Statistical Science
2016, Vol. 31, No. 4, 558–577
DOI: 10.1214/16-STS572

Chaos Communication: A Case of Statistical Engineering

Anthony J. Lawrance

Abstract. The paper gives a statistically focused selective view of chaos-based communication which uses segments of noise-like chaotic waves as carriers of messages, replacing the traditional sinusoidal radio waves. The presentation concerns joint statistical and dynamical modelling of the binary communication system known as "chaos shift-keying", representative of the area, and leverages the statistical properties of chaos. Practically, such systems apply to both wireless and optical laser communication channels. Theoretically, the chaotic waves are generated iteratively by chaotic maps, and practically, by electronic circuits or lasers. Both single-user and multiple-user systems are covered. The focus is on likelihood-based decoding of messages, essentially estimation of binary-valued parameters and efficiency of the system is in terms of the probability of bit decoding error. The emphasis is on exact theoretical results for bit error rate, their structured approximations and engineering interpretations. Design issues, optimality of performance, interference and fading are other topics considered. The statistical aspects of chaotic synchronization are involved in the modelling of optical systems. Empirical illustrations from an experimental laser-based system are presented. The overall aim is to show the use of statistical methodology in unifying and advancing the area.

Key words and phrases: Bit error rate, chaos communications, chaos modelling, chaos shift-keying, communications engineering, decoding as likelihood-based statistical inference, empirical communication system analysis, Gaussian approximations, laser chaos, nonlinear dynamics, optical noise, statistical modelling, statistical time series, synchronization error.

1. INTRODUCTION TO CHAOS COMMUNICATIONS

Communications involving chaotic waves is an area in which there is much synergy between engineering and statistical modelling. Chaos communication systems use segments of noise-like chaotic waves, rather than traditional sinusoidal waves, to carry messages. The chaotic waves are generated mathematically by chaotic maps or physically by electronic means, such as electrical circuits or lasers. The area started in physics some 25 years ago with remote synchronization of chaotic waves, the phenomenon that two remotely generated chaotic sequences can proceed in unison; it then travelled in a multi-disciplinary way into mathematical chaos theory, communications engineering and statistical modelling. The value of synchronization is that it provides remote although imperfect information at the receiver about a segment generated at a transmitter. This implies in the communication setting that for each received wave segment carrying a message there is also available at the receiver an imperfect version of its segment without the message, sometimes referred to as the reference segment. In electrical circuit systems, the reference segment is known from transmission rather than by synchronization. In either case, there is then enough information in the two segments for the transmitted message to be decoded by statistical estimation on the basis of the system model. This involves statistical properties of

Anthony J. Lawrance is Professor, Department of Statistics, University of Warwick, Coventry, United Kingdom (e-mail: a.j.lawrance@warwick.ac.uk).

the chaotic waves, transmission noise and, additionally, synchronization error for laser-based systems. Transmission noise is either electronic or optical. The models used are mathematical extractions of the engineering communication system, but nevertheless allow theoretical investigation of performance and optimal design.

The presentation here particularly concerns statistical and mathematical modelling of the binary communication system known as *antipodal chaos shift-keying* (CSK) which has typical characteristics of systems in the area. Shift-keying is a variously used communications description and here refers to the multiplicative modulation of a chaotic wave segment by a binary bit b, with antipodal implying $b = \pm 1$. One aim of the paper is to mathematically specify the antipodal chaos shift-keying system and assert the importance of its statistical features, thereby clarifying and unifying earlier approaches and their differing interdisciplinary styles. An important communication concern is bit error rate, essentially a statistical measure of performance, and so there is emphasis on its theoretical calculation by exact and approximate mathematics—not found in much of the engineering literature. Initial focus is on single-user systems, but extensions to more realistic multi-user situations are also covered.

In this area, the statistical aspects of chaos and synchronization are more central than the dynamical ones. Kohda and Murao (1990) gave an early account of the statistical properties of chaotic maps based on the Frobenius–Perron operator. Berliner (1992) made a strong connection between statistics and chaos in a wide ranging and at times philosophical discussion. An accessible statistical introduction to the useful invariant statistical distributions of chaos is given in the early chapters of Lasota and Mackey (1994) while Lawrance and Balakrishna (2008) elaborate the dependency aspects most relevant in the present context. The topic of chaos synchronization was initially explored by Pecora and Carroll (1990), Carroll and Pecora (1992), who realized that it could be applied in communications. The earliest work on antipodal CSK appears to be by Parlitz and Ergezinger (1994) in the so-called coherent case when the carrying wave segment is known exactly at the receiver, and by Kolumbán et al. (1996) in the so-called non-coherent or differential case when the carrying wave segment is also transmitted; continuing work was published in two special issues of the Proceedings of the IEEE, Kolumbán and Kennedy (2000) and Hasler et al. (2002). Papers in these collections can be seen as the foundational work in the area. Subsequently there has been a linked pair of monographs consolidating much of the earlier work, Lau and Tse (2003) and Tam, Lau and Tse (2007) and paper collections by Kennedy, Rovatti and Setti (2000), Larson, Liu and Tsimring (2006). While these developments were mainly based on electronic circuit generation of chaos, at about the same time there was independent and parallel experimental work concerning the communications use of remotely synchronized chaotic lasers, beginning with Mirasso, Colet and Garcia-Fernandez (1996) and which was later overviewed in a featured section of the IEEE Journal of Quantum Electronics, Donati and Mirasso (2002). A comprehensive account of laser communication from an engineering perspective is given by Uchida (2012).

The antipodal chaos shift-keying system is illustrated in the block diagram of Figure 1. A chaotic so-called *spreading segment* $X = \{X_i\}_{i=1}^N$ is taken from a chaotic wave of mean zero and variance σ_X^2 generated at the transmitter, either mathematically or by physically; N is called the *spreading factor*. The bit to transmitted b, multiplicatively modulates the spreading

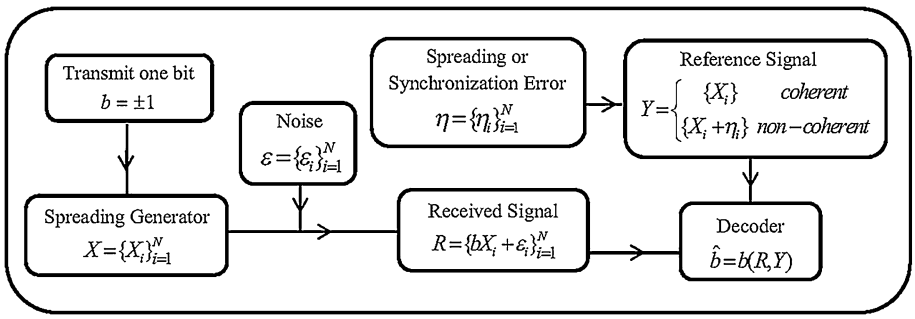


FIG. 1. *Block diagram of antipodal chaos shift-keying communication systems in both the coherent and non-coherent forms. In the coherent form, the reference signal is known exactly at the receiver, and in the non-coherent form it is known inexactly at the receiver by transmission or synchronization.*

segment as $bX = \{bX_i\}_{i=1}^N$ which then passes through a noisy channel and becomes the *message segment* $R = \{bX_i + \varepsilon_i\}_{i=1}^N$ at the receiver where $\{\varepsilon_i\}_{i=1}^N$ is the associated *message noise*. Also, the spreading segment at the transmitter X, may be transmitted to the receiver or synchronized there, where it is known as the *reference segment* $Y = \{X_i + \eta_i\}_{i=1}^N$; here $\{\eta_i\}_{i=1}^N$ is the associated *reference* or *synchronization error*, as applicable. If there is no reference or synchronization error, the system is called *coherent*; one such way would be by perfect synchronization at the receiver. Otherwise, the system is said to be *non-coherent*. The terminology here is not consistent in the communications literature; sometimes non-coherent implies a system without the need for a reference segment; sometimes coherent is used if there is one, whether it is known exactly or not.

The system model equations for antipodal chaos shift-keying communication concern the transmission of a single binary message. The standard assumption is that the system is *memoryless*, with no dependency between bits. From Figure 1, the equation for the received message segment for a single bit is seen to be

$$R_i = bX_i + \varepsilon_i, \quad i = 1, 2, \ldots, N, \tag{1}$$

where $\{\varepsilon_i\}$ is independent and identically distributed channel noise of mean zero and variance σ_ε^2. In a coherent system, covered in Sections 2–4, the reference segments are known, although generated chaotically, and this is the only equation required. Section 5 deals with the non-coherent system in which the reference segments are either transmitted to the receiver or imperfectly known there by synchronization. In these cases, there is a second model equation for the reference segments

$$Y_i = X_i + \eta_i, \quad 1 = 1, 2, \ldots, N. \tag{2}$$

Equations (1) and (2) provide the theoretical model of the system, called a base-band model in the communications literature, and lead to the statistical basis of decoding the received bits and the bit error rate. Independent Gaussian assumptions are usually made for the noise and error terms.

2. MAXIMUM LIKELIHOOD DECODING OF MESSAGES AND BIT ERROR THEORY

In the statistical approach to decoding, a received bit b is treated as an unknown statistical parameter in the system model and is estimated from the available information, which in the case of coherent antipodal CSK is the message segment R and its spreading segment X. Intuitively, it can be seen that the correlation between R and X should have the same sign as b, and thus provide an estimate of b. Verifying this as the maximum likelihood decoder in the simplest coherent case of antipodal coherent CSK with Gaussian channel noise is a useful way to unfold the primary theoretical approach to decoding. The accuracy of the estimate is naturally assessed by its error rate, usually the most important communication measure of performance, and more appropriate than the variance. The exact Gaussian theory (EGT) is of central relevance and will be given, somewhat extended from the original result in Lawrance and Ohama (2003).

Knowing the chaotic segment $X^T \equiv (X_1, X_2, \ldots, X_N)$ exactly, the Gaussian likelihood of a transmitted bit b based on received message segment $R^T \equiv (R_1, R_2, \ldots, R_N)$, is

$$\begin{aligned} &l\big(b, \sigma_\varepsilon^2 \mid R, X\big) \\ &\quad = (\sigma_\varepsilon\sqrt{2\pi})^N \exp\left\{-\frac{1}{2\sigma_\varepsilon^2}\sum_{i=1}^N (R_i - bX_i)^2\right\}. \end{aligned} \tag{3}$$

This is first maximized over σ_ε^2 as a function of b when only interested in estimating b, giving

$$\sigma_\varepsilon^2(b) = N^{-1}\sum_{i=1}^N (R_i - bX_i)^2 \tag{4}$$

and hence the marginal likelihood for b is

$$\begin{aligned} &l\big(b, \sigma_\varepsilon^2(b) \mid R, X\big) \\ &\quad = (2\pi e/N)^{N/2}\left\{\frac{1}{N}\sum_{i=1}^N (R_i - bX_i)^2\right\}^{-N/2}. \end{aligned} \tag{5}$$

Maximum likelihood estimation implies choosing the value of b which has the greater likelihood, or the greater log likelihood, so the appropriate log likelihood difference condition is

$$\begin{aligned} &\log l\big(+1, \sigma_\varepsilon^2(+1) \mid R, X\big) - \log l\big(-1, \sigma_\varepsilon^2(-1) \mid R, X\big) \\ &\quad = \sum_{i=1}^N R_i X_i \equiv C_{R,X} > 0 \end{aligned} \tag{6}$$

in which the middle term is the covariance of R, X segments, assuming a chaotic wave of mean zero, as can always be arranged. Hence, $\hat{b}$, the maximum likelihood estimate of b, is given by

$$\hat{b} = \operatorname{sign}\{C_{R,X}\} \tag{7}$$

and is known in communication theory as a correlation decoder. From a statistical point of view, it is just an

estimate of a regression coefficient b which can only take one of the two values ± 1 according to the sign of the covariance, as previously surmised more generally. Note that its form does not depend on the type of spreading, and whether it is chaotic or not, but does assume knowledge of spreading. The accuracy of $\hat{b}$ will be considered presently, but as a maximum likelihood estimate, the correlation decoder is optimum in this statistical sense for this particular CSK system when there is no prior knowledge of b.

If it is known that the *proportion* of ± 1 transmitted bits is in the ratio $p : 1-p$, then a simple Bayesian regression estimate is obviously an improvement. This implies a class-room twist on the usual Bayesian regression scene in that the regression parameter can only take the values ± 1, and should have binary prior probabilities $(p, 1-p), 0 \le p \le 1$. With the conjugate gamma prior probability density for σ_ε^{-2}, say, with shape g and scale h and thus probability density function proportional to $x^g \exp(hx)$, the Bayesian decoder becomes in the coherent CSK case

$$\hat{b} = \operatorname{sign}\left\{\sum_{i=1}^{N}(R_i + X_i)^2 + 2h - \left(\frac{1-p}{p}\right)^{\frac{2}{N+2g+2}}\left(\sum_{i=1}^{N}(R_i - X_i)^2 + 2h\right)\right\}. \tag{8}$$

This simplifies with the noninformative prior $g = h = 0$. The posterior distribution for σ_ε^2 and its mean or median will be somewhat complicated. In communications practice, an equal proportion of ± 1's are usually transmitted, and then the Bayesian and non-Bayesian decoders are identical.

Bit error rate (BER) is the most important measure of communication performance for binary systems and statistically is the average probability of bit error, and for CSK systems is over spreading segments; there are others, such as outage rate, Lawrance and Ohama (2005). For BER of chaos-based systems, there are early approximate engineering results, thought to be exact, in several key papers, such as Kolumbán (2000) and Abel, Schwarz and Gotz (2000). They were carried over from other communication engineering systems such as *binary phase shift keying* (BPSK), Proakis (2001). The emphasis here is on exact BER results for chaos shift-keying systems and their use in providing engineering insights and structured approximations.

The BER of $\hat{b}$ from (7) needs to be considered conditional on the binary bit value transmitted and then averaged over their relative proportions to get the overall rate. Fortunately, from symmetry considerations, the conditional BERs are equal in most systems, including coherent CSK, and so it is sufficient to consider BER in

$$\begin{aligned}\mathrm{BER} &= P(\hat{b} = -1 \mid b = 1) = P(\hat{b} = 1 \mid b = -1)\\ &= P(C_{R,X} < 0 \mid b = 1)\\ &= P(C_{R,X} > 0 \mid b = -1),\end{aligned} \tag{9}$$

where the third and fourth equalities are from (6). With the received message R given by (1), the covariance in (9) when $b = 1$ is given by

$$C_{R,X} \equiv \sum_{i=1}^{N} X_i R_i = \sum_{i=1}^{N} X_i \varepsilon_i + \sum_{i=1}^{N} X_i^2. \tag{10}$$

Hence, the probability of a bit error by the correlation decoder, conditional on X, can be expressed as

$$\begin{aligned}\mathrm{BER} &= P(C_{R,X} < 0 \mid b = 1)\\ &= P\left(\sum_{i=1}^{N} X_i \varepsilon_i + \sum_{i=1}^{N} X_i^2 < 0\right).\end{aligned} \tag{11}$$

This can be evaluated exactly since the distribution of the linear combination of Gaussian noise terms is Gaussian with mean zero and variance $\sigma_\varepsilon^2 \sum_{i=1}^{N} X_i^2$. Then a simple Gaussian distribution calculation leads to the preliminary conditional result

$$\mathrm{BER}(X) = \Phi\{-(X^T X/\sigma_\varepsilon^2)^{1/2}\}, \tag{12}$$

where $X^T X$ represents an important engineering quantity called *bit energy*, the sum-of-squares of the spreading segment values, and $\Phi(\cdot)$ is the cumulative distribution function of a standardized Gaussian variable. The distribution of the probability of bit error from (12) has been pursued in Lawrance and Ohama (2005) in connection with *outage*, another measure of performance. Unconditionally, over the joint invariant distribution of the spreading sequences, the final exact result becomes

$$\mathrm{BER} = E_X \Phi\{-[(X^T X/N\sigma_X^2)\mathrm{SNR}]^{1/2}\}, \tag{13}$$

where $\mathrm{SNR} = N\sigma_X^2/\sigma_\varepsilon^2$ is the fundamental communications transmission measure, the *per bit signal-to-noise ratio*; the first term in (13) will be called the *standardized bit energy*. The result (13) indicates that BER is in the range $(0, 0.5)$; it is also clear that the expectation over X could be replaced by that over the distribution of bit energy.

The general definition of SNR in communications literature is *bit-energy-to-noise-power-spectral-density*

ratio for continuous signals, for example, Lau and Tse (2003), page 39, but the given form is more easily applicable in the present discrete base-band model.

Of considerable communication interest is the application of Jensen's inequality to (13), giving the lower bound result as

$$\text{BER} \geq \Phi\{-(\text{SNR})^{1/2}\}. \tag{14}$$

This is the result which was originally assumed to be exact. In the CSK setting, it can be reached by extensive spreading, $N \to \infty$. A key statistical observation from (13) and (14) is that bit error rate approaches its lower bound as bit energy reduces in variability, a desirable communication property. Thus, where possible, chaotic sequences should be designed with this in mind, as to be discussed in Section 3. The CSK lower bound is the exact result for the conventional binary phase shift keying (BPSK). This has led to some engineering views that CSK cannot improve on BPSK, but it is better in some respects, such as the steganographic security of its apparently noisy transmissions.

The independence of the Gaussian noise result in (13) can be relaxed to autocorrelated Gaussian channel noise and the generalized EGT result for the correlation decoder is

$$\text{BER} = E_X \Phi\left\{-\left[\left(X^T X/N\sigma_X^2\right)\text{SNR} \Big/ \left(1 + 2\sum_{k=1}^{N-1}\left(1 - \frac{i}{N}\right)\hat{\rho}_X(k)\rho_\varepsilon(k)\right)\right]^{1/2}\right\}, \tag{15}$$

where $\rho_\varepsilon(i)$ is the ith autocorrelation of the noise and $\hat{\rho}_X(k)$ is the kth within-segment empirical autocorrelation

$$\hat{\rho}_X(k) = (N-k)^{-1}\sum_{i=1}^{N-k} X_i X_{i+i} \Big/ N^{-1}\sum_{i=1}^{N} X_i^2 \tag{16}$$

for the zero-mean segment intervals. The within-segment autocorrelations could be replaced by the theoretical ones as a simplifying approximation. The effect of noise autocorrelation is seen to depend on the signs of the autocorrelations of the within-segment spreading and those of the noise. If noise and spreading are oppositely correlated, the square-rooted term is less than one, thus reducing BER toward zero; otherwise, the square-rooted term is greater than one, thus increasing BER toward one-half. There are similar conclusions for the lower bound, also only available from the exact statistical approach.

The calculation of BER via (13) for coherent CSK is seen to involve an expectation over a segment of the spreading sequence. This might seem to require a daunting N-dimensional integral but in the case of chaotic map spreading it is one-dimensional since all spreading values are iterations of the map from an initial random variable X_1. With the chaotic map being denoted by $\tau(x)$ and its iterates by $\tau^{(i)}(x)$, the BER result (13) becomes

$$\text{BER} = E_{X_1}\Phi\left\{-\left[\left(\sum_{i=1}^{N}\tau^{(i-1)}(X_1)^2 \Big/ N\sigma_X^2\right)\cdot \text{SNR}\right]^{1/2}\right\} \tag{17}$$

which is equivalently the one-dimensional integral

$$\int_{x=-c}^{c}\left\{\Phi\left(-\left[\left(\sum_{i=1}^{N}\tau^{(i-1)}(x)^2 \Big/ N\sigma_X^2\right)\cdot \text{SNR}\right]^{1/2}\right)\right\}\varphi(x)\,dx, \tag{18}$$

where $(-c, +c)$ is the range of the map and $\varphi(x)$ is the p.d.f. of its natural invariant distribution. The integral can be numerically evaluated for maps with explicit iterated forms, a key point. Incidentally, note that the chaotic nature had not been leveraged previously since only the random variable stationary aspect of spreading has been assumed.

For logistic spreading, the form taken is

$$\begin{aligned}\tau^{(i)}(x) &= x, \quad i = 1\\ &= \cosh\left(2^{i-1}\operatorname{arccosh}(x)\right),\\ &\qquad -1 \leq x < 1, i = 2, 3, \ldots.\end{aligned} \tag{19}$$

A computational disadvantage of using the logistic map is that for extensive spreading, the function to be integrated in (18) is a polynomial with many turning points and requires delicate evaluation, soon becoming impractical and needing approximation. However, for the Bernoulli-shift map

$$\begin{aligned}\tau^{(i)}(x) &= x, \quad i = 1\\ &= 2^{i-1}x - 1 - 2k, \quad \frac{k}{2^{i-2}} \leq x < \frac{k+1}{2^{i-2}},\\ &\qquad k = -2^{i-2}, \ldots, 2^{i-2} - 1, i = 2, 3, \ldots,\end{aligned} \tag{20}$$

the piece-wise linearity is a computational advantage, although as will be seen in Figure 2, is not always as effective in spreading.

Figure 2 gives some illustrations of the BER results (13) and (14) using (19) and (20) for logistic and Bernoulli-shift map spreading; the scales are

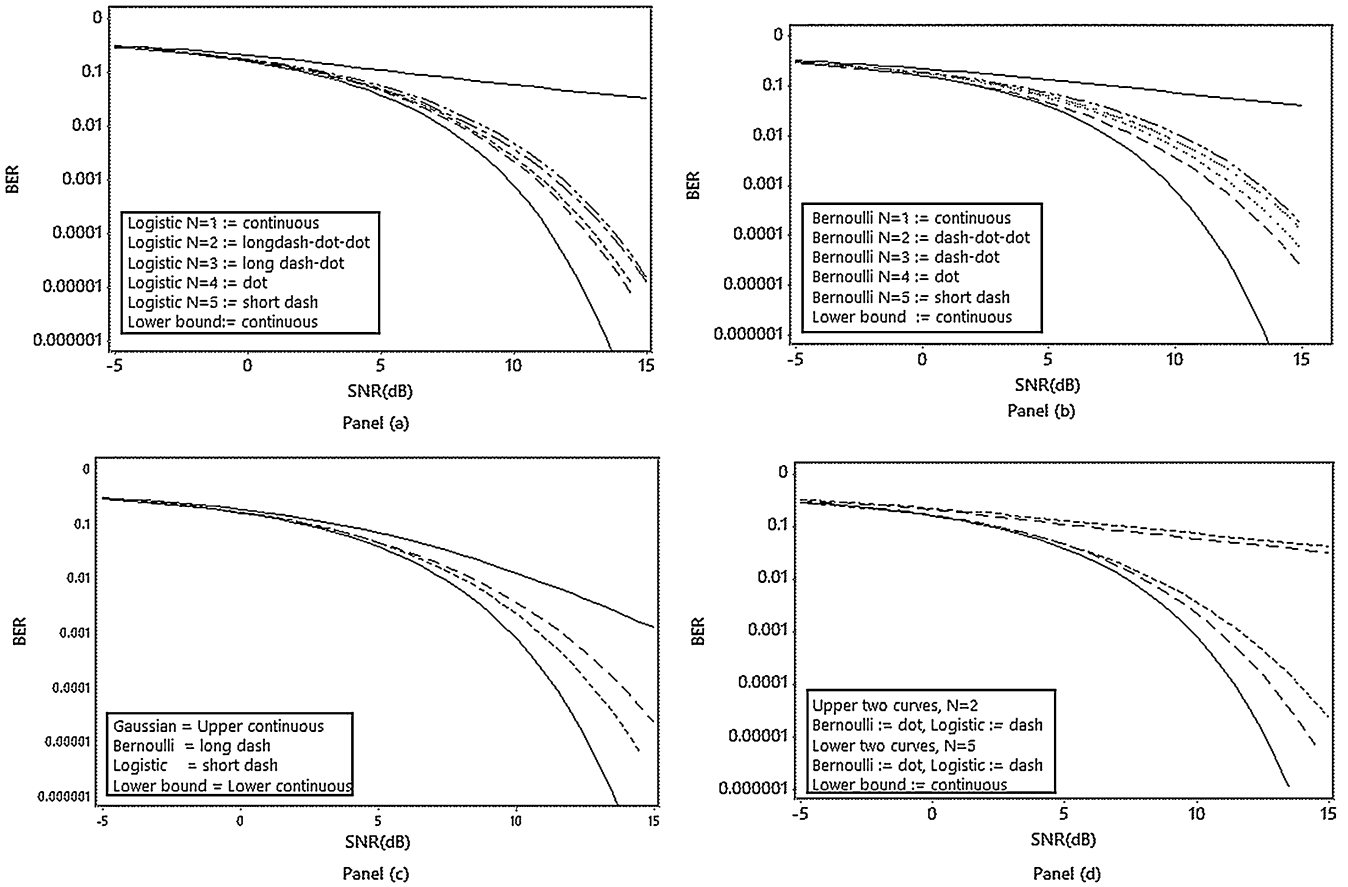


FIG. 2. *BER plotted against* $\mathrm{SNR_{db}}$ *for antipodal CSK*; *order of curves from panel variables*; *solid lowest curve in each panel is the BER Jensen lower bound. Panel* (a): *logistic map spreading*, *spreading factors* $N = 1, 2, \ldots, 5$. *Panel* (b): *Bernoulli-shift map spreading*, *spreading factors* $N = 1, 2, \ldots, 5$. *Panel* (c): *Bernoulli-shift*, *logistic and Gaussian spreading*, *spreading factor*$N = 5$. *Panel* (d): *Bernoulli-shift and logistic spreading*, *spreading factor* $N = 2$ *and Bernoulli-shift and logistic spreading*, *spreading factor* $N = 5$.

logarithmic for BER and decibel for SNR, $\mathrm{SNR_{db}} = 10\log_{10}(\mathrm{SNR})$. The main conclusions are, from Panels (a) and (b), the overall superiority of logistic spreading to Bernoulli-shift spreading. Panel (c) illustrates that independent Gaussian spreading is inferior to chaotic Bernoulli-shift spreading and logistic spreading. Panel (d) shows the advantage of increasing the spreading factor in both of the Bernoulli-shift and logistic cases with the latter being superior although not very close to the lower bound by $N = 5$ for $\mathrm{SNR_{db}} > 10$. Also shown by the top curves of Panels (a), (b) and (c) are the upper bounds of no spreading, $N = 1$. It is conjectured that independent spreading is generally inferior to negatively dependent spreading and superior to positively dependent spreading.

There are many approximate results in the chaos communication literature for the BERs of correlation decoders, mostly following the early approaches in which results were transferred from other systems or obtained by simple Gaussian approximation (SGA), such as to $C_{R,X}$ in (10). By comparisons with the exact results, they have been found inaccurate for moderate N and small BERs, as will be illustrated in Figure 2. The inaccuracy comes from approximating the lower tail of the distribution of $C_{R,X}$, which is skewed, by a Gaussian distribution with the mean and variance of $C_{R,X}$. Nevertheless, the results can be structured to identify influential terms. The SGA approach begins with the equalities

$$\begin{aligned}\mathrm{BER} &= P\{\hat{b} = -1 \mid b = 1\} = P\{C_{R,X} < 0 \mid b = 1\} \\ &= P\left\{\frac{C_{R,X} - E(C_{R,X})}{[\mathrm{var}(C_{R,X})]^{1/2}} < \frac{-E(C_{R,X})}{[\mathrm{var}(C_{R,X})]^{1/2}}\right\},\end{aligned} \tag{21}$$

and making the standardized Gaussian assumption for the first term of the last equality gives

$$\begin{aligned}\mathrm{BER_{sga}} &= P\{C_{R,X} < 0 \mid b = 1\} \\ &\cong \Phi\left\{-\frac{E(C_{R,X})}{[\mathrm{var}(C_{R,X})]^{1/2}}\right\},\end{aligned} \tag{22}$$

with $b = 1$ being assumed in the expectation and variance terms. After some calculation, there is the explicit SGA result

$$\text{BER}_{\text{sga}} \simeq \Phi\left\{-\left[\frac{1}{\text{SNR}} + N^{-1}\frac{\sigma_{X^2}^2}{\sigma_X^4} \cdot \left\{1 + 2\sum_{k=1}^{N-1}\left(1 - \frac{k}{N}\right)\rho_{X^2}(k)\right\}\right]^{-\frac{1}{2}}\right\} \tag{23}$$

as first derived in Lawrance and Balakrishna (2001). Here, $\sigma_{X^2}^2$ is the variance of X_t^2, $\rho_{X^2}(k)$ is the within-segment lag-k quadratic autocorrelation of (X_i, X_{i+k}), that is the lag-k linear autocorrelation of (X_t^2, X_{t+k}^2), and the variances ratio is the kurtosis of X. The approximation agrees with the exact result (13) as $\text{SNR} \to 0$ and also when $N \to \infty$. The second term indicates the influential roles of the kurtosis of the spreading and its quadratic autocorrelations, anther credit to the exact result. Positivity of the quadratic autocorrelation sum is seen to increase BER, as does large kurtosis of the spreading. Negativity of the sum correspondingly reduces BER. Notice that the result does not depend on the actual type of the spreading, only on its marginal and correlational properties, and thus emphasises a qualitative deficiency in the approximation. A variety of such approximations, although less structured, have been reported in Lau and Tse (2003) and Tam, Lau and Tse (2007). Figure 3 illustrates the approximation. The upper set of curves are from the SGA result (23) calculated for $N = 1, 2, \ldots, 6$ while the separated lower set, not including the lower bound curve (14), are from the corresponding exact result (13). It is evident that

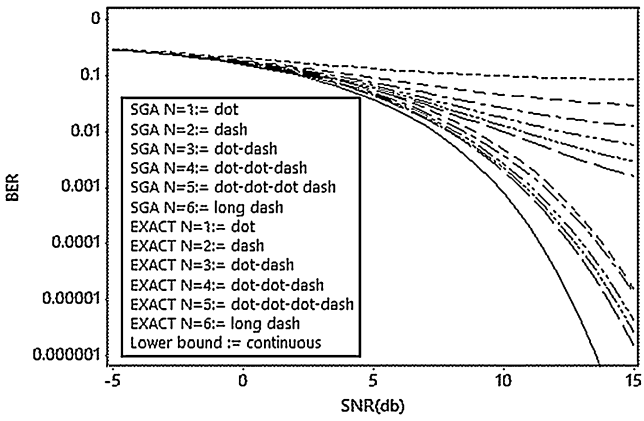


FIG. 3. *Comparison of using the exact BER result for the lower set of curves with using the SGA approach for the upper set of curves, for coherent CSK with logistic map spreading, spreading factors $N = 2, 3, \ldots, 6$, and the lower bound.*

the SGA result gives curves which are much different to the exact ones.

For qualitative information about the effect of autocorrelation in the channel noise, the SGA approximation to the exact BER result (15) can be obtained by reworking of (23) as

$$\text{BER}_{\text{sga}} \simeq \Phi\left\{-\left[\frac{1}{\text{SNR}}\left\{1 + \sum_{k=1}^{N}\left(1 - \frac{k}{N}\right)\rho_X(k)\rho_\varepsilon(k)\right\} + N^{-1}\frac{\sigma_{X^2}^2}{\sigma_X^4}\left\{1 + 2\sum_{k=1}^{N-1}\left(1 - \frac{k}{N}\right)\rho_{X^2}(k)\right\}\right]^{-\frac{1}{2}}\right\}. \tag{24}$$

The first term of the sum is seen as an approximation in the exact result (15), (16) and the second is found in the uncorrelated noise SGA expression (23).

These results have all been for coherent chaos shift keying, but more practically, spreading sequences will have to be transmitted to the receiver or synchronized at the receiver, in order to decode the message. Then the coherent case is best-case modelling. The necessary bit decoding and error rate theory for non-coherent modelling will be covered in Section 5 after considerations of design.

3. DESIGN OF CHAOS COMMUNICATION SYSTEMS

This is not a topic which has received much attention in the engineering literature and draws strength from the exact statistical theory approach to bit error. Homer et al. (2004) discussed in a somewhat ad hoc way the choice among piece-wise linear chaotic map generators. More conceptually, Yao (2004) noted from the exact result (13), that there is a condition for determining the type of spreading which brings BER down to its lower bound (14); this is when the bit energy sum-of-squares term in (13) is equal to its expectation $N\sigma_X^2$. Actually, the condition is an exact impossibility if the spreading segments are to be chaotic but it does suggest a strategy toward this aim. Thus, first consider the variance of the bit energy sum-of-squares

$$\text{var}\left(\sum_{i=1}^{N} X_i^2\right) = N\left\{1 + 2\sum_{k=1}^{N-1}\left(1 - \frac{k}{N}\right)\rho_{X^2}(k)\right\}\sigma_{X^2}^2 \tag{25}$$

and require it to be zero, so X_i^2 is equal to its expectation σ_X^2, by having

$$1 + 2\sum_{k=1}^{N-1}\left(1 - \frac{k}{N}\right)\rho_{X^2}(k) = 0. \tag{26}$$

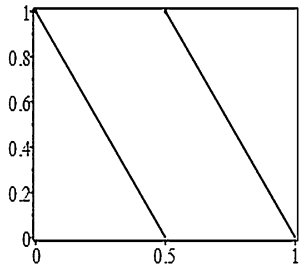
This is too intractable to take forward in general and exactly, but it is possible to consider the case $N = 2$ for adjacent pairs and this yields

$$1 + \rho_{X^2}(1) = 0, \tag{27}$$

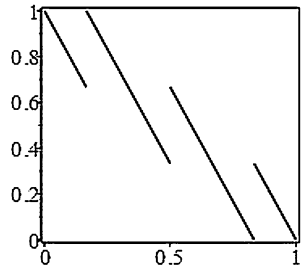
suggesting that (X_i^2, X_{i+1}^2) be linearly related with slope -1 and thus that (X_i, X_{i+1}) lies on a circle. With N even, constant spreading energy per bit is then assured. However, a circle is not a chaotic map, and thus a chaotic circular map approximately resembling a circle is required. First, working in terms of the linearly related squared variables, a map with slope -1 is required but this cannot be chaotic; an approximate chaotic form can be visualized as set of $2m$ sublines, say, as illustrated in Figure 4, Panels (a) and (b). The simplest case of two branches, the Bernoulli map, translates into the 1st-order circular map of Figure 4, Panel (c) given by

$$\begin{aligned}\tau(x) = \{&-\sqrt{-2x^2+2}, -1 \le x < -1/\sqrt{2};\\ &\sqrt{-2x^2+1}, -1/\sqrt{2} \le x < 1/\sqrt{2};\\ &-\sqrt{-2x^2+2}, 1/\sqrt{2} \le x < 1\},\end{aligned} \tag{28}$$

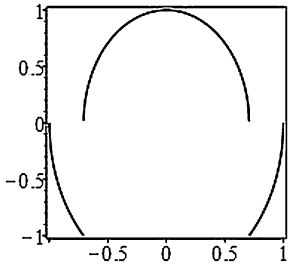
and the case of 4 branches translates into the 2nd-order circular map of Figure 4, Panel (d). More generally, the class of Yao's circular maps is produced which more and more resemble a circle as the number of branches increase. Circular maps such as (28) have a natural invariant distribution with v-shaped p.d.f. $|x|$;



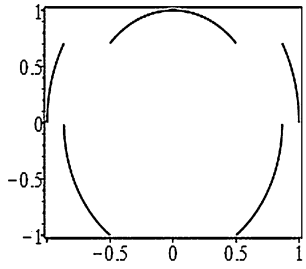
(a) The broken line map of 1st-order



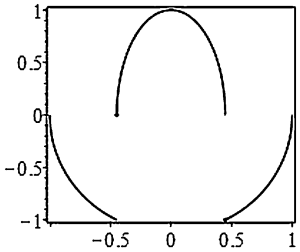
(b) The broken line map of 2nd-order



(c) The circular map of 1st-order



(d) The circular map of 2nd-order



(e) The deformed circular map of 1st-order for r =0.2

FIG. 4. *Circular maps involved with optimal spreading in coherent CSK theory.*

this map has $\rho_{X^2}(1) = -1/2$, so only about halfway to the Frechet lower bound of -0.968. With m linear branches, the map has the same invariant distribution, but now has $\rho_{X^2}(1) = -1 + 1/2(m-1)^2$ which does approach the lower bound -1 as the number of branches increases.

Papamarkou and Lawrance (2007) improved on Yao's map (28) by transforming a non-central Bernoulli map with branch division at r producing a deformed circular map as illustrated in Figure 4, Panel (e) for $r = 0.2$. The minimum lag 1 quadratic autocorrelation is now -0.722, when $r = 0.42$, a considerable improvement on -0.5. The invariant distribution associated with the deformed circular map now has a distribution with p.d.f. $\{-2(1-r)x, -1 \le x \le 0; 2rx, 0 < x < 1\}$. The culmination of this work is that of paired Bernoulli circular spreading, Papamarkou and Lawrance (2013), in which a type of random circular map is developed which actually attains the lower bound (14), admittedly at the cost of some complication and which is not chaotic.

This brief analysis has, hopefully, indicated that a start from the statistically exact bit error result (13) enables theoretical and design insight about CSK. The family of chaotic circular and deformed circular maps bring the BER within close proximity to the theoretical lower bound, without actually reaching it. They are likely to be more useful in practice than theoretically since their explicit convolutions seem intractable.

4. MODELLING INTERFERENCE AND MULTI-PATH FADING CHANNELS

The performance of a communication system can be degraded by external circumstances, such as innocent interference, intentional jamming, multi-path fading and by multi-user activity. The present concern will be limited to the statistical modelling of the second and third of these. Intentional jamming can take many forms, and to illustrate, the effect of *wide band on-off pulsed additive jamming* signals on the BER of the correlation decoder will be considered. The presentation is based on reformulating work in Lau and Tse (2003) to introduce exact and informative structural results which reveal those features of the system controlling susceptibility to jamming. Several new descriptive terms are introduced which emphasize the main aspects.

The on-off jamming signal for a typical spreading segment will be represented by $U_i, i = 1, 2, \ldots, N$ and is modelled as

$$U_i = \{0, i = 0, A_i, i = 1, 2, \ldots, M; \quad 0, i = M+1, \ldots, N\}, \tag{29}$$

where $M, 0 \le M \le N$ is the *jamming number* and $A_i, 1 \le A_i \le M$ give *the jamming strengths* which are modelled as independent Gaussian $N(0, \sigma_a^2)$. The received but jammed message segment is now

$$R_i = bX_i + U_i + \varepsilon_i, \quad i = 1, 2, \ldots, N, \tag{30}$$

where the interference effectively increases the message noise. The covariance of the correlation decoder as in (10) now becomes

$$\begin{aligned} C_{R,X} &= \sum_{i=1}^{N} R_i X_i \\ &= b\sum_{i=1}^{N} X_i^2 + M\sum_{i=1}^{M} X_i(A_i + \varepsilon_i) + \sum_{i=M+1}^{N} X_i \varepsilon_i. \end{aligned} \tag{31}$$

Conditionally on the spreading values $X_1, X_2, \ldots, X_N$ and Gaussian noise, $C_{R,X}$ is a linear combination of independent Gaussian variables, thus with a Gaussian distribution which has mean and variance

$$b\sum_{i=1}^{N} X_i^2, \quad (\sigma_a^2 + \sigma_\varepsilon^2)\sum_{i=1}^{M} X_i^2 + \sigma_\varepsilon^2 \sum_{i=M+1}^{N} X_i^2, \tag{32}$$

respectively. A calculation paralleling (11) and some creative formula display then give the exact BER of the correlation decoder as

$$\mathrm{BER} = E_X \Phi\left\{-\left[((X^T X)_N / N\sigma_X^2)\mathrm{SNR} \Big/ \left(1 + \mathrm{JF}_{M,N}\frac{M^{-1}(X^T X)_M}{N^{-1}(X^T X)_N}\right)\right]^{1/2}\right\}, \tag{33}$$

where $(X^T X)_i$ is the sum of squares of the first i spreading values and $\mathrm{JF}_{M,N} = M\sigma_a^2 / N\sigma_\varepsilon^2$ is termed the *jamming factor*. The SNR term in (33) gives the no-jamming result (13). For increasingly large spreading, but a constant *jamming-spread fraction* (JSF) $\gamma = M/N$, the limiting form of (33) is seen as

$$\mathrm{BER} \to \Phi\{-[\mathrm{SNR}/(1 + \mathrm{JF})]^{1/2}\}, \tag{34}$$

where $\mathrm{JF} = \gamma\{\sigma_a^2/\sigma_\varepsilon^2\}$. Thus, increases in the jamming factor bring BER closer to its maximum value of one-half. Figure 5 gives illustrations of jamming in respect of the exact and limiting results with $M = 2, N = 4$ and several jamming factors JF. Panel (a) shows the exact results for jamming factors $\mathrm{JF} = 0, 1, 3, 5$ with

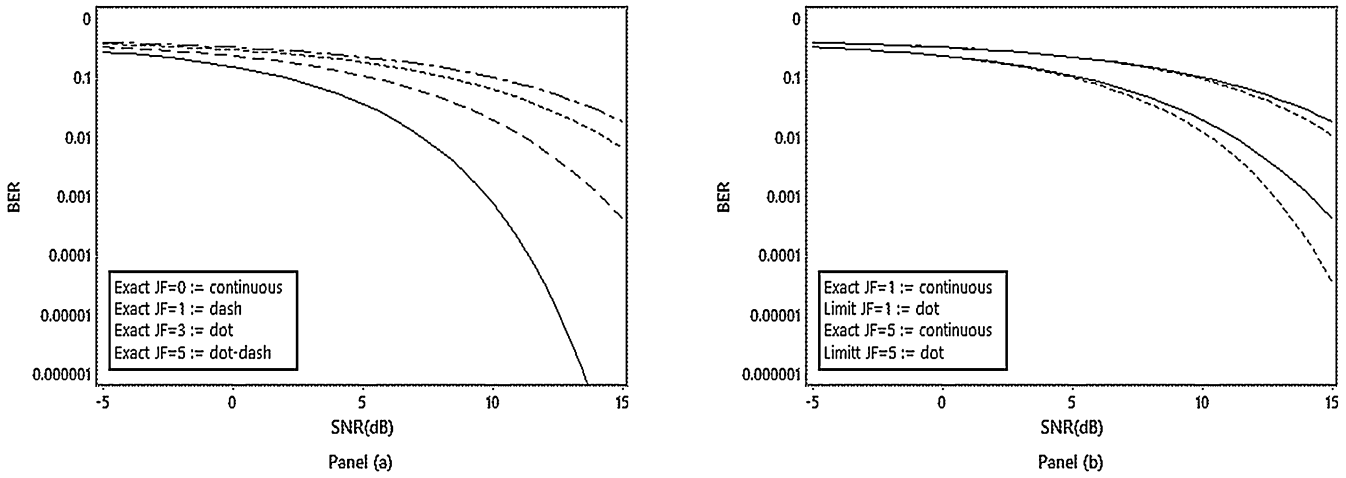


FIG. 5. *Wide band on-off pulsed jamming of coherent CSK with logistic spreading, spreading factor* $N = 4$ *and jamming number* $M = 2$. *Panel* (a): *Exact BER curves for the coherent lower bound and jamming factors* $\mathrm{JF} = 0, 1, 3, 5$. *Panel* (b): *Comparison of limiting and exact BER curves for jamming factors* $\mathrm{JF} = 1, 5$.

$\mathrm{JF} = 0$ indicating no jamming, and emphasizes that jamming is much more effective at high SNR values. Panel (b) compares the exact and limiting curves for $\mathrm{JF} = 1, 5$ and suggests that the limiting result (34) may be a lower bound, but this has not been verified, and that it is a rather optimistic view of BER for large SNR. The results here thus enable engineering judgements about the effect of jamming, and suggest the increased spreading factors and SNR levels needed to overcome unwanted effects on BER.

Multi-path fading is a type of unintentional and often unavoidable transmission disturbance, for instance, the reflecting of signals off buildings. Concern in the literature related to the effect on the BER of chaos-based systems is represented by Zhou, Wang and Ye (2010), Kaddoum et al. (2010) and Kaddoum and Gagnon (2013a). Fading is modelled by the transmitted message segment being multiplicatively distorted by a random variable fading factor, $V (V \geq 0)$, effecting the transmitted message so the received signal according to (1) is now of the form

$$R_i = V_i b X_i + \varepsilon_i, \quad i = 1, 2, \ldots, N. \tag{35}$$

The distribution of V depends on the type of fading. For multi-path scattering from clusters of reflected waves, the Nakagami distribution is often used, with its p.d.f.

$$f_V(v) = \frac{2}{\Gamma(m)} \left(\frac{m}{\varphi}\right)^m v^{2m-1} \exp\left(-m V^2/\varphi\right), \quad v \geq 0, \tag{36}$$

where $\varphi = E(V^2), m = \varphi^2/\operatorname{var}(V^2)$. For other types of fading, Raleigh and Rice distributions are considered more appropriate.

5. NON-COHERENT CHAOS SHIFT-KEYING COMMUNICATION

The difficulty in some circumstances of exactly knowing the reference segments at the receiver in coherent chaos-based systems focussed engineering investigation of non-coherent antipodal CSK; with these the reference sequence has to be either transmitted to the receiver or available there by imperfect synchronization. An empirical illustration of what can be achieved practically by synchronization of two lasers is given in Figure 6. This is taken from the output of an experimental laser-based antipodal CSK system analysed in Lawrance, Papamarkou and Uchida (2017) where the exact laser wave is exceptionally available for comparison with the synchronized version. The tracking between the two waves is seen to be close although not perfect.

There have been several versions of non-coherent CSK systems in which the reference segments are transmitted, with *differential chaos shift-keying* (DCSK) being the most popular. The statistical modelling theory has been treated in a minimal way using correlation decoders and their bit error results have been approximate. However, correlation decoders can be justified by extending the likelihood (3) to include terms from the reference segments equation (2) and then obtaining a partial likelihood. Here, the emphasis is on the performance of the correlation decoder

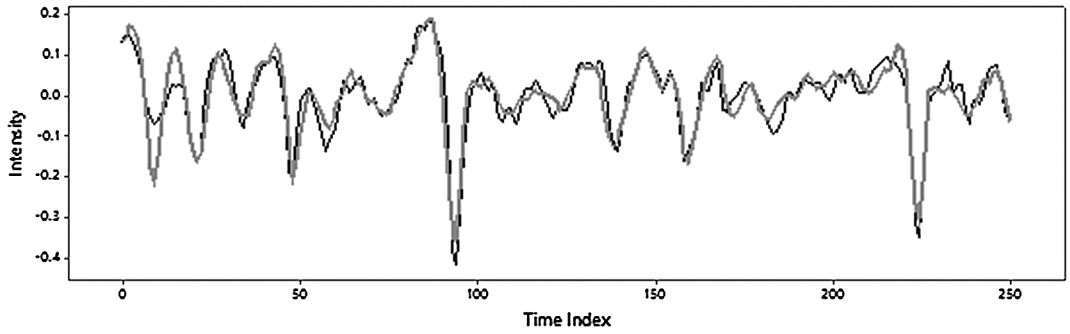


FIG. 6. *Comparison at* 250 *time points of a laser generated chaotic wave* (*thin black line*) *and its remotely synchronized version* (*thick grey line*) *from an experimental CSK communication system.*

via exact theory of its BER; both EGT and SGA approaches are used by expanding on Lawrance and Ohama (2003).

Following its introduction in Section 2, the *spreading segment* $(X_1, X_2, \ldots, X_N)$ is modelled at the receiver as the *reference segment*

$$Y_i = X_i + \eta_i, \quad \operatorname{var}(\eta_i) = \sigma_\eta^2, i = 1, 2, \ldots, N. \tag{37}$$

Similarly, from Section 1, the *message segment* containing the binary bit b is modelled at the receiver as

$$R_i = bX_i + \varepsilon_i, \quad \operatorname{var}(\varepsilon_i) = \sigma_\varepsilon^2, i = 1, 2, \ldots, N. \tag{38}$$

In this formulation there are now two signal-noise ratios,

$$\text{SNR} = \frac{N\sigma_X^2}{\sigma_\varepsilon^2}, \quad \text{SER} = \frac{N\sigma_X^2}{\sigma_\eta^2}, \tag{39}$$

where SNR is the *signal-to-noise ratio* used previously, and SER is the newly termed *spreading-to-error ratio*.

The correlation decoder takes the usual covariance form

$$C_{R,Y} = \sum_{i=1}^{N} R_i Y_i \tag{40}$$

and substituting (37) and (38) into (40) gives

$$\begin{aligned} \text{BER} &= P(C_{R,Y} < 0 \mid b = 1) \\ &= P\left\{\sum_{i=1}^{N}(X_i + \varepsilon_i)(X_i + \eta_i) < 0 \,\middle|\, b = 1\right\}. \end{aligned} \tag{41}$$

The probability in (41) is calculated as in Lawrance and Ohama (2003) under the tractable assumptions of independent Gaussian message noise and segment or synchronization errors, but assumes different variances. It will be mathematically convenient to set $\varepsilon_i = \sigma_\varepsilon u_i, \eta_i = \sigma_\eta v_i$ where $\{u_i\}$ and $\{v_i\}$ are independent standardized Gaussian sequences, and then

$$\begin{aligned} \text{BER} &= P\left\{\sum_{i=1}^{N}(X_i + \sigma_\varepsilon u_i)(X_i + \sigma_\eta v_i) \le 0\right\} \\ &= P\left\{\sum_{i=1}^{N}(X_i/\sigma_\varepsilon + u_i)(X_i/\sigma_\eta + v_i) \le 0\right\}. \end{aligned} \tag{42}$$

This probability will be obtained in terms of the statistical non-central F-distribution. First, define the following variables:

$$z_{1i} = \frac{1}{\sqrt{2}}(u_i + v_i), \quad z_{2i} = \frac{1}{\sqrt{2}}(u_i - v_i) \tag{43}$$

which are also independent Gaussian and give u_i and v_i as

$$u_i = \frac{1}{\sqrt{2}}(z_{1i} + z_{2i}), \quad v_i = \frac{1}{\sqrt{2}}(z_{1i} - z_{2i}). \tag{44}$$

These allow the previous BER expression to be written

$$\begin{aligned} \text{BER} = P\Bigg\{&\sum_{i=1}^{N}\left[(X_i/\sigma_\varepsilon) + \frac{1}{\sqrt{2}}(z_{1i} + z_{2i})\right] \\ &\cdot\left[(X_i/\sigma_\eta) + \frac{1}{\sqrt{2}}(z_{1i} - z_{2i})\right] \le 0\Bigg\}. \end{aligned} \tag{45}$$

The next step is to multiply out the inner terms and separate them into groups involving either z_{1i} or z_{2i} and then complete the squares in these variables, finding there are no resulting $z_{1i}z_{2i}$ product terms. These operations lead to the key result

$$\begin{aligned} \text{BER} = P\Bigg\{&\sum_{i=1}^{N}\left[z_{1i} + \frac{1}{\sqrt{2}}(\sigma_\varepsilon^{-1} + \sigma_\eta^{-1})X_i\right]^2 \\ &- \sum_{i=1}^{N}\left[z_{2i} + \frac{1}{\sqrt{2}}(\sigma_\varepsilon^{-1} - \sigma_\eta^{-1})X_i\right]^2 \le 0\Bigg\} \end{aligned} \tag{46}$$

which shows that the sum and difference of the inverse standard deviations of the optical noise and synchronization error are the important quantities. Writing the previous expression as

$$\text{BER} = P\left\{\frac{\sum_{i=1}^{N}[z_{1i} + \frac{1}{\sqrt{2}}(\sigma_\varepsilon^{-1} + \sigma_\eta^{-1})X_i]^2}{\sum_{i=1}^{N}[z_{2i} + \frac{1}{\sqrt{2}}(\sigma_\varepsilon^{-1} - \sigma_\eta^{-1})X_i]^2} \le 1\right\} \tag{47}$$

is a key statistical step for computation, and actually this result does not depend on Gaussian distribution assumptions. With Gaussian noise and error assumptions, and conditional on spreading X, the ratio term is a standard non-central F variable $F_{N,N}(\nu_{1,X}, \nu_{2,X})$ with degrees of freedom (N, N) and non-centrality parameters

$$\begin{aligned}(\nu_{1,X}, \nu_{2,X}) &\equiv \left\{\frac{1}{2}(\sigma_\varepsilon^{-1} \pm \sigma_\eta^{-1})^2 X^T X\right\} \\ &= \left\{\frac{1}{2}\left(\frac{1}{\{\text{SNR}\}^{1/2}} \pm \frac{1}{\{\text{SER}\}^{1/2}}\right)^2 (X^T X/N\sigma_X^2)\right\}.\end{aligned} \tag{48}$$

The last equality shows how the communication quantities of bit energy, the signal-noise and spreading-error ratios determine bit error rate. Returning to the previous BER expression (47), it can now be written as an average over the spreading sequence as

$$\text{BER} = E_X\big[P\big\{F_{N,N}(\nu_{1,X}, \nu_{2,X}) \le 1\big\}\big]. \tag{49}$$

Notice via (48) that since this only involves spreading through the bit energy sum of squares, its expectation can be over the distribution of bit energy, a considerable simplification. With a particular chaotic map for the spreading, an exact numerical calculation from (49) is theoretically possible via a one-dimensional integral.

A tractable approximate result for any chaotic map and extensive spreading N comes from replacing the bit energy sum-of-squares by $N\sigma_X^2$ to now give the non-centrality parameters $(\nu_{1,X}, \nu_{2,X})$ as

$$(\nu_1, \nu_2) \equiv \frac{1}{2}\left(\frac{1}{\{\text{SNR}\}^{1/2}} \pm \frac{1}{\{\text{SER}\}^{1/2}}\right)^2. \tag{50}$$

If this approximation is used in (49), there is the approximate result

$$\text{BER} \cong P\big\{F_{N,N}(\nu_1, \nu_2) \le 1\big\}. \tag{51}$$

The limiting $N \to \infty$ behaviour of the doubly non-central variable $F_{N,N}(\nu_1, \nu_2)$ does not seem to be available to indicate whether this is a lower bound on BER as a function of SNR and SER. There are somewhat simpler versions of these results when the message noise and reference or synchronization errors have equal variances, as for the DCSK system.

Some further insight into the model is afforded by the SGA result which can be derived in terms of its component statistical features as

$$\begin{aligned}\text{BER}_{\text{sga}} &= \Phi\left\{-\left(\frac{1}{\text{SNR}}\left\{1 + \sum_{k=1}^{N}\left(1 - \frac{k}{N}\right)\rho_X(k)\rho_\varepsilon(k)\right\}\right.\right. \\ &\quad + \frac{1}{\text{SER}}\left\{1 + \sum_{k=1}^{N}\left(1 - \frac{k}{N}\right)\rho_X(k)\rho_\eta(k)\right\} \\ &\quad + \frac{N}{\text{SNR}\cdot\text{SER}}\left\{1 + \sum_{k=1}^{N}\left(1 - \frac{k}{N}\right)\rho_\varepsilon(k)\rho_\eta(k)\right\} \\ &\quad \left.\left. + N^{-1}\frac{\sigma_{X^2}^2}{\sigma_X^4}\left\{1 + 2\sum_{k=1}^{N-1}\left(1 - \frac{k}{N}\right)\rho_{X^2}(k)\right\}\right)^{-\frac{1}{2}}\right\}\end{aligned} \tag{52}$$

and covers the case when both message noise and reference or synchronization errors are autocorrelated. The structured form allows the effects of the important engineering quantities SNR and SER to be seen, as well as the statistical distributional and correlations quantities. Results of this type without the explicit quadratic autocorrelation terms are reported in Tam, Lau and Tse (2007) from a computational angle.

Figure 7 continues the initial illustration in Figure 6 of a laser-based experimental system; correlation decoding and bit error rate are empirically illustrated in Panels (a) and (b), respectively. Panel (a) gives a typical scatterplot of received message-carrying segments and their synchronized segments using a spreading factors of 3–20; it shows a strongly positive correlation (0.889), correctly decoding the transmitted +1 bit. Panel (b) gives the BER performance of the non-coherent system plotted against spreading factor, rather than SNR as is customary, since both optical noise and synchronization error were fixed in the experiment. Additionally, there is experimental knowledge of the spreading wave to enable a similar calculation for the coherent system. The spreading factor and the degree of synchronization were high enough for the bit error performance of the non-coherent system to be very near that of the coherent system, its lower bound.

Another distinguishing feature of non-coherent CSK is that for given values of SNR and SER there may be an optimum extent of spreading, intuitively due to

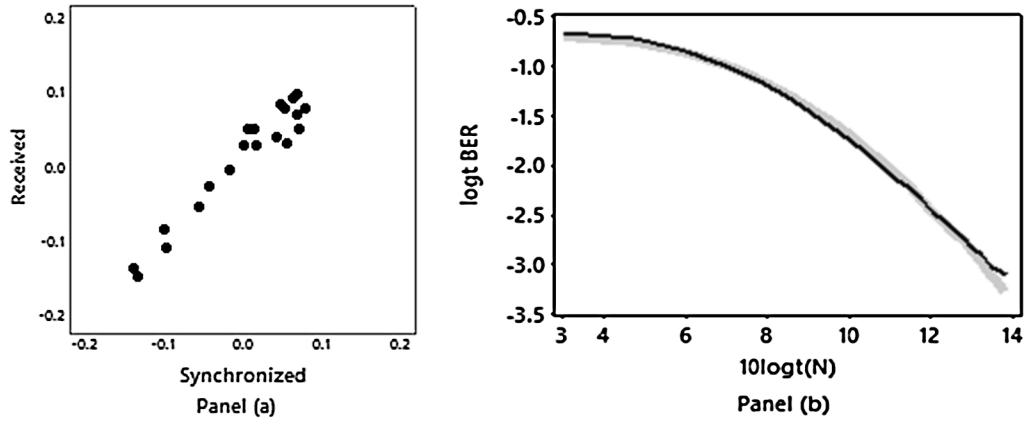


FIG. 7. *Panel* (a): *Scatterplot of a received message segment and its synchronized version for* $N = 20$ *spreading*. *Panel* (b): *BER curves of the experimental non-coherent shift-keying communication system* (*light curve*) *and of the corresponding coherent system* (*dark curve*) *in which there is no synchronization error.*

the balance of information in the message and spreading segments. This was first investigated by Sushchik, Tsimring and Volkovskii (2000). One case is when noise and error are each linearly uncorrelated and spreading is quadratically uncorrelated, as is so with logistic map spreading. Treating N as continuous, it can be shown using (52) that the spreading factor $\tilde{N}$ giving the minimum BER value and this value itself $\widetilde{\mathrm{BER}}$, are given by

$$\tilde{N} \simeq \kappa^{1/2}\sqrt{\mathrm{SNR}\cdot\mathrm{SER}},$$

$$\widetilde{\mathrm{BER}} = \Phi\left\{-\left[\mathrm{SNR}\Big/\left(1+\frac{\mathrm{SNR}}{\mathrm{SER}}+2\kappa_X\sqrt{\frac{\mathrm{SNR}}{\mathrm{SER}}}\right)\right]^{1/2}\right\}, \tag{53}$$

respectively, where $\kappa_X = \sigma_{X^2}^2/\sigma_X^4$. Note that BER is increased relative to the Jensen lower bound (14) of the coherent case, and moreover, that it is reached for very large SER. However, because (52) is an SGA approximation, $\tilde{N}$ may not accord with the exact result via (49).

6. MULTI-USER COHERENT CHAOS SHIFT-KEYING COMMUNICATION

This section treats antipodal chaos shift-keying communication with more than one pair of participants under the coherent assumption. The key new aspect for a user decoding a bit is that of interference from other users. As far as an individual user is concerned, the other users constitute additional non-independent and non-Gaussian noise, and not just simply the addition of more independent Gaussian noise. There has been relatively little published work in this area, apart from Tam et al. (2002), Tam, Lau and Tse (2003) employing correlation decoders and SGA. Development of corresponding EGT results were given in Tam et al. (2004) following Lawrance and Ohama (2003). The presentation here focusses on providing a likelihood-based decoder and its EGT, and initially develops from Lawrance and Yao (2008). Following Lau and Tse (2003), Tam, Lau and Tse (2007) also investigate correlation and other decoders in multi-user chaos shift-keying systems, although not the newer likelihood-based decoder; they present computational BER results, mainly using an SGA approach. The focus here continues to be on the EGT approach giving exact and structured results.

Suppose there are now L users ($L \geq 2$), and the system is concerned with the simultaneous transmission of single binary bit messages, $b = \pm 1$ by the lth user, $l = 1, 2, \ldots, L$, in a single time slot. The focus is on the so-called *active-user*, in the presence of so-called *other-users* who have interfering effects. A block diagram of the model is given in Figure 8. The modelling assumption is that each user l has their own chaotic generator which provides a spreading segment $X_l^T = (X_{l1}, \ldots, X_{lN})$; for simplicity, the same spreading factor N is taken by all users. As in the single-user coherent case, the spreading segment of an active-user is assumed to be known exactly by the intended receiver. To transmit a binary bit b_l, the lth user modulates the spreading segment X_l to become $b_l X_l$, a procedure which is followed by all users. Their modulated spreading segments are transmitted additively in the same time slot, and received through individual channels to each intended receiver, with the

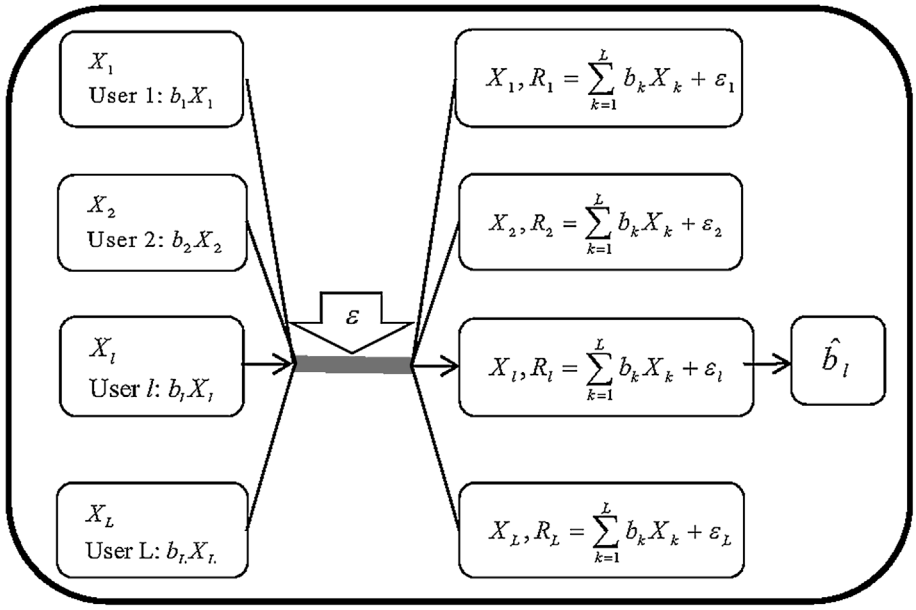


FIG. 8. *Block diagram of an antipodal multi-user coherent CSK communication system and transmission by the active lth user.*

lth receiver attracting individual IID-Gaussian channel noise $\varepsilon_l^T = (\varepsilon_{l1}, \varepsilon_{l2}, \ldots, \varepsilon_{lN})$. The lth receiver has to decode the lth message bit, ignoring interference from the others. From the active lth user transmitting bit b_l, the message segment arriving at the lth receiver $R_l^T = (R_{l1}, R_{l2}, \ldots, R_{lN})$ can be written

$$R_l = b_l X_l + \sum_{k \neq l=1}^{L} b_k X_k + \varepsilon_l. \tag{54}$$

The summation term represents interference by the message bits of other-users. Decoding of the message bit b_l in the received signal uses knowledge of the lth reference segment, but without any knowledge of the reference segments and binary bit messages of other-users; they are thus treated here as random variables, not as unknown statistical parameters to be estimated. This mixture of statistical inferential assumptions reflects practical realities.

This section is next concerned with developing a likelihood-based decoder for multi-user coherent antipodal CSK systems, following Lawrance and Yao (2008); although the exact result is not tractable enough for practical use, a multivariate Gaussian assumption yields an approximate maximum likelihood decoder which is an attractive generalization and an improvement on the correlation decoder. The multi-user system requires a theoretical analysis which significantly departs from that of the single-user system given in Sections 2 and 5. The approach demonstrates further benefits flowing from the use of statistical theory in communications engineering.

As in the single-user case, the message bit b_l is regarded as a parameter to be estimated using the received data R_l and known spreading segment X_l. The likelihood of b_l is thus formed from the joint probability density function of the other-message and noise terms of (54), acting as dependent noise. The optimum decoder is the value of b_l which maximises this other-user likelihood. In the derivation, the channel noise variance σ_ε^2 is assumed known; this is not a disadvantage in two important cases where it is absent from the decoder. To obtain the other-user likelihood, the receiver equation (54) first needs to be expressed in terms of the other-user terms as

$$\begin{gathered} R_l = b_l X_l + R_l' \\ \text{where } R_l' = \sum_{k \neq l=1}^{L} \tilde{b}_k X_k + \varepsilon_l, i = 1, 2, \ldots, N. \end{gathered} \tag{55}$$

Then $R_l' = R_l - b_l X_l$ and the required likelihood of b_l based on R_l' follows from (55) as the joint probability density function

$$f_{R_l'}(R_l \mp X_l \mid b_l = \pm 1, X_l). \tag{56}$$

The likelihood decoder of b_l is given by whichever of $b_l = \pm 1$ makes (56) the largest, and thus can be written

$$\hat{b}_l = \operatorname{sign}\{\log(f_{R_l'}(R_l - X_l)/f_{R_l'}(R_l + X_l))\}, \tag{57}$$

where $f_{R_l'}(\cdot)$ is the joint probability density of R_l'. This is the most general result which can be achieved. More particular results need specific expressions for the joint distribution of R_l' and follow.

The approach to be adopted is that of approximating the distribution of R_l' by a multivariate Gaussian, as is at least plausible since (55) includes a Gaussian noise term. However, this approach has not been formally justified but will nevertheless be seen to produce decoders which are advantageous in their BER performance. The required mean vector of R_l' is one of all zeros due to the zero mean of the spreading, and the covariance components in this case can be written in terms of $\sigma_\varepsilon^2, \sigma_X^2, L$ and the linear autocorrelations $\{\rho_X(1), \rho_X(2), \ldots, \rho_X(N)\}$ of the spreading segment, as

$$\begin{aligned} \operatorname{var}(R_{li}') &\equiv \sigma_{R_l'}^2 = (L-1)\sigma_X^2 + \sigma_\varepsilon^2, \\ \operatorname{cov}(R_{li}', R_{lj}') &= (L-1)\sigma_X^2 \rho_X(|i-j|), \quad i \neq j. \end{aligned} \tag{58}$$

Thus, the auto-covariance matrix of R_l' is thus

$$\begin{aligned} \Sigma_{R_l'} = (L-1)\sigma_X^2 &\cdot \begin{bmatrix} 1+\sigma_\varepsilon^2/(L-1)\sigma_X^2 & \rho_X(1) \\ \rho_X(1) & 1+\sigma_\varepsilon^2/(L-1)\sigma_X^2 \\ \vdots & \vdots \\ \rho_X(N-1) & \rho_X(N-2) \end{bmatrix} \\ &\cdot \begin{bmatrix} \cdots & \rho_X(N-1) \\ \cdots & \rho_X(N-2) \\ \vdots & \vdots \\ \cdots & 1+\sigma_\varepsilon^2/(L-1)\sigma_X^2 \end{bmatrix}. \end{aligned} \tag{59}$$

Then by invoking the multivariate Gaussian density with these results, so assuming

$$\begin{aligned} &f_{R_l'}(r \mid b_l = \pm 1, X_l^T) \\ &\quad = \frac{1}{(2\pi)^{N/2}|\Sigma_{R_l'}|^{1/2}} \exp\left[-\frac{1}{2} r^T \Sigma_{R_l'}^{-1} r\right] \end{aligned} \tag{60}$$

and using (57), there is the approximate likelihood decoder

$$\hat{b}_l = \operatorname{sign}\{X_l^T \Sigma_{R_l'}^{-1} R_l\}. \tag{61}$$

The pleasing statistical structure to this result is from its obvious generalization of the correlation decoder (7) for the single-user case. A key new point is that the decoder gains advantage from leveraging the type of spreading, while a negative point is that in general it is not immediately useable because $\Sigma_{R_l'}^{-1}$ contains the unknown channel noise variance.

However, there are there are two useful special cases applying to multi-user CSK and which have simple forms of $\Sigma_{R_l'}^{-1}$ in which the noise variance is absent, and they suggest a generally useable form. The first is when the spreading is uncorrelated, for example, by a logistic map, and then by (59) $\Sigma_{R_l'}^{-1}$ is the identity matrix and so does not involve σ_ε^2. In this case, the likelihood decoder (61) reduces to an ordinary correlation decoder of the form (7). The second and generally more useful case is when there is a large number of other users, and σ_ε^2 is negligible relative to $(L-1)\sigma_X^2$, the total variance of the other-users. Then, apart from the initial $(L-1)\sigma_X^2$, $\Sigma_{R_l'}^{-1}$ is the inverse matrix of the $N-1$ autocorrelations of the spreading segment, given by the circulant matrix

$$\Sigma_X \equiv \begin{bmatrix} 1 & \rho_X(1) & \cdots & \rho_X(N-1) \\ \rho_X(1) & 1 & \cdots & \rho_X(N-2) \\ \vdots & \vdots & \vdots & \vdots \\ \rho_X(N-1) & \rho_X(N-2) & \cdots & 1 \end{bmatrix}. \tag{62}$$

This suggests the decoder

$$\hat{b}_l = \operatorname{sign}\{X_l^T \Sigma_X^{-1} R_l\} \tag{63}$$

which will be called the *likelihood-based decoder* and is easily available with chaotic spreading after the one-time inversion of Σ_X and is a natural generalization of the correlation decoder (7). Importantly, knowing σ_ε^2 is not necessary. Although a significant improvement on the correlation decoder, it is not fully optimal because of the derivation including approximating assumptions. However, in single-user systems with autocorrelated or non-Gaussian noise there could be advantage relative to the ordinary correlation decoder.

When the autocorrelations are powers of ρ, as they are for Bernoulli-shift map spreading with $\rho = \frac{1}{2}$, the inversion of (62) is explicit and the decoder appears as

$$\begin{aligned} \hat{b}_l = \operatorname{sign}\Bigg\{ &(X_{l1} - \rho X_{l2}) R_{l1} \\ &+ \sum_{i=2}^{N-1} \{-\rho X_{li-1} + (1+\rho^2) X_{li} - \rho X_{li+1}\} R_{li} \\ &+ (-\rho X_{lN-1} + X_{lN}) R_{lN} \Bigg\}. \end{aligned} \tag{64}$$

The form (64) is known from other areas of communication engineering as a type of *rake decoder*. The statistical point to be made is that it can be seen as a likelihood-based.

The exact approach is next used to study the BER of the likelihood-based decoder (63); the structured

mathematical development allows useful communication engineering insights. By writing

$$\begin{aligned} C^l_{R,X} &= X_l^T \Sigma_X^{-1} R_l \\ &= X_l^T \Sigma_X^{-1} (b_l X_l + X_{[l]} b_{[l]} + \varepsilon_l) \end{aligned} \tag{65}$$

the active-user bit error rate is first obtained as

$$\begin{aligned} \mathrm{BER}_l &= P\big(C^l_{R,X} < 0 \mid b_l = 1\big) \\ &= P\big\{X_l^T \Sigma_X^{-1} \varepsilon_l < -X_l^T \Sigma_X^{-1} X_l b_l \\ &\qquad - X_l^T \Sigma_X^{-1} X_{[l]} b_{[l]} \mid b_l = 1\big\}. \end{aligned} \tag{66}$$

Conditional on the spreading segments, and after using the Gaussian distribution of the noise term $X_l^T \Sigma_X^{-1} \varepsilon_l$, which has mean zero and variance $X_l^T \Sigma_X^{-1} \Sigma_X^{-1} X_l \sigma_\varepsilon^2$, there is the result

$$\mathrm{BER}_l(X) = \Phi\left\{-\frac{X_l^T \Sigma_X^{-1} X_l + X_l^T \Sigma_X^{-1} X_{[l]} b_{[l]}}{\sigma_\varepsilon [X_l^T \Sigma_X^{-1} \Sigma_X^{-1} X_l]^{1/2}}\right\}. \tag{67}$$

The argument of Φ can next be regarded as a function of two dependent variables, the lth active-user spreading segment X_l and the other-user interference variable

$$\Psi_l \equiv X_l^T \Sigma_X^{-1} X_{[l]} b_{[l]}. \tag{68}$$

Thus, the unconditional probability from (67) for the EGT result given $b_l = 1$, is

$$\mathrm{BER}_l = E_{X_l, \Psi_l} \Phi\left\{-\frac{X_l^T \Sigma_X^{-1} X_l + \Psi_l}{\sigma_\varepsilon [X_l^T \Sigma_X^{-1} \Sigma_X^{-1} X_l]^{1/2}}\right\}. \tag{69}$$

It is still difficult to proceed exactly, but the joint expectation (69) can be evaluated approximately by noting that $\Psi_l \mid X_l = b_{[l]}^T X_{[l]}^T \Sigma_X^{-1} X_l$ is the sum of $L-1$ conditionally independent variables $b_k X_k^T \Sigma_X^{-1} X_l$, $k = 1, \ldots, [l], \ldots, L$, and thus, at least for several users, can be assumed conditionally to have a Gaussian distribution, by the so-called *theoretical Gaussian approximation* (TGA). Immediate observation of $\Psi_l \mid X_l$ shows that the conditional mean is zero and for the exact conditional variance, taken over the spreading of other-users, there is

$$\begin{aligned} &\mathrm{var}(\Psi_l \mid X_l) \\ &\quad = \mathrm{var}\big(b_{[l]}^T X_{[l]}^T \Sigma_X^{-1} X_l\big) = \sum_{k \neq l, =1}^{L} \mathrm{var}\big(b_k X_k^T \Sigma_X^{-1} X_l\big) \\ &\quad = \sum_{k \neq l, =1}^{L} X_l^T \Sigma_X^{-1} \mathrm{var}\big(b_k X_k^T\big) \Sigma_X^{-1} X_l \\ &\quad = (L-1) X_l^T \Sigma_X^{-1} \mathrm{var}\big(b_k X_k^T\big) \Sigma_X^{-1} X_l. \end{aligned} \tag{70}$$

This simplifies by calculation of $\mathrm{var}(b_k X_k^T)$ and since $b_k^2 \equiv 1$ and leads to the result

$$\mathrm{var}(\Psi_l \mid X_l) = \sigma^2_{\Psi_l \mid X_l} = (L-1) X_l^T \Sigma_X^{-1} X_l \sigma_X^2. \tag{71}$$

Thus, with $\Psi = \sigma_{\Psi_l \mid X_l} Z$, Z a standardized Gaussian variable, (69) conditional on X_l becomes

$$\begin{aligned} &\mathrm{BER}_l(X_l) \\ &\quad \simeq \int_{-\infty}^{\infty} \Phi\left\{-\frac{X_l^T \Sigma_X^{-1} X_l + \sigma_{\Psi_l \mid X_l} z}{\sigma_\varepsilon [X_l^T \Sigma_X^{-1} \Sigma_X^{-1} X_l]^{1/2}}\right\} \phi(z)\, dz, \end{aligned} \tag{72}$$

where $\phi(\cdot)$ is the standardized Gaussian probability density function. By applying the Gaussian identity

$$\begin{aligned} &\int_{-\infty}^{+\infty} \Phi\{-(a+bz)/c\} \phi(z)\, dz \\ &\quad = \Phi\{-a/(b^2+c^2)^{1/2}\}, \end{aligned} \tag{73}$$

to (72), the BER of the likelihood-based decoder becomes

$$\begin{aligned} \mathrm{BER}_l \simeq E_{X_l} \Phi\{&-(X_l^T \Sigma_X^{-1} X_l)\big[(X_l^T \Sigma_X^{-1} \Sigma_X^{-1} X_l)\sigma_\varepsilon^2 \\ &+ (L-1)(X_l^T \Sigma_X^{-1} X_l)\sigma_X^2\big]^{-1/2}\}. \end{aligned} \tag{74}$$

Communication engineering insight is provided by separating out the channel noise and other-user effects in (74). Then a structured form of the approximate BER of the multi-user coherent CSK system with likelihood-based decoding is

$$\begin{aligned} \mathrm{BER}_l \simeq E_{X_l} \Phi\Bigg\{\Bigg[&-(X_l^T \Sigma_X^{-1} X_l / N\sigma_X^2) \\ &\Big/\Big(\frac{X_l^T \Sigma_X^{-1} \Sigma_X^{-1} X_l}{X_l^T \Sigma_X^{-1} X_l} \frac{1}{\mathrm{SNR}} + \frac{1}{\mathrm{SIR}}\Big)\Bigg]^{1/2}\Bigg\}. \end{aligned} \tag{75}$$

Here, SIR is a new and generally useful multi-user quantity, the s*preading to other-user interference ratio* (SIR), defined as $N/(L-1)$. For a very large number of other-users relative to N, SIR approaches 0 and BER approaches its worst value 0.5; and similarly for a very low SNR. A very large SNR overpowers transmission noise, and then BER is due only to SIR other-user interference and becomes

$$\begin{aligned} \mathrm{BER}_l &\simeq E_X \Phi\{-\big[(X^T \Sigma_X^{-1} X / N\sigma_X^2)\mathrm{SIR}\big]^{1/2}\} \\ &\geq \Phi\{-\sqrt{\mathrm{SIR}}\}, \end{aligned} \tag{76}$$

with the lower bound by application of Jensen's inequality.

Extensive spreading relative to other-users, oppositely overpowers other-user interference and BER is

due only to channel SNR. Then there is the earlier result

$$\text{BER}_l \geq \Phi(-\sqrt{\text{SNR}}) \tag{77}$$

as obtained at (14) for the single-user coherent case.

When spreading is uncorrelated, as with the logistic map, there is the particularly simple case of (75) which reduces to

$$\text{BER}_l \simeq E_{X_l}\Phi\left\{\left[-(X_l^T X_l/N\sigma_X^2)\Big/\left(\frac{1}{\text{SNR}}+\frac{1}{\text{SIR}}\right)\right]^{1/2}\right\} \tag{78}$$

and this gives the even simpler lower bound

$$\text{BER}_l \geq \Phi\left\{-\left(\frac{1}{\text{SNR}}+\frac{1}{\text{SIR}}\right)^{-1/2}\right\}. \tag{79}$$

The uncorrelated feature is only significant in multi-user systems. The lower bounds are reached by extensive spreading.

Exact computations using (75) leverages the chaotic spreading assumptions, as in the corresponding single-user calculation at (18). Without it, the expectation in (75) is an intractable N-dimensional integral.

Omitting full presentation, there is an SGA result corresponding to (75) which takes the form

$$\begin{aligned}\text{BER}_l \simeq \Phi\Bigg\{-\Bigg[&N^{-1}\frac{\sigma_{X^2}^2}{\sigma_X^4}\left\{1+2\sum_{k=1}^{N-1}\left(1-\frac{k}{N}\right)\rho_{X^2}(k)\right\}\\ &+\left\{1+2\sum_{k=1}^{N-1}\left(1-\frac{k}{N}\right)\rho_\varepsilon(k)\rho_X(k)\right\}\frac{1}{\text{SNR}}\\ &+\left\{1+2\sum_{k=1}^{N-1}\left(1-\frac{k}{N}\right)\{\rho_X(k)\}^2\right\}\frac{1}{\text{SIR}}\Bigg]^{-\frac{1}{2}}\Bigg\}.\end{aligned} \tag{80}$$

Communication engineering interest in (80) is the implication that low BER is realized when its negative-square-rooted term is large giving a very negative argument to the Gaussian distribution function. The first two terms are as in the single-user result (24) while the third term is due to the other-users. Statistical interest is in the involvement of two types of autocorrelation of the chaotic spreading sequence, with the linear autocorrelations not appearing in single-user results. Low BER occurs when the quadratic autocorrelations are less or equal to zero and linear autocorrelations of either sign are as small as possible; the former was found for single-user systems in Section 3 concerning design issues. The value of these results is qualitative, since as for single-user systems, the accuracy of the SGA may be poor. Tam, Lau and Tse (2007), Chapter 3, obtained a much less structured general SGA expression and a simplified more specific form for uncorrelated spreading.

Similar results to those given in this section are available for standard correlation decoders and in non-coherent multi-user systems, but for reasons of space are not included and will be reported elsewhere.

The section concludes with illustrations in Figure 9 of BER results for a small multi-user system with 3 users each using Bernoulli map spreading segments of length 4. The first and uppermost curve is from the omitted SGA approximation to the BER of correlation decoding. The second curve is from the omitted TGA approach for correlation decoding. The third curve is the TGA approximation (75) for likelihood-based decoding. This compares very favourably with the fourth curve for the lower bound (79) from uncorrelated spreading. The final curve is from the single-user lower bound (14), so ignoring other users. The horizontal line is the lower bound from (76) for no channel noise, with BER only due to the interference by the other two users. The main point is the superiority of likelihood-based decoding over correlation decoding, and achieved practically with little computational cost. The curves generally support the previous theoretical discussions. However, for unrealistically low SNR, where the TGA approach is not claimed to be theoretically accurate, there is no evidence of improvement over the correlation decoder.

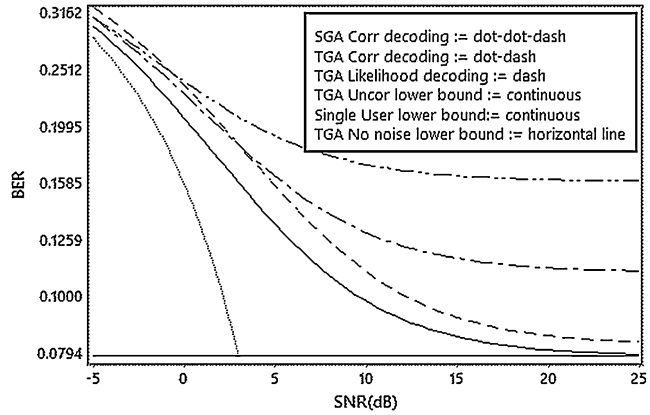


FIG. 9. *Three-user* ($L = 3$) *coherent CSK with Bernoulli spreading*, $N = 4$ *and* SIR $= 2$; *comparisons of BER calculations. SGA for correlation decoding* (*dot-dot-dash curve*), *TGA for correlation decoding* (*dot-dash curve*), *TGA for likelihood decoding* (*dash-curve*), *TGA Jensen lower bound for likelihood-based decoding with uncorrelated spreading* (*continuous curve*), *single-user lower bound* (*continuous-curve*), *theoretical lower bound for likelihood-based decoding with no channel noise at* $\Phi(-\sqrt{\text{SIR}}) = 0.07865$ (*horizontal line*).

The single-user lower bound emphasizes the strong interference effect of other users on BER. Although not shown, Bernoulli-shift map spreading is actually superior to logistic spreading for high SNR, contrary to that in single-user systems. These illustrations are not intended to represent a realistic engineering-size application, rather to be an exemplar of exact results which are not too computationally demanding.

7. FINAL OVERVIEW

More could be said about most of the topics covered and many topics and contributions in the area have not been mentioned. Further information is given in the References, particularly the engineering publications, their references and the list of chaos-based communication models in the Appendix. Among omitted topics are multi-type symbol systems, systems employing different generators for each transmitted symbol type and many models designed to suit to particular practical circumstances. The flexibility of statistical modelling should allow further developments.

This account has focussed on the modelling and statistical aspects of chaos-based antipodal shift-keying systems for transmitting binary messages, with mention of extensions to the basic supporting structures to include non-Gaussian and autocorrelated noise and errors. There has been emphasis on exact statistical modelling approaches and structured results for bit error. This is in contrast to the treatment of quite a large number of other models in the engineering literature for particular communication architectures, mostly with the little or no concern for exact treatments. A distinction has been made between systems involving electronic circuit generation of chaos and those involving laser-based optical generation. Developments of non-coherent models applying to imperfectly synchronized systems, as found in laser-based optical systems, have been given. The possible synchronization of laser-based chaos is usefully employed here because it provides the means of remotely synchronizing spreading sequences as required in non-coherent systems. As far as the non-coherent CSK systems are concerned, likelihood-based decoders, together with their BER performance, have been given, a previously open area. Design analysis has indicated that negative quadratic autocorrelation of spreading reduces bit error. It is worth emphasizing that the chaotic map spreading assumption is mainly theoretically useful because it allows exact or very accurate theoretical calculations of bit error rates in coherent systems and provides insights as to optimal spreading, and further is the basis of steganographic security. In itself, the chaos assumption is not essential. The mathematical structure of chaos-based systems is actually free of chaotic assumptions. There are still theoretical challenges with exact likelihood decoders for multi-user CSK systems and computational challenges in calculating extremely small BERs for very large spreading factors. But, hopefully, this account has demonstrated that statistical-based modelling and communications engineering can work together to provide new insights and useful results.

APPENDIX: CHAOS COMMUNICATION SYSTEMS NOT COVERED

A partial list of other chaos-based models in the literature includes the following:

Differential chaos shift-keying (DCSK), Kolumbán et al. (1996), Frequency modulated differential chaos shift-keying (FM-DCSK), Kolumbán et al. (1998), Correlation delay shift-keying (CD-CSK), Sushchik, Tsimring and Volkovskii (2000), Symmetric chaos shift-keying (SCSK), Sushchik, Tsimring and Volkovskii (2000), On-Off shift-keying (ON-OF CSK), Uchida et al. (2001), Heil et al. (2002), Quadrature CSK (Q-CSK), Galias and Maggio (2001), Chaotic phase shift-keying (CP-CSK), Hasler and Schimming (2002), Permutation-based DCSK (P-DCSK), Lau, Cheong and Tse (2003), High-data-rate code shifted CSK (CSDCSK), Kaddoum and Gagnon (2012), High efficiency CSK (HE-CSK), Hua and Ping (2012), Space-time block CSK (STBC-CSK), Kaddoum and Gagnon (2013a), Decode-and Forward CSK (DF-CSK), Kaddoum and Gagnon (2013b); Improved non-coherent DCSK (I-DCSK), Kaddoum, Soujeri and Arcila (2015).

ACKNOWLEDGMENTS

I am very grateful for my collaborators in this area, in chronological order, to Professor N. Balakrishna, Dr. R. Hilliam, Dr. J. Yao, Dr. T. Papamarkou, Dr. G. Kaddoum and Professor A. Uchida, but I am responsible for any deficiencies. For early encouragements from the communications engineering community, I thank Professor M. Hasler, Dr. F. C. M. Lau, Professor G. Mazzini, Professor R. Rovatti, Dr. T. Schimming, Professor W. Schwarz, Professor G. Setti and Professor C. K. Tse. I am particularly grateful to Professor T. Kohda

for stimulating my initial interest in the statistical properties of chaos. Parts of this paper were revised during a visit to the University of Western Australia, Perth, and supported by the Institute of Advanced Studies and Professor M. Small.

REFERENCES

ABEL, A., SCHWARZ, W. and GOTZ, M. (2000). Noise performance of chaotic communication systems. *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.* **47** 1726–1732.

BERLINER, L. M. (1992). Statistics, probability and chaos. *Statist. Sci.* **7** 69–122. MR1173418

CARROLL, T. L. and PECORA, L. M. (1992). A circuit for studying the synchronization of chaotic systems. *Internat. J. Bifur. Chaos Appl. Sci. Engrg.* **2** 659–667. MR1192702

DONATI, S. and MIRASSO, C. M. (2002). Introduction to the feature section on optical chaos and applications to cryptography. *IEEE J. Quantum Electron.* **38** 1138–1140.

GALIAS, Z. and MAGGIO, G. M. (2001). Quadrature chaos-shift keying: Theory and performance analysis. *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.* **48** 1510–1519. MR1873101

HASLER, M. and SCHIMMING, T. (2002). Optimal and suboptimal chaos receivers. *Proc. IEEE* **50** 733–746.

HASLER, M., MAZZINI, G., OGORZALEK, M., ROVATTI, M. and SETTI, G. (2002). Scanning the special issue—special issue on applications of nonlinear dynamics to electronic and information engineering. *Proc. IEEE* **90** 631–640.

HEIL, T., MULET, J., FISCHER, I., MIRASSO, C. R., PEIL, M., COLET, P. and ELSABER, W. (2002). ON/OFF phase shift keying for chaos-encrypted communication using external-cavity semiconductor lasers. *IEEE J. Quantum Electron.* **38** 1162–1170.

HOMER, M. E., HOGAN, S. J., BERNADO, M. D. and WILLIAMS, C. (2004). The importance of choosing attractors for optimizing chaotic communications. *IEEE Trans. Circuits Syst. II, Express Briefs* **51** 511–516.

HUA, Y. and PING, J. G. (2012). High-efficiency differential-chaos-shift-keying scheme for chaos-based non-coherent communication. *IEEE Trans. Circuits Syst. II, Express Briefs* **59** 312–316.

KADDOUM, G. and GAGNON, F. (2012). Design of a high-data-rate chaos-shift system. *IEEE Trans. Circuits Syst. II, Express Briefs* **59** 448–452.

KADDOUM, G. and GAGNON, F. (2013a). Performance analysis of STBC-CSK communication system over slow fading channel. *Signal Process.* **93** 2055–2060.

KADDOUM, G. and GAGNON, F. (2013b). Lower bound on the BER of a decode-and-forward relay network under chaos shift keying communication system. *IET Commun.* **COM-2013-0421.R2** 1–16.

KADDOUM, G., SOUJERI, E. and ARCILA, C. A. (2015). I-DCSK: An improved non-coherent communication system architecture. *IEEE Transactionson Circuits, Systems II: Express Briefs* 1–5.

KADDOUM, G., COULON, M., ROVIRAS, D. and CHARGE, P. (2010). Theoretical performance for asychronous multi-user chaos-based communication systems on fading channels. *Signal Processing* **90** 2923–2933.

KENNEDY, M. P., ROVATTI, R. and SETTI, G., eds. (2000). *Chaotic Electronics in Telecommunications*. CRC Press, Boca Raton.

KOHDA, T. and MURAO, K. (1990). Approach to time sereis analysis for one-dimnsionaal chaos based on Frobenius–Perron operator. *Transactions Institute of Electronics, Information and Communication Engineers of Japam* (*IEICE*), *E* **73** 793–800.

KOLUMBÁN, G. (2000). Theoretical noise performance of correlator-based chaotic communications schemes. *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.* **47** 1692–1701. MR1816148

KOLUMBÁN, G. and KENNEDY, M. P. (2000). Special Issue on "Non-coherent chaotic communications". *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.* **47** 1661–1662.

KOLUMBÁN, G., VIZVARI, B., SCHWARZ, W. and ABEL, A. (1996). Differential chaos shift keying: A robust coding for chaos communication. In *Proc. NDES'*96 87–92. Seville, Spain.

KOLUMBÁN, G., KIS, G., JAKO, Z. and KENNEDY, M. P. (1998). FM-DCSK: A robust modulation scheme for chaos communication. *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences E* **81-A** 1798–1802.

LARSON, L. E., LIU, J.-M. and TSIMRING, L. S., eds. (2006). *Digital Communications Using Chaos and Nonlinear Dynamics*. Springer, New York.

LASOTA, A. and MACKEY, M. C. (1994). *Chaos, Fractals, and Noise: Stochastic Aspects of Dynamics*, 2nd ed. *Applied Mathematical Sciences* **97**. Springer, New York. MR1244104

LAU, F. C. M., CHEONG, K. Y. and TSE, C. K. (2003). Permutation-based DCSK and multiple-access DCSK systems. *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.* **50** 733–742. MR2003801

LAU, F. C. M. and TSE, C. K. (2003). *Chaos-Based Digital Communications Systems*. Springer, Heidelberg.

LAWRANCE, A. J. and BALAKRISHNA, N. (2001). Statistical aspects of chaotic maps with negative dependence in a communications setting. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **63** 843–853. MR1872070

LAWRANCE, A. J. and BALAKRISHNA, N. (2008). Statistical dependency in chaos. *Internat. J. Bifur. Chaos Appl. Sci. Engrg.* **18** 3207–3219. MR2487910

LAWRANCE, A. J. and OHAMA, G. (2003). Exact calculation of bit error rates in communication systems with chaotic modulation. *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.* **50** 1391–1400. MR2024566

LAWRANCE, A. J. and OHAMA, G. (2005). Bit error probability and bit outage rate in chaos communication. *Circuits Systems Signal Process.* **24** 519–533. MR2187036

LAWRANCE, A. J., PAPAMARKOU, T. and UCHIDA, A. (2017). Synchronized laser choas communication: Statistical investigation of an experimental system. *IEEE J. Quantum Electronics* **53**. To appear.

LAWRANCE, A. J. and YAO, J. (2008). Likelihood-based demodulation in multi-user chaos shift keying communication. *Circuits Systems Signal Process.* **27** 847–864. MR2466043

MIRASSO, C. R., COLET, P. and GARCIA-FERNANDEZ, P. (1996). Synchronization of chaotic semiconductor lasers: Application to encoded communications. *IEEE Photonics Technol. Lett.* **8** 299–301.

PAPAMARKOU, T. and LAWRANCE, A. J. (2007). Optimal spreading sequences for chaos-based communications. In 2007 *International Symposium on Nonlinear Theory and Its Applications*

(*NOLTA'*07). *Research Society of Nonlinear Theory and Its Applications* 208–211. IEICE, Vancouver.

PAPAMARKOU, T. and LAWRANCE, A. J. (2013). Paired Bernoulli circular spreading: Attaining the BER lower bound in a CSK setting. *Circuits Systems Signal Process*. **32** 143–166. MR3018767

PARLITZ, U. and ERGEZINGER, S. (1994). Robust communication based on chaotic spreading sequences. *Phys*. *Lett*. *A* **188** 146–150.

PECORA, L. M. and CARROLL, T. L. (1990). Synchronization in chaotic systems. *Phys*. *Rev. Lett*. **64** 821–824. MR1038263

PROAKIS, J. G. (2001). *Digital Communication*. McGraw Hill, Boston, MA.

SUSHCHIK, M., TSIMRING, L. S. and VOLKOVSKII, A. R. (2000). Performance analysis of correlation-based communication schemes utilizing chaos. *IEEE Trans*. *Circuits Syst*. *I*, *Fundam*. *Theory Appl*. **47** 1684–1691.

TAM, W. M., LAU, F. C. M. and TSE, C. K. (2003). The analysis of bit error rates for multiple access CSK and DCSK communication. *IEEE Trans*. *Circuits Syst*. *I*, *Fundam*. *Theory Appl*. **50** 702–707.

TAM, W. M., LAU, F. C. M. and TSE, C. K. (2007). *Digital Communications with Chaos*. Elsevier, Oxford.

TAM, W. M., LAU, F. C. M., TSE, C. K. and YIP, M. (2002). An approach to calculating the bit-error rate of a coherent chaos-shift-keying communication system under a noisy multiuser environment. *IEEE Trans*. *Circuits Syst*. *I*, *Fundam*. *Theory Appl*. **49** 210–233.

TAM, W. M., LAU, F. C. M., TSE, C. K. and LAWRANCE, A. J. (2004). Exact analytical of bit error rates for multiple access chaos-based communication systems. *IEEE Trans*. *Circuits Syst*. *I* **51** 473–481.

UCHIDA, A. (2012). *Optical Communications with Lasers—Applications of Nonlinear Dynamics and Synchronization*. Wiley-VCH, Weinheim.

UCHIDA, A., YOSHIMORI, S., SHINOZUKA, M., OGAWA, T. and KANNARI, F. (2001). Chaotic on off keying for secure communications. *Opt*. *Lett*. **26** 866–868.

YAO, J. (2004). Optimal chaos shift keying communications with correlation decoding. In *IEEE International Symposium on Circuits and Systems* (*ISCAS*2004) 593–596, Vancouver, Canada.

ZHOU, Z., WANG, J. and YE, Y. (2010). Exact BER analysis of differential chaos shift keying communication system in fading channels. *Wirel*. *Pers*. *Commun*. **53** 299–310.