

UNIVERSIDAD CENTRAL DE VENEZUELA
FACULTAD DE CIENCIAS
ESCUELA DE FÍSICA



**ESTIMACIONES COMPUTACIONALES DE TRANSFERENCIA DE
INFORMACIÓN EN SISTEMAS COMPLEJOS**

Trabajo Especial de Grado presentado por
Sarah Lares
ante la Facultad de Ciencias de la
Ilustre Universidad Central de Venezuela
como requisito parcial para optar al título
de: **Licenciado en Física**
Con la tutoría de: Prof. Rosa Mujica

Julio-2019
Caracas-Venezuela

Escuela de Física



VEREDICTO

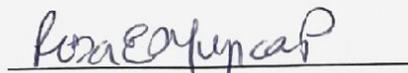
Quienes suscriben, miembros del Jurado designado por el Consejo de la Escuela de Física de la Facultad de Ciencias de la Universidad Central de Venezuela, para examinar el Trabajo Especial de Grado presentado por **Sarah Lares Gallardo**, Cédula de Identidad **V-22667474**, bajo el título “**Estimaciones computacionales de transferencia de información en sistemas complejos**”, a los fines de cumplir con el requisito legal para optar al grado de **Licenciada en Física**, dejan constancia de lo siguiente:

1. Leído como fue dicho trabajo por cada uno de los miembros del Jurado, éste fijó el día 31 de julio de 2019, a Las 9:30 am, para que el autor lo defendiera en forma pública, lo que éste hizo en la Sala de Seminarios Guillermo Ruggeri de la Escuela de Física, mediante un resumen oral de su contenido, luego de lo cual respondió satisfactoriamente a las preguntas que le fueron formuladas por el jurado; todo ello conforme a los artículos 20, 21, 22, 25, 26 y 28 de la Normativa de Trabajo Especial de Grado de la Licenciatura en Física de la Facultad de Ciencias de la UCV vigente.
2. Finalizada la defensa pública del trabajo, el jurado decidió declararlo aprobado por considerar que se ajusta a lo dispuesto y exigido en la Normativa de Trabajo Especial de Grado de la Licenciatura en Física de la Facultad de Ciencias de la UCV vigente en sus artículos 1, 5 y 6.

Se levanta la presente acta a los 31 días del mes de julio de 2019, dejándose también constancia de que, conforme a la normativa jurídica vigente, actuó como coordinadora del jurado la tutora del Trabajo Especial de Grado Profa. Rosa Elena Mujica Paredes.

Firma del jurado evaluador


Prof. Esteban Segundo Álvarez
UCV


Profa. Rosa Elena Mujica
Paredes
UCV


Lic. Gustavo Leandro Ramírez
UCV

UNIVERSIDAD CENTRAL DE VENEZUELA
FACULTAD DE CIENCIAS
ESCUELA DE FÍSICA



**ESTIMACIONES COMPUTACIONALES DE TRANSFERENCIA DE
INFORMACIÓN EN SISTEMAS COMPLEJOS**

Trabajo Especial de Grado presentado por
Sarah Lares
ante la Facultad de Ciencias de la
Ilustre Universidad Central de Venezuela
como requisito parcial para optar al título
de: **Licenciado en Física**
Con la tutoría de: Prof. Rosa Mujica

Julio-2019
Caracas-Venezuela

*La cosa más sencilla
se vuelve inmediatamente complicada
e incomprensible al hablar de ella.*

Hermann Hesse

Agradecimientos

Fue la suma del trabajo incansable y la perseverancia lo que llevaron a la culminación de este trabajo del que siento tanto orgullo, por lo que primeramente debo agradecer a mi tutora, la profesora Rosa Mujica, quien me empujó hacia la excelencia a lo largo del desarrollo de este proyecto, igualmente a los profesores Esteban Álvarez, Pedro García y Fermín Dalmagro, quienes siempre estuvieron dispuestos a resolver mis dudas, y me dieron ánimos para seguir adelante, a la Escuela de Física, a la Facultad de Ciencias y a la Ilustre Universidad Central de Venezuela que me han permitido crecer en conocimientos y en calidad humana.

Quiero agradecer también a mis amigos con los que he vivido tantas buenas experiencias, a mi familia por su apoyo y al universo que; con todos sus fenómenos, entropía y caos; me ha llevado a este hermoso punto de mi vida y a la culminación de una meta.

RESUMEN

La Transferencia de Información es importante en el estudio de sistemas complejos y de interés práctico en muchas áreas, como física, medicina, economía, entre otros. Esta permite el estudio de la causalidad en diversos tipos de sistemas, y es importante por ejemplo, en toma de decisiones sobre cual sistema debe ser controlado para poder regular a otro. Para calcular el intercambio de información, se utilizará la transferencia de entropía de Schreiber, que es una medida efectiva de causalidad y además no hace ninguna suposición sobre la naturaleza del acople entre los sistemas. En este trabajo se exploran dos métodos diferentes para obtener las probabilidades necesarias en la estimación de la transferencia de entropía entre sistemas, y se propone un método novedoso y eficiente basado en el modelado de series temporales a través del algoritmo de los vecinos más cercanos, con el cual se obtienen excelentes resultados al ser empleado en datos generados sintéticamente y en sistemas reales. Se estudiarán tres sistemas reales diferentes, pero el principal aporte de este trabajo es aplicar la metodología propuesta para calcular la transferencia de entropía y estudiar el acople cardio respiratorio, donde es posible inferir cual sistema envía más información, lo cual puede ser de gran ayuda en el área de la medicina.

Palabras claves:

Información Mutua, Transferencia de Información o Entropía, Sistemas de Mapas Acoplados, Sistemas Complejos.

Índice general

Lista de figuras	11
Lista de tablas	13
1. Conceptos Básicos	17
1.1. Sistemas dinámicos	17
1.2. Clasificaciones generales de los sistemas dinámicos	19
1.2.1. De acuerdo a la representación de \mathbf{x} , \mathbf{s} y t	19
1.2.2. De acuerdo a la función de evolución f	21
1.3. Caos Determinista	21
1.3.1. Puntos fijos y estabilidad local	23
1.3.2. Sensibilidad a las condiciones iniciales y exponente de Lyapunov	24
1.4. Probabilidades	25
1.4.1. Probabilidad Marginal, Condicional y Conjunta	26
2. Información y Entropía	29
2.0.1. Información Según Shannon	29
2.0.2. Entropía de Shannon	30
2.0.3. Información Mutua	32
2.1. Causalidad de Wiener-Granger	33
2.1.1. Definición	33
2.2. Transferencia de Entropía de Schreiber	34
2.3. Mapas	36
2.3.1. Mapa Logístico	36
2.3.2. Mapa Triangular	38
2.3.3. Mapa de Hénon	38
2.4. Sistemas Complejos	39
2.5. Sistemas de Mapas Acoplados	40

3. Discretización del Espacio de Estados	43
3.1. Método de Discretización 1:	
Reticulado (MD1)	43
3.2. Método de Discretización 2:	
Representación Simbólica (MD2)	44
3.3. Experimentos Realizados	45
3.3.1. Sistemas de Mapas Acoplados	45
3.3.2. Mapa de Henón	51
3.4. Número de Operaciones	51
4. Modelado de series temporales	53
4.1. Densidad de Probabilidad Condicional	53
4.2. Densidad de Probabilidad Marginal	55
4.3. Algoritmo de los Vecinos más Cercanos	56
4.4. Experimentos Realizados	57
4.4.1. Sistemas de Mapas Acoplados	57
4.4.2. Mapa de Hénon	60
4.4.3. Sistemas Reales	61
4.4.4. Sistema Cardio Respiratorio	65
4.5. Número de Operaciones	69
4.6. Comparación del Tiempo de Cómputo.	69
5. Conclusiones	71

Índice de figuras

1.1.	Intersección entre los conjuntos A y B	26
2.1.	Relación entre entropías e información mutua.	32
2.2.	Representación geométrica del Mapa Logístico con $r = 3,99$	37
2.3.	Diagrama de bifurcaciones del mapa Logístico.	37
2.4.	Iteración del Mapa triangular para $a = 0,5$	38
2.5.	Atractor de Hénon con $a = 1,4$ y $b = 0,3$	39
2.6.	f =Mapa Logístico, g =Mapa Triangular, $r = 3,99$, $a = 0,5$, $\mu = 0,2$ y $\epsilon = 0,4$. Acople 1 (izquierda) y Acople 2 (derecha).	41
3.1.	Transferencias de entropía para el acople 1 con $\mu = 0,4$ y $\epsilon = 0,2$ para $N = 2000$, y $N = 3000$, respectivamente.	46
3.2.	Transferencias de entropía para el acople 1 con $N = 10000$, $\mu = 0,4$ y $\epsilon = 0,2$	46
3.3.	Transferencias de entropía con $N = 3000$, $\mu = 0,2$ y $\epsilon = 0,4$ para el acople 1 y el acople 2 respectivamente.	47
3.4.	Transferencias de entropía con $N = 3000$, $\mu = 0,6$ y $\epsilon = 0,0$ para el acople 1 y el acople 2 respectivamente.	48
3.5.	Transferencias de entropía con $N = 3000$, $\mu = 0,0$ y $\epsilon = 0,6$ para el acople 1 y el acople 2 respectivamente.	48
3.6.	Transferencias de entropía para el acople 1 con $N = 3000$, $\mu = 0,6$ y $\epsilon = 0,0$ y $N = 3000$, $\mu = 0,0$ y $\epsilon = 0,6$ para dos mapas triangulares respectivamente.	49
3.7.	Barrido en los parámetros de acople μ y ϵ para el acople 1 y el MD1 con $M = 128$	50
3.8.	Barrido en los parámetros de acople μ y ϵ para el acople 1 con el MD2 y $m = 7$	50
4.1.	Distribución de los errores de predicción para $P(x_{i+1} x_i, y_i)$ (izquierda) y $P(x_i y_i)$ (derecha), ajustadas con una función densidad de probabilidad Gaussiana.	55

4.2. Barrido en los parámetros de acople para dos mapas triangulares con $a_1 = 0,63$ y $a_2 = 0,65$, con $k = 10$, acople 1 (izquierda), acople 2 (derecha).	58
4.3. Barrido en los parámetros de acople para dos mapas triangulares con $a_1 = 0,63$ y $a_2 = 0,65$, con $k = 12$, acople 1 y acople 2 respectivamente.	58
4.4. Desviación estándar para el promedio de la transferencias de entropía calculada 10 veces (4.2) con $\mu = 0,35$ para el acople 2.	59
4.5. Barrido en los parámetros de acople para f el mapa logístico con $r = 3,99$ y g el mapa triangular con $a = 0,5$, con $k = 10$, acople 1.	60
4.6. Barrido en los parámetros de acople para f el mapa logístico con $r = 3,99$ y g el mapa triangular con $a = 0,65$, con $k = 10$, acople 1 y acople 2 respectivamente.	60
4.7. Circuito de Chua	62
4.8. Series temporales y atractor de Chua.	62
4.9. Diagrama para implementar la transferencia de entropía en el circuito de Chua.	63
4.10. Series temporales de la primera derivación del ECG y de la presión arterial, transferencias de entropía para entre ECG y presión arterial respectivamente.	64
4.11. Esquema del acople Cardio Respiratorio.	66
4.12. Series temporales electrocardiográfica (derivación I) y de frecuencia respiratoria de los datos sin submuestrear y submuestreados ($n = 100$), respectivamente.	67
4.13. Transferencia de entropía entre la primera derivación del ECG (I) y la frecuencia respiratoria (RES) para el estado basal y durante los 50 minutos de música clásica respectivamente.	68
4.14. Transferencia de entropía entre la primera derivación del ECG (I) y la frecuencia respiratoria (RES) después de los 50 minutos de música.	68

Índice de cuadros

3.1. Transferencia de Entropía obtenidas en el mapa de Hénon para MD1 y MD2.	51
4.1. Transferencia de Entropía por pares de series en el circuito de Chua	62
4.2. Transferencia de entropía total para cada elemento del circuito de Chua . .	63
4.3. Comparación del tiempo de cómputo real para el cálculo de la transferencia de entropía con los 2 métodos implementados.	70



Introducción

Generalmente, los sistemas reales están compuestos por elementos que interactúan entre sí a través de acoples y presentan comportamientos complejos. Este tipo de sistemas, llamados sistemas complejos, pueden encontrarse en áreas como la física, biología, economía, ciencias sociales, medicina, entre otros. El estudio de la detección de causalidad en sistemas complejos ha ganado gran importancia recientemente gracias a sus múltiples aplicaciones [1, 2, 3, 4, 5].

Por ejemplo una de las aplicaciones, es detectar entre el sistema respiratorio y el cardíaco cual de ellos comparte mas información, para saber sobre cual se debe ejercer control para regular al otro. De forma similar sucede con el sistema cardiovascular, que ha sido estudiado con la finalidad de conocer cuál subsistema controlar y así administrar la medicación pertinente.

Existen diversos métodos para detectar causalidad en series temporales, muchos de estos sólo funcionan para algunos tipos de sistemas [6]. Uno de los mas utilizados por su capacidad de detectar causalidad en diferentes tipos de sistemas y con eficacia es la transferencia de entropía de Schreiber [7]. Por lo cual fue utilizada en este trabajo para medir la dirección del acople en los sistemas bajo estudio.

Se utilizó la transferencia de entropía para estudiar dos tipos de sistemas diferentes, el primero será data generada sintéticamente, en el cual podamos controlar la dirección del acople y por último se empleó en sistemas reales como por ejemplo la frecuencia cardíaca y la respiración.

Para estimar la transferencia de entropía de Schreiber se implementaron 2 métodos diferentes para calcular las probabilidades necesarias para esta medida (ver sección 2.2). El primero consiste en hallar las probabilidades mediante la discretización del espacio de estados y posteriormente la observación directa de la serie temporal para así hallar la frecuencia relativa de cada uno de los valores discretos (a su vez se utilizaron dos estrategias diferentes para discretizar el espacio de estados). El segundo método implementado está basado en el modelado de series temporales a través de la estrategia de los vecinos más cercanos, este permite entre otras cosas disminuir el tiempo de cómputo en la estimación

de la transferencia de entropía, lo cual permite la aplicación en sistemas reales. Para finalizar, se compararon ambos métodos y se calculó el número de operaciones necesarias para la realización de cada uno de ellos y el tiempo de cómputo.

El método de discretización del espacio de estados fue aplicado sólo a series temporales sintéticas obtenidas de sistemas de mapas acoplados ya que no pudo implementarse a sistemas reales por su alto costo computacional. Por otra parte, el método del modelado pudo ser implementado en ambos tipos de series temporales.

Este trabajo está compuesto por 5 capítulos, en el primero se introducirán los conceptos básicos necesarios para el entendimiento del trabajo, en el segundo capítulo se explica el concepto de transferencia de entropía de Schreiber, los sistemas de mapas acoplados y los mapas que serán utilizados. En el capítulo 3 se discuten los métodos de discretización del espacio de estados utilizados y los experimentos realizados; en el cuarto capítulo, se explica el método del modelado con sus respectivos experimentos y por último en el capítulo 5 se exponen las conclusiones del trabajo.

Capítulo 1

Conceptos Básicos

1.1. Sistemas dinámicos

La palabra sistema puede tener muchas definiciones según el ámbito en el que se utilice. En física un sistema es una colección de elementos los cuales interactúan entre sí y que conforman un todo. Un sistema se nos manifiesta como un aspecto de la realidad dotado de cierta complejidad precisamente por estas partes en interacción.

Cuando los elementos del sistema producen patrones de cambio en el tiempo se les llama sistemas dinámicos. La característica del patrón es que en un tiempo t' está interrelacionada con las de otros tiempos t . Casi todos los fenómenos observados en el mundo científico tienen importantes aspectos dinámicos.

Entiéndase por sistema dinámico a la formulación matemática de un proceso determinista (es decir, el azar no está involucrado en el desarrollo de los futuros estados del sistema). El estado de un sistema dinámico descrito por n variables $x_i(t)$ con $i = 1, \dots, n$, en un instante t se puede representar por un vector definido en un espacio euclidiano n -dimensional (x_1, x_2, \dots, x_n) , donde cada dimensión representa una variable, denominado espacio de fase del sistema, tal que

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

En este tipo de sistemas los estados evolucionan de acuerdo a reglas determinadas. Las reglas especifican cómo cambia el estado del sistema a partir de un estado dado. Estas reglas pueden consistir en ecuaciones diferenciales, funciones iterativas, o en un algoritmo [8].

La teoría que los agrupa ofrece una manera unificada de representar y caracterizar la evolución de sistemas que aparecen en diversos contextos. Se pueden representar a través de la notación vectorial [9], que provee un eficiente marco de trabajo para el desarro-

llo computacional y teórico. Por su propia naturaleza, la notación vectorial suprime los detalles, pero permite su recuperación cuando es necesario.

Dicha representación viene dada en términos de la cuarteta:

$$\{\mathbf{s}, \mathbf{x}, t, f\} \quad (1.1)$$

donde

- \mathbf{s} es el espacio de fase y que representa el conjunto de todos los estados accesibles al sistema.
- \mathbf{x} es el espacio (conjunto de variables) que puede ser continuo o discreto dependiendo del sistema.
- t es el tiempo y según su aplicación puede ser tomado como continuo o discreto.
- f es una función o conjunto de funciones que mapea el espacio de estados en sí mismo. La evolución del estado $\mathbf{x}(t)$ en un subespacio $U \subseteq R^n$ en muchos sistemas se puede describir mediante ecuaciones diferenciales de la forma:

$$\frac{d\mathbf{x}}{dt} = f(\mathbf{x}), \quad (1.2)$$

donde

$$f(\mathbf{x}) = \{f_1(x), f_2(x), \dots, f_n(x)\}. \quad (1.3)$$

La Ec. 1.2 equivale al sistema de n ecuaciones diferenciales de primer orden

$$\begin{aligned} \dot{x}_1 &= f_1(x_1, x_2, \dots, x_n) \\ &\vdots \quad \quad \quad \vdots \\ \dot{x}_n &= f_n(x_1, x_2, \dots, x_n) \end{aligned}$$

y si el tiempo es discreto obtendremos una ecuación de diferencia y no una ecuación diferencial:

$$\begin{aligned} x_1(t+1) &= f_1(x_1(t), x_2(t), \dots, x_n(t)) \\ &\vdots \quad \quad \quad \vdots \\ x_n(t+1) &= f_n(x_1(t), x_2(t), \dots, x_n(t)) \end{aligned}$$

La solución del sistema de n ecuaciones de primer orden (ecuación 1.2) para $x(t) \in U \subseteq R^n$ requiere el conocimiento de n condiciones iniciales:
 $x(0) = (x_1(0), x_2(0), \dots, x_n(0))$.

La evolución del estado de un sistema en su espacio de fase es análogo al movimiento de una partícula en un espacio euclidiano n -dimensional, cuya posición instantánea es $x(t)$ y su velocidad está dada por $\frac{dx}{dt} = f(x(t))$, en el caso discreto $x(t+1) = f(x(t))$, podemos observar que la evolución del sistema vendrá dada por los estados anteriores.

1.2. Clasificaciones generales de los sistemas dinámicos

Los sistemas dinámicos pueden clasificarse según dos criterios, el primero es de acuerdo a la representación de los estados, el espacio y el tiempo, y el segundo es de acuerdo a la regla de evolución.

1.2.1. De acuerdo a la representación de \mathbf{x} , \mathbf{s} y t

Sistemas Totalmente Continuos

Un sistema totalmente continuo es donde un conjunto de ecuaciones diferenciales describen la evolución del sistema en tiempo continuo. La formulación de modelos basados en ecuaciones diferenciales se inicia con la Mecánica Newtoniana y actualmente sigue siendo la manera más utilizada en física para representar sistemas. Este término engloba una gran variedad de sistemas cuyo comportamiento puede ser descrito mediante ecuaciones diferenciales. En este caso, tanto \mathbf{s} como t son continuos. En este apartado se tienen:

- **Sistemas de Ecuaciones en Derivadas Parciales (EDP)**

Son aquellos en los cuales el espacio de fases \mathbf{s} es continuo y de dimensión infinita (debido a que es un espacio de funciones). \mathbf{x} y t también continuos. Demuestran la relación entre una función f desconocida, ciertas variables independientes y sus derivadas parciales asociadas. Generalmente, en conjunto con condiciones iniciales o de frontera. Su resolución provee herramientas para reconstruir sistemas reales con alto detalle, usualmente enlazando con otras EDPs previamente derivadas. Ejemplos incluyen la ecuación de onda, la ecuación de difusión y la ecuación de Schrödinger.

- **Sistemas de Ecuaciones Diferenciales Ordinarias (EDO)**

Poseen \mathbf{s} de dimensión finita y t continuo. Una EDO de orden k relaciona los valores de una función desconocida de una sola variable f , y sus derivadas hasta el orden

k . Representan sistemas localizados ya que el estado no depende de las variables espaciales. Ejemplos incluyen la ecuación de Bessel (de segundo orden, $k = 2$) y la ecuación de Ricatti (de primer orden, $k = 1$).

Sistemas Totalmente Discretos

■ Autómatas Celulares

Poseen \mathbf{s} , \mathbf{x} y t discretos. Son sistemas extendidos compuestos de un arreglo de L celdas dispuestas de forma regular. Cada una es actualizada de forma sincronizada por un conjunto de k reglas f_k que toman en cuenta tanto el estado de la propia celda como la configuración de su vecindad v , de la forma:

$$f_k : \mathbf{s} \otimes \mathbf{x}^v \rightarrow \mathbf{s}$$

Uno de los ejemplos mas famosos es el juego de la vida de Conway [10].

Sistemas Parcialmente Discretos

Representan el punto intermedio de los casos anteriores, algunos modelos con \mathbf{x} y t discretos, y \mathbf{s} continuo son:

■ Mapas

Los mapas, que se conocen también como ecuaciones diferenciales finitas, relaciones de recursión, mapas iterados o mapas dinámicos; son sistemas localizados cuya regla de evolución viene dada por una función determinista f , siendo la trayectoria del sistema:

$$x_{t+1} = f(x_t) \tag{1.4}$$

Entonces si aplicamos la regla de evolución f a un x_0 se obtendrá el siguiente resultado:

$$f^t(x_0) \equiv \underbrace{f \circ f \circ f \circ \dots \circ f}_{t}(x_0) = f(f(\dots(f(x_0)))) = x_t$$

La secuencia $x_0, x_1, x_2, \dots, x_t$ se llama trayectoria u órbita correspondiente a la condición inicial x_0 [11].

Por ejemplo el mapa logístico [12] es uno de los mapas unidimensionales mas utilizados, de este y de otros mapas se hablará mas extensamente en la sección 2.3.

Los mapas sirven como modelos en áreas como electrónica digital, economía, caos y también se usan como representaciones discretas de ecuaciones diferenciales (ver [13] y [14]).

- **Redes de Mapas Acoplados**

Se componen de una secuencia de L mapas acoplados a sus vecinos inmediatos mediados por un parámetro ϵ , el cual es el encargado de modular la intensidad del acople.

La trayectoria del sistema es dada por la iteración de una función vector valuada:

$$x_{t+1} = f(x_t)$$

Los RMAs son capaces de generar múltiples comportamientos dependiendo de los parámetros internos de la función f , como turbulencia, intermitencia, ondas viajeras, segregación de fases, formación de patrones y otros comportamientos colectivos emergentes que no pueden ser explicados a partir del comportamiento de uno de los elementos del sistema [15] [16].

1.2.2. De acuerdo a la función de evolución f

Si f es Lineal

Un sistema dinámico es lineal si se cumple

$$f(ax + by) = af(x) + bf(y) \quad (1.5)$$

es decir, es lineal si la función f que relaciona la tasa de incremento de las variables de estado con sus valores actuales cumple con el principio de superposición.

Si f es no Lineal

Si la Ec.1.5 falla en un sistema, éste se dice es no lineal. El hecho de ser no lineal hace que su análisis sea mucho más complejo (ya que no se puede simplificar el problema a instancias más sencillas). En la mayoría de las ocasiones no se podrá encontrar soluciones analíticas exactas a los problemas no lineales, por lo tanto la representación de la dinámica del sistema se auxilia mucho de técnicas geométricas de visualización y análisis [9].

1.3. Caos Determinista

A pesar de que tradicionalmente se ha prestado una mayor importancia al estudio de los sistemas lineales, conceptualmente más sencillos y matemáticamente más fáciles de resolver, una gran mayoría de los problemas de interés en Física involucran ecuaciones no lineales, cuyas soluciones exhiben una rica variedad de comportamientos dinámicos. Esto

incluye la posibilidad de movimiento caótico, incluso en sistemas de pocos grados de libertad, en los cuales las trayectorias se entremezclan entre sí de forma errática y turbulenta, de manera que resulta imposible toda predicción detallada para tiempos largos.

La característica más importante de la dinámica caótica es su extrema sensibilidad a pequeñas variaciones de las condiciones iniciales, de forma que dos trayectorias que empiezan muy próximas se separan de forma exponencial en el tiempo. Esto conlleva a limitaciones prácticas en la capacidad de predicción del comportamiento de sistemas caóticos.

Hay que destacar, que a pesar de su naturaleza impredecible, el caos es un fenómeno determinista, en el que la evolución del sistema viene dada en todo momento por un conjunto de reglas de evolución temporal que están bien definidas.

El caos se puede definir como el comportamiento aperiódico a tiempos largos que tiene lugar en un sistema dinámico (que como ya se mencionó anteriormente es determinista) que presenta sensibilidad extrema a las condiciones iniciales (es decir, el exponente de Lyapunov de la órbita es positivo) [11, 17].

Las cantidades conservadas constituyen primeras integrales del movimiento de un sistema; son funciones de las coordenadas y de sus velocidades (primera derivadas temporales). Un sistema con s grados de libertad es integrable si posee, al menos, s cantidades conservadas. Si el sistema tiene menos de s cantidades conservadas, se denomina no integrable. Un sistema para el cual existen más cantidades conservadas que grados de libertad, se llama superintegrable.

La integrabilidad puede ser considerada como un tipo de simetría presente en varios sistemas dinámicos, y que conduce a soluciones con comportamiento regular (periódico o estacionario) en el tiempo. Sin embargo, la existencia de integrabilidad no es lo más común; muchos sistemas dinámicos no son integrables. Un sistema no integrable puede exhibir comportamiento caótico para algún rango de valores de sus parámetros. La no integrabilidad es una condición necesaria, pero no suficiente, para la existencia de caos en un sistema [8].

Los sistemas con comportamiento caótico se caracterizan porque las ecuaciones que rigen su dinámica poseen funciones no lineales de las variables.

Las características que se han mencionado anteriormente sobre los sistemas caóticos ponen de manifiesto que el comportamiento caótico no se debe a la presencia de fuentes de ruido, ni a un elevado número de grados de libertad; la irregularidad del comportamiento caótico es exclusivamente resultado de la propia dinámica interna del sistema.

1.3.1. Puntos fijos y estabilidad local

Sea el mapa $x_{t+1} = f(x_t)$. Se dice que x^* es un punto fijo si

$$f(x^*) = x^* \quad (1.6)$$

Es decir, si la órbita cae en un punto fijo, seguirá en él durante las sucesivas iteraciones, puesto que, dado $x_t = x^*$, se tiene $x_{t+1} = f(x_t) = f(x^*) = x^*$ para todo t .

Para estudiar la estabilidad de x^* , consideremos una órbita próxima, $x_t = x^* + \eta_t$ y observemos si esta es atraída o repelida por x^* , esto es, si η_t crece o decrece al aumentar t .

De la definición de los mapas (Ec. 1.4) se tiene

$$x_{t+1} = x^* + \eta_{t+1} = f(x_t) = f(x^* + \eta_t) \quad (1.7)$$

$$= f(x^*) + f'(x^*)\eta_t + O(\eta_t^2) \quad (1.8)$$

tomando en cuenta la ecuación 1.6,

$$\eta_{t+1} = f'(x^*)\eta_t + O(\eta_t^2). \quad (1.9)$$

Supongamos que puede despreciarse el término $O(\eta_t^2)$. Obtenemos el mapa linealizado $\eta_{t+1} = f'(x^*)\eta_t$. Se define su valor propio o multiplicador como la pendiente en el punto fijo

$$\lambda \equiv f'(x^*). \quad (1.10)$$

Entonces, la estabilidad dependerá del valor de λ . En efecto, los primeros términos del mapa linealizado son:

$$\begin{aligned} \eta_1 &= \lambda\eta_0 \\ \eta_2 &= \lambda\eta_1 = \lambda^2\eta_0 \\ &\dots \\ \eta_t &= \lambda^t\eta_0 \end{aligned}$$

En consecuencia,

- Si $|\lambda| < 1$, $\eta_t \rightarrow 0$ para $t \rightarrow \infty$, el punto fijo es linealmente estable y se le dice que x^* es un punto atractor.

- Si $\lambda = 0$ se dice que el punto es superestable.
- Si $|\lambda| > 1$, el punto fijo es inestable.
- Si $|\lambda| = 1$ el punto fijo es crítico.

Si el parámetro de control del mapa varía, generalmente, también cambia el valor propio del punto fijo llegando para ciertos valores a perderse la condición de estabilidad $|\lambda| < 1$. Para este valor del parámetro de control la trayectoria deja de tender hacia el atractor x^* .

1.3.2. Sensibilidad a las condiciones iniciales y exponente de Lyapunov

Dos órbitas con puntos iniciales cercanos que se encuentren dentro de un intervalo periódico suelen mantenerse aproximadamente a la misma distancia. Si se encuentran en una zona caótica no hay puntos atractores, por lo tanto las órbitas de los dos puntos iniciales tienen evoluciones totalmente diferentes. A esto se le llama sensibilidad a las condiciones iniciales y está relacionado al valor absoluto de la derivada del mapa como veremos a continuación.

Haciendo la expansión de Taylor de una función suave cualquiera f , tenemos que:

$$f(x+h) = f(x) + f'(x)h + O(h), \quad (1.11)$$

es decir

$$f(x+h) - f(x) = f'(x)h + O(h). \quad (1.12)$$

Si dos puntos están en un principio a una distancia $|h|$ en la siguiente iteración estarán a una distancia $|f'(x)h|$ si se desprecia $O(h)$, esto es, si $|f'(x)| > 1$ entonces la distancia que los separe será mayor que $|h|$. Por lo tanto el que un punto $x+h$ se acerque o se aleje de x depende del valor de la derivada $f'(x)$.

Sea una condición inicial x_0 y otra extremadamente próxima $x_0 + \delta_0$. Supongamos que la separación después de n iteraciones es $|\delta_n| = |\delta_0|e^{n\lambda}$. Despejando λ y tomando en cuenta que $\delta_n = f^n(x_0 + \delta_0) - f^n(x_0)$, se tiene que:

$$\lambda = \frac{1}{n} \ln \left| \frac{\delta_n}{\delta_0} \right| = \frac{1}{n} \ln \left| \frac{f^n(x_0 + \delta_0) - f^n(x_0)}{\delta_0} \right| = \frac{1}{n} \ln |(f^n)'(x_0)|, \quad (1.13)$$

Si esta expresión tiene límite para $n \rightarrow \infty$, entonces a este límite,

$$\lambda = \lim_{n \rightarrow \infty} \frac{1}{n} \ln |(f^n)'(x_0)| \quad (1.14)$$

se le llama exponente de Liapunov para la órbita que comienza en x_0 . La definición dada depende de x_0 pero se espera el mismo resultado para cualquier órbita en la cuenca de atracción de un atractor y se tiene que $\lambda > 0$ para los atractores caóticos y $\lambda < 0$ para puntos fijos y ciclos estables [11, 17].

1.4. Probabilidades

Muchos de los acontecimientos que suceden a diario no son deterministas. Por ejemplo, ¿Cual será el precio del barril de petróleo?, ¿Lloverá o no?. Este tipo de acontecimientos se denominan experimentos aleatorios. Es decir, un **experimento** es **aleatorio** cuando no podemos predecir exactamente su resultado, sino que sólo podemos predecir con que frecuencia se obtienen los distintos resultados en un gran número de intentos. Los juegos de azar brindan ejemplos clásicos de experimentos aleatorios.

El conjunto de todos los posibles resultados asociados a un experimento aleatorio es llamado **espacio muestral** denominado comúnmente Ω .

Para el análisis de un experimento aleatorio también son de interés los **eventos** que son subconjuntos del espacio muestral Ω que corresponderá a la realización de algunos resultados particulares.

Una medida de **probabilidad** es una función que asigna a cada evento en el espacio muestral un número $P(A) \in [0, 1]$ que nos dice cual es la probabilidad de que ocurra dicho evento al observar un experimento aleatorio, es decir, esta brinda la posibilidad de “medir” la incertidumbre inherente al evento.

Si se le asignan a los eventos que aparecen con poca o ninguna frecuencia probabilidad 0 y a los eventos que ocurren con mayor frecuencia la probabilidad 1, entonces una medida de probabilidad es una función que debe satisfacer las siguientes propiedades:

$$P(\Omega) = 1 \quad (1.15)$$

$$P(\emptyset) = 0 \quad (1.16)$$

Si se tiene un evento A cuyos elementos son disjuntos $a_1, a_2 \dots a_N$, es decir que no pueden producirse simultáneamente ($a_i \cap a_j = \emptyset, \forall i \neq j$) entonces

$$P(A) = \sum_{n=1}^N P(a_n) \quad (1.17)$$

Y si además se tienen dos eventos A y B que son disjuntos

$$P(A \cup B) \leq P(A) + P(B) \quad (1.18)$$

1.4.1. Probabilidad Marginal, Condicional y Conjunta

A menudo en el cálculo de las probabilidades se tiene información a priori acerca del resultado de un experimento, lo cual debe modificar el cálculo de las probabilidades de cada evento.

Probabilidad Conjunta

Dado un espacio muestral Ω y dos eventos A y B , se llama a un evento nuevo como conjunción de A y B ($A \cap B$) cuando ambos eventos ocurren simultáneamente. Ver figura 1.1

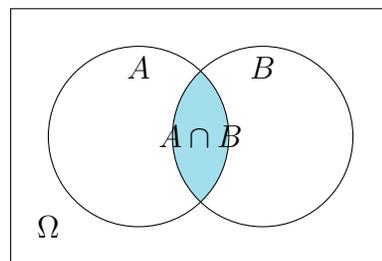


Figura 1.1: Intersección entre los conjuntos A y B .

Entonces, la probabilidad conjunta que puede ser representada como $P(A \cap B)$ o $P(A, B)$ (en este trabajo será representada de la segunda forma), mide la frecuencia con la que suceden dos eventos al mismo tiempo.

Probabilidad Condicional

Como su nombre lo indica se trata de determinar la probabilidad de que ocurra un evento A (aposteriori) dado que ya aconteció un evento B (apriori), y se representa mediante $P(A|B)$, se lee probabilidad de A dado B o probabilidad de A condicionada a B .

En la probabilidad condicional, se considera que de un espacio muestral Ω se conoce únicamente el evento B , que constituye un espacio muestral reducido.

Como únicamente conocemos el evento B , la probabilidad de que exista A está dada por la posible intersección del evento A con el evento B , ver figura 1.1.

Por lo tanto la expresión para la probabilidad condicional quedaría

$$P(A|B) = \frac{n(A \cap B)}{n(B)} = \frac{P(A, B)}{P(B)} \quad (1.19)$$

donde $n(A \cap B)$ es el número de elementos en la intersección de A con B y $n(B)$ es el número de elementos en el evento B .

Probabilidad Marginal

Esta es la más simple de todas y la más utilizada, mide la cantidad de veces que aparece el evento A en N observaciones, y se calcula sumando las probabilidades de los eventos conjuntos, de los cuales el evento A forma parte, por lo tanto:

$$P(A) = \sum_{k=1}^N P(A, B_k) \quad (1.20)$$

En otra palabras, la probabilidad de un evento A es igual a la suma de probabilidades conjuntas del evento B_k y los eventos A , la suma se realiza sobre los eventos B_k .

Teorema de Bayes

Considerando $P(A_i)$ como la probabilidad a priori de los eventos A_i , y se requiere conocer una probabilidad a posteriori de cada uno de ellos, dado que ya conocemos el evento B , A_k representa a cualquiera de los eventos A_i .

$$P(A_k|B) = \frac{P(A_k, B)}{P(B)}$$

como $P(A_k, B) = P(A_k)P(B|A_k)$ y la probabilidad total de B es $P(B) = \sum_{i=1}^N P(A_i, B)$, tenemos:

$$P(A_k|B) = \frac{P(A_k)P(B|A_k)}{\sum_{i=1}^N P(A_i)P(B|A_i)} \quad (1.21)$$

La expresión 1.21 se conoce como Teorema de Bayes, que establece la probabilidad de un evento particular A_k de los eventos A_i , dado que ya sucedió el evento B , expresada en términos de probabilidad condicional. (ver [18])

Capítulo 2

Información y Entropía

En la vida cotidiana el concepto de información tiene gran importancia, por ejemplo, nos basamos en esta para tomar decisiones. Aunque en muchos casos el término información es mal utilizado, ya que, se llama información a cualquier mensaje independientemente de que éste nos informe o no. Una comprensión precisa del concepto de información es crucial para varias disciplinas científicas. A la inversa, una comprensión vaga del concepto puede conducir a profundos malentendidos, en este capítulo se responderán las siguientes preguntas relacionadas con la información: ¿Cuál es realmente el significado del término?. ¿Cómo saber si algo es información o no?. ¿Cuál es la diferencia entre la información y la entropía?

En 1948, Claude Shannon publicó un artículo titulado *Una teoría matemática de la comunicación* [19], en el cual se hizo evidente que la información es una cantidad bien definida y, sobre todo, medible. El artículo de Shannon describe una teoría sutil que nos dice algo fundamental acerca de la forma en que funciona el universo.

2.0.1. Información Según Shannon

La teoría de la información de Shannon proporciona una definición matemática de la información y describe con precisión cuánta información puede comunicarse entre diferentes elementos de un sistema.

Supongamos un evento con k estados posibles y equiprobables, esto quiere decir que $P(x) = \frac{1}{k}$. La Información I de este evento puede ser representada como:

$$I(x) = \log_b(k) = \log_b\left(\frac{1}{P(x)}\right) \quad (2.1)$$

$$I(x) = -\log_b(P(x)) \quad (2.2)$$

donde b determina la unidad en que se medirá la información. Si $b = 2$ la unidad es llamada bit, si $b = 10$ la unidad se denomina bans y por último, si $b = e$ la unidad es llamada nat.

La definición de la información se deriva de las propiedades que se quiere que esta cumpla:

- La información es una cantidad positiva $I(a) \geq 0$.
- Si un evento tiene probabilidad 1 entonces no se tiene información de la ocurrencia del evento $I(1) = 0$.
- Si dos eventos independientes ocurren, entonces la información que se obtiene al observarlos es la suma de la información de cada uno $I(a_1 a_2) = I(a_1) + I(a_2)$.
- Que sea una función continua de la probabilidad.

Es fácil notar en la expresión 2.2, que la probabilidad de un evento es equivalente a su información. Considere una moneda que cae cara el 90 % del tiempo (es decir, $p(x_c) = 0,9$). Cuando se lanza esta moneda, se espera que caiga cara arriba ($x = x_c$), así que cuando lo hace, esto sorprende menos que cuando cae el sello hacia arriba ($x = x_s$), por ello, este no proporciona ningún tipo de información; por el contrario, cuando cae sello hacia arriba, esto si proporciona información valiosa sobre el sistema. Cuanto más improbable sea un resultado particular, más sorprendido está el observador. Es decir, si la probabilidad de un evento es alta/baja el contenido de su información es bajo/alto, esto es conocido como sorpresa del evento. Se puede decir, que un evento transmite más información, mientras este reduzca la incertidumbre del sistema.

2.0.2. Entropía de Shannon

Si se tiene una variable aleatoria \mathbf{x} que toma valores x_i con probabilidades P_i , $i = 1, \dots, N$, su entropía será la suma ponderada de la cantidad de información:

$$H(\mathbf{x}) = - \sum_{i=1}^N P_i \log_2(P_i) \quad (2.3)$$

definiendo $0 \log(0) := 0$ dado que $\lim_{n \rightarrow 0} n \log(n) = 0$.

La expresión 2.3 es llamada entropía de \mathbf{x} , también conocida como entropía de Shannon [19] y esta cuantifica la incertidumbre involucrada en la predicción de una variable aleatoria.

En las ecuaciones 2.2 y 2.3 se puede observar que de acuerdo a Shannon la entropía de una distribución de probabilidades es justo el valor medio de la información de esa distribución.

Análogamente, para dos variables aleatorias \mathbf{x} e \mathbf{y} con sus distribuciones respectivas asociadas se extiende el concepto de entropía:

- Entropía Conjunta [20]: Es escrita en término de la distribución de probabilidad conjunta.

$$H(\mathbf{x}, \mathbf{y}) = - \sum_{x,y} P(x, y) \log_2(P(x, y)) \quad (2.4)$$

donde $P(x, y)$ es la probabilidad conjunta.

- Entropía Condicional [20]: Se define en término del promedio ponderado de probabilidad condicional local.

$$H(\mathbf{x}|\mathbf{y}) = - \sum_{x,y} P(y)P(x|y) \log_2(P(x|y)) \quad (2.5)$$

donde $P(x|y)$ es la probabilidad condicional de x dado y y $P(y)$ es la probabilidad marginal de y .

Si un emisor desea enviar cierta información \mathbf{x} proveniente de una fuente se puede interpretar la entropía como la cantidad de información promedio que se está enviando. Similarmente el receptor cuantifica en la entropía la incertidumbre asociada a la variable \mathbf{x} . A partir de esta definición intuitiva de entropía se pueden entender las propiedades que se derivan de 2.3:

- $H(\mathbf{x})$ es una función continua porque a pequeños cambios en la probabilidad hay pequeños cambios en la información contenida.
- $H(\mathbf{x}) \geq 0$.
- $H(\mathbf{x})$ es máximo si $P(x_i) = P(x_j)$ para todo i, j . Cuando la distribución es uniforme la incertidumbre sobre \mathbf{x} es máxima.
- Si $P_i = 1$ para algún i entonces $H(\mathbf{x}) = 0$. Si existe total certeza de la ocurrencia de un evento, la incertidumbre sobre \mathbf{x} es 0.
- $H(\mathbf{x}|\mathbf{y}) \leq H(\mathbf{x}) + H(\mathbf{y})$, con la igualdad cuando \mathbf{x} y \mathbf{y} son independientes. Esta propiedad se conoce como subaditividad de la entropía de Shannon.

2.0.3. Información Mutua

La independencia entre dos variables \mathbf{x} e \mathbf{y} puede ser medida de la siguiente forma (ver figura 2.1)

$$\begin{aligned}
 I(\mathbf{x}; \mathbf{y}) &= H(\mathbf{x}) - H(\mathbf{x}|\mathbf{y}) \\
 &= -\sum_x P(x) \log_2(P(x)) - \left(-\sum_{x,y} P(y)P(x|y) \log_2(P(y|x)) \right) \\
 &= -\sum_{x,y} P(x,y) \log_2(P(x)) + \sum_{x,y} P(x,y) \log_2(P(x|y)) \\
 &= \sum_{x,y} P(x,y) \log_2 \left(\frac{P(x,y)}{P(x)P(y)} \right) \tag{2.6}
 \end{aligned}$$

La información que comparten dos eventos \mathbf{x} e \mathbf{y} se denomina información mutua [21] y se escribe $I(\mathbf{x}; \mathbf{y})$.

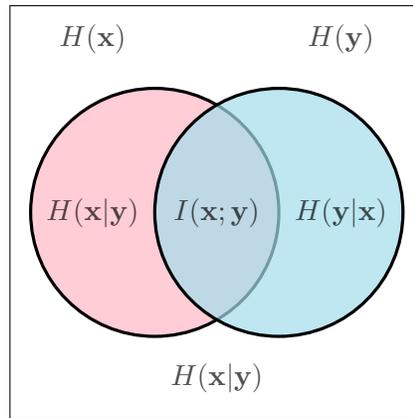


Figura 2.1: Relación entre entropías e información mutua.

Es fácil notar en la ecuación 2.6 que si los eventos \mathbf{x} e \mathbf{y} son independientes $P(x,y) = P(x)P(y)$ entonces, el logaritmo es cero y cada término en la suma será cero, entonces $I(\mathbf{x}; \mathbf{y}) = 0$, como es de esperarse.

Por otro lado, si \mathbf{x} e \mathbf{y} están relacionados (es decir, los eventos comparten información) $I(\mathbf{x}; \mathbf{y}) \neq 0$, esto quiere decir, que la información mutua es una medida útil para determinar si dos eventos están relacionados o no, sin embargo no expresa el grado de relación entre estos eventos, ni la dirección en la que fluye la información, ya que $I(\mathbf{x}; \mathbf{y}) = I(\mathbf{y}; \mathbf{x})$.

2.1. Causalidad de Wiener-Granger

La causalidad en series de tiempo consiste en determinar causa y efecto a partir de la observación conjunta de los datos. Existe una amplia gama de diferentes métodos y técnicas que permiten determinar la causalidad en series temporales, como por ejemplo la transferencia de entropía [7] o la causalidad de Wiener[22]-Granger [23].

Para poder realizar un análisis causal de dos series temporales se requiere la afirmación de que, una de la series puede interpretarse como la conducción de la otra, es decir, la serie “causa” está impulsando la serie de “efectos” donde el significado técnico de “conducir” depende de las herramientas específicas que se utilizan pero, en general, implica que algunos datos o estructura de datos (la “causa”) en el pasado de una serie predice, o aumenta la probabilidad de, algunos datos o estructura de datos (es decir, el “efecto”) en el presente o futuro de otra serie.

El análisis causal exploratorio de series temporales se define por el enfoque del análisis más que por el uso de cualquier conjunto específico de herramientas o suposiciones, ver [24].

Entre 1959-1960 como estudiante posdoctoral, Granger tuvo que dedicar su atención al concepto de causalidad, por lo cual tuvo que leer un trabajo de Norbert Wiener (ver [22]), el cual contenía una definición de la causalidad. Esencialmente, fue esta definición, la que Granger analizó e implementó en 1969 [23]. La afirmación respecto a la causalidad de Wiener se basa en dos componentes:

1. La causa ocurre antes del efecto.
2. La causa contiene información sobre el efecto que es única y no está en otra variable.

Una consecuencia de esta afirmación es que la variable causal puede contribuir a la predicción de la variable efecto.

2.1.1. Definición

Considere dos series temporales $\mathbf{x} = \{x_t\}$ e $\mathbf{y} = \{y_t\}$ con $t = 1, \dots, n$, donde $t = n$ es considerado el tiempo presente y todo el conocimiento disponible en el universo en todo momento $t \leq n$ será denotado como Ω_n . Granger introdujo los siguientes dos axiomas, los cuales se asume que siempre se cumplen:

Axioma 2.1 *El pasado y el presente pueden causar el futuro, pero el futuro no puede causar el pasado.*

Axioma 2.2 Ω_n no contiene información redundante, de modo que si alguna variable \mathbf{z} está relacionada funcionalmente a alguna de las otras variables, de manera determinista, entonces \mathbf{z} debe ser excluida de Ω_n .

Dado esto, Granger introduce la siguiente definición: dado un cierto conjunto A , \mathbf{y} causa \mathbf{x} si

$$P(x_{n+1} \in A | \Omega_n) \neq P(x_{n+1} \in A | \Omega_n - \mathbf{y}) \quad (2.7)$$

donde $\Omega_n - \mathbf{y}$ se define como todo el conocimiento disponible en el universo, excepto \mathbf{y} ; es decir, \mathbf{y} causa \mathbf{x} si la probabilidad de que un valor futuro de la serie \mathbf{x} esté en algún conjunto A es diferente dependiendo de si se da o no todo el conocimiento en el universo hasta el presente o solo se da ese conocimiento que no está en \mathbf{y} . Intuitivamente, esta diferencia 2.7 implica que el conocimiento de \mathbf{y} proporciona cierto conocimiento de \mathbf{x} .

La definición operacional de la causalidad de Granger se centra en construir modelos de predicción en series temporales.

Considere un modelo de vectores autoregresivos para el sistema de \mathbf{x} e \mathbf{y} ,

$$\begin{pmatrix} x_t \\ y_t \end{pmatrix} = \sum_{i=1}^n \begin{pmatrix} A_{xx}^i & A_{xy}^i \\ A_{yx}^i & A_{yy}^i \end{pmatrix} \begin{pmatrix} x_{t-i} \\ y_{t-i} \end{pmatrix} + \begin{pmatrix} \epsilon_{x,t} \\ \epsilon_{y,t} \end{pmatrix} \quad (2.8)$$

donde $\epsilon_{x,t}$ son términos de ruido no correlacionado.

Supongamos que la matriz de coeficientes A es encontrada de tal manera que este modelo es óptimo en algún sentido, por ejemplo, el error de la predicción se minimiza, entonces “ \mathbf{y} Grangercausa \mathbf{x} ” sólo si $A_{xy} \neq 0$ y “ \mathbf{x} Grangercausa \mathbf{y} ” sólo si $A_{yx} \neq 0$. Una prueba estadística puede ser formulada bajo la hipótesis nula para la no causalidad, es decir, “ \mathbf{y} Grangercausa \mathbf{x} ” o viceversa si $A_{xy} = 0$ o $A_{yx} = 0$.

El uso de este esquema en el caso de subsistemas no lineales puede conducir a conclusiones erradas ya que, requiere generalizaciones para permitir la investigación de procesos no lineales además de que Granger no proporciona definiciones formales de “información”, en contraste con las medidas de la teoría de la información. Por lo cual, se han desarrollado otras estrategias para la detección de causalidad en series temporales.

2.2. Transferencia de Entropía de Schreiber

En el año 2000, Schreiber introdujo la transferencia de entropía como una medida teórica de la información que cuantifica la coherencia estadística entre sistemas que evolucionan en el tiempo [7] esta se ha convertido en una de las herramientas más utilizadas para detectar causalidad entre series temporales.

Como se observó anteriormente, la información mutua (ver sección 2.0.3) es simétrica y no es una medida efectiva para determinar en que dirección fluye la información, por ello Schreiber propone la transferencia de entropía que es una medida no simétrica y que está basada en el cambio de entropía del sistema. Esta medida, con suposiciones mínimas sobre la dinámica del sistema y la naturaleza de su acople es capaz de cuantificar el intercambio de información entre sistemas.

Supongamos dos sistemas que generan eventos, ahora definamos una tasa de entropía H_1 , que es la cantidad de información adicional que se requiere para representar el valor de la siguiente observación de uno de los sistemas. Esta tasa de entropía tiene la siguiente forma:

$$H_1 = - \sum_{x_{n+1}} P(x_{n+1}, x_n, y_n) \log_2 (P(x_{n+1}|x_n, y_n))$$

Si además se define una tasa de entropía H_2 en la cual se supone que la observación x_{n+1} es independiente de y_n , es decir, que $P(x_{n+1}|x_n, y_n) = P(x_{n+1}|x_n)$:

$$H_2 = - \sum_{x_{n+1}} P(x_{n+1}, x_n, y_n) \log_2 (P(x_{n+1}|x_n))$$

La transferencia de entropía se obtiene haciendo $H_2 - H_1$:

$$\begin{aligned} H_2 - H_1 &= - \sum_{x_{n+1}, x_n, y_n} P(x_{n+1}, x_n, y_n) \log_2 (P(x_{n+1}|x_n)) \\ &\quad + \sum_{x_{n+1}, x_n, y_n} P(x_{n+1}, x_n, y_n) \log_2 (P(x_{n+1}|x_n, y_n)) \\ &= \sum_{x_{n+1}, x_n, y_n} P(x_{n+1}, x_n, y_n) \log_2 \left(\frac{P(x_{n+1}|x_n, y_n)}{P(x_{n+1}|x_n)} \right) \end{aligned}$$

Podemos observar que esta expresión es asimétrica, por lo tanto obtenemos dos ecuaciones para la transferencia:

$$T_{\mathbf{x} \rightarrow \mathbf{y}} = \sum_{y_{n+1}, x_n, y_n} P(y_{n+1}, x_n, y_n) \log_2 \left(\frac{P(y_{n+1}|x_n, y_n)}{P(y_{n+1}|y_n)} \right) \quad (2.9)$$

$$T_{\mathbf{y} \rightarrow \mathbf{x}} = \sum_{x_{n+1}, x_n, y_n} P(x_{n+1}, x_n, y_n) \log_2 \left(\frac{P(x_{n+1}|x_n, y_n)}{P(x_{n+1}|x_n)} \right) \quad (2.10)$$

Esta es la transferencia de entropía de Schreiber [7].

$P(x_{n+1}|x_n, y_n)$ y $P(y_{n+1}|x_n, y_n)$ pueden ser escritas en término de las probabilidades conjuntas y marginales utilizando la expresión 1.19:

$$T_{\mathbf{x} \rightarrow \mathbf{y}} = \sum_{y_{n+1}, x_n, y_n} P(y_{n+1}, x_n, y_n) \log_2 \left(\frac{P(y_{n+1}, x_n, y_n) P(y_n)}{P(y_{n+1}, y_n) P(x_n, y_n)} \right) \quad (2.11)$$

$$T_{\mathbf{y} \rightarrow \mathbf{x}} = \sum_{x_{n+1}, x_n, y_n} P(x_{n+1}, x_n, y_n) \log_2 \left(\frac{P(x_{n+1}, x_n, y_n) P(x_n)}{P(x_{n+1}, x_n) P(x_n, y_n)} \right) \quad (2.12)$$

Es fácil notar, que la transferencia de entropía también puede ser escrita en término de las probabilidades condicionales y marginales.

$$T_{\mathbf{x} \rightarrow \mathbf{y}} = \sum_{y_{n+1}, x_n, y_n} P(y_{n+1}|x_n, y_n) P(y_n|x_n) P(x_n) \log_2 \left(\frac{P(y_{n+1}|x_n, y_n)}{P(y_{n+1}|y_n)} \right) \quad (2.13)$$

$$T_{\mathbf{y} \rightarrow \mathbf{x}} = \sum_{x_{n+1}, x_n, y_n} P(x_{n+1}|x_n, y_n) P(x_n|y_n) P(y_n) \log_2 \left(\frac{P(x_{n+1}|x_n, y_n)}{P(x_{n+1}|x_n)} \right) \quad (2.14)$$

2.3. Mapas

Este tipo de sistemas dinámicos ya fueron explicados brevemente en la sección 1.2.1. La virtud de los mapas, es que estos son relativamente simples de tratar, fáciles de entender y analizar porque son gobernados por ecuaciones finitas de diferencia en vez de por una ecuación diferencial no lineal. En áreas como la biología, economía y otras áreas de las ciencias, muchas veces el tiempo debe considerarse como discreto en vez de continuo.

En esta sección se hablará un poco de los mapas utilizados.

2.3.1. Mapa Logístico

El mapa Logístico [12] es uno de los mapas más utilizados por su simpleza y porque este tiene una dinámica muy rica. Puede interpretarse como un modelo de población en un sistema en el que los recursos son limitados. Su comportamiento viene dado por la siguiente ecuación:

$$f(x_i) = rx_i(1 - x_i) \quad (2.15)$$

El crecimiento dado por el término $f(x_i) = rx_i$ resulta limitado por la no linealidad de $-rx_i^2$, donde x_i representa el número de la población correspondiente a la i -ésima generación. r es llamado el parámetro de crecimiento y debe cumplir con $0 \leq r \leq 4$ para garantizar que la función se mantenga dentro del intervalo $[0, 1]$. En la figura 2.2 podemos observar el espacio de fase del mapa Logístico.

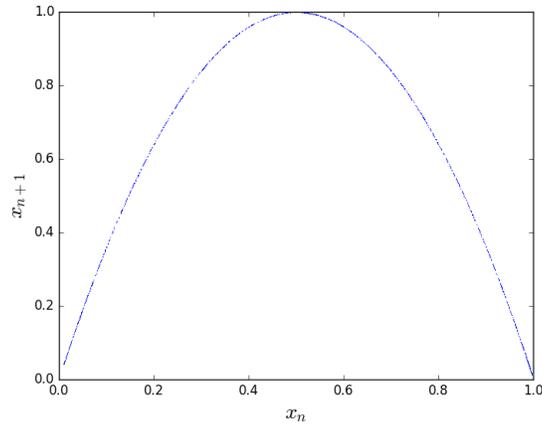


Figura 2.2: Representación geométrica del Mapa Logístico con $r = 3,99$.

Según el valor que se le adjudique a r , se observan diferentes comportamientos en la función.

En la figura 2.3 se puede observar claramente que para el mapa logístico se va doblando el período como ruta al caos, también es fácil ver que la función tiene ventanas de periodicidad que se vuelven más pequeñas hasta que después de $r = 3,6$ las bifurcaciones se vuelven tan impredecibles que la población puede visitar cualquier valor (pseudoaleatorios). Esto se conoce como el período que duplica el camino hacia el caos. El período de duplicación aumenta de 16 a 32 y así sucesivamente, para obtener este efecto de caos. En el punto $r = 3,9$. La función de Mapa Logístico parece determinista a partir de los parámetros iniciales, pero causa un conjunto aparentemente aleatorio de números es decir, para $r > 3,6$ el mapa logístico es una función caótica.

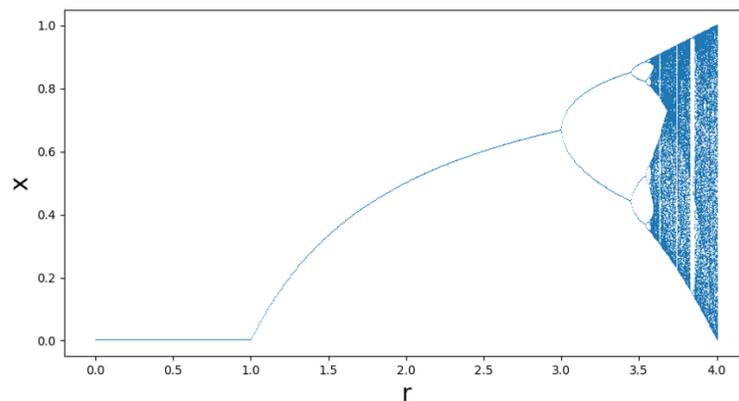


Figura 2.3: Diagrama de bifurcaciones del mapa Logístico.

2.3.2. Mapa Triangular

El mapa Triangular [25], que como el mapa logístico también es un mapa unidimensional. Es una transformación no invertible del intervalo de la unidad a sí mismo.

$$g(x_i) = \begin{cases} \frac{x_i}{a} & \text{si } 0 \leq x_i \leq a \\ \frac{1-x_i}{1-a} & \text{si } a < x_i \leq 1 \end{cases} \quad (2.16)$$

Como se puede observar en la ecuación (2.16), este depende del parámetro $a \in [0, 1]$. Al igual que en el mapa logístico, el comportamiento de la función dependerá del valor del parámetro a (ver figura 2.4).

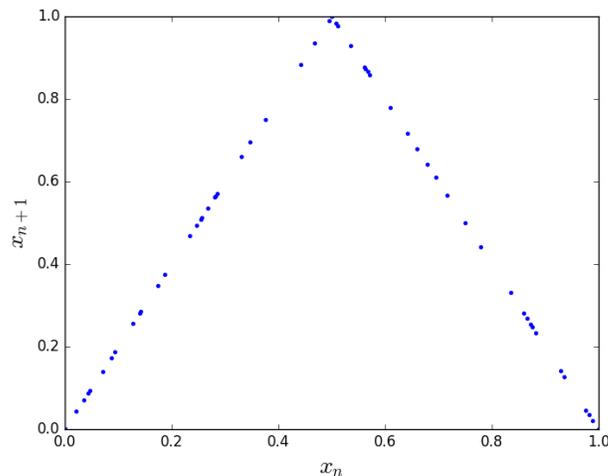


Figura 2.4: Iteración del Mapa triangular para $a = 0,5$.

2.3.3. Mapa de Hénon

El mapa de Hénon [26] es un sistema dinámico discreto en el tiempo. Es uno de los ejemplos de sistemas dinámicos más estudiado que muestra comportamiento caótico. El mapa de Hénon toma un punto (x_n, y_n) en el plano y lo mapea a un punto nuevo.

$$x_{n+1} = y_n + 1 - ax_n^2 \quad (2.17)$$

$$y_{n+1} = bx_n \quad (2.18)$$

El mapa depende de dos parámetros, a y b que son constantes reales y positivas, el

mapa clásico de Hénon tiene valores de $a = 1,4$ y $b = 0,3$, en los cuales el mapa es caótico (ver figura 2.5).

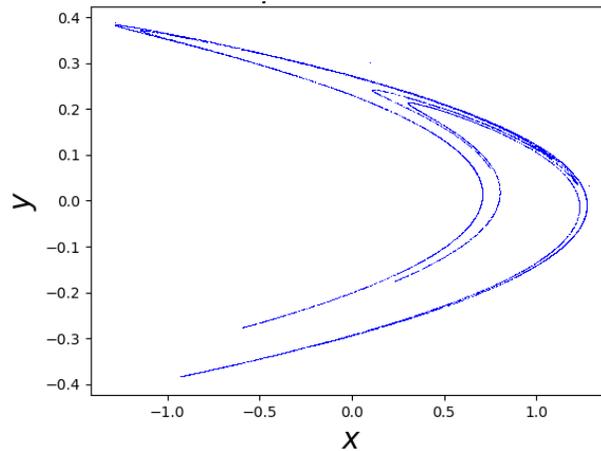


Figura 2.5: Atractor de Hénon con $a = 1,4$ y $b = 0,3$.

2.4. Sistemas Complejos

Un sistema complejo es una colección de elementos que interactúan conformando un todo [27]. Además, estos cumplen con las siguientes características:

- Tienen un gran número de elementos que interactúan entre sí.
- Estos elementos exhibirán comportamientos emergentes, es decir, un comportamiento colectivo de auto-organización.
- Su comportamiento emergente no resulta de la existencia de un controlador central.

En la naturaleza se pueden encontrar una gran cantidad de ejemplos de sistemas complejos que se extienden desde la física hasta la economía. Esta clase de sistemas no constituye un caso raro ni excepcional, sino que se manifiesta en la inmensa mayoría de los fenómenos que se observan a diario. Sin embargo, y a pesar de su gran diversidad y abundancia, se pueden identificar conductas dinámicas genéricas, no importa su naturaleza (física, química, biológica o social). Entre ellas, las leyes de crecimiento, la autoorganización y los procesos colectivos emergentes. Como ejemplos de sistemas complejos se pueden mencionar una célula, un cerebro, un ecosistema, una sociedad de insectos, un sistema inmunológico o una economía de mercado, entre muchos otros.

La mayoría de los sistemas complejos son inestables y de la misma forma que en los sistemas caóticos, cualquier variación mínima entre sus elementos puede modificar,

de forma imprevisible, las interrelaciones y por lo tanto, el comportamiento de todo el sistema. La evolución de esta clase de sistemas se caracteriza por la intermitencia, es decir, una situación en la que el orden y el desorden se alternan constantemente. Los estados evolutivos en los sistemas complejos no transcurren a través de procesos continuos y graduales, sino que suceden por medio de reorganizaciones y saltos.

Estos sistemas nunca llegan a un óptimo global, es decir, al estado de mínima energía. En general, crecen progresivamente hasta que llegan al límite de su desarrollo potencial. En ese instante, sufren un desorden, una especie de ruptura que induce una fragmentación del orden preexistente, pero después, comienzan a surgir regularidades que organizan al sistema de acuerdo con nuevas leyes, produciendo otra clase de desarrollo. Este comportamiento es típico en los sistemas naturales. En consecuencia, la organización de los sistemas complejos se da en diferentes niveles. Las leyes que gobiernan la causalidad de un nivel dado, pueden ser totalmente diferentes a las de un nivel superior [28].

2.5. Sistemas de Mapas Acoplados

Uno de los objetivos de este trabajo es detectar causalidad entre mapas acoplados, esto se debe a que una gran variedad de sistemas complejos pueden ser representados de esta manera.

La dinámica de los sistemas de mapas acoplados consiste en dos estados independientes: las dinámicas locales y las dinámicas acopladoras. Para ello se utilizan dos mapas locales unidimensionales, $f(x)$ y $g(x)$, y se utilizaron dos tipos de dinámicas acopladoras, que serán explicadas a continuación:

Acople 1 [25]

$$x_{t+1} = f(x_t + \mu(y_t - x_t)) \quad (2.19)$$

$$y_{t+1} = g(y_t + \epsilon(x_t - y_t)) \quad (2.20)$$

Acople 2 [29]

$$x_{t+1} = (1 - \mu) f(x_t) + \mu f(y_t) \quad (2.21)$$

$$y_{t+1} = (1 - \epsilon) g(y_t) + \epsilon g(x_t) \quad (2.22)$$

Podemos observar que ambos acoples poseen los parámetros μ y ϵ los cuales tienen como finalidad modular la intensidad de los acoples entre \mathbf{x} e \mathbf{y} , es decir, estos regulan la cantidad de información que comparte cada uno. Esto nos permite saber con anterioridad en que sentido fluye la información, y basándonos en esto determinar si los métodos propuestos nos conducen a resultados correctos. Por lo cual se variaron dichos parámetros ya que a medida que μ y ϵ se acercan a 1 el acoplamiento se hace más fuerte.

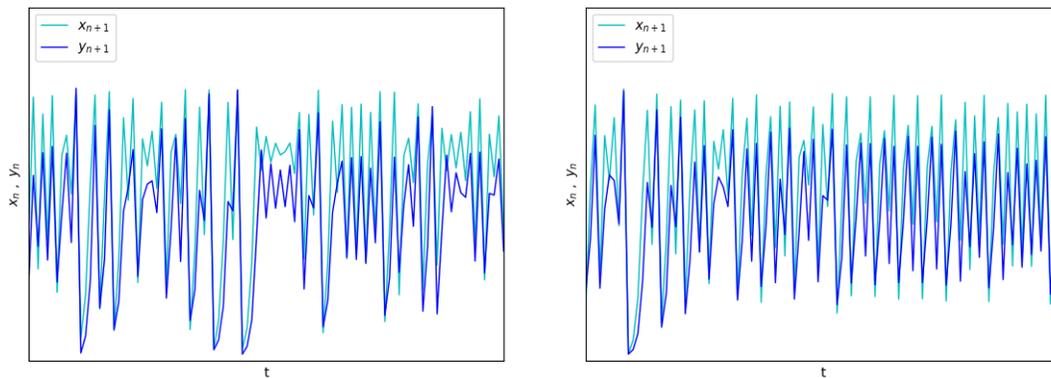


Figura 2.6: f =Mapa Logístico, g =Mapa Triangular, $r = 3,99$, $a = 0,5$, $\mu = 0,2$ y $\epsilon = 0,4$. Acople 1 (izquierda) y Acople 2 (derecha).

En la figura 2.6 se pueden observar las series temporales obtenidas con f el Mapa Logístico y g el Mapa Triangular y con los parámetros de acople $\mu = 0,2$ y $\epsilon = 0,4$. Para ambos acoples podemos observar que las series temporales \mathbf{x} e \mathbf{y} siguen dinámicas parecidas, esto es debido a los parámetros μ y ϵ que son lo que modulan la intensidad del acople.

Capítulo 3

Discretización del Espacio de Estados

En la sección 2.2 se presentó la transferencia de entropía de Schreiber, esta herramienta se utilizó para medir el intercambio de información en los sistemas de mapas acoplados, ya que esta hace suposiciones mínimas sobre la dinámica del sistema y la naturaleza de su acople, siendo una medida eficiente para detectar la transferencia de información y la dirección del acople.

En este caso (y es lo más común en los sistemas reales) la transferencia de entropía deberá ser estimada a partir de los datos, ya que en la mayoría de los casos se desconoce la distribución de probabilidades de las series temporales medidas.

El problema para calcular la transferencia de entropía consiste en hallar la distribución de las probabilidades de la data.

Un método directo para estimar las densidades de probabilidad consiste en discretizar el proceso continuo y luego hallar las probabilidades necesarias mediante la observación de la frecuencia relativa de los nuevos valores discretos.

En este trabajo se utilizaron dos métodos de discretización, el reticulado y la representación simbólica.

3.1. Método de Discretización 1: Reticulado (MD1)

Sea \mathbf{x} una serie temporal parcialmente discreta con longitud finita N , es decir

$$x_1, x_2, x_3, \dots, x_N$$

donde s es continuo y t es discreto.

El método de reticulado consiste en divisiones del espacio de estados \mathbf{s} en M conjuntos

$$s_1, s_2, s_3, \dots, s_M$$

la evolución de la serie temporal \mathbf{x} será representada por una secuencia

$$\sigma_1, \sigma_2, \dots, \sigma_N$$

con $\sigma_i \in \mathbf{s}$.

Es decir, a cada valor de la serie temporal \mathbf{x} se le asignará un valor correspondiente con la nueva representación discreta:

$$x_i \rightarrow s_i \text{ si } x_i \in \mathbf{s}$$

■ Ejemplo:

Consideremos la siguiente serie temporal obtenida de la iteración del mapa Logístico con $r = 3,99$ y $x_0 = 0,670$

$$\mathbf{x} = [0,670; 0,032; 0,990; 0,574; 0,336; 0,519; 0,933; 0,892; 0,277; 0,183] \quad (3.1)$$

Si decidimos dividir \mathbf{s} en 3 conjuntos, es decir

$$\sigma(\mathbf{x}) = \begin{cases} 0 & \text{si } 0 \leq x_i \leq 0,333 \\ 1 & \text{si } 0,333 < x_i \leq 0,666 \\ 2 & \text{si } 0,666 < x_i \leq 1 \end{cases}$$

Por lo tanto \mathbf{x} será representada de la siguiente manera:

$$\sigma(\mathbf{x}) = [2, 0, 2, 1, 1, 1, 2, 2, 0, 0]$$

3.2. Método de Discretización 2: Representación Simbólica (MD2)

Sea \mathbf{x} una serie temporal parcialmente discreta con longitud finita N , es decir

$$x_1, x_2, x_3, \dots, x_N$$

donde \mathbf{s} es continuo y t es discreto.

Este método consiste en realizar una partición binaria de la serie temporal \mathbf{x} , esto puede ser expresado con la función de Heaviside:

$$\Theta(\mathbf{x}) = \begin{cases} 0 & \text{si } 0 \leq x_i \leq 0,5 \\ 1 & \text{si } 0,5 < x_i \leq 0,1 \end{cases} \quad (3.2)$$

A partir de esta nueva serie binaria se formarán cadenas de 0 y 1 de tamaño m y se hallará en cada caso el valor en la recta de los números enteros positivos de este número binario y estos serán las nuevas representaciones de \mathbf{x} .

Se debe tener en cuenta que para realizar la conversión de un número binario a un entero, debe realizarse lo siguiente:

1. Desde el extremo derecho del número binario se multiplica cada dígito por 2 elevado a la potencia consecutiva (comenzando por la potencia 0, 2^0).
2. Después de realizar cada una de las multiplicaciones, se suman todos los valores y el número resultante será el valor equivalente del número binario en los números enteros. Por ejemplo, para el número binario 01101:

0	1	1	0	1
0×2^4	1×2^3	1×2^2	0×2^1	1×2^0

$$= 2^3 + 2^2 + 2^0 = 13$$

▪ Ejemplo:

Considerando la misma serie temporal del ejemplo anterior (ver expresión 3.1).

Primero transformamos \mathbf{x} en una serie binaria utilizando la ecuación 3.2, se obtiene:

$$\Theta(\mathbf{x}) = [1, 0, 1, 1, 0, 1, 1, 1, 0, 0]$$

y si por último se escogen cadenas de dos dígitos, es decir, $m = 2$ la nueva representación de \mathbf{x} será:

$$\mathbf{s}(\mathbf{x}) = [2, 3, 1, 2, 0]$$

3.3. Experimentos Realizados

3.3.1. Sistemas de Mapas Acoplados

Para poder estudiar el desempeño de los dos métodos de discretización explicados anteriormente, se implementaron los acoples mencionados en la sección 2.5 y con estos se realizaron diversos experimentos, los cuales consistían en calcular la transferencia de

entropía de Schreiber al variar diferentes parámetros en el sistema de mapas acoplados (ver sección 2.5), tales como la longitud de las series temporales (N), los parámetros de acople (μ y ϵ) y el número de particiones del espacio de estados (M para el MD1 y m para MD2), además para todas las pruebas se dejó un transiente de 500 iteraciones, todos los experimentos fueron realizados con f el mapa logístico y g el mapa triangular y los parámetros de los mapas permanecieron constantes ($r = 3,99$ y $a = 0,5$).

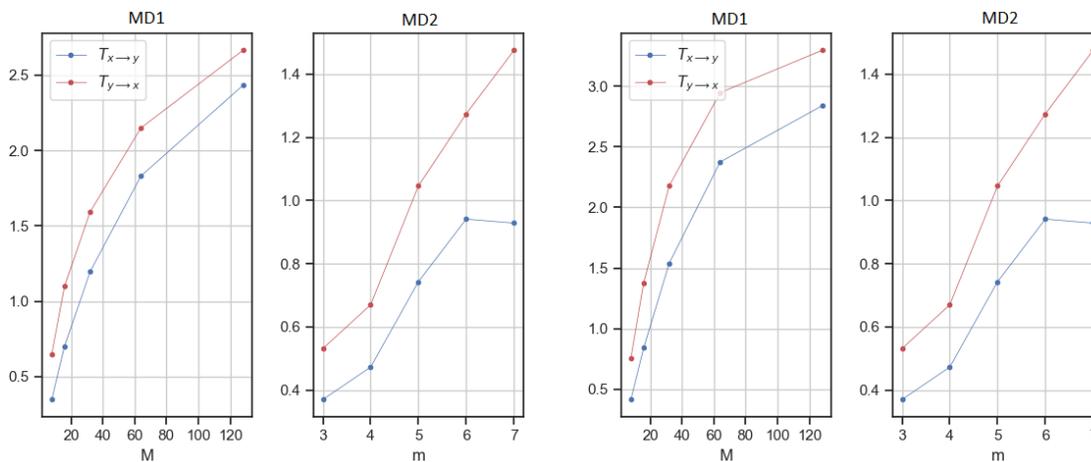


Figura 3.1: Transferencias de entropía para el acople 1 con $\mu = 0,4$ y $\epsilon = 0,2$ para $N = 2000$, y $N = 3000$, respectivamente.

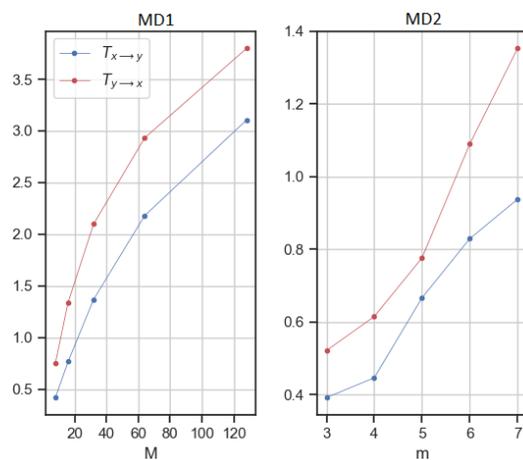


Figura 3.2: Transferencias de entropía para el acople 1 con $N = 10000$, $\mu = 0,4$ y $\epsilon = 0,2$.

El primer experimento puede ser observado en las figuras 3.1 y 3.2, en este sólo fue variado el número de datos, en el cual podemos notar que en ambos métodos de discretización predomina la transferencia de entropía en el sentido de $\mathbf{y} \rightarrow \mathbf{x}$, esto quiere decir

que el mapa triangular envía mas información al mapa logístico, esto es debido a que el parámetro de acople μ es mayor a ϵ , esto puede verse detallado en la sección 2.5.

Vale la pena destacar, que los valores de $T_{x \rightarrow y}$ y $T_{y \rightarrow x}$ obtenidos con el MD1 son ligeramente mayores que los obtenidos con el MD2.

Lo más significativo de este experimento es el hecho de que al variar las longitudes de las series temporales (N), la transferencia de entropía 2.11 y 2.12 son similares, en cada uno de los métodos de discretización, esto quiere decir que N no tiene gran influencia en los valores de la transferencia de entropía, por lo cual todas las pruebas siguientes se realizarán para un sólo valor de N con el fin de disminuir el tiempo de cómputo, ya que, la cantidad de procesos para calcular la transferencia de entropía siempre depende de la longitud de las series temporales.

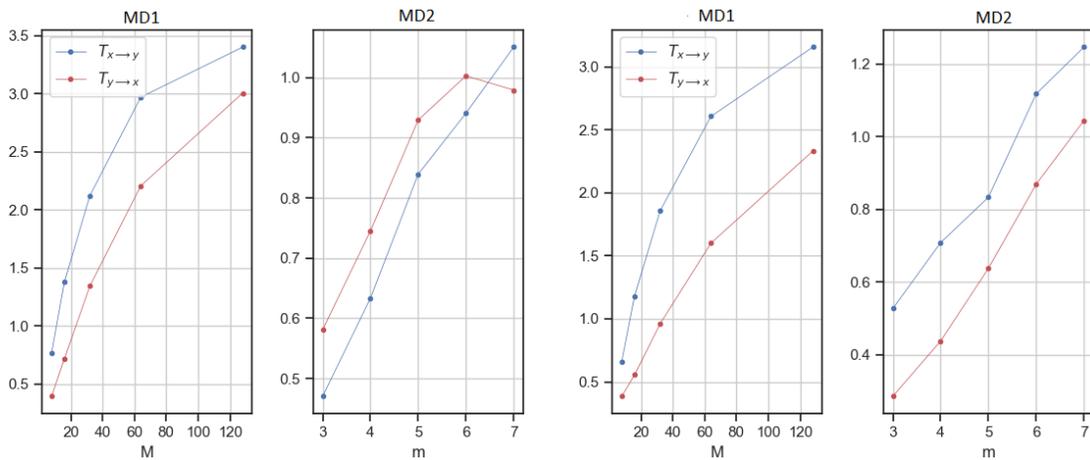


Figura 3.3: Transferencias de entropía con $N = 3000$, $\mu = 0,2$ y $\epsilon = 0,4$ para el acople 1 y el acople 2 respectivamente.

En las gráficas de la figura 3.3, como se esperaba, predomina $T_{x \rightarrow y}$, esto es debido a los valores de los parámetros de acople. Como podemos observar en la sección 2.5, si ϵ es mayor que μ en ambos acoples, \mathbf{x} envía mas información a \mathbf{y} .

En la figura 3.4, al considerar $\mu = 0,6$ y $\epsilon = 0,0$ podemos observar que, para ambos métodos de discretización y acoples, la transferencia de entropía es mayor en la dirección $\mathbf{y} \rightarrow \mathbf{x}$. Para el acople 1 podemos observar que $T_{x \rightarrow y}$ en el caso del MD1 crece ligeramente conforme M aumenta y en el MD2 inicialmente $T_{x \rightarrow y}$ es ligeramente mayor que $T_{y \rightarrow x}$ y después al aumentar m ambos valores se cruzan y predomina $\mathbf{y} \rightarrow \mathbf{x}$. Para el acople 2 en el MD1 $T_{x \rightarrow y}$ tiene el mismo comportamiento que en el acople 1 pero, para el MD2 dicho valor tiende a cero.

En este caso un sólo mapa se encuentra acoplado, es por ello que se observa la notoria diferencia entre los valores de $T_{x \rightarrow y}$ y $T_{y \rightarrow x}$, lo cual puede compararse con la figura 3.3

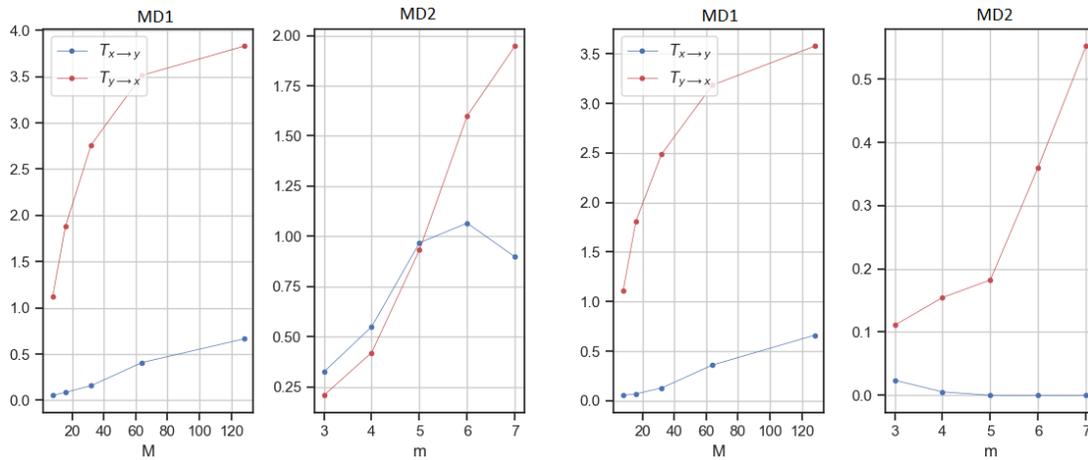


Figura 3.4: Transferencias de entropía con $N = 3000$, $\mu = 0,6$ y $\epsilon = 0,0$ para el acople 1 y el acople 2 respectivamente.

donde $\mu < \epsilon$ pero la diferencia entre estos es menor, por lo tanto la diferencia entre las transferencias también.

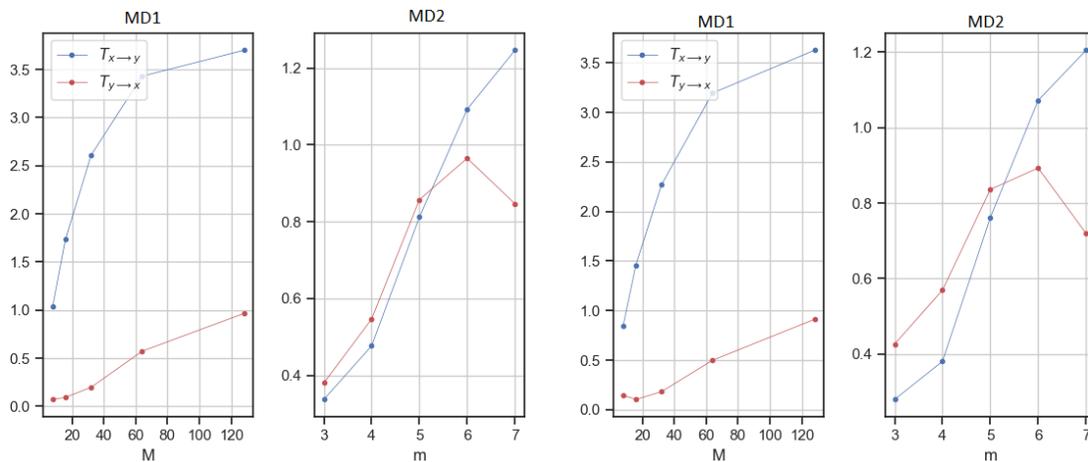


Figura 3.5: Transferencias de entropía con $N = 3000$, $\mu = 0,0$ y $\epsilon = 0,6$ para el acople 1 y el acople 2 respectivamente.

En la figura 3.5 podemos apreciar que para ambos métodos y acoples, la información fluye mayoritariamente en el sentido $\mathbf{x} \rightarrow \mathbf{y}$. Esto se debe a que si $\mu = 0$ en las ecuaciones 2.19 y 2.21 entonces $x(t+1) = f(x(t))$, esto quiere decir que $y(t)$ no envía información a $x(t+1)$.

Si se coloca $\epsilon = 0$ en las ecuaciones 2.20 y 2.22, obtenemos para ambos acoples que

$y(t+1) = g(y(t))$ es decir, obtenemos el caso contrario al de la figura 3.5 (ver figura 3.4).

En todas las figuras podemos observar que con el MD2 es un poco más difícil distinguir como es el flujo de información de un sistema al otro, si m no es escogido correctamente. El MD1 logra identificar con retículas pequeñas la dirección del acople, se puede observar que al aumentar el número de divisiones aumenta el valor de la transferencia y para los últimos puntos aumenta con menor pendiente, es decir parece aproximarse a un valor constante.

También se puede observar que aunque el acople 1 es lineal y el acople 2 es no lineal, se obtienen comportamientos similares en las transferencias en la mayoría de los casos. Cabe destacar que los comportamientos obtenidos concuerdan con lo que se esperaba, esto quiere decir que ambos métodos son útiles en el estudio de la transferencia de información.

En la figura 3.6 se muestran los experimentos realizados con dos mapas triangulares con $a_1 = 0,45$ y $a_2 = 0,5$, y con parámetros de acople $\mu = 0,6$ y $\epsilon = 0,0$ (izquierda) y $\mu = 0,0$ y $\epsilon = 0,6$ (derecha). En ambos casos la información fluye en la dirección esperada. Además, en ambas figuras podemos observar que con el MD2 se necesitan cadenas más grandes de dígitos para detectar la dirección del acople.

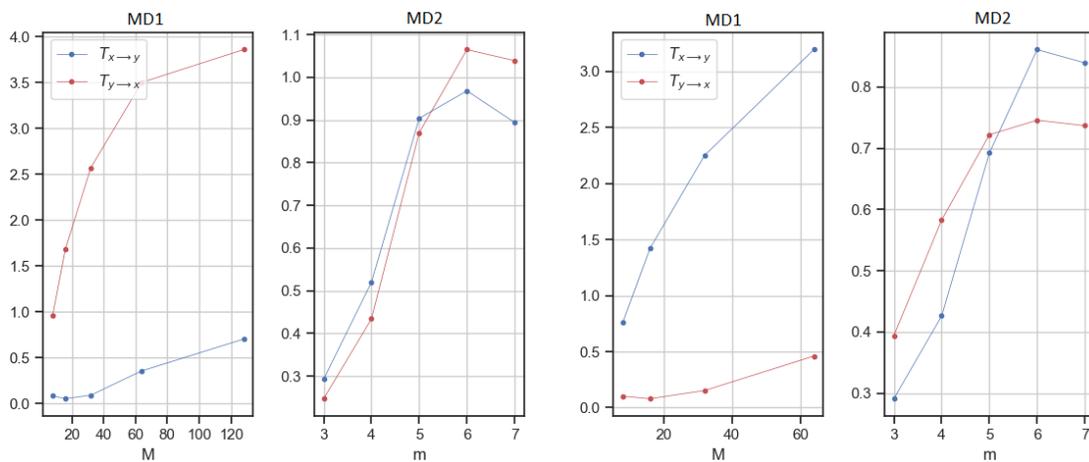


Figura 3.6: Transferencias de entropía para el acople 1 con $N = 3000$, $\mu = 0,6$ y $\epsilon = 0,0$ y $N = 3000$, $\mu = 0,0$ y $\epsilon = 0,6$ para dos mapas triangulares respectivamente.

También, se realizó un barrido en los parámetros de acople μ y ϵ y se calculó la transferencia de entropía para el acople 1 con ambos métodos de discretización. Para el MD1 se escogió $M = 128$ y para el MD2 $m = 7$, debido a que al aumentar estos no se observaron grandes cambios en los valores de la transferencia de entropía, además de que el tiempo de cómputo para valores mayores a $M = 128$ y $m = 7$ era muy grande y no era posible realizar mayor número de particiones por la capacidad del computador empleado.

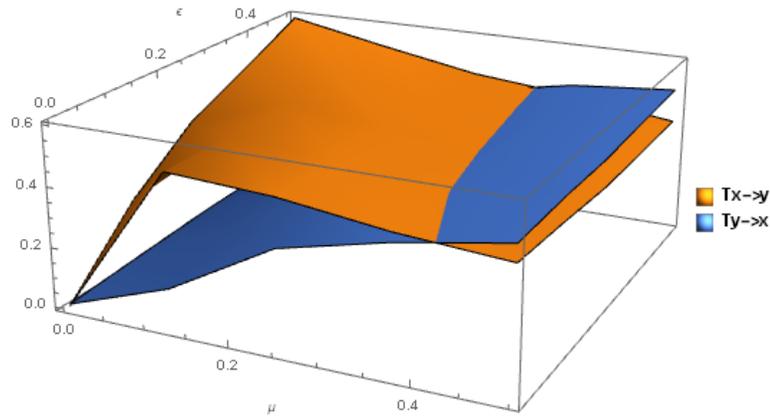


Figura 3.7: Barrido en los parámetros de acople μ y ϵ para el acople 1 y el MD1 con $M = 128$.

En la figura 3.7 podemos observar el barrido realizado para el MD1 siendo f el mapa logístico y g el mapa triangular, los parámetros de los mapas fueron $a = 0,5$ y $r = 3,99$, para $\mu < \epsilon$ predomina $T_{x \rightarrow y}$ y a medida que el parámetro μ va aumentando predomina $T_{y \rightarrow x}$ como es de esperarse.

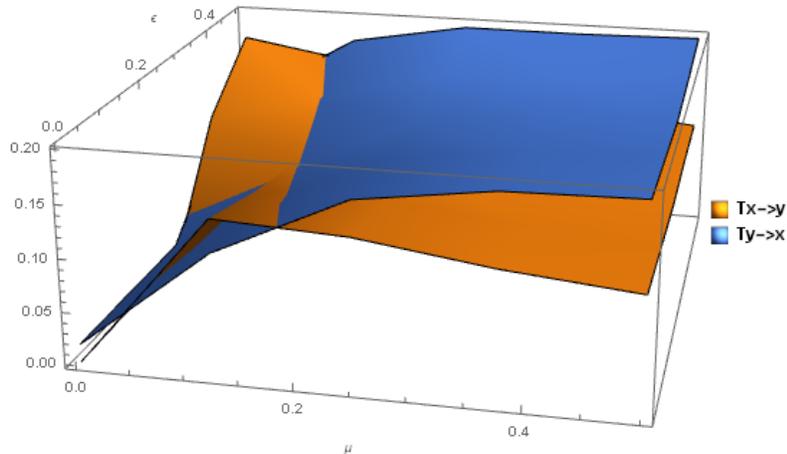


Figura 3.8: Barrido en los parámetros de acople μ y ϵ para el acople 1 con el MD2 y $m = 7$.

De forma similar, en la figura 3.8 se muestra el barrido en los parámetros de acople con las mismas características anteriores y para el MD2. Podemos observar que inicialmente

cuando los valores de μ y ϵ son pequeños y cercanos, a este método le cuesta distinguir cual es la serie que está enviando más información, para $\mu < 0,2$ predomina $T_{x \rightarrow y}$, a partir de $\mu \approx 0,2$ los valores de las transferencias se cruzan y predomina $T_{y \rightarrow x}$ y podemos observar que la distancia entre ambas transferencias aumenta.

3.3.2. Mapa de Henón

Con la finalidad de validar ambos métodos de discretización, se estimó la transferencia de entropía para datos generados de forma sintética con en el mapa de Henón (ver sección 2.5) utilizando los parámetros $a = 1,4$, $b = 0,3$, $M = 128$ y $m = 7$, de tal forma que intuitivamente ya se sabía en que dirección fluye la información.

Los resultados obtenidos fueron los siguientes:

	$T_{x \rightarrow y}$	$T_{y \rightarrow x}$	ΔT
MD1	2,03	0,82	1,21
MD2	2,46	1,33	1,13

Cuadro 3.1: Transferencia de Entropía obtenidas en el mapa de Hénon para MD1 y MD2.

Podemos observar en el cuadro 3.1 que en el mapa de Henón con los parámetros utilizados y en ambos métodos de discretización predomina $T_{x \rightarrow y}$, que era lo que se esperaba. Por otra parte, los valores obtenidos en el MD2 son mayores que los obtenidos en el MD1, sin embargo al calcular la diferencia entre las transferencias para cada método, es decir, $\Delta T = T_{x \rightarrow y} - T_{y \rightarrow x}$ podemos observar que estas son cercanas para ambos métodos de discretización.

3.4. Número de Operaciones

Otro aspecto importante en este trabajo será el número de operaciones necesarias para poder calcular la transferencia de entropía con ambos métodos de discretización (ecuaciones 2.11 y 2.12). Para esto debemos conocer cuántos procesos se realizan para hallar cada una de las probabilidades que se necesitan para dichas ecuaciones.

En el caso de $p(x_n)$ y $p(y_n)$ se necesitan aN procesos, siendo N la longitud de la data y a la cantidad de estados posibles en el sistema.

Para las probabilidades conjuntas $p(x_n, y_n)$, $p(x_{n+1}, x_n)$ y $p(y_{n+1}, y_n)$ deben recorrerse ambas series temporales buscando las posibles b combinaciones de ambos elementos, por lo tanto los procesos necesarios serán bN^2 .

Por último para $p(y_{n+1}, x_n, y_n)$ y $p(x_{n+1}, x_n, y_n)$ se necesitarán cN^3 siendo c el número de posibles combinaciones para cada una de estas probabilidades.

Por lo tanto, para hallar cada transferencia de entropía se necesitarán $S(aN + 2bN^2 + cN^3)$, con S los estados accesibles del sistema discretizado, donde $S \leq N$. Es decir, la transferencia de entropía será del orden de SN^3 procesos, por consiguiente el número de operaciones se verá incrementado al aumentar el número de datos. Podemos observar que en el caso de que se pudiera dividir S en N retículas, la transferencia sería del orden de N^4 , esto quiere decir que el número de retículas también influirá en el tiempo de cómputo.

Por último, el cálculo de la transferencia de entropía a través de la discretización del espacio de estados no es eficiente, ya que, puede perderse detalle de la dinámica del sistema y además, es muy costosa computacionalmente. Por lo cual, no se realizaron otras pruebas con mayor número de divisiones en ambos métodos o con otro tipo de datos.

Capítulo 4

Modelado de series temporales

Como se mencionó en el capítulo anterior el método de la discretización del espacio de estados necesario para el cálculo directo de las probabilidades presenta un gran costo computacional, por lo cual en este trabajo se implementó una estrategia basada en el modelado local, inspirado en la causalidad de Granger e implementado en la idea de Schreiber, que consiste en el modelado de la distribución de probabilidad condicional usada en la estimación de la transferencia de información (ver ecuaciones 2.13 y 2.14) mediante el método de los vecinos más cercanos [30].

4.1. Densidad de Probabilidad Condicional

Dada una serie temporal multivariada $\mathbf{x} = \{x_i\}_{i=1}^N$ proveniente de un sistema dinámico $f(x_i) = x_{i+1}$, sabemos que puede construirse un modelo aproximado de f a partir de los valores medidos de \mathbf{x} .

Así, dado el estado i -ésimo de los datos x_i y un conjunto de los k estados más cercanos $\{x_c^{(k)}\}_{k=1}^v$, si se desarrolla f en serie de Taylor, hasta el primer orden de aproximación, alrededor del estado más cercano a x_i (que es x_c),

$$x_{i+1} \cong f(x_c^{(1)}) + Df(x_c^{(1)}) \cdot (x_i - x_c^{(1)}) \quad (4.1)$$

para predecir x_{i+1} . Donde $Df(x_c)$ es la matriz Jacobiana de f .

De esta manera, en orden cero de aproximación la evolución del estado i -ésimo está dado por $f(x_c^{(1)}) = x_{i+1}$, es decir, el vector que resulta de la evolución del vecino más cercano a x_i .

Se ha demostrado en otros trabajos como en [31] que los errores de predicción se distribuyen de forma gaussiana.

Si tomamos k vecinos x_c cercanos a x_i siendo $|x_{i+1} - x_{c+1}^{(1)}|$ el mínimo error de predicción y $|x_{i+1} - x_{c+1}^{(k)}|$ el máximo error del modelado, estos pueden ser relacionados con la densidad de probabilidad condicional:

$$P(x_{i+1}|x_i) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x_{i+1} - f(x_i))^2}{2\sigma^2}} \quad (4.2)$$

donde $f(x_i)$ es una función que halla el promedio de la evolución de los k vecinos más cercanos, esta expresión ha sido propuesta en [31], donde se considera solamente el primer vecino más cercano, además que la desviación estándar es calculada de forma global, en este trabajo se utilizarán el promedio de los k vecinos más cercanos y la desviación estándar será tomada localmente en la vecindad k .

Entonces la ecuación 4.2 queda escrita de la siguiente manera:

$$P(x_{i+1}|x_i) = \frac{1}{\sigma_{i+1}\sqrt{2\pi}} e^{-\frac{\left(x_{i+1} - \frac{1}{k} \sum_{c=1}^k x_{c+1}\right)^2}{2\sigma_{i+1}^2}} \quad (4.3)$$

donde σ es la desviación estándar de los errores del modelado en la vecindad que contiene k vecinos dentro. Es decir,

$$\sigma_{i+1} = \sqrt{\frac{1}{(k-1)} \sum_{c=1}^k (|x_{i+1} - x_{c+1}| - \mu)^2} \quad (4.4)$$

donde μ será,

$$\mu_{i+1} = \frac{1}{k} \sum_{c=1}^k |x_{i+1} - x_{c+1}| \quad (4.5)$$

En la figura 4.1 podemos observar la distribución de los errores del modelado, donde el error para $P(x_{i+1}|x_i, y_i)$ (izquierda) es $x_{i+1} - \frac{1}{k} \sum_{c=1}^k x_{c+1}$ y para $P(x_i|y_i)$ (derecha) el error será $x_i - \frac{1}{k} \sum_{c=1}^k x_c$. Los histogramas de los errores fueron ajustados con una curva gaussiana, y podemos observar que para ambas distribuciones de probabilidades los parámetros de la gaussiana son prácticamente los mismos, esto quiere decir, que la propuesta de asumir que los errores se distribuyen de forma gaussiana es válida.

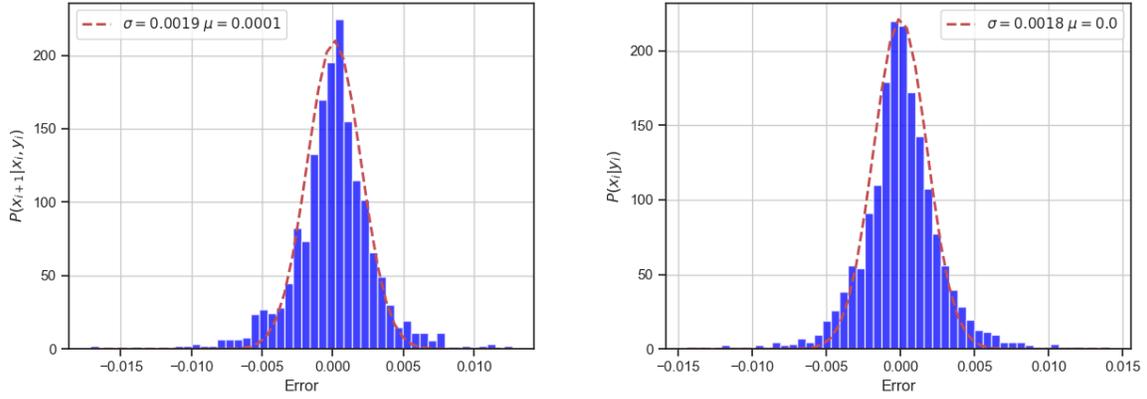


Figura 4.1: Distribución de los errores de predicción para $P(x_{i+1}|x_i, y_i)$ (izquierda) y $P(x_i|y_i)$ (derecha), ajustadas con una función densidad de probabilidad Gaussiana.

4.2. Densidad de Probabilidad Marginal

En el caso de la distribución de probabilidad marginal se utilizó un método de estimación en el que no se asume de antemano la forma de esta [32].

La posibilidad de que un nuevo valor x_i caiga dentro de una región del espacio R será:

$$p(x_i) = \int_R P(x') dx' \quad (4.6)$$

Si se dispone de N muestras se puede tener una buena aproximación de p a partir de la fracción media de muestras k que caen en R

$$p(x_i) \approx \frac{k}{N}$$

Si asumimos que $P(x_i)$ es continua y que no varía sobre R se puede aproximar la ecuación 4.6 por:

$$p(x_i) = P(x_i)V$$

donde V es el volumen de R y x_i es el patrón incluido en R , esto quiere decir que

$$\frac{k}{N} \approx P(x_i)V \Rightarrow P(x_i) \approx \frac{k/N}{V} \quad (4.7)$$

La ecuación 4.7 es la forma genérica para la estimación de una densidad de probabilidad cualquiera. Para esta expresión se puede elegir un V fijo y calcular k a partir de los datos o viceversa. En este caso se fijará el valor de k y se halla el valor de V , que contiene k puntos dentro, lo cual permite que V se escoja en función de la data, ya que este método centra

la celda en x_i y esta crece hasta calcular k muestras. k es llamado el k -ésimo vecino más cercano a x_i , por esto este método es llamado el del k -ésimo vecino más cercano. Según la ecuación 4.7 la función de densidad de probabilidad marginal será de la forma

$$P(x_i) = \frac{k}{N|x_i - x_c^{(k)}|} \quad (4.8)$$

4.3. Algoritmo de los Vecinos más Cercanos

Este algoritmo consiste en estimar el valor de un dato desconocido a partir de las características del dato más próximo, según una medida de similitud o distancia [33]. Esta regla tiene propiedades estadísticas bien establecidas y facilidad de aplicación a sistemas reales [34]. El método del vecino más cercano se puede extender utilizando no uno, sino un conjunto de datos más cercanos para predecir el valor de los nuevos datos, en lo que se conoce como los k -vecinos más cercanos (k -NN o k -Nearest Neighbors). Al considerar más de un vecino, se brinda inmunidad ante ruido y se suaviza la curva de estimación [35].

El método de los k -vecinos más cercanos se adapta fácilmente a la regresión de funciones con valores continuos [35]. El algoritmo asume que todos los datos pertenecen a R^N y mediante una medida de distancia (distancia euclídea) en ese espacio se determinan los k datos más cercanos al dato x_i para aproximar una función $f : R^N \rightarrow R$, a partir de los k valores ya seleccionados. Esta función corresponde al promedio de los k valores más cercanos; si se considera el promedio aritmético (todos los datos dentro del grupo tienen igual relevancia).

A continuación se explicará la estrategia utilizada a través de un ejemplo para mayor simplicidad.

En la ecuación 2.13 para hallar $T_{\mathbf{x} \rightarrow \mathbf{y}}$ se necesitan las siguientes probabilidades condicionales $P(y_{i+1}|x_i, y_i)$, $P(y_i|x_i)$ y $P(y_{i+1}|y_i)$. A su vez éstas serán expresadas de la siguiente manera (ver ecuación 4.2):

$$P(y_{i+1}|x_i, y_i) = \frac{1}{\sigma_1 \sqrt{2\pi}} e^{-\frac{(y_{i+1} - f_1(x_i, y_i))^2}{2\sigma_1^2}} \quad (4.9)$$

$$P(y_i|x_i) = \frac{1}{\sigma_2 \sqrt{2\pi}} e^{-\frac{(y_i - f_2(x_i))^2}{2\sigma_2^2}} \quad (4.10)$$

$$P(y_{i+1}|y_i) = \frac{1}{\sigma_3 \sqrt{2\pi}} e^{-\frac{(y_{i+1} - f_3(y_i))^2}{2\sigma_3^2}} \quad (4.11)$$

donde f_1 , f_2 y f_3 actúan de la siguiente forma:

- $f_1(x_i, y_i)$ evalúa la distancia euclídea d (es decir $d = \sqrt{(x_i - x_n)^2 + (y_i - y_n)^2}$) entre cada punto en ambas series temporales (x_n, y_n) y el punto (x_i, y_i) . Luego, toma los k puntos más parecidos a ese par, que llamaremos (x_c, y_c) . Y finalmente, la función f_1 devolverá el promedio de la evolución temporal de y_c (es decir, $\frac{1}{k} \sum_{c=1}^k y_{c+1}$). Funciona de forma similar con $P(x_{i+1}|x_i, y_i)$ (ecuación 2.14).
- $f_2(x_i)$ evalúa la serie temporal \mathbf{x} y halla los k valores más cercanos a x_i que serán llamados x_c . Finalmente devuelve el promedio de los y_c correspondientes a cada x_c , es decir $\frac{1}{k} \sum_{c=1}^k y_c$. De forma similar para $P(x_i|y_i)$ presente en la ecuación 2.14.
- Por último, $f_3(y_i)$ actúa de forma similar a f_2 pero en este caso halla los k valores más cercanos a y_i que serán llamados y_c en la serie temporal y devuelve el promedio de los y_{c+1} , es decir $\frac{1}{k} \sum_{c=1}^k y_{c+1}$. Y de igual manera para $P(x_{i+1}|x_i)$ en la ecuación 2.14.

Finalmente, σ_1 , σ_2 y σ_3 necesarias en las expresiones 4.9, 4.10 y 4.11, representan la desviación estándar de los errores del modelado en cada punto, en una vecindad con k vecinos y estas serán calculadas según la ecuación 4.4.

4.4. Experimentos Realizados

Para comprobar que este método para estimar las densidades de probabilidad condicional funciona correctamente se realizaron diferentes tipos de experimentos. Primero se probó el método con sistemas de mapas acoplados y después se compararon los resultados obtenidos en el capítulo anterior. Luego se procedió a implementar el método en el Mapa de Hénon y por último se aplicó a datos provenientes de sistemas reales.

4.4.1. Sistemas de Mapas Acoplados

De igual manera que en la sección 3, se empleó el método en datos generados sintéticamente a través de sistemas de mapas acoplados (ver sección 2.5) en los que podemos suponer de antemano como se comparte la información en el sistema.

Cada uno de los experimentos de esta sección se realizaron 10 veces para condiciones iniciales diferentes y se graficó el promedio de los valores obtenidos para las transferencias.

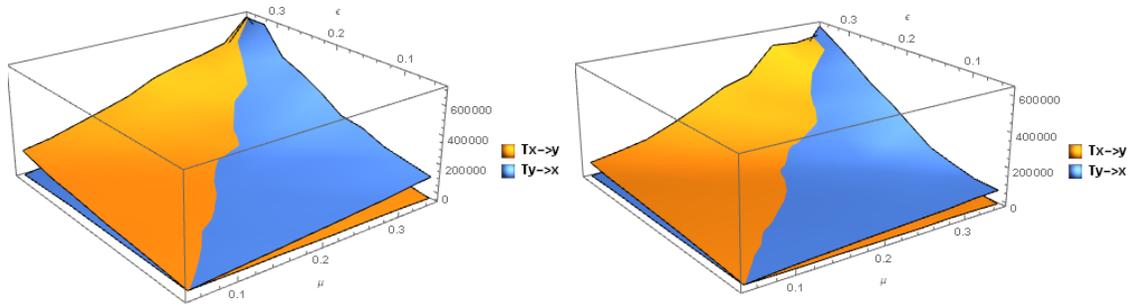


Figura 4.2: Barrido en los parámetros de acople para dos mapas triangulares con $a_1 = 0,63$ y $a_2 = 0,65$, con $k = 10$, acople 1 (izquierda), acople 2 (derecha).

En la figura 4.2 donde se ha considerado una vecindad de $k = 10$ vecinos, podemos observar en ambos acoples que si $\mu < \epsilon$ predomina $T_{x \rightarrow y}$ y si $\mu > \epsilon$ predomina $T_{y \rightarrow x}$, también podemos observar que cuando $\mu = \epsilon$ las transferencias se cruzan y existe una competencia entre ambas funciones. Otro aspecto importante que vale la pena destacar es que podemos observar que mientras mayor sea la diferencia entre los parámetros de acople, mayor será la diferencia entre ambas transferencias de entropía.

De igual manera en la figura 4.3 con $k = 12$, obtenemos un comportamiento similar al de la figura anterior pero los valores para la transferencia de entropía son menores que con $k = 10$ para el acople 2. Sin embargo hay que destacar que no se debe utilizar un número muy grande de vecinos, ya que esto puede disminuir la calidad de la predicción. De estos resultados, para dos mapas triangulares el método logra distinguir de forma eficaz la direccionalidad del acople entre los dos sistemas.

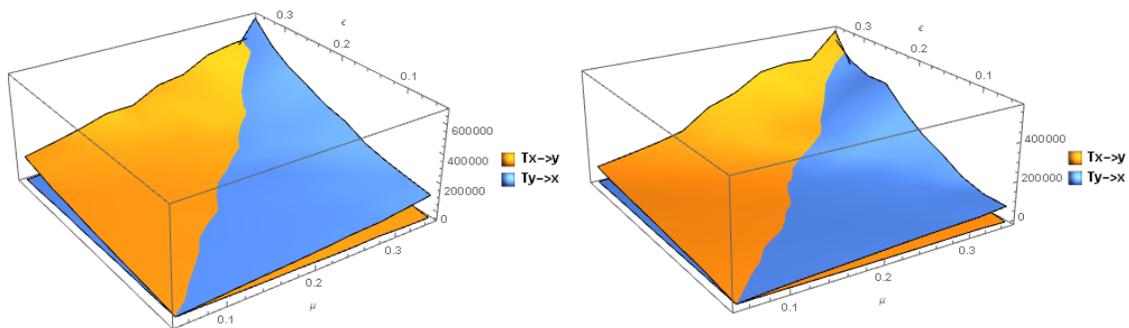


Figura 4.3: Barrido en los parámetros de acople para dos mapas triangulares con $a_1 = 0,63$ y $a_2 = 0,65$, con $k = 12$, acople 1 y acople 2 respectivamente.

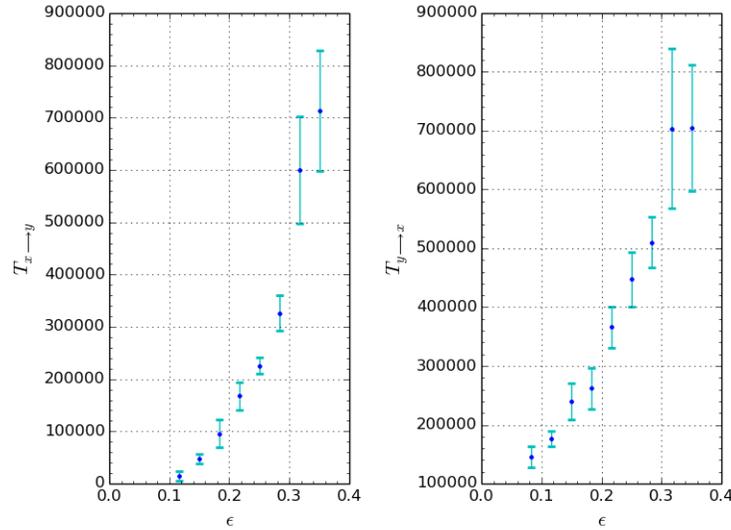


Figura 4.4: Desviación estándar para el promedio de la transferencias de entropía calculada 10 veces (4.2) con $\mu = 0,35$ para el acople 2.

Como se mencionó anteriormente, cada experimento de esta sección fue realizado 10 veces es por ello que se procedió a calcular la desviación estándar para cada punto (se puede observar un ejemplo de esto en la figura 4.4). Posteriormente se calculó el valor en porcentaje de estas desviaciones para cada valor y luego, se realizó el promedio de estos porcentajes para cada transferencia. Finalmente se obtuvo que las desviaciones se encontraban en el rango de 9% – 12% de los valores de la transferencia de entropía. Esto quiere decir que los valores no están tan alejados entre sí, por lo cual se puede decir que el método es confiable.

En la figura 4.5 se muestra un barrido en los parámetros de acople realizado con los mismos parámetros que en las figuras 3.7 y 3.8, pero en este caso se calculó la transferencia de entropía con la estrategia presentada en este capítulo. Cabe destacar que para los parámetros de los mapas empleados el mapa logístico es caótico (ver figura 2.3), mientras el mapa triangular tiene un comportamiento periódico [36]. Podemos observar que con funciones diferentes, al método le cuesta más distinguir la dirección del acople.

Por último, en la figura 4.6 podemos observar el mismo experimento anterior pero fue cambiado el parámetro a del mapa triangular a uno donde este presenta caos. En las figuras 4.5 y figura 4.6 para el acople 1 se puede apreciar que al método le cuesta un poco distinguir la direccionalidad del acople. Sin embargo, al observar la figura 4.6 para el acople 2 es fácil notar, que el método funciona mejor con este acople y en este caso logra distinguir mejor la dinámica del sistema.

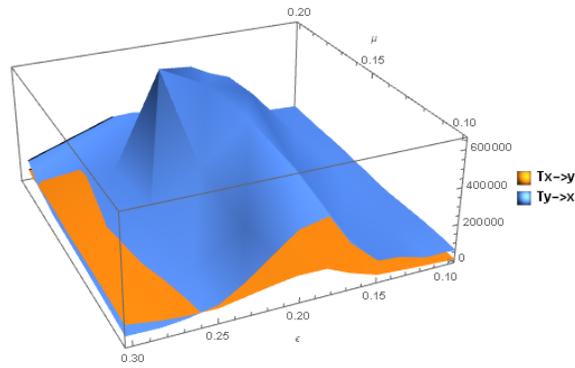


Figura 4.5: Barrido en los parámetros de acople para f el mapa logístico con $r = 3,99$ y g el mapa triangular con $a = 0,5$, con $k = 10$, acople 1.

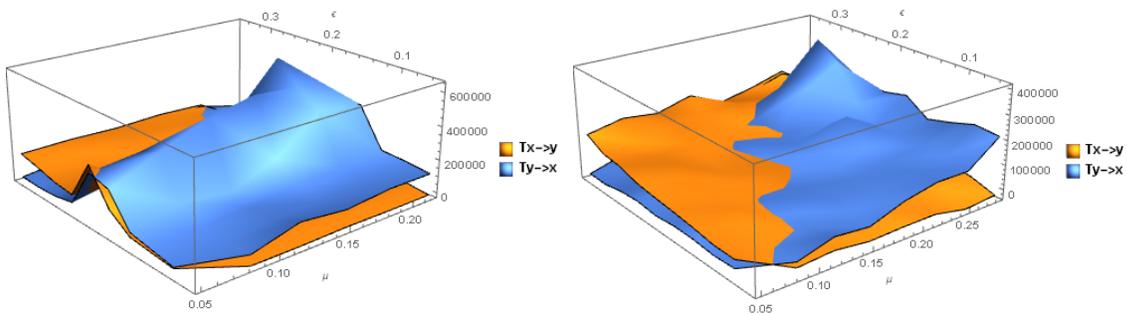


Figura 4.6: Barrido en los parámetros de acople para f el mapa logístico con $r = 3,99$ y g el mapa triangular con $a = 0,65$, con $k = 10$, acople 1 y acople 2 respectivamente.

4.4.2. Mapa de Hénon

En la sección 2.3.3 se presentó el mapa de Hénon, las ecuaciones que lo definen son las siguientes:

$$x_{n+1} = y_n + 1 - ax_n^2 \quad (4.12)$$

$$y_{n+1} = bx_n \quad (4.13)$$

Con estas se generaron datos de forma sintética utilizando los parámetros $a = 1,4$ y $b = 0,3$ [26], para así poder implementar el método expuesto en esta sección.

Para este sistema se obtuvieron los siguientes resultados:

$$T_{x \rightarrow y} = 14215263,40$$

$$T_{y \rightarrow x} = 280574,97$$

al comparar estos resultados con los obtenidos en la sección 3.3.2 se puede apreciar que en ambos casos predomina $T_{\mathbf{x} \rightarrow \mathbf{y}}$, lo cual era de esperarse, ya que en las ecuaciones 4.12 y 4.13 es fácil notar que \mathbf{x} comparte mas información con \mathbf{y} que viceversa.

En vista de los resultados obtenidos con datos generados de forma sintética se procedió a implementar el método en datos provenientes de sistemas reales, los primeros experimentos se llevarán a cabo en sistemas en los cuales hay reportes sobre la direccionalidad entre los componentes de estos usando transferencia de entropía, como el circuito de Chua y el sistema cardiovascular, y por último se implementará en el sistema cardio respiratorio en el cual no ha sido utilizada esta técnica, para determinar que sistema envía mas información.

4.4.3. Sistemas Reales

Circuito de Chua

El circuito de Chua ha sido estudiado extensamente desde su propuesta inicial [37], es uno de los modelos más populares que exhiben caos puesto que es el circuito autónomo más simple capaz de mostrar este comportamiento debido a que contiene 4 elementos no lineales: el diodo de Chua que básicamente es una resistencia no lineal y dos condensadores, una bobina y una resistencia. Por la riqueza en cuanto a su comportamiento, este circuito ha sido y es objeto de mucha investigación científica, convirtiéndose en un paradigma universal para el caos. El sistema de ecuaciones diferenciales que representa el sistema es:

$$\frac{dV_{C_1}}{dt} = \alpha[V_{C_2} - V_{C_1} - g(V_{C_1})] \quad (4.14)$$

$$R_{C_2} \frac{dV_{C_2}}{dt} = V_{C_1} - V_{C_2} + Ri_L \quad (4.15)$$

$$\frac{di_L}{dt} = -\beta V_{C_2} \quad (4.16)$$

donde $g(V_{C_1})$, es la función lineal a trozos que representa el diodo de Chua.

Los datos utilizados fueron obtenidos de la implementación del circuito (ver figura 4.7) en el sistema Educational Laboratory and Virtual Instrumentation Suite de National Instruments, NI ELVIS por sus siglas en inglés, con una tasa de muestreo de 5ms implementado en el Laboratorio de Control y Simulación de la Facultad de Ingeniería en Ciencias Aplicadas de la Universidad Técnica del Norte (Ecuador) [31].

Para la implementación experimental del circuito, se usaron amplificadores operacionales para simular el diodo de Chua y el inductor. La corriente en el inductor fue obtenida por medida indirecta.

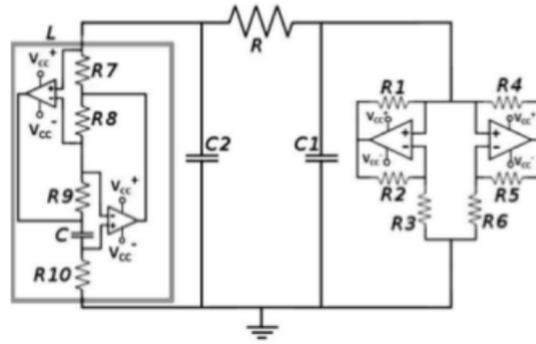


Figura 4.7: Circuito de Chua

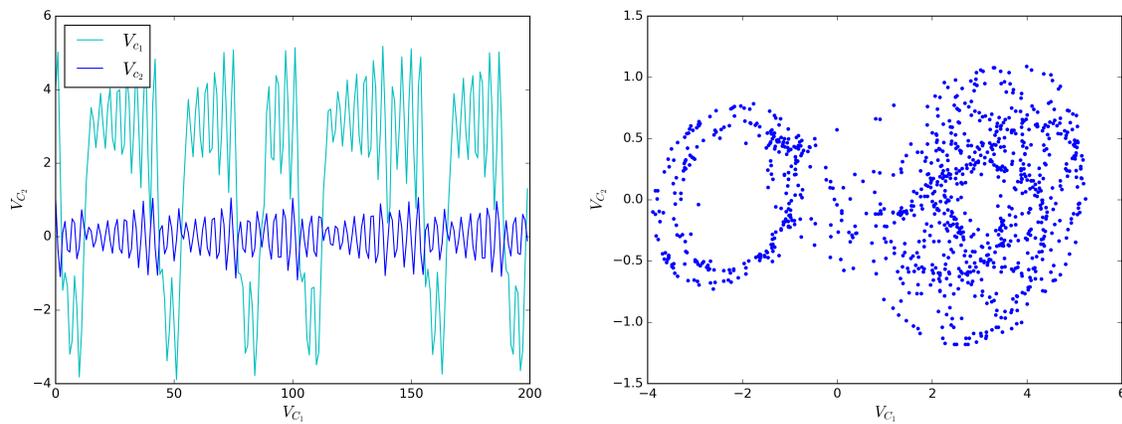


Figura 4.8: Series temporales y atractor de Chua.

En la figura 4.8 podemos observar las series temporales de los voltajes en los condensadores C_1 y en C_2 y en el lado derecho el atractor perteneciente a ambas series temporales, que es un atractor extraño, similar al atractor de Rössler [38, 39] y al de Lorenz [40].

En este caso, como se tienen 3 series temporales se emplearon las ecuaciones de la transferencia de entropía 2.13 y 2.14 por pares de series, como se muestra en la figura 4.9.

Se obtuvieron los siguientes resultados al calcular la transferencia de entropía por pares de series:

$T_{C_1 \rightarrow C_2}$	875,66	$T_{C_2 \rightarrow C_1}$	4857,75
$T_{C_2 \rightarrow i_l}$	16691,49	$T_{i_l \rightarrow C_2}$	14015,50
$T_{i_l \rightarrow C_1}$	4768,43	$T_{C_1 \rightarrow i_l}$	1443,76

Cuadro 4.1: Transferencia de Entropía por pares de series en el circuito de Chua

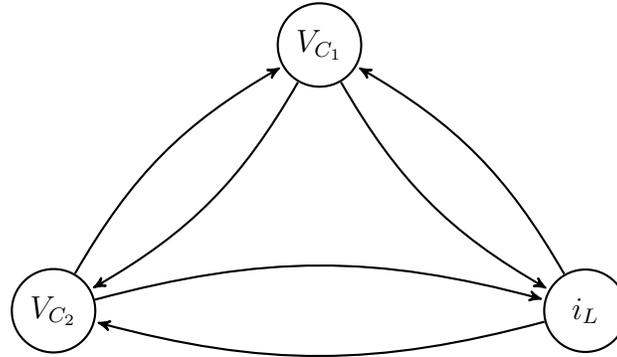


Figura 4.9: Diagrama para implementar la transferencia de entropía en el circuito de Chua.

Para hallar la transferencia de entropía neta correspondiente a cada serie temporal se realizó el siguiente procedimiento:

$$T_i = \sum_{j=1}^3 T_{i \rightarrow j} - T_{j \rightarrow i} \quad (4.17)$$

tomando en cuenta que, si $i = j$ $T_{i \rightarrow j} = T_{j \rightarrow i} = 0$ y siendo $T_1 = T_{C_1}$, $T_2 = T_{C_2}$ y $T_3 = T_{i_L}$.

Finalmente, la transferencia de entropía (según la Ec. 4.17) para cada serie temporal será:

T_{C_1}	T_{C_2}	T_{i_L}
-7306,76	6658,07	648,69

Cuadro 4.2: Transferencia de entropía total para cada elemento del circuito de Chua

Los resultados obtenidos sugieren que hay más influencia del capacitor 2 sobre el capacitor 1 y el inductor. Lo cual coincide con resultados obtenidos acerca del control del circuito de Chua [31, 37, 41], en los que se demuestra que el V_{C_2} es el que envía más información en el sistema.

Sistema Cardiovascular

En el sistema circulatorio humano, la contracción del corazón (o sístole) bombea la sangre, produciendo un pulso en la presión arterial. A continuación, ocurre una relajación del sistema, y el valor de la presión disminuye suavemente. Cuando ésta baja por debajo de cierto umbral, el sistema baro-reflejo envía un estímulo al corazón para que se efectúe un nuevo latido.

Por medio de este proceso, un aumento o una disminución, en la presión arterial promedio. Igualmente si por algún motivo la presión arterial es muy alta, o muy baja, la frecuencia cardíaca es regulada correspondientemente. Todo este mecanismo de control funciona de manera muy robusta para conservar el flujo y la irrigación sanguínea que son requeridos por el cuerpo, y es un buen ejemplo de un sistema complejo con interacciones bidireccionales.

Para este experimento se emplearon registros electrocardiográficos y de presión arterial, obtenidos a través de Physionet [42], esta base de datos contiene registros de sujetos en cuidados intensivos pero que se encontraban estables. Con estos datos se busca conocer la relación entre la serie temporal del ritmo cardíaco y la serie temporal de presión arterial medidos simultáneamente.

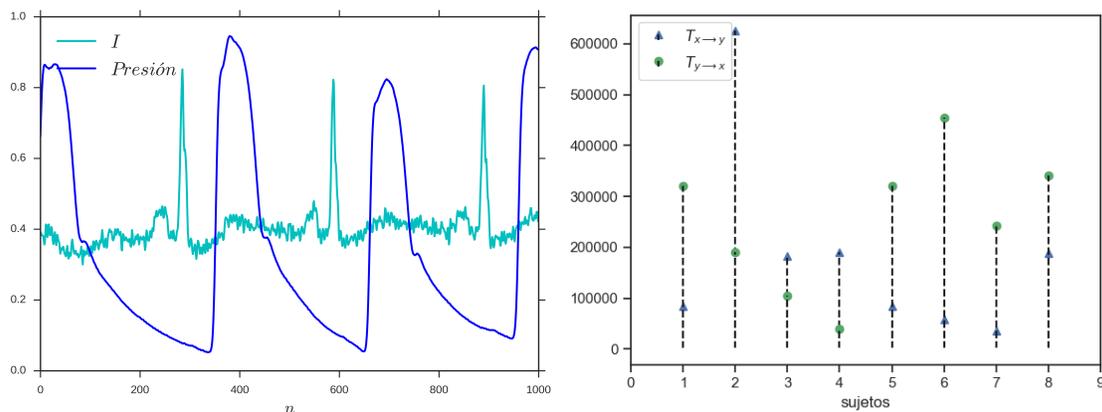


Figura 4.10: Series temporales de la primera derivación del ECG y de la presión arterial, transferencias de entropía para entre ECG y presión arterial respectivamente.

En la figura 4.10 podemos apreciar las series temporales (primera derivación del ECG (I) y presión arterial) pertenecientes a uno de los sujetos de estudio, a su lado podemos observar las transferencias de entropía, siendo $x=I$ e $y=Presión$. Para este experimento fueron considerados 8 sujetos de cuyos registros se tomó 1 minuto el cual corresponde a 10000 puntos. En la figura 4.10 derecha se puede observar que en el 63% de los sujetos la transferencia de entropía arrojó los resultados esperados, es decir, la serie de presión arterial envía mas información a la del ECG. Cabe destacar que este experimento fue realizado con la finalidad de comprobar que el método del modelado funciona en sistemas reales. Este sistema fue escogido, ya que, existen diversos trabajos en los cuales se ha estudiado la transferencia de información entre ambos sistemas, ver [30, 43, 44].

4.4.4. Sistema Cardio Respiratorio

Los sistemas cardiovascular y cardio respiratorio están caracterizados por una interacción compleja de varios subsistemas lineales y no lineales [45]. La interacción de estos subsistemas fisiológicos dentro del sistema cardiovascular pueden describirse como bucles cerrados con mecanismos de compensación y retroalimentación.

En 1733, el Rev. Stephen Hales reportó que la respiración modula la frecuencia cardíaca y la presión arterial. Esta observación fue confirmada por Carl Ludwing (1847), quien midió el incremento de la frecuencia cardíaca y la presión sanguínea en la inspiración.

Las interacciones dentro del sistema respiratorio están principalmente reflejadas en la arritmia sinusal respiratoria que es el aumento de la frecuencia cardíaca durante la inspiración y el aumento de la presión arterial. La frecuencia cardíaca y la presión arterial se modulan neurológicamente, y tanto los sistemas simpático (sistema encargado de las respuestas de estrés) como parasimpático (sistema encargado del reposo) tienen patrones de actividad respiratoria modulada [46, 47, 48, 49].

Múltiples factores, incluido el acoplamiento mecánico, subyacen a los aumentos en la frecuencia cardíaca y la presión arterial. El acoplamiento mecánico en el sistema cardio-respiratorio se debe a la ubicación de los pulmones y el corazón en la cavidad torácica. La inspiración se basa en una disminución de la presión pleural torácica que lleva la sangre al corazón y aumenta el retorno venoso, lo que aumenta la frecuencia cardíaca y el gasto cardíaco. Sin embargo, la frecuencia cardíaca y la presión arterial pueden aumentar durante la inspiración en una preparación en la que el tórax está muy abierto y se expanden los pulmones.

El acople cardiorespiratorio engloba diversos fenómenos que son resultados de entradas compartidas, ritmos comunes y funciones complementarias. Los ritmos del sistema nervioso autónomo (sistema nervioso simpático y parasimpático) y el sistema respiratorio se expresan en la actividad neuronal del otro [50]. Los latidos del corazón y la respiración están estrechamente vinculados, tanto funcional como anatómicamente. Dentro del sistema nervioso central, encontramos redes del tronco cerebral altamente superpuestas que controlan estas fisiologías autónomas. Esto lleva a concebir la interacción recíproca entre los sistemas de control respiratorio y el sistema nervioso autónomo en función del intercambio de gases (ver figura 4.11).

La figura 4.11 ha sido extraída de [51] y podemos observar un esquema en el que las flechas entre cada círculo; azul que representa al sistema respiratorio y roja correspondiente al cardíaco, representan el acoplamiento recíproco entre respiración y ritmos cardíacos, que se describen como armónicos. La relación entre el ritmo cardíaco y el respiratorio es 4:1 [51], esto podrá observarse mas adelante en las series temporales a

considerar. Múltiples mecanismos median la influencia respiratoria en el ciclo cardíaco, mientras que un solo mecanismo media la influencia de la actividad cardíaca simpática en el ciclo respiratorio. El mecanismo propuesto es a través de barorreceptores y su aumento de actividad latido a latido. Es decir, además de la influencia respiratoria bien conocida en la actividad autónoma, el sistema autónomo tiene una influencia en la formación del patrón respiratorio [52].

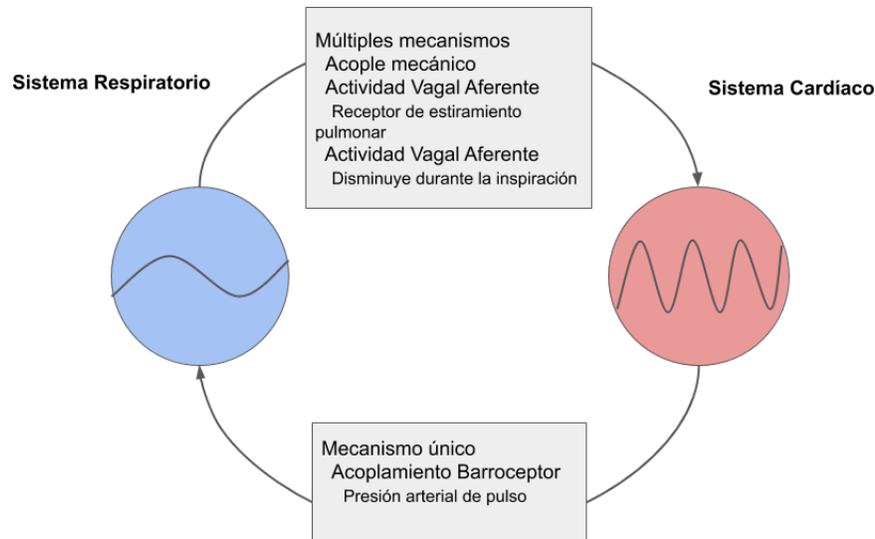


Figura 4.11: Esquema del acople Cardio Respiratorio.

La base de datos a utilizar fue desarrollada por Miguel Ángel García Gonzalez y Ariadna Argelagos Palou del grupo de investigación de Instrumentación Electrónica y Biomédica (IEB) del Departamento de Ingeniería Electrónica (DEE) de la Universitat Politècnica de Catalunya (UPC) y fueron obtenidos a través de Physionet [53], esta base de datos fue desarrollada con propósitos diferentes al de este trabajo [54, 55].

Para construir la base de datos, se consideraron 15 voluntarios presuntamente sanos. Durante la recolección de los datos, se pidió a los sujetos que estuvieran muy quietos en posición supina en una cama individual y despiertos.

Los datos fueron obtenidos utilizando un sistema de adquisición de datos Biopac MP36 (Santa Barbara, CA, EE. UU.). Los canales 1 y 2 del sistema se dedicaron a medir el electrocardiograma (ECG) convencional que es la representación gráfica de la actividad eléctrica del corazón en función del tiempo (derivaciones I (mide la diferencia de potencial entre el electrodo del brazo derecho y el izquierdo) y II (del brazo derecho a la pierna izquierda) respectivamente) y el canal 3 se empleó para medir la señal respiratoria obtenida de una banda piezorresistiva torácica (sensor SS5LB de Biopac, Santa Barbara, CA, EE.

UU.).

Después de conectar los sensores, se registró el estado basal de los sujetos durante 5 minutos (registros b001 a b015). Después de eso, los sujetos comenzaron a escuchar música clásica durante aproximadamente 50 minutos (registros m001 a m015). Finalmente, todos los sujetos fueron monitoreados 5 minutos más después de que la música terminó (registros p001 a p015) [53].

Cada una de las series temporales representa 1 minuto de registro del electrocardiograma (I) y de la frecuencia respiratoria que corresponde a 300000 datos, los cuales fueron submuestreados para obtener 3000 datos (es decir $n = 100$, es decir $300000/n = 3000$), esto fue debido a que la resolución de los sensores utilizados es muy alta y muchos datos se repetían, lo que ocasionaba que si se escogía una vecindad pequeña (ver secciones 4.1 y 4.2) la distancia entre todos los vecinos más cercanos era cero al considerar todos los datos, lo cual representaba un problema para el cómputo de la transferencia de información. En la figura 4.12 podemos observar primero, las series temporales de uno de los registros considerados y junto a esta las mismas series submuestreadas mostrando que no se pierde la dinámica del sistema y que se conserva la relación entre el ritmo cardíaco y respiratorio, tal como se mencionó anteriormente y se mostró en el diagrama de la figura 4.11.

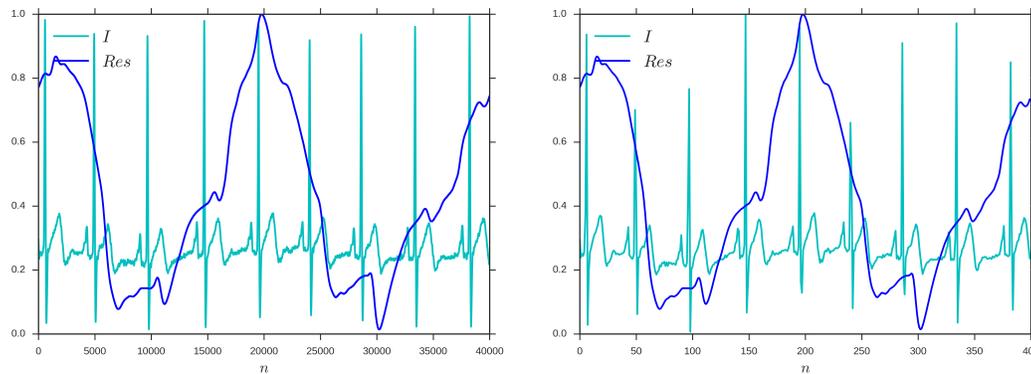


Figura 4.12: Series temporales electrocardiográfica (derivación I) y de frecuencia respiratoria de los datos sin submuestrear y submuestreados ($n = 100$), respectivamente.

Con la base de datos mencionada anteriormente se implementó el método del modelado de series temporales con el mismo número de vecinos considerado en los experimentos anteriores, para estimar las probabilidades necesarias en la transferencia de entropía, los resultados obtenidos son expresados en las figuras 4.13 y 4.14. El principal aporte de este trabajo es precisamente identificar que sistema envía mas información a través de la transferencia de entropía.

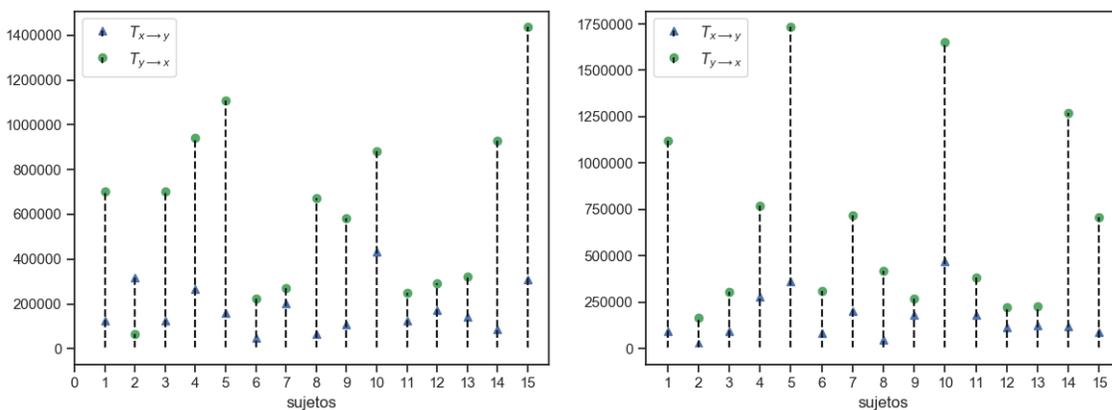


Figura 4.13: Transferencia de entropía entre la primera derivación del ECG (I) y la frecuencia respiratoria (RES) para el estado basal y durante los 50 minutos de música clásica respectivamente.

Al comparar la figura 4.13 (izquierda) que contiene el estado inicial de los sujetos con la figura 4.13 (derecha) que corresponde a un minuto de los 50 minutos de música clásica, podemos observar que los valores de las transferencias para la figura de la derecha son mayores que en el estado inicial.

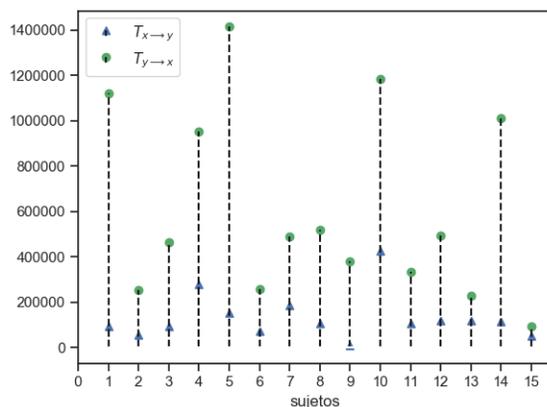


Figura 4.14: Transferencia de entropía entre la primera derivación del ECG (I) y la frecuencia respiratoria (RES) después de los 50 minutos de música.

Por otra parte, al comparar ambas figuras con la figura 4.14 que es posterior a los 50 minutos de música clásica podemos observar que los valores obtenidos para la transferencias son mayores que los obtenidos en el estado inicial, pero a su vez menores que los obtenidos durante los 50 minutos de música clásica.

Podemos observar que los comportamientos obtenidos en las tres etapas (b, m y p) son los mismos, es decir, se evidencia claramente en los resultados que ambos sistemas (cardíaco y respiratorio) comparten información. Pero, el sistema que envía más información es el sistema respiratorio, lo cual, concuerda con los resultados obtenidos en [51, 52].

4.5. Número de Operaciones

Al igual que en la discretización del espacio de estados (ver sección 3.4), se estimará el número de operaciones necesarias para calcular la transferencia de entropía con este método.

Como se mencionó anteriormente, en esta sección se utilizaron las ecuaciones de la transferencia de entropía escritas en función de las probabilidades condicionales y marginales (ver ecuaciones 2.13 y 2.14).

Para cada una de las probabilidades necesarias se realizan la misma cantidad de operaciones. Primero se recorre la serie temporal correspondiente una vez para calcular las distancias, luego una segunda vez para organizar cada uno de los valores de menor a mayor y por último se toman k valores de cada serie, es decir, por cada probabilidad se necesitan $2N + k$ procesos, por lo tanto para cada transferencia de entropía se necesitan $4N(2N + k)$.

Al comparar el número de operaciones necesarias para obtener la transferencia de entropía con el método de la discretización (del orden de N^4) y con el método del modelado (del orden de N^2) podemos observar que para el método de la discretización el número de operaciones es dos ordenes de magnitud mayor, esta diferencia puede verse reflejada en el tiempo de cómputo de ambos métodos, lo cual es lo que se necesita disminuir ya que, la finalidad es poder trabajar con datos reales y estos en su mayoría poseen gran cantidad de datos.

4.6. Comparación del Tiempo de Cómputo.

Finalmente, para comparar la eficiencia de los 2 métodos implementados en este trabajo, se realizó el mismo experimento con el MD1 (ya que fue el que arrojó mejores resultados) y con el método del modelado, y se midió el tiempo de cómputo en cada uno.

Para comparar la eficiencia de ambos programas se midió el **tiempo real transcurrido** (o simplemente el tiempo real) que es el tiempo tomado desde el inicio de un programa de computadora hasta su finalización según lo medido por un reloj ordinario. El tiempo real transcurrido incluye el tiempo de la entrada/salida y todos los otros tipos

de esperas incurridas por el programa, este fue medido a través de la implementación de la función `time` perteneciente a la librería `time` de Python3.

El experimento fue realizado con datos generados de forma sintética con el acople 1, donde fueron escogidos f el mapa logístico, g el mapa triangular, $\mu = 0,6$, $\epsilon = 0,2$, $N = 3000$ y con un transiente de 500. Los parámetros de los mapas fueron para el mapa logístico $r = 3,99$ y para el triangular $a = 0,65$.

Con el MD1 se utilizaron $M = 128$ divisiones de \mathbf{s} y con el método del modelado se utilizó $k = 10$.

	MD1	Modelado
Tiempo (min)	75,39	19,37

Cuadro 4.3: Comparación del tiempo de cómputo real para el cálculo de la transferencia de entropía con los 2 métodos implementados.

En la tabla 4.3 podemos observar que el tiempo de cómputo obtenido con el MD1 fue aproximadamente cuatro veces mayor que el obtenido con el método del modelado. Eso quiere decir, que el cálculo de la transferencia de entropía a través del método del modelado de series temporales es beneficioso en diversos aspectos, por ejemplo, como el tiempo de cómputo es menor pueden utilizarse series temporales mas largas.

Capítulo 5

Conclusiones

Al finalizar este trabajo se obtuvieron las siguientes conclusiones:

- En ambos métodos de discretización del espacio de estados (MD1 y MD2) los valores obtenidos de la transferencia de entropía no dependen de la longitud de las series temporales utilizadas.
 - El MD1 es capaz de distinguir la dirección del flujo de información de forma mas precisa que el MD2.
 - El cálculo de la transferencia de entropía a través de la discretización del espacio de estados no es eficiente ya que, implica un costo computacional muy alto y si no se cuentan con las herramientas computacionales necesarias y si las series temporales a estudiar son largas, como las pertenecientes a los sistemas reales, esto requerirá un tiempo de cómputo mayor. Además de que se puede perder detalle de la dinámica del sistema si no se hacen las divisiones del espacio de estado necesarias.
 - El método para estimar las probabilidades necesarias para la transferencia de entropía mediante la implementación del método del modelado que emplea el algoritmo de los vecinos más cercanos es eficiente y logra detectar la dirección del acople tanto como para series temporales generadas sintéticamente como para series reales.
 - En el sistema cardio respiratorio, el sistema que comparte mayor información es el sistema respiratorio.
 - El número de operaciones necesarias para el cálculo de la transferencia de entropía a través de el método del modelado es dos ordenes de magnitud menor que con ambos métodos de discretización del espacio de estados.
-

- El tiempo de cómputo real empleado en el cálculo de la transferencia de entropía para el MD1 es aproximadamente cuatro veces mayor que el empleado en el método del modelado.

El estudio de la transferencia de entropía en este trabajo fue realizado de forma exploratoria, es decir, solo se esperaba ver la direccionalidad del acople en los sistemas. Por otra parte, los resultados demuestran que el método propuesto en este trabajo para el cálculo de la transferencia de entropía con el método del modelado de series temporales, es eficiente y confiable.

Bibliografía

- [1] P. Duan, F. Yang, S. Shah, and T. Chen. Transfer zero-entropy and its application for capturing cause and effect relationship between variables. *IEEE Transactions on Control Systems Technology*, 2015.
 - [2] S. Ito, M. Hansen, R. Heiland, A. Lumsdaine, A. Litke, and J. Beggs. Extending transfer entropy improves identification of effective connectivity in a spiking cortical network model. *PLoS one*, 2011.
 - [3] B. Gourévitch and J. Eggermont. Evaluating information transfer between auditory cortical neurons. *Journal of Neurophysiology*, 206.
 - [4] O. Stetter, D. Battaglia, J. Soriano, and T. Geisel. Model-free reconstruction of excitatory neuronal connectivity from calcium imaging signals. *PLoS Comput Biol*, 2012.
 - [5] J. Runge, J. Heitzig, V. Petoukhov, and J. Kurths. Escaping the curse of dimensionality in estimating multivariate transfer entropy. *Physical review letters*, 2012.
 - [6] M. Chvosteková D. Coufal N. Jajcay A. Krakovská, J. Jakubík and M. Paluš. Comparison of six methods for the detection of causality in a bivariate time series. *PHYSICAL REVIEW E*, 97, 2018.
 - [7] T. Schreiber. Measuring Information Transfer. *Physical review letters*, page 85:461, 2000.
 - [8] M. Cosenza. *Mecánica Clásica*. Universidad de Los Andes, 2016.
 - [9] D. Luenberger. *Introduction to Dynamic Systems: Theory, Models, and Applications*. Wiley, 1979.
 - [10] J. Conway. The game of life. *Scientific American*, 1970.
-

- [11] K. Alligood, T. Sauer, and J. Yorke. *An Introduction to Dynamical Systems*. Springer-Verlag, 1996.
 - [12] R. May. Simple Mathematical Models with Very Complicated Dynamics. *Nature*, 261:459–467, 1976.
 - [13] W. Melo and V. van Strien. *One-Dimensional Dynamics*. Springer-Verlag, 1993.
 - [14] P. Collet and J. Eckmann. *Iterated Maps on the Interval as Dynamical Systems*. BIRKHAUSER, 1980.
 - [15] K. Kaneko. Lyapunov analysis and information flow in coupled map lattices. *Physica*, 1986.
 - [16] J. Crutchfield and K. Kaneko. Directions in chaos. *World Scientific*, 1987.
 - [17] E. Ott. *Chaos in dynamical systems*. Cambridge University Press, 1993.
 - [18] G. Pereira. *Probabilidad y Estadística Matemática un Primer Encuentro*. Falta, 2001.
 - [19] C. Shannon. Mathematical Theory of Communication. *Bell System Technical*, 27(3):379–423, 1948.
 - [20] G. Klir. *Uncertainty and Information Foundations of Generalized Information Theory*, volume 1. OHN WILEY & SONS, INC, 2006.
 - [21] T. Cover, J. Thomas, and J. Wiley. *Elements of information theory*, volume 306. Wiley Online Library, 1991.
 - [22] N. Wiener. The theory of prediction. *McGraw-Hill*, Vol 1, 1956.
 - [23] C. Granger. Investigating causal relations by econometric models and crossspectral methods. *Econometrica*, 37(3):424–438, 1969.
 - [24] J. McCracken. *Exploratory Causal Analysis with Time Series Data*. Morgan & Claypool Publishers, 2016.
 - [25] M. Hasler and Y-L. Maistrenko. An Introduction to the Synchronization of Chaotic Systems: Coupled Skew Tent Maps. *IEEE Transactions on Circuits and Systems : Fundamental Theory and Applications*, 44(10):856–866, 1997.
 - [26] M. Hénon. A two-dimensional mapping with a strange attracto. *Communications in Mathematical Physics*, 1976.
-

-
- [27] N. Boccara. *Modeling Complex Systems*. Springer-Verlag, 2004.
- [28] S. Strogatz. *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering*. Perseus Books, 1994.
- [29] S. Jalan and R. Amritkar. Self-organized and driven phase synchronization in coupled maps. *Physical review letters*, 90:1, 2003.
- [30] P. García and R. Mujica. A Local Approach to Information Transfer. *Open Systems & Information Dynamics*, 2017.
- [31] P. García, L. Rosa, J. Chiza, P. Méndez, C. Otero, Y. Fernández, A. Silva, and R. Mujica. Transferencia de información en redes basada en modelado. *Editorial UTN*, 21:223–229, 2016.
- [32] J. Parzen. On estimation of a probability density function and mode. *The Annals of Mathematical Statistics*, 1962.
- [33] T. Cover and P. Hart. Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, 1967.
- [34] D. Aha, D. Kibler, and M. Albert. Instance-based learning “instance-based learning algorithms. *Machine Learning*, 1991.
- [35] F. Moreno. Clasificadores eficaces basados en algoritmos rápidos de búsqueda del vecino más cercano. *Ph.D. dissertation, Universidad de Alicante. Departamento de lenguajes y sistemas informáticos.*, 2004.
- [36] W. Hassan and M. Abdulkawi. Some dynamical properties of the family of tent maps. *Int. Journal of Math. Analysis*, 2013.
- [37] T. Matsumoto. A chaotic attractor from chua’s circuit. *IEEE Trans. Circuits Syst*, 1984.
- [38] O. Rössler. An equation for continuous chaos. *Physics Letters*, 1976.
- [39] O. Rössler. An equation for hyperchao. *Physics Letters*, 1979.
- [40] E. Lorenz. Deterministic nonperiodic flow. *J. Atmos. Sci*, 1963.
- [41] G. Chen. On some controllability conditions control for chaotic dynamics. *Chaos*, 1997.
-

- [42] J. Welch, P. Ford, R. Teplick, and R. Rubsamen. The mgh/mf waveform database. *Physionet*, 1992.
- [43] L. Cisneros and J. Jiménez. Medidas del flujo de información e interdependencia no lineal. *Revista Mexicana de Física*, 2003.
- [44] A. Montalto, L. Faes, and D. Marinazzo. A matlab toolbox to compare established and novel estimators of the multivariate transfer entropy. *PLoS ONE*, 2014.
- [45] A. Voss, S. Schulz, R. Schroeder, M. Baumert, and P. Caminal. Methods derived from nonlinear dynamics for analysing heart rate variability. *Phil. Trans. R. Soc.*, 2009.
- [46] D. Baekey, T. Dick, and J. Paton. Pontomedullary transection attenuates central re- spiratory modulation of sympathetic discharge, heart rate and the baroreceptor reflex in the in situ rat preparation. *Exp. Physiol.*, 2008.
- [47] D. Baekey, Y. Molkov, J. Paton, I. Rybak, and T. Dick. Effect of baroreceptor stimulation on the respiratory pattern: insights into respiratory-sympathetic interactions. *Respir. Physiol. Neurobiol.*, 2010.
- [48] T. Dick, D. Baekey, J. Paton, B. Lindsey, and K Morris. Cardio-respiratory coupling depends on the pons. *Respir. Physiol. Neurobiol.*, 2009.
- [49] C. Julien, M. Parkes, S. Tzeng, P. Sin, P. Ainslie, P. Van de Borne, J. Fortrat, M. Cus- taud, C. Gharib, A. Porta, F. Vallais, G. Baselli, M. Pagani, D. Lucini, R. Hugh- son, J. Taylor, C. Tan, D. Baekey, T. Dick, J. Paton, and B. Taha. Comments on point:counterpoint: respiratory sinus arrhythmia is due to a central mechanism vs. respiratory sinus arrhythmia is due to the baroreflex mechanism. *J. Appl. Physiol.*, 2009.
- [50] T. Dick, R. Shannon, B. Lindsey, S. Nuding, L. Segers, D. Baekey, and K. Morris. Arterial pulse modulated activity is expressed in respiratory neural output. *J. Appl. Physiol.*, 2005.
- [51] T. Dick, Y. Hsien, R. Dhingra, D. Baekey, R. Galán, R. Wehrwein, and K. Morris. Cardiorespiratory coupling: Common rhythms in cardiac, sympathetic, and respira- tory activities. *Elsevier*, 2014.
- [52] S. Schulz, F. Adochiei, I. Edu, R. Schroeder, H. Costin, K. Bär, and A. Voss. Car- diovascular and cardiorespiratory coupling analyses: a review. *Phil. Trans. R. Soc.*, 2013.
-

-
- [53] A. Goldberger, L. Amaral, L. Glass, J. Hausdorff, P. Ivanov, R. Mark, J. Mietus, G. Moody, C. Peng, and H. Stanley. Physiobank, physiotoolkit, and physionet: Components of a new research resource for complex physiologic signals. *circulation* 101(23):e215-e220 [circulation electronic pages; <http://circ.ahajournals.org/cgi/content/full/101/23/e215>]. 2000.
- [54] M. García, A. Argelagós, M. Fernández, and J. Ramos. A comparison of heartbeat detectors for the seismocardiogram. *Computing in Cardiology Conference (CinC)*, 2013.
- [55] M. García, A. Argelagós, M. Fernández, and J. Ramos. Differences in qrs locations due to ecg lead: Relationship with breathing. *XIII Mediterranean Conference on Medical and Biological Engineering and Computing 2013, IFMBE Proceedings*, 2014.
-