9

Asymptotically Optimal Estimators for Chaotic Digital Communications

David Luengo

Universidad Politécnica de Madrid (Spain)

Ignacio Santamaría

Universidad de Cantabria (Spain)

CONTENTS

Chaotic signals and systems offer the potential of increased security in digital communications. However, most of the approaches proposed either lack robustness at low SNRs (due to the difficulty of performing chaotic synchronization) or provide a much worse performance than classical techniques based on sinusoidal carrier functions. In this chapter we show how asymptotically optimal estimators, developed for the estimation of chaotic signals generated by discrete chaotic maps and corrupted by additive white Gaussian noise (AWGN), can be applied to improve the performance of digital chaotic communication schemes. First of all, after

a brief review of discrete-time chaotic maps and sequences, we derive the optimal maximum likelihood (ML) estimator for this problem. Unfortunately, its computational cost grows exponentially with the length of the chaotic sequence, thus rendering it unfeasible for moderate/large sequences. Therefore, asymptotically optimal estimators, based on well-known signal processing techniques, such as censoring approaches or the Viterbi algorithm (VA), with a reduced computational cost, are developed. Finally, we show how these methods can be applied to improve the performance of digital chaotic communications schemes based on the iteration of discrete-time chaotic maps, focusing on a recently proposed symbolic coding technique based on backward iteration.

9.1 Introduction: Chaotic Maps and Sequences

Chaotic signals are signals generated by purely deterministic systems that possess features typical of random signals. The dual deterministic/random nature of this type of signals renders them very interesting for a wide range of engineering applications: communications, time series modeling, criptography, watermarking, pseudorandom number generation, etc.

In this chapter we focus on chaotic sequences generated by the discrete-time iteration of unidimensional (1D) piecewise linear (PWL) chaotic maps. Although this choice may look too restritive, one-dimensional discrete-time chaotic maps, described by a non-linear difference equation, seem to possess all the interesting features of higher-dimensional continuous-time systems, defined through non-linear differential equations [3, 5, 37]. Furthermore, it has been shown that 1D PWL maps exhibit the same types of dynamic behaviours as any other one-dimensional chaotic maps, and a particular type of 1D PWL maps (Markov PWL maps) can approximate a wide range of non-PWL maps with an arbitrary accuracy [10, 9]. Hence, many works focus on the one-dimensional case, which is much easier to analyze, and has been exploited also in many practical applications. In the remaining of this section we briefly review the type of chaotic signals and systems that we consider in this chapter: discrete-time chaotic sequences obtained iterating piecewise linear maps.

9.1.1 One-Dimensional Piecewise Linear Chaotic Maps

Here we define a *unidimensional chaotic map* as a non-linear application from an interval I onto the same interval,[1] $f : I \to I$, that fulfills the following three conditions [6]:

1. f has *sensitive dependence on initial conditions*, i.e., there exists a $\delta > 0$ such that for every $x \in I$ and neighbourhood of x, $N_\delta(x) = \{x' : d(x, x') \leq \delta\}$ with $d(\cdot, \cdot)$ denoting any appropriate distance function, there is some $x' \in N_\delta(x)$ and $n > 0$ fulfilling that $f^n(x') \notin N_\delta(f^n(x))$.

2. f is *topologically transitive*, i.e., for any pair of open sets $U, V \in I$ there exists an $n > 0$ such that $f^n(U) \cap V \neq \emptyset$.

3. The *periodic points* of f are *dense* in I, i.e., given the set of periodic points of f inside I, $P = \{x : f^n(x) = x \text{ for } n = 1, 2, 3, \ldots\}$, then $\overline{P} = I$ with $\overline{P}$ denoting the closure of P. [2]

A *piecewise linear (PWL) map* is a particular class of chaotic map which is defined by a different affine transformation inside each of the M intervals into which I is subdivided. Mathematically,

$$y = f(x) = \sum_{i=1}^{M} (a_i x + b_i)\chi_{E_i}(x), \tag{9.1}$$

where $I = [e_0, e_M]$ is the domain of the map (usually $I = [-1, 1]$ in our case), $E_i = [e_{i-1}, e_i)$ for $1 \leq i \leq M - 1$ and $E_M = [e_{M-1}, e_M]$ (with $e_0 < e_1 < \ldots < e_M$) are the M linear intervals of the map, a_i and b_i are the slope and offset of the line that characterizes the map inside the i-th interval respectively, and

$$\chi_R(x) = \begin{cases} 1, & x \in R; \\ 0, & x \notin R. \end{cases} \tag{9.2}$$

is the *characteristic function* of region R, that indicates whether a point x belongs to it or not.

PWL maps are probably the class of 1D chaotic maps most widely used, due to their simplicity and mathematical tractability. As a first example of a PWL map, we consider the skew tent-map (SK-TM), which

[1] A function $f : I \to I$ is *onto* if for every $y \in I$ there is an $x \in I$ such that $f(x) = y$ [6].

[2] Let P be an open subset of I. Its closure, $\overline{P}$, is defined as the set containing all the points in P altogether with all the limit points of P [6].

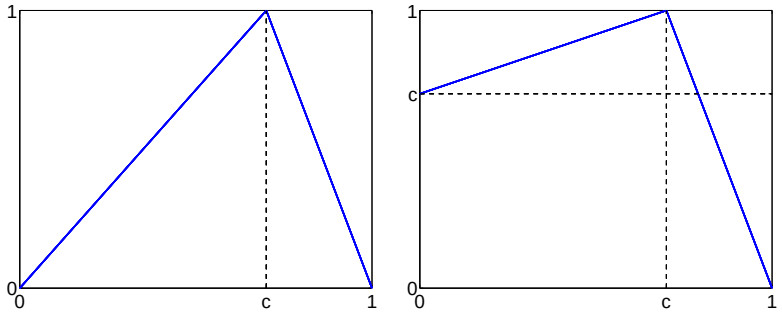
is one of the most popular 1D chaotic maps, having been used for cryptography [21], watermarking [22], digital communications [7], etc. The SK-TM has a single parameter, $0 < c < 1$, a phase space $I = [0, 1]$ and two intervals: $E_1 = [0, c)$ with $a_1 = 1/c$ and $b_1 = 0$, and $E_2 = [c, 1]$ with $a_2 = -1/(1-c)$ and $b_2 = 1/(1-c)$. Mathematically, the equation defining the SK-TM can be expressed as

$$f(x) = \begin{cases} \frac{x}{c}, & 0 \le x < c; \\ \frac{1-x}{1-c}, & c \le x \le 1. \end{cases} \tag{9.3}$$

The shape of the SK-TM is shown in Figure 9.1(a). As a second example, we consider the following unimodal map introduced in [29]:

$$f(x) = \begin{cases} \frac{1-c}{c}x + c, & 0 \le x < c; \\ \frac{1-x}{1-c}, & c \le x \le 1. \end{cases} \tag{9.4}$$

This map still has a single parameter, $0 < c < 1$, a phase space $I = [0, 1]$ and two intervals: $E_1 = [0, c)$ with $a_1 = (1-c)/c$ and $b_1 = c$, and $E_2 = [c, 1]$ with $a_2 = -1/(1-c)$ and $b_2 = 1/(1-c)$. The shape of this unimodal map is shown in Figure 9.1(b). We note that the second interval is identical to that of the SK-TM, but the line in the first one has a non-zero offset and a smaller slope than the corresponding line in the SK-TM.



(a) SK-TM (b) Unimodal PWL map

FIGURE 9.1
Examples of PWL maps: skew tent-map (SK-TM) and unimodal PWL map introduced in [29].

9.1.2 Chaotic Sequences: Forward and Backward Iteration

Chaotic sequences are obtained by the repeated application of the nonlinear function $f(x)$ on the output of the previous iteration. Hence, the n-th sample of the chaotic sequence is given by

$$x[n] = f(x[n-1];\boldsymbol{\theta}) = f^2(x[n-2];\boldsymbol{\theta}) = \ldots = f^n(x[0];\boldsymbol{\theta}), \tag{9.5}$$

where $\boldsymbol{\theta}$ is the vector of parameters that characterizes the chaotic map, $x[0] \in I$ is the initial condition that defines the evolution of the whole sequence, and

$$f^k(x) = \underbrace{f \circ f \circ \cdots \circ f}_{k \text{ times}}(x) = \underbrace{f(f(\cdots f(f(}_{k \text{ times}} x[n])) \cdots)) \tag{9.6}$$

denotes the k-th functional composition of $f(x)$, with $f^0(x) = x$. Chaotic sequences generated using (9.5) and (9.6) are said to be obtained through *forward iteration* starting from an initial condition $x[0] \in I$. Now, let us assume that we want to generate N samples from a chaotic map starting from an initial condition $x[0] \in I$. Making use of (9.5), the *forward orbit* or *trajectory* of length N of $x[0]$ is the ordered set of $N+1$ points generated including $x[0]$, i.e.,[3]

$$O^+_{f,N}(x[0]) = \{x[0],\ f(x[0]),\ f^2(x[0]),\ \ldots,\ f^N(x[0])\}. \tag{9.7}$$

Examples of chaotic orbits generated by the two maps displayed in Figure 9.1 are shown in Figure 9.2. The irregular aspect of both signals can be clearly appreciated, although each one shows different characteristic patterns that repeat themselves in an approximate way.

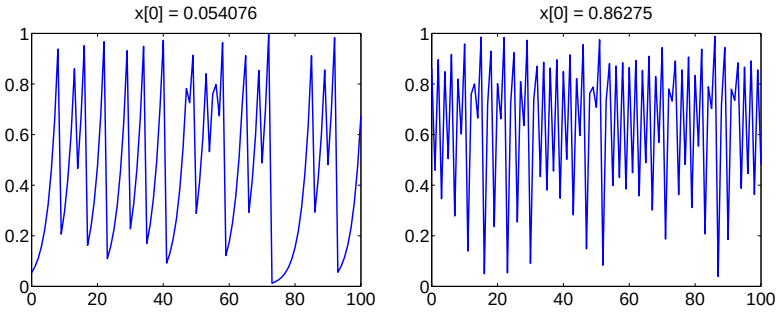
Alternatively, a length $N+1$ chaotic sequence may be generated through *backward iteration* starting from a final condition $x[N]$. However, since 1D chaotic maps must be non-invertible,[4] in order to define the backward iteration of the map we must introduce first the concept of symbolic sequences. Let us consider a 1D chaotic map $f(x)$ and a *partition* of its domain or phase space, $\mathcal{P} = \{J_1,\ J_2,\ \cdots,\ J_P\}$, such that $\cup_{i=1}^{P} J_i = I$ and $J_i \cup J_j = \emptyset$ for $1 \leq i, j \leq P$ and $i \neq j$. Now, given an initial condition $x[0] \in I$, we may define the length N *symbolic sequence* associated to $x[0]$ as

$$S^+_{f,N}(x[0]) = \{s[0],\ s[1],\ \ldots,\ s[N-1]\} \tag{9.8}$$

$$= \{\sigma(x[0]),\ \sigma(x[1]),\ \ldots,\ \sigma(x[N-1])\}, \tag{9.9}$$

[3] For the sake of simplicity, in the sequel we often remove the dependence on $\boldsymbol{\theta}$ of $f(x)$.

[4] Invertible one-dimensional maps lead only to simple dynamical behaviours and can never produce chaotic dynamics [6].



(a) SK-TM (b) Unimodal PWL map

FIGURE 9.2
Examples of chaotic orbits generated by two PWL maps: skew tent-map (SK-TM) and unimodal PWL map introduced in [29].

where $s[n] = \sigma(x[n])$ denotes the n-th symbol for $0 \leq n \leq N-1$, and $\sigma : I \rightarrow \Sigma$ is the function that relates the phase space of the chaotic map, $I = [e_0, e_M]$, to its symbolic space, $\Sigma = \{1,\ 2,\ \ldots,\ P\}$, which is known as a *symbolic dynamics*, and can be expressed as

$$s[n] = \sigma(x[n]) = \sum_{i=1}^{P} i \cdot \chi_{J_i}(x[n]). \tag{9.10}$$

Obviously, infinitely many different symbolic dynamics can be defined for a given 1D chaotic map. The most frequently used partition is the so-called *natural partition*, $\mathcal{P}_N$, which is composed of the minimum number of regions where $f(x)$ is monotonic and continuous, and thus also invertible. Using the natural partition, any 1D chaotic map can be expressed as

$$y = f(x) = \sum_{i=1}^{P} f_i(x)\chi_{J_i}(x), \tag{9.11}$$

where $f_i(x)$ is the monotonic and continuous function that describes the map inside the i-th interval of the natural partition. For a given chaotic map, the natural partition is unique, can be easily obtained, and allows us to define a piecewise inverse function as

$$x = f^{-1}(y) = \sum_{i=1}^{P} f_i^{-1}(y)\chi_{J_i}(x), \tag{9.12}$$

with $f_i^{-1}(y)$ denoting the inverse of $f_i(x)$ inside the i-th interval of $\mathcal{P}_N$. For the particular class of chaotic maps considered here, PWL maps defined by (9.1), the natural partition is clearly given by $\mathcal{P}_N = \{E_1,\ E_2,\ \ldots,\ E_M\}$, and the one-step backward iteration can be expressed as

$$x = f^{-1}(y) = \sum_{i=1}^{M} \frac{y - b_i}{a_i} \chi_{E_i}(x), \tag{9.13}$$

which corresponds to another PWL map with slopes $1/a_i$ and offsets $-b_i/a_i$. Note that this backward iteration requires prior knowledge of the region of the natural partition to which the generated sample belongs. Denoting $s = \sigma(x)$, (9.13) can be expressed more compactly as

$$x = f_s^{-1}(y) = \frac{y - b_s}{a_s}. \tag{9.14}$$

Unfortunately, the natural partition is too restrictive for the digital communications application considered in this chapter. Consequently, we define a generalization of the natural partition, that we call the *invertible partition* of the map, $\mathcal{P}_I$, which is composed of the minimum number of connected intervals inside which $f(x)$ is invertible. Although for many chaotic maps (e.g. continuous maps) the invertible partition is identical to the natural partition, it allows us to consider partitions with regions inside which the map is not continuos or even not monotonic as long as the map is still invertible. We remark that, even though the invertible partition cannot be expressed as compactly as the natural partition for a generic PWL map, it can also be easily found, since each interval will be formed by the union of one or more consecutive intervals of the PWL map. For example, for the second map used in Section 9.3, given by (9.45), the natural partition is composed of $M = 5$ intervals, $\mathcal{P}_N = \{E_1,\ E_2,\ E_3,\ E_4,\ E_5\}$, whereas the invertible partion contains only three intervals, $\mathcal{P}_I = \{E_1 \cup E_2,\ E_3,\ E_4 \cup E_5\}$.

Using the symbolic sequence induced either by the natural or the invertible partition, the n-th sample of a chaotic sequence is alternatively given by

$$x[n] = f_{s[n]}^{-1}(x[n+1];\boldsymbol{\theta}) = f_{s[n],s[n+1]}^{-2}(x[n+2];\boldsymbol{\theta}) = \ldots = f_{s[n],\ldots,s[N-1]}^{-(N-n)}(x[N];\boldsymbol{\theta}), \tag{9.15}$$

where the subindex of f indicates the portion of the symbolic sequence required for the backward iteration and $f^{-k}(x)$ denotes the k-th functional composition of $f^{-1}(x)$. Similarly, making use of (9.15) we can define the *backward orbit* or *trajectory* of length N of $x[N]$ as the ordered set of $N+1$ points generated including $x[N]$, i.e.,

$$O_{f,N}^{-}(x[N]) = \{x[N],\ f_{s[N-1]}^{-1}(x[N]),\ \ldots,\ f_{s[0],\ldots,s[N-1]}^{-N}(x[N])\}, \tag{9.16}$$

which corresponds to the forward trajectory given by (9.7) in reverse order.

Finally, we remark again that the symbolic sequence must also be specified when determining the backward orbit of length N of $x[N]$. Therefore, a PWL map with M intervals may have up to M^N backward orbits for a given value of $x[N]$. In fact, the number of backward orbits can actually be much less than M^N, since some symbolic sequences may be invalid. For instance, for the chaotic map shown in Figure 9.1(b), since $f(E_1) = E_2$ (i.e., the first interval is mapped onto the second) we must necessarily have $s[n+1] = 2$ whenever $s[n] = 1$, implying that all the symbolic sequences containing $\{\ldots, 1, 1, \ldots\}$ will be invalid. Denoting the set of valid symbolic sequences of length N for a given map as $\mathcal{S}_N$, the number of possible backward orbits for a given value of $x[N]$ is $\Gamma_s(N) = |\mathcal{S}_N|$, with $|\mathcal{A}|$ indicating the cardinality of set $\mathcal{A}$.

9.2 Optimal and Suboptimal Estimation of Chaotic Sequences

9.2.1 Problem Formulation

The problem considered in this section is estimating a sequence of samples generated iterating a 1D PWL chaotic map, given observations corrupted by additive white Gaussian noise (AWGN). Mathematically, let us consider an $N+1$ column vector containing the $N+1$ samples of the forward orbit of length N of $x[0]$, or equivalently the $N+1$ samples of the backward orbit of $x[N]$ in reverse order,[5]

$$\begin{aligned}\boldsymbol{x} &= [x[0],\ x[1],\ \ldots,\ x[N-1],\ x[N]]^\top \\ &= [x[0],\ f(x[0]),\ f^2(x[0]),\ \ldots,\ f^{N-1}(x[0]),\ f^N(x[0])]^\top \\ &= [f_{\boldsymbol{s}}^{-N}(x[N]),\ f_{\boldsymbol{s}}^{-(N-1)}(x[N]),\ \ldots,\ f_{s[N-1]}^{-1}(x[N]),\ x[N]]^\top, \quad (9.17)\end{aligned}$$

where $\boldsymbol{s} = [s[0],\ s[1],\ \ldots,\ s[N-1]]^\top$ is the length N column vector with the symbolic sequence associated to the orbit of length N of $x[N]$,[6] and $\boldsymbol{s}_{n:N} = [s[n-1],\ s[n],\ \ldots,\ s[N-1]]^\top$ for $1 \leq n \leq N$ is the column

[5] As shown in [14], any point in the chaotic sequence can be used as a reference, mixing forward and backward iteration to obtain the remaining samples. However, in order to simplify the discussion, here we only consider the first and the last samples of the chaotic sequence, $x[0]$ and $x[N]$ respectively.

[6] Note that $s[N]$ is not included in $\boldsymbol{s}$ since it is not required for the backward iteration.

vector containing the last $N-n+1$ samples from the symbolic sequence. The observed sequence is $\boldsymbol{y} = [y[0],\ y[1],\ \ldots,\ y[N-1],\ y[N]]^\top$ with $y[n] = x[n] + w[n]$ and $w[n] \sim \mathcal{N}(0, \sigma^2)$ for $0 \leq n \leq N$.[7] The goal is obtaining an accurate (i.e., unbiased) and efficient (i.e., with a variance decreasing as a function of N as fast as possible) estimator of the original chaotic sequence.

In the following section we develop the optimal maximum likelihood estimator (MLE) for this problem.[8] Unfortunately, the computational cost of the MLE increases exponentially with the length of the sequence. Consequently, in Section 9.2.3 we describe two simple and asymptotically efficient estimators that attain the Cramer-Rao lower bound (CRLB) as the signal to noise ratio (SNR) tends to infinity. The CRLB, described in the Appendix, provides us with a lower bound on the variance on any unbiased estimator, allowing us to quantify precisely the concept of an efficient estimator as the one attaining the CRLB [11]. Hence, the two estimators described in Section 9.2.3 can be considered asymptotically optimal in the sense of making the most efficient use of all the information available, at least for large values of SNR.

9.2.2 Maximum Likelihood Estimator

The maximum likelihood estimator (MLE) of chaotic sequences corrupted by AWGN was formulated originally in [23], where two suboptimal approaches for finding the MLE, based on dynamic programming and the Kalman filter respectively, were proposed. Then, Papadopoulos and Wornell were the first ones to provide an algorithm that achieved the exact MLE for a particular chaotic map, the tent-map (TM) with $\beta = 2$ [30, 31]. Unfortunately, although this method can be easily applied to the TM with other values of β or to other chaotic maps, in general the estimator obtained will not be the MLE. The exact MLE for generic PWL maps was developed independently by Schimming *et al.* [36, 35] and Pantaleón *et al.* [26]. Finally, an algorithm to attain the MLE for non-PWL maps based on Markov chain Monte Carlo (MCMC) methods was proposed in [16]. In this section we develop the exact MLE following the description of [26, 14].

[7] We use the notation $w[n] \sim \mathcal{N}(\mu, \sigma^2)$ to indicate that $w[n]$ is a sample from a Gaussian distribution with mean μ and variance σ^2.

[8] Bayesian estimators have also been developed for this problem [25, 29, 14], but they will not be discussed here, since they do not provide any advantage for the chaotic communications approach described in this chapter.

9.2.2.1 Standard Problem Formulation

Formally, the MLE is obtained solving the following maximization problem [11, 38]:

$$\begin{aligned} &\hat{\boldsymbol{x}}_{\mathrm{ML}} = \arg\max_{\boldsymbol{x}} p(\boldsymbol{y};\boldsymbol{x}), \\ &\text{s.t.} \quad x[0] \in [e_0, e_M],\ x[n] = f(x[n-1]) \quad \text{for } 1 \le n \le N, \end{aligned} \tag{9.18}$$

where the constraints are imposed by the nature of the chaotic sequence, and $p(\boldsymbol{y};\boldsymbol{x})$ is the *likelihood*, i.e., the probability density function (PDF) of the observations conditioned on the parameters to be estimated, in this case the underlying noiseless chaotic sequence. Since the noise samples are Gaussian and independent, the likelihood is a multivariate Gaussian PDF,

$$p(\boldsymbol{y};\boldsymbol{x}) = (2\pi\sigma^2)^{-(N+1)/2} \exp\left(-\frac{1}{2\sigma^2}(\boldsymbol{y}-\boldsymbol{x})^\top(\boldsymbol{y}-\boldsymbol{x})\right). \tag{9.19}$$

Moreover, since all the samples of the chaotic sequence can be obtained from $x[0]$ through forward iteration, it is straightforward to show that the MLE of the whole sequence can be reformulated in terms of the MLE of the initial condition,

$$\begin{aligned} &\hat{x}_{\mathrm{ML}}[0] = \arg\min_{x[0]} J(x[0]), \\ &\text{s.t.} \quad x[0] \in [e_0, e_M], \end{aligned} \tag{9.20}$$

with

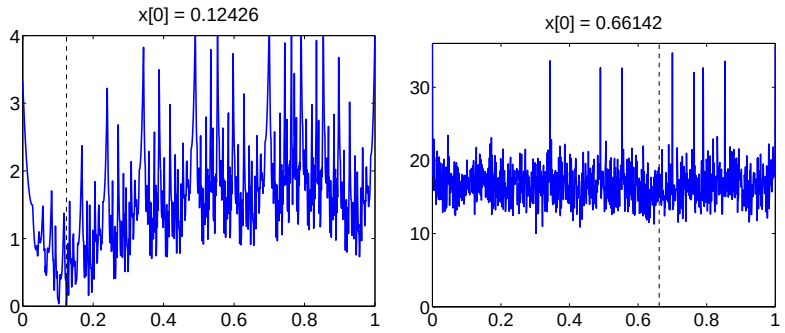
$$J(x[0]) = \sum_{n=0}^{N} \left(y[n] - f^n(x[0])\right)^2 \tag{9.21}$$

indicating the quadratic error between the observations and the chaotic sequence. Hence, in this case the MLE of $x[0]$ and its least squares (LS) estimator are equivalent, a result which is well-known for parameters observed in AWGN. The remaining samples of the chaotic sequence can be obtained by forward iteration from $\hat{x}_{\mathrm{ML}}[0]$, thanks to the invariance property of the MLE [11], resulting in an ML estimate of the whole sequence

$$\hat{\boldsymbol{x}}_{\mathrm{ML}} = [\hat{x}_{\mathrm{ML}}[0],\ f(\hat{x}_{\mathrm{ML}}[0]),\ \ldots,\ f^{N-1}(\hat{x}_{\mathrm{ML}}[0]),\ f^N(\hat{x}_{\mathrm{ML}}[0])]^\top. \tag{9.22}$$

Unfortunately, although the formulation of this problem looks deceptively simple, solving it efficiently is extremely complicated, due to the highly non-linear dependence of the samples of the chaotic sequence on the initial condition. In fact, the cost function given by (9.21) is extremely rugged, with fractal characteristics (see [14] for a more detailed

discussion of this issue) and multiple minima and maxima. As an example, Figure 9.3 shows the cost function for the SK-TM with $N = 10$ and $N = 100$ and two randomly selected values of $x[0]$. Obviously, given the shape of the cost function, any iterative or grid search approach is doomed to get stuck in a local minimum, especially for large values of N. Furthermore, although the exact MLE of $x[0]$ can be obtained as shown in [25, 26], it requires searching for a set of symbolic regions, $R_{\boldsymbol{s}_i}$, associated to the portion of the phase space where the initial condition must lie in order to have $\boldsymbol{s} = \boldsymbol{s}_i$. Finding these regions becomes more involved as N grows, since their size decreases exponentially and the forward iteration of chaotic maps is numerically unstable. Hence, in the following section we introduce an alternative formulation based on backward iteration that avoids all these problems: the cost function is quadratic in $x[N]$, backward iteration is numerically stable and, whenever $\Gamma_s(N) = M^N$ as it happens for most of the maps used in practice, the symbolic regions agree with the phase space of the chaotic map.



(a) SK-TM with $N = 10$ (b) SK-TM with $N = 100$

FIGURE 9.3
Examples of $J(x[0])$ for the SK-TM with $N = 10$ and $N = 100$.

9.2.2.2 Alternative Problem Formulation

For chaotic sequences, since all the samples are related to $x[N]$ through backward iteration with the appropriate symbolic sequence, the MLE

can be reformulated in a more practical way:

$$[\hat{x}_{\mathrm{ML}}[N],\ \hat{\boldsymbol{s}}_{\mathrm{ML}}]^{\top} = \underset{x[N],\boldsymbol{s}}{\arg\min}\ J(x[N],\boldsymbol{s}),$$
$$\text{s.t.}\quad x[N] \in [e_0, e_M],\ \boldsymbol{s} \in \mathcal{S}_N, \tag{9.23}$$

where $\mathcal{S}_N$ is again the set of valid symbolic sequences of length N, and

$$J(x[N],\boldsymbol{s}) = \sum_{n=0}^{N} \big(y[N-n] - f_{\boldsymbol{s}}^{-n}(x[N])\big)^2. \tag{9.24}$$

Hence, the MLE of the chaotic sequence can actually be obtained by estimating the last sample of the sequence, $x[N]$, and the N samples of the symbolic sequence: $s[0],\ s[1],\ \ldots,\ s[N-2],\ s[N-1]$.

Note that, although this formulation may look more complicated than the standard one, for PWL maps it leads to much simpler expressions for the MLE. Indeed, for PWL maps it can be shown that the n-th iteration backwards starting from $x[N]$ for $1 \le n \le N$ is given by [15, 14][9]

$$x[N-n] = f_{\boldsymbol{s}}^{-n}(x[N]) = B^{1,n}_{\boldsymbol{s}_{N-n+1:N}} x[N] - \sum_{m=1}^{n} B^{m,n}_{\boldsymbol{s}_{N-n+1:N-m+1}} b_{s[N-m]}, \tag{9.25}$$

with $B^{1,0}_{\boldsymbol{s}} = 1$ and

$$B^{m,n}_{\boldsymbol{s}_{N-n+1:N}} = \prod_{\ell=m}^{n} a^{-1}_{s[N-\ell]}, \tag{9.26}$$

for $1 \le n \le N$ and $1 \le m \le n$. Inserting (9.26) into (9.24), the cost function for the MLE of a PWL map finally becomes

$$J(x[N],\boldsymbol{s}) = \sum_{n=0}^{N} \big(\gamma_{\boldsymbol{s}}[N-n] - B^{1,n}_{\boldsymbol{s}} x[N]\big)^2, \tag{9.27}$$

where, instead of the precise samples of the symbolic sequence involved, we have used $\boldsymbol{s}$ in all cases to simplify the notation, and

$$\gamma_{\boldsymbol{s}}[N-n] = y[N-n] + \sum_{m=1}^{n} B^{m,n}_{\boldsymbol{s}} b_{s[N-m]}, \qquad \text{for}\quad 0 \le n \le N. \tag{9.28}$$

[9] Closed-form equations for the forward iteration of the tent-map (TM) were developed originally in [25], and extended to a generic PWL map in [26]. Regarding, backward iteration, analytical expressions were developed first for the polar SK-TM in [27] and extended to generic PWL maps in [29]. All these formulas for the forward and backward iteration of a generic PWL map starting from an arbitrary reference sample, $x[n]$, have been compiled in [14]. Furthermore, efficient algorithms for their implementation are also provided in [14].

In the next section we show how, using this alternative formulation and the analytical expressions for the backward iteration of PWL maps, we can get rid of the fractal cost functions that appear in the forward iteration and obtain a consistent estimator of $x[N]$ that can be expressed in a closed-form.

9.2.2.3 Exact Maximum Likelihood Estimator

Looking back at the cost function of the maximum likelihood estimator, we notice that, for a given symbolic sequence, $\boldsymbol{s} = \boldsymbol{s}_i \in \mathcal{S}$, (9.24) is quadratic in $x[N]$. Hence, taking the derivative of (9.24) w.r.t. $x[N]$ and equating it to zero we can easily obtain the estimate of $x[N]$ associated to $\boldsymbol{s} = \boldsymbol{s}_i$:

$$\hat{x}_i[N] = \frac{\sum_{n=0}^{N} B_{\boldsymbol{s}_i}^{1,n} \gamma_{\boldsymbol{s}_i}[N-n]}{\sum_{n=0}^{N} \left(B_{\boldsymbol{s}_i}^{1,n}\right)^2}. \tag{9.29}$$

Note that (9.29) is not the ML estimate of $x[N]$ for the i-th symbolic sequence yet, since there is no guarantee that $\hat{x}_i[N]$ belongs to $I = [e_0, e_M]$. However, assuming that all the symbolic sequences are valid (i.e., $\Gamma_s(N) = |\mathcal{S}_N| = M^N$), the ML estimator of $x[N]$ corresponding to the i-th symbolic sequence is obtained simply by thresholding (9.29):[10]

$$\hat{x}_{\mathrm{ML}}^{i}[N] = \begin{cases} e_0, & \hat{x}_i[N] < e_0; \\ \hat{x}_i[N], & e_0 \leq \hat{x}_i[N] \leq e_M; \\ e_M, & \hat{x}_i[N] > e_M. \end{cases} \tag{9.30}$$

Finally, thanks to the invariance property of the ML estimator [11], the ML estimate of the rest of the sequence for the i-th symbolic sequence can be obtained simply by iterating backwards from $\hat{x}_{\mathrm{ML}}^{i}[N]$ using the i-th symbolic sequence, $\boldsymbol{s}_i$, i.e.,

$$\hat{x}_{\mathrm{ML}}^{i}[N-n] = f_{\boldsymbol{s}_i}^{-n}(\hat{x}_{\mathrm{ML}}^{i}[N]), \qquad \text{for} \quad 1 \leq n \leq N. \tag{9.31}$$

Unfortunately, the estimation of the optimal symbolic sequence is an NP hard problem in the general case, implying that algorithms for obtaining the exact MLE in polynomial time cannot be developed except for some particular cases, such as the tent-map (TM) with $\beta = 2$ [31]. Therefore, the only solution that guarantees that the MLE of the symbolic sequence is achieved for a generic PWL map is an *exhaustive search* or *brute force approach*: testing all the valid symbolic sequences and selecting the one that minimizes the cost function. Mathematically, the ML

[10] When some symbolic sequences are invalid the same approach can be followed, but the limits for the thresholding operation may depend on the symbolic sequence [14].

estimators of $x[N]$ and the symbolic sequence are $\hat{x}_{\mathrm{ML}}[N] = \hat{x}^r_{\mathrm{ML}}[N]$ and $\hat{\boldsymbol{s}}_{\mathrm{ML}} = \boldsymbol{s}_r$ respectively, with $\hat{x}^r_{\mathrm{ML}}[N]$ being the ML estimator associated to the r-th valid symbolic sequence, $\boldsymbol{s}_r$, as given by (9.30), and

$$r = \arg\min_i \; J(\hat{x}^i_{\mathrm{ML}}[N], \boldsymbol{s}_i) \tag{9.32}$$

is the index to the MLE of the symbolic sequence. The MLE of the remaining samples of the chaotic sequence can be obtained again through backwards iteration, resulting in the following MLE of the whole sequence:

$$\hat{\boldsymbol{x}}_{\mathrm{ML}} = [f^{-N}_{\hat{\boldsymbol{s}}_{\mathrm{ML}}}(\hat{x}_{\mathrm{ML}}[N]),\; f^{-(N-1)}_{\hat{\boldsymbol{s}}_{\mathrm{ML}}}(\hat{x}_{\mathrm{ML}}[N]),\; \ldots,\; f^{-1}_{\hat{\boldsymbol{s}}_{\mathrm{ML}}}(\hat{x}_{\mathrm{ML}}[N]),\; \hat{x}_{\mathrm{ML}}[N]]^\top. \tag{9.33}$$

9.2.3 Asymptotically Optimal Estimators

Due to the computational complexity of the MLE, many suboptimal algorithms have been proposed for estimating chaotic signals corrupted by AWGN. As already mentioned before, Myers *et al.* were the first ones to develop suboptimal algorithms that try to approach the MLE [23]. A method for modeling chaotic systems based on hidden Markov models (HMMs) that could be applied to obtain approximate ML and Bayesian estimators was also proposed [24, 33]. Kay also proposed two suboptimal estimators based on topological conjugacy (the halving method) and dynamic programming with good asymptotic performance [13]. During the following years many other authors proposed several simple and suboptimal methods based on the symbolic sequence that attained the CRLB asymptotically [4, 39, 26]. Iterative approaches based either on the E-M algorithm [28] or the Viterbi algorithm [1, 2, 19, 20, 18] have also been proposed. In this section we review the simple hard-censoring approach proposed in [26] and the more elaborate algorithm based on the Viterbi algorithm as described in [18].

9.2.3.1 Hard Censoring Estimator

The hard censoring maximum likelihood (HC-ML) estimator, proposed in [26], is probably the simplest approximate MLE. The idea behind the HC-ML is simply applying a threshold to the noisy observations to obtain an estimate of the symbolic sequence and using it to compute the MLE for that particular symbolic sequence. Mathematically, the HC-ML estimator is given by

$$\hat{x}_{\text{HC-ML}}[N] = \hat{x}^r_{\mathrm{ML}}[N], \tag{9.34}$$

where $\hat{x}^r_{\mathrm{ML}}[N]$ is the MLE associated to the r-th symbolic sequence, given by (9.30) for $\boldsymbol{s}_r = \hat{\boldsymbol{s}} = [\hat{s}[0],\ \hat{s}[1],\ \ldots,\ \hat{s}[N-1],\ \hat{s}[N]]^\top$, with the symbols of this symbolic sequence being obtained as

$$\hat{s}[n] = \begin{cases} 1, & y[n] < e_0; \\ \sigma(y[n]), & e_0 \leq y[n] \leq e_M; \\ M, & y[n] > e_M; \end{cases} \tag{9.35}$$

where $\sigma(y[n])$ is the symbolic dynamics associated to the natural partition of the map, given by (9.10), applied to the noisy observations. The remaining samples of the chaotic sequence are obtained through backward iteration using (9.31) for the r-th MLE.

This HC-ML estimator, denoted as HC-ML(0) in [14], can be improved by locating the k symbols ($1 \leq k \leq N$) most likely to be erroneous (i.e., those associated to observations closer to the borders separating the regions of the natural partition), changing them and checking whether the modified symbolic sequence provides better results or not [14]. Obviously, for $k = 0$ we get the basic HC-ML estimator as described in [26], whereas for $k = N$ we obtain the exact MLE.

9.2.3.2 Estimator based on the Viterbi Algorithm

A more sophisticated approximate MLE than the HC-ML can be obtained using the Viterbi algorithm (VA) as a computationally efficient estimator of $\boldsymbol{s}$. Note that, unlike the well-known cases of decoding of convolutional codes and detection in channels with intersymbol interference (ISI), where the VA provides us with the exact MLE [32], here the VA is a suboptimal estimator that will not achieve the MLE in general. In fact, the VA is able to obtain the exact ML estimator, but that would require a trellis with $\Gamma_s(N)$ states, thus not providing any computational advantage w.r.t. the exact MLE implementation based on an exhaustive search. Hence, following the approach of [18, 14], we propose to use a trellis with a reduced number of states, $R = M^r$, as an approximate MLE. In order to describe the algorithm, we focus on the simplest case: $M = 2$ and $r = 1$. Using $r = 1$ is equivalent to assuming that the next output depends only on the current symbol of the itinerary and the estimated value of $x[n]$. For $M = 2$ and $r = 1$ the basic butterfly of the trellis, shown in Figure 9.4, only has two states.

The transition cost for each branch is given, as usual, by the difference between the observations and the expected signals,

$$c_{ij}[n] = |y[N-n] - \hat{x}_{ij}[n]|\,, \tag{9.36}$$

with $\hat{x}_{ij}[n] = f^{-1}_{s_j[n]}(\hat{x}_i[N-(n-1)])$ for $i, j \in \{1, 2\}$, and $1 \leq n \leq N$. The

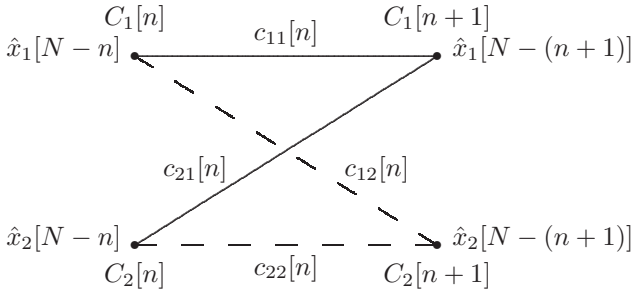


FIGURE 9.4
Basic butterfly for the VA using only two states per iteration of the chaotic sequence and a chaotic map with $M = 2$.

cost of each node is the minimum cost accumulated in all the branches arriving to it,

$$C_j[n] = C_r[n-1] + c_{rj}[n], \tag{9.37}$$

with

$$r = \underset{i\in\{1,2\}}{\arg\min} \left\{ C_i[n-1] + c_{ij}[n] \right\}, \tag{9.38}$$

and the current estimate of $x[n]$ for the j-th node is

$$\hat{x}_j[n] = \hat{x}_{rj}[n] = f^{-1}_{s_j[n]}(\hat{x}_r[n-1]). \tag{9.39}$$

Finally, the initial cost for each state is $C_j[0] = |y[0] - \hat{x}_j[0]|$, where $\hat{x}_j[0]$ is the estimate of $x[N]$. When $x[N]$ is known in the receiver, then $\hat{x}_j[0] = x[N]$. Otherwise, the initial two samples required to start the recursion can can be obtained applying a threshold to $y[0]$:

$$\hat{x}_j[0] = \max\left(e_{j-1}, \min(y[0], e_j)\right). \tag{9.40}$$

Hence, we guarantee that $\hat{x}_1[0] \in E_1 = [e_0, e_1]$ and $\hat{x}_2[0] \in E_2 = [e_1, e_2]$.

Extending this construction to higher values of r is straightforward, as shown in [14], and a slight improvement in the performance of the estimator is observed for low signal to noise ratios (SNRs), as shown in the following section. However, although for $r = 1$ there is a drastic reduction in the number of states of the trellis, the performance of the VA is very close to that of the actual ML estimator [18]. The reason is simple: backward iteration of two different initial conditions using the same itinerary leads to very similar trajectories. As a result, in every iteration we only need to store one estimate for each possible symbol,

since in general all the other paths do not provide very different chaotic sequences in the long term. In [17] it was proved, using map 1, that the distance between the orbits of two different initial states with the same itinerary decreases by a factor $2/(1-c)$ per iteration. Following the same reasoning, it is easy to show that, using map 2, this factor becomes 2 for any value of c, i.e. different initial states converge even faster than for map 1.

9.2.3.3 Comparison of Optimal and Suboptimal Estimators

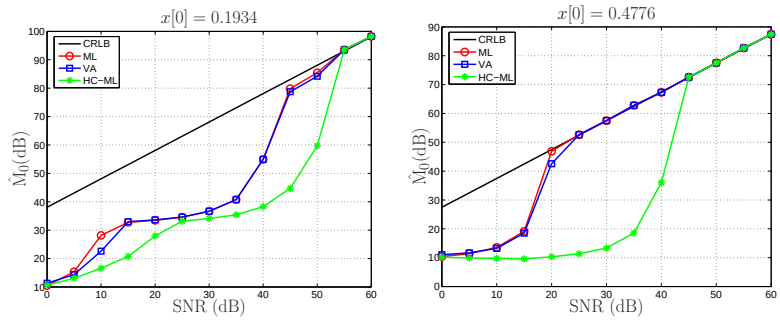
In this section we present some performance results of the exact ML estimator and the two suboptimal estimators described (HC-ML and VA) for two chaotic maps and short sequences. The first map considered is the SK-TM, whereas the second one is another popular chaotic map, the binary shift map (BSM), which is given by

$$f(x) = \begin{cases} 2x, & 0 \leq x < 0.5; \\ 2x - 1, & 0.5 \leq x \leq 1. \end{cases} \tag{9.41}$$

The reference sample used in all cases is $x[N]$, but results are presented for the estimation of $x[0]$, since this is the hardest sample of the sequence to be estimated.[11] The performance measure used is the mean square error (MSE) of the estimate of $x[0]$, $\hat{\mathrm{M}}_0(\mathrm{dB}) = -10\log_{10}\mathrm{MSE}(\hat{x}[0])$, which is in fact equivalent to the variance, since all the estimators considered are unbiased.

First of all, Figure 9.5 shows the value of $\hat{\mathrm{M}}_0(\mathrm{dB})$ obtained for both maps and different values of SNR using the HC-ML, the VA and the exact MLE for $N = 4$. It can be appreciated that all the estimators attain the CRLB for an SNR above a threshold that depends on the map (e.g. a much larger SNR is required to attain the CRLB for the SK-TM than for the BSM) and on the estimator. On the one hand we observe the excellent performance of the VA, which attains the CRLB at the same SNR than the exact MLE and provides very similar MSE values with only a fraction of its computational cost. On the other hand, the HC-ML provides a much worse performance, even requiring a much larger SNR than the VA and exact MLE to attain the CRLB for the BSM. Then, in Table 9.2.3.3 we compare the average performance of the exact MLE, the VA (with $r = 1,\ 2,$ and 3) and the HC-ML for several values of SNR. Each input in the table has been obtained averaging the results of 1000 simulations with a randomly chosen initial condition. Note how the VA provides very good results, achieving virtually the same performance as the MLE for $r = 3$.

[11] In fact, the MLE of $x[N]$ always attains the CRLB, whereas the MLE of $x[0]$ needs a minimum SNR to attain it [14].



(a) SK-TM with $c = 0.2$ (b) BSM

FIGURE 9.5
Comparison of the MSE obtained with the HC-ML, the VA and the exact MLE for $N = 4$.

9.3 Inverse Symbolic Chaotic Encoding

In this section we show how the asymptotically optimal estimators developed in the previous section can be applied to a chaotic digital communications system. Although all of these estimators can be applied to the well-known chaos shift keying (CSK) schemes, in this section we focus on their application to the inverse symbolic chaotic encoding proposed in [17].

9.3.1 Symbolic Encoder and Decoder

Let us consider the transmission of a vector of N information bits, $\boldsymbol{b} = [b[0], \ldots, b[N-1]]^T$. The basic idea of the inverse symbolic chaotic encoding scheme proposed in [17] is iterating backwards from a known final condition, $x[N]$, using those information bits to construct the symbolic sequence. The structure of the chaotic encoder is shown in Fig. 9.6. First of all, a one-to-one correspondence between the sequence of N information bits, and the symbols of the itinerary, $\tilde{s}[n] = s[N-n] = g(b[n])$ is established:

$$\tilde{s}[n] = s[N-n] = 1 + 2b[n], \tag{9.42}$$

for $1 \leq n \leq N$. Then, $\tilde{s}[n]$ is used to generate the chaotic sequence according to (9.15), starting from a given $\tilde{x}[0] = x[N]$ previously fixed

SNR (dB)	$\hat{\mathrm{M}}_0$(dB)					
	HC-ML(0)	VA ($r=1$)	VA ($r=2$)	VA ($r=3$)	ML	CRLB
0	10.5	10.9	11.3	11.4	11.6	20.6
5	14.2	14.2	15.1	15.4	15.3	25.6
10	19.8	19.5	20.6	21.0	21.0	30.6
15	24.7	30.0	30.9	31.2	30.6	35.6
20	31.9	40.6	40.6	40.6	40.6	40.6
25	45.5	45.5	45.5	45.5	45.5	45.6
60	80.4	80.4	80.4	80.4	80.4	80.6

TABLE 9.1
Comparison of MSE obtained by the HC-ML(0), the VA (with $r = 1,\ 2,$ and 3) and the exact MLE for the BSM with $N = 4$.

or randomly chosen, i.e.,

$$\tilde{x}[n] = x[N-n] = f^{-1}_{\tilde{s}[n]}(\tilde{x}[n-1]) = f^{-1}_{s[N-n]}(x[N-n]) == f^{-1}_{1+2b[n]}(x[N-n]) \tag{9.43}$$

for $1 \leq n \leq N-1$. Finally, these samples can be directly transmitted through the channel in baseband after a digital to analog conversion (DAC), as shown in Figure 9.7, or up converted to the desired frequency band. In the receiver, a standard equalizer is required to compensate the effect of the channel prior to the chaotic demodulation using the Viterbi algorithm as described in Section 9.2.3.2.

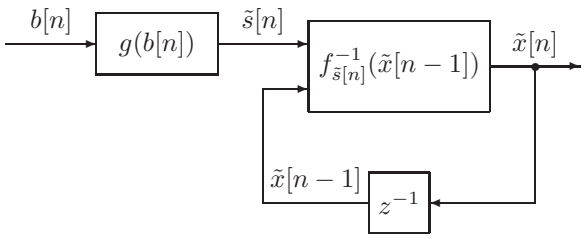


FIGURE 9.6
Block diagram of the inverse symbolic chaotic encoder for a generic map.

The choice of the chaotic map is crucial to achieve a good performance with this scheme. In [17] a PWL map with $M = 3$ regions (*map*

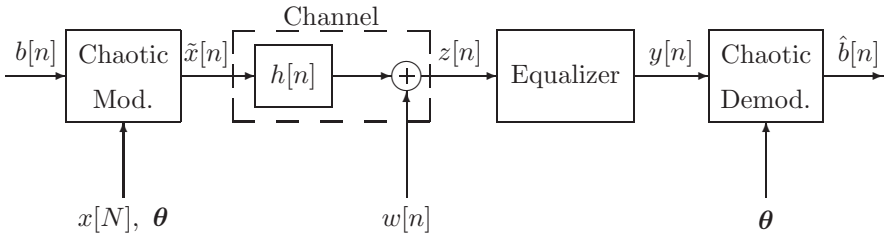


FIGURE 9.7
Baseband chaotic communications system with parameters $x[N]$ and $\boldsymbol{\theta}$: modulator, channel, and receiver.

1) was proposed:

$$f(x) = \begin{cases} \frac{2x+(1+c)}{1-c}, & -1 \le x \le -c; \\ \frac{x}{c}, & -c < x < c; \\ \frac{2x-(1+c)}{1-c}, & c \le x \le 1. \end{cases} \tag{9.44}$$

The two outer intervals, $E_1 = [-1, -c]$ and $E_3 = [c, 1]$ (with E_1 associated to $b[n] = 0$ and E_2 to $b[n] = 1$), are used for coding, whereas the inner one, $E_2 = (-c, c)$, acts as a guard interval, ensuring a minimum distance between the signals associated to $b[n] = 0$ and $b[n] = 1$. This map shows a good performance in terms of bit error rate (BER), but it has an important drawback: $\tilde{x}[n]$ cannot belong to any region of the state space that maps into E_2 after $1 \le k \le n+1$ iterations. This creates exclusion regions for $\tilde{x}[n]$ and a clustering of the samples of the chaotic sequence (see [14] for a detailed description of this issue), thus reducing the security of the system. In order to solve this problem, an alternative map (*map 2*) with $M = 5$ intervals has been introduced in [14], which avoids the innermost interval, $E_3 = (-c, c)$, which is used again as a guard region:

$$f(x) = \begin{cases} 2x + 1, & -1 \le x \le -(1+c)/2; \\ 2x + (1+2c), & -(1+c)/2 \le x \le -c; \\ x/c, & -c < x < c; \\ 2x - (1+2c), & c \le x \le (1+c)/2; \\ 2x - 1, & (1+c)/2 \le x \le 1. \end{cases} \tag{9.45}$$

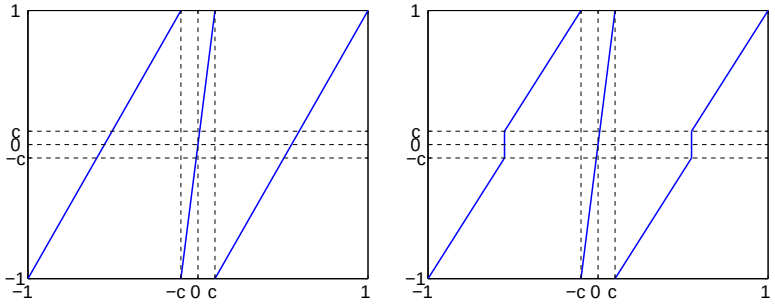
The shape of these two maps is plotted in Figure 9.8. Note that in both cases we use the invertible partition, which is composed of $M_I = 3$ regions, when iterating backwards. For map 1 the invertible partition is

the same as the natural partition, $\mathcal{P}_I = \mathcal{P}_N = \{E_1,\ E_2,\ E_3\}$, whereas for map 2 the invertible partition is $\mathcal{P}_I = \{E_1 \cup E_2,\ E_3,\ E_4 \cup E_5\}$. In this second case, $E_1 = [-1, -(1+c)/2]$ and $E_2 = [-(1+c)/2, -c]$ are associated to $b[n] = 0$, whereas $E_4 = [c, (1+c)/2]$ and $E_5 = [(1+c)/2, 1]$ to $b[n] = 1$. The inverse function, used for coding, is straightforward to obtain for the first map:

$$f(x) = \begin{cases} \frac{(1-c)y-(1+c)}{2}, & s = 1; \\ cy, & s = 2; \\ \frac{(1-c)y+(1+c)}{2}, & s = 3. \end{cases} \tag{9.46}$$

For the second map, the inverse function can also be easily obtained, but now we have to take into account the region to which y belongs in addition to the corresponding symbol:

$$x = f_s^{-1}(y) = \begin{cases} (y-1)/2, & s = 1,\ y \in E_1 \cup E_2; \\ (y-(1+2c))/2, & s = 1,\ y \in E_4 \cup E_5; \\ (y+(1+2c))/2, & s = 3,\ y \in E_1 \cup E_2; \\ (y+1)/2, & s = 3,\ y \in E_4 \cup E_5. \end{cases} \tag{9.47}$$



(a) Map 1 (b) Map 2

FIGURE 9.8
PWL maps used for the inverse symbolic chaotic encoding system.

The good performance of the proposed chaotic communications system is shown in Figure 9.9 for map 2 (the performance of map 1, not shown, is very similar) and several values of the parameter c. The bit error rate (BER) obtained follows the curve of the binary phase shift keying (BPSK) modulation, with approximately 1.5 dB loss for a 10^{-5}

probability of error when $c = 0$ and hardly any loss in performance for values of $c > 0.5$.

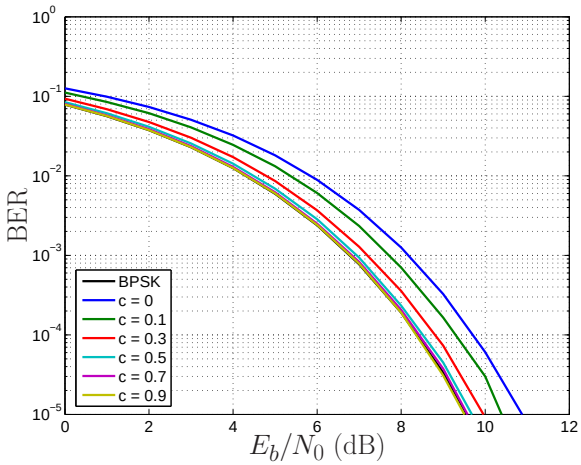


FIGURE 9.9
Bit error rate (BER) for the inverse symbolic chaotic encoding approach using map 2 with several values of c.

9.3.2 Combination with OFDM for Multipath Channels

The coding approach described in Section 9.3.1 provides very good results for the Gaussian channel. However, it suffers from a severe degradation in performance in the presence of multipath interference (unavoidable in wireless communications), just as most chaotic and conventional modulation schemes. In order to provide a certain degree of protection against the distortion introduced by the channel, in this section we propose to combine the chaotic coding with a modulation format robust against multipath fading: OFDM.

The structure of the modulator is shown in Figure 9.10. The idea is simple: substitute the conventional coding used in each subcarrier of the OFDM system (BPSK, QPSK or M-QAM usually) by the chaotic coding described in Section 9.3.1. These coded symbols are then serial-to-parallel converted (S/P) to form blocks of $N_c = N_b$ symbols which, altogether with N_p pilots (used to estimate the channel) and N_z zeros (used as guard intervals), serve to generate the OFDM symbol by means of an $N = N_c+N_p+N_z$ points inverse discrete Fourier transform (IDFT). A parallel-to-serial (P/S) conversion is performed next, a cyclic prefix of

M samples is added to avoid intersymbol and intercarrier interference (ISI and ICI), and the signal is finally transmitted through the channel.

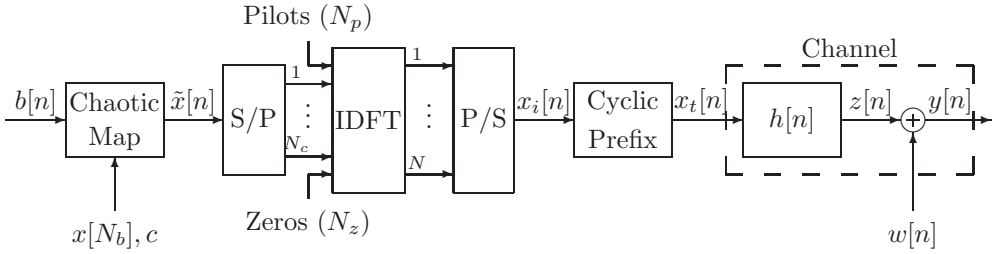


FIGURE 9.10
Block diagram for the proposed OFDM communications system with chaotic coding in the sub-carriers: transmitter and channel.

The receiver, shown in Figure 9.11, is simply the dual of the transmitter. First, the cyclic prefix is discarded, followed by an S/P conversion and an N point DFT, from which we discard the N_z guard zeros, and use the N_p pilots to update the estimate of the channel. Then, the N_c carriers that contain the useful information are equalised in frequency using the current estimate of the channel. A P/S convertion follows, and finally the N_c equalised samples are passed to the chaotic demodulator to estimate the transmitted information bits.

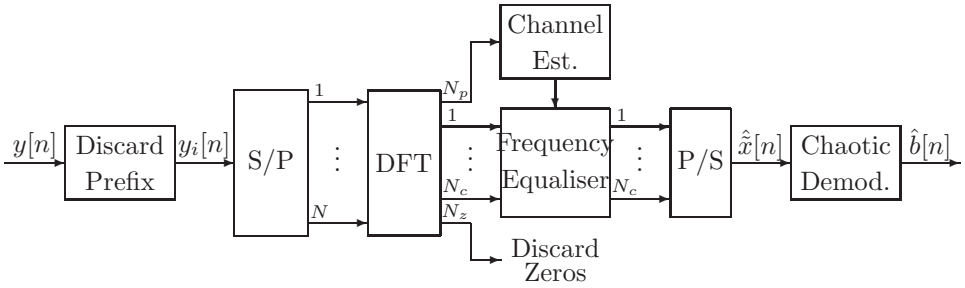
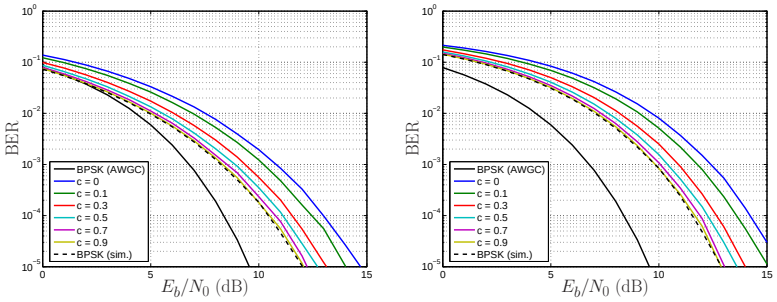


FIGURE 9.11
Block diagram for the proposed OFDM communications system with chaotic coding in the sub-carriers: receiver.

The good performance of the combined inverse symbolic chaotic encoding with OFDM is shown in Figure 9.12 for two different channels with perfect equalization. In Figure 9.12(a) a simple two-ray minimum phase channel, $h_1[n] = \delta[n] - 0.5\delta[n-1]$, is used, whereas in

Figure 9.12(b), a more complex non-minimum phase channel, $h_2[n] = -0.3\delta[n-1] + 0.7\delta[n-2] + 0.4\delta[n-3] + 0.1\delta[n-4]$, is used. In both cases the performance of the proposed scheme is similar to the one for the AWGN channel.



(a) Channel 1 (b) Channel 2

FIGURE 9.12
Bit error rate (BER) for the inverse symbolic chaotic encoding approach plus OFDM for two different channels using map 2 with several values of c.

9.4 Appendix: Cramer-Rao Lower Bound

The Cramer-Rao lower bound (CRLB) is a lower bound on the variance that can be attained by an unbiased estimator [11]. In general, the CRLB is not a strong limit, implying that often it cannot be achieved by any unbiased estimator. However, if an estimator attaining the CRLB is found, then we can ensure that it is the optimum estimator, i.e., the minimum variance unbiased estimator. Otherwise, it can be used as a reference to evaluate the quality of the estimators considered.

The CRLB for the estimation of $x[N]$, is given by

$$\mathrm{Var}(\hat{x}[N]) \geq \left[-\mathrm{E}\left(\frac{\partial^2 \ln p(\boldsymbol{y}; x[N], \boldsymbol{s})}{\partial x[N]^2} \right) \right]^{-1}, \tag{9.48}$$

where $p(\boldsymbol{y}; x[N], \boldsymbol{s})$ is the likelihood, as given by (9.19), but expressed as

a function of $x[N]$ and $\boldsymbol{s}$. After some algebra, it can be shown (see [14]) that the CRLB can be expressed as

$$\begin{aligned}\mathrm{Var}(\hat{x}[N]) &\geq \frac{\sigma^2}{1+\sum_{n=1}^{N}\left(\dot{f}_{\boldsymbol{s}}^{-n}(x[N])\right)^2} \\ &= \frac{\sigma^2}{1+\sum_{n=1}^{N}\prod_{k=0}^{n-1}\left(\dot{f}_{s[N-k-1]}^{-1}(x[N-k])\right)^2},\end{aligned} \tag{9.49}$$

where $\dot{f}^{-n}$ is used to denote the derivative of f^{-n} $(1 \leq n \leq N)$ and the chain rule has been used to obtain the last expression. Finally, using the analytical expressions for the backward iteration of a generic PWL map given by (9.25) and (9.26), the CRLB for this class of maps becomes

$$\begin{aligned}\mathrm{Var}(\hat{x}[N]) &\geq \frac{\sigma^2}{1+\sum_{n=1}^{N}\left(B_{\boldsymbol{s}}^{1,n}\right)^2} \\ &= \frac{\sigma^2}{1+\sum_{n=1}^{N}\prod_{k=0}^{n-1}a_{s[N-k-1]}^{-2}}.\end{aligned} \tag{9.50}$$

The CRLB can be obtained similarly for any other sample of the sequence, as shown in [14]. In particular, the CRLB for the initial condition, $x[0]$, is

$$\begin{aligned}\mathrm{Var}(\hat{x}[0]) &\geq \frac{\sigma^2}{1+\sum_{n=1}^{N}\left(A_{\boldsymbol{s}}^{0,n}\right)^2} \\ &= \frac{\sigma^2}{1+\sum_{n=1}^{N}\prod_{k=0}^{n-1}a_{s[k]}^{2}},\end{aligned} \tag{9.51}$$

where $A_{\boldsymbol{s}}^{0,n}$ are the coefficients for the forward iteration of the PWL map, given by [26, 14]

$$A_{\boldsymbol{s}}^{0,n} = \prod_{k=0}^{n-1} a_{s[k]}. \tag{9.52}$$

Finally, note that simpler closed-form expression for the CRLB cannot be developed in general for most PWL maps. However, for some particular cases, such as the BSM used in Section 9.2.3.3, simple analytical expressions can be provided, since the slope is identical in both regions of the natural particion (i.e., $a_1 = a_2 = 2$). In this case, the CRLB of the initial condition is [14]

$$\mathrm{Var}(\hat{x}[0]) \geq \frac{3\sigma^2}{4^{N+1}-1}, \tag{9.53}$$

whereas the CRLB of the final condition is given by

$$\mathrm{Var}(\hat{x}[N]) \geq \frac{3 \cdot 4^N \sigma^2}{4^{N+1} - 1}. \tag{9.54}$$

These two limits show that $\mathrm{Var}(\hat{x}[0]) \to 0$ as $N \to \infty$, confirming that a consistent estimator for $x[0]$ can theoretically be found, whereas $\mathrm{Var}(\hat{x}[N]) \to 3\sigma^2/4$ as $N \to \infty$, showing the unconsistency in the estimation of $x[N]$. Unfortunately, several authors have proved that the MLE of $x[0]$ is inconsistent for low SNRs, thus not being able to attain the CRLB below a certain SNR threshold that depends on the chaotic map, its parameters and even the initial condition [12, 8, 34]. Finally, note that the CRLB does not depend on the sample of the chaotic sequence chosen as a reference. Hence, all the equations in this appendix are valid regardless of whether $x[0]$, $x[N]$ or $x[n]$ $(1 \leq n \leq N-1)$ is used as a reference point for the estimation.

Bibliography

[1] M. Ciftci and D. B. Williams. Optimal estimation and sequential channel equalization algorithms for chaotic communications systems. *EURASIP Journal on Applied Signal Processing*, 4:249–256, 2001.

[2] M. Ciftci and D. B. Williams. Optimal estimation for chaotic sequences using the viterbi algorithm. In *IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, 2001.

[3] P. Collet and J. P. Eckmann. *Iterated Maps on the Interval as Dynamical Systems*. Birkhauser, 1980.

[4] L. Cong, W. Xiaofu, and S. Songgeng. A general efficient method for chaotic signal estimation. *IEEE Trans. on Signal Processing*, 47(5):1424–1428, 1999.

[5] W. de Melo and S. van Strien. *One-Dimensional Dynamics*. Springer-Verlag, 1993.

[6] Robert L. Devaney. *An Introduction to Chaotic Dynamical Systems*. Addison Wesley, Redwood City, CA, 2nd edition edition, 1989.

[7] Martin Hasler and Yuri L. Maistrenko. An introduction to the synchronization of chaotic systems: Coupled skew tent maps. *IEEE Trans. on Circuits and Systems I*, 44(10):856–866, 1997.

[8] I. Hen and N. Merhav. On the threshold effect in the estimation of chaotic sequences. *IEEE Trans. on Information Theory*, 50(11):2894–2904, 2004.

[9] S. H. Isabelle and G.W.Wornell. *The Digital Signal Processing Handbook*, chapter Nonlinear maps. CRC Press & IEEE Press, 1998.

[10] S.H. Isabelle and G.W. Wornell. Statistical analysis and spectral estimation techniques for one-dimensional chaotic signals. *IEEE Transactions on Signal Processing*, 45(6):1495–1506, Jun 1997.

[11] S. M. Kay. *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice Hall, 1993.

[12] S. M. Kay. Asymptotic maximum likelihood estimator performance for chaotic signals in noise. *IEEE Trans. on Signal Processing*, 43(4):1009–1012, Apr. 1995.

[13] S. M. Kay and V. Nagesha. Methods for chaotic signal estimation. *IEEE Trans. on Signal Processing*, 43(8):2013–2016, Aug. 1995.

[14] D. Luengo. *Estimación Óptima de Secuencias Caóticas con Aplicación en Comunicaciones*. PhD thesis, Red (TDR), 2006. Available at http://www.tesisenred.net/handle/10803/10663 (in Spanish).

[15] D. Luengo, C. Pantaléon, and I. Santamaría. Competitive chaotic AR(1) model estimation. In *Proc. XI IEEE Neural Networks for Signal Processing (NNSP) Workshop*, pages 83–92, 2001.

[16] D. Luengo, C. Pantaléon, and I. Santamaría. Bayesian estimation of discrete chaotic signals by MCMC. In *Proc. European Signal Processing Conference (EUSIPCO)*, pages 333–336, 2002.

[17] D. Luengo and I. Santamaría. Asymptotically optimal maximum-likelihood estimation of a class of chaotic signals using the viterbi algorithm. In *Proc. European Signal Processing Conference (EUSIPCO)*, Antalya (Turkey), 2005.

[18] D. Luengo and I. Santamaría. Secure communications using OFDM with chaotic modulation in the subcarriers. In *Proc. IEEE 61st Semiannual Vehicular Tech. Conf. (VTC2005-Spring)*, Stockholm (Sweden), 2005.

[19] G. M. Maggio and L. Reggiani. Applications of symbolic dynamics to UWB impulse radio. In *Proc. IEEE Int. Symp. on Circuits and Systems (ISCAS), volume III*, pages 153–156, Sydney (Australia), 2001.

[20] G. M. Maggio, N. Rulkov, and L. Reggiani. Pseudo-chaotic time hopping for UWB impulse radio. *IEEE Trans. on Circuits and Systems I: Fundamental Theory and Applications*, 48(12), Dec. 2001.

[21] N. Masuda, G. Jakimoski, K. Aihara, and L. Kocarev. Chaotic block ciphers: From theory to practical algorithms. *IEEE Trans. on Circuits and Systems I*, 53(6):1341–1352, June 2006.

[22] A. Mooney, J. G. Keating, and I. Pitas. A comparative study of chaotic and white noise signals in digital watermarking. *Chaos, Solitons and Fractals*, 35:913–921, 2008.

[23] C. Myers, S. Kay, and M. Richard. Signal separation for nonlinear dynamical systems. In *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP)*, pages 129–132, San Francisco, CA (USA), 1992.

[24] C. Myers, A. C. Singer, B. Shin, and E. Church. Modeling chaotic systems with hidden markov models. In *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP)*, pages 23–26, San Francisco, CA (USA), 1992.

[25] C. Pantaleón, D. Luengo, and I. Santamaría. Bayesian estimation of a class of chaotic signals. In *In Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, pages 193–196, 2000.

[26] C. Pantaleón, D. Luengo, and I. Santamaría. Optimal estimation of chaotic signals generated by piecewise-linear maps. *IEEE Signal Processing Letters*, 7(8):235–237, 2000.

[27] C. Pantaléon, D. Luengo, and I. Santamaría. Chaotic AR(1) model estimation. In *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, pages 3477–3480, 2001.

[28] C. Pantaléon, D. Luengo, and I. Santamaría. Estimation of a certain class of chaotic signals: An EM-based approach. In *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP)*, pages 1129–1132, 2002.

[29] C. Pantaléon, L. Vielva, D. Luengo, and I. Santamaría. Bayesian estimation of chaotic signals generated by piecewise-linear maps. *Signal Processing*, 83:659–664, Mar. 2003.

[30] H. C. Papadopoulos and G. W. Wornell. Optimal detection of a class of chaotic signals. In *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP)*, pages 117–120, 1993.

[31] H. C. Papadopoulos and G. W. Wornell. Maximum-likelihood estimation of a class of chaotic signals. 41(1):312–317, January 1995.

[32] J. G. Proakis. *Digital Communications*. McGraw-Hill, Singapore, 1995.

[33] M. D. Richard. Properties and discrimination of chaotic maps. In *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, pages 141–144, 1993.

[34] I. Rosenhouse and A. J. Weiss. Consistent estimation of symmetric tent chaotic sequences with coded itineraries. *IEEE Trans. on Signal Processing*, 56(11):5580–5588, Nov 20098.

[35] T. Schimming and M. Hasler. Chaos communication in the presence of channel noise. *Journal of Signal Processing*, 4(1):21–28, 2000.

[36] T. Schimming and J. Schweizer. Chaos communication from a maximum likelihood perspective. In *Proc. Int. Workshop on Nonlinear Dynamics of Electronic Systems (NDES),*, pages 179–182, Rome (Denmark), 1999.

[37] S. Smale. Differentiable Dynamical Systems I. Diffeomorphisms. *Bull. Am. Math. Soc.*, 73:747–817, 1967.

[38] H. L. Van Trees. *Detection, Estimation and Modulation Theory*. John Wiley and Sons, 1968.

[39] S. Wang, P. C. Yip, and H. Leung. Estimating initial conditions of noisy chaotic signals generated by piece-wise linear markov maps using itineraries. *IEEE Trans. on Signal Processing*, 47(12):3289–3302, 1999.