Universidad Nacional de Mar del Plata - Facultad de Ingeniería
Departamento de Ingeniería Electrónica

CAOS APLICADO A SISTEMAS DE ESPECTRO ESPARCIDO

Por
Luciana De Micco

Tesis Doctoral para optar al grado académico de Doctor en Ingeniería, orientación Electrónica

Directores de Tesis

Dra. Hilda Ángela Larrondo
Dr. Constancio M. Arizmendi

Mar del Plata, Argentina. Diciembre de 2008.

a mi hija Malena

Índice General

Abstract

Spread Spectrum (SS) Systems are the subject of these thesis. In this systems the useful signal is masked to get a broad band spectrum, it means a spectrum wider than that of the signal. This masking procedure uses a pseudo random digital signals (pseudo noise) and may be motivated by different reasons. Each particular case has specific statistical requirements to be fulfilled by the pseudo noise.

In this thesis chaos is used to generate the pseudo random noise. Three application fields are mainly studied: digital communications (spread spectrum CDMA), improvement of electromagnetic compatibility (EMC) and digital filtering in power sources.

Agradecimientos

Agradezco a mis directores de tesis Hilda Larrondo y Miguel Arizmendi por su inestimable apoyo, dedicación y sobre todo por ser un ejemplo a seguir.

También a mis compañeros de laboratorio Gustavo Zabaleta y Daniel Zarlenga. A los profesores que tanto me ayudaron Claudio González, Jorge Castiñeira Moreira y Mónica Liberatori.

A todos los excelentes profesores del Departamento de Electrónica, y al Mg. Manuel González, que lideró ese departamento durante el desarrollo de mi doctorado.

A Roxana Vandenberghe por su buena onda y su eficiencia.

También a mis compañeros de ayudantía E. Blotta, M. Revuelta, J. Domecq, P. Marcone, D. Petruzzi y J. Bonadero.

Agradezco a mi mamá por ayudarme a cuidar a Malena ... y a Malena por dejarme escribir la tesis. A mis hermanas a toda mi familia y amigos.

Mar del Plata, Argentina — Luciana De Micco

Diciembre de 2008

Síntesis de la Tesis

0.1. Caos y Espectro Esparcido

Los sistemas de espectro esparcido (SS de Spread Spectrum), objeto de estudio de esta tesis, son una clase de sistemas digitales en los que las señales útiles están enmascaradas, de modo que la señal resultante cubre un espectro mucho más amplio en el dominio de la frecuencia, que el de la señal útil original. Varias son las razones que llevan a utilizar este tipo de enmascaramiento y cada una de ellas exige propiedades diferentes de la señal que se emplea para realizar el proceso.

En particular en esta tesis doctoral se investigan tres campos de aplicación: el de las comunicaciones digitales, el de la mejora de la interferencia electromagnética y el del filtrado digital de señales de potencia. La denominación señal útil se refiere entonces a un mensaje, si se trata de un sistema de comunicaciones, a una señal de clock, si se trata de un problema de interferencia electromagnética o a la señal que define los tiempos de muestreo, si se trata de un muestreo no uniforme en un proceso de filtrado.

El énfasis de esta tesis está puesto en la utilización del caos determinista en estas aplicaciones. El caos determinista ha representado desde hace unos 20 aõ una revolución en campos básicos y aplicados. Esa revolución se produjo como consecuencia de estudios pioneros de Lorenz, Feigenbaum y otros investigadores, y a partir del uso de

las computadoras para la solución numérica de sistemas dinámicos. La razón que hace al caos determinista un fenómeno de tanta trascendencia es que quedó demostrado que el paradigma de considerar que las señales estocásticas debían ser generadas por sistemas complejos es falso: esas mismas señales pueden ser generadas por sistemas deterministas muy simples.

En el caso de un sistema determinista cada una de las variables de estado tiene una evolución predecible que surge de la solución de las ecuaciones diferenciales o discretas que constituyen el modelo matemático del sistema. En el caso de un sistema estocástico se estudia la evolución temporal de distribuciones de probabilidades de las variables del sistema y, en el mejor de los casos, se cuenta con ecuaciones estocásticas que modelan esa evolución temporal. Cuando el modelo se prevé como muy complejo o muy difícil de obtener, es común un modelado estocástico. De allí el paradigma mencionado en el párrafo anterior.

El caos determinista es entonces un fenómeno intermedio entre un sistema determinista y uno estocástico, en el que el modelo es simple y conocido, pero sin embargo la inestabilidad de las soluciones y la resolución en aritmética discreta (¡el número real no existe!) impiden obtener la evolución temporal de las variables de estado para tiempos largos. La capacidad de predicción se reduce a tiempos muy cortos. Esta propiedad conocida como "sensibilidad a las condiciones iniciales" es la clave de la apariencia estocástica que presentan las señales generadas por sistemas caóticos deterministas.

A medida que este fenómeno se fue comprendiendo con mayor profundidad y que mayor número de problemas físicos pudieron ser explicados con modelos caóticos, surgió el interés de utilizar el caos determinista en la ingeniería. Se lo puede, en principio,

aplicar a todas aquellas situaciones en las que se emplean señales estocásticas.

Así es como surge el uso del caos en diversos métodos de modulación y encriptado de los sistemas de comunicaciones electrónicos, tanto analógicos como digitales. También aparece el caos como fuente de ruido.

Si bien un sistema caótico es modelado por un sistema dinámico continuo o discreto (también llamado mapa), desde el punto de vista estadístico es posible suponer una distribución de condiciones iniciales de las variables de estado y estudiar su evolución temporal. La evolución temporal de las distribuciones podrá ser también discreta o continua. Para obtener una forma explícita de esa evolución es necesario determinar el operador de Perron-Frobenius asociado al sistema dinámico.

La teoría cualitativa de las ecuaciones diferenciales es otra herramienta matemática importante para poder realizar un análisis de las soluciones de un sistema caótico. En ese caso se trata de obtener los estados finales estacionarios del sistema en forma cualitativa (puntos fijos, ciclos límites, atractores caóticos, etc.)

A la tan útil transformada de Fourier se le deben adicional nuevas herramientas basadas en las propiedades particulares de los sistemas caótico. Los exponentes de Lyapunov, las dimensiones fractales, la función correlación no lineal, los diagramas de recurrencia, las medidas de complejidad estadística, las entropías, son ejemplos de métodos típicos de sistemas no lineales.

0.2. Espectro esparcido

Los fundamentos teóricos sobre el uso de la técnica de espectro esparcido son conocidos desde hace tiempo, pero recién en los últimos años se han vuelto factibles las implementaciones prácticas. Inicialmente, las implementaciones eran excesivamente

costosas y por lo tanto eran utilizadas exclusivamente para fines militares.

Pero gracias a nuevos avances tecnológicos y a las técnicas de procesamiento de señales se hizo posible el desarrollo de equipamientos más económicos para uso civil.

Un sistema SS es, como se indicó al inicio, un sistema que produce una señal de salida con un ancho de banda mucho mayor que el ancho de banda de la señal de entrada. Como el sistema SS distribuye la energía de la señal sobre un ancho de banda mayor, la relación señal a ruido resultante es baja. Sin embargo, es posible revertir el efecto del esparcimiento y recuperar la señal útil correlacionando la señal recibida con una réplica del código PN (código pseudoaleatorio, PN de pseudo noise). Por lo tanto, esto será posible siempre que se puede reproducir la forma de onda que esparció la señal. Como este código es, de hecho, una señal determinista, puede ser reproducido.

En resumen, todos los sistemas de SS deben satisfacer dos criterios básicos:

- El ancho de banda de la señal final debe ser mucho mayor que el de la señal original.
- Este ancho de banda debe ser determinado por alguna función conocida, independiente de la señal útil y además, para que sea posible la recuperación del mensaje, debe ser conocida por transmisor y receptor.

0.3. Espectro Esparcido en Comunicaciones

La primera técnica de SS desarrollada en comunicaciones digitales fue la llamada técnica de salto de frecuencia (FHSS por Frequency-Hopping Spread Spectrum). La idea principal es variar la frecuencia de la información según una secuencia (código

PN). Para recuperar el mensaje el receptor debería cambiar de frecuencias en sincronía con el transmisor. El mensaje es recuperado únicamente cuando la secuencia de frecuencias que se utilizó en la modulación es conocida.

Otro método de SS fue la técnica de secuencia directa (DSSS por Direct Sequence Spread Spectrum). En este caso, la información es multiplicada por un código PN y nuevamente, sólo será posible recuperar correctamente la información, si se utiliza un receptor para el que la secuencia es conocida. Como cada transmisor emplea una secuencia distinta, es posible que varios transmisores-receptores compartan una misma área geográfica sin interferirse.

En la mayoría de los sistemas de comunicaciones el factor más importante a tener en cuenta es la eficiencia con la que los sistemas utilizan tanto la energía de la señal como el ancho de banda. En algunos casos, existen situaciones en las que es necesario que es sistema soporte las interferencias externas, que provea la capacidad de múltiple acceso sin control externo, y que suministre canales seguros, es decir inaccesibles para quienes no estén involucrados en la comunicación. Para obtener estas características es a menudo necesario sacrificar parte de la eficiencia. Sin embargo la técnica de espectro esparcido permite lograr tales objetivos manteniendo la eficiencia.

Las ventajas principales de los sistemas de comunicaciones de SS son:

- Resistencia a la interferencia intencional y no intencional, esta es una cualidad muy importante cuando se transmite en áreas congestionadas.
- Habilidad de eliminar o atenuar el efecto de la propagación multicamino, que es un gran obstáculo en las comunicaciones urbanas.

- Los usuarios pueden compartir la misma banda de frecuencia con otros usuarios, debido a la similitud con una señal de ruido.
- En cualquier situación se aprovecha todo el ancho de banda.
- Privacidad debido a los códigos aleatorios desconocidos, los códigos aplicados son, en principio, desconocidos para un usuario no deseado.
- Posibilidad de acceso aleatorio, los usuarios pueden iniciar su transmisión en cualquier instante de tiempo.
- Reducción de la potencia de transmisión incrementando la vida de las baterías y reduciendo el tamaño de los transmisores y receptores.

En esta tesis se investiga específicamente la tecnología CDMA (Code Division Multiple Access), que es una técnica de acceso múltiple en la que cada comunicación se codifica digitalmente utilizando una clave de encriptado (código PN) conocida únicamente por los terminales involucrados en el proceso de comunicación. En estos sistemas todos los usuarios transmiten en el mismo ancho de banda simultáneamente. Lo esencial es ser capaz de extraer la señal deseada mientras se rechaza la máscara de ruido aleatorio.

Que el sistema sea exitoso depende principalmente de las secuencias PN utilizadas, hay tres características principales que deben analizarse para la eleccion de estas:

1. Cada secuencia debe ser fácilmente distinguible de cualquier versión desplazada temporalmente de sí misma (auto correlación). Esto facilita la sincronización en el receptor y reduce la interferencia por la propagación multicamino.

2. Cada secuencia debe ser fácilmente distinguible del resto de las secuencias código de una misma familia, con independencia del desplazamiento temporal entre ellas (correlación cruzada). Esto permite una mayor capacidad de separación de señales de distintos usuarios y minimiza la interferencia.

3. Mayor cantidad de códigos por familia de secuencias para poder tener muchos usuarios compartiendo el mismo canal de comunicaciones.

Estas tres características son contradictorias entre sí, en el sentido que no se las puede optimizar en forma simultánea: el hecho de querer conseguir, por ejemplo, una correlación cruzada nula entre secuencias PN, para reducir al máximo la interferencia mutua entre diferentes usuarios, puede perjudicar las propiedades de auto correlación y por lo tanto la facilidad para sincronizar transmisor con receptor.

Del mismo modo la cobertura de toda la banda espectral escogida también es un aspecto importante a ser considerado.

Es decir que debe buscarse una solución de compromiso potenciando aquella cualidad que resulte más importante en la aplicación específica que se esté considerando.

En la literatura se han utilizado diferentes conjuntos de secuencias PN que en esta tesis denominaremos en forma genérica secuencias clásicas, en contraposición con aquellas que se han investigado, basadas en el caos determinista y que denominaremos secuencias caóticas.

En particular, como se demuestra en la tesis, la principal desventaja de las secuencias clásicas frente a las caóticas es la baja cantidad de códigos por familia de secuencias, esto implica una limitación en el número de usuarios permitidos.

En esta tesis se estudiaron secuencias caóticas obtenidas truncando, cuantificando e iterando series caóticas que podemos llamar en forma genérica secuencias caóticas randomizadas. Se las comparó con las secuencias clásicas.

Los resultados se ejemplifican con los siguientes casos de estudio:

- Secuencias caóticas: Mapas Logísticos, de Bernoulli (Three Way y Four Way), Tent y Skew Tent.
- Secuencias clásicas: M-secuencias, secuencias Gold, secuencias Kasami, secuencias Barker, Códigos Walsh y Códigos Gold ortogonales.

Para comparar las secuencias se evaluaron, además de la correlación cruzada, la autocorrelación y el esparcimiento del espectro, otras características como la aleatoriedad la constante de tiempo de mezcla, la distribución de valores y momentos (cumulantes), las entropías de histograma y de permutación, las complejidades estadísticas y de zipping, entre otras.

0.4. Caos, SS y mejora de la compatibilidad electromagnética

El caso de la reducción de la interferencia electromagnética (EMI) es de sumo interés para lograr sistemas electromagnéticamente compatibles (EMC). La interferencia electromagnética es la perturbación que ocurre en cualquier circuito, componente o sistema electrónico a causa de una fuente externa al mismo. Esta perturbación puede interrumpir, degradar o limitar el rendimiento de ese sistema.

La fuente de la interferencia puede ser cualquier objeto, ya sea artificial o natural,

que posea corrientes eléctricas que varíen rápidamente, como un circuito eléctrico, el Sol o las auroras borealis.

Un caso muy frecuente en un gran número de circuitos es la presencia de algún tipo de oscilador local, que genera una señal de clock, y cuya función es mantener el sincronismo de otros componentes. Debido a las características que posee esta señal (flancos abruptos y periodicidad) su espectro presenta toda la energía concentrada en ciertas frecuencias (la del oscilador y sus armónicas impares). Por lo tanto, radiará picos de energía en estas frecuencias, produciendo un espectro que puede exceder los límites aceptables de la interferencia electromagnética.

Ejemplo de esta situación son las plaquetas que cuentan con microprocesadores y señales de clock o las mixtas, que tienen componentes tanto analógicos como digitales.

Una solución clásica para la mejora de EMC se basa en la adopción de filtros, cables y conectores blindados. Estas metodologías que podemos denominar *a posteriori* tienen varias desventajas especialmente el incremento del costo de los aparatos, en especial en sistemas de potencia.

Una estrategia para evitar los filtros, es introducir una pequeña modulación en la frecuencia de la señal de clock, de forma tal que el espectro se esparza en una banda alrededor de la frecuencia deseada, destruyendo así la periodicidad: aquí es donde se utiliza la técnica de FHSS.

El objetivo es obtener una señal cuyo espectro sea lo más constante posible y así conseguir un nivel lo suficientemente bajo como para que no interfiera con el resto de los componentes pero que no altere la función de sincronización del clock

enmascarado. Estas señales son llamadas señales CEW (de Constant Envelope Wideband Signals). Al reducirse los picos de energía irradiada se logra cumplir con las regulaciones de EMC.

La forma del espectro de la señal modulada dependerá principalmente de la señal modulante, si esta es una señal periódica la potencia de la señal modulada estará aún densamente concentrada alrededor de frecuencias especificas y por lo tanto la interferencia producida seguirá siendo relativamente alta no lográndose el objetivo de mejora de EMC.

Una forma de conseguir la mejora es usar una señal de modulación no periódica, tal como las generadas por señales aleatorias o, en nuestro caso, por señales caóticas.

En esta tesis se estudió el caso de utilizar como señal modulante una secuencia caótica. Se encontró que bajo ciertas condiciones las señales moduladas con secuencias caóticas pueden comportarse con calidad similar a las basadas en modelos aleatorios, presentando además la ventaja de su fácil implementación y la posibilidad de generar una cantidad ilimitada de secuencias.

0.5. Muestreo caótico para el filtrado digital

El muestreo no periódico es un método efectivo para el filtrado de señales de baja frecuencia inmersas en ruido de alta frecuencia. Es particularmente útil cuando existen limitaciones en la frecuencia de muestreo tales como las originadas por el tiempo requerido para la conversión Analógico Digital (AD). En ese caso el uso de filtros antialiasing es la solución standard para eliminar el ruido de alta frecuencia. Sin embargo, el muestreo no periódico reduce el efecto de aliasing en forma natural permitiendo eliminar los filtros físicos.

Según cuál sea la función de muestreo empleada se han definido en la literatura dos tipos básicos de muestreo no periódico: el Additive Asynchronous Sampling (AAS) y el Jitter Added Sampling (JAS). En el AAS la falta de periodicidad se logra sumando muestra a muestra un lapso variable. De esta forma obtiene un sistema asíncrono. En el JAS se adiciona un tiempo variable a un marco sincrónico.

Una ventaja del caos determinista, además de su propiedad de generar señales quasi estocásticas empleando mapas no lineales simples, es la posibilidad de diseñar esos mapas de modo que su salida posea una Función Densidad de Probabilidad Invariante (IPDF) deseada. También es posible regular la constante de tiempo del transitorio requerido por el sistema para llegar a esa distribución de salida.

Mediante técnicas de skipping (uso de mapas iterados) o bien mediante dinámica simbólica se logra obtener generadores que, a pesar de ser pseudo aleatorios, superan bancos de tests estadísticos exigentes.

Las simulaciones realizadas muestran que los resultados en el filtrado de señales de baja frecuencia con ruido de alta frecuencia, por medio de JAS y AAS, combinado con filtros digitales de respuesta finita al impulso (FIR) con muestreo caótico, son equivalentes a los que pueden obtenerse mediante muestreo aleatorio.

Las series temporales más convenientes para esta aplicación son las que poseen IPDF uniforme y constante de tiempo de mezclado baja. Sin embargo, los resultados obtenidos demuestran que el efecto de la constante de tiempo sólo interviene cuando se trabaja con filtros FIR excesivamente cortos, y estos en la práctica son inadecuados debido a que el número de elementos del FIR también afecta la atenuación en alta frecuencia.

Un resultado importante obtenido es la posibilidad de mejorar en forma simultánea las propiedades de mezclado y la uniformidad de la IPDF. Es posible entonces extender los resultados a un conjunto muy amplio de mapas caóticos.

0.6. Organización y contribuciones de la Tesis

La tesis comienza con un capítulo introductorio en el que se presentan los sistemas caóticos, tanto desde el punto de vista determinista como del punto de vista estadístico. En este capítulo se incluyen además algunas contribuciones originales básicas de esta tesis que serán utilizadas en todas las aplicaciones: en especial la demostración de que es posible asignar distintas funciones distribución de probabilidad a una misma secuencia, en base a la utilización de diferentes dinámicas simbólicas.

En el segundo capítulo se presentan los sistemas de comunicaciones de espectro esparcido, especialmente aquellos que serán objeto de aplicación de secuencias PN caóticas. Se definen y comparan nuevos cuantificadores de las secuencias PN empleadas en CDMA. Esos cuantificadores permiten optar por la familia de códigos más adecuada según las características que se deseen optimizar (correlación cruzada, auto correlación, esparcimiento del espectro, etc.). Se propone además un nuevo cuantificador genérico basado en la Complejidad de Zipping como cuantificador genérico de las distintas familias.

En el tercer capítulo se describen técnicas de randomización desarrolladas y evaluadas mediante cuantificadores estudiados en esta tesis, con el objeto de mejorar las características estadísticas de las secuencias caóticas de mapas simples, para poder asemejarlas a una señal aleatoria (ruido blanco). Se definen distintos planos de representación donde es posible evaluar la calidad del proceso de randomización en forma

cuantitativa.

En los capítulos 4 y 5 se describen las otras dos aplicaciones ya mencionadas: la mejora de la compatibilidad electromagnética y el filtrado digital de ruido de alta frecuencia en señales de baja frecuencia. Se presentan resultados originales publicados empleando secuencias caóticas y se los compara con los obtenidos con secuencias estocásticas.

El capítulo 6 presenta resultados preliminares obtenidos en cuanto a implementación de generadores de ruido digital, tanto caóticos como estocásticos, empleando lógicas programables, especialmente Field Programmable Gate Arrays (FPGA) para su uso en diversas aplicaciones.

La tesis finaliza con la descripción de problemas abiertos, las líneas de trabajo a encarar en una próxima etapa y la bibliografía.

Capítulo 1

Metodología para el estudio de Sistemas Caóticos y Estocásticos

1.1. Introducción

El desarrollo del análisis matemático iniciado en el siglo XVII por Newton y Leibniz entre otros, fue la herramienta clave para la investigación de sistemas en ciencia y en ingeniería hasta hace unos 30 años. Durante ese extenso lapso muchos problemas fueron resueltos exitosamente en la mecánica y el electromagnetismo. Por otra parte el desarrollo de la mecánica estadística de Boltzmann y Gibbs permitió la comprensión de sistemas complejos, en especial de los gases.

Se fue consolidando en la ciencia la idea que los sistemas pueden ser clasificados en deterministas y estocásticos. Los primeros tienen un modelo matemático expresado en la forma de: 1) un sistema de ecuaciones diferenciales ordinarias o a derivadas parciales, denominado sistema dinámico continuo ó 2) un sistema de ecuaciones expresadas en variables discretas, denominado sistema dinámico discreto o mapa. Los sistemas estocásticos, por otra parte están caracterizados por funciones densidad de

probabilidad asignadas a sus variables de estado. En muchos de ellos es también posible plantear ecuaciones diferenciales estocásticas, que gobiernan la evolución temporal y cuyas funciones solución son precisamente las funciones densidad de probabilidad.

Los sistemas no lineales y en especial el caos determinista comenzaron a estudiarse hacia el 1900. Ya en esa época Poincaré describió la sensibilidad a las condiciones iniciales. Sin embargo, recién hace unos 30 años emerge la teoría del caos determinista en forma consistente, a partir de los trabajos fundamentales de Lorenz [1], Feigenbaum [2] y otros. Surge una nueva visión de la naturaleza en la que el caos determinista es el protagonista principal. La gran repercusión del caos determinista se debe a que el amplio espectro de problemas en los que aparece lo muestran como un fenómeno habitual en sistemas no lineales. El gran desarrollo de la electrónica digital hizo posible descubrir y estudiar en profundidad este tipo de sistemas en los que, en general, falla el cálculo analítico.

Hasta la aparición en escena del caos determinista la creencia era que la naturaleza estadística de los sistemas estocásticos provenía de una alta dimensión, es decir de un número muy elevado de variables requeridas en su representación. Pero un sistema caótico tiene la particularidad de tener un modelo determinista simple y conocido, a pesar de lo cual la evolución temporal de sus variables tiene la apariencia de una señal estocástica.

Para sistemas continuos se demuestra que es suficiente que el sistema tenga dimensión 3. Es decir, si el sistema está modelado por 3 ó más ecuaciones diferenciales ordinarias es posible la generación de caos. En el caso de sistemas discretos, los mapas no lineales de una única variable (mapas unidimensionales) pueden presentar este comportamiento [2].

Es decir que el paradigma aceptado antes de la teoría del caos, acerca de la correspondencia entre señales estocásticas y sistemas cuyos modelos son demasiado complejos para poder ser obtenidos quedó refutado por la existencia del caos determinista y la conclusión es que **cuando hay nolinealidades se pueden obtener señales de aspecto estocástico de modelos deterministas simples**.

Las similitudes y diferencias entre sistemas caóticos y estocásticos han estado en el centro del debate científico de los últimos años. Se ha modificado la visión tradicional que hacía énfasis en el desarrollo analítico, para nutrir la investigación de los resultados de una rama de la matemática, la matemática discreta en gran expansión.

Este capítulo introduce la metodología para estudiar los sistemas no lineales discretos, el caos determinista discreto (pseudo caos) y los sistemas estocásticos discretos (pseudo estocásticos). Se presentan un conjunto de herramientas de tipo estadístico que se basan en el estudio de las señales creadas por un sistema más que en el estudio del modelo que lo describe.

1.2. Descripción estadística de series temporales

La existencia de valores reales en la naturaleza (es decir valores en el continuo) es un problema filosófico que está fuera de la discusión de esta tesis. Es bien conocido que el mundo físico que nos rodea es en realidad discreto, si bien ese carácter discreto se hace notable al ir aumentando el grado de detalle con que se lo observa. Además a los fines de la medición de sus variables los valores son discretizados, ya sea por la exactitud de los instrumentos de medida (si se trata de instrumentos analógicos), ya sea por el sistema de adquisición de datos (si se trata de instrumentos digitales).

La "hipótesis del continuo" utilizada en física y en ingeniería es una herramienta valiosa que ha permitido el estudio analítico de los sistemas. Pero, dado que el número real no es representable en una computadora o en un dispositivo lógico programable (paradójicamente el número real es en realidad una idealización!), desde el punto de vista de esta tesis una serie temporal será en todos los casos un conjunto de N valores numéricos $\{x_i,\ i = 1, ...N\}$, enteros o bien expresados en una representación IEEE punto flotante normalizada. No se tomarán en cuenta otros sistemas de representación tales como la aritmética de residuos [3]. Se considerará en lo que sigue que los números representados en formato flotante satisfacen los teoremas que se han demostrado para el caso de números reales.

La inferencia estadística lineal es un campo muy desarrollado. A partir de una serie temporal permite estimar parámetros del sistema que la genera. Esa estimación se ve modificada por las suposiciones de las que se parta. Dado que las ecuaciones de movimiento deterministas lineales con coeficientes constantes sólo tienen soluciones exponenciales o periódicas, los sistemas lineales necesitan entradas irregulares para producir señales irregulares acotadas. El sistema más simple que produce señales no periódicas es un "proceso estocástico lineal".

Las principales herramientas provenientes del campo de la inferencia estadística lineal, de uso también en sistemas no lineales, y que serán empleadas en esta tesis son:

1. Espectro de Potencia

2. Autocorrelación lineal

3. Histograma y cálculo de sus cumulantes (valor medio, varianza, etc.)

Estas herramientas son ampliamente conocidas y desarrolladas en la literatura.

Las herramientas provenientes del campo de los sistemas no lineales caóticos no son en cambio tan difundidas por lo que en las subsecciones siguientes se da una breve descripción de cada una de ellas. Al emplearse cada herramienta específica en los capítulos posteriores se dará un mayor detalle.

Un aspecto importante es verificar que las series estudiadas son estacionarias. La estacionariedad de una serie temporal proveniente de un sistema no lineal no puede comprobarse en forma analítica. El método usual entonces es construir series subrogadas. Estas series pueden ser en algunos casos subseries de la serie total, pero en otros se las obtiene a partir del procesamiento de los datos. Por ejemplo, si se desea cuantificar la aleatoriedad de una serie temporal pueden construirse series surrogadas permutando los valores numéricos, y los cuantificadores utilizados (la complejidad de zipping por ejemplo) de las series surrogadas deben ser iguales (dentro de las fluctuaciones estadísticas) a los correspondientes a la serie original.

1.2.1. Espacios de embedding

Una medición escalar es la proyección de variables "internas" (es decir no medibles) del sistema, sobre el eje real. El proceso de proyección es en general no lineal y puede mezclar el efecto de distintas variables internas.

Es imposible reconstruir el espacio de estados de un sistema a partir de una de sus proyecciones (datos tomados de una variables). Sin embargo en general no es necesaria una reconstrucción completa sino que es suficiente construir un nuevo espacio conocido como "espacio de embedding". En 1936 Whitney [4] probó el Teorema de Embedding que establece que toda variedad D-dimensional suave se puede embeber en el espacio

$\mathbb{R}^{2D+1}$ y que el conjunto de mapas que producen el embedding forman un conjunto abierto y denso en el espacio de los mapas $\mathcal{C}^1$.

Este teorema tiene como consecuencia práctica otro teorema, demostrado por Takens [5], aplicable especialmente a los atractores de los sistemas dinámicos. La situación de reconstrucción de un atractor $\mathfrak{A}$ es muy particular. El atractor es un subconjunto del espacio de estados del sistema. Eso garantiza que el mapa que convierte el punto $\overrightarrow{s}_i$, en el siguiente punto $\overrightarrow{s}_{i+1}$, es único. Entonces la dinámica del sistema constituye un mapeo único, independiente del tiempo. Entonces es posible demostrar que puede generarse el espacio de embedding mediante un "vector retardado" $\{x_i, x_{i+1}, ...x_{i+m}\}$, asociado a cada valor x_i de la serie temporal. Sauer resumió muchos resultados importantes referentes a la construcción de espacios de embedding [6].

1.2.2. El operador de Perrón-Frobenius

Supongamos que una serie temporal es generada por un mapa M, 1-dimensional, sobre el intervalo $[0, 1]$. Partiendo de distintas condiciones iniciales se obtienen diferentes series temporales. En el caso en que el mapa es caótico y la serie es suficientemente larga, los valores obtenidos recorren todo el intervalo y su histograma normalizado puede considerarse una distribución de probabilidades (PDF natural o invariante del sistema). Entonces, es posible describir la dinámica del sistema evaluando la evolución temporal de una distribución de valores iniciales $\rho_0(x)$ en el espacio de probabilidades [7]. El operador llamado de Perron-Frobenius es el que gobierna esta evolución temporal [7, 8, 9].

En esta tesis se aplicará esta metodología al estudiar, en el capítulo 2, los generadores de números pseudo aleatorios (PRNGs) basados en mapas caóticos.

En especial interesa analizar el espectro de este operador, el cual refleja las características del mapa. El autovector del autovalor unidad corresponde a la función densidad de probabilidades invariante del mapa. El autovalor de módulo máximo, dentro del círculo unidad define la "constante de mezcla", esta constante es proporcional a cuán rápido una PDF inicial converge a la invariante. Se hace un análisis sistemático de distintos cuantificadores de aleatoriedad, y su relación con la uniformidad de la autofunción del autovalor unidad y la constante de mezcla del operador PF del mapa.

1.2.3. Dinámica simbólica

Dinámica simbólica es el modelado de un sistema dinámico mediante un espacio discreto formado por secuencias de longitud infinita de símbolos abstractos, cada uno de los cuales corresponde a un estado del sistema. La idea data de un trabajo publicado por Hadamard [10] sobre geodésicas en superficies de curvatura negativa. El primer tratamiento formal fue desarrollado por Morse y Hedlund [11]. Shannon usó secuencias simbólicas en su célebre trabajo que diera origen a la teoría de la información [12]. La teoría se desarrolló notablemente desde 1960. Una aplicación muy importante fue el teorema de Sharkovskii acerca de órbitas periódicas de mapas continuos del intervalo [13].

La dinámica simbólica se caracteriza porque tanto el tiempo como los valores son discretos. Es decir corresponde a sistemas digitales discretos. Si los estados del sistema no son discretos la dinámica simbólica representa una descripción de "grano grueso" del mismo.

En esta tesis la dinámica simbólica es empleada para el análisis y la randomización de generadores de números pseudo aleatorios. El empleo de distintas dinámicas simbólicas da origen a distintas distribuciones de probabilidad correspondientes a una misma serie temporal.

1.2.4. Complejidad de Zipping

En sistemas de comunicaciones, el problema de la codificación eficiente es bastante antiguo. Un ejemplo de ello lo constituye el código Morse que codifica un texto con dos caracteres: el punto y la línea. La idea es determinar la mejor forma (más corta) de codificar los caracteres del idioma inglés con secuencias de puntos y líneas. Morse codificó los caracteres más frecuentes con el mínimo número de símbolos. Por lo tanto la letra *e*, que es la letra del idioma inglés mas frecuente, se codificó con un punto, y el número 0 se codificó con 5 líneas. Aunque este tipo de codificación está bien definida, su uso no es demasiado práctico dado que están sustentados en la existencia de una fuente ergódica que permite calcular *a-priori* las probabilidades de aparición de las N palabras.

Los denominados compresores o zipeadores son programas concebidos para encontrar el archivo más pequeño que representa una secuencia dada. La idea es sustituir de alguna manera un número de caracteres repetidos por un código que indique que "el carácter X se repite N veces". Éste es el principal objetivo de estos algoritmos: reducir las cadenas de caracteres idénticos a una indicación del carácter y la longitud de la repetición.

LZ77

Uno de los primeros algoritmos de compresión, el de Lempel-Ziv fue ideado por Jacob Ziv y Abraham Lempel, y publicado en [14, 15]. El algoritmo LZ77 mantiene un registro de los últimos caracteres procesados de la entrada pero no construye un diccionario explícito. En cada instante el algoritmo está procesando un punto de la entrada, los n caracteres anteriores forman la historia del algoritmo o ventana, lo que equivale al diccionario. Los caracteres posteriores al punto actual constituyen lo que se denomina el *lookahead buffer* o buffer de adelanto. En cada paso, la secuencia de símbolos que comienza en el punto actual de la entrada se busca en la ventana. Si se encuentra una coincidencia en la ventana que se considere lo suficientemente larga, en la salida se sustituye la cadena coincidente por un par que indica el desplazamiento hacia atrás y la longitud de la cadena coincidente. Como los pares desplazamiento-longitud ocupan menos que la cadena coincidente, se obtiene compresión. Si no se encuentra una coincidencia, la salida es una copia literal de la entrada. A continuación, se avanza la posición actual (y consecuentemente la ventana) la longitud de la coincidencia si la hubo, o bien un símbolo si no hubo coincidencia. El hecho de ir desplazando la ventana sobre la entrada hace que estos algoritmos se denominen de *ventana deslizante*. Cuanto mayor sea el tamaño de la ventana, mayor será la compresión que se obtenga. Esto es así porque también será mayor la historia sobre la que se buscan las posibles coincidencias, y por tanto, se tiene una mayor probabilidad de encontrar una coincidencia más larga. Sin embargo, un tamaño de ventana grande implica que se necesite más espacio para codificar los valores de los desplazamientos. Por otro lado, las coincidencias más cortas (dos o tres símbolos) son las más probables y es deseable que se ahorre espacio cuando se codifican los pares desplazamiento-longitud

de las coincidencias más probables.

Este algoritmo tiene la particularidad que si codifica un texto de longitud L extraído de una fuente ergódica, cuya entropía por carácter es s entonces la longitud del archivo zipeado dividido por la entropía del texto converge a 1 a medida que la longitud del texto se incrementa. Esto significa que la efectividad de la compresión es mejor cuanto mayor sea la longitud del archivo a comprimir y que la entropía de una cadena de caracteres puede calcularse mediante un proceso de compresión.

Si un sistema es caótico, su horizonte de predecibilidad es temporalmente limitado. El límite está asociado con su exponente positivo de Lyapunov. La secuencia temporal, generada a partir de una de sus trayectorias caóticas, no puede comprimirse por un factor arbitrario, es decir es algorítmicamente compleja. En un sistema caótico, una pequeña perturbación produce una separación, creciente con el tiempo, entre una trayectoria y la misma perturbada. Se tiene de esta manera una rápida amplificación (exponencial) de un error en las condiciones iniciales. Por el contrario, una trayectoria regular puede comprimirse fácilmente.

La caracterización de la información contenida en una secuencia puede enfocarse desde dos puntos de vista. El primero es de índole estadística, es decir no considera la transmisión de un mensaje específico, sino las propiedades estadísticas del mensaje emitido por la fuente. Esto la hace un método poderoso para caracterizar sistemas caóticos. El segundo punto de vista considera el problema de caracterizar una secuencia particular y conduce a la teoría de la complejidad algorítmica. La complejidad algorítmica de Chaitin-Kolmogorov es la longitud (en bits) del programa más pequeño que produce como salida esa cadena. Esta definición es extremadamente abstracta y, en principio, es imposible encontrar tal programa. Sin embargo la complejidad

de Zipping, medida a partir de algoritmos de compresión, es una aproximación a la complejidad algorítmica.

En esta tesis la complejidad de zipping se emplea para definir un parámetro que caracterice las secuencias pseudo aleatorias que se emplean en los sistemas de comunicaciones de espectro esparcido (en especial los que utilizan CDMA).

1.2.5. Complejidad Estadística

La adaptación de Kolmogorov-Sinaí de la Teoría de la Información de Shannon permitió la caracterización estadística de fuentes deterministas caóticas. Estos esfuerzos para describir la impredicibilidad de sistemas dinámicos condujeron a la definición de cantidades tales como la entropía métrica, los Exponentes de Lyapunov y las dimensiones fractales que pueden emplearse para detectar la presencia y para cuantificar el comportamiento caótico determinista.

No obstante su utilidad, estas medidas no logran capturar adecuadamente las estructuras correlacionadas en el comportamiento de estos sistemas.Por ejemplo la entropía, definida a partir del histograma de la serie temporal, no tiene en cuenta el orden de los valores, sino únicamente la cantidad de veces que cada uno de ellos se repite dentro de la serie.La existencia de una estructura implica que existe alguna relación entre sus componentes. Cuanto mayores y más intrincadas son las correlaciones entre los constituyentes del sistema, más estructurada es la distribución subyacente. Sin embargo, las estructuras y correlaciones no son completamente independientes de la aleatoriedad. Un modo simple de observar cualitativamente la presencia de estructuras es representar la serie temporal en un diagrama $2D$ o $3D$ colocando en un eje el valor x_n en el otro el x_{n+1} y en el tercero (para el caso $3D$) x_{n+2}.

Un sistema periódico produce un conjunto de puntos. Un sistema en que un estado es función de los anteriores produce una curva. Un sistema aleatorio produce una nube de puntos. El problema con este criterio cualitativo es que sólo es aplicable hasta $3D$, empleando las herramientas gráficas disponibles en la actualidad. La Figs.1.1 muestra el diagrama $3D$ para dos series temporales, una sin estructura (Fig.1.1a) y otra con estructura (Fig.1.1b).

En años recientes se viene realizando un considerable esfuerzo para desarrollar una medida general que cuantifique el grado de estructura o patrón presente en un proceso. Existen muchos métodos ad hoc para detectar estructuras, pero ninguno ha logrado tener una aplicabilidad tan amplia como la de la entropía para indicar aleatoriedad. Las cantidades que se han propuesto como medidas generales de estructuras se denominan "medidas de complejidad estadística". Las definiciones corrientes caen dentro de tres amplias categorías:

a) Clase I: Se considera la complejidad como una función monótona creciente del desorden, ejemplo de esto son las complejidades algorítmicas [16, 17] y las diversas entropías [12, 18, 19]. En esta concepción una distribución de probabilidades uniforme tiene máxima complejidad. Este criterio no concuerda con la idea cualitativa generalmente aceptada, de no considerar complejas las series temporales que producen una nube uniforme de puntos en el método cualitativo expuesto arriba y ejemplificado con las Figs.1.1.

b) Clase II: corresponde a las definiciones en que la complejidad es una función convexa del desorden; es decir la complejidad tiene un mínimo para los sistemas completamente ordenados (que producen unos pocos puntos en el diagrama explicado arriba) o completamente desordenados (que producen una nube uniforme de puntos

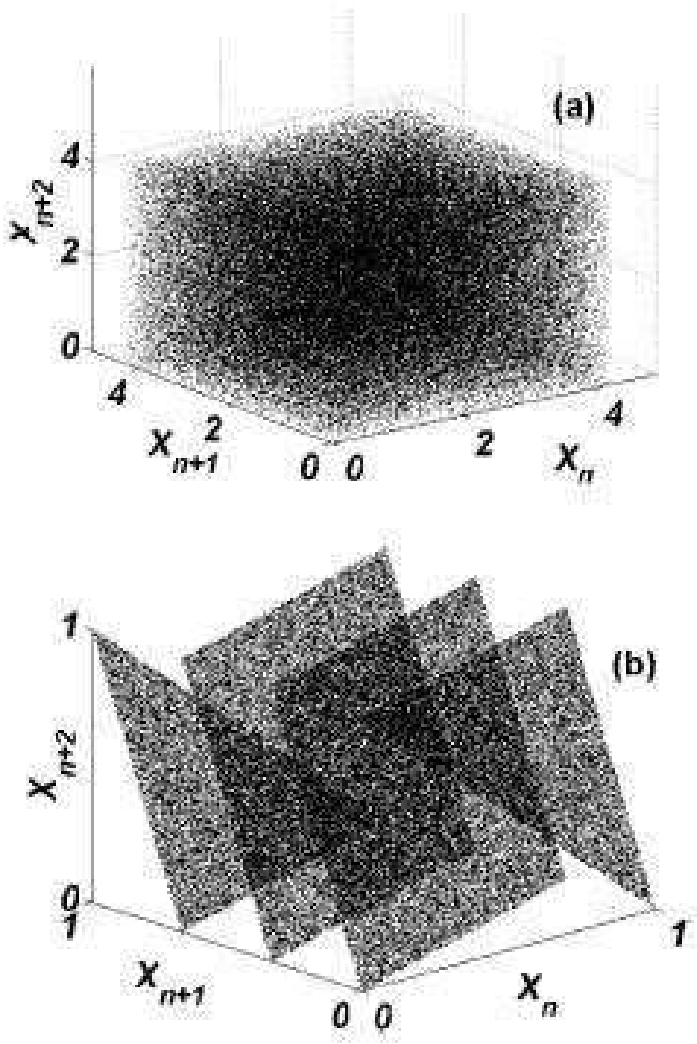

Figura 1.1: Representación $3D$ de dos series temporales: (a) sin estructura, generada por el PRNG conocido como Mother, (b) con estructura, generada por el mapa CAT discretizado

en el diagrama explicado arriba). Las medidas de la clase II tienen un máximo en algún nivel intermedio. A esta categoría pertenecen las profundidades lógica y termodinámica [20, 21] así como la complejidad que se utiliza en este trabajo y sobre la que profundizamos más adelante.

c) Clase III: Engloba las definiciones que asocian la complejidad con el orden, estas identifican complejidad con auto-organización y la auto-organización con el orden [22].

En esta tesis se utilizará una medida de complejidad estadística clase II, por ser la más consistente con el método gráfico cualitativo. Se detalla a continuación la evolución de estas medidas de complejidad. La forma funcional fue introducida por López-Ruiz, Mancini y Calbet (LMC) en un trabajo seminal [23]. En su definición la complejidad es un funcional de la distribución de probabilidades dado por:

$$\mathcal{C}[P] \;=\; \mathcal{H}[P] \cdot \mathcal{Q}[P] \;. \tag{1.2.1}$$

El desequilibrio Q representa una distancia entre la distribución de probabilidades del sistema y la distribución de probabilidades de equilibrio y puede tener diferentes expresiones según la métrica que se adopte en el espacio de las probabilidades. Es decir:

$$\mathcal{Q}[P] = \mathcal{Q}_0 \cdot \mathcal{D}[P, P_e] \;, \tag{1.2.2}$$

donde $\mathcal{Q}_0$ es una constante de normalización que hace que $0 \leq \mathcal{Q} \leq 1$.

Cuando el sistema produce una serie temporal que en el método gráfico luce como una nube uniforme de puntos, el desequilibrio resulta ser nulo y la forma producto hace que la complejidad también sea nula, a pesar que la entropía normalizada tiene valor 1 en este caso. Por otra parte, el desequilibrio es 1 para el caso de una serie periódica (que produce sólo puntos aislados en el diagrama cualitativo). Pero en ese

caso es la entropía normalizada la que resulta nula y la forma producto hace que la complejidad también lo sea.

La entropía normalizada H también puede adoptar distintas formas. Por mencionar algunas de las más importantes:

(a) Shannon, $\mathcal{H}_1^{(S)} = \mathrm{S}_1^{(S)}[P] \,/\, \mathrm{S}_1^{(S)}[P_e]$,

$$\mathrm{S}_1^{(S)}[P] \;=\; -\sum_{j=1}^{N} p_j \;\ln(p_j) \;; \tag{1.2.3}$$

(b) Tsallis, $\mathcal{H}_q^{(T)} = \mathrm{S}_q^{(T)}[P] \,/\, \mathrm{S}_q^{(T)}[P_e]$,

$$\mathrm{S}_q^{(T)}[P] \;=\; \frac{1}{(q-1)} \sum_{j=1}^{N} \left[\; p_j \;-\; (p_j)^q \;\right] ; \tag{1.2.4}$$

(c) escort-Tsallis $\mathcal{H}_q^{(G)} = \mathrm{S}_q^{(G)}[P] \,/\, \mathrm{S}_q^{(G)}[P_e]$,

$$\mathrm{S}_q^{(G)}[P] \;=\; \frac{1}{(q-1)} \left\{ 1 \;-\; \left[\sum_{j=1}^{N} \left(\; p_j \;\right)^{1/q} \right]^{-q} \right\} ; \tag{1.2.5}$$

(d) Rényi $\mathcal{H}_q^{(R)} = \mathrm{S}_q^{(R)}[P] \,/\, \mathrm{S}_q^{(R)}[P_e]$,

$$\mathrm{S}_q^{(R)}[P] \;=\; \frac{1}{(1-q)} \ln \left\{ \sum_{j=1}^{N} (p_j)^q \right\} . \tag{1.2.6}$$

La cantidad q se conoce como parámetro de deformación ($q = 1$ para la entropía de Shannon). En el límite $q \to 1$ todas las expresiones de arriba coinciden con la de Shannon.

En relación con la métrica y su distancia inducida $\mathcal{D}$ que forma parte del factor $\mathcal{Q}$, existe una diversidad de opciones. Para $P_i \equiv \{p_1^{(i)}, \cdots, p_N^{(i)}\}$, con distribuciones discretas $i = 1, 2$ nos limitaremos a los casos estudiados en [24, 25]:

a) Norma Euclideana $\mathcal{D}_E$ en $\mathbb{R}^N$ [23],

$$\mathcal{D}_E[P_1, P_2] \;=\; \|P_1 \;-\; P_2\|_E \;=\; \sum_{j=1}^{N} \;\left\{\; p_j^{(1)} \;-\; p_j^{(2)} \;\right\}^2 \;; \tag{1.2.7}$$

b) Distancia de Wootters's $\mathcal{D}_W$ [26],

$$\mathcal{D}_W[P_1, P_2] \;=\; \cos^{-1}\left\{ \;\sum_{j=1}^{N} \;\left(\; p_j^{(1)} \;\right)^{1/2} \cdot \left(\; p_j^{(2)} \;\right)^{1/2} \;\right\} \;; \tag{1.2.8}$$

c) Entropías relativas de Kullback $K_q^{(\kappa)}$

$$\mathrm{K}_1^{(S)}[P_1|P_2] \;=\; \sum_{j=1}^{N} \; p_j^{(1)} \cdot \ln\left(\frac{p_j^{(1)}}{p_j^{(2)}}\right) \;, \tag{1.2.9}$$

$$\mathrm{K}_q^{(T)}[P_1|P_2] \;=\; \frac{1}{(q-1)} \;\sum_{j=1}^{N} \;\left(p_j^{(1)}\right)^q \;\left\{\left(p_j^{(2)}\right)^{1-q} - \left(p_j^{(1)}\right)^{1-q}\right\} \;, \tag{1.2.10}$$

$$\begin{aligned} \mathrm{K}_q^{(G)}[P_1|P_2] \quad = \quad & \frac{1}{(q-1)} \;\sum_{j=1}^{N} \;\frac{p_j^{(1)}}{(A[P_1])^q} \;\cdot \\ & \left\{ \left[\frac{\left(p_j^{(2)}\right)^{1/q}}{A[P_2]}\right]^{1-q} - \left[\frac{\left(p_j^{(1)}\right)^{1/q}}{A[P_1]}\right]^{1-q} \right\} \;, \end{aligned} \tag{1.2.11}$$

$$\mathrm{K}_q^{(R)}[P_1|P_2] \;=\; \frac{1}{(q-1)} \;\ln\left\{\sum_{j=1}^{N} \;\left(p_j^{(1)}\right)^q \;\left(p_j^{(2)}\right)^{1-q}\right\} \;, \tag{1.2.12}$$

para $q \neq 1$ y $A[P] = \sum_{j=1}^{N} \;(p_j)^{1/q}$;

d) Divergencias de Jensen $\mathcal{J}_q^{(\kappa)}$

$$\mathcal{J}_q^{(\kappa)}[P_1, P_2] \;=\; \frac{1}{2}\left\{\mathrm{K}_q^{(\kappa)}[P_1 \mid (\; P_1 + \;P_2)/2] \;+\; \mathrm{K}_q^{(\kappa)}[P_2 \mid (\; P_1 + \;P_2)/2]\right\} \;, \tag{1.2.13}$$

con $\kappa =$ Shannon (S) ($q = 1$), Tsallis (T), escort-Tsallis (G) y R´enyi (R).

Tomando como base la forma funcional LMC obtenemos una familia de SCMs para cada una de las medidas de desorden y desequilibrio mencionadas arriba, es decir,

$$\mathcal{C}_{\nu,q}^{(\kappa)}[P] \;=\; \mathcal{H}_q^{(\kappa)}[P] \cdot \mathcal{Q}_q^{(\nu)}[P] \; . \tag{1.2.14}$$

con $\kappa =$ S, T, G, R para un valor fijo de q. En el caso Shannon ($\kappa =$ S) tenemos por supuesto $q = 1$. El índice $\nu = E, W, K_q^{(\kappa)}, \mathcal{J}_q^{(\kappa)}$ nos dice que el desequilibrio debe medirse con una medida de distancia adecuada.

Nótese que para $\nu = K_q^{(\kappa)}$, la familia de SCM's se convierte en $\mathcal{C}_q^{(\kappa)}[P] = (1 - \mathcal{H}_q^{(\kappa)}[P]) \cdot \mathcal{H}_q^{(\kappa)}[P]\cdot$, que tiene la forma funcional generalizada propuesta por Shiner, Davison y Lansberg [27]. Debe tenerse en cuenta que para este caso particular la complejidad se convierte en una función de la entropía H de modo que el desequilibrio Q y la complejidad C no aportan nueva información sobre el sistema que la ya contenida en la propia entropía H. En cambio los restantes miembros de la familia $\mathcal{C}_{\nu,q}^{(\kappa)}$ (con $\nu \neq K_q^{(\kappa)}$) no son funciones de la entropía. Por el contrario, para una dada entropía $\mathcal{H}_q^{(\kappa)}$, existe un rango amplio de valores para la SCM, desde un valor mínimo $\mathcal{C}_{min}$ hasta un valor máximo $\mathcal{C}_{max}$. Entonces la complejidad $\mathcal{C}_{\nu,q}^{(\kappa)}$ contiene nueva información que toma en cuenta las peculiaridades de la distribución de probabilidades. En [26] se explica el procedimiento general para obtener los límites $\mathcal{C}_{min}$ y $\mathcal{C}_{max}$ correspondientes a la familia de SCM's $\mathcal{C} = \mathcal{H} \cdot \mathcal{Q}$.

Si $\kappa = S$ ($q = 1$) y $\nu = E$ recuperamos la SCM de Lopez-Ruiz, Mancini y Calbet (LMC) [23], $\mathcal{C}_{LMC} = \mathcal{C}_{E,1}^{(S)}$. Crutchfield y colaboradores destacaron [28] que la LMC tiene algunas características indeseables que se listan a continuación:

• no es extensiva ni intensiva.

• se desvanece exponencialmente en el límite termodinámico para todos los sistemas uni dimensionales de rango finito.

Los autores mencionados recalcan que una SCM razonable debería

• se capaz de distinguir entre distintas periodicidades y

• desvanecerse sólo para período 1.

Por último, y en relación con la capacidad de la LMC para capturar adecuadamente aspectos dinámicos esenciales, se encontraron dificultades reportadas en [29]. Por ejemplo, el uso de la forma funcional producto hace imposible eliminar la segunda deficiencia apuntada arriba.

Realizando algunos cambios en la definición de desequilibrio (utilizar la distancia de Wootters' [26] o la de Jensen's [30]), se puede finalmente obtener una SCM generalizada que satisfaga:

(i) rescata detalles esenciales de la dinámica,

(ii) es intensiva, y

(iii) es capaz de discernir entre distintos grados de periodicidad.

En definitiva en esta tesis se utilizará la medida de complejidad estadística dada por la ecuación 1.2.14 con $\kappa = S$, $\nu = J$ y $q = 1$, que por simplicidad de notación designaremos en lo que sigue como $\mathcal{C}^{MPR}[P]$

1.2.6. Diagramas de Recurrencia

Los diagramas de recurrencia (Recurrence Plots RP) fueron introducidos por Eckmann *y colaboradores* [31] para visualizar la repetición de estados en la evolución en el espacio de estados. Un diagrama de recurrencia es una representación bidimensional en la que ambos ejes son temporales. La repetición o recurrencia de un estado es

su aparición en dos instantes diferentes t_i, t_j, y es representada por un punto en el diagrama. Por supuesto una recurrencia exacta sólo es posible en sistemas periódicos continuos. En cualquier otro caso sólo es posible detectar recurrencias aproximadas a un error ϵ. La denominada función de recurrencia se puede expresar matemáticamente:

$$\mathbf{R}(i,j) = \Theta\left(\varepsilon - \|\overrightarrow{x}(i) - \overrightarrow{x}(j)\|\right), \tag{1.2.15}$$

con $\overrightarrow{x}(i) \in \Re^m$ y $i,j = 1,\cdots,N$. N es el número de estados discretos $\overrightarrow{x}(i)$ considerados, $\|\bullet\|$ es una norma, y $\Theta(\bullet)$ es la función escalón de Heaviside.

Cuantificadores basados en diagramas de recurrencia

Estos cuantificadores se basan en el estudio de la función recurrencia 1.2.15.

Consideraremos dos clases de medidas sobre diagramas de recurrencia a saber:

1. Medidas basadas en la densidad de recurrencia (medida por el número de puntos en el Diagrama de Recurrencia). En particular utilizaremos *Recurrence Rate* (RR), dado por:

$$RR(\varepsilon) = \frac{1}{N^2}\sum_{i,j=1}^{N}\mathbf{R}_{ij}(\varepsilon)\,. \tag{1.2.16}$$

   Nótese que en el límite $N \to \infty$, RR es la probabilidad de que un estado se repita dentro de una vecindad ε en el espacio de estados. Para los PRNGs el valor ideal sería $RR = 0$. Pero en la práctica si no se encuentra ningún punto en el diagrama de recurrencia es necesario incrementar el valor de ε adoptado para que este cuantificador tenga sentido.

2. Medidas diagonales: estas medidas están relacionadas con el histograma $P(\varepsilon, l)$

de las longitudes de líneas diagonales sobre el RP:

$$P(\varepsilon, l) = \sum_{i,j=1}^{N} \left[1 - \mathbf{R}_{i-1,j-1}(\varepsilon)\right] \left[1 - \mathbf{R}_{i+l,j+l}(\varepsilon)\right] \cdot \prod_{k=0}^{l-1} \mathbf{R}_{i+k,j+k}(\varepsilon) \; . \tag{1.2.17}$$

Los procesos con comportamiento no correlacionado o bien débilmente correlacionado no originan diagonales en el RP (o bien diagonales muy cortas) en tanto que los procesos deterministas originan diagonales "largas" y un número menor de puntos de recurrencia aislados. Consideraremos tres medidas diagonales:

a) *Cuantificador de Determinismo* DET, es el cociente entre el número de puntos que forman estructuras diagonales de longitud igual o superior a l_{min} y el total de puntos de recurrencia:

$$DET = \frac{\sum_{l=l_{min}}^{N} l \cdot P(\varepsilon, l)}{\sum_{l=1}^{N} l \cdot P(\varepsilon, l)} \; ; \tag{1.2.18}$$

Esta medida da idea al grado de "predictibilidad´´ del sistema. Para $l_{min} = 1$ el determinismo es uno. Al realizar la elección de l_{min} debe tenerse en cuenta que un valor de l_{min} demasiado grande generará un histograma $P(l)$ demasiado disperso, esto hara que la fiabilidad de DET disminuya. Una línea diagonal de longitud l significa que un segmento de la trayectoria permanece "cercano´´ a otro segmento de la trayectoria durante l pasos de tiempo; por lo tanto estas líneas están relacionadas a la divergencia de los segmentos de la trayectoria.

b) *Longitud media de las diagonales* L

$$L = \frac{\sum_{l=l_{min}}^{N} l \cdot P(\varepsilon, l)}{\sum_{l=l_{min}}^{N} P(\varepsilon, l)} \; ; \tag{1.2.19}$$

esta medida es el tiempo promedio en que dos segmentos de la trayectoria permanecen "cercanos´´, y puede ser interpretado como el tiempo de predicción promedio.

c) *Entropía de la distribución de líneas diagonales ENTR*

$$ENTR = -\sum_{l=l_{min}}^{N} P(\varepsilon, l) \ln P(\varepsilon, l) \; . \tag{1.2.20}$$

La medida entropía se refiere a la entropía de Shannon de la probabilidad $p(l) = P(l)/N_l$ de encontrar una línea diagonal de longitud l en el RP. ENTR refleja la complejidad del RP con respecto a las líneas diagonales, por ejemplo para ruido no correlacionado el valor de ENTR sera pequeño, indicando una baja complejidad.

En esta tesis los diagramas de recurrencia se utilizan para estudiar la aleatoriedad en PRNGs generados por mapas caóticos $1\mathcal{D}$ pero el método puede extenderse a espacios de dimensión $\mathcal{D}$ o bien a espacios de embedding. En el capítulo 2 se dan resultados de estos cuantificadores aplicados a distintas series temporales.

1.2.7. Cuantificadores basados en computación intrínseca

Los cuantificadores introducidos en [32], el *ritmo de entropía* h_μ y el *exceso de entropía* $\mathbf{E}$. Están definidos para una serie temporal con un alfabeto finito $\mathcal{A}$, lo que no constituye una limitación ya que nuestras series temporales son *números de punto flotante* con un número finito de valores permitidos. Una subsecuencia $s^L = \{x_i, x_{i+1}, \cdots, x_{i+L}\}$ se denomina bloque-L. Si $P(s^L)$ define la probabilidad de un

bloque-L en especial, entonces la entroía $H(L)$ está dada por:

$$H(L) \equiv -\sum_{s^L} P(s^L)\, log_2\, P(s^L)\ . \tag{1.2.21}$$

La sumatoria se realiza sobre todos los bloques posibles de longitud $L > 0$ y $H(0) \equiv 0$ por definición. Para procesos estacionarios y valores de L suficientemente grandes, $H(L) \sim L$. Además el *ritmo de entropía* h_μ se define como

$$\begin{aligned} h_\mu(L) &= \frac{H(L)}{L}\ , \\ h_\mu &= \lim_{L\to\infty} \quad h_\mu(L)\ . \end{aligned} \tag{1.2.22}$$

El *ritmo de entropía* también se conoce como *entropía métrica* en la teoría de sistemas dinámicos, y es equivalente a la *densidad de entropía termodinámica* familiar en los problemas de equilibrio de la mecánica estadística [32]. El ritmo de entropía proporciona una medida confiable y bien comprendida de la aleatoriedad o desorden intrínseco de un proceso. Sin embargo nos dice muy poco de la organización de las estructuras o las correlaciones. Puede obtenerse una medida de la organización del sistema viendo la forma en que $h_\mu(L)$ converge hacia su valor asintótico h_μ. Si se toman en cuenta sólo los bloques de una dada longitud L el proceso tiene un ritmo de entropía $h_\mu(L)$ que es superior al valor asintótico h_μ. Esa diferencia, sumada para todas las longitudes L, es lo que se denomina *exceso de entropía* [33]:

$$\mathbf{E} \equiv \sum_{L=1}^{\infty} \left[h_\mu(L) - h_\mu\right]. \tag{1.2.23}$$

Capítulo 2

Randomización de sistemas caóticos mediante dinámica simbólica

2.1. Introducción

Dada una serie temporal generada por un sistema físico se denomina dinámica simbólica a la evolución temporal obtenida asignando un símbolo a uno o varios valores de la serie.

En el caso de señales generadas por sistemas deterministas el teorema de Takens demuestra que si el sistema tiene dimensión n es posible reconstruirlo a partir del conocimiento de la serie temporal generada por una única variable de estado. El método para realizar esa reconstrucción se conoce como "reconstrucción en el espacio de embedding". Con este método puede reconstruirse un sistema aún sin conocer su dimensión e incluso obtener cuál es esa dimensión, a partir de esa única variable. A partir del sistema reconstruido es posible obtener una serie de parámetros invariantes a la transformación de embedding y que brindan información sobre el sistema generador.

Los sistemas caóticos generados por mapas unidimensionales simples poseen un

atractor en el espacio de embedding que claramente muestra la no aleatoriedad de la serie temporal. En las aplicaciones de espectro esparcido que se investigan en esta tesis las exigencias de aleatoriedad y distribución espectral de las señales a ser utilizadas no permite en general adoptar esas señales en forma directa y es necesario someterlas a un proceso que denominamos de "randomización".

Todos los procesos de randomización propuestos en esta tesis constan de la reconstrucción en un espacio de embedding seguida de una asignación de símbolos a cada punto de ese espacio. El objetivo es que la secuencia de símbolos obtenida tenga mejores características estadísticas que la serie original, en relación a las exigencias ya mencionadas.

En este capítulo se presentan dos aplicaciones básicas: (a) la clasificación de sistemas en caóticos o estocásticos[34, 35]; (b) el estudio de generadores de números pseudo aleatorios basados en mapas caóticos y las técnicas de randomización aplicables para mejorar su calidad estadística [36, 37].

2.2. Sistemas caóticos y estocásticos

A pesar de provenir de fuentes muy distintas, las series temporales que provienen de sistemas caóticos (CS) comparten con aquéllas generadas por procesos estocásticos (SP) varias propiedades que las hacen casi indistinguibles:

- un espectro de potencias (PS) de banda ancha,
- una función autocorrelación tipo delta y
- un comportamiento temporal irregular

De hecho, esta similitud es la que ha hecho posible reemplazar SP por CS en muchas aplicaciones prácticas. Aquí intentamos distinguir entre SP y CS por medio de una representación adecuada, que destaca el papel de la así denominada medida de complejidad. Se trabaja con modelos bien conocidos que generan series temporales conforme a reglas preestablecidas. Es decir, la situación bajo estudio es diferente de la situación con datos reales en que no está probado que exista un modelo determinista (aunque es un presupuesto que sí dado que en física creemos que siempre existen leyes que gobiernan los fenómenos). Además los datos reales siempre poseen una componente de ruido [38, 39].

De hecho, Wold demostró [38] que cualquier serie temporal estacionaria se puede descomponer en la suma de dos partes. La primera (que es determinista) se puede describir por medio de una combinación lineal del pasado, en tanto que la segunda parte es un promedio móvil (moving average) de orden finito. Con esta idea parece superfluo preguntarse si una serie temporal generada por "procesos naturales" es determinista, caótica o estocástica. No obstante, tomando en cuenta el teorema de Wold [39] tiene sentido preguntarse respecto a la parte determinista, si *(i)* es dominante respecto de la parte impredecible y *(ii)* es de naturaleza regular o caótica. *CS siempre producen series temporales con estructura física.* El poder descubrir estas estructuras escondidas es la clave de un método de clasificación que permita separar los sistemas pseudo caóticos de los pseudo estocásticos.

Para descubrir estas estructuras se utiliza en este capítulo uno de los miembros de la familia de complejidades estadísticas generalizadas definidas en el capítulo 1. Por simplicidad de notación designaremos esta complejidad con la sigla $C_{JS}[P]$, que es una funcional de la distribución de probabilidad P asociada a la serie temporal.

La expresión de C_{JS} es:

$$C_{JS}[P] \;=\; Q_J[P, P_e] \cdot H_S[P] \;, \tag{2.2.1}$$

y está asociada con la distribución de probabilidad $P = \{p_j \;;\; j = 1, \cdots, N\}$, con la entropía normalizada

$$H_S[P] = S[P]/S_{max} \tag{2.2.2}$$

en la que $S_{max} = S[P_e] = \ln N$, $(0 \leq H_S \leq 1)$.

Aquí se ha designado $P_e = \{1/N, \cdots, 1/N\}$ a la *distribución uniforme* en tanto que $S[P] = -\sum_{j=1}^{N} \; p_j \; \ln(p_j)$ es la entropía de Shannon.

El espacio de probabilidades puede ser estructurado como espacio métrico utilizando distintos tensores métricos, cada uno de ellos da lugar a una definición diferente de desequilibrio Q (distancia a la distribución equiprobable). Para el análisis de sistemas estocásticos la distancia conocida como divergencia de Jensen-Shannon induce una métrica cuadrática en el espacio de las distribuciones de probabilidad, a diferencia de la divergencia de Kullback-Leiber [30]) definida en 2.3.9. El desequilibrio Q_J resulta entonces una magnitud intensiva que es diferente de cero sí y sólo si existen estados "privilegiados", o "más probables", en el espacio de estados accesibles [23, 26, 30], y permite distinguir entre diferentes grados de periodicidad.

Un punto importante en la evaluación de la SCM C_{JS} es elegir la distribución de probabilidades P asociada a la serie temporal. Este aspecto no ha sido analizado en la literatura pero merece una consideración especial debido a que la distribución de probabilidades no es única y está ligada estrechamente al espacio explorado (Ω, P) (con Ω espacio de probabilidades).

Se han empleado varias aproximaciones en la literatura de forma de "extraerla" de una dada serie temporal. Sólo para mencionar algunos procedimientos de extracción

comúnmente utilizados: *a)* distribuciones basadas en el histograma [40], *b)* en la dinámica simbólica binaria [41], *c)* en el análisis de Fourier [42], *d)* en la transformada wavelet [43, 44], *e)* entropías de partición [45], *f)* entropías de permutación [46, 47], *g)* entropías discretas [48], etc.

PDF basada en el histograma:

Para obtener la PDF basada en el histograma se divide el intervalo en un número finito n_{bin} de subintervalos sin intersección A_i: $[0,1] = \bigcup_{i=1}^{n_{bin}} A_i$ y $A_i \bigcap A_j = \emptyset \; \forall i \neq j$. Se emplea luego el método usual basado en contar la frecuencia relativa de los distintos valores dentro de cada subintervalo. Esta PDF carece que toda información que tenga que ver con la evolución temporal del sistema .

PDF basada en una dinámica simbólica binaria:

Para cada valor del parámetro r se reduce la secuencia a una secuencia binaria (0 si $x \leq \frac{1}{2}$; 1 si $x > \frac{1}{2}$). Luego se forman cadenas de $L = 12$ bits sin superposición y se realiza el histograma de las cadenas obtenidas.

En este caso la PDF retiene cierta información temporal media, porque al realizar la asignación del bit 0 ó 1 (paso 1 del procedimiento se pierde gran parte de la evolución temporal pero al utilizar cadenas de L bits sin superposición (paso 2 del procedimiento) se recupera cierta información temporal de orden de recorrido del intervalo.

Pero su aplicabilidad depende de características particulares de la serie temporal tales como su estacionariedad, la longitud de la serie, la variación de los parámetros, el nivel de ruido, etc. Es importante notar que la elección de la distribución ha sido

en casi todos los casos arbitraria.

Cuando lo que se desea es diferenciar sistemas caóticos de sistemas estocásticos no es posible utilizar distribuciones de probabilidad que no tengan en cuenta la dinámica del sistema. Por ejemplo una serie temporal consistente en los números 1 a 256 repetidos en forma periódica produce un histograma plano cuando se toman 256 bins. Y también una serie generada por números al azar produce un histograma plano. Es decir el histograma sólo cuenta el número de veces en que aparece cada elemento de la serie, pero no tiene en cuenta el orden de los mismo.

PDF de Bandt y Pompe:

Bandt y Pompe [49] hicieron notar que es conveniente utilizar un espacio de embedding D-dimensional y además generar una serie simbólica que tenga en cuenta el orden de aparición de los valores.

Dada la serie temporal $\{x_t : t = 1, \cdots, M\}$ y una dimensión de embedding $D > 1$, interesa obtener los "patrones de orden" para dimensión D [49, 50] generados por:

$$(s) \;\mapsto\; \left(x_{s-(D-1)},\; x_{s-(D-2)},\; \cdots,\; x_{s-1},\; x_s \right) , \tag{2.2.3}$$

donde a cada instante s asignamos un vector D-dimensional de valores correspondientes a los instantes $s, s-1, \cdots, s-(D-1)$. La expresión "patron de orden correspondiente al instante s" se denomina a la permutación $\pi = (r_0, r_1, \cdots, r_{D-1})$ de $(0, 1, \cdots, D-1)$ definida por:

$$x_{s-r_{D-1}} \;\leq\; x_{s-r_{D-2}} \;\leq\; \cdots \;\leq\; x_{s-r_1} \;\leq\; x_{s-r_0}. \tag{2.2.4}$$

Con el objeto de obtener un resultado único se considera que $r_i < r_{i-1}$ si $x_{s-r_i} = x_{s-r_{i-1}}$. Así, para todas las $D!$ permutaciones posibles, π de orden D, la distribución

de probabilidad $P = \{p(\pi)\}$ está definida por $(\mathcal{Y} = M - D + 1)$

$$p(\pi) \; = \; \sharp\{s | s \leq \mathcal{Y};\; (s) \texttt{ es de tipo } \pi\} \; / \; \mathcal{Y} \; . \tag{2.2.5}$$

En esta expresión, el símbolo $\sharp$ significa "número". El método de Bandt y Pompe [49] para evaluar la distribución de probabilidad P está basado en detalles del procedimiento de reconstrucción del atractor y por lo tanto *contiene información causal* que hace que $P \in \Omega$ [51]. Un resultado notable de Bandt y Pompe es la clara mejora que produce en los cuantificadores estadísticos obtenidos empleando su distribución de probabilidad P_{BP}.

Se debe suponer que el sistema cumple con un requisito muy débil de estacionariedad: para $k \leq D$, la probabilidad de que $x_t < x_{t+k}$ debe ser independiente de t [49] y se debe contar con suficiente número de datos para una correcta reconstrucción del atractor en la dimensión D elegida. Las ventajas del método de Bandt y Pompe residen en su simplicidad, que lleva a un cálculo muy rápido, y su robustez frente a transformaciones monótonas no lineales. Se lo puede aplicar a cualquier tipo de serie temporal, regular, caótica, ruidosa, etc. [49]. Finalmente es importante remarcar que los cálculos realizados con la prescripción de Bandt y Pompe son robustos en presencia de ruido dinámico o de ruido experimental [49]. Por supuesto la dimensión de embedding elegida D juega un papel importante en la evaluación adecuada de la distribución de probabilidades ya que D determina el número de estados accesibles $D!$ y nos informa acerca de la longitud M de la serie temporal necesaria para poder trabajar en forma confiable. En relación con este último punto, todos los cálculos reportados aquí satisfacen la condición $M \gg D!$. En particular Bandt y Pompe sugieren por razones prácticas trabajar con $3 \leq D \leq 7$, y eso es lo que se ha hecho en esta tesis. La mayoría de los cálculos han sido realizados con $D = 6$. Si se permite que

D crezca sin límites surgen consecuencias significativas e importantes desde el punto de vista teórico.

Para cualquier sistema caótico la entropía de BP tiende a la de Kolmogorov-Sinai [49] y se puede probar que en el límite,para dimensión de embedding infinita, la Complejidad tiende a cero.

Se debe destacar el hecho que la complejidad estadística definida arriba no sólo cuantifica la aleatoriedad sino también el grado de correlación entre estructuras y en consecuencia *no es una función trivial de la entropía*, en el sentido que, para un dado valor de H_S, existe un rango de valores posibles de C_{JS} comprendidos entre un valor mínimo $C_{\min}$ y un valor máximo $C_{\max}$ [23]. Por lo tanto la evaluación de C_{JS} proporciona información *adicional* en función de las particularidades de la distribución de probabilidades. Esta información adicional no aparece si únicamente se trabaja con la PDF basada en el histograma y con la correspondiente entropía. Un procedimiento general para obtener los límites $C_{\min}$ y $C_{\max}$ para la familia de complejidades estadísticas está desarrollado en [26].

Se podría asimismo trabajar con Q_J, en lugar de hacerlo con C_{JS} pero C_{JS}tiene las siguientes ventajas: (1) *valor cero* tanto para series regulares como para series aleatorias y (2) *máxima* para sistema que presentan estructuras "inmersas" (o escondidas) [26].

Para estudiar la evolución temporal de C_{JS}, se puede utilizar un diagrama C_{JS} versus H_S (en este plano H_S puede ser considerado como el eje de tiempo [52]).

Esta clase de diagramas también ha sido empleado para estudiar los cambios en la dinámica de un sistema al modificarse algunos de sus parámetros [23, 30, 53, 54, 25, 55]. Los procesos que se estudian en esta sección son ejemplos ilustrativos de *(a)*

CS y *(b)* SP, dos clases de procesos que puede pensarse que son diferentes en función de lo dicho al inicio de la sección.

Entre los CS se estudian a continuación:

(1) El mapa logístico [56] dado por:

$$x_{n+1} \;=\; r\; x_n\; (1 - x_n)\;. \tag{2.2.6}$$

Para $r = 4$ este mapa tiene una PDF invariante natural que no es uniforme.

(2) El mapa Skew Tent (carpa inclinada): [56]

$$\begin{cases} x/\omega & \text{for } x \in [0, \omega] \\ (1-x)/(1-\omega) & \text{for } x \in [\omega, 1] \end{cases}\;. \tag{2.2.7}$$

Para cualquier valor de ω este mapa tiene una PDF invariante uniforme. Los resultados se muestran para ($\omega = 0{,}1847$).

(3) El mapa de Henon: es una extensión $2D$ del mapa logístico [56] dada por:

$$\begin{cases} x_{n+1} \;=\; 1 \;-\; a\; x_n^2 \;+\; y_n \\ y_{n+1} \;=\; b\; x_n \end{cases}\;. \tag{2.2.8}$$

Los valores usados ahí son, $a = 1{,}4$ y $b = 0{,}3$, y corresponden a un atractor caótico con una PDF no suave.

(4) El mapa de Lorenz del Oscilador de Rossler: el oscilador continuo $3D$ de Rossler [56] está dado por:

$$\begin{cases} \dot{x} \;=\; -\,y \;-\; z \\ \dot{y} \;=\; x \;+\; a\; y \\ \dot{z} \;=\; b \;+\; z\; (x - c) \end{cases}\;. \tag{2.2.9}$$

donde $a = 0{,}2$, $b = 0{,}2$, y $c = 5{,}7$ corresponde a un atractor caótico. El mapa de Lorenz se obtiene almacenando únicamente los valores mínimos locales de la variable x [56].

(5) Schuster Maps: Schuster y sus colaboradores [56] introdujeron una clase de mapas que genera señales intermitentes, con ráfagas caóticas y que también producen ruido de tipo $1/f^z$:

$$x_{n+1} \;=\; x_n \;+\; x_n^z, \qquad \text{Mod 1.} \tag{2.2.10}$$

En particular, se estudian series correspondientes a $z = 5/2$, 2 y $3/2$.
En cuanto a los ruidos estocásticos considerados (SP) ellos son:

1. Ruidos con espectro de potencia (PS) de la forma f^{-k} generados del siguiente modo:

   - Mediante la función *RAND* de MATLAB© se genera una secuencia de números pseudo aleatorios en el intervalo $(-0{,}5, 0{,}5)$ con: *(a)* un PS casi plano, *(b)* una PDF uniforme, y *(c)* valor medio cero.
   - Se calcula la transformada rápida de Fourier (FFT) a la serie, obteniendo el vector y_k^1. Se multiplica el vector por $f^{-k/2}$, obteniendo así el vector y_k^2;
   - Luego y_k^2 es simetrizado para obtener una función real al aplicar la transformada inversa IFFT, para obtenerse x_i , luego de descartar los pequeños valores imaginarios producidos por las aproximaciones numéricas. La serie final x_i tiene el PS deseado y es representativa de ruidos no gaussianos.

2. Movimientos brownianos fraccionarios (fBm) y ruidos Gaussianos fraccionarios (fGn): fBm es la única familia de procesos estocásticos que es:

   - gaussiana,
   - auto-similar,

- dotado de incrementos estacionarios (ver Ref. [55] y la referencias incluidas en este trabajo).

La familia normalizada de estos procesos Gaussianos, $\{B^{\mathcal{H}}(t), t > 0\}$, está dotada de estas propiedades:

- *i)* $B^{\mathcal{H}}(0) = 0$ con probabilidad 1 ,
- *ii)* $\mathbb{E}[B^{\mathcal{H}}(t)] = 0$ (valor medio cero), y
- *iii)* covarianza dada por

$$\mathbb{E}[B^{\mathcal{H}}(t)B^{\mathcal{H}}(s)] \;=\; (t^{2\mathcal{H}} + s^{2\mathcal{H}} - |t-s|^{2\mathcal{H}}) \;/\; 2 \tag{2.2.11}$$

  para $s, t \in \mathbb{R}$.

Aquí $\mathbb{E}[\cdot]$ se refiere al valor medio computado con una PDF Gaussiana. El exponente $0 < \mathcal{H} < 1$ se conoce generalmente como parámetro o exponente de Hurst. Estos procesos presentan "*memoria*" para cualquier valor del parámetro de Hurst *excepto para* $\mathcal{H} = 1/2$, como se ve de la Eq. (2.2.11). El caso $\mathcal{H} = 1/2$ corresponde al movimiento browniano clásico, donde los movimientos son incrementos sucesivos que tienen igual probabilidad de mantener el mismo signo como de cambiarlo (es decir no hay correlación entre ellos). Por lo tanto el parámetro de Hurst define dos regiones diferentes en el intervalo $(0, 1)$. Cuando $\mathcal{H} > 1/2$, los incrementos consecutivos tienden a tener el mismo signo y los procesos se denominan *persistentes*. Por otro lado si $\mathcal{H} < 1/2$, es más probable que se produzca el cambio de signo en incrementos consecutivos, y entonces se dice que los procesos son *anti-persistentes*. En la siguiente Eq. 2.2.12 se define

una cantidad $\{W^{\mathcal{H}}(t), t > 0\}$ ("incrementos" fBm)

$$W^{\mathcal{H}}(t) \;=\; B^{\mathcal{H}}(t+1) - B^{\mathcal{H}}(t), \tag{2.2.12}$$

permitiendo expresar el ruido Gaussiano en la forma

$$\begin{aligned} \rho(k) \;&=\; \mathbb{E}[W^{\mathcal{H}}(t) W^{\mathcal{H}}(t+k)] \\ &=\; \frac{1}{2}\left[(k+1)^{2\mathcal{H}} - 2k^{2\mathcal{H}} + |k-1|^{2\mathcal{H}}\right],\ k > 0. \end{aligned} \tag{2.2.13}$$

Nótese que para $\mathcal{H} = 1/2$ todas las correlaciones para retardos distintos de cero y para $\{W^{1/2}(t), t > 0\}$ representan *ruido blanco*. Los fBm y los fGn son procesos continuos no diferenciables en sentido clásico. Como son procesos no estacionarios no se puede definir su espectro en la forma usual; sin embargo es posible definir un *espectro de potencia generalizado* de la forma $\Phi \;\propto\; |f|^{-\alpha}$, con $\alpha = 2\mathcal{H} + 1$, $1 < \alpha < 3$ para los fBm y, $\alpha = 2\mathcal{H} - 1$, $-1 < \alpha < 1$, para los fGn. Debido a su naturaleza Gaussiana, y a las otras características arriba enumeradas, las ideas de Bandt-Pompe son aplicables a estos procesos dinámicos [57].

Para evaluar las series temporales generadas por los fBm y los fGn adoptamos el algoritmo de Davies-Harte [58], con las mejoras introducidas por Wood y Chan [59]. Estos algoritmos son exactos y rápidos.

Para todos los casos estudiados se emplean 10 series temporales de 2^{15} datos cada una. Cada serie comienza con una condición inicial diferente. Los valores medios de H_S y C_{JS} son graficados en la Fig.2.1.

Todos los CS analizados presentan entropías en la región entre 0,45 y 0,7, en el plano CH. Están ubicados cerca de la curva máxima C_{JS}. La razón es que valores

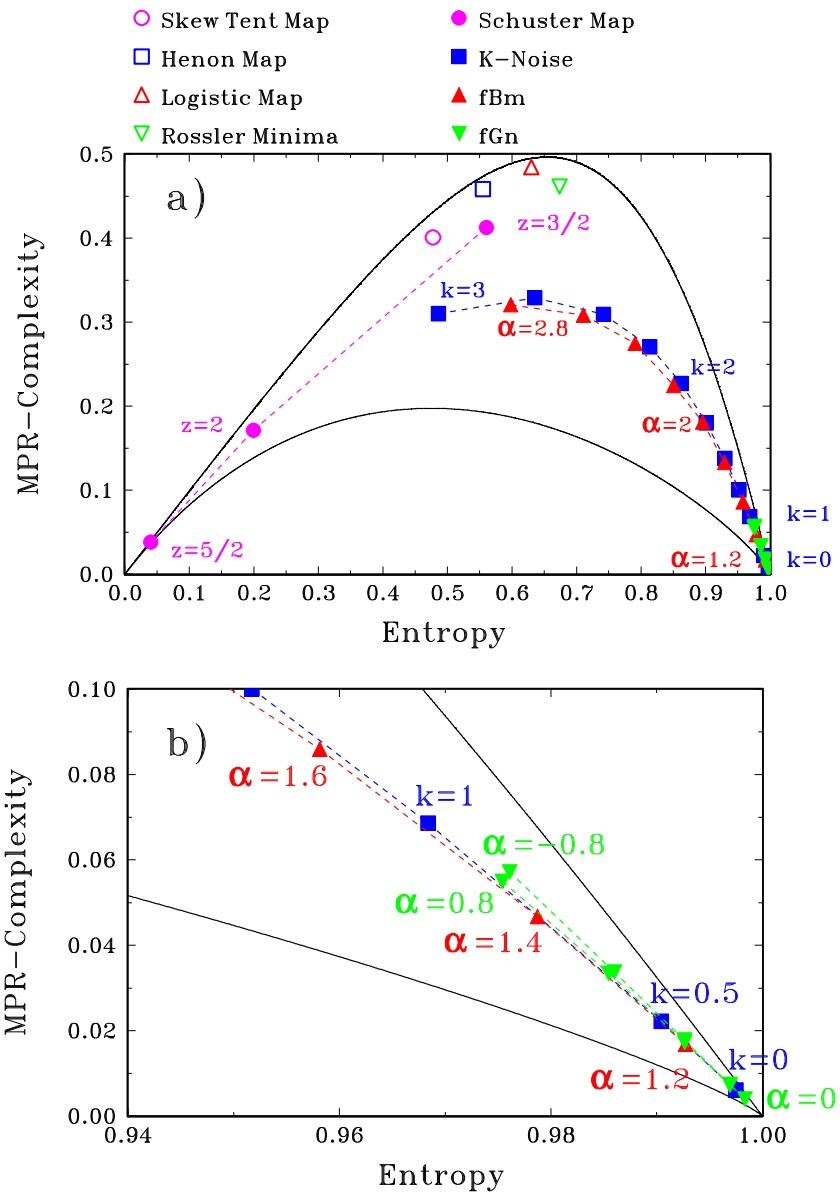

Figura 2.1: Las líneas continuas representan los valores mínimos C_{min} y máximo C_{max} para la complejidades a una entropía dada. El área encerrada entre ellas es el plano CH. *(a)* Ubicación de los distintos CS y SP en el plano CH. *(b)* Ampliación cerca del punto $H_S = 1$, $C_{JS} = 0$. Se trabajó con un espacio de embedding de dimensión $D = 6$. El gráfico ilustra el hecho que, en el caso de todos los modelos estudiados son clasificados en la bibliografía como deterministas o estocásticos [56]. Y esos procesos quedan ubicados en regiones separadas del plano CH

altos de C_{JS} corresponden a series temporales con estructuras inmersas, tal como ocurre con las series caóticas. Para incrementar estos valores de H_S se deben utilizar técnicas de randomización, que son estudiadas en este capítulo. Estas técnicas de randomización tienen por objeto incrementar el mixing y destruir las estructuras de correlación.

En el caso del mapa logístico podemos ver en el plano CH el efecto del parámetro de control $r < 4$. Los puntos se ubican siempre en el sector de baja entropía, y en todos los casos cerca de la curva de complejidad máxima, de acuerdo con lo que ocurre en el caso de la PDF binaria [30]. Para el mapa de Hénon ambas coordenadas, X e Y, tienen la misma estructura de orden como lo muestra el hecho que ocupan el mismo punto en el plano CH. En cuanto a los mapas de Schuster también presentan bajas entropías en la región $H_S < 0{,}6$. En este caso la causa es que estos mapas muestran regiones de comportamiento laminar separadas por ráfagas caóticas. Su complejidad es menor que la de los sistemas caóticos. A medida que disminuye el parámetro z aumenta H_S debido al ensanchamiento en el tiempo de las ráfagas caóticas. La trayectoria se mantiene siempre por debajo de los sistemas caóticos pero cercana a ellos.

Los ruidos con PS f^{-k} y $0 \leq k \leq 3$, muestran valore intermedio de entropía ($0{,}45 < H_S < 1$) en tanto que C_{JS} tiene valores casi equidistantes entre los de las curvas de máxima y mínima complejidad. En particular, su C_{JS} es mucho menor que la de los ruidos deterministas (como los ruidos de Schuster y los CS). Para valores pequeños de k ($k = 0$ y $k = 1$) se convierten en ruidos casi ideales con $H_S \simeq 1$ y $C_{JS} \simeq 0$. A medida que se incrementa k aparecen correlaciones y eso produce una disminución de la entropía H_S. fBm ($1 < \alpha < 3$) tienen entropías cercanas a la de los ruidos de PS f^{-k}, pero con una menor C_{JS} si se los compara con los procesos no Gaussianos. Los

fGn asociados ($-1 < \alpha < 1$) tiene mayores valores de entorpía ($0{,}97 < H_S < 1$) y sus valores de C_{JS} se encuentran entre 0 y 0,1. Además estos valores son mayores que los de los procesos con PS f^{-k} (ver Fig. 2.1.b). Se puede asociar este comportamiento a la naturaleza Gaussiana y No Gaussiana de los respectivos procesos.

El movimiento Browniano ordinario ($\alpha = 2$) se caracteriza por una entropía relativamente baja y una complejidad relativamente alta ($H_S \simeq 0{,}9$ y $C_{JS} \simeq 0{,}18$). Los fBm persistentes de memoria larga ($2 < \alpha < 3$) son más complejos que los antipersistentes de memoria corta ($1 < \alpha < 2$) en concordancia con el comportamiento esperado desde el punto de vista intuitivo. Los valores de complejidad de los fGn son mayores que los correspondientes a ruidos con PS f^{-k}. Los fGn, persistentes y antipersistentes presentan valores muy similares (ver Fig. 2.1.b). Los valores de máxima entropía y mínima complejidad de todos los ruidos observados son los correspondientes a $\alpha = 0$, es decir al ruido blanco Gaussiano. Es interesante notar que en el plano CH esta situación corresponde al punto que se encuentra por debajo del correspondiente al ruido con PS de tipo f^{-k} y $k = 0$.

En resumen, en el plano CH en que ambos ejes se han obtenido utilizando la distribución de probabilidad generada por la dinámica simbólica de BP, las señales generadas por modelos bien conocidos han quedado clasificados de modo que ruido (o SP) y caos (o CP) se ubican en *posiciones diferenciadas en el plano*. Esta propiedad de la representación CH puede ser útil al trabajar con *datos reales* en los que se espera encontrar siempre una mezcla de determinismo y ruido de modo de medir el "grado de estocasticidad".

Es también interesante que la representación CH permite también distinguir *a)* procesos Gaussianos de no Gaussianos y, *b)* variaciones en el grado de correlación (ruidos

coloreados).

2.2.1. Resultados para el mapa logístico

El mapa logístico es un ejemplo paradigmático empleado a menudo para verificar nuevos conceptos que se aplican a los sistemas dinámicos. Es interesante entonces analizar el resultado de calcular las distintas SCMs (ver Eq. (1.2.14)) para el caso del mapa logístico.

Revisemos brevemente (ver figuras 2.2.a y 2.2.b) algunos resultados bien conocidos para este mapa que nos permiten poner en perspectiva las propiedades de nuestra familia de SCM's. Para valores del parámetro de control $1 < r < 3$ sólo existe una solución estacionaria. Al incrementar el parámetro de control por encima de $r = 3$ el sistema sufre una bifurcación de duplicación de período. Aparecen ciclos de períodos 8, 16, 32, $\cdots$ y si r_n indica el valor de r donde nace el ciclo de período 2^n, la sucesión de valores r_n converge a un valor límite $r_\infty \cong 3{,}57$ [60, 61]. A medida que r crece por encima de este valor aparece una estructura rica y bien conocida. Para poder apreciar a golpe de vista el comportamiento del mapa logístico, una vez pasado el transitorio, se grafica en la Fig. 2.2.a los valores recorridos por la serie temporal generada para todos los valores de r entre 3,5 y 4,0. Se nota de modo inmediato una cascada de duplicaciones de período hasta que en r_∞ el mapa se vuelve caótico y las series pasan a tener un número infinito de puntos, cubriendo un intervalo. Para $r > r_\infty$ el diagrama de bifurcación revela una extraña mezcla de orden y caos. La ventana que comienza cerca de $r = 3{,}83$ contiene una órbita de período-3 estable. En la Fig. 2.2.b se ve la evolución del exponente de Lyapunov Λ, que permanece negativo para $r_\infty \cong 3{,}57$. Nótese que Λ tiende a cero en los puntos de duplicación de doble

período. La aparición de caos cerca de $r \cong 3{,}57$, se pone en evidencia porque Λ se vuelve positivo por primera vez por encima de este valor. Como se explicó arriba para $r > 3{,}57$ el exponente de Lyapunov crece salvo por profundos pozos correspondientes a las ventanas de comportamiento periódico. La ventana de período-3 corresponde a un pozo especialmente notable ubicado cerca de $r = 3{,}83$.

De todas las opciones mencionadas y resumidas en la Eq. 1.2.14 se muestra en las figuras los resultados correspondientes a *(i)* histogramas, *(ii)* representación binaria y *(iii)* técnica de Bandt-Pompe. Nótese que si la evaluación de la PDF se realiza en el espacio de frecuencias se obtienen resultados similares. En el caso en que se emplea la transformada discreta wavelet se incluye en forma promediada información temporal y frecuencial (más detalles pueden verse en [25, 62]).

Los resultados que se muestran corresponden a series temporales con $N = 10^7$ *datos* generados con el parámetro r en el rango $3{,}4 \leq r \leq 4{,}0$ en pasos $\Delta r = 0{,}0003$. Para la evaluación del histograma se utilizó $n_{bin} = 2^{12} = 4096$ [36]. En el caso binario se utilizaron cadenas de $L = 12$ *bits* con el que el número de estados es 2^{12} y para el caso de la PDF de Bandt y Pompe, la dimensión del espacio de embedding fue $D = 6$ con lo que el número posible de estados es $6! = 720$.

En la Fig. 2.3 se muestra la entropía normalizada de Shannon , $\mathcal{H}_1^{(S)}$ como función del parámetro r para las tres distribuciones de probabilidad elegidas. Se observa que en todas las instancias se produce un crecimiento abrupto de la entropía cerca de $r > r_\infty \cong 3{,}57$. Luego de pasar este punto las entropías tienden a incrementarse y adquieren su valor máximo en $r = 4$. Las ventanas periódicas se corresponden con descensos bruscos en los valores de entropía y aparecen en $r_\infty < r < 4$ (ver Fig. 2.2). Es interesante observar que el caso de adoptarse la PDF del histograma o

la PDF binaria, la entropía adopta el valor $\mathcal{H}_1^{(S)} \cong 1$ para $r = 4$, que corresponde a caos totalmente desarrollado. Es decir que las PDFs obtenidas son uniformes. Por el contrario si se emplea la prescripción de Bandt y Pompe el valor de entropía, si bien es máximo se reduce a $\mathcal{H}_1^{(S)} \cong 0{,}6$. La razón es que aún en caos desarrollado la serie temporal tiene secuencias de orden que no aparecen (forbidden patterns) con lo que empleando la prescripción de Bandt y Pompe se obtiene una PDF no uniforme.

En la Fig. 2.4 se muestra la complejidad estadística $\mathcal{C}_{\mathcal{J},1}^{(S)}$ evaluada utilizando las tres diferentes metodologías de asignación de la PDF explicadas arriba. En todos los casos existe un crecimiento abrupto en $r > r_\infty$. Luego de superar este punto las tres complejidades se comportan de modo diferente. Si bien la complejidad decrece, permanece constante dentro de la ventanas periódicas (considere por ejemplo la ventana de período 3 $r \in [3{,}58,\ 3{,}62]$). Para ventanas periódicas diferentes se obtienen diferentes valores de complejidad de acuerdo al grado de la periodicidad. Si se emplea el histograma para evaluar la PDF se observa que para $r > r_\infty$, se produce una tendencia decreciente con un mínimo en $\mathcal{C}_{histo} \equiv \mathcal{C}_{\mathcal{J},1}^{(S)}\Big|_{histo} \cong 0$ para $r = 4$. Los numerosos picos que aparecen en la región $r_\infty < r < 4$ indican un crecimiento local de complejidad. Si se comprar con el diagrama de bifurcación (ver Fig. 2.2) puede notarse que estos picos se corresponden con las ventanas periódicas. Es decir señalan la transición entre comportamientos dinámicos, es decir de caótico a periódico.

Si se utilizan la PDF obtenida por la dinámica simbólica binaria se obtiene un comportamiento parabólico con el parámetro $r_\infty < r < 4$, con un valor $\mathcal{C}_{binary} \equiv \mathcal{C}_{\mathcal{J},1}^{(S)}\Big|_{binary} \cong 0$ para $r = 4$. Globalmente, el decaimiento de $\mathcal{C}_{binary}$ es el más rápido cuanto más nos aproximamos al valor $r = 4$. Nótese además que para esta región la complejidad estadística crece en las regiones inter-ventana y cae rápidamente dentro

de las ventanas periódicas.

Si se utiliza la PDF evaluada con la metodología de Bandt y Pompe $\mathcal{C}_{BP} \equiv \mathcal{C}^{(S)}_{\mathcal{J},1}\Big|_{BP}$ muestra un crecimiento general con un máximo para $r = 4$, que corresponde al caso de caos totalmente desarrollado. Nótese que en este caso puede distinguirse la periodicidad de las distintas ventanas periódicas [34]. Es interesante notar que la metodología de Bandt y Pompe permite distinguir que "caos no es ruido" aún cuando tienen características en común. El determinismo subyacente influye en el orden que se definen las probabilidades en este método. En la Fig. 2.6 se representa el plano CH para las tres opciones en la evaluación de la PDF y para r en el rango $3{,}4 \leq r \leq 4$. Las dos curvas continuas representan $\mathcal{C}_{max}$ y $\mathcal{C}_{min}$ para el correspondiente número de estados permitidos de cada representación. En el caso de las ventanas periódicas, si $\mathcal{H} < \mathcal{H}^* \approx 0{,}3$ podemos asegurar que el exponente de Lyapunov $\Lambda < 0$, mientras que para $\mathcal{H} > \mathcal{H}^*$ se observa que $\Lambda > 0$, indicando comportamiento caótico. Como se evidencia en la Fig. 2.6, en todas las instancias de evaluación de la PDF, los comportamientos periódicos presentan valores menores que los caóticos.

Conclusiones

Concluimos que es importante resaltar los siguientes aspectos si se utiliza la medida de complejidad estadística para caraterizar series temporales:

- Se requiere un cuidado especial en la selección de la distribución de probabilidades. Si es importante detectar el deterministo, la prescripción de Bandt y Pompe tiene ventajas sobre otras opciones.

- Una medida de complejidad estadística depende también de la selección de una distancia o desequilibrio $\mathcal{Q}$. Trabajos previos publicados por otros autores [30]

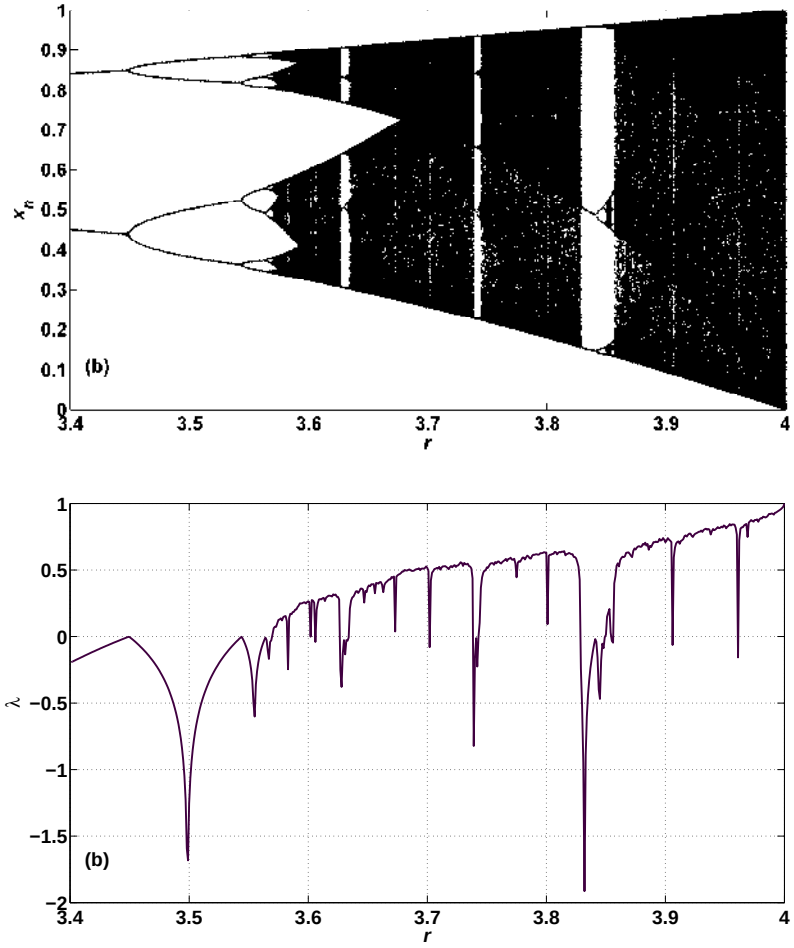

Figura 2.2: *a)* Diagrama de órbitas *b)* Exponente de Lyapunov (λ) para el mapa logístico como función del parámetro r con pasos $\Delta r = 0{,}0003$. Serie temporal con largo total de $M = 10^7$ *datos*.

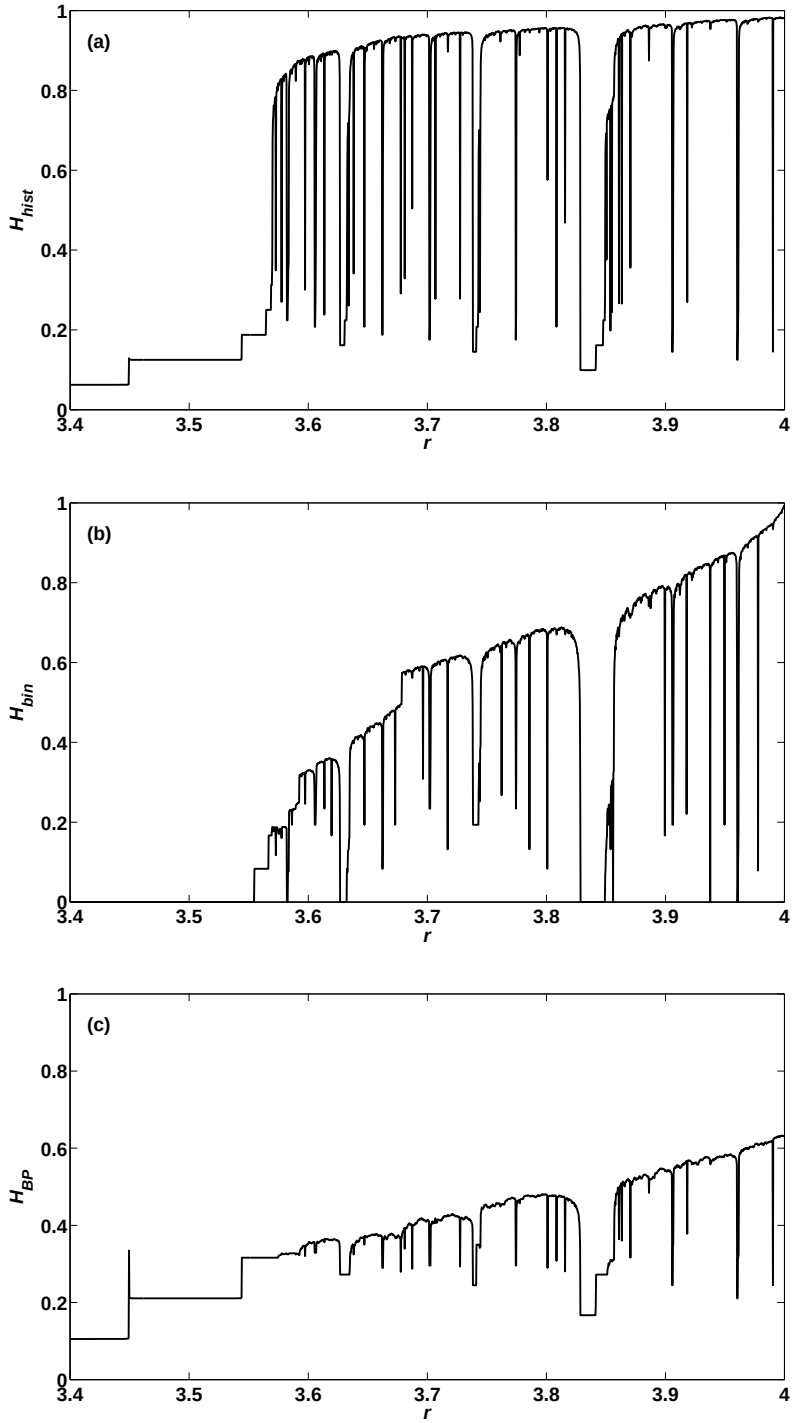

Figura 2.3: Entropía normalizada de Shannon para el mapa logístico como función del parámetro r con pasos $\Delta r = 0{,}0003$, (series temporales con longitud total $M = 10^7$ *datos*) calculada según: *a)* PDF basada en el histograma ($n_{bin} = 2^{12} = 4096$), *b)* PDF-binaria ($L = 12$), *c)* PDF-Bandt y Pompe ($D = 6$).

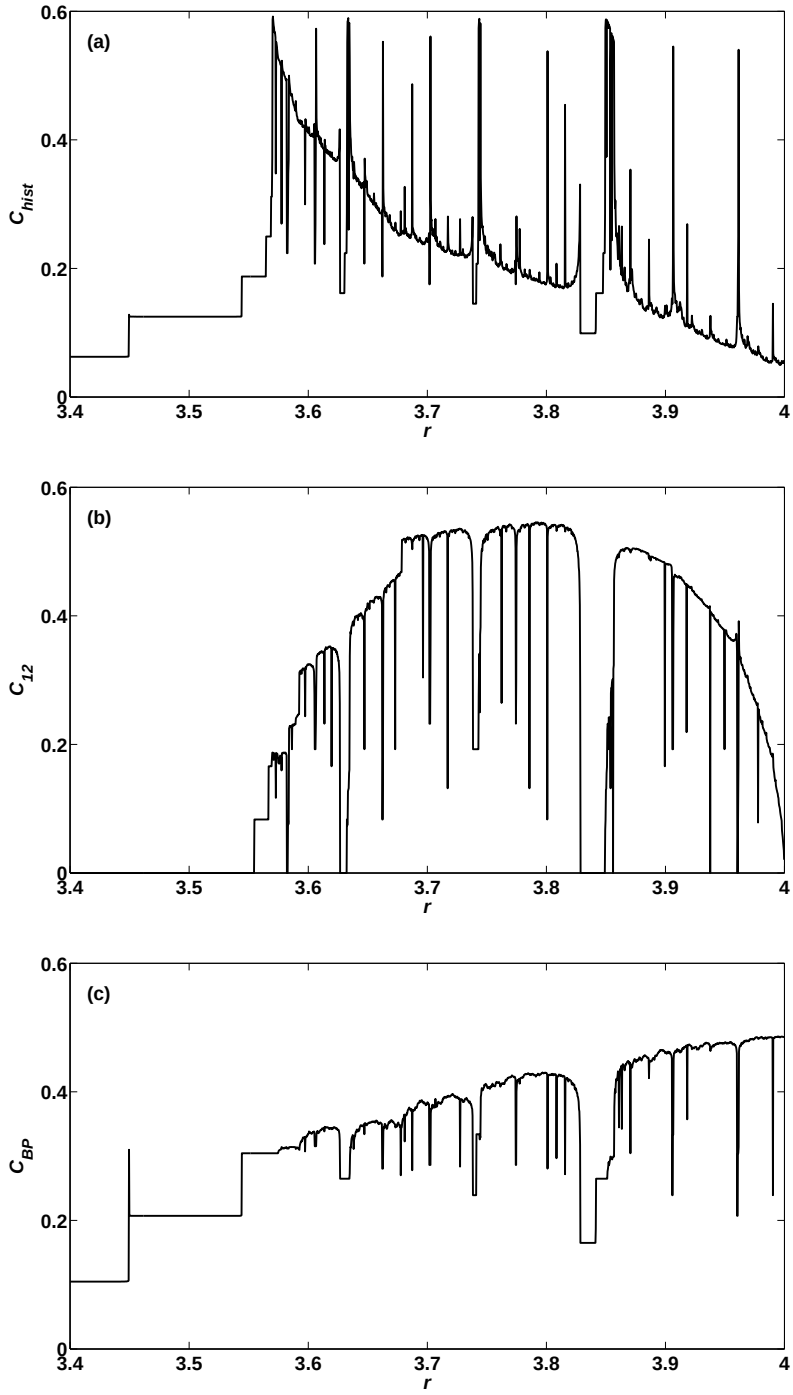

Figura 2.4: Complejidad estadística MPR para el mapa logístico como función del parámetro r con pasos $\Delta r = 0{,}0003$, (series temporales de longitud total $M = 10^7$ *datos*) evaluados según lo siguiente: *a)* histograma de PDF ($n_{bin} = 2^{12} = 4096$), *b)* PDF-binaria ($L = 12$), *c)* PDF-Bandt y Pompe, ($D = 6$).

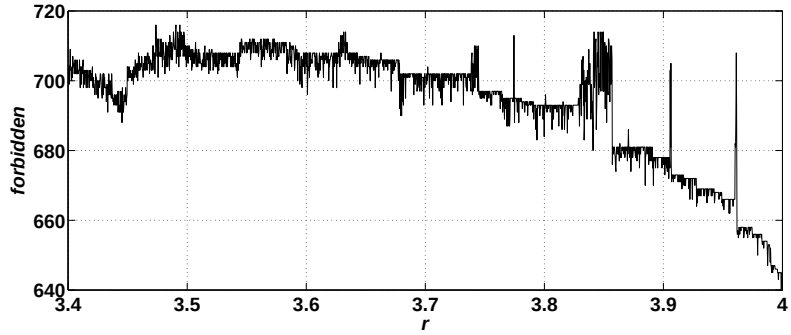

Figura 2.5: Patrones prohibidos para el mapa logístico como función del parámetro r con pasos $\Delta r = 0{,}0003$, (series temporales de longitud total $M = 10^7$ *datos*). Dimensión de embedding $D = 6$.

muestran que la divergencia de Jensen Shannon tiene las mejores propiedades, como ya se resaltó bajo la Eq. 2.2.1.

- Las series temporales caóticas quedan ubicadas cerca de la curva $\mathcal{C}_{max}$ en cualquiera de los planos de representación $\mathcal{H} \times \mathcal{C}$. Pero la prescripción de Bandt y Pompe ubica el caos más cerca de la cima, con elevada complejidad y valores de $\mathcal{H}$ cercanos a 0,5. Las otras prescripciones "no causalesübican a los sistemas deterministas más cerca de los aleatorios (cerca de $\mathcal{C} = 0$ y $\mathcal{H} = 1$).

- El número de patrones de orden prohibidos es otro cuantificador que detecta el comportamiento determinista dentro del caso completamente caótico $r = 4$. Sin embargo no exhibe una correspondencia con el diagrama de bifurcación como lo hace $\mathcal{C}_{BP}$.

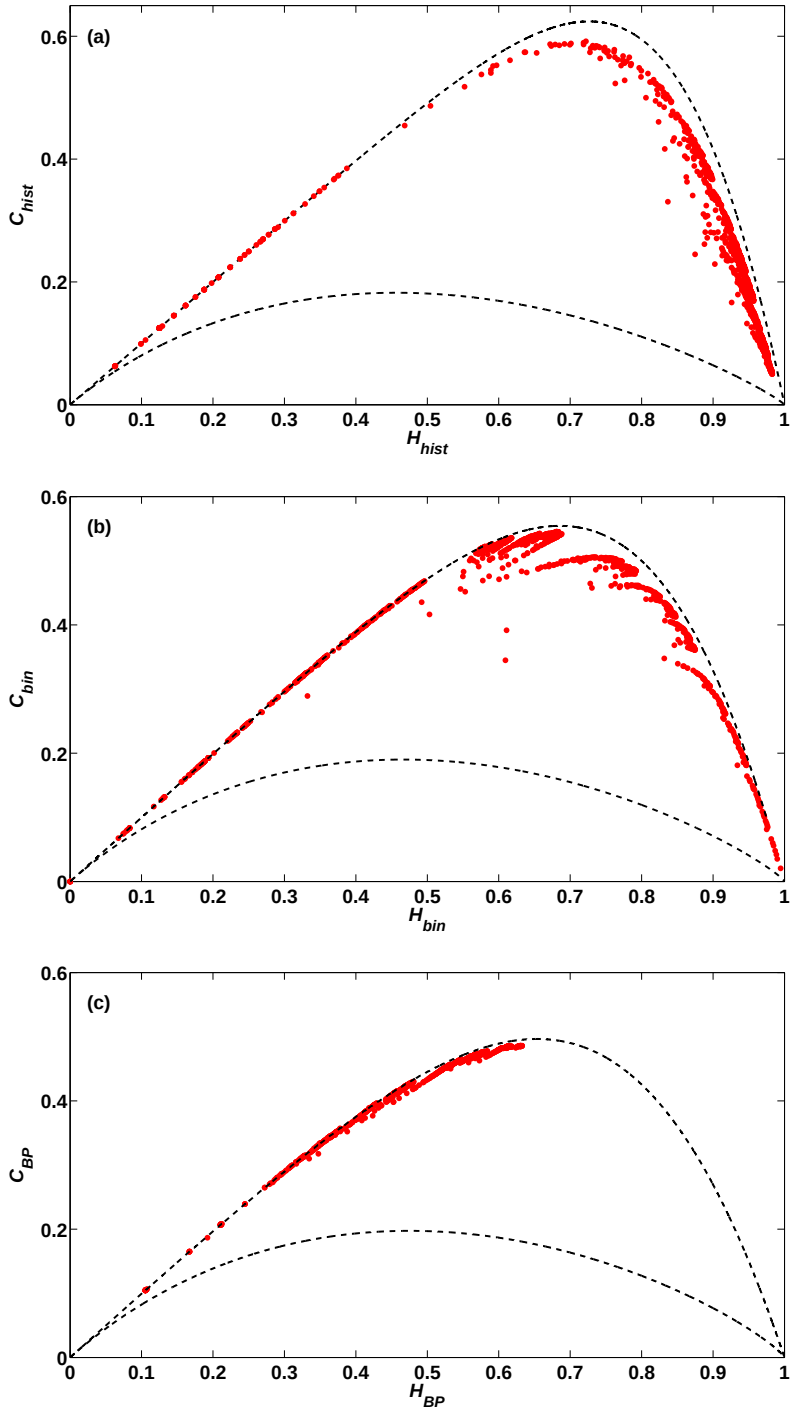

Figura 2.6: Plano Entropía-Complejidad para el mapa logístico (parámetro r con pasos $\Delta r = 0{,}0003$, series temporales de longitud total $M = 10^7$ *datos*) para: *a)* histograma de PDF ($n_{bin} = 2^{12} = 4096$), *b)* PDF-binaria ($L = 12$), *c)* PDF-Bandt y Pompe, ($D = 6$). También mostramos los posibles valores máximo y mínimo de complejidad estadística (línea entrecortada).

2.3. Generadores de números pseudo aleatorios y procesos de randomización

Tal como se mostró en las secciones anteriores el caos muestra que un comportamiento complejo puede aparecer de reglas deterministas simples cuando hay no-linealidades. Dado que los mapas caóticos (el mapa logístico mostrado en la sección anterior por ejemplo) son capaces de generar señales de aspecto estocástico y las implementaciones prácticas basadas en sistemas caóticos son más simples que las requeridas para sistemas estocásticos (ver 6) es natural que se procure utilizar el caos como generador de ruido en diversas aplicaciones [63, 64, 65].

Los PRNGs son usados no solo en criptografía y en el método de Monte Carlo sino también en aplicaciones menos obvias [66, 67, 68, 69, 70]. En esta tesis se analizan precisamente algunas de esas aplicaciones: 1) en las técnicas de espectro esparcido estudiadas en los capítulos 3, 4 y 5 una señal binaria es mezclada con una secuencia de números aleatorios para esparcir el espectro sobre un rango de frecuencias más amplio que el ocupado por la señal pura. Usando distintas secuencias es posible además compartir un canal de comunicaciones entre varios usuarios [71, 72, 73, 74]. La reducción de la interferencia electromagnética es otro beneficio importante del espectro esparcido [75, 76]; 2) Consideremos una señal de baja frecuencia inmersa en ruido digital (de alta frecuencia). La situación aparece típicamente en las fuentes conmutadas y los convertidores de potencia. El muestreo a intervalos definidos por una secuencia de números aleatorios permite filtrar la señal sin necesidad de utilizar bobinas o capacitores que son complejos, ocupan volumen y tienen alto costo [77].

Los números aleatorios no pueden obtenerse mediante una computadora y es probable que tampoco podamos obtenerlos de fuentes "naturales", si es cierta la suposición usual de que todo sistema es gobernado por reglas físicas y consecuentemente es determinista. Una estrategia exitosa para construir un PRNG es partir de una serie temporal generada por un mapa caótico simple y aplicarle un *procedimiento de randomización* adecuado, de modo de reforzar la naturaleza estocástica. El uso de una estrategia de randomización requiere contar con un método cuantitativo de evaluación de la mejora obtenida luego de aplicar el procedimiento. En [78] se empleó la SCM propuesta por Lamberti et al. [30] para cuantificar la efectividad de la randomización de sistemas continuos como el oscilador caótico de Lorenz. Se demostró asimismo que el procedimiento ampliamente empleado de mezclar dos señales caóticas para mejorar la calidad no es efectivo en este caso.

Dos preguntas básicas cuando se necesita usar un PRNG son: 1) *¿qué mapa caótico, entre varios disponibles, genera la mejor serie temporal?*; 2) *¿cuál es la mejor estrategia para randomizar un mapa caótico y mejorar sus propiedades?* Para obtener respuestas objetivas a estas dos preguntas se necesita definir cuantificadores adecuados.

Si bien existen en la literatura bancos de prueba de propósito general para el testeo de PRNGs [79] no están diseñados tomando en cuenta las características particulares de los mapas caóticos. Como queda demostrado en la sección anterior la naturaleza determinista del caos deja su marca en la serie temporal y deben buscarse cuantificadores adecuados para descubrir esa firma.

2.3.1. Análisis del operador de Perrón-Frobenius de un mapa caótico

Sea una serie temporal caótica (CHTS) $S_{IN} = \{x_0, x_1, ..., x_\infty\}$ generada por un mapa 1-D, f sobre el intervalo $[0, 1]$. El histograma normalizado nos da la frecuencia relativa de los valores iniciales, distribuidos en n_{bin} subintervalos $A \subset [0, 1]$ sin intersección, es decir la probabilidad $\mu_0(A)$ de tener un valor inicial $x_0 \in A$ (también denominada medida de probabilidad de A está dada por.

$$\mu_0(A) = \int_A \rho_0(x)\, dx \ . \tag{2.3.1}$$

La Eq. (2.3.1) define una densidad de probabilidad ρ_0 sobre todo el espacio de estados (que es el intervalo $[0, 1]$) . Definimos ahora $\mu_n(A)$ como la probabilidad de encontrar un iterado x_n en el subset A. La densidad apropiada es ahora

$$\mu_n(A) = \int_A \rho_n(x)\, dx \ . \tag{2.3.2}$$

Entonces, en lugar de analizar la evolución de una condición inicial particular, podemos considerar la evolución de ρ_0 en el espacio de probabilidades (para ver detalles matemáticos ver [7]). Este proceso es generado por el llamado operador Perron-Frobenius $L_f : PDF \to PDF$ [7, 8, 9] y cumple:

$$\rho_{n+1} = L_f \rho_n \ . \tag{2.3.3}$$

Los autovalores mayores, en valor absoluto,($\eta_0 \quad y \ \eta_1$) adquieren una relevancia especial. La conservación de la probabilidad lleva a la siguiente condición (trivial) para un subset A arbitrario.

$$\mu_{n+1}\,(A) = \mu_n\,(f^{-1}(A))\;, \tag{2.3.4}$$

donde $f^{-1}(A)$ es la pre-imagen de A, i.e., el conjunto de todos los puntos que son mapeados en A al iterar una vez. La Ec. (2.3.4) nos dice que la frecuencia relativa de la iteración de x_{n+1} en el subconjunto A debe ser igual a la frecuencia relativa de la iteración x_n en el subset $f^{-1}(A)$.

Interesa particularmente la medida de probabilidad invariante (también llamada medida invariante) que satisface:

$$\mu_{n+1}\,(A) = \mu_n\,(A) \tag{2.3.5}$$

y sus correspondientes PDFs invariantes, también llamadas densidades invariantes. Un mapa es llamado ergódico si para cualquier función de testeo integrable $\mathcal{T}(x)$ su promedio temporal es igual a su promedio en el ensamble.

$$\overline{T} = \langle T \rangle\;. \tag{2.3.6}$$

La ecuación 2.3.6 es consecuencia del famoso teorema de ergodicidad de Birkhoff [80] Para tales mapas el promedio temporal no depende del valor inicial x_0 [7]. Pueden existir varias medidas invariantes para un mapa ergódico, pero sólo una de ellas es realmente importante en el sentido de que, si iteramos un punto inicial aleatoriamente elegido, las iteraciones estarán "casi seguro" distribuidas de acuerdo a su medida. La expresión "casi seguro" significa que como máximo existe un conjunto de condiciones iniciales, de medida nula en el intervalo, cuya evolución no conduce a la medida invariante. Por ello esta medida es llamada *medida natural invariante* μ_{inv} y su correspondiente PDF es llamada *densidad natural invariante* ρ_{inv}.

Un mapa f es *mixing* si a partir de una densidad de probabilidad suave inicial ρ_0 converge a ρ_{inv} [81, 82, 83, 84, 85]. Como cada elemento del set de condiciones iniciales (con PDF ρ_0) evoluciona en el tiempo independientemente del resto, el operador L_f es lineal a pesar de la nolinealidad del mapa. ρ_{inv} es el L_f-avo autovector correspondiente al autovalor $\eta_0 = 1$. Puede obtenerse analíticamente sólo en algunos pocos casos, sin embargo, puede aproximarse numéricamente. El segundo autovalor con mayor valor absoluto, η_1, tiene un significado especial: su valor absoluto da la "velocidad" con la cual una distribución inicial se aproxima a la densidad natural invariante, llamada *constante de mixing* r_{mix} del mapa caótico. Cuanto más chico es el valor de r_{mix}, más rápido sera el proceso de mixing. Se ha mostrado que los mapas $f^2, f^3, \cdots$ obtenidos mediante la interacción de un dado f comparten una ρ_{inv} común, sin embargo el r_{mix} decrece a medida que la iteración aumenta (ver Tabla 2.3). En otras palabras, *los mapas iterados presentan mejores propiedades de mixing* [9].

Una *técnica de randomización* es un procedimiento que mejora el grado de aleatoriedad de un PRNG. En esta subsección se demostrará que el plano CH presentado en la sección 2.2 es efectivo para analizar la calidad de generadores provenientes de mapas caóticos pero sólo si se utiliza una distribución de probabilidad diferente para cada eje. Los resultados se ejemplifican para el Mapa Logístico y el Mapa Three Way Bernoulli como ejemplos típicos de la dinámica caótica, demostrando que esta metodología permite seleccionar la técnica de randomización más conveniente en cada caso.

Usualmente la dinámica simbólica es utilizada para realizar descripciones de sistemas dinámicos reales continuos en el tiempo [45, 50], el objetivo es obtener una descripción muy simple del sistema, y poder comprender cómo se comporta el objeto

de estudio. Este punto de vista estimula la investigación orientada a la generación de particiones adecuadas a un espacio de fase relevante, de modo de producir una descripción sin pérdida de información.

En nuestro enfoque la dinámica simbólica es utilizada como una herramienta para randomizar un Generador Caótico Pseudo Aleatorio. Esto es, partiendo de una serie caótica, nuestro logro es transformarla en una serie temporal simbólica con propiedades estadísticas más convenientes que las de la serie original.

La CHTS generada por el mapa f es la entrada del proceso de randomización por lo que la designaremos como $\mathcal{S}_{IN} = \{x_0, x_1, \cdots\}$. Asumamos que x_i es un número en punto flotante (en la representación normalizada de IEEE). En este caso, consideramos valores restringidos al intervalo $[0, 1]$. Luego de varios pasos, la salida de cada uno de estos procesos será una nueva serie temporal (STS) $\mathcal{S}_{OUT}$ obtenidas como se describe a continuación (ver Tablas 2.1 y 2.2).

Discretización

El proceso de discretización es realizado de acuerdo a los siguientes pasos:

1. Cada valor de $\mathcal{S}_{IN}$ es primero discretizado usando N-bits. Esta operación se puede expresar como $x'_n = floor[x_n(2^N)]$. Esto es análogo a mapear el intervalo continuo (real) $(0, 1)$ en el intervalo discreto $[0, 2^N - 1]$. La función floor devuelve el valor del argumento redondeado por defecto, es decir, el mayor número entero menor o igual al argumento. Se debe redondear para convertir un número de mayor precisión en un número de menor precisión. En nuestro caso estamos convirtiendo un número en punto flotante (Estándar IEEE-754 precisión simple) que utiliza 1 bit de signo, 8 bits de exponente y 23 bits de mantisa, a un

entero de 16 bits.

Generándose la nueva serie $\mathcal{S}_1$. En la Tabla 2.1 se ilustra el caso $N = 4$ (ver columnas $\mathcal{S}_1$, $\mathcal{S}_2$, etc.). Cabe destacar que este primer paso es ineludible en el caso de un sistema caótico que este implementado en hardware, usando por ejemplo, field-programmable gate-arrays, cuando se utiliza aritmética de punto fijo. Luego de este primer paso el espacio de representación simbólica contiene 2^N símbolos diferentes. Nótese que en algunos casos *patológicos* (por ejemplo el mapa Tent) este procedimiento genera una serie en la cual desaparece el comportamiento caótico. En estos casos se pueden usar varias aproximaciones para evitar este problema [86, 87, 88].

2. Luego se traduce esta serie temporal $\mathcal{S}_1$ en una nueva $\mathcal{S}_2-$ (formato binario). En la columna $\mathcal{S}_2$ se resaltaron en negrita los dígitos más significativos (MSB), ya que son los que nos interesan para realizar la conversión en los pasos siguientes.

3. Se agrupan de a N $\mathcal{S}_2$-miembros consecutivos como "coordenadas" de un espacio N-dimensional $\mathcal{E}$ de embedding. Los "puntos" de $\mathcal{E}$ son representados por números binarios. Así se forma una nueva serie temporal $\mathcal{S}_3$ reteniendo únicamente el bit más significativo (MSB) de cada coordenada del espacio de embedding. Nótese que esta representación es equivalente al proceso comúnmente empleado en dinámica simbólica que consiste en la división del intervalo $[0, 1]$ en dos subintervalos $A = [0, 0{,}5)$ y $B = [0{,}5, 1]$, asignándoles un "0" si $x_i \in A$ o un "1" si $x_i \in B$.

4. Se re-convierten $\mathcal{S}_3$ en números naturales de N-bits $\mathcal{S}_4$ en el intervalo $[0, 2^N - 1]$.

5. Después se normaliza $\mathcal{S}_4$ (dividiendo cada miembro por $2^N - 1$).

6. La serie resultante es la serie de salida $\mathcal{S}_{OUT}$, nuevamente una serie de números decimales en el intervalo $[0, 1]$.

Del mismo modo, puede emplearse el bit *menos* significativo (LSB), en vez del MSB. Por supuesto, el procedimiento es idéntico al recién detallado, sólo que ahora cada símbolo (número) en $\mathcal{S}_3$ representará la secuencia de paridad de N miembros consecutivos de la CHTS original. Nótese que debido a la discretización existe una pérdida de información ya que muchos puntos del espacio de embedding N-dimensional compartirán el mismo símbolo ($\mathcal{S}_{OUT}$) en la serie temporal final.

Skipping

Esta técnica es una aproximación completamente diferente (ver Tabla 2.2). El proceso consta de dos etapas:

1. Se particiona la CHTS original $\mathcal{S}_{IN}$ en grupos de longitud d, sin superposición, cada uno de estos grupos será un punto en un espacio de embedding d-dimensional $\mathcal{E}$. Esto implica que cada punto en el espacio de embedding representará ahora un vector de longitud d de números en formato de punto flotantes (IEEE).

2. Se adjunta a cada punto en $\mathcal{E}$ un símbolo que consiste únicamente en una coordinada (por ejemplo el valor d-avo), generando la serie temporal de salida (STS) S_{OUT} de la Tabla 2.2.

Notese que $d-1$ valores de $\mathcal{S}_{IN}$ fueron "salteados" (skipped) para obtener la STS $\mathcal{S}_{OUT}$ lo cual origina el nombre **Skipping** para esta técnica. En otras palabras, se

emplea, en lugar del mapa original f, su d-ava iteración f^d. Esta técnica de randomización es generalmente (y exitosamente) utilizada con mapas lineales por tramos en muchas aplicaciones [9]. En la Tabla 2.2, se muestra un ejemplo para diferentes valores de d.

El plano de representación de "dos-probabilidades" (ver abajo), que utiliza H_S y $C_{MPR} = Q \cdot H$ como coordenadas. En Ref. [26] el desequilibrio Q es calculado usando la distancia estadística de Wootters y tomando H como la entropía normalizada de Shannon (ver [26]). En el sentido termodinámico este SCM resultante no es una cantidad intensiva ni tampoco extensiva, sin embargo provee resultados útiles. Una mejora de SCM es darle una característica intensiva, como se logró en Ref. [30]. Esta versión de SCM es *(i)* capaz de mostrar detalles esenciales, *(ii)* posee cualidades intensivas y, *(iii)* es capaz de discernir entre distintos niveles de periodicidad y caos [89]. Esta medida, llamada complejidad estadística intensiva, es una $C_{MPR}[P]$ funcional que caracteriza a las funciones densidad de probabilidad P asociadas a las series temporales de longitud M, generadas por el sistema dinámico estudiado. Se escribe:

$$C_{MPR}[P] \;=\; Q_J[P, P_e] \cdot H_S[P] \;, \tag{2.3.7}$$

donde,

$$H_S[P] \;=\; S[P]/S_{max} \;=\; \left[-\sum_{j=1}^{N} \; p_j \;\; \ln(p_j) \right] \;\Big/\; S_{max} \;, \tag{2.3.8}$$

con $S_{max} = S[P_e] = \ln N$, $(0 \leq H_S \leq 1)$, mientras que N representa el número total de estados del sistema en el espacio de fase. Definimos $P_e = \{1/N, \cdots, 1/N\}$ la distribución *uniforme*, mientras que S representa la entropía de Shannon. Q_J se refiere al "desequilibrio", definido en términos de la divergencia extensiva de Jensen-Shannon [30] que introduce una métrica cuadrada, en contraste con la divergencia de

Kullback- Leiber y se escribe:

$$Q_J[P, P_e] \;=\; Q_0 \cdot \{S[(P + P_e)/2] - S[P]/2 - S[P_e]/2\}, \tag{2.3.9}$$

siendo Q_0 una constante de normalización ($0 \leq Q_J \leq 1$) esto es,

$$Q_0 \;=\; -2\left\{\left(\frac{N+1}{N}\right)\ln(N+1) - 2\ln(2N) + \ln N\right\}^{-1} . \tag{2.3.10}$$

Se vio que el desequilibrio Q_J es una cantidad intensiva que refleja la "arquitectura" del sistema, es distinta de cero únicamente si existen estados "privilegiados", o "más probables", entre los estados accesibles. $C_{MPR}[P]$ cuantifica la presencia de estructuras correlacionadas [26, 30]. Los extremos opuestos que serían orden perfecto y máxima aleatoriedad no poseen estructura y, en consecuencia, $C_{MPR}[P] = 0$. Entre estos dos casos especiales existe un amplio rango con distintos grados de estructura física, grados que deben ser reflejados por las características de la distribuciones de probabilidad correspondientes.

Ambos cuantificadores $H_S[P]$ y $C_{MPR}[P]$ pueden ser calculados para cualquier función de distribución P. Como se vio en la sección 2.2 P no esta únicamente definido. La elección del procedimiento a utilizarse para realizar la extracción de la correspondiente P es un aspecto esencial. De hecho, en la representación de este plano trabajamos con **dos PDFs diferentes**: una basada en estadísticas de amplitud y la otra ideada via el procedimiento de reconstrucción del atractor propuesto por Bandt y Pompe [46], usualmente llamada en la literatura *entropía de permutación*. La razón de haber elegido dos PDFs es que una de ellas, al basarse en estadísticas de amplitud, refleja los cambios producidos por cada técnica de randomización en la medida invariante de los mapas caóticos, mientras que el procedimiento de Bandt-Pompe refleja la calidad de mixing del mapa bajo análisis.

En este trabajo a los cuantificadores obtenidos mediante la PDF basada en el histograma se los llamó $H_S^{(hist)}$ y $C_{MPR}^{(hist)}$. Para series temporales de longitud finita M es relevante considerar un valor óptimo de subintervalos no solapados en los que se divide el intervalo $[0, 1]$, n_{bin}, como se explicará más adelante.

Los cuantificadores obtenidos mediante el procedimiento de Bandt y Pompe son llamados $H_S^{(BP)}$ y $C_{MPR}^{(BP)}$. Un resultado importante de Bandt-Pompe es la obtención de una clara mejora de la Teoría de la Información basada en cuantificadores utilizando sus algoritmos de generación de P [90, 51, 91, 34, 92, 89, 55, 93].

Nuestro plano de representación utiliza $H_S^{(hist)}$ como la coordenada x y $C_{MPR}^{(BP)}$ como la coordenada y. Note que se han utilizado dos distribuciones de probabilidad. Esta es una característica esencial que permite obtener buenos resultados que luego serán reportados.

2.3.2. Aplicaciones

El siguiente ejemplo ilustra las consideraciones precedentes.

- El Mapa Three Way Bernoulli (TWBM) dado por

$$x_{n+1} = \begin{cases} 3x_n & \text{si } 0 \leq x_n \leq 1/3 \\ 3x_n - 1 & \text{si } 1/3 < x_n \leq 2/3 \\ 3x_n - 2 & \text{si } 2/3 < x_n \leq 1 \end{cases} . \tag{2.3.11}$$

  Este mapa comparte con muchos otros (Four Way Tailed shift, Three Way Tailed shift, Skew Tent, etc.) una densidad invariante uniforme ρ_{inv} en el intervalo $[0, 1]$ mientras la constante de mixing r_{mix} de la familia entera de mapas f^j esta dada por $r_{mix}^j = (1/3)^j$ (ver Table 2.3).

- Luego se consideró el mapa logístico (LOG) dado por

$$x_{n+1} \;=\; 4\; x_n\; (1 - x_n)\;, \tag{2.3.12}$$

cuya densidad natural invariante esta exactamente determinada, siendo expresada por

$$\rho_{inv}(x) \;=\; \frac{1}{\pi\sqrt{x(1-x)}}\;. \tag{2.3.13}$$

Los valores de r_{mix} correspondientes se muestran en la Tabla 2.3. Estos han sido obtenidos mediante el Método Transfer Operator, descripto en [7].

Como ya se resaltó, deben tomarse algunas consideraciones con respecto a la cantidad de subintervalos del histograma. Figura 2.7 muestra $H_S^{(hist)}$ como función de $n_{bin} = N$ para LOG (Fig. 2.7.a) y TWBM (Fig. 2.7.b). Se consideraron dos longitudes diferentes M= $1 \cdot 10^6$ y $5 \cdot 10^7$. Para ambos valores de M la entropía $H_S^{(hist)}$ primero aumenta a medida que se incrementa n_{bin} (ó N) hasta alcanzar un valor máximo, y luego comienza a disminuir. En un principio se podría pensar que n_{bin} debe tender a ∞ para obtener la máxima entropía $H_S^{(hist)}$, pero la Fig. 2.7 muestra que existe un valor óptimo (este es $5 \cdot 10^4$ para $M = 1 \cdot 10^6$). La razón para este comportamiento es que la longitud de las series temporales es finita $M < \infty$. Por lo tanto, la cantidad de puntos en cada subintervalo se decrementa a medida que n_{bin} se incrementa, y consecuentemente no podemos obtener una buena estadística con un valor de n_{bin} mayor que el óptimo. La meseta donde $H_S^{(hist)}$ se mantiene casi constante crece a medida que aumentamos el tamaño del archivo.

Otra cuestión importante a tener en cuenta es asegurarse que las características particulares del set de datos que estamos utilizando para el cálculo no tengan influencia en el valor de entropía obtenido. A fin de solucionar este tema se generan, para

una dada longitud M, varias secuencias "subrogadas", estas se obtuvieron iterando el mapa y comenzando de distintas condiciones iniciales. Luego, se evalúa el valor medio (sobre los subrogados) del cuantificador (por ejemplo $\widehat{H_S} \equiv \langle H_S \rangle$). En este caso, se han realizado tests de convergencia empleando 8 subrogados con $M = 50 \cdot 10^6$ de datos cada uno; $\widehat{H_S^{(hist)}}$ es tomada como el valor de entropía en todos los casos. Hemos verificado que la cantidad de subrogados tomados, así como el tamaño de cada uno de estos, sea suficiente para que el valor de $\widehat{H_S^{(hist)}}$ obtenido no varíe sus cinco dígitos más significativos, $\widehat{H_S^{(hist)}}$ corresponde a la coordenada x de nuestro plano de representación. Por simplicidad de notación a partir de aquí se omitirá el símbolo del sombrero ancho.

En la Fig. 2.8 puede verse cómo varía $H_S^{(hist)}$ cuando a una secuencia inicial se le aplican las dos técnicas de randomización ya discutidas. Allí se ilustra $H_S^{(hist)}$ como función de n_{bin} para el caso del mapa LOG. En la Fig. 2.8(a) la línea punteada es obtenida con números en formato de punto flotante IEEE y la línea continua con pequeños círculos muestra el efecto de la **Discretización** utilizando 16 bits. El máximo valor (válido) de n_{bin} es $2^{16} - 1 = 65535$, y la máxima entropía es obtenida precisamente para este valor. Luego de aplicar el tratamiento del bit más significativo (MSB) en el armado de la serie temporal simbólica deseada (STS) (línea continua con pequeños cuadrados en el gráfico), obtuvimos que los valores de entropía aumentaron considerablemente, como se esperaba.

Además, la representación de $n-$bits tiene la misma entropía que la de la entropía de la secuencia con números en punto flotante sólo cuando el número de valores posibles 2^n es un múltiplo exacto de n_{bin} (véase en la Fig. 2.8(a) la coincidencia para los valores 65536, 65536/2, 65536/3, etc). Esta es una situación usual cuando se

realiza el histograma de una distribución discreta. De hecho, la grilla del histograma es equivalente a colapsar cada subintervalo de longitud $1/n_{bin}$ en un único valor. Por otro lado, una discretización de $n-$bits colapsa cada subintervalo de longitud $1/2^n$ en un único valor natural y consecuentemente induce otra grilla. Esta es la razón por la que la entropía de series discretas sea menor que la de la serie de números flotantes si $n_{bin} \neq 2^n/k \; k = (1, 2, ...)$ como se ve en la Fig. 2.8(a).

En la Fig. 2.8(b) se ilustran los efectos de **Skipping**. La línea punteada es nuevamente obtenida utilizando el mapa original f y representando los valores en formato de punto flotante IEEE. Los círculos pequeños corresponden a la segunda iteración del mapa , f^2, y los cuadrados a f^4. Note que la misma entropía $H_S^{(hist)}$ es obtenida en todos los casos. Este comportamiento confirma que **Skipping** no afecta la distribución de probabilidades cuando se utiliza estadística de amplitud [9]. Como punto final en cuanto a esta figura notamos que la entropía de Shannon de la STS obtenida usando el bit menos significativo (LSB), en el proceso de discretización, aparecen valores de entropía casi idénticos a los de MSB (no se muestran en Fig. 2.8(a)). El resultado más significativo puede verse en las Figs. 2.9, para ambas técnicas de randomización, el Mapa LOG *(a)* y el Mapa TWBM *(b)*: el plano $H_S^{(hist)} \times C_{MPR}^{(BP)}$.

El objetivo de toda técnica de randomización es lograr aproximarse al punto ideal $(1, 0)$ en este plano, ya que la randomización allí es óptima. La representación plana ayuda claramente a mostrar las características de calidad de las aproximaciones de randomización. En el Mapa LOG la CHTS inicial presenta un punto aproximadamente en (0,98, 0,5) en el plano $H_S^{(hist)} \times C_{MPR}^{(BP)}$. Este punto corresponde a la entropía de un histograma no uniforme con un valor alto de complejidad lo que indica la existencia de estructuras geométricas. Mediante **Skipping** se destruyen estas estructuras

desplazando la coordinada y hacia el valor ideal 0, sin embargo el histograma no es modificado, consecuentemente, la coordenada x no cambia. A pesar del hecho de que el STS (indicado como S_{OUT}^{skip}) presenta propiedades estadísticas mejores que la CHTS (indicado como S_{IN}), no logra alcanzar el punto ideal $(1, 0)$. Por otro lado, **Discretización** disminuye la coordenada y y aumenta la coordenada x, así la STS alcanza el punto ideal $(0, 1)$. La razón de esto es que la CHTS original sigue la ecuación logística pero la cantidad de valores que pertenecen al intervalo $[0; 0{,}5)$ es casi idéntico a la cantidad de valores en $[0{,}5; 1]$, por lo que al aplicar la técnica de MSB se logra alcanzar este punto (lo mismo se aplica para la alternativa LSB).

Notese que el CHTS generado por el mapa TWBM presenta el valor ideal $H_S^{(hist)} = 1$ desde el comienzo. Cuando se le aplica la técnica de **Discretización** en el plano vemos como se decrementa la coordenada y hacia el valor ideal $C_{MPR}^{(BP)} = 0$ y, al mismo tiempo, disminuye ("empeora")la coordenada x a el valor $H_S^{(hist)} = 0{,}96$. Se concluye luego que la mejor opción es **Skipping**, la cual mejora la coordenada y sin ningún cambio de la coordenada x, permitiendo que la STS alcance el punto ideal $(1, 0)$.

Luego de realizar muchas corridas de simulación con distintos mapas caóticos, caracterizados por medidas naturales invariantes uniformes y no uniformes, se puede concluir que: la técnica de **Skipping** es mejor que la técnica de **Discretización** para el caso de mapas que posean medidas naturales invariantes uniformes, y viceversa en el caso no uniforme.

La razón de esto es clara, **Skipping** decrementa $C_{MPR}^{(BP)}$ sin cambiar $H_S^{(hist)}$. Por lo tanto en los mapas con distribuciones no-uniformes (el mapa LOG es un ejemplo) **Skipping** disminuye la complejidad $C_{MPR}^{(BP)}$ sin cambiar $H_S^{(hist)}$ y, consecuentemente, nunca alcanza el punto ideal $(1, 0)$. En cambio, mediante la técnica de **Discretización**

$\mathcal{S}_{IN}$ ($\mathbb{R}$)	$\mathcal{S}_1$ ($\mathbb{N}$)	$\mathcal{S}_2$ (*binario*)	$\mathcal{S}_3$ — MSB embedding 4	$\mathcal{S}_4$ ($\mathbb{N}$)	$\mathcal{S}_{OUT}$ ($\mathbb{R}$)
0.010559404	0	**0**000			
0.041791613	0	**0**000			
0.160180296	2	**0**010			
0.538090276	8	**1**000	**0001**	1	0.066666667
0.994196523	14	**1**110			
0.023079185	0	**0**000			
0.090186145	1	**0**001			
0.328210418	4	**0**100	**1000**	8	0.533333333
0.881953358	13	**1**101			
0.416446528	6	**0**110			
0.972075269	14	**1**110			
0.10857976	1	**0**001	**1010**	10	0.666666667
0.387160782	5	**0**101			
0.949069244	14	**1**110			
0.193347257	2	**0**010			
0.62385638	9	**1**001	**0101**	5	0.333333333

Cuadro 2.1: Procedimiento de **Discretización**.

se decrementa $C_{MPR}^{(BP)}$ y, a la vez, se logra aumentar $H_S^{(hist)}$, pudiéndose llegar al punto ideal. Los mapas que poseen distribuciones naturales invariantes uniformes tienen un valor de $H_S^{(hist)}$ ideal (aproximadamente uno) desde un comienzo, entonces, sólo pueden ser mejoradas las propiedades de mixing del mapa. Esto es exitosamente realizado mediante **Skipping**.

$\mathcal{S}_{IN}$	$\mathcal{S}_{OUT} \equiv f^d$				
$(\mathbb{R})$	$d=2$	$d=3$	$d=4$	$d=5$	$d=6$
0.010559404					
0.041791613	0.041791613				
0.160180296		0.160180296			
0.538090276	0.538090276		0.538090276		
0.994196523				0.994196523	
0.023079185	0.023079185	0.023079185			0.023079185
0.090186145					
0.328210418	0.328210418		0.328210418		
0.881953358		0.881953358			
0.416446528	0.416446528			0.416446528	
0.972075269					
0.10857976	0.10857976	0.10857976	0.10857976		0.10857976
0.387160782					
0.949069244	0.949069244				
0.193347257		0.193347257		0.193347257	
0.62385638	0.62385638		0.62385638		

Cuadro 2.2: Procedimiento de **Skipping**.

2.3.3. Otros planos de representación

La expresión analítica de la medida invariante $\mu(x)$ de un mapa f en general no es conocida. El mapa logístico con caos completo y los mapas lineales por tramos son excepciones. La constante de mezcla r_{mix} ha sido obtenida analíticamente sólo para los mapas lineales por tramos. Para otros mapas se la puede obtener numéricamente mediante una aproximación lineal por tramos del mapa [8, 7].

Es entonces conveniente contar con cuantificadores para medir tanto la uniformidad de la medida invariante $\mu(x)$, como la constante de mezcla r_{mix}, para mapas caóticos arbitrarios. La entropía H^{hist} y la complejidad $C^{(BP)}_{MPR}$ son precisamente dos de tales cuantificadores. En esta sección se investigan otras posibilidades.

La medida invariante del mapa iterado f^d es idéntica a la del mapa original f.

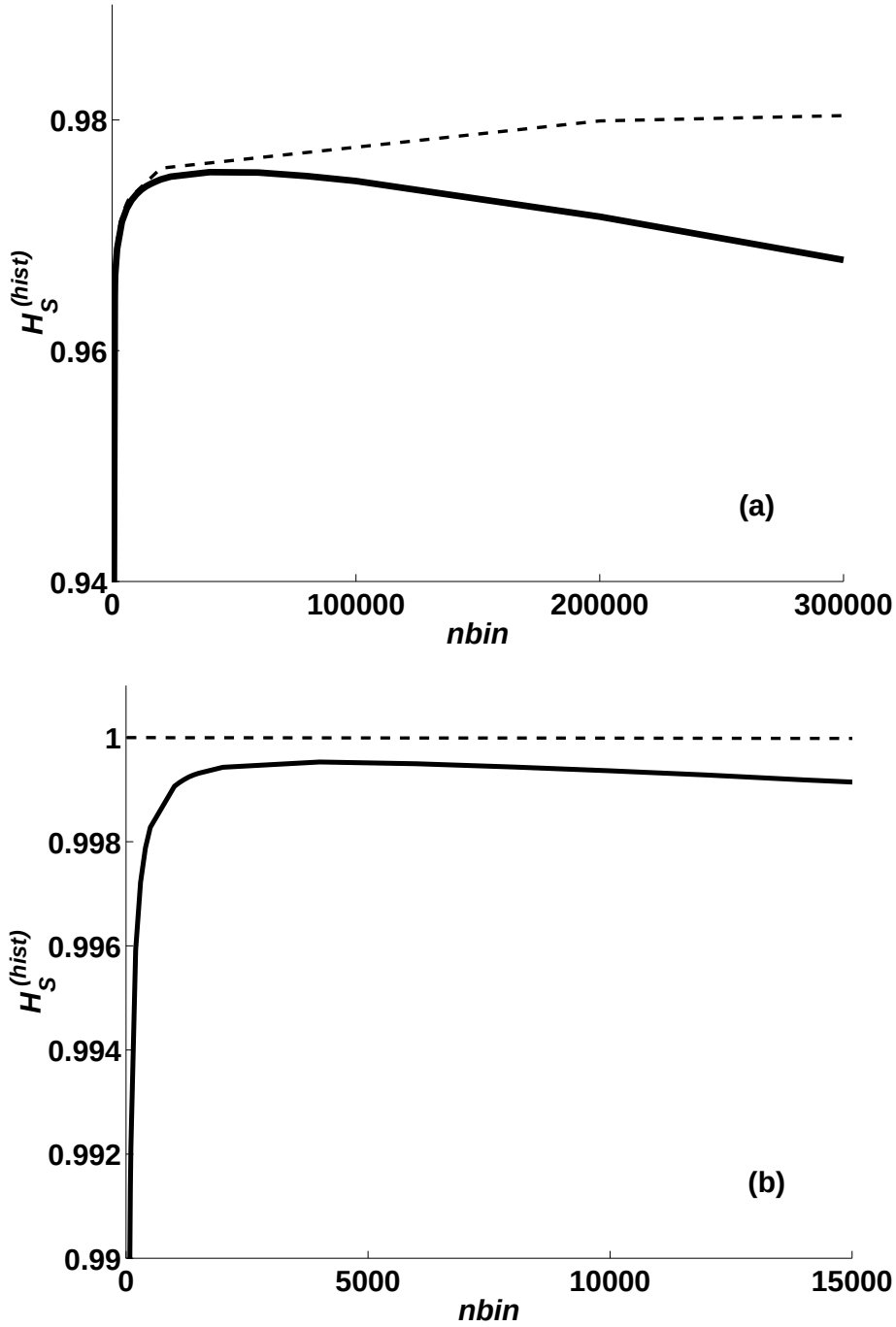

Figura 2.7: Entropía de Shannon $H_S^{(hist)}$ como función de la cantidad de bins n_{bin} para diferentes cantidades de datos (tamaños de archivos) $M = 10^6$ (línea continua) y $M = 5 \cdot 10^6$ (línea punteada). *(a)* mapa LOG, *(b)* mapa TWBM.

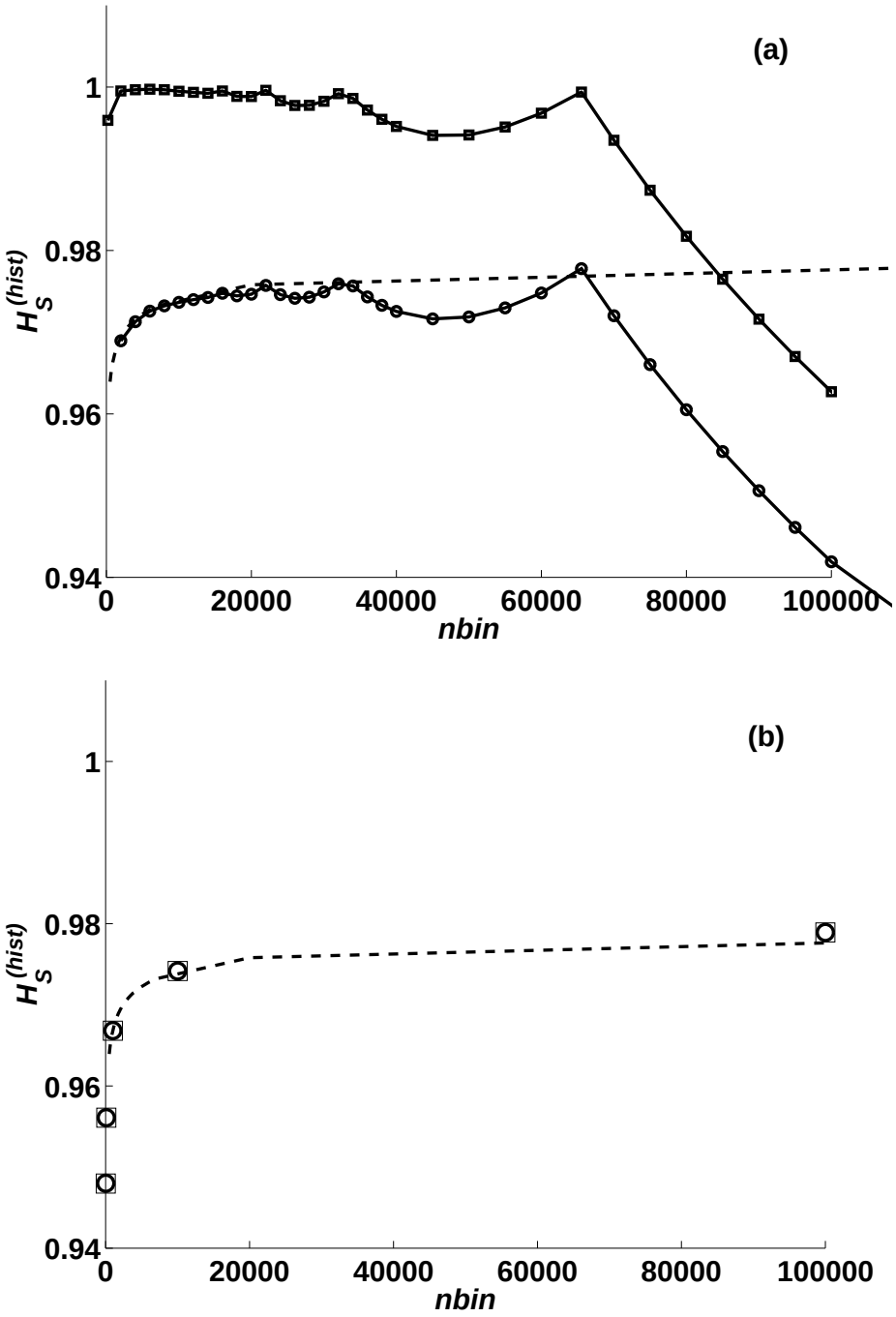

Figura 2.8: Entropía de Shannon $H_s^{(hist)}$ como función de la cantidad de bins n_{bin} para ambos procesos de randomización para el mapa LOG. *(a) Discretización:* IEEE punto flotante (línea punteada); 16 bits (línea continua con círculos); MSB (línea continua con cuadrados). *(b) Skipping:* IEEE punto flotante — mapa original f (línea punteada); segunda iteración del mapa f^2 (círculos); cuarta iteración del mapa f^4 (cuadrados).

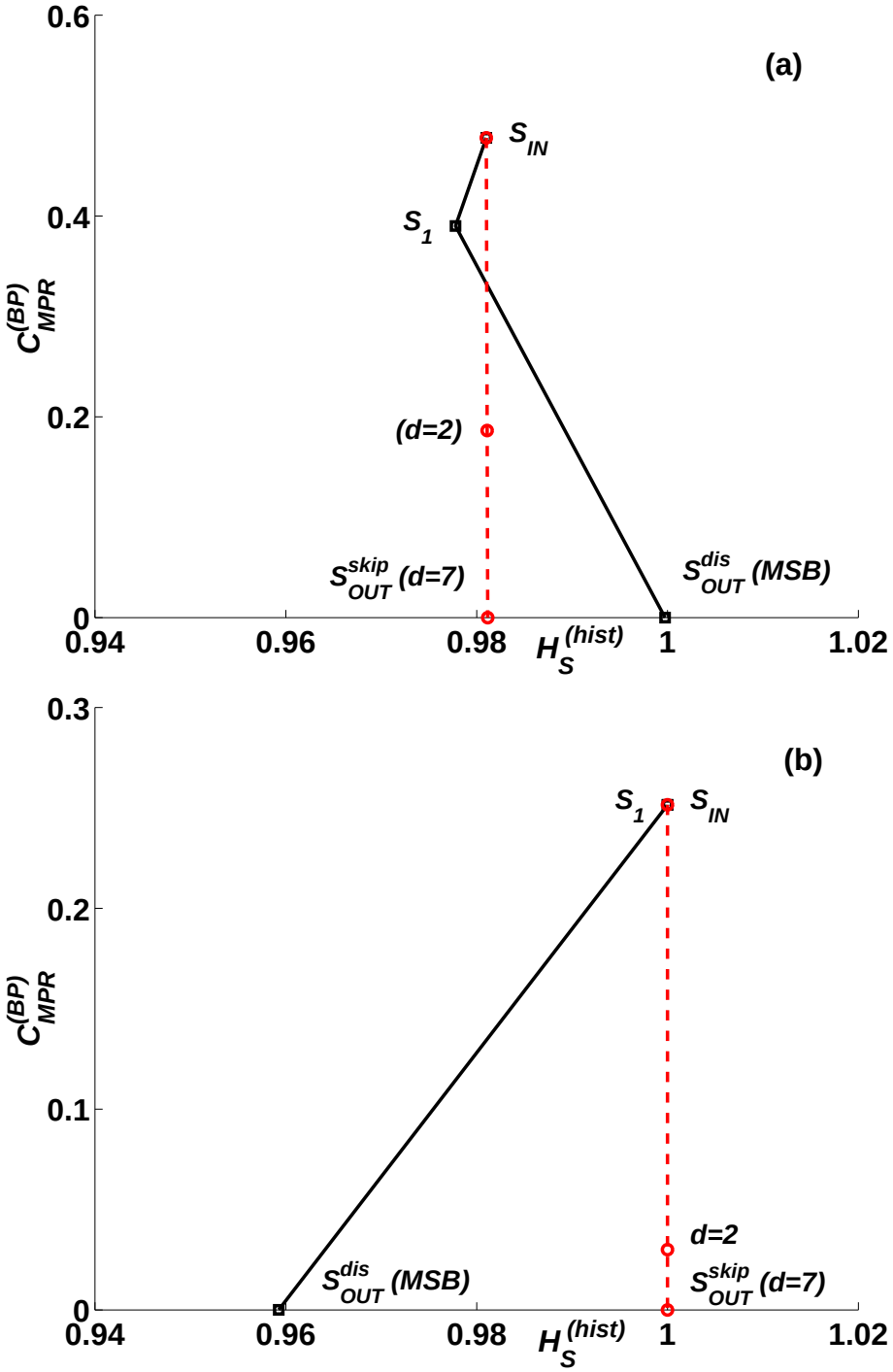

Figura 2.9: Plano de representación $H_S^{(hist)}$—$C_{MPR}^{(BP)}$ para ambos procesos de randomización. *(a)* Para el mapa LOG **Discretización** produce una STS con coordenadas ideales $(1, 0)$ pero **Skipping** no es capaz de mejorar $H_S^{(hist)}$ y la STS posee coordenadas $(0{,}98, 0)$; *(b)* Para el mapa TWBM **Discretización** decrementa la coordenada y al valor ideal $C_{MPR}^{(BP)} = 0$ pero también decrementa la coordenada x a $H_S^{(hist)} = 0{,}96$ mientras que **Skipping** mejora la coordenada y y no cambia el valor de la entropía. Luego, la STS alcanza el valor ideal $(1, 0)$.

j	LOG	TWBM
1	0.56789	0.333333333
2	0.31848	0.111111111
3	0.13290	0.037037037
4	0.05788	0.012345679
5	0.03646	0.004115226
6	0.01791	0.001371742
7	0.01152	0.000457247
8	0.00515	0.000152416

Cuadro 2.3: r_{mix} como función del orden de iteración j para los mapas TWBM y LOG , respectivamente.

En cambio, la constante de mezcla r_{mix} para el mapa iterado f^d es menor que la del mapa original f. Es decir que la iteración de un mapa (que es equivalente a la técnica de **Skipping** explicada en la subsección anterior) genera nuevas series temporales con r_{mix} decreciente. En esta subsección aprovecharemos esta propiedad para comparar distintos cuantificadores, propuestos en la literatura como cuantificadores de aleatoriedad. Los cuantificadores explorados pueden dividirse en tres clases: (1) Cuantificadores basados en la Teoría de la información [23, 34, 30]; (2) cuantificadores basados en diagramas de recurrencia [31, 94]; (3) cuantificadores basados en computación intrínseca [32].

Los *cuantificadores basados en la teoría de la información* son las entropías y las complejidades estadísticas ya explicadas previamente.

Los *cuantificadores basados en diagramas de recurrencia*, como se vio en el capítulo 1 se basan en el estudio de la función recurrencia [31]. Los PRNGs, presentan estructuras de "pequeña escala" structures que pueden medirse [94] mediante funcionales de la función recurrencia (Eq. (**??**)).

Los *Cuantificadores basados en computación intrínseca*, fueron detallados en la

sección 1.2.7, en el capítulo 1.

En resumen es interesante comparar los siguientes cuantificadores: $H^{(hist)}$, $C^{(hist)}$, $H^{(BP)}$, $C^{(BP)}$, RR, DET, $ENTR$, L, h_μ, and $\mathbf{E}$. Estas cantidades deberían decirnos cuan bueno es un PRNG con respecto al ideal $\mu(x) = const$, $r_{mix} = 0$. $H^{(hist)}$ es el cuantificador natural para medir el apartamiento de la uniformidad de $\mu(x)$, y su valor es 1 para el PRNG ideal. La entropía no depende del orden de aparición de los valores en la serie sino sólo del número de veces que aparecen. Es decir que $H^{(hist)}$, no es capaz de brindarnos información sobre los valores de r_{mix}. Entonces un buen plano de representación tiene que utilizar un segundo eje con un cuantificador que cambie con r_{mix} pero no con $\mu(x)$. Para encontrar tal cuantificador estudiamos $H^{(hist)}$, $C^{(hist)}$, $H^{(BP)}$, $C^{(BP)}$, RR, DET, $ENTR$, L, h_μ, and $\mathbf{E}$ como funciones de r_{mix}. Una familia de mapas iterados f^d puede ser usada para este propósito dado que comparten la misma medida invariante y sólo difieren en r_{mix}, que es una función decreciente de d. El mejor cuantificador para r_{mix} será aquél que sea monótono y tenga máxima variación sobre la familia de mapas.

La constante de mezcla de las familias iteradas del mapa logístico y del TWB se muestra en la tabla 2.3. En el caso de TWB $r^d_{mix} = (1/3)^d$.

Para la evaluación de los diferentes cuantificadores usamos archivos de $M = 50{\cdot}10^6$ números de punto flotante. En el caso de la aproximación de Bandt-Pompe consideramos $D = 6$ y para los histogramas tomamos $n_{bin} = 2^{16} = 65536$. Las medidas basadas en RP dependen de varios parámetros:

- la dimensión D_e del espacio de embedding fue adoptada $D_e = 1$.
- ε, parámetro esencial para definir si se ha producido o no una recurrencia. Adoptamos $\varepsilon = 1/(2^{16} - 1)$ correspondiente a números de 16-bits para poder

analizar serie discretizadas.

- l_{min} es la longitud mínima de las líneas diagonales. Excepto para el cálculo de L, para el que tomamos $l_{min} = 1$, en el resto de los cuantificadores se utilizó $l_{min} = 2$.

- N es el número de valores empleados en cada realización. Trabajamos con 10 series subrogadas de $N = 10000$ datos cada una.

Las figuras 2.10, 2.11, y 2.12 muestran el comportamiento de todos los cuantificadores para el mapa LOG (Figs. (a)) y para el mapa TWB (Figs. (b)), respectivamente. Estas figuras muestran que los siguientes cuantificadores son utilizables para medir r_{mix}: $C^{(BP)}$, DET, $ENTR$ y L. Por otro lado los siguientes cuantificadores dependen de la densidad invariante pero no de r_{mix}: $H^{(hist)}$, $C^{(hist)}$, y RR.

Los cuantificadores de computación intrínseca muestran un comportamiento diferente para LOG y para TWB. En LOG ambos cuantificadores son independientes de r_{mix}, pero en TWB, h_μ disminuye al aumentar r_{mix} en tanto que $\mathbf{E}$ se incrementa con r_{mix}. Por lo tanto estos cuantificadores no parecen adecuados para nuestros planos de representación.

Si se compara LOG con TWB empleando estos parámetros se ve que TWB es ligeramente superior a LOG. El problema con TWB y con otros mapas lineales por tramos es que no son realistas y su implementación es más compleja que con mapas con derivada continua como LOG.

El estudio de métodos de randomización de la subsección anterior puede realizarse con cualquier plano de representación que utilice un cuantificador dependiente $\mu(x)$ como eje x (se eligió $H^{(hist)}$ en la Fig. 2.13) y un cuantificador dependiente de r_{mix}

como eje y (se eligió DET en la Fig. 2.13). El plano de representación muestra nuevamente que **Skipping** es un procedimiento mejor que **Discretización** para el caso de TWB (y lo mismo ocurre para todos los mapas que tienen medida invariante constante y sólo necesitan que r_{mix} sea reducida). La Fig 2.13 (b) muestra en cambio que **Discretización** es un procedimiento mejor para LOG dado que el punto ideal $[1, 0]$ no puede alcanzarse con **Skipping**.

En resumen:

1. Para la evaluación de calidad de un PRNG son necesarios dos tipos de cuantificadores: *a) cuantificadores dependientes de r_{mix} únicamente (pero no de $\mu(x)$)*, como $C^{(BP)}$, $H^{(BP)}$, DET, L and $ENTR$ y *b) cuantificadores dependientes de $\mu(x)$ únicamente (pero no de r_{mix})*, como $H^{(hist)}$, $C^{(hist)}$, y RR.

2. Los cuantificadores de computación intrínseca dependen de ambos $\mu(x)$ y r_{mix} y por lo tanto no son convenientes para el análisis de PRNG's con nuestra metodología.

3. Los planos de representación adecuados deben tener un cuantificador de cada clase. Estos planos de representación permiten comparar PRNG's entre sí y además elegir el método de randomización más adecuado.

4. La representación en el plano de la Fig. 2.1 no es adecuada pues utiliza una única distribución de probabilidades en ambos ejes (la distribución obtenida mediante la prescripción de Bandt y Pompe). En ese plano tanto las series sin randomizar como las randomizadas se encuentran muy cercanas al punto $Complexity = 0$, $Entropy = 1$. es

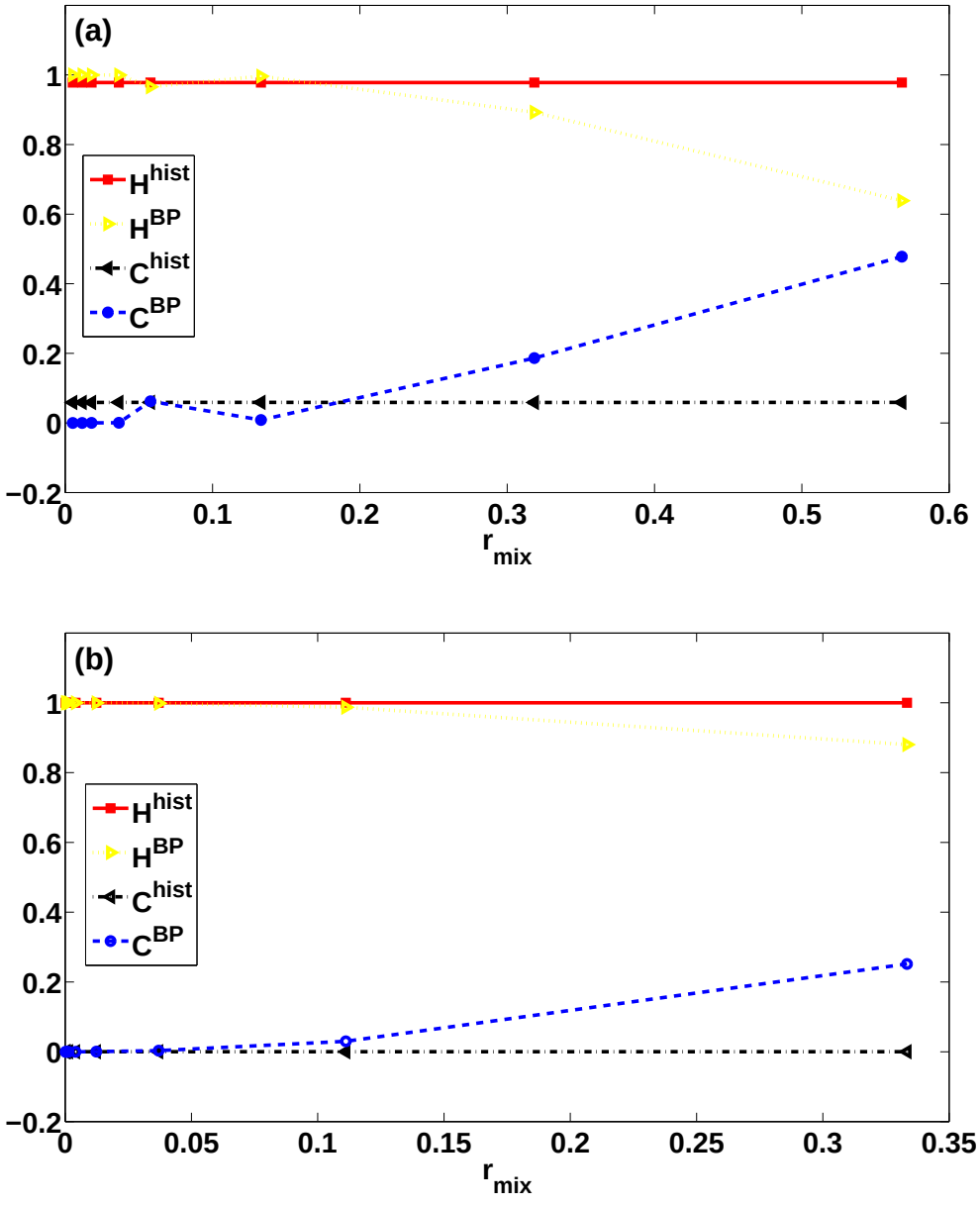

Figura 2.10: Information Theory quantifiers as functions of r_{mix} for: *(a)* LOG map, *(b)* TWB map.

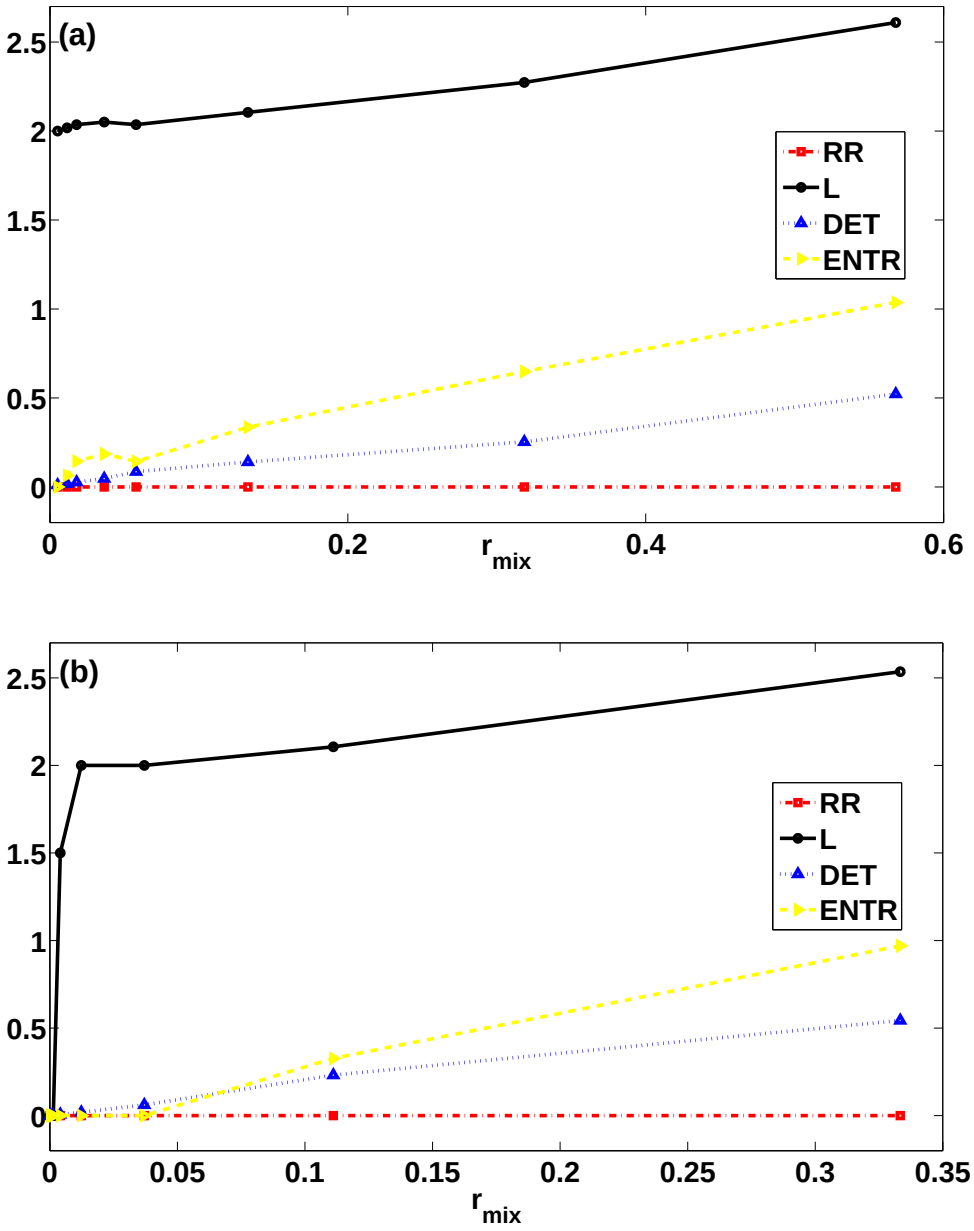

Figura 2.11: Recurrence Plots quantifiers as functions of r_{mix} for: *(a)* LOG map, *(b)* TWB map.

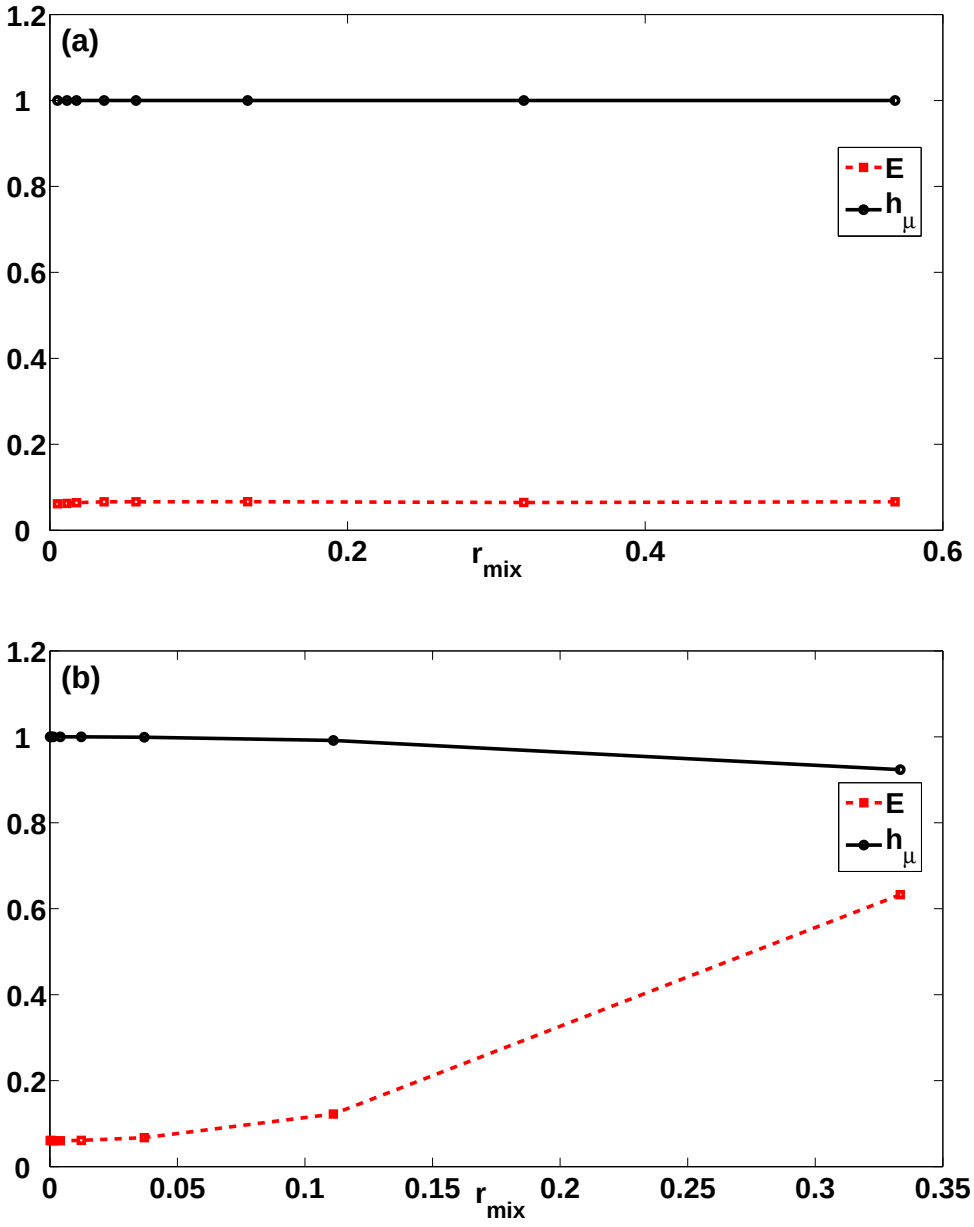

Figura 2.12: Intrinsic Computation quantifiers as functions of r_{mix} for: *(a)* LOG map, *(b)* TWB map.

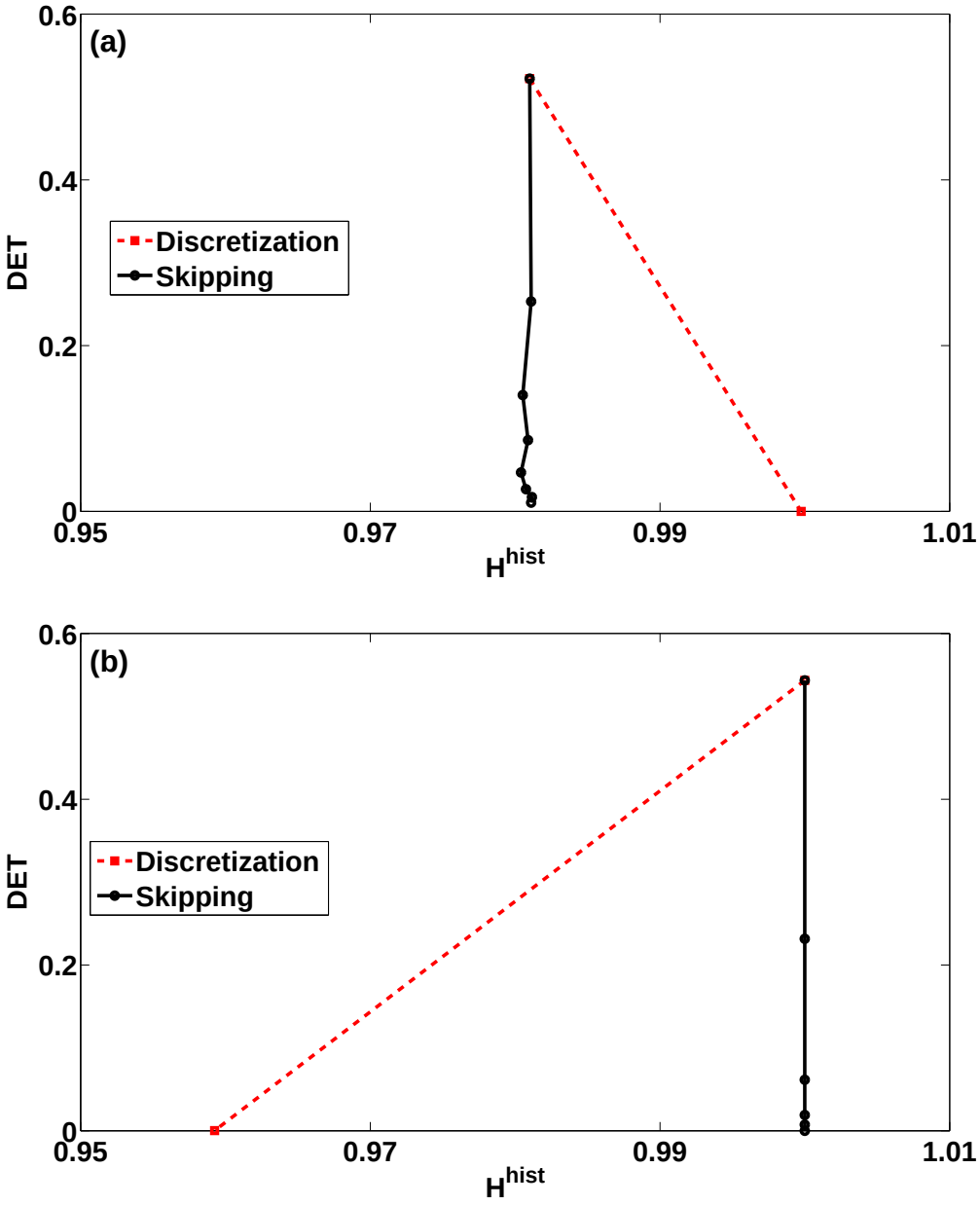

Figura 2.13: DET as a function of H^{hist}, as evaluated in [36], for both randomization procedures applied to: *(a)* LOG map, *(b)* TWB map.

Capítulo 3

Sistemas de comunicaciones de Espectro Esparcido (SS)

3.1. Introducción

El espectro esparcido (también llamado espectro ensanchado, espectro disperso, spread spectrum o SS) es una técnica por la cual el espectro en frecuencia de una señal se ensancha a lo largo de una banda de frecuencias mayor que el que la propia señal posee. Hay múltiples razones por las que puede ser importante ensanchar el espectro de una señal. Por mencionar sólo algunas:

- Lograr que la señal coexista con otras sin interferir. Si la potencia no está concentrada en una única frecuencia se logra disminuir la influencia de esa señal sobre otros sistemas situados en sus proximidades.

- Lograr que varios usuarios de un sistema de comunicaciones de distribución de tiempo, compartan el mismo espectro de frecuencias.

- Encriptar una señal de modo que sólo quien conoce la clave de encriptado, que es la que produce el ensanchamiento del espectro, pueda invertir el proceso y

recuperar la información transmitida a través de un canal de comunicación.

- Generar una señal que simule el ruido producido en un canal de comunicaciones en la banda de frecuencias en la que el sistema trabaja.

En este capítulo se presentan las técnicas básicas de espectro esparcido. En especial se investiga el uso del caos determinista como generador de secuencias pseudo aleatorias o pseudo códigos a ser empleados en el proceso de codificación de los sistemas DS-CDMA (Direct Sequence Code Division Multiple Access). La principal contribución original presentada es un estudio de un conjunto importante de pseudo códigos tradicionales y su comparación con las secuencias generadas por medio de mapas caóticos simples.

3.2. Técnica de Espectro Esparcido

La técnica de espectro esparcido consiste en el procesamiento de una señal para lograr que la energía media de esta se reparta sobre un ancho de banda mucho mayor que el de la señal original.

El ensanchamiento de la banda se realiza a partir de una señal pseudo aleatoria (código PN por Pseudo Noise) utilizada como forma de onda modulante. Esta señal posee una apariencia de ruido pero es, de hecho, una señal determinista. La señal final tendrá, por lo tanto, características pseudo aleatorias y su espectro dependerá de las características del código PN.

Luego, la señal podrá ser "recuperada" sólo si se cuenta con una réplica sincronizada de este código PN.

Existen dos tipos de técnicas:

La primera, y más simple, que fue desarrollada en comunicaciones digitales es la llamada técnica de salto de frecuencia (FHSS por Frequency-Hopping Spread Spectrum). Con esta técnica la información modula en frecuencia a una señal portadora, sin embargo, en vez de operar a una frecuencia fija, el sintetizador cambia la frecuencia muchas veces por segundo de acuerdo con una secuencia de canales preprogramada. Esta secuencia es el código PN. Para recuperar el mensaje, el receptor debería cambiar de frecuencias en sincronía con el transmisor. El mensaje es recuperado únicamente cuando la secuencia de frecuencias que se utilizó en la modulación es conocida.

Otro método de SS es la técnica de secuencia directa (DSSS por Direct Sequence Spread Spectrum). En este tipo de técnica la información es multiplicada por un código PN y nuevamente sólo será posible recuperar correctamente la información por un receptor para el que dicha secuencia sea conocida. Como cada transmisor emplea una secuencia distinta, es posible que varios transmisores-receptores compartan una misma área geográfica sin interferirse.

Estos sistemas esencialmente intercambian un mayor ancho de banda por una densidad espectral de potencia más baja. Esta reducción de la densidad espectral, expresada en watts por hertz de ancho de banda, da como resultado una relación señal a ruido menor que uno (es decir, la potencia de la señal en cualquier intervalo de frecuencia es menor que la potencia de ruido en el mismo ancho de banda). A primera vista podría parecer que esto haría imposible detectar la señal, lo cual es cierto a menos que se utilicen técnicas especiales para “recuperar” la señal mientras que, al mismo tiempo, se dispersa la energía de las señales interferentes. De hecho, la baja densidad de potencia promedio de las señales de espectro esparcido tiene que

ver con su inmunidad relativa a la interferencia y a la intercepción de mensajes.

Por otra parte, un parámetro que debe ser considerado en los sistemas SS es la ganancia de procesamiento (G_P), que como puede verse en la ecuación 3.2.1, es la razón en decibelios del ancho de banda de transmisión (BW_T) y el ancho de banda de la información (BW_I).

$$G_P = 10log(BW_T/BW_I) \tag{3.2.1}$$

El parámetro de ganancia es necesario para determinar el número de usuarios que pueden ser permitidos en un sistema; además debe ser también considerado el total de reducción por efecto multicamino y la dificultad de detección de las señales. De esta forma al ser nuestra ganancia mayor nos permite que más usuarios utilicen la técnica de SS esto ante la poca atenuación que le implica el resto de señales esparcidas en un canal.

El parámetro de ganancia refleja la pérdida de amplitud de la señal durante el proceso de "ensanchamiento", su valor depende de qué tanto sea esparcida la información. Sin embargo, la cantidad de energía es la misma antes y después del proceso.

La técnica de SS ofrece, por lo tanto, la posibilidad de compartir el espectro con sistemas convencionales de banda estrecha, debido a su bajo nivel de potencia.

3.2.1. Múltiple Acceso por división de Código (CDMA)

El multiplexado, o control de acceso al medio, es uno de los temas más importantes a resolver en comunicaciones de datos cuando se tienen varios usuarios, un único canal y se desean realizar varias comunicaciones al mismo tiempo.

Existen métodos de organización para permitir estas comunicaciones y evitar la aparición de interferencias que pueden degradar e inclusive impedir la comunicación. Algunos de estos esquemas de multiplexado utilizados son: división en frecuencia (FDMA) utilizado comúnmente en la mayoría de los sistemas de comunicación analógicos, división en el tiempo (TDMA) utilizado por sistemas digitales, división en el espacio (SDMA) y division por código (CDMA). Este último es el caso estudiado en esta tesis.

CDMA emplea la tecnología de espectro esparcido por secuencia directa. A cada transmisor se le asigna un código único, escogido de forma que sea ortogonal respecto al del resto; el receptor capta las señales emitidas por todos los transmisores al mismo tiempo, pero gracias al esquema de codificación (que emplea códigos ortogonales entre sí) puede seleccionar la señal de interés si se conoce el código empleado, y el resto de las transmisiones las desecha como ruido. De esta forma, múltiples usuarios transmiten simultáneamente sobre la misma banda espectral y se diferencian únicamente por su código PN.

El procesamiento realizado a los datos a transmitir es simplemente aplicar la función lógica AND con el código de transmisión, que es único para ese usuario y se emite con un ancho de banda significativamente mayor que el de los datos. Multiplicar dos señales en tiempo implica hacer una convolución en frecuencia lo cual esparce o dispersa el espectro, ya que el ancho de banda resultante luego de una convolución es equivalente a la suma de los anchos de los espectros convolucionados. Como el area total en el espectro se mantiene (ya que se mantiene la potencia enviada) la señal a transmitirse tendrá en el plano de las frecuencias una amplitud mucho menor, comparable con el ruido del canal.

Como ya se dijo, cada usuario de un sistema CDMA emplea un código de transmisión distinto. La selección del código a emplear para la modulación es vital para el buen desempeño de los sistemas CDMA, porque de él depende que el receptor capte la señal de interés correcta, esto se realiza mediante la correlación cruzada de la señal captada con el código del usuario de interés, así como el rechazo del resto de señales y de las interferencias multicamino (producidas por los distintos rebotes de señal).

Las propiedades de correlación de los códigos PN son esenciales para una exitosa decodificación. El receptor realiza la correlación entre la secuencia correspondiente al receptor particular y la señal recibida. Cuando esta correlación supera un cierto nivel de umbral el receptor se sincroniza y decodifica la señal.

El mejor caso se presenta cuando existe una buena separación entre la señal del usuario deseado (la señal de interés) y las del resto. Cuando la señal captada es la buscada, el receptor obtendrá un resultado muy alto de la correlación, y el sistema podrá extraer la señal. En cambio, si la señal recibida no es la de interés, como el código empleado por cada usuario es distinto, la correlación debería ser muy pequeña, idealmente tendiendo a cero (y por tanto eliminando el resto de señales). Y además, si la correlación se produce con cualquier retardo temporal distinto de cero, la correlación también debería tender a cero. A esto se le denomina autocorrelación y se emplea para rechazar las interferencias multicamino.

La división por código se emplea en múltiples sistemas de comunicación por radiofrecuencia, tanto de telefonía móvil (como IS-95, CDMA2000, FOMA o UMTS), transmisión de datos (WiFi) o navegación por satélite (GPS).

3.2.2. Performance del Sistema

La performance de un sistema de comunicacioenes se carateriza principalmente por:

- Relación Señal a Ruido (SNR).
- Error binario (BER).
- Sincronización del código.
- Rechazo a señales multicamino.
- Capacidad del sistema (cantidad de usuarios simultáneos).

Las propiedades estadísticas de los códigos PN juegan un rol fundamental en la implementación de los sistemas CDMA y la selección del código influye directamente en la performance del sistema.

3.2.3. Performance de las secuencias PN

Cualquier sistema de comunicaciones CDMA utiliza una o varias familias de códigos PN. Cada secuencia en la familia es utilizada por un único usuario y el número máximo de usuarios simultáneos esta limitado por la cantidad de códigos disponibles.

En vistas al proceso de decodificación las propiedades de correlación de cada familia de secuencias PN juega un rol muy importante en la eficiencia del sistema CDMA, ya que determina el nivel de interferencia por múltiple acceso, la auto-interferencia debida a la propagación por multicamino, y el tiempo requerido para realizar la sincronización del código.

Las familias de secuencias PN deben cumplir los siguientes requisitos:

- Disponibilidad de una gran cantidad de códigos por familia, de forma de permitir múltiples usuarios simultáneos.

- Autocorrelación: cada código en la familia debe ser fácilmente distinguible de una versión desplazada en el tiempo de él mismo para acelerar el proceso de sincronización. Esta propiedad es también importante para la robustez multicamino; esto significa para reducir el efecto de propagación por multicamino. Por lo tanto, para t=0 debe haber un pico lo más acusado posible (se obtiene mediante códigos más largos),mientras que el nivel de los lóbulos secundarios debe ser lo más bajo posible (el código será mejor cuanto más aleatorio sea).

- Correlación cruzada: cada secuencia en la familia debe ser fácilmente distinguible de (una versión posiblemente desplazada en el tiempo de) cada otra señal de la misma familia. Esto es importante para minimizar la interferencia por multiusuario y para incrementar la capacidad de separación de señales de distintos usuarios.

- Esparcimiento del espectro: la información modulada debe tener un espectro uniforme sobre el ancho de banda del sistema.

El problema es que la optimización de las propiedades de la autocorrelación de las secuencias de una familia se consigue, en general, a expensas de propiedades de correlación cruzada peores, y viceversa.

Por lo tanto, los sets de secuencias que tienen muy buenas propiedades de correlación cruzada usualmente tienen propiedades de autocorrelación malas. Una situación similar ocurre con las propiedades de esparcimiento de la secuencia. Esta es la razón

por la que existen muchas familias y cada familia optimiza alguno de los requisitos detallados anteriormente.

3.2.4. Familias clásicas de Códigos de Spreading

En general, en CDMA se distinguen dos categorías básicas: CDMA asíncrono (mediante secuencias pseudo aleatorias) y sincrónico (mediante códigos ortogonales).

El primer tipo, basado en secuencias pseudo aleatorias, presenta mejores prestaciones cuando los usuarios estan no sincronizados entre sí. Estos códigos tienen valores bajos de correlación cruzada distintos de cero, por lo tanto la interferencia entre usuarios no es nula. Se utilizan en sistemas asíncronos (los transmisores no transmiten sincronizados). También producen interferencia por acceso multiple, determinada por las propiedades de correlación de las secuencias.

El segundo tipo de secuencias son útiles cuando todos los usuarios o señales se hallan perfectamente sincronizados. La correlación cruzada de estas secuencias es cero para desplazamiento cero. Sin embargo, las propiedades de autocorrelación usualmente no son buenas. No existe interferencia por acceso múltiple pero el número de canales esta limitado dependiendo de la cantidad de códigos disponibles. Necesitan un sincronismo muy preciso. Ejemplo de estos códigos son las secuencias Walsh-Hadamard.

Con la introducción de la sincronización del caos [67] se han desarrollado muchas aplicaciones de éste a sistemas prácticos[95, 96, 97, 72, 98, 99]. En esta tesis se analiza la performance de un sistema de comunicaciones de SS con secuencia directa utilizando secuencias caóticas como códigos de spreading. Entre las ventajas de las

series caóticas se destacan la posibilidad ilimitada de generación de secuencias, su facilidad de implementación, y su mejora inherente en cuanto a la seguridad de la comunicación.

- Secuencias de Máxima Longitud.
- Códigos Gold.
- Códigos Kasami.
- Códigos Walsh-Hadamard.
- Códigos Caóticos.

Secuencias de Máxima longitud (M-secuencias)

Las secuencias de máxima longitud o M-secuencias son las más conocidas. La denominación de máxima longitud se debe a que las secuencias generadas tienen la máxima longitud posible con m registros de desplazamiento, por lo tanto son secuencias de longitud $n = 2^m - 1$. Cada valor m define una familia diferente.

Son generadas mediante m-etapas de registros de desplazamiento con una realimentación lineal de acuerdo a un polinomio primitivo.

Se generan por registros de desplazamiento linealmente realimentados según $g(D)$ que es un polinomio primitivo ($g(D)$ de grado r, es primitivo si es factor de $(D^{2^r-1}-1)$ y no es factor de ningún otro polinomio de la forma $(D^N + 1)$ para $N < 2^r - 1$, ej: $g(D) = D^5 + D^2 + 1$).

Dado un arreglo de m registros, el número de secuencias de longitud $2^m - 1$ es limitado. Por ejemplo para $m = 4$, solo hay 2 secuencias de longitud $n = 15$; para

$m = 5$ y $m = 6$ hay solo 6 secuencias. Habrá más secuencias si m es un número primo.

Cada condición inicial distinta resulta en la misma M-secuencia desfasada.

Su principal característica es que el esparcimiento del espectro que presenta es casi óptimo y la autocorrelación es ideal, ya que tiene un único pico correspondiente a cuando ambas señales están perfectamente alineadas.

Otro punto interesante de considerar es el comportamiento de la correlación cruzada entre las secuencias posibles, ya que si el generador PN se quiere utilizar para diferenciar usuarios, es deseable que la correlación cruzada sea baja. Las M-secuencias no tiene un buen comportamiento respecto a este parámetro, ya que la correlación cruzada entre cualquiera de las secuencias de esta familia es una función periódica con picos altos.

Otra característica de estos códigos es que cada secuencia contiene un uno más que cero. Esto sirve para limitar la componente de DC en el espectro de la señal esparcida.

Estas familias de secuencias no se utilizan en la práctica en CDMA principalmente por limitaciones en el número de secuencias por familia. Sin embargo, son la base de otros códigos con familias más numerosas.

Códigos Gold

Las secuencias Gold son muy útiles gracias a la gran cantidad de códigos que poseen por familia. Su generación se basa en dos M-secuencias de una misma familia, las que son elegidas de forma de que los códigos generados posean correlación cruzada uniforme y limitada.

Los códigos Gold se generan mediante la operación XOR (o suma) de dos registros de desplazamiento de m etapas cada uno determinados por las dos M-secuencias. Estas M-secuencias deben ser *preferidas*, esto significa que son dos M-secuencias que presentan mínima correlación cruzada. Incluyendo el par de secuencias preferidas se obtienen un total de $p = 2^m + 1$ códigos Gold.

La correlación cruzada entre dos M-secuencias puede ser tri-valuada, cuatri-valuada, ..., multi-valuada. Si son pares preferidos la correlación es tri-valuada. Para ser pares preferidos deben cumplir:

1. $r \neq 0 mod 4$; (r es impar ó $r = 2 mod 4$).
2. q es impar y $q = 2^k + 1$ ó $q = 2^{2k} - 2^k + 1$.
3. mdc(r,k) = 1 para r impar; o mcd(r,k)= 2 para $r = 2(mod 4)$.

Los códigos Gold generados tienen dos secuencias de longitud máximas, cuyo autocorrelación toma dos valores (un valor pico coincidente con n cuando están en fase, y un valor muy bajo para cualquier corrimiento), pero el resto de las secuencias de la familia tiene una autocorrelación que toma los tres valores de la correlación cruzada.

Secuencias Kasami

Estas secuencias son muy importantes ya que presentan valores de correlación cruzada muy bajos. Su generación se basa, al igual que las Gold, en M-secuencias. Hay dos set diferentes de secuencias Kasami: set de secuencias cortas y largas.

El **set de secuencias cortas** se genera mediante un método similar al utilizado para generar las secuencias Gold, este método genera un set pequeño de $p = 2^{m/2}$

secuencias binarias de periodo $n = 2^m - 1$, con m par. El procedimiento comienza con una M-secuencia a con la que se genera la secuencia a' realizando la decimación por $q = 2^{r/2}+1$, esto es muestrear a a cada q símbolos, obteniéndose $a' = a(2^{r/2}+1)$. Puede verificarse que la secuencia a' generada es una M-secuencia con período $n = 2^{r/2} - 1$. Por ejemplo, si $r = 10$ el período de a es $n = 1023$ y el de a' es 31. Luego se obtiene b repitiendo q veces a a' [71].

La familia contiene las secuencias: a y $a + b$ cíclicas desplazadas $2^{r/2} - 2$.

La desventaja de estos códigos es la implementación práctica, ya que para realizar la decimación se requieren clocks más rápidos.

Las funciones autocorrelación y correlación cruzada de estas secuencias toman los valores -1, $\{-(2^{r/2} + 1,\ 2^{r/2} - 1)\}$

El **set largo** de secuencias Kasami consiste en secuencias de período $n = 2^r - 1$, con r par, y contiene las secuencias Gold y el set pequeño Kasami. Sean a' y a'' M-secuencias obtenidas de la decimación de a por $q' = 2^{r/2} + 1$ y $q'' = 2^{(r+2)/2} + 1$ respectivamente, se toman todas las secuencias obtenidas mediante la adición (XOR) a, a' y a'' con distintos desplazamientos de a' y a''. La cantidad de secuencias que se obtiene es $p = 2^{3r/2}$ si $r = 0(mod4)$, y par mayor, o se obtienen $p = 2^{3r/2} + 2^{r/2}$, si $r = 2(mod4)$.

Las secuencias Kasami se pueden implementar por la multiplicación de la salida de tres registros de desplazamiento (u, v y w) con apropiadas realimentaciones. Dos registros de desplazamiento forman un par de secuencia preferidas, mientras que la tercer secuencia (w) resulta de decimar la primer secuencia (u). Todos los registros de desplazamiento crean M-secuencias.

Para obtener todos los códigos Kasami posibles las tres M-secuencias se deben sumar con desplazamientos relativos distintos una de otra.

La autocorrelación y correlación cruzada de este set toman cinco valores $\{-1, -1 \pm 2^{r/2}, -1 \pm 2^{r/2} + 1\}$.

Códigos Walsh-Hadamard

La principal ventaja de estas secuencias es que tienen correlación cruzada nula siempre y cuando el desfasaje entre secuencias sea nulo, pero para desplazamientos distintos de cero el valor de la correlación cruzada depende fuertemente en las secuencias particulares utilizadas. Son óptimas para comunicaciones sincrónicas. En cuanto a la autocorrelación, esta no es buena ya que no presenta un pico único. El esparcimiento del espectro no cubre todo el ancho de banda disponible, sólo presenta algunas componentes de frecuencia. Para estas familias hay $p = n$ miembros.

Se generan mediante matrices Hadamard, se obtienen 2^r códigos unipolares de longitud $N = 2^r$; con $r \geq 4$. H1=[0] H2=[0 0].. .. [0 1] H2m=[Hm Hm] [Hm Hmc] ;con $m = 2^j$;para $j \geq 1$

Se utilizan en *multicarrier CDMA* y *IS95 cellular*. Donde los usuarios esta sincronizados entre ellos.

Códigos Caóticos

Las secuencias de códigos caóticos estudiadas en esta tesis son obtenidas mediante la iteración de dos mapas caóticos: el mapa Three-Way Bernoulli (TWBM) y el mapa Four-Way Tailed Shift (FWTSM)[72, 98, 99]. Las secuencias generadas por cada mapa fueron luego procesadas mediante la siguiente regla de dinámica simbólica: ($[0, 0{,}5) \rightarrow$

$0, (0{,}5, 1] \rightarrow 1)$ de modo de obtener secuencias binarias. La cantidad de secuencias p por familia no está limitada en este caso.

3.3. Cuantificadores de Performance Propuestos

Se proponen tres nuevos cuantificadores para la comparación de la performance que presentan las familias de códigos PN. El primer cuantificador C tiene en cuenta la calidad de correlación cruzada y la autocorrelación de las secuencias.

El segundo cuantificador propuesto S mide el esparcimiento del espectro sobre el ancho de banda disponible. El tercer cuantificador propuesto Z es un parámetro que cuantifica al sistema completo.

Todos estos cuantificadores son globales, de forma de comparar diferentes familias como un todo, esto significa que incluyen a todos los miembros de cada familia, en lugar de las aproximaciones tradicionales, las cuales utilizan miembros representativos de cada familia elegidos aleatoriamente[100, 101, 102].

Las comparaciones son realizadas entre familias con miembros de igual longitud n. Además, la definición de cuantificadores globales es un primer paso hacia la optimización del proceso.

3.3.1. Cuantificador de performance de Correlación: C.

Este cuantificador incluye las propiedades de autocorrelación y correlación cruzada de cada familia.

Para una detección y sincronización perfectas se requiere que la autocorrelación tenga la forma de una función delta y la correlación cruzada entre los diferentes

miembros de la familia debe ser cero. Este es un caso ideal, sin embargo, lo mejor que podemos obtener es una autocorrelación "tipo delta" y una baja correlación cruzada. Por esta razón se considera una *buena familia* a aquella que presenta valores de correlación cruzada menores al nivel de umbral utilizado por el receptor en el *paso de detección*. Este valor de umbral es un porcentaje del valor de autocorrelación para desplazamiento cero.

El método consiste en:

1. Llamamos $\{x^{(i)}, (i = 1, ...p)\}$ a los miembros de una dada familia, cada miembro es un vector de longitud n de valores binarios. Calcular todas las correlaciones entre todos los pares de la familia, para todos los corrimientos de tiempo circulares.

$$c^{ij}(s) = \begin{cases} \sum_{k=0}^{n-|s|-1} x_k^i \, x_{k+s}^j & s \geq 0 \\ \\ c^{ij}(-s) & s < 0 \end{cases} \tag{3.3.1}$$

   El número de valores es $N_{corr} = 2\, p^2\, n$.

2. Elegir el nivel de threshold c_{th} definido como una fracción del valor de autocorrelación para corrimiento cero A_0.

3. Contar la cantidad de elementos mayores que el nivel de umbral n_{corr} y normalizar de la siguiente forma:

$$\tilde{n}_{corr} = n_{corr}/N_{corr} \tag{3.3.2}$$

   La Tabla 3.1 muestra que $\tilde{n}_{corr}$ es una función decreciente de c_{th}.

			$\tilde{n}_{corr}$		
c_{th}	M-seq	Gold	FWTSM	TWBM	Walsh
1	0	0	0	0	0.0004
0.9	0	0	0	0	0.0005
0.8	0	0	0	0	0.0010
0.7	0	0	0	0	0.0019
0.6	0	0	0	0	0.0029
0.5	0	0	0	0	0.0057
0.4	0	0	0	0.0001	0.0090
0.3	0.0028	0	0.0024	0.0023	0.0169
0.2	0.0028	0	0.0380	0.0384	0.0310
0.1	0.4164	0.4961	0.3401	0.3412	0.0703
0.0	1.0000	1.0000	1.0000	1.0000	1.0000

Cuadro 3.1: $\tilde{n}_{corr}$ como función del valor de threshold para diferentes familias con $n = 127$.

4. Finalmente C es el mínimo valor de threshold c_{th} que da $\tilde{n}_{corr} = 0$.

En la Fig. 3.1 se ven los valores para las familias estudiadas. En el eje x se ordenan las familias desde la que presenta mejor performance (izquierda) a la peor (derecha).

3.3.2. Cuantificador Global del Espectro: S.

Idealmente se desea que la señal modulada presente un espectro constante sobre toda la banda disponible. Pero, como los códigos PN son periódicos tienen un espectro discreto. Si la señal modulada esta concentrada en un pequeño grupo de componentes de frecuencias discretas el esparcimiento no es muy eficiente. El cuantificador propuesto analiza esta situación, y se calcula de la siguiente forma:

1. Evaluar la magnitud de la Transformada Rápida de Fourier (FFT) para cada código PN de la familia (ver Fig.3.2a).

2. Normalizar cada magnitud dividiéndolas por su valor medio.

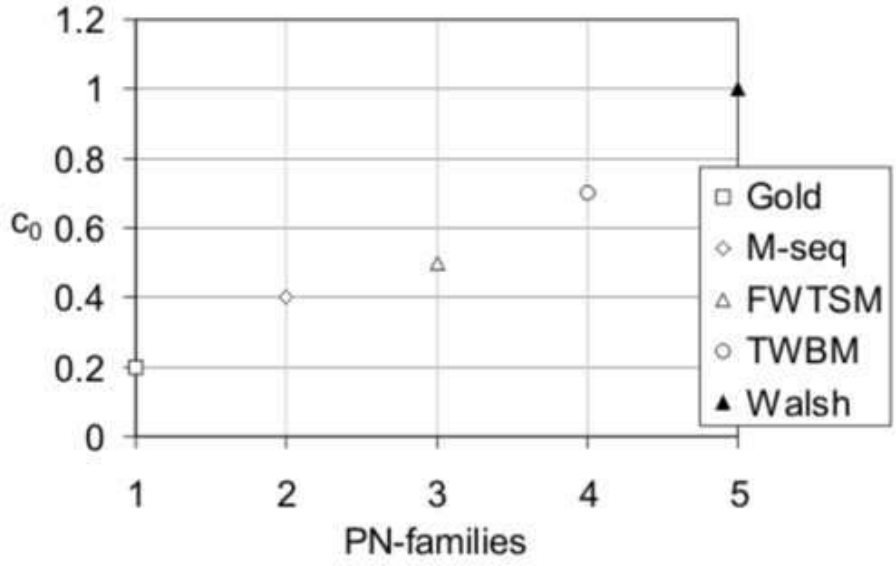

Figura 3.1: Cuantificador global de Correlación C como función del valor de threshold c_{th} para varias familias PN con $n = 127$, menos para la familia Walsh donde $n = 128$.

3. Calcular la varianza σ_i, de cada vector de magnitud de FFT normalizado (ver la etiqueta de cada panel en la Fig. 3.2a).

4. El cuantificador del espectro S es el valor medio sobre todos los miembros de la familia, $S =< \sigma_i >$.

Cada gráfico de la Fig. 3.2b muestra σ_i para una dada familia, como función de $n/2$ frecuencias discretas. El orden de las familias es de arriba hacia abajo: M-secuencias, Gold, FWSTM, TWBM, y Walsh. S es mostrada como etiqueta y también con una línea horizontal en cada subgráfico.

La familia de las M-secuencias presentan el menor S y por lo tanto el espectro más uniforme. Las familias Gold, TWTSM y FWBSM dan resultados similares. Finalmente, la familia Walsh presenta el mayor valor de S mostrando que el esparcimiento

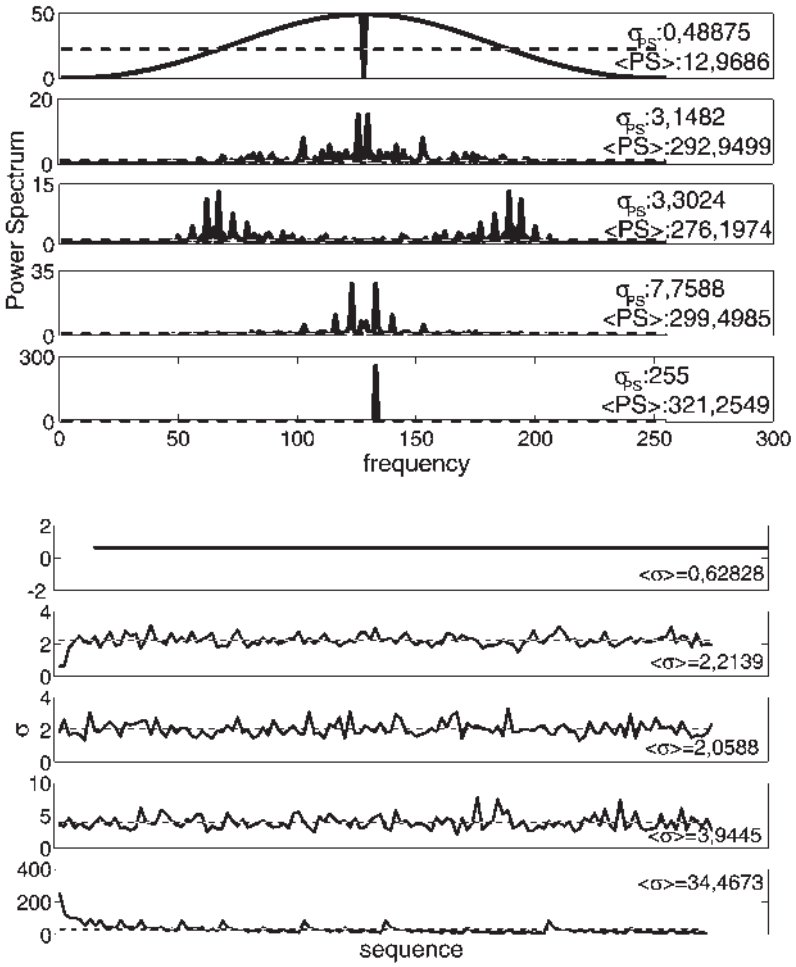

Figura 3.2: (a) Espectro típico de un código PN para diferentes familias, se muestran con etiquetas el valor medio y la varianza σ_i en cada subgráfico; (b) Varianzas σ_i para códigos PN de diferentes familias. Se muestra como etiqueta el valor de S en cada subgráfico.

del espectro es el peor ya que se concentra en pocas componentes de frecuencias.

Es posible eliminar de cada familia los códigos con valores altos de σ_i con el objeto de mejorar la performance de toda la familia. Por supuesto, esta decisión disminuye el valor de p, que es el número permitido de usuarios simultáneos.

3.3.3. Cuantificador de complejidad de Zipping: Z.

Para una cadena de caracteres la complejidad algorítmica es definida como la longitud en bits del menor programa que produce a la string como salida [103]. El problema con esta definición es que es imposible, por lo menos en principio, encontrar tal programa. Sin embargo, los zippers o compresores de archivos son algoritmos concebidos para realizar ese trabajo, por lo menos aproximadamente. Los algoritmos de Lempel y Ziv son utilizados por la mayoría de los zippers y es uno de los mejores compresores de archivos conocidos [15, 14].

Para una dada cadena finita la complejidad de zipping puede ser definida como l_z/l donde l, es la longitud de la cadena y l_z es la longitud de la cadena comprimida. El problema es que la habilidad de compresión del zipper depende de la longitud de la cadena y también del orden de los códigos PN dentro de la cadena. Luego, para esta aplicación particular es necesario construir series subrogadas para obtener parámetros independientes de estos dos factores. El cuantificador de complejidad Z es evalúa de la siguiente forma:

1. Construir una cadena binaria consistente de todos los miembros de la familia, un código seguido de otro. El tamaño de la cadena queda definido por la familia PN con mayor valor de p $l = Np$. Las familias con valores de p menores se completan repitiendo sus propios códigos.

familia	N	p	l	Z
Gold	31	33	2,112	0.1416
FWTSM	31	33	2,112	0.1330
TWBM	31	33	2,112	0.1311
Walsh	32	32	2,112	0.0971
M-seq	63	6	768	0.1575
Gold	63	65	8,320	0.1174
FWTSM	63	520	66,560	0.1111
TWBM	63	520	66,560	0.1113
Walsh	64	64	8,320	0.0534
M-seq	127	18	4,608	0.1220
Gold	127	129	33,024	0.1096
FWTSM	127	129	33,024	0.1075
TWBM	127	129	33,024	0.1075
Walsh	128	128	33,024	0.0032

Cuadro 3.2: Global Quantifier Z for different families and different lengths

2. Generar $p!$ subrogados cambiando el orden de los códigos PN dentro de la cadena original con todas las combinaciones posibles.

3. Comprimir cada subrogado para obtener z_i. El cuantificador global Z esta dado por el valor medio de todos los subrogados, $Z =< z_i >$.

La Tabla 3.2 muestra los resultados obtenidos para las familias clásicas y las familias caóticas. Este parámetro no es completamente independiente de los dos definidos anteriormente ya que:

- La cadena de una familia PN con altos valores de correlación cruzada entre sus miembros puede ser más comprimida.

- Una familia PN con espectro coloreado presenta periodicidades y por lo tanto la cadena podrá comprimirse más.

familia	N	p	C	S	Z
Gold	127	129	0.2	3.1148	0.1174
FWTSM	127	129	0.4	3.3024	0.1109
TWBM	127	129	0.5	7.7588	0.1104
M-seq	127	19	0.4	0.6202	0.0072
Walsh	128	128	1	255.00	0.0363

Cuadro 3.3: Proposed global quantifiers for different classical and chaotic PN-families

- Una familia que presente menos miembros debe repetir sus códigos más veces para completar la cadena, luego, podrá ser más comprimida.

Luego el valor de Z disminuye cuando los valores de la correlación aumentan, el espectro no será plano y la cantidad de miembros disminuye como muestra la Tabla 3.3.

3.4. Conclusiones

Se definieron tres cuantificadores. Estos son el cuantificador de correlación C, el cuantificador de spreading S y el cuantificador de zipping Z. Estos cuantificadores son globales en el sentido de que existe un valor para cada familia PN. El cuantificador de zipping aparece como un posible cuantificador para comparar distintas familias, ya que Z sera mayor en cadenas consistentes de códigos PN correlacionados y disminuye si el espectro no es plano. Se estudiaron las familias clásicas y caóticas con los cuantificadores propuestos.

Los resultados presentados muestran que las familias caóticas pueden tener una mayor cantidad de miembros que las familias convencionales, con propiedades de spreading y correlación equivalentes. Además son fácilmente implementados y consecuentemente son buenos candidatos para ser utilizados en sistemas reales.

Capítulo 4

Caos y la reducción de la interferencia electromagnética (EMI)

4.1. Introducción

Debido a la presencia de aparatos eléctricos y electrónicos en la vida diaria, es muy importante el diseño de sistemas compatibles electromagnéticamente (EMC por Electro Magnetic Compatibility). Las técnicas usuales para incrementar el EMC de los sistemas se basan en la adopción de filtros, cables y conectores blindados. Estas metodologías *a-posteriori* tienen varias desventajas. La más notable es el incremento del costo de los aparatos, además, en algunos casos, estas metodologías no pueden aplicarse.

La mayoría de los sistemas digitales cuentan con algún tipo de oscilador local que genera una señal de clock para mantener el sincronismo de otros componentes. Debido a las características de la señal, periódica y con flancos abruptos, presenta un espectro de potencia con toda su energía concentrada en pocas frecuencias, la del oscilador y sus armónicas impares. Por lo tanto, radiará picos de energía en estas frecuencias,

resultando en un espectro que puede exceder los límites regulares de la interferencia electromagnética. Ejemplo de esto son las plaquetas con microprocesadores que utilizan señales de clock o las mixtas, que tienen componentes tanto analógicos como digitales, las señales de clock de los componentes digitales pueden interferir con la parte analógica del circuito.

Por lo tanto, estos "tonos" que generan las señales de clock se emiten en forma de interferencia al resto de los componentes que se encuentren en el área. Existen regulaciones sobre la potencia de los clocks en aparatos electrónicos para prevenir estas interferencias.

En este capítulo se investiga una solución para evitar este problema. La técnica que se presenta se basa en la introducción de una pequeña modulación en la señal de clock [75, 104, 105, 106], de forma tal que presente una pequeña variación, en el sentido que continúe cumpliendo eficientemente su función de sincronizar pero con una frecuencia variable, muy cercada a la deseada, y así romper su periodicidad. El objetivo es obtener una señal cuyo espectro sea lo más constante posible dentro de una banda de frecuencias para conseguir un nivel lo suficientemente bajo para que no interfiera con el resto de los componentes. Las señales con estas características son llamadas señales CEW (Constant Envelop Wideband signals).

Tradicionalmente la señal de clock se modula con una secuencia pseudo-aleatoria, en esta tesis se investiga la utilización de secuencias caóticas randomizadas en reemplazo de las pseudo-aleatorias ya que las caóticas tienen la ventaja de ser más sencillas de implementar en hardware.

La principal contribución de este capítulo es evaluar el uso de secuencias generadas por mapas caóticos y randomizadas según los lineamientos del capítulo 2 para la

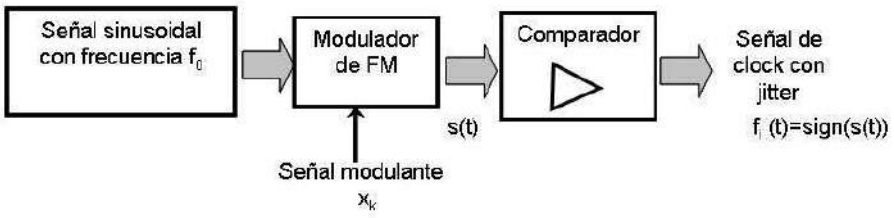

Figura 4.1: Esquema de generación de señales de clock CEW.

generación de señales de clock CEW. Los resultados muestran que la calidad obtenida con sistemas caóticos es buena y comparable a la obtenida con sistemas pseudo-estocásticos.

4.2. Generación de señales de *clock* CEW mediante modulación en FM

Una forma sencilla de implementar esta técnica es mostrada en la Fig. 4.1. En ella se puede ver un diagrama en bloques simplificado, el primer paso es modular en FM una señal sinusoidal cuya frecuencia sea igual a la del clock deseado, f_0. Luego, se aplica un nivel de comparación de forma de obtener un tren de pulsos, $f_i(t) = sign[s(t)]$, que será el nuevo clock del sistema. ver Fig. 4.2. $f_i(t)$ presentara un pequeño "jitter" que debe ser imperceptible para los dispositivos que utilicen el clock.

Por ejemplo, un procesador de $1GHz$ la frecuencia podría ser $999,5MHz$ en un momento y $1,0005GHz$ en otro. Por supuesto, la máxima desviación de frecuencia

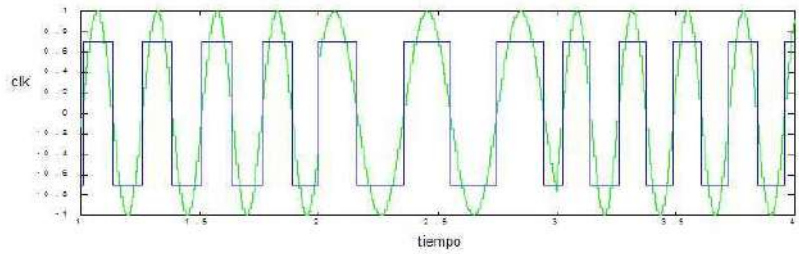

Figura 4.2: Nueva señal de clock con "jitter".

debe ser compatible con los componentes que dependen del clock para una correcta operación. La diferencia de frecuencia del nuevo clock con f_0, $|f_0 - [frecuencia\,de\,s(t)]|$, no debe superar en ningún momento un porcentaje de f_0, generalmente este valor se encuentra entre 5 y 10.

Como consecuencia de esta modulación el espectro se esparce en un ancho de banda mayor (centrado en la frecuencia deseada) y, por lo tanto, se consigue una reducción en los picos de energía irradiada. Con este método se logra cumplir con las regulaciones de EMC.

Muchas PC tienen un seteo en el BIOS para habilitar el clock con espectro esparcido.

Es importante notar que este método no reduce los picos eléctricos de la señal de clock, tampoco reduce la fuerza del campo eléctrico ni magnético emitido por el sistema, por lo tanto no reduce la potencia. Esta técnica distribuye la energía de forma tal que el espectro se presente esparcido y no concentrado en una frecuencia en particular, y así no interfiera con otros sistemas.

Se consideraran tres casos según la señal modulante, x_k :

1. La señal de FM es modulada por una señal PAM (Pulse Amplitud Modulated). La secuencia modulante x_k son muestras uniformemente distribuidas pertenecientes al intervalo $[-1, 1]$. ("Modulación FM uniforme").

2. La señal de FM es también modulada por una señal PAM, pero la secuencia puede tener los valores -1 y 1 únicamente con igual probabilidad ("Modulación FM binaria").

3. La señal modulante es la suma de una señal PAM uniforme y una binaria ("Modulación FM híbrida").

4.2.1. Modulación FM Uniforme

Este caso es muy interesante ya que, como se verá más adelante, bajo ciertas condiciones, el espectro de la señal modulada es proporcional a la función densidad de probabilidad (PDF) de la secuencia modulante. Esta característica permitiría sintetizar una señal CEW conforme a un requerimiento espectral dado, es decir, a partir de requerimientos en frecuencia, es posible sintetizar una secuencia con PDF de acuerdo a las especificaciones dadas. Por lo tanto, la señal resultante de la modulación de una portadora con esta secuencia producirá una señal cuyo espectro cumplirá los requisitos.

Señales CEW

Nos interesa la generación de señales CEW, como su nombre lo indica estas señales presentan un espectro ancho y constante.

La señal CEW es indicada como $s(t)$:

$$s(t) = Re(e^{2\pi i(f_0 t + \Delta f \int_\infty^t \xi(\tau) d\tau)}) \tag{4.2.1}$$

donde f_0 es la frecuencia de portadora, Δf es la desviacion de frecuencia, y $\xi(t)$ es la señal PAM modulante definida:

$$\xi(t) = \Delta f \Sigma_{k=-\infty}^{\infty} x_k g(t - kT) \tag{4.2.2}$$

donde $g(t)$ es un pulso unitario de duración T, Δf es la desviación de frecuencia x_k es la secuencia modulante, con $x_k \, \epsilon \, [-1, 1]$.

El espectro deseado tiene un valor constante en el rango de $[f_0 - \Delta f, f_0 + \Delta f]$.

Para expresar mejor las características de la modulación se define el índice de modulación $m = T\Delta f$, este índice cuantifica la modulación cuando la frecuencia de portadora esta fijada (f_0) y se trabaja con los equivalentes pasa bajo, m contribuye a la forma del espectro, mientras que f_0 sólo lo ubica en el eje de frecuencias.

El la Fig. 4.3 puede verse el efecto de la modulación en FM de la portadora mediante una secuencia x_k modulante, allí se ve como aumenta y disminuye la frecuencia de portadora proporcionalmente a la secuencia. La máxima desviación de frecuencia será $f_{0MAX} = f_0 + \Delta f x_{kMAX}$ y la mínima $f_{0MIN} = f_0 - \Delta f x_{kMAX}$.

Si las leyes de desviación de frecuencias (secuencia modulante) son periódicas, la potencia de la señal modulada estará densamente concentrada alrededor de frecuencias especificas y, por lo tanto, la interferencia que producirá seguirá siendo relativamente alta. Una forma de conseguir densidades de potencia menores es usar una señal de modulación no periódica, tales como las generadas por sistemas de mapas caóticos discretos o señales aleatorias.

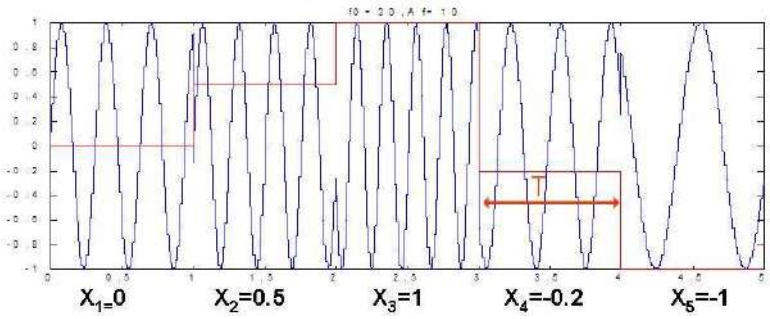

Figura 4.3: Portadora modulada.

Las secuencias aleatorias son muy usadas en aplicaciones de electrónica digital. Estas pueden basarse en fuentes de ruido físicas o en algoritmos matemáticos, pero en ninguno de estos casos se obtienen secuencias verdaderamente aleatorias. Sino que toda implementación real produce números pseudo aleatorios.

Este no es un obstáculo ya que los generadores de números pseudo aleatorios (PRNGs) se han desarrollado de forma de cumplir con las propiedades estadísticas requeridas por las aplicaciones.

Sin embargo, las secuencias generadas por mapas caóticos son buenas candidatas para reemplazar a los PPRNGs, ya que su implementación en hardware es más simple porque se basan en ecuaciones simples con alguna nolinealidad.

Dependiendo en cómo es obtenida la secuencia modulante llamaremos señales FM-aleatoria y FM-caótica.

La eficiencia de este método depende críticamente de las propiedades estadísticas de la señal modulante, es decir de la secuencia x_k.

Existen tres características esenciales que determinan cuán aleatoria es una secuencia:

- El espectro de potencia.
- La función densidad de probabilidad (o la medida invariante , IPDF, en el caso de mapas)
- La calidad de mixing, esta se ve en un embedding n-dimensional (con n=1, 2,...) cuanto más uniforme sea el cubrimiento *mejor* es la secuencia.

En la Fig.4.4 puede verse el espectro de potencia correspondiente a una señal de clock modulada con una secuencia periódica de tipo Diente de Sierra y en la Fig.4.5 se moduló con una secuencia aleatoria; en ambos casos se usó $f_0 = 10$, $T = 1,9905$ y $m = 9,9524$. Se ve que el esparcimiento del espectro no es bueno en el caso de la señal modulante periódica, a pesar de que ambas secuencias comparten una PDF constante.

Calegari et al. [104, 106] demostraron tres teoremas importantes que relacionan la PDF de una secuencia x_k con la densidad espectral de potencia (PDS) de una función sinusoidal modulada en FM por esta secuencia.

Estos teoremas son los siguientes:

Teorema:

$s(t)$ es la señal CEW obtenida modulando en FM con desviación de frecuencia Δf por una secuencia x_k que son muestras independientes con período T, cuya PDF es $\rho(x)$.

Teorema 1: Espectro general de señal FM-aleatoria El espectro de potencia de la señal de FM esta dado por:

$$\Phi_{ss}(f) = E_x[K_1(x,f) + Re(\frac{E_x^2[K_2(x,f)]}{1 - E_x[K_3(x,f)]})] \tag{4.2.3}$$

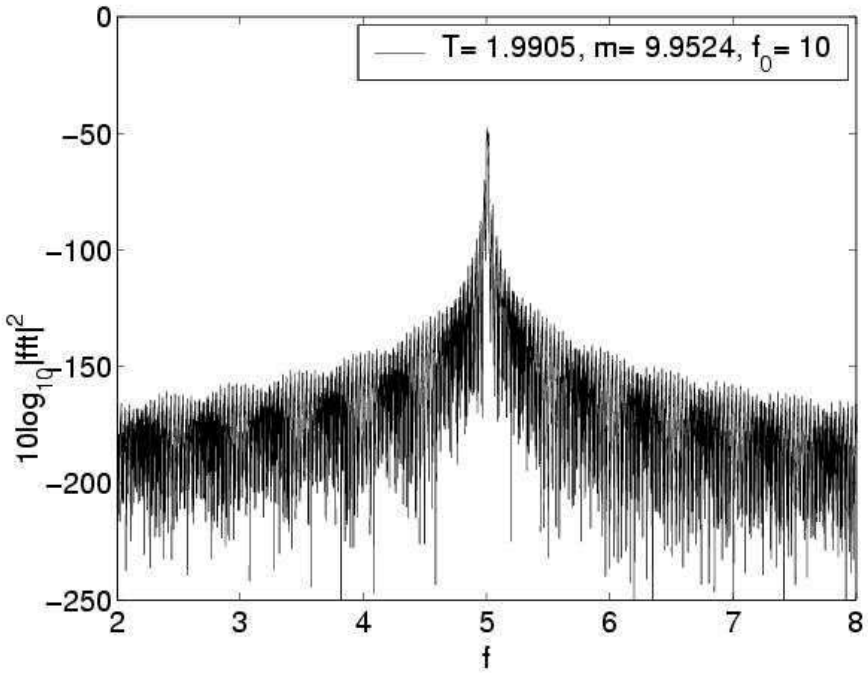

Figura 4.4: Portadora modulada con una señal periódica tipo Diente de Sierra.

Donde:

$K_1(x,f) = \frac{1}{2}Tsinc(\pi T(f-\Delta f_x))$

$K_2(x,f) = i\frac{e^{-i2\pi T(f-\Delta f_x)}}{2\pi\sqrt{T}(f-\Delta f_x)}$

$K_3(x,f) = e^{(-i2\pi T(f-\Delta f_x))}$

Teorema 2: Espectro de señal de FM con modulación lenta

$$\lim_{T\to\infty} \Phi_{ss}(f) = \frac{1}{2\Delta f}\rho(\frac{f}{\Delta f})$$

Teorema 3: Simetría del espectro de señal de FM-aleatoria

$$\rho(x) = \rho(-x) \Rightarrow \Phi_{ss}(f) = \Phi_{ss}(-f)$$

En resumen, la forma de la PDS de una señal de FM tiende a la PDF de la secuencia modulante a medida que T tiende a infinito.

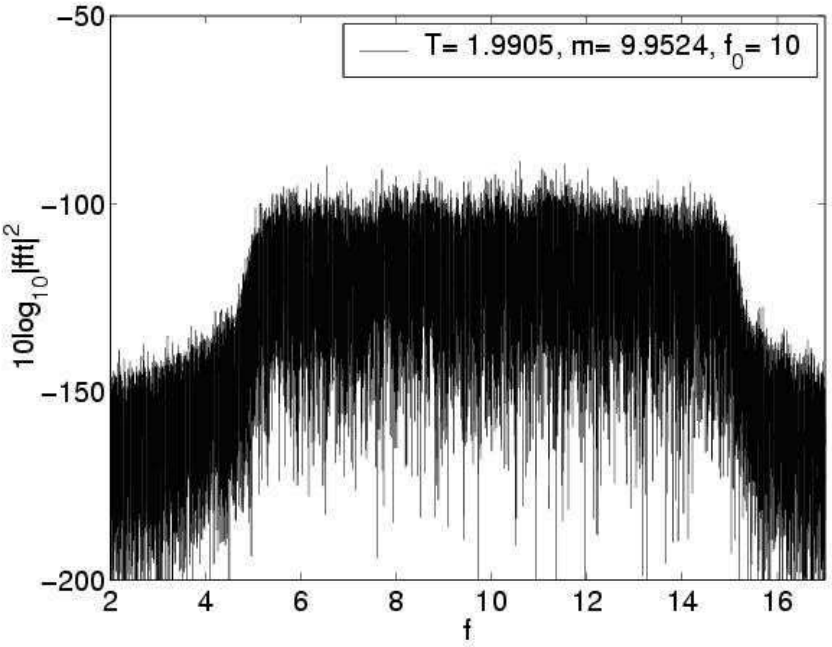

Figura 4.5: Portadora modulada con una señal aleatoria generada mediante la función "rand" de Matlab.

Los mapas caóticos son buenos candidatos para generar PRNGs. Es posible sintetizar mapas caóticos lineales por tramos que presente una densidad invariante deseada, esto es conocido como el problema inverso de Perron-Frobenius (IFPP). Sin embargo, los sistemas basados en mapas lineales por tramos no son simples de implementar en computadoras. Por ejemplo los mapas Tent y Bernoulli tienen puntos de equilibrio parásitos que anulan el comportamiento caótico deseado del sistema. Esto se debe a la diferencia entre la aritmética infinita de los cálculos teóricos y la aritmética finita de la implementación real.

Cualquier mapa no lineal unidimensional puede ser expresado de la forma: $x_k = M(x_{k-1})$

Donde $M : [-1, 1] \rightarrow [-1, 1]$ es una función no lineal.

La eficiencia de mixing de un mapa M es medida por r_{mix}, éste es el segundo autovalor de mayor módulo del operador Perron-Frobenius del Mapa. Cuanto menor es el valor de r_{mix} menos correlacionadas estarán las secuencias producidas por el sistema caótico $1D$.

Mediante la representación en espacio de embedding en $2D$ y $3D$ es posible observar cualitativamente las correlaciones y estimar las propiedades de mixing de las secuencias, ya que r_{mix} es difícil de calcular y en algunos casos imposible.

Las fuentes aleatorias ideales presentarían una distribución uniforme en esos embedding, los mapas caóticos $1D$, en cambio, generan curvas.

En las Figs. 4.6 y 4.7 pueden verse los embedding $3D$ de secuencias generada por la función *rand* de *Matlab* y por el Mapa Logístico $x_k = rx_k(1 - x_k)$ con $r = 4$.

Comparación del embedding $3D$ de una fuente aleatoria y una caótica:

Calegari et al. demostraron que para que se cumpla el *teorema 2* el mapa caótico debe tener $r_{mix} \to 0$, consecuentemente para conseguir la señal modulada en FM con la PDS deseada se debe cumplir:

$$r_{mix} \to 0$$
$$T \to \infty$$

$$\rho(x) = \begin{cases} 1/2 & \text{si } x \in [-1, +1] \\ 0 & \text{para todo otro valor.} \end{cases}$$

Una forma de disminuir el valor de r_{mix} es mediante la técnica de skipping (es equivalente a la utilización de un mapa iterado). El procedimiento de skipping no cambia la PDF de la secuencia pero produce un mejor mixing como se ve en la Fig.4.8.

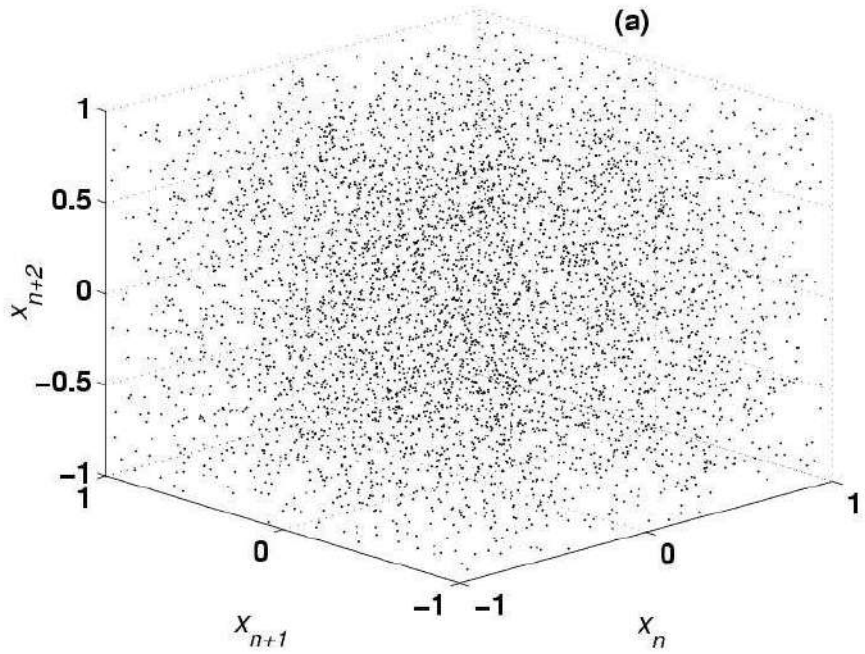

Figura 4.6: Puntos generados por la función "rand" de Matlab

El Mapa Logístico no tiene la IPDF requerida para las aplicaciones de EMI, ya que no es constante (Fig.4.9), por lo tanto, aunque utilizamos la técnica de skipping, disminuiríamos el r_{mix} del mapa, la PDS se aproximaría a la IPDF del mapa, pero esta IPDF no es buena. Por otro lado, hay que utilizar otro método que también modifique la IPDF del mapa. Una buena opción es el método de Discretización propuesto en 2, este método logra modificar simultáneamente el valor de r_{mix} y la IPDF del mapa.

4.3. Aplicación del método de Discretización.

El método, ya detallado en el capítulo 2, consta de los siguientes pasos:

1. Comenzar con un mapa caótico arbitrario. La selección debe realizarse de forma de simplificar la implementación. En nuestro caso utilizamos el mapa Logístico

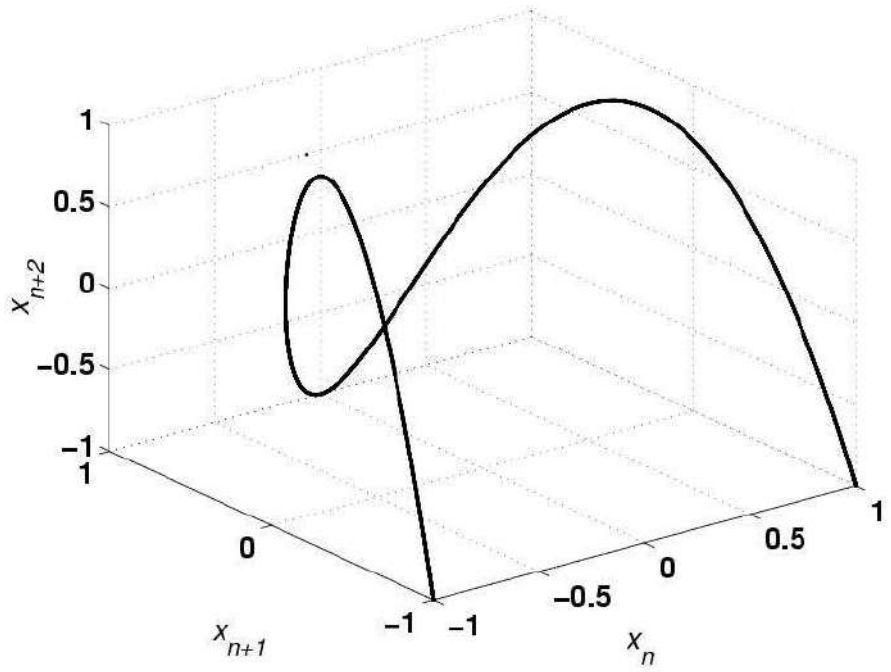

Figura 4.7: Puntos generados con Mapa Logístico con $r = 4$.

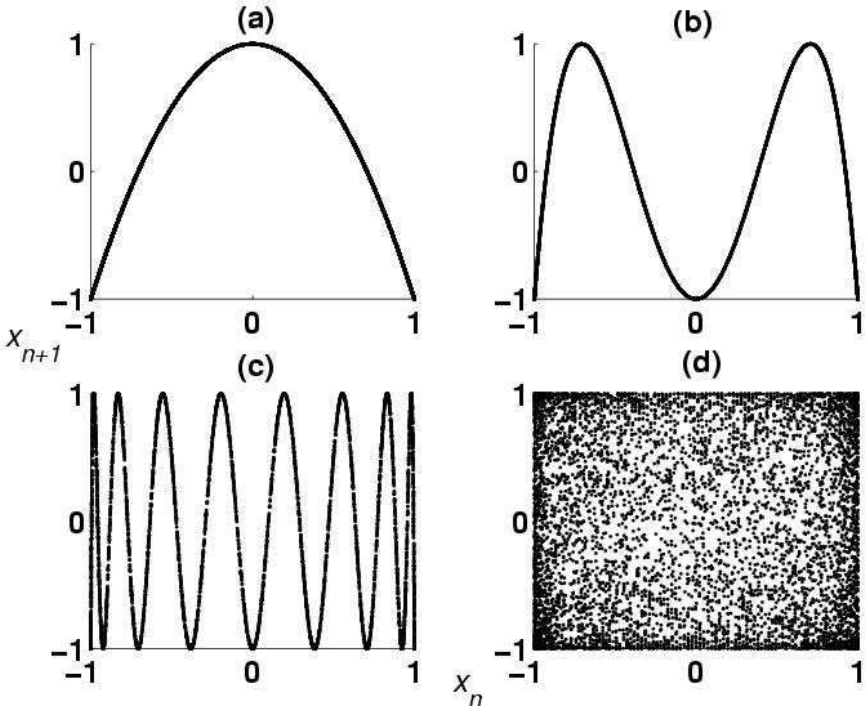

Figura 4.8: Comparación de embedding 2D del Mapa Logístico y sus iterados: a) Mapa Logístico M con r = 4; b) M2; c) M3; d) M4.

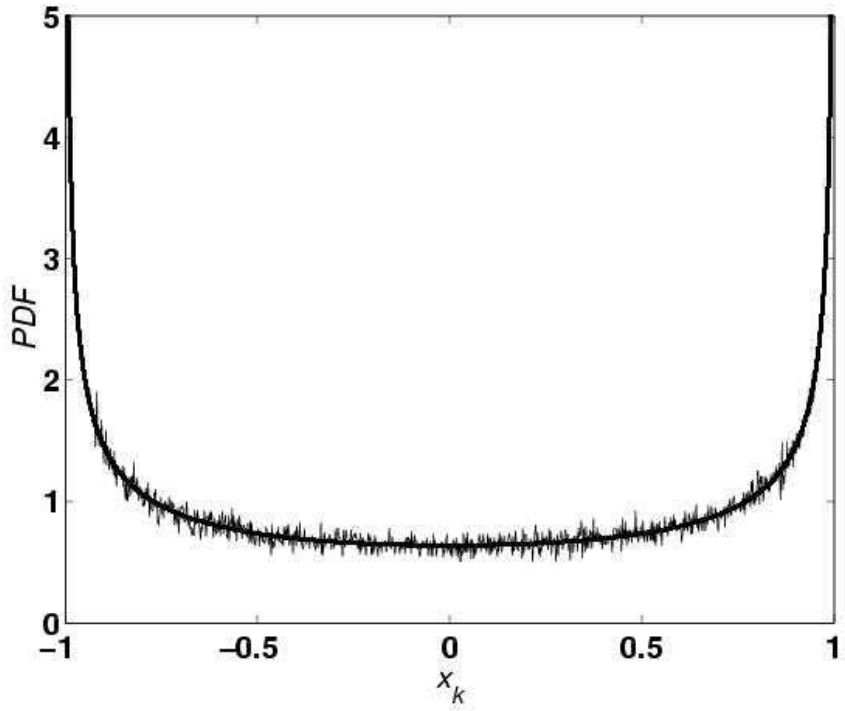

Figura 4.9: IPDF de Mapa Logístico con $r = 4$.

con $r = 4$. La IPDF del Mapa Logístico puede verse en la Fig. 4.9.

2. Normalizar y discretizar la secuencia para convertirla en números naturales de $N-bits$. Esta operación se puede expresar: $x'_n = floor[x_n(2^N)] = floor[x_n 65536]$. En nuestro caso se usaron 16 bits ($N = 16$), lo que significa que los valores generados por el mapa Logístico, que son números decimales en el rango $(0, 1)$, son convertidos a números naturales en el rango $[0, 65535]$.

3. Para cada valor obtenido se descartan todos los bits menos el menos significativo.

4. Se agrupan estos bits de a N para obtener nuevamente números naturales de N bits.

5. Normalizar a números racionales en el rango $[0, 1]$ dividiendo cada número de la secuencia por $2^N - 1$.

6. En este caso, como se requieren números en el rango $[-1, 1]$, a la secuencia anteriormente obtenida se le aplica $x''_n = x'_n 2 - 1$, de forma de llevar la secuencia al rango requerido.

4.3.1. Resultados obtenidos

Se moduló en FM una portadora mediante secuencias generadas por distintos mapas caóticos, iteraciones de estos y secuencias randomizadas mediante el método de Discretización. A continuación se muestran algunos resultados obtenidos.

Las figuras 4.10 a 4.12 muestran el espectro obtenido con la secuencia original del mapa (Fig. 4.10), el espectro obtenido con la secuencia modificada mediante el

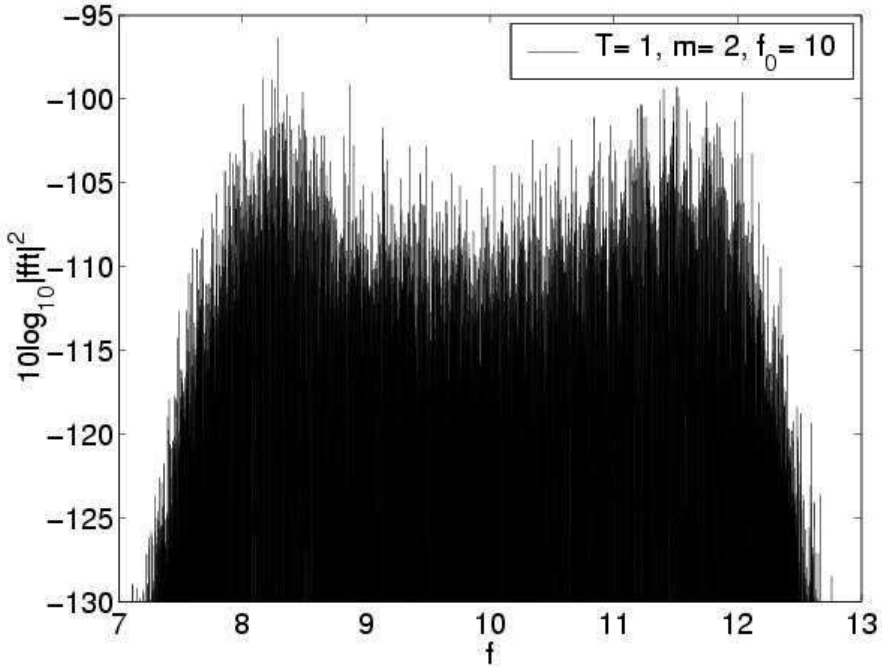

Figura 4.10: Espectro de potencia de portadora modulado en FM con secuencia de Mapa Logístico.

método de skipping, tomando la 8−va iteración del mapa (Fig. 4.11) y con el método de randomización propuesto (Fig. 4.12).

En esta figura pude verse claramente que el espectro obtenido modulando con la $8-va$ iteración es idéntico a la IPDF del mapa Logístico, como predicen los teoremas de Calegari et al.

Pero esta IPDF no cumple con las condiciones requeridas, por lo tanto, el espectro obtenido tampoco. Con el método propuesto (ver Fig. 4.12) obtenemos un espectro plano muy similar al obtenido con la secuencia pseudo-aleatoria.

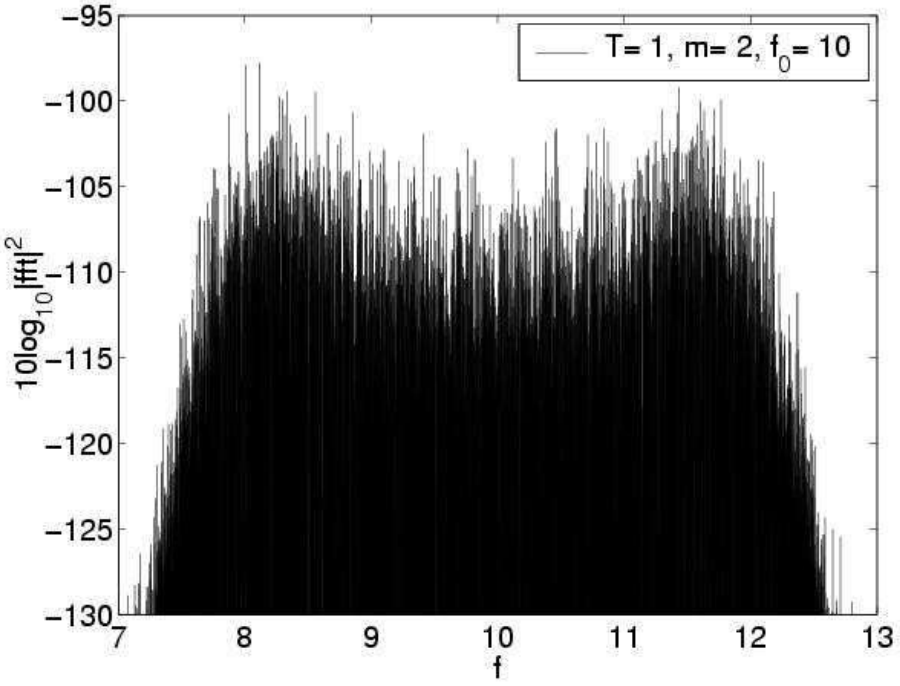

Figura 4.11: Espectro de potencia de portadora modulado en FM con secuencia de la $8 - va$ iteración del Mapa Logístico (Skipping).

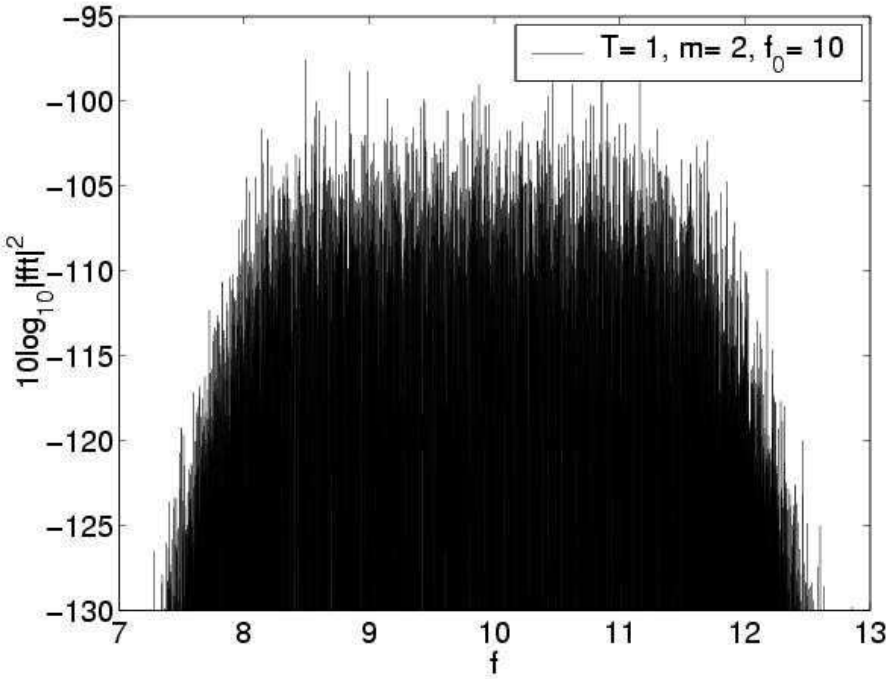

Figura 4.12: Espectro de potencia de portadora modulado en FM con secuencia del Mapa Logístico randomizada mediante la técnica de Discretización.

Capítulo 5

Muestreo caótico para la Adquisición de Señales de baja frecuencia inmersas en Ruido de alta frecuencia

5.1. Introducción

Un sistema de adquisición de señales es un equipo que permite tomar señales físicas del entorno y convertirlas en datos para luego ser procesadas. Para realizar esto la señal a adquirirse es primero acondicionada y luego muestreada.

Desde el punto de vista frecuencial, cuando una señal limitada en banda es muestreada por un tren de deltas (muestreo ideal), en la frecuencia se produce la convolución del espectro de ambas señales. Estas son, el espectro de la señal de interés con un

tren de deltas (transformada de la señal se muestreo).

La convolución da como resultado el espectro de la señal de interés repetido en cada una de las deltas, estas se encuentran separadas a una distancia coincidente con la frecuencia de muestreo. Luego, medante un filtro pasa bajos, se recupera la señal original.

Esto es posible siempre y cuando la separación de las deltas, o sea la frecuencia de muestreo, sea mayor que el doble de la frecuencia máxima de la señal. De lo contrario habrá solapamiento de los espectros, y no sera posible recuperar la señal. Este efecto es llamado aliasing.

En ele spacio temporal puede entenderse como mediciones demasiado espaciadas lo que provoca que se pierdan "detalles" de la señal original y así no puede ser reconstruida de forma unívoca a partir de las muestras digitales.

Este requerimiento impone límites tecnológicos para el procesamiento de señales de alta frecuencia, ya que para muestrear a más del doble de la frecuencia máxima se requieren velocidades de conversión y de procesamiento muy altas.

En la práctica, muchas veces, debido a limitaciones de los circuitos, no es posible cumplir con Nyquist. Entonces es necesario buscar otros métodos para evitar o disminuir el aliasing. La solución estandar, en estos casos, es el uso de filtros antialiasing. La función de estos filtros es atenuar lo más posible el contenido armónico superior a la frecuencia de Nyquist. Para que luego al ser muestreada no haya solapamiento de espectros, y luego con el filtro pasa bajos sea posible recuperar la señal.

Los filtros antialiasing son analógicos y se insertan en la etapa de acondicionamiento de la señal, previas a la conversión analógico-digital. La implementación de estos filtros implica una superficie de la placa electrónica que puede ser significativa si la

cantidad de canales es importante.

Otra posibilidad es el muestreo no periódico, éste reduce el efecto de aliasing en forma natural permitiendo eliminar los filtros físicos.

En este capítulo se investiga el uso de muestreo caótico para adquirir señales de baja frecuencia inmersas en ruido de alta frecuencia. El método reduce el aliasing provocado por el ruido sin utilizar los típicos filtros antialiasing. Se realizaron simulaciones con varios mapas caóticos que poseen diferentes IPDFs y diferente constante de tiempo de evolución en el espacio de las PDFs. Se reportan resultados con diferentes tipos de filtros de respuesta finita al impulso (FIR) y se los compara con los obtenidos en trabajos previos con muestreo aleatorio [107].

5.2. Muestreo no periódico

Varios autores han tratado el tema del muestreo no periódico como por ejemplo [108, 109, 110] en estos trabajos se estudia el problema de recuperar una señal a partir de una secuencia de muestras tomadas en instantes de muestreo irregulares. Típicamente, son utilizados algoritmos iterativos para recuperar la señal.

En [111] se estudia el caso de la utilización del muestreo aleatorio para determinar la potencia espectral de un proceso estacionario aleatorio, mediante la eliminación del aliasing.

En los trabajos [112, 113, 114, 115] se utiliza para analizar las distorciones producidas por el jitter aleatorio que ocurre en los sistemas que requieren muestreo uniforme.

En especial, el muestreo no periódico ha demostrado ser un método efectivo para

el filtrado de señales de baja frecuencia inmersas en ruido de alta frecuencia como se ve en [116, 107].

Se analiza el caso de una señal de baja frecuencia, la que presenta acoplado un ruido de alta frecuencia, siendo esta frecuencia superior a la permitida por Nyquist.

Una forma intuitiva de mitigar el aliasing producido por la señal de alta frecuencia interferente, consistiría en esparcir el espectro de esta, lo cual se puede lograr, por ejemplo, muestreando en forma aleatoria en lugar de uniforme.

Este muestreo no tendrá efecto sobre la señal de continua, ni tampoco en una señal de frecuencia baja, sin embargo logra "destruir" la periodicidad de la señal indeseada y así esparcir su esepctro. Luego, una vez reducido el aliasing, el ruido provocado por la propia aleatoriedad del muestreo podría ser atenuado por medio de un filtro de reconstrucción tipo pasa-bajos

Se han definido en la literatura dos tipos de muestreo no periódico según cuál sea la función de muestreo empleada, estos son: el Additive Asynchronous Sampling (AAS) y el Jitter Added Sampling (JAS). En el AAS la falta de periodicidad se logra sumando muestra a muestra un lapso variable. De esta forma se obtiene un sistema asincrónico [111]. En el JAS se adiciona un tiempo variable a un marco sincrónico [117].

5.2.1. Muestreo no periódico con secuencias caóticas

El caos determinista permite generar señales con características estocásticas empleando mapas no lineales simples [118]. Es posible diseñar mapas que generen secuencias con una dada IPDF, y además, que alcancen dicha distribución con una constante de tiempo transitorio de unas pocas iteraciones [119]. Mediante técnicas de Skipping

(uso de mapas iterados) [118] o bien mediante dinámica simbólica (Discretización) [51] se logra obtener generadores que, a pesar de ser pseudo aleatorios, superan los tests estadísticos usuales [79] y pueden utilizarse para simular sistemas aleatorios en múltiples aplicaciones.

En este capítulo se analiza la aplicación de estos generadores pseudo aleatorios "mejorados" para emplearse con éxito en un esquema de filtrado de señales de baja frecuencia con ruido de alta frecuencia, por medio de JAS y filtros FIR.

Se analiza la influencia de la IPDF del mapa elegido así como de la constante de tiempo de evolución (r_{mix}) requerida con que una PDF inicial evoluciona hacia la PDF invariante. En particular se utilizan el mapa logístico (con IPDF no uniforme) y mapas lineales por tramos (Three Way Bernoulli, Three Way Tailed Shift, Four Way Tailed Shift, etc) que comparten una IPDF uniforme. Empleando los iterados se logra mantener la IPDF, reduciendo la constante de tiempo r_{mix}, es decir mejorando las propiedades de mixing del mapa [8]. Para medir la calidad del método se utiliza el concepto de Power Frequency Response (PFR) introducido en [116, 107] para muestreo aleatorio.

5.2.2. Fundamentación teórica

Sea $y(t)$ la señal ruidosa que debe adquirirse:

$$y(t) = z(t) + n(t) \tag{5.2.1}$$

donde $z(t)$ es la señal y $n(t)$ es un ruido aditivo dado por:

$$n(t) = \sum_{i=1}^{\infty} A_i cos(\omega_i t + \varphi_i) \tag{5.2.2}$$

.

Las frecuencias ω_i son mayores que la frecuencia de corte ω_c de $z(t)$. Se muestrea la señal en un esquema JAS y se filtran los datos mediante filtros FIR (en la Sección 5.3 se analizan distintos tipos de filtros FIR).

Para cuantificar la calidad de los resultados se emplea el indicador Power Frequency Response [107].

La PFR se define como el cociente entre la potencia media de salida y la potencia media de entrada con una entrada es sinusoidal. La PFR de una señal muestreada en un esquema JAS con muestreo caótico se desarrolla a continuación. Los instantes de muestreo están dados por:

$$t_k = T + (\tau_k - 0{,}5)T_S \tag{5.2.3}$$

donde T es el periodo de muestreo promedio y τ_k, una variable pseudo aleatoria generada por un mapa caótico en el intervalo $[0, 1]$ y una constante T_S que determina el rango de variación de los intervalos de muestreo. El valor de T_S establece la correlación entre la señal de entrada y la de muestreo. Si T_S es cero el período de muestreo es uniforme. En el caso de muestreo aleatorio [116], se supone que la PDF de la variable aleatoria τ_k es uniforme en el intervalo $[0, 1]$. Sin embargo en el caso de mapas caóticos es posible evaluar mapas con otras IPDF.

Las Fig.5.1 y Fig.5.3 muestran dos de los mapas empleados en este trabajo: el mapa logístico Fig.5.1, cuya IPDF tiene la expresión analítica, Ec. (5.2.4), (Fig.5.2) y el mapa Three Way Bernoulli Shift (TWB) (ver Fig.5.3) cuya IPDF es uniforme (ver Fig.5.4). Otros mapas como el Three Way Tailed Shift (TWTS), el Four Way Tailed

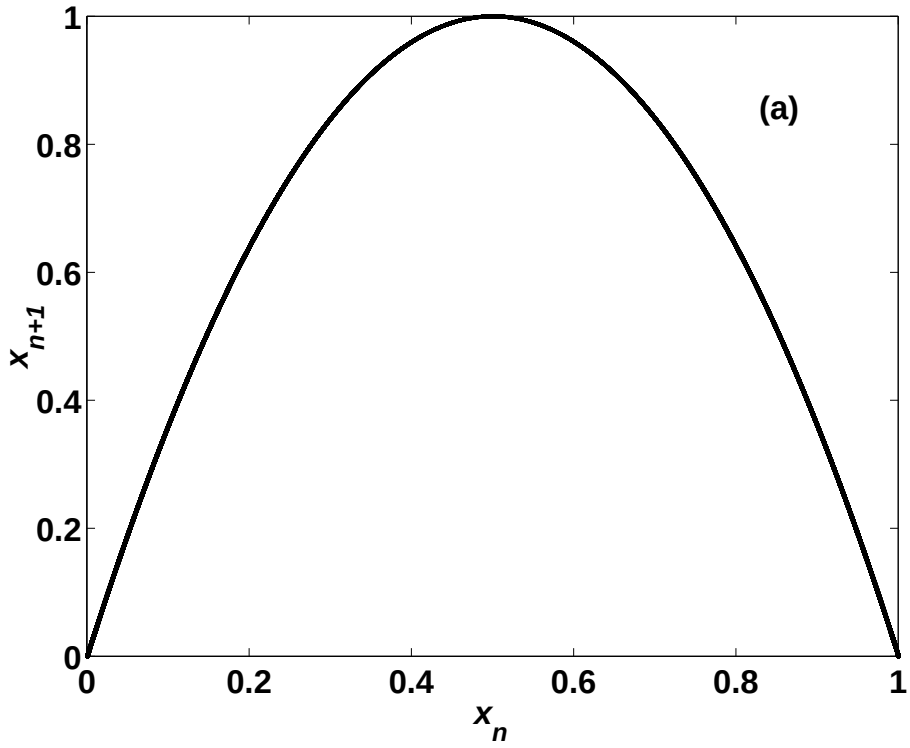

Figura 5.1: Mapa Logístico.

Shift (FWTS) comparten con el TWB la IPDF uniforme.

$$\rho(x) = \begin{cases} \frac{1}{\pi\sqrt{x(1-x)}} & x \,\epsilon\, [0,1] \\ 0 & other \quad x \end{cases} \tag{5.2.4}$$

,

Es importante destacar que la técnica de Skipping propuesta por Mazzini et al [118] consiste en saltear valores en la secuencia del mapa caótico, lo que es equivalente a trabajar con el mapa iterado. Esta técnica reduce la constante de tiempo r_{mix} con la que la PDF del mapa evoluciona hacia la IPDF. En cambio no produce alteración en la PDF invariante.

La técnica de Discretización, empleadas en [51] permite mediante dinámica simbólica modificar en forma conjunta la IPDF y el r_{mix} del mapa.

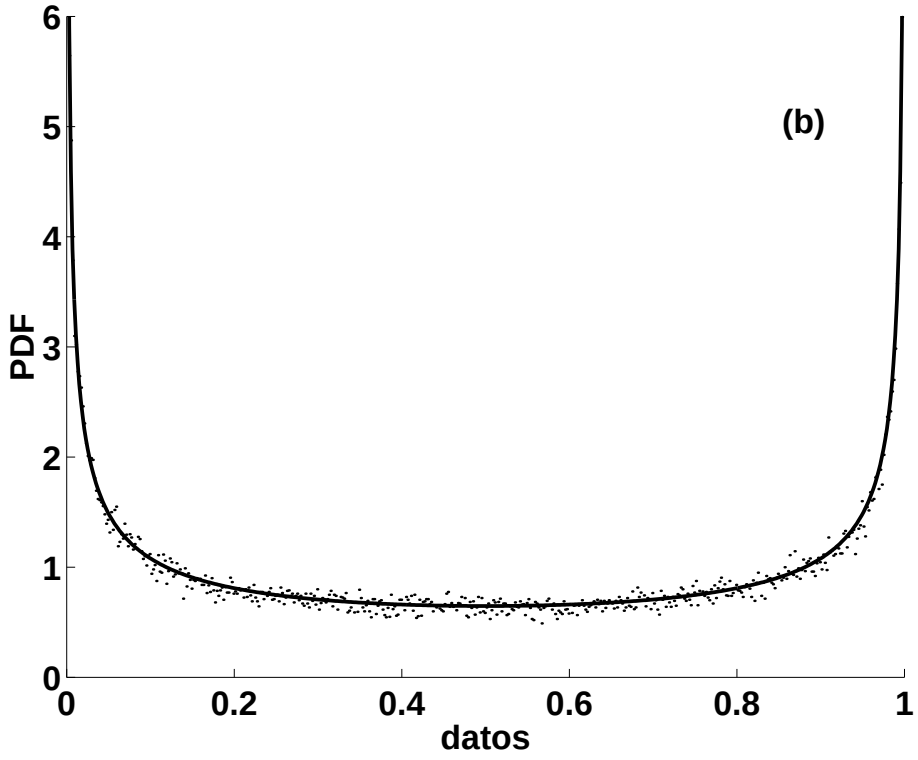

Figura 5.2: IPDF de Mapa Logístico.

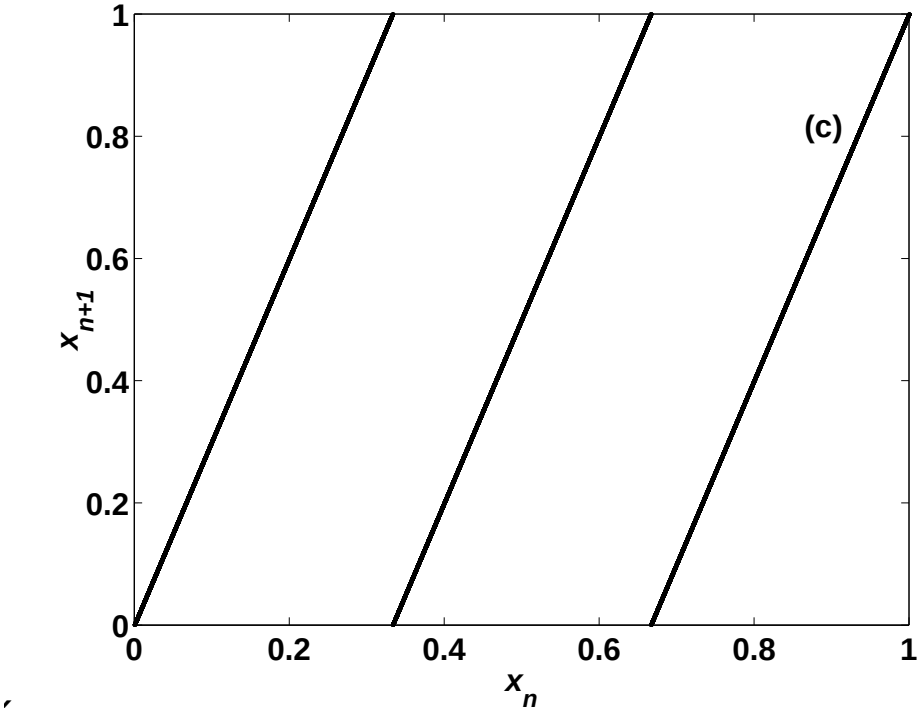

Figura 5.3: Mapa TWB.

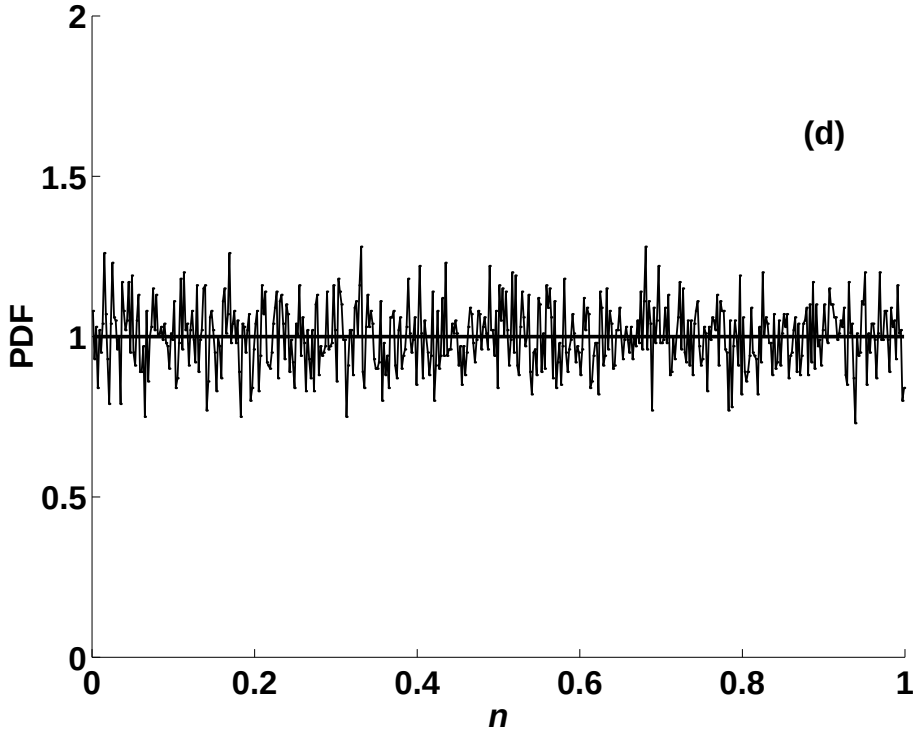

Figura 5.4: IPDF de Mapa TWB.

Sea entonces $x(t)$ una señal senoidal dada por:

$$x(t) = A\cos(\omega t + \phi) \tag{5.2.5}$$

donde ω es la frecuencia angular y ϕ es una variable aleatoria que expresa el corrimiento de fase entre el instante de muestreo inicial y la señal de entrada $x(t)$. Se supone que el proceso de muestreo es independiente de la señal muestreada y que la PDF de la variable aleatoria ϕ es uniforme en el intervalo $[0, 2\pi]$.

Las muestras obtenidas por JAS están dadas por:

$$\widehat{x}_k = A\cos\left[\omega(kT + \tau_k) + \phi\right] \tag{5.2.6}$$

El resultado del filtrado con un filtro FIR es:

$$\mu = \sum_{k=1}^{n} h_{n-k}\widehat{x}_k \tag{5.2.7}$$

donde $h_0 ... h_{n-1}$ son los coeficientes del filtro FIR y n es el orden del filtro.

Suponiendo que el proceso es ergódico y que la IPDF es uniforme (tal como ocurre en los mapas TWB, TWTS y FWTS), la varianza de la salida, que es igual a la potencia media, está dada por

$$E[\mu^2] = \frac{A^2}{2}\|\mathbf{h}\|_2^2 + A^2 sinc^2(fT_S)\sum_{k=1}^{n} h_{n-k}^{\sharp}\cos[\omega\, Tk] \tag{5.2.8}$$

donde $\|\mathbf{h}\|_2^2$ es la norm2 del FIR $\mathbf{h}$ y $\mathbf{h}^{\sharp} = \mathbf{B}\cdot\mathbf{h}$ es un vector obtenido del siguiente modo:

$$\mathbf{h}^{\sharp} = \begin{pmatrix} h_0^{\sharp} \\ h_1^{\sharp} \\ \cdots \\ h_{n-1}^{\sharp} \end{pmatrix}, \quad \mathbf{h} = \begin{pmatrix} h_0 \\ h_1 \\ \cdots \\ h_{n-1} \end{pmatrix} \tag{5.2.9}$$

$$\mathbf{B} = \begin{pmatrix} 0 & 0 & \cdots & \cdots & \cdots & 0 & h_0 \\ 0 & \cdots & \cdots & \cdots & 0 & h_0 & h_1 \\ \cdots & \cdots & \cdots & 0 & h_0 & h_1 & h_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & 0 & h_0 & h_1 & \cdots & h_{n-4} & h_{n-3} \\ 0 & h_0 & h_1 & \cdots & h_{n-4} & h_{n-3} & h_{n-2} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{5.2.10}$$

Finalmente bajo condiciones estacionarias la Ec. (5.2.8) se convierte en:

$$E[\mu^2]_{ss} = \frac{A^2}{2}\|\mathbf{h}\|_2^2 + A^2 sinc^2(fT_S)|\mathbf{H}^\sharp(z)|_{z=e^{(j\omega t)}} \cos[\omega T n + \theta] \tag{5.2.11}$$

donde

$$|\mathbf{H}^\sharp(z)| = \sum_{k=1}^{n} h_k^\sharp z^{-k} \tag{5.2.12}$$

$$\theta = Arg\left[\mathbf{H}^\sharp(e^{j\omega T})\right] \tag{5.2.13}$$

La PFR queda definida por

$$PFR(f) = \frac{E[\mu^2]_{ss}}{A^2/2} = \|\mathbf{h}\|_2^2 + 2\ sinc^2(fT_S) \times|\mathbf{H}^\sharp(z)|_{z=e^{(j\omega t)}} \cos[\omega T n + \theta] \tag{5.2.14}$$

La Ec. (5.2.14) es una expresión general aplicable para un muestreo aleatorio con PDF uniforme, para un muestreo caótico generado por un mapa con IPDF uniforme, y a un muestreo uniforme. Como puede verse el efecto del muestreo caótico aparece sólo en $sinc^2(fT_S)$. Para un muestreo uniforme $T_S = 0$, ese término desaparece

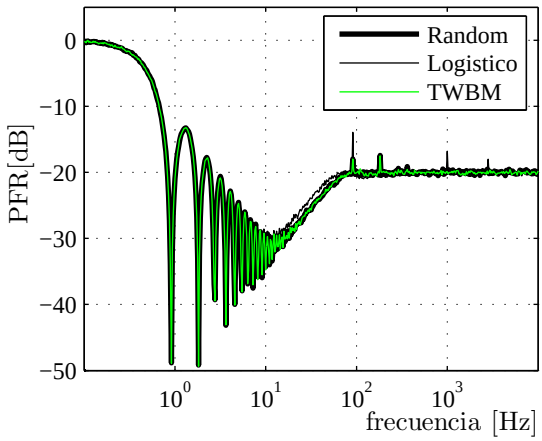

Figura 5.5: Filtro
muestreo caótico

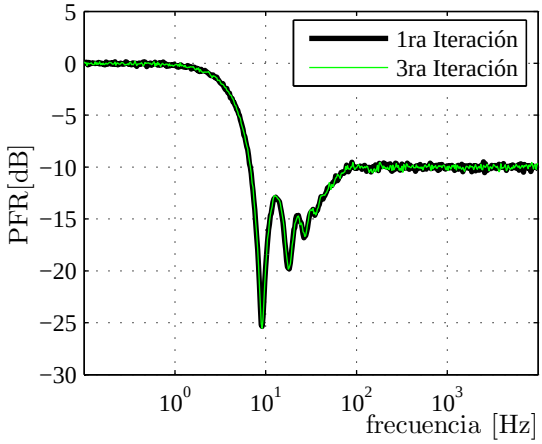

Figura 5.6: Filtro MA con $T_S = 10ms$, $T = 11ms$, $n = 10$ para muestreo caótico con la primera y la tercera iteraciones del mapa TWB

.

5.3. Resultados

Se desarrollaron experimentos estadísticos extensivos para validar la propuesta. Cada experimento consistió en obtener un juego de salidas del FIR μ_i con $i \in \{1, 2 \cdots m\}$. La entrada al filtro fue una señal senoidal de frecuencia constante y fase aleatoria la cual fue muestreada utilizando la Ec (5.2.3). Cada salida μ_i se obtuvo utilizando la Ec (5.2.7). La varianza de μ en un juego de $m = 1000$ experimentos fue calculada por medio de la Ec. 5.3.1.

$$\Gamma_\mu^2 = \frac{1}{m} \sum_{i=1}^{m} (\mu_i - \overline{\mu})^2 \tag{5.3.1}$$

donde $\overline{\mu}$ es el valor medio de μ_i. Los resultados típicos mostrados en las Figs. 5.5-5.7 se obtuvieron barriendo en frecuencia.

En la Fig. 5.5, por ejemplo, se comparan los resultados para un muestreo aleatorio empleando la función *rand* de MATLAB®, con los correspondientes a dos mapas caóticos: el logístico y el TWB. Los resultados para el mapa TWBM cuya IPDF es uniforme son similares al caso de la función *rand* de MATLAB® y concuerdan con la Ec. (5.2.14).

En Fig. 5.6 se utiliza un FIR Moving Average (MA) y se emplean la primera y la tercera iteraciones de un mapa TWB. La constante de tiempo de evolución de este mapa fue determinada en forma teórica como función de la iteración ($r_{mix} = (1/3)^j$ donde j es el orden de iteración). Los resultados muestran que en tanto el número de elementos del filtro sea suficiente ($n = 10$ en la Fig. 5.6), el efecto de r_{mix} no es relevante.

En la Fig. 5.7 se comparan resultados para tres FIR diferentes empleando el mapa FWTS. El mejor resultado se obtiene con la función fir1 de MATLAB®.

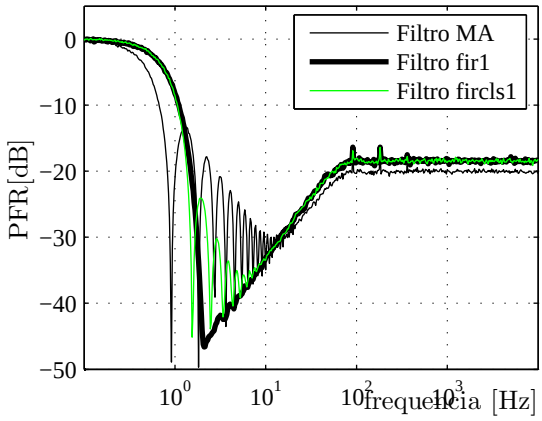

Figura 5.7: Tres diferentes FIR: filtro MA, filtro utilizando el método de ventana (función fir1 de MATLAB®) y filtro utilizando método de mínimos cuadrados (función fircls1 de MATLAB®). Muestreo caótico con mapa FWTS

.

5.4. Conclusiones

las simulaciones realizadas muestran que los resultados con muestreo caótico son equivalentes a los obtenidos mediante muestreo aleatorio, si se emplean mapas con IPDF uniforme y buenas propiedades de mixing ($r_{mix} \to 0$). Sin embargo para observar una influencia del valor de r_{mix} debieron emplearse filtros FIR excesivamente cortos, que en la práctica son inadecuados debido a que el número de elementos del FIR también afecta la atenuación en alta frecuencia.

Luego puede concluirse que, a los fines prácticos, r_{mix} no es un factor relevante. Es interesante notar que las técnicas de aleatorización empleando dinámica simbólica permiten simultáneamente mejorar las propiedades de mixing de un mapa y volver uniforme su IPDF, a la vez que la implementación puede hacerse en aritmética entera [120]. Es posible entonces extender los resultados a un conjunto muy amplio de mapas caóticos.

Capítulo 6

Generador en FPGA de Ruido Estocástico Coloreado

En ingeniería es de gran interés analizar el comportamiento de sistemas y circuitos electrónicos en presencia de ruido no deseado, ya que, a excepción de temperatura cero, el ruido esta presente en todo circuito electrónico [121]. Para esto, es necesario disponer de señales ruidosas de testeo cuyas propiedades estadísticas sean conocidas con exactitud.

Es necesario entonces contar con generadores físicos de ruido con características controlables (espectro, función densidad de probabilidad, autocorrelación, etc.). La generación mediante programación en software, con la utilización de aritmética de punto flotante, es relativamente sencilla. Ha habido gran desarrollo de generadores de este tipo por software, pero su contraparte en hardware está mucho menos desarrollada.

En este capítulo se propone la implementación en hardware de un generador de ruido estocástico coloreado, con espectro de potencia de tipo f^{-d}. La implementación se ejemplifica para los casos $d = 1, 2$. La compilación preliminar permite estimar que la arquitectura propuesta se puede implementar en una FPGA $EP2S15F484C3$ de

la familia Stratix II de *Altera*©.

6.1. Fundamentación Teórica

El efecto del ruido sobre los sistemas físicos es un tema de investigación que ha ganado importancia recientemente a partir de la posibilidad de generar ruidos con características específicas tanto mediante su simulación por computadora (implementación en software) como en hardware; en especial mediante el uso de dispositivos FPGA (Field Programmable Gate Array).

Todo sistema electrónico es afectado por muchas fuentes de ruido. Estas pueden ser clasificadas de la siguiente forma:

- Transmitido: Recibido con, y superpuesto a la señal de interés.
- Intrínseco: Proveniente de los dispositivos electrónicos del mismo sistema.
- Interferente: Externo al sistema, ingresa al mismo por entradas no designadas.

En este caso nos interesa la generación de ruido intrínseco. Este ruido es clasificado según cómo varía su intensidad en frecuencia y es conocido como ruido coloreado. Dos patrones muy conocidos son el el ruido blanco, que presenta un espectro de potencia constante, y el ruido rosa, cuyo espectro de potencias es alto en frecuencias bajas y más débil en las altas.

En general, estos ruidos presentan un espectro de potencias de la forma f^{-d}, para el ruido blanco $d = 0$, ruido rosa $d = 1$, ruido rojo $d = 2$, entre otros. La principal contribución en este capítulo es el diseño, con vistas a la implementación en hardware, de un generador con espectro de la forma f^{-d} partiendo de una señal de ruido blanco

como entrada, que puede generarse en una FPGA a partir de un sistema caótico discreto [122].

6.2. Implementación por Software

El ruido de tipo f^{-d} se genera por software en forma relativamente simple utilizando las funciones de librería de Matlab y el procedimiento que se indica a continuación:

1. Mediante la función rand de Matlab generar números pseudoaleatorios, $y_{(n)}$, en el intervalo (-0.5, 0.5), con espectro de potencia plano, función densidad de probabilidad uniforme y valor medio cero.

2. Calcular la FFT de la secuencia generada obteniéndose el vector complejo $Y_{[k]}$.

3. Multiplicar el vector complejo $Y_{[k]}$ por la secuencia $k^{-d/2}$ obteniéndose $W_{[k]}$.

4. Luego, simetrizar $W_{[k]}$, de forma de que cumpla $W_{[-k]} = conj(W_{[k]})$ y represente una función real del tiempo, esto es, $Re(W_{[-k]}) = Re(W_{[k]})$ y $Im(W_{[-k]}) = Im(-W_{[k]})$.

5. Finalmente antitransformar la secuencia $W_{[k]}$ y descartar las pequeñas componentes imaginarias que surgen del error cometido por la utilización de precisión finita, se obtiene la serie temporal, $w_{(n)}$ cuyo espectro de potencia tiene la forma k^{-d}. Este método constructivo produce ruidos estocásticos.

La metodología utilizada para la implementación en hardware sigue los mismos lineamientos expuestos como se describe en la sección siguiente.

6.3. Implementación por Hardware

Se decidió realizar la implementación en una plaqueta FPGA de *Altera*© [123]. Estas plaquetas son programables a nivel hardware, así, proporcionan las ventajas de un procesador de propósito general y un circuito especializado que se puede reconfigurar las veces que sea necesario para depurar su funcionalidad. Las FPGAs son más lentas y tienen un mayor consumo de potencia que los ASICs, sin embargo tienen la ventaja de ser re-programables, sus costes de desarrollo y adquisición son mucho menores para pequeñas cantidades de dispositivos y su ciclo de diseño es más corto.

Para realizar la programación del dispositivo se utilizó un entorno visual para flujos de datos. En los últimos años han aparecido dos tendencias en entornos de diseño a nivel de sistema, estos son, lenguajes de alto nivel (como SystemC, VHDL y Verilog) que son lenguajes de descripción de hardware; y entornos visuales para flujo de datos.

Los entornos visuales para flujos de datos son adecuados para modelar sistemas que se representan de forma más fácil a través de diagramas de flujos.

Estos entornos visuales son similares a las tradicionales herramientas de esquemáticas, disponen de librerías de bloques funcionales que permiten componer gráficamente el modelo de un sistema.

Mediante la utilización de las librerías de bloques y los entornos de simulación en un entorno visual para flujos de datos se consigue un alto nivel de abstracción funcional, permitiéndose la utilización de diferentes tipos de datos y operadores para modelar aritmética trabajando con variables enteras, de punto fijo y flotantes.

Dos ejemplos de entornos visuales para flujo de datos son las herramientas System

Generator de *Xilinx*© y DSP Builder de *Altera*© junto con Matlab/Simulink de *Mathwork*.

En este caso se utilizó la herramienta de *Altera*©.

6.3.1. Simulink de Matlab

Simulink es una herramienta gráfica que sirve para modelar, simular y analizar sistemas dinámicos. Soporta sistemas lineales y no lineales, modelados en tiempo continuo, discreto o un híbrido de los dos. Al igual que la herramienta Matlab se apoya en un conjunto de librerías algunas de las cuáles son de otros fabricantes. Para realizar la síntesis del hardware es necesario disponer de los entornos o toolboxes System Generator (de Xilinx) o Dsp Builder (de *Altera*).

La principal ventaja de programar en el entorno Simulink de Matlab es la simulación del sistema utilizando todas las herramientas disponibles tales como generadores de señales e instrumentos virtuales de medición que permiten tener acceso a cada etapa del circuito.

Matlab además de tener toolboxes propias, integra toolboxes de otros fabricantes (third party) que sirven para desarrollar aplicaciones específicas. En el ámbito de la lógica programable tiene acuerdos con los fabricantes más importantes, Xilinx y *Altera*.

6.3.2. DSP Builder

DSP Builder de *Altera* permite realizar diseños de algoritmos en Matlab, la integración del sistema en Simulink y llevar el diseño a un lenguaje de descripción hardware (HDL) para sintetizar el código HW a FPGAs mediante Quartus II, este es

una herramienta de software producido por *Altera* para el análisis y la síntesis de los diseños en HDL.

DSP Builder genera automáticamente un diseño a nivel de transferencias de registros y un archivo de estímulos para Simulink. Estos archivos están optimizados para ser utilizados con Quartus II para verificar los resultados de síntesis.

Dentro de una FPGA se puede incluir la funcionalidad de varios circuitos integrados. Esta funcionalidad puede ser desarrollada o adquirida a través de un tercero. Debido a que estas funcionalidades son como componentes electrónicos, pero sin su parte física, se los suele llamar componentes virtuales. En la industria se los conoce como bloques de propiedad intelectual o cores IP.

La implementación es más eficiente a través del uso de estos cores IP, ya que proporcionan un rango de funcionalidad que va desde operaciones aritméticas básicas hasta algoritmos complejos para DSP.

En nuestro sistema particular se utilizó uno de estos cores IP para realizar uno de los bloques más importantes, el encargado de realizar la transformada (y antitransformada) de Fourier de la secuencia de entrada. Esto fue implementado mediante el core IP *MegaCore FFT*.

Transformada discreta de Fourier

La Transformada Discreta de Fourier (DFT), calcula N muestras de la transformada de Fourier de una secuencia $x(n)$ de N puntos, de acuerdo con:

$$X[k] = \sum_{n=1}^{N} x_{(n)} e^{(-j2\pi nk/N)} \tag{6.3.1}$$

con $k = 0, 1, ..., N-1$.

En la ecuación (6.3.2) se muestra la DFT inversa de longitud N, donde se incluye el factor de escala $1/N$, de acuerdo a la convención adoptada por Matlab.

$$x(n) = (1/N) \sum_{k=0}^{N-1} X_{[k]} e^{(-j2\pi nk/N)} \tag{6.3.2}$$

con $n = 0, 1, ..., N-1$.

La complejidad computacional de la DFT puede reducirse significativamente usando algoritmos rápidos que realizan una descomposición anidada de las sumatorias de las ecuaciones (6.3.1) y (6.3.2), además explotan varias simetrías inherentes a los exponentes. Uno de estos algoritmos es el llamado Cooley-Tukey base-r con Decimación en Frecuencia (DIF). Este algoritmo divide en forma recursiva la secuencia de entrada en N/r secuencias de longitud r, hasta obtener transformadas de r puntos. Estas operaciones básicas son llamadas mariposas. Normalmente el valor de r es un número de base 2,4 ó 16 (base-2, base-4 ó base-16 respectivamente). De esta forma, se logra reducir las etapas de cálculo a $log_r(N)$. Estos algoritmos se conocen con el nombre genérico FFT.

Como ya se dijo, para realizar la operación de FFT se utilizó la función MegaCore FFT de *Altera* que es un core IP optimizado para trabajar con los dispositivos de *Altera*.

6.3.3. MegaCore FFT

La utilización de cores IP reduce el tiempo de diseño y desarrollo del proyecto ya que permiten "saltear" el proceso de diseño de funciones estandarizadas. Así, es posible realizar una programación en un nivel más alto.

Este core IP es un procesador altamente parametrizable capaz de implementar la FFT compleja y la FFT inversa (IFFT) para aplicaciones de alta performance, realiza el algoritmo FFT empleando DIF base 2 ó 4. El valor de N debe corresponder a la m-ésima potencia de 2 ($N = 2^m$) con $6 \leq m \leq 14$.

Cada etapa de la descomposición comparte el mismo hardware, los datos son leídos de una memoria, atraviesan la mariposa y son escritos nuevamente en memoria. El aumento del valor de la base disminuye la cantidad requerida de "pasadas" por la mariposa, a expensas de recursos del dispositivo y velocidad de cálculo.

Este core permite elegir que el tamaño total de la secuencia a transformar, N, sea fijo (streaming fijo) o modificado en tiempo de ejecución (streaming variable). En este trabajo se optó por streaming fijo con $N = 2048$. La entrada a este bloque son vectores de datos complejos de longitud N en formato de complemento a dos. La función genera un vector complejo en el dominio transformado. El tipo de transformada (directa o inversa) se especifica por medio de un puerto de entrada.

Si consideramos una relación señal a ruido (SNR) tomando como ruido al error generado por la utilización de precisión finita, MegaCore maximiza esta relación representando los datos en Punto Flotante en Bloque (BFP) que es una combinación de punto fijo y punto flotante.

En una arquitectura de punto fijo, la precisión de los datos debe ser lo suficientemente grande para poder representar adecuadamente todos los valores intermedios de los cálculos del algoritmo. Para tamaños grandes de transformaciones de FFT, la implementación de punto fijo puede provocar un crecimiento excesivo de los recursos o demasiada pérdida de precisión.

En una arquitectura de punto flotante cada número es representado por una mantisa y un exponente individual - esto produce una gran mejora en la precisión - sin embargo las operaciones de punto flotante consumen muchos recursos del dispositivo.

En una arquitectura BFP todos los valores tienen una mantisa independiente pero comparten un mismo exponente por cada bloque de N datos. Esto es, los datos son "escalados", es decir divididos por un factor común, esta división se realiza con l desplazamientos hacia la derecha de los bits, lo que genera una división por 2^l.

La escala se ajusta de acuerdo a la medida del rango dinámico del bloque de los N datos. La cantidad de desplazamientos es acumulada en l y luego proveída en una salida como un valor de exponente válido para los N valores de salida. Este desplazamiento asegura que la cantidad de bits menos significativos (LSBs) descartada sea la mínima. La representación de punto flotante en bloque actúa en efecto, como un control automático de ganancia digital. Para obtener el valor final de la señal transformada, tanto la salida real como la salida imaginaria, deben afectarse por el exponente de ajuste de escala según la ecuación (6.3.3).

$$X[k] = sal_{real}[k],2^{sal_{exp}} + j.sal_{imaj}[k],2^{sal_{exp}} \tag{6.3.3}$$

donde $sal_{real}[k]$ y $sal_{imaj}[k]$ son las salidas real e imaginaria respectivamente, y sal_{exp} es el valor del exponente para ese bloque de N datos, que corresponde a los l desplazamientos efectuados.

Análisis del funcionamiento del bloque FFT MegaCore

Para verificar el correcto funcionamiento del bloque se realizaron simulaciones con diferentes señales de entrada cuyos espectros de potencia son conocidos. En la

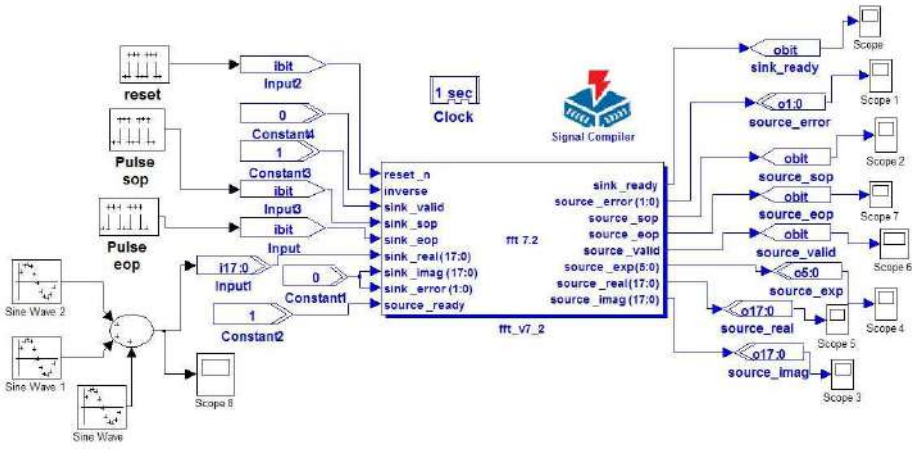

Figura 6.1: Banco de prueba.

Fig. 6.1 se muestra el esquema en Simulink del banco de prueba utilizado. En este gráfico puede verse el bloque de la FFT MegaCore al cual se le conectó a la entrada tres fuentes senoidales sumadas de distintas frecuencias. Las salidas son analizadas mediante los scope proveídos por Simulink, mediante los cuales se almacenaron los datos de salida para su posterior procesamiento con Matlab.

Se presentan algunos resultados obtenidos. La frecuencia de muestreo utilizada es $T_{samp} = 1seg.$ y la longitud de transformada es $N = 2048$. Con estos parámetros la resolución en frecuencia que se obtiene en la FFT es $f_{step} = (T_{samp}.N)^{-1} = 1/2048 = 0,00048828125Hz$, por lo tanto el máximo error posible de frecuencia es $f_{step}/2$. El rango de frecuencias representadas va desde $0,00048828125Hz a f_{samp}/2 = 0,5Hz$.

La señal de entrada es una sumatoria de tres senos con frecuencias $0,0078125Hz$; $0,00390625Hz$ y $0,001953125Hz$. En la Fig. 6.2 se representa la señal temporal resultante.

En la Fig.6.3 se observa el espectro obtenido mediante el bloque FFT MegaCore, podemos ver que las componentes espectrales (bins) no tienen error de amplitud y

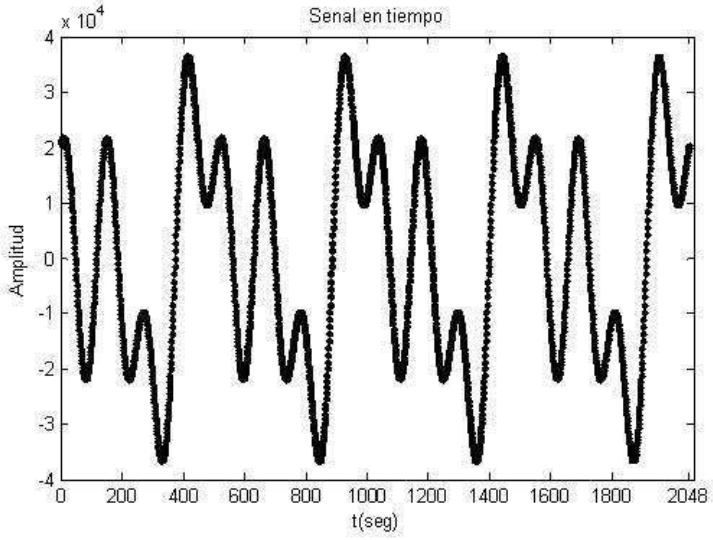

Figura 6.2: Señal de entrada a transformar .

son solo las tres que corresponden a cada una de las sinusoides de entrada.

A continuación se presenta un caso más real en el cual la señal a transformar no entra un número entero de veces en la ventana N. Nuevamente la señal de prueba es la suma de tres sinusoides cuyas frecuencias son $0,01Hz$; $0,00666Hz$ y $0,005Hz$. En la Fig. 6.4 se presenta la señal en el tiempo y la Fig. 6.5 el espectro de potencia correspondiente.

Podemos ver que aparecen más bins que los esperados. Además, las amplitudes están afectadas por una envolvente *sinc* ya que se utilizó una ventana rectangular ($N = 2048$) y no entraron períodos enteros de las sinusoides en esta ventana. Con la frecuencia de muestreo utilizada en este caso ($1Hz$), por ejemplo, de la primera sinusoide de frecuencia $0,01Hz$ entraron 100 muestras por período por lo tanto en 2048 puntos entran $2048/100 = 20,48$ períodos. Por lo que el algoritmo de la FFT toma como señal a 20 períodos y un poco más.

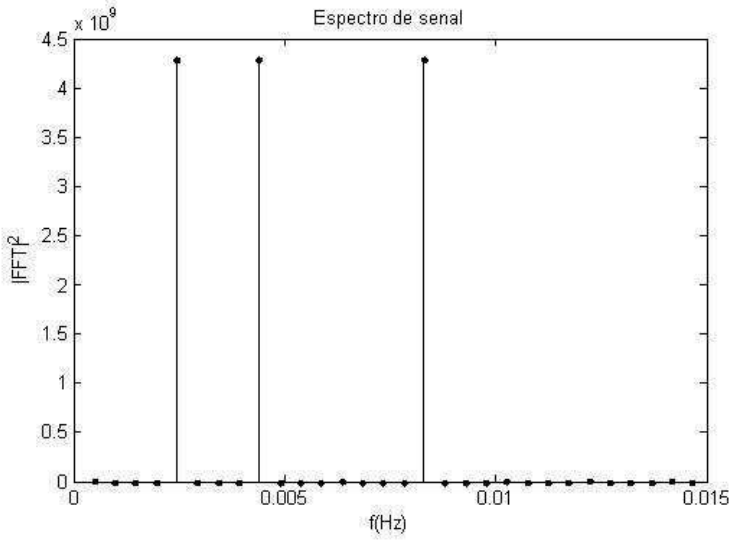

Figura 6.3: Espectro .

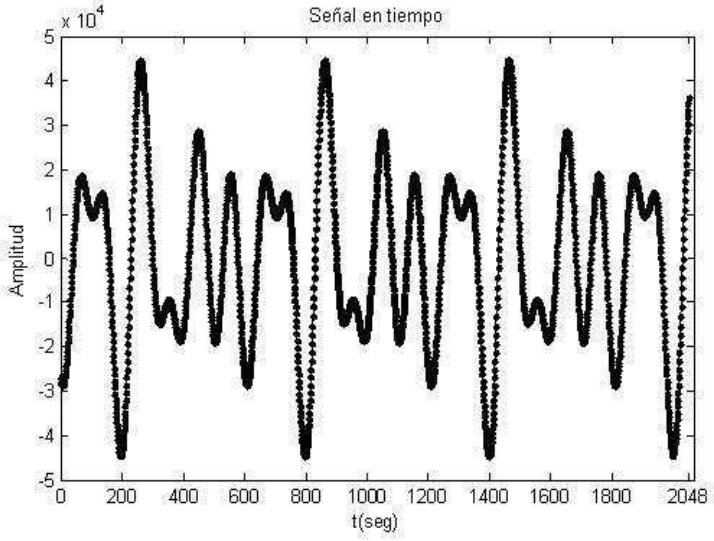

Figura 6.4: Señal de entrada a transformar.

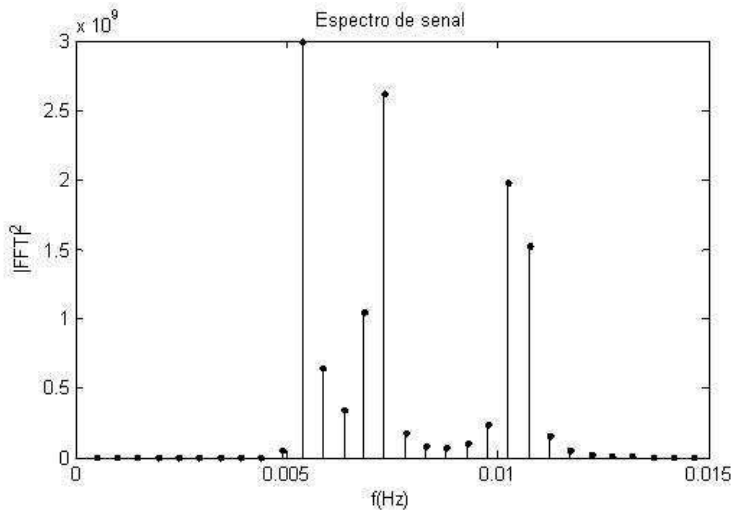

Figura 6.5: Espectro con error.

En el caso real de una señal muestreada, si se tiene un valor fijo de N se elige una frecuencia de muestreo de forma tal que entre una cantidad exacta de períodos de la señal. Si lo que se tiene fija es la frecuencia de muestreo, se debe encontrar el N que mejor se ajuste. En el caso en el que no exista un N potencia de dos que cumpla con los requisitos antes mencionados se utilizará el mayor N posible, a expensas de un incremento de tiempo de cálculo y recursos. Cuando no es posible solucionar de esta forma se recurre a utilización de ventanas ya no rectangulares tales como Hamming, Hanning, etc., con el objetivo de suavizar las transiciones. Otra opción es la ventana auto ajustable, la cual ajusta la cantidad de muestras automáticamente por medio de un algoritmo [124].

También se utilizó como señal de prueba una onda cuadrada con período 128 segundos, que por ser potencia de dos asegura no presentar error en amplitud. Fig.6.6.

Puede verse en el espectro de potencia en la Fig. 6.7 las deltas ubicadas en las

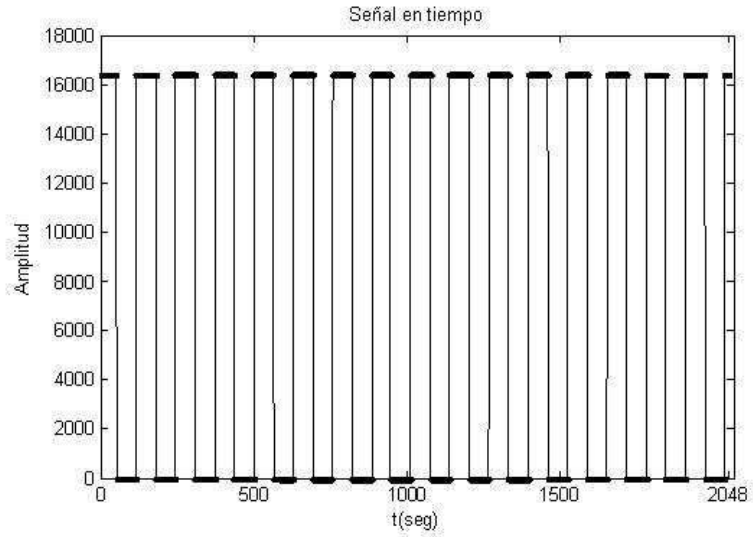

Figura 6.6: Señal cuadrada.

armónicas impares de la frecuencia fundamental de la señal, y la correspondiente al valor de continua ya que la señal tiene valor medio distinto de cero.

Arquitectura del diseño

La arquitectura adoptada consta de cuatro etapas básicas. La primera etapa (ver Fig. 6.8) realiza la FFT de un ruido blanco uniforme con media cero. Existen dos opciones para ingresar el ruido blanco, la primera es utilizar un generador externo a la placa (como se hizo en este caso), y la segunda es generar el ruido blanco en la misma FPGA empleando secuencias caóticas randomizadas. [122, 120].

El segundo bloque (ver Fig. 6.9) se encarga de "colorear" el espectro recibido, es decir convertir el espectro del ruido blanco en uno de ruido de tipo f^{-d}, con $d = 1, 2, 3, 4$. Desde el punto de vista de las aplicaciones no es de interés utilizar valores mayores de d pues el espectro se vuelve demasiado concentrado en las bajas

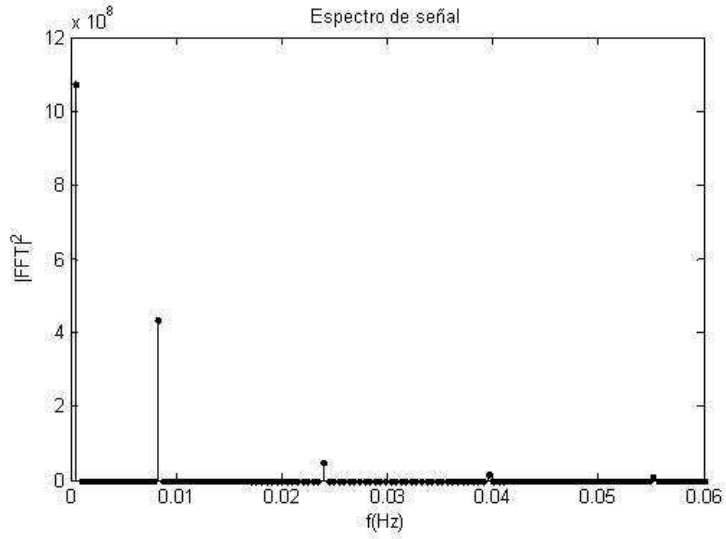

Figura 6.7: Espectro de señal cuadrada.

frecuencias.

En la tercera etapa (ver Fig. 6.10) se realiza la IFFT de la señal coloreada y finalmente la cuarta etapa prepara los datos para la salida, Fig. 6.11.

En todo el diseño se utilizó aritmética de punto fijo, con resolución de 18 bits, menos el bloque FFT MegaCore que utiliza BFP internamente.

Para ilustrar el funcionamiento del sistema las Figs. 6.8 a 6.11 muestran las etapas del generador. El dispositivo propuesto permite la selección de d mediante una entrada externa.

Primeramente se ingresa al bloque FFT, por el bus *sink_real*, con una señal de ruido blanco, generada por ejemplo, con la fuente Uniform Random Number de Simulink que presenta un espectro de potencia plano. En este caso, los niveles utilizados en la entrada *sink_real* van desde -2^{12} a 2^{12}, y son digitalizados a 18 bits. Por su parte la entrada *sink_imag* es seteada a cero (es decir que la señal de entrada es real pura).

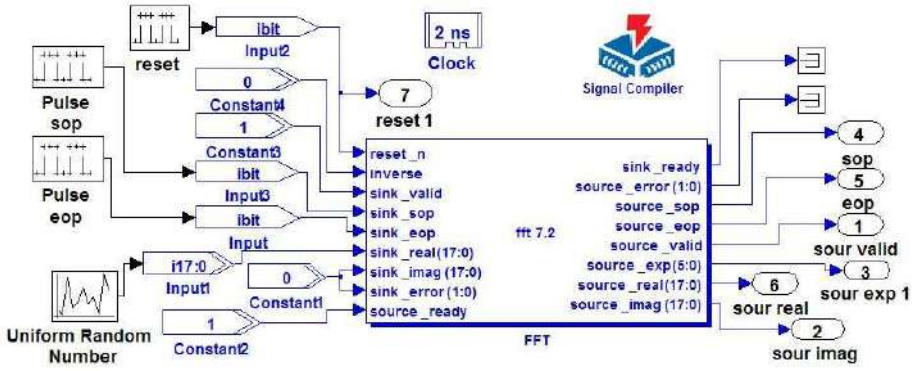

Figura 6.8: Bloque que realiza la FFT.

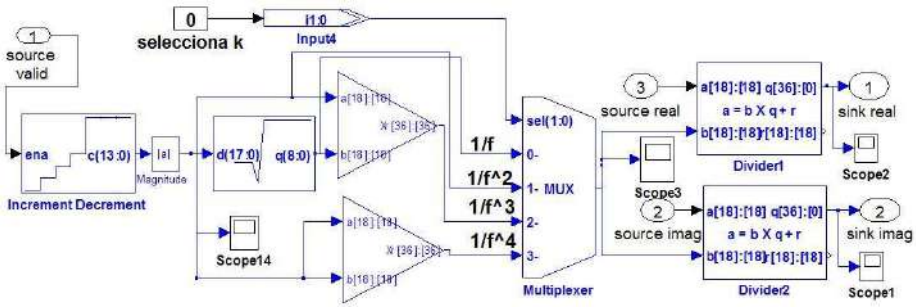

Figura 6.9: Bloque generador de vector f^k, divisor.

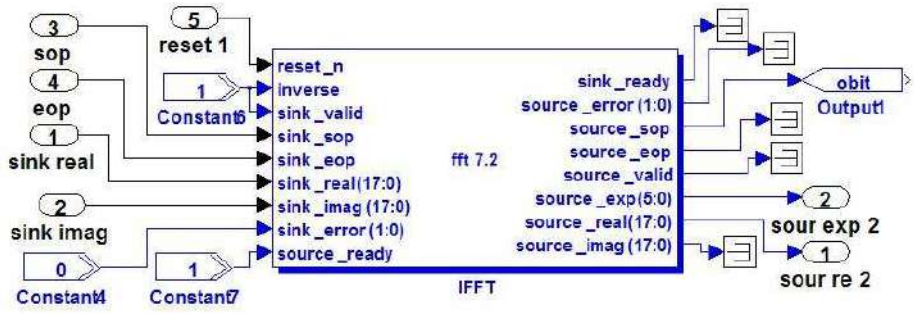

Figura 6.10: Bloque que realiza la IFFT.

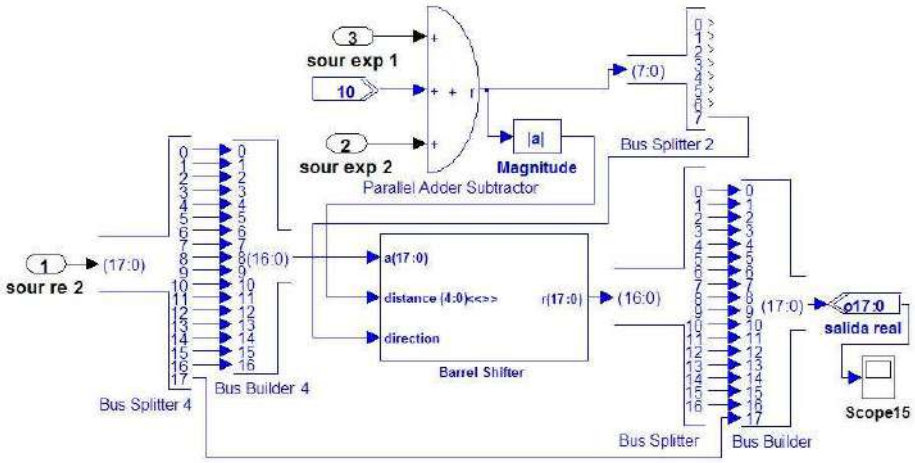

Figura 6.11: Bloque que realiza el Ajuste de la Ganancia a la salida total.

El bloque FFT fue previamente parametrizado del siguiente modo:

1. longitud de transformada $N = 2^{14}$,

2. 18 bits de resolución,

3. dispositivo perteneciente a la familia *StratixII* de *Altera*.

El resto de los parámetros se mantuvieron con los valores que vienen por defecto. Las partes real e imaginaria de la FFT son entregadas por las salidas *source_real* y *source_imag* del bloque FFT. Ambas salidas se dividen por el vector $f^{1/2}$ (en el caso general se divide por $f^{1/2 \cdot d}$).

Cuando la salida *source_valid* es igual a uno, el bloque fft_v7_2 indica que hay salida válida en los buses *source_real* y *source_imag*, en ese instante comienza a incrementarse un contador para generar un vector de frecuencia a pasos f_{paso} dados por la expresión 6.3.4, donde T_{ck} es el período de reloj. A excepción del caso $d = 2$,

la salida del contador ingresa a un bloque que realiza la operación de raíz cuadrada (ver Fig.6.9). Luego, la salida de ese bloque pasa por un sistema compuesto de multiplicadores para generar el vector $f^{d/2}$.

$$f_{paso} = 1/(T_{ck},2^{14}) \tag{6.3.4}$$

La entrada d que selecciona el ruido corresponde a la entrada de un multiplexor. El vector $f^{d/2}$ divide al espectro complejo ($source\_real$ y $source\_imag$) generando así a la salida de los divisores las componentes espectrales real e imaginaria del ruido elegido.

Una vez procesados los datos transformados, ingresan a un segundo bloque $MegaCore$, fft_v7_2 esta vez se establece que realice la transformada inversa de Fourier mediante el seteo de la entrada $inverse$.

Finalmente, un último bloque ajusta la ganancia total del proceso. En primer lugar se debe convertir nuevamente a punto fijo tomando en cuenta el valor de las salidas $source\_exp$ de los dos bloques $MegaCores$, fft_v7_2 (el que realiza la FFT y el que realiza la IFFT). Además, la función $MegaCore$ no aplica internamente el factor de escala $1/N = 2^{-n}$ requerido por la IFFT. Por lo tanto se lo aplica externamente (en este caso $n = 11$).

6.3.4. Resultados

En las Figs.6.12 y 6.13 se muestran los ruidos $1/f$ y $1/f^2$ generados mediante la arquitectura descrita. El valor de m indicado en esas figuras corresponde a la pendiente estimada de los puntos ajustados a una recta mediante la función $polyfit$ de Matlab. También se indica el período de muestreo (T_{sam}) y el valor de d elegido.

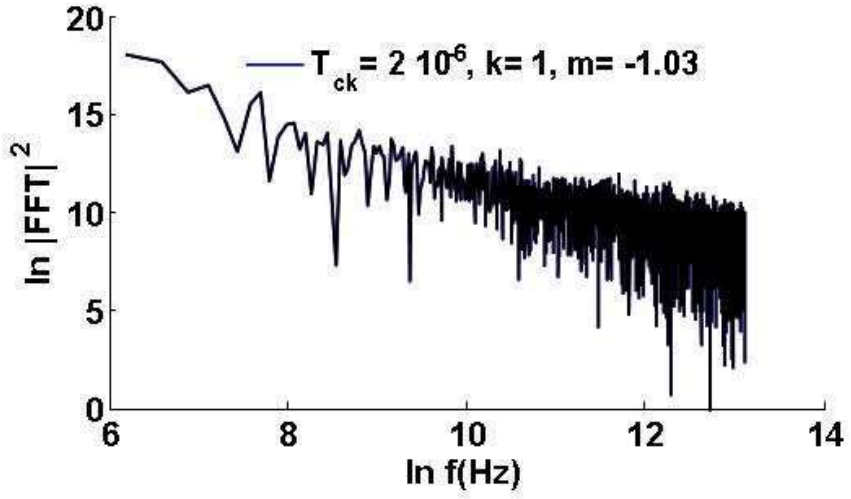

Figura 6.12: .

Puede verse que el espectro obtenido presenta la forma de ruido coloreado esperado en cada caso.

6.3.5. Conclusiones

Se diseñó, un generador de ruido estocástico coloreado, del tipo f^{-d}, con vistas a su implementación en una placa FPGA. El sistema permite seleccionar el tipo de ruido deseado dentro del rango $d = 1$ a 4. Para realizar el algoritmo de la DFT se utilizó de la librería $Altera\ IP - MegaCores$ el bloque $FFTMegaCore$, el cual se analizó con detalle.

En esta primera etapa no fue importante la optimización de recursos si no el definir una metodología general de diseño que fuera aplicable no sólo al ruido f^{-d} sino también a otros ruidos estocásticos de interés, tales como los gaussianos fraccionarios. No obstante, en una segunda etapa debe realizarse la optimización de recursos utilizados

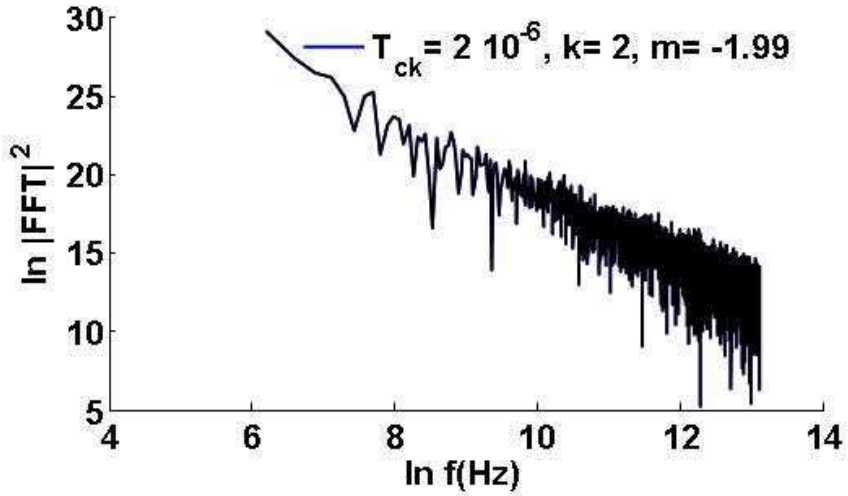

Figura 6.13: .

en cada caso [123, 125, 126].

La principal conclusión en esta etapa de diseño es la ventaja de la utilización del entorno Simulink de Matlab para el diseño y simulación, dado que permite en forma relativamente simple una futura ampliación. Por otra parte la misma metodología se prevé la implementación de otros tipos de ruido estocástico.

Capítulo 7

Conclusiones y líneas de trabajo futuro

Las principales conclusiones de la tesis son:

1. Conclusiones generales sobre secuencias caóticas pseudo aleatorias

   - Se demostró que es posible reemplazar las señales pseudo aleatorias empleadas en los sistemas de Espectro Esparcido por señales generadas por sistemas caóticos sencillos, randomizadas de modo adecuado.
   - Se ha además encontrado la manera de cuantificar la calidad de las series generadas por distintos mapas caóticos así como la calidad de los métodos de randomización utilizados para mejorarlas.
   - Se concluyó que para la evaluación de calidad de las secuencias son necesarios dos tipos de cuantificadores: *a) cuantificadores dependientes de* r_{mix} *únicamente (pero no de la* $IPDF$*)*, como $C^{(BP)}$, $H^{(BP)}$, DET, L and $ENTR$ y *b) cuantificadores dependientes de la* $IPDF$ *únicamente (pero no de* r_{mix}*)*, como $H^{(hist)}$, $C^{(hist)}$, y RR.
   - Se analizó la utilización de cuantificadores de computación intrínseca, se

comprobó que estos dependen de ambos $IPDF$ y r_{mix} y por lo tanto no son convenientes para el análisis de secuencias con nuestra metodología.

- Los planos de representación adecuados deben tener un cuantificador de cada clase. Estos planos de representación permiten comparar secuencias entre sí y además elegir el método de randomización más adecuado.
- También se verificó que la representación en el plano C^{BP}-H^{BP} no es adecuada pues utiliza una única distribución de probabilidades en ambos ejes (la distribución obtenida mediante la prescripción de Bandt y Pompe). En ese plano tanto las series sin randomizar como las randomizadas se encuentran muy cercanas al punto $C^{BP} = 0$, $H^{BP} = 1$.

2. En cuanto a la aplicación de secuencias caóticas en CDMA:

   - Se definieron tres cuantificadores. Estos son el cuantificador de correlación C, el cuantificador de spreading S y el cuantificador de zipping Z. Estos cuantificadores son globales en el sentido de que existe un valor para cada familia PN.
   - El cuantificador de zipping aparece como un posible cuantificador para comparar distintas familias, ya que Z sera mayor en cadenas consistentes de códigos PN correlacionados y disminuye si el espectro no es plano.
   - Se estudiaron las familias clásicas y caóticas con los cuantificadores propuestos.
   - Los resultados presentados muestran que las familias caóticas pueden tener una mayor cantidad de miembros que las familias convencionales, con

propiedades de spreading y correlación equivalentes. Además son fácilmente implementados y consecuentemente son buenos candidatos para ser utilizados en sistemas reales.

3. En cuanto a la reducción del EMI mediante secuencias caóticas:

   - Se moduló en FM una portadora mediante secuencias generadas por distintos mapas caóticos y secuencias randomizadas mediante el método de Discretización y Skipping.
   - Se comprobó claramente que el espectro obtenido modulando con las respectivas iteraciones del mapa caótico es idéntico a la $IPDF$ del mismo, como predicen los teoremas de Calegari et al.
   - Cuando la $IPDF$ no cumple con las condiciones requeridas, el espectro obtenido tampoco. Con el método de randomización Discretización es posible obtener un espectro plano muy similar al obtenido con la secuencia pseudo aleatoria.
   - Es posible utilizar mapas caóticos con $IPDF$ arbitrarias y luego randomizar las secuencias con el método de Discretización. De este modo se consigue extender los resultados a un conjunto muy amplio de mapas caóticos.

4. Respecto al muestreo caótico:

   - Los resultados con muestreo caótico son equivalentes a los obtenidos mediante muestreo aleatorio, si se emplean mapas con $IPDF$ uniforme y buenas propiedades de mixing ($r_{mix} \to 0$).

- Sin embargo para observar una influencia del valor de r_{mix} debieron emplearse filtros FIR excesivamente cortos, que en la práctica son inadecuados debido a que el número de elementos del FIR también afecta la atenuación en alta frecuencia.
- Luego puede concluirse que, a los fines prácticos, r_{mix} no es un factor relevante en este caso.

5. En cuanto a la implementación en Hardware de ruido estocástico del tipo f^{-d}:
   - El sistema diseñado permite seleccionar el tipo de ruido deseado dentro del rango $d = 1$ a 4.
   - En esta primera etapa no fue importante la optimización de recursos si no el definir una metodología general de diseño que fuera aplicable no sólo al ruido f^{-d} sino también a otros ruidos estocásticos de interés, tales como los gaussianos fraccionarios.
   - La principal conclusión en esta etapa de diseño es la ventaja de la utilización del entorno Simulink de Matlab para el diseño y simulación.

Las contribuciones de esta tesis se encuentran parcialmente publicadas en los siguientes trabajos:

1. "Randomizing nonlinear maps via symbolic dynamics". L. De Micco, C. M. González, H. A. Larrondo, M. T. Martin, A. Plastino, O. A. Rosso. Physica A, vol 387, Issue 14, pp 3373-3383 (2008) 10.1016/j.physa.2008.02.037.[36]

http://www3.fi.mdp.edu.ar/fc3/paperspdf/2008/physa08/PHYSA11361.pdf

2. “Zipping characterization of chaotic sequences used in spread spectrum communication systems”. L. De Micco, C. M. Arizmendi y H. A. Larrondo. American Institute of Physics Conference Proceedings 913, ISBN 913 978-0-7354- 0421-2, 238 páginas, pp 139-144. (2007).[74]

http://www3.fi.mdp.edu.ar/fc3/paperspdf/2006/aipCP913-2007-139.pdf

3. “Acquisition of Low Frequency Signals Immersed in Noise by Chaotic Sampling and FIR Filters. R. A. Petrocelli, L. De Micco, D. O. Carrica y H. A. Larrondo. Proceedings of WISP2007 (IEEE International Symposium on Intelligent Signal Processing) 351-356 (2007) ISBN 1-4244-0829-6/07/IEEE.[77]

http://www3.fi.mdp.edu.ar/fc3/paperspdf/2007/wisp07/paper_wisp07.pdf

4. “Muestreo caótico para la adquisición de señales de baja frecuencia con ruido de alta frecuencia”. L. De Micco, R. A. Petrocelli, D. O. Carrica y H. A. Larrondo. Proceedings de la XII Reunión de Trabajo en Procesamiento de la Información y Control, 16 al 18 de octubre de 2007 ISBN 978-987-1242-23-8. Trabajo 335.[127]

http://www3.fi.mdp.edu.ar/fc3/paperspdf/2007/rpic07/paper_rpic07_335final.pdf

5. “Constant Envelope Wideband Signals using Arbitrary Chaotic Maps. L. De Micco, R. A. Petrocelli y H. A. Larrondo. Proceedings de la XII Reunión de Trabajo en Procesamiento de la Información y Control, 16 al 18 de octubre de 2007 ISBN 978-987-1242-23-8, trabajo 381.[128]

http://www3.fi.mdp.edu.ar/fc3/paperspdf/2007/rpic07/paper_rpic07_381final.pdf

6. “Stochatic colored noise generator in FPGA”. O. G. Zabaleta, L. De Micco, C. M. González, C. M. Arizmendi y H. A. Larrondo. Proceedings of Designer’s

Forum. IV Southern Programmable Logic Conference (SPL08) ISBN 978-84-612-2376-3, 69-73 (2008).[129]

http://www3.fi.mdp.edu.ar/fc3/paperspdf/2008/spl08/PID550944.pdf

7. "Quantifiers for randomness of chaotic pseudo random number generators. L. De Micco, H. A. Larrondo, A. Plastino and O. A. Rosso aceptado para su publicación en Philosophical Transactions of the Royal Society A. (2008) [35]

http://arxiv.org/0812.2250v1

8. "Generalized Statistical Complexity Measure: a new tool for dynamical systems". O. A. Rosso, L. De Micco, H. A. Larrondo, M. T. Martin y A. Plastino. Enviado a International Journal of Bifurcation and Chaos (2008) [37]

Trabajo futuro

Sobre la base de estos resultados surge la necesidad de profundizar la implementación en lógicas programables:

1. Implementar en lógicas programables y en especial en FPGA, prototipos de distintos generadores caóticos y estocásticos. Se procurará lograr implementaciones modulares para su uso en aplicaciones de espectro esparcido.

2. Aplicar los cuantificadores de calidad desarrollados para distintas aritméticas, en especial aritméticas enteras y de punto flotante.

3. Desarrollar cuantificadores de calidad de implementación que midan aspectos como área ocupada, potencia requerida, velocidad de generación.

4. Implementar prototipos en las tres líneas investigadas: sistemas de comunicaciones, mejora de la compatibilidad electromagnética y filtrado digital mediante muestreo caótico.

Bibliografía

[1] E.Ñ. Lorenz. Deterministic non periodic flow. *Journal of the Atmospheric Sciences*, 20:130 – 141, 1963.

[2] M. J. Feigenbaum. Presentation functions and scaling function theory for circle maps. *Nonlinearity*, 1:577–602, 1988.

[3] C. A. Gayoso. *Generadores de números pseudoaleatorios en aritmética de residuos: teoría e implementación en FPGAs.* PhD thesis, Universidad Nacional de Mar del Plata, 2008.

[4] H. Whitney. Differentiable manifolds. *Ann. Math.*, 37:645, 1936.

[5] F. Takens. Detecting strange attractors in turbulence. *Lecture Notes in Mathematics*, 898, 1981.

[6] T. Sauer, J. A. Yorke, and M. Casdagli. Embedology. *Journal of statistical Physics*, 65:579, 1993.

[7] C. Beck and F. Schlögl. *Thermodynamics of chaotic systems: an introduction.* Cambridge University Press, 1997.

[8] A. Lasota and M. C. Mackey. *Chaos, Fractals, and Noise: Stochastic Aspects of Dynamics.* Applied Mathematical Sciences 97. Springer Verlag, 2nd. edition. edition, 1994.

[9] G. Setti, G. Mazzini, R. Rovatti, and S. Callegari. Statistical modeling of discrete-time chaotic processes: Basic finite-dimensional tools and applications. *Proceedings of the IEEE*, 90(5):662–689, Mayo 2002.

[10] J. Hadamard. Les surfaces à courbures opposées et leurs lignes géodésiques. *J. Math. Pures et Appl.*, 4:27–73, 1898.

[11] M. Morse and G. A. Hedlund. Symbolic dynamics. *American Journal of Mathematics*, 60:815–866, 1938.

[12] C. E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27:379–423 y 623–656., 1948.

[13] A.Ñ. Sharkovskii. Coexistence of cycles of a continuous map of a line into itself. *Ukrainian. Math. J.*, 16:61–71, 1964.

[14] J. Ziv and A. Lempel. A universal algorithm of sequential data compression. *IEEE Transactions on Information Theory*, IT23(3):337–343, 1977.

[15] J. Ziv and A. Lempel. On the complexity of finite sequences. *IEEE Trans. Inf. Theory*, 22:75–81, 1976.

[16] A.Ñ. Kolmogorov. Problems of information and transmission. *Three approaches to the quantitative definition of information.*, 1(1):1–7, 1965.

[17] G.J. Chaitin. On the length of programs for computing finite binary sequences. *Journal of the Association for Computing Machinery*, 13(4):547–569, 1966.

[18] C. E. Shannon and W. Weaver. The mathematical theory of communication. *Urbana, Illinois: University of Illinois Press*, 1949.

[19] A. Rényi. On measures of entropy and information. *Proc. Fourth Berkeley Symp. Math. Stat. and Probability*, 1:547–561, 1961.

[20] S. Lloyd and Pagels. Complexity as thermodynamic depth. *Annals of Physics*, 1988.

[21] C.H. Bennett. Emerging syntheses of science. *Information, Dissipation, and the Definition of Organization.*, 1987.

[22] D. McShea. Complexity and evolution: what everybody knows. *Biology and Philosophy*, 6(3):303–3024, 1991.

[23] R. López-Ruiz, H. L. Mancini, and X. Calbet. A statistical measure of complexity. *Phys. Lett. A*, 209:321–326, 1995.

[24] M. T. Martín and A. Plastino. Generalized statistical complexity measures: Geometrical and analytical properties. *Physica A*, 369:439–462, 2006.

[25] O. A. Rosso, M. T. Martin, A. Figliola, K. Keller, and A. Plastino. Wavelet-based informational tools: application to electroencephalogram record analysis. *Journal on Neuroscience Methods*, 153:163–182, 2006.

[26] M. T. Martín, A. Plastino, and O. A. Rosso. Statistical complexity and disequilibrium. *Phys. Lett. A*, 311:126–132, 2003.

[27] S. Shiner, M. Davison, and T. Landsberg. Simple measure for complexity. *Physical Review E*, 59:1459–1464, 1999.

[28] D. P. Feldman and J. P. Crutchfield. Measures of statistical complexity: why? *Physics Letters A*, 238:244–252, 1998.

[29] C. Anteneodob and A. R. Plastino. Some features of the lópez-ruiz-mancini-calbet (lmc) statistical measure of complexity. *Phys. Lett. A*, 223(5):348–354, Diciembre 1996.

[30] P. W. Lamberti, M. T. Martín, A. Plastino, and O. A. Rosso. Intensive entropic non-triviality measure. *Physica A*, 334:119–131, 2004.

[31] J. Eckmann, S. Oliffson Kamphorst, and D. Ruelle. Recurrence plots of dynamical systems. *Europhys. Lett.*, 4:973–977, 1987.

[32] D. P. Feldman, C. S. McTague, and P. Crutchfield. The organization of intrinsic omputation: complexity-entropy diagrams and the diversity of natural information processing. *arxiv.org:0866.4789[nlin.CD]*, pages 1–18, junio 2008.

[33] J. P. Crutchfield and N. H. Packard. Symbolic dynamics of noisy chaos. *Physica D*, 7:201–223, 1983.

[34] O. A. Rosso, H. A. Larrondo, M. T. Martin, A. Plastino, and M. A. Fuentes. Distinguishing noise from chaos. *Phys. Rev. Lett.*, 99:154102–154106, 2007.

[35] L. De Micco, H. A. Larrondo, A. Plastino, and O. A. Rosso. Quantifiers for randomness of chaotic pseudo random number generators. *arxiv.org/0812.2250v1*, preprint, 2008.

[36] L. De Micco, C. M. González, H. A. Larrondo, M. T. Martin, A. Plastino, and O. A. Rosso. Randomizing nonlinear maps via symbolic dynamics. *Physica A*, 387:3373–3383, 2008.

[37] O. A. Rosso, L. De Micco, H. A. Larrondo, M. T. Martin, and A. Plastino. Generalized statistical complexity measure: a new tool for dynamical systems. *submitted to International Journal of Bifurcation and Chaos*, preprint, 2008.

[38] H. Wold. *A Study in the Analysis of Stationary Time Series*. Ph D. dissertation, University of Stokholm, 1938.

[39] J. Kurths and H. Herzel. An attractor in a solar time series. *Physica D*, 25:165–172, 1987.

[40] M. T. Martin. *Ph.D. Thesis, Department of Mathematics,*. PhD thesis, Faculty of Sciences, University of La Plata, 2004.

[41] K. Mischaikow, M. Mrozek, J. Reiss, and A. Szymczak. Construction of symbolic dynamics from experimental time series. *Phys. Rev. Lett.*, 82:1114–1147, 1999.

[42] G. E. Powell and I. C. Percival. A spectral entropy method for distinguishing regular and irregular motion of hamiltonian systems. *J. Phys. A: Math. Gen.*, 12:2053–2071, 1979.

[43] S. Blanco, A. Figliola, R. Quian Quiroga, O. A. Rosso, and E. Serrano. Time-frequency analysis of electroencephalogram series (iii): Wavelet packets and information cost function. *Phys. Rev. E*, 57:932–940., 1998.

[44] O. A. Rosso, S. Blanco, J. Jordanova, V. Kolev, A. Figliola, M. Schürmann, and E. Başar. Wavelet entropy: a new tool for analysis of short duration brain electrical signals. *Journal of Neuroscience Methods*, 105:65–75, 2001.

[45] W. Ebeling and R. Steuer. Partition-based entropies of deterministic and stochastic maps. *Stochastics and Dynamics*, 1(1):1–17, 2001.

[46] C. Bandt and B. Pompe. Permutation entropy: a natural complexity measure for time series. *Phys. Rev. Lett.*, 88:174102–1, 2002.

[47] K. Keller and M. Sinn. Ordinal analysis of time series. *Physica A*, 356:114–120, 2005.

[48] J. M. Amigó, L. Kocarev, and I. Tomovski. Discrete entropy. *Physica D*, 228:77–85., 2007.

[49] C. Bandt and B. Pompe. Permutation entropy: a natural complexity measure for time series. *Phys. Rev. Lett.*, 88:174102–1, 2002.

[50] K. Keller and H. Lauffer. Symbolic analysis of high-dimensional time series. *Int. J. Bifurcation and Chaos*, 13:2657–2668, 2003.

[51] H. A. Larrondo, M. T. Martin, C.M. González, A. Plastino, and O. A. Rosso. Random number generators and causality. *Phys. Lett. A*, 352((4- 5)):421–425, Abril 2006.

[52] A. R. Plastino and A. Plastino. Symmetries of the fokker-planck equation and the fisher-frieden arrow of time. *Phys. Rev. E*, 54, 1996.

[53] R. López-Ruiz Calbet. Tendency towards maximum complexity in a nonequilibrium isolated system. *Phys. Rev. E*, 63:066116, 2001.

[54] A. M. Kowalski, M. T. Martín, A. Plastino, and O. A. Rosso. Entropic non-triviality, the classical limit and geometry-dynamics correlations. *Int. J. Mod. Phys. B*, 19:2273, 2005.

[55] L. Zunino, D.G. Pérez, M. T. Martín, A. Plastino, M. Garavaglia, and O. A. Rosso. Characterization of gaussian self-similar stochastic processes using wavelet-based informational tools. *Phys. Rev. E*, 75:021115, 2007.

[56] H. G. Schuster. *Deterministic Chaos*. 2nd edition, 1988.

[57] C. Bandt and F. Shiha. Order patterns in time series. *Journal of Time Series Analysis*, 28:646–665, 2007.

[58] R. B. Davies and D. S. Harte. Tests for hurst effect. *Biometrika*, 74:95, 1987.

[59] A. T. A. Wood and G. Chan. Simulation of stationary gaussian vector fields. *J. Comput. Graph. Stat.*, 3:409–432, 1994.

[60] J. C. Sprott. *Chaos and Time Series Analysis*. Oxford University Press, 2004.

[61] E. Ott, T. Sauer, and J. A. Yorke. *Coping with Chaos*. Wiley, New York, 1994.

[62] O. A. Rosso and M. L. Mairal. Characterization of time dynamical evolution of electroencephalographic records. *Physica A*, 312:469–504, 2002.

[63] T. Stojanovski and L. Kocarev. Chaos-based random number generators - part i: Analysis. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, 48(3):281–288, Marzo 2001.

[64] T. Stojanovski, J. Pihl, and Kocarev. Chaos-based random number generators - part ii: Practical realization. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, 48(3):382–386, March 2001.

[65] L. Kocarev and G. Jakimoski. Pseudorandom bits generated by chaotic maps. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, 50(1):123–126, Enero 2003.

[66] P. L'Ecuyer. Uniform random number generation. *Annals of Operations Research*, 53:77–120, 1994.

[67] M. Pecora, L. Carroll, and L. Thomas. Synchronization in chaotic systems. *Phys. Rev. Lett.*, 64(8):821–824, Febrero 1990.

[68] L. Kocarev and U. Parlitz. General approach for chaotic synchronization with applications to communication. *Physical Review Letters*, 74(25):5028–5031, 1995.

[69] R. M. Hidalgo, J. G. Fernández, R. R. Rivera, and H. A. Larrondo. Versatile dsp-based chaotic communication system. *Electronic Letters*, 37:1204–1205, 2001.

[70] J. G. Fernández, H. A. Larrondo, H. A. Slavin, D. G. Levin, R. M. Hidalgo, and R. R Rivera. Masking properties of apd communication systems. *Physica A*, 328:351–359, 2003.

[71] E. H. Dinan and B. Jabbari. Spreading codes for direct sequence cdma and wideband cdma cellular networks. *IEEE Communications Magazine*, 36(9):48–54, September 1998.

[72] G. Mazzini, G. Setti, and R. Rovatti. Chaotic complex spreading sequences for aynhchronous ds-cdma-part 1: System modeling and results. *IEEE Trans. Circuits Sys. 1*, 44(10):937–947, 1997.

[73] X. Shan, Y. Xia, Y. Ren, and J. Yuan. Spatiotemporal chaotic spreading sequences for cdma communications. In *Communication Technology Proceedings*, volume 1, pages 530–535, Agosto 2006.

[74] L. De Micco, C. M. Arizmendi, and H. A. Larrondo. Zipping characterization of chaotic sequences used in spread spectrum communication systems. *Institute of Physics Conference Proceedings 913*, pages 139–144, 2007.

[75] G. Setti, M. Balestra, and R. Rovatti. Experimental verification of enhanced electromagnetic compatibility in chaotic fm clock signals. In *Proceedings of ISCAS'00*, pages III–229–232. IEEE Circuits and Systems Society, 2000.

[76] S. Callegari, R. Rovatti, and G. Setti. Chaotic modulations can outperform random ones in electromagnetic interference reduction tasks. *Electronics Letters*, 38(12):543–544, Junio 2002.

[77] R. A. Petrocelli, L. De Micco, D. O. Carrica, and H. A. Larrondo. Acquisition of low frequency signals immersed in noise by chaotic sampling and fir filters. *Proceedings of WISP2007 (IEEE proceedings ISBN 1-4244-0830-X)*, pages 351–356, 2007.

[78] C. M. González, H. A. Larrondo, and O. A. Rosso. Statistical complexity measure of pseudorandom bit generators. *Physica A*, 354:281–300, Agosto 2005.

[79] G. Marsaglia. The marsaglia random number cdrom including the diehard battery of tests of randomness. *http://www.stat.fsu.edu/pub/diehard/*, 1995.

[80] P. Cornfeld, S. V. Fomin, and Ya. G. Sinai. *Ergodic Theory.* Springer New York, 1982.

[81] M. Dellnitz, G. Froyland, and S. Sertl. On the isolated spectrum of the perron-frobenius operator. *Nonlinearity*, 13:1171–1188, 2000.

[82] A. Lasota and J. A. Yorke. On the existence of invariant measure for piecewise monotonic transformations. *Trans. Amer. Math. Soc.*, 186:481–488, 1973.

[83] J. Ding and A. Zhou. Finite approximations of frobenius-perron operators. a solution to ulam's conjecture to multi-dimensional transformations. *Physica D*, 92:61–68, 1996.

[84] A. Rogers, R. Shorten, and D. M. Heffernan. Synthesizing chaotic maps with prescribed invariant densities. *Physics Letters A*, 330(6):435–441, 2004.

[85] D. Pingel and P. Schmelcher. Theory and examples of the inverse frobenius-perron problem for complete chaotic maps. *Chaos*, 9(3):357–366, 1999.

[86] S. Callegari, G. Setti, and P. J. Langlois. A cmos tailed tent map for the generation of uniformly distributed chaotic sequences. In IEEE Cyrcuits and Systems Society, editors, *Proceedings of ISCAS'97*, volume 1, pages 781–784. IEEE, 2003.

[87] M. Jessa. The period of sequences generated by tent-like maps. *IEEE Trans. on Circuits and Systems I: Fundamental theory and Applications*, 49:84–89, 2002.

[88] C. Beck and G. Röpstorff. Effects of phase space discretizacion on the long-time behavior of dynamical systems. *Physica D*, 25:173–180, 1987.

[89] O. A. Rosso, L. Zunino, D. G. Pérez, A. Figliola, H. A. Larrondo, M. Garavaglia, Martín M. T., and A. Plastino. Extracting features of gaussian selfsimilar stochastic processes via the bandt & pompe approach. *Phys. Rev. E*, 76(6):061114, 2007.

[90] H. A. Larrondo, C. M. González, M. T. Martin, A. Plastino, and O. A. Rosso. Intensive statistical complexity measure of pseudorandom number generators. *Physica A*, 356:133–138, 2005.

[91] A. M. Kowalski, M. T. Martín, A. Plastino, and O. A. Rosso. Bandt-pompe approach to the classical-quantum transition. *Physica D*, 233:21–31, 2007.

[92] O. A. Rosso, R. Vicente, and C.R. Mirasso. Encryption test of pseudo-aleatory messages embedded on chaotic laser signals: an information theory approach. *Phys. Lett. A*, 372:1018–1023, 2008.

[93] L. Zunino, D.G. Pérez, M. T. Martín, M. Garavaglia, A. Plastino, and O. A. Rosso. Permutation entropy of fractional brownian motion and fractional gaussian noise. *Physics Letters A*, 372(27-28):4768–4774, June 2008.

[94] N. Marwan, M. C. Romano, M. Thiel, and J. Kurths. Recurrence plots for the analysis of complex systems. *Physics Reports*, 438:237–329, 2007.

[95] M. Sushchik, L. S. Tsimring, and A. R. Volkovskii. Performance analysis of correlation-based communication schemes utilizing chaos. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, 47(12):1684–1691, Diciembre 2000.

[96] R. Rovatti, G. Mazzini, and G. Setti. On the ultimate limits of chaos-based asynchronous ds-cdma- i: Basic definitions and results. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, 51(7):1336–1347, Julio 2004.

[97] R. Rovatti, G. Mazzini, and G. Setti. On the ultimate limits of chaos-based asynchronous ds-cdma-ii: Analytical results and asymptotics. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, 51(7):1348–1364, Julio 2004.

[98] G. Mazzini. Ds-cdma systems using q-level m sequences: coding map theory. *IEEE Transactions on Communications*, 45(10):1304–1313, Octubre 1997.

[99] G. Mazzini, G. Setti, and R. Rovatti. Chaotic complex spreading sequences for asynchronous ds-cdma - part i: System modeling and results. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, 44(10):937–947, 1997.

[100] M. B. Pursley. Performance evaluation for phase coded spread spectrum multiple access communication, part i: System analysis. *IEEE Transactions on Communications*, COM-25(8):795–799, 1977.

[101] D. V. Sarwate and M. B. Pursley. Crosscorrelation properties of pseudo-random and related sequences. *Proc. of the IEEE.*, 68(5):593–619, 1980.

[102] L. P. Welch. Lower bounds on the maximum cross correlation of signals. *IEEE. Trans. Inform. Theory*, 20(3):397–399, 1974.

[103] J. R. Sanchez, F. Family, and C. M. Arizmendi. Complexity of thermal ratchet motion. *Phys. Lett. A*, 249:281–285, 1998.

[104] S. Callegari, R. Rovatti, and G. Setti. Chaos-based fm ignals: application and implementation issues. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, 50(8):1141–1147, Agosto 2003.

[105] S. Santi, R. Rovatti, and G. Setti. Advanced chaos-based frequency modulations for clock signals emc tuning. In IEEE Cyrcuits and Systems Society, editors, *Proceedings of ISCAS'03*, volume 3, pages 116–119. IEEE, 2003.

[106] S. Callegari, R. Rovatti, and G. Setti. Spectral properties of chaos-based fm signals: theory and simulation results. *IEEE Trans. Circuits Sys. I*, 50(1):3–15, 2003.

[107] D. O. Carrica, R. Petrocelli, M. Benedetti, and M. A. Funes. Acquisition of low-frequency signals immersed in noise by random sampling and finite impulse response filters. *Review of Scientific Instruments*, 78:044702/1–8, 2007.

[108] F. J. Beutler. Alias-free randomly timed sampling of stochastic processes. *IEEE Trans. Information Theory*, 16(1):147–152, Marzo 1970.

[109] A. K. Ivars Bilinskis, Mikelson. *Randomized Signal Processing*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1992.

[110] I. Bilinskis, G. Cain, I. Mednieks, and A. Mikelsons. Nonuniform sampling based signal generation. In Norway Loen, editor, *Proceedings of the 1999 International Workshop on Sampling Theory and Application*, pages 180–183, 1999.

[111] H. S. Shapiro and R. A. Silverman. Alias-free sampling of random noise. *SIAM J. Appl. Math.*, 8:245–248, 1960.

[112] A. Balakrishnan. On the problem of time jitter in sampling. *IEEE Trans. Information Theory*, 8(1):226–236, Abril 1962.

[113] A. Tarczynski. Fir filters for systems with input clock jitter. *IEEE International Symposium on Circuits and Systems, ISCAS 2001*, 2(1):617–620, Mayo 2001.

[114] D. Mirri, G. Iuculano, F. Pasini, and Filicori. The effect of time-jitter in equispaced sampling wattmeters. *IEEE Trans. Instrum. Meas.*, 47(1):720–727, 1998.

[115] N. Da Dalt, M. Harteneck, C. Sandner, and A. Wiesbauer. On the jitter requirements of the sampling clock for analog-to-digital converters. *IEEE Trans. Circuits Syst. I*, 49(9):1354–1360, Septiembre 2002.

[116] D. Carrica, M. Benedetti, and R. Petrocelli. Random sampling applied to the measurement of a dc signal immersed in noise. *IEEE Transactions on Instrumentation and Measurement*, 50(5):1319–1323, Octubre 2001.

[117] D. Bland and A. Tarczynski. The effect of sampling jitter in a digitized signal. *IEEE International Symposium on Circuits and Systems ISCAS97*, 4:2685–2688, 1997.

[118] G. Mazzini, R. Rovatti, and G. Setti. Capacity of chaos-based asynchronous ds-cdma systems with exponentially vanishing autocorrelations. *Electronics Letters*, 35(25):1717–1718, Diciembre 2002.

[119] A. R. Rogers. *Synthesis and Applications of Chaotic Maps*. PhD thesis, 2005.

[120] C. M. González, H. A. Larrondo, C. A Gayoso, and L. J. Arnone. Generación de secuencias binarias pseudo aleatorias por medio de un mapa caótico 3d. In *Proceedings del IX Workshop de IBERCHIP*, 2003.

[121] Bernard M. Oliver. *Electronic Measurements and Instrumentation*. McGraw-Hill Inc.,US, 1971.

[122] C. M. González, H. A. Larrondo, C. A Gayoso, and L. J. Arnone. Secuencias binarias pseudo aleatorias generadas por un mapa caótico 2d. In *Proceedings del X Workshop de IBERCHIP*, 2004.

[123] ALTERA. ip-basesuite.html. *http://www.altera.com/products/ip/design/basesuite/ip-basesuite.html.*, 2008.

[124] R. M. Hidalgo, J. G. Fernandez, R. R. Rivera, and H. A. Larrondo. A simple adjustable window algorithm to improve fft measurements. *IEEE Transactions on instrumentation and measurement*, 51(1):31–36, 2002.

[125] E. Boemo and G. Sutter. Permutación de los datos de entrada como estrategia de diseño de bajo consumo: Algunos ejemplos en fpgas. In *Computación Reconfigurable y FPGAs*, 2003.

[126] S. Lopez-Buedo, J. Garrido, and E. Boemo. Thermal testing on reconfigurable computers. *IEEE Design and Test of Computers*, pages 84–90, 2000.

[127] L. De Micco, R. A. Petrocelli, D. O. Carrica, and H. A. Larrondo. Muestreo caótico para la adquisición de señales de baja frecuencia con ruido de alta frecuencia. *Proceedings de la XII Reunión de Trabajo en Procesamiento de la Información y Control*, 2007.

[128] L. De Micco, R. A. Petrocelli, and H. A. Larrondo. Constant envelope wideband signals using arbitrary chaotic maps. *Proceedings of XII RPIC*, 2007.

[129] O. G. Zabaleta, L. De Micco, C. M. González, C. M. Arizmendi, and H. A. Larrondo. Stochatic colored noise generator in fpga. *Proceedings of Designer's Forum. IV Southern Programmable Logic Conference*, 2008.