ON OPTIMAL DETECTION OF NONCOHERENT CHAOS-SHIFT-KEYING SIGNALS IN A NOISY ENVIRONMENT

FRANCIS C. M. LAU* and CHI K. TSE†

Department of Electronic and Information Engineering,
The Hong Kong Polytechnic University, Hong Kong, China
*encmlau@polyu.edu.hk
†encktse@polyu.edu.hk

Recently, an optimal noncoherent detection technique for chaos-shift-keying digital communication system has been proposed. It has been stated that computational intensity showed increases that were exponential with the spreading factor. In this Letter, we show that the implementation of the optimal detector can be made independent of the chaotic maps used, and that the computational intensity will increase almost linearly with the spreading factor. In particular, we use a tent map as an example to illustrate the decoding algorithm. The bit error performance of the system is then evaluated by computer simulations for a range of spreading factors. Further, we extend the optimal decoding algorithm for maps of higher dimension. The bit error performance for the case of simple 2-D maps are compared with that obtained using the tent map. Finally, the effect of increasing spreading factor on the bit error performance is studied for the case of 2-D maps.

Keywords: Chaos-shift-keying; noncoherent; optimal detection.

1. Introduction

Chaos-based communications has attracted a lot of attention in the past decade and a number of chaos-based digital communication systems have been proposed [Dedieu *et al.*, 1993; Hasler & Schimming, 2000; Kennedy & Kolumbán, 2000; Kolumbán *et al.*, 1998a, 1998b; Kolumbán & Kennedy, 2000; Sushchik *et al.*, 2000; Wu & Chua, 1993]. Among the communication systems proposed, chaos-shift-keying (CSK) and differential CSK (DCSK) are the most widely studied. Recently, an optimal non-coherent detector for CSK digital communications has been proposed [Hasler & Schimming, 2000]. It has been pointed out, however, that the optimal detector suffers from exponentially increasing computational complexity with the spreading factor. As a consequence, the applicability of the optimal classifier is limited to spreading factors below 20. Also, whether the optimal classifier would also suffer from a decrease in performance for large spreading factors is uncertain.

In this Letter, we show that in the implementation of the optimal detector, the detecting algorithm can be made independent of the chaotic maps being used. We also show that the computational intensity will only increase slowly with the spreading factor. In particular, we use a tent map as an example to illustrate the decoding algorithm. The bit error performance of the system is then evaluated by computer simulations for a range of spreading factors. Further, we extend the optimal decoding algorithm for maps of higher dimension. The bit error performance of the system using simple 2-D maps are compared with that obtained using the tent map.

Finally, for the case of using 2-D maps, the effect of increasing spreading factor will also be studied.

2. System Overview

We consider a discrete-time binary CSK communication system, as shown in Fig. 1. In the transmitter, a pair of chaotic sequences, denoted by $\{\hat{x}_k\}$ and $\{\check{x}_k\}$, are generated by two chaotic maps. If the symbol "+1" is sent, $\{\hat{x}_k\}$ is transmitted during a bit period, and if "−1" is sent, $\{\check{x}_k\}$ is transmitted. Further, we assume that "+1" and "−1" occur with equal probabilities.

Suppose a spreading factor of N is used, i.e. N chaotic samples are transmitted for each binary symbol. Without loss of generality, we consider the signal block transmitted during the first symbol period and we denote it by the vector $\mathbf{s} = (s_1 \ s_2 \ \cdots \ s_N)$. Assuming that the channel is additive white Gaussian with power-spectral-density $N_0/2$ and denoting the noise vector by $\mathbf{n} = (n_1 \ n_2 \ \cdots \ n_N)$, the received signal block is then given by $\mathbf{r} = (r_1 \ r_2 \ \cdots \ r_N) = \mathbf{s} + \mathbf{n}$. Denote the transmitted symbol and decoded symbol by q and $\tilde{q} \in \{-1, +1\}$, respectively. Based on the received signal block, detection techniques are designed so as to optimize the system performance, e.g. minimize the probability of an incorrect decision with a simple receiver structure.

3. Review of Hasler–Schimming's Optimal Noncoherent Detector

It has been shown by Hasler and Schimming [2000] that the optimal detector selects the symbol q such that the *a posteriori* probability given $\mathbf{r}$ is maximized, i.e.

$$\tilde{q} = \arg\max_{q} \text{Prob}(q \text{ is sent} \,|\, \mathbf{r}). \tag{1}$$

As the *a posteriori* probability is not convenient to calculate, the Bayes' rule is applied to $\text{Prob}(q \text{ is sent} \,|\, \mathbf{r})$ to obtain

$$
\begin{aligned}
&\text{Prob}(q \text{ is sent} \,|\, \mathbf{r}) \\
&= \frac{p(\mathbf{r} \,|\, q \text{ is sent}) \times \text{Prob}(q \text{ is sent})}{p(\mathbf{r})},
\end{aligned} \tag{2}
$$

where $p(.)$ denotes the probability density function. Hence, (1) can be rewritten as

$$\tilde{q} = \arg\max_{q} p(\mathbf{r} \,|\, q \text{ is sent}) \tag{3}$$

because $\text{Prob}(\text{"+1" is sent}) = \text{Prob}(\text{"−1" is sent}) = 1/2$ and $p(\mathbf{r})$ is independent of q.

Example: *Skew Tent Map*

Suppose the chaotic signals are generated from the following skew tent map

$$
f(x) = \begin{cases}
\dfrac{2x + 1 - a}{1 + a} & \text{for } -1 \leq x \leq a \\[2mm]
\dfrac{-2x + 1 + a}{1 - a} & \text{for } a \leq x \leq 1
\end{cases} \tag{4}
$$

with $-1 < a < 1$ and $-1 < x < 1$. The chaotic signals for the symbols "+1" and "−1", i.e. $\{\hat{x}_k\}$ and $\{\check{x}_k\}$, are then generated by the iterations of f and $-f$, respectively. It can be shown that the natural invariant probability density of f and $-f$, denoted by $\rho(\cdot)$, equals $1/2$ in the interval $[-1, +1]$. Assume that the initial value s_1 of each chaotic signal block $\mathbf{s}$ is chosen randomly according to the natural invariant probability density of f and $-f$. For a spreading factor of 2, i.e. $N = 2$, it has been shown that the conditional probability densities are given by
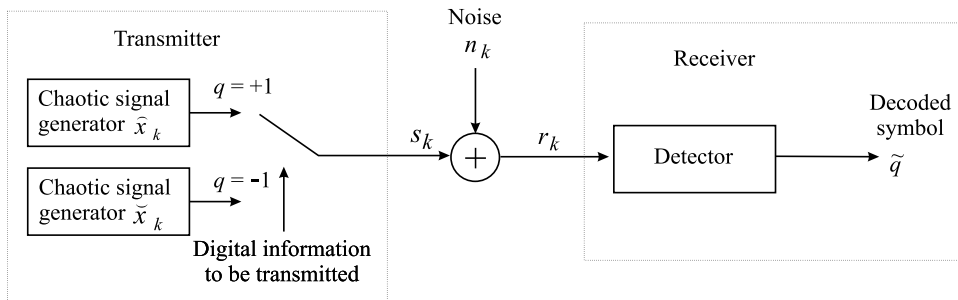

Fig. 1.   Block diagram of a discrete-time binary chaos-shift-keying communication system.

$$p(\mathbf{r}|\text{``}+1\text{''} \text{ is sent}) = \frac{1}{2\pi\sigma^2} \int_{-1}^{+1} \rho(s_1) \exp\left(-\frac{(r_1 - s_1)^2 + (r_2 - f(s_1))^2}{2\sigma^2}\right) ds_1 \tag{5}$$

$$= \frac{1}{4\pi\sigma^2} \int_{-1}^{a} \exp\left(-\frac{(r_1 - x)^2 + \left(r_2 - \dfrac{2x + 1 - a}{1 + a}\right)^2}{2\sigma^2}\right) dx$$

$$+ \frac{1}{4\pi\sigma^2} \int_{a}^{+1} \exp\left(-\frac{(r_1 - x)^2 + \left(r_2 - \dfrac{-2x + 1 + a}{1 - a}\right)^2}{2\sigma^2}\right) dx \tag{6}$$

$$p(\mathbf{r}|\text{``}-1\text{''} \text{ is sent}) = \frac{1}{2\pi\sigma^2} \int_{-1}^{+1} \rho(s_1) \exp\left(-\frac{(r_1 - s_1)^2 + (r_2 + f(s_1))^2}{2\sigma^2}\right) ds_1 \tag{7}$$

$$= \frac{1}{4\pi\sigma^2} \int_{-1}^{a} \exp\left(-\frac{(r_1 - x)^2 + \left(r_2 + \dfrac{2x + 1 - a}{1 + a}\right)^2}{2\sigma^2}\right) dx$$

$$+ \frac{1}{4\pi\sigma^2} \int_{a}^{+1} \exp\left(-\frac{(r_1 - x)^2 + \left(r_2 + \dfrac{-2x + 1 + a}{1 - a}\right)^2}{2\sigma^2}\right) dx \tag{8}$$

where σ^2 denotes the variance of noise and equals $N_0/2$. The integrals in (6) and (8) can all be expressed in terms of the error functions. Then, for a given received signal vector $\mathbf{r}$, (6) and (8) are evaluated by computing the corresponding error functions. The decoded symbol will be "+1" or "−1" depending upon $p(\mathbf{r}|\text{``}+1\text{''}$ is sent) being larger or smaller than $p(\mathbf{r}|\text{``}-1\text{''}$ is sent).

For the aforementioned example in which $N = 2$, two error functions need to be computed for each of the two conditional probability density functions. Thus, a total of $2 \times 2 = 4$ error functions have to be evaluated to decode one symbol. It therefore appears that in general, for a spreading of N, each probability density function has to be split into 2^{N-1} integrals, implying that a total of 2^N integrals (error functions) have to be evaluated in

order to decode one symbol. Thus, it has been concluded [Hasler & Schimming, 2000] that "the applicability of the optimal classifier and its possible variants that take into account all linear regions, is limited to values of N that do not exceed, say, 15–20."

4. Implementation of Optimal Detector

When implementing the optimal detecting algorithm, we find that it is unnecessary to split the integrals (5) and (7) into a number of integrals if we apply numerical integration directly (instead of trying to represent them in terms of complementary error functions). That means, for a spreading factor of N, the two conditional probabilities are given by

$$p(\mathbf{r}|\text{``}+1\text{''} \text{ is sent}) = \frac{1}{(2\pi\sigma^2)^{N/2}} \int_{-1}^{+1} \rho(s_1) \exp\left(-\frac{\displaystyle\sum_{i=1}^{N}(r_i - f^{(i-1)}(s_1))^2}{2\sigma^2}\right) ds_1 \tag{9}$$

$$= \lim_{K\to\infty} \frac{1}{2(2\pi\sigma^2)^{N/2}} \sum_{k=1}^{K} \exp\left(-\frac{\sum_{i=1}^{N}(r_i - f^{(i-1)}(x_k))^2}{2\sigma^2}\right) \delta x_k \qquad (10)$$

$$p(\mathbf{r} \mid \text{``} -1\text{''} \text{ is sent}) = \frac{1}{(2\pi\sigma^2)^{N/2}} \int_{-1}^{+1} \rho(s_1) \exp\left(-\frac{\sum_{i=1}^{N}(r_i - g^{(i-1)}(s_1))^2}{2\sigma^2}\right) ds_1 \qquad (11)$$

$$= \lim_{K\to\infty} \frac{1}{2(2\pi\sigma^2)^{N/2}} \sum_{k=1}^{K} \exp\left(-\frac{\sum_{i=1}^{N}(r_i - g^{(i-1)}(x_k))^2}{2\sigma^2}\right) \delta x_k \qquad (12)$$

where $\delta x_k = 2/K$; $x_k = -1 + (k-1)\delta x_k$; $f^{(j)}(\alpha)$ denotes the iteration of the function f, j times with the initial condition α and $f^{(0)}(\alpha) = \alpha$. In this particular example, $g = -f$. In any case, (9) and (11) are computed directly by numerical integration using (10) and (12). Note that the accuracy of the numerical integration will depend on the choice of the value K. Larger K implies higher accuracy and also more resources in terms of computational time and memory. From (10) and (12), it can be observed that when the spreading factor increases, only the number of operations (subtraction, squaring and addition) within the exponential function will increase proportionally. Other operations remain unchanged. Thus, it is expected that the computational time will increase almost linearly with the spreading factor. The decoded symbol will be "+1" or "−1" depending upon $p(\mathbf{r} \mid \text{``} +1\text{''} \text{ is sent})$ being larger or smaller than $p(\mathbf{r} \mid \text{``} -1\text{''} \text{ is sent})$. In general, f and g are arbitrary maps which are used to represent the two symbols.

5. Application of the Optimal Detector to Higher-Dimensional Maps

Suppose f and g are L-dimensional maps. The corresponding chaotic sequences are generated by $\hat{x}_{n+1} = f(\hat{\mathbf{x}}_n)$ and $\check{x}_{n+1} = g(\check{\mathbf{x}}_n)$, respectively, where $\hat{\mathbf{x}}_n = (\hat{x}_n \ \hat{x}_{n-1} \ \cdots \ \hat{x}_{n-L+1})$ and $\check{\mathbf{x}}_n = (\check{x}_n \ \check{x}_{n-1} \ \cdots \ \check{x}_{n-L+1})$. Assume a spreading factor of $N \geq L$. For each transmitted symbol, we first select the initial condition $\mathbf{x}_L = (x_L \ x_{L-1} \ \cdots \ x_1)$ according to some pre-specified criteria. If a "+1" is to be sent, we assign $\hat{\mathbf{x}}_L = \mathbf{x}_L$ and apply f to generate the chaotic signals that follow. The transmitted symbol "+1" is then represented by the block $\mathbf{s} = (\hat{x}_1 \ \hat{x}_2 \ \cdots \ \hat{x}_N)$. Likewise, when a "−1" is to be sent, we assign $\check{\mathbf{x}}_L = \mathbf{x}_L$ and g is applied. The transmitted signal block is now given by $\mathbf{s} = (\check{x}_1 \ \check{x}_2 \ \cdots \ \check{x}_N)$. Corrupted by noise, the received signal block is $\mathbf{r} = (r_1 \ r_2 \ \cdots \ r_N) = \mathbf{s} + \mathbf{n}$. Applying the optimal detecting technique, the two conditional probabilities are given by

$$p(\mathbf{r} \mid \text{``} +1\text{''} \text{ is sent}) = \frac{1}{(2\pi\sigma^2)^{N/2}} \underbrace{\int\int\cdots\int}_{L \text{ integrals}} \rho(\mathbf{x}_L) \exp\left(-\frac{\sum_{i=1}^{N}(r_i - f^{(i-1)}(\mathbf{x}_L))^2}{2\sigma^2}\right) dx_1\, dx_2 \cdots dx_L \qquad (13)$$

$$p(\mathbf{r} \mid ``-1\text{'' is sent}) = \frac{1}{(2\pi\sigma^2)^{N/2}} \underbrace{\int\int\cdots\int}_{L \text{ integrals}} \rho(\mathbf{x}_L) \exp\left(-\frac{\sum_{i=1}^{N}(r_i - g^{(i-1)}(\mathbf{x}_L))^2}{2\sigma^2}\right) dx_1 dx_2 \cdots dx_L \quad (14)$$

where $\rho(\mathbf{x}_L)$ denotes the joint probability density function of $\mathbf{x}_L$.

Example: Two-Dimensional Map

Suppose the following 2-D maps (i.e. $L = 2$) are used to generate the chaotic sequences.

$$x_{n+1} = f(x_n, x_{n-1}) = 1 + 0.3x_n - 1.37x_{n-1}^2 \quad (15)$$

$$x_{n+1} = g(x_n, x_{n-1}) = -1 + 0.3x_n + 1.37x_{n-1}^2 \quad (16)$$

The corresponding attractors are shown in Fig. 2. Further, the initial values x_2 and x_1 are selected randomly and uniformly within $[-1, +1]^2$. As a consequence, $\rho(\mathbf{x}_L) = 1/4$. Hence, (13) and (14) can be rewritten as

$$p(\mathbf{r} \mid ``+1\text{'' is sent}) = \frac{1}{4(2\pi\sigma^2)^{N/2}} \int_{-1}^{+1}\int_{-1}^{+1} \exp\left(-\frac{\sum_{i=1}^{N}(r_i - f^{(i-1)}(x_2,\, x_1))^2}{2\sigma^2}\right) dx_1 \, dx_2 \quad (17)$$

$$= \lim_{L\to\infty}\lim_{K\to\infty} \frac{1}{4(2\pi\sigma^2)^{N/2}} \sum_{l=1}^{L}\sum_{k=1}^{K} \exp\left(-\frac{\sum_{i=1}^{N}(r_i - f^{(i-1)}(y_l,\, x_k))^2}{2\sigma^2}\right) \delta y_l \delta x_k \quad (18)$$

$$p(\mathbf{r} \mid ``-1\text{'' is sent}) = \frac{1}{4(2\pi\sigma^2)^{N/2}} \int_{-1}^{+1}\int_{-1}^{+1} \exp\left(-\frac{\sum_{i=1}^{N}(r_i - g^{(i-1)}(x_2,\, x_1))^2}{2\sigma^2}\right) dx_1 \, dx_2 \quad (19)$$

$$= \lim_{L\to\infty}\lim_{K\to\infty} \frac{1}{4(2\pi\sigma^2)^{N/2}} \sum_{l=1}^{L}\sum_{k=1}^{K} \exp\left(-\frac{\sum_{i=1}^{N}(r_i - g^{(i-1)}(y_l,\, x_k))^2}{2\sigma^2}\right) \delta y_l \delta x_k \quad (20)$$

where $\delta x_k = 2/K$; $x_k = -1 + (k-1)\delta x_k$; $\delta y_l = 2/L$; $y_l = -1 + (l-1)\delta y_l$. Similar to the skew tent map case, the accuracy of the numerical integrations (18) and (20) depends on the choice of L and K. Also, when the spreading factor increases, only the number of processes within the exponential function will be increased proportionally. Thus, the computational time should increase with a slower rate than the spreading factor. For the decoded symbol, it will be "+1" or "−1" depending upon $p(\mathbf{r} \mid ``+1\text{'' is sent})$ being larger or smaller than $p(\mathbf{r} \mid ``-1\text{'' is sent})$.
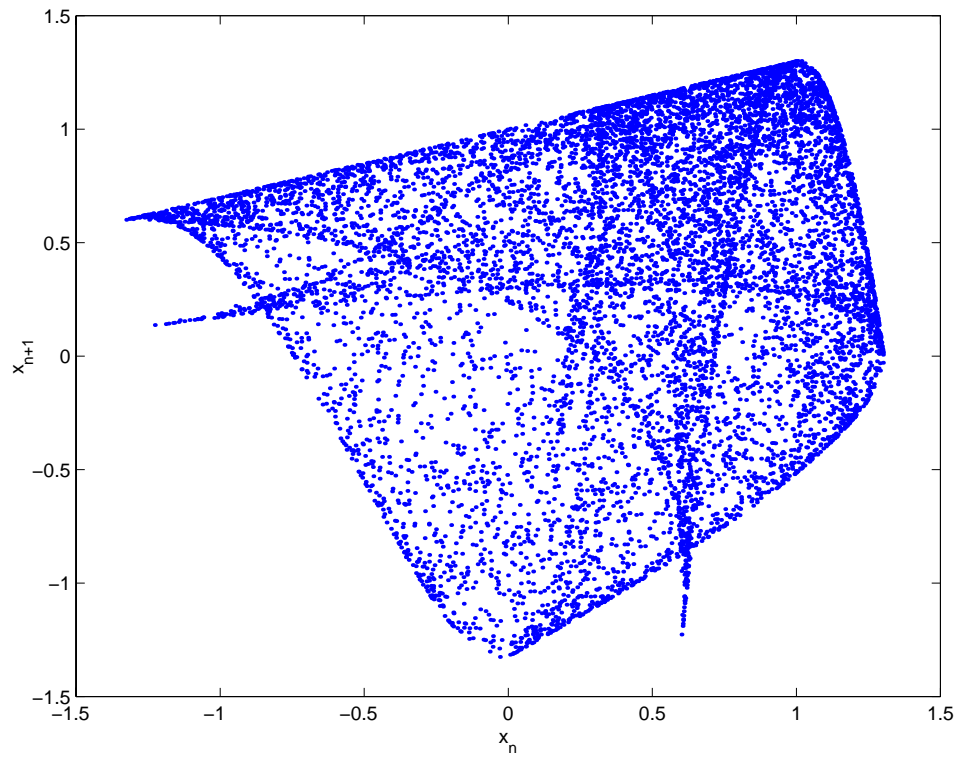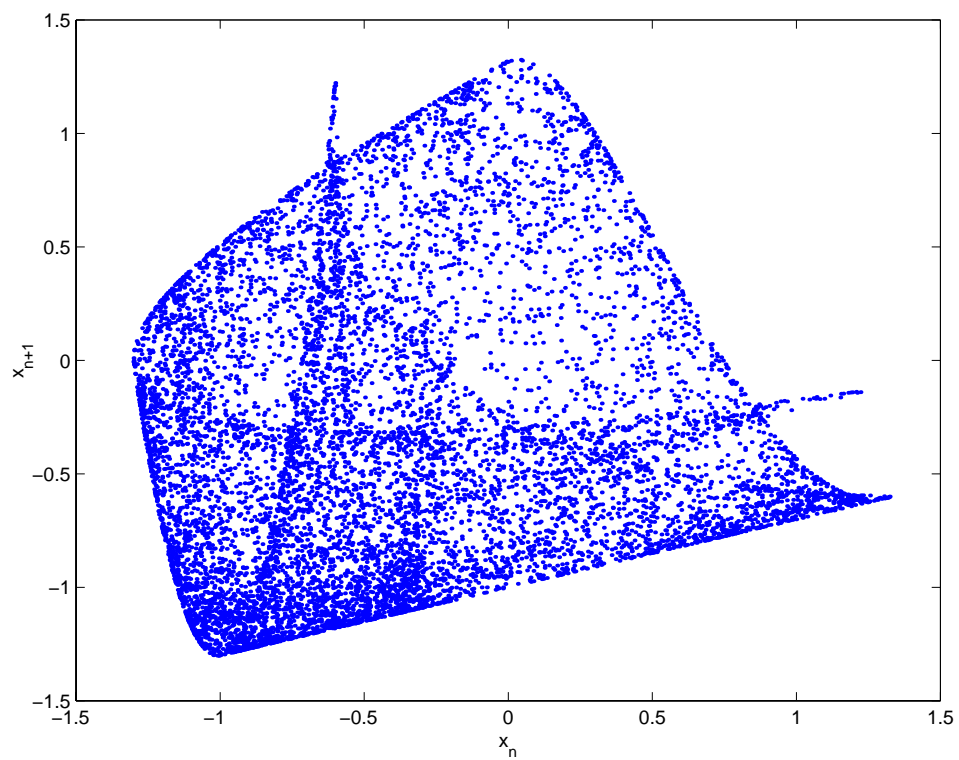
Fig. 2.   Attractors of the sample 2-D chaotic maps. (a) 2-D map f; (b) 2-D map g.

6. Simulations and Discussions

We have developed computer programs to compute (10) and (12) for the skew tent map, and (18) and (20) for the 2-D maps (15) and (16). Spreading factors ranging from 4 to 128 are used. In each simulation, 10^4 symbols are transmitted and the bit error rate (BER) is recorded for various E_b/N_0 values, where E_b denotes the average energy per bit. The simulation time (cpu-time) spent is also noted in each case.

In Table 1, the simulation time spent is tabulated against the spreading factor. For the skew tent map, $K = 2000$ has been used whereas in the case of the 2-D maps, $K = L = 100$ are used. It can be seen that the simulation time increases with the spreading factor. Further, the system employing the 2-D maps requires a longer simulation time compared with the case of the tent map. Figure 3 plots the simulation time versus spreading factor in logarithmic scales. The slope of both lines is found to be approximately 0.9. That is to say, the computing time T and the

Table 1. Simulation time spent using skew tent map and sample 2-D maps with various spreading factors.

Spreading Factor (N)	Simulation Time in Seconds (T)	
	Skew Tent Map ($K = 2000$)	2-D Maps ($K = L = 100$)
4	527	2806
6	736	3697
8	992	4818
10	1254	5893
12	1553	7090
14	1868	8293
16	2206	9070
32	3796	16203
64	6954	31200
128	13201	61616

spreading factor N are related by

$$T = c \times N^{0.9} \tag{21}$$

where c is some constant. Therefore, using numerical integration, the computing time required increases very slowly with the spreading factor.
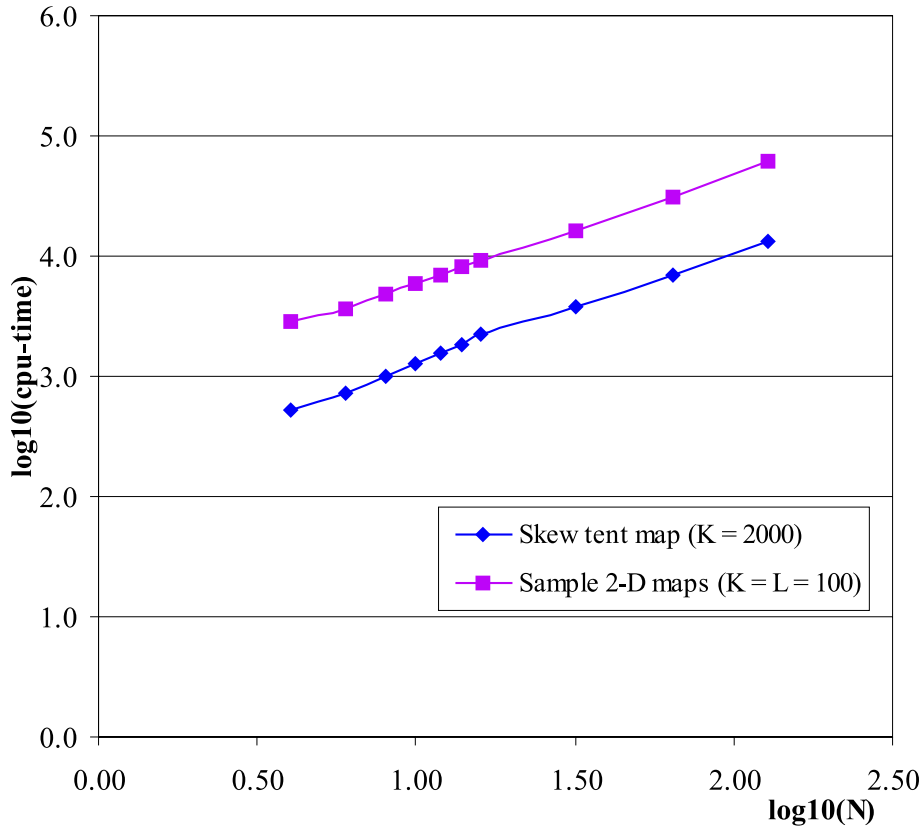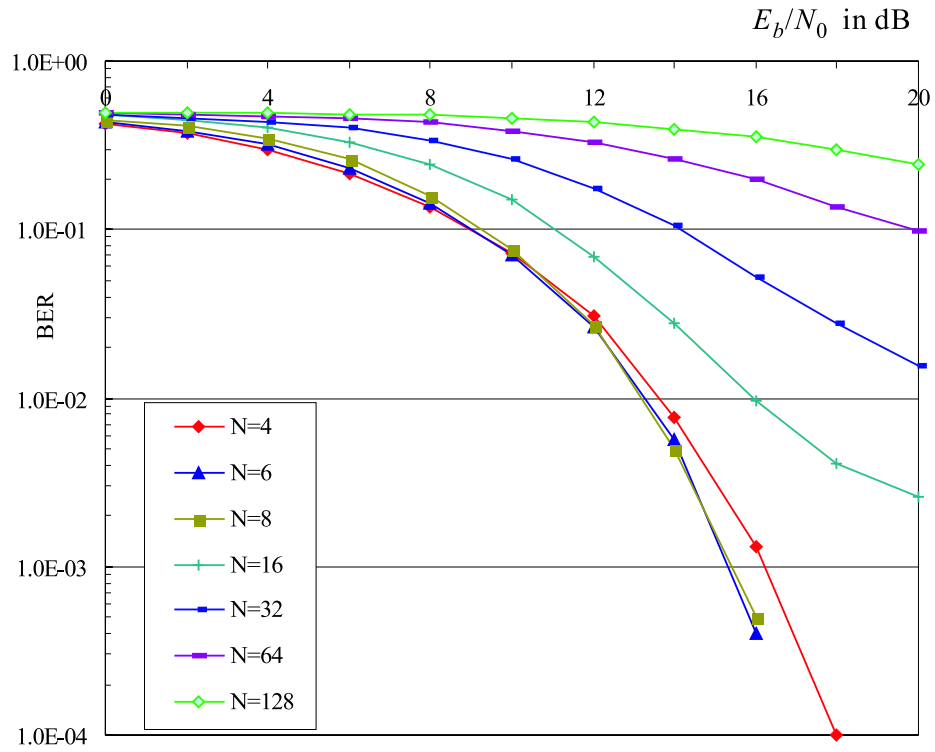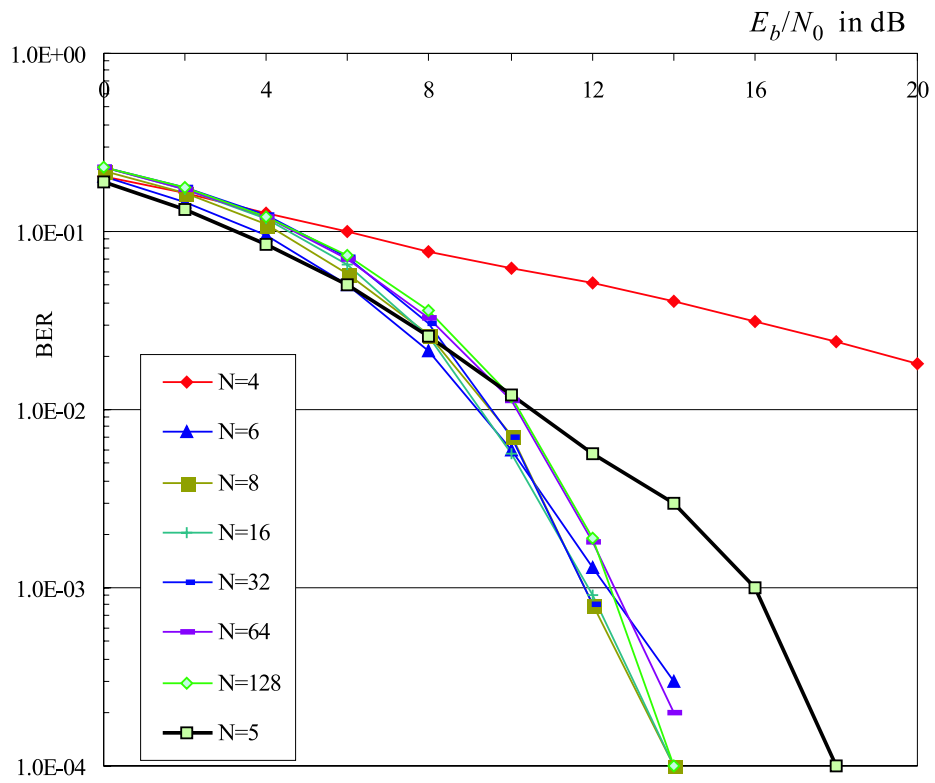

Fig. 3. Simulation time versus spreading factor in logarithmic scales.
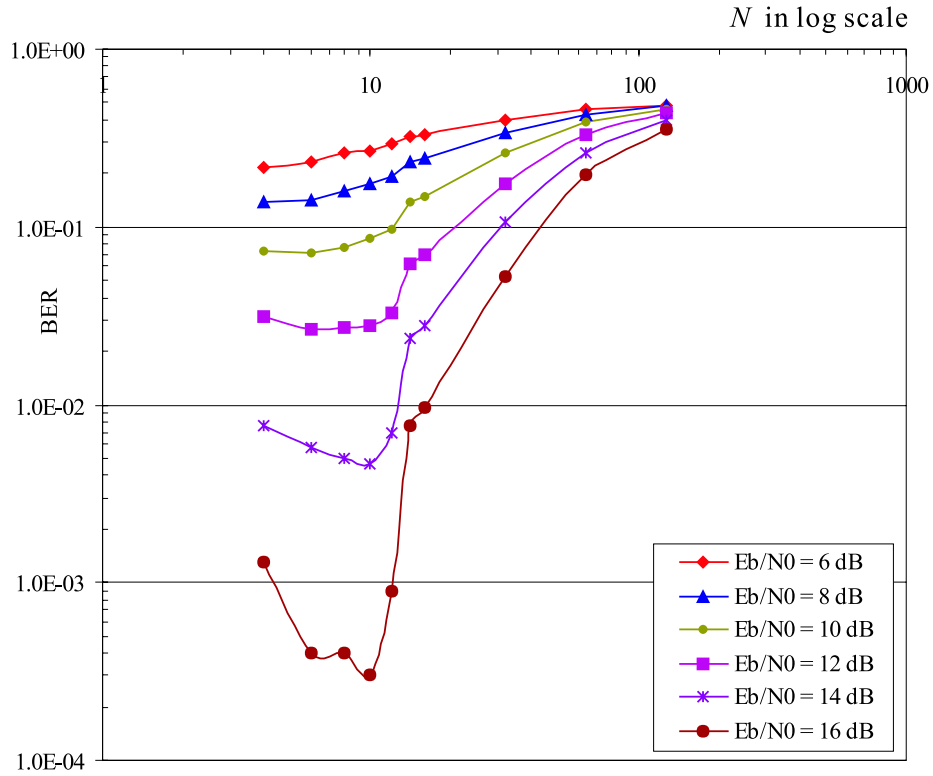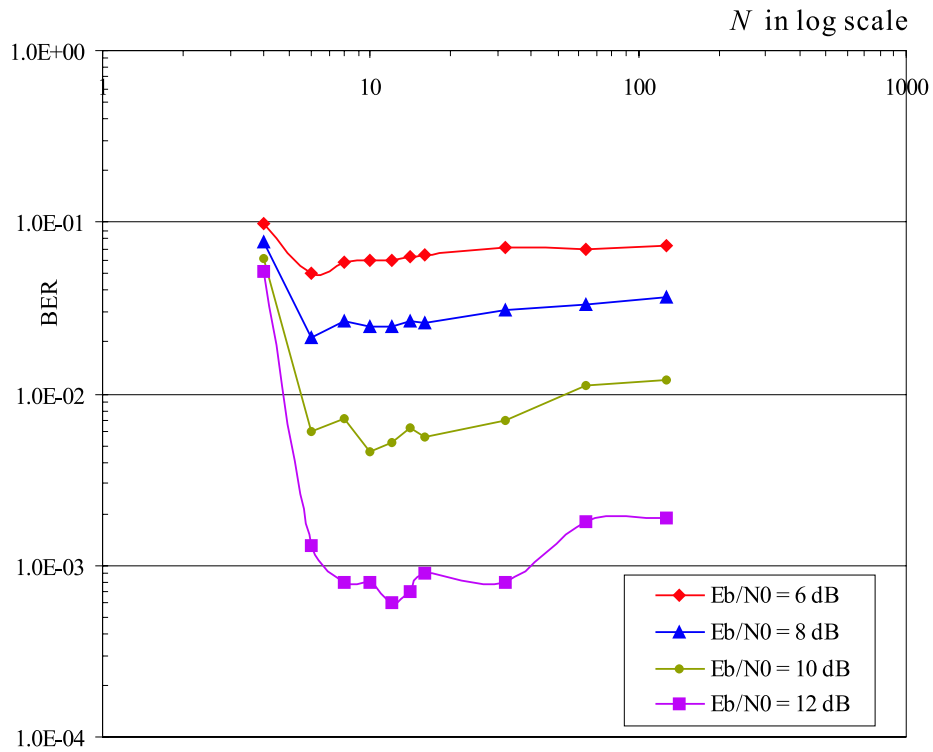
Fig. 4.   BERs versus E_b/N_0 for spreading factor 4 to 128 for a CSK system employing an optimal noncoherent detector. (a) Skew tent map ($K = 2000$); (b) sample 2-D maps ($K = L = 100$).
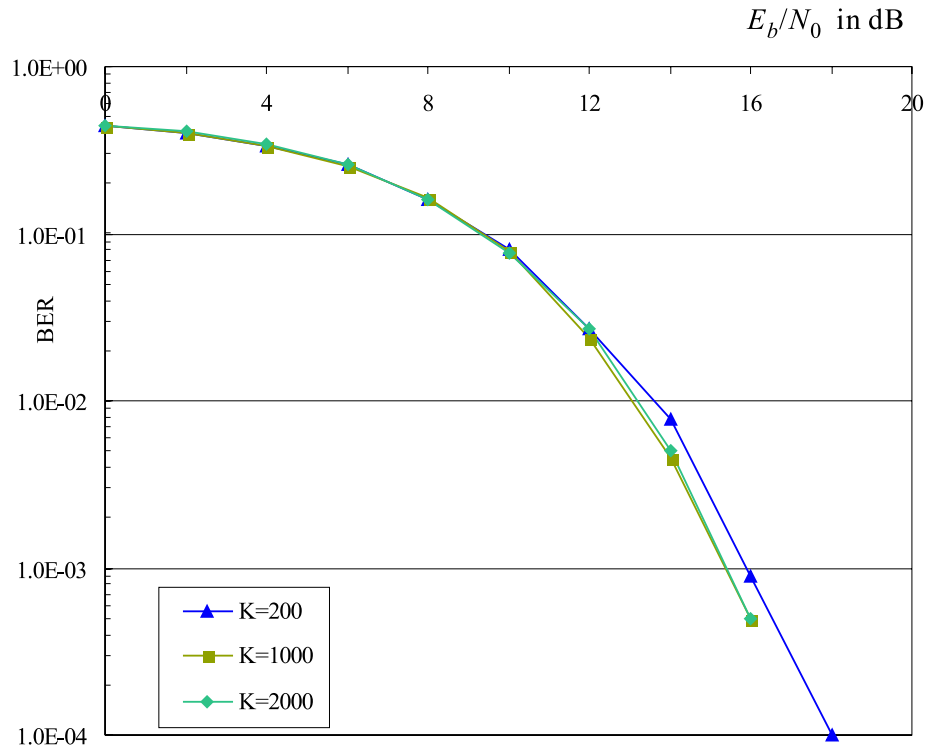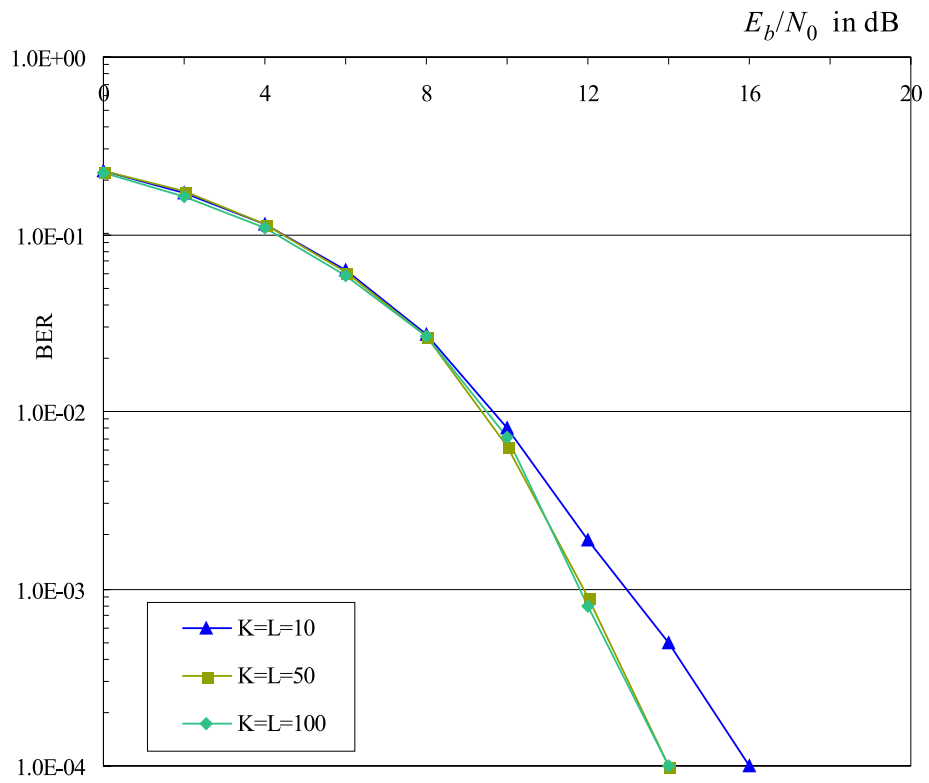
Fig. 5. BERs versus spreading factor (log scale) for different E_b/N_0 values for a CSK system employing an optimal noncoherent detector. (a) Skew tent map ($K = 2000$); (b) sample 2-D maps ($K = L = 100$).

Fig. 6.   BERs versus E_b/N_0 for different values of K and L for a CSK system employing an optimal noncoherent detector. Spreading factor $N = 8$. (a) Skew tent map; (b) sample 2-D maps.

Figures 4(a) and 4(b) plot the BERs versus E_b/N_0 for different spreading factors. It can be observed that the performance of the optimal detector improves as E_b/N_0 increases. Comparing the two figures, except when $N = 4$, the system using the 2-D maps outperforms the one using the skew tent map in all cases. Note that when $N = 4$, only two iterations have been performed by the 2-D maps, and the iterated sequence may not have converged to the attractors. As a consequence, the received signal block is not so easily distinguished, giving a higher error rate at $N = 4$. When N is increased to 5, Fig. 4(b) indicates that the system using the 2-D maps has a substantial improvement in BER. The results show that after three iterations, the chaotic sequences generated by the 2-D maps have converged to the attractors, making the receiving signal block more distinguishable.

In Fig. 5, the BER is plotted against the spreading factor (in log scale) for different E_b/N_0 values. In both systems, the BER reaches a minimum. For the tent map, it can be found that like other noncoherent detection techniques, the BER performance degrades as N becomes large [Sushchik *et al.*, 2000; Tse & Lau, 2003]. Moreover, the system using the 2-D maps degrades less substantially for large N. Note that in Fig. 5(b), the curves with E_b/N_0 values larger than 12 dB are not plotted since the BER is zero for some values of N.

Figures 6(a) and 6(b) plot the BERs versus E_b/N_0 for different values of K and L. Spreading factor of 8 is used. It can be observed that the BER improves slightly when K increases from 200 to 1000 for the skew tent map. When the value of K increases further to 2000, little BER enhancement is found. For the 2-D maps, improvement in BER is observed when the values of K and L are increased from 10 to 50. Further increases in K and L give little improvement.

7.   Conclusion

In this Letter, we apply numerical integration directly in the optimal detection of CSK signals. The merit of using numerical integration is that the detecting algorithm becomes independent of the chaotic maps being used. Also, the computational intensity will increase almost linearly as the spreading factor increases. Thus, the detection technique

can be applied even for large spreading factors, making the detector more practical. Finally, we extend the optimal decoding algorithm for maps of higher dimension. From the simulation results, it is observed that with the sample 2-D maps, a better bit error performance is achieved compared with that obtained from the tent map.

References

Dedieu, H., Kennedy, M. P. & Hasler, M. [1993] "Chaos shift keying: Modulation and demodulation of a chaotic carrier using self-synchronizing Chua's circuit," *IEEE Trans. Circuits Syst.-II* **40**, 634–642.

Hasler, M. & Schimming, T. [2000] "Chaos communication over noisy channels," *Int. J. Bifurcation and Chaos* **10**, 719–735.

Kennedy, M. P. & Kolumban, G. [2000] "Digital communication using chaos," in *Controlling Chaos and Bifurcation in Engineering Systems*, ed. Chen, G. (CRC Press, Boca Raton, FL), pp. 477–500.

Kolumbán, G., Kennedy, M. P. & Chua, L. O. [1998a] "The role of synchronization in digital communications using chaos — Part II: Chaotic modulation and chaotic synchronization," *IEEE Trans. Circuits Syst.-I* **45**, 1129–1140.

Kolumbán, G., Kis, G., Jáko, Z. & Kennedy, M. P. [1998b] "A robust modulation scheme for chaotic communications," *IEICE Trans. Fundam.* **E81-A**, 1798–1802.

Kolumbán, G. & Kennedy, M. P. [2000] "The role of synchronization in digital communications using chaos — Part III: Performance bounds for correlation receivers," *IEEE Trans. Circuits Syst.-I* **47**, 1673–1683.

Sushchik, M., Tsimring, L. S. & Volkovskii, A. R. [2000] "Performance analysis of correlation-based communication schemes utilizing chaos," *IEEE Trans. Circuits Syst.-I* **47**, 1684–1691.

Tse, C. K. & Lau, F. C. M. [2003] "A return-map regression approach for noncoherent detection in chaotic digital communications," *Int. J. Bifurcation and Chaos* **13**, 685–690.

Wu, C. W. & Chua, L. O. [1993] "Transmission of digital signals by chaotic synchronization," *Int. J. Bifurcation and Chaos* **3**, 1619–1627.