THE ROLE OF CODING IN MULTIPLE-ACCESS SATELLITE COMMUNICATION SYSTEMS

MICHAEL B. PURSLEY

APPROVED FOR PUBLIC RELEASE. DISTRIBUTION UNLIMITED.

UNIVERSITY OF ILLINOIS – URBANA, ILLINOIS
# The Role of Coding in Multiple-Access Satellite Communication Systems

<table>
<thead>
<tr>
<th>Title (and Subtitle)</th>
<th>The Role of Coding in Multiple-Access Satellite Communication Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Authors</td>
<td>Michael B. Pursley</td>
</tr>
<tr>
<td>Organization</td>
<td>Coordinated Science Laboratory</td>
</tr>
<tr>
<td>Address</td>
<td>University of Illinois at Urbana-Champaign Urbana, Illinois 61801</td>
</tr>
<tr>
<td>Controlling Office</td>
<td>Joint Services Electronics Program</td>
</tr>
<tr>
<td>Report Date</td>
<td>April, 1976</td>
</tr>
<tr>
<td>Number of Pages</td>
<td>72</td>
</tr>
<tr>
<td>Distribution</td>
<td>Approved for public release; distribution unlimited</td>
</tr>
<tr>
<td>Security Class</td>
<td>UNCLASSIFIED</td>
</tr>
<tr>
<td>Key Words</td>
<td>Multiple-Access Communications; Spread-Spectrum Communications; Satellite Communication Systems; Coding; Information Theory; Correlation Properties of Binary Sequences; Random Access; Computer-Communication Networks</td>
</tr>
</tbody>
</table>

It has been demonstrated that coding can play a major role in multiple-access satellite communication systems. In addition to providing error-control capability as in single-user communication systems, coding actually provides the multiple-access capability in many multi-user systems. The subject of this report is the use of coding techniques in a multiple-access satellite communication system in which the communication is to be shared by several asynchronous radio-frequency signals that occupy the same bandwidth. However, most of the concepts and techniques discussed...
in the paper apply to a wide variety of multiple-access communication systems in addition to satellite systems. Results applicable to multiple-access communication systems from the fields of information theory, algebraic coding theory, and communication theory are presented along with a discussion of important mathematical models of multiple-access satellite systems and practical considerations and constraints in the design of such systems. Analytical results on the performance of phase-coded spread-spectrum multiple-access systems are obtained and some new multiple-access communication schemes for a satellite link in a computer-communication network are discussed.
THE ROLE OF CODING IN MULTIPLE-ACCESS SATELLITE COMMUNICATION SYSTEMS

by

Michael B. Pursley

This work was supported in part by the Joint Services Electronics Program (U.S. Army, U.S. Navy and U.S. Air Force) under Contract DAAB-07-72-C-0259.

Reproduction in whole or in part is permitted for any purpose of the United States Government.

Approved for public release. Distribution unlimited.
The Role of Coding in Multiple-Access Satellite Communication Systems

by

Michael B. Pursley

April 15, 1976

Abstract

It has been demonstrated that coding can play a major role in multiple-access satellite communication systems. In addition to providing error-control capability as in single-user communication systems, coding actually provides the multiple-access capability in many multi-user systems. The subject of this report is the use of coding techniques in a multiple-access satellite communication system in which the communication capacity is to be shared by several asynchronous radio-frequency signals that occupy the same bandwidth. However, most of the concepts and techniques discussed in the paper apply to a wide variety of multiple-access communication systems in addition to satellite systems. Results applicable to multiple-access communication systems from the fields of information theory, algebraic coding theory, and communication theory are presented along with a discussion of important mathematical models of multiple-access satellite systems and practical considerations and constraints in the design of such systems. Analytical results on the performance of phase-coded spread-spectrum multiple-access systems are obtained and some new multiple-access communication schemes for a satellite link in a computer-communication network are discussed.

*This work is supported in part by the Joint Services Electronics Program (U.S. Army, U.S. Navy, and U.S. Air Force) under Contract DAAB-07-72-C-0259.
I. Introduction

Multiple-access communication theory is a relatively young research area even when compared with related disciplines like information theory and coding theory which are only twenty-eight years old now. Multiple-access communication theory was perhaps first recognized as a major branch of communications in 1964 when the Institute for Defense Analysis held its summer study on multiple-access techniques for satellite communications. The proceedings of that study \[49,6\] formulated many important theoretical and practical problems that are unique to multiple-access communications.

One of the important conclusions of this report and a number of other publications that appeared in the mid-1960's was that coding can play a dual role in satellite multiple-access systems. In addition to its usual error-control function in single-user systems, coding can actually provide the multiple-access capability in multi-user systems. We will follow the convention of referring to any multiple-access technique in which coding provides the primary multiple-access capability as code-division multiple access (CDMA). The important feature of a CDMA system is that each user's signal is coded in such a way that the various signals can be separated at the receiver even though they may occupy the same bandwidth at the same time. Thus, unlike traditional time-division multiple access (TDMA) and frequency-division multiple access (FDMA) there is no requirement for
precise time and frequency coordination between the various transmitters in the system. We should mention that according to our convention, CDMA includes certain multiple-access schemes which are sometimes viewed as hybrid techniques [102].

A specific example of a CDMA system which will be discussed in much greater detail in Section III is the spread-spectrum multiple-access (SSMA) technique in which the code sequence is modulated onto the carrier along with the digital data. The various transmitted signals may occupy the full channel bandwidth at all times. We will see that the code sequences should be selected to have good cross-correlation properties in order to enable the receiver to distinguish between the signals. Additional coding can be employed to correct message errors that occur during transmission.

In recent years CDMA techniques have been considered for a variety of satellite systems including the NASA tracking and data-relay system [90, 14,44], systems to provide communication to aircraft and other mobile users [55], and air traffic control systems [91]. In addition CDMA techniques are very attractive for military satellite communication systems since they can provide additional encryption and anti-jam benefits. In certain multiple-access satellite communication systems CDMA techniques can be designed to simultaneously provide multiple-access capability and to reduce the effects of multipath distortion.
In this paper we concentrate on those multiple-access methods in which coding techniques can be most profitably employed to improve system performance. In addition we restrict attention to systems in which several asynchronous radio-frequency signals occupy the same bandwidth. Thus, for example, we do not discuss FDMA where each user is assigned a different frequency slot or synchronous TDMA in which each user is assigned a different set of time slots. This is not to say that such systems are less important or that coding cannot be used to improve performance in FDMA or TDMA. However, coding is not the major element in providing multiple-access capability in FDMA and TDMA. For applications like FDMA the selection and performance of a code for a given user is essentially independent of the coding method employed by other users in the system. The coding theory needed is essentially the same as for single-user systems. This is not the case in the SSMA system mentioned above where the codes for all users must be selected as a set in order to minimize the mutual interference. For example, if orthogonal codes are employed in a completely synchronized SSMA system this mutual interference is eliminated. In practice, however, the SSMA system is not completely synchronized, mutual interference cannot be eliminated, and there are better choices of codes than orthogonal codes. Coding theory can be employed in the selection of sets of sequences to minimize the mutual interference as well as in the selection of codes to correct errors resulting from this interference.

Coding also plays a major role in a particular class of multiple-access techniques that we will refer to as random access. We will include
under the heading "random access" those multiple-access methods which utilize
coding together with the statistical properties of low duty-factor sources
to permit the communication capacity to be shared by several asynchronous
RF signals which occupy the same frequency band. Included in this category
are techniques which are related to asynchronous multiplexing [101,93],
random-access discrete address (RADA) [87,17], random time-division multiple-
access [16], and asynchronous time-division multiplexing [3, Chapter 7]. In
addition many of the pulse-address multiple-access (PAMA) techniques [81]
fall in this category.

Random-access schemes have been used for computer-communication net-
works such as the ALOHA system which is described in [1] and [3, Chapter
14]. The use of a satellite channel as part of the ALOHA system is dis-
cussed in [2]. In Section IV we review the past role of coding in random-
access systems like ALOHA and we present some new suggestions for future
application of more advanced communication techniques for random-access
satellite communication systems.

One of the important steps in the design of new methods of
communication via satellite is to obtain tractable mathematical models of
the overall communication system. The amount of detail needed or desired
in the system model increases as the system progresses through the various
design phases. In Section II we discuss gross system models for prelimi-
nary design and analysis. Here the attempt is to include essential features
of the overall communication network to give a basis for the selection of
an appropriate multiple-access technique.

Finally in Section V we briefly review recent results in multi-user information theory which pertain to multiple-access satellite communication systems.
II. System Models for Multiple-Access Communications

In a given satellite communication system the selection of the multiple-access technique to be employed depends on such factors as the overall system configuration, the number of transmitters and their data rates, the duty factors of the various transmitters, the availability of bandwidth in the system, the ability to have precise time and frequency coordination, the transmission characteristics of the channel, the acceptable error rate, and the acceptable amount of time delay. For instance, contrast the requirements for a system in which a few high power, high data rate signals are to be simultaneously transmitted over a single wideband satellite channel to a ground station with the requirements imposed by a system with a large number of inexpensive terminals, each of which transmits infrequently to other terminals in the system over a computer-communication network. Clearly these two situations call for completely different models and different multiple-access techniques.

There are a number of different system configurations that arise in multiple-access satellite communications. Since the choice of multiple-access technique will depend to a great extent on the system configuration, it is worthwhile to give block diagram models for a few of the more common configurations for satellite communication systems. The emphasis here is on gross system models and not on detailed models of the actual components of the system.
An important system configuration which arises frequently in multiple-access systems is shown in Figure 1. In this model there are K transmitters which share a single satellite communication channel. The channel output consisting of the K signals is delivered to each of J receivers. It may be that there is a single receiver (J = 1) which wishes to receive the data of all K sources as in the problem of satellite data relay from remote sources to a single ground station. On the other hand there may be several receivers each of which is interested in the data from a subset of the K users as is the case in certain computer-communication networks. As shown, the model does not allow any communication or cooperation between the K transmitters such as when a timing signal for synchronization purposes is supplied to all of the transmitters. Such external control signals can be added to the model for specific applications as needed. However, for the majority of applications involving remote users, such control signals are not available and asynchronous multiple-access techniques must be employed.

An important feature of this model is that all relevant communication is in one direction, from the source to the receiver. Since there is no feedback, schemes like error detection and retransmission are not applicable and only forward-acting error-control coding schemes can be considered.

This model is appropriate for such systems as the data-relay satellite system in which several low-altitude satellites simultaneously transmit data to a few ground stations (typically only one) via a data-relay satellite.
FIGURE 1. MULTIPLE-ACCESS DATA RELAY SYSTEM MODEL.
Although there is usually a return communication link to the satellite for command and tracking purposes, it is not suitable for feedback communication purposes. Furthermore, in such systems the low-altitude satellites are transmitting real-time sensor data continuously over the channel at a high rate (the data rate may often be as high as $10^7$ to $10^8$ bits per second) so that error-detecting codes are of little value since retransmission is not possible.

The model for the channel in Figure 1 will depend on the application. A suitable channel model for many systems is one in which the incoming signals are summed together with noise to given the channel output. That is, the channel output $v(t)$ is given by $v(t) = \sum_{k=1}^{K} s_k(t) + n(t)$ where $n(t)$ is the channel noise process. This model is appropriate for a data-relay satellite channel, for instance, if the satellite repeater is operated in linear mode. The model can be modified to include the effects of nonlinearities such as for hard-limiting satellite repeater [4,6,10]. In a satellite channel the summing of the signals actually takes place in space and the satellite communication subsystem "sees" a composite signal. We refer to this as a sum-in-space channel. For an ideal linear frequency-translating repeater, the only noise introduced by the satellite is thermal noise in the communication subsystem, which is both additive and Gaussian. Thermal noise will also be added at the ground station receiver. For satellite communication systems operating in certain frequency bands,
multipath and other forms of fading should be included in the model since they may greatly influence the selection of a multiple-access method. On the other hand, for most satellite systems such second-order effects as intersymbol interference, phase jitter, long-term frequency drifts, and minor gain and time-delay variations have little impact on the choice of multiple-access technique except perhaps to rule out certain impractical schemes which require extremely precise time, phase, or frequency coordination and synchronization.

For certain applications, there may be $K$ remote users in the system each of which has a transmitter (consisting of a data source, an encoder, and a modulator) and a receiver (consisting of a demodulator, a decoder, and a data sink). In such cases the $k^{\text{th}}$ transmitter and $k^{\text{th}}$ receiver together will be called the $k^{\text{th}}$ terminal. The model of Figure 1 should be modified by adding a subsystem which connects the $k^{\text{th}}$ transmitter with the $k^{\text{th}}$ receiver. This would produce the model of Figure 2 which is applicable for a variety of multiple-access systems including a satellite system which provides for communication between several ground stations, computer facilities, aircraft, etc. This typically involves two-way communication via an earth-coverage synchronous satellite so that, as indicated, the individual terminal has the ability to receive all of the transmitted signals (including its own).

Since several terminals can simultaneously transmit information to all of the terminals in the system, we refer to this as a multiple-access broadcast system. The individual terminals may demodulate, decode, and use the
FIGURE 2. MULTIPLE-ACCESS BROADCAST SYSTEM MODEL.
data from a particular transmitter or simply ignore it. They may also ask for a retransmission if an error was detected during decoding or they may acknowledge receipt of the message. Note that due to receiver thermal noise in the terminals themselves, a terminal will not be able to verify correct reception by another terminal merely by monitoring its own signal at the channel output. However, in certain multiple-access systems the transmitting terminal may want to monitor the channel output to detect the presence of significant interference due to other users and automatically retransmit the message after a short delay if such interference is detected.*

The key features of the multiple-access broadcast system model are that two-way communication is possible between any pair of terminals and each terminal receives all signals transmitted over the channel. Again, the model of the channel itself depends on the specific system. For example, the sum-in-space additive noise channel is of interest for linear frequency-translating satellite repeaters. The satellite multiple-access broadcast system has been considered for computer-communication applications [53].

We will mention one more of the many models of interest. In this model neither of the extremes of the previous models is valid because a feedback channel is available for retransmission requests but it may be a different channel than that used for the original transmission. This is illustrated in Figure 3 for K + J terminals and two channels. Note that this model results when we simply add a feedback channel and the necessary

*This idea is discussed further in Section IV.
FIGURE 3. TWO-WAY MULTIPLE-ACCESS SYSTEM MODEL.
transmitters and receivers to the basic model shown in Figure 1. This model may represent a satellite system with two narrow-beam antennas for relaying data between the east coast and the west coast for example. Of course the model can be generalized to any number of terminals and channels.

Before proceeding further with our discussion, we should distinguish between the terms "multiple access" and multiplexing". It has been suggested (e.g., [82]) that this can be accomplished by referring to schemes in which communications capacity is shared by several radio-frequency (RF) signals as multiple access while referring to schemes in which the capacity is shared by several baseband signals as multiplexing. While this division into baseband vs. RF signals has certain deficiencies, it accurately distinguishes between multiple access and multiplexing in most cases. In particular, the schemes that are suitable for communication systems in which the communication capacity is to be shared by several asynchronous RF signals which occupy the same frequency band are correctly classified as multiple access according to these definitions. These are the schemes that will be discussed further in this paper. Just to give one example where these definitions do not accurately portray the nature of the system, we mention FDMA which is really more of a multiplexing technique. In fact, FDMA has been used under the name of "frequency-division multiplexing" for many years but it involves RF signals and not baseband signals. Since in this paper we do not consider FDMA or related techniques, this example does not invalidate the definition for our purposes. However, it is perhaps more accurate to view multiplexing as a special case of multiple access since
in practice, when designing a multiple-access system, we should feel free to use multiplexing techniques whenever they are applicable. This would be the case when it is possible to completely synchronize the transmitters or coordinate their use of particular frequency bands. It is this synchronization and coordination that most accurately characterizes multiplexing techniques and determines the extent of their usefulness in a given system. In particular, multiplexing techniques are not appropriate for satellite communication systems in which the communications capacity is to be shared by several RF signals which occupy the same bandwidth and are not time-synchronous.

The main distinguishing feature between coding for multiple-access systems and coding for multiplexing is that in the former case each user has a separate encoder and the encoding process (not the code itself) for a given user is independent of all other users. In the latter case a single encoder processes the data from all of the users prior to modulation and transmission. The model for the latter case shown in Figure 4 should be contrasted with our basic model shown in Figure 1. Note that Figure 4 is not a valid model for direct-access satellite systems with remote users unless the uplink to the satellite is noise-free. In addition the satellite must have a signal processing repeater which demodulates, decodes, and synchronizes the uplink signals and then performs the encoding operation on the resulting data streams. Repeaters of this type have been investigated by Sommer [89], Huang and Hooten [43], and Horwood and Gagliardi [42]. Since the channel capacity is shared by several baseband digital
FIGURE 4. MULTIPLEXING SYSTEM MODEL
signals, the schemes considered in these papers are multiplexing techniques according to the definitions above. We will primarily be concerned with multiple-access systems like those shown in Figures 1 though 3 which do not require precise time and frequency coordination between users and do not require any sophisticated signal processing in the repeater.
III. Spread-Spectrum Multiple-Access Systems

Two of the most commonly used multiple-access techniques for data-relay satellite systems are phase-coded SSMA and frequency-hopped SSMA. The latter technique was employed for example in the TATS modulation system for the Lincoln Experimental Satellites and is described in detail in [55] and [23]. In this section we concentrate on phase-coded SSMA. We will first review methods for the selection of codes for phase-coded SSMA systems and then present recent results on analysis of such systems.

A. Sequences with Good Periodic Correlation Properties

In this section we will summarize some of the important work relating to the problem of selecting codes for phase-coded SSMA communication systems. We will concentrate on maximal length linear feedback shift-register sequences (m-sequences) and related codes since their autocorrelation properties and ease of generation make them attractive for practical digital communication systems [38]. Nearly all of the early work on SSMA systems was concerned with systems which used m-sequences to distinguish between the various users. The goal in choosing a set of sequences for SSMA applications is to make the magnitude of the cross-correlation between any two sequences in the set as small as possible. Although it is the aperiodic or partial cross-correlation that must be made small in magnitude, nearly all techniques available for selecting these sequences are based on the periodic or full cross-correlation.
In the most common form of phase-coded SSMA, each of the K users is assigned a different sequence of positive and negative pulses. The sequence of pulses for the k\textsuperscript{th} user \((k = 1, 2, \ldots, K)\) is represented by a periodic infinite sequence \(a_{j}^{(k)}\) of elements from \(+1, -1\). The K-sequences must be easily distinguishable, which amounts to choosing a family of sequences with small pairwise cross-correlation function. Given two periodic sequences \(a_{j}^{(k)}\) and \(a_{j}^{(i)}\) over \(+1, -1\) of common period \(p\), their \textit{full} or \textit{periodic} cross-correlation is given by the formula:

\[
\theta_{k,i}(\tau) = \frac{1}{p} \sum_{j=0}^{p-1} a_{j}^{(k)} a_{j+\tau}^{(i)} .
\]

Note that \(\theta_{k,i}(\tau)\) is just \(p\) minus twice the Hamming distance between codewords \((a_{0}^{(k)}, a_{1}^{(k)}, \ldots, a_{p-1}^{(k)})\) and \((a_{\tau}^{(i)}, a_{\tau+1}^{(i)}, \ldots, a_{\tau+p-1}^{(i)})\). Thus, minimization of the magnitude of the correlation corresponds to selecting codewords with Hamming distances as close to \(\frac{p}{2}\) as possible. This latter definition of \(\theta_{k,i}(\tau)\) is also suitable for the equivalent sequences over GF(2) (the binary alphabet \(\{0, 1\}\) with addition modulo 2) in which +1 is replaced by 0 and -1 is replaced by 1. The GF(2) representation is commonly employed in the literature and will be used throughout Section III.A. We will first discuss results on the problem of selection of m-sequences with uniformly small (in magnitude) periodic cross-correlation and then consider the selection of sequences with good aperiodic cross-correlation properties in Section III.C.

*This definition will also be used if \(p\) is the least common multiple of the periods of the two sequences.*
Throughout the remainder of the paper, \( p \) will denote the period of a linear feedback shift-register sequence and \( n \) will denote the length of the shift register that generates the sequence, so \( p = 2^n - 1 \) for m-sequences. The integer part of the real number \( x \) is denoted by \( \lfloor x \rfloor \). In order to present analytical and experimental results on the periodic cross-correlation function for linear feedback shift-register sequences it is necessary to establish some terminology and notation commonly employed in algebraic coding theory.

We will represent the \( n \)-stage linear feedback shift-register by the polynomial

\[
 f(x) = \sum_{i=0}^{n} c_i x^{n-1}
\]

where \( c_0 = c_n = 1 \) and for \( 0 < i < n \), \( c_i = 1 \) if there is a feedback tap connected to the \( i \)th stage of the register and \( c_i = 0 \) if not. There is always such a connection for the \( n \)th stage. The outputs of the register satisfy the linear recurrence relation

\[
 a_j = \sum_{i=1}^{n} c_i a_{j-i}
\]

where \( a_j \) represents the present output and \( a_{j-1} \) represents the output at \( i \) time units in the past. The convention that we are following in defining \( f(x) \) is consistent with error-control coding texts [69, Section 7.4] where \( f(x) \) represents the parity check polynomial for the code generated by the shift register. Especially in the literature on shift-register sequences, some authors prefer to represent the register by the polynomial
\[ f^*(x) = \prod_{i=0}^{n} c_i x^i \]

where the \( c_i \) are as defined above. The polynomial \( f^*(x) \) is just the reciprocal polynomial of \( f(x) \) [69]; that is, \( f^*(x) = x^nf(x^{-1}) \). Thus the parity check polynomial for the code can also be written as \( x^n f^*(x^{-1}) \).

Two sequences \( (a_j) \) and \( (\hat{a}_j) \) are in the same equivalence class (with respect to the shift transformation) if there exists an integer \( k \) such that \( \hat{a}_j = a_{j+k} \) for each \( j \). The set of sequences generated by \( f(x) \) consists of one member of each equivalence class in the collection of all sequences that can be obtained as outputs of the shift register with no more than \( (n-1) \) zeroes initially in the register. The sequence \( (a_{j+k}) \) is called the \( k \)th shift of the sequence \( (a_j) \).

The polynomials which generate \( m \)-sequences of period \( p = 2^n - 1 \) are the primitive polynomials of degree \( n \) over GF(2). (A polynomial of degree \( n \) is primitive if it divides \( x^m - 1 \) for \( m = 2^n - 1 \) but not for any \( m < 2^n - 1 \).) A complete discussion of primitive polynomials can be found in standard texts on algebraic coding [13, 69] and a table which gives all primitive polynomials of degree \( n \) for \( n \leq 16 \) is available in [69]. The roots of primitive polynomials of degree \( n \) are primitive elements of the extension field GF(\( 2^n \)). (\( \alpha \) is a primitive element of GF(\( 2^n \)) if \( \alpha^m = 1 \) for \( m = 2^n - 1 \) but \( \alpha^m \neq 1 \) for all \( m < 2^n - 1 \). Every nonzero element of GF(\( 2^n \)) can be written as a power of the primitive element \( \alpha \).) If \( \alpha \) is a primitive element of GF(\( 2^n \)) let \( f_1(x) \) be the primitive polynomial of degree \( n \) which has \( \alpha \) as a root. For each \( k < 2^n - 1 \) let \( f_k(x) \) be the minimal polynomial of
\(\alpha^k\) (i.e., the polynomial of smallest degree for which \(\alpha^k\) is a root). The polynomial \(f_k(x)\) is a primitive polynomial of degree \(n\) (and \(\alpha^k\) is a primitive element of \(GF(2^n)\)) if and only if the greatest common divisor of \(k\) and \(2^n - 1\) is 1. If \(k\) and \(j\) are related by \(k2^i \equiv j \mod 2^n - 1\) for some \(i, 0 \leq i \leq n - 1\), then \(f_j(x) = f_k(x)\). A common convention is to denote the minimal polynomial of \(\alpha^k\) by \(f_j(x)\) where \(j\) is the smallest integer for which \(k2^i \equiv j \mod 2^n - 1\) for some \(i, 0 \leq i \leq n - 1\). For example, if \(n = 5\), \(\alpha^3\) is a primitive element with minimal polynomial \(f_9(x) = f_{18}(x) = f_5(x) = f_{10}(x) = f_{20}(x)\) which is denoted by \(f_5(x)\). If \(f_1(x)\) is selected to be the primitive polynomial \(x^5 + x^2 + 1\) and if \(\alpha\) is a root of this polynomial, then according to [69, Table C.2] the minimal polynomial of \(\alpha^3\) is \(f_5(x) = x^5 + x^4 + x^2 + x + 1\).

The mathematical study of m-sequences and related codes seems to have started in the mid-1950's. According to Golomb [39] the theory of linear shift-register sequences was being developed and applied around this time by Gilbert, Golomb, Welch, and Zierler. This theory is presented in detail in [39] and [108]. Much of the early research was concerned with the autocorrelation properties and the "noise-like" aspects of the sequences. In fact, m-sequences were commonly referred to as pseudo-noise sequences. However, some attention was given to the problem of selecting sets of m-sequences with good cross-correlation properties and by the late 1960's several theoretical and experimental results were available. In 1965 Gold and Koptizke [32] experimentally determined sets of m-sequences called maximal connected sets which are the largest possible subsets of m-sequences.
of period \( p \) for which any two sequences in the same set have a preferred three-valued cross-correlation function (any pair of sequences for which the only periodic cross-correlation values are \(-2\lfloor \frac{n+2}{2} \rfloor - 1, 2\lfloor \frac{n+2}{2} \rfloor - 1, \) and \(-1\) is said to have a preferred three-valued cross-correlation function).

The advantage of choosing sequences from a maximal connected set as illustrated for the case \( p = 127 \) where any two sequences from the same maximal connected set have a cross-correlation function with the values \(-1, -17, \) and \(15\). By contrast, nearly all other pairs of \( m \)-sequences of period 127 have cross-correlation values as large as 41 and all pairs of \( m \)-sequences of period 127 which are not in the same maximal connected set have cross-correlation functions \( \theta(\tau) \) for which \(|\theta(\tau)| \geq 21 \) for at least 7 different values of \( \tau \). Let \( M_n \) denote the size of the maximal connected set of \( m \)-sequences of period \( p = 2^n - 1 \). Gold and Kopitzke [32] have determined the values of \( M_n \) for \( 5 \leq n \leq 13 \) which are as follows: \( M_7 = 6, M_{13} = M_{11} = 4, M_{10} = M_5 = 3, M_9 = M_6 = 2, \) and \( M_{12} = M_8 = 0 \). For completeness we record that \( M_3 = 2 \) and \( M_4 = 0 \). There is no interest in \( n = 1 \) or \( n = 2 \) since there is only one \( m \)-sequence for each of these two values (thus, \( M_1 = M_2 = 0 \)).

The first analytical results on the selection of sequences with low periodic cross-correlation were obtained independently by several researchers including Gold [33,35,37] and Kasami [50,51]. In 1966, Gold established the following result [33,35].
Theorem 1: Let $f_1(x)$ be a primitive polynomial of degree $n$ and let $\alpha$ be a root of $f_1(x)$ in the field $GF(2^n)$. If $f_t(x)$ is the minimal polynomial of $\alpha^t$ where $t = 2^\left\lfloor \frac{n+2}{2} \right\rfloor + 1$ then the magnitude of the cross-correlation between the sequences generated by $f_1(x)$ and $f_t(x)$ is not greater than $t$.

The theorem is valid for all positive integers $n$ but it is especially significant when $n$ is not a multiple of 4 since in this case $f_t(x)$ will be a primitive polynomial and hence both sequences will be m-sequences.

Gold [35,37], Kasami [50,51] and several others have independently established theorems which are stronger and more general than Theorem 1. One such result is the following which is stated in [68] where it is credited to Gold, Kasami, and Solomon.

Theorem 2: Suppose $f_1(x)$ and $\alpha$ are as in Theorem 1 and $f_t(x)$ is the minimal polynomial of $\alpha^t$ where $t = 2^k + 1$ and $0 < k < n$. Let $e$ denote the greatest common divisor of the two integers $n$ and $k$. If $n/e$ is odd, the cross-correlation function for the pair $\{f_1(x), f_t(x)\}$ takes on only three values. The value $-1$ occurs for $2^n - 2^{n-e} - 1$ different shifts, $-1 - 2^{(n+e)/2}$ occurs for $2^{n-e-1} - 2^{(n-e-2)/2}$ different shifts, and $-1 + 2^{(n+e)/2}$ occurs for $2^{n-e-1} + 2^{(n-e-2)/2}$ different shifts.
Note that $2^k(2^{n-k} + 1) = (2^n - 1) + (2^k + 1)$ so that $t = 2^k + 1$ and $s = 2^{(n-k)} + 1$ implies $f_t(x) = f_s(x)$. Furthermore, the greatest common divisor of $n$ and $k$ is equal to the greatest common divisor of $n$ and $n - k$. Therefore, in generating pairs $\{f_1(x), f_t(x)\}$ via Theorem 2 we only need to consider $0 < k \leq n/2$. In all that follows, any pair of primitive polynomials which have the preferred three-valued cross-correlation function will be referred to as a preferred pair of primitive polynomials. Theorem 2 will yield preferred pairs of primitive polynomials only for those values of $n$ and $k$ for which $n$ is odd and $e = 1$ or $n \equiv 2 \mod 4$ and $e = 2$.

In particular if $k = \lfloor (n + 2)/2 \rfloor$ then $e = 2$ for $n \equiv 2 \mod 4$ and $e = 1$ for $n \not\equiv 2 \mod 4$. Hence, for this value of $k$, $n/e$ is odd if $n \not\equiv 0 \mod 4$ and $n/e$ is even if $n \equiv 0 \mod 4$. For $n \not\equiv 0 \mod 4$ and $k = \lfloor (n + 2)/2 \rfloor$, Theorem 2 implies that the pair $\{f_1(x), f_t(x)\}$ obtained from Theorem 1 is a preferred pair of primitive polynomials.

As an example, let $n = 5$ so that $e = 1$ for all $k$. Theorem 1 gives the preferred pair $\{f_1(x), f_9(x)\}$. Since $9(2^2) \equiv 5 \mod 31$, $f_9(x) = f_5(x)$. Theorem 2 provides the additional pair $\{f_1(x), f_3(x)\}$ for $k = 1$ (or $k = 4$) but it generates the same pair as Theorem 1 for $k = 2$ (or $k = 3$). Note that $(\alpha^3)^3 = \alpha^9$ so that if $\alpha$ is replaced by $\alpha^3$ in Theorem 2 the conclusion is that $\{f_3(x), f_5(x)\}$ is also a preferred pair of primitive polynomials. Thus, $\{f_1(x), f_3(x), f_5(x)\}$ is a set which has the property that any two polynomials from the set have a preferred three-valued cross-correlation function.
Since $M_5 = 3$ this set must be a maximal connected set. According to the tables of Gold and Koptizke [32], $t = 11$ and $t = 7$ also give the preferred three-valued cross-correlation function relative to $f_1(x)$. Although it may not be obvious at this point, we will see later that these two values of $t$ are predicted by Theorem 2. The value $t = 11$ is also predicted by the next theorem.

The following result was observed to be true for $n < 13$ by Gold [33] in 1966 and later reported as a theorem for arbitrary values of $n$ by Golomb [40] who credits its proof to Welch.

**Theorem 3**: Theorem 2 is also true if $t = 2^k + 1$ is replaced by $t = 2^{2k} - 2^k + 1$.

From the identity

$$
(2^{2(n-k)} - 2^{n-k} + 1) + (2^n - 1) = 2^{n-2k(2^{n-2k} + 2^2k)}
$$

it follows that if $k \leq n/2$, $t = 2^{2k} - 2^k + 1$, and $s = 2^{2(n-k)} - 2^{n-k} + 1$ then $f_t(x) = f_s(x)$. Thus, only $0 < k < n/2$ need be considered in Theorem 3.

For the example $n = 5$, Theorem 3 yields $\{f_1(x), f_{11}(x)\}$ for $k = 2$ since $13(2^2) \equiv 11 \mod 31$. Also, since $(a^{11})^5 = a^{31}a^{24} = a^{24}$ and $24(2^2) \equiv 3 \mod 31$, replacement of $a$ by $a^{11}$ in Theorem 2 (for $k = 2$) shows $\{f_{11}(x), f_3(x)\}$ is a preferred pair of primitive polynomials and therefore $\{f_1(x), f_3(x), f_{11}(x)\}$ is a maximal connected set. The case $t = 7$ is not predicted by Theorem 3.
The following proposition can be used to extend Theorems 2 and 3.

**Proposition 1:** If \( a \) and \( a^t \) are primitive elements of \( GF(2^n) \) and if

\[
rt \equiv 2^i \mod 2^n - 1
\]

for some \( i, 0 \leq i \leq n - 1 \), then \( a^r \) is also primitive and

\[
(a^r)^t = a^{2^i}.
\]

It follows that \( rt \equiv 2^i \mod 2^n - 1 \) implies \( f_{rt}(x) = f_1(x) \). Because of this fact, we can replace \( a \) by \( a^r \) in Theorems 2 and 3 to show that the conclusion of these theorems will also be true for the pair \( \{f_1(x), f_r(x)\} \). For values of \( n \) of greatest interest, use of Proposition 1 doubles the number of pairs obtained from a given theorem. A simple example for which nothing new is added is the case \( n = 3, t = 3, \) and \( r = 3 \). Henceforth, when we refer to pairs obtained from these theorems, we will include those pairs obtained via Proposition 1.

Returning to the example in which \( n = 5 \) we see that an application of Proposition 1 to Theorem 2 (for \( k = 2, t = 5, r = 7 \)) furnishes the remaining pair \( \{f_1(x), f_7(x)\} \). This same pair is also predicted by a conjecture of Welch that for any odd \( n \), \( \{f_1(x), f_r(x)\} \) has the preferred three-valued cross-correlation function if \( t = 2^{(n-1)/2} + 3 \). This conjecture has been verified for \( n \leq 17 \) [68]. The remaining maximal connected sets containing \( f_1(x) \) can now be determined. Since \( (a^5)^5 = a^{25} \) and \( f_{25}(x) = f_7(x) \), \( \{f_1(x), f_5(x), f_7(x)\} \) is a maximal connected set. Also \( (a^7)^3 = a^{21} \) and
f_{21}(x) = f_{11}(x) \implies \{f_1(x), f_7(x), f_{11}(x)\} is a maximal connected set.

Although Welch's conjecture gives a pair \{f_1(x), f_t(x)\} which can also be obtained from Theorem 2 for \(n = 5\), this is not true in general. For instance, if \(n = 9\), Welch's conjecture (with Proposition 1) gives \(t = 19\) and 27. Theorem 2 furnishes \(t = 3, 5, 17, 31, 103,\) and 171. Theorem 3 gives \(t = 3, 13, 47, 59, 87,\) and 171. According to the tables in [32], these are all of the values of \(t\) for which \(\{f_1(x), f_t(x)\}\) is a preferred pair of primitive polynomials. In addition, Theorem 2 (for \(k = 3\)) provides the pairs \(\{f_1(x), f_9(x)\}\) and \(\{f_1(x), f_{57}(x)\}\) which have a three-valued cross-correlation function with values 63, -1, and -65 rather than the preferred values 31, -1, and -33. The same is true for Theorem 3 (for \(k = 3\)).

For many shift-register lengths \(n > 10\) there are preferred pairs of primitive polynomials which are not included in Theorems 1 and 2 and are not covered by Welch's conjecture. Some examples obtained from [32] are as follows: \(t = 25, 41, 59,\) and 107 for \(n = 10\); \(t = 107\) and 249 for \(n = 11\); and \(t = 71\) and 347 for \(n = 13\). Additional examples for larger values of \(n\) are provided in [68].

These analytical and experimental results make it possible to construct sets of m-sequences of common period \(p = 2^n - 1\) which have pairwise cross-correlation bounded in magnitude by \(2^{\left[\frac{(n+2)}{2}\right]} + 1\) provided \(n\) is not a multiple of 4. The size of these sets will not be greater than the size of the maximal connected set of sequences of the same period which contain no more than six sequences for \(p \leq 4095\). Thus, for many applications these sets are too small.
There are a number of alternatives that one might consider for obtaining larger sets. First, the bound $2^\left\lceil (n+2)/2 \right\rceil + 1$ could be increased. For most cases of interest, however, this bound must be increased by a factor of two or more in order to obtain a set that is significantly larger than the maximal connected set. For example, for $n = 7$ the bound on cross-correlation for the maximal connected set is 17 and the number of sequences in the set is 6. In order to obtain a larger set, the cross-correlation bound must be increased to 41 at which point the set contains all m-sequences of period 127.

A second approach that one might consider is to select sequences for the set based on a criterion other than the maximum magnitude of the cross-correlation function. This approach may yield some useful results for certain criteria. One criterion that comes to mind is to minimize the mean-squared value of the cross-correlation function. However, if $\theta(\tau)$ is the cross-correlation function between any two m-sequences of period $p = 2^n - 1$, Gold [37] has shown that

$$\sum_{\tau=0}^{p-1} [\theta(\tau)]^2 = 2^{2n} - 2^n - 1.$$  \hspace{1cm} (2)

Thus all cross-correlation functions for pairs of m-sequences of a given period have the same mean-squared value. (They also have the same average value.)

A third approach is to drop the requirement that all of the sequences in the set must be m-sequences. For certain applications, such as communication over multipath channels, the fact that some sequences in the set will not have
the good autocorrelation properties possessed by m-sequences is a serious
disadvantage. However, if large sets of linear feedback shift-register
sequences with good cross-correlation properties are needed, there seems
to be no alternative for the majority of applications. The following
theorem, which was one of the first results available for obtaining these
large sets, is due to Gold [35] and Kasami [50].

**Theorem 4:** Let \( f_1(x) \) and \( f_3(x) \) be a preferred pair of
primitive polynomials of degree \( n \) (\( n \) is not a multiple
of 4). The shift register with the product \( f_1(x)f_3(x) \)
as its characteristic polynomial will generate a set of
\( 2^n + 1 \) different sequences of period \( 2^n - 1 \). Any pair
of sequences in this set has a preferred three-valued
cross-correlation function and any sequence in the set
has an autocorrelation function with out-of-phase values
bounded by \( 2\lfloor(n+2)/2\rfloor + 1 \).

These sequences are commonly referred to in the literature as Gold sequences.

The \( 2^n + 1 \) different sequences of Theorem 4 are obtained by finding an
appropriate collection of \( 2^n + 1 \) different binary words of length \( 2n \) to
serve as initial conditions for the shift register. Such a collection can
be found by the following "brute-force" technique. Start with a list of all
\( 2^{2n} - 1 \) nonzero binary words of length \( 2n \). Choose a word from the list and
use it as the initial condition for the shift-register. Observe the register
contents for all \( 2^n - 1 \) shifts. Each of the \( 2^n - 1 \) different binary words
that appears in the register is eliminated from the list since it will not generate a new sequence. A binary word is selected from the reduced list and the above process is repeated. Sarwate [76] has suggested a considerably simpler method for generating the collection of initial conditions based on the observation that if \((a_i)\) is the sequence generated by \(f_1(x)\) and \((b_i)\) is generated by \(f_t(x)\) then all but two of the sequences generated by \(f_1(x)f_t(x)\) are of the form \(a_i \oplus b_{i+k}\) for \(0 \leq k < 2^n - 1\) (where \(\oplus\) denotes summation modulo 2). The two sequences not of this form are the original sequences \((a_i)\) and \((b_i)\). Thus to specify the first \(2n\) elements of the sequences generated by \(f_1(x)f_t(x)\) we only need to know the sequence \((b_i)\) and the first \(2n\) elements of \((a_i)\). These \(2^n - 1 + 2n\) binary digits specify the \(2^n + 1\) binary words of length \(2n\) needed to generate the Gold sequences.

A number of results similar to Theorem 4 have been obtained by Kasami [50,51] who has also investigated smaller and larger sets of sequences and has obtained sequences for the case \(n \equiv 0 \mod 4\). Many additional sets of sequences have been considered for SSMA applications. Examples are the primitive root codes of Anderson [9], the Frank-Heiniller sequences [24, 25, 41], and the sequences of Chu [19]. As in the case of the sequences of Gold and Kasami, however, these have been designed to have good autocorrelation or periodic cross-correlation properties as opposed to good aperiodic cross-correlation properties. In the next section we will discuss the role of the aperiodic cross-correlation function in the design and analysis of phase-coded SSMA systems.
B. Performance of Phase-Coded SSMA Systems

The SSMA system model that will be employed in this section is shown in Figure 5 for $K$ users. The $k$th user's data signal $b_k(t)$ is a sequence of unit amplitude, positive and negative, rectangular pulses of duration $T$. This signal represents the $k$th user's binary information sequence. The $k$th user is assigned a code waveform $a_k(t)$ which consists of a periodic sequence of unit amplitude, positive and negative rectangular pulses of duration $T_c$. If $(a_j^{(k)})$ is a sequence of elements of $\{+1,-1\}$ then we can represent $a_k(t)$ as

$$a_k(t) = \sum_{j=-\infty}^{\infty} a_j^{(k)} p_{T_c}(t - jT_c)$$

where $p_{T_c}(t) = 1$ for $0 \leq t < T_c$ and $p_{T_c}(t) = 0$ otherwise. We assume that each sequence $(a_j^{(k)})$ has period $p = T/T_c$ so that there is one code period $a_0^{(k)}, a_1^{(k)}, \ldots, a_{p-1}^{(k)}$ per data symbol. The results presented can easily be generalized to multiple code periods per data symbol.

The data signal $b_k(t)$ is modulated onto the phase-coded carrier $c_k(t)$ given by

$$c_k(t) = \sqrt{2P} \sin (\omega_c t + \theta_k + (\pi/2)a_k(t))$$

$$= \sqrt{2P} a_k(t) \cos (\omega_c t + \theta_k)$$

which produces the signal
FIGURE 5. PHASE-CODED SPREAD-SPECTRUM MULTIPLE-ACCESS SYSTEM MODEL.
In the above expressions \( \theta_k \) represents the phase of the \( k^{th} \) carrier, \( \omega_c \) represents the common center frequency, and \( P \) represents the common signal power. The results that follow can easily be modified for unequal center frequencies and power levels.

The term "spread spectrum" is derived from the fact that the bandwidth of \( s_k(t) \) is approximately \( p \) times that of the corresponding uncoded RF signal. However, the actual increase in channel bandwidth can be somewhat less than \( p \) for most practical systems. The type of SSMA coding described above has been referred to as block sub-baud coding by Massey and Uhran [62,63] (who have also investigated convolutional sub-baud coding).

If the SSMA system is completely synchronized then the time delays \( \tau_k \) shown in the model of Figure 5 can be ignored (i.e., \( \tau_k = 0 \) for \( k = 1,2,\ldots,K \)). This would require a common timing reference for the \( K \) transmitters and it would necessitate compensation for delays in the various transmission paths. For the majority of SSMA communication systems such compensation is not feasible and hence the transmitters are not time-synchronous. For asynchronous systems the received signal \( r(t) \) in Figure 5 is given by

\[
r(t) = n(t) + \sum_{k=1}^{K} \sqrt{2P} a_k(t-\tau_k)b_k(t-\tau_k) \cos (\omega_c t + \phi_k)
\]
where $\phi_k = \theta_k - \omega_c \tau_k$ and $n(t)$ is the channel noise process which we assume to be a white Gaussian process with two-sided spectral density $N_0/2$.

Since we are concerned with relative phase shifts modulo $2\pi$ and relative time delays modulo $T$, there is no loss in generality in assuming $\theta_i = 0$ and $\tau_i = 0$ and considering only $0 \leq \tau_k < T$ and $0 \leq \theta_k < 2\pi$ for $k \neq i$.

If the received signal $r(t)$ is the input to a correlation receiver matched to $s_i(t)$, the output is

$$Z_i = \int_0^T r(t) a_i(t) \cos \omega_c t \, dt.$$ 

In all that follows we assume $\omega_c T/2\pi$ is an integer. However, the results obtained are also valid if $\omega_c \gg T^{-1}$ since the frequency response of a realistic hardware implementation of the correlation receiver is such that the double frequency component of $r(t) \cos \omega_c t$ will be negligible under this condition. The condition $\omega_c \gg T^{-1}$ is always satisfied in a practical SSMA communication system. The data signal $b_k(t)$ can be expressed as

$$b_k(t) = \sum_{\ell=-\infty}^{\infty} b_k,\ell p_T(t - \ell T)$$

where $b_k,\ell \in \{+1,-1\}$. The output of the correlation receiver is given by

$$Z_i = \sqrt{P/2} \left\{ b_{i,0} T + \sum_{k=1}^{K} b_{k,-1} R_{k,i}(\tau_k) + b_{k,0} \hat{R}_{k,i}(\tau_k) \cos \phi_k \right\} + \int_0^T n(t) a_i(t) \cos \omega_c t \, dt$$
where $R_{k,i}$ and $\hat{R}_{k,i}$ are the continuous-time partial cross-correlation functions defined by

$$R_{k,i}(\tau) = \int_0^T a_k(t - \tau)a_i(t) \, dt, \quad \hat{R}_{k,i}(\tau) = \int_0^\tau a_k(t - \tau)a_k(t) \, dt$$

for $0 \leq \tau \leq T$. It is easy to show that for $0 \leq \ell T_c \leq \tau \leq (\ell + 1)T_c \leq T$, these two cross-correlation functions can be written as

$$R_{k,i}(\tau) = \rho_{k,i}(\ell)T_c + [\rho_{k,i}(\ell + 1) - \rho_{k,i}(\ell)](\tau - \ell T_c) \quad (3)$$

and

$$\hat{R}_{k,i}(\tau) = \hat{\rho}_{k,i}(\ell)T_c + [\hat{\rho}_{k,i}(\ell + 1) - \hat{\rho}_{k,i}(\ell)](\tau - \ell T_c) \quad (4)$$

where the discrete partial cross-correlation functions $\rho_{k,i}$ and $\hat{\rho}_{k,i}$ are defined by

$$\rho_{k,i}(\ell) = \sum_{j=0}^{\ell-1} a_k(j)a_j, \quad \ell = 1,2,\ldots,p$$

$$\hat{\rho}_{k,i}(\ell) = \sum_{j=\ell}^{p-1} a_k(j)a_j, \quad \ell = 0,1,\ldots,p-1$$

and $\rho_{k,i}(0) = \delta_{k,i}(p) = 0$. Notice that the periodic cross-correlation function defined in equation (1) is given by $\theta_{k,i}(\ell) = \rho_{k,i}(\ell) + \rho_{k,i}(\ell)$. We also define $\hat{\theta}_{k,i}(\ell) = \hat{\rho}_{k,i}(\ell) - \rho_{k,i}(\ell)$, which is called the odd cross-correlation function by Massey and Uhran [62,63]. It is easy to see that

$\rho_{k,i}(p - \ell) = \rho_{\ell,k}(\ell)$ and $\hat{\rho}_{k,i}(p - \ell) = \rho_{\ell,k}(\ell)$ and therefore $\theta_{k,i}(p - \ell) = \theta_{\ell,k}(\ell)$ and $\hat{\theta}_{k,i}(p - \ell) = -\hat{\theta}_{\ell,k}(\ell)$. 

The cross-correlation parameters can also be represented as a single function defined on the larger set \{-p, -p+1, 0, 1, ..., p-1\}. One such function \( \tilde{\rho}_{k,1} \), the aperiodic cross-correlation function, is defined by:

\[
\tilde{\rho}_{k,1}(\ell) = \left\{ \begin{array}{ll}
p - \ell & \\
p + \ell & \\
\sum_{j=0}^{p-1} a_j a_{j+\ell} &
\end{array} \right.
\]

\( \ell = 0, 1, \ldots, p-1 \)

\( \ell = 0, -1, \ldots, -(p-1) \).

The aperiodic cross-correlation function is related to the partial cross-correlation functions as follows. For nonnegative values of \( \ell \), \( \tilde{\rho}_{k,1}(\ell) = \tilde{\beta}_{k,1}(\ell) \) and for negative values of \( \ell \), \( \tilde{\rho}_{k,1}(\ell) = \rho_{k,1}(p - |\ell|) \).

If \( v_{k,1}(\tau_k) \) is defined by

\[
v_{k,1}(\tau_k) = [b_{k,-1} R_{k,1}(\tau_k) + b_{k,0} \tilde{R}_{k,1}(\tau_k)] \cos \phi_k
\]

then \( \sqrt{P/2} v_{k,1}(\tau_k) \) is the contribution of the \( k \)th signal to the output \( Z_i \) of the correlation receiver matched to \( s_i(t) \). For fixed \( \tau_k \), \( v_{k,1}(\tau_k) \) depends only on \( \phi_k \), the data symbols \( b_{k,-1} \) and \( b_{k,0} \), and the discrete cross-correlation functions \( \rho_{k,1} \) and \( \tilde{\beta}_{k,1} \) (or, equivalently, \( \theta_{k,1} \) and \( \tilde{\theta}_{k,1} \)).

Specifically, if \( \ell_k \) is an integer for which \( \ell_k T_c \leq \tau_k \leq (\ell_k + 1) T_c \) and \( b_{k,0} = b_{k,-1} \) then

\[
v_{k,1}(\tau_k) = b_{k,0} \{ \theta_{k,1}(\ell_k) T_c + [\theta_{k,1}(\ell_k + 1) - \theta_{k,1}(\ell_k)](\tau_k - \ell_k T_c) \} \cos \phi_k.
\]

On the other hand, if \( b_{k,0} \neq b_{k,-1} \) then

\[
v_{k,1}(\tau_k) = b_{k,0} \{ \tilde{\theta}_{k,1}(\ell_k) T_c + [\tilde{\theta}_{k,1}(\ell_k + 1) - \tilde{\theta}_{k,1}(\ell_k)](\tau_k - \ell_k T_c) \} \cos \phi_k.
\]
Up to this point we have not explicitly indicated which parameters of the partial cross-correlation functions should be optimized. The ideal situation would be to find a code for which the error probabilities $\Pr(Z_i > 0|b_{i,0} = -1)$ and $\Pr(Z_i < 0|b_{i,0} = +1)$ are small for all values of the parameters $\tau_k, \theta_k, b_{k,-1},$ and $b_{k,0}$. It is clear from symmetry considerations that for any code, the set of values that one of the two probabilities takes on as the parameters are varied is the same as the corresponding set for the other probability. In particular, the two probabilities have the same maximum value $P_e^{(i)}$ for any given code. One code selection procedure that is often suggested is to choose the code that gives the smallest value of $P_e^{(i)}$; that is, the maximum value of the error probability is minimized. This approach is open to the usual criticism of minimax methods which is that too much emphasis is placed on the worst-case parameter values. However, the minimax approach is warranted for certain systems so we will pursue it further before suggesting an alternate criterion.

If $b_{i,0} = -1$, $P_e^{(i)}$ depends on the maximum value of the sum of the $v_k(\tau_k)$ over all $k \neq i$. From equations (5) and (6) it is clear that the maximum value of $v_k(\tau_k)$ is achieved for some value of $\tau_k$ in the set $\{0, T_c, 2T_c, \ldots, (p - 1)T_c\}$ and $\phi_k = 0$. That is, the maximum value of $v_k(\tau_k)$ is of the form

$$[b_{k,-1}^p k, i(\ell) + b_{k,0}^p k, i(\ell)]T_c$$
for $\ell \in \{0,1,2,\ldots,p-1\}$, $b_{k,-1} \in \{+1,-1\}$, and $b_{k,0} \in \{+1,-1\}$. For a fixed $\ell$, this quantity has four possible values, $\pm \theta_{k,i}(\ell)$ and $\pm \hat{\theta}_{k,i}(\ell)$. The maximum of these values over all values of $\ell$ is $\lambda_{k,i} = \max\{u_{k,i}, \hat{u}_{k,i}\}$ where

$$u_{k,i} = \max\{|\theta_{k,i}(\ell)| \mid \ell = 0,1,2,\ldots,p-1\}$$

and

$$\hat{u}_{k,i} = \max\{|\hat{\theta}_{k,i}(\ell)| \mid \ell = 0,1,2,\ldots,p-1\}.$$

From the above discussion we conclude that if $b_{i,0} = -1$, the maximum error probability for the $i$th receiver corresponds to the maximum value of $v_k(\tau_k)$ for each $k \neq i$ and that this maximum value is $\lambda_{k,i}$. The same argument can be applied for $b_{i,0} = +1$ in which case the maximum error probability corresponds to the minimum value of $v_k(\tau_k)$ for each $k \neq i$ and $\min v_k(\tau_k) = -\lambda_{k,i}$. Thus, $P_e^{(1)}$ is minimized if the quantity

$$\Lambda_i = \sum_{k \neq i} \lambda_{k,i}$$

is minimized.

Of course, the goal in choosing a code is to simultaneously minimize the error probability for all $K$ receivers. Hence we might wish to minimize $P_{e,m} = \max P_e^{(i)}$, which is equivalent to minimizing $\Lambda = \max \Lambda_i$ since

$$P_e^{(i)} = 1 - \Phi([1 - (\Lambda_i/p)]\sqrt{2E/N_0})$$

where $\Phi$ is the standard (i.e., zero mean, unit variance) Gaussian cumulative distribution function and $E = P/T$ is the energy per data bit.
It follows that

\[ P_{e,m} = 1 - \Phi([1 - (\lambda/p)]\sqrt{2E/N_0}) \] \hspace{1cm} (8a)

\[ \leq 1 - \Phi([1 - (K - 1)(\lambda/p)]\sqrt{2E/N_0}) \] \hspace{1cm} (8b)

where \( \lambda \) is the code parameter defined by

\[ \lambda = \max\{\lambda_{k,i} \mid 1 \leq i < k \leq K\}. \]

The property \( \lambda_{k,i} = \lambda_{i,k} \) was used to obtain the inequality (8b).

There are a few additional code parameters that are of interest. These are the maximum magnitude of the periodic cross-correlation functions for the set of sequences

\[ \theta_C = \max\{\mu_{k,i} \mid 1 \leq i < k \leq K\}, \]

the maximum magnitude of the odd cross-correlation

\[ \hat{\theta}_C = \max\{\theta_{k,i} \mid 1 \leq i < k \leq K\}, \]

and the maximum magnitude of the aperiodic cross-correlation

\[ \tilde{\rho}_C = \max\{\tilde{\rho}_{k,i}(\ell) \mid -(p - 1) \leq \ell \leq p - 1, 1 \leq i < k \leq K\}. \]

The following relationships between these parameters are easily established:

\[ \lambda = \max\{\theta_C, \hat{\theta}_C\} \]

\[ \tilde{\rho}_C \leq \lambda \leq 2\tilde{\rho}_C. \] \hspace{1cm} (9)

Inequalities (8b) and (9) imply

\[ P_{e,m} \leq 1 - \Phi([1 - (K - 1)(2\tilde{\rho}_C/p)]\sqrt{2E/N_0}). \] \hspace{1cm} (10)

It should be noted that it is impossible to obtain an upper bound on the error probability if we know only the periodic cross-correlation function.
If the communication engineer were forced to base his design on minimax results alone he would be seriously handicapped. In many practical systems the worst-case error probability is nearly \( \frac{1}{2} \) even though the conditions that lead to this occur only rarely. This is particularly true in the case of SSMA communication systems. The maximum cross-correlation values occur for only a few values of the delay parameters, \( \tau_1, \tau_2, \ldots, \tau_K \) but, when these values do occur, the resulting error probability is unacceptable. In fact, the noise and other user interference will typically exceed the desired signal by more than a factor of 10 when these worst-case parameter values occur. Thus, selection of codes for SSMA systems on the basis of the maximum cross-correlation may not be the best approach.

An alternate approach to SSMA system analysis and code evaluation is to treat the phase shifts, time delays, and data symbols as mutually independent random variables. In this approach the interference terms are random and are treated as additional noise. The signal-to-noise ratio, \( \text{SNR}_i \), at the output of the \( i \)th correlation receiver is the most important performance measure that can be obtained with a reasonable amount of computation. Furthermore, bounds and approximations to the error probability can be given in terms of \( \text{SNR}_i \). The complete derivation of \( \text{SNR}_i \) is given in [72] and [73].

As before, there is no loss of generality in assuming \( \phi_k = 0 \) and \( \tau_k = 0 \) for \( k \neq i \) when considering \( Z_i \), the output of \( i \)th correlation receiver.
Also, due to the symmetry involved we need consider only \( b_{i,0} = +1 \). The desired signal component of \( Z_i \) is then \( \sqrt{P/2} T \) while the variance of the noise component of \( Z_i \) is

\[
\text{Var}(Z_i) = \frac{P}{4T} \sum_{k=1}^{K} \int_0^T R_{k,i}^2(\tau) + \sigma_{k,i}^2(\tau) \, d\tau + \frac{N_0 T}{4}
\]  

where we have assumed \( \phi_k \) is uniformly distributed on \([0,2\pi)\) and \( \tau_k \) is uniformly distributed on \([0,T)\) for \( k \neq i \). In addition, the data symbols \( b_{k,i} \) are assumed to take values \(+1\) or \(-1\) with equal probability. Equations (3) and (4) are then employed to show that (11) is equivalent to

\[
\text{VAR}(Z_i) = \frac{PT}{12p^3} \sum_{k=1}^{K} r_{k,i} + \frac{N_0 T}{4}
\]  

where

\[
r_{k,i} = \sum_{\ell=0}^{p-1} \left\{ \rho_{k,i}(\ell) + \rho_{k,i}(\ell) \rho_{k,i}(\ell + 1) + \rho_{k,i}^2(\ell + 1)
\right. \\
+ \delta_{k,i}^2(\ell) + \delta_{k,i}^2(\ell) \delta_{k,i}(\ell + 1) + \delta_{k,i}^2(\ell + 1) \}
\]  

The code parameter \( r_{k,i} \) can also be written as

\[
r_{k,i} = \sum_{\ell=0}^{p-1} \left\{ \rho_{k,i}(\ell) [2\rho_{k,i}(\ell) + \rho_{k,i}(\ell + 1)] + \delta_{k,i}(\ell) [2\delta_{k,i}(\ell) + \delta_{k,i}(\ell + 1)] \right\}.
\]  

The signal-to-noise ratio is just \( \sqrt{P/2} T \) divided by \( \text{VAR}(Z_i) \) which is

\[
\text{SNR}_i = \left\{ \frac{(6p^3)^{-1}}{k=1} \sum_{k=1}^{K} r_{k,i} + \frac{N_0}{2E} \right\} - \frac{1}{2}.
\]
Note that for $K = 1$ this gives $SNR_1 = \sqrt{2E/N_0}$ which has associated error probability $P_e = 1 - \Phi(\sqrt{2E/N_0})$. In general the error probability is not $1 - \Phi(SNR_1)$ but this is a good approximation* for large values of $p$ and $K$. A more thorough discussion of the signal-to-noise ratio as a performance measure and a numerical evaluation of (14) for Gold sequences of period 511 can be found in [73].

C. Sequences with Good Aperiodic Correlation Properties

It is clear from Section III.B. that the partial cross-correlation functions $\rho_{k,i}$ and $\hat{\rho}_{k,i}$ and the odd cross-correlation function $\hat{\delta}_{k,i}$ are at least as important as the usual periodic cross-correlation function $\theta_{k,i}$. Although this has been recognized for some time by communication engineers, the functions $\rho_{k,i}$, $\hat{\rho}_{k,i}$, and $\hat{\delta}_{k,i}$ have received very little attention in the literature. This is primarily because the theory of linear cyclic codes, which provides the periodic cross-correlation results given in Section III.A., is considerably more difficult to apply to the design of sequences with good partial cross-correlation properties.

Although $|\theta_{k,i}(\ell)| = |\hat{\delta}_{k,i}(\ell) + \rho_{k,i}(\ell)|$ can be made small by selecting the sequences $(a_j^{(k)})$ and $(a_j^{(l)})$ as in Section III.A., the quantities $|\hat{\delta}_{k,i}(\ell)|$ and $|\rho_{k,i}(\ell)|$ that result from such a selection are often quite large. As a result, $|\hat{\delta}_{k,i}(\ell)|$ is large and the case $b_{k,0} \neq b_{k,-1}$ (which leads to equation (6)) will have much greater impact on the system performance than the case $b_{k,0} = b_{k,-1}$ (which leads to equation (5)).

*For some recent results on the accuracy of this approximation for $K << p$, see [107].
One of the first detailed investigations of SSMA system performance which included the effects of the partial cross-correlation functions was published in 1969 by Anderson and Wintz [10] who obtained a bound on the signal-to-noise ratio at the output of the correlation receiver for a SSMA system with a hard-limiter in the channel. The need for considering the partial cross-correlation properties of the code sequences is discussed on page 286 of that paper. Gold had previously investigated partial correlation functions but his work was published only in company reports. In particular, Gold determined the mean and variance of the partial autocorrelation function and the mean of the partial cross-correlation function where the randomness is in the starting position and the phase shift [36]. A thorough investigation of the odd cross-correlation function $\hat{\theta}_{k,1}$ was undertaken by Massey and Uhran in 1969 as part of a study for the NASA tracking and data relay satellite system. The role of the odd autocorrelation function $\hat{\theta}_{i,i}$ in the analysis and design of SSMA systems for multipath channels was one of the main topics of their final report [62]. Other results on the partial autocorrelation and cross-correlation properties of sequences and related topics can be found in [14,24,29,58,60,63,67,72,73,80,83,92,94, and 99].

Equation (8a) shows that if we wish to minimize the maximum error probability, then an optimum code is one for which the parameter $A$ is smallest. Inequalities (8b) and (10) suggest suboptimum code selection
procedures based on minimization of the parameters $\lambda$ or $\tilde{\rho}_C$. In practice it is common to select a general class of codes which have good periodic cross-correlation properties (as in Section III.A.). Typically, if the class has such properties then $\mu_{k,i} \leq \theta_{k,i}$ for each $k \neq i$ or at least $\theta_C \leq \hat{\theta}_C$. If $\theta_C \leq \hat{\theta}_C$ then $\Lambda \leq (K - 1)\hat{\theta}_C$. We conjecture that this inequality is valid for maximal connected sets of m-sequences and Gold sequences. This suggests another suboptimum method for selecting code sequences for phase-coded SSMA systems. A family of sequences for which $\theta_C$ is relatively small (e.g., the Gold sequences) is searched in hope of finding a subfamily for which $\hat{\theta}_C$ is also small. This is basically the approach of Massey and Uhran [62,63] and Sywyk [92] who consider a family of sequences which consists of all distinct shifts of a set of $K$ sequences of period $p$ for which $\theta_C$ is small. The goal is to minimize the parameter $\hat{\theta}_C$ by finding the best shift for each of the $K$ sequences in the set.

For many applications (e.g., multipath channels) it is necessary to consider the partial or aperiodic autocorrelation functions. It will be sufficient for our purposes to consider the parameters

$$\tilde{\rho}_A = \max\{\tilde{\rho}_{k,k}(\ell) \mid 0 < |\ell| \leq p - 1, 1 \leq k \leq K\},$$
$$\theta_A = \max\{\theta_{k,k}(\ell) \mid 0 < \ell \leq p - 1, 1 \leq k \leq K\},$$
$$\hat{\theta}_A = \max\{\hat{\theta}_{k,k}(\ell) \mid 0 < \ell \leq p - 1, 1 \leq k \leq K\},$$
$$\tilde{\rho}_C = \max\{\tilde{\rho}_A, \tilde{\rho}_C\}, \theta_C = \max\{\theta_A, \theta_C\}, \text{ and } \hat{\theta}_C = \max\{\hat{\theta}_A, \hat{\theta}_C\}. $$

As part of the previously mentioned NASA study, Massey and Uhran have investigated minimization of the odd autocorrelation parameter $\hat{\theta}_A$ for
a set of sequences for which \( \theta_A \) is small \([62,63]\). They found the parameter \( \hat{\theta}_A \) was very sensitive to the shifts of the sequences in the set. In particular they discovered there is unique optimum shift for most of the 18 different m-sequences of period 127 \([63]\). This gave a set of 18 sequences for which \( \hat{\theta}_A = 19 \) and, since they are m-sequences, \( \theta_A = 1 \). Massey and Uhran computed the cross-correlation functions for these 18 sequences and found \( \hat{\theta}_C = \hat{\theta}_C = 41 \). We should point out that some of these parameters can be improved (perhaps at the expense of others) if a smaller set of sequences is acceptable. For instance, any maximal connected set of m-sequences of period 127 has 6 sequences with periodic cross-correlation parameter \( \theta_C = 17 \). Roefs \([75]\) has examined the partial cross-correlation properties of some of these maximal connected sets and has found a set for which \( \hat{\theta}_C = 29 \) and \( \hat{\theta}_A = 23 \). These two values can probably be decreased further if an exhaustive search is carried out. Any set consisting of more than 6 m-sequences of period 127 must have \( \theta_C = 41 \). However, Massey and Uhran found a set of 14 m-sequences of period 127 with odd autocorrelation parameter \( \hat{\theta}_A = 17 \). Additional results on codes with good odd autocorrelation properties can be found in \([83]\) and \([94]\).

Two important analytical results on aperiodic cross-correlation properties have recently appeared in the literature. The first of these, which was obtained by Welch \([99]\), provides lower bounds on the maximum correlation for a set of sequences of complex numbers. We mention only the following special case.
Theorem 5: For a set of $K$ binary sequences of period $p$

\[ \theta_{\text{max}} \geq p [(K - 1)/(pK - 1)]^{\frac{1}{2}} \]

and

\[ \tilde{\rho}_{\text{max}} \geq p [(K - 1)/(2pK - K - 1)]^{\frac{1}{2}}. \]

Note that for large values of $K$ and $p$ the lower bound on $\theta_{\text{max}}$ is approximately $\sqrt{p}$ and the lower bound on $\tilde{\rho}_{\text{max}}$ is approximately $\sqrt{p/2}$.

For the $K = 2^n + 1$ Gold sequences of period $p = 2^n - 1$ the lower bound on $\theta_{\text{max}}$ is approximately $2^{n/2}$ for large $n$ whereas the actual values of $\theta_{\text{max}}$ is $2\lfloor (n+2)/2 \rfloor + 1$ which is greater than the bound by a factor of $\sqrt{2}$ for $n$ odd and by a factor of 2 for $n$ even.

The second result, which is due to Massey and Uhran [63], is perhaps even more significant in that it provides a method for constructing a set of sequences with a known bound on the odd cross-correlation function.

Furthermore, as a special case, it gives a bound on the odd cross-correlation function for the Gold sequences.

Theorem 6: Suppose $(1 + x)$ does not divide $f(x)$ and let $h(x) = (1 + x)f(x)$. Suppose the sequences generated by $h(x)$ have maximum periodic cross-correlation $c$. Then the set of sequences generated by $f(x)$ has correlation parameters $\theta_{\text{max}} \leq c$ and $\hat{\theta}_{\text{max}} \leq (p + c + 4)/2$ where $p$ is the least common multiple of the periods of the sequences generated by $f(x)$.

Massey and Uhran [63] have also given a method of constructing an infinite sequence of codes of increasing block length for which
$p^{-1} \hat{\theta}_{\text{max}}$ and $p^{-1} \hat{\theta}_{\text{max}}$ can be made arbitrarily small for large $p$. Note that even if we know $p^{-1} c \to 0$, the bound on $\hat{\theta}_{\text{max}}$ obtained from Theorem 6 implies only $\limsup_{p \to \infty} p^{-1} \hat{\theta}_{\text{max}} \leq 1/2$. The stronger result $p^{-1} \hat{\theta}_{\text{max}} \to 0$ is obtained by requiring $f(x) = f_1(x)f_t(x)$ in Theorem 6 where $f_1(x)$ is the minimal polynomial of $\alpha$, a primitive element of $\text{GF}(2^n)$, and $f_t(x)$ is the minimal polynomial of $\alpha^t$ for $t = Kp[2(K-1)]^{-1}$. It is required that $p = 2^n - 1$ and that the number of codewords $K$ is $2^k$ where $k$ is a nontrivial divisor of $n$, so the codes have block lengths $p = 2^{2k} - 1, 2^{3k} - 1, 2^{4k} - 1, \ldots$. The difficulty with this code is that the bandwidth expansion coefficient $pK^{-1} \to \infty$ as the block length increases. Schneider and Orr [79] have demonstrated the existence of sequences of codes which do not suffer from this difficulty but still achieve small values of $p^{-1} \theta_{\text{max}}$ and $p^{-1} \hat{\theta}_{\text{max}}$ only as $p \to \infty$. An important feature of the multiple-access problem is that the number of users and the available bandwidth are both constrained. Thus, unlike the error-control coding problem where $p$ can become large provided the rate $p^{-1} \log_2 K$ is fixed, the SSMA coding problem requires that $K$ be fixed and $pK^{-1}$ be not greater than some constant $W_0$ determined by the bandwidth constraint. Thus asymptotic results for $p \to \infty$ (and especially for $pK^{-1} \to \infty$) are considerably less useful in SSMA coding than in error-control coding. Thus, it would seem that the most important contribution of a result such as Theorem 6 is the code construction technique and the bound provided for moderate values of $p$ rather than the information supplied about asymptotic performance as $p \to \infty$. 
Theorem 6 is the only analytical result that we have found on the construction of sets sequences with a specified bound on both the odd autocorrelation and odd cross-correlation functions. For Gold codes generated by \( f(x) = f_1(x)f_2(x) \) where \( \{f_1, f_2\} \) is a preferred pair of primitive polynomials of degree \( n \), the bounds become

\[
\theta_{\text{max}} \leq 2^{\left\lfloor \frac{n+2}{2} \right\rfloor + 1}
\]

and

\[
\hat{\theta}_{\text{max}} \leq 2^{n-1} + 2^{\left\lfloor \frac{n}{2} \right\rfloor + 2}.
\]

Since \( \tilde{\rho}_{\text{max}} \leq \max\{\theta_{\text{max}}, \hat{\theta}_{\text{max}}\} \), the upper bound on \( \tilde{\rho}_{\text{max}} \) for Gold codes is greater than \( \rho/2 \) which is approximately the square of the lower bound provided by Theorem 5.

There seems to be no justification for considering only the maximum correlation when selecting sequences for SSMA applications. For instance, the number of different sets of values for \( \tau_1, \tau_2, \ldots, \tau_K \) which yield the maximum correlation would seem to be as important as the maximum value itself. In the remainder of this section we discuss criteria that are related to the signal-to-noise ratio. These criteria offer a number of advantages over the maximum correlation which will be pointed out in the discussion.

The analysis of Section III B shows that the signal-to-noise ratio at the output of the \( i \)th correlation is maximized if the parameter

\[
\bar{r}_i = \sum_{k=1}^{K} r_{k,i} \]

is minimized. If we focus attention on \( \min \{\text{SNR}_i | 1 \leq i \leq K \} \)

then \( \bar{r}_{\text{max}} = \max \{\bar{r}_i | 1 \leq i \leq K \} \) is the parameter to be minimized. However,
it may be easier in certain cases to minimize the parameter
\[ r_{\text{max}} = \max \{ r_{k,i} \mid 1 \leq i < k \leq K \} \]  
(notice that \( r_{i,k} = r_{k,i} \)) and we expect the results will be nearly the same. Either of these two parameters is obtained by averaging over the various relative time shifts which is the main feature that we wished to introduce by considering random time delays. In practice, the set of sequences selected on the basis of either \( r_{\text{max}} \) or \( r_{\text{max}} \) is seldom the same as the set selected by considering peak cross-correlation parameters.

We have observed experimentally, and it can be shown analytically for random sequences, that for large values of \( p \)

\[
\left| \sum_{\ell=0}^{p-1} \tilde{\rho}_{k,i}(\ell) \rho_{k,i}(\ell + 1) + \tilde{\rho}_{k,i}(\ell) \tilde{\rho}_{k,i}(\ell + 1) \right| < \sum_{\ell=0}^{p-1} \rho_{k,i}^2(\ell) + \tilde{\rho}_{k,i}^2(\ell).
\]

Therefore, (14) implies \( r_{k,i} \leq \tilde{r}_{k,i} \) where

\[
\tilde{r}_{k,i} = 2 \sum_{\ell=0}^{p-1} \rho_{k,i}^2(\ell) + \tilde{\rho}_{k,i}^2(\ell) = 2 \sum_{\ell=1-p}^{p-1} \rho_{k,i}^2(\ell).
\]  

(16)

An attractive feature of using \( \tilde{r}_{k,i} \) as a code selection procedure is the computational saving due to the fact that \( \tilde{r}_{k,i} \) can be computed from the aperiodic autocorrelation functions. Since this result appears to be new, a proof is included.*

*All of the results in the remainder of this section originated in a joint research project involving the author and D. V. Sarwate, also of the Coordinated Science Laboratory. A joint report is currently being prepared on these and related results [74].
Theorem 7: Let \((x_j)\) and \((y_j)\) be any (not necessarily binary) sequences for which \(x_{j+p} = x_j\) and \(y_{j+p} = y_j\). The aperiodic cross-correlation function \(\rho_{x,y}\) is related to the aperiodic autocorrelation functions \(\rho_{x,x}\) and \(\rho_{y,y}\) by

\[
\sum_{\ell=1-p}^{p-1} \rho_{x,y}(\ell) = \sum_{\ell=1-p}^{p-1} \rho_{x,x}(\ell) \rho_{y,y}(\ell)
\]

Proof: Following Welch [99] we define sequences \((x'_j)\) and \((y'_j)\) by \(x'_j = x_j\) and \(y'_j = y_j\) for \(0 \leq j < p\) and \(x'_j = y'_j = 0\) for \(p \leq j < 2p-2\). We then let \(x'_j+m(2p-1) = x'_j\) and \(y'_j+m(2p-1) = y'_j\) for \(0 \leq j < 2p-2\) and all integers \(m\). The periodic cross-correlation of \((x'_j)\) and \((y'_j)\) can then be related to the aperiodic cross-correlation of the original sequences \((x_j)\) and \((y_j)\). From the definitions of the periodic and aperiodic cross-correlation functions we have

\[
\theta_{x',y'}(\ell) = \sum_{j=0}^{2p-2} x'_j y'_{j+\ell} = \begin{cases} 
\sum_{j=0}^{p-1-\ell} x_j y_{j+\ell}, & 0 \leq \ell \leq p-1 \\
\sum_{j=p-1}^{p-1+\ell} x_j y_{j-\ell}, & 1-p \leq \ell < 0.
\end{cases}
\]

and

\[
\tilde{\rho}_{x,y}(\ell) = \begin{cases} 
\sum_{j=0}^{p-1-\ell} x_j y_{j+\ell}, & 0 \leq \ell \leq p-1 \\
\sum_{j=p-1}^{p-1+\ell} x_j y_{j-\ell}, & 1-p \leq \ell < 0.
\end{cases}
\]

Therefore, we have

\[
\tilde{\rho}_{x,y}(\ell) = \begin{cases} 
\theta_{x',y'}(\ell), & 0 \leq \ell < p-1 \\
\theta_{x',y'}(2p+\ell-1), & 1-p \leq \ell < 0.
\end{cases}
\]
or, equivalently,

\[ \theta_{x',y'}(\xi) = \begin{cases} \beta_{x,y}(\xi), & 0 \leq \xi \leq p-1 \\ \beta_{x,y}(\xi+1-2p), & p \leq \xi < 2p-1 \end{cases} \]

Gold [37] has shown that for sequences \((x'_j)\) and \((y'_j)\) of period \(p'\),

\[ \sum_{\xi=0}^{p'-1} (x'_\xi)(y'_{\xi+1-2p}) = \sum_{\xi=0}^{p'-1} \beta_{x,y}(\xi) \theta_{y',y'}(\xi). \]

Letting \(p' = 2p-1\) and substituting for \(\theta_{x',y'}(\xi)\), \(\theta_{x',x'}(\xi)\), and \(\theta_{y',y'}(\xi)\) in terms of \(\beta_{x,y}(\xi), \beta_{x,x}(\xi), \) and \(\beta_{y,y}(\xi)\) gives the desired result.

An application of Cauchy's inequality gives the following bound on the sum of the squares of the aperiodic cross-correlation values in terms of the sum of the squares of the aperiodic autocorrelation values.

**Corollary 7.1:** For sequences \((x_j)\) and \((y_j)\) as in Theorem 7

\[ \sum_{\xi=1-p}^{p-1} \beta_{x,y}^2(\xi) \leq \left\{ \sum_{\xi=1-p}^{p-1} \beta_{x,x}^2(\xi) \right\}^{1/2} \left\{ \sum_{\xi=1-p}^{p-1} \beta_{y,y}^2(\xi) \right\}^{1/2}. \]

We also get the following result for binary sequences \((\{+1, -1\}\)-valued sequences).

**Corollary 7.2:** If the sequences \((x_j)\) and \((y_j)\) in Theorem 7 are binary sequences then

\[ \sum_{\xi=1-p}^{p-1} \beta_{x,y}^2(\xi) \leq \left\{ \sum_{\xi=1-p}^{p-1} \beta_{x,x}^2(\xi) \right\}^{1/2} \left\{ \sum_{\xi=1-p}^{p-1} \beta_{y,y}^2(\xi) \right\}^{1/2}. \]

We get an interesting result if we apply these bounds to the Barker sequence.
of period 7 obtained by setting \((x_0, x_1, \ldots, x_6) = (1, -1, 1, 1, -1, -1, -1)\)

If \((y_0, y_1, \ldots, y_6) = (x_6, x_5, \ldots, x_0)\) then

\[
\sum_{\ell=0}^{6} \theta^2_{x,y}(\ell) = 55 = \left\{ \sum_{\ell=-6}^{6} \beta^2_{x,x}(\ell) \right\}^{1/2} \left\{ \sum_{\ell=6}^{6} \beta^2_{y,y}(\ell) \right\}^{1/2}.
\]

Thus, for these two sequences the inequalities in Corollaries 7.1 and 7.2 are actually equalities. Furthermore, it is easy to see that

\[
\sum_{\ell=0}^{6} \theta^2_{x,y}(\ell) = 55 = \left\{ \sum_{\ell=0}^{6} \theta^2_{x,x}(\ell) \right\}^{1/2} \left\{ \sum_{\ell=0}^{6} \theta^2_{y,y}(\ell) \right\}^{1/2},
\]

since the sequences above are also m-sequences and, according to (2), the sum of the squares of the periodic cross-correlation values for any two m-sequences of period \(p\) is always \(p^2 + p - 1\) which is also equal to the sum of the squares of the autocorrelation function of any m-sequence of period \(p\).

We note that the right side of (18) is larger than \(p^2 + p - 1\) unless the sequences are Barker sequences. Furthermore, since equality holds in the Cauchy inequality if and only if there exists a number \(\beta\) such that

\[
\beta_{x,x}(\ell) = \beta_{y,y}(\ell) \quad \text{for} \quad 1 - p \leq \ell \leq p - 1 \quad \text{and since} \quad \beta_{x,x}(0) = \beta_{y,y}(0) = p
\]

for any binary sequences \((x_1)\) and \((y_1)\) of period \(p\), then equality holds in (18) if and only if \(\beta_{x,x}(\ell) = \beta_{y,y}(\ell)\) for \(1 - p \leq \ell \leq p - 1\). This will be the case when \((x_0, x_1, \ldots, x_{p-1}) = (y_p, y_{p-1}, \ldots, y_0)\) for example.
IV. Random-Access Systems

Random-access schemes appear to be particularly attractive for systems such as computer-communication networks where there are many data sources each of which is actually transmitting data only a small fraction of the time. The ALOHA system random-access technique [1] is an example of the type of scheme we have in mind. In ALOHA-type systems several terminals can communicate with each other via a single communication channel. At each terminal the packets are generated so infrequently that with very high probability the channel will be clear when a given terminal is ready to transmit. Channel usage is therefore not allocated in any fixed way as it is in FDMA or TDMA, since fixed allocation is known to be extremely inefficient for systems of this type.

Occasionally two or more terminals will attempt to transmit packets simultaneously which, at least with very high probability, produces errors in all of the packets involved. Error-control coding is essential in a system of this type to at least detect the presence of such errors. An error-detecting code is employed in the ALOHA system [1]. In most random-access systems, the receiving terminal acknowledges correct reception of a packet if no errors are detected. If errors are detected, no acknowledgement is sent and the transmitting terminal automatically retransmits the packet after a certain period of time. For a synchronous satellite system, the propagation time for the complete round trip (from the transmitter to the receiver and back) is about 1/2 second. The transmitter must wait somewhat longer than this before concluding the message was not received correctly and should therefore be retransmitted.
Actually, random-access schemes were used in communication and navigation systems in the 1940's, long before their application to communication networks. The study of random-access techniques was carried out under the name asynchronous multiplexing during the 1940's. One of the first published theoretical investigations was a paper by White [101] which, like most of the early work on multiple-access systems, was concerned with low duty cycle, on-off, pulse communication. In such systems each user can turn the channel "on" but all users must be transmitting an "off" signal for the channel to be in the "off" state. Thus, when a given user transmits an "on" he will interfere with any other user who is transmitting an "off". This scheme requires the use of low duty factor signals (i.e., an "on" is transmitted with much lower probability than an "off"), and for practical applications the resulting interference must be reduced via coding methods. This multiple-access technique has been discussed more recently by Taylor [93] and Sommer [87-89].

Investigations of random-access methods continued in the 1960's under such names as pulse-address multiple-access (PAMA) and random-access discrete-address (RADA). As defined in [81], RADA can be considered as a special case of PAMA. Both use addresses consisting of time-frequency coded pulses. In RADA, each receiver has a set of addresses which are permanently assigned and are used only for transmitting to that receiver. In PAMA, no such fixed assignment is made. For the general RADA technique, each receiver is allowed to have more than one address. The disadvantage of the single address per user, on-off RADA systems is that they achieve low error rates only with very low channel utilization. Channel utilization
is defined as the total data rate for all transmitters divided by the total channel bandwidth. The RADA system with many addresses per user can achieve low error probability with channel utilization up to 0.69 [17]. Of course, for a system with many users, this still requires a low duty factor for each of the individual transmitters. In fact, all of the PAMA techniques require low duty factor signals in order to achieve reliable communication [81].

A specific coding method for the random-access schemes described above has been proposed [20]. This coding method uses orthogonal convolutional codes and the Viterbi decoding algorithm to correct errors introduced by the multiple-access interference. This scheme is efficient whenever the multiple-access interference is the dominant source of noise in the system. However, this does not appear to be an efficient coding method whenever there is significant channel interference due to other sources such as receiver thermal noise. The effects of thermal noise were not considered in [20].

The computer-communication network appears to be ideally suited for random-access techniques. Information transmission in such a network is accurately characterized as burst communication as opposed to the continuous communication encountered in the data relay satellite system described in Section II. This is because the users in a computer-communication network are typically transmitting only a small fraction of the time. Thus, if a communication satellite is to be employed as part of the network, the usual TDMA and FDMA techniques are very inefficient.
There are a number of benefits that will be obtained by including a communication satellite in a computer-communication network. These are to a large extent due to the data transmission capabilities of communication satellites as compared with telephone lines. In particular, let us compare the multiple-access broadcast satellite system shown in Figure 2 with the telephone cable system. At present, for long distance communication, higher data rates can be transmitted via satellite. For example, one Intelsat IV channel can handle about 56,000 bits per second which is more than five times greater than the maximum rate for a private leased telephone line. Furthermore, in the future, and to some extent even at present for military satellites, data rates in excess of $10^8$ bits per second are possible for wideband satellite channels. The upper limit on data rate for future wideband telephone facilities appears to be about $10^5$ bits per second. The broadcast feature of the satellite system of Figure 2 permits each user to transmit a single signal which will be received by all of the terminals in the network. In addition, for an appropriately chosen multiple-access technique (e.g., SSMA), the satellite system permits all terminals to simultaneously transmit over the single channel. Neither of these two features is available in telephone line systems. Finally, as pointed out by Abramson [2], the multiple-access broadcast satellite system allows the transmitting terminal to monitor the transmitted message as it is retransmitted from the satellite to the receivers. Abramson calls this perfect information feedback. However, if the noise in the terminals' receivers is taken into account, it is clear that this information feedback is not perfect. Information feedback of this type is not available in telephone line networks.
Even though the feedback in the multiple-access broadcast satellite system is not perfect, it does supply information that could be used to improve the error rate and to decrease the delay in the system. For systems of interest, we would expect to be able to detect the presence of an interfering signal by monitoring the message as it is retransmitted from the satellite. After detecting an interfering signal, the terminal could retransmit the message. For example, we might simply count the number of errors present in the monitored signal. If the monitoring receiver detects more than L errors, the message is automatically retransmitted. If the number of errors detected by the monitoring receiver is no more than L, the transmitter waits for an acknowledgement as usual. It is quite possible that L = 0 would be optimum for certain systems, depending on the error rate and delay that can be tolerated. We expect that such a scheme can be used to decrease the delay or the error rate (or both). Before this can be confirmed, we need a better understanding of the relationship between L, the delay, and the eventual error rate.* The motivation for pursuing this topic further stems from the fact that in a satellite system the propagation time is significant. As a result, large delays are introduced by waiting to receive an acknowledgement or a request for retransmission (1/4 second for a synchronous satellite). The role of error-correcting codes for this scheme will be to permit larger values of L and hence reduce the average number of retransmissions required to achieve a given error rate.

* By "eventual error rate" we mean the probability of error for the decoded message at the final destination, after all retransmissions (if any) have been received.
V. Application of Information Theory to Multiple-Access Communications

The first applications of information theory to multiple-access communications were concerned with the low duty cycle, on-off, pulse communication systems discussed in Section IV. This includes the work of White [101]; Taylor [93]; Sommer [87-89]; Chesler [17-18]; and Cohen, Heller and Viterbi [20]. Most of the recent work has addressed the general multi-user communication system with no restrictions on the type of digital signaling to be employed. One of the first important results on this topic was presented in a paper by Shannon [85] on two-way communication between a pair of terminals. The model is roughly that of Figure 3 with \( K = 1 \) and \( J = 1 \).

There has been a great deal of very recent information-theoretic work on multiple-access communications. For instance Ahlswede [7] and Liao [59] have investigated the multiple-access system in which several terminals transmit over a common discrete memoryless channel to a single receiver. This system was also considered by Slepian and Wolf [86] and extended to permit more than one receiver by Ulrey [95]. Similar investigations for multiple-access communication over the additive white Gaussian noise channel have been carried out by Wyner [106] and by Cover [21].

The theory of information transmission for multiple-access channels is not nearly as complete as for the single-input, single-output channels. For the single-user communication system the information transmission theorem states roughly that for \( D > 0 \) a source with rate-distortion function \( R(D) \) can be encoded, transmitted over a channel with capacity \( C \), and decoded at the receiver in such a way that the average distortion is at most \( D \) provided that \( R(D) < C \). For a single user system this theorem can be
divided into two parts: the source coding theorem and the channel coding theorem. These individual theorems have been established for quite general sources [12,22,28] and channels [28,47,70]. For the general multi-user system (such as a multiple-access system), the separation of the information transmission theorem into a channel coding theorem and a source coding theorem is considerably more difficult. Because of this, there are almost no results on the information transmission theorem for the multi-user systems, even though there are a number of results on each of the two individual theorems. A more thorough discussion of the present status of multi-user information theory can be found in two excellent survey articles by Wyner [106] and by Cover [21].

It may be worthwhile at this point to again mention the problem of modeling the satellite multiple-access system. The following statements are valid for the majority of CDMA satellite systems that have been proposed:

1. The channel has memory.
2. The various transmitters in the system are not synchronized.
3. The information sources are independent.

Almost none of the models used to date in the work on multi-user information theory apply to a system with the first two characteristics. The third characteristic suggests that the work on multiple-access systems with correlated sources is of less interest for satellite systems than for some other multiple-access systems. These characteristics can be incorporated into the models discussed in Section II. In fact the analysis...
of the phase-coded SSMA system given in Section III.B is for a system with independent, unsynchronized sources. The work of Massey and Uhran [62] suggests this analysis can be extended to systems in which the channel memory is due to multipath distortion. It is conceptually straightforward to extend the analysis to systems in which the memory is due to linear filtering in the channel. Thus, some of the important channel models which have the three characteristics listed can be analysed via the methods of Section III.B. However, all such models appear to be very difficult to handle from an information-theoretic point of view. For example, even for a completely synchronized multiple-access system, the information transmission theorem for channels with memory is an open problem [86]. Other open problems in multiple-access information theory are described in [21,86, and 106].

In Section IV we mentioned the possibility of using information feedback in a multiple-access broadcast satellite system. It is known that for a single user communication system, the use of feedback cannot increase the capacity of the channel [84,48]. However Gaardner and Wolf [26] have recently demonstrated that the capacity of a multiple-access channel can be increased by using a feedback link. This provides some additional incentive for studying the information feedback problem discussed in Section IV.
Acknowledgement

The author wishes to thank Professors F. P. Preparata and D. V. Sarwate and Mr. H. F. A. Roefs, all of the Coordinated Science Laboratory, for many helpful discussions on various aspects of the multiple-access problem. In particular, the comments and unpublished notes of Professor Sarwate were very beneficial in the preparation of portions of Section III.A.
References and Selected Bibliography


