Chaos Communication: An Overview of Exact, Optimum and Approximate Results Using Statistical Theory

Anthony J. Lawrance

Department of Statistics
University of Warwick
Coventry CV4 7AL, UK
(email: a.j.lawrance@warwick.ac.uk)

Abstract: This paper overviews exact, optimum and approximate decoding and performance results for antipodal chaos shift-keying (CSK) in which a bit is transmitted by modulating a chaotic segment and decoded by use of the corresponding unmodulated reference segment. Both single- and multiple- user versions with both known- and transmitted- reference segments are considered, the so-called coherent and non-coherent cases. There are three main themes to the paper (i) the use of statistical likelihood theory for deriving optimum or improved decoders, (ii) the availability of mathematically exact theory for BER performance of decoders, (iii) qualitative statistical insights provided by simple Gaussian approximations to BER.

Keywords: Communication systems, Chaos shift keying, Correlation decoding, Likelihood-based decoding, Exact calculation of bit error rate, Gaussian approximations, Statistical theory

1. Introduction

Chaos-based communication involves using segments of chaotic waves to carry messages rather than the traditional sinusoidal waves. There are perceived advantages of chaos-based communication in terms of spread spectrum, security and robustness. Research activity in the area can be traced back to before 2000, but the set of papers in a special issue of the IEEE proceedings, Hasler, Mazzini *et al.*, 2002[3], was definitive at the time. Since then there has been much activity in the journals, a pair of linked monographs, Lau and Tse, 2003[9], Tam, Lau *et al.*, 2007[16], and three edited collections of chaos applications in telecommunications, Kennedy, Rovatti *et al.*, 2000[5], Stavroulakis, 2006[14] and Larson, Liu *et al.*, 2006[8]. The aim here is to give an overview of exact, optimum and approximate results from the point of view of statistical theory. There is much engineering leverage to be obtained from this perspective. Earlier results were transferred from conventional communication systems or derived with simple Gaussian approximations which did not fully incorporate the new joint chaotic-statistical interplay and sometimes tended to under-represent or over-complicate theoretical matters. A particular feature of chaos communications systems is the interplay of dynamical and statistical behaviour, and more particularly, the interplay of the

Figure 1. Block diagram of multi-user chaos shift-keying communication system.

statistical aspects of chaotic spreading and channel noise. The exact results give useful insights as well as opening computational paths, while the reworked and additional Gaussian approximations, although inaccurate, direct attention to key quantities determining performance. Modelling in the paper is restricted to employing discrete rather than continuous time, and thus is discrete time baseband. The models used are mathematical extractions of real systems, without electronic intricacies, but never-the-less allow theoretical investigation of performance and design issues.

2. Chaos Shift Keying Communication

Chaos shift-keying (CSK) is perhaps the most developed form of chaos-based communication, probably inspired by earlier spread spectrum systems, and can be traced back to Dedieu, Kennedy *et al.*, 1993[2] and Kocarev and Parlitz, 1995[6]. An engineering perspective on the area is given in the monographs by Lau and Tse, 2003[9] and Tam, Lau *et al.*, 2007[16].

To appreciate the structure of a CSK communication system, the block diagram in Figure 1 illustrates for L -users. The binary messages $b_l = \pm 1, l = 1, 2, \dots, L$ are to be simultaneously transmitted in a memoryless channel by the L users to L receivers. The focus is on a single active user, in the presence of interference by other-users. To transmit a binary message, user l has a chaotic generating map $\tau(\cdot)$ defined over $(-c, c)$, such as logistic or Bernoulli-shift, which provides a spreading segment, given by the row vector $x_l^T = (x_{l1}, \dots, x_{lN})$, where

$$x_{lt} = \tau(x_{l,t-1}) = \tau^{(t-1)}(x_{l1}), \quad \tau^{(t)}(\cdot) = \tau \left\{ \tau^{(t-1)}(\cdot) \right\}, \quad t = 1, 2, \dots, N. \quad (1)$$

The binary bit b_l is multiplied into each value of the spreading segment and is then transmitted as the message segment, $b_l x_{lj}, j = 1, 2, \dots, N$.

All the L bit-modulated spreading segments are transmitted additively in the same time slot and arrive at all receivers but with additive individual

channel noise, ε_{lj} , $j = 1, 2, \dots, N$ for the l^{th} user; this is white Gaussian noise with variance σ_ε^2 . In the simplest model, described as coherent, additionally the individual spreading segment of the l^{th} user is available without noise at the l^{th} receiver, either by chaotic synchronization or by identical chaotic generation, and called the reference segment. This segment is compared with message segment in order to decode b_l , or in statistical language, to estimate b_l . In the non-coherent or differential case, the spreading segments are separately and additionally transmitted and are received with channel noise or synchronization error. The main motivation of all the subsequent modeling is calculation of the bit error rate.

For the l^{th} and active-user transmitting a single bit, the message segment containing the bit information arriving at the l^{th} receiver is

$$r_{lj} = b_l x_{lj} + \sum_{k \neq l=1}^L b_k x_{kj} + \varepsilon_{lj}, \quad j = 1, 2, \dots, N \quad (2)$$

for $l = 1, 2, \dots, L$. The summation term represents interference by the message bits of other users. The decoding of the bit b_l in the received l^{th} message segment is done in conjunction with knowledge of the l^{th} spreading reference segment, but without knowledge of the other spreading segments at the receiver. The message bit b_l is treated as a statistical parameter and the optimum decoder is the likelihood-based estimate of b_l ; on the other hand, the bits and spreading segments of the other users are not of interest to the l^{th} user and are treated as unknown random variables.

A statistically-based view of bit error rate is that it is the sum of the conditional probabilities of decoding a +1 as -1 and *vice versa* weighted by the ± 1 transmission proportions. For most systems and decoders, by symmetry the conditional probabilities are equal, and so no weighting is necessary.

3. Single-User CSK: Likelihood Decoding

Developing the single-user case $L = 1$ is a useful way to unfold the primary approach without the complication of interference by other users; the suffix $l = 1$ is henceforth omitted in this case. The received message segment is, by taking $L = 1$ in (2),

$$r_j = b x_j + \varepsilon_j, \quad j = 1, 2, \dots, N. \quad (3)$$

With the Gaussian white noise channel, the likelihood of the transmitted bit b and the unknown noise variance σ_ε^2 is obviously

$$l(b, \sigma_\varepsilon^2 | r) = \left(\sigma_\varepsilon \sqrt{2\pi} \right)^N \exp \left\{ -\frac{1}{2\sigma_\varepsilon^2} \sum_{j=1}^N (r_j - b x_j)^2 \right\}. \quad (4)$$

Maximum likelihood estimation implies choosing the value of b which has the greater likelihood, and the appropriate likelihood ratio simplifies as

Lawrance, A. J.

$$\sum_{j=1}^N x_j r_j \equiv C_{rx}, \quad (5)$$

Hence $\hat{b}$, the maximum likelihood estimate of b , is given by

$$C_{rx} \geq 0 \Rightarrow \hat{b} = +1, C_{rx} < 0 \Rightarrow \hat{b} = -1. \quad (6)$$

This is known as a correlation decoder. From a statistical point of view, it is proportional to the standard estimate of the regression coefficient b in the line through the origin (3); the twist is that the regression coefficient can only take one of the two values ± 1 . Thus, the form (6), positive if $\hat{b}$ is $+1$, negative if $\hat{b}$ is -1 , is intuitively sensible; it does not depend on the type of spreading, chaotic or not, but is conditional on spreading.

4. Single-User CSK: Exact and Approximate Bit Error Rate

This is particularly a topic where careful statistical argument yields simple insightful and computational results which were not otherwise available, first derived in the author's paper Lawrance and Ohama, 2003[11]. The emphasis is on an exact result for the bit error rate. Unlike the approximate results quoted in several key early papers, such as Kolumban, Kennedy *et al.*, 2002[7], Sushchik, Tsimring *et al.*, 2000[15], which strangely did not involve the type of spreading, the exact result involves all aspects of the system, in particular, the bit energy sum-of-squares. Interestingly, its lower bound for extensive spreading yields the corresponding BPSK result. By involving the spreading, the exact result gives a pathway to design of spreading for maximum performance, Yao, 2004[18] and Papamarkou Chaos, 2009[1].

Following earlier remarks in Section 2, the overall bit error rate can be taken as the conditional probability

$$BER = P(\hat{b} = -1 | b = 1) \quad (7)$$

and so by (6) - (7),

$$BER = P(C_{rx} < 0 | b = 1) = P\left(\sum_{j=1}^N \varepsilon_j x_j + \sum_{j=1}^N x_j^2 < 0\right) \quad (8)$$

which involves both the statistical noise and the chaotic spreading as random variables.

In the exact approach, this probability is evaluated exactly, while in the simple Gaussian approximation approach (SGA), the random variables in the probability are first approximated with a Gaussian distribution.

The exact approach begins by conditioning on the spreading values and using the independent Gaussian distributions of the ε_j and the Gaussian distribution of their linear combination in (8); then continues by averaging over the chaotic spreading segments. The exact BER result becomes

$$BER = E_x \left\{ \Phi \left(-\sqrt{SNR \sum_{j=1}^N x_j^2 / N \sigma_x^2} \right) \right\} \quad (9)$$

where E_x is the expectation over the chaotic spreading segments, marginally with means zero and variances σ_x^2 , $\Phi(\cdot)$ is the cumulative distribution function of a standardized Gaussian distribution, SNR is the signal-to-noise ratio defined as $N\sigma_x^2/\sigma_e^2$, and the second part of the square root is the standardized bit energy.

In order to calculate BER the expectation in (9) has to be evaluated numerically. In the case of chaotic spreading this requires a one-dimensional integral; with chaotic map $\tau(\cdot)$ according to (1) and its invariant distribution $\varphi(\cdot)$, the result (9) becomes

$$\int_{-c}^c \left\{ \Phi \left(-\sqrt{SNR \sum_{j=1}^N \tau^{(j-1)}(x)^2 / N \sigma_x^2} \right) \right\} \varphi(x) dx. \quad (10)$$

This can be exactly numerically evaluated in cases where the map has a explicit and computable iterated form. Illustrations of BER as a function of SNR are given in Figure 2 for logistic, Bernoulli-shift and Gaussian spreading. Logistic is better than Bernoulli-shift and Gaussian is by far the worst.

For large N , the standardized bit energy term in (9) converges to 1, leaving the result for extensive spreading as

$$\lim_{N \rightarrow \infty} BER = \Phi(-\sqrt{SNR}). \quad (11)$$

This result is in fact the lower bound of (9) for all N , the proof being from applying Jensen's expectation inequality to (9), see Figure 2. The exact result can be used until the result (11) is acceptably accurate. Note that the spreading limit (11) and lower bound do not depend on the type of spreading, chaotic or not. In early work, (11) was used as the exact BER result. It is exact for binary phase shift-keying (*BPSK*) in which spreading takes one of two oppositely signed values with fixed probability and thus bit energy is constant. This has led to the view that *CSK* cannot improve on *BPSK*, but it is very different, e.g., with the advantages of a spread spectrum and security.

When the chaotic spreading map does not have an explicit iterated form, exact numerical computations from (9) are intractable. The approach suggested in Lawrance and Ohama, 2005[12] was to approximate the bit energy sum of squares by a chi-squared distribution with matching mean and variance. In current work, Kaddoum, Charge *et al.*, 2007[4], Chaos, 2009[1], have developed an approach in which the bit energy distribution is either simulated or matched by a Rician distribution, leading to analytical results.

The SGA approach when applied to (8) yields

$$BER \approx \Phi \left\{ -\left[\frac{1}{SNR} + N^{-1} \frac{\sigma_{x^2}^2}{\sigma_x^4} \left\{ 1 + 2 \sum_{j=1}^{N-1} \left(1 - \frac{k}{N} \right) \rho_{x^2}(j) \right\} \right]^{-\frac{1}{2}} \right\}, \quad (12)$$

Lawrance, A. J.

as first derived in Lawrance and Balakrishna, 2001[10]. While this result is seriously inaccurate, partly shown by its incorrect limiting form as $SNR \rightarrow \infty$, the roles of kurtosis and quadratic autocorrelation of the spreading are revealed. They are illustrated by Figure 2 in that logistic spreading with its zero quadratic autocorrelations is better than Bernoulli-shift spreading with its geometric one-quarter quadratic autocorrelations. The corresponding SGA for the non-optimum correlation decoding in non-coherent CSK is

$$BER \approx \Phi \left\{ - \left[\frac{2}{SNR} + \frac{N}{SNR^2} + N^{-1} \frac{\sigma_{x^2}^2}{\sigma_x^4} \left\{ 1 + 2 \sum_{j=1}^{N-1} \left(1 - \frac{k}{N} \right) \rho_{x^2}(j) \right\} \right]^{\frac{1}{2}} \right\}. \quad (13)$$

Figure 2. BER plotted against SNR in dB for coherent CSK and Jensen lower bound. Comparison of logistic (long dashes), Bernoulli-shift (short dashes), and Gaussian (dot-dashes) spreading, spreading factor $N = 5$.

Both (12) and (13) give statistical structure and simplicity to the computational results in Lau and Tse, 2003[9]; the additional terms in (13) indicate the loss in performance due the transmission of reference segments in place of their exact replication at the receiver. An exact BER result for correlation decoding in single-user non-coherent CSK is given in Lawrance and Ohama, 2003[11] in terms of non-central F-distributions, and optimal likelihood-based decoding is has been developed by T. Papamarkou, Chaos, 2009[1].

An alternative direction for the non-coherent case avoids the transmission of a reference segment by use of a particular chaotic attractor which itself carries the necessary demodulation information; this approach is currently being developed, Xu, Charge *et al.*, 2008[17], Chaos, 2009[1].

5. Multi-User CSK: Coherent and Non-Coherent

In any multi-user communication system the new feature relative to the corresponding single-user system is the interference of other users when decoding a bit, and the effect this has on the *BER*. As far as an individual user is concerned, the other-users constitute additional non-Gaussian noise. A

correlation decoder could still be used but would not be as good in BER terms as a likelihood-based decoder, as will be seen. The likelihood decoder of the multi-user CSK system requires a careful analysis which significantly departs from that of the single-user case, Lawrance and Yao, 2008[13], and leads to the generalized correlation decoder

$$\hat{b}_i = \begin{cases} +1 & x_i^T \Sigma_x^{-1} r > 0 \\ -1 & \text{otherwise} \end{cases}. \quad (14)$$

where Σ_x is the circulant autocorrelation matrix of the spreading segment. There is a pleasing statistical simplicity to this result from its obvious generalization of the correlation decoder; it comes from a Gaussian approximation to the likelihood. A key new point is that the decoder now makes use of any autocorrelation in the spreading.

In the interests of brevity, theoretical developments leading to multi-user BER results are omitted; they involve conditional Gaussian approximation (CGA) and give near exact results. Accurate Gaussian approximations are made during calculation of bit error probability which are conditional on spreading. Concise expressions can be obtained for both correlation decoders and generalized correlation decoders, and have been given in Lawrance and Yao, 2008[13] for coherent CSK. For the generalized correlation decoder, its bit error has been derived as

$$BER_{gcd} = E_x \left\{ \Phi \left(- \sqrt{ \left(x^T \Sigma_x^{-1} x / N \sigma_x^2 \right) \left[\left(\frac{x^T \Sigma_x^{-1} \Sigma_x^{-1} x}{x^T \Sigma_x^{-1} x} \right) \frac{1}{SNR} + \frac{1}{SIR} \right]^{-1} } \right) \right\} \quad (15)$$

where SIR is a new quantity, the ‘spreading to other user interference ratio’ (SIR), defined as $N/(L-1)$; this will be important in all multi-user results.

For the ordinary correlation decoder there is correspondingly

$$BER_{cd} = E_x \left\{ \Phi \left(- \sqrt{ \left(\frac{N^{-1} (x^T x)^2}{(x^T \Sigma_x X) \sigma_x^2} \right) \left[\left(\frac{x^T x}{x^T \Sigma_x x} \right) \frac{1}{SNR} + \frac{1}{SIR} \right]^{-1} } \right) \right\}. \quad (16)$$

The advantage of the generalized correlation decoder is clear from Figure 3.

When spreading is uncorrelated, as it is for logistic spreading, both (15) and (16) become much simpler in terms of standardized bit energy as

$$BER_{cd}^{uncor} = E_x \left\{ \Phi \left(- \sqrt{ \left(x^T x / N \sigma_x^2 \right) \left(\frac{1}{SNR} + \frac{1}{SIR} \right)^{-1} } \right) \right\} \quad (17)$$

and consequently have the lower bound

$$BER_{sprdg}^{uncor} \geq \Phi \left(- \sqrt{ \left(\frac{1}{SNR} + \frac{1}{SIR} \right)^{-1} } \right). \quad (18)$$

The result may be a lower bound over correlated spreading as well, as suggested in Figure 3, but this has not been proved. Note also that the BER

Figure 3. Three-user ($L = 3$) coherent CSK with spreading factor $N = 4$ and $SIR = 2$. BER comparison of SGA approximation for correlation decoding with Bernoulli-shift spreading (long-short dashes), with near exact CGA results for Bernoulli-shift spreading with correlation decoding (long-dashes), Bernoulli-shift spreading with optimum decoding (short-dashes), Jensen lower bound for uncorrelated spreading (solid curve) and universal lower bound (horizontal line at 0.07865).

curves in Figure 3 level out rather than continue to decrease; this effect is evident from (15) - (18) by setting SNR to infinity, effectively saying there is only interference noise.

There are also SGA results for multi-user CSK with correlation decoding which are informative as to the key statistical quantities involved; under less general assumptions and presented more computationally, they have been used in the approach adopted by Lau and Tse, 2003[9] and Tam, Lau *et al.*, 2007[16]. In the coherent case,

$$BER \approx \Phi \left\{ - \left[\frac{1}{SNR} + N^{-1} \frac{\sigma_{x^2}^2}{\sigma_x^4} \left\{ 1 + 2 \sum_{k=1}^{N-1} \left(1 - \frac{k}{N} \right) \rho_{x^2}(k) \right\} + \frac{1}{SIR} \left\{ 1 + 2 \sum_{k=1}^{N-1} \left(1 - \frac{k}{N} \right) \{\rho_x(k)\}^2 \right\} \right]^{\frac{1}{2}} \right\} \quad (19)$$

and in the non-coherent case

$$BER \approx \Phi \left\{ - \left[\frac{2}{SNR} + \frac{N}{SNR^2} + N^{-1} \frac{\sigma_{x^2}^2}{\sigma_x^4} \left\{ 1 + 2 \sum_{k=1}^{N-1} \left(1 - \frac{k}{N} \right) \rho_{x^2}(k) \right\} + \frac{N}{SNR \cdot SIR} + \frac{1}{SIR} \left\{ 1 + 2 \sum_{k=1}^{N-1} \left(1 - \frac{k}{N} \right) \{\rho_x(k)\}^2 \right\} \right]^{\frac{1}{2}} \right\}. \quad (20)$$

The aspects which are not involved in the corresponding single-user results (12) and (13) are SIR and linear autocorrelations, and both are seen to increase BER, an increase being intuitively inevitable because of

interference. The results (19) and (20) may insightfully be compared, with the extra terms in (20) suggesting increased BER, again intuitively inevitable.

Exact results for correlation decoding are much more complicated than SGA results, and involve the distribution function of a doubly non-central-F random variable; some computational progress has been made in Yao and Lawrance, 2006[19]. SGA results for generalized correlation decoding have not yet been developed.

6. Conclusions

This paper has been a selective and compacted overview of performance modelling in chaos shift-keying communication. The emphasis has been to present mathematical statistically-based exact and approximate structural results which are useful for computation and insightful for qualitative understanding. The paper could not cover the whole field, even for chaos shift-keying, and hence topics such as optimality in choice of chaotic map, interference jamming, optimality in spreading factor for non-coherent systems, and laser-based chaotic systems, were omitted. Some of these and others too may be covered elsewhere in Chaos, 2009[1].

References

- [1] Chaos, Presented at *The 2nd Chaotic Modelling and Simulation International Conference*, Chania, Crete, Greece, 2009.
- [2] H. Dedieu, M.P. Kennedy and M. Hasler, "Chaos shift keying: modulation and demodulation of a chaotic carrier using self-synchronizing Chua's circuits," *IEEE Transactions on Circuits and Systems-II: Analog and Digital Signal Processing*, vol. 40, pp. 634-642, 1993.
- [3] M. Hasler, G. Mazzini, M. Ogorzalek, M. Rovatti and G. Setti, "Scanning the special issue: Applications of nonlinear dynamics and information engineering," *Proceedings of the IEEE*, vol. 90, pp. 10, 2002.
- [4] G. Kaddoum, P. Charge, D. Roviras and D. Fournier-Prunaret, "Comparison of chaotic sequences in a chaos based DS-CDMA system," in *International Symposium on Nonlinear Theory and its Applications (NOLTA)*. Vancouver, Canada, 2007.
- [5] M.P. Kennedy, R. Rovatti and G. Setti, *Chaotic Electronics in Telecommunications*. London: CRC Press, 2000.
- [6] I. Kocarev and U. Parlitz, "General approach for chaotic synchronization with application to communication," *Physical Review Letter*, vol. 74, pp. 5028-5031, 1995.
- [7] G. Kolumban, P. Kennedy, Z. Jako and G. Kis, "Chaotic communications with correlator receivers: theory and performance limits," *Proceedings of the IEEE*, vol. 50, pp. 711-732, 2002.

Lawrance, A. J.

- [8] L.E. Larson, J.-M. Liu and L.S. Tsimring, "Digital Communications Using Chaos and Nonlinear Dynamics." New York: Springer, 2006.
- [9] F.C.M. Lau and C.K. Tse, *Chaos-based Digital Communications Systems*. Heidelberg: Springer Verlag, 2003.
- [10] A.J. Lawrance and N. Balakrishna, "Statistical aspects of chaotic maps with negative dependence in a communications setting," *Journal Royal Statistical Society, Series B*, vol. 63, pp. 843-853, 2001.
- [11] A.J. Lawrance and G. Ohama, "Exact calculation of bit error rates in communication systems with chaotic modulation," *IEEE Trans. Circuits Systems I: Fund. Theory Appl.*, vol. 50, pp. 1391-1400, 2003.
- [12] A.J. Lawrance and G. Ohama, "Bit error probability and bit outage rate in chaos communication," *Circuits, Systems, and Signal Processing*, vol. 24, pp. 519-534, 2005.
- [13] A.J. Lawrance and J. Yao, "Likelihood-based demodulation in multi-user chaos shift keying communication," *Circuits, Systems, and Signal Processing*, vol. 27, pp. 847-864, 2008.
- [14] P. Stavroulakis, *Chaos Applications in Telecommunications*. New York: CRC Press, 2006.
- [15] M. Sushchik, L.S. Tsimring and A.R. Volkovskii, "Performance analysis of correlation-based communication schemes utilizing chaos," *IEEE Transactions on Circuits and Systems-1: Fundamental Theory and Applications*, vol. 47, pp. 1684-1691, 2000.
- [16] W.M. Tam, F.C.M. Lau and C.K. Tse, *Digital Communications with Chaos*. Oxford: Elsevier, 2007.
- [17] Y. Xu, P. Charge and D. Fournier-Prunaret, "Chaotic cyclic attractors shift keying," in *International Conference Neural Networks and Signal Processing*, Zhenjiang, China, 2008, pp. 62-66.
- [18] J. Yao, "Optimal chaos shift keying communications with correlationa decoding," in *IEEE International Symposium on Crcuits and Systems (ISCAS2004)*, vol. IV. Vancouver, Canada, 2004, pp. 593-596.
- [19] J. Yao and A.J. Lawrance, "Performance analysis and optimization of multi-user differential chaos shift-keying communication systems," *IEEE Transactions on Circuits and Systems-1: Regular Papers*, vol. 53, pp. 2075-2091, 2006.