

**UNIVERSIDAD DE CANTABRIA**



**DEPARTAMENTO DE INGENIERÍA DE  
COMUNICACIONES**

**TESIS DOCTORAL**

**ESTIMACIÓN ÓPTIMA DE SECUENCIAS  
CAÓTICAS CON APLICACIÓN EN  
COMUNICACIONES**

**Autor : David Luengo García  
Directores : Carlos Pantaleón Prieto  
Ignacio Santamaría Caballero**

**Grupo de Tratamiento Avanzado de Señal**

**Septiembre de 2006**

# Capítulo 4

## Estimación Bayesiana de Secuencias Caóticas

### 4.1. Introducción

En el Capítulo 3 se ha resuelto el problema de la estimación de una secuencia caótica mediante el procedimiento de máxima verosimilitud (ML). Aunque este método alcanza el CRLB asintóticamente cuando la SNR tiende a infinito, y proporciona buenos resultados en general, su rendimiento para bajas relaciones señal a ruido puede mejorarse recurriendo a estimadores Bayesianos, que tienen en cuenta el conocimiento a priori disponible acerca de la secuencia. En este capítulo se plantea el marco de estimación Bayesiana general para mapas caóticos unidimensionales, y se proponen fórmulas explícitas para una clase amplia de funciones: los mapas PWL.

La estructura del capítulo es similar a la del Capítulo 3. El modelo matemático y las expresiones para la iteración de un mapa PWL ya se han mostrado en dicho capítulo, de modo que en este se comienza planteando el marco de estimación Bayesiana general para el problema en la Sección 4.2. Se propone una función de densidad de probabilidad (FDP) a priori para las muestras de la secuencia caótica, se describen la FDP de las observaciones y la FDP a posteriori, y se muestra la forma general de los dos estimadores Bayesianos más comunes: el estimador de máxima probabilidad a posteriori (MAP) y el estimador de mínimo error cuadrático medio (MMSE o MS).

A continuación, en la Sección 4.3 se particularizan los estimadores MAP y MS para el caso de mapas PWL, en el que es posible obtener expresiones cerradas y algoritmos explícitos que resuelven el problema de manera exacta. En el caso de un mapa genérico no es posible encontrar dichas expresiones, al igual que ocurría con el estimador ML. No obstante, en la Sección 4.4 se discute brevemente el problema de la estimación Bayesiana de secuencias generadas por mapas no PWL, y se presenta un nuevo mecanismo que permite obtener estimadores Bayesianos para cualquier mapa caótico: los métodos de Monte Carlo basados en muestreo sobre cadenas de Markov (métodos MCMC). Finalmente, el capítulo se cierra mostrando los principales resultados obtenidos en la Sección 4.5, y la discusión de los mismos y las conclusiones en la Sección 4.6.

## 4.2. El Problema de la Estimación Bayesiana de Secuencias Caóticas

El estimador ML trata los parámetros a estimar como variables desconocidas pero deterministas. Los estimadores Bayesianos intentan obtener una mejor estima modelando los parámetros como variables aleatorias con una determinada FDP a priori, y planteando un criterio de optimalidad basado en algún estadístico de la FDP de los parámetros que se desean estimar, la secuencia caótica en este caso, condicionados por las observaciones disponibles [VanTre1968, Kay1993]. Utilizando el teorema de Bayes, esta *FDP a posteriori* de los parámetros, que mide la probabilidad de que una determinada secuencia caótica,  $\mathbf{x}$ , haya sido la generada realmente por el mapa dada la secuencia de observaciones ruidosas,  $\mathbf{y}$ , se puede escribir como

$$p(\mathbf{x}|\mathbf{y}) = \frac{p(\mathbf{y}|\mathbf{x})p(\mathbf{x})}{p(\mathbf{y})}, \quad (4.1)$$

Respecto al resto de FDPs,  $p(\mathbf{y}|\mathbf{x})$  es la FDP de las observaciones cuando se conoce la secuencia caótica original,  $\mathbf{x}$ ;  $p(\mathbf{x})$  es la *FDP a priori* de generar una determinada secuencia caótica (es decir, la información disponible sobre los parámetros antes de realizar el experimento); y  $p(\mathbf{y})$  es la FDP de las observaciones, que no depende de la secuencia caótica real, y que es simplemente un término de normalización.

Los estimadores Bayesianos encuentran el vector  $\mathbf{x}$  óptimo con respecto a algún criterio estadístico formulado sobre la FDP a posteriori de la secuencia caótica: el valor más probable (esto es, el máximo de la FDP) en el caso del *estimador de máxima probabilidad a posteriori* (MAP), y el valor medio en el caso del *estimador de mínimo error cuadrático medio* (MMSE o MS). No obstante, para plantear dichos estimadores es necesario encontrar previamente las FDPs que aparecen en (4.1). Todos los estimadores requieren las expresiones de la FDP condicional de las observaciones,  $p(\mathbf{y}|\mathbf{x})$ , y la FDP a priori,  $p(\mathbf{x})$ , y algunos (como el MS por ejemplo) requieren adicionalmente el cálculo del factor de normalización,  $p(\mathbf{y})$  [VanTre1968, Kay1993]. En los siguientes apartados se muestra la forma genérica de estas tres FDPs, así como las expresiones generales de los estimadores MAP y MS para un mapa caótico unidimensional arbitrario.

### 4.2.1. FDP Condicional de las Observaciones

En este caso la FDP está compuesta simplemente por una colección de  $N + 1$  variables aleatorias Gaussianas independientes e idénticamente distribuidas (IID) con la misma varianza aunque distinta media, de modo que la FDP es idéntica a la que se utiliza para el estimador ML, dada por (3.6):

$$p(\mathbf{y}|\mathbf{x}) = (2\pi\sigma^2)^{-(N+1)/2} \exp\left(-\frac{1}{2\sigma^2}(\mathbf{y} - \mathbf{x})^T(\mathbf{y} - \mathbf{x})\right). \quad (4.2)$$

La diferencia de notación entre (3.6) y (4.2),  $p(\mathbf{y}; \mathbf{x})$  frente a  $p(\mathbf{y}|\mathbf{x})$ , es debida simplemente a que en el primer caso los parámetros se consideran deterministas aunque

desconocidos, mientras que en el segundo caso son variables aleatorias. Es decir, en el primer caso se tiene una familia de FDPs parametrizada por un vector  $\mathbf{x}$  que se desea estimar, mientras que en el segundo caso se trata de una FDP condicional [Kay1993].

### 4.2.2. FDP a Priori de la Secuencia Caótica

La FDP a priori de la secuencia caótica completa es la FDP conjunta de todas las muestras de la secuencia. Tomando como referencia la muestra  $n$ -ésima,  $x[n]$ , esta FDP se puede expresar como

$$\begin{aligned}
 p(\mathbf{x}) &= p(x[0], \dots, x[N]) \\
 &= p(x[0], \dots, x[n-1], x[n+1], \dots, x[N] | x[n]) p(x[n]) \\
 &= p(x[0], \dots, x[n-1] | x[n], \dots, x[N]) p(x[n+1], \dots, x[N] | x[n]) p(x[n]) \\
 &= p(x[0], \dots, x[n-1] | x[n]) p(x[n+1], \dots, x[N] | x[n]) p(x[n]), \quad (4.3)
 \end{aligned}$$

donde el último paso es posible gracias a que la secuencia de muestras  $\mathbf{x}_{0:n-1} = [x[0], \dots, x[n-1]]^T$  depende únicamente del valor de  $x[n]$  y en ningún caso de  $\mathbf{x}_{n+1:N} = [x[n+1], \dots, x[N]]^T$ . Esto es, la iteración hacia atrás es un *proceso de Markov*, aunque no con una correspondencia unívoca entre el estado actual y el futuro, como en el caso de la iteración hacia delante, ya que en general los mapas usados no son invertibles, de modo que siempre existen varias preimágenes.

Ahora debe hallarse la expresión de cada una de estas tres FDPs para obtener la FDP a priori de la secuencia caótica. Aunque presenten características propias de señales aleatorias, las señales caóticas con las que se trabaja en esta Tesis son puramente deterministas. Por lo tanto, la iteración hacia delante del mapa a partir de  $x[n]$  se encuentra completamente especificada, y la segunda FDP de (4.3) resulta

$$p(x[n+1], \dots, x[N] | x[n]) = \prod_{k=1}^{N-n} \delta(x[n+k] - f^k(x[n])),$$

donde  $\delta(\cdot)$  es la *función impulso* o *delta de Dirac* unidimensional, que indica que la probabilidad vale uno en los puntos en que las muestras de la secuencia coinciden con la iteración hacia delante a partir de  $x[n]$ , y cero en el resto.

Respecto a las muestras previas, dado que los mapas con los que se trabaja en general no son invertibles, no se hallan completamente determinadas por  $x[n]$ . Es necesario conocer también la secuencia simbólica:  $\mathbf{s}_{0:n-1} = [s[0], \dots, s[n-1]]^T$ . Puesto que en un caso general dicha secuencia simbólica es desconocida,  $p(x[0], \dots, x[n-1] | x[n])$  depende de la probabilidad de tener un determinado itinerario, y esta FDP se debe expresar como un sumatorio a lo largo de todos los posibles itinerarios de longitud  $n$ :

$$p(x[0], \dots, x[n-1] | x[n]) = \sum_{i=1}^{P(n)} \Pr(\mathbf{s}_{0:n-1}^i) \prod_{k=1}^n \delta(x[n-k] - f_{\mathbf{s}_{n-k:n-1}^i}^{-k}(x[n])).$$

Donde,  $\mathbf{s}_{0:n-1}^i$  denota la  $i$ -ésima secuencia simbólica de longitud  $n$  correspondiente a las  $n$  primeras muestras de la secuencia caótica,  $\mathbf{s}_{0:n-1}^i = [s_i[0], \dots, s_i[n-1]]^T$ ;  $\Pr(\mathbf{s}_{0:n-1}^i)$  es la probabilidad de que ocurra dicha secuencia simbólica;  $P(n)$  es el número de secuencias simbólicas de longitud  $n$  válidas para un cierto  $x[n]$ ;  $\delta(x)$  es nuevamente la delta de Dirac; y el subíndice  $\mathbf{s}_{n-k:n-1}^i$  en  $f$  indica que las iteraciones hacia atrás de  $f$  se obtienen utilizando la secuencia simbólica  $\mathbf{s}_{n-k:n-1}^i$  correspondiente. Juntando las dos expresiones anteriores, la ecuación (4.3) para la FDP a priori se convierte en

$$\begin{aligned}
p(\mathbf{x}) &= \left( \sum_{i=1}^{P(n)} \Pr(\mathbf{s}_{0:n-1}^i) \prod_{k=1}^n \delta(x[n-k] - f_{\mathbf{s}_{n-k:n-1}^i}^{-k}(x[n])) \right) \\
&\quad \times \prod_{k=1}^{N-n} \delta(x[n+k] - f^k(x[n])) p(x[n]) \\
&= p(x[n]) \sum_{i=1}^{P(n)} \Pr(\mathbf{s}_{0:n-1}^i) \prod_{k=-n}^{N-n} \delta(x[n+k] - f_{\mathbf{s}_{n+k:n-1}^i}^k(x[n])). \\
&= p(x[n]) \sum_{i=1}^{P(n)} \Pr(\mathbf{s}_{0:n-1}^i) \delta(\mathbf{x} - \mathbf{x}_i^n), \tag{4.4}
\end{aligned}$$

con

$$\mathbf{x}_i^n = [f_{\mathbf{s}_{0:n-1}^i}^{-n}(x[n]), \dots, f_{\mathbf{s}_{i[n-1]}^i}^{-1}(x[n]), x[n], f(x[n]), \dots, f^{N-n}(x[n])]^T.$$

La ecuación (4.4) muestra la FDP más general posible para un mapa caótico unidimensional. Los dos casos particulares que presentan un mayor interés son, como siempre, elegir como referencia la última muestra de la secuencia, y elegir la primera muestra. En caso de seleccionarse como referencia la última muestra,  $x[N]$ , todas las demás muestras deben obtenerse mediante iteración hacia atrás, y (4.4) resulta

$$p(\mathbf{x}) = p(x[N]) \sum_{i=1}^{P(N)} \Pr(\mathbf{s}_{0:N-1}^i) \prod_{k=0}^N \delta(x[N-k] - f_{\mathbf{s}_{N-k:N-1}^i}^{-k}(x[N])). \tag{4.5}$$

Si, por el contrario, se escoge como muestra de referencia la primera de la secuencia,  $x[0]$ , la evolución de toda la secuencia se encuentra perfectamente determinada, y en consecuencia se evita el inconveniente de la dependencia de la FDP a priori con respecto a la secuencia simbólica observada en (4.4) y (4.5). En este caso, (4.4) se reduce a

$$p(\mathbf{x}) = p(x[0]) \prod_{k=0}^N \delta(x[k] - f^k(x[0])). \tag{4.6}$$

Para caracterizar completamente la FDP a priori únicamente resta proponer una FDP para la muestra de referencia, es decir, definir  $p(x[n])$ . En general se va a trabajar

con mapas caóticos estacionarios y ergódicos. Esto significa que puede considerarse que las secuencias generadas se encuentran dentro de su atractor y sus muestras tienen una FDP que coincide con la FDP invariante natural del mapa. En consecuencia, la elección lógica para la FDP a priori de la muestra de referencia es

$$p(x[n]) = p(x) \quad \text{con } 0 \leq n \leq N,$$

siendo  $p(x)$  la FDP invariante natural del mapa. Esta FDP condensa toda la información disponible a priori sobre el comportamiento estadístico del mapa, e indica el grado de conocimiento que se tiene acerca de su evolución.

Nótese que, si no se dispone de ninguna información relativa a la estadística de las secuencias generadas por el mapa, se puede utilizar una *FDP no informativa*, que típicamente va a ser una FDP uniforme en todo el espacio de fases correspondiente al atractor del mapa caótico [Box1973]. En estas circunstancias todas las regiones del atractor presentan la misma probabilidad, y por lo tanto el estimador obtenido es básicamente el estimador ML. Aunque la FDP invariante de algunos mapas caóticos es uniforme (véase el Apéndice C), la FDP de la mayoría se aleja mucho de la misma. En estos casos, al utilizar la FDP invariante natural como FDP a priori se están favoreciendo ciertas regiones del atractor dentro de las cuales se sabe que la secuencia caótica tiene mayor probabilidad de encontrarse, en detrimento de otras que la señal visita con menor frecuencia. En consecuencia, es razonable suponer que el rendimiento de los estimadores va a mejorar para la mayor parte de las secuencias típicas, aunque en general resulta inevitable que empeore para algunas secuencias atípicas.

Para finalizar, téngase en cuenta que, aunque en principio cualquier FDP a priori es igualmente válida dentro del contexto de la filosofía de estimación Bayesiana, para conseguir un resultado óptimo de los estimadores se debe buscar una FDP lo más ajustada posible a la que presenta la secuencia real. En realidad, si no se puede utilizar una FDP realista no parece razonable considerar estimadores Bayesianos, ya que presentan una complejidad mayor que el estimador ML, y en este caso proporcionarían seguramente peores resultados en promedio [Box1973, Kay1993]. Por consiguiente, en esta Tesis se considera únicamente la FDP invariante natural (o una aproximación constante a tramos (PWC) de la misma cuando la FDP real no se puede obtener de manera analítica, como se muestra en la Sección 4.3.1.1) como FDP de  $x[n]$ .

### 4.2.3. FDP de las Observaciones

El último elemento necesario para obtener la FDP deseada y plantear los estimadores Bayesianos es la FDP de las observaciones,  $p(\mathbf{y})$ . Esta FDP constituye simplemente un factor de normalización (necesario para que la probabilidad total a lo largo de todo el espacio de fases valga uno) que no depende de la secuencia caótica real, y que se puede obtener integrando  $p(\mathbf{x} | \mathbf{y})$  e igualando a uno [VanTre1968, Kay1993]:

$$\int_{\mathbf{x}} p(\mathbf{x} | \mathbf{y}) d\mathbf{x} = \int_{\mathbf{x}} \frac{p(\mathbf{y} | \mathbf{x})p(\mathbf{x})}{p(\mathbf{y})} d\mathbf{x} = 1.$$

Donde  $X$  indica el espacio de fases de  $\mathbf{x}$ , que en general va a ser simplemente una proyección  $N$ -dimensional del atractor. Esto es, si el atractor es  $A \subseteq \mathbb{R}$ , entonces  $X = A^N \subseteq \mathbb{R}^N$ . Puesto que  $p(\mathbf{y})$  no depende de la variable de integración, se puede despejar en la ecuación anterior, obteniéndose

$$p(\mathbf{y}) = \int_X p(\mathbf{y} | \mathbf{x}) p(\mathbf{x}) d\mathbf{x}. \quad (4.7)$$

Respecto a la secuencia de muestras que se desea estimar,  $\mathbf{x}$ , (4.7) es una constante. Por consiguiente, para algunos estimadores Bayesianos, como el MAP por ejemplo, (4.7) no influye en el resultado final, y no es necesario calcularla. Sin embargo, para otros, como el MS por ejemplo, es imprescindible su concurso e influye decisivamente en el resultado final.

#### 4.2.4. FDP a Posteriori

La expresión general para la FDP a posteriori de la secuencia caótica condicionada por las observaciones se obtiene juntando las ecuaciones (4.2) y (4.7):

$$p(\mathbf{x} | \mathbf{y}) = \frac{p(\mathbf{x}) \exp\left(-\frac{1}{2\sigma^2}(\mathbf{y} - \mathbf{x})^T(\mathbf{y} - \mathbf{x})\right)}{(2\pi\sigma^2)^{(N+1)/2} \int_X p(\mathbf{y} | \mathbf{x}) p(\mathbf{x}) d\mathbf{x}}. \quad (4.8)$$

Sin embargo, en el Capítulo 3 y en los apartados anteriores se ha mostrado que todas las FDPs se pueden expresar en función de una muestra de referencia,  $x[n]$ , de modo que, usando (4.4), (4.8) se puede escribir como

$$\begin{aligned} p(\mathbf{x} | \mathbf{y}) &= K p(x[n]) \sum_{i=1}^{P(n)} \Pr(\mathbf{s}_{0:n-1}^i) \prod_{k=0}^N \delta\left(x[k] - f_{\mathbf{s}_{k:n-1}^i}^{k-n}(x[n])\right) \\ &\times \exp\left(-\frac{1}{2\sigma^2} \sum_{k=0}^N \left(y[k] - f_{\mathbf{s}_{k:n-1}^i}^{k-n}(x[n])\right)^2\right), \end{aligned} \quad (4.9)$$

donde  $K$  es una constante que engloba todos aquellos términos que no dependen de la secuencia caótica que se desea estimar:

$$K = [p(\mathbf{y})(2\pi\sigma^2)^{(N+1)/2}]^{-1}.$$

Se puede observar que (4.9) depende fundamentalmente de la muestra de referencia utilizada, así como del itinerario de la señal hasta la muestra de referencia. Esto es, conocidos  $x[n]$  y  $\mathbf{s}_{0:n-1}$  el resto de la secuencia se encuentra determinada de manera inequívoca. Por lo tanto, el problema de nuevo consiste en estimar  $x[n]$  y  $\mathbf{s}_{0:n-1}$ , y se puede resolver de un modo similar al seguido en el Capítulo 3 para el estimador ML: calculando el estimador Bayesiano deseado para cada posible secuencia simbólica de longitud  $n$ , y seleccionando el mejor de todos ellos a posteriori (estimador MAP) o promediándolos (estimador MS). La principal diferencia entre el estimador ML y los estimadores Bayesianos se encuentra en la modificación del peso asociado a cada itinerario debido a la inclusión de la información a priori disponible.

### 4.2.5. Estimadores MAP y MS

En este punto ya se pueden plantear las ecuaciones genéricas de los dos estimadores Bayesianos que se van a considerar: el MAP y el MS o MMSE. El estimador MAP de  $x[n]$ , para una determinada secuencia simbólica,  $\mathbf{s}_{0:n-1}^i$ , se obtiene buscando aquel valor de  $x[n]$  en el que la FDP condicional de la secuencia, dada por (4.9), sea máxima [VanTre1968, Kay1993]:

$$\hat{x}_{\text{MAP}}^i[n] = \arg \max_{x[n]} \left\{ p(x[n]) \exp \left( -\frac{1}{2\sigma^2} \sum_{k=0}^N \left( y[k] - f_{\mathbf{s}_{k:n-1}^i}^{k-n}(x[n]) \right)^2 \right) \right\}.$$

Donde se han eliminado todos aquellos términos que no dependen de  $x[n]$ . Y el estimador MAP global de  $x[n]$  es aquel de entre los  $P(n)$  estimadores MAP asociados a los distintos itinerarios (estimadores MAP locales) que maximiza la FDP a posteriori. Es decir,

$$\hat{x}_{\text{MAP}}[n] = \hat{x}_{\text{MAP}}^r[n],$$

siendo

$$r = \arg \max_{i=1, \dots, P(n)} \left\{ \Pr(\mathbf{s}_{0:n-1}^i) p(\hat{x}_{\text{MAP}}^i[n]) \exp \left( -\frac{1}{2\sigma^2} \sum_{k=0}^N \left( y[k] - f_{\mathbf{s}_{k:n-1}^i}^{k-n}(\hat{x}_{\text{MAP}}^i[n]) \right)^2 \right) \right\}.$$

La estima MAP para la secuencia completa se obtiene iterando a partir de  $\hat{x}_{\text{MAP}}[n]$ : hacia delante utilizando la ecuación conocida del mapa caótico, y hacia atrás utilizando el mapa caótico y la secuencia simbólica que produce el máximo de la FDP,  $\hat{\mathbf{s}}_{\text{MAP}} = [\hat{s}_{\text{MAP}}[0], \dots, \hat{s}_{\text{MAP}}[n-1]]^T$ .

El estimador MS se obtiene simplemente calculando la esperanza matemática de (4.9) [VanTre1968, Kay1993]. Por consiguiente, el estimador MS de  $x[n]$  para una determinada secuencia simbólica,  $\mathbf{s}_{0:n-1}^i$ , viene dado por

$$\hat{x}_{\text{MS}}^i[n] = K \int_{\mathbf{X}} p(x[n]) \exp \left( -\frac{1}{2\sigma^2} \sum_{k=0}^N \left( y[k] - f_{\mathbf{s}_{k:n-1}^i}^{k-n}(x[n]) \right)^2 \right) dx[n].$$

Y el estimador MS de  $x[n]$  es el promedio de los estimadores MS asociados a cada secuencia particular ponderados de acuerdo con la probabilidad de dicha secuencia:

$$\hat{x}_{\text{MS}}[n] = \sum_{i=1}^{P(n)} \Pr(\mathbf{s}_{0:n-1}^i) \hat{x}_{\text{MS}}^i[n].$$

La estima del resto de la señal caótica se debería obtener promediando igualmente las secuencias obtenidas iterando a partir de  $\hat{x}_{\text{MS}}^i[n]$  para cada posible secuencia simbólica. Esto es, la estima MS  $i$ -ésima de  $x[n+k]$  vendría dada por

$$\hat{x}_{\text{MS}}^i[n+k] = \begin{cases} f^k(\hat{x}_{\text{MS}}^i[n]), & k \geq 0; \\ f_{\mathbf{s}_{n-k:n-1}^i}^{-|k|}(\hat{x}_{\text{MS}}^i[n]), & k < 0. \end{cases}$$



Y el estimador MS de  $x[n+k]$  sería nuevamente el promedio de las  $P(n)$  estimas locales:

$$\hat{x}_{\text{MS}}[n+k] = \sum_{i=1}^{P(n)} \Pr(\mathbf{s}_{0:n-1}^i) \hat{x}_{\text{MS}}^i[n+k].$$

El problema de esta aproximación es que dicha secuencia en general no va a ser una secuencia válida del mapa caótico, y el estimador MS en realidad vendría dado por la mejor proyección de la estima obtenida sobre el espacio de todas las secuencias válidas, que puede resultar muy difícil de encontrar.

No obstante, aunque estas expresiones resultan complicadas, existe un caso particular en el que las ecuaciones de los estimadores MAP y MS se simplifican mucho: la utilización de  $x[0]$  como muestra de referencia de la secuencia. En este caso no existen diferentes combinaciones de signos que deban explorarse, y el estimador MAP resulta:

$$\hat{x}_{\text{MAP}}[0] = \arg \max_{x[0]} \left\{ p(x[0]) \exp \left( -\frac{1}{2\sigma^2} \sum_{k=0}^N (y[k] - f^k(x[0]))^2 \right) \right\}.$$

Respecto al estimador MS, este viene dado por:

$$\hat{x}_{\text{MS}}[0] = K \int_X p(x[0]) \exp \left( -\frac{1}{2\sigma^2} \sum_{k=0}^N (y[k] - f^k(x[0]))^2 \right) dx[0].$$

Nótese que, en este caso, a partir de la estima MS de  $x[0]$  únicamente existe una secuencia caótica válida, que se obtiene simplemente iterando a partir de  $\hat{x}_{\text{MS}}[0]$ , ya que cualquier otro valor de  $x[k]$  distinto de los obtenidos mediante la iteración de  $x[0]$  tiene probabilidad cero.

En las secciones posteriores se consideran los estimadores Bayesianos para diversas familias de mapas caóticos. En concreto, en la Sección 4.3 se analizan los estimadores MAP y MS de secuencias generadas iterando mapas PWL, para los que se pueden encontrar ecuaciones explícitas de los estimadores. Posteriormente, en la Sección 4.4 se intenta extender dicho estudio a mapas no PWL genéricos, para los que no se pueden plantear fórmulas cerradas de los estimadores, pero se puede resolver el problema recurriendo a los métodos MCMC.

### 4.3. Estimadores Bayesianos de Secuencias Generadas por Mapas PWL

En esta sección se muestra la formulación detallada para los estimadores Bayesianos de secuencias generadas por mapas PWL. Al igual que sucede en el caso del estimador ML, el punto más controvertido es la elección de la muestra de referencia,  $x[n]$ . Aunque los estimadores Bayesianos no poseen en general la propiedad de invariancia [Kay1993], se ha observado el mismo comportamiento que para el estimador ML: su rendimiento

es esencialmente independiente de la posición de la muestra de referencia. Este aspecto se discute mediante un ejemplo en la Sección 4.3.2.3 para el estimador MAP. En el resto de los apartados se asume que la muestra de referencia es  $x[0]$  para simplificar la formulación. En la Sección 4.3.1 se propone una aproximación PWC para la FDP a priori de un mapa PWL y se obtiene una expresión cerrada para la FDP a posteriori de secuencias generadas por esta clase de mapas. En la Sección 4.3.2 se desarrolla el estimador MAP, tanto en el caso en que la secuencia simbólica es conocida como en el que es desconocida. Por último, en la Sección 4.3.3 se sigue un desarrollo paralelo para el estimador MS.

### 4.3.1. FDPs A Priori y A Posteriori

En esta sección se van a desarrollar expresiones explícitas para las FDPs a priori, condicional y a posteriori de mapas PWL. Puesto que se utiliza siempre como muestra de referencia  $x[0]$ , se va a simplificar la notación del Capítulo 3, eliminando los superíndices en los que aparecen  $n$  y  $N$ , y sustituyendo  $\mathbf{s}_i$  por  $i$ . Así, la región del espacio de fases que engloba las condiciones iniciales que generan una determinada secuencia simbólica de longitud  $N$ ,  $R_{\mathbf{s}_i}^0 = [\eta_{\mathbf{s}_i}^0, \kappa_{\mathbf{s}_i}^0]$ , ahora se va a denotar como  $R_i = [\eta_i, \kappa_i]$ . Del mismo modo, la función característica de dicha región,  $\chi_{R_{\mathbf{s}_i}^0}(x)$ , ahora se denota mediante  $\chi_i(x)$ . Por último, resulta conveniente definir la anchura de la región  $i$ -ésima,  $\Delta_i = \kappa_i - \eta_i$ , y su punto central,  $\rho_i = (\eta_i + \kappa_i)/2$ .

#### 4.3.1.1. FDP A Priori de Mapas PWL

En la Sección 4.2 se ha obtenido una expresión cerrada para la FDP de la secuencia caótica dadas las observaciones,  $p(\mathbf{x}|\mathbf{y})$ . Sin embargo, dicha ecuación depende de la FDP a priori de la secuencia caótica, que va a estar indefectiblemente dada por la FDP invariante natural del mapa caótico,  $p(x)$ , y para la cual no siempre se puede encontrar una fórmula analítica. En el Apéndice C se demuestra que muchos de los mapas usados en la Tesis presentan una FDP invariante natural uniforme en todo su espacio de fases: los mapas de Bernoulli, el SK-TM, el BSK-TM, o el TM y el S-TM para  $\beta = 2$ . Además, es bien conocido que la FDP invariante natural de los mapas PWL de Markov es constante a tramos (PWC) con intervalos que coinciden con las regiones de su partición natural [Isabel1997, Isabel1998]:

$$p(x) = \sum_{i=1}^M p_{E_i} \chi_{E_i}(x[n]), \quad (4.10)$$

siendo  $M$  el número de intervalos del mapa y  $p_{E_i}$  la densidad de probabilidad en cada intervalo, cuyo valor se encuentra en el Apéndice C para el mapa de Markov con cuatro intervalos de la Figura 2.4, y para el mapa con dos intervalos de la Figura 2.5(a).

Desafortunadamente, existen muchos otros mapas importantes cuya FDP invariante natural es desconocida, y para los que no puede asumirse que sea PWC: los mapas no

markovianos. Ejemplos de esta clase de mapas son el TM y el S-TM para  $\beta \neq 2$ , o el mapa de la Figura 2.5(b). En estos casos, así como en aquellos en los que la FDP invariante sea conocida pero resulte excesivamente complicada, se propone una aproximación constante a tramos (PWC) de la FDP invariante natural dentro de cada una de las regiones  $R_i$  [Panta2000a, Panta2003]:

$$p(x) \simeq \sum_{i=1}^{P(N)} p_i \chi_i(x). \quad (4.11)$$

Las constantes que aparecen en cada región,  $p_i$ , están relacionadas con la probabilidad total de que una muestra pertenezca a dicha región, y puesto que

$$\Pr(x \in R_i) = \int_{R_i} p(x) dx = p_i \Delta_i,$$

entonces su valor viene dado por

$$p_i = \frac{\Pr(x \in R_i)}{\Delta_i} = \frac{1}{\Delta_i} \int_{R_i} p(x) dx. \quad (4.12)$$

Nótese que (4.11) es una generalización de (4.10) en la que aparecen  $P(N)$  intervalos en lugar de los  $M$  correspondientes a la partición natural, y que sigue siendo válida para un mapa de Markov tomando simplemente

$$p_i = \prod_{n=0}^{N-1} p_{E_s[n]}.$$

Así pues, a modo de resumen, a partir de este momento se va a trabajar en general con FDPs invariantes naturales PWC. En el caso de mapas PWL de Markov esta va a ser su verdadera FDP invariante natural, existiendo únicamente  $M$  valores de  $p_i$  distintos como mucho. Sin embargo, para otra clase de mapas se va a tratar únicamente de una aproximación conveniente, ya que va a facilitar en gran medida la resolución analítica del problema considerado. Téngase en cuenta que en este caso pueden aparecer hasta  $P(N)$  valores distintos de densidad que deben ser obtenidos para cada valor de  $N$ . Por último, nótese que este es un modelo razonable de la FDP invariante, puesto que conserva la probabilidad de la FDP real para cada una de las regiones en que se puede subdividir el espacio de fases del mapa en función del itinerario.

En un caso práctico, en el que no se conoce  $p(x)$  o pueda resultar difícil de integrar, estas constantes se pueden calcular a partir de largas secuencias de datos. Debido a la ergodicidad de los mapas caóticos utilizados, la forma más sencilla de encontrar estas constantes consiste en generar una muestra inicial aleatoria, iterar un cierto número de veces para garantizar que se llega a una secuencia típica, y posteriormente generar  $N_T$  muestras adicionales y construir su histograma de acuerdo a los límites conocidos de  $R_i$ , obteniéndose la constante  $p_i$  como

$$p_i = \frac{\Pr(x \in R_i)}{\Delta_i} = \frac{N_i}{\Delta_i N_T},$$

donde  $N_T$  es el número total de muestras del histograma, y  $N_i$  es el número de muestras pertenecientes a su intervalo  $i$ -ésimo, con  $i = 1, \dots, P(N)$ .

Como ejemplo de la estimación práctica de la FDP, en la Figura 4.1 se muestra la FDP invariante natural obtenida mediante este procedimiento para dos mapas caóticos diferentes: el mapa markoviano de la Figura 2.4 (Markov1), y el mapa no markoviano de la Figura 2.5(b). En ambos casos se ha generado una condición inicial aleatoria de acuerdo con una FDP uniforme en todo el espacio de fases, se ha iterado el mapa 1000 veces para evitar posibles periodos transitorios iniciales, y se ha tomado un millón de muestras posteriormente para construir el histograma usando como intervalos las regiones de la partición obtenida en función del itinerario para  $N = 5$ .

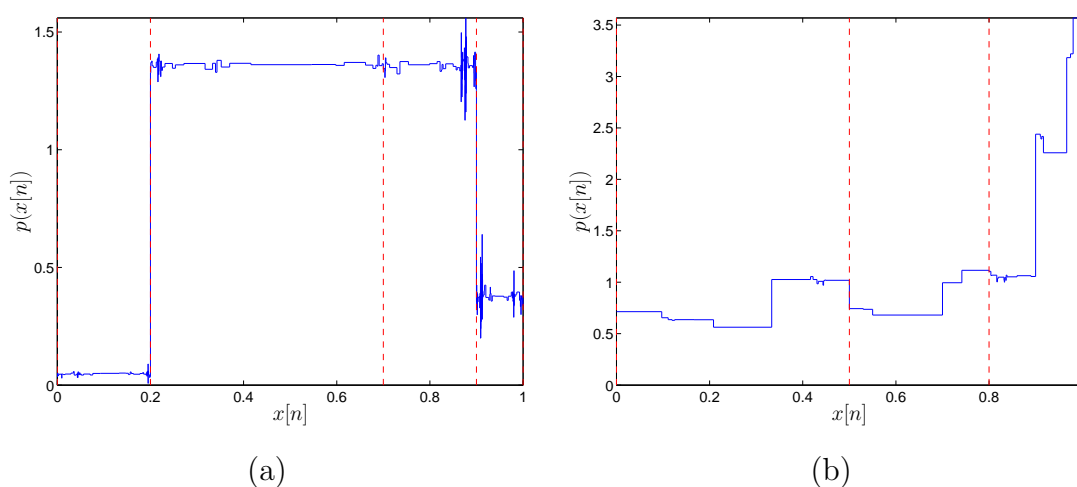


Figura 4.1: Comparación de la FDP invariante para un mapa de Markov y un mapa no markoviano. (a) FDP invariante del mapa Markov1 de la Figura 2.4. (b) FDP invariante del mapa PWL1 de la Figura 2.5(b).

Se puede apreciar cómo en el primer caso a pesar de existir un total de 272 regiones (véase la Tabla 3.8), la FDP de  $x[n]$  se estabiliza alrededor de tan sólo 3 valores distintos (las fluctuaciones de la figura se deben a malas estimas obtenidas dentro de regiones muy estrechas a causa de la utilización de un número insuficiente de muestras) en función de la región de la partición natural a la que pertenece  $x[n]$  (las regiones 2 y 3 de este mapa presentan el mismo valor de  $p_i$ , como se muestra en el Apéndice C). Por otro lado, en el segundo caso el número de regiones es mucho más reducido (sólo 45), pero se puede distinguir claramente la variación de la FDP entre cada par de regiones consecutivas.

Para finalizar, un breve apunte en relación con el número de regiones definidas por la partición. Puesto que la muestra de referencia en este caso siempre va a ser  $x[0]$ , la partición, obtenida del modo indicado en la Sección 3.4.3.2, va a estar determinada exclusivamente por el itinerario posterior,  $\mathbf{s}_{0:N-1}$ . En dicha sección se discute igualmente el crecimiento del número de regiones,  $P(N)$ , en función de  $N$ , aunque sin tener en cuenta la dimensión del atractor, que puede ocupar únicamente una porción del espacio

de fases. Esta situación aparece por ejemplo para el TM y el S-TM con  $\beta < 2$ , dando lugar a regiones del espacio de fases con probabilidad nula, que pueden descartarse a la hora de plantear los estimadores Bayesianos.

Teniendo esto en cuenta, el número de regiones típicas (esto es, el número de regiones dentro del atractor cuya probabilidad es no nula) para el TM se muestra en la Tabla 4.1 para diferentes valores de  $\beta$ . Comparando los valores de la Tabla 4.1 con los de la Tabla 3.6 puede apreciarse que el porcentaje de regiones típicas disminuye conforme  $N$  aumenta o  $\beta$  disminuye. Por ejemplo, para  $\beta = 1,5$  el 75 % de las regiones son típicas para  $N = 2$ , disminuyendo este porcentaje hasta el 45,83 % para  $N = 5$  y hasta el 39,35 % para  $N = 10$ . En cambio, para  $\beta = 1,8$  todas de las regiones son típicas para  $N = 2$ , el 80 % para  $N = 5$  y el 72,77 % para  $N = 10$ . Por último, cuando  $\beta = 2$  el atractor es el espacio de fases completo, de modo que todas las regiones son típicas.

$N$	$P(N)$									
	1	2	3	4	5	6	7	8	9	10
$\beta = 1,5$	2	3	5	7	11	16	25	37	57	85
$\beta = 1,6$	2	3	5	8	13	21	34	55	88	141
$\beta = 1,7$	2	4	7	12	21	36	61	105	179	304
$\beta = 1,8$	2	4	7	13	24	43	78	141	253	457
$\beta = 1,9$	2	4	8	15	29	55	105	199	379	720
$\beta = 2,0$	2	4	8	16	32	64	128	256	512	1.024

Tabla 4.1: Número de regiones típicas de la partición del espacio de fases del TM,  $P(N)$ , en función del itinerario para distintos valores de  $\beta$  y  $N$ .

#### 4.3.1.2. FDP A Posteriori cuando el Itinerario es Conocido

Cuando la muestra de referencia es  $x[0]$ , y la secuencia simbólica asociada a la señal caótica es conocida, (4.9) se puede poner como

$$\begin{aligned}
 p(\mathbf{x}|\mathbf{y}, \mathbf{s}_i) &= K_i p(x[0]) \chi_i(x[0]) \exp\left(-\frac{1}{2\sigma^2} \sum_{k=0}^N (y[k] - f_{\mathbf{s}_i}^k(x[0]))^2\right) \\
 &\times \prod_{k=1}^N \delta(x[k] - f_{\mathbf{s}_i}^k(x[0])). \tag{4.13}
 \end{aligned}$$

Nótese que (4.13) se puede factorizar en tres términos: una constante de normalización que no depende de la señal caótica que se desea estimar,  $K_i$ ; un término que engloba la dependencia con la muestra de referencia,  $x[0]$ ,  $p(x[0]) \chi_i(x[0]) \exp(-J(x[0], \mathbf{s}_i)/2\sigma^2)$ ; y un producto de términos que expresan la dependencia con cada muestra del resto de la secuencia,  $x[k]$ ,  $\delta(x[k] - f_{\mathbf{s}_i}^k(x[0]))$ , para  $k = 1, \dots, N$ . Nótese además que se ha incluido la dependencia explícita con la secuencia simbólica en la notación de la FDP, y que el itinerario  $\mathbf{s}_i$ , debe ser uno de los  $P(N)$  válidos para el mapa en cuestión. Por

último, la función característica que aparece en (4.13) indica que  $x[0]$  debe pertenecer a  $R_i$ , ya que de lo contrario el itinerario de la señal caótica generada no coincidiría con el dado a priori.

Siguiendo la metodología Bayesiana habitual, se puede integrar (4.13) para obtener la FDP marginal de  $x[0]$ , aislando de este modo el parámetro de interés, puesto que el resto de la secuencia caótica se encuentra especificada de manera única por el valor de la condición inicial. Esta FDP marginal resulta

$$\begin{aligned} p(x[0]|\mathbf{y}, \mathbf{s}_i) &= \int_{X_{1:N}} p(x[0], \mathbf{x}_{1:N}|\mathbf{y}, \mathbf{s}_i) d\mathbf{x}_{1:N} \\ &= K_i p(x[0]) \chi_i(x[0]) \exp\left(-\frac{1}{2\sigma^2} \sum_{k=0}^N (y[k] - f_{\mathbf{s}_i}^k(x[0]))^2\right). \end{aligned} \quad (4.14)$$

Tanto (4.13) como (4.14) son expresiones perfectamente válidas para cualquier mapa caótico, no sólo para mapas PWL. Recordando ahora la forma de  $f_{\mathbf{s}}^k(x[0])$  para un mapa PWL, que viene dada por (3.15),

$$f_{\mathbf{s}_i}^k(x[0]) = A_{\mathbf{s}_i}^{0,k} x[0] + \sum_{m=0}^{k-1} A_{\mathbf{s}_i}^{m+1,k} b_{\mathbf{s}_i}[m],$$

la ecuación (4.14) se puede particularizar para esta clase de mapas, obteniéndose

$$p(x[0]|\mathbf{y}, \mathbf{s}_i) = K_i p(x[0]) \chi_i(x[0]) \exp\left(-\frac{1}{2\sigma^2} J(x[0], \mathbf{s}_i)\right), \quad (4.15)$$

donde  $J(x[0], \mathbf{s}_i)$  es la función de coste cuadrática en  $x[0]$  que aparecía en el cálculo del estimador ML en el Capítulo 3, dada por (3.31) para  $n = 0$ :

$$J(x[0], \mathbf{s}_i) = \sum_{k=0}^N (\gamma_i[k] - \alpha_i[k] x[0])^2, \quad (4.16)$$

con  $\alpha_i[k]$  y  $\gamma_i[k]$  dadas por (3.27) y (3.32) respectivamente, ambas de nuevo para  $n = 0$ .

La única diferencia entre (4.15), cuya maximización proporciona el estimador MAP de  $x[0]$ , y la función de verosimilitud,  $p(\mathbf{y}; x[0], \mathbf{s}_i)$ , usada para obtener el estimador ML en el Capítulo 3, es la inclusión de la FDP a priori de  $x[0]$ . No obstante, esto provoca que la FDP ya no se pueda maximizar simplemente minimizando  $J(x[0], \mathbf{s}_i)$ . Afortunadamente, cuando se utiliza una FDP a priori PWC, (4.16) no es más que una versión normalizada de  $p(\mathbf{y}; x[0], \mathbf{s}_i)$ . En estas circunstancias, resulta sencillo demostrar que (4.15) es una Gaussiana truncada y escalada, que se puede expresar de manera alternativa como

$$p(x[0]|\mathbf{y}, \mathbf{s}_i) = K_i q_i \chi_i(x[0]) \exp\left(-\frac{(x[0] - \mu_i)^2}{2\sigma_i^2}\right). \quad (4.17)$$

En la ecuación (4.17), los nuevos parámetros que aparecen ( $\mu_i$ ,  $\sigma_i^2$  y  $q_i$ ) son el resultado de desarrollar (4.16) y completar cuadrados. De este modo, la media de la Gaussiana sin truncamiento resulta

$$\mu_i = \frac{\sum_{k=0}^N \alpha_i[k] \gamma_i[k]}{\sum_{k=0}^N (\alpha_i[k])^2}, \quad (4.18)$$

su varianza, de nuevo en ausencia de truncamiento, viene dada por

$$\sigma_i^2 = \frac{\sigma^2}{\sum_{k=0}^N (\alpha_i[k])^2}, \quad (4.19)$$

y el factor  $q_i$ , que proviene del término independiente de  $x[0]$  que falta para completar un cuadrado perfecto, toma el valor

$$q_i = p_i \exp\left(-\frac{c_i^2 - \mu_i^2}{2\sigma_i^2}\right), \quad (4.20)$$

siendo  $p_i$  el valor de la FDP invariante dentro de la región  $i$ -ésima, y

$$c_i^2 = \frac{\sum_{k=0}^N (\gamma_i[k])^2}{\sum_{k=0}^N (\alpha_i[k])^2}. \quad (4.21)$$

La ecuación de  $q_i$  se puede expresar de un modo más interesante. Desarrollando la expresión de la función de coste dada por (4.16), es fácil ver que

$$J(x[0], \mathbf{s}_i) = (x[0]^2 - 2\mu_i x[0] + c_i^2) \sum_{k=0}^N (\alpha_i[k])^2. \quad (4.22)$$

Y particularizando esta expresión en  $x[0] = \mu_i$  se llega a la conclusión de que

$$J(\mu_i, \mathbf{s}_i) = (c_i^2 - \mu_i^2) \sum_{k=0}^N (\alpha_i[k])^2 = \frac{(c_i^2 - \mu_i^2)\sigma^2}{\sigma_i^2}. \quad (4.23)$$

Por último, sustituyendo (4.23) en (4.20) se obtiene la ecuación definitiva para el factor de escala de la Gaussiana truncada,  $q_i$ :

$$q_i = p_i \exp\left(-\frac{J(\mu_i, \mathbf{s}_i)}{2\sigma^2}\right). \quad (4.24)$$

En este caso, puesto que el valor de  $q_i$  es independiente del de  $x[0]$ , se trata únicamente de un factor de escala que se puede agrupar junto con la constante  $K_i$ . El valor de  $K_i \cdot q_i$

puede hallarse fácilmente imponiendo la condición de que la integral de  $p(x[0]|\mathbf{y}, \mathbf{s}_i)$  sea uno a lo largo de todo su espacio de fases. Puesto que dicha FDP es distinta de cero únicamente dentro de  $R_i$  (esto es, entre  $\eta_i$  y  $\kappa_i$ ),  $K_i \cdot q_i$  se puede despejar, obteniéndose una expresión cerrada para este término de normalización:

$$K_i q_i = \sqrt{\frac{2}{\pi \sigma_i^2}} \left[ \operatorname{erf} \left( \frac{\kappa_i - \mu_i}{\sqrt{2\sigma_i^2}} \right) - \operatorname{erf} \left( \frac{\eta_i - \mu_i}{\sqrt{2\sigma_i^2}} \right) \right]^{-1}, \quad (4.25)$$

donde  $\operatorname{erf}(\cdot)$  es la conocida *función de error*, que se define como [Abramo1965]

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp(-t^2) dt, \quad (4.26)$$

#### 4.3.1.3. FDP A Posteriori cuando el Itinerario es Desconocido

Cuando la secuencia simbólica asociada a la señal caótica es desconocida, entonces la FDP a posteriori se puede expresar como

$$\begin{aligned} p(\mathbf{x}, \mathbf{s}|\mathbf{y}) &= K p(x[0]) \exp \left( -\frac{1}{2\sigma^2} \sum_{k=0}^N (y[k] - f_{\mathbf{s}}^k(x[0]))^2 \right) \\ &\times \prod_{k=1}^N \delta(x[k] - f_{\mathbf{s}}^k(x[0])), \end{aligned} \quad (4.27)$$

Nótese que, aunque (4.27) es muy similar a (4.13), existen dos diferencias importantes. En primer lugar, ahora  $\mathbf{s}$  no se corresponde con una secuencia simbólica concreta,  $\mathbf{s}_i$ , sino que es un vector aleatorio desconocido que se desea estimar, y puede tomar un valor cualquiera de los  $P(N)$  posibles, lo que se indica mediante la inclusión de  $\mathbf{s}$  entre las variables condicionadas en lugar de entre los condicionantes, como en (4.13). Y en segundo lugar, ya no aparece el factor  $\chi_i(x[0])$  en la FDP, puesto que en este caso  $x[0]$  puede pertenecer a cualquier región del espacio de fases del mapa caótico, y la limitación ya la impone la FDP a priori, que vale cero fuera del rango del atractor del mapa PWL,  $A \subseteq I = [e_0, e_M]$ .

De nuevo, integrando (4.27) se obtiene la marginal, en este caso de la condición inicial y el itinerario, cuyo valor es

$$\begin{aligned} p(x[0], \mathbf{s}|\mathbf{y}) &= \int_{X_{1:N}} p(x[0], \mathbf{x}_{1:N}, \mathbf{s}|\mathbf{y}) d\mathbf{x}_{1:N} \\ &= K p(x[0]) \exp \left( -\frac{1}{2\sigma^2} \sum_{k=0}^N (y[k] - f_{\mathbf{s}}^k(x[0]))^2 \right) \\ &= K p(x[0]) \exp \left( -\frac{1}{2\sigma^2} J(x[0], \mathbf{s}) \right). \end{aligned} \quad (4.28)$$

Donde  $J(x[0], \mathbf{s})$  está definida por (4.16), al igual que en el caso en que el itinerario es conocido. La diferencia es que, cuando la secuencia simbólica es conocida, esta delimita



una región del espacio de fases, y en consecuencia el rango de valores que puede tomar  $x[0]$ . Cuando la secuencia simbólica es desconocida, entonces  $x[0]$  puede moverse a lo largo de todo su atractor, siendo la posición dentro del mismo la que determina el valor del itinerario, asociado de manera única a cada valor de  $x[0]$ .

La ecuación (4.28) de la FDP marginal no es adecuada para desarrollar los estimadores Bayesianos. Nos interesa disponer de una FDP que se pueda expresar como el sumatorio de  $P(N)$  Gaussianas truncadas similares a la de (4.17), cada una asociada a uno de los  $P(N)$  itinerarios válidos. Para efectuar esta transformación de la FDP, se comienza por notar que [Panta2000a, Panta2000d]

$$J(x[0], \mathbf{s}) = \sum_{i=1}^{P(N)} J(x[0], \mathbf{s}_i) \chi_i(x[0]).$$

En consecuencia, (4.17) se puede poner como

$$\begin{aligned} p(x[0], \mathbf{s}|\mathbf{y}) &= Kp(x[0]) \exp \left( -\frac{1}{2\sigma^2} \sum_{i=1}^{P(N)} J(x[0], \mathbf{s}_i) \chi_i(x[0]) \right) \\ &= Kp(x[0]) \prod_{i=1}^{P(N)} \exp \left( -\frac{1}{2\sigma^2} J(x[0], \mathbf{s}_i) \chi_i(x[0]) \right) \\ &= Kp(x[0]) \sum_{i=1}^{P(N)} \chi_i(x[0]) \exp \left( -\frac{1}{2\sigma^2} J(x[0], \mathbf{s}_i) \right). \end{aligned} \quad (4.29)$$

Donde el último paso es posible gracias a que multiplicar por términos de valor unitario (esto es, aquellos términos para los que la función indicador vale cero, de modo que la exponencial toma valor uno) es equivalente a sumar términos de valor nulo. Finalmente, considerando un modelo PWC para la FDP invariante, (4.29) se transforma en

$$\begin{aligned} p(x[0], \mathbf{s}|\mathbf{y}) &= K \sum_{i=1}^{P(N)} p_i \chi_i(x[0]) \exp \left( -\frac{1}{2\sigma^2} J(x[0], \mathbf{s}_i) \right) \\ &= K \sum_{i=1}^{P(N)} q_i \chi_i(x[0]) \exp \left( -\frac{(x[0] - \mu_i)^2}{2\sigma_i^2} \right), \end{aligned} \quad (4.30)$$

que, a la vista de (4.17), se puede expresar como un sumatorio de las FDPs a posteriori que se tienen para cada una de las  $P(N)$  posibles secuencias simbólicas:

$$p(x[0], \mathbf{s}|\mathbf{y}) = K \sum_{i=1}^{P(N)} K_i^{-1} p(x[0]|\mathbf{y}, \mathbf{s}_i). \quad (4.31)$$

Igual que antes, el valor de la constante  $K$  se puede obtener sabiendo que la integral de la FDP a posteriori,  $p(x[0], \mathbf{s}|\mathbf{y})$ , a lo largo de todo el espacio en que se encuentra

definida, debe ser igual a la unidad. En este caso habría que llevar a cabo una integral en  $x[0]$  y un sumatorio en  $\mathbf{s}$  para todas las posibles secuencias simbólicas. Sin embargo, puesto que cada condición inicial tiene asociado un solo itinerario válido de longitud  $N$ , para cada valor de  $x[0]$  únicamente existe un término del sumatorio en  $\mathbf{s}$  distinto de cero, y en la práctica solamente es necesario realizar la integral, resultando

$$K = \sqrt{\frac{2}{\pi}} \left\{ \sum_{i=1}^{P(N)} \sigma_i q_i \left[ \operatorname{erf} \left( \frac{\kappa_i - \mu_i}{\sqrt{2\sigma_i^2}} \right) - \operatorname{erf} \left( \frac{\eta_i - \mu_i}{\sqrt{2\sigma_i^2}} \right) \right] \right\}^{-1}. \quad (4.32)$$

La ecuación (4.32) es similar a la expresión obtenida cuando la secuencia simbólica es conocida, con la particularidad de que ahora no se puede incluir el término  $q_i$  dentro de la constante, y que aparece un sumatorio ponderado a lo largo de todos los posibles itinerarios, ya que se desconoce el itinerario real. Notando la relación existente entre la función de coste,  $J(x[0], \mathbf{s}_i)$ , y las variables  $\mu_i$  y  $\sigma_i^2$ ,

$$\frac{J(x[0], \mathbf{s}_i) - J(\mu_i, \mathbf{s}_i)}{2\sigma^2} = \frac{(x[0] - \mu_i)^2}{2\sigma_i^2},$$

ya esbozada en la Sección 4.3.1.2 en las ecuaciones (4.22) y (4.23), esta constante de normalización se puede expresar alternativamente como

$$K = \sqrt{\frac{2}{\pi}} \left\{ \sum_{i=1}^{P(N)} \sigma_i q_i \left[ \operatorname{erf} \left( \sqrt{\frac{J(\kappa_i, \mathbf{s}_i) - J(\mu_i, \mathbf{s}_i)}{2\sigma^2}} \right) - \operatorname{erf} \left( \sqrt{\frac{J(\eta_i, \mathbf{s}_i) - J(\mu_i, \mathbf{s}_i)}{2\sigma^2}} \right) \right] \right\}^{-1}.$$

Obviamente, también se puede llegar a una expresión similar cuando la secuencia simbólica es conocida. El motivo de buscar expresiones cerradas para esta constante es que, aunque no resulta necesaria para plantear el estimador MAP, sí que va a ser imprescindible para la obtención del estimador MS, así como la de otros posibles estimadores Bayesianos genéricos que se pudieran desarrollar.

Por último, al igual que ocurre con la FDP,  $K$  se puede expresar en función de los  $P(N)$  términos de normalización de cada región,  $K_i$ , como

$$K = \left( \sum_{i=1}^{P(N)} K_i^{-1} \right)^{-1}, \quad (4.33)$$

y sustituyendo (4.33) en (4.31) la FDP a posteriori cuando el itinerario es desconocido resulta finalmente

$$p(x[0], \mathbf{s}|\mathbf{y}) = \frac{\sum_{i=1}^{P(N)} K_i^{-1} p(x[0]|\mathbf{y}, \mathbf{s}_i)}{\sum_{i=1}^{P(N)} K_i^{-1}}. \quad (4.34)$$

### 4.3.2. Estimador de Máxima Probabilidad A Posteriori (MAP)

En este apartado se desarrolla el estimador de máxima probabilidad a posteriori (MAP) de la secuencia caótica. En primer lugar, se supone que la secuencia simbólica es conocida y se llega a la conclusión de que el estimador es el mismo que el ML. A continuación se considera el caso en que la secuencia simbólica es desconocida, y se obtiene el estimador, que en este caso es diferente del ML siempre que la FDP a priori no sea uniforme en todo el espacio de fases.

#### 4.3.2.1. Secuencia Simbólica Conocida

En este caso, habiendo escogido como muestra de referencia  $x[0]$ , obtener el estimador MAP equivale simplemente a encontrar el valor que maximice  $p(x[0]|\mathbf{y}, \mathbf{s}_i)$ . Puesto que la FDP siempre toma valores mayores o iguales que cero y el logaritmo es una función continua y monótona estrictamente creciente, se puede transformar el problema en el equivalente de hallar el valor de  $x[0]$  que maximice

$$\ln p(x[0]|\mathbf{y}, \mathbf{s}_i) = \ln(K_i q_i) - \frac{(x[0] - \mu_i)^2}{2\sigma_i^2} + \ln \chi_i(x[0]). \quad (4.35)$$

Nótese que  $\chi_i(x[0]) = 1$  si  $x[0]$  se encuentra dentro de  $R_i$ , mientras que fuera de ella  $\chi_i(x[0]) = 0$ . Esto implica que  $\ln \chi_i(x[0]) = 0$  si  $x[0] \in R_i$  y  $\ln \chi_i(x[0]) = -\infty$  si  $x[0] \notin R_i$ , y en consecuencia la búsqueda del máximo sólo debe llevarse a cabo entre  $\eta_i$  y  $\kappa_i$ . Al tratarse  $\ln p(x[0]|\mathbf{y}, \mathbf{s}_i)$  de una función unimodal y continua (al menos dentro del rango de interés, la región  $i$ -ésima,  $R_i$ ) se podría derivar para hallar su único mínimo. No obstante, en este caso no va a resultar necesario, ya que evidentemente el máximo de (4.35) con respecto a  $x[0]$  viene dado por  $\hat{x}[0] = \mu_i$ , siempre que  $\mu_i$  se encuentre dentro de  $R_i$ , y por el límite más próximo de la región en caso contrario. Es decir, el estimador MAP de  $x[0]$  dentro de la región  $i$ -ésima es

$$\hat{x}_{\text{MAP}}^i[0] = \begin{cases} \eta_i, & \mu_i < \eta_i; \\ \mu_i, & \eta_i \leq \mu_i \leq \kappa_i; \\ \kappa_i, & \mu_i > \kappa_i. \end{cases} \quad (4.36)$$

Comparando (4.36) con el estimador ML de  $x[0]$  cuando el itinerario es conocido, dado por (3.34) para  $n = 0$ , resulta evidente que son idénticos. Es decir, cuando se conoce la secuencia simbólica asociada a la señal caótica no se obtiene ninguna mejora utilizando el estimador MAP. Por consiguiente, el estimador MAP de la secuencia completa en este caso viene dado por la Proposición 3.1, y se puede encontrar mediante el Algoritmo 3.3.

En la Figura 4.2 se muestran dos ejemplos de la FDP a posteriori de  $x[0]$  dadas las observaciones cuando el itinerario es conocido, para el TM con  $\beta = 2$ . En ambos casos, la parte representada con línea discontinua es la porción de la Gaussiana que se encuentra fuera de la región  $i$ -ésima, y que por lo tanto se ha truncado. En el

primer ejemplo el máximo de la Gaussiana se halla dentro de la región, siendo este en consecuencia la estima MAP de  $x[0]$ , mientras que en el segundo el máximo se encuentra a la derecha de la región, de modo que el estimador MAP es simplemente el punto de la región más cercano al máximo: en esta ocasión su límite superior,  $\kappa_i$ .

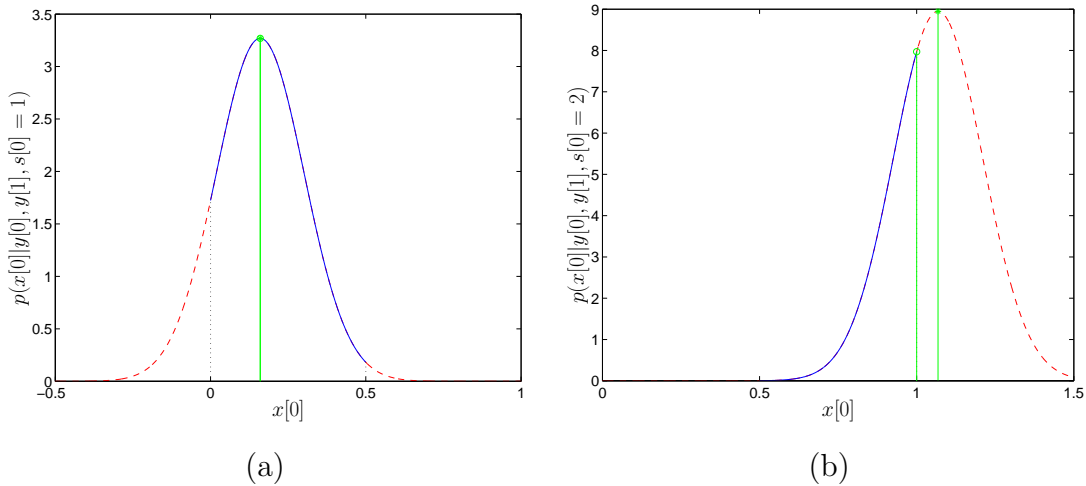


Figura 4.2: Ejemplos de FDP a posteriori de  $x[0]$  cuando el itinerario es conocido,  $p(x[0]|\mathbf{y}, \mathbf{s}_i)$ . En ambos casos se usa el TM con  $\beta = 2$  y  $N = 1$ , y se representa con línea discontinua la porción de la Gaussiana situada fuera de los límites de  $R_i$ . (a) El máximo de la Gaussiana se encuentra dentro de la región  $i$ -ésima ( $s[0] = 1$ ,  $y[0] = 0,6$ ,  $y[1] = 0,1$  y  $\hat{x}_{\text{MAP}}^1[0] = \hat{x}_{\text{ML}}^1[0] = \mu_1 = 0,16$ ). (b) El máximo de la Gaussiana se encuentra fuera de la región  $i$ -ésima ( $s[0] = 2$ ,  $y[0] = 0,4$ ,  $y[1] = -0,47$ ,  $\mu_2 = 1,068$  y  $\hat{x}_{\text{MAP}}^2[0] = \hat{x}_{\text{ML}}^2[0] = \kappa_2 = 1$ ).

#### 4.3.2.2. Secuencia Simbólica Desconocida

La situación es muy diferente cuando se desconoce la secuencia simbólica. En este caso, se debe obtener el máximo de  $p(x[0], \mathbf{s}|\mathbf{y})$ , no sólo con respecto a  $x[0]$ , sino también con respecto a  $\mathbf{s}$ . Asumiendo que la FDP a priori sigue el modelo PWC presentado en la Sección 4.3.1.1, la FDP a posteriori, dada por (4.30), se puede reescribir como

$$\begin{aligned}
 p(x[0], \mathbf{s}|\mathbf{y}) &= K \sum_{i=1}^{P(N)} q_i \chi_i(x[0]) \exp\left(-\frac{(x[0] - \mu_i)^2}{2\sigma_i^2}\right) \\
 &= K \sum_{i=1}^{P(N)} \chi_i(x[0]) \exp\left(\ln q_i - \frac{(x[0] - \mu_i)^2}{2\sigma_i^2}\right) \\
 &= K \chi_I(x[0]) \prod_{i=1}^{P(N)} \exp\left(\left[\ln q_i - \frac{(x[0] - \mu_i)^2}{2\sigma_i^2}\right] \chi_i(x[0])\right), \quad (4.37)
 \end{aligned}$$

donde se ha empleado el mismo truco que en (4.29) para convertir el productorio en un sumatorio, aunque a la inversa, y  $\chi_I(x[0])$  es la función indicadora de la pertenencia o no al espacio de fases del mapa de  $x[0]$ ,  $I = [e_0, e_M]$ , necesaria para garantizar que la FDP a posteriori sea nula fuera del mismo. Nuevamente, en lugar de tratar de maximizar (4.37) se va a utilizar su logaritmo neperiano, obteniéndose

$$\ln p(x[0], \mathbf{s}|\mathbf{y}) = \ln K + \ln \chi_I(x[0]) + \sum_{i=1}^{P(N)} \left( \ln q_i - \frac{(x[0] - \mu_i)^2}{2\sigma_i^2} \right) \chi_i(x[0]). \quad (4.38)$$

A la vista de (4.38), resulta evidente que dentro de cada región la FDP es continua y unimodal, ya que  $q_i$  no depende de  $x[0]$  para una región dada. En consecuencia, existen un total de  $P(N)$  máximos, que coinciden con los de  $p(x[0]|\mathbf{y}, \mathbf{s}_i)$  para cada una de las  $P(N)$  posibles secuencias simbólicas. Esto significa que el estimador MAP se obtendrá de manera similar al ML cuando el itinerario es desconocido: hallando el estimador óptimo para cada secuencia simbólica y seleccionando finalmente el más probable de entre todos ellos. La única diferencia en este caso es la presencia del factor  $q_i$ , que va a ponderar los estimadores de las diferentes regiones, dando más peso a aquellas en las que la densidad de probabilidad es mayor. El estimador MAP cuando el itinerario es desconocido se describe formalmente en la Proposición 4.1, y su implementación práctica se lleva a cabo mediante el Algoritmo 4.1, mostrado al final de esta sección.

**Proposición 4.1** *Sea una secuencia de muestras desconocida,  $\{x[k]\}_{k=0}^N$ , generada mediante la iteración, usando (3.11) o (3.12), de un mapa PWL con  $M$  intervalos definido por (2.14), y sea  $\{y[k]\}_{k=0}^N$  la secuencia de observaciones disponibles, obtenidas sumando ruido blanco Gaussiano a cada muestra  $x[k]$ . El estimador MAP de la primera muestra de la secuencia,  $x[0]$ , cuando la secuencia simbólica es desconocida, viene dado por*

$$\hat{x}_{\text{MAP}}[0] = \hat{x}_{\text{MAP}}^r[0]. \quad (4.39)$$

Donde  $\hat{x}_{\text{MAP}}^r[0]$  es el estimador MAP local dado por (4.36) para la región  $r$ -ésima, que es el que maximiza la FDP marginal de  $x[0]$  y  $\mathbf{s}$  condicionada por las observaciones, dada por (4.37), o su logaritmo, dado por (4.38), de entre todos los estimadores MAP posibles, y se obtiene mediante la siguiente ecuación:

$$\begin{aligned} r &= \arg \max_i q_i \exp \left( -\frac{(\hat{x}_{\text{MAP}}^i[0] - \mu_i)^2}{2\sigma_i^2} \right) \\ &= \arg \max_i \left\{ \ln q_i - \frac{(\hat{x}_{\text{MAP}}^i[0] - \mu_i)^2}{2\sigma_i^2} \right\}, \end{aligned} \quad (4.40)$$

con  $i = 1, \dots, P(N)$ . El resto de muestras de la secuencia caótica se obtienen iterando a partir de  $\hat{x}_{\text{MAP}}[0]$ :

$$\hat{x}_{\text{MAP}}[k] = f^k(\hat{x}_{\text{MAP}}[0]), \quad (4.41)$$

para  $k = 1, \dots, N$ . Y por consiguiente, la estima MAP del vector completo que contiene la secuencia caótica es:

$$\hat{\mathbf{x}}_{\text{MAP}} = [\hat{x}_{\text{MAP}}[0], f(\hat{x}_{\text{MAP}}[0]), f^2(\hat{x}_{\text{MAP}}[0]), \dots, f^N(\hat{x}_{\text{MAP}}[0])]^T.$$

Nótese que como consecuencia del proceso también se obtiene la estima MAP de la secuencia simbólica,  $\hat{\mathbf{s}}_{\text{MAP}} = \mathbf{s}_r$ .

Nótese que, utilizando la función de coste,  $J(x[0], \mathbf{s}_i)$ , en lugar de  $\mu_i$  y  $\sigma_i^2$ , el índice del estimador MAP también se puede expresar como

$$\begin{aligned} r &= \arg \max_i p_i \exp \left( -\frac{1}{2\sigma^2} J(\hat{x}_{\text{MAP}}^i[0], \mathbf{s}_i) \right) \\ &= \arg \max_i \left\{ \ln p_i - \frac{1}{2\sigma^2} J(\hat{x}_{\text{MAP}}^i[0], \mathbf{s}_i) \right\} \\ &= \arg \min_i \left\{ J(\hat{x}_{\text{MAP}}^i[0], \mathbf{s}_i) - 2\sigma^2 \ln p_i \right\}. \end{aligned} \quad (4.42)$$

Comparando la última expresión de (4.42) y (3.47) para  $n = 0$ , se aprecia que la única diferencia entre ambas es el término  $2\sigma^2 \ln p_i$ , que aparece restando a la función de coste del estimador en cada región. Este término puede ser positivo (si  $p_i > 1$ ) o negativo (si  $p_i < 1$ ), pero en cualquier caso, cuanto menor sea  $p_i$  mayor va a ser el coste del estimador MAP  $i$ -ésimo, ya que se trata de un término fijo e independiente de  $x[0]$  para cada región. En consecuencia, este término tiene el efecto deseado: penaliza a las regiones con baja densidad de puntos del atractor (esto es, las regiones poco probables), provocando que sea menos factible seleccionar el estimador correspondiente a las mismas. Obviamente, si la FDP es uniforme en todo el espacio de fases se puede prescindir de dicho término para calcular  $r$ , y de nuevo el estimador MAP coincide con el estimador ML incluso cuando el itinerario es desconocido.

Como ejemplo de la aplicación del estimador MAP, en la Figura 4.3 se muestran dos posibles FDPs a posteriori de  $x[0]$  dadas las observaciones cuando el itinerario es desconocido, así como los correspondientes estimadores MAP y ML. En el primer caso se utiliza el S-TM con  $\beta = 2$ , de modo que la FDP invariante del mapa es uniforme y los estimadores ML y MAP coinciden siempre. En el segundo caso se usa el S-TM con  $\beta = 1,5$ , y en consecuencia la FDP a priori de  $x[0]$  ya no es uniforme, y la estima MAP en promedio se halla más próxima al valor real de la condición inicial que la estima ML. Evidentemente, van a existir condiciones iniciales y vectores de observación para los que el estimador ML funcione mejor que el MAP, pero en promedio la inclusión de información a priori veraz va a mejorar en el rendimiento de los estimadores, como se muestra en la Sección 4.5. En la Figura 4.3 también se muestran los logaritmos de las FDPs a posteriori, con el fin de que se pueda apreciar claramente la distribución de probabilidad en las regiones donde aquella es menor.

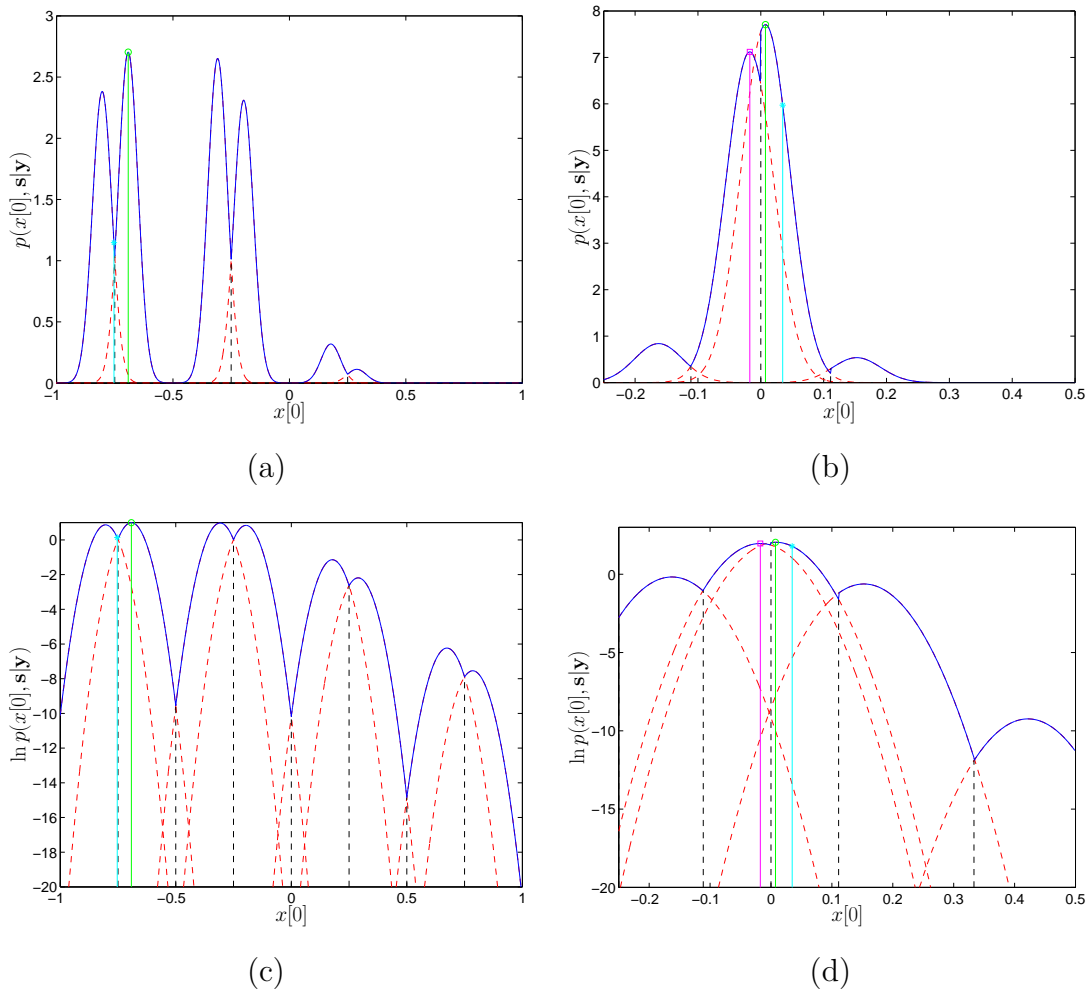


Figura 4.3: Ejemplos de FDP a posteriori de  $x[0]$  cuando el itinerario es desconocido,  $p(x[0], \mathbf{s}|\mathbf{y})$ , con una SNR = 10 dB. (a) Señal caótica generada con el S-TM y  $\beta = 2$ . La secuencia caótica es  $\mathbf{x} = [-0,7533, -0,5067, -0,0134, 0,9733]^T$ , y la secuencia de observaciones es  $\mathbf{y} = [-0,7867, 0,1396, -0,2686, 0,4063]^T$  ( $\sigma^2 = 0,1486$ ). En este caso el estimador ML y el MAP siempre coinciden:  $\hat{x}_{\text{MAP}}[0] = \hat{x}_{\text{ML}}[0] = \mu_2 = -0,6921$ . (b) Señal caótica generada con el S-TM y  $\beta = 1,5$ . La secuencia caótica es  $\mathbf{x} = [0,0349, 0,4477, -0,1716, 0,2427]^T$ , y la secuencia de observaciones es  $\mathbf{y} = [-0,0984, 0,3170, -0,1898, 0,0763]^T$  ( $\sigma^2 = 0,0290$ ). En este caso el estimador ML es  $\hat{x}_{\text{ML}}[0] = \mu_2 = -0,0175$ , mientras que el estimador MAP es  $\hat{x}_{\text{MAP}}[0] = \mu_3 = 0,0075$ . (c) Logaritmo de la FDP a posteriori mostrada en (a). (d) Logaritmo de la FDP a posteriori mostrada en (b).

Como ejemplo de la mejora introducida por el estimador MAP, en la Figura 4.4 se muestran la función de verosimilitud, utilizada por el estimador ML para determinar el valor óptimo de  $x[0]$ , y la FDP a posteriori con la inclusión de información a priori, para el mismo ejemplo que en la Figura 4.3(b). En este caso se aprecia claramente cómo

la información a priori incrementa la probabilidad de la tercera región en detrimento de la segunda, de modo que la estima MAP de  $x[0]$  se desplaza a la tercera región (desde la segunda en que se encuentra la estima ML), acercándose al valor real de  $x[0]$ .

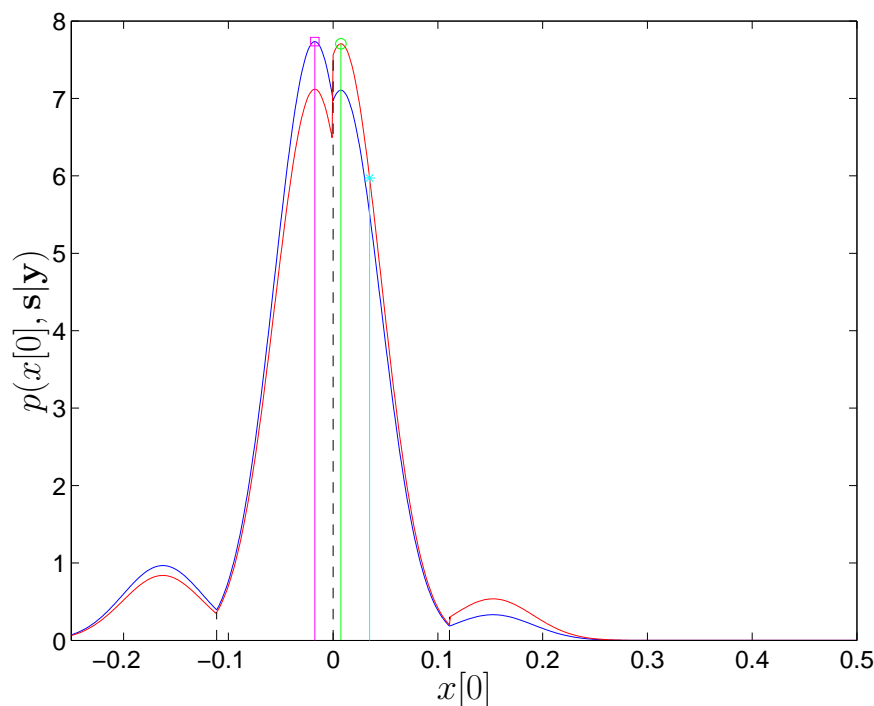


Figura 4.4: Comparación de la función de verosimilitud (en azul), utilizada para obtener el estimador ML, y la FDP a posteriori (en rojo) obtenida con la inclusión de información a priori, utilizada para obtener el estimador MAP. La señal caótica ha sido generada con el S-TM,  $\beta = 1,5$  y  $x[0] = 0,0349$ . La secuencia de observaciones es  $\mathbf{y} = [-0,0984, 0,3170, -0,1898, 0,0763]^T$  ( $\sigma^2 = 0,0290$ ). En este caso el estimador ML es  $\hat{x}_{\text{ML}}[0] = \mu_2 = -0,0175$ , mientras que el estimador MAP es  $\hat{x}_{\text{MAP}}[0] = \mu_3 = 0,0075$ .

#### 4.3.2.3. Invariancia del Estimador MAP con Respecto a la Muestra de Referencia

En todos los estimadores de la secuencia caótica planteados, la primera elección que se debe realizar es la de la muestra de referencia. En el Capítulo 3 se mostró mediante simulaciones que la muestra de referencia escogida no afecta al rendimiento del estimador ML, gracias a su propiedad de invariancia. En esta sección se va a mostrar, mediante un ejemplo sencillo, cómo esta muestra es esencialmente irrelevante también para el estimador MAP desarrollado en la Sección 4.3.2.

Considérese el caso más sencillo posible: un mapa caótico con dos intervalos ( $M = 2$ ) y una única iteración ( $N = 1$ ). Tomando como muestra de referencia  $x[0]$ , su estimador



1. Para cada secuencia simbólica válida,  $\mathbf{s}_i$  ( $i = 1, \dots, P(N)$ ), se repiten los pasos siguientes:
  - 1.1. Obtener el mínimo de la función de coste dada por (4.16) mediante (4.18).
  - 1.2. Hallar la región del espacio de fases,  $R_i = [\eta_i, \kappa_i]$ , formada por el conjunto de valores de  $x[0]$  que pueden generar el itinerario actual dado por  $\mathbf{s}_i$  aplicando el Algoritmo 3.5.
  - 1.3. Encontrar el estimador MAP de  $x[0]$  usando (4.36). Si  $x[0] \in R_i$ , entonces  $\hat{x}_{\text{MAP}}^i[0] = \mu_i$ . En caso contrario, el estimador MAP es el punto más cercano de  $R_i$ :  $\eta_i$  ó  $\kappa_i$ .
  - 1.4. Evaluar el error cometido por esta estima MAP local calculando el valor de la función de coste para la misma:

$$\xi_{\text{MAP}}^i = \frac{(\hat{x}_{\text{MAP}}^i[0] - \mu_i)^2}{2\sigma_i^2} - \ln q_i.$$

2. Una vez obtenidas todas las estimas MAP locales, seleccionar aquella cuyo error,  $\xi_{\text{MAP}}^i$ , sea menor,  $\hat{x}_{\text{MAP}}[0] = \hat{x}_{\text{MAP}}^r[0]$ , y encontrar el estimador MAP del resto de la secuencia iterando hacia delante a partir de  $\hat{x}_{\text{MAP}}[0]$ .

**Algoritmo 4.1:** Estimador MAP de una secuencia caótica cuando la secuencia simbólica es desconocida.

ML viene dado por el valor de  $x[0]$  que minimiza la siguiente función de coste:

$$J(x[0]) = \sum_{i=1}^M J_i(x[0]) \chi_{E_i}(x[0]),$$

donde  $E_i = [e_{i-1}, e_i]$ , como de costumbre, y

$$J_i(x[0]) = (y[0] - x[0])^2 + (y[1] - a_i x[0] - b_i)^2.$$

La estima local  $i$ -ésima se obtiene derivando  $J_i$  con respecto a  $x[0]$  e igualando a cero, y viene dada por

$$\hat{x}_i^0[0] = \frac{y[0] + a_i(y[1] - b_i)}{1 + a_i^2}, \quad (4.43)$$

donde el superíndice indica el punto de referencia, e  $i \in \{1, 2\}$ . A continuación, la estima MAP/ML local se obtiene simplemente limitando (4.43) para garantizar que dicha estima pertenece a  $E_i$ :

$$\hat{x}_{\text{ML}}^{0,i}[0] = \hat{x}_{\text{MAP}}^{0,i}[0] = \begin{cases} e_{i-1}, & \hat{x}_i^0[0] < e_{i-1}; \\ \hat{x}_i^0[0], & e_{i-1} \leq \hat{x}_i^0[0] \leq e_i; \\ e_i, & \hat{x}_i^0[0] > e_i. \end{cases}$$

Por último, el estimador ML global es aquel de los dos estimadores ML locales que

proporciona una menor función de coste,

$$\hat{x}_{\text{ML}}^0[0] = \begin{cases} \hat{x}_{\text{ML}}^{0,1}[0], & J_1(\hat{x}_{\text{ML}}^{0,1}[0]) \leq J_2(\hat{x}_{\text{ML}}^{0,2}[0]); \\ \hat{x}_{\text{ML}}^{0,2}[0], & J_2(\hat{x}_{\text{ML}}^{0,2}[0]) < J_1(\hat{x}_{\text{ML}}^{0,1}[0]); \end{cases}$$

mientras que para hallar el estimador MAP global hay que tener en cuenta la densidad de probabilidad dentro de cada región,  $p_1$  y  $p_2$ , obteniéndose

$$\hat{x}_{\text{MAP}}^0[0] = \begin{cases} \hat{x}_{\text{MAP}}^{0,1}[0], & \ln p_1 - J_1(\hat{x}_{\text{ML}}^{0,1}[0])/2\sigma^2 \leq \ln p_2 - J_2(\hat{x}_{\text{ML}}^{0,2}[0])/2\sigma^2; \\ \hat{x}_{\text{MAP}}^{0,2}[0], & \ln p_2 - J_2(\hat{x}_{\text{ML}}^{0,2}[0])/2\sigma^2 < \ln p_1 - J_1(\hat{x}_{\text{ML}}^{0,1}[0])/2\sigma^2. \end{cases}$$

Por otro lado, si se escoge como muestra de referencia  $x[1]$  el desarrollo a seguir resulta análogo. En esta ocasión la función de coste es

$$J(x[1]) = \sum_{i=1}^M J_i(x[1])\chi_{f(E_i)}(x[1]),$$

con

$$J_i(x[1]) = (y[0] - a_i^{-1}(x[1] - b_i))^2 + (y[1] - x[1])^2.$$

Y ahora el estimador ML/MAP local de  $x[1]$  resulta:

$$\hat{x}_i^1[1] = \frac{y[1] + a_i^{-1}y[0] + a_i^{-2}b_i}{1 + a_i^{-2}}, \quad (4.44)$$

siendo nuevamente necesario aplicar un umbral a (4.44) para obtener la estima ML/MAP dentro de cada región, y seleccionar la mejor de las dos de acuerdo con el criterio adecuado para obtener el estimador ML o MAP global.

Con el fin de comparar las estimas obtenidas con ambos puntos de referencia, veamos cuanto vale  $\hat{x}_i^0[1]$ . En primer lugar, asumiendo que  $\hat{x}_i^0[0]$  se encuentra dentro de  $E_i$  (esto es, que no resulta necesario recortarla), entonces la estima ML/MAP local  $i$ -ésima de  $x[1]$  sería

$$\hat{x}_{\text{ML}}^{0,i}[1] = \hat{x}_{\text{MAP}}^{0,i}[1] = f_i(\hat{x}_i^0[0]) = a_i\hat{x}_i^0[0] + b_i = \frac{a_i^2y[1] + a_iy[0] + b_i}{1 + a_i^2}. \quad (4.45)$$

Obviamente, (4.45) es igual que (4.44), de modo que es evidente que la muestra de referencia no influye en el estimador obtenido en este caso, y que  $\hat{x}_i^1[1] = \hat{x}_i^0[1] = f_i(\hat{x}_i^0[0]) = \hat{x}_{\text{ML}}^{1,i}[1] = \hat{x}_{\text{MAP}}^{1,i}[1]$ .

Por consiguiente, únicamente resta por considerar lo que ocurre cuando a  $\hat{x}_i^0[0]$  se le debe aplicar un umbral para obtener la estima ML/MAP local. En estas circunstancias existen dos posibilidades:

1.  $\hat{x}_i^0[0] > e_i$ : En este caso  $\hat{x}_{\text{MAP}}^{0,i}[0] = e_i$ , de modo que  $\hat{x}_{\text{MAP}}^{0,i}[1] = f_i(e_i) = a_ie_i + b_i$ . Y ahora existen dos posibilidades en función de  $a_i$ :

- 1.1.  $a_i > 0$ : En esta ocasión  $\hat{x}_i^1[1] = \hat{x}_i^0[1] = f_i(\hat{x}_i^0[0]) > f_i(e_i)$ , resultando evidente que  $\hat{x}_i^1[1] \notin R_i^1 = [f_i(e_{i-1}), f_i(e_i)]$ . En consecuencia,  $\hat{x}_{\text{MAP}}^{1,i}[1] = f_i(e_i) = \hat{x}_{\text{MAP}}^{0,i}[1]$ , lo que implica igualmente que  $\hat{x}_{\text{MAP}}^{1,i}[0] = \hat{x}_{\text{MAP}}^{0,i}[0]$ .
- 1.2.  $a_i < 0$ : Ahora  $\hat{x}_i^1[1] = \hat{x}_i^0[1] = f_i(\hat{x}_i^0[0]) < f_i(e_i)$ , de modo que nuevamente  $\hat{x}_i^1[1] \notin R_i^1 = [f_i(e_i), f_i(e_{i-1})]$ , y  $\hat{x}_{\text{MAP}}^{1,i}[1] = f_i(e_i) = \hat{x}_{\text{MAP}}^{0,i}[1]$ , conllevando que  $\hat{x}_{\text{MAP}}^{1,i}[0] = \hat{x}_{\text{MAP}}^{0,i}[0]$ .
2.  $\hat{x}_i^0[0] < e_{i-1}$ : En este caso  $\hat{x}_{\text{MAP}}^{0,i}[0] = e_{i-1}$ , de modo que  $\hat{x}_{\text{MAP}}^{0,i}[1] = f_i(e_{i-1}) = a_i e_{i-1} + b_i$ . Y de nuevo existen dos posibilidades en función de  $a_i$ :
  - 2.1.  $a_i > 0$ : En esta oportunidad  $\hat{x}_i^1[1] = \hat{x}_i^0[1] = f_i(\hat{x}_i^0[0]) < f_i(e_{i-1})$ , así que una vez más  $\hat{x}_i^1[1] \notin R_i^1 = [f_i(e_{i-1}), f_i(e_i)]$ , y  $\hat{x}_{\text{MAP}}^{1,i}[1] = f_i(e_{i-1}) = \hat{x}_{\text{MAP}}^{0,i}[1]$ , resultando  $\hat{x}_{\text{MAP}}^{1,i}[0] = \hat{x}_{\text{MAP}}^{0,i}[0]$ .
  - 2.2.  $a_i < 0$ : Ahora  $\hat{x}_i^1[1] = \hat{x}_i^0[1] = f_i(\hat{x}_i^0[0]) > f_i(e_{i-1})$ , y por consiguiente  $\hat{x}_i^1[1] \notin R_i^1 = [f_i(e_i), f_i(e_{i-1})]$ , y  $\hat{x}_{\text{MAP}}^{1,i}[1] = f_i(e_{i-1}) = \hat{x}_{\text{MAP}}^{0,i}[1]$ , lo que supone que  $\hat{x}_{\text{MAP}}^{1,i}[0] = \hat{x}_{\text{MAP}}^{0,i}[0]$ .

Así pues, queda demostrada la invarianza con respecto a la muestra de referencia para el ejemplo considerado. Nótese que esta demostración es fácilmente extensible por inducción a cualquier cantidad de intervalos del mapa,  $M$ , así como a un número arbitrario de iteraciones,  $N$ . En consecuencia, se llega a la conclusión de que los estimadores ML y MAP no dependen de la muestra de referencia escogida.

### 4.3.3. Estimador de Mínimo Error Cuadrático Medio (MS)

En este apartado se muestra el estimador de mínimo error cuadrático medio (MS o MMSE) de la secuencia caótica. Se realiza un desarrollo idéntico al del estimador MAP: primero se asume que el itinerario es conocido y se obtiene el estimador MS local (que en este caso es distinto del ML en general), y a continuación se extiende este resultado al caso en el que la secuencia simbólica es desconocida, obteniéndose el estimador MS global.

#### 4.3.3.1. Secuencia Simbólica Conocida

Cuando la secuencia simbólica es conocida, el estimador MS se obtiene calculando la esperanza matemática de  $x[0]$  con respecto a la FDP condicional dada por (4.17):

$$p(x[0]|\mathbf{y}, \mathbf{s}_i) = \tilde{K}_i \chi_i(x[0]) \exp\left(-\frac{(x[0] - \mu_i)^2}{2\sigma_i^2}\right), \quad (4.46)$$

donde  $\tilde{K}_i = K_i \cdot q_i$ . El estimador MS se encuentra integrando (4.46) entre los límites de la región  $i$ -ésima:

$$\hat{x}_{\text{MS}}^i[0] = \tilde{K}_i \int_{\eta_i}^{\kappa_i} x[0] \exp\left(-\frac{(x[0] - \mu_i)^2}{2\sigma_i^2}\right) dx[0].$$

Esta integral se puede solucionar fácilmente mediante el cambio de variable  $v = (x[0] - \mu_i)/\sqrt{2\sigma_i^2}$ , que genera dos nuevas integrales: la primera inmediata, y la segunda que se resuelve mediante un nuevo cambio de variable,  $t = -v^2$ . Como resultado de este proceso de integración, el estimador MS obtenido finalmente es

$$\hat{x}_{\text{MS}}^i[0] = \mu_i + \sigma_i^2 \tilde{K}_i(\hat{\xi}_i) g_i(\hat{\xi}_i), \quad (4.47)$$

siendo  $\hat{\xi}_i$  la desviación del pico de la Gaussiana (estimadores ML y MAP locales) con respecto al punto central de la región  $i$ -ésima,

$$\hat{\xi}_i = \mu_i - \rho_i,$$

$g_i(\hat{\xi}_i)$  una función de ponderación local, dada por

$$g_i(\hat{\xi}_i) = \exp\left(-\frac{(\hat{\xi}_i + \Delta_i/2)^2}{2\sigma_i^2}\right) - \exp\left(-\frac{(\hat{\xi}_i - \Delta_i/2)^2}{2\sigma_i^2}\right), \quad (4.48)$$

que vale cero si  $\mu_i = \rho_i$  (esto es, si  $\hat{\xi}_i = 0$ ), y  $K_i(\hat{\xi}_i)$  el factor de normalización correspondiente a la región  $i$ -ésima, que viene dado por (4.25). No obstante, teniendo en cuenta que el límite inferior de la región,  $\eta_i$ , se puede poner como

$$\eta_i = \rho_i - \Delta_i/2 = \mu_i - \hat{\xi}_i - \Delta_i/2,$$

y que igualmente el límite superior se puede formular como

$$\kappa_i = \rho_i + \Delta_i/2 = \mu_i - \hat{\xi}_i + \Delta_i/2,$$

este término de normalización también se puede expresar en función de  $\hat{\xi}_i$  y  $\Delta_i$ :

$$\tilde{K}_i(\hat{\xi}_i) = \sqrt{\frac{2}{\pi\sigma_i^2}} \left[ \operatorname{erf}\left(\frac{\hat{\xi}_i + \Delta_i/2}{\sqrt{2\sigma_i^2}}\right) + \operatorname{erf}\left(\frac{-\hat{\xi}_i + \Delta_i/2}{\sqrt{2\sigma_i^2}}\right) \right]^{-1}. \quad (4.49)$$

A continuación se proporciona la definición formal del estimador MS cuando el itinerario es conocido en la Proposición 4.2, mientras que su implementación se muestra en el Algoritmo 4.2. Posteriormente, se analiza de manera breve la forma de los dos términos que intervienen en el mismo,  $\tilde{K}_i(\hat{\xi}_i)$  y  $g_i(\hat{\xi}_i)$ , así como la del propio estimador en función de  $\hat{\xi}_i$ , comparándola con la de los estimadores ML y MAP.

**Proposición 4.2** *Sea una secuencia de muestras desconocida,  $\{x[k]\}_{k=0}^N$ , generada mediante la iteración, usando (3.11) o (3.12), de un mapa PWL con  $M$  intervalos definido por (2.14), sea  $\mathbf{s} = [s[0], \dots, s[N-1]]^T$  el vector que contiene el itinerario conocido asociado a la secuencia caótica, y sea  $\{y[k]\}_{k=0}^N$  la secuencia de observaciones disponibles, obtenidas sumando ruido blanco Gaussiano a cada muestra  $x[k]$ . El estimador Bayesiano de mínimo error cuadrático medio (MS o MMSE) de la condición inicial de la secuencia,  $x[0]$ , viene dado por (4.47):*

$$\hat{x}_{\text{MS}}^i[0] = \mu_i + \sigma_i^2 \tilde{K}_i(\hat{\xi}_i) g_i(\hat{\xi}_i).$$

Donde  $\mu_i$  es la posición de la moda de la FDP a posteriori, que es una Gaussiana truncada, dada por (4.18);  $\sigma_i^2$  es su varianza en ausencia de truncamiento, dada por (4.19);  $\tilde{K}_i(\hat{\xi}_i)$  es un factor de normalización, dado por (4.49);  $\hat{\xi}_i$  es la desviación de la media de la Gaussiana con respecto al centro de la región definida por la secuencia simbólica conocida,  $\hat{\xi}_i = \mu_i - \rho_i$ ; y  $g_i(\hat{\xi}_i)$  es una función de ponderación, dada por (4.48). Respecto al resto de muestras de la señal caótica, se puede obtener una proyección adecuada de su estimador MS sobre el espacio de secuencias caóticas válidas iterando hacia delante a partir de  $\hat{x}_{\text{MS}}^i[0]$ . Es decir,

$$\hat{x}_{\text{MS}}^i[k] \simeq f^k(\hat{x}_{\text{MS}}^i[0]), \quad (4.50)$$

con  $k = 1, \dots, N$ . En consecuencia, una buena aproximación del estimador MS de la secuencia completa viene dada por

$$\hat{\mathbf{x}}_{\text{MS}}^i \simeq [\hat{x}_{\text{MS}}^i[0], f(\hat{x}_{\text{MS}}^i[0]), f^2(\hat{x}_{\text{MS}}^i[0]), \dots, f^N(\hat{x}_{\text{MS}}^i[0])]^T. \quad (4.51)$$

Nótese que, a diferencia de los estimadores ML y MAP, no se puede garantizar que (4.51) proporcione la estima MS de la secuencia caótica completa, puesto que para el estimador MS no se cumple la propiedad de invariancia. Como se ha discutido anteriormente, habría que calcular la estima MS de cada muestra de la señal caótica y posteriormente encontrar la mejor proyección de este conjunto de estimas, que en general no concuerdan con ninguna secuencia caótica generada por el mapa, sobre el espacio de secuencias caóticas válidas. Sin embargo, tomando como estado inicial  $\hat{x}_{\text{MS}}^i[0]$  únicamente existe una secuencia caótica válida, dada por (4.51), y puesto que usando la misma se obtienen buenos resultados, esta es la única que se va a considerar en lo sucesivo.

1. Obtener el mínimo de la función de coste dada por (4.16) derivando con respecto a  $x[0]$ , y cuyo valor viene dado por  $\mu_i$  (4.18).
2. Hallar la región del espacio de fases,  $R_i = [\eta_i, \kappa_i]$ , formada por el conjunto de valores de  $x[0]$  que pueden generar el itinerario conocido,  $s_i$ , mediante el Algoritmo 3.5, y obtener su punto medio,  $\rho_i = (\kappa_i + \eta_i)/2$ , y su anchura,  $\Delta_i = \kappa_i - \eta_i$ . Calcular también  $\hat{\xi}_i = \mu_i - \rho_i$ .
3. Encontrar el estimador MS de  $x[0]$  usando (4.47), para lo que es necesario calcular previamente el factor de normalización,  $\tilde{K}_i(\hat{\xi}_i)$ , usando (4.49), y la función de ponderación,  $g_i(\hat{\xi}_i)$ , mediante (4.48).
4. Obtener el estimador MS del resto de la secuencia iterando hacia delante a partir de  $\hat{x}_{\text{MS}}^i[0]$ .

**Algoritmo 4.2:** Estimador MS de una secuencia caótica cuando la secuencia simbólica es conocida.

En la Figura 4.5 se presenta la forma de las diferentes funciones de corrección que intervienen en el cálculo del estimador MS. En la primera gráfica se muestra el valor

del factor de normalización de la FDP a posteriori,  $\tilde{K}_i(\hat{\xi}_i)$ , multiplicado por la varianza dentro de la región,  $\sigma_i^2$ . Se puede apreciar cómo este término crece exponencialmente cuando el valor absoluto de  $\hat{\xi}_i$  es mayor que  $\Delta_i/2$ . La segunda función que se representa es  $g_i(\hat{\xi}_i)$ , que se mantiene entre -1 y 1 dentro de los límites de la región (esto es, cuando  $|\hat{\xi}_i| \leq \Delta_i/2$ ), y que luego tiende a cero (nótese que en la Figura 4.5(b) se representa en valor absoluto). Aunque esta gráfica puede parecer un tanto extraña, se va a demostrar posteriormente que esta función da lugar a un estimador suave de la señal, que va a tender hacia  $\eta_i$  y  $\kappa_i$  en el límite, aunque sin sufrir un corte abrupto (es decir, un cambio brusco de la derivada) en los límites de la región como ocurre en el caso de los estimadores ML y MAP. Esto es debido a que el término  $\tilde{K}_i(\hat{\xi}_i)$  contrarresta la tendencia hacia cero de  $g_i(\hat{\xi}_i)$  fuera de la región  $i$ -ésima. El factor de corrección completo se presenta en la tercera gráfica, constatándose que se trata de una función continua y suave.

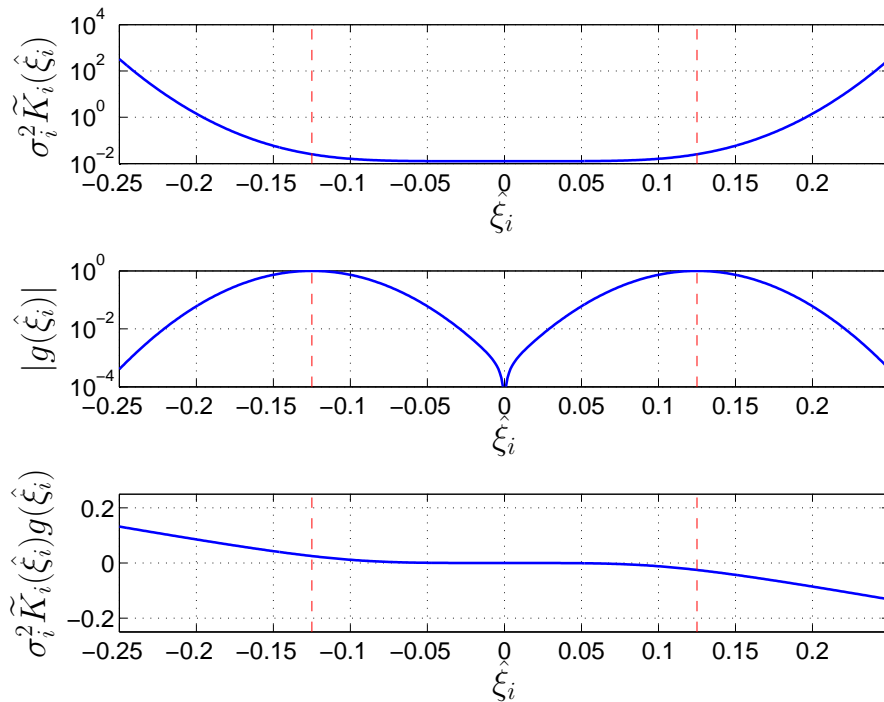


Figura 4.5: Funciones de corrección para el estimador MS con secuencia simbólica conocida. La anchura de la región es  $\Delta_i = 0,25$ , y la varianza del ruido es  $\sigma_i^2 = 10^{-3}$ . (a) Constante de normalización,  $\tilde{K}_i(\hat{\xi}_i)$ , multiplicada por la varianza en la región,  $\sigma_i^2$ . (b) Función de ponderación,  $g(\hat{\xi}_i)$ . (c) Desviación del estimador MS respecto a la moda de la Gaussiana en la región,  $\sigma_i^2 \tilde{K}_i(\hat{\xi}_i)g(\hat{\xi}_i)$ .

Para analizar el estimador MS resulta conveniente desarrollar (4.47) usando (4.48) y (4.49), de tal modo que  $\hat{x}_{\text{MS}}^i[0]$  quede únicamente en función del punto medio de la región,  $\rho_i$ , su anchura,  $\Delta_i$ , la varianza de la Gaussiana sin truncamiento,  $\sigma_i^2$ , y la

desviación del pico de la Gaussiana con respecto al punto central de la región,  $\hat{\xi}_i$ , que es el único parámetro que depende de las observaciones:

$$\begin{aligned}
\hat{x}_{\text{MS}}^i[0] &= \rho_i + \hat{\xi}_i + \sigma_i^2 \tilde{K}_i(\hat{\xi}_i) g_i(\hat{\xi}_i) \\
&= \rho_i + \hat{\xi}_i + \sqrt{\frac{2\sigma_i^2}{\pi}} \cdot \frac{\exp\left(-\frac{(\hat{\xi}_i + \Delta_i/2)^2}{2\sigma_i^2}\right) - \exp\left(-\frac{(\hat{\xi}_i - \Delta_i/2)^2}{2\sigma_i^2}\right)}{\operatorname{erf}\left(\frac{\hat{\xi}_i + \Delta_i/2}{\sqrt{2\sigma_i^2}}\right) + \operatorname{erf}\left(\frac{-\hat{\xi}_i + \Delta_i/2}{\sqrt{2\sigma_i^2}}\right)} \\
&= \rho_i + \hat{\xi}_i + \sqrt{\frac{2\sigma_i^2}{\pi}} \cdot \frac{N(\hat{\xi}_i)}{D(\hat{\xi}_i)}, \tag{4.52}
\end{aligned}$$

siendo  $N(\hat{\xi}_i)$  y  $D(\hat{\xi}_i)$  el numerador y el denominador de la última fracción que aparece en la segunda expresión de (4.52) respectivamente.

Ahora ya se puede analizar el comportamiento del estimador MS en sus puntos límite ( $\hat{\xi}_i = 0$ ,  $\hat{\xi}_i = \pm\Delta_i/2$  y  $\hat{\xi}_i = \pm\infty$ ) y extraer conclusiones acerca de su comportamiento. Para ello, en primer lugar nótese que  $g_i$  es antisimétrica con respecto a  $\hat{\xi}_i$ , mientras que  $\tilde{K}_i$  es simétrica. En consecuencia, el estimador MS es antisimétrico en relación con  $\hat{\xi}_i$ : si para un cierto valor de  $\hat{\xi}_i$  se obtiene  $\hat{x}_{\text{MS}}^i[0]$ , para  $-\hat{\xi}_i$  se va a obtener  $-\hat{x}_{\text{MS}}^i[0]$ . Por lo tanto, únicamente es necesario estudiar el comportamiento del estimador para valores positivos de  $\hat{\xi}_i$ :

1.  $\hat{\xi}_i = 0$ : En este caso  $N(\hat{\xi}_i) = 0$ , de modo que  $\hat{x}_{\text{MS}}^i[0] = \rho_i$ , puesto que la moda de la Gaussiana se encuentra en el medio de  $R_i$  y por consiguiente la FDP es simétrica. Es decir, en esta ocasión coinciden los tres estimadores,  $\hat{x}_{\text{MS}}^i[0] = \hat{x}_{\text{MAP}}^i[0] = \hat{x}_{\text{ML}}^i[0]$ , algo que no ocurre para ningún otro valor de  $\hat{\xi}_i$  excepto cuando  $\hat{\xi}_i \rightarrow \pm\infty$  o  $\sigma_i^2 \rightarrow 0$ .
2.  $\hat{\xi}_i = \Delta_i/2$ : En este caso, sustituyendo el valor de  $\hat{\xi}_i$  y operando un poco obtenemos

$$\hat{x}_{\text{MS}}^i[0] = \rho_i + \frac{\Delta_i}{2} - \sqrt{\frac{2\sigma_i^2}{\pi}} \cdot \frac{1 - \exp\left(-\frac{\Delta_i^2}{2\sigma_i^2}\right)}{\operatorname{erf}\left(\frac{\Delta_i}{\sqrt{2\sigma_i^2}}\right)},$$

que se encuentra situado entre 0 y  $\Delta_i/2$ , y cuyo valor exacto depende de la relación entre  $\Delta_i$  y  $\sigma_i$ , como se muestra en la Figura 4.6.

3.  $\hat{\xi}_i = \infty$ : En este último caso, al aplicar el límite se tiene una indeterminación del tipo 0/0 en el último término de  $\hat{x}_{\text{MS}}^i[0]$ , que se resuelve aplicando la regla de L'Hôpital [Apost1967]. La derivada de  $N(\hat{\xi}_i)$  es

$$\dot{N}(\hat{\xi}_i) = -\frac{\hat{\xi}_i + \Delta_i/2}{\sigma_i^2} \exp\left(-\frac{(\hat{\xi}_i + \Delta_i/2)^2}{2\sigma_i^2}\right) + \frac{\hat{\xi}_i - \Delta_i/2}{\sigma_i^2} \exp\left(-\frac{(\hat{\xi}_i - \Delta_i/2)^2}{2\sigma_i^2}\right),$$

mientras que la derivada de  $D(\hat{\xi}_i)$  es

$$\dot{D}(\hat{\xi}_i) = \sqrt{\frac{2}{\pi\sigma_i^2}} \left[ \exp\left(-\frac{(\hat{\xi}_i + \Delta_i/2)^2}{2\sigma_i^2}\right) - \exp\left(-\frac{(\hat{\xi}_i - \Delta_i/2)^2}{2\sigma_i^2}\right) \right],$$

ya que la derivada de  $\text{erf}(x)$  viene dada por [Abramo1965]

$$\frac{d\text{erf}(x)}{dx} = \frac{2}{\sqrt{\pi}} \exp(-x^2).$$

Juntando ambas expresiones, el estimador MS de  $x[0]$  resulta:

$$\begin{aligned} \lim_{\hat{\xi}_i \rightarrow \infty} \hat{x}_{\text{MS}}^i[0] &= \rho_i + \lim_{\hat{\xi}_i \rightarrow \infty} \left\{ \hat{\xi}_i + \frac{(\hat{\xi}_i - \Delta_i/2) \exp\left(-\frac{(\hat{\xi}_i - \Delta_i/2)^2}{2\sigma_i^2}\right)}{\exp\left(-\frac{(\hat{\xi}_i + \Delta_i/2)^2}{2\sigma_i^2}\right) - \exp\left(-\frac{(\hat{\xi}_i - \Delta_i/2)^2}{2\sigma_i^2}\right)} \right. \\ &\quad \left. - \frac{(\hat{\xi}_i + \Delta_i/2) \exp\left(-\frac{(\hat{\xi}_i + \Delta_i/2)^2}{2\sigma_i^2}\right)}{\exp\left(-\frac{(\hat{\xi}_i + \Delta_i/2)^2}{2\sigma_i^2}\right) - \exp\left(-\frac{(\hat{\xi}_i - \Delta_i/2)^2}{2\sigma_i^2}\right)} \right\} \\ &= \rho_i + \lim_{\hat{\xi}_i \rightarrow \infty} \left\{ \hat{\xi}_i + \frac{(\hat{\xi}_i - \Delta_i/2) - (\hat{\xi}_i + \Delta_i/2) \exp\left(-\frac{\Delta_i \hat{\xi}_i}{2\sigma_i^2}\right)}{\exp\left(-\frac{\Delta_i \hat{\xi}_i}{2\sigma_i^2}\right) - 1} \right\} \\ &= \rho_i + \frac{\Delta_i}{2} + \lim_{\hat{\xi}_i \rightarrow \infty} \frac{1 + \exp\left(-\frac{\Delta_i \hat{\xi}_i}{2\sigma_i^2}\right)}{1 - \exp\left(-\frac{\Delta_i \hat{\xi}_i}{2\sigma_i^2}\right)} = \rho_i + \frac{\Delta_i}{2} = \kappa_i, \end{aligned}$$

donde para pasar de la primera a la segunda expresión sencillamente se han dividido el numerador y el denominador por  $\exp(-(\hat{\xi}_i - \Delta_i/2)^2/(2\sigma_i^2))$ , y posteriormente se han simplificado términos en la siguiente ecuación.

Para finalizar este apartado, en la Figura 4.6 se muestra la desviación del estimador MS con respecto al punto central de la región  $i$ -ésima en función de  $\hat{\xi}_i$ , comparándola con la de los estimadores ML y MAP (idénticos cuando el itinerario es conocido). Se puede apreciar claramente su forma suave, con una transición no lineal parsimoniosa entre  $\eta_i$  y  $\kappa_i$ , más rápida cuanto mayor sea  $\Delta_i/\sigma_i$ . Por el contrario, en esta ocasión los estimadores ML y MAP presentan una expresión lineal en  $\hat{\xi}_i$  dentro de la región,

$$\hat{x}_{\text{ML}}^i[0] = \hat{x}_{\text{MAP}}^i[0] = \mu_i = \rho_i + \hat{\xi}_i,$$

ya que siempre coinciden y son iguales al valor de la moda de la FDP a posteriori. Fuera de la región simplemente se produce una saturación, tomando el estimador el valor de los límites de la misma,  $\eta_i$  ó  $\kappa_i$ , lo que provoca un cambio brusco de la derivada (aunque se mantiene la continuidad), dando lugar a una función total PWL. Por último, nótese que estos estimadores resultan parecidos a los estimadores conocidos de una constante en ruido con FDP uniforme [Schwar1975].



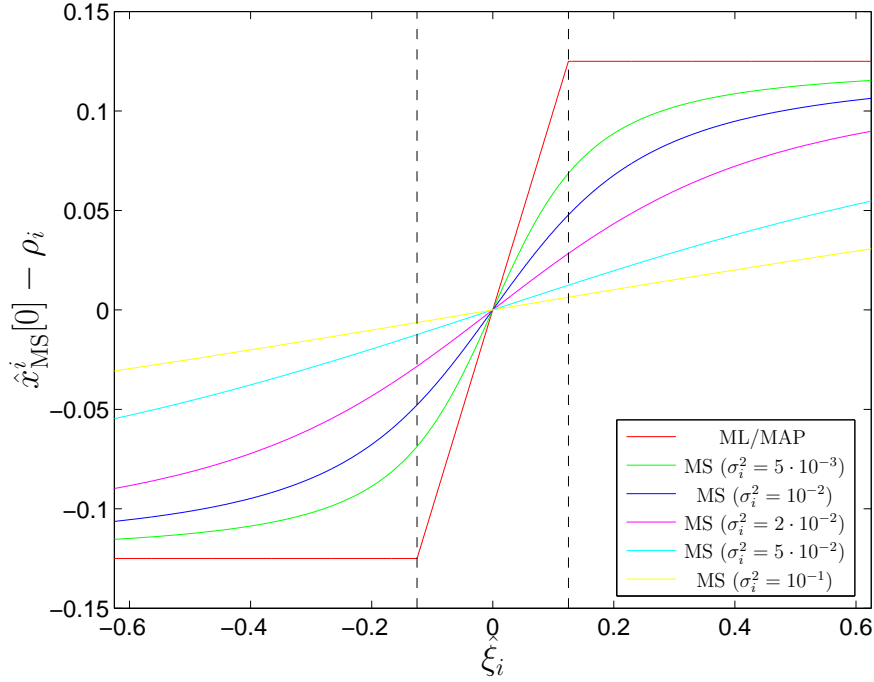


Figura 4.6: Desviación de los estimadores MS y ML/MAP (en rojo) de  $x[0]$  con respecto al punto central de la región,  $\rho_i$ , en función del desplazamiento de la media de la FDP a posteriori con respecto a  $\rho_i$ ,  $\hat{\xi}_i$ , para un itinerario conocido. La anchura de la región es  $\Delta_i = 0,25$ , y la varianza del ruido varía entre  $\sigma_i^2 = 5 \cdot 10^{-3}$  y  $\sigma_i^2 = 10^{-1}$ .

#### 4.3.3.2. Secuencia Simbólica Desconocida

El estimador MS cuando la secuencia simbólica es desconocida es similar al estimador para un itinerario conocido. En este caso la FDP a posteriori viene dada por (4.30), y el estimador MS se obtiene nuevamente calculando la esperanza matemática de  $x[0]$  con respecto a (4.30):

$$\hat{x}_{\text{MS}}[0] = K \sum_{i=1}^{P(N)} q_i \int_{\eta_i}^{\kappa_i} x[0] \exp\left(-\frac{(x[0] - \mu_i)^2}{2\sigma_i^2}\right) dx[0].$$

Esta integral se resuelve del mismo modo que en la Sección 4.3.3.1: se realiza el cambio de variable  $v = (x[0] - \mu_i)/\sqrt{2\sigma_i^2}$ , que da lugar a dos integrales, una de ellas inmediata, y la otra fácil de resolver mediante el cambio de variable  $t = -v^2$ . Siguiendo este proceso, se llega a la expresión del estimador MS:

$$\hat{x}_{\text{MS}}[0] = K \sum_{i=1}^{P(N)} \left( K_i^{-1}(\hat{\xi}_i)\mu_i + \sigma_i^2 g(\hat{\xi}_i) \right).$$

Esta ecuación se puede poner de forma alternativa haciendo uso de la relación existente entre la constante  $K$  y los factores  $K_i$ , dada por (4.33). Utilizando esta expresión, la

ecuación del estimador MS se puede reformular de manera que se obtenga una ecuación del estilo a las presentadas en [Panta2000a, Panta2001b, Panta2003]:

$$\hat{x}_{\text{MS}}[0] = \frac{\sum_{i=1}^{P(N)} K_i^{-1}(\hat{\xi}_i) \left( \mu_i + \sigma_i^2 K_i(\hat{\xi}_i) g(\hat{\xi}_i) \right)}{\sum_{i=1}^{P(N)} K_i^{-1}(\hat{\xi}_i)}. \quad (4.53)$$

Y por último, reconociendo la expresión de  $\hat{x}_{\text{MS}}^i[0]$  en el numerador, el estimador MS global se puede poner como

$$\hat{x}_{\text{MS}}[0] = \frac{\sum_{i=1}^{P(N)} K_i^{-1}(\hat{\xi}_i) \hat{x}_{\text{MS}}^i[0]}{\sum_{i=1}^{P(N)} K_i^{-1}(\hat{\xi}_i)}. \quad (4.54)$$

Esto es, el estimador MS global es simplemente una suma ponderada de los estimadores MS locales asociados a cada región, realizada de tal modo que se otorga un mayor peso a aquellos situados en regiones donde la FDP invariante del mapa es mayor. Nótese que las ecuaciones (4.53) y (4.54) se pueden obtener igualmente a partir de la formulación alternativa de  $p(x[0], \mathbf{s}|\mathbf{y})$  proporcionada en la ecuación (4.34). En este caso, llevando a cabo la integración de  $x[0]$  con respecto a (4.34) se obtiene directamente (4.53). El estimador MS se define formalmente en la Proposición 4.3, y la forma de obtenerlo en la práctica se describe en el Algoritmo 4.3.

**Proposición 4.3** *Sea una secuencia de muestras desconocida,  $\{x[k]\}_{k=0}^N$ , generada mediante la iteración, usando (3.11) o (3.12), de un mapa PWL con  $M$  intervalos definido por (2.14), y sea  $\{y[k]\}_{k=0}^N$  la secuencia de observaciones disponibles, obtenidas sumando ruido blanco Gaussiano a cada muestra  $x[k]$ . El estimador Bayesiano de mínimo error cuadrático medio (MS o MMSE) de la condición inicial de la secuencia,  $x[0]$ , cuando la secuencia simbólica es desconocida, viene dado por (4.54):*

$$\hat{x}_{\text{MS}}[0] = \frac{\sum_{i=1}^{P(N)} K_i^{-1}(\hat{\xi}_i) \hat{x}_{\text{MS}}^i[0]}{\sum_{i=1}^{P(N)} K_i^{-1}(\hat{\xi}_i)}.$$

Donde  $P(N)$  es el número de itinerarios válidos de longitud  $N$ ;  $\hat{x}_{\text{MS}}^i[0]$  es la estima MS local de  $x[0]$  dentro de la región  $i$ -ésima (esto es, la estima MS de  $x[0]$  suponiendo que la secuencia simbólica es  $\mathbf{s}_i$ ), dada por (4.47); y  $K_i$  es una constante de normalización, dada por (4.49), que actúa como factor de ponderación de los distintos estimadores MS locales, otorgando mayor peso a aquellos situados en regiones de mayor densidad de la FDP a posteriori. Respecto al resto de muestras de la señal caótica, se puede obtener

una proyección adecuada de su estimador MS sobre el espacio de secuencias caóticas válidas iterando hacia delante a partir de  $\hat{x}_{\text{MS}}[0]$ . Es decir,

$$\hat{x}_{\text{MS}}[k] \simeq f^k(\hat{x}_{\text{MS}}[0]), \quad (4.55)$$

con  $k = 1, \dots, N$ . En consecuencia, una buena aproximación del estimador MS de la secuencia completa viene dada por

$$\hat{\mathbf{x}}_{\text{MS}} \simeq [\hat{x}_{\text{MS}}[0], f(\hat{x}_{\text{MS}}[0]), \dots, f^N(\hat{x}_{\text{MS}}[0])]^T.$$

Nótese que, como consecuencia del proceso, también se ha obtenido una buena aproximación de la estima MS de la secuencia simbólica, que viene dada por  $S_f^+(\hat{x}_{\text{MS}}[0])$ .

1. Para cada posible secuencia simbólica, hallar el estimador MS local vinculado a la misma mediante el Algoritmo 4.2.
2. Encontrar el estimador MS global de  $x[0]$  mediante la suma ponderada de los  $P(N)$  estimadores locales, normalizada por la constante  $K$ , utilizando (4.54).
3. Obtener el estimador MS del resto de la secuencia iterando hacia delante a partir de  $\hat{x}_{\text{MS}}[0]$ .

**Algoritmo 4.3:** Estimador MS de una secuencia caótica cuando la secuencia simbólica es desconocida.

El efecto conseguido al obligar al estimador a cumplir el criterio Bayesiano de mínimo error cuadrático medio es el mismo que en el caso en que la secuencia simbólica es conocida: forzar a que el estimador obtenido sea una función suave de las observaciones. Sin embargo, en esta ocasión resulta más complicado comprobarlo, ya que en general el estimador es una función de  $P(N)$  variables:  $\hat{\xi}_1, \dots, \hat{\xi}_{P(N)}$ .

## 4.4. Estimadores Bayesianos de Secuencias Generadas por Mapas No PWL

Al igual que ocurría en el caso del estimador ML, no se pueden presentar estimadores cerrados para mapas no PWL, ya que sería necesario disponer de una expresión analítica para su iteración  $k$ -ésima, que resulta imposible de obtener en la mayoría de los casos. Sin embargo, las mismas técnicas utilizadas para obtener el estimador ML se pueden usar en general para obtener estimadores Bayesianos de mapas no PWL: métodos de rejilla y algoritmos iterativos locales. Estos métodos se describen brevemente en la Sección 4.4.1, y a continuación, en la Sección 4.4.2 se presenta una nueva clase de técnicas que permiten plantear estimadores Bayesianos para cualquier mapa caótico: los métodos de Monte Carlo basados en muestreo sobre cadenas de Markov (métodos MCMC).

#### 4.4.1. Métodos de Rejilla y Algoritmos Iterativos Locales

Cuando el mapa caótico no es PWL, entonces no es posible plantear estimadores cerrados. No obstante, recurriendo a técnicas de rejilla o a métodos iterativos locales, resulta sencillo desarrollar técnicas de estimación Bayesiana para esta clase de mapas. Recuérdese que la FDP a posteriori de  $x[0]$  viene dada por

$$p(x[0]|\mathbf{y}) = Kp(x[0]) \exp \left( -\frac{1}{2\sigma^2} \sum_{k=0}^N (y[k] - f^k(x[0]))^2 \right), \quad (4.56)$$

y nótese que, aunque no se disponga de ninguna ecuación cerrada para  $f^k(x[0])$ , su valor se puede calcular fácilmente para cualquier condición inicial simplemente iterando  $k$  veces a partir de  $x[0]$ , ya que la expresión del mapa,  $f(x)$ , es perfectamente conocida. En consecuencia, (4.56) puede evaluarse para cualquier condición inicial, obteniéndose el estimador MAP como el valor de  $x[0]$  que maximiza  $p(x[0]|\mathbf{y})$ , y el estimador MS como su esperanza matemática. Obviamente, si el itinerario de la señal es conocido se debe considerar únicamente la región de  $x[0]$  asociada a dicho itinerario, mientras que si es desconocido debe explorarse todo el espacio de fases del mapa.

En primer lugar, las *técnicas de rejilla* se pueden emplear del mismo modo que en el caso del estimador ML, aunque teniendo en cuenta que ahora el objetivo es distinto. Si se define una rejilla de  $N_r$  puntos por región,  $x^{i,j}[0]$  con  $1 \leq i \leq P(N)$  y  $0 \leq j \leq N_r$ , del modo descrito en la Sección 3.5, el estimador MAP de  $x[0]$  sería

$$\hat{x}_{\text{MAP}}[0] = \arg \max_{x^{i,j}[0]} p(x^{i,j}[0]|\mathbf{y}) = \arg \min_{x^{i,j}[0]} J_{\text{MAP}}(x^{i,j}[0]),$$

donde  $J_{\text{MAP}}(x[0])$  es la función de coste habitual modificada debido a la inclusión de la información a priori disponible:

$$J_{\text{MAP}}(x[0]) = J(x[0]) - 2\sigma^2 \ln p(x[0]). \quad (4.57)$$

Por otro lado, el estimador MS se obtendrá integrando numéricamente la expresión deseada (el producto de  $x[0]$  y  $p(x[0]|\mathbf{y})$ ) a lo largo de todos los puntos de la rejilla. Suponiendo que el espacio de fases del mapa sea  $I = [e_0, e_M]$ , entonces el estimador MS de  $x[0]$  se puede calcular numéricamente mediante una aproximación de orden cero:

$$\hat{x}_{\text{MS}}[0] = \sum_{i=1}^{P(N)} \frac{\kappa_i - \eta_i}{N_r} \sum_{j=0}^{N_r-1} x^{i,j}[0] p(x^{i,j}[0]|\mathbf{y}).$$

Respecto a los *métodos iterativos*, se pueden usar para obtener el estimador MAP realizando una modificación similar a la empleada en los métodos de rejilla: considerar como función a maximizar  $p(x[0]|\mathbf{y})$ , dada por (4.56), y aplicar un algoritmo de *ascenso de gradiente*, o alternativamente utilizar la función de coste modificada equivalente, dada por (4.57), y aplicar un algoritmo iterativo cualquiera de minimización local.

En cuanto al estimador MS, los métodos iterativos no resultan adecuados, ya que se limitan a buscar el máximo o el mínimo de una cierta función, mientras que para obtener el estimador MS se necesita explorar todo el espacio de fases del mapa (o al menos aquella porción del mismo donde  $x[0] \cdot p(x[0]|\mathbf{y})$  toma valores significativos) con un número suficientemente elevado de muestras, algo que en general no puede garantizarse empleando un algoritmo de minimización/maximización iterativo local.

## 4.4.2. Métodos de Monte Carlo Basados en Muestreo sobre Cadenas de Markov (Métodos MCMC)

### 4.4.2.1. Introducción a los métodos MCMC

Los métodos de Monte Carlo basados en muestreo sobre cadenas de Markov (métodos MCMC) fueron desarrollados originalmente para aplicaciones de mecánica estadística en la década de los 50. El propósito de estos primeros algoritmos era investigar las propiedades de una sustancia cualquiera compuesta por un elevado número de partículas (moléculas) interactuando entre sí. Puesto que este tipo de problemas resultaban intratables analíticamente, los métodos MCMC ofrecían una solución aproximada consistente en simular un sistema con unas pocas partículas (inicialmente unos pocos centenares) descrito probabilísticamente mediante la *FDP canónica* [Neal1993]. Posteriormente, esta clase de técnicas se han generalizado, adoptándose como un mecanismo eficiente para obtener muestras de acuerdo con una FDP a posteriori complicada en toda clase de problemas de inferencia estadística, y encontrando numerosas aplicaciones en las áreas de procesamiento de señal y comunicaciones: diseño de detectores e igualadores ciegos en sistemas de comunicaciones digitales (para canales FIR con AWGN [Chen1995a], sistemas CDMA [Wang2000], turbo igualación [Wang2001], sistemas MIMO [Zhu2005], etc.), deconvolución ciega [Cheng1996, Andrie2001], estimación espectral [Andrie1999, Davy2004], seguimiento de múltiples blancos (filtros de partículas basados en técnicas MCMC, MCMC-PF) [Tanaka2004, Khan2005], etc.

Los métodos MCMC se basan en construir una cadena de Markov cuya FDP invariante sea la FDP deseada. De esta manera, una vez que se supere un periodo transitorio inicial, *periodo de "burn-in"*, se puede considerar que la FDP de las muestras generadas por la cadena se ajusta a la FDP deseada. En un problema de estimación, esta FDP es la de los parámetros a inferir dadas las observaciones,  $p(\boldsymbol{\theta}|\mathbf{y})$ , y los métodos MCMC proporcionan realizaciones de  $\boldsymbol{\theta}$  de acuerdo con dicha FDP. En consecuencia, las muestras obtenidas una vez superado el periodo transitorio inicial ( $\boldsymbol{\theta}^1, \dots, \boldsymbol{\theta}^{N_m}$ ) se pueden usar para obtener cualquier estimador deseado. La gran ventaja de este tipo de técnicas frente a otros métodos de búsqueda global como los algoritmos genéticos es que cuentan con una sólida base teórica: la de las cadenas y procesos de Markov [Neal1993, Ruanai1996].

Dentro de esta clase de métodos, el algoritmo conceptualmente más sencillo es el *muestreador de Gibbs*, propuesto originalmente por Geman y Geman en el contexto de la restauración de imágenes sometidas a ruido y distorsión no lineal [Geman1984], y

revisado y extendido posteriormente por Gelfand y Smith para su uso en otros problemas de inferencia estadística [Gelfan1990]. Sin embargo, el primer método MCMC propuesto fue el *algoritmo de Metropolis*, planteado inicialmente por Metropolis et al. en 1953 para estudiar las propiedades de una sustancia [Metro1953]. Posteriormente, Hastings generalizó el algoritmo de Metropolis, estudiando su aplicación en problemas genéricos de inferencia estadística y dándole su forma actual [Hastin1970], por lo que con frecuencia se habla de *algoritmo de Metropolis-Hastings* (algoritmo MH) [Ruanai1996]. Numerosas técnicas MCMC han sido propuestas desde entonces, entre las que cabe destacar los métodos de dinámica molecular [Alder1959, Ander1980] o el método híbrido de Monte Carlo [Duane1987]. No obstante, el algoritmo MH sigue siendo uno de los más utilizados en la actualidad debido a su sencillez conceptual y su bajo coste computacional en comparación con el resto. Además, el algoritmo MH es el que se va a usar en la presente Tesis, de modo que es el único considerado a partir de este momento.

#### 4.4.2.2. Aplicación del Algoritmo de Metropolis-Hastings a la Estimación de Señales Caóticas

En esta sección se desarrolla un algoritmo para la estimación de señales caóticas generadas por un mapa unidimensional cualquiera, PWL o no PWL, basado en el algoritmo MH. Este método es simplemente una extensión del presentado en [Luengo2002a] para la estimación Bayesiana (MAP y MS) de señales generadas por el mapa logístico con  $\lambda = 4$ , de tal modo que resulte aplicable a cualquier mapa unidimensional. La idea básica del mismo consiste en usar como muestra de referencia  $x[N]$  (en lugar de  $x[0]$ , como en el resto del capítulo), de modo que el vector de parámetros a estimar esté compuesto por  $x[N]$  y la secuencia simbólica completa:

$$\boldsymbol{\theta} = [x[N], s[0], s[1], \dots, s[N-1]]^T = [x[N], \mathbf{s}^T]^T.$$

En cada iteración del algoritmo se actualiza un único parámetro (esto es, se trata de un algoritmo MH local), empezando por la muestra final del mapa,  $\boldsymbol{\theta}(1) = x[N]$ , y continuando con el resto de parámetros que identifican su itinerario:  $\boldsymbol{\theta}(2) = s[0], \dots, \boldsymbol{\theta}(N+1) = s[N-1]$ .

Para desarrollar el algoritmo es necesario disponer de la FDP a posteriori de los parámetros dadas las observaciones,

$$p(\boldsymbol{\theta}|\mathbf{y}) = p(x[N], \mathbf{s}|\mathbf{y}) = K p(\mathbf{y}|x[N], \mathbf{s})p(x[N]|\mathbf{s})p(\mathbf{s}), \quad (4.58)$$

donde  $K = p(\mathbf{y})^{-1}$  es una constante que no depende de  $\boldsymbol{\theta}$ , y cuyo cálculo no es necesario para el algoritmo. La FDP de las observaciones dados los parámetros presenta la forma habitual: una Gaussiana,

$$p(\mathbf{y}|x[N], \mathbf{s}) = (2\pi\sigma^2)^{-(N+1)/2} \exp\left(-\frac{J(x[N], \mathbf{s})}{2\sigma^2}\right). \quad (4.59)$$

La FDP de  $\mathbf{s}$  también resulta sencilla de obtener: se trata simplemente de un sumatorio ponderado de impulsos centrados en cada uno de los  $P(N)$  posibles itinerarios,

$$p(\mathbf{s}) = \sum_{i=1}^{P(N)} \Pr(\mathbf{s} = \mathbf{s}_i) \delta(\mathbf{s} - \mathbf{s}_i). \quad (4.60)$$

Para un mapa de Markov  $\Pr(\mathbf{s} = \mathbf{s}_i)$  se puede expresar como el producto de la probabilidad de cada uno de los  $N$  símbolos de la secuencia,

$$\Pr(\mathbf{s} = \mathbf{s}_i) = \prod_{k=0}^{N-1} \Pr(s[k] = s_i[k]) = \prod_{k=0}^{N-1} p_{E_{s_i[k]}},$$

mientras que en el caso de un mapa no markoviano tiene que calcularse su valor empíricamente a partir de una secuencia caótica suficientemente larga, tal y como se ha descrito en la Sección 4.3.1.1. Por último, la FDP de  $x[N]$  condicionada por el itinerario,  $\mathbf{s}$ , viene dada simplemente por  $p(x[N])$ , con la restricción impuesta por  $\mathbf{s}$  de que  $x[N]$  pertenezca a la región  $R_{\mathbf{s}}^N = [\eta_{\mathbf{s}}^N, \kappa_{\mathbf{s}}^N]$ , y multiplicada por una constante de normalización dependiente de la secuencia simbólica:

$$p(x[N]|\mathbf{s}) = C_{\mathbf{s}} p(x[N]) \chi_{R_{\mathbf{s}}^N}(x[N]), \quad (4.61)$$

siendo

$$C_{\mathbf{s}} = \left[ \int_{\eta_{\mathbf{s}}^N}^{\kappa_{\mathbf{s}}^N} p(x[N]) dx[N] \right]^{-1}.$$

Suponiendo que se utiliza la aproximación PWC de la FDP a priori dada por (4.11), y que los límites inferior y superior de  $R_{\mathbf{s}}^N$  se encuentran dentro de las regiones  $l$ -ésima y  $r$ -ésima ( $r \geq l$ ) de la partición de  $x[0]$  en función del itinerario respectivamente (es decir, que  $\eta_{\mathbf{s}}^N \in R_{\mathbf{s}_l}^0$  y  $\kappa_{\mathbf{s}}^N \in R_{\mathbf{s}_r}^0$ ), entonces

$$C_{\mathbf{s}}^{-1} = \Pr(\mathbf{s} = \mathbf{s}_l) \frac{\kappa_{\mathbf{s}}^N - \eta_{\mathbf{s}}^N}{\kappa_{\mathbf{s}_l}^0 - \eta_{\mathbf{s}_l}^0}$$

cuando  $r = l$ , mientras que si  $r > l$ , se tiene que

$$C_{\mathbf{s}}^{-1} = \Pr(\mathbf{s} = \mathbf{s}_l) \frac{\kappa_{\mathbf{s}_l}^0 - \eta_{\mathbf{s}}^N}{\kappa_{\mathbf{s}_l}^0 - \eta_{\mathbf{s}_l}^0} + \sum_{j=l+1}^{r-1} \Pr(\mathbf{s} = \mathbf{s}_j) + \Pr(\mathbf{s} = \mathbf{s}_r) \frac{\kappa_{\mathbf{s}}^N - \eta_{\mathbf{s}_r}^0}{\kappa_{\mathbf{s}_r}^0 - \eta_{\mathbf{s}_r}^0}.$$

Finalmente, sustituyendo (4.59), (4.60) y (4.61) en (4.58), y agrupando todos aquellos términos que no dependen de  $x[N]$  ni  $\mathbf{s}$  en  $K$ , la FDP a posteriori buscada resulta

$$p(x[N], \mathbf{s}|\mathbf{y}) = K p(x[N]) \sum_{i=1}^{P(N)} C_{\mathbf{s}_i} \Pr(\mathbf{s} = \mathbf{s}_i) \exp\left(-\frac{J(x[N], \mathbf{s}_i)}{2\sigma^2}\right) \chi_{R_{\mathbf{s}_i}^N}(x[N]) \delta(\mathbf{s} - \mathbf{s}_i). \quad (4.62)$$

Para un mapa en el que cada una de sus  $M$  regiones del espacio de fases,  $E_i = [e_{i-1}, e_i]$  con  $1 \leq i \leq M$ , se mapea en el espacio de fases completo,  $I$ , se tiene que  $R_s^N = I = [e_0, e_M]$ , y por lo tanto  $p(x[N]|\mathbf{s}) = p(x[N])$ . En este caso, (4.62) se reduce a

$$p(x[N], \mathbf{s}|\mathbf{y}) = Kp(x[N]) \sum_{i=1}^{P(N)} \Pr(\mathbf{s} = \mathbf{s}_i) \exp\left(-\frac{J(x[N], \mathbf{s}_i)}{2\sigma^2}\right) \delta(\mathbf{s} - \mathbf{s}_i), \quad (4.63)$$

y si además se considera el mapa logístico con  $\lambda = 4$ , cuya FDP invariante viene dada por (C.4), y para el cual todos los itinerarios son válidos (esto es,  $P(N) = 2^N$ ) y equiprobables (es decir,  $\Pr(\mathbf{s} = \mathbf{s}_i) = 2^{-N}$  para  $i = 1, \dots, 2^N$ ), se obtiene la expresión particular mostrada en [Luengo2002a]:

$$p(x[N], \mathbf{s}|\mathbf{y}) = \frac{K}{\sqrt{x[N](1-x[N])}} \sum_{i=1}^{2^N} \exp\left(-\frac{J(x[N], \mathbf{s}_i)}{2\sigma^2}\right) \delta(\mathbf{s} - \mathbf{s}_i).$$

Para completar el algoritmo es necesario definir otros tres aspectos relacionados con el mismo: el estado inicial, la FDP usada para modificar cada uno de los parámetros, y la función de aceptación. Respecto al *estado inicial*, en lugar de obtenerlo aleatoriamente como en la forma estándar del algoritmo MH, en esta ocasión se va a obtener a partir de las observaciones ruidosas truncadas de tal modo que se encuentren dentro del espacio de fases del mapa. Es decir, a partir del vector de observaciones,  $\mathbf{y} = [y[0], \dots, y[N]]^T$ , se va a construir un vector de observaciones recortadas,  $\mathbf{y}_T = [y_T[0], \dots, y_T[N]]^T$ , que contiene para cada observación,  $y[k]$  con  $0 \leq k \leq N$ , una nueva observación limitada al espacio de fases del mapa,  $I = [e_0, e_M]$ , cuya expresión es

$$y_T[k] = \begin{cases} e_0, & y[k] < e_0; \\ y[k], & e_0 \leq y[k] \leq e_M; \\ e_M, & y[k] > e_M. \end{cases} \quad (4.64)$$

Este conjunto de observaciones es el que se utiliza para definir el estado inicial,

$$\boldsymbol{\theta}^0 = [y_T[N], s^0[0], s^0[1], \dots, s^0[N-1]]^T = [y_T[N], (\mathbf{s}^0)^T]^T, \quad (4.65)$$

siendo  $\mathbf{s}^0$  la secuencia de símbolos asociados a  $\mathbf{y}_T$ :  $s^0[k] = i \Leftrightarrow y_T[k] \in E_i$ , para  $1 \leq i \leq M$  y  $k = 0, \dots, N-1$ .

Para aquellos mapas en los que no todas las secuencias simbólicas sean válidas y  $\mathbf{s}^0$  resulte ser una secuencia inválida, el itinerario inicial se obtiene modificando uno por uno aquellos símbolos con mayor probabilidad de ser erróneos hasta encontrar una secuencia simbólica válida. Además, debe comprobarse que  $y_T[N]$  pertenece a la región correspondiente de  $\mathbf{s}^0$  (o su versión modificada), recortándose su valor en caso contrario. En consecuencia, el valor inicial de  $\boldsymbol{\theta}(1)$  es realmente

$$\boldsymbol{\theta}^0(1) = \begin{cases} \eta_{\mathbf{s}^0}^N, & y_T[N] < \eta_{\mathbf{s}^0}^N; \\ y_T[N], & \eta_{\mathbf{s}^0}^N \leq y_T[N] \leq \kappa_{\mathbf{s}^0}^N; \\ \kappa_{\mathbf{s}^0}^N, & y_T[N] > \kappa_{\mathbf{s}^0}^N. \end{cases}$$



Respecto a la FDP empleada para actualizar los parámetros, en el caso de  $x[N]$  se va a usar simplemente la FDP invariante natural del mapa, cuando se conozca, o una aproximación PWC de la misma en caso contrario. En este sentido, existen cuatro posibilidades a la hora de generar las muestras en función del tipo de FDP:

1. Cuando la FDP invariante es uniforme, como en el caso del TM y el S-TM con  $\beta = 2$ , el SK-TM, el BSK-TM o los mapas de Bernouilli, las muestras se pueden generar fácilmente usando cualquier método estándar.
2. Cuando la FDP invariante es PWC, como ocurre en general para cualquier mapa PWL de Markov [Isabel1997, Isabel1998], entonces cada muestra se puede obtener generando primero una región aleatoria de acuerdo con la probabilidad de cada una de las  $M$  regiones, y posteriormente escogiendo el nuevo valor de  $x[N]$  de manera uniforme dentro de la región.
3. Cuando el mapa es topológicamente conjugado con un mapa PWL de Markov, como por ejemplo el mapa logístico con  $\lambda = 4$  [Devan1989], el proceso consiste en generar muestras de acuerdo con la FDP del mapa PWL correspondiente (el TM con  $\beta = 2$  en el caso del mapa logístico con  $\lambda = 4$ ) y aplicar la conjugación topológica para obtener muestras con la FDP deseada. Este es el método seguido en [Luengo2002a], donde para la actualización  $n$ -ésima se genera, en primer lugar  $z^n[N] \sim \mathcal{U}(0, 1)$ , y a continuación la muestra deseada se obtiene como

$$x^n[N] = \sin^2(\pi z^n[N]/2).$$

4. Para otra clase de mapas, como los mapas PWL no markovianos o los mapas no PWL no conjugados con algún mapa PWL, se obtiene una aproximación PWC empírica de su FDP invariante y se generan las muestras siguiendo el procedimiento descrito en el punto 2, teniendo en cuenta que ahora existen  $P(N)$  regiones en lugar de  $M$ .

En el caso de los símbolos del itinerario, las sucesivas modificaciones se llevan a cabo nuevamente siguiendo su FDP invariante, aunque eliminando la posibilidad de seleccionar el símbolo actual en esta ocasión. En consecuencia, la FDP para la actualización  $n$ -ésima del símbolo  $k$ -ésimo viene dada por

$$p(s^n[k]) = \sum_{\substack{i=1 \\ i \neq s^{n-1}[k]}}^M \frac{\Pr(s^n[k] = s_i)}{\Pr(s^n[k] \neq s^{n-1}[k])} \delta(s^n[k] - s_i), \quad (4.66)$$

donde  $\Pr(s^n[k] \neq s^{n-1}[k])$  es un factor de normalización necesario para garantizar que la suma de  $p(s^n[k])$  a lo largo de todos los posibles símbolos vale uno. Nótese que en el caso de un mapa unimodal sólo existen dos símbolos posibles, de modo que si  $s[k] \in \{1, 2\}$  entonces  $s^n[k] = 3 - s^{n-1}[k]$ , mientras que si  $s[k] \in \{-1, 1\}$ , como en [Luengo2002a], resulta que  $s^n[k] = -s^{n-1}[k]$ .

En relación con la *función de aceptación*, únicamente se ha considerado aquella usada con mayor frecuencia por los métodos MCMC: la *función de aceptación de Metropolis* [Neal1993, Ruanai1996]. Suponiendo que  $\boldsymbol{\theta}^{n-1}$  representa el estado actual del vector de parámetros, y que  $\tilde{\boldsymbol{\theta}}^n$  es el vector de parámetros candidato, entonces la función de aceptación se puede escribir como

$$A(\boldsymbol{\theta}^{n-1}, \tilde{\boldsymbol{\theta}}^n) = \min\left(1, Q(\boldsymbol{\theta}^{n-1}, \tilde{\boldsymbol{\theta}}^n)\right), \quad (4.67)$$

siendo  $Q(\cdot)$  el cociente de las FDPs a posteriori de los dos vectores de parámetros, el actual y el candidato, dadas por (4.62) o (4.63):

$$Q(\boldsymbol{\theta}^{n-1}, \tilde{\boldsymbol{\theta}}^n) = \frac{p(\tilde{\boldsymbol{\theta}}^n | \mathbf{y})}{p(\boldsymbol{\theta}^{n-1} | \mathbf{y})} = \frac{p(\tilde{x}^n[N], \tilde{\mathbf{s}}^n | \mathbf{y})}{p(x^{n-1}[N], \mathbf{s}^{n-1} | \mathbf{y})}. \quad (4.68)$$

En consecuencia, si el valor de la FDP a posteriori evaluada en el estado candidato es mayor que la del estado actual, aquel será siempre aceptado. Mientras que, incluso cuando sea menor existe una cierta probabilidad de aceptarlo, ya que un nuevo estado peor que el actual puede dar lugar mejores estados futuros.

En el Algoritmo 4.4 se muestra de un modo detallado el método seguido para generar muestras de acuerdo con la FDP a posteriori de los parámetros mediante el algoritmo MH. En la práctica los procesos de aceptación o no de los nuevos candidatos se llevan a cabo generando una serie de variables aleatorias independientes uniformes entre 0 y 1, y comparándolas con la función de aceptación correspondiente. Esto es, el nuevo estado en el punto 2.3 del Algoritmo 4.4 se obtiene como

$$\boldsymbol{\Phi}_0^n = \begin{cases} \tilde{\boldsymbol{\theta}}_0^n, & u_0^n \leq A(\boldsymbol{\theta}^{n-1}, \tilde{\boldsymbol{\theta}}_0^n); \\ \boldsymbol{\theta}^{n-1}, & u_0^n > A(\boldsymbol{\theta}^{n-1}, \tilde{\boldsymbol{\theta}}_0^n); \end{cases} \quad (4.69)$$

con  $u_0^n \sim \mathcal{U}(0, 1)$ . De igual modo, en el punto 2.4.2 del Algoritmo 4.4, el nuevo estado resultante es

$$\boldsymbol{\Phi}_k^n = \begin{cases} \tilde{\boldsymbol{\theta}}_k^n, & u_k^n \leq A(\boldsymbol{\Phi}_{k-1}^n, \tilde{\boldsymbol{\theta}}_k^n); \\ \boldsymbol{\Phi}_{k-1}^n, & u_k^n > A(\boldsymbol{\Phi}_{k-1}^n, \tilde{\boldsymbol{\theta}}_k^n); \end{cases} \quad (4.70)$$

con  $u_k^n \sim \mathcal{U}(0, 1)$ , e independiente de las decisiones anteriores.

Nótese que, a la vista del Algoritmo 4.4, en (4.69) el único cambio posible en el estado es en  $\boldsymbol{\theta}^{n-1}(1) = x^{n-1}[N]$ , cuya nueva expresión viene dada por

$$x^n[N] = \begin{cases} \tilde{x}^n[N], & u_0^n \leq A(\boldsymbol{\theta}^{n-1}, \tilde{\boldsymbol{\theta}}_0^n); \\ x^{n-1}[N], & u_0^n > A(\boldsymbol{\theta}^{n-1}, \tilde{\boldsymbol{\theta}}_0^n). \end{cases}$$

1. Obtener el vector de observaciones recortadas dentro de los límites del espacio de fases del mapa ( $I = [e_0, e_M]$ ),  $\mathbf{y}_T = [y_T[0], \dots, y_T[N]]^T$ , con  $y_T[k]$  dadas por (4.64) para  $0 \leq k \leq N$ , y construir el vector de parámetros inicial a partir de  $\mathbf{y}_T$  aplicando (4.65).

2. Para  $1 \leq n \leq N_m + N_b$ , obtener nuevas muestras llevando a cabo los pasos siguientes:

2.1. Generar una nueva muestra del mapa caótico,  $\tilde{x}^n[N]$ , de acuerdo con su FDP invariante, o de una aproximación PWC de la misma cuando la FDP invariante sea desconocida.

2.2. Construir el candidato inicial para la iteración  $n$ -ésima del algoritmo como

$$\tilde{\boldsymbol{\theta}}_0^n = [\tilde{x}^n[N], \boldsymbol{\theta}^{n-1}(2 : N+1)^T]^T = [\tilde{x}^n[N], (\mathbf{s}^{n-1})^T]^T.$$

2.3. Aceptar dicho candidato de acuerdo con (4.67),

$$A(\boldsymbol{\theta}^{n-1}, \tilde{\boldsymbol{\theta}}_0^n) = \min\left(1, Q(\boldsymbol{\theta}^{n-1}, \tilde{\boldsymbol{\theta}}_0^n)\right),$$

donde la función  $Q$ , que viene dada por (4.68), resulta en este caso

$$Q(\boldsymbol{\theta}^{n-1}, \tilde{\boldsymbol{\theta}}_0^n) = \frac{p(\tilde{x}^n[N])}{p(x^{n-1}[N])} \exp\left(\frac{J(x^{n-1}[N], \mathbf{s}^{n-1}) - J(\tilde{x}^n[N], \mathbf{s}^{n-1})}{2\sigma^2}\right) \chi_{R_{\mathbf{s}^{n-1}}}(\tilde{x}^n[N]),$$

obteniéndose un nuevo estado,  $\boldsymbol{\Phi}_0^n$ , dado por (4.69).

2.4. Para  $1 \leq k \leq N$ , modificar el símbolo  $(k-1)$ -ésimo de la siguiente manera:

2.4.1. Construir un nuevo candidato,

$$\tilde{\boldsymbol{\theta}}_k^n = [\boldsymbol{\Phi}_{k-1}^n(1 : k), \tilde{\phi}_{k-1}^n(k+1), \boldsymbol{\Phi}_{k-1}^n(k+2 : N+1)]^T,$$

donde  $\tilde{\phi}_{k-1}^n(k+1)$  denota la modificación del componente  $(k+1)$ -ésimo de  $\boldsymbol{\Phi}_{k-1}^n$  (esto es, el símbolo  $(k-1)$ -ésimo del itinerario) siguiendo la FDP dada por (4.66).

2.4.2. Aceptar este candidato de acuerdo con (4.67),

$$A(\boldsymbol{\Phi}_{k-1}^n, \tilde{\boldsymbol{\theta}}_k^n) = \min\left(1, Q(\boldsymbol{\Phi}_{k-1}^n, \tilde{\boldsymbol{\theta}}_k^n)\right),$$

donde la función  $Q$ , que de nuevo viene dada por (4.68), es ahora

$$Q(\boldsymbol{\Phi}_{k-1}^n, \tilde{\boldsymbol{\theta}}_k^n) = \frac{C_{\tilde{\mathbf{s}}^{n,k}} \Pr(\mathbf{s} = \tilde{\mathbf{s}}^{n,k}) \chi_{R_{\tilde{\mathbf{s}}^{n,k}}}(\mathbf{x}^n[N])}{C_{\mathbf{s}^{n,k-1}} \Pr(\mathbf{s} = \mathbf{s}^{n,k-1}) \chi_{R_{\mathbf{s}^{n,k-1}}}(\mathbf{x}^n[N])} \times \exp\left(\frac{J(\mathbf{x}^n[N], \mathbf{s}^{n,k-1}) - J(\mathbf{x}^n[N], \tilde{\mathbf{s}}^{n,k})}{2\sigma^2}\right),$$

donde  $\mathbf{s}^{n,k-1} = \boldsymbol{\Phi}_{k-1}^n(2 : N+1)$  es el itinerario actual, y  $\tilde{\mathbf{s}}^{n,k} = \tilde{\boldsymbol{\theta}}_k^n(2 : N+1)$  es el itinerario candidato. El nuevo estado obtenido,  $\boldsymbol{\Phi}_k^n$ , está dado por (4.70).

2.5. Establecer el estado final de la iteración  $n$ -ésima,  $\boldsymbol{\theta}^n = \boldsymbol{\Phi}_N^n$  y  $\mathbf{s}^n = \mathbf{s}^{n,N} = \boldsymbol{\Phi}_N^n(2 : N+1)$ .

3. Descartar las muestras correspondientes a las  $N_b$  primeras iteraciones ( $\boldsymbol{\Phi}_k^1, \dots, \boldsymbol{\Phi}_k^{N_b}$  con  $0 \leq k \leq N$ ) como transitorio (periodo de "burn-in") y usar el resto ( $\boldsymbol{\Phi}_k^{N_b+1}, \dots, \boldsymbol{\Phi}_k^{N_m+N_b}$  con  $0 \leq k \leq N$ ) para implementar el estimador deseado: ML mediante (4.71), MAP usando (4.72), o MS a partir de (4.73) o (4.74).

**Algoritmo 4.4:** Método basado en el algoritmo MH para estimar señales generadas por un mapa caótico unidimensional cualquiera.

De igual modo, en (4.70) el único cambio potencial del estado se da en  $\boldsymbol{\Phi}_{k-1}^n(k+1) = \mathbf{s}^{n,k-1}[k-1]$ , que ahora va a ser

$$\mathbf{s}^{n,k}[k-1] = \begin{cases} \tilde{\mathbf{s}}^{n,k}[k-1], & u_k^n \leq A(\boldsymbol{\Phi}_{k-1}^n, \tilde{\boldsymbol{\Theta}}_k^n); \\ \mathbf{s}^{n,k-1}[k-1], & u_k^n > A(\boldsymbol{\Phi}_{k-1}^n, \tilde{\boldsymbol{\Theta}}_k^n); \end{cases}$$

permaneciendo el resto del itinerario inalterado. Esto es,  $\mathbf{s}^{n,k}[l] = \mathbf{s}^{n,k-1}[l]$  para  $0 \leq l \leq N-1$  con  $l \neq k-1$ .

Una vez generadas las  $N_m + N_b$  muestras, se deben descartar las correspondientes a las  $N_b$  primeras iteraciones ( $\boldsymbol{\Phi}_k^1, \dots, \boldsymbol{\Phi}_k^{N_b}$  con  $0 \leq k \leq N$ ), ya que se corresponden con el periodo transitorio inicial durante el que la cadena aún no ha convergido a su FDP invariante (periodo de “burn-in”), y usarse el resto ( $\boldsymbol{\Phi}_k^{N_b+1}, \dots, \boldsymbol{\Phi}_k^{N_m+N_b}$  con  $0 \leq k \leq N$ ) para implementar el estimador deseado. La duración de este periodo de “burn-in” se encuentra directamente relacionada con la tasa de convergencia del algoritmo. Aunque existen numerosas cotas de la misma que permiten estimar  $N_b$  de manera aproximada, con frecuencia son difíciles de calcular en la práctica o son demasiado holgadas para resultar útiles [Neal1993]. En consecuencia, en lugar de realizar un análisis teórico de la convergencia del algoritmo, se ha seguido una de las aproximaciones más habituales en la literatura: estudiar mediante simulación el número de iteraciones inicial necesario para garantizar la convergencia de la cadena de Markov a la FDP deseada.

Por último, una vez que se dispone del conjunto total de muestras, los diferentes estimadores se obtienen como se indica a continuación:

1. Si se utiliza una FDP para  $x[N]$  uniforme en todo su espacio de fases,  $p(x[N]) \sim \mathcal{U}(e_0, e_M)$ , en lugar de su verdadera FDP invariante, entonces se obtiene un estimador equivalente al de máxima verosimilitud (ML) seleccionando la muestra que minimiza la función de coste habitual:

$$\hat{\boldsymbol{\Phi}}_k^n(\text{ML}) = \arg \min_{\boldsymbol{\Phi}_k^n} J(\boldsymbol{\Phi}_k^n), \quad (4.71)$$

con  $N_b + 1 \leq n \leq N_m + N_b$  y  $0 \leq k \leq N$ .

2. Usando la FDP invariante real del mapa (o una aproximación PWC en su defecto), el estimador de máxima probabilidad a posteriori (MAP) viene dado por

$$\hat{\boldsymbol{\Phi}}_k^n(\text{MAP}) = \arg \max_{\boldsymbol{\Phi}_k^n} p(\boldsymbol{\Phi}_k^n | \mathbf{y}) = \arg \min_{\boldsymbol{\Phi}_k^n} J_{\text{MAP}}(\boldsymbol{\Phi}_k^n), \quad (4.72)$$

con  $N_b + 1 \leq n \leq N_m + N_b$  y  $0 \leq k \leq N$ .

3. Por último, usando de nuevo la FDP invariante del mapa y realizando una integración de Monte Carlo se obtendría el estimador de menor error cuadrático medio (MMSE o MS). El problema en este caso es análogo al existente para mapas PWL: si la integración se lleva a cabo sobre todas las muestras de la secuencia,

entonces la señal caótica resultante no va a ser válida en general, debiéndose buscar una proyección adecuada sobre el espacio de todas las posibles secuencias caóticas. Aunque existen infinitas proyecciones posibles, aquí únicamente se consideran las dos propuestas en [Luengo2002a]:

3.1. Estimador MH-MS1: La estima MS de  $x[0]$  se obtiene como

$$\hat{x}_{\text{MH-MS1}}[0] = \frac{1}{N_m(N+1)} \sum_{n=N_b+1}^{N_m+N_b} \sum_{k=0}^N f_{s^{n,k}}^{-N}(x^n[N]), \quad (4.73)$$

y la estima MS del resto de la secuencia iterando hacia delante:  $\hat{x}_{\text{MH-MS1}}[k] = f^k(\hat{x}_{\text{MH-MS1}}[0])$  con  $k = 1, \dots, N$ .

3.2. Estimador MH-MS2: Se obtiene la estima MS de la condición final,  $x[N]$ , promediando todas las muestras generadas,

$$\hat{x}_{\text{MH-MS2}}[N] = \frac{1}{N_m} \sum_{n=N_b+1}^{N_m+N_b} x^n[N], \quad (4.74)$$

y la estima MS del resto de la secuencia iterando hacia atrás usando el itinerario más probable:

$$\hat{x}_{\text{MH-MS2}}[k] = f_{\hat{s}_{\text{MH-MS2}}}^{k-N}(\hat{x}_{\text{MH-MS2}}[N]),$$

siendo

$$\hat{s}_{\text{MH-MS2}}[k] = \arg \max_i N_i[k],$$

con

$$N_i[k] = \sum_{n=N_b+1}^{N_m+N_b} \delta(s^{n,k+1}[k] - i),$$

y  $k = 1, \dots, N$ .

## 4.5. Resultados

En esta sección se presentan los resultados más relevantes obtenidos con los estimadores Bayesianos considerados (MAP y MS), tanto para mapas PWL como no PWL, haciendo especial énfasis en su comparación con los logrados con el estimador ML, mostrados en la Sección 3.6. En primer lugar, en la Sección 4.5.1 se discute brevemente el mecanismo seguido para evaluar los estimadores. A continuación, en la Sección 4.5.2 se muestran numerosos resultados para mapas PWL, cuya estimación es el principal objetivo de este capítulo. Por último, en la Sección 4.5.3 se presentan algunos resultados relativos a la estima de secuencias generadas por mapas no PWL obtenidos mediante los métodos MCMC.

### 4.5.1. Evaluación de los Estimadores Bayesianos

El criterio utilizado para la evaluación de la calidad de los estimadores Bayesianos es fundamentalmente el mismo que se emplea en el Capítulo 3 para analizar el estimador ML: el MSE. En la práctica, dado que los estimadores son básicamente insesgados, esto es equivalente a considerar su varianza, que va a ser el factor limitante de su rendimiento, y que coincide aproximadamente con el MSE.

Además, en general se va a considerar como criterio de comparación entre los distintos estimadores el MSE de  $x[0]$ , ya que se trata de la muestra de la secuencia más difícil de estimar, y los resultados obtenidos para dicha muestra son extrapolables a cualquier otro punto de la secuencia e incluso a la secuencia completa. Esto es, si un estimador mejora el rendimiento de otro para  $x[0]$  (proporcionando un menor MSE para una cierta SNR o alcanzando antes el CRLB por ejemplo), también lo va a mejorar para cualquier otro  $x[k]$  ( $1 \leq k \leq N$ ), así como para  $\mathbf{x}$ , aunque probablemente el grado de mejora sea menor que para  $x[0]$ . Como excepción, en la Sección 4.5.3 se va a considerar tanto el MSE de  $x[0]$  como el de  $\mathbf{x}$ , ya que, debido a las múltiples proyecciones posibles del estimador MS sobre el espacio de secuencias caóticas válidas, el estimador MS que proporciona un mejor rendimiento para  $x[0]$  es distinto del estimador óptimo para la secuencia completa,  $\mathbf{x}$ .

En cuanto a la muestra de referencia, en la Sección 4.5.2 se va a considerar exclusivamente  $x[0]$ . Aunque para los estimadores Bayesianos en general no se cumple la propiedad de invariancia, en la Sección 4.3.2.3 se ha mostrado mediante un sencillo ejemplo que los resultados obtenidos son independientes de la muestra de referencia para el estimador MAP. Respecto al estimador MS, éste no satisface la propiedad de invariancia, de modo que la elección de la muestra de referencia puede modificar los resultados obtenidos. Sin embargo, es de esperar que esta diferencia no sea sustancial, resultando mucho más importante obtener una proyección adecuada del estimador MS sobre el espacio de secuencias válidas. Por otro lado, en la Sección 4.5.3 se considera únicamente  $x[N]$  como muestra de referencia, analizándose cómo la selección de una u otra proyección puede dar lugar a estimas diferentes. La exploración de otros puntos intermedios de la secuencia caótica como muestra de referencia se deja como una posible línea futura.

Para finalizar esta sección, nótese que en este capítulo se introducen tres novedades a la hora de medir el rendimiento de los diferentes estimadores:

1. Para valorar la mejora introducida por los estimadores Bayesianos con respecto al estimador ML se introduce el concepto de *ganancia*, que se define simplemente como la diferencia en dBs entre el MSE de los estimadores MAP/MS y el ML.
2. La manera más justa de determinar la calidad de un estimador no es utilizando una secuencia concreta, sino considerando un conjunto amplio de secuencias generadas de acuerdo con la FDP invariante del mapa y promediando los resultados obtenidos. En la Sección 4.5.2 se muestra una tabla en la que se compara la ganancia en promedio de los estimadores MAP y MS.

- Además del MSE, otra medida de calidad interesante es la tasa de aciertos en los símbolos del itinerario con respecto a la secuencia simbólica real. En esta sección se muestran resultados tanto para los estimadores MAP/MS como para el estimador ML.

#### 4.5.2. Resultados para Mapas PWL

Para empezar, en la Figura 4.7 se muestran dos ejemplos del rendimiento de los estimadores MAP y MS, en comparación con el ML, para una secuencia generada usando el S-TM con  $x[0] = 0,34$ ,  $\beta = 1,5$  y  $N = 8$  [Panta2000a], así como para el mapa Markov2 de la Figura 2.5(a) con  $x[0] = 0,83$ ,  $c = 0,75$  y  $N = 6$  [Panta2003]. En esta ocasión el estimador MAP apenas mejora el rendimiento del ML para el S-TM, encontrándose su ganancia comprendida entre  $-0,37$  y  $1,84$  dB, con un valor promedio de tan sólo  $0,22$  dB. Para este mapa y esta condición inicial el estimador MS proporciona un resultado claramente superior, con una ganancia mínima y máxima de  $-0,13$  y  $3,74$  dB respectivamente, y una media de  $1,29$  dB. Para el segundo ejemplo los resultados son bastante mejores. En este caso el estimador MAP proporciona una ganancia media de  $1,59$  dB (con un mínimo de  $-0,71$  dB y un máximo de  $10,80$  dB), mientras que con el estimador MS se consigue una ganancia media de  $3,91$  dB (con un mínimo de  $-0,04$  dB y un máximo de  $13,44$  dB).

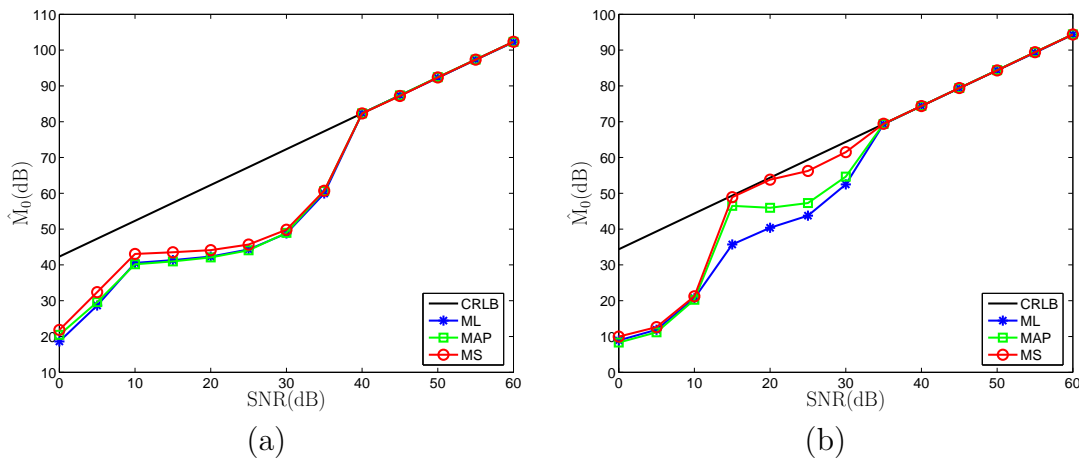


Figura 4.7: Comparación del MSE de los estimadores ML, MAP y MS de dos secuencias caóticas generadas usando el S-TM y el mapa Markov2. (a) S-TM con  $x[0] = 0,34$ ,  $\beta = 1,5$  y  $N = 8$ . (b) Mapa Markov2 de la Figura 2.5(a) con  $x[0] = 0,83$ ,  $c = 0,75$  y  $N = 6$ .

Los dos ejemplos de la Figura 4.7 son ilustrativos de la variabilidad de la ganancia obtenida con los estimadores Bayesianos con respecto al estimador ML. Al igual que en el caso del estimador ML, el rendimiento de los estimadores Bayesianos va a depender en gran medida del mapa, de sus parámetros, de la longitud de la secuencia caótica,

y de la condición inicial. En consecuencia, la ganancia conseguida va a variar en gran medida con dichos factores, aunque en general se va a obtener una mayor ganancia con el estimador MS que con el MAP.

Nótese que para aquellos mapas cuya FDP invariante es uniforme (como el TM y el S-TM para  $\beta = 2$ , el SK-TM, el BSK-TM o los mapas de Bernouilli) los estimadores ML y MAP coinciden siempre, y la ganancia obtenida utilizando el estimador MS es mínima. Por consiguiente, los resultados presentados en este capítulo se concentran única y exclusivamente en mapas cuya FDP invariante es no uniforme. Además, la ganancia obtenida con los estimadores Bayesianos va a depender en gran medida de la desviación con respecto a la uniforme de dicha FDP, así como de si la condición inicial a estimar se encuentra en una región de alta o baja probabilidad. En general, cuanto menos uniforme sea la FDP y cuanto más típica sea la secuencia caótica (es decir, cuanto más se ajuste a su FDP invariante natural) mayor ganancia se va a obtener.

Para ilustrar esta variabilidad en el comportamiento de los estimadores Bayesianos, en la Figura 4.8 se muestran ejemplos adicionales para otros tres mapas caóticos: el mapa Markov1 de la Figura 2.4, y los mapas Markov2 y PWL1 de la Figura 2.5. En las gráficas se puede apreciar cómo los estimadores MAP y MS mejoran el rendimiento del estimador ML para la mayor parte de los casos y valores de SNR, siendo superior en general el grado de mejora del estimador MS. Sin embargo, en algunos ejemplos (como la Figura 4.7(a) o la Figura 4.8(c)) el estimador MAP se comporta básicamente como el estimador ML, y el estimador MS proporciona una ganancia mínima. Incluso es posible encontrar ejemplos en los que el estimador MAP o el MS funcionan peor que el ML a lo largo de un cierto rango de valores de SNR (casos (d) para el MAP y (f) para el MS de la Figura 4.8).

Una posible explicación de este comportamiento se encuentra en el hecho de que, para una secuencia caótica dada, el rendimiento de los estimadores Bayesianos depende en gran medida de la región a la que pertenece  $x[0]$ . Supóngase que  $x[0]$  pertenece a la región  $i$ -ésima, esto es,  $x[0] \in R_i$ . En el caso del estimador MAP, cuando la densidad de probabilidad en dicha región,  $p_i$ , es pequeña, su comportamiento puede ser peor que el del estimador ML, especialmente si  $p_i$  es sensiblemente inferior a la densidad de las regiones adyacentes. Por el contrario, cuando  $p_i$  es grande en comparación con el valor de las regiones colindantes, entonces la ganancia del estimador MAP es potencialmente alta. En cuanto al estimador MS, la variación de su rendimiento se puede explicar de un modo similar, aunque ahora este no depende únicamente de  $p_i$ , sino de  $\Pr(x[0] \in R_i)$  (es decir, de  $p_i \cdot \Delta_i$ ), ya que el estimador MS realiza una integral a lo largo de las distintas regiones. Como prueba de las afirmaciones anteriores, en la Figura 4.9 se muestran  $p(x)$  y  $\Pr(x \in R_i)$  (con  $i = 1, \dots, P(N)$ ) para los tres mapas y las seis condiciones iniciales de la Figura 4.8.

Todas las gráficas mostradas en las Figuras 4.7 y 4.8 se han obtenido promediando 10.000 simulaciones de Monte Carlo para cada valor de SNR, con el fin de obtener valores precisos para la ganancia de los estimadores MAP y MS. En la Tabla 4.2 se muestra la ganancia mínima, máxima y media para todos los ejemplos de ambas figuras.



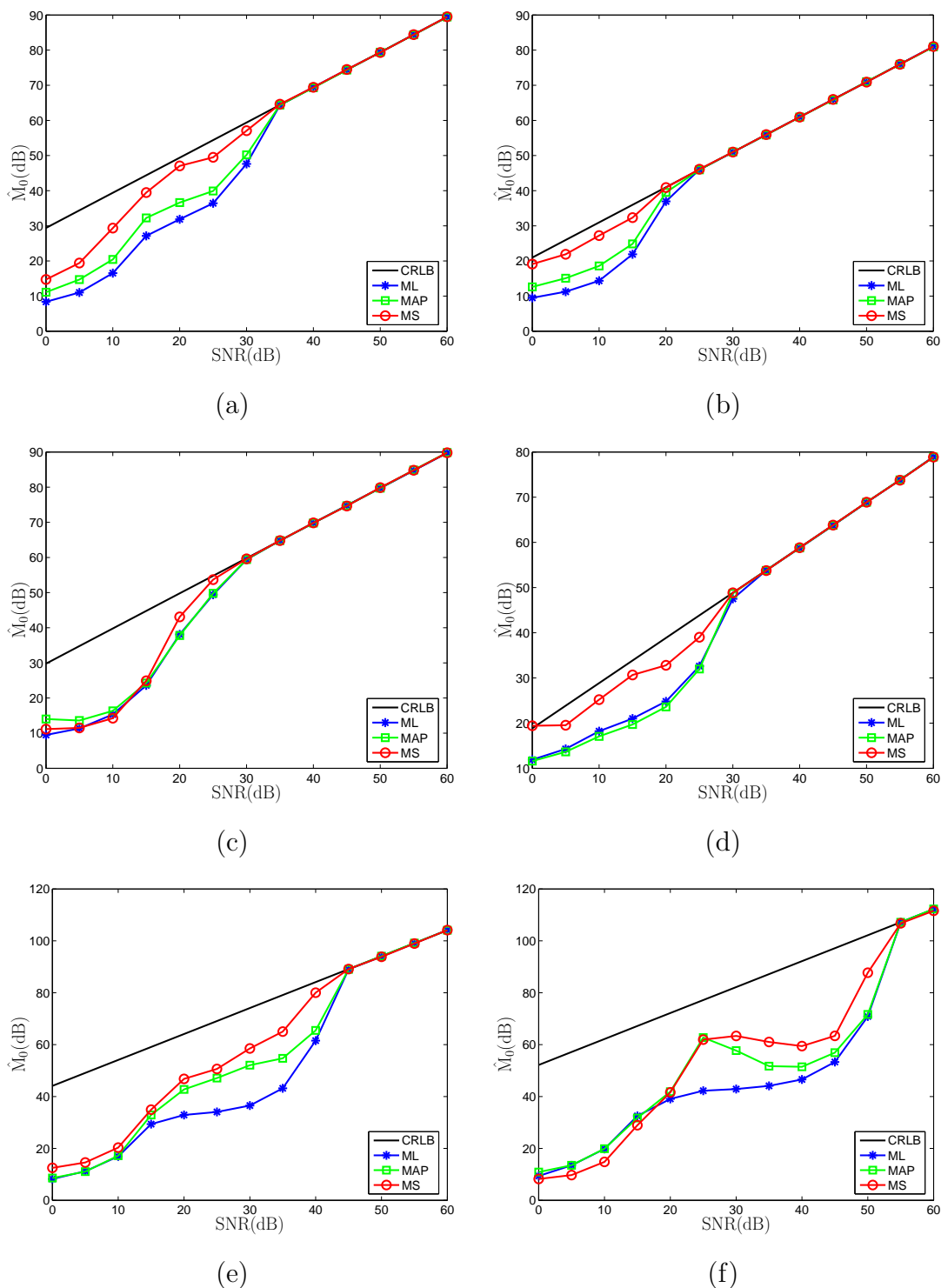


Figura 4.8: Comparación del rendimiento de los estimadores ML, MAP y MS para secuencias generadas usando diferentes mapas caóticos y condiciones iniciales. (a) Mapa Markov2 con  $x[0] = 0,2389$ ,  $c = 0,75$  y  $N = 10$ . (b) Mapa Markov2 con  $x[0] = 0,4040$ ,  $c = 0,75$  y  $N = 10$ . (c) Mapa PWL1 con  $x[0] = 0,88131$  y  $N = 6$ . (d) Mapa PWL1 con  $x[0] = 0,65251$  y  $N = 6$ . (e) Mapa Markov1 con  $x[0] = 0,22853$  y  $N = 4$ . (f) Mapa Markov1 con  $x[0] = 0,89164$  y  $N = 4$ .

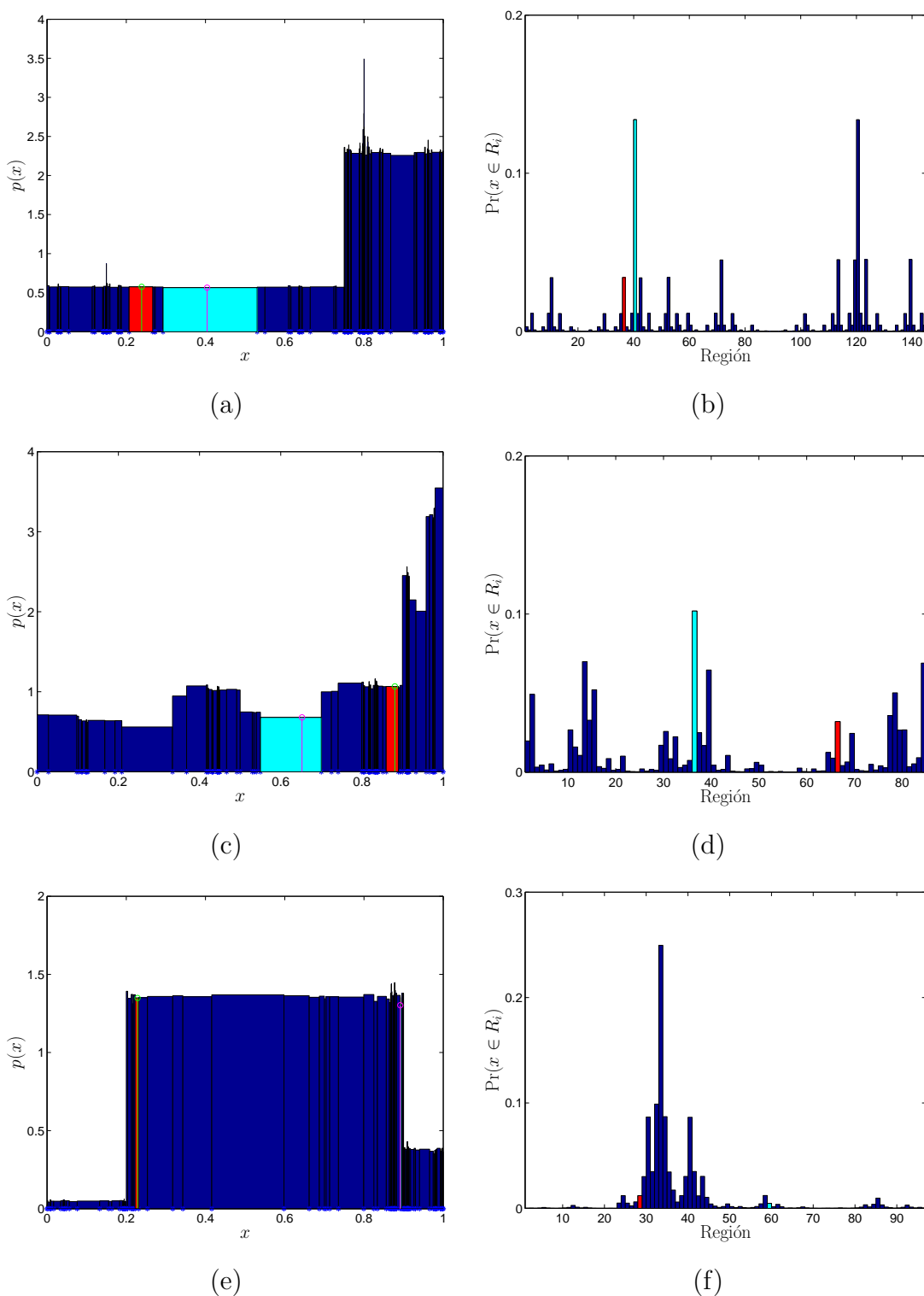


Figura 4.9: Relación entre el rendimiento de los estimadores Bayesianos y la región a la que pertenece  $x[0]$ . (a) y (b)  $p(x)$  y  $\Pr(x \in R_i)$  para el mapa Markov2 con  $N = 10$ , y regiones correspondientes a  $x[0] = 0,2389$  (en rojo) y  $x[0] = 0,4040$  (en azul cian). (c) y (d)  $p(x)$  y  $\Pr(x \in R_i)$  para el mapa PWL1 con  $N = 6$ , y regiones correspondientes a  $x[0] = 0,88131$  (en rojo) y  $x[0] = 0,65251$  (en azul cian). (e) y (f)  $p(x)$  y  $\Pr(x \in R_i)$  para el mapa Markov1 con  $N = 4$ , y regiones correspondientes a  $x[0] = 0,22853$  (en rojo) y  $x[0] = 0,89164$  (en azul cian).

Estimador	mín{ $g$ }	máx{ $g$ }	E( $g$ )
MAP Fig. 4.7(a)	-0,37 dB	1,84 dB	0,22dB
MS Fig. 4.7(a)	-0,13 dB	3,74 dB	1,29 dB
MAP Fig. 4.7(b)	-0,71 dB	10,80 dB	1,59 dB
MS Fig. 4.7(b)	-0,04 dB	13,44 dB	3,91 dB
MAP Fig. 4.8(a)	-0,13 dB	5,11 dB	2,00 dB
MS Fig. 4.8(a)	-0,05 dB	15,17 dB	6,00 dB
MAP Fig. 4.8(b)	-0,13 dB	4,23 dB	1,26 dB
MS Fig. 4.8(b)	-0,06 dB	12,86 dB	3,66 dB
MAP Fig. 4.8(c)	-0,32 dB	4,55 dB	0,67 dB
MS Fig. 4.8(c)	-1,08 dB	4,98 dB	0,90 dB
MAP Fig. 4.8(d)	-1,30 dB	1,19 dB	-0,31 dB
MS Fig. 4.8(d)	-0,08 dB	9,66 dB	3,48 dB
MAP Fig. 4.8(e)	-0,16 dB	15,53 dB	4,47 dB
MS Fig. 4.8(e)	-0,12 dB	22,00 dB	8,41 dB
MAP Fig. 4.8(f)	-0,79 dB	20,50 dB	4,31 dB
MS Fig. 4.8(f)	-5,00 dB	20,47 dB	6,54 dB

Tabla 4.2: Ganancia mínima, máxima y media de los estimadores MAP y MS con respecto al estimador ML para los diferentes mapas de las Figuras 4.7 y 4.8.

Como puede apreciarse, en todos los casos existe algún valor de SNR para el que el estimador ML es ligeramente mejor que los estimadores Bayesianos, aunque generalmente se da para muy baja SNR, y la diferencia es inferior a 1 dB. Por el contrario, existen valores intermedios de SNR para los cuales la ganancia de los estimadores MAP y MS es elevada, consiguiéndose en general una ganancia apreciable en promedio.

Al igual que ocurría con el estimador ML, el rendimiento de los estimadores Bayesianos viene determinado fundamentalmente por su varianza, ya que pueden considerarse insesgados en la práctica para cualquier SNR salvo quizás para valores muy bajos. Como ejemplo, en la Figura 4.10 se muestran el MSE y la varianza por un lado, y el sesgo por el otro, para la secuencia generada por el mapa PWL1 con el mismo valor de  $x[0]$  y  $N$  que en la Figura 4.8(d), y para la secuencia generada por el mapa Markov1 bajo las mismas condiciones que en la Figura 4.8(e). Los resultados para otros mapas y condiciones iniciales son muy similares: excepto para valores muy bajos de SNR, el MSE y la varianza son prácticamente idénticos.

De manera adicional a las gráficas presentadas del MSE, resulta interesante analizar el número de aciertos en los símbolos del itinerario para cada uno de los tres estimadores. En la Figura 4.11 se muestra la tasa de aciertos en los símbolos de la secuencia estimada con respecto al itinerario real para todas las secuencias de la Figura 4.8. Comparando ambas figuras se obtiene una conclusión de vital importancia a la hora de plantear estimadores eficientes computacionalmente en el Capítulo 5: los estimadores desarrollados alcanzan el CRLB cuando el número de errores en los símbolos del itinerario tiende a cero. En consecuencia, se deduce que para obtener una buena estima

de la señal caótica resulta fundamental conseguir una buena estima de la secuencia simbólica.

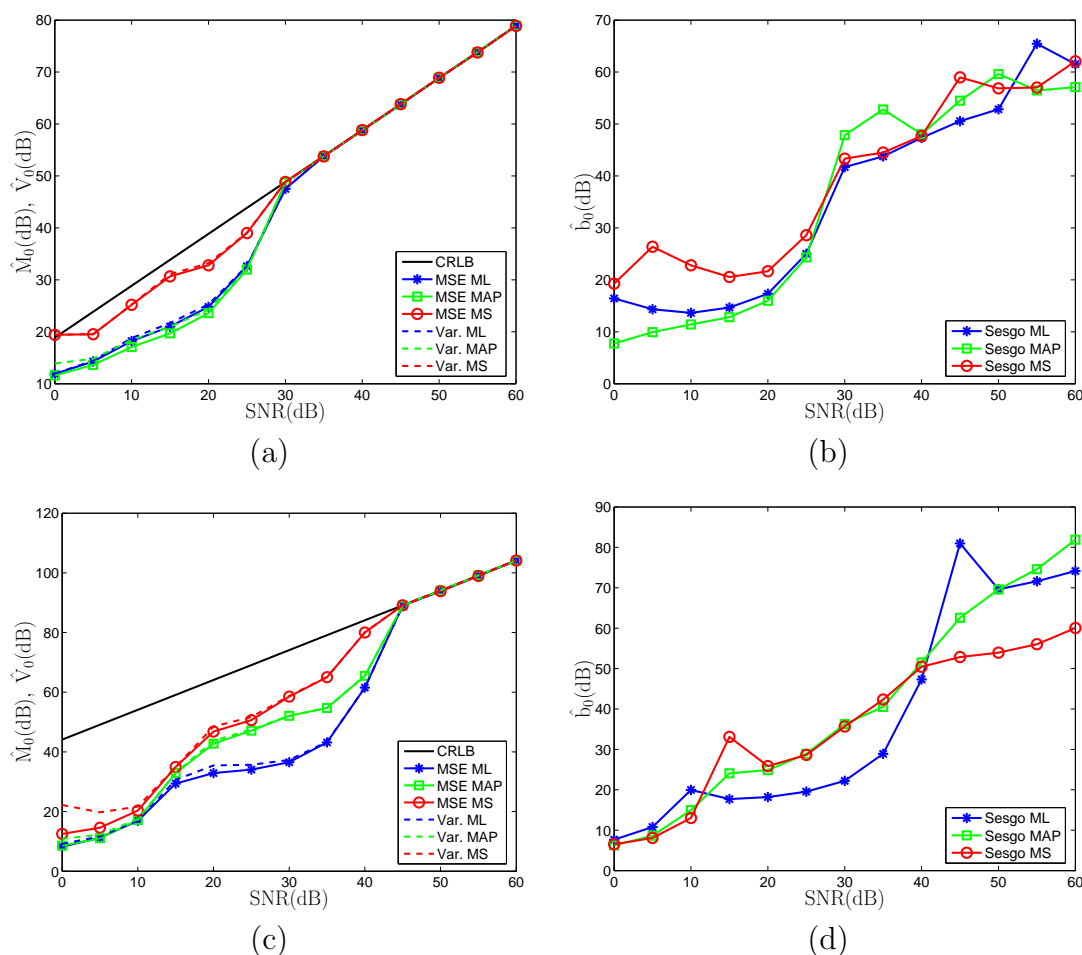


Figura 4.10: Sesgo, Varianza y MSE de los estimadores MAP, MS y ML de una secuencia caótica generada usando los mapas PWL1 y Markov1. (a) Varianza y MSE para el mapa PWL1 con  $x[0] = 0,65251$  y  $N = 6$ . (b) Sesgo para el mapa PWL1 con  $x[0] = 0,65251$  y  $N = 6$ . (c) Varianza y MSE para el mapa Markov1 con  $x[0] = 0,22853$  y  $N = 4$ . (d) Sesgo para el mapa Markov1 con  $x[0] = 0,22853$  y  $N = 4$ .

A pesar de que las gráficas presentadas son ilustrativas del comportamiento típico de los estimadores desarrollados, un análisis riguroso de los mismos exige la comprobación de su rendimiento medio para un cierto mapa con unos parámetros fijos. En la Tabla 4.3 se muestra el MSE promedio de los estimadores ML, MAP y MS en función de la SNR para dos secuencias caóticas de longitudes cinco y nueve ( $N = 4$  y  $N = 8$  respectivamente) generadas usando el S-TM con  $\beta = 1,5$ . Para conformar dicha tabla se han seleccionado aleatoriamente 1.000 condiciones iniciales (para cada valor de  $N$ ) de acuerdo con la FDP invariante del mapa, y se han realizado 1.000 simulaciones de Monte Carlo para cada condición inicial y valor de SNR [Panta2000a].

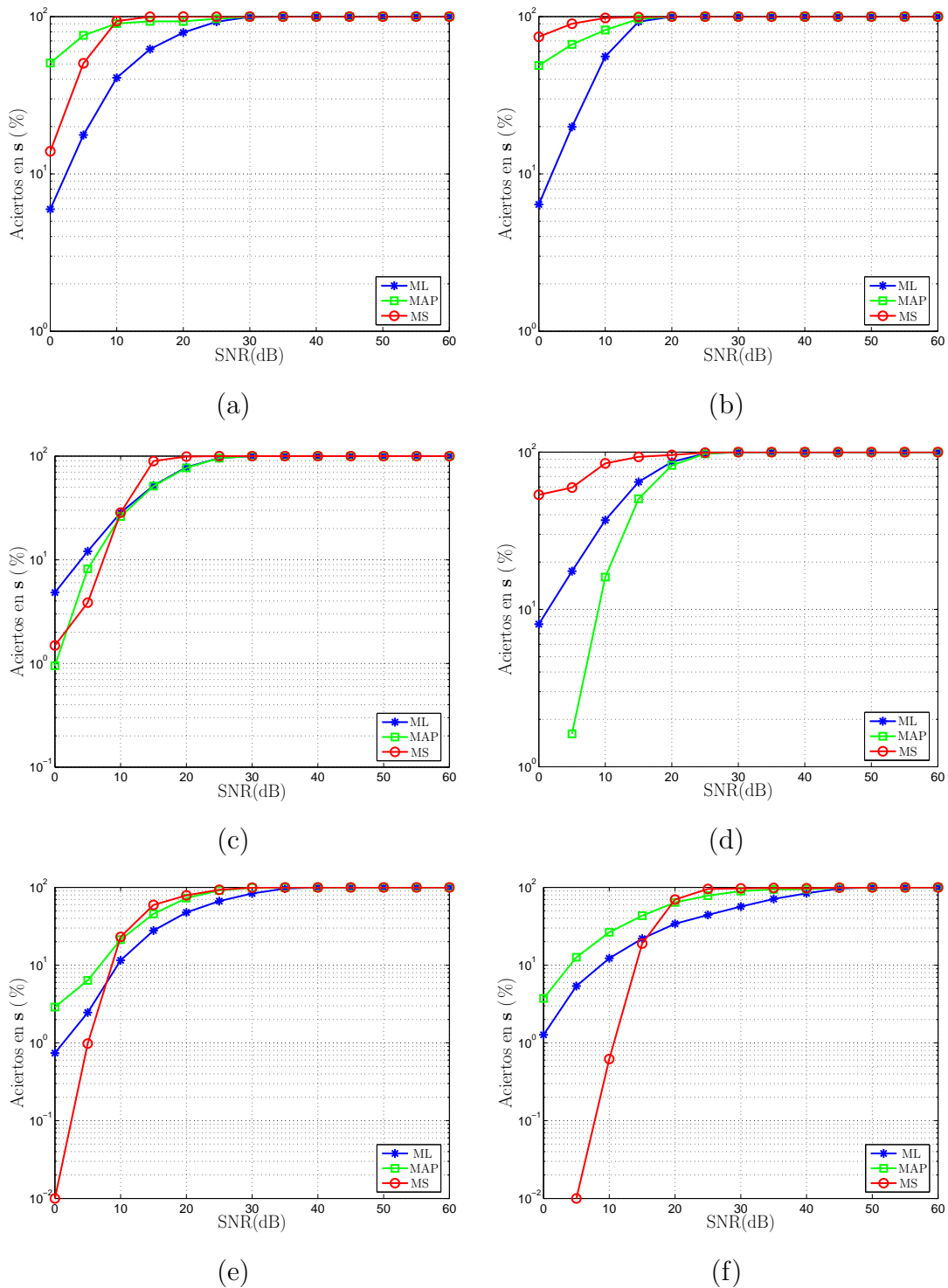


Figura 4.11: Tasa de aciertos en los símbolos del itinerario de los estimadores ML, MAP y MS para diferentes mapas caóticos y condiciones iniciales. (a) Mapa Markov2 con  $x[0] = 0,2389$ ,  $c = 0,75$  y  $N = 10$ . (b) Mapa Markov2 con  $x[0] = 0,4040$ ,  $c = 0,75$  y  $N = 10$ . (c) Mapa PWL1 con  $x[0] = 0,88131$  y  $N = 6$ . (d) Mapa PWL1 con  $x[0] = 0,65251$  y  $N = 6$ . (e) Mapa Markov1 con  $x[0] = 0,22853$  y  $N = 4$ . (f) Mapa Markov1 con  $x[0] = 0,89164$  y  $N = 4$ .

SNR (dB)	$-10 \log_{10}(\text{MSE})$					
	$N = 4$			$N = 8$		
	ML	MAP	MS	ML	MAP	MS
0	13,5	16,1	17,9	13,6	16,3	18,1
5	19,9	21,0	22,6	19,9	21,0	22,6
10	27,2	27,4	28,7	27,2	27,4	28,8
15	34,1	34,1	35,2	34,5	34,5	36,0
20	41,6	41,7	42,7	42,0	42,0	43,5
25	48,6	48,6	49,5	50,6	50,6	52,0
60	87,4	87,4	87,6	95,0	95,0	96,1

Tabla 4.3: Rendimiento (MSE) promedio de los estimadores ML, MAP y MS para el S-TM con  $\beta = 1,5$  y 1.000 condiciones iniciales seleccionadas de acuerdo a la aproximación PWC de su FDP invariante.

Se observa que ambos estimadores Bayesianos mejoran el rendimiento del estimador ML. Además, debido a la inconsistencia de los tres estimadores, el incremento en su rendimiento es muy pequeño al doblar el número de puntos, aunque se aprecia una mejora cada vez mayor conforme aumenta la SNR. Aunque en el aspecto cuantitativo los resultados mostrados en la Tabla 4.3 pueden variar en gran medida dependiendo del mapa caótico y sus parámetros, en el aspecto cualitativo las conclusiones obtenidas se pueden considerar generales. Un segundo ejemplo, en este caso para el mapa Markov2 con  $c = 0,75$  y  $N = 6$ , puede verse en [Panta2003].

Por último, reseñar brevemente que el comportamiento de los estimadores Bayesianos del resto de muestras de la secuencia es idéntico al observado en el caso del estimador ML: el CRLB empeora conforme se avanza en la secuencia, pero los estimadores se hallan más próximos al mismo, alcanzándolo en general para menores valores de SNR. Asimismo, el efecto de saturación apreciado en el caso del estimador ML aparece también en el caso de los estimadores Bayesianos. En consecuencia, estos estimadores son igualmente inconsistentes, a pesar de alcanzar el CRLB asintóticamente cuando la SNR tiende a infinito.

### 4.5.3. Resultados para Mapas no PWL

Aunque los estimadores MCMC propuestos en la Sección 4.4.2 son válidos para cualquier mapa caótico, en esta sección se llevan a cabo simulaciones únicamente para el mapa logístico con distintos valores de  $x[0]$ ,  $\lambda$  y  $N$ . No obstante, antes de evaluar el rendimiento de los estimadores basados en métodos MCMC es necesario determinar el número de muestras iniciales descartadas (longitud del periodo de “burn-in”),  $N_b$ , y el número de muestras usadas posteriormente para estimar la secuencia,  $N_m$ . Como ya se mencionó en la Sección 4.4.2.2, dicha tarea no se va a llevar a cabo analíticamente, sino mediante simulación.

En primer lugar, el valor de  $N_b$  parece no tener ningún efecto sobre el MSE final de los estimadores, excepto quizás para valores de SNR muy altos, y aún así es pe-

queño. Esto puede ser debido a que tanto  $x^0[0]$  como  $\mathbf{s}^0$  se generan de acuerdo con sus respectivas FDPs invariantes. En consecuencia, el algoritmo MH proporciona muestras de la FDP deseada desde un principio, y no es necesario ningún periodo de “burn-in”. Respecto al valor de  $N_m$ , en la Figura 4.12 se muestra el MSE de  $\hat{x}_{\text{MH-MS1}}[0]$ ,  $\hat{M}_0(\text{dB})$ , en función de  $N_m$  para distintos valores de SNR. Se puede apreciar cómo para SNRs bajas el valor de  $N_m$  apenas influye en la calidad de las estimas obtenidas. Sin embargo, conforme aumenta la SNR, el valor de  $N_m$  va cobrando cada vez una mayor importancia. Por ejemplo, para SNR=60 dB la ganancia por pasar de  $N_m = 1.000$  a  $N_m = 50.000$  muestras es de tan sólo 1,90 dB, pero al incrementar la SNR a 80 dB, la ganancia en  $N_m$  aumenta considerablemente hasta llegar a 10,40 dB.

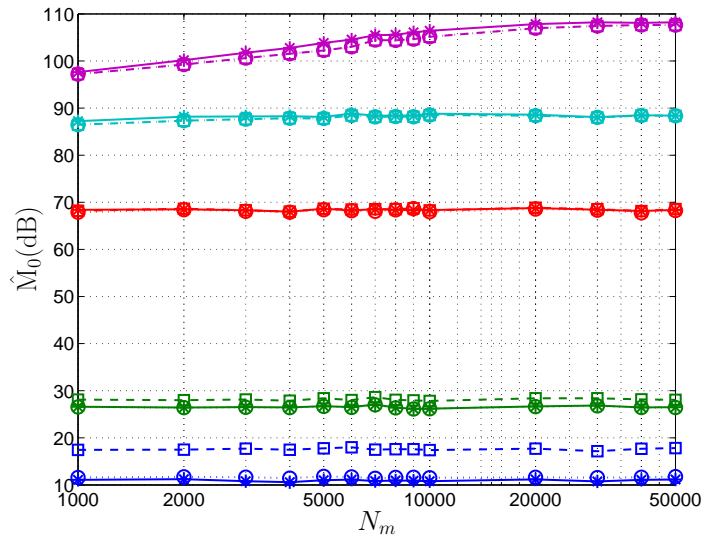


Figura 4.12: MSE de los distintos estimadores MCMC (MAP, ‘\*’; MS1, ‘□’; MS2, ‘o’) en función del número de muestras usadas para estimar  $x[0]$ ,  $N_m$ , para  $x[0] = 0,55$ ,  $\lambda = 4$ ,  $N = 5$ , y 1.000 simulaciones por cada valor de SNR. Cada color se corresponde con una determinada SNR: azul, SNR=0 dB; verde, SNR=20 dB; rojo, SNR=40 dB; cian, SNR=60 dB; magenta, SNR=80 dB.

Por lo tanto, como conclusión, a la vista de la Figura 4.12 se deduce que en general es suficiente con tomar  $N_m = 1.000$  para obtener buenos resultados para SNRs inferiores a 60 dB aproximadamente. Sin embargo, si se desea conseguir el rendimiento óptimo (esto es, alcanzar el CRLB) para valores de SNR superiores, es necesario incrementar  $N_m$  hasta valores de 10.000, 20.000 o incluso 50.000 muestras. Respecto al periodo de “burn-in”, en general puede tomarse  $N_b = 0$  sin ninguna pérdida de calidad de las estimas obtenidas, siendo mucho más importante utilizar un valor de  $N_m$  suficientemente elevado para conseguir un buen rendimiento.

Una vez estudiados los valores que deben tomar  $N_b$  y  $N_m$ , se comienza analizando en primer lugar el MSE de  $x[0]$  cuando  $\lambda < 4$ . En la Figura 4.13 se muestra  $\hat{M}_0(\text{dB})$  para dos valores de  $\lambda$ ,  $\lambda = 3,7$  y  $\lambda = 3,9$ , con diferentes condiciones iniciales, usando

$N_m = 1.000$ . Se puede apreciar cómo los estimadores MCMC alcanzan siempre el CRLB (las pequeñas discrepancias que aparecen en algunos casos para  $\text{SNR} \geq 40$  dB son achacables precisamente al reducido valor de  $N_b$  y  $N_m$ ), mejorando siempre el rendimiento del estimador HC-ML (que estima  $x[0]$  básicamente iterando hacia atrás a partir de la última muestra de la secuencia recibida usando el itinerario obtenido aplicando un umbral duro a la señal ruidosa). Como de costumbre, el punto exacto en que se alcanza el CRLB depende de los valores concretos de  $\lambda$  y  $x[0]$ , situándose entre 25 y 40 dB en general. La ganancia con respecto al estimador HC-ML también es muy variable, aunque en los dos ejemplos mostrados se consigue una buena ganancia para valores de SNR superiores a 35 dB.

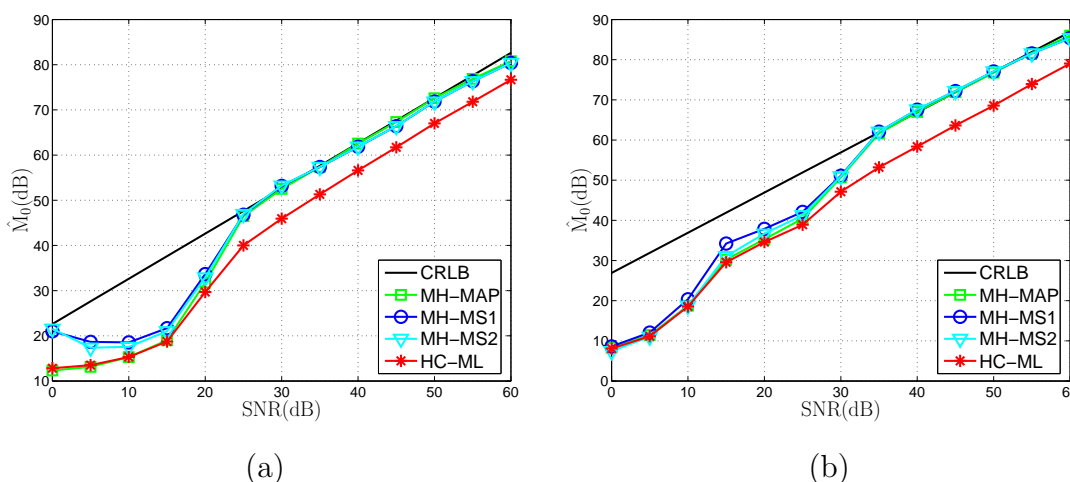


Figura 4.13: Comparación del rendimiento de los estimadores MCMC para secuencias generadas usando diferentes valores de  $x[0]$  y  $\lambda$  con  $N = 5$  y  $N_m = 1.000$ . (a)  $\lambda = 3,7$  y  $x[0] = 0,6917$ . (b)  $\lambda = 3,9$  y  $x[0] = 0,1334$ .

A continuación, en la Figura 4.14, se estudia el MSE de  $x[0]$  para  $\lambda = 4$ . Este valor de  $\lambda$  es el único para el que se conoce la FDP invariante natural del mapa, y además todas las secuencias simbólicas son válidas, presentan la misma probabilidad y la región de  $x[N]$  asociada a cada itinerario es el espacio de fases completo,  $[0, 1]$ . Todos estos factores permiten simplificar mucho el Algoritmo 4.4, contribuyendo a que su coste computacional sea aproximadamente un orden de magnitud inferior al del resto de valores de  $\lambda$ , lo que posibilita el uso de  $N_m = 10.000$  y la consiguiente obtención de mejores resultados. No obstante, las conclusiones son similares a las obtenidas para valores de  $\lambda$  inferiores: los estimadores basados en métodos MCMC alcanzan el CRLB a partir de una SNR que oscila entre 25 y 40 dB dependiendo del valor de  $x[0]$ , proporcionando una ganancia muy variable con respecto al estimador HC-ML. Por lo tanto, en lo sucesivo se va a utilizar únicamente  $\lambda = 4$  para el resto de las simulaciones.

En algunas ocasiones el objetivo puede ser estimar únicamente  $x[0]$ , pero en muchas otras se pretende estimar la secuencia completa,  $\mathbf{x}$ . Para medir la calidad de  $\hat{\mathbf{x}}$ , en lugar



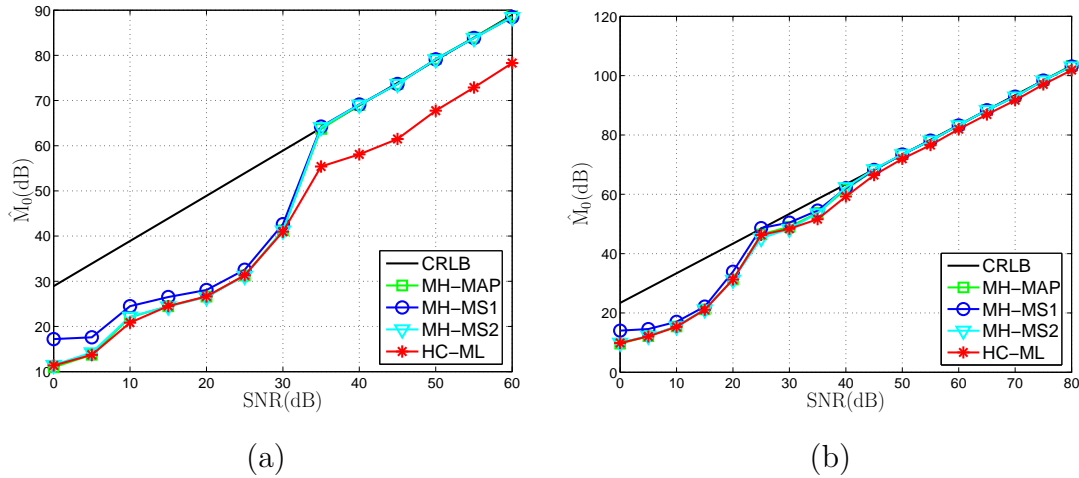


Figura 4.14: Comparación del MSE de  $x[0]$  de los estimadores MCMC para secuencias generadas usando diferentes valores de  $x[0]$  con  $\lambda = 4$ ,  $N = 5$ , y  $N_m = 10.000$ . (a)  $x[0] = 0,4485$ . (b)  $x[0] = 0,6880$ .

del MSE total se va a utilizar la SNR de la secuencia estimada, que se define como

$$\text{SNR}_{\hat{\mathbf{x}}}(\text{dB}) = \frac{\mathbf{x}^T \mathbf{x}}{\frac{1}{N_m} \sum_{i=1}^{N_m} J(\hat{\mathbf{x}}_i)} = \frac{\mathbf{x}^T \mathbf{x}}{\frac{1}{N_m} \sum_{i=1}^{N_m} (\mathbf{y}_i - \hat{\mathbf{x}}_i)^T (\mathbf{y}_i - \hat{\mathbf{x}}_i)},$$

donde  $\hat{\mathbf{x}}_i$  hace referencia a cada una de las  $N_m$  simulaciones de Monte Carlo. En la Figura 4.15 se muestra  $\text{SNR}_{\hat{\mathbf{x}}}(\text{dB})$  para los mismos ejemplos de la Figura 4.14. Se puede observar cómo los estimadores MCMC en general presentan una SNR de salida similar a la de entrada, mientras que la del estimador HC-ML es igual o inferior. Sin embargo, la mejora de los estimadores MCMC se reduce mucho con respecto a la obtenida para  $x[0]$ , no obteniéndose apenas ganancia alguna en el segundo de los casos.

Las simulaciones anteriores se han realizado para registros muy cortos ( $N = 5$ ). Aunque el coste computacional de los estimadores basados en el algoritmo MH es elevado, no es tan alto como el de los estimadores ML y Bayesianos exactos para mapas PWL presentados anteriormente, que requieren una búsqueda exhaustiva de todos los posibles itinerarios. En consecuencia, se pueden llevar a cabo simulaciones para secuencias más largas, como se muestra en la Figura 4.16, en la que se utiliza  $N = 29$ . En estos casos la ganancia de los estimadores MH-MS con respecto al HC-ML parece ser mínima, aunque sería necesario llevar a cabo un mayor número de simulaciones para confirmarlo. Nótese que para  $\text{SNR} \geq 40$  dB el CRLB de  $x[0]$  toma valores para los cuales pueden empezar a aparecer problemas de precisión numérica en el programa de simulación.

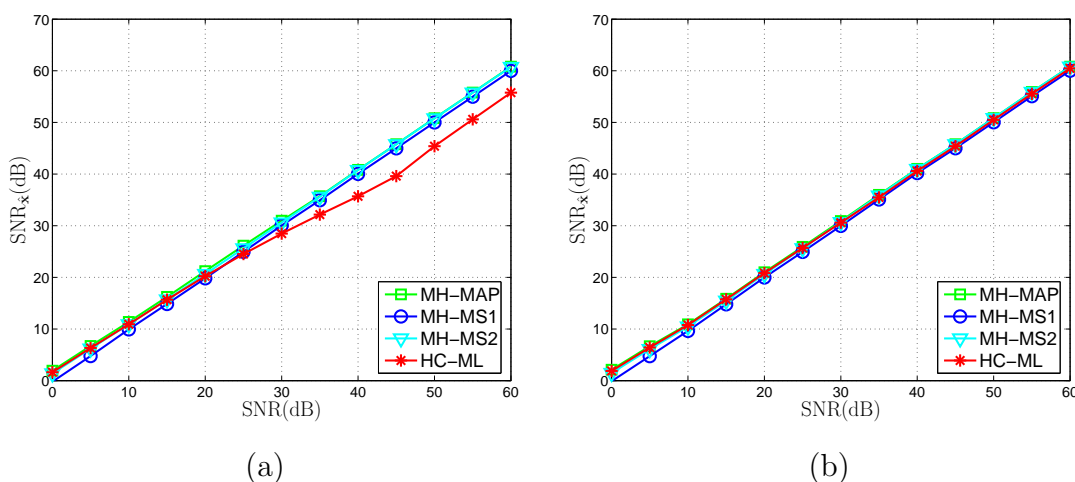


Figura 4.15: Comparación de la SNR de la secuencia estimada completa de los estimadores MCMC para secuencias generadas usando diferentes valores de  $x[0]$  con  $\lambda = 4$ ,  $N = 5$ , y  $N_m = 10.000$ . (a)  $x[0] = 0,4485$ . (b)  $x[0] = 0,6880$ .

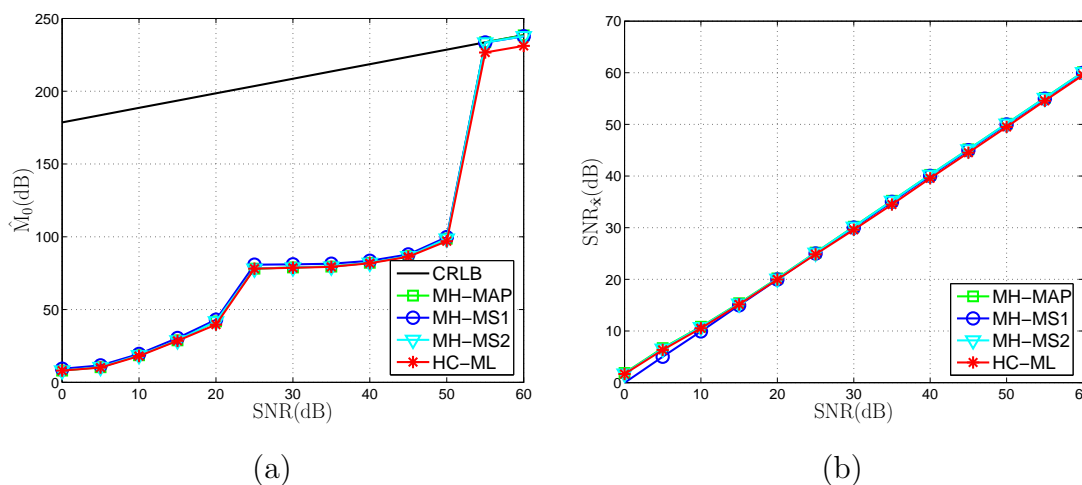


Figura 4.16: Comparación del MSE de  $x[0]$  y la SNR de  $\hat{\mathbf{x}}$  para los estimadores MCMC con una secuencia “larga” generada usando  $x[0] = 0,09849$ ,  $\lambda = 4$ ,  $N = 29$ , y  $N_m = 1.000$ . (a)  $\hat{M}_0$ (dB). (b)  $\text{SNR}_{\hat{\mathbf{x}}}$ (dB).

Para finalizar, al igual que en el caso de los estimadores Bayesianos de mapas PWL, el modo más razonable de evaluar la calidad de los estimadores MCMC consiste en llevar a cabo un promediado para un amplio conjunto de condiciones iniciales. En la Tabla 4.4 se muestran el MSE de  $x[0]$  y la SNR de  $\hat{\mathbf{x}}$  promediando 1.000 condiciones iniciales seleccionadas de acuerdo con la FDP invariante conocida del mapa logístico para  $\lambda = 4$ ,  $N_m = 1.000$ , y usando 1.000 simulaciones de Monte Carlo para cada valor de SNR [Luengo2002a].

Se puede ver cómo todos los estimadores basados en el algoritmo MH mejoran nota-

SNR (dB)	$\hat{M}_0(\text{dB}) / \text{SNR}_{\hat{x}}(\text{dB})$				CRLB
	HC-ML	MCMC			
		MH-MAP	MH-MS1	MH-MS2	
10	17,5 / 5,4	17,9 / 7,9	18,8 / 5,3	17,5 / 7,2	41,7
15	25,1 / 8,5	25,4 / 12,5	26,0 / 10,4	25,0 / 11,8	46,7
20	32,9 / 11,1	33,4 / 17,4	33,9 / 15,5	32,9 / 16,7	51,7
25	40,2 / 13,2	42,0 / 22,3	42,4 / 20,5	41,1 / 21,8	56,7
30	44,9 / 14,8	50,4 / 27,3	50,5 / 25,6	49,1 / 26,9	61,7
35	47,5 / 16,2	59,5 / 32,3	59,2 / 30,6	57,6 / 32,0	66,7
40	49,1 / 17,8	67,1 / 37,2	66,8 / 35,7	65,2 / 37,0	71,7
60	55,3 / 23,9	90,6 / 57,0	89,4 / 55,9	89,3 / 56,7	91,7

Tabla 4.4: MSE promedio de  $x[0]$ ,  $\hat{M}_0(\text{dB})$ , y SNR de la secuencia estimada completa,  $\text{SNR}_{\hat{x}}(\text{dB})$ , para el mapa logístico con  $\lambda = 4$ ,  $N = 5$  y 1.000 valores de  $x[0]$  obtenidos de acuerdo con su FDP invariante natural, (C.4). Se comparan los estimadores MCMC (MH-MAP, MH-MS1 y MH-MS2) con  $N_m = 1.000$ , el estimador HC-ML, y el CRLB promedio de  $x[0]$ .

blemente el rendimiento del HC-ML, cuyas prestaciones se ven gravemente penalizadas por unos pocos casos malos. El rendimiento de los estimadores MCMC se acerca cada vez más al CRLB conforme aumenta la SNR, aunque no llega a alcanzarlo para  $\text{SNR} = 60$  dB debido de nuevo a unos pocos valores de  $x[0]$  que requieren una SNR mayor, así como posiblemente al reducido valor de  $N_m$  usado. Por último, el hecho de que para valores altos de SNR el MH-MAP funcione mejor que los estimadores MH-MS es achacable de nuevo a la mayor sensibilidad de estos últimos frente al bajo valor de  $N_m$ .

## 4.6. Discusión

En este capítulo se ha considerado el problema de la estimación Bayesiana de secuencias generadas por mapas caóticos. Para mapas PWL se han podido obtener expresiones analíticas para los estimadores MAP y MS gracias a la utilización de una aproximación constante a tramos (PWC) de la FDP invariante del mapa, que se usa como FDP a priori, así como de las fórmulas cerradas para la iteración del mapa desarrolladas en el Capítulo 3. Para mapas no PWL la aproximación PWC de la FDP invariante sigue siendo válida como FDP a priori, pero en esta ocasión no se dispone de una expresión cerrada para la iteración del mapa. En consecuencia, al igual que ocurría con el estimador ML, no se puede encontrar una regla cerrada para los estimadores Bayesianos de mapas no PWL. En este caso, se puede recurrir a técnicas de rejilla y algoritmos iterativos locales, o utilizarse métodos MCMC, que permiten obtener muestras de la FDP a posteriori, y por lo tanto implementar cualquier estimador Bayesiano.

La evaluación del rendimiento de los estimadores Bayesianos se ha llevado a cabo del

mismo modo que la del estimador ML en el Capítulo 3: mediante simulaciones de Monte Carlo. Además, en este capítulo se han presentado resultados obtenidos promediando diversas condiciones iniciales, así como la probabilidad de error en los símbolos del itinerario. A partir de dichas simulaciones se han extraído las siguientes conclusiones:

1. Los estimadores Bayesianos mejoran en promedio el rendimiento del estimador ML (aunque para algún caso y algún valor de SNR pueden ofrecer peores prestaciones), alcanzando el CRLB asintóticamente cuando la SNR tiende a infinito.
2. El grado de mejora (ganancia) depende en gran medida del mapa caótico usado, sus parámetros, la longitud de la secuencia, y la condición inicial.
3. Los estimadores Bayesianos resultan potencialmente más beneficiosos (mayor margen de ganancia) para mapas con FDP invariante muy alejada de una FDP uniforme y  $x[0]$  situada dentro de una región de alta probabilidad, especialmente cerca del límite con una región de baja probabilidad.
4. Los estimadores Bayesianos no proporcionan ninguna ventaja con respecto al ML cuando la FDP invariante del mapa usado para generar la señal caótica es uniforme, y pueden dar lugar a un peor rendimiento cuando  $x[0]$  se halla situada dentro de una región de baja probabilidad y cercana al límite con una región de alta probabilidad.
5. Al igual que en el caso del estimador ML, el rendimiento de los estimadores Bayesianos se encuentra limitado fundamentalmente por su varianza, resultando asintóticamente insesgados. El sesgo puede empeorar el rendimiento únicamente para SNRs muy bajas.
6. En este caso sí que existe una diferencia importante entre mapas PWL markovianos y no markovianos: la FDP invariante de mapas PWL markovianos es constante dentro de las regiones de su partición natural, mientras que para mapas PWL no markovianos la FDP constante a tramos es únicamente una aproximación, que se debe calcular para cada longitud de la señal caótica y dentro de cada una de las regiones inducidas por el itinerario.
7. Se ha observado que los estimadores alcanzan el CRLB básicamente cuando la probabilidad de tener un error en los símbolos de su itinerario tiende a cero. En consecuencia, resulta evidente que para obtener una buena estima de una secuencia caótica resulta imprescindible disponer de una buena estima de su secuencia simbólica. Este hecho se va a explotar en el Capítulo 5 para plantear estimadores computacionalmente eficientes del itinerario a partir de los cuales se puedan encontrar buenos estimadores de la secuencia caótica.
8. Para mapas no PWL se puede plantear cualquier estimador Bayesiano mediante los métodos MCMC. Esta técnica proporciona muy buenos resultados, alcanzándose el CRLB asintóticamente, aunque presenta la misma desventaja que los estimadores Bayesianos de mapas PWL: un elevado coste computacional.

Finalmente, la extensión de los estimadores Bayesianos a mapas multidimensionales plantea esencialmente el mismo problema que en el caso del estimador ML: el coste computacional. Por consiguiente, el planteamiento de estimadores Bayesianos de mapas de mayor dimensión no se aborda en esta Tesis, considerándose nuevamente como una posible línea futura de investigación.