentes procesos fisiológicos que se llevan a cabo en el cerebro generan señales en diferentes partes de la corteza cerebral. La actividad colectiva de las neuronas produce un campo eléctrico lo suficientemente fuerte para ser registrado por medio de electrodos posicionados en el cuero cabelludo [40]. El electroencefalograma (EEG) consiste en el registro de señales eléctricas, provenientes principalmente de la corteza cerebral, mediante electrodos adecuadamente posicionados en el cuero cabelludo. La morfología de las señales de EEG es por lo general oscilatoria y repetitiva, por este motivo se les conoce también como ritmos [40]. La amplitud y el rango de frecuencias de estos ritmos están estrechamente relacionadas con la actividad neuronal y pueden ser utilizadas para caracterizar diferentes estados mentales. Por ejemplo, oscilaciones de alta frecuencia y baja amplitud están relacionadas con estados de alerta, mientras que señales de baja frecuencia y alta amplitud se asocian a estados de reposo [43].

La amplitud del EEG, que oscila generalmente entre 0 y $100\mu\text{V}$, describe el grado de sincronización a la cual las neuronas de la corteza interaccionan [40]. La sincronía entre neuronas causa que la actividad eléctrica se sume de forma coherente. En este sentido, los ciclos de excitación sincrónica generan señales rítmicas de gran amplitud. Por el contrario, la excitación asincrónica de grupos distintos de neuronas genera señales de EEG irregulares y de baja amplitud como se muestra en la figura 1.6a [44].

Las patologías del sistema nervioso central también se reflejan en la señal de EEG. Esta señal ha sido indispensable en la práctica clínica para detectar crisis epilépticas. Su importancia radica en su gran resolución temporal y que puede ser adquirida de forma económica y no invasiva [35]. Las señales eléctricas registradas durante una crisis epiléptica se denominan EEGs ictales [40]. En general, estos contiene oscilaciones anormales con patrones de onda tipo picos y con incrementos repenti-

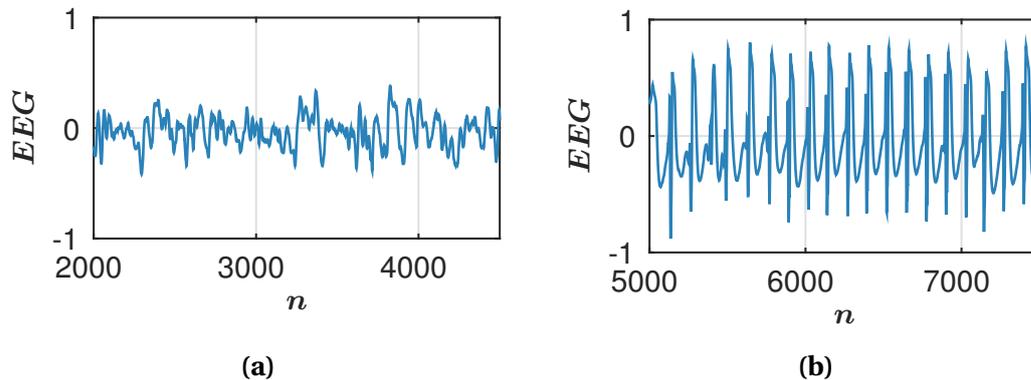


Figura 1.6. Señal de electroencefalograma. Se presentan segmentos de EEG (a) normal y (b) ictal. El segmento ictal corresponde a una crisis epiléptica del tipo petit mal

nos de amplitud. El EEG registrado al comienzo de los episodios ictales se caracteriza por un aumento súbito de la frecuencia [40]. Los episodios ictales de la epilepsia de tipo grand mal se distinguen por descargas sincrónicas de gran amplitud sobre la totalidad de la corteza cerebral. Por otro lado, el EEG de una epilepsia petit mal presenta un patrón de espiga-onda, como se muestra en la figura 1.6b. La porción de las oscilaciones en forma de espiga son muy parecidas a las de la epilepsia grand mal, sin embargo la onda que le prosigue es su sello distintivo.

Debido a que la señal electroencefalográfica se obtiene de electrodos posicionados en el cuero cabelludo presenta dos problemas: baja relación señal a ruido y baja resolución espacial. Existen otras técnicas para medir las señales eléctricas de la corteza cerebral que mejoran las deficiencias de la señal de EEG pero a costa de ser invasivas. La electrocorticografía utiliza arreglos de electrodos que son implantados en la corteza. Esto mejora en gran medida la relación señal a ruido. De forma similar, la estéreo-encefalografía utiliza pequeños electrodos implantados en regiones focalizadas del cerebro. Esta técnica presenta un excelente resolución espacial y alta relación señal a ruido, lo que la hace ideal para detectar focos epilépticos [35].

Por todo lo mencionado con anterioridad, vemos que el sistema nervioso central y en especial el cerebro es un sistema altamente complejo desde el punto de vista biológico. Si bien se han realizado muchos descubrimientos acerca de su comportamiento mediante el uso de técnicas lineales, la información que las mismas proveen es limitada. Esto radica principalmente en que la dinámica cerebral es altamente no lineal. La no linealidad es introducida a nivel celular ya que el mecanismo de activación de la neurona involucra operaciones de integración, umbralado y saturación. Existen quienes abogan por la idea de que un arreglo de millones de unidades no lineales puede modelarse a gran escala de forma lineal, sin embargo, es claro

que la dinámica de patologías como la epilepsia es no lineal [35]. Por estos motivos muchos investigadores han dirigido su atención al uso de técnicas no lineales para analizar la complejidad del cerebro a partir de señales de EEG [30, 36, 45-48]. Por ejemplo, Kelso y col. mostraron que existe una gran similitud entre señales de EEG registradas durante crisis epilépticas del tipo petit mal y dinámicas provenientes del sistema caótico de Shilnikov [49]. Este tipo de avances ha dado lugar para que otros investigadores diseñaran algoritmos para la detección automática de episodios ic-tales a partir de EEG, combinando técnicas no lineales y teoría de reconocimientos de patrones y clasificación [50-56].

1.4.2 Sistema fonador

La voz humana es una onda de presión sonora que viaja a través del aire hasta el sistema auditivo del oyente, donde es procesada para luego formar una idea del mensaje recibido. El aparato fonador puede ser descrito como un sistema que transforma energía muscular en energía acústica [57]. En la figura 1.7 se presenta una imagen anatómica del aparato fonador y debajo una representación fisiológica de la producción de la voz. Este sistema se puede dividir en tres subsistemas interconectados llamados: respiratorio, laríngeo y tracto vocal [58].

El subsistema respiratorio es el encargado de llevar a cabo el proceso de respiración. Este proceso visto desde la mecánica consiste en el ingreso de aire desde el ambiente hacia el interior de los pulmones (inspiración) y el egreso de aire desde el interior de los pulmones hacia el ambiente circundante (expiración). Durante la inspiración, el aire ingresa por un descenso de presión intrapulmonar, consecuencia del incremento de la capacidad volumétrica de la cavidad torácica. Este efecto es logrado a partir de la acción conjunta de los músculos intercostales y el diafragma. Su función más importante es la de provocar el descenso del límite inferior de la cavidad torácica durante la inspiración, así como también regular la cantidad de aire expulsada en el momento de la producción vocal propiamente dicha [59].

La laringe es un órgano superficial que se puede palpar a través de la piel. Está situado en la parte media anterior del cuello, debajo del hueso hioides y por delante de la faringe con la cual se continúa hacia arriba. La variabilidad de sus dimensiones entre los distintos individuos redundan en la emisión de voces más agudas o más graves. Dentro de la laringe se encuentran las cuerdas vocales que internamente están compuestas por ligamentos y tejido muscular. Esto les confiere propiedades de elasticidad y contractilidad, indispensables para efectuar la fonación.

De acuerdo al estado de contracción o relajación, entre las cuerdas vocales queda

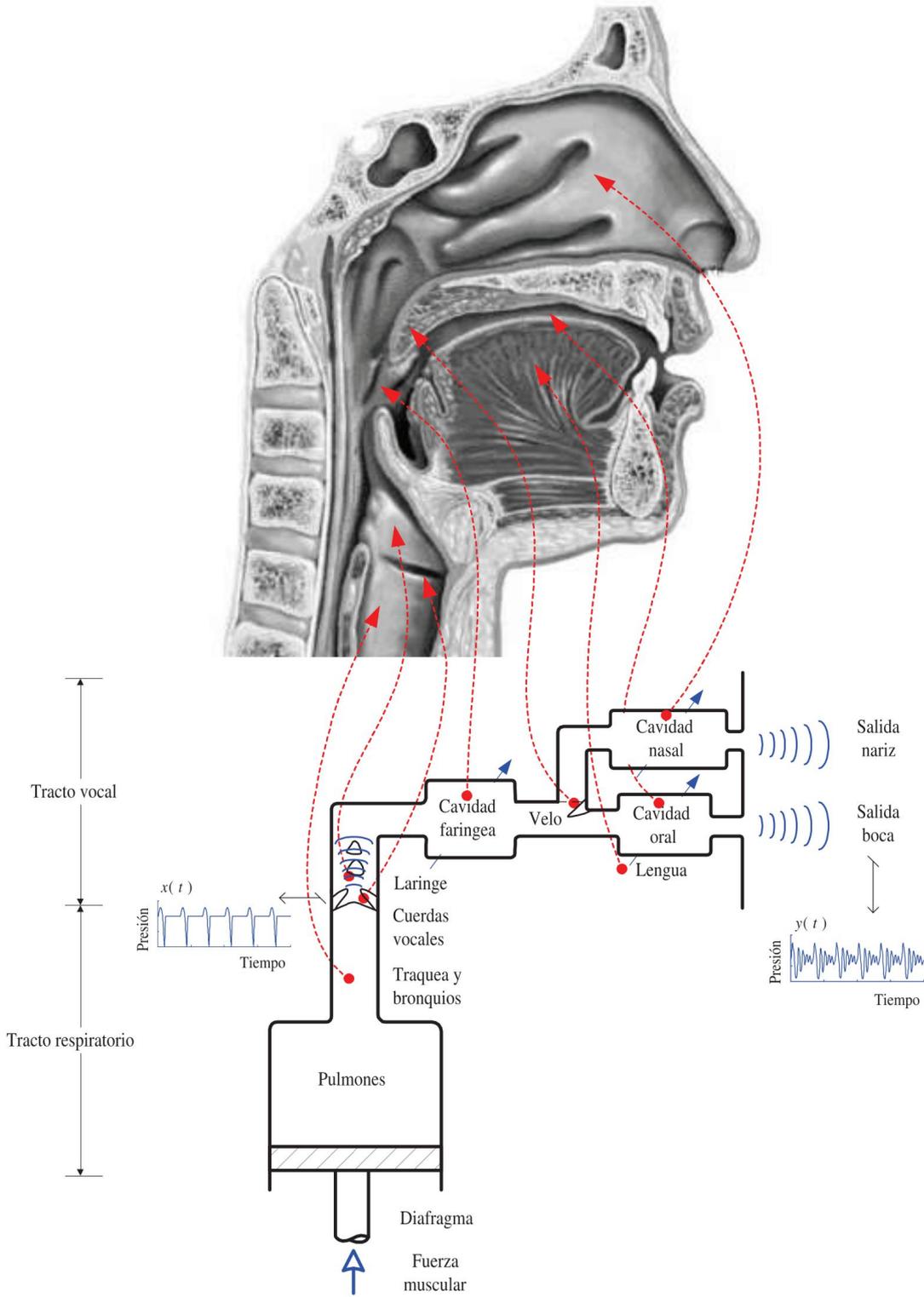


Figura 1.7. Se presenta un esquema anatómico y fisiológico del aparato fonador. Imagen tomada de [57].

1.4. Sistemas fisiológicos y dinámicas no lineales

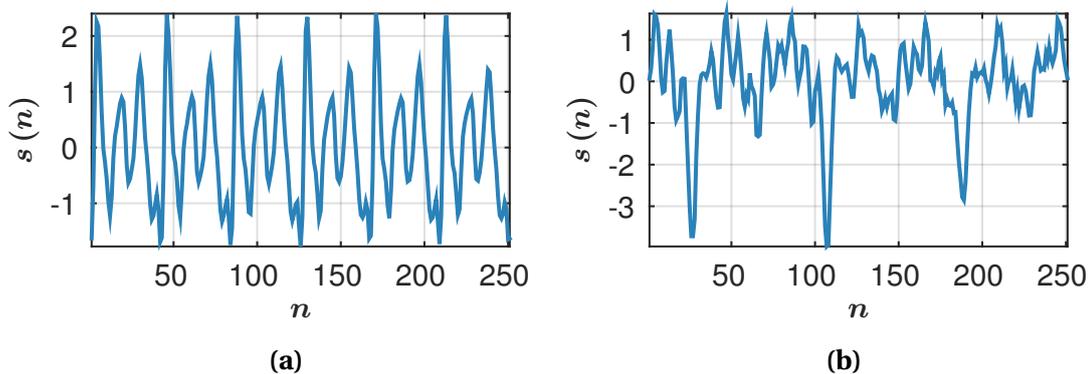


Figura 1.8. Señal de voz. Se presentan segmentos de voz correspondientes a una vocal /a/ sostenida. (a) Voz sana. (b) Voz patológica. Estas señales provienen de la base de datos “Voice disorders database” de Kay Elemetrics Corp. [62].

comprendido un espacio de forma triangular y dimensiones variables denominado glotis [60]. La laringe tiene un esqueleto y músculos intrínsecos y extrínsecos que, de acuerdo con las circunstancias, le permiten cerrar o abrir su luz (glotis), tensar las cuerdas vocales y/o desplazarse verticalmente en el cuello, según lo requiere la deglución, la respiración o la fonación [59]. En relación con el proceso de la voz, la laringe actúa como la fuente de energía acústica a partir del flujo de aire provisto por el subsistema respiratorio [58].

El tracto vocal está conformado por las estructuras que se hallan “por encima” de la laringe. Estas comprenden las vías aéreas superiores, la faringe y las cavidades bucal y nasal [60]. Durante la producción del habla el tracto vocal actúa como un filtro acústico variable en el tiempo que modifica la onda sonora producida en la glotis (ver figura 1.7). La variabilidad es causada por la modificación de la geometría de este tracto al hablar [58], es decir, por el movimiento del paladar blando, la lengua, los labios y la mandíbula. Se denomina articulación al proceso de “ajustar” el tracto vocal para producir los diferentes sonidos del habla.

La producción de la voz normal posee cuatro atributos principales: tono apropiado según la edad y el género; la sonoridad suficiente para la situación en la cual se desenvuelve el individuo; resonancia oral y nasal apropiada y calidad agradable al ser escuchada. La carencia de alguno de estos atributos es considerado un trastorno vocal que según Lehto se define como la “posible causa de cambios en la estructura o en la función del mecanismo laríngeo de la producción de la voz, a un grado tal que no permite al hablante cubrir las necesidades básicas funcionales de uso de su voz” [61].

Los trastornos vocales se pueden clasificar según su etiología en trastornos estructurales y neurológicos [63]. Los trastornos estructurales son causados por una alteración de la anatomía y/o la fisiología normal de la laringe en sí misma o de otra estructura o mecanismo nervioso que altere, su forma o función [64]. Los trastornos orgánicos se clasifican a su vez en estructurales o neurogénicos. Los estructurales son aquellos trastornos asociados a alguna anomalía física de la laringe en sí misma, ya sea por una lesión traumática o la presencia de una formación histológica anormal que interfiera directamente con el movimiento de las cuerdas vocales. Como ejemplos de este tipo de trastornos se encuentran los nódulos, los pólipos, la laringitis, entre otros. Los síntomas más comunes son la disfonía, ronquera con voz áspera, tendencia al uso de tonos graves. Por otra parte, los trastornos neurogénicos son aquellos causados por una alteración en el sistema de inervación de la laringe, ya sea a nivel de sistema nervioso periférico o central [63]. Puede decirse que la afectación de las neuronas motoras laríngeas a nivel periférico puede ocasionar cuadros de parálisis de las cuerdas vocales. Por ejemplo, la disfonía espasmódica se caracteriza por una producción vocal excesivamente forzada, debida a una falta de control motor del mecanismo laríngeo [65].

Señal de voz

Con motivo de llevar a cabo el proceso de detección de patologías vocales, es preciso contar con señales provenientes del aparato fonador que provean información sobre su estado. Desde un punto de vista biomédico, la voz es el resultado de la generación, propagación, modificación y emisión de una onda sonora, desde el interior del cuerpo hacia el medio circundante [66]. La voz, como onda sonora, produce en el aire una ligera modificación alternante de la presión atmosférica. Esta variación puede ser capturada y transducida mediante el empleo de un micrófono que es un elemento transductor acústico-eléctrico capaz de producir un correlato eléctrico a partir de una señal de presión acústica. El registro de la señal de voz a partir de micrófonos ha sido, la técnica de medición más empleada por los patólogos del sistema fonador [67]. En la figura 1.8 se presentan dos señales de voz correspondientes a una vocal /a/ sostenida. Es posible ver en la figura 1.8a cómo para una persona sana la señal de voz presenta una estructura cuasi-periódica, en el caso de un fonema sonoro. Por el contrario en la figura 1.8b se puede apreciar cómo para una persona que padece una patología vocal, la morfología cuasi-periódica se ha perdido y la señal muestra una morfología mucho más irregular que el caso no patológico.

El estudio de las diferentes patologías vocales mediante técnicas no lineales ha da-

do lugar a numerosos trabajos. En [68] Jiang y col. caracterizaron voces patológicas y normales utilizando la dimensión de correlación (ver subsección 2.2.2). Ellos concluyeron que el sistema vocal es caracterizado por una baja dimensión, tanto para voces sanas como patológicas. Este estudio y muchos otros abrieron las puertas para que otros investigadores propusieran detectores automáticos de patologías vocales, combinando medidas no lineales con teoría de reconocimiento de patrones y clasificación. Tal es el caso de Vaziri y col. [34] quienes, combinando características no lineales como la dimensión de correlación o el máximo exponente de Lyapunov (ver sección 2.4), con un clasificador basado en máquinas de soporte vectorial, lograron tasas de reconocimiento de patologías del 94.4 % y 88.89 % respectivamente. Esta línea de trabajo ha dado lugar a muchos más estudios como [69-71].

1.5 Conclusiones

En este capítulo repasamos algunos conceptos importantes en las teorías de las dinámicas no lineales y del caos. Se presentó cómo a partir del teorema de reconstrucción es posible reconstruir el atractor de un sistema a partir de la medición indirecta de una de sus variables. Gracias a esta herramienta podemos estudiar diferentes sistemas dinámicos, entre ellos los fisiológicos. En el siguiente capítulo revisaremos algunos conceptos que permiten caracterizar el atractor de un sistema dinámico. El resto de capítulos mostraremos los avances que esta tesis doctoral ha aportado a la teoría del análisis de las dinámicas no lineales y al estudio de las señales que los sistemas biológicos proveen.

A very small cause which escapes our notice determines a considerable effect that we cannot fail to see, and then we say that that effect is due to chance. If we knew exactly the laws of nature and the situation of the universe at the initial moment, we could predict exactly the situation of that same universe at a succeeding moment. But, even if it were the case that the natural laws had no longer any secret for us, we could still only know the initial situation approximately. It may happen that small differences in the initial conditions produce very great ones in the final phenomena. Prediction becomes impossible, and we have the fortuitous phenomenon.

Poincaré

2

Medidas de complejidad

Es materia de este capítulo el revisar los conceptos más importantes acerca de las diferentes medidas de complejidad definidas en la literatura. Describiremos tres tipos principales de invariantes. Comenzaremos por aquellos basados en conceptos geométricos, tales como las dimensiones. Luego, abordaremos aquellos que encuentran su fundamento en la teoría de la información, como por ejemplo las entropías. Por último, discutiremos los invariantes basados en el concepto de divergencia de trayectorias cercanas. También describiremos los algoritmos que permiten estimar estos invariantes a partir de series temporales.

2.1 Introducción

Como mencionamos en el capítulo [1](#), la dinámica de un sistema determinístico disipativo tiende a un conjunto de puntos llamado atractor del sistema [\[14\]](#). La estructura geométrica de este conjunto es materia de estudio, ya que refleja la complejidad de un sistema. Por ejemplo, la acción de la transformación triangular (transformación de la “Tienda” o *Tent map*):

$$x_{n+1} = \begin{cases} 2x_n & \text{si } 0 \leq x < 0.5 \\ 2 - 2x_n & \text{si } 0.5 \leq x \leq 1 \end{cases} \quad (2.1)$$

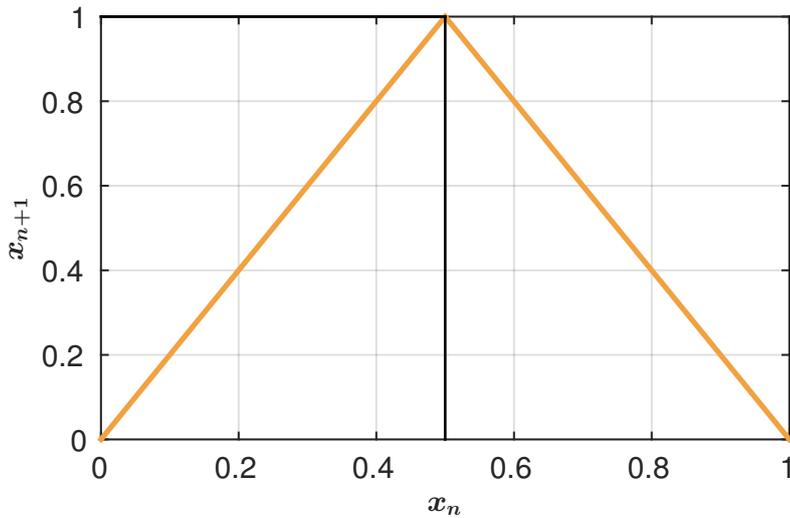


Figura 2.1. Transformación triangular.

sobre el intervalo $[0, 1]$ nos muestra el tipo de alteraciones geométricas esenciales para crear caos [17]: el estiramiento y el doblado sobre sí mismo. En la figura 2.1 se aprecia cómo esta transformación toma el intervalo $[0, 0.5]$ y lo alarga llevándolo al intervalo $[0, 1]$. También puede verse cómo el intervalo $[0.5, 1]$ es plegado en $x_n = 0.5$ y redirigido al intervalo $[0, 0.5]$. Las alteraciones de alargamiento derivan en sensibilidad a las condiciones iniciales y los pliegues son necesarios para mantener acotadas las soluciones. Estas dos propiedades le confieren a un conjunto de puntos el carácter de atractor.

Existen dos puntos de vista complementarios para estudiar la evolución de funciones o transformaciones sobre un atractor. El primero es un enfoque geométrico ya que mucha información sobre la dinámica se refleja en la dimensión del atractor. Por ejemplo un atractor cuya dimensión sea un número no entero es llamado fractal y su dinámica casi siempre es caótica. Por otro lado, la dimensión de un atractor también provee una cota para la cantidad de información necesaria para reconstruir el atractor y su dinámica.

Desde un punto de vista estadístico, la sensibilidad a las condiciones iniciales de un sistema caótico afecta la predictibilidad de valores de la dinámica distantes en el tiempo. Sin embargo, esta dinámica goza de cierta regularidad estadística basada en una única medida de probabilidad, físicamente relevante, sobre el atractor. Dicha medida de probabilidad es denominada medida natural del sistema y describe la frecuencia relativa con la cual la dinámica visita las diferentes partes del atractor, es decir, describe cómo la dinámica a medida que el tiempo avanza va poblando el

atractor. Por ejemplo, para uno muy complejo (con alta dimensión) la probabilidad de que una trayectoria visite una porción determinada de éste es muy baja, por tanto se deberá esperar mucho tiempo para que esto ocurra. En conclusión el estudio de la dinámica de un sistema puede ser simplificado al estudio de las propiedades de su medida natural. En este sentido es necesario formalizar un poco este concepto.

Una medida de probabilidad π sobre un conjunto $M \subseteq \mathbb{R}^d$ es invariante (estacionaria) con respecto a una función invertible $f: M \rightarrow M$ si para cada conjunto de Borel¹ $A \subseteq M$, $\pi(f(A)) = \pi(A)$ [17], es decir, la función f preserva una distribución estacionaria. Por ejemplo, si escogemos $X_0 \in M$ de manera aleatoria de una distribución π e iteramos $\{X_n = f^n(X_0), n \in \mathbb{Z}^+\}$ (f^n denota la composición de f con ella misma n veces), podemos observar que la distribución marginal de X_n será igual a π . Adicionalmente, la distribución conjunta de $\{X_{n_1}, X_{n_2}, \dots, X_{n_m}\}$ y de $\{X_{n_1+c}, X_{n_2+c}, \dots, X_{n_m+c}\}$, donde $m, c \in \mathbb{Z}^+$ será la misma. Esto nos dice que la distribución conjunta de un número finito de X_n es invariante ante corrimientos del índice temporal [17].

Además de ser estacionaria, π también puede ser ergódica con respecto a la dinámica f . Decimos que una medida de probabilidad estacionaria π es ergódica con respecto a f , si para cada conjunto invariante A , es decir $A = f^{-1}A$, $\pi(A) = 0$ o $\pi(A) = 1$ [74]. Gracias a esta propiedad, podemos afirmar que si π es ergódica y ϕ es una función tal que $\int |\phi(x)|\pi(x) dx < \infty$, entonces se cumple que [17, 75]:

$$\frac{1}{N} \sum_{n=0}^{N-1} \phi(X_n) = \int \phi(x)\pi(x) dx \quad \text{para } N \rightarrow \infty.$$

El lado izquierdo de esta ecuación es la media muestral del observable ϕ sobre la trayectoria X_n . La parte derecha es la media poblacional de ϕ bajo la distribución π [17]. En el contexto de un sistema dinámico, esto nos dice que la frecuencia relativa (asintóticamente hablando) con la que una trayectoria visita un conjunto de puntos A pertenecientes al atractor M es igual a la probabilidad $\pi(A)$ donde π es la medida de probabilidad asociada al atractor [17].

Otra consecuencia de esta ecuación es que si definimos ϕ como la función delta de

¹Un conjunto de Borel es un elemento de una σ -álgebra de Borel. Dado un conjunto X , una σ -álgebra de Borel es la σ -álgebra más pequeña que contiene todos los subconjuntos medibles de X [72, 73].

Dirac:

$$\phi = \delta_{X_n} = \begin{cases} 1 & \text{si } x = X_n \\ 0 & \text{en otro caso,} \end{cases}$$

entonces la distribución empírica dada por $\frac{1}{N} \sum \delta_{X_n}$ tomará gradualmente la forma de la distribución π .

Un atractor M puede admitir varias medidas de probabilidad ergódicas. Aún más, este tipo de medidas es común, y se obtienen fácilmente si la dinámica del sistema es continua. Sin embargo, de todas las posibles medidas ergódicas admisibles la más relevante es aquella que concuerda con las observaciones físicas (medida natural). Para definir dicha medida existen dos propuestas. Una planteada por Kolmogorov, en la cual, al sistema dinámico se le agrega una pequeña cantidad de ruido hasta cierto nivel, controlado por el parámetro ϵ . Para este sistema estocástico es casi seguro que exista una medida invariante única π_ϵ . Entonces la medida natural del sistema determinístico será el $\lim_{\epsilon \rightarrow 0} \pi_\epsilon$, si este existe y es único [17, 75]. La segunda propuesta está basada en la medida de Lebesgue. Este enfoque dice que si existe un conjunto W con medida de Lebesgue positiva entonces para todo $x \in W$ la distribución empírica inducida por la trayectoria que comienza en x converge a la medida natural del sistema [74].

En conclusión la medida natural de un sistema dinámico es una medida de probabilidad única, ergódica y físicamente observable. Esta es inherente al atractor del sistema y da cuenta de la frecuencia con que una trayectoria visita las diferentes regiones de éste.

Determinar la medida natural de un sistema puede ser una tarea difícil. Sin embargo puede ser caracterizada mediante varias cantidades, tales como las dimensiones y las entropías. Gracias al teorema de reconstrucción (ver sección 1.3) sabemos que estas cantidades son invariantes ante este procedimiento y, por tal motivo, son muy utilizadas para caracterizar sistemas dinámicos [11, 15]. En el resto de este capítulo revisaremos varias de estas medidas de complejidad.

2.2 Dimensiones

La complejidad de un sistema está reflejada parcialmente por su dimensión. De esta forma una dinámica proveniente de un sistema de alta dimensión es más compleja que aquella extraída de uno con baja dimensión. En esta sección se explorarán dos medidas de dimensión comúnmente utilizadas: la dimensión de conteo por cajas (en inglés “Box-Counting Dimension”) y la dimensión de correlación (D).

2.2.1 Dimensión de conteo por cajas

En la geometría Euclídea clásica un punto tiene dimensión cero, una línea tiene dimensión uno y un plano dimensión dos. Sin embargo, existen en la naturaleza y en la matemática objetos (conjuntos) cuya dimensión no es un número entero, estos los llamamos fractales. En la geometría clásica la dimensión de un objeto se traduce como la cantidad de coordenadas necesarias para posicionar un punto dentro del objeto [76]. Esto tiene sentido si la dimensión del objeto en estudio es un número entero. Para los fractales la definición de dimensión clásica no tiene mucho sentido. Por este motivo es necesario una definición de dimensión con la cual podamos caracterizar los conjuntos fractales y a su vez conservar la interpretación clásica para conjuntos con dimensión entera [14].

La definición que buscamos encuentra su base en el concepto de autosimilaridad que es una propiedad de los conjuntos fractales. Desde un punto de vista geométrico la palabra “similar” significa “diferente en tamaño o en posición mas no en forma”. Tomemos entonces un segmento de línea de tamaño l . Podemos pensar que este segmento está compuesto por 2 copias de él mismo pero de tamaño $l/2$ cada uno. Dos cantidades caracterizan esta forma de ver el segmento inicial como un segmento autosimilar, la primera el número de copias $N(\epsilon)$ y el segundo es el tamaño ϵ de estas copias. Podemos entonces escribir la siguiente relación [17]: $N(\epsilon) = l\epsilon^{-1}$. Ésta da cuenta del número de segmentos de tamaño ϵ necesarios para cubrir un segmento de tamaño l . Este concepto se puede extender fácilmente a objetos en cualquier dimensión como $N(\epsilon) \sim \epsilon^{-d_0}$, donde la constante de proporcionalidad dependerá de la forma del objeto en estudio. Si el $\lim_{\epsilon \rightarrow 0} -\frac{\ln N(\epsilon)}{\ln \epsilon}$ existe y es igual a d_0 entonces llamamos a d_0 la *dimensión de conteo por cajas* [17].

Consideremos el ejemplo anterior con un segmento de línea de tamaño $l = 1$. Vemos que si dividimos el segmento en dos partes se cumple que $N(1/2) = 2$. De forma general, si dividimos el segmento en k partes tendremos que $N(1/k) = k$. Se puede demostrar entonces que la dimensión por conteo de cajas para este segmento es $\lim_{\epsilon \rightarrow 0} -\frac{\ln k}{\ln 1/k} = 1$. La anterior deducción puede hacerse para formas geométricas clásicas como un cuadrado o un cubo, dando como resultado 2 y 3 respectivamente. Esto demuestra que esta definición de dimensión es congruente con el enfoque clásico.

Por otro lado consideremos un objeto fractal como el conjunto de Cantor. Este conjunto se construye tomando el segmento $[0, 1]$, dividiéndolo en 3 partes iguales y removiendo la parte central. Este proceso puede ser iterado dando como resultado el conjunto presentado en la figura 2.2.

Capítulo 2. Medidas de complejidad

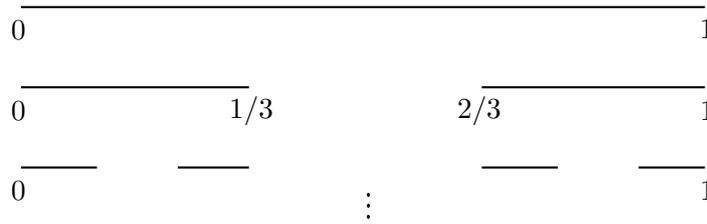


Figura 2.2. Conjunto de Cantor obtenido tomando el segmento $[0, 1]$, dividiéndolo en 3 partes iguales y removiendo la parte central.

Aplicando la anterior metodología, vemos cómo este conjunto se puede cubrir con dos segmentos $[0, 1/3] \cup [2/3, 1]$ de tamaño $1/3$. De esta misma forma lo podemos cubrir con 4 intervalos de tamaño $1/9$. De forma general podemos decir que el conjunto de Cantor puede ser cubierto con 2^m intervalos de tamaño 3^{-m} . En consecuencia $N(3^{-m}) = 2^m$ y su dimensión es igual a $d_0 = \lim_{\epsilon \rightarrow 0} \frac{\ln 2}{\ln 3} = 0.6309$. Como la dimensión calculada no es un número entero clasificamos al conjunto de Cantor como un objeto fractal. La propiedad de autosimilaridad hace fácil el cómputo de la dimensión de conteo por cajas para este tipo de conjuntos. No obstante para conjuntos más irregulares se debe utilizar la siguiente metodología [14]:

Algoritmo 1. Algoritmo de conteo por cajas.

- 1: Cubrir todo el objeto con cajas de tamaño ϵ_0 . Si el objeto es una imagen entonces las cajas serán cuadrados y si el objeto yace en un espacio tridimensional, entonces las cajas serán cubos. Luego se debe contar cuántas de estas cajas fueron necesarias para cubrir el objeto
- 2: Repetir el paso 1 utilizando cajas con tamaño $\epsilon_i = \epsilon_0/2^i$ con $i = 1, 2, \dots$ para formar la función $P(\epsilon)$.
- 3: Calcular:

$$d_0 = \frac{\ln[P(\epsilon_{i+1})/P(\epsilon_i)]}{\ln[\epsilon_{i+1}/\epsilon_i]}.$$

Las mayores dificultades de utilizar este algoritmo son su alto costo computacional, la necesidad de grandes cantidades de datos para obtener resultados fiables y la alta dependencia al valor del parámetro i [14].

2.2.2 Dimensión de correlación

Un problema de la dimensión de conteo por cajas es que estima la dimensión de un atractor de forma geométrica, lo que no brinda información sobre cómo la dinámica del sistema va poblando las diferentes partes del atractor. El concepto de dimensión de correlación suple parcialmente las dificultades presentadas por la

dimensión de conteo por cajas. Esta medida de dimensión está ligada a la medida natural del sistema, por tanto brinda información de su dinámica. A diferencia de la dimensión de conteo por cajas, puede ser calculada de manera eficiente y no requiere un número muy grande de datos para su convergencia.

Para describir la idea básica detrás de esta medida de dimensión es necesario considerar los vectores aleatorios \mathbf{x} e \mathbf{y} , cuya función de probabilidad es la medida natural π del sistema dinámico. Se define entonces la variable aleatoria $\tilde{z} = \|\mathbf{x} - \mathbf{y}\|$ como la distancia entre estos vectores. Comúnmente se utiliza la distancia euclídea o la distancia infinito. Generalmente para sistemas dinámicos caóticos se cumple que para $h > 0$ la $\Pr\{\tilde{z} \leq h\}$ se comportará como h^D . Por tanto, si

$$\lim_{h \rightarrow 0} \frac{\ln(\Pr\{\tilde{z} \leq h\})}{\ln(h)},$$

existe y es igual a D entonces llamamos a D la *dimensión de correlación* del atractor con respecto a la medida natural del sistema.

La dimensión de correlación se relaciona con el comportamiento local de la medida natural π del sistema. Sea $B_{\mathbf{x}}(h)$ la bola de radio h centrada en el vector de estado \mathbf{x} . Entonces:

$$\Pr\{\tilde{z} \leq h\} = \Pr\{\|\mathbf{x} - \mathbf{y}\| \leq h\} = E[\pi(B_{\mathbf{x}}(h))] \sim h^D,$$

donde $E[\cdot]$ denota el valor esperado. Como D es el exponente en la ley de escalamiento del primer momento de $\pi(B_{\mathbf{x}}(h))$, entonces un valor alto de dimensión de correlación implica que al sistema le tomará un largo tiempo (en promedio) visitar una determinada región del atractor. Por tanto, una dimensión alta indica que la complejidad del sistema es alta [17].

Todos los aspectos importantes en la estimación de esta medida de dimensión a partir de una serie temporal serán tratados en la sección 2.5.

2.3 Entropía

La entropía de un sistema dinámico es un invariante que está estrechamente ligado con la cantidad de información que contiene una serie temporal acerca del sistema del cual proviene. Motivado en parte por el trabajo de Shannon sobre la teoría de la información en la década de los 40, Kolmogorov propuso en 1958 una extensión de esta teoría para sistemas dinámicos ergódicos [77, 78]. La entropía de un sistema se define como la tasa de generación de nueva información o la tasa de destrucción de vieja información. Es decir, la entropía es una medida de cuan rápido la informa-

ción relacionada con un estado del sistema se hace irrelevante para predecir futuros estados [76].

Para explicar esto consideremos las trayectorias $\{x_n\}$ y $\{y_n\}$ de un sistema dinámico que comienzan cerca y permanecen cerca. Bajo esta condición, la información que se puede obtener sobre el estado $\{x_{n+i}\}$ a partir del estado $\{y_{n+i}\}$ es aproximadamente constante [15]. Por el contrario, para sistemas con sensibilidad a las condiciones iniciales, las trayectorias $\{x_n\}$ y $\{y_n\}$ divergen rápidamente. Esto significa que la información compartida entre trayectorias decae a una alta velocidad a medida que el tiempo avanza. Otra forma de ver el mismo fenómeno es considerar que a medida que avanza el tiempo se está generando nueva información [15]. Existe una conexión entre el concepto de entropía usado en el contexto de las dinámicas no lineales y la termodinámica. En ambos contextos la entropía está relacionada con la información. En termodinámica es utilizado para cuantificar la cantidad de desorden de un sistema, es decir, la cantidad de información necesaria para especificar con cierta precisión un estado del sistema, lo que concuerda con la definición de entropía para sistemas dinámicos [11].

2.3.1 Entropía de Kolmogorov-Sinai

Como se mencionó antes, la entropía de un sistema dinámico guarda relación con la información $I(h, T)$ obtenida al observar una trayectoria del sistema con precisión h durante un tiempo T . Se define entonces la entropía de Kolmogorov-Sinai como [79]:

$$K = \lim_{h \rightarrow 0} \lim_{T \rightarrow \infty} \frac{I(h, T)}{T}.$$

La precisión h también representa la incertidumbre en la medida de alguna de las coordenadas del vector de estado. Observar que en la anterior ecuación el límite temporal se toma antes que el límite relacionado con la precisión. Para sistemas caóticos, dada una precisión fija lo suficientemente cercana a 0, K será una constante mayor que cero, indicando que $I(h, T)$ crece linealmente a medida que $T \rightarrow \infty$. Para sistemas determinísticos no caóticos $K = 0$ y para sistemas puramente estocásticos $K \rightarrow \infty$.

Con el fin de calcular K es necesario encontrar una forma de cuantificar la información $I(h, T)$. Esta será definida mediante una partición del espacio-tiempo de la serie temporal en cajas de lado h , tal como se muestra en la figura 2.3. Se define entonces $P_{\{i_1, i_2, \dots, i_T\}}$ como la probabilidad conjunta de que una solución esté en la caja i_1 para un tiempo $t = \tau$, en la caja i_2 para un tiempo $t = 2\tau$ y en la caja i_T para un tiempo $t = T\tau$. A partir de esta probabilidad podemos definir la entropía de orden

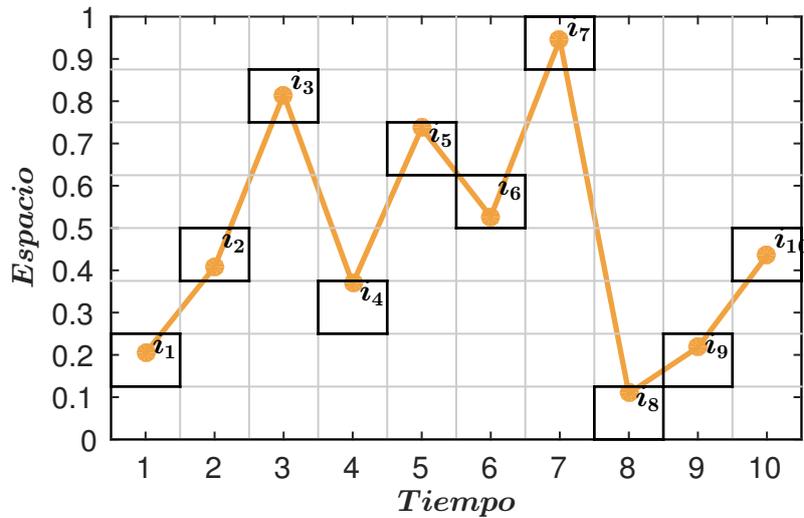


Figura 2.3. Partición del espacio-tiempo para una solución de la transformación triangular. Imagen reproducida de [79].

q como [79]:

$$K_q = \lim_{\tau \rightarrow 0} \lim_{h \rightarrow 0} \lim_{T \rightarrow \infty} \frac{1}{\tau T} \frac{1}{1-q} \ln \left(\sum_{i_1, i_2, \dots, i_T} P_{\{i_1, i_2, \dots, i_T\}}^q \right). \quad (2.2)$$

Para $q \geq 0$, esta entropía es invariante ante transformaciones suaves como el proceso de inmersión. Además es una función que decrece de forma monótona con respecto a q . La entropía de Kolmogorov-Sinai puede ser calculada como $K = \lim_{q \rightarrow 1} K_q$. El mayor inconveniente para estimar K es que, dado un sistema dinámico, su cálculo requiere observar un número grande de trayectorias con resolución y precisión infinitesimales por un periodo muy largo de tiempo. Esto es computacionalmente muy costoso. Más aún, la presencia de una pequeña cantidad de ruido en las series temporales puede distorsionar los resultados. Existen varias simplificaciones de esta entropía. Una en particular ha sido ampliamente utilizada en varios campos de las ciencia y por eso le dedicaremos la sección 2.6.

2.3.2 Entropía de correlación

Existe otro caso particular de la entropía generalizada cuyo cómputo puede realizarse de una manera más simple que la entropía de Kolmogorov-Sinai. Para $q = 2$ la ecuación 2.2 se puede escribir como [80]:

$$K_2 = \lim_{\tau \rightarrow 0} \lim_{h \rightarrow 0} \lim_{T \rightarrow \infty} \frac{1}{\tau T} \ln \left(\sum_{i_1, i_2, \dots, i_T} P_{\{i_1, i_2, \dots, i_T\}}^2 \right).$$

Esta entropía es llamada entropía de correlación (K_2). $P_{\{i_1, i_2, \dots, i_T\}}^2$ es el cuadrado de la probabilidad de que la solución del sistema visite una partición determinada del espacio-tiempo. Lo que hace que esta entropía pueda ser calculada fácilmente es que, para un h dado, la probabilidad $P_{\{i_1, i_2, \dots, i_T\}}^2$ es aproximadamente igual a la $\Pr\{\|\mathbf{x}_i - \mathbf{x}_j\| \leq h\}$ donde \mathbf{x}_i y $\mathbf{x}_j \in \mathbb{R}^T$ son vectores de estado formados a partir de los segmentos de solución x_{i+k} e x_{j+k} con $k = 1, 2, \dots, T$. En otras palabras, el cuadrado de la probabilidad de que dos segmentos de solución pasen por la misma partición del espacio-tiempo es aproximadamente igual a la probabilidad de que los vectores de estado formados a partir de ellos sean vecinos en el espacio de estados. Es decir que el problema se reduce a contar vecinos en el espacio de estados [79].

Como se mencionó antes, la entropía generalizada K_q es una función que decrece de forma monótona para q creciente. Esto implica que $K > K_2$ y por tanto un valor de $0 < K_2 < \infty$ implica que la serie temporal proviene de un sistema dinámico caótico.

2.4 Divergencia entre trayectorias: Exponentes de Lyapunov

Los exponentes de Lyapunov son una clase de invariantes relacionados con la evolución dinámica de las trayectorias de un sistema. Estos son una medida indirecta del error de predicción no lineal, ya que cuantifican la divergencia (o convergencia) de trayectorias que en un principio son cercanas. Para un sistema d -dimensional, existen d exponentes de Lyapunov $\lambda_1, \lambda_2, \dots, \lambda_d$ [15]. Con el fin de ilustrar el fundamento de los exponentes de Lyapunov consideremos una transformación unidimensional $y = f(x)$. Esta transformación al ser iterada producirá la serie temporal $x_{n+1} = f(x_n)$, con $t = 1, 2, \dots$. Ahora consideremos dos soluciones del sistema, con condiciones iniciales x_0 y $x = x_0 + dx$ que difieren infinitesimalmente. Podemos calcular la diferencia entre las soluciones después de una iteración como [17]:

$$\begin{aligned} dx_1 &= f(x) - f(x_0) \\ &= [f(x) - f(x_0)] \frac{x - x_0}{x - x_0} \\ &= f'(x_0) dx, \end{aligned} \tag{2.3}$$

donde f' indica la primera derivada de f respecto a x . Utilizando la regla de la cadena es posible demostrar que la diferencia entre las soluciones después de m ite-

2.4. Divergencia entre trayectorias: Exponentes de Lyapunov

raciones es:

$$dx_m = f'(x_0)f'(x_1)\dots f'(x_{m-1})dx.$$

Esta ecuación indica que la diferencia inicial entre soluciones dx es amplificada por un factor igual al producto $f'(x_0)f'(x_1)\dots f'(x_{m-1})$. Es más conveniente estudiar el promedio del logaritmo natural de la magnitud de este producto, es decir $\sum_{t=0}^{m-1} \ln|f'(x_t)|/m$. Esta operación equivale a calcular el valor esperado del $\ln|f'(x_t)|$ a lo largo de la trayectoria. Por tanto, para un sistema ergódico con medida natural $\pi(x)$ vemos que:

$$\sum_{t=0}^{m-1} \ln|f'(x_t)|/m \rightarrow \lambda = \int \ln|f'(x)|\pi(x) dx.$$

Tomemos como ejemplo la transformación triangular. De la ecuación (2.1) podemos ver que su primera derivada es $f'(x) = 2$ o $f'(x) = -2$ (no está definida para $x = 0.5$). Teniendo en cuenta la discusión anterior es posible deducir que para este sistema el exponente de Lyapunov es $\lambda = \sum_{t=0}^{m-1} \ln|f'(x_t)|/m = \ln 2$. Por tanto, si dos trayectorias cercanas divergen con una velocidad de $e^{\lambda m}$, entonces la cantidad de iteraciones m necesarias para que $dx_m = 2dx$ es $m = \ln(2)/\lambda = 1$. En otras palabras, para esta transformación sólo se necesita de una iteración para que dos trayectorias cercanas se encuentren al doble de su distancia inicial. Esto indica que errores pequeños de medición en la condición inicial pueden ser magnificados en pocas iteraciones de la transformación. Una metodología similar puede ser aplicada para calcular los exponentes de Lyapunov para sistemas de mayor dimensión [15, 17].

Es importante resaltar que los exponentes de Lyapunov no dependen de las condiciones iniciales elegidas. Más importante aún es el hecho de que un sistema dinámico cerrado y acotado es llamado caótico si y sólo si tiene al menos un exponente de Lyapunov positivo [17]. Entonces, para sistemas con sensibilidad a las condiciones iniciales existe por lo menos un exponente positivo. Sistemas con más de un exponente positivo los llamamos hipercaóticos [15].

La suma de todos los exponentes de Lyapunov es una medida de la expansión o contracción de un volumen finito bajo la acción de la dinámica del sistema. Por tanto, para sistemas disipativos esta suma es negativa y para los conservativos es igual a cero. Es de resaltar que, si el sistema es continuo, por lo menos uno de sus exponentes debe ser igual a cero. Esto implica que para estos sistemas el caos requiere de por lo menos de tres dimensiones, para asegurar que el sistema sea disipativo [15].

Una propiedad muy útil del espectro de Lyapunov $(\lambda_1, \lambda_2, \dots, \lambda_d)$ es que puede ser-

vir para calcular la entropía de Kolgomorov-Sinai mediante la identidad de Pesin [76]:

$$K = \sum_{i=1}^d \lambda_i \quad \text{si } \lambda_i > 0.$$

Esta relación entre la entropía de Kolgomorov-Sinai y los exponentes de Lyapunov era esperable ya que ambos miden la tasa con la que se crea nueva información.

2.5 Integral de Correlación

La integral de correlación es la función de la cual se pueden estimar la D y K_2 de un sistema dinámico. La definición de la integral de correlación involucra en primer lugar la reconstrucción del espacio de estados del sistema a partir de una serie temporal (ver sección 1.3). Estos vectores de estado pueden ser interpretados como muestras aleatorias tomadas a partir de la distribución de probabilidad ligada al atractor del sistema (medida natural del sistema) [12]. Dados dos vectores aleatorios \mathbf{x} e $\mathbf{y} \in \mathbb{R}^m$ distribuidos conforme a la medida natural $f_{\mathbf{x}}^{(m)}$ y un parámetro $h \in \mathbb{R}^+$ se puede definir, de forma general, la integral de correlación $G_m(h)$ como [16]:

$$\begin{aligned} G_m(h) &= \int \int g(\|\mathbf{x}-\mathbf{y}\|/h) f_{\mathbf{x}}^{(m)}(\mathbf{x}) f_{\mathbf{x}}^{(m)}(\mathbf{y}) d\mathbf{x} d\mathbf{y} \\ &= \int g(\bar{z}/h) f_{\bar{z}}^{(m)}(\bar{z}) d\bar{z}, \end{aligned} \quad (2.4)$$

donde $\bar{z} = \|\mathbf{x}-\mathbf{y}\|$ es una medida de distancia entre los vectores aleatorios \mathbf{x} e \mathbf{y} , $f_{\bar{z}}^{(m)}(\bar{z})$ su función de distribución de probabilidad (FDP) y m es la dimensión de inmersión. Puede verse en la ecuación (2.4) que la integral de correlación puede interpretarse como el valor esperado de una función $g(\bar{z}; h)$ que llamaremos función núcleo.

La integral de correlación propuesta por Grassberger y Procaccia $C_m(h)$ (estándar) puede obtenerse a partir de la ecuación (2.4) si utilizamos como núcleo la función escalón $H(1 - \bar{z}/h)$ [81, 82]:

$$C_m(h) = \int H(1 - \bar{z}/h) f_{\bar{z}}^{(m)}(\bar{z}) d\bar{z}. \quad (2.5)$$

Para sistemas dinámicos caóticos en ausencia de ruido se comportará como:

$$C_m(h) = \phi e^{-\tau m K_2} h^D \quad \text{para } m \rightarrow \infty, h \rightarrow 0, \quad (2.6)$$

donde $\phi \in \mathbb{R}$ es una constante de normalización y τ es el retardo de inmersión. Partiendo de la ecuación (2.6) es posible determinar que la dimensión de correlación y la entropía de correlación pueden ser calculadas mediante los funcionales:

$$\frac{d \ln C_m(h)}{d \ln h} = D \quad (2.7)$$

y

$$-\frac{1}{\tau} \ln \frac{C_{m+1}(h)}{C_m(h)} = K_2, \quad (2.8)$$

respectivamente.

Un resultado muy importante fue obtenido por Ghez y col. en [83]. Ellos demostraron que si una función $g(t)$ cumple con dos propiedades, entonces la integral de correlación que la use como función núcleo escalará como indica la ecuación (2.6). Estas propiedades son: (i) $g(t)$ debe decrecer de forma monótona para $t > 0$ y (ii) ésta debe hacerlo más rápido que una potencia, es decir:

$$\lim_{h \rightarrow 0} h^{-p} g(t/h) = 0 \quad \text{para } t > 0 \text{ y cualquier } p \geq 0. \quad (2.9)$$

Esto nos abre la posibilidad para explorar otras funciones núcleo además de la función escalón.

En aplicaciones reales, debido a la presencia de ruido o al número finito de datos en la serie temporal, la estimación de invariantes siempre está sesgada (se desvía del régimen de escala de la ecuación (2.6)). A pesar de que la integral de correlación estándar (ecuación (2.5)) estima correctamente los invariantes de atractores de baja dimensión, es prácticamente inútil cuando las series temporales provenientes de estos sistemas están inmersas en ruido [11, 15, 16, 84, 85]. Esto se debe a que las hipótesis de esta metodología involucran que el sistema observado tiene una dimensión finita y sin presencia de ruido [15].

Cuando la desviación de la ley de escalas se da por causa de ruido, el investigador tiene dos opciones. La primera es intentar reducir el nivel de ruido (σ) presente en la señal. Esta actividad no siempre es posible, además, un nivel de ruido muy pequeño (2%) puede afectar de forma dramática la estimación de invariantes mediante la integral de correlación estándar [86]. La segunda opción es intentar modelar la influencia que tiene σ en la integral de correlación. Desafortunadamente, este tipo de análisis es particularmente difícil para la integral de correlación estándar, debido a que su función núcleo (función escalón) no es derivable [16].

Capítulo 2. Medidas de complejidad

Como alternativa, Diks propone cambiar la función núcleo por $g(\tilde{z}; h) = e^{-\tilde{z}^2/4h^2}$ dando lugar a la integral de correlación de núcleo Gaussiano (ICG) [16, 87]:

$$T_m(h) = \int e^{-\tilde{z}^2/4h^2} f_{\tilde{z}}^{(m)}(\tilde{z}) d\tilde{z} \quad (2.10)$$

donde, a diferencia de la ecuación (2.5), $f_{\tilde{z}}^{(m)}(\tilde{z})$ denota la distribución de la distancia entre vectores de estado reconstruidos en presencia de ruido blanco Gaussiano de media cero varianza σ^2 . Por tanto, esta integral de correlación escalará como [16, 87, 88]:

$$T_m(h) = \phi h^m (h^2 + \sigma^2)^{\frac{D-m}{2}} e^{-\tau m K_2} \quad \text{para } m \rightarrow \infty, \sqrt{h^2 + \sigma^2} \rightarrow 0. \quad (2.11)$$

Se puede ver que en la ecuación (2.11) está contemplada la acción del ruido en el cálculo de esta integral de correlación. Además, si $\sigma \rightarrow 0$, la ICG escala de igual forma que la integral de correlación estándar (ecuación (2.6)), es decir, $\lim_{\sigma \rightarrow 0} T_m(h) = C_m(h)$. En la práctica σ es un valor fijo distinto de cero, por tanto la expresión $\sqrt{h^2 + \sigma^2}$ en la ecuación (2.11) no puede tender a cero. Sin embargo se espera que dicha ecuación tenga validez en un rango valores de h cercano a cero, contando con que el nivel de ruido no es muy grande.

De forma similar a las ecuaciones (2.7) y (2.8) podemos definir funcionales para D y K_2 a partir de la ICG. De la ecuación (2.11), tenemos que el funcional para la dimensión de correlación se puede escribir como [85]:

$$\begin{aligned} \frac{d \ln T_m(h)}{d \ln h} &= h \frac{\frac{d}{dh} T_m(h)}{T_m(h)} \\ &= h \left(\frac{m}{h} + \frac{h^2(D-m)}{h^2 + \sigma^2} \right) \\ &= \frac{m\sigma^2 + Dh^2}{h^2 + \sigma^2}, \end{aligned} \quad (2.12)$$

y el funcional para la entropía de correlación como [85]:

$$\begin{aligned} -\frac{1}{\tau} \ln \frac{T_{m+1}(h)}{T_m(h)} &= -\frac{1}{\tau} \ln \left[h (h^2 + \sigma^2)^{-1/2} e^{-\tau K_2} \right] \\ &= \frac{1}{2\tau} \ln \left[\frac{h^2 + \sigma^2}{h^2} \right] + K_2. \end{aligned} \quad (2.13)$$

Se puede observar que los funcionales de dimensión (ecuación (2.12)) y de entropía

de correlación (ecuación [2.13](#)) presentan un sesgo que depende del valor de nivel de ruido. Si $\sigma \rightarrow 0$, entonces estos funcionales tienden a \mathbf{D} y a \mathbf{K}_2 respectivamente. Sin embargo, para valores altos de ruido, $\sigma \gg h$, el funcional de dimensión de correlación tiende a m (la dimensión de inmersión) y el funcional de entropía de correlación tiende a sobrestimar \mathbf{K}_2 .

De igual forma que existen funcionales para \mathbf{D} y \mathbf{K}_2 , es posible definir un funcional para σ a partir de la integral de correlación de núcleo Gaussiano. Schreiber encontró que, para valores de m grandes, σ puede ser estimado mediante el funcional de nivel de ruido [\[84\]](#):

$$\begin{aligned} \Delta_m^T(h) &= \frac{d \ln T_{m+1}(h)}{d \ln h} - \frac{d \ln T_m(h)}{d \ln h} \\ &= \frac{\sigma^2}{\sigma^2 + h^2}, \end{aligned} \tag{2.14}$$

donde el superíndice T denota que este funcional se calcula a partir de $T_m(h)$. Podemos observar que $\Delta_m^T(h) \rightarrow 1$ para $h \rightarrow 0$ y $\Delta_m^T(h) \rightarrow 0$ para $h \rightarrow \infty$. Es decir, este funcional decrece de forma monótona de 1 a 0 a una velocidad proporcional a σ .

Es importante aclarar que para las simulaciones correspondientes a la integral de correlación, en este trabajo de tesis, las series temporales utilizadas han sido normalizadas para tener varianza unitaria. Por tanto σ es el nivel de ruido después de la normalización, es decir:

$$\sigma = \frac{\sigma_\eta}{\sqrt{\sigma_s^2 + \sigma_\eta^2}},$$

donde σ_η es la varianza del ruido y σ_s es la varianza de la serie temporal limpia. En consecuencia un $\sigma = 0$ correspondería a una serie temporal limpia y un $\sigma = 1$ implicaría una serie temporal compuesta solo por muestras de ruido idénticamente distribuidas.

Podemos establecer una concordancia entre σ_η y la relación señal a ruido. Partiendo de la ecuación [\(2.5\)](#) podemos establecer que:

$$\text{RSR} = 10 \log_{10} \left(\frac{1 - \sigma^2}{\sigma^2} \right) \text{dB},$$

en este sentido un nivel de ruido $\sigma = 0.2$ equivale a una RSN ≈ 13.8 dB.

La ICG estima de forma más precisa los invariantes \mathbf{D} y \mathbf{K}_2 que la integral de correlación estándar [\[15\]](#). Esto es porque modela el efecto que el nivel de ruido tiene sobre la integral de correlación y los funcionales derivados de ésta. Además, poder

estimar el nivel de ruido es de vital importancia a la hora de concluir sobre los resultados de una investigación. Lo anterior es consecuencia de que el valor de σ sesga la estimación de los otros dos invariantes. Por tanto, en presencia de mucho ruido se debe tener cuidado en la interpretación de los resultados. Más aún, el valor de σ debe ser reportado para que la comunidad científica pueda juzgar las conclusiones de una investigación.

Es de resaltar que aunque la ICG es mucho mejor que la integral de correlación estándar, ésta presenta serios problemas en la estimación de la entropía de correlación [15]. Estos problemas radican en que es necesario un valor de dimensión de inmersión muy alto para poder converger, incluso en ausencia de ruido [15]. Un valor de dimensión de inmersión muy alto implica que se necesitan muchos datos para poder de estimar de forma adecuada la integral de correlación. Por tanto, para series temporales cortas, la estimación precisa de K_2 es una tarea casi imposible.

Como comentario final debemos decir que existen en la literatura otras expresiones analíticas que modelan la influencia del ruido tanto en la integral de correlación estándar como en la ICG. De estas expresiones se han deducido otro tipo de funcionales para D , K_2 y σ . Por ejemplo, para la integral de correlación estándar se puede consultar [89, 90] y para la ICG [11, 91, 92].

2.5.1 Estimadores dependientes de la escala

Para estimar los invariantes D , K_2 y σ de un sistema dinámico, es necesario elegir un tipo de integral de correlación y su modelo de escala. Luego se debe estimar la integral de correlación utilizando el algoritmo de correlación (ver subsección 2.5.2). Una vez obtenidos los perfiles de la integral de correlación se cuenta con dos opciones. La primera es encontrar los valores de D , K_2 y σ haciendo una regresión no lineal sobre el modelo elegido [16, 87]. Esta metodología tiene como desventaja que es altamente dependiente del rango de escalas elegido para la regresión, además no existe un consenso general de cómo elegir este rango.

La segunda opción es utilizar estimadores dependientes de la escala (EsDE). Estos son expresiones analíticas de D , K_2 y σ en función de h y m [11]. Los EsDE son muy útiles para estudiar integrales de correlación en casos donde éstas solo muestran un escalamiento aproximado, típico en casos con presencia de ruido. La desviación del régimen de escalamiento puede ser fácilmente observada en la evolución de estos estimadores como funciones de h . Se han reportado en la literatura varios estimadores de este tipo [11, 89-92], sin embargo se destacan aquellos EsDE propuestos

por Nolte y col. [88]:

$$D_m^T(h) = \frac{d \ln T_m(h)}{d \ln h} - \frac{\left[m - \frac{d \ln T_m(h)}{d \ln h} \right] \Delta_m^T(h)}{1 - \Delta_m^T(h)}, \quad (2.15)$$

$$K_m^T(h) = \frac{1}{\tau} \left(\ln \frac{T_m(h)}{T_{m+1}(h)} + \frac{1}{2} \ln(1 - \Delta_m^T(h)) \right), \quad (2.16)$$

y

$$\sigma_m^T(h) = \sqrt{\frac{h^2 \Delta_m^T(h)}{1 - \Delta_m^T(h)}}, \quad (2.17)$$

donde $D_m^T(h)$, $K_m^T(h)$ y $\sigma_m^T(h)$ son los estimadores dependientes de la escala para \mathbf{D} , \mathbf{K}_2 y $\boldsymbol{\sigma}$ respectivamente. $\Delta_m^T(h)$ está dado por la ecuación (2.14). Debemos aclarar que el superíndice T denota que estos estimadores son calculados en base a $T_m(h)$.

Estos EsDE tienen la ventaja de que no requieren grandes recursos computacionales. Además, dependen sólo de integrales de correlación de núcleo Gaussiano calculadas para valores consecutivos de m . Es decir, no necesitan la manipulación de ningún parámetro externo por parte del usuario para su estimación. Por estos méritos los hemos elegido para comparar nuestros resultados.

2.5.2 Algoritmo de correlación

La integral de correlación se estima utilizando la suma de correlación $\widehat{G}_m(h)$. Dada una serie temporal $\{x_n\}_{n=1}^N$ y los parámetros $m \in \mathbb{N}$ y $\tau \in \mathbb{N}$ se deben formar los vectores de estado m -dimensionales $\mathbf{x}_i = \{x_i, x_{i+\tau}, x_{i+2\tau}, \dots, x_{i+(m-1)\tau}\}$ con $1 \leq i \leq L = N - (m-1)\tau$. Se define entonces la distancia euclídea entre estos vectores como: $\tilde{z}_\omega = \|\mathbf{x}_i - \mathbf{x}_j\|$, donde $\omega = \{(i, j) / i \neq j, i = 1, 2, \dots, L, j = 1, 2, \dots, L\}$ y $Q = L(L-1)$ (número total de distancias entre vectores de estado). Entonces la suma de correlación puede calcularse como [16]:

$$\begin{aligned} \widehat{G}_m(h) &= \frac{1}{L(L-1)} \sum_{i=1}^L \sum_{\substack{j=1 \\ j \neq i}}^L g(\|\mathbf{x}_i - \mathbf{x}_j\|/h) \\ &= \frac{1}{Q} \sum_{\omega=1}^Q g(\tilde{z}_\omega/h), \end{aligned} \quad (2.18)$$

donde $g(t)$ es la función núcleo. Es importante notar que la segunda sumatoria en la primera ecuación no cuenta los valores $j = i$. Esto quiere decir que la distancia de un vector consigo mismo no es tenida en cuenta en el cálculo de la suma de co-

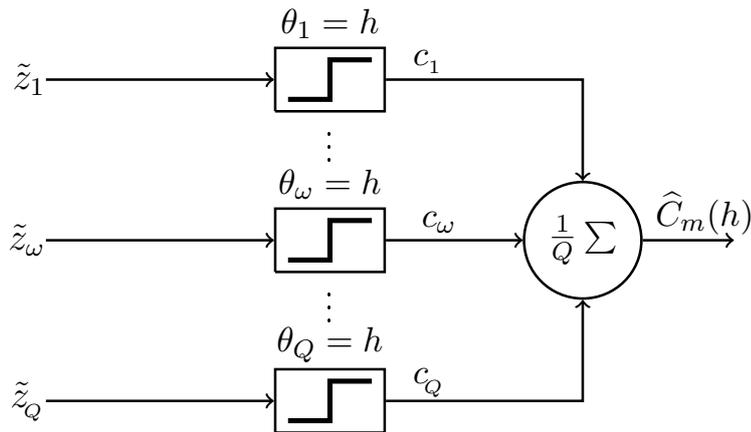


Figura 2.4. Algoritmo de correlación: Las distancias entre todos los vectores de estado \tilde{z}_ω son comparadas contra un umbral de valor $\theta = h$ para producir una salida binaria c_ω , estas son luego promediadas para producir la suma de correlación $\hat{C}_m(h)$.

rrelación.

Este procedimiento puede ser traducido en un algoritmo llamado “algoritmo de correlación” que consiste en cinco etapas. La primera es la conformación de los vectores de estado m -dimensionales a partir de la serie temporal. La segunda involucra el cálculo de la distancia entre cada uno de estos vectores. La tercera etapa requiere obtener el valor de la función núcleo $g(\tilde{z}_\omega/h)$ para cada valor de ω con h fijo. A continuación, la suma de correlación se calcula como el promedio de los valores de la función núcleo a través ω . Finalmente los cuatro primeros pasos deben repetirse para diferentes valores de h .

La suma de correlación es un estimador de la integral de correlación, la cual es un funcional de la medida natural del sistema en estudio. A partir de una muestra de vectores de estado m -dimensionales finita, dicho estimador será insesgado sólo si los vectores de estado están descorrelacionados (respecto a la medida natural) en el espacio de inmersión (de dimensión m). Los datos provenientes de la serie temporal, de la cual se forman los vectores de estado, están correlacionados temporalmente. Esto entra en contradicción con el requerimiento antes mencionado. Sin embargo, si dicha correlación es “débil” como en el caso de señales provenientes de un sistema altamente caótico (con un exponente de Lyapunov grande), el sesgo introducido en la estimación de la integral de correlación será mínimo. Por el contrario, si esta misma serie temporal se muestrea con una frecuencia de muestreo (f_m) muy alta (en comparación con la tasa de divergencia del sistema), la correlación temporal entre muestras puede tener una influencia catastrófica en la estimación

Algoritmo 2. Algoritmo de correlación de Grassberger y Procaccia.

- 1: Formar los vectores de estado m -dimensionales a partir de la serie temporal y calcular todas las distancias \tilde{z}_ω con $1 \leq \omega \leq Q = L(L-1)$ entre pares de vectores de estado.
- 2: Fijar el valor de h y calcular la variable binaria:

$$c_\omega(h) = \begin{cases} 0, & \text{si } \tilde{z}_\omega - h \geq 0, \\ 1, & \text{en otro caso.} \end{cases}$$

- 3: Obtener la suma de correlación como:

$$\widehat{C}_m(h) = \frac{1}{Q} \sum_{\omega=1}^Q c_\omega(h).$$

- 4: Repetir los pasos 1–4 para todos los valores de h y m .

de la integral de correlación. Para solucionar este inconveniente basta con conocer el tiempo máximo t_{max} para el cual los datos están correlacionados temporalmente y rechazar los $P = 2t_{max}f_m$ vecinos más cercanos a cada vector de estado al calcular la suma de correlación clásica.

Un caso muy especial de este algoritmo ocurre cuando la función núcleo es la función escalón. Esta versión del algoritmo de correlación fue propuesta por Grassberger y Procaccia [82] (ver algoritmo [2]). Éste busca comparar cada una de las distancias \tilde{z}_ω con un umbral $\theta = h$ para producir una salida binaria c_ω . Luego estas salidas son promediadas para producir la suma de correlación $\widehat{C}_m(h)$. Es importante resaltar que esta versión del algoritmo de correlación puede ser vista como un conjunto de comparadores en paralelo (ver figura [2.4]). Esta configuración se asemeja al modelo de funcionamiento del bien conocido conversor análogo-digital (CAD).

Con el fin de ilustrar el funcionamiento de la suma de correlación y cómo puede ser utilizada para la estimación de los distintos invariantes, hemos hecho algunas simulaciones utilizando una serie temporal proveniente de la transformación triangular. Como modelo hemos elegimos la integral de correlación de núcleo Gaussiano.

En la figura [2.5] se muestra el logaritmo de la ICG en función del $\ln h$ para varios valores de m y $\tau = 1$. Esta fue calculada a partir de una serie temporal en ausencia de ruido con la cual se formaron $L = 5000$ vectores de estado. Es posible observar que las curvas para diferentes valores de m presentan todas la misma pendiente para valores de $\ln h \leq -3$. La pendiente de estas porciones de recta es la dimen-

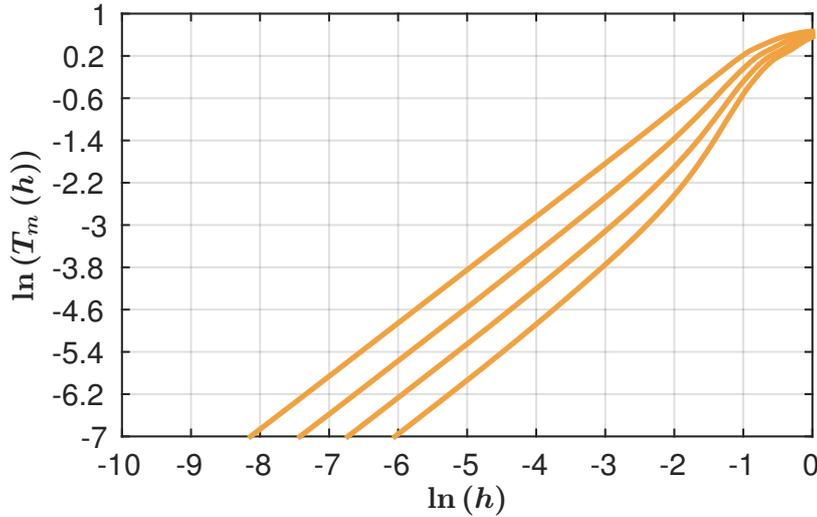


Figura 2.5. Gráfica log-log de la integral de correlación para la transformación triangular sin ruido. Se presentan las curvas calculadas con $m = \{2, 3, 4, 5\}$ ($m = 2$ curva superior y $m = 5$ curva inferior) para la ICG.

sión de correlación. Lo anterior se puede verificar observando la figura 2.6 donde se presenta el funcional para D (ecuación (2.12)) versus el $\ln h$. Se aprecia que para valores de h pequeños esta derivada oscila alrededor del valor teórico de la dimensión de correlación $D = 1$ [93]. Al avanzar sobre h las curvas para diferentes valores de m tienden a ser aproximadamente constantes, esto es evidencia de la presencia de una regla de escalamiento. Finalmente estas curvas comienzan a diverger para valores grandes de h .

La estimación de K_2 puede observarse en la figura 2.7 donde se muestra el funcional de entropía de correlación (ecuación 2.13) de la ICG, calculado para varios valores de m y en función del $\ln h$. Es fácil observar cómo este funcional se aproxima al valor teórico de la entropía de correlación de esta transformación $K_2 = \ln 2$ [17] para el rango de valores $-8 \leq \ln h \leq -5$, salvo para $m = 2$ (curva superior).

En la figura 2.8 se presenta el comportamiento de la ICG en presencia de ruido. Se muestra una gráfica del logaritmo de la ICG en función del $\ln h$ para $m = \{2, 3, 4, 5\}$. Esta integral de correlación fue calculada a partir de una serie temporal proveniente de la transformación triangular con un nivel de ruido $\sigma = 0.05$. Es posible observar cómo la ICG presenta diferentes valores de pendiente al avanzar sobre h . Para valores chicos de h ($\ln h \leq -2$) la pendiente de la ICG tiene valor elevado. Esta es una característica de la presencia del ruido, que domina las bajas escalas de la dinámica y cuya dimensión tiende a la dimensión de inmersión m . Sin embargo, para un rango de valores de h ($-1.8 \leq \ln h \leq -0.4$) aun se puede observar que las

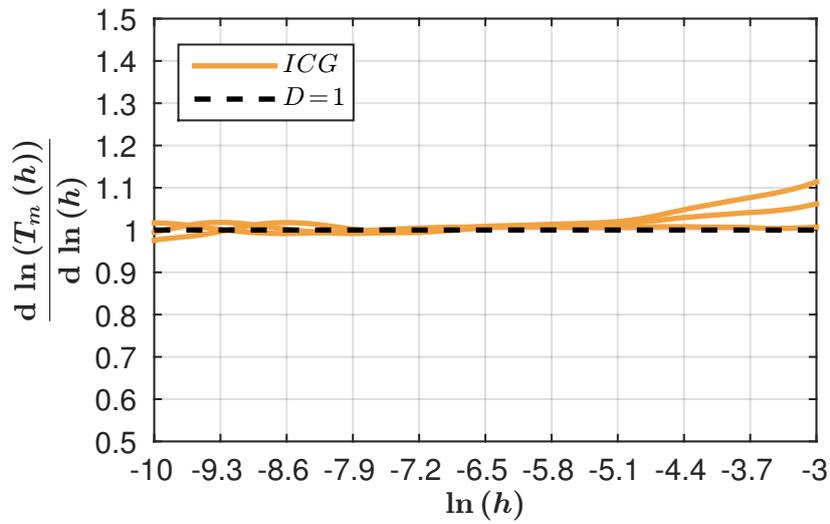


Figura 2.6. Estimación de D para la transformación triangular en ausencia de ruido. Se presenta la derivada logarítmica de la ICG para $m = \{2, 3, 4\}$. El valor teórico de D se muestra en línea negra discontinua.

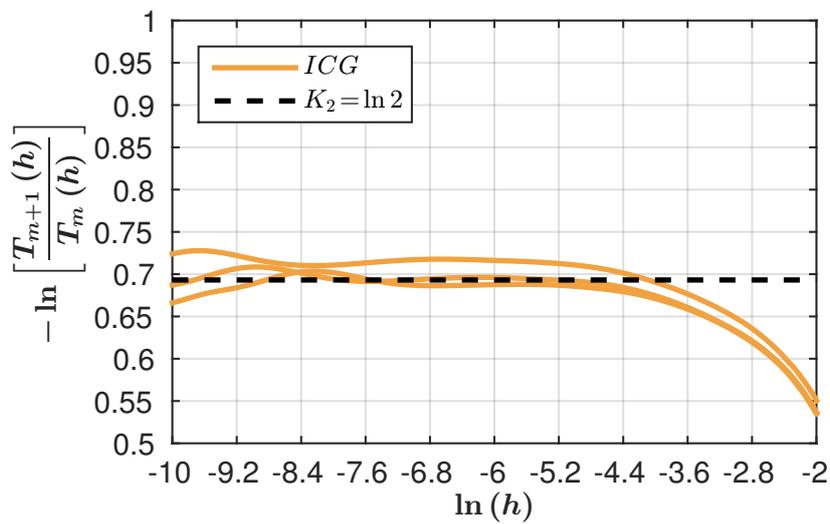


Figura 2.7. Estimación de K_2 para la transformación triangular en ausencia de ruido. Se presenta el funcional de entropía de la ICG para $m = \{2, 3, 4\}$. El valor reportado de K_2 se muestra en línea negra discontinua.

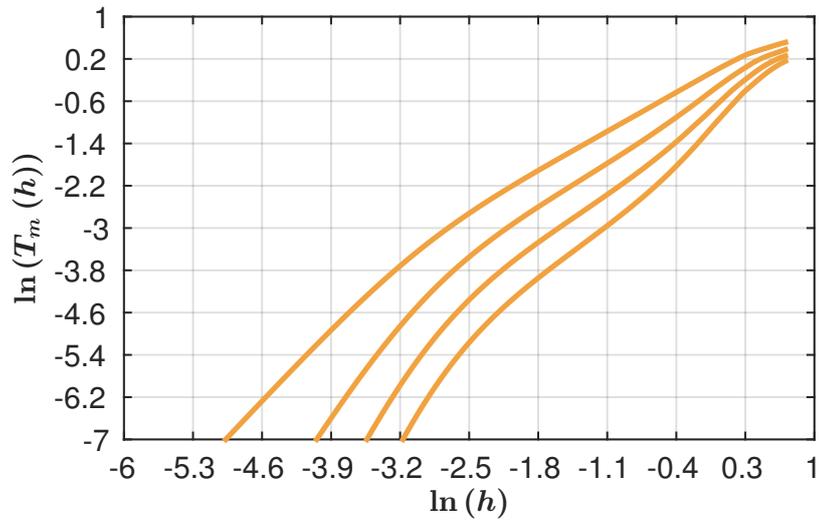


Figura 2.8. Gráfica log-log de la integral de correlación para la transformación triangular con ruido $\sigma = 0.05$. Se presentan las curvas calculadas con $m = \{2, 3, 4, 5\}$ ($m = 2$ curva superior y $m = 5$ curva inferior) para la ICG.

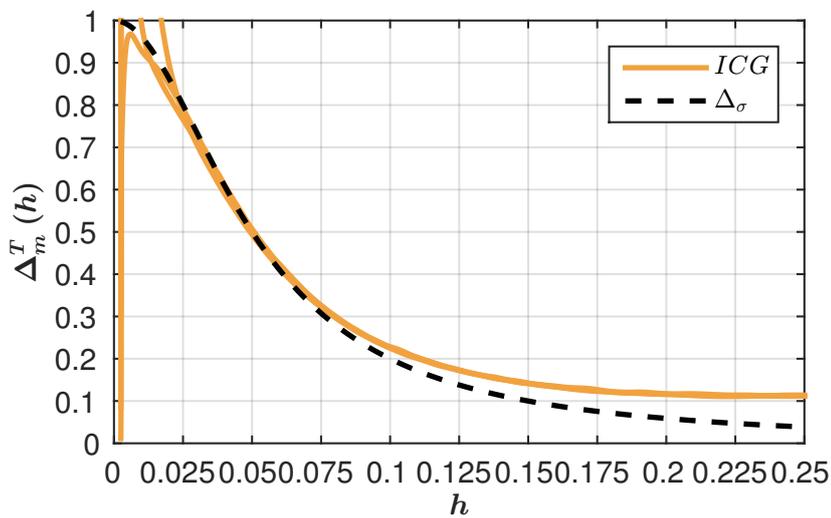


Figura 2.9. Estimación del nivel de ruido para la transformación triangular con ruido ($\sigma = 0.05$). Se presenta el funcional de nivel de ruido de la ICG (naranja) y la función $\Delta_\sigma(h) = \sigma^2 / (\sigma^2 + h^2)$ (línea negra discontinua).

curvas de ICG escalan de la misma forma que en ausencia de ruido. Este rango de valores de h se hace más corto a medida que el nivel de ruido aumenta, haciendo más difícil la estimación de cualquier invariante. En la figura 2.9 se presenta el funcional para el nivel de ruido $\Delta_m^T(h)$ (ecuación 2.14) en función de h . Es posible observar cómo las curvas para los diferentes valores de m decrecen a medida que h crece, además, para $0.01 \leq h \leq 0.05$ estas curvas siguen la tendencia de la función $\Delta_\sigma(h) = \sigma^2 / (\sigma^2 + h^2)$, tal como lo indica la ecuación 2.14. Es en este rango de valores de h donde es posible estimar σ mediante una regresión no lineal.

2.6 Entropía aproximada

La entropía aproximada (ApEn) encuentra su origen en la entropía de Kolgomorov-Sinai [77]. En [76] Eckmann y Ruelle introdujeron una forma de estimar la entropía de Kolgomorov-Sinai a partir de una serie temporal finita. Modificando el algoritmo de correlación de Grassberger y Procaccia (ecuación 2.18) ellos propusieron que K puede ser estimado mediante [76]:

$$K = \frac{1}{\tau} \lim_{h \rightarrow 0} \lim_{m \rightarrow \infty} \lim_{L \rightarrow \infty} [\phi^{m+1}(h) - \phi^m(h)], \quad (2.19)$$

donde

$$\phi^m(h) = \frac{1}{L} \sum_{i=1}^L \ln G_i^m(h),$$

$$G_i^m(h) = \frac{1}{L-1} \sum_{\substack{j=1 \\ j \neq i}}^L g(\|\mathbf{x}_i - \mathbf{x}_j\|/h),$$

y $g(t)$ es una función núcleo. Los límites presentes en esta ecuación nos indican que para su convergencia necesitamos series temporales con un número de datos que tiende a infinito y sin ruido. Esto hace que la estimación de la entropía de Kolgomorov-Sinai para series temporales reales sea una tarea muy difícil de llevar a cabo.

En los años 90, Pincus, inspirado en el trabajo de Eckmann y Ruelle propuso la ApEn como un estadístico para medir la regularidad de una serie temporal [94]. Su idea básicamente fue relajar algunos límites de la ecuación 2.19. En este sentido definió la ApEn como:

$$ApEn(m, h) = \frac{1}{\tau} \lim_{L \rightarrow \infty} [\phi^{m+1}(h) - \phi^m(h)], \quad (2.20)$$

donde

$$\phi^m(h) = \frac{1}{L} \sum_{i=1}^L \ln G_i^m(h),$$
$$G_i^m(h) = \frac{1}{L} \sum_{j=1}^L g(\|\mathbf{x}_i - \mathbf{x}_j\|/h).$$

Existen dos diferencias principales entre las ecuaciones (2.19) y (2.20). La primera es que los límites respecto a las variables h y m en la ecuación (2.19) han sido removidos en la ecuación (2.20). El segundo es que en la definición de la cantidad $G_i^m(h)$ en la ecuación (2.20) la sumatoria permite que $i = j$ lo que no pasa en la ecuación (2.19). Estas modificaciones nos dicen que la ApEn es una familia de funciones dependientes de los parámetros m y h . Además, para evitar el $\ln(0)$ en la ecuación (2.20), es decir $G_i^m(h) = 0$, es permitido el “autoconteo” [95]. En otras palabras, la distancia de un vector de estado consigo mismo es tenida en cuenta en el cálculo de la ApEn. Gracias a estas modificaciones la ApEn posee la capacidad para cuantificar la regularidad de una serie temporal de pocas muestras e inmersa en poco ruido. Es importante destacar que un valor positivo de ApEn no garantiza que la dinámica provenga de un sistema caótico [96].

La ApEn mide el logaritmo de la verosimilitud de que dos vectores de estado \mathbf{x}_i y $\mathbf{x}_j \in \mathbb{R}^m$ que son vecinos en un espacio m -dimensional, sigan siendo vecinos en un espacio de dimensión $(m + 1)$. Por lo tanto, para sistemas con dinámica irregular el valor de esta verosimilitud es pequeño en comparación con sistemas de dinámica más regular. En consecuencia la ApEn será mayor para el sistema cuya dinámica sea más irregular [97]. Existe una relación estrecha entre el grado de regularidad y la complejidad de una serie temporal. Los sistemas que son muy complejos se caracterizan por una alta dimensión y exponentes de Lyapunov positivos. Esta combinación genera series temporales muy irregulares.

Existen en la literatura dos versiones de la entropía aproximada: uno es la ApEn estándar, propuesta por Pincus, que se obtiene utilizando como función núcleo la función escalón $H(1 - \|\mathbf{x}_i - \mathbf{x}_j\|_\infty / h)$ y la distancia infinito como medida de distancia [97]. La segunda es la entropía aproximada de núcleo Gausiano (GApEn), propuesta por Xu y col. [98], obtenida utilizando la función $e^{-\|\mathbf{x}_i - \mathbf{x}_j\|/h}$ como función núcleo y la distancia euclídea como medida de distancia. La ApEn ha sido ampliamente utilizada en varios campos de la ciencia y ha encontrado en el área biomédica diversas aplicaciones [30, 33]. Se ha utilizado para caracterizar la dinámica de diferentes sistemas fisiológicos como el cardíaco [99-101], el sistema nervioso central [102-104], el sistema productor de la voz [34], entre otros. De forma similar la

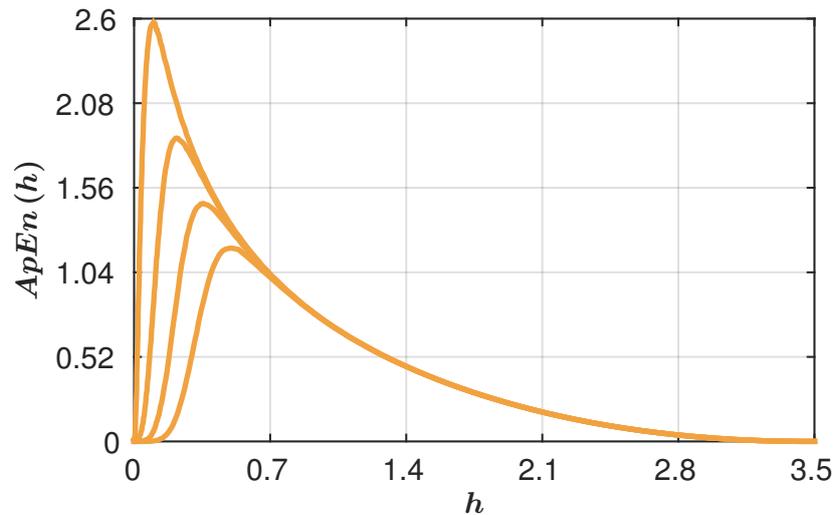


Figura 2.10. Entropía aproximada en función de h calculada con $m = \{2, 3, 4, 5\}$ ($m = 2$ curva superior y $m = 5$ curva inferior). La ApEn fue calculada a partir de $L = 5000$ vectores de estado obtenidos de una serie temporal proveniente de la transformación triangular.

ApEn se ha utilizado como característica para detectar diferentes tipos de patologías como episodios ictales [52, 105, 106], apnea del sueño [107] y voces patológicas [70].

En la figura 2.10 se muestra el comportamiento de la ApEn en función de h para $m = \{2, 3, 4, 5\}$. Se utilizó una serie temporal, proveniente de la transformación triangular, de la cual se formaron $L = 5000$ vectores de estado. Vemos cómo a través de h la ApEn aumenta su valor rápidamente hasta alcanzar su valor máximo y luego decrece de forma monótona hasta cero. De forma general la ApEn es una función decreciente respecto al parámetro m , es decir, $ApEn(m, h, L) \geq ApEn(m + 1, h, L)$. Esto se evidencia en la figura 2.10 donde se ve que la curva de mayor valor pertenece a $m = 2$.

La acción de autoconteo mencionada anteriormente hace que $ApEn(m, h, L)$ sea un estimador sesgado del estadístico $ApEn(m, h)$. Por ello este estadístico resulta inconsistente [95, 96]. Es decir, para dos dinámicas de diferente regularidad el estimador $ApEn(m, h, L)$ puede sugerir que una dinámica es más regular que otra para un valor determinado de los parámetros m y h . Sin embargo, para otros valores, puede sugerir todo lo contrario. En conclusión el estimador $ApEn(m, h, L)$ es altamente dependiente del valor de sus parámetros. En vista de esto y a las limitaciones computacionales de la época, Pincus propuso fijar el valor del parámetro m a un valor de 2 o 3 y el valor de h a un valor entre 0.1 y 0.3 veces la desviación estándar de la serie temporal en estudio [96].

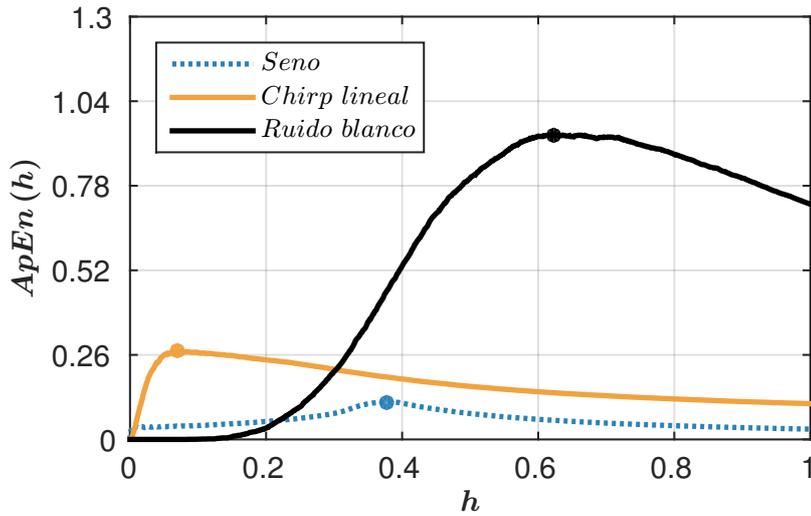


Figura 2.11. Entropía aproximada en función de h y $ApEn_{max}$ para una señal senoidal (azul punteado), un *chirp* lineal (naranja) y ruido blanco Gaussiano (negro). Todas las señales han sido normalizadas para tener varianza unitaria. Se marca con un punto el valor de $ApEn_{max}$ para cada señal. Figura reproducida de [108].

Para comprender las implicaciones de las sugerencias de Pincus debemos entender primero cómo afecta el valor de los parámetros m y h a la ApEn. Recordemos del capítulo 1 que el propósito de la inmersión de una serie temporal es reconstruir el atractor del sistema del cual proviene. Para tener una buena reconstrucción es necesario que el parámetro m sea mayor que dos veces la dimensión por conteo de cajas del sistema ($m > 2d_0$). Por tanto uno podría utilizar cualquier valor de m que cumpla con esta condición, sin embargo, utilizar un valor de m muy grande para calcular la ApEn implica un alto costo computacional. Además, entre más grande sea m más grande es el número de vectores de estado necesarios para poder estimar adecuadamente la cantidad $G_i^m(h)$ en la ecuación (2.20). Es decir si en una dimensión se necesitan N datos en dos dimensiones serán necesarios N^2 . Sumado a esto, una vez que se fija la longitud N de la serie temporal, un valor de m grande disminuye el número de vectores de estado L disponibles para el cálculo de la ApEn. Por este motivo Pincus sugiere utilizar valores de m chicos como 2 o 3. Por otro lado, una reconstrucción pobre del espacio de estados de un sistema lleva a cometer errores en la estimación de la ApEn. Como se mencionó antes, la ApEn cuantifica el logaritmo de la verosimilitud de que dos vectores vecinos en un espacio m -dimensional sean también vecinos en un espacio $(m + 1)$ -dimensional. El problema es que en un atractor mal reconstruido es común la presencia de falsos vecinos (ver subsección 1.3.2), es decir, vectores de estado que son vecinos por causa de una mala reconstrucción mas no por la dinámica del sistema [15, 109]. Este

fenómeno puede llevar a la ApEn a sugerir resultados equivocados. En este sentido siempre se recomienda utilizar valores de m cercanos a la mínima dimensión de inmersión [27], la cual puede calcularse utilizando el algoritmo propuesto por Cao en [29]. Debemos mencionar que a pesar de que la anterior discusión ha sido ampliamente difundida, hay investigadores que aún utilizan las recomendaciones de Pincus para calcular la ApEn. Por ejemplo Aletti y col. [110] utilizaron $m = 2$ para evaluar posibles malformaciones congénitas del corazón en niños a partir de señales de variabilidad de frecuencia cardíaca (VFC). De forma similar Zarjam y col. [104] utilizaron $m = 2$ y $m = 3$ sobre señales de EEG para identificar cambios en las respuestas del cerebro ante tareas con diferentes grados de complejidad cognitiva.

Con respecto al parámetro h , Pincus propuso fijar su valor entre 0.1 y 0.3 veces la desviación estándar de la señal [33, 111]. Esta recomendación se funda en el hecho de que para series temporales con pocos datos un valor muy pequeño de h no permitirá una buena estimación de la cantidad $G_i^m(h)$ en la ecuación (2.20). Además, la influencia del ruido en la serie temporal se manifiesta en las bajas escalas (valores de h pequeños). Por otro lado, dado que los atractores son conjuntos acotados, un valor de h muy grande podría abarcar todo el atractor impidiendo hacer distinciones detalladas [96]. Aunque esta metodología ha tenido sus adeptos [112–114], otros autores aseguran que puede llevar al análisis equivoco de la verdadera regularidad de una señal [74, 108, 115]. En su lugar Chon y col. propusieron el estimador $ApEn_{max}$ que se define como el valor máximo de $ApEn(m, h, L)$ con m y L fijos [108].

En la figura 2.11 se presenta una gráfica de la $ApEn(5, h, 3500)$ en función de h para tres señales de diferente complejidad y varianza unitaria. Es posible observar como para el rango $0.1 \leq h \leq 0.3$ la ApEn sugiere de forma errónea que la señal senoidal y el *chirp* lineal son mucho más irregulares que la señal de ruido blanco. Sin embargo, si se observa el valor del estimador $ApEn_{max}$ vemos que el orden de complejidad queda bien establecido comenzando en orden creciente por la señal sinusoidal, siguiendo con el *chirp* lineal y finalizando con el ruido blanco.

Otra solución al problema de la inconsistencia presentada por la ApEn es la llamada entropía muestral SampEn propuesta por Richman y Moorman [95]. La diferencia entre la ApEn y entropía muestral (SampEn) radica en que esta última no permite el autoconteo ($j \neq i$) en el cálculo de la cantidad $G_i^m(h)$ en la ecuación (2.20). La SampEn es un estimador de regularidad mucho más consistente que la ApEn [74, 95], sin embargo ambas cantidades son altamente influenciadas por el nivel de ruido presente en la serie temporal [116].

2.7 Conclusiones

En este capítulo se revisó el estado del arte en materia de medidas de complejidad y los algoritmos para su cálculo. Se mostró que los distintos grupos de invariantes nos ayudan a caracterizar la geometría de un atractor y la dinámica que lo gobierna. Se mencionaron y describieron medidas de dimensión, de entropía y de divergencia. También se presentaron los distintos algoritmos para el cálculo de estos invariantes y se dio evidencia gráfica de su utilidad. Este material será de gran importancia para abordar el resto de capítulos de este documento.

I'm trying to understand cosmology, why the Big Bang had the properties it did. And it's interesting to think that connects directly to our kitchens and how we can make eggs, how we can remember one direction of time, why causes precede effects, why we are born young and grow older. It's all because of entropy increasing.

Sean M. Carroll

3

Entropía aproximada y cambios de regularidad

A principios de los años 90, Pincus [94] propuso la entropía aproximada (ApEn) como un estadístico capaz de medir la regularidad de una serie temporal de pocas muestras e inmersa en ruido. Desde entonces la ApEn ha sido utilizada en diversas disciplinas científicas ya que permite caracterizar la dinámica de un sistema y diferenciarla de otra. Existe evidencia de una fuerte dependencia de la ApEn con el valor de sus parámetros m , h y N , lo cual crea inconsistencias al estimar la regularidad de una señal. Es decir, para un valor determinado de sus parámetros la ApEn puede sugerir que una dinámica es más irregular que otra, pero al cambiar el valor de estos parámetros puede sugerir todo lo contrario. Como solución se ha propuesto el estimador $ApEn_{max}$ (valor máximo de $ApEn(m, h, N)$ con m y N fijos) como un estimador de complejidad más consistente [108]. A pesar de esto, el nivel de ruido presente en la serie temporal deteriora significativamente el poder de discriminación de $ApEn_{max}$.

Teniendo en cuenta estas limitaciones, en este capítulo se tratará un nuevo enfoque basado en el estimador h_{max} , que se define como el valor de h para el cual $ApEn(m, h, N) = ApEn_{max}$. Esta nueva propuesta no solo mejora la capacidad de discriminación entre dinámicas de diferente complejidad, sino que también es más

robusta ante la presencia de ruido. Analizaremos esta nueva metodología con señales provenientes de sistemas dinámicos de baja y alta dimensión para luego extenderla a señales de EEG y de voz.

3.1 Introducción

La ApEn [94] depende de tres parámetros principalmente: la dimensión de inmersión (m), la escala o umbral (h) y la longitud de la serie temporal (N)¹ (ver sección 2.6). Es conocido que el estimador $ApEn(m, h, N)$ es altamente dependiente del valor de los parámetros elegidos para su cálculo [95, 96]. Esto hace que dicho estimador sea muy inconsistente y que los resultados basados en él sean difíciles de reproducir. Con el fin de dar solución a este inconveniente Pincus propuso que el valor de la dimensión de inmersión sea $m = 2$ o $m = 3$ [96, 117]. Esta sugerencia encuentra su justificación en que una vez que N se fija, un valor muy alto de m produce estimaciones pobres de la ApEn. Esto se debe a que para un valor de m alto son necesarios muchos vectores de estado para estimar la probabilidad condicional de que dos vectores que son vecinos en un espacio m -dimensional lo sigan siendo en un espacio $(m + 1)$ -dimensional. A lo anterior se le suma que el número L de vectores disponibles decrece con el valor de m . Dicha metodología puede ser acertada para analizar sistemas dinámicos de baja dimensión. Sin embargo, el utilizar valores de m pequeños en sistemas de alta dimensión puede provocar malas estimaciones de $ApEn(m, h, N)$ debido a una pobre reconstrucción del espacio de estados [15, 109].

Con respecto al valor del parámetro h , Pincus sugirió que este fuese fijado a un valor entre 0.1 o 0.3 veces la desviación estándar de la señal analizada [33, 111]. Esto se fundamenta en que para escalas pequeñas las trayectorias de un sistema están gobernadas principalmente por ruido. Además, un valor muy pequeño de h no permitirá una estimación adecuada de la distribución de probabilidades del atractor [117]. Por otro lado, un valor de h muy grande abarcará todo el espacio de estados del sistema haciendo imposible cualquier cálculo. Esta metodología ha tenido un vasto número de adeptos [112-114]. Pese a esto, Chon y col. mostraron con varios ejemplos que el enfoque propuesto por Pincus era incapaz de analizar la verdadera complejidad de una señal [108]. En su lugar propusieron utilizar el estimador $ApEn_{max}$ (el máximo valor de $ApEn(m, h, N)$ para m y N fijos) como una medida más consistente de regularidad [74, 108, 115] (ver sección 2.6).

¹En realidad la ApEn depende del número de vectores de estado L disponibles para su cálculo. Sin embargo L depende en gran medida de la longitud de la serie temporal $L = N - (m - 1)\tau$.

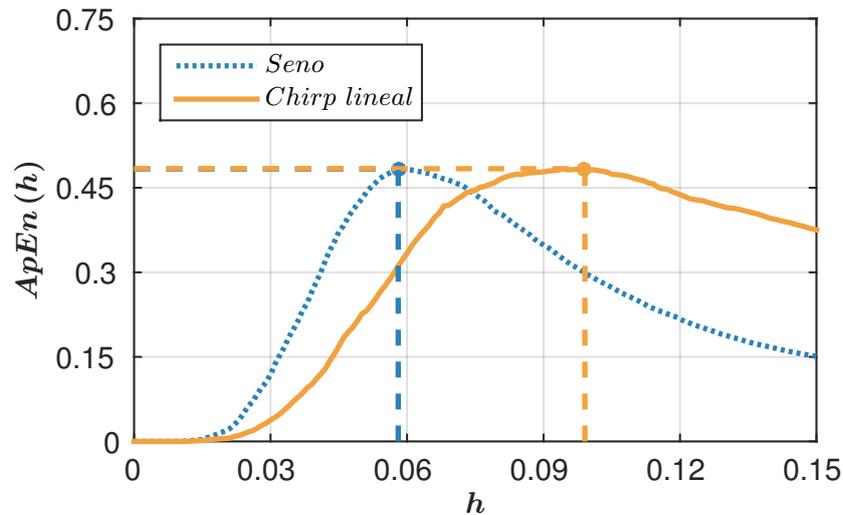


Figura 3.1. Entropía aproximada en función de h , $ApEn_{max}$ y h_{max} para una señal senoidal (azul punteado) y un *chirp* lineal (naranja) inmersos en ruido blanco Gaussiano (RSR= 30 dB). Las dos señales han sido normalizadas para tener varianza unitaria. Para cada una se marca el valor de $ApEn_{max}$ y h_{max} con líneas horizontales y verticales respectivamente. Los valores próximos de $ApEn_{max}$ erróneamente sugieren que ambas dinámicas poseen casi la misma regularidad.

En base a resultados experimentales con señales de VFC Castiglioni y col. [118] concluyen que el uso de $ApEn_{max}$ permite cuantificar mejor la complejidad de una señal, en comparación con cualquier otro valor de $ApEn$. Por otro lado, Liu y col. [100] observaron que $ApEn_{max}$ era incapaz de distinguir entre un grupo de pacientes saludables y otro con insuficiencia cardíaca en experimentos con señales de VFC. Recientemente, Boskovic y col. [119], presentaron evidencia de la inestabilidad del estimador $ApEn_{max}$. Ellos observaron que para dos series temporales de diferente complejidad, este estimador sugería resultados opuestos al disminuir la longitud de datos. Es decir, para cierta longitud de datos $ApEn_{max}$ sugería que una dinámica era más irregular que otra y al disminuir N este estimador indicaba todo lo contrario.

El nivel de ruido presente en la serie temporal tiene una influencia muy importante en la estimación de la $ApEn(m, h, N)$ y por tanto también afecta en gran medida a $ApEn_{max}$. Pincus asegura que la confiabilidad de la $ApEn$ se compromete seriamente cuando la relación señal a ruido (RSR) está por debajo de 3 dB [117]. Más aún, la presencia de ruido disminuye significativamente la capacidad de discriminación de $ApEn_{max}$ haciéndola una herramienta inútil. En la figura 3.1 se muestra el estimador $ApEn(5, h, 3500)$ en función de h para una señal senoidal y un *chirp* lineal, ambos normalizados para tener varianza unitaria e inmersos en ruido blanco

Gausiano (RSR= 30 dB). Es posible observar cómo los valores de $ApEn_{max}$ para ambas señales son muy próximos, indicando que ambas señales tienen casi la misma regularidad. Es decir, el ruido deteriora la capacidad de discriminación de $ApEn_{max}$ (comparar con la figura [2.11](#)). Por otro lado vemos que existe una diferencia marcada entre los valores de h para los cuales cada señal alcanza su valor máximo de entropía.

Entendiendo las limitaciones del estimador $ApEn_{max}$, en este capítulo proponemos el uso de un nuevo estimador denominado h_{max} , que es el valor de h para el cual $ApEn(m, h, N) = ApEn_{max}$ con m y N fijos. Demostraremos que este estimador brinda información útil que puede ser utilizada con fines de clasificación. Veremos también como el uso conjunto de $ApEn_{max}$ con h_{max} provee una metodología robusta para diferenciar entre dinámicas de diferente complejidad, incluso en presencia de ruido.

3.2 $ApEn_{max}$, h_{max} y cambios de regularidad

En presencia de ruido, el estimador $ApEn_{max}$ puede ser incapaz de diferenciar dinámicas con diferente complejidad. La hipótesis a tratar entonces es que el estimador y h_{max} brinda valiosa información que puede ser útil para el proceso de clasificación.

Con el fin de evaluar esta hipótesis realizamos 3 simulaciones utilizando series temporales provenientes de sistemas de baja y alta dimensión como lo son los modelos de Shilnikov y de Mackey-Glass respectivamente. En la tabla [3.1](#) se presentan estos modelos y el valor de los parámetros utilizados para obtener las series temporales. También se muestran el valor de los parámetros utilizados para calcular la $ApEn(m, h, N)$.

3.2.1 Simulaciones con señales artificiales

Comenzaremos estudiando las señales obtenidas a partir del modelo de Mackey-Glass. Este modelo fue seleccionado debido a la necesidad de observar el comportamiento de $ApEn_{max}$ y h_{max} en series temporales provenientes de un sistema de alta dimensión ($D \approx 8$) [\[120\]](#). Además, este sistema ha sido ampliamente utilizado para estudiar otro tipo de invariantes [\[80\]](#) y para modelar la dinámica de sistemas fisiológicos como el sistema neurológico [\[120\]](#), el sistema respiratorio [\[121\]](#) y el sistema hematopoyético [\[121\]](#).

3.2. $ApEn_{max}$, h_{max} y cambios de regularidad

Modelo	Ecuación	Parámetros del Modelo	Parámetros de la Simulación
Mackey-Glass	$\dot{x} = ax \frac{t-\lambda}{1+[x(t-\lambda)]^c} - bx(t)$	$c = \{4.5, 6\}$ $a = 0.9$ $b = 0.3$ $\lambda = 80$ $\Delta t = 1$	$m = \{2, 3, \dots, 20\}$ $h = \{0, 5 \times 10^{-5}, \dots, 0.035\}$ $\tau = 83$ $N = 5000$
Shilnikov	$\dot{x} = y$ $\dot{y} = z$ $\dot{z} = \mu x - y - \varepsilon z - ax^2 - bx^3$	$a = \{0.008, 0.2217\}$ $\varepsilon = 0.55$ $\mu = 0.65$ $b = 0.65$ $\Delta t = 0.2$	$m = \{2, 3, \dots, 20\}$ $h = \{0, 5 \times 10^{-5}, \dots, 0.035\}$ $\tau = 10$ $N = 5000$

Tabla 3.1. Modelos y parámetros utilizados en las simulaciones. Δt es el paso utilizado para obtener las soluciones numéricas. La cota superior del intervalo de valores de h es igual al desvío estándar de la serie temporal normalizada.

Construimos dos conjuntos de 240 señales cada uno. Para el primero utilizamos el parámetro $c = 4.5$ y para el segundo elegimos $c = 6$, con el fin de simular dinámicas con diferente grado de complejidad. Cada realización contó con una condición inicial distinta escogida de forma aleatoria de una distribución $\mathcal{U}(0; 0.01)$.

De cada realización de 20 000 puntos descartamos los primeros 15 000 puntos y formamos series temporales de $N = 5000$ puntos que luego normalizamos para tener energía unitaria. Para calcular el valor de τ calculamos la función de información mutua (ver ec. (1.2)) para dos series temporales, una de cada grupo. Elegimos el valor de τ correspondiente al primer mínimo local de cada función, y seleccionamos como retardo de inmersión al mayor de estos valores ($\tau = 83$). Luego calculamos el estimador $ApEn(m, h, N)$ para cada serie temporal con los valores de m y h listados en la tabla 3.1. Finalmente, para cada serie temporal registramos los valores de $ApEn_{max}$ y h_{max} . De forma adicional estimamos la mínima dimensión de inmersión de cada señal utilizando el algoritmo de Cao [29].

Para evaluar el comportamiento de $ApEn_{max}$ y h_{max} en presencia de ruido, a cada serie temporal le adicionamos ruido blanco Gaussiano hasta alcanzar relaciones de señal a ruido de 5 dB y de 0 dB. Luego, cada serie temporal ruidosa la normalizamos para tener energía unitaria y $ApEn_{max}$ y h_{max} los calculamos como mencionamos antes.

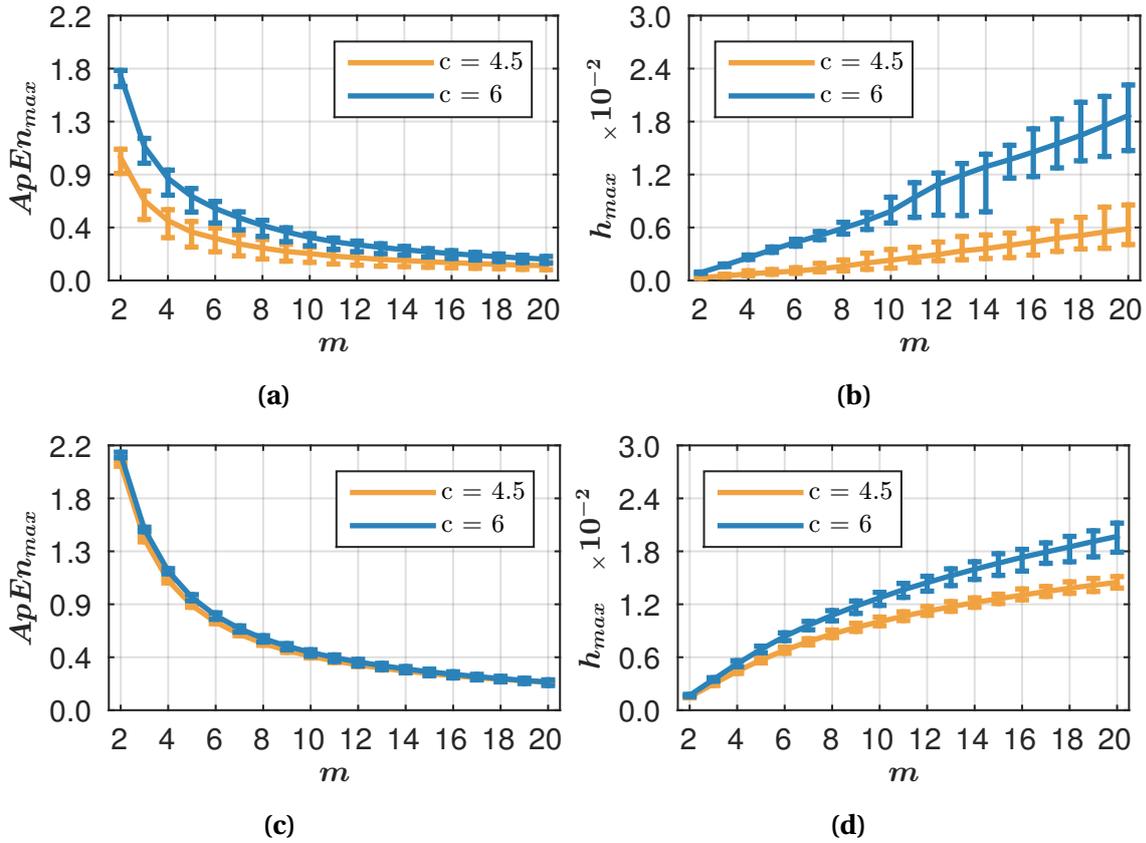


Figura 3.2. Sistema de Mackey-Glass. Intervalo de confianza del 95%. Sin ruido: (a) $ApEn_{max}$ y (b) h_{max} . RSR= 5dB: (c) $ApEn_{max}$ y (d) h_{max} .

En la figura 3.2 se presentan los resultados de la simulación con el sistema de Mackey-Glass para dos grupos de dinámicas con distinto valor del parámetro c . Se muestran los intervalos de confianza del 95% para $ApEn_{max}$ y h_{max} calculados con distintos valores de $m = \{2, 3, \dots, 20\}$. Estos intervalos de confianza fueron hallados de forma empírica. Primero se ordenaron de mayor a menor las estimaciones de $ApEn_{max}$ y h_{max} para las 240 realizaciones. Luego se tomaron los percentiles 2.5 y 97.5 como los límites inferior y superior del intervalo respectivamente. En la figura 3.2a se puede observar que, en ausencia de ruido, las curvas de $ApEn_{max}$ para los dos grupos de dinámicas diferentes se hacen más cercanas a medida que crece m . Estas curvas alcanzan su máxima diferencia cuando $m = 2$. Por el contrario, en la figura 3.2b se puede apreciar que la diferencia entre las curvas de h_{max} pertenecientes a diferentes dinámicas aumenta a medida que m crece.

El efecto que tiene la presencia de ruido en la serie temporal (RSR= 5dB) en el comportamiento de $ApEn_{max}$ y h_{max} puede observarse en las figuras 3.2c y 3.2d respectivamente. En comparación con el caso sin ruido, tanto el valor medio de $ApEn_{max}$

3.2. $ApEn_{max}$, h_{max} y cambios de regularidad

como el de h_{max} se han incrementado y los intervalos de confianza en ambos casos se han reducido. En segundo lugar podemos observar en la figura 3.2c que el estimador $ApEn_{max}$ ha perdido la capacidad de discriminación entre las diferentes dinámicas para valores de $m \geq 6$. Por el contrario, el estimador h_{max} conserva su poder de discriminación para un rango mayor de valores de m (ver figura 3.2d).

Estos resultados sugieren que el estimador h_{max} puede ser utilizado para discernir entre dinámicas de diferente complejidad aún en presencia de ruido.

Para comprender mejor el funcionamiento de estos estimadores, estudiamos su utilidad para el caso de series temporales provenientes del sistema de Shilnikov. Escogimos este sistema ya que su dinámica para valores diferentes del parámetro a asemeja a la dinámica de señales de EEG normales y durante una crisis epiléptica de tipo petit mal [122] (ver subsección 1.4.1). La simulación contó con 240 realizaciones de $N = 5000$ datos de la variable x . Cada realización la calculamos partiendo de una condición inicial obtenida aleatoriamente de una distribución $\mathcal{U}(0;0.01)$ y luego normalizamos cada serie para tener energía unitaria. Para cada serie temporal calculamos las curvas de $ApEn(m, h, N)$ con $m = \{2, 3, \dots, 20\}$ y $\tau = 10$. De ellas obtuvimos los estimadores $ApEn_{max}$ y h_{max} . Practicamos el mismo procedimiento con series temporales con relación señal a ruido de 5 dB y 0 dB.

En la figura 3.3 se presentan los intervalos de confianza del 95% para $ApEn_{max}$ y h_{max} en función de m . Las curvas corresponden a grupos diferentes de dinámicas provenientes del sistema de Shilnikov con y sin ruido. En ausencia de ruido se aprecia en la figura 3.3a que el estimador $ApEn_{max}$ solo puede discernir entre los dos tipos de dinámicas para un rango limitado de valores de m ($4 \leq m \leq 12$). De la misma forma es posible observar en la figura 3.3b que el estimador h_{max} solo diferencia estas dos dinámicas para $m \geq 13$. Se puede concluir que existe una relación de complementariedad entre $ApEn_{max}$ y h_{max} a lo largo de los distintos valores de m .

En presencia de ruido (RSR = 5 dB) en las figuras 3.3c y 3.3d se puede observar que tanto la distancia entre las curvas de $ApEn_{max}$ como entre las de h_{max} es mucho más corta que en el caso sin ruido, incluso es más corta que la distancia entre las curvas de $ApEn_{max}$ y h_{max} correspondientes al sistema de Mackey-Glass (ver figura 3.2). Se aprecia que la varianza de ambos estimadores también se ha reducido considerablemente. Estos resultados nos llevan a pensar que la utilización en conjunto de los estimadores $ApEn_{max}$ y h_{max} podría permitir una mejor detección de dinámicas con diferente complejidad.

Para probar la anterior hipótesis, en la figura 3.4 se presenta un gráfico de disper-

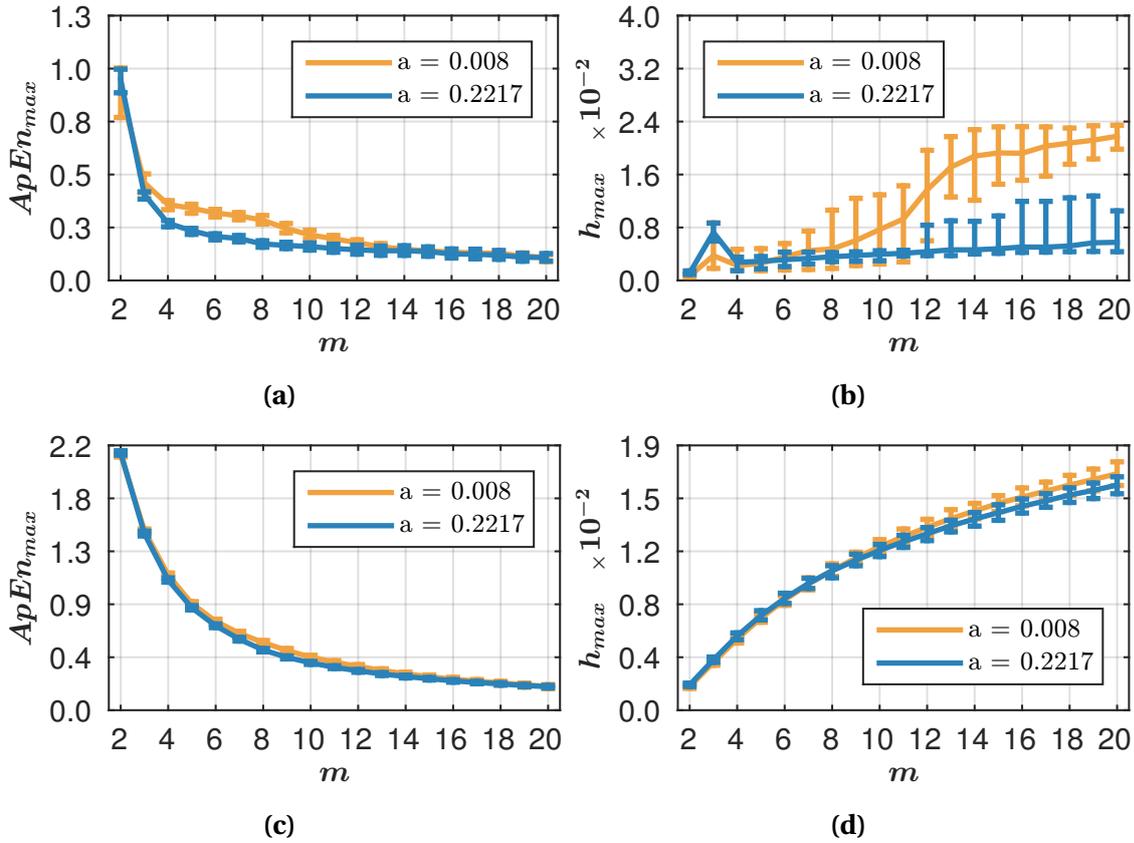


Figura 3.3. Sistema de Shilnikov. Intervalo de confianza del 95%. Sin ruido: (a) $ApEn_{max}$. (b) h_{max} . RSR= 5dB: (c) $ApEn_{max}$. (d) h_{max} .

sión de $ApEn_{max}$ versus h_{max} para ambos modelos con ruido (RSR= 5dB), usando $m = 2$ y $m = 12$. En la figura 3.4a se muestra que para el sistema de Mackey-Glass en presencia de ruido, basta utilizar el estimador $ApEn_{max}$ con $m = 2$ para diferenciar correctamente ambas dinámicas. Sin embargo, es evidente que h_{max} provee información útil para facilitar el proceso de clasificación. Una realidad diferente se puede observar en la figura 3.4c para el sistema de Shilnikov. Es posible apreciar que no es posible diferenciar las diferentes dinámicas utilizando sólo el estimador $ApEn_{max}$ calculado con $m = 2$. No obstante, con la información que aporta h_{max} estas dos clases pueden ser separadas de una forma que, si bien no es óptima, es más conveniente.

Como mencionamos anteriormente, ante la presencia de ruido en la serie temporal, la estimación de $ApEn_{max}$ puede ser más precisa si se utiliza un valor de m mayor o igual a la mínima dimensión de inmersión que si se utiliza $m = 2$. Para estos dos sistemas inmersos en ruido, el algoritmo de Cao [29] sugiere que la mínima dimensión de inmersión es $m = 12$. Este valor tan alto de m es el resultado de la pre-

3.2. $ApEn_{max}$, h_{max} y cambios de regularidad

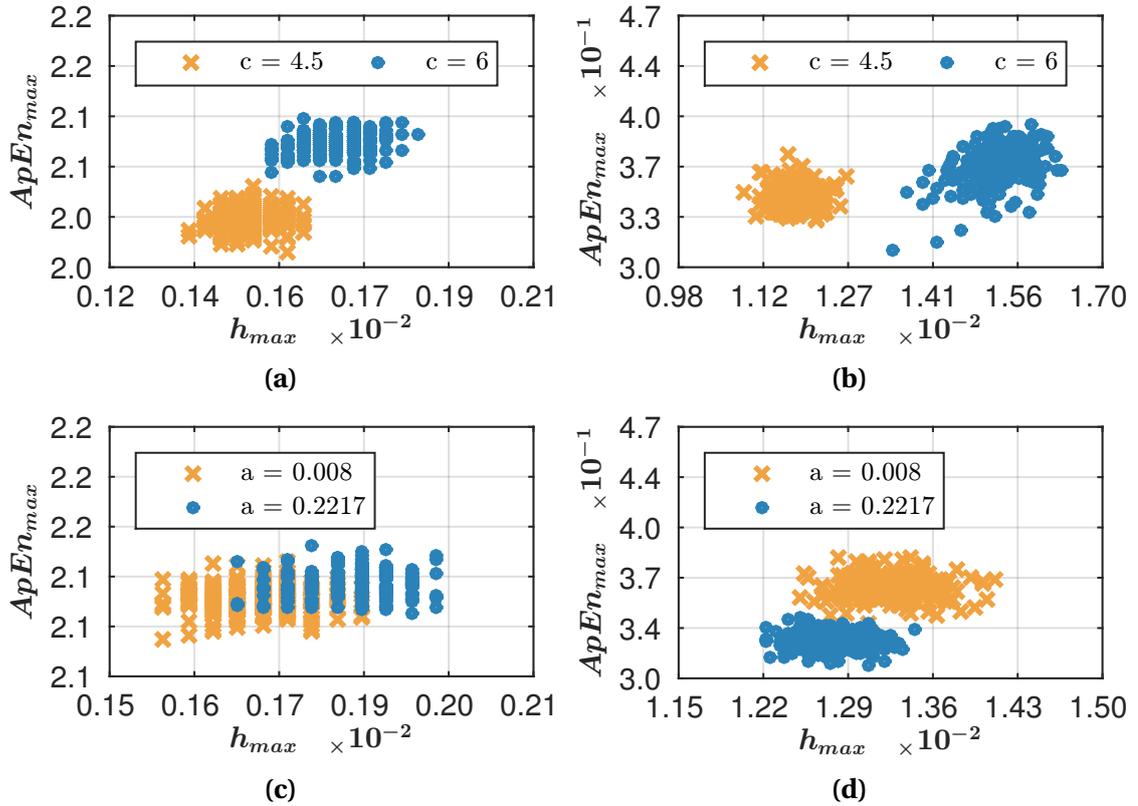


Figura 3.4. Gráfica de dispersión de $ApEn_{max}$ vs h_{max} en presencia de ruido (RSR= 5dB). Sistema de Mack-Glass: (a) $m = 2$. (b) $m = 12$. Sistema de Shilnikov: (c) $m = 2$. (d) $m = 12$.

sencia de ruido. Podemos ver en la figura [3.4b](#) que para el sistema de Mackey-Glass la capacidad de discriminación de $ApEn_{max}$ calculado con $m = 12$ es casi nula. Por otro lado, vemos que con la ayuda de h_{max} ambos grupos de dinámicas pueden ser adecuadamente separados. Esto se debe al comportamiento complementario de $ApEn_{max}$ y h_{max} a lo largo de los valores del parámetro m . De forma similar se puede apreciar en la figura [3.4d](#) que en el caso del sistema de Shilnikov, el uso conjunto de ambos estimadores permite que los datos se agrupen de manera conveniente para su clasificación. Lo anterior refuerza la idea de que los estimadores $ApEn_{max}$ y h_{max} deberían ser utilizados conjuntamente para aumentar la capacidad de discriminación entre dinámicas de diferente complejidad.

Análisis de selección de características

Para verificar cuantitativamente la capacidad de discriminación de los estimadores $ApEn_{max}$ y h_{max} , realizamos una validación cruzada de 10 particiones utilizando máquinas de soporte vectorial (MSV) de tipo lineal [\[123\]](#) como clasificador. En este

tipo de validación cruzada, el conjunto de datos es dividido en 10 subconjuntos, de los cuales nueve son utilizados para entrenar el clasificador y el conjunto restante es dejado como grupo de prueba. Este procedimiento es repetido 10 veces, dejando siempre un conjunto de prueba diferente [124]. Con respecto al clasificador, la idea básica de las MSV lineales es separar las clases utilizando un hiperplano óptimo, es decir, una función de decisión lineal que maximice la distancia entre los puntos más cercanos de cada clase al hiperplano [123]. Es importante que el clasificador sea lineal ya que dentro de los clasificadores basados en MSV este es el más simple. Esto nos ayudará a revelar el verdadero poder de los estimadores $ApEn_{max}$ y h_{max} para discriminar entre diferentes dinámicas.

En la figura 3.5 se presenta la tasa de error de clasificación (TEC) en función de m . Esta tasa se calculó como el número de señales clasificadas incorrectamente sobre el total de señales para tres clasificadores diferentes. El primer clasificador fue entrenado sólo con el estimador $ApEn_{max}$, el segundo sólo con h_{max} y el tercero cuenta con la información provista por ambos estimadores en conjunto. Estos estimadores fueron calculados con series temporales provenientes del sistema de Mackey-Glass (figuras 3.5a, 3.5b y 3.5c), el sistema de Shilnikov (figuras 3.5d, 3.5e y 3.5f) y la transformación logística (figuras 3.5g, 3.5h y 3.5i) bajo diferentes niveles de ruido (sin ruido, RSR = 5dB y RSR = 0dB). De forma adicional, se probó esta metodología con series temporales provenientes de la transformación logística (ecuación (1.1)). Los dos grupos de dinámicas se obtuvieron con $\mu = 3.75$ y $\mu = 3.95$ respectivamente.

En la figura 3.5 se muestran comparativamente los resultados de la simulación para cada caso. Para el sistema de Mackey-Glass en ausencia de ruido se aprecia en la figura 3.5a que la TEC del clasificador basado en $ApEn_{max}$ se incrementa con m , alcanzando su máximo (0.065) para $m = 19$. Por su parte, el clasificador basado en h_{max} presenta una TEC diferente de cero solo para $m \geq 15$. Sin embargo la TEC del clasificador entrenado con ambos estimadores es cero para $2 \leq m \leq 14$ con un valor máximo de 0.006 para $m = 15$. Un resultado parecido puede observarse para este sistema inmerso en ruido (RSR = 5 dB). En contraste con el caso sin ruido, en la figura 3.5b se puede apreciar que la TEC del clasificador que cuenta solo con la información de $ApEn_{max}$ muestra un notable incremento. También se puede ver que utilizando solo la información de h_{max} el clasificador tiene una TEC diferente de cero para $2 \leq m \leq 4$. No obstante, la TEC del clasificador que utiliza ambos estimadores en conjunto permanece igual a cero para todo valor de m . Al aumentar el nivel de ruido a RSR = 0 dB (figura 3.5c), se puede ver que la TEC del clasificador con los dos estimadores siempre es menor o igual que la TEC de los otros dos clasificadores. Este resultado es una clara evidencia de que en conjunto $ApEn_{max}$ y h_{max} proveen características robustas ante el ruido.

3.2. $ApEn_{max}$, h_{max} y cambios de regularidad

Con respecto a los resultados con el sistema de Shilnikov, se aprecia en la figura [3.5d](#) que, para $3 \leq m \leq 11$ la TEC del clasificador basado en $ApEn_{max}$ es menor que la TEC del clasificador entrenado con h_{max} . Por el contrario, para $m = 2$ y $12 \leq m \leq 20$ la TEC de clasificador basado en h_{max} es menor que la del basado en $ApEn_{max}$. Sin embargo, la TEC del clasificador basado en ambos estimadores siempre se encuentra por debajo de la TEC de los otros dos, siendo cero para $4 \leq m \leq 11$ y $16 \leq m \leq 20$. De esta misma forma, en presencia de ruido (RSR=5 dB) se aprecia en la figura [3.5e](#) que para todos los valores de m la TEC del clasificador basado en $ApEn_{max}$ y h_{max} siempre está por debajo o es igual a la menor TEC entre los otros dos clasificadores, siendo cero para $8 \leq m \leq 12$ y alcanzando el valor de 0.004 para $m = 12$. Para niveles de ruido mayores (RSR=0 dB) se observa en la figura [3.5f](#) que para $m \geq 8$ se pueden alcanzar tasas de error de clasificación muy bajas utilizando solo $ApEn_{max}$ o en conjunto h_{max} y $ApEn_{max}$, siendo cero para $10 \leq m \leq 15$.

Los resultados obtenidos para la transformación logística se muestran en las figuras [3.5g](#), [3.5h](#) y [3.5i](#). Se puede ver que la TEC del clasificador que usa en conjunto a $ApEn_{max}$ y h_{max} es cero para todos los valores de m en el caso sin ruido y para $m \geq 3$ en el caso con RSR=5 dB. De la figura [3.5i](#) se puede ver que la TEC del clasificador basado en $ApEn_{max}$ y h_{max} está por debajo de la de los otros dos para todos los valores de m . Esta es cero para $8 \leq m \leq 9$ y tiene un valor muy bajo para valores de $m \geq 5$. Estos hallazgos ilustran la utilidad de esta metodología para discriminar entre series temporales provenientes de sistemas dinámicos discretos.

Es importante resaltar que cuando las series temporales provenientes de estos sistemas (Mackey-Glass, Shilnikov y la transformación logística) están inmersas en ruido con una baja RSR (ver figuras [3.5c](#), [3.5f](#) y [3.5i](#)), los peores resultados fueron obtenidos para valores bajos de m (especialmente $m = 2$). Esto sugiere que incrementar la dimensión de inmersión puede traer beneficios a la hora de discriminar entre dinámicas. Por tanto, concluimos que para alcanzar tasas de error pequeñas, se deben utilizar de forma conjunta $ApEn_{max}$ y h_{max} calculados para valores de m cercanos a la mínima dimensión de inmersión.

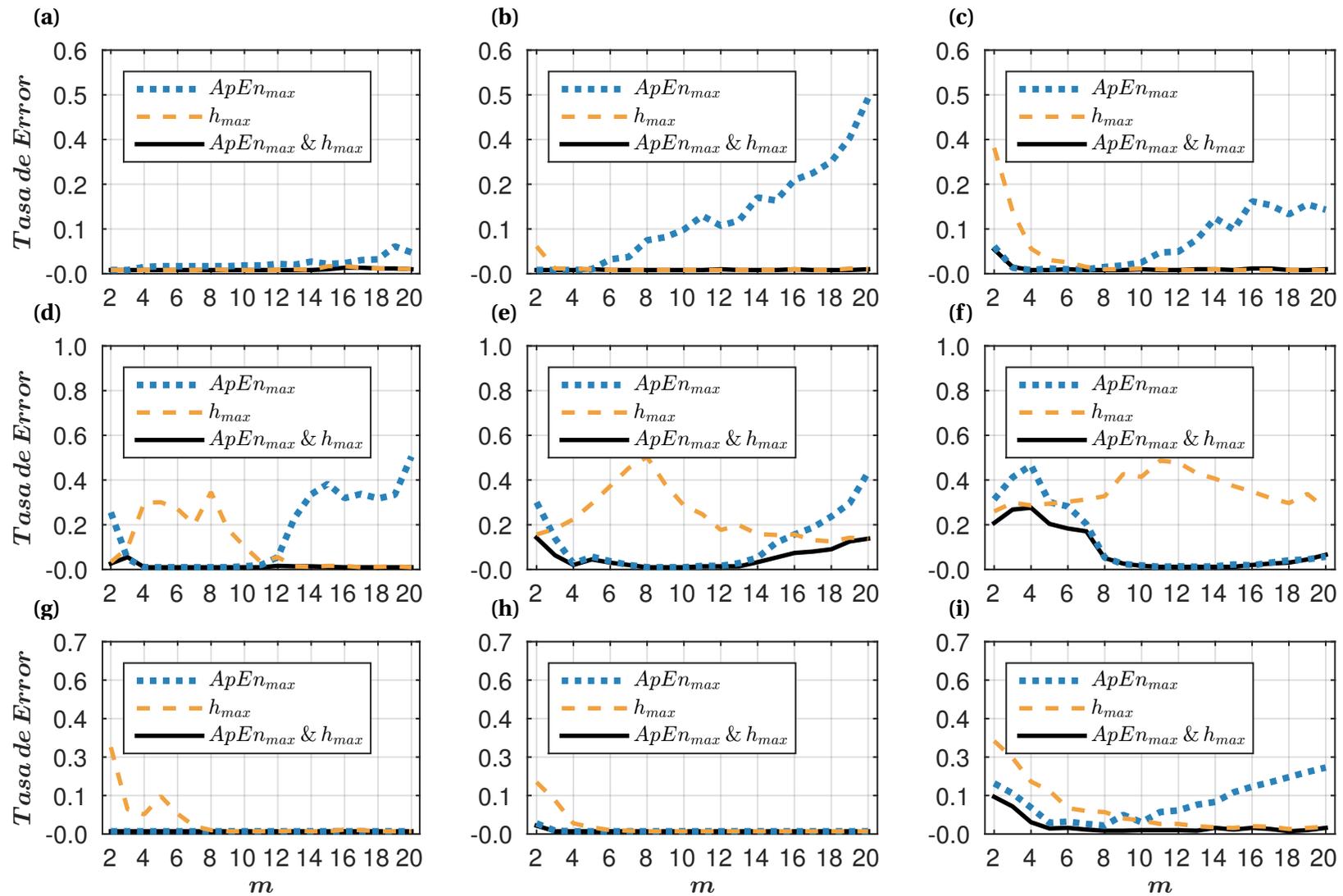


Figura 3.5. Validación cruzada de 10 particiones para la tasa de error de clasificación para 3 clasificadores diferentes. El primer clasificador fue entrenado solo con el estimador $ApEn_{max}$ (azul punteado), el segundo sólo con h_{max} (naranja discontinuo) y el tercero con ambos estimadores en conjunto (negro). Sistema de Mackey-Glass: (a) Sin Ruido. (b) RSR=5dB. (c) RSR=0dB. Sistema de Shilnikov: (d) Sin Ruido. (e) RSR=5dB. (f) RSR=0dB. Transformación logística: (g) Sin Ruido. (h) RSR=5dB. (i) RSR=0dB.

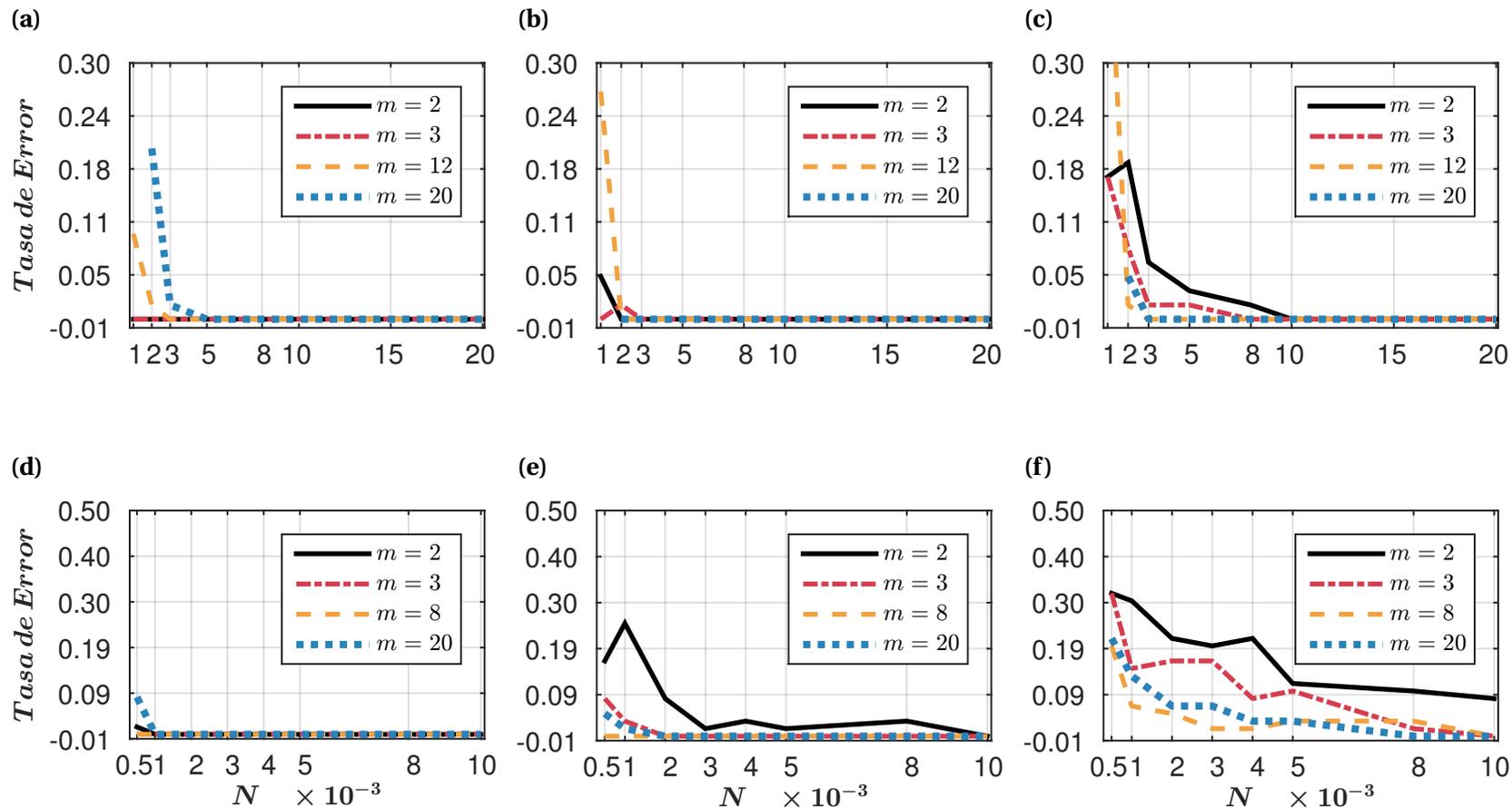


Figura 3.6. Validación cruzada de una iteración, en función de N , para la tasa de error de clasificación para un clasificador entrenado con $ApEn_{max}$ y h_{max} en conjunto. Estos estimadores fueron calculados con $m = 2$ (negro), $m = 3$ (rojo discontinuo), $m = [8, 12]$ (naranja discontinuo) y $m = 20$ (azul punteado). Sistema de Mackey-Glass: (a) Sin ruido. (b) RSR=5dB. (c) RSR=0dB. Transformación logística: (d) Sin Ruido. (e) RSR=5dB. (f) RSR=0dB.

Capítulo 3. Entropía aproximada y cambios de regularidad

Con el propósito de evaluar nuestra metodología ante las variaciones en el tamaño de la serie temporal N , realizamos una simulación contando con dinámicas de un sistema de baja dimensión como la transformación logística y uno de alta dimensión como el de Mackey-Glass. Se construyeron dos grupos de 30 realizaciones. En cada grupo se utilizaron valores distintos del parámetro R para la transformación logística y del parámetro c para el sistema de Mackey-Glass. Se normalizó cada serie temporal, para tener energía unitaria. Para la transformación logística $ApEn_{max}$ y h_{max} se estimaron con $m = \{2, 3, 8, 20\}$, $\tau = 1$ y $N = \{0.5, 1, 2, 3, 4, 5, 8, 10\} \times 10^3$. Luego se calculó la TEC de un clasificador, basado en MSV lineales, a través de un proceso de validación cruzada de k -particiones (“Leave one out”). Este clasificador está alimentado por las estimaciones de $ApEn_{max}$ y h_{max} en conjunto. La misma metodología fue aplicada a señales contaminadas con ruido blanco (RSR=5 dB y RSR=0 dB). Es importante mencionar el valor de $m = 8$ corresponde a la estimación de la mínima dimensión de inmersión arrojada por el algoritmo de Cao para este sistema bajo RSR=0 dB. De esta forma se repitió la misma simulación para el sistema de Mackey-Glass con $m = \{2, 3, 12, 20\}$, $\tau = 83$ y $N = \{1, 2, 3, 5, 8, 10, 15, 20\} \times 10^3$. La mínima dimensión de inmersión para las series temporales provenientes de este sistema con RSR=0 dB es $m = 12$. Finalmente debemos aclarar que las estimaciones para $m = 20$ con el sistema de Mackey-Glass se calculó para todos los valores de N excepto para $N = 1000$, ya que esa longitud de datos es muy corta para poder calcular la ApEn.

En la figura 3.6 se muestra la tasa de error de clasificación para un clasificador basado en MSV de tipo lineal, calculada a través de una validación cruzada de k particiones. En el eje de las abscisas se encuentra la longitud de las series temporales utilizadas. Estas provienen del sistema de Mackey-Glass (figuras 3.6a, 3.6b y 3.6c) y de la transformación logística (figuras 3.6d, 3.6e y 3.6f).

Para el sistema de Mackey-Glass en ausencia de ruido, se puede observar en la figura 3.6a que, para longitudes de datos cortas, la mayor TEC se alcanza utilizando $m = 20$ ($N = 2000$) y $m = 12$ ($N = 1000$). Esto ocurre porque para estos valores de m la cantidad de vectores de estado disponibles es pequeña, lo que repercute en malas estimaciones. No obstante, al incrementar el tamaño de la serie temporal, la TEC es igual a cero sin importar el valor de m . Es de resaltar que para $m = 2$ y $m = 3$ la TEC es cero para todo N . Por otro lado, en la figura 3.6b se ve que, comparado con el caso sin ruido, para $m = 2$ ($N = 1000$) y $m = 3$ ($N = 2000$) la tasa de error se ha incrementado, mientras que la TEC para $m = 12$ y $m = 20$ se ha reducido a cero para $N = 2000$. Es de resaltar que para $N \geq 3000$ la TEC es cero para todos los valores de m . Se muestra en la figura 3.6c que al incrementar la RSR a 0 dB, las mayores tasa de error son alcanzadas para $m = 2$ y $m = 3$ (excluyendo el error 0.48

para $m = 12$ en $N = 1000$). Esto nos lleva a pensar que ante la presencia de ruido es importante incrementar la dimensión de inmersión. Sin embargo hay que tener en cuenta que para longitudes de datos cortas, utilizar un valor de m muy grande puede resultar perjudicial.

En el caso de la transformación logística en ausencia de ruido se puede ver en la figura 3.6d que para $N = 500$ la mayor TEC (0.08) pertenece al clasificador entrenado con $ApEn_{max}$ y h_{max} calculados con $m = 20$. La segunda tasa de error más alta (0.017) se obtiene con $m = 2$. Por otro lado, para $N \geq 1000$ la TEC es cero para todos los valores de m . Al incrementar el nivel de ruido RSR= 5 dB, se ve en la figura 3.6f que los peores resultados son obtenidos con $m = 2$, excepto cuando $N = 10000$. Sin embargo, hay que resaltar que la TEC para $m = 8$ es igual a cero independientemente del valor N . Además el error obtenido utilizando $m = 20$ siempre está por debajo o es igual al error obtenido con valores de m más chicos. Finalmente de la figura 3.6f se puede observar que los peores resultados son obtenidos para $m = 2$ con independencia del tamaño de la serie temporal. Para $m = 3$ la situación anterior solo mejora para valores de $N \geq 8000$. No obstante, para longitudes de datos cortas, la TEC calculada para $m = 8$ y $m = 20$ está siempre por debajo de aquella calculada con valores menores de m .

Los resultados obtenidos con esta simulación nos llevan a concluir que para señales temporales limpias, el utilizar en conjunto los estimadores $ApEn_{max}$ y h_{max} calculados con $m > 2$ permite una adecuada clasificación entre dinámicas de diferente complejidad en sistemas de baja y alta dimensión. Para mantener los buenos resultados en condiciones de ruido vemos que es necesario incrementar la dimensión de inmersión por encima de $m = 3$. Se debe recordar que para un N fijo, utilizar un valor de m muy grande puede llevar estimaciones inconsistentes. Por tanto se sugiere trabajar con una estimación de la mínima dimensión de inmersión.

3.2.2 Aplicaciones con señales reales

Gracias a las señales sintéticas generadas a partir de modelos no lineales pudimos estudiar el comportamiento de los estimadores $ApEn_{max}$ y h_{max} ante diferentes niveles de ruido y longitud de datos. Es menester ahora estudiar el comportamiento de estos estimadores en señales fisiológicas reales.

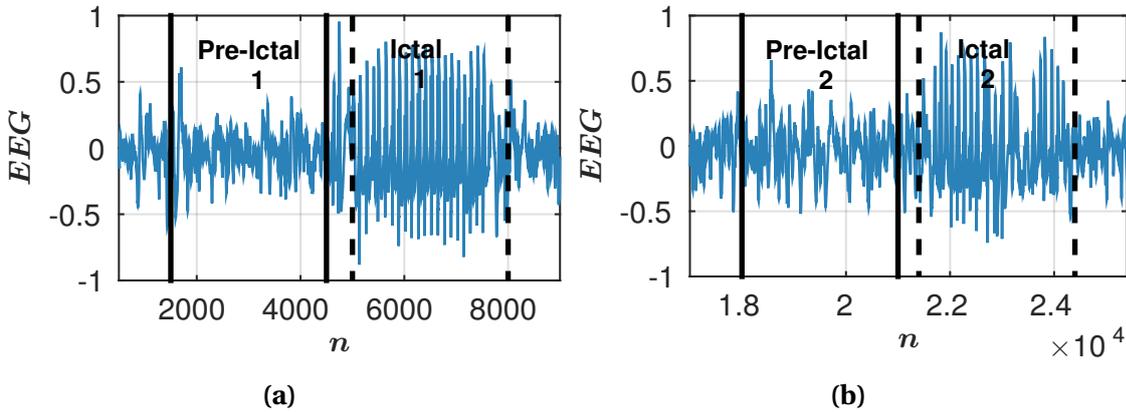


Figura 3.7. Señal de EEG con episodios pre-ictales e ictales: (a) Primeros episodios pre-ictales e ictales. (b) Segundos episodios pre-ictales e ictales.

Detector automático de episodios ictales

En la figura 3.7 se muestran dos porciones de una señal de estereo EEG (ver subsección 1.4.1) obtenida mediante ocho electrodos implantados en el hipocampo de un paciente [1]. La señal fue amplificada y filtrada (entre 1 – 40 Hz) mediante un filtro pasa banda. Se utilizó un filtro pasa bajos de tipo Butterworth de cuarto orden como filtro anti-aliasing. Esta señal fue digitalizada a 250 Hz a través de un CAD de 10 bits. En las figuras 3.7a y 3.7b se presentan dos episodios pre-ictales y dos ictales, cada uno con 3000 datos de longitud. El primer episodio pre-ictal se encuentra en el intervalo $n = [1500, 4500]$ y su consecutivo episodio ictal esta ubicado en el intervalo $n = [5000, 8000]$. Los segundos episodios pre-ictales e ictales se encuentran en los intervalos $n = [18000, 21000]$ y $n = [21400, 24400]$ respectivamente. Según el análisis visual de un médico, el paciente presenta un foco epileptogénico en el área del hipocampo con propagación inmediata a la circunvolución del cíngulo y el área motora suplementaria del hemisferio izquierdo.

Cada episodio fue normalizado para tener energía unitaria. El valor de τ fue elegido utilizando el criterio de la función de información mútua entre las cuatro señales, es decir, se calculó un valor de τ para los cuatro segmentos y luego se eligió el mayor para la simulación ($\tau = 10$). Los estimadores $ApEn_{max}$ y h_{max} fueron calculados para $2 \leq m \leq 20$. De forma adicional, se le agregó a cada episodio ruido blanco Gaussiano hasta alcanzar una RSR= 5 dB (la RSR de la señal original es desconocida).

En la figura 3.8 se presenta el comportamiento de los estimadores $ApEn_{max}$ y h_{max} en función del parámetro m para la señal de EEG con diferentes niveles de ruido. En la figura 3.8a se puede observar que, para el caso de la señal original sin ruido agregado, al utilizar el estimador $ApEn_{max}$ las diferencias entre episodios pre-ictales e

3.2. $ApEn_{max}$, h_{max} y cambios de regularidad

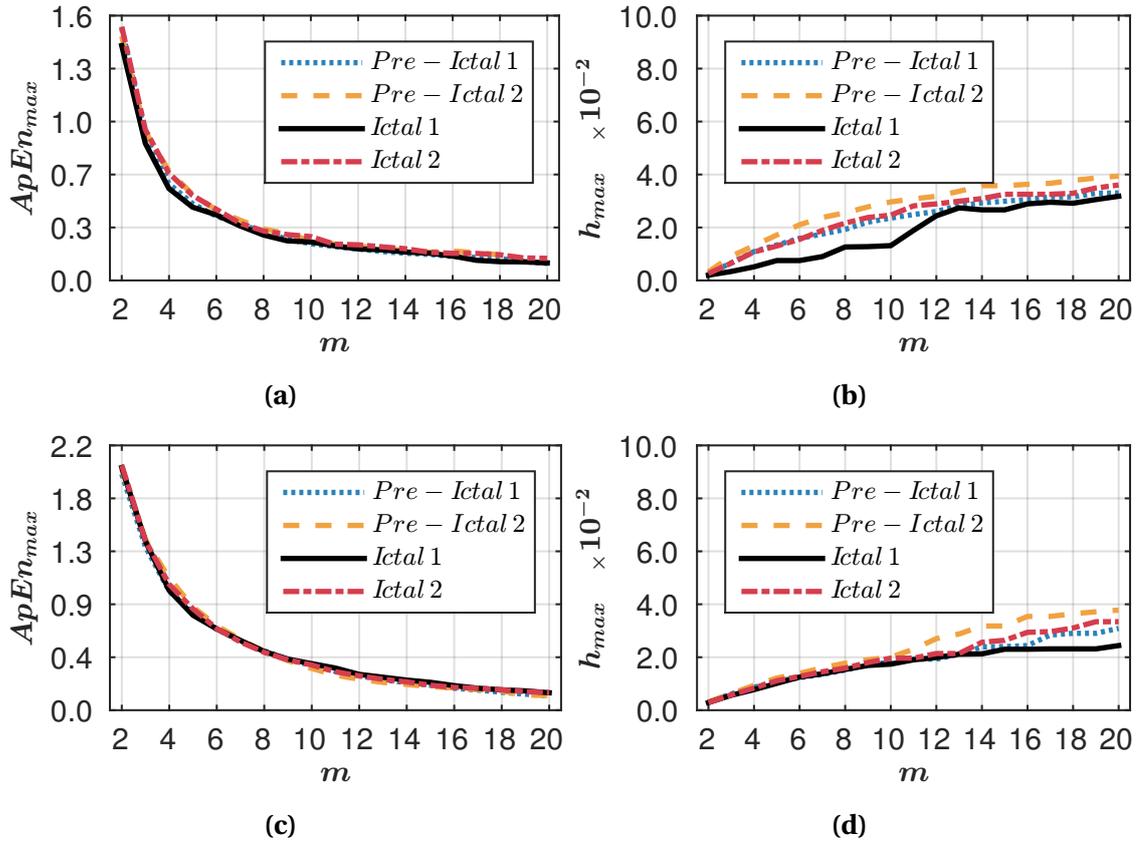


Figura 3.8. Comportamiento de $ApEn_{max}$ y h_{max} en función de m para dos episodios pre-ictales y dos ictales. Señal original: (a) $ApEn_{max}$. (b) h_{max} . RSR= 5 dB (c) $ApEn_{max}$. (d) h_{max} .

ictales son mínimas. Sin embargo, utilizando h_{max} para valores de $m > 2$ se pueden distinguir con claridad entre los episodios pre-ictales y sus respectivos episodios ictales (figura 3.8b). Para valores menores de señal a ruido se observa en las figuras 3.8c y 3.8d que la distancia entre curvas de $ApEn_{max}$ se ha reducido aún más. No obstante utilizando la información provista por h_{max} para valores de $m > 14$ se puede diferenciar los episodios pre-ictales de sus respectivos episodios ictales.

Dado que el método aquí propuesto permite separar los episodios pre-ictales de sus episodios ictales consecutivos, podemos concluir que una posible forma de detectar estos episodios sea evaluando la evolución temporal de los estimadores $ApEn_{max}$ y h_{max} .

Para evaluar esta idea contaminamos la señal de EEG con ruido blanco Gaussiano (RSR= 5 dB) y consideramos una ventana temporal móvil de $N = 1000$ y desplazamiento igual a 128 datos. Normalizamos cada ventana para tener energía unitaria y los estimadores $ApEn_{max}$ y h_{max} los calculamos para $2 \leq m \leq 6$ con $\tau = 10$. Forma-

mos las matrices \mathbf{C}^A y \mathbf{C}^H , donde la entrada $C_{k,i}^A$ es el valor del estimador $ApEn_{max}$ calculado con el i -ésimo valor de m para la k -ésima ventana. De igual manera formamos la matriz \mathbf{C}^H con las estimaciones de h_{max} . Cada matriz la normalizamos por columnas para tener media cero y varianza unitaria. De esta forma la evolución temporal de los estimadores $ApEn_{max}$ y h_{max} calculados con el i -ésimo valor de m se puede evidenciar en la i -ésima columna de las matrices \mathbf{C}^A y \mathbf{C}^H respectivamente. Conformamos una tercera matriz llamada \mathbf{C}^{AH} con la concatenación horizontal de las dos matrices anteriores, es decir, $\mathbf{C}^{AH} = (\mathbf{C}^A | \mathbf{C}^H)$. Vemos que cada matriz cuenta con información de ambos estimadores calculados para varios valores de m . Esto puede dar lugar a información redundante, por tal motivo realizamos un análisis de componentes principales [124]. De este procedimiento se obtuvo la primera componente principal de cada matriz, las cuales representan más del 94 % de la variabilidad de los datos. La misma puede ser vista como un buen resumen de la información recolectada por estos estimadores a través de todos los valores de m . Finalmente para detectar los diferentes episodios ictales utilizamos un algoritmo de detección de cambios en valores medios (CUSUM) [125] sobre la primera componente principal de cada matriz. Para este algoritmo la media objetivo y el valor de referencia los fijamos como el promedio y dos veces el desvío estándar de los primeros 20 datos, respectivamente.

Los resultados de este estudio se presentan en la figura 3.9. Se muestran cuatro episodios ictales (figura 3.9a) marcados entre líneas verticales y los resultados del algoritmo de CUSUM aplicado a la primera componente principal de cada matriz (figura 3.9b). Se puede ver de la figura 3.9b que los cuatro episodios ictales han sido detectados utilizando la información contenida en cada matriz. Se observa que para algunos episodios es mejor utilizar el estimador $ApEn_{max}$ en lugar de h_{max} y en otros se prefiera lo inverso. Un detector de episodios ictales más consistente puede ser implementado utilizando la información provista por ambos estimadores en conjunto. Este resultado sugiere que la habilidad para discriminar entre diferentes dinámicas se incrementa considerablemente al utilizar ambos estimadores en conjunto. Esto se debe a que, eventos que no pueden ser detectados en la evolución temporal de uno de los estimadores, pueden aparecer distinguibles en la evolución temporal del otro.

Finalmente es importante resaltar que los anteriores hallazgos solo sugieren la posibilidad de utilizar $ApEn_{max}$ y h_{max} en la detección de episodios ictales a partir de señales de EEG. Queda como trabajo futuro el evaluar esta metodología estadísticamente.

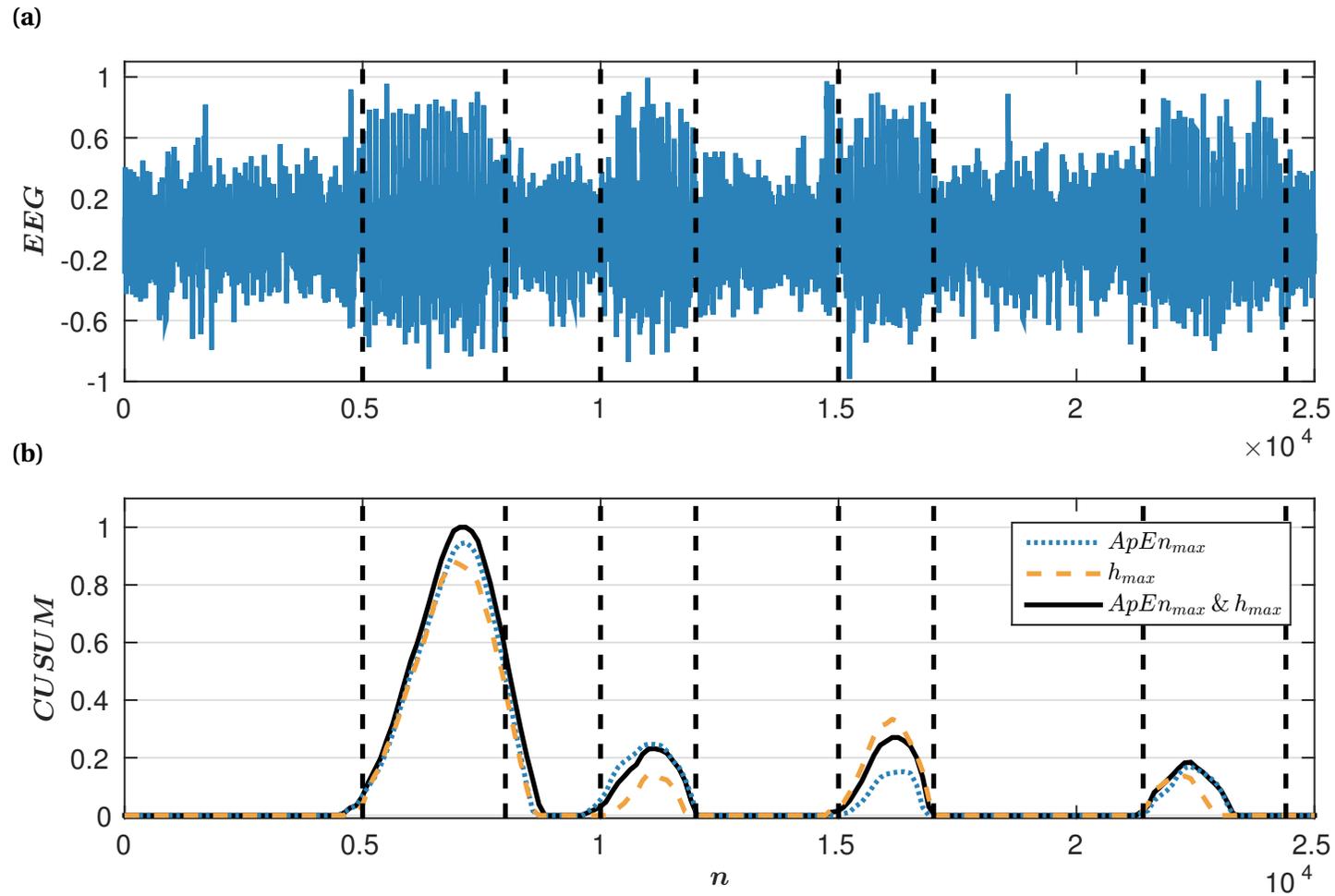


Figura 3.9. Detección de episodios ictales utilizando $ApEn_{max}$ y h_{max} a partir de una señal de EEG con ruido aditivo (RSR= 5 dB). (a) Señal de EEG. (b) Algoritmo de CUSUM aplicado a la primera componente principal de la matriz C^A (azul punteado), a la primera componente de la matriz C^H (naranja discontinuo) y a la primera componente de la matriz C^{AH} (negro). Los episodios ictales se marcan entre líneas verticales.

Detector automático de voces patológicas

Debido a que la mayoría de los trastornos de la laringe causan cambios notables en la señal acústica de la voz, su estudio ha sido ampliamente utilizado por los fonodólogos para diagnosticar y tratar patologías vocales. Gracias a que la adquisición de la señal de voz se hace de forma no invasiva y a muy bajo costo, muchos investigadores han dedicado sus esfuerzos a mejorar el diagnóstico de patologías vocales a partir de esta señal [67]. El complejo comportamiento de la señal de voz durante un estado patológico ha guiado a diferentes investigadores a buscar respuestas en el área de las técnicas de procesamiento no lineal, con el fin de extraer características que permitan una clasificación adecuada entre voces sanas y voces patológicas [68]. Por ejemplo, Torres y col. [126] concluyeron que se pueden establecer diferencias significativas entre sujetos sanos y patológicos utilizando medidas de complejidad tales como: la ApEn, la complejidad de Lempel-Ziv y medidas de información como la q -entropía y la entropía relativa.

Esta ha sido un área fructífera de la cual existe en la literatura científica una amplia gama de trabajos. Debido a esto Sáenz-Lechón y col. [69] han propuesto una serie de consejos metodológicos a tener en cuenta en el diseño de sistemas automáticos de detecciones de voces patológicas, en un intento por unificar aspectos técnicos para poder reproducir y comparar resultados entre las diferentes metodologías. En [34] Vaziri y col. propusieron una metodología para detectar voces patológicas a partir de invariantes como la D , el máximo exponente de Lyapunov, la ApEn, la dimensión de conteo por cajas (dimensión fractal) y la complejidad de Lempel-Ziv. Como resultado ellos reportaron una tasa de reconocimiento del 94% utilizando MSV como clasificador. En [70] Arias-Londoño y col. evaluaron diferentes medidas de complejidad sobre la base de datos de trastornos vocales del “Massachusetts Eye & Ear Infirmary”, distribuida por la “Kay Elemetrics” [62]. En la tabla 3.2 se muestran los resultados de sensibilidad, especificidad y exactitud reportados para diferentes grupos de características, entre los que se destacan la combinación entre D y el máximo exponente de Lyapunov y la combinación entre ApEn y entropía muestral. Es necesario recordar que la sensibilidad es definida como la razón entre el número de voces patológicas clasificadas correctamente (verdadero positivo) y el número total de voces clasificadas como patológicas. La especificidad es la razón entre el número de voces patológicas clasificadas como normales (falso negativo) y el número total de voces clasificadas como normales. La exactitud se define como el número de clasificaciones correctas dividido el número total de clasificaciones realizadas.

En voces patológicas, la señal de voz correspondiente a la emisión de vocales sos-

3.2. $ApEn_{max}$, h_{max} y cambios de regularidad

	Sensibilidad [%]	Especificidad [%]	Exactitud [%]
MEL	76.30	77.40	82.10
D	89.00	84.90	92.30
ApEn	76.30	86.80	84.10
SampEn	84.40	94.40	91.20
GApEn	72.80	79.30	80.00
GSampEn	78.60	75.40	83.32
MEL/D	88.44	98.11	90.71
ApEn/SampEn	87.30	90.71	92.70

Tabla 3.2. Resultados de Arias-Londoño y col. en la clasificación de voces patológicas utilizando diferentes medidas de complejidad: máximo exponente de Lyapunov (MEL), dimensión de correlación (D), entropía aproximada (ApEn), entropía muestral (SampEn), entropía aproximada de núcleo Gaussiano (GApEn), entropía muestral de núcleo Gaussiano (GSampEn). Esta tabla fue extraída de los datos publicados en [70].

tenidas está caracterizada por una irregularidad prominente en comparación con emisiones hechas por sujetos normales [127]. En este ámbito tiene sentido el proponer un detector de voces patológicas basado en los estimadores $ApEn_{max}$ y h_{max} .

Con el fin de incrementar la velocidad en el cálculo de estos estimadores vamos a utilizar un algoritmo de optimización global llamado optimización por enjambre de partículas, en inglés “Particle Swarm Optimization” [128]. La idea detrás de este algoritmo consiste en un grupo (enjambre) de puntos (partículas) que son posicionados aleatoriamente en el dominio de búsqueda de la función objetivo. En cada iteración, cada partícula ajustará su posición y velocidad tomando en cuenta la mejor posición histórica alcanzada por ella misma y la mejor posición alcanzada dentro del enjambre entero. Al final, se conserva la mejor posición alcanzada por el enjambre entero. Para nuestra aplicación se utilizó una versión de este algoritmo propuesta por Zheng y col. [129], en donde la regla de actualización de la velocidad (v_i^k) y la dirección (y_i^k) de la i -ésima partícula en la k -ésima iteración está dada por:

$$\begin{aligned}
 w^k &= \frac{k_{max} - k}{k_{max}} (w_{min} - w_{max}) + w_{max}, \\
 v_i^{k+1} &= w^k v_i^k + \phi_1 u_1^k (pb_i^k - y_i^k) + \phi_2 u_2^k (gb^k - y_i^k), \\
 y_i^{k+1} &= y_i^k + v_i^{k+1},
 \end{aligned}$$

donde pb_i^k es la mejor posición que ha alcanzado la i -ésima hasta la iteración k . La variable gb^k denota la mejor posición alcanzada por el enjambre entero hasta la k -ésima iteración. El coeficiente de inercia w^k se incrementa en cada iteración

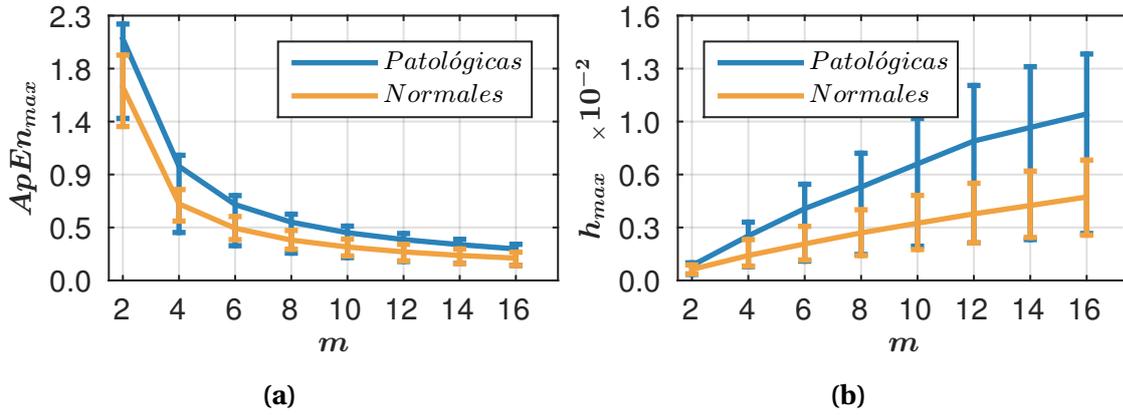


Figura 3.10. Comportamiento de $ApEn_{max}$ y h_{max} en función de m para voces normales y patológicas. (a) $ApEn_{max}$. (b) h_{max} .

desde $w_{min} = 0.4$ hasta $w_{max} = 0.9$. Finalmente, u_1^k y u_2^k son números aleatorios con distribución $\mathcal{U}(0, 1)$ y los parámetros ϕ_1 y ϕ_2 fueron fijados a un valor de 2.

Para esta simulación utilizamos la base de datos de trastornos vocales del “Massachusetts Eye & Ear Infirmary”, distribuida por la “Kay Elemetrics” [62]. Siguiendo [130], escogimos 226 emisiones (53 normales y 173 patológicas) de una emisión de vocal /a/ sostenida. De acuerdo con Parsa y Jamieson [130], las edades y géneros en este subconjunto están uniformemente distribuidos entre ambas clases. Cada grabación fue sub-muestreada a 10 kHz. Cuatro registros fueron excluidos de este estudio por tener menos de 9000 muestras, dejando un total de 222 grabaciones (53 normales y 169 patológicas) para analizar.

Siguiendo la metodología propuesta por nosotros con anterioridad, tomamos una muestra de $N = 9000$ datos de cada registro y la normalizamos para tener energía unitaria. Luego calculamos los estimadores $ApEn_{max}$ y h_{max} para $m = \{2, 3, \dots, 17\}$ usando $\tau = 15$.

La figura 3.10 muestra el comportamiento de los estimadores $ApEn_{max}$ y h_{max} en función del parámetro m para voces normales y patológicas. En la figura 3.10a se presenta la media y un intervalo de confianza del 95% para el estimador $ApEn_{max}$. Se puede apreciar que las voces patológicas presentan un valor medio de entropía más alto que las voces normales. Este resultado concuerda con lo publicado en [127]. La dispersión $ApEn_{max}$ para voces patológicas es mayor que para voces sanas, además, dicha dispersión tiende a decrecer al avanzar sobre m . En la figura 3.10b se presenta la media y un intervalo de confianza del 95% para estimador h_{max} . Es posible observar que este estimador presenta un valor medio mayor para voces patológicas que para voces normales y su dispersión aumenta con m para ambas cla-

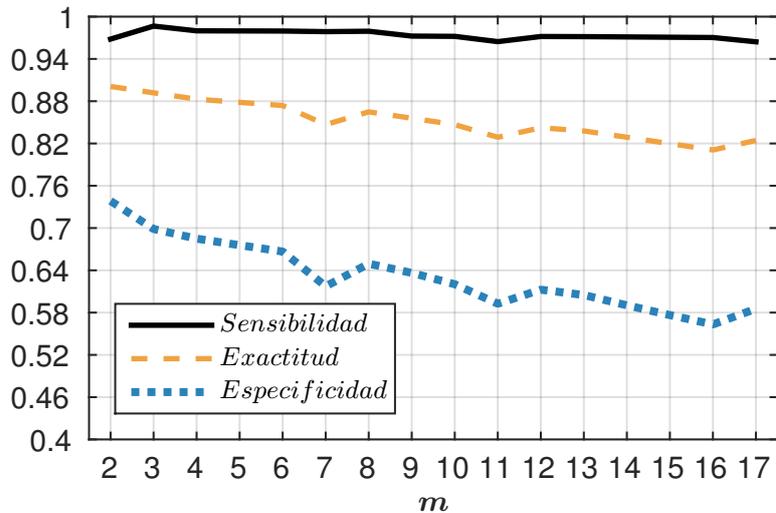


Figura 3.11. Sensibilidad (negro), exactitud (naranja discontinuo) y especificidad (azul punteado) en función de m para un clasificador de voces sanas y patológicas basado en MSV. El clasificador fue entrenado con los estimadores $ApEn_{max}$ y h_{max} en conjunto y probado mediante una validación cruzada de k -particiones.

ses. Es importante resaltar que tanto para $ApEn_{max}$ como para h_{max} , existe un gran solapamiento entre clases, esto es debido a la dispersión de los estimadores calculados para voces patológicas. Lo anterior puede ser debido a la gran heterogeneidad de las patologías presentes en la base de datos.

Con estos resultados en mente hemos diseñado otra simulación para estudiar la capacidad de discriminación de estos estimadores. Mediante una validación cruzada de k -particiones se calculó la sensibilidad, la especificidad y la exactitud de un clasificador basado en MSV de tipo lineal, entrenado con ambos estimadores en conjunto para cada valor de m .

En la figura 3.11 se presenta la sensibilidad, la exactitud y la especificidad en función de m . Estas cantidades fueron calculadas mediante una validación cruzada de k -particiones. Es posible observar que la sensibilidad está por encima del 95% para todos los valores de m , alcanzando su máximo (98.7%) para $m = 3$. Por otro lado, la especificidad decrece al incrementarse el valor de m alcanzando un valor máximo de 73.6% y un valor mínimo de 56.3% para $m = 2$ y $m = 16$ respectivamente. El resultado anterior se ve reflejado en la exactitud, que decrece al incrementarse m , alcanzando un máximo de 90.1% en $m = 2$ y un mínimo de 81.1% para $m = 16$. Estos resultados sugieren que la mejor forma para clasificar voces patológica con $ApEn_{max}$ y h_{max} es utilizar toda la información provista por estos estimadores para todos los valores de m .

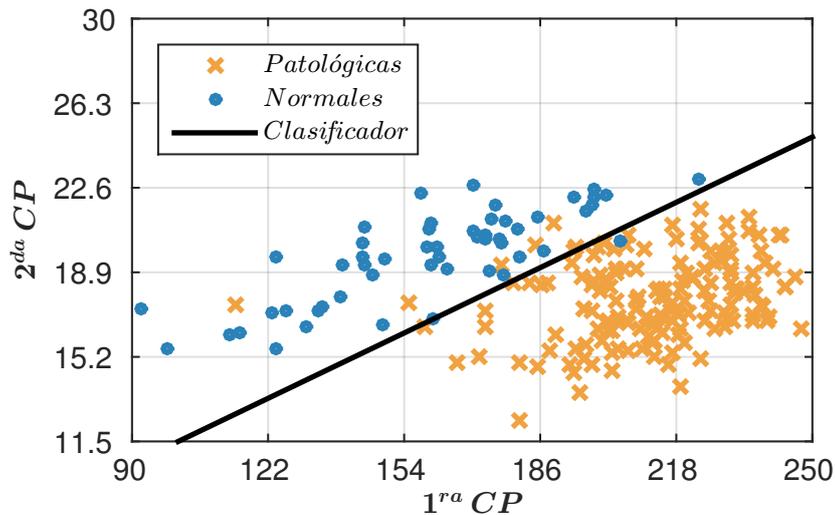


Figura 3.12. Detector automático de voces patológicas. Gráfico de dispersión para voces patológicas y sanas luego de una reducción de dimensionalidad mediante el discriminante lineal de Fisher y una proyección en dirección de las dos primeras componentes principales.

En [70] y [71] Arias-Londoño y colaboradores (y col.) analizaron la capacidad de discriminación de 11 estimadores de complejidad sobre la misma base de datos. Entre estos se encuentran el máximo exponente de Lyapunov, D , la $ApEn$, la $SampEn$ y la entropía aproximada de núcleo Gaussiano. De sus resultados surge que el mayor valor de sensibilidad (89%) y de exactitud (88%) se logró utilizando la dimensión de correlación. De este estudio, también se observa que el mayor valor para la especificidad (94.3%) se alcanza utilizando la $SampEn$. Los resultados alcanzados utilizando en conjunto los estimadores $ApEn_{max}$ y h_{max} sugieren que el uso de estos estimadores tienen un mejor desempeño (por lo menos en términos de sensibilidad y exactitud) que cada uno de los estimadores evaluados por separado en [70] y [71].

Gracias a las simulaciones anteriores hemos podido estudiar y comprender el comportamiento de los estimadores $ApEn_{max}$ y h_{max} sobre señales de voz de sujetos sanos y con alguna patología. Con esta información se diseñó un detector de voces patológicas. Este consta de tres etapas principales: la primera comprende el cálculo de ambos estimadores para diferentes valores del parámetro m . En este estudio se obtuvo un total de 32 características (un valor de $ApEn_{max}$ y uno de h_{max} por cada valor de m) para cada señal de voz. El siguiente paso consiste en una reducción de dimensionalidad utilizando el discriminante lineal de Fisher [124] seguido de una proyección en la dirección de las dos primeras componentes principales. Para

		Condición		Precisión
		Patológicas	Normales	
Test	Patológicas	160	9	0.947
	Normales	3	50	
		0.982	0.847	0.946
		Sensibilidad	Especificidad	Exactitud

Tabla 3.3. Matriz de confusión para el detector automático de voces normales y patológicas aquí propuesto.

finalizar se utilizó un MSV de tipo lineal como clasificador.

Se presenta en la figura 3.12 una gráfica de dispersión para los grupos patológicos y sanos. Es posible ver como, en general, con la metodología descrita podemos lograr que ambas clases se agrupen de una forma conveniente para su clasificación. La evaluación cuantitativa se resume en la tabla 3.3 donde se presenta la matriz de confusión calculada a partir de una validación cruzada de una iteración.

Se observa en la tabla 3.3 que con esta metodología podemos alcanzar una sensibilidad final del 98.2%, una especificidad total del 84.7% y una exactitud total del 94.6%. Debemos resaltar que, comparado con los resultados de la figura 3.11, cada una de estas cantidades ha incrementado su valor. Esto sugiere que es muy conveniente utilizar la información provista por $ApEn_{max}$ y h_{max} en conjunto calculados para varios valores de m . En comparación con otros estudios sobre la misma base de datos (ver tabla 3.2), es posible concluir que en términos de sensibilidad y exactitud el uso de $ApEn_{max}$ y h_{max} presenta un mejor desempeño que otros estimadores de complejidad.

3.3 Conclusiones

La entropía aproximada es reconocida por su habilidad para diferenciar distintas dinámicas a partir de series temporales de corta longitud e inmersas en ruido de amplitud moderada. Sin embargo, la presencia de grandes cantidades de ruido y una incorrecta selección de los parámetros m y h puede afectar gravemente su capacidad de discriminación. Para superar estas dificultades hemos propuesto el uso conjunto de los estimadores $ApEn_{max}$ y h_{max} como características para discernir entre dinámicas. Mediante simulaciones con sistemas dinámicos de baja y alta dimensión hemos estudiado el comportamiento de estos estimadores en función de

Capítulo 3. Entropía aproximada y cambios de regularidad

la dimensión de inmersión y la longitud de datos en escenarios con diferente cantidad de ruido. De los resultados podemos concluir que h_{max} provee valiosa información, útil para fines de clasificación. Gracias a la relación de complementariedad que presentan estos estimadores al variar el parámetro m , el uso conjunto de $ApEn_{max}$ con h_{max} proporciona una metodología robusta frente a cambios en dicho parámetro. Esto es confirmado por los resultados de clasificación obtenidos mediante validación cruzada.

En presencia de ruido, estos estimadores conservan su relación de complementariedad permitiendo la clasificación de series temporales con baja relación señal a ruido. Es importante resaltar que bajo estas condiciones los mejores resultados son obtenidos utilizando los estimadores $ApEn_{max}$ y h_{max} calculados con valores de m mayores a 3. Esto contradice lo reportado en la literatura donde se aconseja que $m = 2$ o $m = 3$ y que h sea igual a un valor entre 0.1 – 0.3 veces la desviación estándar de la señal.

En base a los resultados en series temporales sintéticas, hemos diseñado un detector automático de episodios ictales a partir de señales de EEG y un detector automático de patologías vocales a partir de señales de voz. El núcleo central de estos algoritmos es el uso conjunto de los estimadores $ApEn_{max}$ y h_{max} calculados para varios valores de m . Los resultados de estos algoritmos dan testimonio de la utilidad de estos estimadores en aplicaciones biomédicas con señales fisiológicas.

Aún queda por investigar la relación del estimador h_{max} con la complejidad de un sistema dinámico. Queda entonces abierta la pregunta de si es posible que este estimador pueda ser utilizado por sí solo para saber si una dinámica es más compleja que otra. La respuesta a esta pregunta nos ayudará a revelar el verdadero significado físico de h_{max} .

Los resultados mostrados en este capítulo fueron publicados en [1-4].

El heroísmo no consiste en carecer de miedo, sino en superarlo.

Roberto Gómez Bolaños

4

Integral de correlación asistida por ruido

En este capítulo presentaremos las bases teóricas que soportan el desarrollo de la integral de correlación asistida por ruido (ICAR) y demostraremos cómo la integral de correlación estándar [81] y la ICG [16] pueden ser consideradas como casos particulares de la ICAR.

4.1 Introducción

En la subsección 2.5.2 describimos el algoritmo de correlación propuesto por Grassberger y col. [82] (ver figura 2.4). Mencionamos además que este algoritmo puede ser visto como un arreglo de comparadores en paralelo, similar al implementado en los CADs [131]. Se sabe que el desempeño de los CADs puede ser mejorado a través de la técnica de “dithering” que consiste en agregar una señal aleatoria (generalmente ruido blanco Gaussiano) a la señal de entrada antes de ser digitalizada [132]. El mejoramiento del desempeño de la conversión analógico digital a causa de la adición de ruido se conoce como resonancia estocástica [133, 134]. La resonancia estocástica es un fenómeno no lineal por medio del cual la adición de ruido puede mejorar el desempeño de un sistema de procesamiento de señales [135]. En nuestro

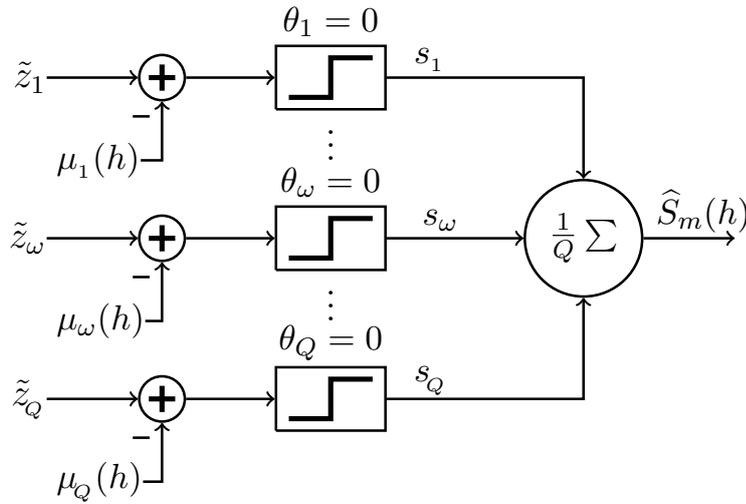


Figura 4.1. Algoritmo de correlación asistido por ruido: Consiste en Q comparadores con umbral $\theta = 0$. La entrada a cada comparador está compuesta por la resta de una realización de distancia \tilde{z}_ω entre vectores de estado m -dimensionales y una realización de ruido blanco μ_ω cuya FDA depende del parámetro h . La salida binaria de todos los comparadores s_ω se promedia para producir la suma de correlación asistida por ruido $\hat{S}_m(h)$.

caso, consideramos cómo el algoritmo de correlación se relaciona estrechamente con el modelo de resonancia estocástica supraumbral [136].

El modelo de resonancia estocástica supraumbral consiste en un arreglo de comparadores en paralelo que producen una salida binaria. A cada comparador le ingresa como entrada una copia de una señal (puede ser determinística [135] o estocástica [137]) más una realización de ruido independiente e idénticamente distribuida (iid), también independiente de la señal de entrada. La respuesta global del modelo puede ser calculada como el promedio de la salida de todos los comparadores. El fenómeno de resonancia estocástica es observado debido a que el ruido agregado produce diferentes respuestas en cada comparador ante una misma entrada. Esto, a su vez, induce una mejor representación de la información contenida en la señal de entrada [135, 137, 138].

Chapeau-Blondeau y col. demostraron en [137] que la estimación de parámetros estadísticos de una señal aleatoria es optimizada mediante el modelo de resonancia estocástica supraumbral. Esto nos indujo a pensar en la posibilidad de mejorar la estimación de los invariantes \mathbf{D} , \mathbf{K}_2 y σ a través del fenómeno de resonancia estocástica.

4.2 Deducción analítica

Para explorar esta idea planteamos una versión modificada del algoritmo de correlación inspirada en el modelo de resonancia estocástica supraumbral. En la figura 4.1 mostramos el algoritmo de correlación asistido por ruido. Este consiste en Q comparadores binarios con umbral $\theta = 0$. A cada comparador le ingresa como entrada una realización de una variable aleatoria $\tilde{z}_\omega \sim f_{\tilde{z}}^{(m)}(\tilde{z})$, donde \tilde{z}_ω representa una realización de la distancia entre dos vectores de estado m -dimensionales, menos una realización de ruido iid μ_ω con función de distribución acumulada (FDA) F_μ , también independiente de \tilde{z}_ω . La salida de cada cuantizador s_ω está regida por:

$$s_\omega = \begin{cases} 0, & \text{si } \tilde{z}_\omega - \mu_\omega \geq \theta_\omega, \\ 1, & \text{en otro caso.} \end{cases} \quad (4.1)$$

La salida global del algoritmo $\widehat{S}_m(h)$ se calcula como el promedio entre la salida de todos los comparadores.

Podemos estimar la probabilidad de que la salida del ω -ésimo comparador sea igual a cero dada la entrada \tilde{z}_ω como:

$$\begin{aligned} \Pr\{s_\omega = 0 | \tilde{z}_\omega\} &= \Pr\{\tilde{z}_\omega - \mu_\omega \geq \theta_\omega\} \\ &= \Pr\{\tilde{z}_\omega - \mu_\omega \geq 0\} \\ &= \Pr\{\mu_\omega \leq \tilde{z}_\omega\} \\ &= F_\mu(\tilde{z}_\omega), \end{aligned}$$

y la probabilidad de que la salida sea igual a uno dada la entrada \tilde{z}_ω como:

$$\begin{aligned} \Pr\{s_\omega = 1 | \tilde{z}_\omega\} &= 1 - \Pr\{s_\omega = 0 | \tilde{z}_\omega\} \\ &= 1 - F_\mu(\tilde{z}_\omega). \end{aligned}$$

Si \widehat{S} denota el número de comparadores encendidos (salida igual a uno) entonces la probabilidad de que $\widehat{S} = k$ dada la entrada \tilde{z}_ω está dada por la distribución binomial:

$$\Pr\{\widehat{S} = k | \tilde{z}_\omega\} = \binom{Q}{k} [1 - F_\mu(\tilde{z}_\omega)]^k F_\mu(\tilde{z}_\omega)^{Q-k}.$$

Por tanto, la probabilidad marginal de tener k comparadores encendidos es:

$$\Pr\{\widehat{S} = k\} = \int_0^\infty \binom{Q}{k} [1 - F_\mu(\tilde{z}_\omega)]^k F_\mu(\tilde{z}_\omega)^{Q-k} f_{\tilde{z}}^{(m)}(\tilde{z}_\omega) d\tilde{z}_\omega,$$

y el valor esperado de \widehat{S} se puede calcular como:

$$E[\widehat{S}] = Q \int_0^{\infty} [1 - F_{\mu}(\tilde{z}_{\omega})] f_{\tilde{z}}^{(m)}(\tilde{z}_{\omega}) d\tilde{z}_{\omega}.$$

Definiendo el estimador $S_m(h) = E[\widehat{S}]/Q$, haciendo el cambio de variable $\tilde{z} = \tilde{z}_{\omega}$ y suponiendo que la distribución del ruido agregado depende de un parámetro h , es decir $F_{\mu}(\tilde{z}_{\omega}) = F_{\mu}(\tilde{z}_{\omega}; h)$, podemos escribir la ecuación anterior como:

$$\begin{aligned} S_m(h) &= E[\widehat{S}]/Q \\ &= \int_0^{\infty} [1 - F_{\mu}(\tilde{z}; h)] f_{\tilde{z}}^{(m)}(\tilde{z}) d\tilde{z} \\ &= \int_0^{\infty} \tilde{F}_{\mu}(\tilde{z}; h) f_{\tilde{z}}^{(m)}(\tilde{z}) d\tilde{z}, \end{aligned} \quad (4.2)$$

donde $\tilde{F}_{\mu}(t; h)$ es la función de distribución acumulada complementaria (FDAC) del ruido agregado. Llamaremos la ecuación (4.2) la *integral de correlación asistida por ruido*.

Es posible observar que la ecuación (4.2) es similar a la ecuación (2.4) del capítulo 2. La diferencia radica en que ahora la función núcleo es $g(\tilde{z}; h) = \tilde{F}_{\mu}(\tilde{z}; h)$. En otras palabras, el ruido agregado al algoritmo de correlación induce una integral de correlación cuya función núcleo depende de la naturaleza del ruido utilizado.

Recordemos que en la sección 2.5 presentamos las condiciones que debe cumplir una función núcleo para poder estimar invariantes a partir la integral de correlación. Esta claro que toda FDAC es monótona decreciente lo cual cumple con la primera de las condiciones. La segunda condición (ecuación (2.9)) es cumplida por distribuciones con soporte compacto como la distribución uniforme o por distribuciones cuya FDAC decaiga igual o más rápido que la función exponencial. Quedan excluidas entonces funciones de distribución de cola pesada como la distribución normal logarítmica o la distribución de Pareto, entre otras.

Algo muy interesante es que tanto la integral de correlación estándar (propuesta por Grassberger y Procaccia) como la ICG (propuesta por Diks) pueden ser consideradas casos particulares de la ICAR.

4.2.1 Caso particular 1: Integral de correlación estándar

La integral de correlación estándar [81] puede ser calculada mediante el algoritmo de correlación asistido por ruido si agregamos ruido μ_ω con distribución determinística o degenerada. Esta es la distribución que tendría una variable aleatoria discreta cuyo valor es constante sin importar el evento, por ejemplo un dado con todas sus caras marcadas con el mismo número. La FDP de esta variable aleatoria es $f_\mu(\tilde{z}; h) = \delta[\tilde{z}/h - 1]$ y su FDA es $F_\mu(\tilde{z}; h) = H(\tilde{z}/h - 1)$. Por tanto es posible reescribir la ecuación (4.2) como:

$$\begin{aligned} S_m(h) &= \int_0^\infty [1 - H(\tilde{z}/h - 1)] f_{\tilde{z}}^{(m)}(\tilde{z}) d\tilde{z} \\ &= \int_0^\infty H(1 - \tilde{z}/h) f_{\tilde{z}}^{(m)}(\tilde{z}) d\tilde{z} \\ &= C_m(h). \end{aligned}$$

Esta es la integral de correlación estándar (ver ecuación (2.5)).

4.2.2 Caso particular 2: Integral de correlación de núcleo Gaussiano

La función núcleo de esta integral de correlación puede ser inducida utilizando ruido con distribución de Rayleigh $\mu_\omega \sim R(\tilde{z}; \sqrt{2}h)$. Esta es la distribución que sigue una variable aleatoria definida como la distancia euclídea entre dos vectores aleatorios bidimensionales iid con distribución normal multivariada de media cero y varianza $2h^2$ [139]. Dado que su FDAC es $\tilde{F}_\mu(\tilde{z}; \sqrt{2}h) = e^{-\tilde{z}^2/4h^2}$, podemos reescribir la ecuación (4.2) como:

$$\begin{aligned} S_m(h) &= \int_0^\infty e^{-\tilde{z}^2/4h^2} f_{\tilde{z}}^{(m)}(\tilde{z}) d\tilde{z} \\ &= T_m(h). \end{aligned}$$

Es muy importante resaltar que estas dos resultados analíticos demuestran que tanto la $C_m(h)$ (ecuación (2.5)) como la $T_m(h)$ (ecuación (2.10)) son casos particulares de la ICAR, lo que indica que pueden ser inducidas agregando ruido μ_ω al algoritmo de correlación. En la siguiente sección ilustraremos de forma numérica estos resultados.

4.3 Simulaciones

Para ilustrar numéricamente los resultados analíticos obtenidos en la sección anterior, estimaremos los invariantes \mathbf{D} y \mathbf{K}_2 a partir de la ICG y de la ICAR con ruido de Rayleigh. Presentamos a continuación el algoritmo para calcular la suma de de correlación asistida por ruido:

Utilizaremos el algoritmo [3] con series temporales provenientes de la transformación de Hénon (ecuación [1.3]) con $a = 4$ y $b = 0.3$. Este es uno de los sistema discretos de dos variables más simples que muestran comportamiento caótico. Gracias a esto, ha sido ampliamente utilizado para comparar resultados en el cálculo de invariantes [16, 82, 87].

Algoritmo 3. Algoritmo de correlación asistido por ruido.

- 1: Formar los vectores de estado m -dimensionales a partir de la serie temporal y calcular todas las distancias \tilde{z}_ω con $1 \leq \omega \leq Q = L(L-1)$ entre pares de vectores de estado.
- 2: Fijar el valor de h y obtener Q realizaciones de ruido μ_ω con distribución $f_\mu(\mu; h)$.
- 3: Calcular la variable binaria:

$$s_\omega(h) = \begin{cases} 0, & \text{si } \tilde{z}_\omega - \mu_\omega \geq 0, \\ 1, & \text{en otro caso.} \end{cases}$$

- 4: Obtener la suma de correlación asistida por ruido como:

$$\hat{S}_m(h) = \frac{1}{Q} \sum_{\omega=1}^Q s_\omega(h).$$

- 5: Repetir los pasos 1–4 para todos los valores de h y m .
-

En la figura [4.2] se muestra una gráfica del logaritmo de la ICAR con ruido de Rayleigh y de la ICG versus el $\ln(h)$. Tanto la ICG como la ICAR se calcularon a partir de una serie temporal de longitud $N = 5000$ datos, en ausencia de ruido, con retardo de inmersión $\tau = 1$ y para diferentes valores de $m = \{2, 4, 6, 8, 10\}$. La serie temporal fue normalizada para tener media cero y varianza unitaria. Es posible observar como la ICAR oscila alrededor de la ICG para valores pequeños de h . Sin embargo para valores mayores de h ambas integrales de correlación convergen a los mismos valores. Tanto las curvas correspondientes a la ICG como aquellas correspondientes a la ICAR presentan la misma pendiente. Esto significa que ambas deben converger al mismo valor de \mathbf{D} . Esto puede verificarse en la figura [4.3] donde se presenta la de-

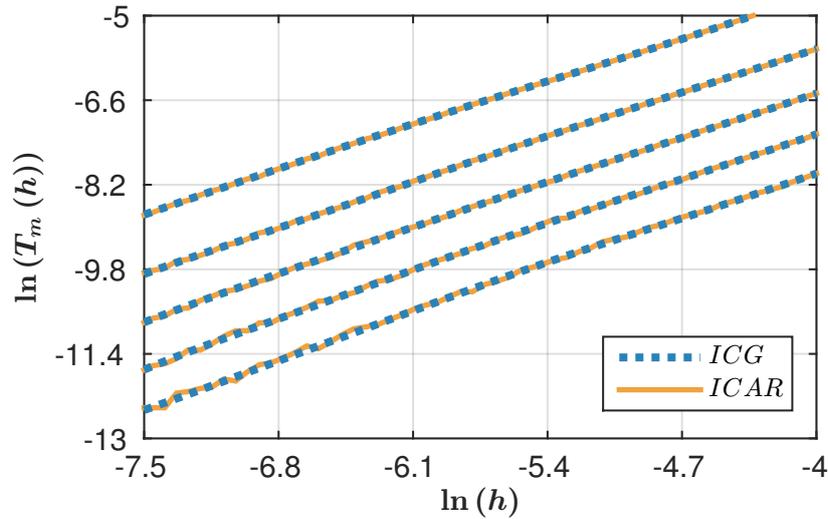


Figura 4.2. Gráfica log-log de la integral de correlación para la transformación de Hénon sin ruido. Se presentan las curvas calculadas con $m = \{2, 4, 6, 8, 10\}$ ($m = 2$ curva superior y $m = 10$ curva inferior) para la ICG (azul punteado) y la ICAR con ruido de Rayleigh (naranja).

rivada logarítmica de la ICG y de la ICAR en función de $\ln(h)$. Vemos como ambos estimadores oscilan alrededor del valor reportado de $D = 1.22$ [16, 87]. Además, se aprecia como esta oscilación es mayor para valores pequeños de h que para valores más grandes.

La estimación de K_2 se muestra en la figura 4.4. Es posible observar que las curvas pertenecientes a K_2 calculadas a partir de la ICAR oscilan alrededor de aquellas calculadas a partir de la ICG. También se aprecia como ninguno de los dos estimadores converge al valor reportado $K_2 = 0.3$ [16, 87]. Esto muestra que, para estimar K_2 a partir de la integral de correlación de núcleo Gaussiano, es necesario trabajar con valores muy grandes de m .

Las oscilaciones presentadas en la estimación de los invariantes a partir de la ICAR son causadas porque el número de vectores de estado disponibles para su cálculo es limitado. Por tanto, éstas pueden ser reducidas de dos maneras distintas. La primera es aumentar la longitud de la serie temporal, lo cual no siempre es posible. La segunda es crear copias de cada distancia \tilde{z}_ω y agregarles una realización de ruido diferente. Esto incrementará el número de comparadores Q . Luego se debe proceder con el umbralado para finalmente obtener la integral de correlación asistida por ruido $\hat{S}_m(h)$ como el promedio de la salida de todos los comparadores. Hemos observado que la amplitud de las oscilaciones decrece al incrementar el número de copias.

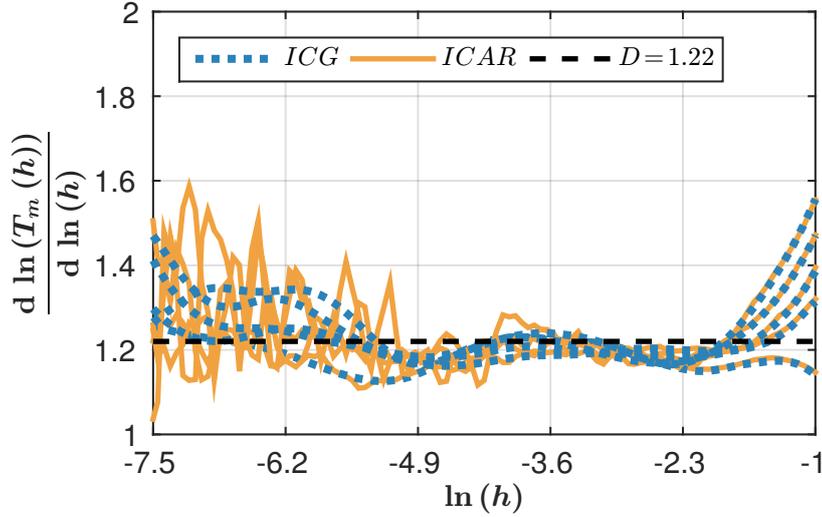


Figura 4.3. Estimación de D para la transformación de Hénon en ausencia de ruido. Se presenta la derivada logarítmica de la ICG (azul punteado) y de la ICAR con ruido de Rayleigh (naranja) para $m = \{2, 4, 6, 8, 10\}$. El valor reportado de D se muestra en línea negra discontinua.

Como mencionamos en la sección [2.5](#) las series temporales utilizadas en todas las simulaciones de este documento han sido normalizadas para tener varianza unitaria. Por tanto σ es el nivel de ruido después del escalado:

$$\sigma = \frac{\sigma_n}{\sqrt{\sigma_s^2 + \sigma_n^2}},$$

donde σ_n es la varianza del ruido y σ_s es la varianza de la serie temporal limpia. En consecuencia $\sigma = 0$ correspondería a una serie temporal limpia y $\sigma = 1$ implicaría una serie temporal compuesta solo por muestras de ruido independientes e idénticamente distribuidas.

En la figura [4.5](#) se muestra una gráfica del logaritmo de la ICAR y de la ICG versus el $\ln(h)$. Ambas integrales de correlación se calcularon a partir de una serie temporal de longitud $N = 5000$ datos, con un nivel de ruido $\sigma = 0.05$, con $\tau = 1$ y para diferentes valores de $m = \{2, 4, 6, 8, 10\}$. Es posible observar cómo para valores pequeños de h , tanto la ICG y la ICAR presentan una pendiente más elevada que para valores más grandes de h . Esta es la marca de la presencia de ruido en la serie temporal. Además, se puede ver que ambas integrales de correlación convergen a los mismos valores. Esto da más evidencias que demuestran que la ICG puede ser inducida utilizando ruido de Rayleigh en el algoritmo de correlación asistido por ruido.

La mayor diferencia entre el algoritmo de correlación (figura [2.4](#)) y el algoritmo de

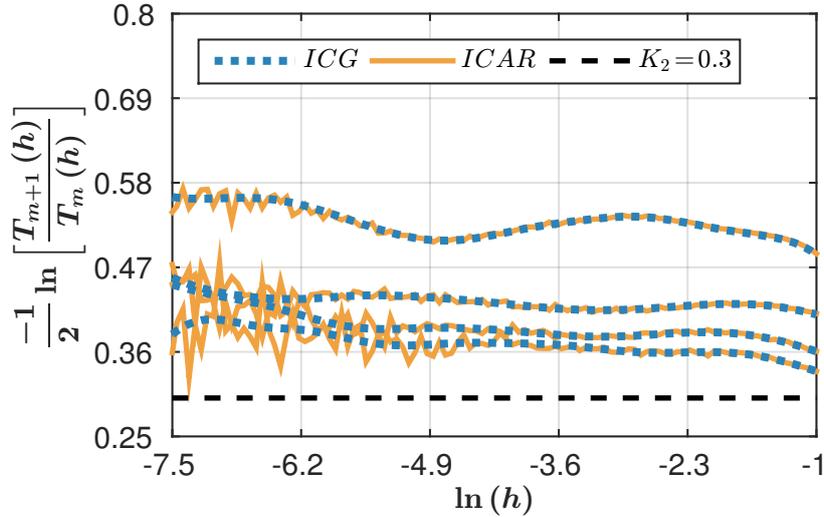


Figura 4.4. Estimación de K_2 para la transformación de Hénon en ausencia de ruido. Se presenta el funcional de entropía de la ICG (azul punteado) y de la ICAR con ruido de Rayleigh (naranja) para $m = \{2, 4, 6, 8\}$. El valor reportado de K_2 se muestra en línea negra discontinua.

correlación asistido por ruido se basa en la naturaleza del umbral utilizado para comparar las distancias \tilde{z}_ω . El primero utiliza un umbral determinístico y el segundo utiliza un umbral estocástico. El ruido μ_ω agregado a las distancias \tilde{z}_ω es equivalente a comparar dichas distancias con un umbral cuya distribución es igual a la del ruido.

Una ventaja del algoritmo asistido por ruido está relacionada con la evaluación de la función núcleo. Con el fin de estimar la integral de correlación es necesario calcular la suma de correlación lo cual demanda muchas evaluaciones de dicha función. Para funciones núcleo que requieran integración numérica para su evaluación, el algoritmo de correlación clásico resulta ser computacionalmente costoso. Sin embargo, si esta función núcleo puede ser escrita como la FDAC de alguna distribución, sería posible utilizar el algoritmo asistido por ruido para calcular la suma de correlación.

Uno de los mayores problemas de la ICG es que para estimar K_2 necesita valores muy altos de m [15]. Esto es por que su función núcleo, $e^{-\tilde{z}^2/4h^2}$, es independiente del valor de m utilizado para reconstruir los vectores de estado. Si la distribución de las distancias entre vectores de estado, $f_{\tilde{z}}^{(m)}(\tilde{z})$, cambia para cada valor de m , también lo debería hacer la función núcleo. Sabemos que la distribución de Rayleigh es un caso particular de la distribución Chi (χ_β) cuando $\beta = 2$. La χ_β es la distribución que sigue la distancia euclídea entre dos vectores aleatorios β -dimensionales iids

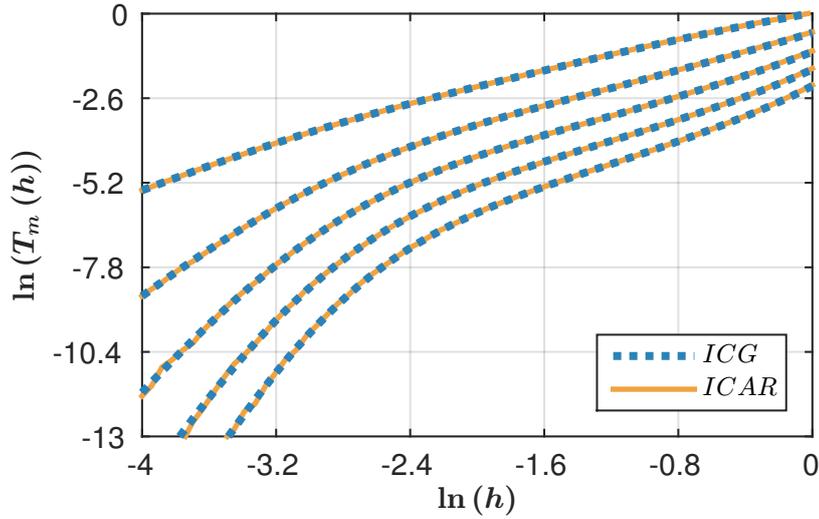


Figura 4.5. Gráfica log-log de la integral de correlación para la transformación de Hénon con ruido ($\sigma = 0.05$). Se presentan las curvas calculadas con $m = \{2, 4, 6, 8, 10\}$ ($m = 2$ curva superior y $m = 10$ curva inferior) para la ICG (azul punteado) y la ICAR con ruido de Rayleigh (naranja).

con distribución $\mathcal{N}(\mathbf{0}, \Sigma^2)$. Esto nos lleva a pensar que es posible generalizar la ICG utilizando ruido con distribución χ_β en el algoritmo de correlación asistido por ruido. Este tipo de ruido nos permitiría inducir una función núcleo con la capacidad de incorporar información de la dimensión de inmersión, lo que no es posible utilizando el núcleo Gaussiano. Esto podría ayudar en la estimación de la entropía de correlación. Exploraremos esta idea en el siguiente capítulo.

4.4 Conclusiones

En este capítulo hemos presentado las bases analíticas de la ICAR. Se fundamentó su origen en el algoritmo de correlación asistido por ruido, inspirado en el modelo de resonancia estocástica supraumbral. Se demostró analíticamente que la integral de correlación estándar y la ICG pueden ser considerados casos particulares de la ICAR. Los hallazgos analíticos han sido verificados experimentalmente utilizando series temporales proveniente de un sistema dinámico caótico. Es importante resaltar que estos resultados muestran que agregar ruido a las distancias entre vectores de estado en el algoritmo de correlación induce una función núcleo. Podemos decir que la ICAR abre la posibilidad de incorporar funciones de núcleo diferentes al Gaussiano y cuya evaluación directa requiere de métodos numéricos. Esto permitirá analizar analíticamente fenómenos físicos que se alejen de la gaussianidad.

Let the future tell the truth, and evaluate each one according to his work and accomplishments. The present is theirs; the future, for which I have really worked, is mine.

Nikola Tesla

5

Integral de correlación U

Introduciremos un caso particular de la ICAR llamado la integral de correlación U (ICU). Deduciremos su expresión analítica y mostraremos que puede ser interpretada como una generalización de la ICG. Deduciremos también los funcionales que permitirán estimar la \mathbf{D} , la \mathbf{K}_2 y el σ a partir de esta integral de correlación. Finalmente los resultados analíticos serán comprobados con evidencia numérica.

5.1 Introducción

En la sección [2.5](#) mostramos que la ICG presenta serios problemas en la estimación de la entropía de correlación. Estos problemas radican en que es necesario un valor de dimensión de inmersión muy alto para poder converger, incluso en ausencia de ruido [\[15\]](#).

La ICU tiene su origen en la ICAR y puede interpretarse como una generalización de la ICG. En el capítulo [4](#) vimos que la diferencia entre el algoritmo de correlación clásico [\[82\]](#) (ver figura [2.4](#)) y el asistido por ruido (ver figura [4.1](#)) radica en la naturaleza del umbral utilizado para comparar las distancias \tilde{z}_ω entre vectores de estado. El primero utiliza un umbral determinístico, mientras que el ruido μ_ω utilizado en el segundo induce un umbral estocástico con igual distribución. Se demostró también

en ese capítulo que la ICG puede ser inducida utilizando ruido con distribución de Rayleigh.

El cálculo de la integral de correlación involucra la función núcleo $g(\tilde{z}; h)$ y la distribución de la distancia euclídea \tilde{z} entre vectores de estado $f_{\tilde{z}}^{(m)}(\tilde{z})$ (ver sección 2.5). Es natural pensar que esta última sea dependiente de la dimensión de inmersión m utilizada para reconstruir dichos vectores. En este sentido el cálculo de la ICG (ecuación 2.10) para diferentes valores del parámetro m involucra una distribución que cambia con m y una función núcleo $g(\tilde{z}; h) = e^{-\tilde{z}^2/4h}$ que es independiente de este parámetro. En consecuencia de esto, la ICG necesita valores muy altos de m para estimar K_2 [15]. El número de vectores necesarios para poder estimar $f_{\tilde{z}}^{(m)}(\tilde{z})$ crece a medida que se incrementa m . Si para $m = 1$ son necesarios L vectores de estado, para $m = 2$ serían necesarios L^2 vectores de estado. El problema es que en aplicaciones reales siempre se cuenta con un número finito de datos.

Podemos pensar que si la distribución de la distancia \tilde{z} es una función de la dimensión de inmersión, así también lo debería ser la función núcleo. Por tanto, la hipótesis a tratar en este capítulo es que la incorporación de información sobre la dimensión de inmersión en la función núcleo puede ayudar a estimar los invariantes D , K_2 y σ .

5.2 Deducción analítica

Para poder trabajar sobre la hipótesis planteada, es necesario primero proponer una función núcleo que permita incorporar información concerniente a la dimensión de inmersión mediante uno de sus parámetros. Dicha función es:

$$g(z; h) = \frac{\Gamma(\beta/2, z/h^2)}{\Gamma(\beta/2)}, \quad (5.1)$$

donde $\Gamma(a, t)$ es la función Gamma Incompleta Superior, $\Gamma(a)$ es la función Gamma y $z = \tilde{z}^2$ es el cuadrado de la distancia euclídea \tilde{z} . Esta función es la FDAC de la distribución Chi cuadrada (χ_{β}^2) con β grados de libertad y varianza h^2 . Se escogió la χ_{β}^2 por que modela la distribución del cuadrado de la distancia euclídea entre dos vectores β -dimensionales iids con distribución $\mathcal{N}(\mathbf{0}, \Sigma^2)$. Por tanto, el parámetro β es el que podría incluir la información de la dimensión de inmersión en la función núcleo.

Debemos resaltar que en este caso optamos por no trabajar con la distancia euclídea \tilde{z} , sino que en su lugar utilizamos su cuadrado ($z = \tilde{z}^2$). Esto es debido a

que computacionalmente es más costoso trabajar con distancias euclídeas que con sus cuadrados. Adicionalmente, es más eficiente generar realizaciones de ruido con distribución χ_{β}^2 que con distribución χ_{β} . Lo anterior también implica que la variable de entrada al algoritmo de correlación asistido por ruido (ver figura 4.1) es $z_{\omega} = \bar{z}_{\omega}^2$.

El último ingrediente para poder deducir una expresión analítica para la ICU es la función de distribución del cuadrado de la distancia entre vectores de estado. Oltmans y col. [140] encontraron una expresión para la función de densidad de probabilidad del cuadrado de la distancia euclídea entre vectores de estado m -dimensionales, formados a partir de una serie temporal proveniente de un sistema dinámico contaminada con ruido aditivo $\mathcal{N}(0, \sigma^2)$. Esta es [140] ec. (18):

$$f_z^{(m)}(\sigma; z) = \frac{z^{m/2-1}}{2(2\sigma)^{m-D}} \frac{\Gamma(D/2)}{\Gamma(m/2)} {}_1F_1\left(\frac{m-D}{2}, \frac{m}{2}; -\frac{z}{4\sigma^2}\right), \quad (5.2)$$

donde ${}_1F_1(a, b; t)$ es la Función Hipergeométrica Confluyente. Utilizando las ecuaciones (5.1) y (5.2) en la ecuación (4.2), podemos definir la ICU ($U_m^{\beta}(h)$) como:

$$U_m^{\beta}(h) = \int_0^{\infty} \frac{\Gamma(\beta/2, z/h^2)}{\Gamma(\beta/2)} \frac{z^{m/2-1}}{2(2\sigma)^{m-D}} \frac{\Gamma(D/2)}{\Gamma(m/2)} {}_1F_1\left(\frac{m-D}{2}, \frac{m}{2}; -\frac{z}{4\sigma^2}\right) dz.$$

Resolviendo se obtiene:

$$U_m^{\beta}(h) = \frac{\widehat{\phi}}{2} (2\sigma)^D e^{-m\tau K_2} \frac{\Gamma(D/2)\Gamma((\beta+m)/2)}{\Gamma(\beta/2)\Gamma(m/2+1)} \left(\frac{h}{2\sigma}\right)^m {}_2F_1\left(\frac{\beta+m}{2}, \frac{m-D}{2}; \frac{m+2}{2}; -\frac{h^2}{4\sigma^2}\right), \quad (5.3)$$

donde $\widehat{\phi}$ es una constante de normalización y ${}_2F_1(a, b; c; t)$ es la Función Hipergeométrica de Gauss y β son los grados de libertad del ruido μ_{ω} , nos referiremos a este parámetro como “la dimensión del ruido” utilizado en el algoritmo de correlación asistido por ruido.

5.3 Caso particular: $ICU^{\beta=2}$

Demostraremos que la ICG puede ser considerada un caso particular de la ICU. Para esto basta con hacer $\beta = 2$ en la ecuación (5.3):

$$U_m^{\beta=2}(h) = \frac{\widehat{\phi}}{2} (4\sigma^2)^{D/2} \Gamma(D/2) e^{-m\tau K_2} \left(\frac{h^2}{4\sigma^2}\right)^{m/2} {}_2F_1\left(\frac{m+2}{2}, \frac{m-D}{2}; \frac{m+2}{2}; -\frac{h^2}{4\sigma^2}\right).$$

Capítulo 5. Integral de correlación U

Teniendo en cuenta que: ${}_2F_1(a, b; a; t) = (1-t)^{-b}$ y utilizando la ecuación (2.11) se verifica que:

$$\begin{aligned} U_m^{\beta=2}(h) &= \widehat{\phi} 2^{D-1} \Gamma(D/2) h^m (h^2 + 4\sigma^2)^{\frac{D-m}{2}} e^{-m\tau K_2} \\ &= \frac{\widehat{\phi}}{\phi} 2^{(2D-1)} \Gamma(D/2) T_m(h/2). \end{aligned}$$

Es fácil ver que la integral de correlación U con $\beta = 2$ (ICU $^{\beta=2}$) es solo una versión escalada de la ICG.

5.4 Caso particular: ICU $^{\beta=m}$

Recordemos que la ecuación (5.1) introduce el parámetro β (llamado dimensión del ruido) en la estimación de la integral de correlación. La idea principal es que β nos permite incorporar información sobre la dimensión de inmersión m en la función núcleo y esto será útil en la estimación de los invariantes.

Comenzaremos entonces haciendo $\beta = m$ en la ecuación (5.1) y definiendo la integral de correlación $U_m^{\beta=m}(h)$ como:

$$U_m^{\beta=m}(h) = \frac{\widehat{\phi}}{2} (2\sigma)^D e^{-m\tau K_2} \frac{\Gamma(D/2)\Gamma(m)}{\Gamma(m/2)\Gamma(m/2+1)} \left(\frac{h^2}{4\sigma^2}\right)^{m/2} {}_2F_1\left(m, \frac{m-D}{2}; \frac{m+2}{2}; \frac{-h^2}{4\sigma^2}\right), \quad (5.4)$$

de esta integral de correlación podemos entonces deducir el funcional para la dimensión de correlación como (ver Apéndice A):

$$\frac{d \ln U_m^{\beta=m}(h)}{d \ln h} = m - \frac{2m(m-D)}{m+2} \left(\frac{h^2}{4\sigma^2}\right) \frac{{}_2F_1\left(m+1, \frac{m-D}{2}+1; \frac{m+2}{2}+1; \frac{-h^2}{4\sigma^2}\right)}{{}_2F_1\left(m, \frac{m-D}{2}; \frac{m+2}{2}; \frac{-h^2}{4\sigma^2}\right)}. \quad (5.5)$$

En el Apéndice A demostramos que este funcional se aproxima a D cuando $\sigma \ll h$ y a m cuando $\sigma \gg h$. Este comportamiento es similar a otros funcionales basados tanto en la integral de correlación estándar como en la ICG [11, 12, 16, 81].

Por otro lado, definimos el funcional para la entropía de correlación como (ver Apéndice B):

$$\ln \frac{U_{m+2}^{\beta=m+2}(h)}{U_m^{\beta=m}(h)} = \ln \left\{ \left(\frac{m+1}{m+2}\right) \left(\frac{h^2}{\sigma^2}\right) \frac{{}_2F_1\left(m+2, \frac{m-D}{2}+1; \frac{m+2}{2}+1; \frac{-h^2}{4\sigma^2}\right)}{{}_2F_1\left(m, \frac{m-D}{2}; \frac{m+2}{2}; \frac{-h^2}{4\sigma^2}\right)} \right\} - 2\tau K_2. \quad (5.6)$$

En el Apéndice B se muestra que cuando $\sigma \ll h$ este funcional se aproxima a:

$$\ln \frac{U_{m+2}^{\beta=m+2}(h)}{U_m^{\beta=m}(h)} = \ln \left(\frac{D}{m} + 1 \right) - 2\tau K_2 \quad (5.7)$$

y, para valores de m mucho más grandes que D , se puede ver que:

$$-\frac{1}{2\tau} \ln \frac{U_{m+2}^{\beta=m+2}(h)}{U_m^{\beta=m}(h)} \approx K_2.$$

Con relación al nivel de ruido σ presente en la serie temporal, recordemos que las series temporales utilizadas en las simulaciones han sido normalizadas para tener varianza unitaria. Por tanto $\sigma = 0$ implica una serie temporal limpia, mientras que $\sigma = 1$ se refiere a una serie temporal compuesta sólo por ruido (ver sección 5.5).

Inspirados en [84], proponemos el funcional para el nivel de ruido presente en la serie temporal (Apéndice C):

$$\begin{aligned} \Delta_m^U(h) &= \frac{1}{2} \left[\frac{d \ln U_{m+2}^{\beta=m}(h)}{d \ln h} - \frac{d \ln U_m^{\beta=m}(h)}{d \ln h} \right] \\ &\approx \frac{4\sigma^2}{4\sigma^2 + h^2}. \end{aligned} \quad (5.8)$$

Podemos observar que $\Delta_m^U(h)(h) \rightarrow 1$ cuando $h \rightarrow 0$ y que $\Delta_m^U(h) \rightarrow 0$ cuando $h \rightarrow \infty$. Esto nos indica que éste es un funcional que decrece de 1 a 0 con una velocidad que solo depende de σ . Este comportamiento es similar a otros estimadores de nivel de ruido como los presentados en [11, 84, 88, 141].

5.5 Simulaciones

En esta sección presentaremos los resultados numéricos obtenidos en la estimación de los invariantes D , K_2 y σ a partir de la integral de correlación U con $\beta = m$ ($ICU^{\beta=m}$). Utilizaremos series temporales provenientes de la transformación de Hénon (ecuación (1.3)) con $a = 4$ y $b = 0.3$. Las integrales de correlación ICG y $ICU^{\beta=m}$ fueron calculadas para $m = \{2, 4, 6, 8, 10\}$ con $\tau = 1$. El número pares de distancias entre vectores de estado fue mantenido constante $Q = 5000$ a través de los diferentes valores de m .

En la figura 5.1 se presenta una gráfica del logaritmo de ICG y de $ICU^{\beta=m}$ en función del $\ln h$, en ausencia de ruido. Como puede observarse, las curvas para $m = 2$ (arri-

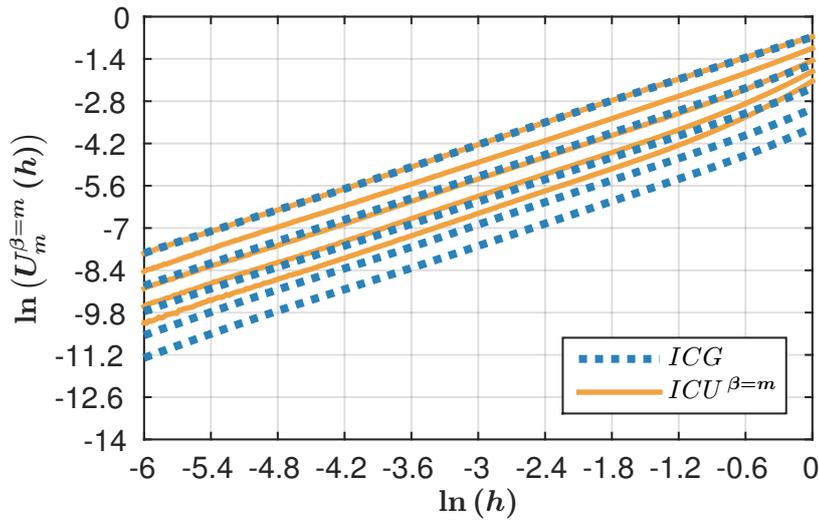


Figura 5.1. Gráfica log-log de la ICG y la $ICU^{\beta=m}$ para la transformación de Hénon sin ruido. Se presentan las curvas calculadas con $m = \{2, 4, 6, 8, 10\}$ ($m = 2$ curva superior y $m = 10$ curva inferior) para la ICG (azul punteado) y la $ICU^{\beta=m}$ (naranja).

ba) de ambas integrales de correlación coinciden. Esto es esperable ya que, como demostramos antes, la ICG es un caso particular de la $ICU^{\beta=m}$ cuando $m = 2$. Se verifica que las curvas de ambas integrales de correlación presentan la misma pendiente, lo que indica que ambas deben converger a la misma D . Podemos verificar lo anterior observando la figura 5.2 donde se presenta la derivada logarítmica de la ICG y de la $ICU^{\beta=m}$ en función de $\ln h$. Vemos cómo ambas integrales de correlación oscilan alrededor del valor reportado para D [16].

En la figura 5.1 se evidencia también una de las principales diferencias entre la ICG y la $ICU^{\beta=m}$. Es posible observar que la distancia entre curvas para diferentes valores de m es más uniforme para la $ICU^{\beta=m}$ que para la ICG. Teniendo en cuenta que la estimación de K_2 se basa en integrales de correlación con distinto valor de m , lo anterior sugiere que la $ICU^{\beta=m}$ sea más adecuada que la ICG para el cálculo de la entropía de correlación. Esto se verifica en la figura 5.3 donde se muestra el funcional de entropía basado en la $ICU^{\beta=m}$ y el funcional de entropía basado en la ICG (ecuación (2.13)) en función del $\ln(h)$ para $m = \{2, 4, 6, 8\}$. Se puede ver fácilmente como las curvas pertenecientes a la $ICU^{\beta=m}$ convergen mejor al valor reportado de la entropía de correlación ($K_2 = 0.3$ [16, 87]) que el funcional basado en la ICG que necesitaría un valor más grande de m para converger.

La presencia de ruido en la serie temporal hace que la integral de correlación se desvíe de su escalamiento como ley de potencia. Este fenómeno es observado en la figura 5.4 donde para valores pequeños de $\ln h$ las curvas de ambas integrales de

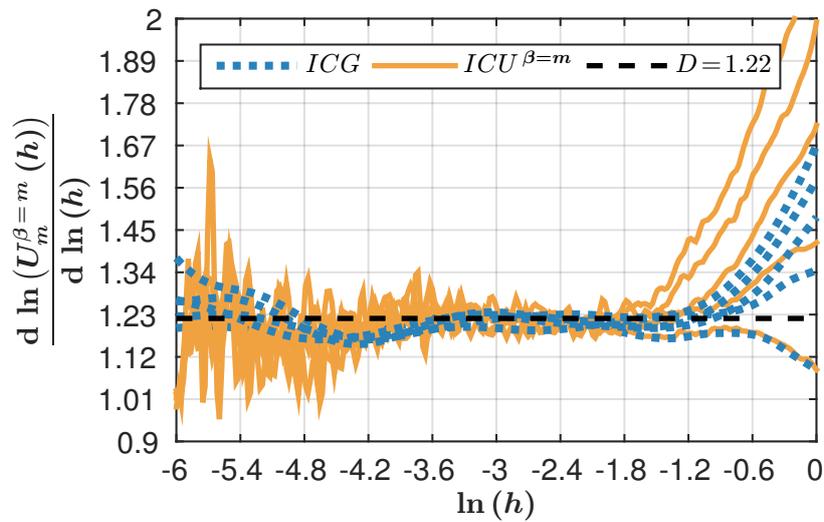


Figura 5.2. Estimación de D para la transformación de Hénon en ausencia de ruido. Se presenta la derivada logarítmica de la ICG (azul punteado) y de la $ICU^{\beta=m}$ (naranja) para $m = \{2, 4, 6, 8, 10\}$. El valor reportado de D se muestra en línea negra discontinua.

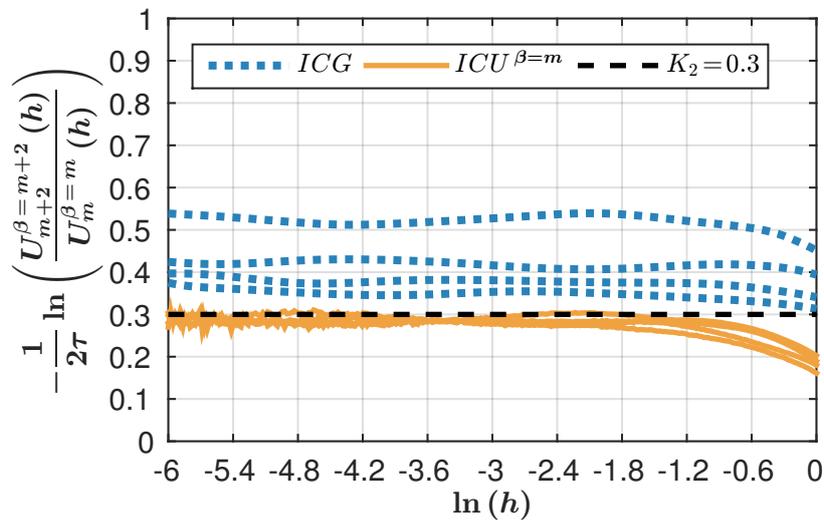


Figura 5.3. Estimación de K_2 para la transformación de Hénon en ausencia de ruido. Se presenta el funcional de entropía de correlación de la ICG (azul punteado) y de la $ICU^{\beta=m}$ (naranja) para $m = \{2, 4, 6, 8\}$. El valor reportado de K_2 se muestra en línea negra discontinua.

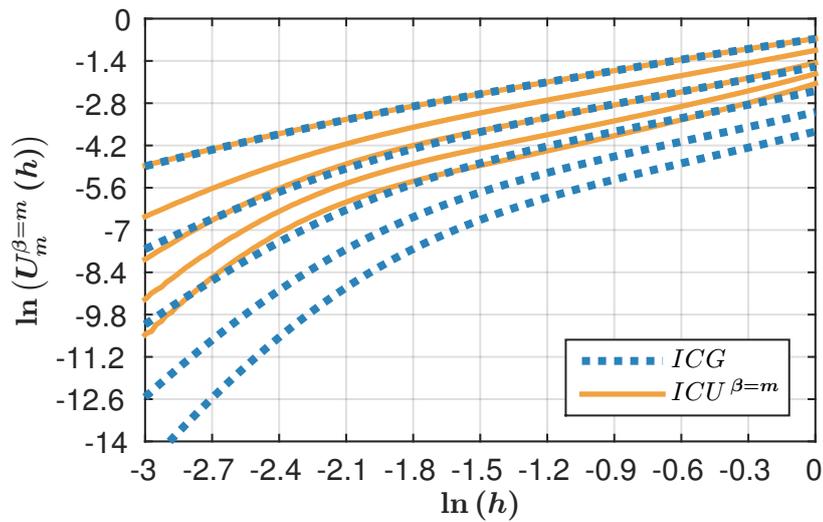


Figura 5.4. Gráfica log-log de la ICG y la $ICU^{\beta=m}$ para la transformación de Hénon con ruido $\sigma = 0.05$. Se presentan las curvas calculadas con $m = \{2, 4, 6, 8, 10\}$ ($m = 2$ curva superior y $m = 10$ curva inferior) para la ICG (azul punteado) y la $ICU^{\beta=m}$ (naranja).

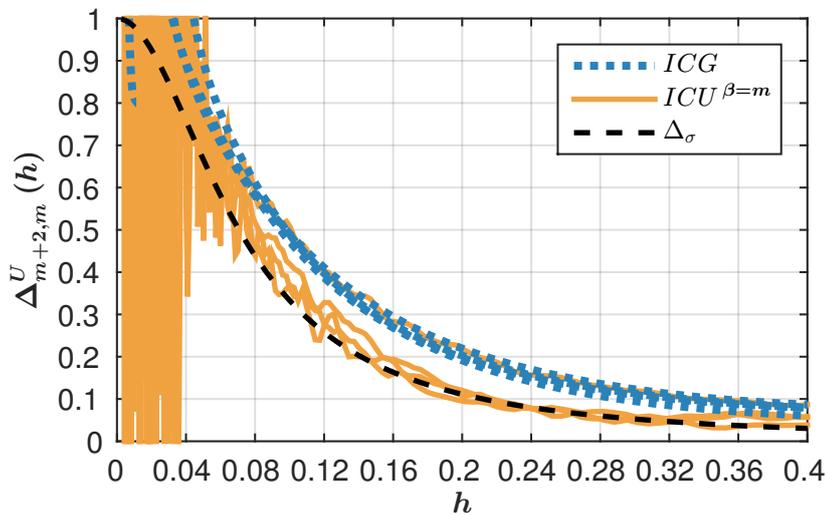


Figura 5.5. Estimación del nivel de ruido para la transformación de Hénon con ruido ($\sigma = 0.05$). Se presenta el funcional de nivel de ruido de la ICG (azul punteado) y de la $ICU^{\beta=m}$ (naranja) para $m = \{2, 4, 6, 8\}$. La función $\Delta_{\sigma}(h) = 4\sigma^2 / (4\sigma^2 + h^2)$ se muestra en línea negra discontinua.

correlación ($ICU^{\beta=m}$ y ICG) presentan una pendiente más alta. Esto es debido a que el ruido tiene dimensión infinita y domina completamente las escalas bajas. Por otro lado, para valores de $\ln h \geq -1.8$ el escalamiento habitual de ambas integrales de correlación es recuperado.

Bajo estas condiciones de ruido es posible analizar el comportamiento del funcional $\Delta_m^U(h)$. En la figura 5.5 se presentan los funcionales de nivel de ruido basados en la $ICU^{\beta=m}$ y en la ICG, con $m = \{2, 4, 6, 8\}$. Se muestra también el valor de la función $\Delta_\sigma = 4\sigma^2/(4\sigma^2 + h^2)$ en negro punteado. Se puede apreciar cómo las curvas para ambos estimadores decaen de 1 a 0 al incrementarse h . Es de resaltar que las curvas pertenecientes a la $ICU^{\beta=m}$ se acercan cada vez más a la función Δ_σ (ecuación 5.8) al incrementar el valor de m . Podemos observar también que las curvas para $m = 2$ (superior) de la $ICU^{\beta=m}$ y la ICG coinciden. Recordemos que esto se da por que la ICG y la ICU coinciden cuando $m = 2$.

5.6 Conclusiones

En este capítulo hemos introducido la ICU (ecuación 5.3). Esta integral de correlación también puede ser vista como un caso particular de la ICAR en la cual la función núcleo está dada por la ecuación 5.1. Dicha función, a diferencia de la función de núcleo Gaussiano, permite incorporar información sobre la dimensión de inmersión en el cálculo de la integral de correlación.

A partir de la $ICU^{\beta=m}$ (ecuación 5.4) hemos propuesto, definido y analizado el comportamiento de un funcional de dimensión de correlación (ecuación 5.5), uno de entropía de correlación (ecuación 5.7) y uno de nivel ruido (ecuación 5.8). Estas deducciones analíticas han sido posibles gracias a la versatilidad de la función Hipergeométrica de Gauss que cuenta con un vasto conjunto de identidades y de la cual hay una documentación extensa.

Hemos realizado simulaciones que corroboran los resultados analíticos sobre cada funcional. En cuanto a la dimensión de correlación podemos decir que en ausencia de ruido nuestro funcional converge a D . También mostramos analíticamente que nuestro funcional de nivel de ruido escala en proporción a σ . El resultado más sorprendente es que el funcional de entropía definido para la $ICU^{\beta=m}$ converge de mejor forma a K_2 que el funcional definido para la ICG. Esto es debido a la incorporación de información sobre la dimensión de inmersión.

No, el éxito no se lo deseo a nadie. Le sucede a uno lo que a los alpinistas, que se matan por llegar a la cumbre y cuando llegan, ¿qué hacen? Bajar, o tratar de bajar discretamente, con la mayor dignidad posible.

Gabriel García Márquez

6

Estimadores basados en la integral de correlación U

En el capítulo [5](#) presentamos la integral de correlación U ($ICU^{\beta=m}$). A partir de esta definimos una serie de funcionales para estimar D , K_2 y σ . En presencia de grandes cantidades de ruido, como cualquier otra integral de correlación, la $ICU^{\beta=m}$ se desvía notablemente del régimen de escalamiento habitual (ecuación [\(2.6\)](#)). En consecuencia, los funcionales presentados en las ecuaciones [\(5.5\)](#), [\(5.6\)](#) y [\(5.8\)](#) son incapaces de estimar con precisión los invariantes D , K_2 y σ respectivamente. Con el fin de solucionar este inconveniente se proponen los estimadores dependientes de la escala (EsDE) que son expresiones analíticas para D , K_2 y σ en función de los parámetros m y h [\[11, 12\]](#).

En este capítulo deduciremos las expresiones analíticas de los EsDE basados en la $ICU^{\beta=m}$ para D , K_2 y σ . Estudiaremos mediante simulaciones sus comportamientos bajo diferentes condiciones de ruido y contrastaremos estos estimadores con EsDE existentes en la literatura.

También propondremos una metodología basada en los estimadores aquí propuestos que nos permitirá estimar de forma automática los valores de D , K_2 y σ a partir de una serie temporal. Para finalizar, esta metodología será evaluada estadística-

mente.

6.1 Introducción

Se demostró en el capítulo 2 que la presencia de ruido en la serie temporal introduce un sesgo en la estimación de la integral de correlación [12]. Este sesgo se traduce en un incremento de la pendiente de la gráfica log-log de la integral de correlación para valores pequeños de h (ver figura 5.4). A su vez, esto reduce el rango de escalas en donde sería posible estimar los invariantes del sistema dinámico.

La solución para este inconveniente consiste en modelar analíticamente la influencia del ruido sobre la integral de correlación. Por ejemplo, bajo la presencia de ruido con distribución $\mathcal{N}(0, \sigma^2)$, la ICG escala como [16]:

$$T_m(h) = \phi h^m (h^2 + \sigma^2)^{\frac{D-m}{2}} e^{-\tau m K_2} \quad \text{para } m \rightarrow \infty, \sqrt{h^2 + \sigma^2} \rightarrow 0.$$

En esta ecuación vemos que la varianza del ruido σ es introducida como un parámetro en la función de escala. De forma similar podemos encontrar en la literatura otros modelos que permiten modelar la influencia de ruido [11, 91, 92]. Recordemos que en el capítulo 5 propusimos la función de escala de la ICU $^{\beta=m}$ (ecuación 5.4).

Una vez escogido el modelo a utilizar y estimada la integral de correlación, es posible determinar los invariantes de un sistema mediante dos caminos. El primero es una regresión no lineal sobre la función de escala elegida [16, 87]. Esta metodología tiene la desventaja de que es altamente dependiente del rango de escalas elegido para la regresión, con el agravante de que no existe un consenso general en la literatura sobre cómo escogerlo. Más importante aún, si se quiere determinar que la serie temporal en estudio proviene de un sistema determinístico, es necesario comprobar que los invariantes no dependen del rango de escalas escogido [12]. Esto solo puede hacerse mediante inspección visual. Con el fin de superar las dificultades de esta metodología surgen los EsDE o “coarse-grained estimators” [11, 12, 88]. Estos son expresiones analíticas para D , K_2 y σ en función de los parámetros m y h .

Como ejemplo tenemos los EsDE propuestos por Nolte y col. en [88] y presentados en el capítulo 2. Los estimadores $D_m^T(h)$ (ecuación 2.15), $K_m^T(h)$ (ecuación 2.16) y $\sigma_m^T(h)$ (ecuación 2.17) cuentan con la ventaja de que no requieren grandes recursos computacionales para su cálculo y lo más importante, estos dependen solo de las ICGs calculadas para valores consecutivos de m , es decir, no necesitan ningún parámetro externo para su estimación.

6.2. Estimador dependiente de la escala para D

En este capítulo presentaremos nuevos EsDE basados en la $ICU^{\beta=m}$. Estudiaremos el comportamiento de estos estimadores ante distintas longitudes de datos y niveles de ruido. Para comparar nuestros resultados con aquellos publicados en la literatura, hemos optado por utilizar los EsDE propuestos por Nolte y col. [88] ya que por sus características no requieren de la manipulación de ningún parámetro por parte del usuario, convirtiéndolos en estimadores robustos para hacer comparaciones.

El análisis de grandes bases de datos requiere una manera eficiente de estimar los invariantes D , K_2 y σ . Por tal motivo, desarrollaremos también en este capítulo una metodología para el cálculo automático de estos invariantes. Esto nos permitirá realizar un análisis estadístico de los nuevos estimadores aquí propuestos.

6.2 Estimador dependiente de la escala para D

En esta sección deduciremos la expresión analítica del EDE basado en la $ICU^{\beta=m}$ para dimensión de correlación ($D_m^U(h)$). Para esto utilizaremos la definición de la ICU y varias identidades de la función hipergeométrica de Gauss [142].

Para encontrar una expresión analítica para $D_m^U(h)$ es necesario estudiar la derivada logarítmica de la $ICU^{\beta=m}$. Partiremos entonces de la forma general de la ICU (ecuación (5.3)):

$$U_m^\beta(h) = \frac{\widehat{\phi}}{2} (2\sigma)^D e^{-m\tau K_2} \frac{\Gamma(D/2)\Gamma((\beta+m)/2)}{\Gamma(\beta/2)\Gamma(m/2+1)} \left(\frac{h}{2\sigma}\right)^m {}_2F_1\left(\frac{\beta+m}{2}, \frac{m-D}{2}; \frac{m+2}{2}; \frac{-h^2}{4\sigma^2}\right). \quad (6.1)$$

Definiendo:

$$\begin{aligned} a &= \frac{\beta+m}{2}; & b &= \frac{m-D}{2}; & c &= \frac{m+2}{2}; \\ z &= -\frac{h^2}{4\sigma^2}; & \frac{dz}{dh} &= -\frac{h}{2\sigma^2} \end{aligned} \quad (6.2)$$

y

$$P = (-1)^{m/2} \frac{\widehat{\phi}}{2} (2\sigma)^D \frac{\Gamma(D/2)\Gamma((m+\beta)/2)}{\Gamma(\beta/2)\Gamma(m/2+1)} e^{-m\tau K_2}.$$

Vemos que podemos escribir la ecuación (6.1) como:

$$U_m^\beta(z) = P z^{m/2} {}_2F_1(a, b; c; z).$$

Capítulo 6. Estimadores basados en la integral de correlación U

La derivada de $U_m^\beta(z)$ con respecto a h es:

$$\frac{d}{dh} U_m^\beta(z) = \frac{d}{dz} [P z^{m/2} {}_2F_1(a, b; c; z)] \frac{dz}{dh}.$$

Usando [143] eq. 15.2.1]:

$$\frac{d}{dh} U_m^\beta(z) = \frac{-Ph}{2\sigma^2} z^{m/2} \left[\frac{m}{2z} {}_2F_1(a, b; c; z) + \frac{ab}{c} {}_2F_1(a+1, b+1; c+1; z) \right].$$

La derivada logarítmica de $U_m^\beta(h)$ se puede obtener como:

$$\begin{aligned} \frac{d \ln U_m^\beta(z)}{d \ln h} &= \frac{h}{U_m^\beta(z)} \frac{d}{dh} U_m^\beta(z) \\ &= m + \frac{2abz}{c} \frac{{}_2F_1(a+1, b+1; c+1; z)}{{}_2F_1(a, b; c; z)}. \end{aligned} \quad (6.3)$$

Para continuar utilizaremos una de las identidades de la función hipergeométrica de Gauss [142] Ec. 9.137-9]:

$$\frac{abz}{c} \frac{{}_2F_1(a+1, b+1; c+1; z)}{{}_2F_1(a, b; c; z)} = \frac{c-a}{1-z} \frac{{}_2F_1(a-1, b; c; z)}{{}_2F_1(a, b; c; z)} - \frac{c-a-bz}{1-z}. \quad (6.4)$$

Vemos como el término de la derecha de la ecuación (6.4) puede ser transformado en la derivada logarítmica de la ecuación (6.3), es decir:

$$\frac{d \ln U_m^\beta(h)}{d \ln h} = m + \frac{2(c-a)}{1-z} \frac{{}_2F_1(a-1, b; c; z)}{{}_2F_1(a, b; c; z)} - \frac{2(c-a-bz)}{1-z}. \quad (6.5)$$

Ahora, es necesario encontrar un estimador para el cociente $\frac{{}_2F_1(a-1, b; c; z)}{{}_2F_1(a, b; c; z)}$. Utilizando las ecuaciones (6.1) y (6.2) definimos el estimador:

$$\frac{U_m^{\beta-2}(z)}{U_m^\beta(z)} = \frac{\beta-2}{2(a-1)} \frac{{}_2F_1(a-1, b; c; z)}{{}_2F_1(a, b; c; z)}. \quad (6.6)$$

Usando las ecuaciones (6.5) y (6.6) encontramos que:

$$\frac{d \ln U_m^\beta(z)}{d \ln h} = m + \frac{4(c-a)(a-1)}{(1-z)(\beta-2)} \frac{U_m^{\beta-2}(z)}{U_m^\beta(z)} - \frac{2(c-a-bz)}{1-z}.$$

6.2. Estimador dependiente de la escala para D

Despejando b :

$$2b = \frac{1-z}{z} \left[\frac{d \ln U_m^\beta(z)}{d \ln h} - m \right] - \frac{4(a-1)(c-a)}{z(\beta-2)} \frac{U_m^{\beta-2}(z)}{U_m^\beta(z)} + \frac{2(c-a)}{z},$$

restaurando los valores de a, b, c, z (ecuación (6.2)), haciendo $\beta = m$, y despejando D , definimos el EDE para la dimensión de correlación $D_m^u(h)$ como:

$$D_m^U(h) = \left(\frac{4\sigma^2 + h^2}{h^2} \right) \frac{d \ln U_m^{\beta=m}(h)}{d \ln h} + \frac{4\sigma^2}{h^2} (2m-2) \left[\frac{U_m^{\beta=m-2}(h)}{U_m^{\beta=m}(h)} - 1 \right]. \quad (6.7)$$

Es necesario aclarar que el estimador $U_m^{\beta=m-2}(h)$ debe entenderse como la ICU calculada con vectores de estado de dimensión m y dimensión de ruido $\beta = m - 2$.

De la ecuación anterior podemos ver que:

$$\lim_{\sigma \rightarrow 0} D_m^U(h) = \frac{d \ln U_m^{\beta=m}(h)}{d \ln h}, \quad (6.8)$$

es decir, que en ausencia de ruido, el $D_m^U(h)$ tiende al funcional de dimensión de correlación (ver ecuación (5.5)). Este tipo de comportamiento lo podemos observar en otros EsDE como el propuesto por Diks [11], por Nolte y col. [88] y por Jayawardena y col. [91].

En la figura 6.1 se muestra el comportamiento del nuestro estimador $D_m^U(h)$ y el propuesto por Nolte y col. $D_m^T(h)$, en función del $\ln h$. Estos se calcularon a partir de series temporales normalizadas (varianza unitaria), provenientes de la transformación de Hénon bajo distintos niveles de ruido $\sigma = \{0, 0.05, 0.2\}$ y distinto número de vectores de estado $L = \{500, 5000, 10000\}$. Se presenta además el valor para la dimensión de correlación ($D = 1.22$) reportado en la literatura [16].

Es posible observar en las figuras 6.1a, 6.1d y 6.1g que, para series temporales sin ruido el estimador $D_m^U(h)$ oscila alrededor del valor reportado para D . La amplitud de dicha oscilación se atenúa al incrementar el número de vectores de estado disponibles para el cálculo de la ICU $^{\beta=m}$. También vemos que para $L = 500$ (figura 6.1a) el estimador $D_m^U(h)$ está más cerca del valor reportado de D que el estimador $D_m^T(h)$. Sin embargo para valores de L mayores los resultados son muy similares (figuras 6.1b y 6.1c).

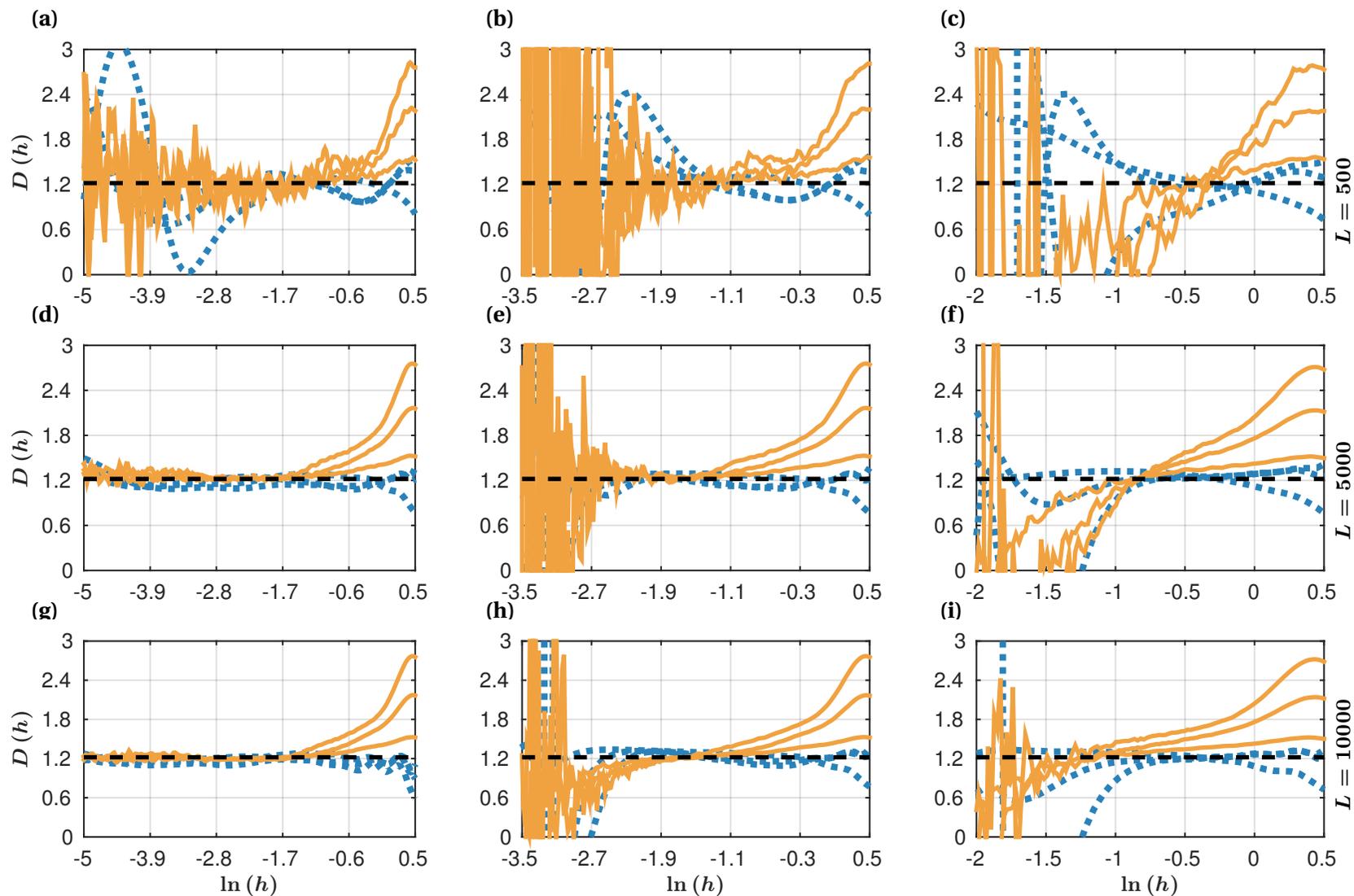


Figura 6.1. Transformación de Hénon. EDE para la dimensión de correlación. $D_m^T(h)$ (azul punteado) y $D_m^U(h)$ (naranja) para $m = [4, 6, 8]$. $L = 500$: (a) Sin ruido, (b) $\sigma = 0.05$ y (c) $\sigma = 0.2$. $L = 5000$: (d) sin ruido, (e) $\sigma = 0.05$ y (f) $\sigma = 0.2$. $L = 10000$: (g) sin ruido, (h) $\sigma = 0.05$ e (i) $\sigma = 0.2$. El valor reportado de la dimensión de correlación $D = 1.22$ se muestra en línea negra discontinua.

6.3. Estimador dependiente de la escala para K_2

En el caso de valores moderados de ruido ($\sigma = 0.05$), se puede apreciar en las figuras 6.1b, 6.1e y 6.1h que el rango de valores de $\ln h$ adecuado para estimar D ha disminuido en comparación al caso sin ruido para $D_m^U(h)$ y para $D_m^T(h)$. Algo muy interesante acerca del estimador $D_m^U(h)$ puede observarse en la figura 6.1b. Este estimador se acerca a D cuando las curvas calculadas para diferentes valores de m están más cerca entre sí ($\ln h \approx -1.3$). En otras palabras, $D_m^U(h)$ se aproxima a la dimensión de correlación cuando la varianza de este estimador a través de m es mínima. Este fenómeno también es evidente en la figura 6.1e para $L = 5000$ ($\ln h \approx -1.9$) y en la figura 6.1h para $L = 10000$ ($\ln h \approx -1.7$).

Las figuras 6.1c, 6.1f y 6.1i muestran cómo para niveles grandes de ruido $\sigma = 0.2$ es difícil encontrar un rango de valores de h el cual se pueda estimar la dimensión de correlación tanto para $D_m^U(h)$ como para $D_m^T(h)$. No obstante, puede observarse una vez más que el estimador $D_m^U(h)$ se acerca a D cuando la distancia entre las curvas para diferentes valores de m se hace mínima ($\ln h \approx -0.5$ para $L = 500$, $\ln h \approx -0.8$ para $L = 5000$ y $\ln h \approx -1.3$ para $L = 10000$).

6.3 Estimador dependiente de la escala para K_2

La deducción analítica del EDE para la entropía de correlación $K_m^U(h)$ involucra las integrales de correlación $U_m^{\beta=m}(h)$ y $U_{m+2}^{\beta=m+2}(h)$. Partiendo de la definición de la ICU (ecuación 6.1) y utilizando la ecuación 6.2) podemos escribir:

$$U_m^\beta(z) = M \frac{\Gamma(a) e^{-m\tau K_2} (-z)^{m/2}}{\Gamma(\beta/2)\Gamma(c)} {}_2F_1(a, b; c; z), \quad (6.9)$$

con

$$M = \frac{\widehat{\phi}}{2} (2\sigma)^D \Gamma(D/2).$$

Utilizando la ecuación 6.9) podemos escribir el cociente $U_{m+2}^{\beta+2}(z)/U_m^\beta(z)$ como:

$$\begin{aligned} \frac{U_{m+2}^{\beta+2}(z)}{U_m^\beta(z)} &= \frac{\Gamma(a+2)\Gamma(\frac{\beta}{2})\Gamma(c)}{\Gamma(a)\Gamma(\frac{\beta+2}{2})\Gamma(c+1)} \frac{(-z) {}_2F_1(a+2, b+1; c+1; z)}{{}_2F_1(a, b; c; z)} e^{-2\tau K_2} \\ &= \frac{2a(a+1)(-z) {}_2F_1(a+2, b+1; c+1; z)}{\beta c {}_2F_1(a, b; c; z)} e^{-2\tau K_2}. \end{aligned}$$

Haciendo $w = a + 1$, la última ecuación se transforma en:

$$\frac{U_{m+2}^{\beta+2}(z)}{U_m^\beta(z)} = \frac{2(w-1)}{\beta} \frac{w(-z) {}_2F_1(w+1, b+1; c+1; z)}{c {}_2F_1(w-1, b; c; z)} e^{-2\tau K_2}. \quad (6.10)$$

Utilizando la identidad [142] ec. 9.137-9]:

$$b(z-1) \frac{w z {}_2F_1(w+1, b+1; c+1; z)}{c {}_2F_1(w-1, b; c; z)} = (c-w) - (c-w-bz) \frac{{}_2F_1(w, b; c; z)}{{}_2F_1(w-1, b; c; z)}, \quad (6.11)$$

podemos escribir la ecuación (6.10) como:

$$\frac{U_{m+2}^{\beta+2}(z)}{U_m^\beta(z)} = \frac{-2(w-1)}{\beta b(1-z)} \left[c-w - (c-w-bz) \frac{{}_2F_1(w, b; c; z)}{{}_2F_1(w-1, b; c; z)} \right] e^{-2\tau K_2}. \quad (6.12)$$

Reemplazando $a = w-1$ y sabiendo que $c-a-1 = -\beta/2$ (ecuación (6.2)), entonces podemos hallar que:

$$\frac{U_{m+2}^{\beta+2}(z)}{U_m^\beta(z)} = \frac{-2a}{\beta b(1-z)} \left[\left(\frac{\beta}{2} + bz \right) \frac{{}_2F_1(a+1, b; c; z)}{{}_2F_1(a, b; c; z)} - \frac{\beta}{2} \right] e^{-2\tau K_2}. \quad (6.13)$$

Usando la ecuación (A.5) en la ecuación (6.13) vemos que:

$$\begin{aligned} \frac{U_{m+2}^{\beta+2}(h)}{U_m^\beta(h)} &= \frac{e^{-2\tau K_2}}{\beta b \left(1 + \frac{h^2}{4\sigma^2}\right)} \left\{ \frac{2a\beta}{2} - \left(\frac{\beta}{2} - \frac{bh^2}{4\sigma^2} \right) \left[\frac{d \ln U_m^\beta(h)}{d \ln h} + \beta \right] \right\} \\ &= \frac{e^{-2\tau K_2}}{\left(1 + \frac{h^2}{4\sigma^2}\right)} \left\{ \frac{1}{2b} \left[2a - \frac{d \ln U_m^\beta(h)}{d \ln h} - \beta \right] + \frac{h^2}{4\sigma^2 \beta} \left[\frac{d \ln U_m^\beta(h)}{d \ln h} + \beta \right] \right\}. \end{aligned}$$

Restaurando los valores de a y b tenemos que:

$$\frac{U_{m+2}^{\beta+2}(h)}{U_m^\beta(h)} = e^{-2\tau K_2} \left[\frac{4\sigma^2}{h^2 + 4\sigma^2} \left(\frac{m - \frac{d \ln U_m^\beta(h)}{d \ln h}}{m - D} \right) + \frac{h^2}{h^2 + 4\sigma^2} \left(\frac{\frac{d \ln U_m^\beta(h)}{d \ln h}}{\beta} + 1 \right) \right]. \quad (6.14)$$

Tomando $\beta = m$, aplicando logaritmo a ambos lados de la ecuación (6.14) y despejando obtenemos:

$$2\tau K_2 = -\ln \frac{U_{m+2}^{\beta=m+2}(h)}{U_m^{\beta=m}(h)} + \ln \left[\frac{4\sigma^2}{h^2 + 4\sigma^2} \left(\frac{m - \frac{d \ln U_m^{\beta=m}(h)}{d \ln h}}{m - D} \right) + \frac{h^2}{h^2 + 4\sigma^2} \left(\frac{\frac{d \ln U_m^{\beta=m}(h)}{d \ln h}}{m} + 1 \right) \right].$$

6.3. Estimador dependiente de la escala para K_2

Substrayendo $\ln(D/m + 1)$ ambos lados:

$$2\tau K_2 - \ln\left(\frac{D}{m} + 1\right) = -\ln\frac{U_{m+2}^{\beta=m+2}(h)}{U_m^{\beta=m}(h)} - \ln\left(\frac{D}{m} + 1\right) + \ln\left[\frac{4\sigma^2}{h^2 + 4\sigma^2}\left(\frac{m - \frac{d\ln U_m^{\beta=m}(h)}{d\ln h}}{m - D}\right) + \frac{h^2}{h^2 + 4\sigma^2}\left(\frac{\frac{d\ln U_m^{\beta=m}(h)}{d\ln h}}{m} + 1\right)\right]. \quad (6.15)$$

Finalmente el EDE para la entropía de correlación $K_m^U(h)$ será:

$$K_m^U(h) = \frac{1}{2\tau} \left\{ \ln\left[\frac{4\sigma^2}{h^2 + 4\sigma^2}\left(\frac{m - \frac{d\ln U_m^{\beta=m}(h)}{d\ln h}}{m - D}\right) + \frac{h^2}{h^2 + 4\sigma^2}\left(\frac{\frac{d\ln U_m^{\beta=m}(h)}{d\ln h}}{m} + 1\right)\right] - \ln\frac{U_{m+2}^{\beta=m+2}}{U_m^{\beta=m}} - \ln\left(\frac{D}{m} + 1\right) \right\}. \quad (6.16)$$

En el Apéndice [A](#) demostramos que si $\sigma \rightarrow 0$, entonces $d\ln U_m^{\beta=m}(h)/d\ln h \rightarrow D$. Por tanto $K_m^U(h) \rightarrow K_2 - \ln(D/m + 1)/2\tau$ y para valores de $m \gg D$ el estimador $K_m^U(h)$ se aproximará a K_2 .

En la figura [6.2](#) se presenta el comportamiento de los EsDE $K_m^U(h)$ y $K_m^T(h)$ en función del $\ln h$ para $m = \{2, 4, 6, 8\}$. Estos estimadores se calcularon a partir de series temporales provenientes de la transformación de Hénon bajo distintos niveles de ruido $\sigma = \{0, 0.05, 0.2\}$ y normalizadas para tener varianza unitaria. El número de vectores de estado fue fijado en los valores de $L = \{500, 5000, 10000\}$. También se muestra el valor reportado en la literatura para la entropía de correlación $K_2 = 0.3$ [\[16\]](#).

Al aumentar el nivel de ruido ($\sigma = 0.05$) se puede ver en las figuras [6.2b](#), [6.2e](#) y [6.2h](#) que el rango de valores de h donde estimar K_2 se ha reducido en comparación con el caso sin ruido. Esto pasa tanto para el estimador $K_m^U(h)$ como para el estimador $K_m^T(h)$. Sin embargo, podemos observar que el estimador $K_m^U(h)$ converge a un valor más cercano a la entropía de correlación en comparación con $K_m^T(h)$.

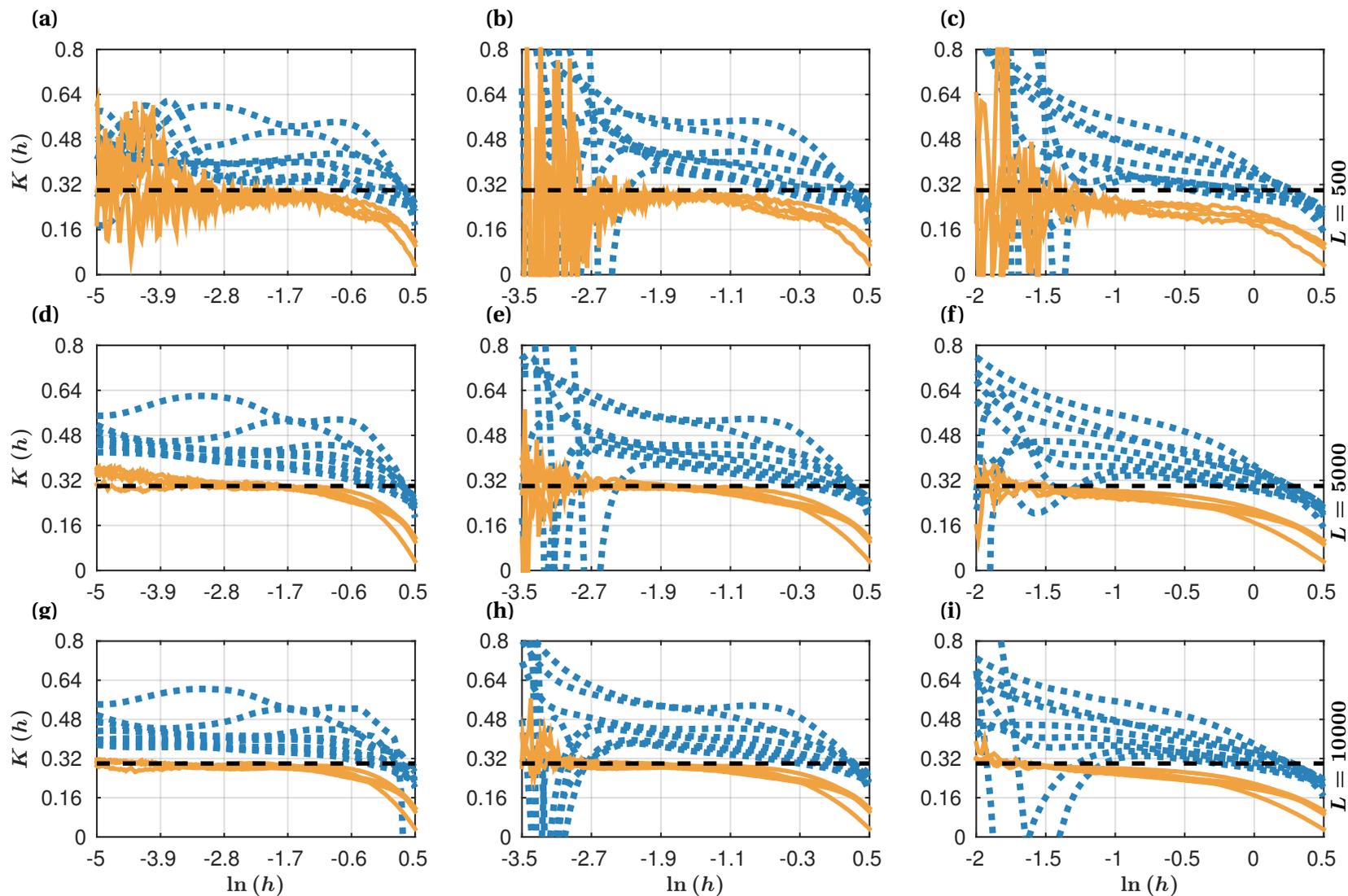


Figura 6.2. Transformación de Hénon. EDE para la entropía de correlación. $K_m^T(h)$ (azul punteado) y $K_m^U(h)$ (naranja) para $m = [2, 4, 6, 8]$. $L = 500$: (a) Sin ruido, (b) $\sigma = 0.05$ y (c) $\sigma = 0.2$. $L = 5000$: (d) sin ruido, (e) $\sigma = 0.05$ y (f) $\sigma = 0.2$. $L = 10000$: (g) sin ruido, (h) $\sigma = 0.05$ e (i) $\sigma = 0.2$. El valor reportado de la entropía de correlación $K_2 = 0.3$ se muestra en línea negra discontinua.

6.4. Estimador dependiente de la escala para σ

Para altos niveles de ruido ($\sigma = 0.2$) se muestra en las figuras [6.2c](#), [6.2h](#) y [6.2i](#) que ya no se puede distinguir un rango de h donde estimar K_2 a partir de $K_m^T(h)$. Por otro lado, para $K_m^U(h)$ es posible distinguir un pequeño intervalo de valores de h donde estimar la entropía de correlación, incluso para bajas longitudes de datos ($L = 500$).

En función de los resultados podemos afirmar que el estimador $K_m^U(h)$ supera a su contra parte basado en la ICG, ya que no depende fuertemente del parámetro m ni de la longitud de la serie temporal. Este comportamiento sigue presente en situaciones con ruido moderado y alto. Es muy interesante ver que para el caso sin ruido y $L = 10000$ (figura [6.2a](#)) el valor al que converge $K_m^U(h)$ ($K_2 \approx 0.28$) es menor que el reportado en la literatura.

6.4 Estimador dependiente de la escala para σ

Recordemos que las series temporales utilizadas en este estudio están normalizadas para tener varianza unitaria. Esta normalización se realiza luego de contaminar con ruido los datos. Por tanto un valor de $\sigma = 0$ implica series temporales limpias y un valor de $\sigma = 1$ hace referencia a que la serie es solo ruido blanco (ver sección [2.5](#)).

Para llegar a la ecuación de nuestro EDE para el nivel de ruido $\sigma_m^U(h)$ partiremos del funcional $\Delta_m^U(h)$ (ver sección [5.4](#)):

$$\Delta_m^U(h) = \frac{1}{2} \left[\frac{d \ln U_{m+2}^{\beta=m}(h)}{d \ln h} - \frac{d \ln U_m^{\beta=m}(h)}{d \ln h} \right],$$

que para valores de m grandes se comporta como (ver Apéndice [C](#)):

$$\Delta_m^U(h) = \frac{4\sigma^2}{4\sigma^2 + h^2}. \quad (6.17)$$

Despejando σ de la ecuación [\(6.17\)](#) definimos el EDE para el nivel de ruido como:

$$\sigma_m^U(h) = \frac{h}{2} \sqrt{\frac{\Delta_m^U(h)}{1 - \Delta_m^U(h)}} \quad (6.18)$$

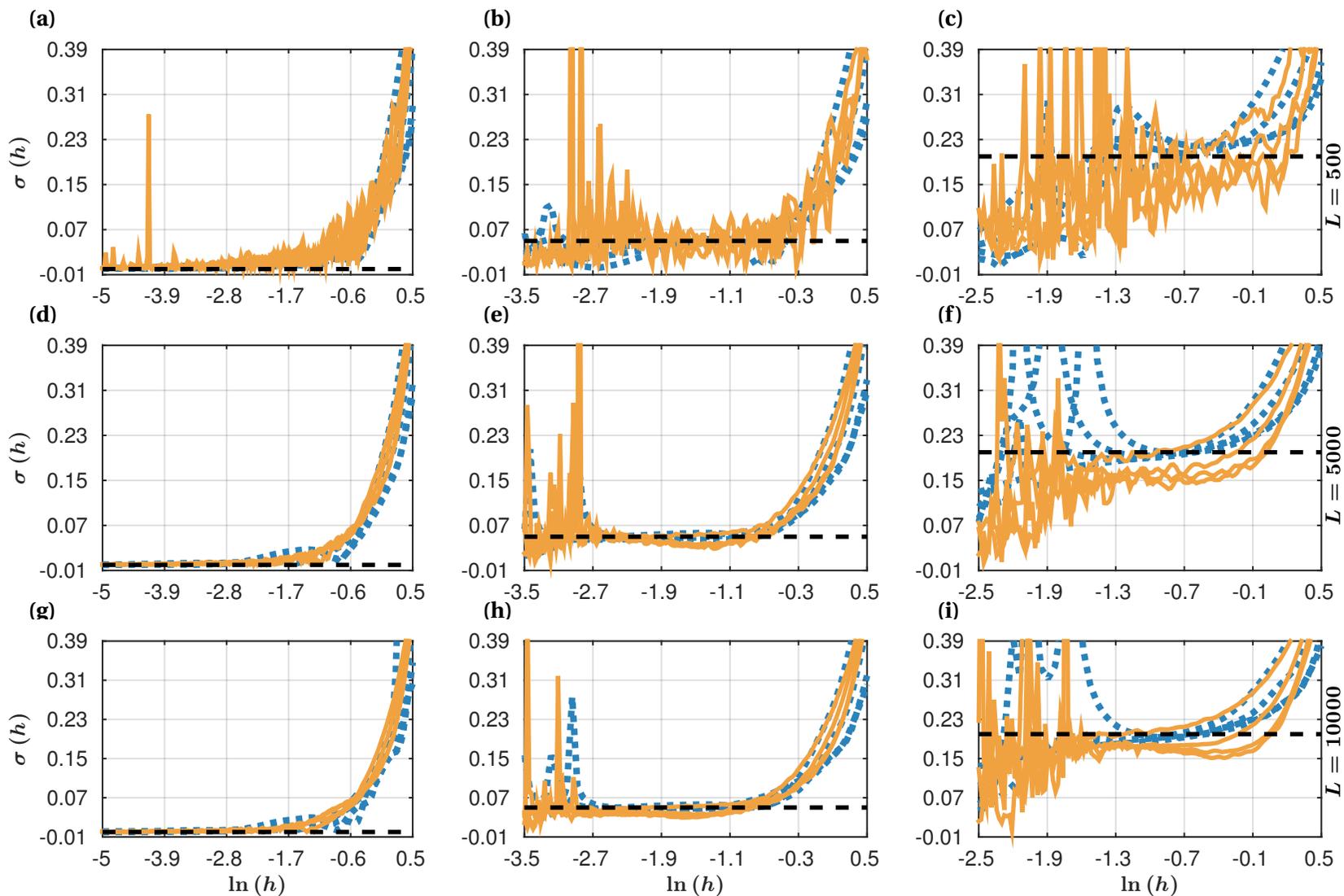


Figura 6.3. Transformación de Hénon. EDE para el nivel de ruido. $\sigma_m^T(h)$ (azul punteado) y $\sigma_m^U(h)$ (naranja) para $m = [2, 4, 6, 8]$. $L = 500$: (a) Sin ruido, (b) $\sigma = 0.05$ y (c) $\sigma = 0.2$. $L = 5000$: (d) sin ruido, (e) $\sigma = 0.05$ y (f) $\sigma = 0.2$. $L = 10000$: (g) sin ruido, (h) $\sigma = 0.05$ e (i) $\sigma = 0.2$. El verdadero valor de σ se muestra en línea negra discontinua.

En la figura [6.3](#) se presenta el comportamiento de los EsDE $\sigma_m^U(h)$ y $\sigma_m^T(h)$ en función del $\ln h$ con $m = \{2, 4, 6, 8\}$. Estos estimadores se calcularon a partir de series temporales provenientes de la transformación de Hénon bajo distintos niveles de ruido $\sigma = \{0, 0.05, 0.2\}$ y distinto número de vectores de estado $L = \{500, 5000, 10000\}$. Además, el valor numérico de σ se presenta en cada figura.

Podemos ver en las figuras [6.3a](#), [6.3d](#) y [6.3g](#) que, en ausencia de ruido, los estimadores $\sigma_m^U(h)$ y $\sigma_m^T(h)$ se acercan al valor verdadero de nivel de ruido ($\sigma = 0$) para valores pequeños de h y sin importar el valor de m . Vemos también que el estimador $\sigma_m^U(h)$ tiene un comportamiento oscilatorio cuya amplitud se ve atenuada al incrementar el número de vectores de estado.

Para niveles de ruido pequeños ($\sigma = 0.05$) se muestra en las figuras [6.3b](#), [6.3e](#) y [6.3h](#) que pueden distinguirse tres tipos de comportamientos diferentes para tres regiones de $\ln h$. Para valores pequeños de h , el estimador $\sigma_m^U(h)$ se caracteriza por un oscilación de gran amplitud (región I). A continuación encontramos la región II donde hay una disminución importante de la varianza del estimador $\sigma_m^U(h)$. Finalmente, en la región III, $\sigma_m^U(h)$ incrementa su valor rápidamente. Se puede observar que la región II es en la que mejor se aproxima $\sigma_m^U(h)$ a σ . Además, este comportamiento se repite para los diferentes valores de m y L .

Para valores de ruido más elevados ($\sigma = 0.2$) la dependencia de $\sigma_m^U(h)$ de h también puede ser dividida en las tres regiones antes mencionadas (figuras [6.3c](#), [6.3f](#) y [6.3i](#)). La región II sigue siendo el mejor lugar para estimar σ , sin embargo ésta es más pequeña en comparación con casos de menor nivel de ruido. Por último se puede ver que el estimador $\sigma_m^T(h)$ tiene un comportamiento similar al de $\sigma_m^U(h)$ pero con menos oscilaciones.

6.5 Estimador automático de invariantes

Con el fin de realizar un análisis estadístico de los estimadores dependientes de la escala aquí propuestos, hemos diseñado un algoritmo que nos permitirá calcular los invariantes σ , D y K_2 a partir de $\sigma_m^U(h)$, $D_m^U(h)$ y $K_m^U(h)$, respectivamente. Esta metodología está basada en el estudio del comportamiento de cada EDE presentado en secciones anteriores.

En la ecuación [\(6.7\)](#) se puede observar que $D_m^U(h)$ depende de σ para ser estimado. A su vez, la ecuación [\(6.16\)](#) nos muestra que $K_m^U(h)$ depende de la estimación de σ y de D . Por otro lado, el estimador $\sigma_m^U(h)$ (ecuación [\(6.18\)](#)) no depende de la estimación de ninguno de los otros dos invariantes. En este sentido, el orden lógico

Capítulo 6. Estimadores basados en la integral de correlación U

para estimar estos invariantes es primero σ , luego D y finalmente K_2 .

La metodología propuesta comienza con el cálculo de las integrales de correlación $U_m^{\beta=m}(h)$ y $U_{m+2}^{\beta=m}(h)$ para valores de $m \geq 2$. Luego, se debe estimar $\sigma_m^U(h)$ utilizando la ecuación (6.18) y estas dos ICUs. El siguiente paso consiste en buscar un rango adecuado de valores de h donde poder estimar σ . Como ya se sabe σ es un invariante, por tanto no debe depender del valor de h . En este sentido, σ debe ser estimado en un rango de valores de h donde $\sigma_m^U(h)$ sea aproximadamente constante. En consecuencia para encontrar este rango de valores de h se usará la derivada de $\sigma_m^U(h)$. Se elegirá como centro de este intervalo al valor de h para el cual el valor absoluto de dicha derivada sea más cercano a cero. De esta forma se estimará el nivel de ruido para cada uno de los valores de m y finalmente σ será igual al promedio de éstos.

Una vez encontrado σ se procede a estimar la dimensión de correlación D . Con este fin se calcula $D_m^U(h)$ mediante la ecuación (6.7), utilizando las ICUs $U_m^{\beta=m}(h)$, $U_m^{\beta=m-2}(h)$ y el valor de σ obtenido con anterioridad. Como se mencionó en la sección 6.2, el estimador $D_m^U(h)$ se aproxima a D cuando la distancia entre las curvas para diferentes valores de m es mínima (ver figura 6.1b). En consecuencia, para encontrar un rango adecuado de valores h donde estimar D se usará la varianza de $D_m^U(h)$ a través de m . Este rango se centrará en el valor de h para el cual esta varianza sea mínima. Luego, se calculará la dimensión de correlación para cada valor de m y D será fijado como el promedio de estos valores.

Por último, la entropía de correlación puede ser estimada utilizando $K_m^U(h)$ (ecuación (6.16)), calculada con $U_m^{\beta=m}(h)$, $U_{m+2}^{\beta=m+2}(h)$ y los valores de σ y D encontrados previamente. Como K_2 es un invariante, esta no debe depender del valor de h , por tanto, el rango de valores de h donde se estimará K_2 se escogerá observando la derivada de $K_m^U(h)$. Dicho rango se centrará en el valor de h para el cual el valor absoluto de esta derivada sea más cercano a cero. Luego, se debe calcular la entropía de correlación para cada valor de m y K_2 será su promedio.

Debemos mencionar que la implementación de esta metodología (ver algoritmo 4) requiere que las derivadas sean aproximadas utilizando la transformada ondata [144] ya que esto mejora su convergencia.

Para estudiar estadísticamente el comportamiento de nuestro algoritmo utilizamos 128 realizaciones de la transformación de Hénon, cada una con una condición inicial distinta. Cada serie temporal fue contaminada con distintos niveles de ruido y normalizada para tener varianza unitaria. Luego las ICUs requeridas por el algoritmo fueron calculadas utilizando $m = \{4, 5, \dots, 8\}$. En la figura 6.4 se presenta

un gráfico de cajas de la estimación automática de los invariantes σ , D y K_2 para diferente número de vectores de estado $L = \{500, 1000, 3000, 5000, 10000\}$ y bajo diferentes niveles de ruido $\sigma = \{0, 0.05, 0.2\}$. Se muestra también el valor reportado de cada invariante.

Algoritmo 4. Detector automático de invariantes utilizando los estimadores dependientes de la escala $\sigma_m^U(h)$, $D_m^U(h)$ y $K_m^U(h)$.

- 1: Calcular $U_m^{\beta=m}(h)$ mediante el algoritmo [3], utilizando el cuadrado de la distancia z_ω entre vectores de estado y ruido $\mu_\omega \sim \chi_m^2(h)$.
 - 2: Calcular $U_{m+2}^{\beta=m}(h)$ mediante el algoritmo [3] formando vectores de estado con dimensión $(m+2)$, calculando el cuadrado de la distancia z_ω entre ellos y utilizando ruido $\mu_\omega \sim \chi_m^2(h)$. Notar que los grados de libertad de la distribución Chi cuadrada es m .
 - 3: Obtener el EDE $\sigma_m^U(h)$ usando la ecuación (6.18). Estimar σ dentro de un rango de valores de h centrado en el h donde la $|\mathrm{d}\sigma_m^U(h)/\mathrm{d}h|$ está más cerca de cero. $|\cdot|$ denota el valor absoluto.
 - 4: Calcular el EDE $D_m^U(h)$ usando la ecuación (6.7) y la estimación de σ obtenida en el paso 3. Estimar D dentro de un rango de valores de h centrado en el h donde la varianza de $D_m^U(h)$ para los diferentes valores de m sea mínima.
 - 5: Obtener el EDE $K_m^U(h)$ usando la ecuación (6.16) y los valores de σ y D obtenidos en los pasos 3 y 4. Estimar K_2 dentro de un rango de valores de h centrado en el h donde la $|\mathrm{d}K_m^U(h)/\mathrm{d}h|$ está más cerca de cero.
-

La estimación de σ se muestra en las figuras [6.4a], [6.4b] y [6.4c]. De forma general vemos que la metodología propuesta logra una buena estimación del nivel de ruido, incluso para grandes cantidades de ruido y poca cantidad de datos (ver $L = 500$ en la figura [6.4c]). En ausencia de ruido es posible observar en la figura [6.4a] que esta metodología sobrestima a σ , sin embargo, para niveles de ruido mayores este algoritmo subestima a σ (figuras [6.4b] y [6.4c]). Vemos también que la varianza de la estimación decrece al incrementar el número de vectores de estado disponibles.

La estimación de la dimensión de correlación se puede observar en las figuras [6.4d], [6.4e] y [6.4f]. Para el caso sin ruido vemos en la figura [6.4d] que la estimación de D por medio de este algoritmo está muy cerca del valor reportado en la literatura ($D = 1.22$). Vemos también que la varianza de la estimación decrece al aumentar L . Para niveles moderados de ruido ($\sigma = 0.05$) se observa en la figura [6.4e] que esta metodología sobrestima ligeramente a D . No obstante, el buen desempeño es difícil de mantener para niveles mayores de ruido ($\sigma = 0.2$). En la figura [6.4f] vemos que aunque la estimación de D no es precisa, la mediana de la estimación no está lejos del valor reportado para la dimensión de correlación.

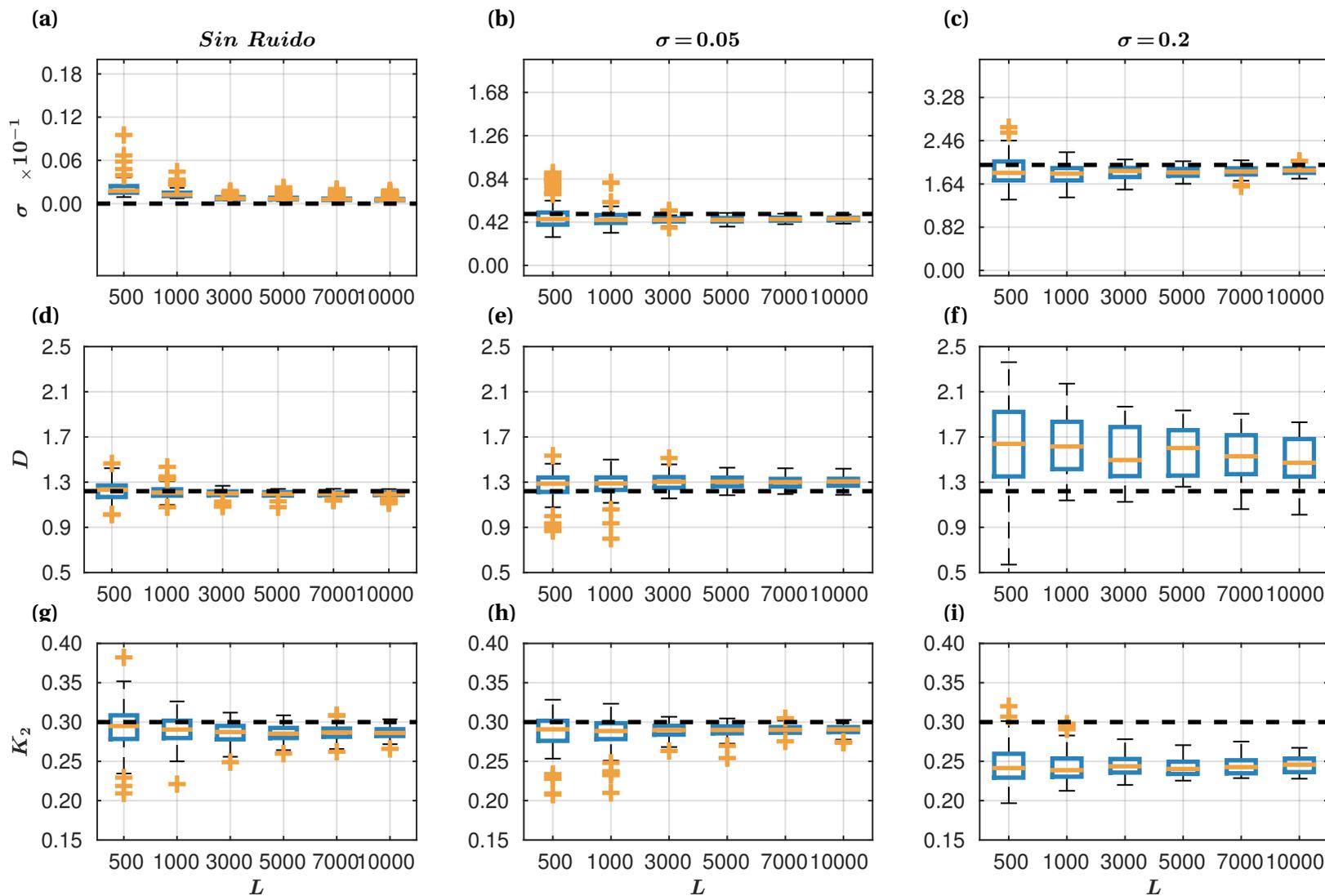


Figura 6.4. Transformación de Hénon. Gráfico de cajas de la estimación automática de los invariantes σ , D y K_2 para diferente número de vectores de estado $L = \{500, 1000, 3000, 5000, 10000\}$. Nivel de ruido: (a) sin ruido, (b) $\sigma = 0.05$ y (c) $\sigma = 0.2$. Dimensión de correlación: (d) sin ruido, (e) $\sigma = 0.05$ y (f) $\sigma = 0.2$. Entropía de correlación: (g) sin ruido, (h) $\sigma = 0.05$ e (i) $\sigma = 0.2$. El valor reportado de cada invariante se muestra en línea negra discontinua.

En las figuras [6.4g](#), [6.4h](#) y [6.4i](#) se presenta la estimación de K_2 . En ausencia de ruido vemos que esta metodología converge a valores muy cercanos al reportado en la literatura ($K_2 = 0.3$). De forma interesante vemos que para $L = 10\,000$ la mediana de la estimación es igual a 0.28, siendo este un valor menor al reportado en la literatura. Esto nos lleva a pensar que tal vez el valor de K_2 para este sistema sea menor a 0.3. Sin embargo estos resultados son numéricos y no existe en la literatura un resultado teórico para esta transformación. Se muestra en la figura [6.4h](#) que al aumentar el valor del ruido a $\sigma = 0.05$ este algoritmo subestima ligeramente al valor reportado de K_2 . Y aunque para valores mayores de ruido ($\sigma = 0.2$) la estimación de K_2 no es precisa, el valor alcanzado por el método aquí propuesto se acerca mucho al valor reportado.

Los resultados anteriores nos permiten afirmar que el algoritmo presentado en esta sección alcanza muy buen desempeño en la estimación de los invariantes σ , D y K_2 . Para grandes niveles de ruido la estimación de D y K_2 no es precisa debido a los errores cometidos en la estimación σ . Sin embargo, los valores estimados no están lejos de los valores reportados en la literatura.

A continuación, analizamos el desempeño de la metodología propuesta utilizando el sistema de Mackey-Glass (ver tabla [3.1](#)) con parámetros $a = 0.2$, $b = 0.1$, $c = 10$ y $\lambda = 23$. Escogimos este sistema ya que sus valores de dimensión y entropía de correlación están disponibles en la literatura [\[79\]](#). Utilizamos 128 realizaciones, cada una con una condición inicial distinta, a las cuales se les agregó ruido hasta alcanzar diferentes niveles ($\sigma = \{0, 0.05, 0.2\}$). Además se normalizaron para tener varianza unitaria. Las $ICU^{\beta=m}$ s fueron calculadas para $m = \{4, 6, \dots, 16\}$ y el retardo de inmersión τ fue elegido utilizando la función de información mutua ($\tau = 20$). Los 15 vecinos más cercanos de cada vector de estado fueron descartados para evitar correlaciones temporales [\[12\]](#).

Los resultados de esta simulación se resumen en la figura [6.5](#) donde se muestran los gráficos de cajas correspondientes a la estimación de σ , D y K_2 para diferente número de vectores de estado $L = \{1\,000, 3\,000, 5\,000, 7\,000, 10\,000\}$ y niveles de ruido. Podemos ver en la figura [6.5a](#) que, en el caso sin ruido, la estimación de σ es menos precisa para este sistema que para la transformación de Hénon para longitudes de datos cortas. No obstante, para $Q \geq 5\,000$ la convergencia al verdadero valor de σ es muy buena. Para niveles de ruido moderado ($\sigma = 0.05$) vemos en la figura [6.5b](#) que σ es ligeramente subestimado. En la figura [6.5c](#) se observa que para grandes niveles de ruido $\sigma = 0.2$ la precisión de este algoritmo para estimar σ , decrece en comparación con lo obtenido para niveles menores de ruido. Sin embargo, el valor de la mediana de cada caja está muy cerca al valor real de σ . Como sería de espe-

rar, la varianza de la estimación decrece al incrementar el número de vectores de estado disponibles.

Los resultados en la estimación de D se muestran en las figuras 6.5d, 6.5e y 6.5f. Para los casos sin ruido (figura 6.5d) y $\sigma = 0.05$ (figura 6.5e) vemos que este algoritmo subestima ligeramente el valor de la dimensión de correlación reportada para este sistema ($D = 2.44$ [79]), aún así, dicha estimación es muy cercana a este valor. Es de destacar que para valores grandes de ruido ($\sigma = 0.2$) la figura 6.5f muestra convergencia al valor reportado de D . Comparando estos resultados con los obtenidos para la transformación de Hénon nos hace pensar en la posibilidad de que el valor reportado en la literatura para este invariante pueda ser un poco menor a 2.3. Sin embargo, cabe la posibilidad de que la estimación de D esté sesgada debido a imprecisiones en la estimación de σ .

Por otro lado, vemos en las figuras 6.5g, 6.5h y 6.5i que la estimación de K_2 por medio de esta metodología es coherente con su valor reportado $K_2 = 0.008$ [79]. Esto se da incluso para altos niveles de ruido y con imprecisiones en la estimación de σ y D .

Un aspecto importante para el buen funcionamiento de método aquí propuesto es el uso de la transformada ondita para aproximar derivadas. Esta implementación incorpora un filtro pasa-bajos el cual ayuda a mitigar las oscilaciones de alta frecuencia características de los estimadores basados en la ICU.

Claramente, otra forma de mitigar estas oscilaciones es incrementando el número de comparadores en el algoritmo de correlación asistido por ruido (ver figura 4.1). Lo anterior puede ser llevado a cabo mediante dos caminos: el primero consiste en incrementar la longitud de la serie temporal, lo que no es siempre posible. El segundo camino consiste en calcular todas las distancias \tilde{z}_ω (o sus cuadrados) entre todos los vectores de estado. Esto incrementará el número de comparadores en el algoritmo. Luego, habría que hacer copias de cada distancia \tilde{z}_ω y a cada copia agregarle una realización de ruido μ_ω diferente. Esta información deberá pasar a los comparadores donde se realiza el proceso de umbralado. Finalmente, la suma de correlación asistida por ruido $\hat{S}(h)$ será calculada como el promedio entre la salida de todos los comparadores. Es necesario aclarar que la suavidad de las curvas y el costo computacional se incrementan con el número de copias.

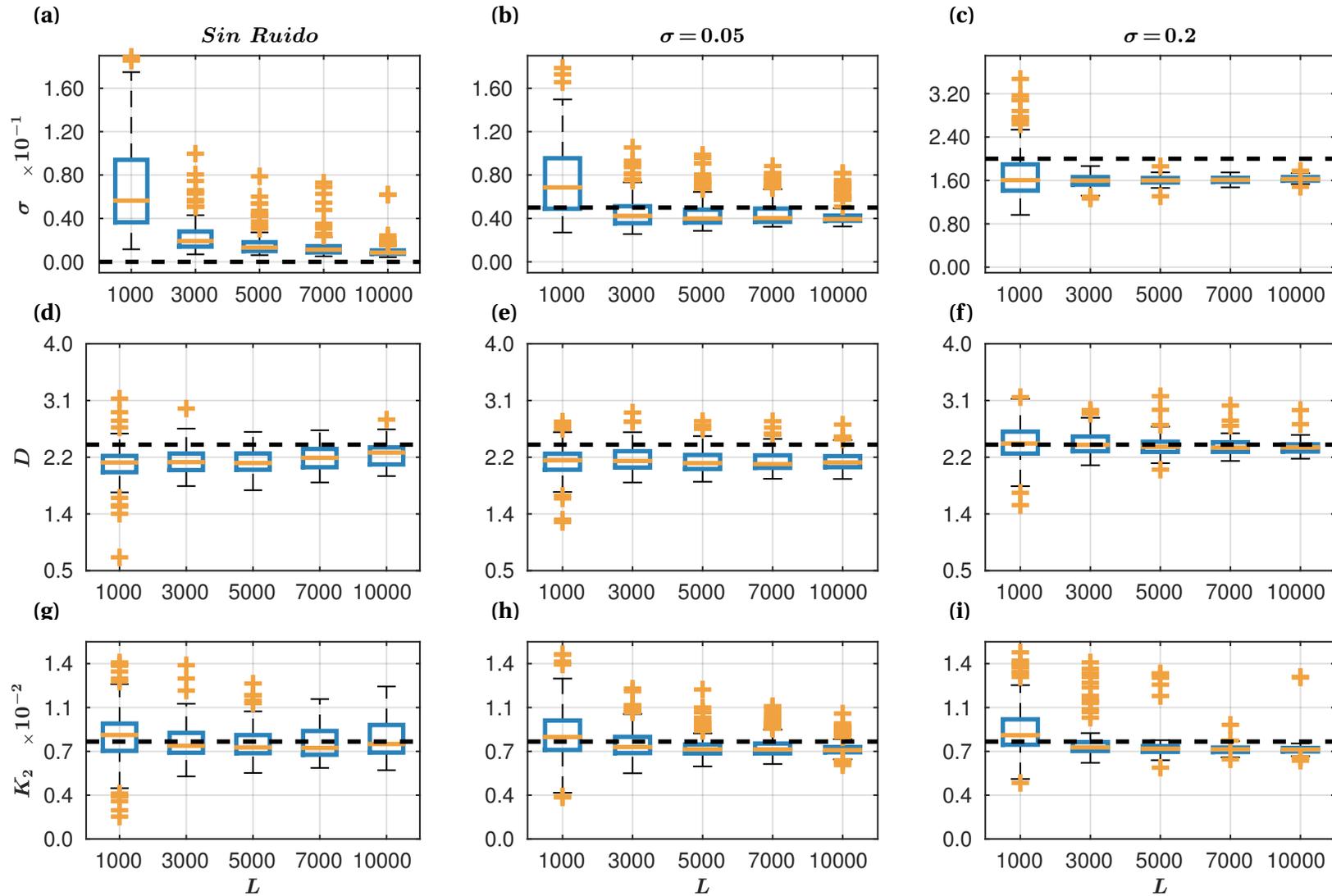


Figura 6.5. Sistema de Mackey-Glass. Gráfico de cajas de la estimación automática de los invariantes σ , D y K_2 para diferente número de vectores de estado $L = \{500, 1000, 3000, 5000, 10000\}$. Nivel de ruido: (a) sin ruido, (b) $\sigma = 0.05$ y (c) $\sigma = 0.2$. Dimensión de correlación: (d) sin ruido, (e) $\sigma = 0.05$ y (f) $\sigma = 0.2$. Entropía de correlación: (g) sin ruido, (h) $\sigma = 0.05$ e (i) $\sigma = 0.2$. El valor reportado de cada invariante se muestra en línea negra discontinua.

6.6 Conclusiones

En este capítulo se dedujeron analíticamente los estimadores dependientes de la escala $\sigma_m^U(h)$, $D_m^U(h)$ y $K_m^U(h)$ para los invariantes σ , D y K_2 respectivamente. Hemos estudiado el comportamiento de estos estimadores ante diversos niveles de ruido y longitudes de datos. En comparación con los EsDE propuestos por Nolte y col. podemos concluir que tanto los estimadores de nivel de ruido $\sigma_m^U(h)$ y $\sigma_m^T(h)$ como los de dimensión de correlación $D_m^U(h)$ y $D_m^T(h)$ se comportan de forma similar. Por otro lado, el estimador $K_m^U(h)$ es muy superior a su contraparte $K_m^T(h)$ para la estimación de K_2 .

En base a los estimadores $\sigma_m^U(h)$, $D_m^U(h)$ y $K_m^U(h)$ hemos propuesto un algoritmo para la detección automática de los invariantes σ , D y K_2 . Mostramos que esta metodología estima de manera confiable cada uno de estos invariantes, incluso ante la presencia de grandes niveles de ruido. Pese a que la precisión de dicha estimación está fuertemente ligada a la precisión en el cálculo de σ , los resultados obtenidos con modelos clásicos usados en la literatura muestran un muy buen desempeño del enfoque aquí propuesto.

*Somewhere, something incredible is waiting
to be known.*

Carl Sagan

7

Conclusiones finales y trabajos futuros

Esta tesis doctoral contiene dos avances diferentes que permiten caracterizar la complejidad de un sistema. Hemos investigado profundamente cada una de estas metodologías y presentamos aquí nuestras conclusiones.

El primer avance está relacionado con la ApEn. Este estadístico presenta dos problemas principales: la alta dependencia al valor de sus parámetros y la incapacidad para discriminar dinámicas de diferente complejidad bajo niveles altos de ruido. El primero de estos problemas es parcialmente solucionado al utilizar el estimador $ApEn_{max}$. Sin embargo altos niveles de ruido también disminuyen la capacidad de discriminación de este estimador. Como solución propusimos el estimador h_{max} . Mediante simulaciones con sistemas dinámicos de baja y alta dimensión, estudiamos el comportamiento tanto de h_{max} como de $ApEn_{max}$ en función de sus parámetros. Observamos que a medida que cambia la dimensión de inmersión estos estimadores se complementan. Es decir, cambios de regularidad que no se reflejan $ApEn_{max}$, pueden reflejarse en h_{max} y viceversa. En este sentido concluimos que h_{max} brinda información útil para el proceso de discriminación entre dinámicas. Más aún, el uso conjunto de h_{max} y $ApEn_{max}$ hace que el proceso de clasificación sea robusto ante grandes cantidades de ruido y poca longitud de datos.

En base a estos resultados diseñamos un detector de episodios ictales a partir de

Capítulo 7. Conclusiones finales y trabajos futuros

señales de EEG. Los resultados muestran que, en la señal estudiada, es posible detectar este tipo de episodios epilépticos utilizando conjuntamente $ApEn_{max}$ y h_{max} . No obstante, es necesario un estudio estadístico profundo con una base de datos completa. También, hemos diseñado un detector de voces patológicas a partir de señales de voz. Este detector fue puesto a prueba utilizando una base de datos. Los resultados de validación cruzada, nos permiten concluir que utilizando los estimadores $ApEn_{max}$ y h_{max} es posible clasificar voces patológicas y normales. También que la clasificación es más robusta cuando se utiliza estimaciones de $ApEn_{max}$ y h_{max} para varios valores de m superiores pero cercanos a la mínima dimensión de inmersión.

El segundo avance contenido en este documento guarda relación con los invariantes D y K_2 . Estas cantidades se estiman a partir de la integral de correlación, que a su vez, se calcula a partir de algoritmo de correlación. En la literatura se reportan dos tipos de integrales de correlación: la estándar y la ICG. En presencia de ruido la integral de correlación estándar es incapaz de estimar adecuadamente D y K_2 . Esta deficiencia es superada por la ICG. Sin embargo, ésta presenta problemas para estimar K_2 . Como solución propusimos la ICAR y la ICU, las cuales pueden ser calculadas mediante el algoritmo de correlación asistido por ruido. Un resultado muy importante de este estudio es que la integral de correlación estándar y la ICG son casos particulares de la ICAR. Concluimos que agregar ruido a las distancias entre vectores de estado en el algoritmo de correlación es equivalente a calcular la integral de correlación con una función núcleo asociada a la naturaleza del ruido. Es decir, el ruido induce una función núcleo. En base a esto, propusimos la ICU cuya función núcleo incorpora información sobre la dimensión de inmersión. Las simulaciones con la ICU nos permiten concluir que ésta es capaz de estimar D y σ de forma similar a la ICG. Sin embargo, es importante resaltar que la estimación de K_2 es mucho mejor a partir de la ICU que de la ICG.

En base a la ICU deducimos estimadores dependientes de escala para D , K_2 y σ y diseñamos una metodología para estimar automáticamente estos invariantes. La validación estadística de esta metodología nos lleva a concluir que mediante los EsDE propuestos es posible aproximar los invariantes D , K_2 y σ de forma confiable, incluso en presencia de ruido.

7.1 Trabajos futuros

Con respecto al estimador h_{max} aún queda por investigar su relación con la regularidad de una señal. Es decir, debemos responder la pregunta de si este estimador puede medir regularidad. Esto nos llevaría a comprender su posible interpretación física.

Los buenos resultados de la ICU nos animan a buscar aplicaciones con señales biomédicas. Queda como trabajo futuro aplicar esta idea al estudio y detección de patologías.

Debido a la estrecha relación entre el algoritmo de correlación y el algoritmo utilizado para calcular la ApEn y la SampEn, cabe la posibilidad de extender la idea general del algoritmo de correlación asistido por ruido al cálculo de estos invariantes.

A

Apéndice A

En este apéndice se presentaran las deducciones de varias expresiones analíticas obtenidas y utilizadas en esta tesis para:

$$\frac{d \ln U_m^{\beta=m}(h)}{d \ln h},$$

que es el funcional que permite encontrar D a partir de la $U_m^{\beta=m}(h)$. También analizaremos su comportamiento con y sin ruido.

A.1 Funcional para D basado en ICU $^{\beta=m}$

Recordemos que en su forma general la ICU es (ecuación (5.3)):

$$U_m^\beta(h) = \frac{\widehat{\phi}}{2} (2\sigma)^D e^{-m\tau\kappa_2} \frac{\Gamma(D/2)\Gamma((\beta+m)/2)}{\Gamma(\beta/2)\Gamma(m/2+1)} \left(\frac{h}{2\sigma}\right)^m {}_2F_1\left(\frac{\beta+m}{2}, \frac{m-D}{2}; \frac{m+2}{2}; \frac{-h^2}{4\sigma^2}\right). \quad (\text{A.1})$$

Apéndice A.

Definiendo:

$$\begin{aligned} a &= \frac{\beta + m}{2}; & b &= \frac{m - D}{2}; & c &= \frac{m + 2}{2}; \\ z &= -\frac{h^2}{4\sigma^2}; & \frac{dz}{dh} &= -\frac{h}{2\sigma^2} \end{aligned} \quad (\text{A.2})$$

y

$$P = (-1)^{m/2} \frac{\widehat{\phi}}{2} (2\sigma)^D \frac{\Gamma(D/2) \Gamma((m + \beta)/2)}{\Gamma(\beta/2) \Gamma(m/2 + 1)} e^{-m\tau K_2}.$$

Vemos que podemos escribir (A.1) como:

$$U_m^\beta(z) = P z^{m/2} {}_2F_1(a, b; c; z).$$

La derivada de $U_m^\beta(z)$ con respecto a h es:

$$\frac{d}{dh} U_m^\beta(z) = \frac{d}{dz} [P z^{m/2} {}_2F_1(a, b; c; z)] \frac{dz}{dh}.$$

Usando [143] ec. 15.2.1]:

$$\frac{d}{dh} U_m^\beta(z) = \frac{-Ph}{2\sigma^2} z^{m/2} \left[\frac{m}{2z} {}_2F_1(a, b; c; z) + \frac{ab}{c} {}_2F_1(a+1, b+1; c+1; z) \right].$$

La derivada logarítmica de $U_m^\beta(h)$ se puede obtener como:

$$\begin{aligned} \frac{d \ln U_m^\beta(z)}{d \ln h} &= \frac{h}{U_m^\beta(z)} \frac{d}{dh} U_m^\beta(z) \\ &= \frac{2P z^{m/2+1}}{P z^{m/2} {}_2F_1(a, b; c; z)} \left[\frac{m}{2z} {}_2F_1(a, b; c; z) + \frac{ab}{c} {}_2F_1(a+1, b+1; c+1; z) \right] \\ &= m + \frac{2abz}{c} \frac{{}_2F_1(a+1, b+1; c+1; z)}{{}_2F_1(a, b; c; z)}, \end{aligned} \quad (\text{A.3})$$

y reemplazando los valores de a , b , c y z obtenemos:

$$\frac{d \ln U_m^\beta(h)}{d \ln h} = m - \frac{2m(m-D)}{m+2} \left(\frac{h^2}{4\sigma^2} \right) \frac{{}_2F_1\left(m+1, \frac{m-D}{2}+1; \frac{m+2}{2}+1; \frac{-h^2}{4\sigma^2}\right)}{{}_2F_1\left(m, \frac{m-D}{2}; \frac{m+2}{2}; \frac{-h^2}{4\sigma^2}\right)}. \quad (\text{A.4})$$

Podemos encontrar otras expresiones para $d \ln U_m^{\beta=m}(h)/d \ln h$ utilizando algunas de las identidades de la ${}_2F_1(a, b; c; t)$. Por ejemplo, a partir de la ecuación (A.3) y

A.2. Comportamiento de $d \ln U_m^{\beta=m}(h)/d \ln h$

utilizando la identidad [142] ec. 9.137.12]:

$$\frac{zb}{c} {}_2F_1(a+1, b+1; c+1; z) = {}_2F_1(a+1, b; c; z) - {}_2F_1(a, b; c; z),$$

podemos escribir:

$$\frac{d \ln U_m^\beta(z)}{d \ln h} = 2a \frac{{}_2F_1(a+1, b; c; z)}{{}_2F_1(a, b; c; z)} - \beta. \quad (\text{A.5})$$

Así mismo, utilizando $(zab/c) {}_2F_1(a+1, b+1; c+1; z) = (c-1)[{}_2F_1(a, b; c-1; z) - {}_2F_1(a, b; c; z)]$ encontramos que:

$$\frac{d \ln U_m^\beta(z)}{d \ln h} = 2(c-1) \frac{{}_2F_1(a, b; c-1; z)}{{}_2F_1(a, b; c; z)}. \quad (\text{A.6})$$

A.2 Comportamiento de $d \ln U_m^{\beta=m}(h)/d \ln h$

Para series temporales muy ruidosas ($\sigma \gg h$), $z \rightarrow 0$. Tomando en cuenta que ${}_2F_1(a, b; c; 0) = 1$ y haciendo $\beta = m$ obtenemos a partir de la ecuación (A.5) que:

$$\frac{d \ln U_m^\beta(h)}{d \ln h} = m.$$

Cuando la series temporales están libres de ruido o contiene una cantidad relativamente pequeña ($\sigma \ll h$) el valor de $z \rightarrow -\infty$. Utilizando [143] ec. 15.3.7] y el hecho de que $1/z \rightarrow 0$:

$$\begin{aligned} {}_2F_1(a, b; c; z) &= \frac{\Gamma(c)\Gamma(b-a)}{\Gamma(b)\Gamma(c-a)} (-z)^{-a} {}_2F_1(a, 1-c+a; 1-b+a; 1/z) \\ &\quad + \frac{\Gamma(c)\Gamma(a-b)}{\Gamma(a)\Gamma(c-b)} (-z)^{-b} {}_2F_1(b, 1-c+b; 1-a+b; 1/z) \\ &= \frac{\Gamma(c)\Gamma(b-a)}{\Gamma(b)\Gamma(c-a)} (-z)^{-a} + \frac{\Gamma(c)\Gamma(a-b)}{\Gamma(a)\Gamma(c-b)} (-z)^{-b}. \end{aligned}$$

Podemos reescribir la ecuación (A.3) como:

$$\frac{d \ln U_m^\beta(h)}{d \ln h} = m - \frac{2ab}{c} \frac{\frac{\Gamma(c+1)\Gamma(b-a)}{\Gamma(b+1)\Gamma(c-a)} (-z)^{b-a} + \frac{\Gamma(c+1)\Gamma(a-b)}{\Gamma(a+1)\Gamma(c-b)}}{\frac{\Gamma(c)\Gamma(b-a)}{\Gamma(b)\Gamma(c-a)} (-z)^{b-a} + \frac{\Gamma(c)\Gamma(a-b)}{\Gamma(a)\Gamma(c-b)}}.$$

Apéndice A.

Finalmente, teniendo en cuenta que $a > b$ y restaurando el valor de b podemos ver que:

$$\begin{aligned}\frac{d \ln U_m^\beta(h)}{d \ln h} &= m - \frac{2ab}{c} \left[\frac{\Gamma(c+1)\Gamma(a-b)}{\Gamma(a+1)\Gamma(c-b)} \right] / \left[\frac{\Gamma(c)\Gamma(a-b)}{\Gamma(a)\Gamma(c+b)} \right] \\ &= m - 2(ab/c)(c/a) \\ &= D.\end{aligned}\tag{A.7}$$

Hemos demostrado que bajo condiciones ideales el funcional $d \ln U_m^{\beta=m}(h)/d \ln h$ es igual a la dimensión de correlación. Por el contrario para señales con una relación señal a ruido pequeña este funcional tiende a la dimensión de inmersión m .

B

Apéndice B

En este apéndice se deducirá la expresión analítica para:

$$\ln \frac{U_{m+2}^{\beta=m+2}(h)}{U_m^{\beta=m}(h)},$$

que es el funcional que permite encontrar K_2 a partir de la $U_m^{\beta=m}(h)$. También analizaremos su comportamiento con y en ausencia de ruido.

B.1 Funcional para K_2 basado en $ICU^{\beta=m}$

Partiremos de la forma general de la ICU (ecuación (5.3)):

$$U_m^\beta(h) = \frac{\widehat{\phi}}{2} (2\sigma)^D e^{-m\tau K_2} \frac{\Gamma(D/2)\Gamma((\beta+m)/2)}{\Gamma(\beta/2)\Gamma(m/2+1)} \left(\frac{h}{2\sigma}\right)^m {}_2F_1\left(\frac{\beta+m}{2}, \frac{m-D}{2}; \frac{m+2}{2}; \frac{-h^2}{4\sigma^2}\right). \quad (\text{B.1})$$

Haciendo:

$$M = \frac{\widehat{\phi}}{2} (2\sigma)^D \Gamma(D/2), \quad (\text{B.2})$$

utilizando las definiciones de la ecuación (A.2) y la ecuación (B.2) podemos escribir

Apéndice B.

la ecuación (B.1) como:

$$U_m^\beta(z) = M \frac{\Gamma(a)}{\Gamma(\beta/2)\Gamma(c)} (-z)^{m/2} {}_2F_1(a, b; c; z) e^{-m\tau K_2}. \quad (\text{B.3})$$

Escribimos el cociente $U_{m+2}^{\beta+2}(z)/U_m^\beta(z)$ como:

$$\begin{aligned} \frac{U_{m+2}^{\beta+2}(z)}{U_m^\beta(z)} &= \frac{\Gamma(a+2)\Gamma(\beta/2)\Gamma(c)}{\Gamma(a)\Gamma((\beta+2)/2)\Gamma(c+1)} (-z)^2 \frac{{}_2F_1(a+2, b+1; c+1; z)}{{}_2F_1(a, b; c; z)} e^{-2\tau K_2} \\ &= \frac{2a(a+1)}{\beta c} (-z)^2 \frac{{}_2F_1(a+2, b+1; c+1; z)}{{}_2F_1(a, b; c; z)} e^{-2\tau K_2}. \end{aligned} \quad (\text{B.4})$$

Restituyendo los valores de a, b, c, z y haciendo $\beta = m$ obtenemos:

$$\ln \frac{U_{m+2}^{\beta=m+2}(h)}{U_m^{\beta=m}(h)} = \ln \left\{ \left(\frac{m+1}{m+2} \right) \left(\frac{h^2}{\sigma^2} \right) \frac{{}_2F_1\left(m+2, \frac{m-D}{2}+1; \frac{m+2}{2}+1; \frac{-h^2}{4\sigma^2}\right)}{{}_2F_1\left(m, \frac{m-D}{2}; \frac{m+2}{2}; \frac{-h^2}{4\sigma^2}\right)} \right\} - 2\tau K_2. \quad (\text{B.5})$$

B.2 Comportamiento de $\ln U_{m+2}^{\beta=m+2}(h)/U_m^{\beta=m}(h)$

Teniendo en cuenta que para series temporales con un elemento de ruido apreciable $\sigma \gg h$ y que ${}_2F_1(a, b; c; 0) = 1$, es fácil ver a partir de la ecuación (B.5) que:

$$\ln \frac{U_{m+2}^{\beta=m+2}(h)}{U_m^{\beta=m}(h)} \rightarrow -\infty.$$

De igual forma en la que procedimos en el Apéndice A. Para casos sin ruido o con una cantidad relativamente poca ($\sigma \ll h$) el valor de $z \rightarrow -\infty$. Partiendo de la ecuación (B.4), utilizando [143, eq. 15.3.7] y el hecho de que $1/z \rightarrow 0$ es posible demostrar que:

$$\lim_{1/z \rightarrow 0} \frac{U_{m+2}^{\beta+2}(z)}{U_m^\beta(z)} = \frac{2a(a-1)}{\beta c} \frac{\frac{\Gamma(c+1)\Gamma(b-a-1)}{\Gamma(b+1)\Gamma(c-a-1)} (-z)^{b-a-1} + \frac{\Gamma(c+1)\Gamma(a-b+1)}{\Gamma(a+2)\Gamma(c-b)}}{\frac{\Gamma(c)\Gamma(b-a)}{\Gamma(b)\Gamma(c-a)} (-z)^{b-a} + \frac{\Gamma(c)\Gamma(a-b)}{\Gamma(a)\Gamma(c-b)}} e^{-2\tau K_2},$$

B.2. Comportamiento de $\ln U_{m+2}^{\beta=m+2}(h)/U_m^{\beta=m}(h)$

y sabiendo que $a > b$:

$$\begin{aligned} \lim_{1/z \rightarrow 0} \frac{U_{m+2}^{\beta+2}(z)}{U_m^\beta(z)} &= \frac{2a(a+1)}{\beta c} \left[\frac{\Gamma(c+1)\Gamma(a-b+1)}{\Gamma(a+2)\Gamma(c-b)} \right] / \left[\frac{\Gamma(c)\Gamma(a-b)}{\Gamma(a)\Gamma(c+b)} \right] e^{-2\tau K_2} \\ &= \frac{a-b}{\beta} e^{-2\tau K_2}. \end{aligned} \quad (\text{B.6})$$

Restaurando los valores de a , b , z y tomando logaritmo a ambos lados de la ecuación (B.6):

$$\ln \frac{U_{m+2}^{\beta+2}(h)}{U_m^\beta(h)} = \ln \left(\frac{D}{\beta} + 1 \right) - 2\tau K_2, \quad (\text{B.7})$$

y haciendo $\beta = m$:

$$\ln \frac{U_{m+2}^{\beta=m+2}(h)}{U_m^{\beta=m}(h)} = \ln \left(\frac{D}{m} + 1 \right) - 2\tau K_2. \quad (\text{B.8})$$

Finalmente, podemos analizar de la ecuación (B.8) que cuando $m \gg D$:

$$\ln \frac{U_{m+2}^{\beta=m+2}(h)}{U_m^{\beta=m}(h)} \rightarrow 2K_2\tau. \quad (\text{B.9})$$

C

Apéndice C

En este apéndice se presentara la deducción analítica para $\Delta_m^U(h)$, que es el funcional que permite encontrar σ a partir de la ICU. Encontraremos una aproximación que facilite su cálculo y analizaremos su comportamiento.

C.1 Funcional para σ basado en ICU $^{\beta=m}$

Nos basaremos en la metodología planteada en [84]. Partimos entonces del estimador

$$\Delta_{n,m}^U(h) = \frac{1}{n, m} \left[\frac{d \ln U_n^{\bar{\beta}}(h)}{d \ln h} - \frac{d \ln U_m^{\beta}(h)}{d \ln h} \right]. \quad (\text{C.1})$$

Utilizando la ecuación (A.5) y haciendo $n = m+2$ podemos escribir la ecuación (C.1):

$$\Delta_m^U(z) = \frac{1}{2} \left\{ 2(m+1) \frac{{}_2F_1(m+2, \bar{b}; \bar{c}; z)}{{}_2F_1(m+1, \bar{b}; \bar{c}; z)} - 2m \frac{{}_2F_1(m+1, b; c; z)}{{}_2F_1(m, b; c; z)} \right\}. \quad (\text{C.2})$$

donde $\bar{b} = (m+2-D)/2$, $\bar{c} = (m+4)/2$, $b = (m-D)/2$, $c = (m+2)/2$ y $z = -h^2/4\sigma^2$.

Analizaremos el estimador $\Delta_m^U(h)$ como una función de h con un σ fijo. Si $h \rightarrow 0$,

Apéndice C.

$z \rightarrow 0$ y ${}_2F_1(a+1, b; c; z)/{}_2F_1(a, b; c; z) = 1$. Por tanto:

$$\begin{aligned} \lim_{z \rightarrow 0} \Delta_m^U(z) &= \frac{1}{2} [2(m+1) - 2m] \\ &= 1. \end{aligned} \quad (\text{C.3})$$

Por otro lado, cuando $h \rightarrow \infty$, $z \rightarrow -\infty$. Utilizando [143, eq.15.3.7] y el hecho de que $1/z \rightarrow 0$ podemos demostrar que:

$$\frac{{}_2F_1(a+1, b; c; z)}{{}_2F_1(a, b; c; z)} = \frac{a-b}{a}, \quad (\text{C.4})$$

y en consecuencia:

$$\begin{aligned} \lim_{z \rightarrow 0} \Delta_m^U(z) &= \frac{1}{2} \left\{ 2 \left[m+1 - \frac{m+2-D}{2} \right] - 2 \left[m - \frac{m-D}{2} \right] \right\} \\ &= 0. \end{aligned} \quad (\text{C.5})$$

Podemos observar entonces que $\Delta_m^U(z)$ es una función que decrece de forma monótona desde 1 hasta 0 a una velocidad determinada solo por el valor de σ .

Para series temporales provenientes de sistemas dinámicos de baja dimensión y para m y n grandes se puede ver que $\bar{b} \approx \bar{c}$ y $b \approx c$. Teniendo en cuenta que ${}_2F_1(a, b; b; z) = (1-z)^{-a}$, la ecuación [C.4] puede reducirse a:

$$\begin{aligned} \Delta_m^U(z) &\approx \frac{1}{2} \left[\frac{2(m+1)}{1-z} - \frac{2m}{1-z} \right] \\ &\approx \frac{1}{1-z}. \end{aligned} \quad (\text{C.6})$$

Es posible analizar que la ecuación [C.6] también cumple con los límites presentados en las ecuaciones [C.3] y [C.5]. Finalmente reemplazando el valor de z tenemos que:

$$\Delta_m^U(h) \approx \frac{4\sigma^2}{4\sigma^2 + h^2}. \quad (\text{C.7})$$

Para finalizar debemos aclarar que para estimar el funcional $\Delta_m^U(h)$ se necesita la integral de correlación $U_{m+2}^{\beta=m}$ que es la ICU con dimensión de inmersión $m+2$ y dimensión de ruido $\beta = m$.

Lista de Publicaciones

- [1] J. F. Restrepo, G. Schlotthauer, M. E. Torres, Maximum approximate entropy and r threshold: a new approach for regularity changes detection, *Physica A: Statistical Mechanics and its Applications* 409 (2014) 97–109.
- [2] J. F. Restrepo, G. Schlotthauer, M. E. Torres, Máxima entropía aproximada y cambios de regularidad, en: XIX Congreso Argentino de Bioingeniería y VIII Jornadas de Ingeniería Clínica (SABI), 2013, p. 122.
- [3] J. F. Restrepo, G. Schlotthauer, M. E. Torres, Regularity changes detection using maximum approximate entropy and particle swarm optimization, en: XV Reunión de Trabajo en Procesamiento de la Información y Control (RPIC), 2013, pp. 475–480.
- [4] J. F. Restrepo, G. Schlotthauer, T. M. E, Maximum approximate entropy for normal and pathological voices classification, en: VI Latin American Congress on Biomedical Engineering (CLAIB), Paraná, Argentina 29, 30 & 31 October 2014, Vol. 49, 2015, pp. 548–551.
- [5] J. F. Restrepo, G. Schlotthauer, Noise-assisted estimation of attractor invariants, *Phys. Rev. E* 94 (2016) 012212.

Bibliografía

- [6] H. Poincaré, Sur le problème des trois corps et les équations de la dynamique, *Acta mathematica* 13 (1) (1890) A3–A270.
- [7] E. N. Lorenz, Deterministic nonperiodic flow, *Journal of the atmospheric sciences* 20 (2) (1963) 130–141.
- [8] D. Ruelle, F. Takens, On the nature of turbulence, *Communications in mathematical physics* 20 (3) (1971) 167–192.
- [9] R. M. May, Biological populations with nonoverlapping generations: stable points, stable cycles, and chaos, *Science* 186 (4164) (1974) 645–647.
- [10] S. Strogatz, *Nonlinear dynamics and chaos: with applications to physics, biology, chemistry and engineering*, Westview Press, 2000.
- [11] C. Diks, *Nonlinear time series analysis: methods and applications*, *Nonlinear Time Series and Chaos*, World Scientific, 1999.
- [12] H. Kantz, T. Schreiber, *Nonlinear time series analysis*, Cambridge University Press, 2004.
- [13] T.-Y. Li, J. A. Yorke, Period three implies chaos, *The American Mathematical Monthly* 82 (10) (1975) 985–992.
- [14] D. Kaplan, L. Glass, *Understanding nonlinear dynamics*, Springer Science & Business Media, 2012.
- [15] M. Small, *Applied nonlinear time series analysis: applications in physics, physiology and finance*, World Scientific, 2005.
- [16] C. Diks, Estimating invariants of noisy attractors, *Physical Review E* 53 (5) (1996) R4263–R4266.

Bibliografía

- [17] K.-S. Chan, H. Tong, *Chaos: a statistical perspective*, Springer Science & Business Media, 2013.
- [18] R. Mañé, On the dimension of the compact invariant sets of certain nonlinear maps, en: D. Rand, L.-S. Young (Eds.), *Dynamical Systems and Turbulence, Warwick 1980: Proceedings of a Symposium Held at the University of Warwick 1979/80*, Springer Berlin Heidelberg, Berlin, Heidelberg, 1981, pp. 230–242.
- [19] N. H. Packard, J. P. Crutchfield, J. D. Farmer, R. S. Shaw, Geometry from a time series, *Physical Review Letters* 45 (9) (1980) 712–716.
- [20] H. Nagashima, Y. Baba, *Introduction to chaos: physics and mathematics of chaotic phenomena*, 1st Edition, Taylor & Francis, 1998.
- [21] F. Takens, Detecting strange attractors in turbulence, en: D. Rand, L.-S. Young (Eds.), *Dynamical Systems and Turbulence, Warwick 1980: Proceedings of a Symposium Held at the University of Warwick 1979/80*, Springer Berlin Heidelberg, Berlin, Heidelberg, 1981, pp. 366–381.
- [22] U. Parlitz, Nonlinear time-series analysis, en: *Nonlinear Modeling*, Vol. 7, Springer US, Boston, MA, 1998, pp. 209–239.
- [23] T. Sauer, J. A. Yorke, M. Casdagli, Embedology, *Journal of Statistical Physics* 65 (3) (1991) 579–616.
- [24] T. Sauer, J. A. Yorke, How many delay coordinates do you need?, *International Journal of Bifurcation and Chaos* 03 (03) (1993) 737–744.
- [25] J. D. Farmer, Chaotic attractors of an infinite-dimensional dynamical system, *Physica D: Nonlinear Phenomena* 4 (3) (1982) 366 – 393.
- [26] F. Kaffashi, R. Foglyano, C. G. Wilson, K. A. Loparo, The effect of time delay on approximate and sample entropy calculations, *Physica D: Nonlinear Phenomena* 237 (23) (2008) 3069–3074.
- [27] M. B. Kennel, R. Brown, H. D. I. Abarbanel, Determining embedding dimension for phase-space reconstruction using a geometrical construction, *Physical Review A* 45 (6) (1992) 3403–3411.
- [28] A. Krakovská, K. Mezeiová, H. Budáčová, Use of false nearest neighbours for selecting variables and embedding parameters for state space reconstruction, *Journal of Complex Systems* 2015.

- [29] L. Cao, Practical method for determining the minimum embedding dimension of a scalar time series, *Physica D: Nonlinear Phenomena* 110 (1–2) (1997) 43–50.
- [30] O. Faust, M. G. Bairy, Nonlinear analysis of physiological signals: a review, *Journal of Mechanics in Medicine and Biology* 12 (04) (2012) 1240015.
- [31] M. Stéphane, *A wavelet tour of signal processing*, third edition Edition, Academic Press, 2009.
- [32] D. E. Vaillancourt, K. M. Newell, Changing complexity in human behavior and physiology through aging and disease, *Neurobiology of Aging* 23 (1) (2002) 1–11.
- [33] S. M. Pincus, Assessing serial irregularity and its implications for health, *Annals of the New York Academy of Sciences* 954 (1) (2001) 245–267.
- [34] G. Vaziri, F. Almasganj, R. Behroozmand, Pathological assessment of patients' speech signals using nonlinear dynamical analysis, *Computers in Biology and Medicine* 40 (1) (2010) 54–63.
- [35] K. Lehnertz, Epilepsy and nonlinear dynamics, *Journal of Biological Physics* 34 (3-4) (2008) 253–266.
- [36] M. M. Añino, M. E. Torres, G. Schlotthauer, Slight parameter changes detection in biological models: a multiresolution approach, *Physica A: Statistical Mechanics and its Applications* 324 (3-4) (2003) 645–664.
- [37] A. Fernández, P. Zuluaga, D. Abásolo, C. Gómez, A. Serra, M. A. Méndez, R. Hornero, Brain oscillatory complexity across the life span, *Clinical Neurophysiology* 123 (11) (2012) 2154–2162.
- [38] A. Voss, S. Schulz, R. Schroeder, M. Baumert, P. Caminal, Methods derived from nonlinear dynamics for analysing heart rate variability, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 367 (1887) (2009) 277–296.
- [39] V. R. Neves, A. C. M. Takahashi, M. D. B. do Santos-Hiss, A. M. Kiviniemi, M. P. Tulppo, S. C. G. de Moura, M. Karsten, A. Borghi-Silva, A. Porta, N. Montano, A. M. Catai, Linear and nonlinear analysis of heart rate variability in coronary disease, *Clinical Autonomic Research* 22 (4) (2012) 175–183.
- [40] L. Sörnmo, P. Laguna, *Bioelectrical signal processing in cardiac and neurological applications*, Academic Press, 2005.

Bibliografía

- [41] D. U. Silverthorn, W. C. Ober, C. W. Garrison, A. C. Silverthorn, B. R. Johnson, Human physiology: an integrated approach, Pearson/Benjamin Cummings, 2009.
- [42] Posit Science, BrainHQ (2016).
URL <http://www.brainhq.com/brain-resources/image-gallery/brain-anatomy-images>
- [43] A. C. K. Guyton, D. Silvia, y col., Anatomía y fisiología del sistema nervioso: neurociencia básica, Editorial médica panamericana, 1994.
- [44] R. M. Rangayyan, Biomedical signal analysis, 2nd Edition, IEEE Press Series in Biomedical Engineering, Wiley, 2015.
- [45] B. Hosseinifard, M. H. Moradi, R. Rostami, Classifying depression patients and normal subjects using machine learning techniques and nonlinear features from EEG signal, Computer Methods and Programs in Biomedicine 109 (3) (2013) 339–345.
- [46] F. Shayegh, S. Sadri, R. Amirfattahi, K. Ansari-Asl, A model-based method for computation of correlation dimension, Lyapunov exponents and synchronization from depth-EEG signals, Computer methods and programs in biomedicine 113 (1) (2014) 323–337.
- [47] N. Kannathal, J. Chee, K. Er, K. Lim, O. H. Tat, Chaotic analysis of epileptic EEG signals, en: The 15th International Conference on Biomedical Engineering: ICBME 2013, 4th to 7th December 2013, Singapore, Springer International Publishing, 2014, pp. 652–654.
- [48] S. A. Akar, S. Kara, F. Latifolu, V. Bilgiç, Analysis of the complexity measures in the EEG of schizophrenia patients, International Journal of Neural Systems (2015) 13.
- [49] J. A. S. Kelso, A. Fuchs, Self-organizing dynamics of the human brain: Critical instabilities and šil'nikov chaos, Chaos: An Interdisciplinary Journal of Non-linear Science 5 (1) (1995) 64–69.
- [50] S. Pravin Kumar, N. Sriraam, P. Benakop, B. Jinaga, Entropies based detection of epileptic seizures with artificial neural network classifiers, Expert Systems with Applications 37 (4) (2010) 3284–3291.
- [51] Q. Yuan, W. Zhou, S. Li, D. Cai, Epileptic EEG classification based on extreme learning machine and nonlinear features, Epilepsy Research 96 (1-2) (2011) 29–38.

- [52] U. R. Acharya, F. Molinari, S. V. Sree, S. Chattopadhyay, K.-H. Ng, J. S. Suri, Automated diagnosis of epileptic EEG using entropies, *Biomedical Signal Processing and Control* 7 (4) (2012) 401–408.
- [53] Y. Song, J. Crowcroft, J. Zhang, Automatic epileptic seizure detection in EEGs based on optimized sample entropy and extreme learning machine, *Journal of Neuroscience Methods* 210 (2) (2012) 132 – 146.
- [54] U. R. Acharya, S. V. Sree, G. Swapna, R. J. Martis, J. S. Suri, Automated EEG analysis of epilepsy: a review, *Knowledge-Based Systems* 45 (2013) 147–165.
- [55] S. M. Akareddy, P. Kulkarni, EEG signal classification for epilepsy seizure detection using improved approximate entropy, *International Journal of Public Health Science (IJPHS)* 2 (1) (2013) 23–32.
- [56] Z. Zhang, Z. Chen, Y. Zhou, S. Du, Y. Zhang, T. Mei, X. Tian, Construction of rules for seizure prediction based on approximate entropy, *Clinical Neurophysiology* 125 (10) (2014) 1959–1966.
- [57] H. L. Rufiner, *Análisis y modelado digital de la voz: técnicas recientes y aplicaciones*, Ediciones UNL, 2009.
- [58] P. Lieberman, S. E. Blumstein, *Speech physiology, speech perception, and acoustic phonetics*, Cambridge University Press, 1988.
- [59] F. Le Huche, A. Allali, *La voz: anatomía y fisiología, patología-terapéutica*, Vol. 3, Elsevier España, 2004.
- [60] M. Latarjet, A. R. Liard, *Anatomía humana*, Vol. 2, Ed. Médica Panamericana, 2004.
- [61] L. Lehto, y col., *Occupational voice: studying voice production and preventing voice problems with special emphasis on call-centre employees*, Helsinki University of Technology, 2007.
- [62] M. Eye, E. Infirmery, *Voice disorders database*, Version. 1.03 (CD-rom), Kay Elemetrics Corp., Lincoln Park, NJ.
- [63] S. Ayers, A. Baum, C. McManus, S. Newman, K. Wallston, J. Weinman, R. West, *Cambridge handbook of psychology, health and medicine*, Cambridge University Press, 2007.
- [64] A. E. Aronson, D. Bless, *Clinical voice disorders*, Thieme, 2011.

Bibliografía

- [65] J. C. Stemple, L. E. Glaze, B. K. Gerdeman, *Clinical voice pathology: theory and management*, Cengage Learning, 2000.
- [66] I. R. Titze, *Principles of voice production*, National Center for Voice and Speech, 2000.
- [67] R. J. Baken, R. F. Orlikoff, *Clinical measurement of speech and voice*, Cengage Learning, 2000.
- [68] J. Jiang, Y. Zhang, Nonlinear dynamic analysis of speech from pathological subjects, *Electronics Letters* 38 (6) (2002) 294–295.
- [69] N. Sáenz-Lechón, J. I. Godino-Llorente, V. Osma-Ruiz, P. Gómez-Vilda, Methodological issues in the development of automatic systems for voice pathology detection, *Biomedical Signal Processing and Control* 1 (2) (2006) 120–128.
- [70] J. D. Arias-Londoño, J. I. Godino-Llorente, N. Sáenz-Lechón, V. Osma-Ruiz, G. Castellanos-Domínguez, Automatic detection of pathological voices using complexity measures, noise parameters, and mel-cepstral coefficients, *IEEE Transactions on Biomedical Engineering* 58 (2) (2011) 370–379.
- [71] J. D. Arias-Londoño, J. I. Godino-Llorente, G. Castellanos-Domínguez, N. Sáenz-Lechón, V. Osma-Ruiz, Complexity analysis of pathological voices by means of hidden Markov entropy measurements, en: *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2009, pp. 2248–2251.
- [72] P. Billingsley, *Probability and measure*, John Wiley & Sons, 2008.
- [73] A. H. Jazwinski, *Stochastic processes and filtering theory*, Courier Corporation, 2007.
- [74] X. Chen, I. Solomon, K. Chon, Comparison of the use of approximate entropy and sample entropy: applications to neural respiratory signal, en: *27th Annual International Conference of the Engineering in Medicine and Biology Society*, 2005, pp. 4212–4215.
- [75] D. Ruelle, Small random perturbations of dynamical systems and the definition of attractors, *Communications in Mathematical Physics* 82 (1) (1981) 137–151.
- [76] J. P. Eckmann, D. Ruelle, Ergodic theory of chaos and strange attractors, *Reviews of Modern Physics* 57 (3) (1985) 617–656.

- [77] A. Kolmogorov, New metric invariant of transitive dynamical systems and endomorphisms of Lebesgue spaces, *Doklady of Russian Academy of Sciences* 119 (N5) (1958) 861–864.
- [78] T. D. Austin, *Entropy and sinai's theorem* (2006).
- [79] P. Grassberger, I. Procaccia, Dimensions and entropies of strange attractors from a fluctuating dynamics approach, *Physica D: Nonlinear Phenomena* 13 (1984) 34–54.
- [80] P. Grassberger, I. Procaccia, Estimation of the Kolmogorov entropy from a chaotic signal, *Physical review A* 28 (4) (1983) 2591–2593.
- [81] P. Grassberger, I. Procaccia, Measuring the strangeness of strange attractors, *Physica D: Nonlinear Phenomena* 9 (1983) 189–208.
- [82] P. Grassberger, I. Procaccia, Characterization of strange attractors, *Physical Review Letters* 50 (1983) 346–349.
- [83] J.-M. Ghez, E. Orlandini, M.-C. Tesi, S. Vaienti, Dynamical integral transform on fractal sets and the computation of entropy, *Physica D: Nonlinear Phenomena* 63 (1993) 282–298.
- [84] T. Schreiber, Determination of the noise level of chaotic time series, *Physical Review E* 48 (1) (1993) R13–R16.
- [85] T. Schreiber, Influence of gaussian noise on the correlation exponent, *Physical Review E* 56 (1) (1997) 274–277.
- [86] T. Schreiber, H. Kantz, Observing and predicting chaotic signals: is 2% noise too much?, en: Y. A. Kravtsov, J. B. Kadtko (Eds.), *Predictability of complex dynamical systems*, Springer, 1996, pp. 43–65.
- [87] D. Yu, M. Small, R. G. Harrison, C. Diks, Efficient implementation of the Gaussian kernel algorithm in estimating invariants and noise level from noisy time series data, *Physical Review E* 61 (4) (2000) 3750–3756.
- [88] G. Nolte, A. Ziehe, K.-R. Müller, Noise robust estimates of correlation dimension and K2 entropy, *Physical Review E* 64 (2001) 1–10.
- [89] D. Kugiumtzis, Correction of the correlation dimension for noisy time series, *International Journal of Bifurcation and Chaos* 07 (06) (1997) 1283–1294.

Bibliografía

- [90] J. Schouten, F. Takens, van den Bleek CM, Estimation of the dimension of a noisy attractor, *Physical review. E, Statistical physics, plasmas, fluids, and related interdisciplinary topics* 50 (3) (1994) 1851–1861.
- [91] a. W. Jayawardena, P. Xu, W. K. Li, A method of estimating the noise level in a chaotic time series, *Chaos* 18.
- [92] a. W. Jayawardena, P. Xu, W. K. Li, Modified correlation entropy estimation for a noisy chaotic time series, *Chaos (Woodbury, N.Y.)* 20 (2) (2010) 023104.
- [93] M. Ding, C. Grebogi, E. Ott, T. Sauer, J. A. Yorke, Estimating correlation dimension from a chaotic time series: when does plateau onset occur?, *Physica D: Nonlinear Phenomena* 69 (3) (1993) 404–424.
- [94] S. M. Pincus, Approximate entropy as a measure of system complexity, *Proceedings of the National Academy of Sciences* 88 (6) (1991) 2297–2301.
- [95] J. S. Richman, J. R. Moorman, Physiological time-series analysis using approximate entropy and sample entropy, *American Journal of Physiology-Heart and Circulatory Physiology* 278 (6) (2000) H2039–H2049.
- [96] S. M. Pincus, A. L. Goldberger, Physiological time-series analysis: what does regularity quantify?, *American Journal of Physiology-Heart and Circulatory Physiology* 266 (4) (1994) H1643–H1656.
- [97] S. M. Pincus, Approximate entropy: a complexity measure for biological time series data, en: *Proceedings of the IEEE Seventeenth Annual Northeast Bioengineering Conference, 1991*, pp. 35–36.
- [98] L.-S. Xu, K.-Q. Wang, L. Wang, Gaussian kernel approximate entropy algorithm for analyzing irregularity of time-series, en: *Proceedings of 2005 International Conference on Machine Learning and Cybernetics, 2005*, Vol. 9, 2005, pp. 18–21.
- [99] D. T. Kaplan, M. I. Furman, S. M. Pincus, S. M. Ryan, L. A. Lipsitz, A. L. Goldberger, Aging and the complexity of cardiovascular dynamics, *Biophysical Journal* 59 (4) (1991) 945–949.
- [100] C. Liu, C. Liu, P. Shao, L. Li, X. Sun, X. Wang, F. Liu, Comparison of different threshold values r for approximate entropy: application to investigate the heart rate variability between heart failure and healthy control groups, *Physiological Measurement* 32 (2) (2011) 167.

- [101] P. Van Leeuwen, D. Cysarz, F. Edelhäuser, D. Grönemeyer, Heart rate variability in the individual fetus, *Autonomic Neuroscience* (2013) 1–5.
- [102] L. a. Manilo, S. S. Volkova, Recognition of the deep anesthesia stage from parameters of the approximated entropy of EEG signal, *Pattern Recognition and Image Analysis* 23 (1) (2013) 92–97.
- [103] F. Liao, J. Wang, P. He, Multi-resolution entropy analysis of gait symmetry in neurological degenerative diseases and amyotrophic lateral sclerosis, *Medical Engineering & Physics* 30 (3) (2008) 299–310.
- [104] P. Zarjam, J. Epps, N. Lovell, F. Chen, Characterization of memory load in an arithmetic task using non-linear analysis of EEG signals, en: *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2012, pp. 3519–3522.
- [105] Y. Song, J. Crowcroft, J. Zhang, Epileptic EEG signal analysis and identification based on nonlinear features, en: *IEEE International Conference on Bioinformatics and Biomedicine*, IEEE, 2012, pp. 1–6.
- [106] C. C. Jouny, G. K. Bergey, Characterization of early partial seizure onset: frequency, complexity and entropy, *Clinical Neurophysiology* 123 (4) (2012) 658–669.
- [107] U. R. Acharya, E. C.-P. Chua, O. Faust, T.-C. Lim, L. F. B. Lim, Automated detection of sleep apnea from electrocardiogram signals using nonlinear parameters, *Physiological Measurement* 32 (3) (2011) 287.
- [108] K. Chon, C. Scully, S. Lu, Approximate entropy for all signals, *IEEE Engineering in Medicine and Biology Magazine* 28 (6) (2009) 18–23.
- [109] A. Wolf, J. B. Swift, H. L. Swinney, J. A. Vastano, Determining Lyapunov exponents from a time series, *Physica D: Nonlinear Phenomena* 16 (3) (1985) 285–317.
- [110] F. Aletti, M. Ferrario, T. Almas de Jesus, R. Stirbulov, A. Borghi Silva, S. Cerutti, L. Malosa Sampaio, Heart rate variability in children with cyanotic and acyanotic congenital heart disease: analysis by spectral and non linear indices, en: *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2012, pp. 4189–4192.
- [111] S. M. Pincus, D. L. Keefe, Quantification of hormone pulsatility via an approximate entropy algorithm, *American Journal of Physiology - Endocrinology And Metabolism* 262 (5) (1992) E741–E754.

Bibliografía

- [112] D. Sapochnikov, M. H. Luria, M. S. Gotsman, Detection of regularities in heart rate variations by linear and non-linear analysis: power spectrum versus approximate entropy, *Computer Methods and Programs in Biomedicine* 48 (3) (1995) 201–209.
- [113] S. M. Pincus, E. F. Gevers, I. C. Robinson, G. v. d. Berg, F. Roelfsema, M. L. Hartman, J. D. Veldhuis, Females secrete growth hormone with more process irregularity than males in both humans and rats, *American Journal of Physiology - Endocrinology And Metabolism* 270 (1) (1996) E107–E115.
- [114] S. M. Pincus, T. Mulligan, A. Iranmanesh, S. Gheorghiu, M. Godschalk, J. D. Veldhuis, Older males secrete luteinizing hormone and testosterone more irregularly, and jointly more asynchronously, than younger males, *Proceedings of the National Academy of Sciences* 93 (24) (1996) 14100–14105.
- [115] S. Lu, X. Chen, J. K. Kanters, I. C. Solomon, K. H. Chon, Automatic selection of the threshold value r for approximate entropy, *IEEE Transactions on Biomedical Engineering* 55 (8) (2008) 1966–1972.
- [116] A. Molina-Picó, D. Cuesta-Frau, P. Miró-Martínez, S. Oltra-Crespo, M. Aboy, Influence of QRS complex detection errors on entropy algorithms. application to heart rate variability discrimination, *Computer Methods and Programs in Biomedicine* 110 (1) (2013) 2–11.
- [117] S. M. Pincus, I. M. Gladstone, R. A. Ehrenkranz, A regularity statistic for medical data analysis, *Journal of Clinical Monitoring* 7 (4) (1991) 335–345.
- [118] P. Castiglioni, M. Di Rienzo, How the threshold “ r ” influences approximate entropy analysis of heart-rate variability, en: *Computers in Cardiology, 2008, 2008*, pp. 561–564.
- [119] A. Boskovic, T. Loncar-Turukalo, O. Sarenac, N. Japundzic-Zigon, D. Bajic, Unbiased entropy estimates in stress: a parameter study, *Computers in Biology and Medicine* 42 (6) (2012) 667–679.
- [120] M. R. Guevara, L. Glass, M. C. Mackey, A. Shrier, Chaos in neurobiology, *IEEE Transactions on Systems, Man, and Cybernetics* 13 (5) (1983) 790–798.
- [121] M. C. Mackey, L. Glass, Oscillation and chaos in physiological control systems, *Science* 197 (4300) (1977) 287–289.
- [122] R. Friedrich, C. Uhl, *Evolution of dynamical structures in complex systems*, Springer Berlin–Heidelberg, NY, 1992.

- [123] V. Vapnik, *The nature of statistical learning theory*, 2nd Edition, Springer, 2000.
- [124] C. M. Bishop, *Pattern recognition and machine learning*, Springer, 2007.
- [125] D. C. Montgomery, G. C. Runger, *Applied statistics and probability for engineers*, John Wiley & Sons, 2010.
- [126] M. E. Torres, L. G. Gamero, H. L. Rufiner, C. Martinez, D. Milone, G. Schlott-hauer, Study of complexity in normal and pathological speech signals, en: *Proceedings of the 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Vol. 3, 2003, pp. 2339–2342.
- [127] P. R. Scalassara, M. E. Dajer, C. D. Maciel, R. C. Guido, J. C. Pereira, Relative entropy measures applied to healthy and pathological voice characterization, *Applied Mathematics and Computation* 207 (1) (2009) 95–108.
- [128] J. Kennedy, R. Eberhart, Particle swarm optimization, en: *Proceedings of IEEE International Conference on Neural Networks*, Vol. 4, 1995, pp. 1942–1948.
- [129] Y.-L. Zheng, L.-H. Ma, L.-Y. Zhang, J.-X. Qian, On the convergence analysis and parameter selection in particle swarm optimization, en: *Proceedings of the International Conference on Machine Learning and Cybernetics*, Vol. 3, 2003, pp. 1802–1807.
- [130] V. Parsa, D. G. Jamieson, Identification of pathological voices using glottal noise measures, *Journal of Speech, Language & Hearing Research* 43 (2) (2000) 469–485.
- [131] W. R. Bennett, Spectra of quantized signals, *Bell System Technical Journal* 27 (3) (1948) 446–472.
- [132] L. Gammaitoni, Stochastic resonance and the dithering effect in threshold physical systems, *Physical Review E* 52 (5) (1995) 4691–4698.
- [133] L. Gammaitoni, P. Hänggi, P. Jung, F. Marchesoni, Stochastic resonance, *Reviews of modern physics* 70 (1) (1998) 223.
- [134] M. D. McDonnell, D. Abbott, What is stochastic resonance? definitions, misconceptions, debates, and its relevance to biology, *PLoS computational biology* 5 (5) (2009) e1000348.
- [135] D. Rousseau, F. Chapeau-Blondeau, Suprathreshold stochastic resonance and signal-to-noise ratio improvement in arrays of comparators, *Physics Letters A* 321 (5-6) (2004) 280–290.

Bibliografía

- [136] M. D. McDonnell, N. G. Stocks, C. E. M. Pearce, D. Abbott, *Stochastic Resonance: from suprathreshold stochastic resonance to stochastic signal quantization*, Cambridge University Press, 2008.
- [137] F. Chapeau-Blondeau, S. Blanchard, D. Rousseau, Fisher information and noise-aided power estimation from one-bit quantizers, *Digital Signal Processing* 18 (3) (2008) 434–443.
- [138] D. Rousseau, F. Chapeau-Blondeau, Noise-improved Bayesian estimation with arrays of one-bit quantizers, *IEEE Transactions on Instrumentation and Measurement* 56 (6) (2007) 2658–2662.
- [139] A. Papoulis, S. U. Pillai, *Probability, random variables, and stochastic processes*, Tata McGraw-Hill Education, 2002.
- [140] H. Oltmans, P. Verheijen, Influence of noise on power-law scaling functions and an algorithm for dimension estimations, *Physical Review E* 56 (1) (1997) 1160–1170.
- [141] G. Çoban, A. H. Büyüklü, A. Das, A linearization based non-iterative approach to measure the Gaussian noise level for chaotic time series, *Chaos, Solitons and Fractals* 45 (3) (2012) 266–278.
- [142] I. S. Gradshteyn, I. M. Ryzhik, *Table of integrals, series, and products*, 5th Edition, Academic Press, 1994.
- [143] M. Abramowitz, I. Stegun, *Handbook of mathematical functions: with formulas, graphs, and mathematical tables*, Applied mathematics series, Dover Publications, 1964.
- [144] J. W. Luo, J. Bai, J. H. Shao, Application of the wavelet transforms on axial strain calculation in ultrasound elastography, *Progress in Natural Science* 16 (9) (2006) 942–947.

A mi familia, amigos y compañeros...