Chaotic Communications Over Radio Channels

Chris Williams

***Abstract*—The real issues of communicating over radio channels with continuous time chaotic signals are addressed. Synchronized and nonsynchronized methods are assessed. The modeling of radio channels (including multipath and Doppler) is reviewed, and then the effect on chaotic signals is investigated. The bandwidth efficiency of chaotic communication systems is discussed, and compared to conventional methods. The practical issues of symbol synchronization and equalization are discussed. It is pointed out that discussing these systems as spread spectrum is often unhelpful, because a noise free reference signal is not available at the receiver to exploit the excess bandwidth.**

***Index Terms*—Bandwidth efficiency, chaos, communications, equalization, synchronization.**

I. Introduction

FUTURE communication systems will be required to operate at higher data rates, be more reliable, and operate in increasingly crowded frequency allocations. The concept of personal communication services put additional constraints on power efficiency (impacting on size and weight), and mobility range. Conventional communication systems have the fundamental constraint that the carrier waveforms are limited to a two signal orthogonal set (sine and cosine waves). This limitation leads to a trade off between power efficiency (for orthogonal signal sets) and bandwidth efficiency (for high level modulation schemes). The emphasis in this paper will be on mobile radio communication systems.

Recent years have seen the emergence of chaotic communication systems. However, it is not always recognized what application these systems would serve. By examining the properties of chaotic systems we can learn why one may want to use them for communications. Being fundamentally broadband, the presence of information does not necessarily change the properties of the transmitted signal. While obviously an advantage from a security viewpoint, a further consequence is that the power output will remain constant regardless of the information content (or lack of it). The term *broadband* needs clarification. The bandwidth of a chaotic signal is infinite, but some large fraction (say 90%) of the signal power will lie in some bandwidth B. When a chaotic signal is used as a carrier, the unmodulated bandwidth B is clearly broadband compared to a conventional carrier (sinusoid), which has a purely impulsive spectra (definitely narrowband). Many papers have proposed using chaotic signals in the spread spectrum context, in which case the transmitted signal bandwidth is significantly wider than the information bandwidth. Whether such an approach has any merit for chaotic communications is discussed within this paper. The discussions here are not limited to this point of view, and in some cases the transmitted signal bandwidth will be comparable to the information bandwidth. The nonperiodic nature would suggest that a large number of waveforms are available. However, these are not orthogonal in the conventional sense, except in the infinite time limit. The nonperiodic nature, and limited predictability, may also suggest that these systems would be secure in some sense. This will not be an over-riding issue in this paper.

High-dimensional systems, and the consequent increase in the number of parameters, may offer a wider variety of waveforms. It is natural to assume that the larger the signal set size, the greater the data-carrying capacity. However, the reliability of the recovered information is of great importance, and will be assessed here. The sensitive dependence property can be used to apply small amplitude control signals to a chaotic system, and, after some latency period, cause a large scale change in the output. This can be used to remove the requirement for frequency conversion and amplification from the transmitter, thus improving power efficiency but also reducing size and weight.

At first sight, this is a very compelling list. However, chaotic-communication systems are not without their problems. While chaotic synchronization has been demonstrated, its robustness to channel distortion and noise is a cause for concern. The basis for using synchronization is discussed. As will be seen bandwidth efficiency can be poor. Direct RF generation and modulation for practical systems have not been readily demonstrated. This paper concentrateS on such continuous time systems, where, unlike discrete time systems, formal results on signal design do not exist. The structured nature of chaotic signals places clear doubt on the level of security they actually offer [1].

Some distinct approaches to communicating using chaotic waveforms are as follows. 1) DS/CDMA using discrete chaotic maps. To solve the synchronization problem, sections of a chaotic trajectory are used to form a periodic signal [2]. 2) Sending a copy of the chaotic signal as a reference signal [3]. 3) Using a continuous-time chaotic system, either directly at RF, or at baseband and using a conventional sinusoidal carrier.

This paper concentrates on the continuous-time approach. Typically, this has involved using chaotic-synchronization techniques, these, and methods to communicate using synchronization are discussed in Section II. Section III reviews the effect of the radio channel. The effects of different channel distortions are then investigated in Section IV, and the issue of bandwidth

Manuscript received February 19, 2001; revised August 13, 2001. This work was supported by the U.K. Ministry of Defence, Corporate Research Programme TG9. This paper was recommended by Guest Editor G. M. Maggio.

The author is with the Communications Department, QinetiQ, Malvern WR14 3PS, U.K. (e-mail: C_Williams@QinetiQ.com).

Publisher Item Identifier S 1057-7122(01)10380-6.

efficiency is presented in Section V. Practical issues of implementing chaotic communications systems are discussed in Section VI. Finally, conclusions are given in Section VII.

II. Review of Synchronization and Communications

When a system is said to be *synchronized*, this typically refers to identical synchronization which is when the state-space trajectories of drive and response systems (only one way coupling is considered here) are identical, and in phase. This paper will be restricted to scalar drive signals. Accurate identical synchronization relies on ideal parameter matching in the two systems and there being no distortion of the drive signal. Any discrepancy from this ideal will desynchronize the systems, the degree of which will depend how far from ideal the systems are. The relaxations in this definition of synchronization have the potential to provide more robust synchronized systems [4], [5]. If instead, there exists some continuous transformation that maps the trajectories in one state space to another, of two coupled systems, then the systems possess generalized synchronization (GS) [6]. The next level in the synchronization hierarchy is phase synchronization, if the difference between some suitably defined phases remains bounded then phase synchronization occurs [7].

The actual method by which synchronization is achieved depends on modifying the response system such that it is stable under driving from the received chaotic signal. The Pecora and Carroll method (PC) of synchronization [8] decomposes the n-dimensional dynamical system into two subsystems, the drive is composed of the full n-dimensional dynamics, whereas the response system is a subset of the full dynamics. In the active-passive decomposition (APD) [9] method the original autonomous system is rewritten as a nonautonomous (driven) system that possesses the desired synchronization properties. The linear feedback method (LF) [10] is derived from control theory. More recent work has considered that synchronization is a state-space observer, and so modern control theory can be applied [11].

The ability to synchronize does not also imply the ability to communicate. There are a number of possible methods that will be described now. The masking method [12] uses the fact that synchronization is robust against small perturbations of the drive signal. The APD method gives a framework where the message becomes part of the dynamics in all the coupled systems (chaotic modulation), due to the greater symmetry between drive and response it is expected that this method will be more robust [9], [13]. The chaos shift keying (CSK) technique [12] encodes the message information onto the attractor by modulating a parameter of the drive system, typically in a binary manner. The receiver's job is to decide which of the possible dynamical systems generated the received signal. The communicating by control approach [14] encodes the information in the symbolic dynamics (SD) of the chaotic system, using small control perturbations. Since the qualitative position on the attractor is more important for message recovery than the actual amplitude, such an approach can be considered as a form of phase synchronization. In the masking, APD and CSK methods synchronization is an obvious way of recovering the information. The synchronization mechanism can be considered as a nonlinear filter, matched to the specified dynamics. In the SD approach synchronization can be used to recover the full chaotic dynamics, and hence the information. However, full recovery of the state space is not required, just sufficient information to recover the symbolic dynamics. Alternative message recovery methods will be discussed.

III. Radio Channel Effects

A key aspect of any communication system is its performance over noisy distorting channels, and the distortion in a radio channel can be particularly severe. There are two classes of propagation models to consider, large scale models consider signal characterization that varies over large distances (many wavelengths), and small scale models that characterize behavior that varies over short distances (of the order of a wavelength, or a few seconds).

Large-scale variations are caused by path loss through the propagation media (free space), loss due to obstructions (buildings) and interaction with the environment (reflection, diffraction, and scattering). In free space, the path loss is characterized by a power reduction as the square of distance, whereas for a two-path model (to include a ground reflection) this is a fourth power law. In general, due to more complex environments, the path-loss exponent is not just limited to 2 or 4, but can vary between 1.5 and 6 [15]. The path loss exponent can also change its value with distance. The exact path loss at any particular location is a random variable depending on the actual environment, a log-normal random process is used to characterize this behavior, the path loss L is given by

$$L(\mathrm{dB}) = \bar{L}(\mathrm{dB}) + X_\sigma(\mathrm{dB}). \tag{1}$$

Where $\bar{L}$ is the mean path loss and X_σ is a Gaussian random variable (in decibels), with standard deviation σ (also in decibels). In the outdoor environment σ is in the range of 6 to 9, whereas in the more variable indoor environment it is 3 to 14. The mean path loss can be obtained from a propagation model or from empirical studies (see [15], [16] for details). While the path loss due to large scale models can be significant, the slow time variation can usually be tracked within the receiver, though it is the large scale variations that limit the coverage area of a transmitter. In contrast the small scale variations due to motion are a more significant problem.

Small-scale fading is caused by interference between multiple received signals form the same transmitter (multipath). The channel transfer function is dependent upon the relative amplitudes, phases, and time delays of the individual paths, and changes over wavelength distances. The effects of multipath are rapid-signal strength fluctuations, random frequency modulation (due to different Doppler shifts on the paths), and time dispersion. The effects are dependent on the scattering environment, speed of transmitter and receiver, and the channel characteristics. The multipath channel has a time varying impulse response, useful parameters include the mean excess delay and RMS delay spread τ (first and second moments) of the impulse response [16]. The effects of the channel are dependent on the relative characteristics of the channel and the signal, in both the

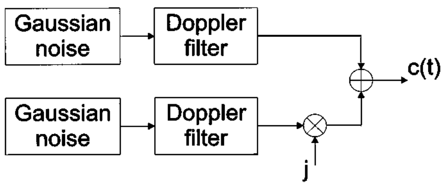

Fig. 1. Method of generating channel model tap coefficients.

time and frequency domains. The coherence bandwidth B_c is defined as the frequency separation required for two frequencies to be correlated to a specified percentage, for 50% correlation this is

$$B_c = \frac{1}{5\tau}. \tag{2}$$

The coherence bandwidth measures the flatness of the frequency response, when the signal bandwidth exceeds this, the channel is frequency selective. Alternatively, the channel is frequency selective if the delay spread is larger than the symbol period, when this is the case the signal is wideband, otherwise it is narrowband. The motion of the transmitter, receiver, or environment causes time variation of the channel, the timescales of this variation is given by the coherence time of the channel T_c. This can be defined as

$$T_c = \frac{1}{f_D}. \tag{3}$$

Where, f_D is the Doppler spread. Frequency dispersion due to Doppler spread causes time selective fading. Time selective fading, and frequency selective fading (due to multipath) are independent processes.

A channel model for small scale variation is a finite impulse response filter (tapped delay line), when the equivalent baseband representation is used [17] the tap coefficients are complex (to represent amplitude and phase variations in the channel). It is common to fix the multipath delays for each simulation, though the amplitudes and phases are time varying. The time variation of the tap coefficients are related to the Doppler spread characteristic, where there is no dominant path the Rayleigh fading distribution gives a baseband Doppler spectrum of

$$S_D(f) = \frac{1}{2\pi f_m \sqrt{1 - \left(\frac{f}{f_m}\right)^2}}. \tag{4}$$

Where f_m is the maximum Doppler shift. The channel–tap coefficients are made by summing two Gaussian sources (real and imaginary) which have been filtered according to (4), Fig. 1. The impulse response of such a filter is given by the inverse Fourier transform of (4). When a dominant line of sight (LOS) path exists the channel fading envelope is Ricean. The model described above can also be used when the additional LOS component is added.

IV. Towards Robust Communication

This section will demonstrate the effect of the radio channel on chaotic communication systems. Results will be presented as bit-error rate (BER) versus the bit energy over noise power density ratio (E_b/N_o), this removes the dependency on timescales and simulation bandwidth. Unless otherwise noted the Lorenz chaotic system has been used

$$\begin{aligned} \dot{x} &= \sigma(y - x) \\ \dot{y} &= rx - y - xz \\ \dot{z} &= -bx + xy. \end{aligned} \tag{5}$$

With parameters $r = 28.0, \sigma = 10.0, b = 2.6667$. An Adams integration routine has been used, with step size 0.01 s. Unless noted otherwise, symbol timing synchronization is assumed known, this will be discussed later. For the SD approach, the mean symbol period (time between section crossings) is 0.75 s, the symbol period for other methods will be as stated. The simulated chaotic signal is a real signal, and will be treated as the transmitted signal, and thus, all the noise power is purely real. It is possible to consider the simulation to be a baseband equivalent of the RF signal, with a heterodyning process separating the two. In this case, the occupied RF bandwidth includes the negative frequency components at baseband. To solve the obvious aliasing problem with the equivalent baseband representation, this signal should actually be considered as complex (with zero imaginary component in this case). The consequence of this is that the noise power is divided equally between the real and imaginary components.

A. Noise

Using additive white Gaussian noise (AWGN), the relative performance of a number of communication methods will be explored, and the effect of system parameters quantified. Due to the large number of degrees of freedom, direct comparison between the different methods is difficult, but trends and general behaviors can be identified.

Fig. 2 demonstrates the effect of AWGN on synchronization-based communication systems. The asymmetry of chaotic masking means that perfect synchronization cannot be achieved with finite message amplitudes, consequently synchronization error bursts lead to an error floor [Fig. 2(a)]. Increasing the symbol period improves performance (steeper BER curve) initially, but eventually due to the large bit energies performance degrades. Increasing the message amplitude can also improve performance (though this is less desirable from an implementation viewpoint), but in some systems (such as Chua's circuit) too large a message can destabilize the synchronization mechanism [18].

CSK has the disadvantage of requiring resynchronization between symbol changes, this also leads to an error floor (Fig. 2(b)). By not including the first part of the symbol in the synchronization error estimate (blanking) the error floor can be reduced. There will clearly be a trade-off between noise performance (long estimation period) and error floor level (long blanking period). Larger parameter shifts improve performance, but as before, this may be undesirable from an implementation viewpoint, the security of the waveform would also be compromised (if that is a concern). An alternative method using generalized synchronization has been explored in [4], it was found that while synchronization error was

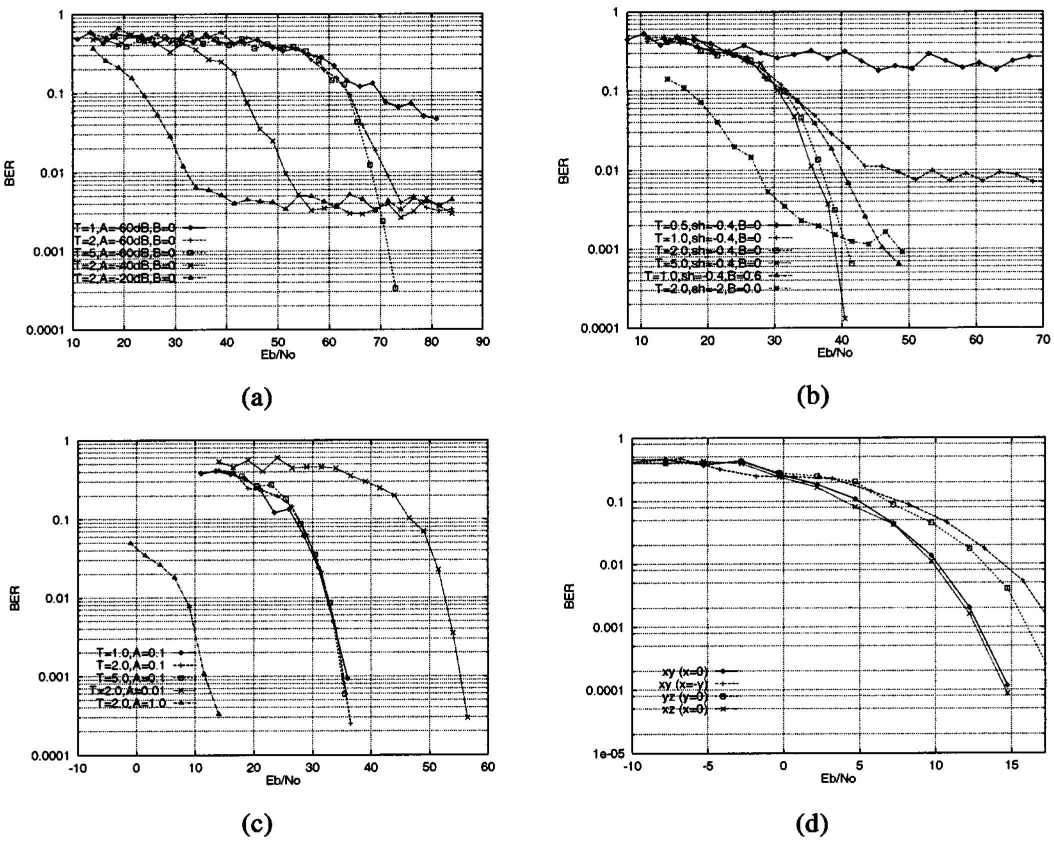

Fig. 2. Performance in AWGN for: (a) Masking; (b) CSK; (c) APD; (d) SD. T-symbol period (0.75 s unless noted otherwise), A-message amplitude, sh- σ parameter shift for encoding, B-blanking period. For SD, decision boundary in 2-D plane is defined.

reduced over a wide range of signal-to-noise ratios, due to slow synchronization times requiring long symbol periods, the performance in terms of E_b/N_o was poor.

The symmetry of the APD technique gives improved performance, and in this case there is much less variation in performance as the symbol period changes (Fig. 2(c)). The results show an improvement as the message amplitude increases, this may be beneficial if security is a concern [1] due to greater mixing of message and transmitter dynamics. However, as the message amplitude increases, we are moving away from the small signal control desired for power efficiency.

The coding by SD results show good performance when the symbol partition is carefully chosen [Fig. 2(d)]. Changing the decision boundary from $x = 0$ to the $x = -y$ line degraded performance. This is explained by considering the noise characteristics on the state variables, Fig. 3 shows the error between transmitter state variables and the equivalent in the receiver along the symbol surface of sections. The y variable errors have a greater variance and a larger mean offset. Also, the correlation between the errors on the x and y state variables has a correlation of 0.88. Consequently, in the xy plane, the noise is more likely to displace the trajectory closer toward the negative slope diagonal than the y-axis. In real systems, this technique will require careful coding, because not only are substitution errors likely to occur, but also data insertion and deletion errors. In terms of synchronization terminology, phase slips and advances may occur.

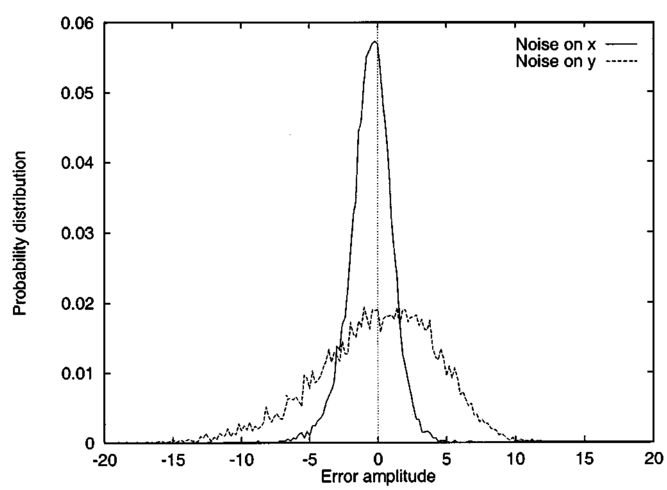

Fig. 3. Noise distribution on surface of section in x and y directions.

As previously noted, synchronization is not strictly necessary to recover the symbolic dynamics, and in the system explored here, the peaks of the received signal give sufficient information. Decisions could be made purely on the received signal (again assuming the symbol timing is known). To be fair, a 2-Hz second-order Butterworth filter was used in the receiver since the total received power over the simulation bandwidth would severely degrade performance with there being no inherent filtering process in the receiver (as in dynamical synchronization).

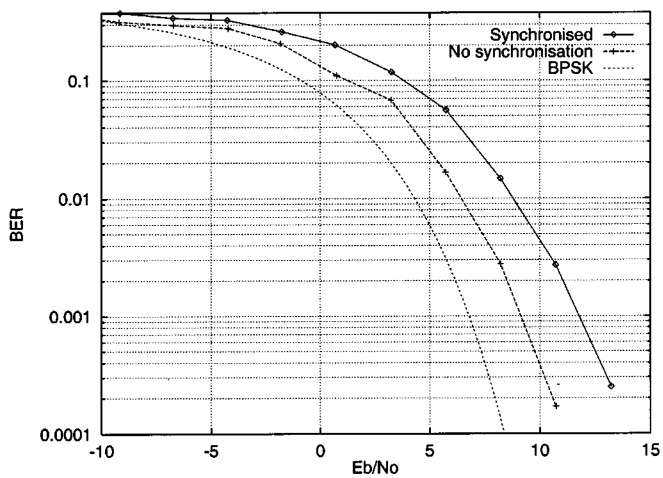

Fig. 4. Comparison of SD with and without chaotic synchronization in the receiver.

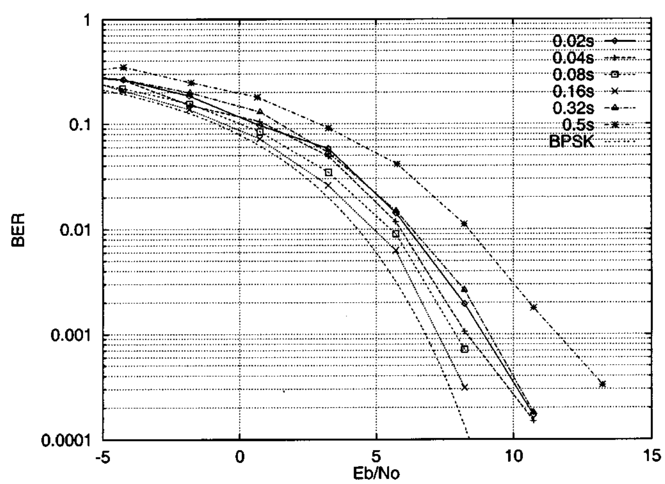

Fig. 5. Performance of SD with delay embedding in the receiver, as a function of sample period.

The improvement in performance (Fig. 4) clearly shows that the synchronization mechanism, while reducing the noise power due to its filtering properties, also distorts the dynamics which corrupts some of the decisions. Does this mean that the synchronization mechanism has little use in message recovery? Since ideal symbol synchronization was assumed up to now, in a practical receiver, the sampling instants will need to be determined. The obvious method is to use synchronization which exploits the known dynamics, and so some combination of synchronization and direct estimation should be used. In higher dimensional systems, with more elements in the symbolic dynamics, a more complete state-space reconstruction will be required and synchronization (or some other state-space observer mechanism [11]) will be required.

A state-space reconstruction using the delay-embedding method [19] could also be used. A two-lag delay vector is not sufficient for a full reconstruction, but does allow the SD to be recovered using $x = -y$ as the decision boundary. As in the previous nonsynchronised example a front-end filter is used to pre-process the incoming signal. Fig. 5 shows the performance against delay time for a 2-Hz preprocessing filter.

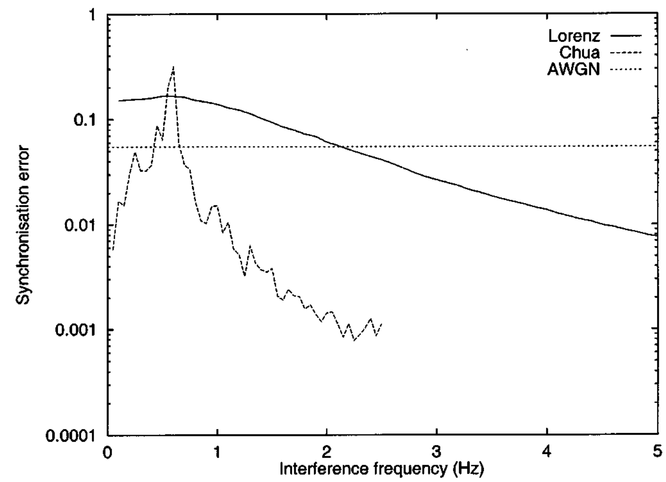

Fig. 6. Synchronization error with interfering sine wave (− 40 dB).

For short delays, the noise samples in each component of the lag vectors are correlated, so noise acts near-perpendicular to the decision boundary, whereas for the longer delays there is less correlation so performance improves. For delays that are too long, the reconstruction is too coarse, and sample points do not accurately reproduce the underlying attractor.

The assumption of AWGN is not always entirely valid, particularly when other interfering signals are present. The variation of sensitivity to frequency components across the operating bandwidth is demonstrated in Fig. 6 for the Lorenz and Chua's circuit. Clearly the response of the receiver to colored noise will depend both on the noise spectrum and this sensitivity characteristic.

B. Bandlimiting

Physical component constraints and intentional filtering to meet radio regulations, will lead to the transmitted signals being bandlimited. Here a second order Butterworth IIR low pass filter is considered. Fig. 7(a) (APD) shows how there is a transition in performance as the filter bandwidth is reduced, CSK and masking show similar behavior. This is a consequence of the synchronization error increasing as more of the chaotic dynamics are distorted [18]. In the SD case the transition is at a much lower frequency [Fig. 7(b)]. This figure shows that while synchronization error can be large the communication performance can be good, since it is the geometry from which the message is recovered, not absolute signal levels. This is clear for bandlimiting with a 3-dB break point of 1 Hz, where performance shows an improvement over the unfiltered case. Here, the removal of noise is more beneficial than the loss of signal (and consequent distortion).

C. Fading

Although different in nature, slow and fast fading can both be modeled using a tapped delay line (finite impulse response) structure. The simplest case is the static nonfrequency selective (flat) fading channel, demonstrated in Fig. 8 for APD and SD. This models the situation where an automatic-gain control (possibly in conjunction with transmitter-power control) system tries to maintain a constant power output to the synchronising circuit,

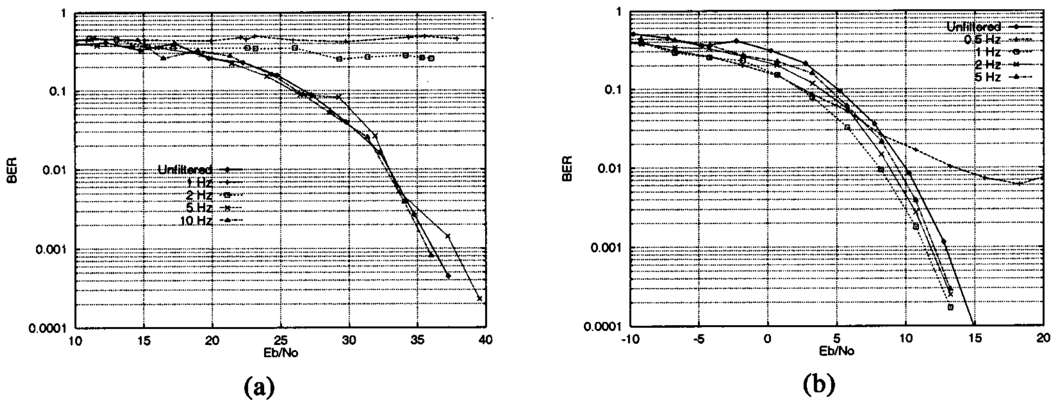

Fig. 7. Effect of a bandlimiting channel. (a) APD, $A = 0.1$. (b) SD.

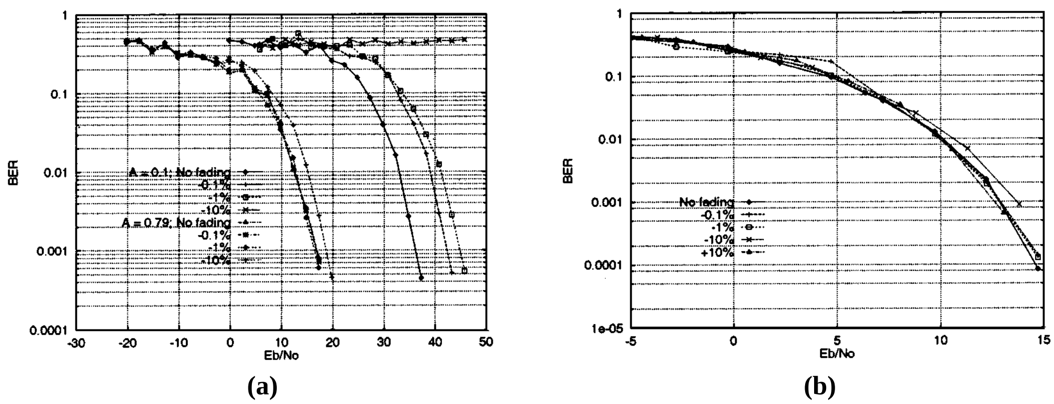

Fig. 8. Effect of Flat fading (expressed as a percentage). (a) APD. (b) SD.

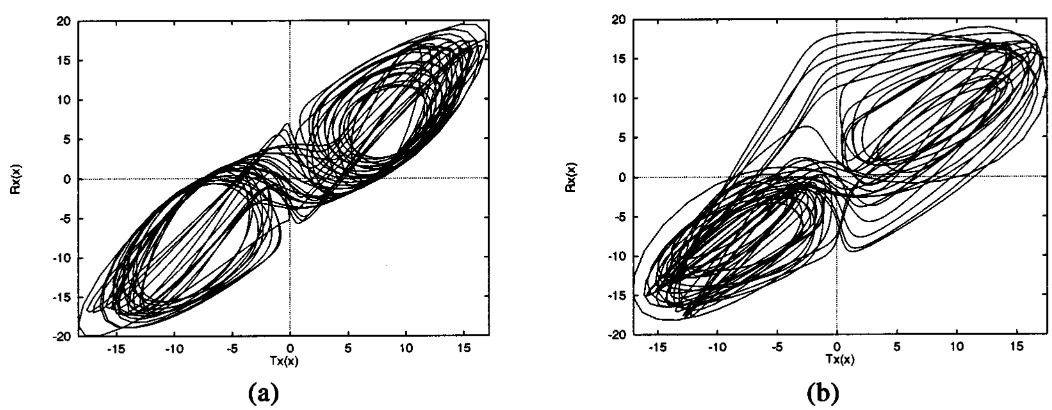

Fig. 9. Comparison of transmitter and receiver state variables (x) in synchronised receiver, two-ray multipath channel, channel tap amplitudes 1.0 and 0.4 (a) Delay = 100 ms. (b) Delay = 1000 ms.

but there will be some error in this process. The APD behavior is representative of IS systems and shows, as commented before, that as the magnitude of the distortion approaches the magnitude of the message, performance degrades. Also, the sign of the fading variation (gain or attenuation) makes little difference, in either case it is just perceived as distortion. In contrast, the SD method is not only more robust but can use the constructive interference with channel gain to maintain (even slightly improve) performance. This is not unreasonable, since in these cases, the reconstructed signal is stretched in phase space, increasing the distance noise must displace the signal to cause an error. It is more desirable to have systems that are independent to amplitude variations, the limitation for the systems just described is the range over which synchronization can be maintained. Generalized synchronization gives lower synchronization errors than IS over a limited range [4], alternatively amplitude independent

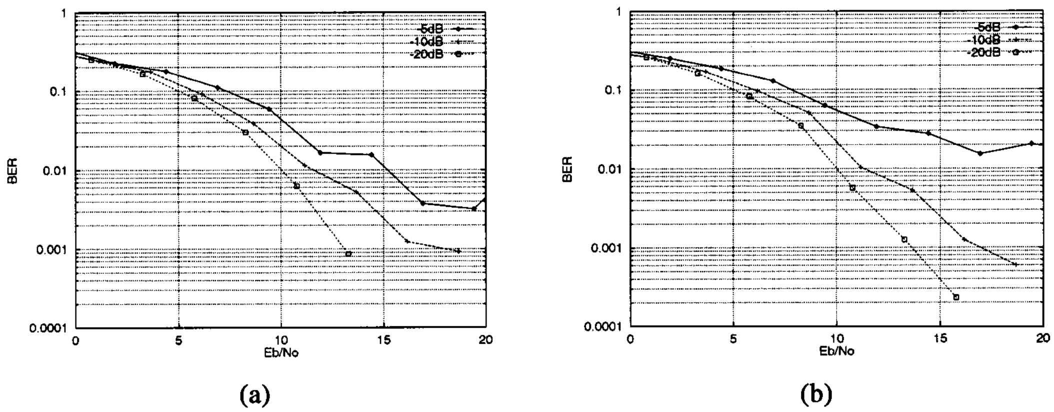

Fig. 10. Effect of Doppler fading two-path channel on SD versus tap amplitude, tap delay 100 ms. (a) Velocity $= 4$ ms^{-1}. (b) Velocity $= 20$ ms^{-1}.

synchronization is presented in [20]. Processing using delay embedding reconstructions are also amplitude independent.

The fast fading case uses multiple taps in the channel filter, each is multiplied by a time variant random signal with Rayleigh statistics (see Section III). The parameters for this model are the tap delays, mean tap power, and velocity. The total received power from all multipath components is included in the bit energy calculation, not just the largest. With so many degrees of freedom, some simplifications will help the understanding. The channel will be limited to just two taps, and the first tap will be the largest with unity amplitude, the channel is thus Ricean.

The effect of a number of channels for synchronized systems is demonstrated in Fig. 9. These show that increasing the mulitpath amplitude increases the reconstruction error, and the structures are changed. This structural change is more obvious for longer tap delays. With the delay-embedding approach, using sufficient dimensions is important to be able to recover the SD after corruption in a multipath channel. It has been shown that finite-impulse–response filters do not change the dimension of the received signal [21], [22], and so it is possible to embed the received signal, without needing to know about the channel (other than it is FIR). Finding the SD from the reconstruction may require the channel parameters to be estimated.

To investigate time-varying channels some rescaling is necessary, so that the time-varying behavior of the channel taps was meaningful. In what follows, the Lorenz system is considered as a baseband equivalent of a signal centred on 900 MHz, operating at a data rate of 10 kbps. Fig. 10 clearly shows performance degrading as the amplitude of the multipath component increases, the effect of velocity is less significant, having the greatest impact for the larger amplitude multipath case. Fig. 11 demonstrates the effect of tap delay. For short delays performance is degraded most because phase inversions of the multipath component will cause significant destructive interference, and so the signal has almost-flat fading. As the delay increases, there is initially less (anti-)correlation between main path and delayed path since maximum turning points do not coincide, so performance improves. As the delay becomes comparable with the average symbol period (0.75 s) again, then performance will degrade. This cyclic behavior will continue for a while, but for mulitpath delays more than a few symbols the variation in performance will decrease, this is because the symbol periods are irregular for the SD system.

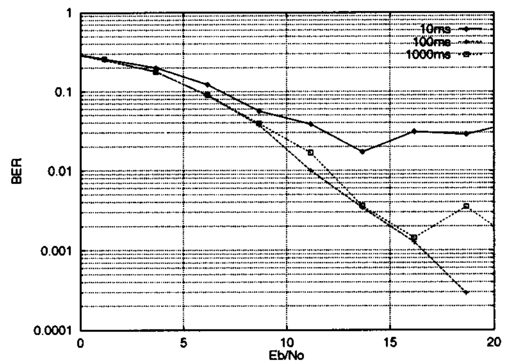

Fig. 11. Effect of Doppler fading two-path channel on SD versus tap delay. Velocity $= 10$ ms^{-1}, second tap amplitude $-$ 10 dB.

The most significant effect of multipath channels is dependent on the power in the multipath components. In IS systems any delayed components (whether causing constructive or destructive interference) are perceived as noise and degrade performance. In the SD case constructive interference does improve performance, though the combined effect of both types of interference is a degradation. Spread-spectrum techniques (discussed further in Section V) are commonly used because the wider bandwidth can improve robustness against multipath effects, however, this assumes knowledge of some reference signal at the receiver. Frequency hopped spread spectrum changes its centre frequency such that the signal does not remain in a fade for significant periods of time (so any losses can be recovered by coding). Alternatively, direct-sequence techniques use a spreading sequence, and correlation at the receiver such that individual multipath components can be resolved. This is clearly not the case here. One approach is to transmit the reference signal, but then noise on the reference will degrade performance. The worst performance degradation was demonstrated for small delay paths, this will have the effect of near-flat fading of the signal. For longer delay paths, relative to the symbol period, performance degraded. Consequently, there is not always a compelling reason to increase the bandwidth of these signals, other than to increase the data rate, and indeed it is more favorable to reduce the bandwidth and maintain bandwidth efficiency.

V. Towards Bandwidth Efficiency

Multiple-access mechanisms are the means by which a number of communication connections can share the available resources. It is important that the total system capacity is maximized, the recent spectrum auctions for 3G services show the real cost of such a resource. While some frequency bands are unlicensed, excess occupied bandwidth will expose the signal to greater interference. Multiple access mechanisms separate users based on allocating orthogonal resources. Frequency-division multiple access (FDMA) allocates a unique frequency channel (of specified bandwidth) to each user. Time-division multiple access (TDMA) divides the radio resource into time slots, each user is allocated time slots in which it can use the whole frequency band. In code-division multiple access (CDMA), signals are spread over a wider bandwidth than the information bandwidth by a coding mechanism independent of the message (so wideband frequency modulation is not included). The orthogonality of these code sequences allows separation of the users. The most common CDMA methods are direct sequence (DS/CDMA) in which the information is multiplied by a spreading sequence, and frequency hopping in which the centre frequency is regularly varied in a pseudo random nature within the channel. For CDMA orthogonality is rarely perfect (particularly in asynchronous and multipath environments), and as the number of users increases the level of interference increases. There is no absolute limit on the number of users accessing the channel (assuming sufficient codes are available, not usually a problem), but performance degrades and so these systems are interference limited. In contrast, TDMA and FDMA techniques do have a hard limit on the number of users that can be supported (number of channels or time slots).

From the above characterization, it could be assumed that chaotic systems fall within the CDMA type classification. However, for the methods being discussed in this paper (i.e., distinct from the DS/CDMA methods such as [2]), this is not the case because different users cannot share the same time and frequency resources, they must be separated by FDMA or TDMA mechanisms. Consequently, the bandwidth efficiency (defined as data rate divided by occupied frequency bandwidth) of many chaotic communication systems is poor. This section will discuss the bandwidth efficiency of chaotic communication systems, and propose how it can be increased.

The first issue to address is to define the bandwidth of a chaotic signal, this is not obvious since the true bandwidth is infinite. There are three approaches to be considered; cumulative distribution function (CDF) based, filtering method and an adjacent channel spacing method. Table I shows the bandwidth required to contain some defined percentage of the total power, For example 99% of the power is contained within 2.56 Hz. An alternative method is to apply a bandlimiting filter to the signal and measure the point at which performance (BER) degrades by some acceptable amount, this was demonstrated in Fig. 7, and a filter with a 1-Hz breakpoint gave good performance. However, the result will depend on the filter characteristics. In an FDMA environment, a more realistic measure is to reduce the channel spacing to an acceptable compromise between occupied bandwidth and performance. This is demonstrated in Fig. 12 (a 2-Hz preprocessing filter was also used), a channel bandwidth of 2.5 Hz is reasonable.

TABLE I
Cumulative Probability Distribution for Lorenz Power Spectrum (SD Encoding)

Cumulative probability	90%	95%	99%	99.9%
Frequency	1.42 Hz	1.86 Hz	2.56 Hz	3.8 Hz

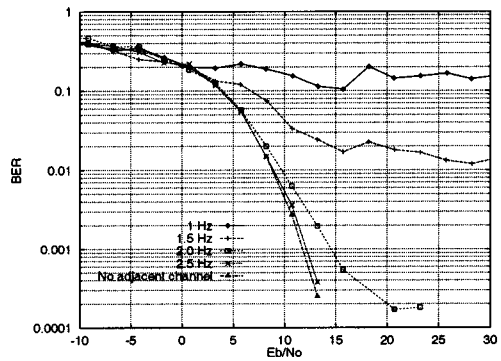

Fig. 12. Degradation in performance versus adjacent channel separation.

The limit on the maximum channel capacity was introduced by Shannon, and the usual limit is given by [17]

$$\frac{C}{W} = \log_2\left(1 + \frac{C}{W}\frac{E_b}{N_o}\right). \tag{6}$$

Where C is the channel capacity (bits/s) and W occupied bandwidth (hertz). The formulation for this equation is based on single channel communication with only AWGN at the output of the demodulator. Another important assumption is that Nyquist sampling theorem is appropriate.

All these assumptions must be questioned. In this work the single-channel assumption is reasonable, but elsewhere, removing this assumption has demonstrated large capacity increases [23]. The previous discussion on bandwidth of chaotic signals has shown that the conventional Nyquist sampling rate cannot apply. The published work on delay embedding demonstrates that a signal can indeed be reconstructed adequately despite the infinite bandwidth. For bandlimited signals, the $\mathrm{sinc}(x)$ function is used as the interpolation function, impulsive chaotic synchronization [24] has demonstrated similar interpolation properties for finite sampling rates. The ability to interpolate with finite sample rates is a consequence of being able to use prior information (the dynamics), whereas the Nyquist theorem assumes no prior knowledge other than the bandlimited restriction. In the case of the Lorenz system (with the specified parameter set) a sample period of 0.08 s was sufficient for synchronization [18], but there is no clear relationship between this and the bandwidth measures previously described.

Fig. 3 clearly shows that the noise in the decision process is non-Gaussian. To calculate the bit error rate a two dimensional

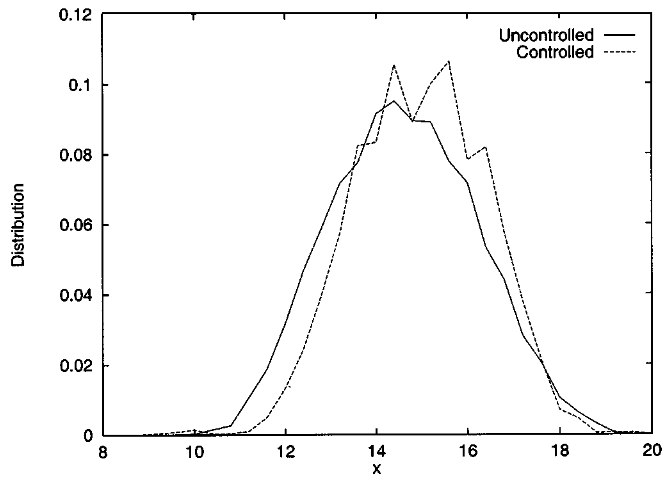

Fig. 13. Probability distribution of surface of section crossing points (x state variable), for uncontrolled and controlled (SD) systems.

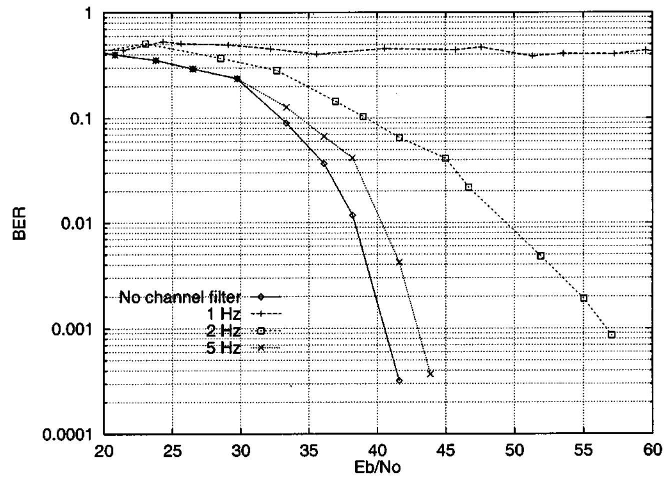

Fig. 15. Performance of the APD method, with bandlimiting channel and complementary filter in the receiver.

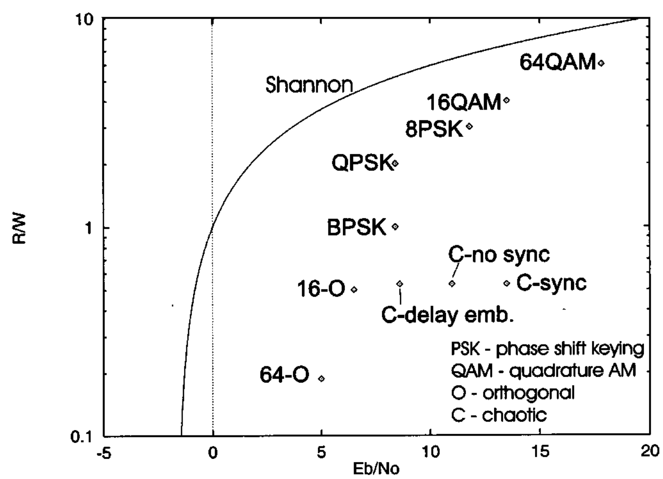

Fig. 14. Comparison of conventional and chaotic (SD method) communication systems to the Shannon limit.

integral needs to be computed (surface of section $x = y$, decision boundary $x = 0$)

$$\text{Probability of error} = \int_{\substack{y=x \\ x>0}} \int_{n=-\infty}^{0} P_N(x,n) P_X(x)\, dn\, dx. \tag{7}$$

Where $P_N(x,n)$ is the noise (n) probability distribution (Fig. 3) with offset of x (along the surface of section), and $P_X(x)$ is the section crossing distribution demonstrated in Fig. 13. It would be expected that compressing the range of section crossings would improve performance, this is demonstrated in [25]. The extreme case of only having a two valued section crossing function $(\pm X)$, is equivalent to a conventional binary communication system. The relative performance of conventional and chaotic communication systems compared to the Shannon limit are shown in Fig. 14 (at $\mathrm{BER} = 10^{-4}$). There is clearly some way to go before chaotic schemes rival the conventional. There are two directions to go, improve noise performance or bandwidth efficiency. Improving noise performance is discussed in the next section.

There are two basic approaches to improving bandwidth efficiency, reduce the occupied bandwidth or increase the number of symbols in the symbol set (thus sending more data per symbol). While it is desirable to generate chaotic signals with most of the energy limited to a narrow band of frequencies, this is not always possible. Consequently filtering the transmitted signal is required. This has already been demonstrated in Fig. 7. However, this introduces distortion into the chaotic carrier, and it is desirable to remove this distortion at the receiver. An inverse filter is not sufficient since this will amplify noise. The complementary filter method of Carroll [26] can indeed improve performance for the APD system, Fig. 15 shows improved performance for a 2 Hz filter whereas before (Fig. 7) this was not the case. This approach can destabilise the receiver (due to the complementary filter), and for this reason filters below 2 Hz were not usable. A disadvantage of this method, as pointed out by Rulkov [27], is the asymmetry between drive and response systems. Some alternative architectures are proposed in [27] which embed the transmit filter within the transmitter dynamics. This changes the chaotic dynamics, and for the cases tested with narrowband filters stable synchronization was not achieved.

The SD system investigated to date has only used a binary symbol set, increasing the number of symbols possible will increase bandwidth efficiency. Using higher dimensional chaotic systems will give the flexibility for this approach. Conventionally using higher order modulation has involved a trade-off between bandwidth efficiency and noise robustness, due to the limited number of signal dimensions (two). Higher dimensionality can be achieved by using more time slots, or more frequency channels, both methods decrease bandwidth efficiency. More general dynamical systems approaches do not suffer this limitation, however while the symbolic partitions give distinct signals, the concept of orthogonality does not apply in the same way. Consequently, the bandwidth efficiency and noise robustness trade-off is still likely to exist. Further work will investigate these techniques.

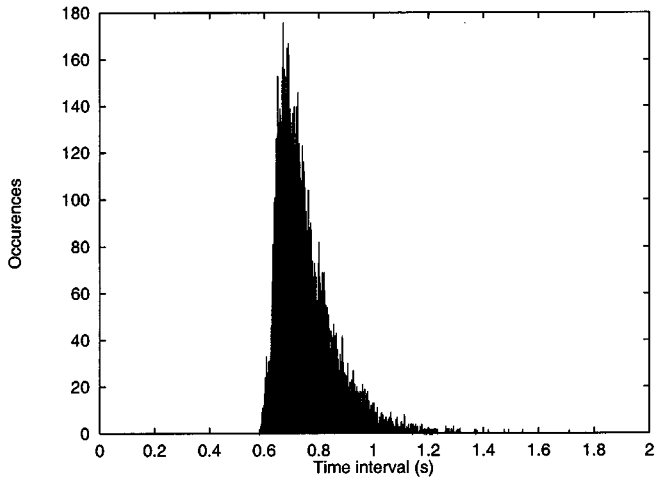

Fig. 16. Surface of section crossing time-interval distribution.

VI. Chaotic Communications in Practice

The previous sections have assumed some idealistic conditions, channel models have been simplified and perfect symbol synchronization was assumed. At the same time no performance enhancing measures have been taken, such as noise reduction or equalization (to compensate for the multipath channel). The masking and APD methods sent the data as an analogue signal (with values ± 1), the recovered signal can then be treated as a conventional demodulated binary sequence. Consequently, standard symbol synchronization techniques such as phase locked loops (tuned to the symbol rate), early-late trackers, etc., can be used [17]. In CSK, because the symbols are periodic and so the output error sequences will contain periodic components, similar techniques could also be used. The SD method requires a different approach because the symbol decision points do not occur regularly in time. In the simulations presented so far a decision "genie" was used that passed information from the transmitter to mark the decision points. In good channel conditions detection of section crossings would suffice, but as noise and distortion increases additional crossing may occur (insertion errors) or the section could be missed (deletion errors). When the error rate is not high a number of approaches can be used, these are; time interval bounding, nonlinear noise reduction and block marking. Time interval bounding relies on the fact that there is only a range of intersymbol intervals that can occur, and so any lying outside this range must show an error has occurred (Fig. 16). Nonlinear noise-reduction techniques have been widely researched [19], and can clean noisy chaotic time series. However, they mostly operate with high signal to noise ratios, where error rate would be low anyway, and there is a significant amount of processing required. Block marking inserts a unique word at regular intervals into the data sequence, knowing how many symbols are in a block allow most insertions and deletions to be detected (but not corrected). Pairs of insertions and deletions will not be detected (since the number of symbols is unchanged). There is clearly a reduction in data rate with this method. In practice some combination of these techniques could be used. It is clear that while nonsynchronised methods give better error performance, because they do not use the known dynamics at the receiver (and may be more coarsely sampled) a high signal to noise ratio may be required for symbol synchronization. A combination of techniques, using chaotic synchronization to estimate symbol detection points will be investigated.

In high noise and multipath environments these methods will be less effective. The optimum symbol detection surface will move and distort (no longer be a flat surface), this will be the case particularly for multipath channels with no (or insufficient) equalization. A maximum likelihood method for symbol detection (and synchronization) is described in [25]. The process of compensating for a multipath channel is known as equalization, the technique used depends on what prior information is available. When the channel is unknown, an adaptation mechanism is required, to control the equalizer, this may or may not involve estimating the channel parameters directly. The most simple equalizers use a linear pre-filter on the received signal, which may be FIR [moving average (MA)] or IIR [autoregressive (AR)]. When the channel itself is FIR then the AR-based equalizer can give perfect equalization when the channel is minimum phase (required to ensure stability). A FIR-based equalizer can only approximate the ideal equaliser in the limit of a large number of taps, but is guaranteed to be stable. Equalization of a chaotic signal when the channel is AR, with a FIR equalizer, is demonstrated in [28]. Designing the equalizer to give perfect equalization (zero forcing) often leads to poor performance in the presence of noise, because nulls in the spectrum lead to high gain regions in the equaliser frequency response. Consequently, least-mean-squares approaches are favored [29]. When the channel is time varying, the equaliser parameters need to be adaptive to follow the channel variations. Coarse equalization can be achieved using spectral estimation [30], and filtering the received signal such that the spectrum matches that from the transmitter (assumed known). This requires a minimum phase channel, and is susceptible to noise when there are spectral nulls as already described. When synchronization is used, the synchronization error can be employed to drive an optimization process [30], [31]. A gradient descent approach in [30] adapts the coefficients of an FIR linear equaliser to compensate for an ARMA channel. The difficulty in finding the local gradient is avoided in [31] where a downhill simplex method is used to estimate the channel, which are then used in a zero forcing AR equaliser.

VII. Conclusion

The use of communication methods requiring identical synchronization show poor robustness to channel noise and distortions. Robustness is improved by increasing the amplitude of the message (or parameter shift), this moves away from the small signal control ideal. Encoding by symbolic dynamics, does not require IS and has been shown to be more robust, in effect the message is equivalent in size to the chaotic carrier, but is controlled by small perturbations. Under the assumption of perfect symbol-timing information, nonsynchronized processing techniques were shown to outperform their synchronised counterparts. However, to acquire timing information, and to give error estimates to an equalizer, the chaotic synchronization mechanism are desirable, and hybrid approaches may be useful. In multipath channels, a narrowband signal will be subject to flat fading, which (if sufficiently slow) could be tracked. Wideband signals showed a degradation in performance that could be compensated by an equaliser or using a noise free reference. Considering the systems described here as spread spectrum is unhelpful, the bandwidth efficiency is degraded when occupying larger bandwidths which is not recovered using a CDMA mechanism. Current work is integrating symbol timing, equalization and noise reduction into a single system. Improving the bandwidth efficiency by exploring higher dimensional state spaces is also being investigated.

REFERENCES

[1] K. M. Short, "Unmasking a modulated chaotic communication scheme," *Int. J. Bifurcation Chaos*, vol. 6, no. 2, pp. 367–375, 1996.

[2] G. Mazzini, G. Setti, and R. Rovatti, "Chaotic complex spreading sequences for asynchronous CDMA-part 1: System, modeling, and results," *IEEE Trans. Circuits Syst. I*, vol. 44, pp. 937–947, Oct. 1997.

[3] G. Kolumban, M. P. Kennedy, G. Kis, and Z. Jako, "FM-DCSK: A novel method for chaotic communications," in *Proc. ISCAS'98*, vol. IV, May 1998, pp. 477–480.

[4] C. Williams, "Robust chaotic communication system based on generalized synchronization," in *Proc. ICECS'98*, vol. 1, Sept. 1998, pp. 243–246.

[5] ——, "Robust chaotic synchronization for communications," in *Proc. ECCTD'99*, vol. 1, Sept. 1999, pp. 441–444.

[6] N. F. Rulkov *et al.*, "Generalized synchronization of chaos in directionally coupled chaotic systems," *Phys. Rev. E*, vol. 52, no. 2, pp. 980–994, 1992.

[7] U. Parlitz *et al.*, "Experimental observation of phase synchronization," *Phys. Rev. E*, vol. 54, no. 2, pp. 2115–2117, 1996.

[8] L. M. Pecora and T. L. Carroll, "Synchronization in chaotic systems," *Phys. Rev. Lett.*, vol. 64, no. 8, pp. 821–824, 1990.

[9] U. Parlitz *et al.*, "Encoding messages using chaotic synchronization," *Phys. Rev. E*, vol. 53, no. 5, pp. 4351–4361, 1996.

[10] K. Pyragas, "Predictable chaos in slightly perturbed unpredictable chaotic systems," *Phys. Lett. A*, vol. 181, pp. 203–210, 1993.

[11] H. Nijmeijer and I. M. Y. Mareels, "An observer looks at synchronization," *IEEE Trans. Circuits Syst. I*, vol. 44, pp. 882–890, Oct. 1997.

[12] K. M. Cuomo, A. V. Oppenheim, and S. H. Strogatz, "Synchronization of Lorenz based chaotic circuits with applications to communications," *IEEE Trans. Circuits Syst I.*, vol. 40, pp. 626–633, Oct. 1993.

[13] K. S. Halle *et al.*, "Spread spectrum communication through modulation of chaos," *Int. J. Bifurcation Chaos*, vol. 3, no. 2, pp. 469–477, 1993.

[14] S. Hayes, C. Grebogi, and E. Ott, "Communicating with chaos," *Phys. Rev. Lett.*, vol. 70, no. 20, pp. 3031–3034, 1993.

[15] T. S. Rappaport, *Wireless Communications: Principles and Practice*. Englewood Cliffs, NJ: Prentice-Hall, 1996.

[16] J. D. Parsons and J. G. Gardiner, *Mobile Communication Systems*: Blackie, 1989.

[17] J. G. Proakis, *Digital Communications*, 2nd ed. New York: McGraw-Hill, 1989.

[18] C. Williams, "Chaotic Synchronization in Wideband Communication Systems," PhD. dissertation, Dept. Elect. Eng., Bristol Univ., Bristol, U.K., 1997.

[19] H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1997.

[20] T. L. Carroll, "Amplitude independent synchronization," *Phys. Rev. E.*, vol. 53, pp. 3117–3122, 1996.

[21] D. S. Broomhead, J. P. Huke, and M. R. Muldoon, "Linear filters and nonlinear systems," *J. R. Stat. Soc. B*, vol. 54, no. 2, pp. 373–382, 1992.

[22] S. H. Isabelle, A. V. Oppenheim, and G. W. Wornell, "Effects of convolution on chaotic signals," in *Proc. ICASSP'92*, vol. IV, 1992, pp. 133–136.

[23] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Commun.*, vol. 6, pp. 311–335, 1998.

[24] R. E. Amritkar and N. Gupte, "Synchronization of chaotic orbits: The effect of a finite time step," *Phys. Rev. E*, vol. 47, no. 6, pp. 3889–3895, 1993.

[25] M. Carroll and C. Williams, "Improving chaotic communication by symbolic dynamics," , 2001, submitted for publication.

[26] T. L. Carroll, "Communicating with the use of filtered, synchronised, chaotic signals," *IEEE Trans. Circuits Syst. I*, vol. 42, pp. 105–110, Mar. 1995.

[27] N. F. Rulkov and L. S. Tsimring, Communication with chaos over band-limited channels, , 1997. pre-print.

[28] H. Leung, "System identification using chaos with application to equalization of a chaotic modulation system," *IEEE Trans. Circuits Syst. I*, vol. 45, pp. 314–320, Mar. 1998.

[29] S. H. Qureshi, "Adaptive equalization," *Proc. IEEE*, vol. 73, pp. 1349–1387, Sept. 1985.

[30] K. M. Kuomo, A. V. Openheim, and R. J. Barron, "Channel equalization for self-synchronising chaotic systems," in *Proc. ICASSP'96*, 1996, pp. 1605–1608.

[31] N. Sharma and E. Ott, "Combating channel distortions in communication with chaotic systems," *Phys. Lett. A*, vol. 248, pp. 347–352, 1998.

Chris Williams, photograph and biography not available at the time of publication.