Information Theory Aspects of Spread-Spectrum Communications

James L. Massey
Signal & Info. Proc. Lab., Swiss Federal Inst. of Tech., Ch-8092 Zürich

Abstract–A novel definition of a spread-spectrum system as a communications system in which the Fourier bandwidth is much greater than the Shannon bandwith (the number of dimensions of signal space used per second) is proposed. Six different communication systems are analyzed in terms of this definition. It is shown that there is a fundamental difference between the bandwidth expansion due to coding and due to "spectrum spreading". It is further shown that spectrum spreading plays no role in increasing the capacity of the system, but can perform other useful roles such as providing low probability of interception of the signal, good electromagnetic compatibility, and a multiple-access capability.

I. INTRODUCTION

The main purpose of this paper is to consider spread-spectrum systems from the fundamental viewpoint of Shannon's information theory [1]. To do this requires that we carefully define what we mean by a spread-spectrum system. This is done in Section II in which we give a rather unconventional definition of a spread-spectrum system, but the only one that we were able to formulate that we ourselves found to be satisfactory. To illustrate the implications of this definition, we consider six different communication systems in Secion III to see which qualify (under our definition) to be called spread-spectrum systems. In Section IV, we consider various reasons why one might wish to use a spread-spectrum system. In Section V, we make a more strictly information-theoretic investigation of spread spectrum systems with one sender where we show that spreading spectrum can never increase capacity but also that it need not decrease capacity significantly. In Section VI, we make some concluding remarks as to what more needs to be done before we have a true information theory of spread-spectrum systems that will allow one to make sound practical judgements and choices on the basis of the theory.

Throughout this paper, we have limited ourselves for simplicity to baseband signals, but the reader should have no difficulty in adapting our approach to passband signals.

II. WHAT IS A SPREAD-SPECTRUM SYSTEM?

In his brilliant treatise [1] that established the field, Shannon called information theory the "mathematical theory of communication". We have often maintained that, in a very real sense, mathematics is definitions. Once the definitions are in place, all the lemmas, theorems and corollaries are determined; one has only to find them and prove them. If we wish to say something about the information theory of spread-spectrum systems, it follows that our unavoidable first task must be to define such systems. This task may well strike the reader as either superfluous or quixotic. Like the U.S. supreme court justice who admitted the difficulty of defining pornography but claimed that he knew it when he saw it, many communication engineers might maintain that a definition is not needed; they know a spread-spectrum system when they see it. One such friend described a spread-spectrum communication system to us as "one that uses much more bandwidth than it needs". There seems to be a certain coarse truth in this description, but it will hardly do for mathematical purposes. After some futile attempts to make this description more precise, our friend concluded that a clear definition of a spread-spectrum system is not possible, which whetted our appetite to make a stab at one.

Every communication engineer is familiar with the ordinary notion of bandwidth, which we will call *Fourier bandwidth* both to honor the French pioneer in this field and to distinguish it from a less familiar but no less important type of bandwidth. The "sinc pulse" m(t) = sinc (2Wt), where sinc (x) = sin(π x)/(π x) , has a Fourier Bandwidth of W Hz, as one sees immediately from its Fourier transform M(f) shown in Fig. 1. For less dichotomous spectra, there are many options for calculating the precise Fourier bandwidth (rms bandwidth, 3 dB bandwidth, 99% energy bandwidth, etc.), but they are all roughly equivalent and any is good enough for our purposes. The notion of Fourier bandwidth extends easily from deterministic signals to stochastic processes (such as modulated signals) in a way familiar to all communication engineers.

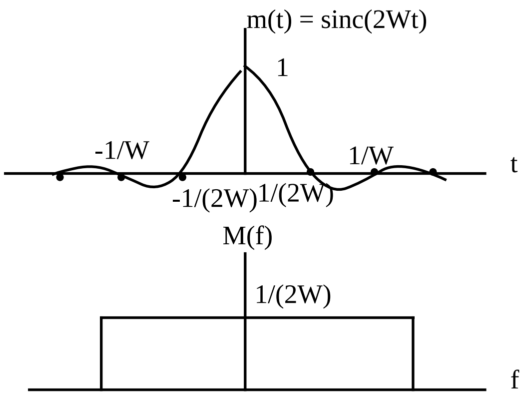

Fig. 1: The Sinc pulse m(t) = sinc(2Wt) and its Fourier Transform M(f)

The second type of bandwidth, which we will call *Shannon bandwidth* because Shannon [2] was the first to appreciate its importance, cannot even be defined for a deterministic signal, i.e., for a single time function. One must consider a "variable" signal (or a stochastic process) such as a modulated signal s(t) that can take on any of a multiplicity of time functions as its value. Then one must consider a signal-space representation of s(t) over some long time interval, say the interval $0 \le t < T$. By this we mean that one must find orthonormal functions $\phi_i(t)$, i = 1, 2, .., N, so that one can represent (or very well approximate) s(t) as

$$s(t) = \sum_{i=1}^{N} s_i \, \phi_i(t) \tag{1}$$

for $0 \le t < T$. One says then that one has a signal-space representation of s(t) as a vector $\mathbf{s} = (s_1, s_2, \ldots, s_N)$ in N-dimensional Euclidean space. When one does this in such a way as to minimize the dimensionality N of the representation, then he has arrived at the *Shannon bandwidth* B, which we now define as

$$B = \frac{1}{2}\frac{N}{T} \text{ (dim/sec).} \tag{2}$$

Equivalently, the Shannon bandwidth is *one-half the minimum number of dimensions per second required to represent the modulated signal in a signal space.* [In earlier papers [3], [4], where we used the notion of Shannon bandwidth, we omitted the division by 2 in (2). Emboldened by Emerson's dictum that "a foolish consistency is the hobgoblin of small minds", we now opt to insert the 2 in (2) in order to avoid many factors of 2 elsewhere.]

We now state what might be called *the fundamental theorem of bandwidth.*

> The Shannon bandwidth B of a modulated signal is at most equal to its Fourier bandwidth W; (rough) equality holds when the orthonormal functions are $\phi_i(t) = \sqrt{2W}$ sinc (2Wt - i) (or any functions whose spectra are nearly flat for $-W < f < W$ and nearly zero elsewhere).

There are many proofs of this theorem, starting with that of Shannon [2]; essentially one shows that one can construct 2WT orthonormal functions of Fourier bandwidth W or less that are confined within the time interval $0 \le t < T$ when $WT >> 1$, and that one can construct no more than this. See the insightful book of Wozencraft and Jacobs [6] and the penetrating paper of Slepian [7] for further discussion of this theorem.

We are now ready to offer our *definition of a spread-spectrum system as a communication system in which the modulated signal has a Fourier bandwidth substantially greater than its Shannon bandwidth.* If one considers the Shannon bandwidth to be the amount of bandwidth that the system *needs* (and we will offer arguments to this effect later) and the Fourier bandwidth to be the amount of bandwidth that the system *uses*, then we are back at our friend's pithy characterization of a spread-spectrum communication system as "one that uses much more bandwidth than it needs".

It is an obvious next step to define the *spreading factor*, γ, of a communication system as the ratio of the Fourier bandwidth of the modulated signal to its Shannon bandwidth, i.e.,

$$\gamma = W/B. \tag{3}$$

For every communication system, $\gamma \ge 1$. A spread-spectrum system is a communication system with "large" γ, say $\gamma \ge 5$, but of course the precise dividing line between a spread-spectrum system and an unspread system is quite arbitrary.

We now have our definitions. It remains to show that they make sense, i.e., that they lead to interesting and useful conclusions.

III. SOME EXAMPLES AND THEIR LESSONS

In digital communication systems (to which we will mostly confine our discussion), the modulated signal in an interval $0 \le t < T$ can assume only finitely many values. In this case, one can in principle always apply the familiar Gram-Schmidt orthogonalization technique [8, p. 277] to these signals to obtain an orthonormal basis $\phi_1(t), \phi_2(t), \ldots, \phi_N(t)$ for the signal space of smallest dimension N containing these signals and can thus determine the Shannon bandwidth according to (2). In most cases of practical interest (as the following examples will illustrate), however, one can find such a basis (and hence find N) by inspection. For an analog system, the modulated signal must generally be treated as a stochastic process. In this case, one can use the Karhunen-Loéve expansion [9, pp. 96-99] to determine N, which will be the number of orthonormal functions in the expansion that have coefficients of non-negligible energy.

We now give several examples of modulated signals, whose analysis will give insight into our definition of a spread-spectrum system.

Example 1: The modulated signal corresponds to that for one of K users in a time-division multiple-access system in which each user sends L data symbols during each TDMA frame of duration T seconds. Choosing sinc pulses to make the Fourier bandwidth unambiguous, we can write the selected user's modulated signal as

$$s(t) = \sum_{i=1}^{n} b_i \operatorname{sinc}\left(\frac{KL}{T} t - i\right), \quad 0 \le t < T \tag{4}$$

where $b_1, b_2, \ldots, b_L$ are the data symbols. We see that the Fourier bandwidth W must satisfy 2W = KL/T and hence that

$$W = \frac{KL}{2T}. \tag{5}$$

Example 2: The modulated signal corresponds to that for one user in a code-division multiple-access (CDMA) system in which each user modulates a user-specific binary (± 1) code sequence of length L with one data symbol in each symbol period of duration T. We can write the selected user's modulated signal as

$$s(t) = b_1 \sum_{i=1}^{L} a_i \,\mathrm{sinc}\left(\frac{L}{T} t - i\right), \quad 0 \le t < T \tag{6}$$

where b_1 is the data symbol and where $(a_1, a_2, ..., a_L)$ is the binary (± 1) code sequence. The Fourier bandwidth W satisfies $2W = L/T$ and hence

$$W = \frac{L}{2T}. \tag{7}$$

Example 3: A user sends an M-ary pulse-position modulation signal in each T second interval, i.e.,

$$s(t) = A \,\mathrm{sinc}\left(\frac{M}{T} t - b_1\right), \quad 0 \le t < T \tag{8}$$

where $b_1 \in \{1, 2, ..., M\}$ is the single data symbol and where A is some fixed amplitude. The Fourier bandwidth satisfies $2W = M/T$ and hence

$$W = \frac{M}{2T}. \tag{9}$$

Example 4: A user employs binary antipodal signalling to transmit random binary (± 1) data symbols in such a manner as to send n such symbols in each T second interval, i.e.,

$$s(t) = \sum_{i=1}^{n} b_i \,\mathrm{sinc}\left(\frac{n}{T} t - i\right), \quad 0 \le t < T. \tag{10}$$

The Fourier bandwidth satisfies $2W = n/T$ and hence

$$W = \frac{n}{2T}. \tag{11}$$

Example 5: Same as example 4 except that now the "data symbols" are the output of a powerful rate $R = 1/n$ (bits/symbol) trellis encoder fed by random "information bits". Equations (10) and (11) apply unchanged.

Example 6: Same as example 5 except that the code is a trivial $R = 1/n$ code with two binary (± 1) codewords, $(a_1, a_2, ..., a_n)$ and its negative. Then

$$s(t) = b_1 \sum_{i=1}^{n} a_i \,\mathrm{sinc}\left(\frac{n}{T} t - i\right), \quad 0 \le t < T \tag{12}$$

where b_1 is the information bit encoded. The Fourier bandwidth is

$$W = \frac{n}{2T}. \tag{13}$$

We now consider which of the above six systems are in fact spread-spectrum systems (by our definition). The task reduces essentially to finding the Shannon bandwidth B of each system.

Example 1 (concluded): By inspection of (4), we see that the signal space is minimally spanned by the orthonormal functions $\phi_i(t) = \sqrt{2W}\,\mathrm{sinc}\,(2Wt - i)$, i = 1, 2, ..., L, where W is given by (5). Thus, its dimension is N = L so that (2) gives the Shannon bandwidth as $B = L/(2T)$. The spreading factor according to (3) is then just $\gamma = K$. This TDMA system is indeed a spread-spectrum system when the number K of users is large.

Example 2 (concluded): By inspection of (6) and because $a_1, a_2, ..., a_L$ are fixed, we see that the signal special is one-dimensional, i.e., N = 1. Thus the Shannon bandwidth is only $B = 1/(2T)$ and the spreading factor is $\gamma = L$. This CDMA system is a spread-spectrum system when L is large (and indeed our definition of a spread-spectrum system would be indefensible if it were not).

Example 3 (concluded): From (8), we see that the signal space is minimally spanned by the orthonormal signals $\sqrt{2W}$ sinc (2Wt - i) for i = 1, 2, ..., M, where W is given by (9). Thus N = M and the Shannon bandwidth is $B = M/(2T)$. The spreading factor is thus $\gamma = 1$, the minimum possible. It follows that such an M-ary PPM communication system is *never a spread-spectrum system*, even though it utilizes a *very large Fourier bandwidth* according to (9), when M is large, to send at most $\log_2 M$ bits of information every T seconds.

Example 4 (concluded): From (19), we see that the signal space is minimally spanned by the orthonormal signals $\sqrt{2W}$ sinc (2Wt - i) for i = 1, 2, ..., n, where W is given by (11). Thus, N = n, $B = n/(2T)$ and $\gamma = 1$. Such binary antipodal modulation, not surprisingly, is never a spread-spectrum system.

Example 5 (concluded): For virtually any nontrivial trellis coding system, the encoded symbols $b_1, b_2, ..., b_n$ that appear in (10) will take on such a variety of different possible binary (± 1) patterns that one cannot imbed the set of possible s(t) in a smaller dimensional space than is required when all n binary symbols can be independently chosen (even though the code, with the mapping $0 \to +1$ and $1 \to -1$ may be linear over some finite field). Thus both the Fourier and Shannon bandwidths are the same as for the uncoded system of example 4 and $\gamma = 1$. A non-trivially coded binary antipodal modulation system is *never a spread-spectrum system*, even though, when the code rate $R = 1/n$ is very low, it follows from (11) that it utilizes *a very large Fourier bandwidth* to send one bit of information every T seconds.

Example 6 (concluded): Because the trivial code consists only of the two codewords $(a_1, a_2, ..., a_n)$ and its negative, we see from (12) that the signal space has collapsed to a one-dimensional space, i.e., N = 1. The Shannon bandwidth is thus $B = 1/(2T)$ and hence $\gamma = n$. Trivial coding of low rate $R = 1/n$ restricts the output of the binary-antipodal modulator in such a way that the system becomes a spread-spectrum one. In fact, comparing (6) and (12) we see that such trivial coding gives us precisely the same modulated signal as in the CDMA system of Example 2.

If one accepts our definition of a spread-spectrum system, then the above examples allow us to draw the following conclusions:

- From examples 3, we see that *a large ratio of Fourier bandwidth to data rate does* ***not*** *imply that a system is spread-spectrum.*

- Example 5 teaches us the, perhaps surprising, lesson that the fact that a coding system expands Fourier bandwidth by a large factor n for a fixed information bit rate does *not* imply that the system is spread-spectrum.

- *The Fourier bandwidth expansion due to nontrivial coding is fundamentally different from the kind of Fourier bandwidth increase that produces a spread-spectrum system,* although example 6 shows that trivial coding can indeed produce this latter type of expansion.

It is common when considering coded CDMA systems for communication engineers to speak of doing part of the spectrum spreading with coding and part with direct-sequence multiplication–we have not infrequently so spoken ourselves. But we now see that such statements are nonsensical. These are not two parts of the overall spectrum spreading, but rather two very different kinds of Fourier bandwidth expansion. Much of the remainder of this paper will be devoted to investigating the reasons that one might wish to do one or both of these kinds of bandwidth expansion.

IV. WHY SPREAD-SPECTRUM?

The original motivation for spread-spectrum systems was a military one, viz., to hide from an enemy the very fact that one is transmitting a signal. Today one speaks of the *low probability of interception* (LPI) of a spread-spectrum signal. The argument that spectrum spreading should provide the possibility for achieving LPI goes as follows. If the signal is confined to a small number N of dimensions within the global signal space of dimension $2WT = N\gamma$ in which all signals of bandwidth W and time-limited to $0 \le t < T$ must lie, and if there are parameters of the signal that can be varied to create a very large number of possible choices for the N-dimensional signal space occupied by the signal, then one can achieve LPI by selecting the value of these parameters at random. [We ignore here the role of the signal power and the intensity of the enemy's receiver noise, which determine essentially how long it takes to search a given number of dimensions simultaneously for the presence of signal.]

For the CDMA signal of example 2, there are 2^L possible choices of the binary (± 1) parameters $a_1, a_2, ..., a_L$, but changing the sign of all parameters leaves the signal in the same one-dimensional signal space. Fixing $a_1 = +1$ leaves us with 2^{L-1} choices of $a_2, a_3, ..., a_L$, each of which gives a different one-dimensional signal space. For large L, the enemy's task of finding the signal space actually used by the sender is thus akin to "looking for a needle in a haystack". A CDMA signal with a large number L of "chips" per symbol period does indeed afford low probability of interception.

For the TDMA signal of example 1, the only parameter that can be varied is the choice of the L consecutive symbol periods (out of the total of KL such periods) in which data symbols will be transmitted. There are only KL possible choices so that a low probability of interception can be achieved only if the product $KL = \gamma L$ is very large. [Spies in World War II frequently used this method to hide their transmission; they transmitted Morse code for only a few seconds and then went silent for long periods.] The point is that *a large spreading factor* γ *alone is not enough to provide LPI capability*. To put it another way, a TDMA signal is much easier to intercept than a CDMA signal.

The twin brother of low probability of interception is *electromagnetic compatibility* (EMC). If it is hard to determine whether a signal is present, then that signal cannot be interfering substantially with other commonly present signals. The excellent EMC capability of a CDMA signal is perhaps the strongest argument that we have today for preferring it to a TDMA signal. [We leave to the reader the task of showing the EMC superiority of a CDMA signal over a frequency-division multiple-access (FDMA) signal.]

The first cousin of low probability of interception is small *inter-user interference* (IUI), which is the prerequisite for a good multiple-accessing capability. If it is difficult to detect the presence of a signal, then shouldn't two such signals hardly interfere with one another? The answer is "yes, at least in a statistical sense!". If the two such signals are selected independently at random, then the probability of substantial IUI will be small. But care must be taken when the total number of signals is large, or when users persist for a long time in using the signal determined by one random choice of parameters. Of course, the K users of a TDMA system with signals as in example 1 will experience zero IUI when they are well synchronized. But K users (say, K on the order of L) of a CDMA system with signals as in example 2 and with the signal parameters frequently varied will experience IUI with roughly the same statistics no matter whether they are well synchronized or not. This robustness of a CDMA system with regard to inter-user interference is, of course, an important practical consideration.

V. CODING, SPREADING AND NOISE

It is time now to take a more strictly information-theoretic look at the advantages, if any, provided by spectrum spreading. To keep matters simple, we consider the single user system shown in Fig. 2. Of fundamental interest are (1) the *information rate*, R , measured in information bits (i.e., random binary digits) per second at the modulator input; (2) the *capacity*, C, also measured in information bits per second, of the channel created by the modulator and the band-limited additive white Gaussian noise (AWGN) waveform channel, which is the upper limit of rates R for which reliable (in the sense of arbitrarily small positive probability of error) recovery of the information bits is possible at the receiver when an appropriate coding system is used; (3) the *average power*, S, of the modulated signal; (4) the *one-sided noise* *power spectral density*, N_o, of the AWGN; (5) the *Fourier bandwidth*, W, of the bandlimited AWGN waveform channel (which we take without loss of essential generality as equal to the Fourier bandwidth of the modulated signal, as there is no point in transmitting anything outside this band and, if one transmits in a smaller bandwidth, then one might as well reduce the channel bandwidth accordingly); and finally (6) the *Shannon bandwidth*, B, of the modulated signal.

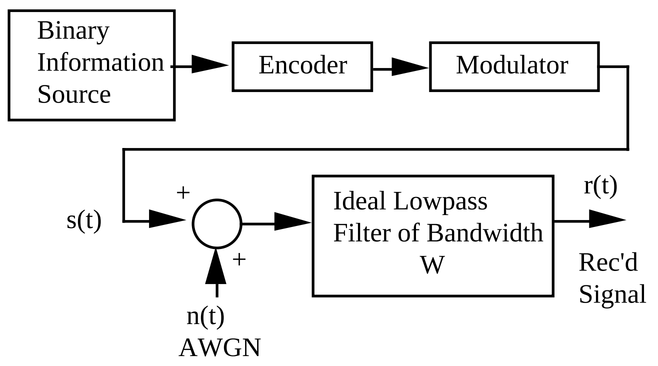

Fig. 2: The single-user communication system under study.

Shannon [1], with his penchant for getting to the heart of the matter, has given us the key relationship among these quantities, namely,

$$C = B \log_2 \left(1 + \frac{S}{N_o B}\right) \quad \text{(bits/sec).} \qquad (14)$$

The reader may be surprised to see the Shannon bandwidth B rather than the Fourier bandwidth W in this expression for C. but he or she will find that (14) is precisely the equation that Shannon derives in [1]. It is easy to check that the right side of (14) increases monotonically with increasing B; because B ≤ W, it follows that

$$C \leq W \log_2 \left(1 + \frac{S}{N_o W}\right) \quad \text{(bits/sec)} \qquad (15)$$

with equality if and only if B = W, i.e., if and only if there is no spectrum spreading! The reason that Shannon wrote (15) with an equality sign, rather than (14), in his final expression for the capacity of the AWGN channel is that he assumed that the choice of the modulator was up to the sender and that thus the sender would choose a modulator with B = W to obtain (maximum) capacity.

Here we must in honesty point out that we have been somewhat cavalier in writing (14) without stating the precise condition for which this expression gives the true capacity. This condition is that all the coefficients $s_1, s_2, \ldots s_N$ in the expansion (1) must be independently controllable by the choice of the modulator input symbols. This is indeed the case for all of the signals in the above six examples with the exception of example 3, PPM modulation. We see from (8) that in fact only one of the N = M coefficients can be non-zero so they are certainly not independently controllable. For such modulation systems, the expression in (14) gives only an upper bound on capacity–which is why PPM modulation is not "energy efficient" except for high "signal-to-noise ratios".

It is important not to draw the wrong conclusion from (14) and (15). The real question is not whether spectrum spreading can increase capacity (it never can!) but whether spectrum spreading, which may be desirable for other reasons such as those considered in the previous section, necessarily entails a substantial loss of capacity for the used Fourier bandwidth W. This time the answer is more subtle and more interesting. As B increases without limit, the right side of (14) tends to $1.44\, S/N_o$. Thus,

$$C \leq 1.44 \frac{S}{N_o} \quad \text{(bits/sec)} \qquad (16)$$

with near equality when the Shannon bandwidth B is large, say, when

$$B \geq 4 \frac{S}{N_o} \qquad (17)$$

since B = 4 S/No in the right side of (14) gives about 90% of the capacity for B = ∞. *As long as the Shannon bandwidth is large enough to satisfy* (17), *then no matter how large a spreading factor is used, the capacity will be at least 90% of the maximum capacity achievable with the used Fourier bandwidth W.* Spectrum spreading cannot increase capacity, but it need not reduce it.

The Shannon bandwidth B is always proportional to what we will call the *data rate*, R_m, at the input to the modulator, which we define as the number of modulation symbols per second that are input to the demodulator. But the information bit rate, R, *cannot exceed the capacity*, C, of the channel if the system is to operate reliably, and C, in turn must satisfy (16). How then with this fixed upper bound on R can we always achieve the necessary large Shannon bandwidth B? The answer is to use a code with sufficiently low *code rate* R, measured in modulation symbols per information bit, because $R = R_m \cdot \mathrm{R}$ and hence

$$R_m = \frac{R}{\mathrm{R}} \quad \text{(symbols/sec).} \qquad (18)$$

Because the Shannon bandwidth is proportional to R_m, we see that *coding increases the Shannon bandwidth by a factor proportional to the reciprocal of the code rate* R. This is the true nature of the "bandwidth expansion" due to coding. *Spectrum spreading*, on the other hand, *causes the Fourier bandwidth to be much greater than the Shannon bandwidth*, which is the true nature of its "bandwidth expansion".

VI. CONCLUDING REMARKS

It should be apparent that we have in this paper barely scratched the surface of the information theory of spread-spectrum systems. At best, we have pointed out the starting direction for a long journey. The important next step would be to consider spread-spectrum multiple-access systems, i.e., multiple-access systems in which each of several users sends a spread-spectrum signal in the same band and the sum of these signals is received. It is hardly a guess that a very interesting quantity will be the Shannon bandwidth of this sum signal, which incidentally will not in general itself be a spread-spectrum signal. (For instance, for a K-user TDMA system with signals as in example 1 and perfect synchronization, this sum signal will have a spreading factor of precisely 1.) Consideration of the Shannon bandwidth B and the Fourier bandwidth W of this sum signal has already permitted us to resolve certain CDMA "paradoxes" such as the apparent increase in system capacity for a two-user system with rectangular modulation pulses when the two users are *not* in synch. We can also use information-theoretic arguments to explain the apparent increase in system capacity when the two users are in synch but, because of imperfect power control, have slightly different received powers. However, we are far from being able to offer a coherent information-theoretic treatment for spread-spectrum multiple-access systems, even when we restrict the channel to be the bandlimited additive white Gaussian channel for the sum signal. And we have not even begun to scratch the surface for considerations of paramount practical interest such as multipath propagation of each signal in the sum, time variation of the multipath channels, unequal user signal powers, and imperfect synchronization. Nonetheless, it is our conviction that until the information theory of spread-spectrum systems is worked out in enough generality to deal with such issues, the many arguments over which type of multiple-access system is better (say, offers greater "spectral efficiency") than another will continue to generate more heat than light.

ACKNOWLEDGEMENT

This paper profited materially from our discussions with Prof. I. Bar-David of the Technion–Israel Institute of Technology.

REFERENCES

[1] C. E. Shannon, "A Mathematical Theory of Communications," Bell System Tech. J., vol. 27, pp. 379-423 and 623-656, July and October, 1948.
[2] C. E. Shannon, "Communication in the Presence of Noise," *Proc. IRE,* vol. 37, pp. 10-21, Jan. 1949.
[3] J. L. Massey, "Coding and Modulation for Code-Division Multiple Accessing", in *Proc. Third Int. Workshop on Digital Signal Processing Techniques Applied to Space Communications*, ESTEC, Noordwijk, 1992, pp. 3.1-3.17.
[4] J. L. Massey, "Deep-Space Communications and Coding: A Marriage Made in Heaven," in *Advanced Methods for Satellite and Deep Space Communications* (Ed. J. Hagenauer), Lecture Notes in Control and Information Sciences No. 182. Heidelberg and New York: Springer 1992, pp. 1-17.
[5] R. W. Emerson, "Self Reliance, in *Essays*, 1841.
[6] J. M. Wozencraft and I. M. Jacobs, *Principles of Communication Engineering*. New York: Wiley, 1965.
[7] D. Slepian, "On Bandwidth," *Proc. IEEE*, vol. 64, pp. 292-300, 1976.
[8] D. L. Kreider, R. G. Kuller, D. R. Ostberg and F. W. Perkins, *An Introduction to Linear Analysis*. Reading, MA: Addison-Wesley, 1966.
[9] W. B. Davenport, Jr., and W. L. Root, *Random Signals and Noise*. New York: McGraw-Hill, 1958.