**Final Report** of ONR Contract N00014-87-K-0054 on the fundamental limits of Code Division Multiple Access communications. It includes the Publications/Patents/Presentations/Honors report; a summary of publications sponsored by the contract and reprints of the key publications sponsored by the contract.
THE MULTIUSER INFORMATION THEORY OF
CODE-DIVISION MULTIPLE-ACCESS CHANNELS

Sergio Verdú, Principal Investigator
Department of Electrical Engineering
Princeton University
Princeton, New Jersey 08544
Phone: (609) 258-5315
Email: Verdu@princeton.edu

Reproduction in whole, or in part, is permitted for any purpose of the United States Government.
Office of Naval Research
Publications/Patents/Presentations/Honors Report
For
1 January 1987 - 31 March 1990

SET Number: 417J010-01

Contract/Grant Number: N00014-87-K-0054

Contract/Grant Title: The Multiuser Information Theory of Code-Division Multiple-Access Channels

Principal Investigator: Sergio Verdú

Mailing Address: Department of Electrical Engineering
Princeton University
Princeton, NJ 08544

Phone Number: (with Area Code) (609) 258-5315

Email Address: verdu@princeton.edu

1. Number of Papers Submitted to Refereed Journal but not yet published: 6

2. Number of Papers Published in Refereed Journals: 6

3. Number of Books or Chapters Submitted but not yet Published: 1

4. Number of Books or Chapters Published: 0

5. Number of Printed Technical Reports & Non-Refereed Papers: 18

6. Number of Patents Filed: 0

7. Number of Patents Granted: 0

8. Number of Invited Presentations at Workshops or Prof. Society Meetings: 5

9. Number of Presentations at Workshops or Prof. Society Meetings: 18

10. Honors/Awards/Prizes for Contract/Grant Employees: 16

11. Total number of Graduate Students and Post-Docs Supported at least 25% this year on this contract/grant: Grad Students 3 and Post-Docs _____.

How many of each are females or minorities?

These 6 numbers are for ONR's EEO/Minority reports; minorities include Blacks, Aleuts, Asians, etc and those of Hispanic or Asian extraction or nationality. The Asians are singled out to facilitate the varying report semantics re "economically disadvantaged".

Grad Student Female
Grad Student Minority
Grad Stu Asian Ext 0
Post-Doc Female
Post-Doc Minority
Post-Doc Asian Ext.
SUMMARY OF WORK SPONSORED BY CONTRACT N00014-87-K-0054

1. Refereed Journal Papers not yet published


2. Papers Published in Refereed Journals


C. Book Chapters not yet published


E. Conference Publications


D. Brady and S. Verdú, “Error Probability of CDMA in the Poisson Channel,” Invited paper, 1989 Workshop on Information Theory, Cornell University,


**J. Honors/Awards/Prizes for Contract Employees**

*S. Verdú, Principal Investigator*

A. Rheinstein Award 1987, Outstanding Junior Faculty Member, School of Engineering and Applied Science, Princeton University.

NSF Presidential Young Investigator Award, 1988

IEEE Senior Member, 1988
promoted to Associate Professor, Princeton University, 1989

IEEE Information Theory Society Board of Governors, Elected 1989

Engineering Council Award for Excellence in Teaching, 1989
School of Engineering and Applied Science, Princeton University.

Appointed Visiting Fellow, Institute of Advanced Studies, Australian National University, Canberra, Australia, Sep-Oct 1989


Delegate of the IEEE Information Theory Society, IEEE Region 10 Tour, Oct. 1989,

Member, Awards Committee, IEEE Technical Activities Board, 1990

Co-Chairman, 1990 IEEE Information Theory Workshop, Veldhoven, The Netherlands, June 1990


J. Ghez, Research Assistant

Vallace Memorial Fellowship in Engineering, Graduate School, Princeton University, 1987-88

L. Cheng, Research Assistant

Meitec Corporation Junior Fellowship Award, 1989

George Van Ness Lothrop Fellowship, Graduate School, Princeton University, 1990-91
Stability Properties of Slotted Aloha with Multipacket Reception Capability

Sylvie Ghez
Sergio Verdú
Stuart C. Schwartz
Stability Properties of Slotted Aloha with Multipacket Reception Capability

TYLVE GHEZ, STUDENT MEMBER, IEEE, SERGIO VERDÜ, MEMBER, IEEE, AND STUART C. SCHWARTZ, SENIOR MEMBER, IEEE

Abstract—The stability of the Aloha random access algorithm in an infinite-user slotted channel with multipacket reception capability is considered. This channel is a generalization of the usual collision channel, in that it allows the correct reception of one or more packets involved in a collision. The number of successfully received packets in each slot is modeled as a random variable which depends exclusively on the number of simultaneous attempted transmissions. This general model includes as special cases channels with capture, noise, and code division multiplexing. It is shown by means of drift analysis that the channel backlog Markov chain is ergodic if the packet arrival rate is less than the expected number of packets successfully received in a collision of size $n$, as $n$ goes to infinity. Finally, the properties of the backlog in the nonergodicity region are examined.

I. INTRODUCTION

One of the main problems in random access communications is the determination of the maximum stable throughput. In particular, an important result is that the Aloha protocol is unstable [1]–[3] in an infinite-user slotted collision channel where a transmission is successful only if no other users attempt simultaneous transmissions. Several strategies have been designed to stabilize this channel, such as collision resolution algorithms [see (4), for example] where transmissions are deferred until the current conflict is solved, and more recently, Aloha-type strategies using decentralized control, where the retransmission probability is updated according to previous channel outcomes. It has been shown [5]–[7] that the maximum stable throughput achievable by such Aloha-type strategies with decentralized control is $e^{-1}$.

However, the collision channel model does not hold in many important practical multiuser communication systems [8]–[21] because simultaneous transmission of several packets does not necessarily result in the destruction of all the transmitted information. For instance, the capture phenomenon is common in local area radio networks [12]–[15]; if the power of one of the received packets is sufficiently large compared to the power of the other packets involved in a collision, then the strongest packet can be correctly decoded, while the other packets are lost. Other examples are multiple-access channels where several users transmit simultaneously in the same frequency band, and a multuser detector demodulates the information transmitted by all active users (e.g., [8]–[11]). Although those systems do not necessarily require a random access protocol, it is sometimes useful to exercise some flow control through such a protocol so as to limit the maximum number of simultaneous transmitters, in order to bound the multuser receiver complexity and guarantee lower bit-error rates.

Previous studies of some of the aforementioned systems [9], [12]–[18] where some of the packets involved in a collision may be correctly received have shown that the performances are noticeably improved with respect to slotted Aloha. However, even in those special cases, no precise stability result is available; either because finite population networks with no buffer space were considered, or because the Poisson approximation of channel traffic was used for infinite population networks. In [19] (see also [20]), upper and lower bounds are derived for the capacity of a multiple access channel where all packets are correctly received if the collision size does not exceed a fixed threshold and otherwise all packets are destroyed.

In this paper, we consider a generalization of the collision channel, where the receiver can demodulate several packets simultaneously. It is assumed that the number of correctly demodulated packets is a random variable, which, given the number of packets simultaneously transmitted, is independent of the backoff and of the number of previous retransmission attempts. This random variable can take any integer value between zero and the collision size. Thus, the channel is described by a matrix of conditional probabilities $(e_k)$ where $e_k$ is the probability that $k$ packets are correctly demodulated given that there were $n$ simultaneous transmissions. We analyze the usual Aloha algorithm with the multipacket reception capability just described. Users are synchronized so that transmissions take place within one slot, and at the end of each slot, stations that did transmit a packet learn whether or not their transmission was successful. Unsuccessful or backlogged packets are retransmitted in each subsequent slot with probability $p$, $0 < p < 1$. It turns out that multipacket reception capability can stabilize Aloha. Our main result states that the maximum stable throughput is equal to the limit of the average number of packets correctly received in collisions of size $n$ when $n$ goes to infinity. To show this, we model the channel backoff as a Markov chain, and then study its properties by using some simple drift analysis techniques.

The last part of this paper is a study of the properties of the backlog in the nonergodicity region. Unlike the backlog Markov chain for slotted Aloha which is always transient [1], the backlog for our model does in general have a null recurrence region of positive length, which depends on the matrix $e_k$ and on the retransmission probability $p$. However, transience in the nonergodicity region can be ensured for a large class of systems, and in particular for channels where the number of successful simultaneous transmissions is bounded.

II. MULTIPACKET RECEPTION MODEL

Let $A_k$ be the number of new packets arriving during time slot $k$. Assume that $(A_k)_{k=0}$ are i.i.d. random variables with probability distribution:

$$P[A_k=n] = \lambda_n \quad (n \geq 0)$$

such that the mean arrival rate $\lambda = \sum_{n=1}^{\infty} n \lambda_n$ is finite. New packets are transmitted with probability one at the beginning of the first slot following their arrival.

Given that $n$ packets are being transmitted in one slot, we define
for \( n \geq 1, 0 \leq k \leq n \).

\[ e_{ak} = P[k \text{ packets are correctly received} \mid n \text{ are transmitted}] \]

The multipacket reception properties of the channel are summarized by the stochastic matrix

\[
E = \begin{bmatrix}
\epsilon_{00} & \epsilon_{01} & \epsilon_{02} & 0 \\
\epsilon_{10} & \epsilon_{11} & \epsilon_{12} & 0 \\
\epsilon_{20} & \epsilon_{21} & \epsilon_{22} & 0 \\
\vdots & \vdots & \vdots & \ddots \\
\epsilon_{n0} & \epsilon_{n1} & \epsilon_{n2} & \cdots & \epsilon_{nn}
\end{bmatrix}
\]

which we refer to as the \textit{reception matrix} of the channel. For instance, the reception matrix for the usual collision channel is

\[
\begin{bmatrix}
0 & 1 & 0 \\
1 & 0 & 0 \\
0 & 1 & 0 \\
\vdots & \vdots & \ddots \\
0 & 1 & 0 \\
0 & 1 & 0 \\
\end{bmatrix}
\]

while for a system with capture it has the form

\[
\begin{bmatrix}
0 & 1 & 1-x_n & x_n & 0 \\
1-x_n & 1 & x_n & 0 & 0 \\
0 & 1-x_n & 1 & x_n & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots \\
0 & 1 & x_n & 0 & 0 \\
0 & 1 & x_n & 0 & 0 \\
\end{bmatrix}
\]

where \( x_n \) is the probability of capture given that the collision size is \( n \). The model studied in [19], [20] can be described by a reception matrix of the form

\[
\begin{bmatrix}
0 & 1 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 1 & 0 \\
\vdots & \vdots & \vdots & \ddots \\
0 & 0 & 1 & 0 \\
0 & 0 & 1 & 0 \\
\end{bmatrix}
\]

Note that by letting \( \epsilon_{00} \neq 0 \) our model allows not only collisions but also background noise to be a source of errors.

Denote by \( X_n \) the number of backlogged packets in the system at the beginning of slot \( n \). The discrete-time process \((X_n)_{n \geq 0}\) is easily seen to be a homogeneous Markov chain. We define the system to be stable if \((X_n)_{n \geq 0}\) is ergodic and unstable otherwise. The average number of packets correctly received in collisions of size \( n \) is denoted by \( C_n = \Sigma_{k=1}^{n} e_{nk} \). We can now state the main result.

\textbf{Theorem 1:} If \( C_n \) has a limit \( C = \lim_{n \to \infty} C_n \), then the system is stable for all arrival distributions such that \( \lambda < C \) and is unstable for \( \lambda > C \). This also holds if \( C \) is infinite: if \( \lim_{n \to \infty} C_n = +\infty \), then the system is always stable.

The proof is given in Section III. In the remainder of this section, we use Theorem 1 to analyze several simple random access channels that fall within the scope of the multipacket reception channel.

1) \textbf{Mobile Users with Pairwise Transmissions:} Consider an infinite number of transmitters \( T_1, T_2, \ldots \), and an infinite number of receivers \( R_1, R_2, \ldots \), whose positions in the plane are i.i.d. random variables. Suppose that transmissions are pairwise

This result holds under the assumption that the Markov chain of the number of backlogged packets is irreducible and aperiodic (for details and sufficient conditions, see Section III).

in the sense that transmitter \( T_n \) sends packets only to receiver \( R_n \), and \( R_n \) is only interested in the packets sent by \( T_n \) (see Fig. 1). Assume also that each receiver can only detect correctly the packet sent by the closest transmitter (in particular, this is the case if there is perfect capture, see Example 3 below). The successes of transmissions occurring at the same time are independent, so that for \( n \geq 2 \)

\[ e_{nk} = {n \choose k} p(n)^k (1-p(n))^{n-k} \]

where \( p(n) \) is the probability that any given transmitter is successful in a collision of size \( n \), which is equal to \( 1/n \) if we assume that all locations are memoryless, i.e., independent from slot to slot. It follows that

\[ C_n = np(n) = 1 \]

and the maximum throughput is 1. More generally, if because of channel noise, the message of the closest transmitter is received correctly with probability \( \alpha \) (in other words \( \epsilon_{11} = \alpha \)), then the throughput is equal to \( \alpha \). The assumption that the locations of the stations are memoryless is equivalent to assuming that they move infinitely fast. If this simplifying assumption is dropped, then the number of successes depends not only on the current number of retransmissions, but also on the previous history of retransmissions, and thus the problem is no longer encompassed by our multipacket reception model. In Fig. 2, the result of a simulation shows that for moderate speeds, the actual throughput is well approximated by the foregoing analysis.

2) \textbf{Frequency Hopping Random Access Channel:} Consider a finite population of \( N \) users transmitting by frequency hopping, as in [11], [22]. For each packet he wants to transmit, a user selects with equal probability one frequency in a fixed set of \( q \) frequencies. A packet is correctly received iff no other packet is transmitted on the same frequency during the same slot. We compute \((e_{nk})_{1 \leq k \leq n, n} \) and \( C = \lim_{n \to \infty} C_n \). If the users have infinite buffer space, then \( C \) can be taken as a good approximation for large \( N \) of the maximum stable throughput of the system, which is unknown. If the users have no buffer space, as is often assumed, the backlog Markov chain is always ergodic, but even then, one should expect reasonable delays in large population problems only for arrival rates below \( C \). The computation of the reception matrix of this channel is a simple combinatorial problem of random assignment of objects to cells (e.g., see [23, App. A]).
IEEE TRANSACTIONS ON AUTOMATIC CONTROL, VOL. 33, NO. 7, JULY 1988

\[ q^N C_N = \sum_{i=1}^{N} a_i \cdot (q-1) \cdots (q-1+i-1)(q-i)^{N-i} \]

\[ = \sum_{n=0}^{N-1} (-1)^n \cdot N! \cdot (N-1)^{n+1} \cdot (N-1)^{N-n} \]

which can be simplified as

\[ q^N C_N = \sum_{i=1}^{N} \left( \frac{N!}{(i-1)(N-i)!} \right) \cdot (q-1)^{N+1} \cdot (q-i)^{N-i} \]

to get the final result

\[ C_N = N \left( 1 - \frac{1}{q} \right)^{N-1} \]

\(b) N > q\): In this case, there can be at most \(q-1\) successes in a collision of size \(N\). The same method applies to get the following probabilities:

\[ \epsilon_{N^q} = \left( \frac{N}{j} \right) \sum_{k=0}^{q-j-1} \left( \frac{N-j}{k} \right) \cdot \left( \frac{-1}{k} \right) \cdot \left( \frac{N-j}{-k} \right) \]

\[ = \left( \frac{q}{j} \right)^{N+1} \cdot (q-j+1)^{N-j} \cdot \left( \frac{N-j-k}{q^N} \right) \]

\[ \epsilon_{N^q} = 0 \quad (q \leq j \leq N) \]

resulting in the same expected number of successes as before

\[ \gamma = N \left( 1 - \frac{1}{q} \right)^{N-1} \]

Now we let the population size \(N\) go to infinity and we apply our result. If we let \(N\) grow to infinity while keeping \(q\) constant, we have \(\lim_{N \to \infty} C_N = 0\), so the system is always unstable. On the other hand, if we let \(N\) go to infinity while keeping \(q\) equal to a fixed percentage of the population size, i.e., \(N/q\) constant, then \(\lim_{N \to \infty} C_N = +\infty\), and the system is always stable. It is easily shown that to get a finite maximum stable throughput, \(q\) has to grow as \(N/\log N\).

3) Mobile Radio Network with Capture: Consider an infinite number of users independently and uniformly distributed in a circle of radius \(R\), whose positions are independent from slot to slot. Users transmit packets to a common receiver located at the center of the network. Denote by \(P_1\) and \(P_2\) the received powers of the strongest and the next to strongest packets involved in a collision. Assume, as in [12]-[14], that the strongest packet is correctly received if \(P_1/P_2 > K\) (\(K\) being a system dependent constant), and that all the other packets involved in the collision are not received successfully. Assume, moreover, that the received power of a packet only depends on the distance \(r\) between the sender and the receiver

\[ P = \text{constant} \cdot \frac{1}{r^\alpha} \quad (\alpha \geq 2) \]

Then there will be capture iff

\[ r_2 > \beta r_1 \]

where \(\beta = K^{1/\alpha}\) is the capture parameter, and \(r_1, r_2\) are the distances of the closest and the next to closest senders from the receiver.
Denote by $D$ the distance between a given user and the receiver. It is easily shown that the pdf of $D$ is given by

$$p_D(r) = 2 \frac{r}{R^2} \left(1 - \left(\frac{r}{R}\right)^2\right) (0 \leq r \leq R).$$

Given $N$ users, denote by $U_N$ the closest from the receiver, and by $D_N$ its distance from the receiver. Computing the cdf of $D_N$ and taking its derivative, we obtain

$$P_D(r) = \frac{2}{N} \left(1 - \left(\frac{r}{R}\right)^2\right)^{N-1} (0 \leq r \leq R).$$

Given $D_N = r$, the other $N-1$ users are uniformly distributed in the annular region $(r, R)$. So if $N$ users collide and $D_N = r$, $U_N$ will be correctly received iff all the other users are in the annular region $(0, r)$, which is empty if $\beta r > R$. Therefore, if we denote by

$$P_N(r) = P[\text{capture} \mid N \text{ collide, } D_N = r] \quad (N \geq 2)$$

we have

$$P_N(r) = \left\{ \begin{array}{ll}
\frac{R^2 - \beta^2 r^2}{R^2 - r^2} & \text{if } r \leq \frac{R}{\beta} \\
0 & \text{if } r \geq \frac{R}{\beta}
\end{array} \right. \quad (6)$$

Thus, the probability of capture in a collision of $N$ ($N \geq 2$) is

$$\epsilon_{i1} = \int_0^{R/\beta} P_N(r) p_{DN}(r) \, dr.$$

Using (5) and (6), and with the change of variable $x = \beta/R$, this is easily computed

$$\epsilon_{i1} = \int_0^{R/\beta} 2Nx(1 - \beta^2 x^2)^{N-1} \, dx = \frac{1}{\beta^2}.$$

It follows that $C = 1/\beta^2$ is the maximum stable throughput. Notice, in particular, that for $\beta = 1$ (perfect capture), we have $C = 1$ and for $\beta \to \infty$ (no capture), we have $C \to 0$.

Under certain conditions, the performances of Aloha in the multipacket channel can be improved by varying the retransmission probability as a function of the channel history, and a maximum stable throughput of $\sup_{\epsilon_{i0}} \epsilon_{i0}^n$ can be reached (see [31]).

### III. Ergodicity Region

The number of backlogged packets in the system at time $n$, $(X_n)_{n \geq 0}$, is a homogeneous Markov chain whose one-step transition probability matrix can be computed as a function of $p$, $(\lambda_n)_{n \geq 0}$ and $E$. Denoting by $B_j(i)$ the probability of having $j$ retransmissions out of $i$ backlogged packets

$$B_i(j) = \binom{i}{j} p^j (1-p)^{i-j}$$

we get

$$\lambda_0 = \lambda_0 + \sum_{i=1}^{\infty} \lambda_i \epsilon_{i0}$$

and for $i \geq 1$

$$P_i(k) = \sum_{n=0}^{\infty} \lambda_n \sum_{j=k}^{i} B_i(j) \epsilon_{n+j,k} \quad (1 \leq k \leq i)$$

$$P_i(k) = \lambda_0 \left[ B_i(0) + \sum_{j=1}^{\infty} B_i(j) \epsilon_{j+k} \right] + \sum_{n=1}^{\infty} \lambda_n \sum_{j=0}^{\infty} B_i(j) \epsilon_{n+j,n} \quad (k \geq 1).$$

Sufficient conditions for $(X_n)_{n \geq 0}$ to be irreducible and aperiodic are as follows:

- if $0 < p < 1$:
  $$\lambda_0 \neq 0$$
  $$\lambda_0 + \sum_{n=1}^{\infty} \lambda_n \epsilon_{i0} < 1$$

- if $p = 1$:
  $$\lambda_0 \neq 0$$
  $$\lambda_0 + \sum_{n=1}^{\infty} \lambda_n \epsilon_{i0} < 1$$

These are only sufficient conditions, but they hold for almost all non-trivial systems. For example, if (9b) does not hold, then zero is an absorbing state, since the left-hand side of (9b) is equal to $P_{i0}$. Also, (9c) simply means that the successful reception of a single packet in the absence of other active users is possible. Assume, for instance, that $0 < p < 1$ and that the arrivals are Poisson distributed. Then we only have to assume (9c), and (9b) is true unless there is perfect reception, that is $\epsilon_{i0} = 1$ for all $i \geq 1$, in which case the system would of course always be stable. The case $p = 1$ gives rise to a number of pathological situations, hence, the much stronger condition (9d). It generally turns out that either (9d) is not necessary or the stability region of the system is obvious. For instance, it is clear from the transition probabilities that slotted Aloha with $p = 1$ is always unstable. In any case, it is assumed in what follows that $(X_n)_{n \geq 0}$ is irreducible and aperiodic.

**Proof of Theorem 1:** The proof is based on drift analysis. Recall that in general, the drift at state $i$ ($i \geq 0$) is defined by

$$d_i = E[X_{i+1} - X_i] | X_i = i].$$

If we denote by $\Sigma_i$ the number of successful transmissions in slot $t$, we have

$$X_{i+1} - X_i = A_i - \Sigma_i$$

and therefore

$$d_i = \lambda - E[\Sigma_i] | X_i = i].$$

Now if $R_i$ is the number of retransmissions in slot $t$, we get

$$P[R_i = k | X_i = i, A_i = n, R_i = j] = \epsilon_{n+j,k}$$

for $0 \leq j \leq i$, $0 \leq k \leq n + j$ and with the convention that $\epsilon_{i0} = \epsilon_{i0}^n$.
C\_0 = 0. Thus,
\[ E[\Sigma_i | X_t = i, A_t = n, R_t = j] = C_{n+j} \]
and
\[ E[\Sigma_i | X_t = i] = \sum_{n=0}^{\infty} \lambda_n \sum_{j=0}^{\infty} B(j) C_{n+j}. \]  \hspace{1cm} (11)

The value of the drifts for our model follows from (10) and (11):
\[ d_t = \lambda - \sum_{n=0}^{\infty} \lambda_n \sum_{j=0}^{\infty} B(j) C_{n+j}. \]  \hspace{1cm} (12)

The idea of the proof is to compute \( \lim_{t \to \infty} d_t \) which will turn out to be a very simple expression, and then apply the results of [3] and [24] to determine the ergodicity region of \( (X_t)_{t \geq 0} \). Let us first recall the two results that will be used in the sequel.

**Lemma A (Pakes [24]):** Let \( (X_t)_{t \geq 0} \) be an irreducible and aperiodic Markov chain having as state space the nonnegative integers, denote by \( (P_t) \) its transition probability matrix, and by \( d_t \) its drift at time \( t \). Then if for all \( |d_t| < \infty \), and if \( \lim_{t \to \infty} d_t = \lambda \), \( (X_t)_{t \geq 0} \) is ergodic.

**Lemma B (Kaplan [3]):** Under the assumptions of Lemma A, if for some integer \( N \geq 0 \) and some constants \( B > 0, c \in [0, 1] \) the following two conditions hold, then \( (X_t)_{t \geq 0} \) is not ergodic:
1) for all \( i \geq N, d_i > 0 \)
2) for all \( i \geq N, \lambda_i \in [c, 1], \lambda_i - \Sigma_j P_i \lambda_j \geq -B(1 - \theta) \).

From (12), it can be seen that \( |d_t| \) is finite since
\[ |d_t| \leq \lambda + \sum_{n=0}^{\infty} \lambda_n \sum_{j=0}^{\infty} B(j) C_{n+j} \leq 2 \lambda + \epsilon p. \]

Next, the drift limit is given by the following lemma.

**Lemma 1:** If \( C_t \) is a finite or not, then \( \lim_{t \to \infty} \sum_{n=0}^{\infty} \lambda_n \sum_{j=0}^{\infty} B(j) C_{n+j} = C \).

**Proof of Lemma 1:** We consider two separate cases depending on whether \( C \) is finite.
1) \( C = \infty \).
Fix \( A > 0 \) and pick \( r \geq 0 \) such that \( \lambda_r \neq 0 \). There exists an integer \( \tilde{M} \) such that for all \( n \geq \tilde{M}, C_n > A \). Fix such an \( \tilde{M} \). Then we have for any \( i \geq \tilde{M} \)
\[ \sum_{n=0}^{\infty} \lambda_n \sum_{j=0}^{\infty} B(j) C_{n+j} > \lambda_r \sum_{j=0}^{\infty} B(j) C_{\tilde{M}+j} > \lambda A \sum_{j=\tilde{M}}^{\infty} B(j) \]
which terminates the proof, since for any fixed \( M \geq 0 \)
\[ \lim_{M \to \infty} \sum_{j=0}^{M} B(j) = 1. \]  \hspace{1cm} (13)

2) \( C < \infty \).
We have for \( i > \tilde{M} \)
\[ \left| \sum_{n=0}^{\infty} \lambda_n \sum_{j=0}^{\infty} B(j) C_{n+j} - C \right| \leq \sum_{j=0}^{M} B(j) \sum_{n=0}^{\infty} \lambda_n |C_{n+j} - C| + \sum_{j=\tilde{M}+1}^{\infty} B(j) \sum_{n=0}^{\infty} \lambda_n |C_{n+j} - C|. \]  \hspace{1cm} (14)

Fix \( \epsilon > 0 \). There exists \( M \) such that for all \( n > M, |C_n - C| < \epsilon/2 \). Fix such an \( M \). Then
\[ \sum_{n=M+1}^{\infty} \lambda_n |C_{n+j} - C| < \frac{\epsilon}{2}. \]

Also, if \( L \) is an upper bound for \( C \),
\[ \sum_{j=0}^{M} B(j) \sum_{n=0}^{\infty} \lambda_n |C_{n+j} - C| \leq 2L \sum_{j=0}^{M} B(j) < \frac{\epsilon}{2}. \]

for \( L \) big enough because (13) holds, which takes care of the first term in (14) and ends the proof of Lemma 1.

Putting together (12) and Lemmas A and 1, we get that 1) if \( \lim_{t \to \infty} C_t = \infty \), then \( \lim_{t \to \infty} d_t = -\infty \), and \( (X_t)_{t \geq 0} \) is ergodic; and 2) if \( \lim_{t \to \infty} C_t = C < \infty \), then \( \lim_{t \to \infty} d_t = \lambda - C \), and \( (X_t)_{t \geq 0} \) is ergodic for \( \lambda > C \). If \( \lambda < C \), we can apply Lemma B and conclude that \( (X_t)_{t \geq 0} \) is not ergodic provided that Kaplan's condition ii) holds. This is the purpose of Lemma 2, which is the last step in the proof of Theorem 1.

**Lemma 2:** If for all \( n \geq 1, C_n < L \) for some \( L \in [0, \infty) \), then Kaplan's condition holds: there exists a constant \( B \), an integer \( N \), and a real \( c \in [0, 1] \) such that
\[ \theta^i - \sum_{j} P_{j} \theta^j \geq -B(1 - \theta) \]
all \( i \geq N, \theta \in [c, 1] \).

**Proof of Lemma 2:** According to [25], it is enough to show that the downward part of the drift, defined as
\[ D(i) = -\sum_{k=1}^{i} k P_{i,k} \]
is bounded below. From the transition probabilities (8), we get
\[ D(i) = -\sum_{k=1}^{i} k \sum_{n=0}^{\infty} \lambda_n \sum_{j=k}^{\infty} B(j) C_{n+j,n+k} \]
which can also be put in the form
\[ D(i) = -\sum_{k=1}^{i} k \sum_{n=0}^{\infty} \lambda_n \sum_{j=k}^{\infty} k C_{n+j,n+k} \]
from which it follows that
\[ D(i) \leq -\sum_{j=1}^{i} \sum_{n=0}^{\infty} \lambda_n C_{n+j} \leq -L. \]

Notice that in the proof of Theorem 1 (and this also holds for Theorem 2 below), the exact expression (7) for \( B(j) \) is never used. The only requirement is that \( B(j) \) is bounded by a probability distribution, and that (13) holds. Therefore, our results are valid for a larger class of retransmission policies than was first assumed. For example, there could be \( K \) priority groups, each with a different retransmission probability.

Although Theorem 1 is quite general, in many practical cases, the reception matrix has a very simple structure and the stability region can be obtained with virtually no computations. This happens for instance in radio networks with capture where all is needed is the limit of the second column of the matrix, or also in the simple case where above a certain collision size \( N \), the transmission is too garbled for the receiver to be able to decode anything correctly, so that \( C_n = 0 \) for \( n > N \).

This last example is a particular case of a noteworthy feature of Theorem 1, namely that the stability region does not depend on any finite number of rows of the reception matrix. In fact, any number of modifications of the matrix that leaves \( \lim_{n \to \infty} C_n \) unchanged does not affect the stability region. Although this may be surprising at first sight, it can be intuitively explained by the fundamental instability of the collision channel: unless the
receiver is perfect (all $x_n$s equal to 1), the backlog will eventually exceed any prefixed value with probability one, thus it is the limit at $C_0$ that determines the stability region.

The stability region is also unchanged if the first transmission of packets is delayed. If new packets are backlogged, that is, transmitted for the first time with probability $p$ in each slot following their arrival (this transmission rule appears in the literature as controlled-access or delayed first transmission); the tricts become $d_i = \lambda - \sum_{j=1}^{i} B(j) C_j$ for $i \geq 1$, and from Lemmas 1 and 2 the ergodicity region remains the same. If $C_i$ does not have a limit, Theorem 1 does not give the stable throughput of the system. Even though in almost all practical cases, and indeed in all the examples of Section II, $C_i$ does have a limit, it is conceptually interesting to examine the case when $\lim_{n \to \infty} C_n = \lim_{n \to \infty} C_{\infty}$ is not the limit. It is worth pointing out that adding constraints as strong as the following on the reception matrix still does not imply that $C_{\infty}$ has a limit:

1. $(\epsilon_{\infty})_{n=1}$ is nondecreasing
2. $(\epsilon_{\infty})_{n=1}$ is nonincreasing for all $k \geq 1$
3. $\epsilon_{n} \leq \epsilon_{n+k+1}$ for $n \geq 2, 1 \leq k \leq n-1$

although the counterexamples we have been able to build are somewhat contrived. Notice that conditions i) and ii) above imply that each column has a limit $\epsilon_{\infty} = \lim_{k \to \infty} \epsilon_{n}(k \geq 0)$, which is very likely to happen in practice. In any case, Theorem 2 below still gives some information on the stability region, although the exact result requires in general the complete knowledge of the sequence $(C_{n+k})$. In fact, given any nonnegative numbers $\alpha < \gamma$, one can construct a reception matrix with $n$th row average $\gamma$ such that:

1. $\liminf_{n \to \infty} C_n = \alpha$
2. $\limsup_{n \to \infty} C_n = \beta$

and such that the maximum stable throughput is $\gamma$.

Theorem 2: The system is stable for $\lambda < \liminf_{n \to \infty} C_n$ and unstable for $\lambda > \limsup_{n \to \infty} C_n$.

Proof:

1. If $\lambda < \liminf_{n \to \infty} C_n$, then $(X_{n})_{n=0}$ is ergodic.
2. If $\liminf_{n \to \infty} C_n = +\infty$, then $\liminf_{n \to \infty} C_n = +\infty$, and the result has already been proved, so assume that $\liminf_{n \to \infty} C_n$ is finite. From Lemma A, it is enough to prove that for all $\epsilon > 0$, there exists $N$ such that:

\[ u(\lambda - \liminf_{n \to \infty} C_n + \epsilon) \quad \forall n \geq N.\]

Recall from (12) that we have:

\[ u = \lambda - \sum_{n=0}^{N} \lambda \sum_{j=0}^{n} B(j) C_{n+j}.\]

It is only needed to prove that for all $\epsilon > 0$ there exists $N$ such that:

\[ \lambda \sum_{j=0}^{n} B(j) C_{n+j} > \liminf_{n \to \infty} C_n - \epsilon \quad \forall n \geq N.\]

Now by definition there exists $M$ such that for all $k \geq M$:

\[ \epsilon > \liminf_{n \to \infty} C_n - \epsilon \]

and therefore for all $i > M$:

\[ \sum_{n=0}^{N} \lambda \sum_{j=0}^{n} B(j) C_{n+j} > (\liminf_{n \to \infty} C_n - \epsilon) \sum_{j=0}^{M} B(j) \]

which completes the proof since (13) holds.

b) If $\lambda > \limsup_{n \to \infty} C_n$, then $(X_{n})_{n=0}$ is not ergodic.

Since $\lambda$ is finite, in this case lim sup $C_{n}$ is necessarily finite. Therefore, $(C_{n})_{n=0}$ is bounded and from Lemma 2, Kaplan's condition holds. Thus, it is enough to show that for all $\epsilon > 0$, there exists $N$ such that:

\[ d_{i} > \lambda - \limsup_{n \to \infty} C_{n} - \epsilon \quad \forall i \geq N.\]

From (15), we only need to show:

\[ \sum_{i=0}^{n} \lambda \sum_{j=0}^{n} B(j) C_{n+j} \leq \limsup_{n \to \infty} C_{n} + \epsilon \quad \forall n \geq N.\]

Since there exists $M$ such that for all $k \geq M$:

\[ C_{k} < \limsup_{n \to \infty} C_{n} + \epsilon \]

then if $L$ is an upper bound for $C_n$, we have for $i \geq M$:

\[ \sum_{n=0}^{N} \lambda \sum_{j=0}^{n} B(j) C_{n+j} \leq L \sum_{j=0}^{M} B(j) + \limsup_{n \to \infty} C_{n} + \epsilon \]

from which the result follows, using (13).

\[ \square \]

V. Behavior of the Backlog Markov Chain in the Nonergodicity Region

In this section, we further investigate the properties of $(K_{n})_{n=0}$, in the case $\lambda > C$, assuming of course that $(C_{n})_{n=0}$ has a finite limit. It has been proved in [1] that the backlog Markov chain for the usual slotted Aloha algorithm is transient, but this result cannot be generalized to our model when $\lambda > C$. We give below an example showing that $(X_{n})_{n=0}$ can be null recurrent when the mean arrival rate $\lambda$ belongs to an interval of positive length. The boundary between the null recurrence and the transience regions generally depends in a rather complicated manner on both the reception matrix and the retransmission probability $p$. We give a sufficient condition for $(X_{n})_{n=0}$ to be transient when $\lambda > C$, as well as bounds on the recurrence region.

Consider the reception matrix defined by

\[ e_{nk} = \frac{1}{n+1} \quad (1 \leq k \leq n) \]

and $s_{n} = \frac{1}{n+1}$ for $n \geq 1$. Then $C_n = \sum_{k=1}^{n} k/n^2 = (n + 1)/2n$, and $C = 1/2$.

Using Lemmas C and D below, we show in [26] that $X_{n}$ is recurrent for $\lambda < 2R(p)$ and transient for $\lambda > 2R(p)$, where $R(p)$ is a function of the retransmission probability $p$ and is given by:

\[ R(p) = \frac{(1-p^2)}{p^2} \ln (1-p^2) \quad (0 < p < 1) \]

$R(1) = 1$.

It is easily seen that $R(p)$ is an increasing function of $p$ for $p \in (0, 1]$, with extrema $\lim_{p \to 0} R(p) = 1/2$ and $\lim_{p \to 1} R(p) = 1$.

Fig. 3 summarizes the behavior of $X_{n}$ for this example.

It is somehow surprising to see that in this case, as well as in all
The other examples we have computed, the recurrence region becomes larger as $p^*$ increases. Intuitively, the recurrence of $X_n$ when $\lambda > C$ seems to be due to the fact that transitions from any state to 0 for some fixed integer $k_0$ are possible and that the probability of such an event, $P_{00}$ (or $P_{0n}$), goes to zero slowly with $i$. It can be checked that these probabilities are increasing functions of $P(1)$ when $i$ is large enough.

Transience is ensured for $\lambda > C$ if the supremum of the elements of the $k$th-column goes to zero faster than $k^2$. This condition holds for all the examples in Section II, as well as for many real life cases, due to the practical limitations on the receiver capabilities. In particular, it is always verified if the reception matrix has only a finite number of nonzero columns (or equivalently, if the backlog Markov chain has uniformly bounded transitions, as defined in [3]) which happens, for instance, if there is capture. Note that the proof of Theorem 3 below is of course valid for the conventional collision channel, as in this case becomes somewhat simpler than the proof in [1].

**Theorem 3:** If $\lim_{k \to \infty} k^2 \sup_{n \geq 0} \epsilon_n = 0$, then $(X_n)_{n \geq 0}$ is transient for $\lambda > C$.

Because of the complexity and lack of structure of the one-step transition probabilities (8), few results on the recurrence and transience of Markov chains can be applied to our model. Before proving Theorem 3, let us introduce the following two criteria from [27].

**Lemma C:** Let $(X_{n=0})$ be an irreducible and aperiodic Markov chain, having as state space the set of nonnegative integers, and with one-step transition probability matrix $P = P_{00}$. $(X_{n=0})$ is recurrent if and only if there exists a sequence $(y_n)_{n \geq 0}$ such that

1) $\lim_{n \to \infty} y_n = +\infty$
2) for some integer $N > 0$ $\sum_{n=1}^{\infty} y_j P_{0j} = y_i$ all $i \geq N$.

We will only use the sufficiency part, which has also been proved in [24].

**Lemma D:** With the same assumptions as in Lemma C, $(X_{n=0})$ is transient if and only if there exists a sequence $(y_n)_{n \geq 0}$ such that

1) $(y_n)_{n \geq 0}$ is bounded
2) for some integer $N > 0$ $\sum_{j=0}^{\infty} y_j P_{0j} = y_i$ all $i \geq N$.

3) for some $k \geq N, y_k < y_0, \ldots, y_{N-1}$.

The idea of the proof is to show that

$$\lim_{i \to \infty} [D'(i) + U'(i)] = -\theta \lim_{i \to \infty} d_i$$

and since it has been proved in Section III that $\lim_{i \to \infty} d_i = \lambda - C$, we will be able to conclude that $(X_{n=0})$ is transient for $\lambda > C$.

1) $\lim_{i \to \infty} [D'(i) + \theta D(i)] = 0$. 

---

**IEEE TRANSACTIONS ON AUTOMATIC CONTROL, VOL. 33, NO. 7, JULY 1988**

Sufficiency under the additional constraints $y_1 > 0$ and $\lim_{i \to \infty} y_i = 0$ has also been proved in [28]. Also, the sufficiency parts of both lemmas are an immediate consequence of [29, Theorems 5 and 6] together with the results in [30].
From (18) and (20)

\[ D'(i)+\theta D(i)=(i+1) \sum_{k=0}^{i} \lambda_n \sum_{j=1}^{i} B(j) \sum_{k=1}^{j} \left( \frac{i+1}{i+1-k} \right)^{\sigma-1} \frac{\theta k}{i+1}, \]

which is more conveniently written as

\[ D'(i)+\theta D(i)=(i+1) \sum_{n=0}^{i} \lambda_n \sum_{j=1}^{i} B(j) \sum_{k=1}^{j} \left( \frac{i+1}{i+1-k} \right)^{\sigma-1} \frac{\theta k}{i+1} \epsilon_{n+j,n+k}. \]

This expression is nonnegative since

\[ \left( \frac{i+1}{i+1-k} \right)^{\sigma-1} - 1 - \frac{\theta k}{i+1} > 0 \quad (1 \leq k \leq i). \]

Define \( \gamma_k = \sup_{n \geq k} \epsilon_{nk} \). Then

\[ 0 \leq D'(i)+\theta D(i) \leq (i+1) \sum_{n=0}^{i} \lambda_n \sum_{j=1}^{i} B(j) \sum_{k=1}^{j} \left( \frac{i+1}{i+1-k} \right)^{\sigma-1} - 1 - \frac{\theta k}{i+1} \gamma_{n+k}. \]

That is

\[ J'(i)+\theta U(i) \leq x(i)+x(i) \quad (22) \]

With a change of variable

\[ J'(i)+\theta U(i)= (i+1) \sum_{j=0}^{i} B(j) \sum_{n=1}^{i} \lambda_n \sum_{k=1}^{j} \left( \frac{i+1}{i+1+k} \right)^{\sigma-1} - 1 - \frac{\theta k}{i+1} \epsilon_{n+j,n+k}. \]

We show that \( x(i) \) and \( x(i) \) go to zero independently. Fix \( \epsilon > 0 \). Define \( 0 < x \leq i \) the function

\[ \rho(x)=\frac{i+1}{x^2} \left[ \left( \frac{i+1}{i+1-x} \right)^{\sigma-1} - 1 - \frac{\theta x}{x} \right]. \]

It is easily proved that for each \( i \geq 1, \rho(x) \) is a positive nondecreasing function of \( x \). Also

\[ r \left( \frac{i+1}{2} \right) = \frac{1}{i+1} \left[ 4(2^\sigma-1)-2\theta \right] = \frac{A}{i+1}, \]

where \( A \) is a positive constant depending only on \( \theta \). From (23)

\[ x_1(i)=\sum_{n=0}^{i} \sum_{k=1}^{i+1/2} k^2 \rho(k) \gamma_{n+k} \leq \frac{A}{i+1} \sum_{n=0}^{i} \lambda_n \sum_{k=1}^{i+1/2} k^2 \gamma_{n+k}. \]

If \( \lim_{n \to \infty} k^2 \gamma_k = 0 \), then \( \lim_{n \to \infty} 1/n \sum_{k=1}^{i} k^2 \gamma_k = 0 \). So we can choose \( i \) large enough so that for \( n \geq (i + 1)/2 \), \( \sum_{k=1}^{i} k^2 \gamma_k < ne \). Then

\[ x_1(i) \leq \epsilon A \sum_{i+1}^{\infty} \lambda_n \left( \frac{n+i+1}{2} \right) = \epsilon A \left( \frac{i+1}{i+1+1/2} \right). \]

Now if we choose \( i \) big enough so that for \( k > (i + 1)/2 \), we have \( \gamma_k < \epsilon/k^2 \), then

\[ x_2(i) \leq \epsilon A \sum_{i+1}^{\infty} \lambda_n \sum_{k=1}^{i+1} \left( \frac{i+1}{i+1+k} \right)^{\sigma-1} - 1 - \frac{\theta k}{i+1} \epsilon_{n+j,n+k}. \]

By bounding the sum in the last equation by integrals, it can be seen that it is upper bounded by a linear function of \( i \).

2) \( \lim_{n \to \infty} \left[ U'(i)+\theta U(i) \right] = 0 \).

From (18) and (20)

\[ U'(i)+\theta U(i)= \sum_{n=0}^{i} \lambda_n \sum_{j=1}^{i} \left( \frac{i+1}{i+1+k} \right)^{\sigma-1} - 1 - \frac{\theta k}{i+1} \epsilon_{n+j,n+k}. \]

With a change of variable

\[ U'(i)+\theta U(i)= \sum_{j=0}^{i} B(j) \sum_{n=1}^{i} \lambda_n \sum_{k=1}^{j} \left( \frac{i+1}{i+1+k} \right)^{\sigma-1} - 1 - \frac{\theta k}{i+1} \epsilon_{n+j,n+k}. \]

We get

\[ 0 \leq U'(i)+\theta U(i) \leq \theta \left( \frac{i+1}{i+1} \right) \sum_{j=0}^{i} B(j) \sum_{n=1}^{i} \lambda_n \sum_{k=1}^{j} \left( \frac{i+1}{i+1+k} \right)^{\sigma-1} - 1 - \frac{\theta k}{i+1} \epsilon_{n+j,n+k}. \]

Fix \( \epsilon > 0 \). Choose \( N \) such that \( \sum_{n=1}^{N} \lambda_n < \epsilon/2 \), and then, \( N \) being fixed, choose \( i \) large enough so that \( 1/(i+1) \sum_{n=1}^{N} k^2 \lambda_k < \epsilon/2 \).

It should be clear at this point that unlike the ergodicity region,
the recurrence region depends in general on the elements of the exception matrix (instead of only the row averages) and on the transmission probability \( p \). For this reason, an exact expression for the recurrence region seems rather difficult to obtain; nonetheless, the method (see [26]) that we used to study the example in Fig. 3 can be generalized to obtain the following upper and lower bounds on the recurrence region.

**Theorem 4:** \( (X_n)_{n \geq 0} \) is recurrent for \( \lambda < L \) and transient for \( \lambda > L \), with \( L = \max \{ \lambda \upsilon, \sup_{0 \leq c < 1} \lambda \upsilon, \sup_{0 \leq c < 1} \lambda \upsilon \} \) and \( U = \min \upsilon_1, \upsilon_2, \min \upsilon_3, \upsilon_4 \) where

\[
\begin{align*}
\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \ln \left( \frac{1}{i} \right)^{n+1} \left( \frac{n+1+k}{i+1+k} \right) & = 0, \\
\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \ln \left( \frac{1}{i} \right)^{n+1} \left( \frac{n+1+k}{i+1+k} \right) & = 0, \\
\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \ln \left( \frac{1}{i} \right)^{n+1} \left( \frac{n+1+k}{i+1+k} \right) & = 0.
\end{align*}
\]

We are assuming that the limits above exist, which indeed happens in most practical cases. The theorem is valid if any of these limits is infinite. In particular, if \( L = +\infty \), then \( X_n \) is always recurrent. Note that usually, it is not necessary to carry out all the computations, because one of the three terms in the definition of \( L \) is equal to one of the terms in the definition of \( U \); in fact, in most cases, we have \( \sup_{0 \leq c < 1} \lambda \upsilon = \inf_{0 \leq c < 1} \lambda \upsilon \) if \( 0 < p < 1 \), and \( \upsilon_1 = \lambda \upsilon \) if \( p = 1 \). The proof of Theorem 4 can be found in [26].

**REFERENCES**


Sylvie Ghez (S’88) was born in H. She received the engineering diploma from the École Nationale Supérieure des Télécommunications, Paris, in 1984, and the M.S. degree in electrical engineering from Princeton University, Princeton, NJ, in 1987, where she is currently a doctoral candidate. She is the recipient of the Wallace Memorial Fellowship for the year 1987–1988. Her research interests include multiple access protocols, network routing problems, queueing theory, and applied probability theory.

Sergio Verdú (S’80–M’84) received the Ph.D. degree in electrical engineering from the University of Illinois at Urbana-Champaign in 1984.

Upon completion of his Ph.D., he joined the faculty of Princeton University, Princeton, NJ, where he is an Assistant Professor of Electrical Engineering. His recent research contributions are in the areas of multiuser communication and information theory, and statistical signal processing.

Dr. Verdú is a recipient of the National University Prize of Spain, the Rheinstein Outstanding Junior Faculty Award of the School of Engineering and Applied Sciences, Princeton University, and the NSF Presidential Young Investigator Award. He is currently serving as Associate Editor of the IEEE TRANSACTIONS ON AUTOMATIC CONTROL.

Sergio Verdú (S’80–M’84) received the Ph.D. degree in electrical engineering from the University of Illinois at Urbana-Champaign in 1984.

Upon completion of his Ph.D., he joined the faculty of Princeton University, Princeton, NJ, where he is an Assistant Professor of Electrical Engineering. His recent research contributions are in the areas of multiuser communication and information theory, and statistical signal processing.

Dr. Verdú is a recipient of the National University Prize of Spain, the Rheinstein Outstanding Junior Faculty Award of the School of Engineering and Applied Sciences, Princeton University, and the NSF Presidential Young Investigator Award. He is currently serving as Associate Editor of the IEEE TRANSACTIONS ON AUTOMATIC CONTROL.

Stuart C. Schwartz (S’64–M’66–SM’83) was born in H. He received the B.S. and M.S. degrees in aeronautical engineering from the Massachusetts Institute of Technology, Cambridge, in 1961, and the Ph.D. degree from the Information and Control Engineering Program, University of Michigan, Ann Arbor, in 1966.

While at M.I.T., he was associated with the Naval Supersonic Laboratory and the Instrumentation Laboratory. From 1964 to 1962 he was at the Jet Propulsion Laboratory, Pasadena, CA, working on problems in orbit estimation and telemetry. He is presently a Professor of Electrical Engineering and Chairman of the department at Princeton University, Princeton, NJ. He served as Associate Dean of the School of Engineering from July 1977 to June 1980. During the academic year 1972–1973, he was a John S. Guggenheim Fellow and a Visiting Associate Professor at the Department of Electrical Engineering, Technion—Israel Institute of Technology, Haifa, Israel. During the academic year 1980–1981, he was a member of the Technical Staff at the Radio Research Laboratory, Bell Telephone Laboratories, Crawford Hill, NJ. His principal research interests are in the application of probability and stochastic processes to problems in statistical communication and system theory.

Dr. Schwartz is a member of Sigma Gamma Tau, Eta Kappa Nu, and Sigma Xi. He has served as an Editor for the SIAM Journal on Applied Mathematics and as Program Chairman for the 1986 ISIT.
Performance Analysis of an Asymptotically Quantum-Limited Optical DPSK Receiver

David Brady
Sergio Verdú
Performance Analysis of an Asymptotically Quantum-Limited Optical DPSK Receiver

DAVID BRADY, STUDENT MEMBER, IEEE, AND SERGIO VERDÚ, SENIOR MEMBER, IEEE

Abstract—In this paper, we analyze an optical, direct-detection DPSK receiver whose error probability is quantum-limited as the transmitting laser linewidth vanishes. The design of the receiver is based on a binary hypothesis test with a delay-and-sum (DAS) optical preprocessor followed by a photodetector and an integrate-and-dump circuit. Under the assumption of a binary hypothesis test with point process observations and on the condition that the rates of the point process under each hypothesis are equal, the optimal receiver is shown to be a DAS receiver. The DAS receiver is optimal in the sense that it minimizes the error probability. The optimal receiver is then obtained as a limit of the DAS receiver. The error probability of the optimal receiver is shown to converge to the true error probability as the transmitting laser linewidth decreases. Bounds on the power penalty are derived by developing bounds on the conditional rates of the point process, and it is shown that the error probability bounds converge to the true value as the transmitting laser linewidth decreases. Bounds on the power penalty are derived for parameters corresponding to existing semiconductor injection lasers, and are seen to be less than the limiting power penalty for a balanced DPSK receiver.

I. INTRODUCTION

In this paper, we analyze an optical, direct-detection DPSK receiver whose error probability is quantum-limited as the transmitting laser linewidth vanishes. The design of the receiver is based on a binary hypothesis test with point process observations and on the condition that the rates of the point process under each hypothesis are equal, the optimal receiver is shown to be a DAS receiver. The DAS receiver is optimal in the sense that it minimizes the error probability. The optimal receiver is then obtained as a limit of the DAS receiver. The error probability of the optimal receiver is shown to converge to the true error probability as the transmitting laser linewidth decreases. Bounds on the power penalty are derived for parameters corresponding to existing semiconductor injection lasers, and are seen to be less than the limiting power penalty for a balanced DPSK receiver.

In binary DPSK modulation, the transmitted scalar field is amplitude-modulated by a bit stream derived from the information sequence. Denoting the information sequence by \( b_n \), the transmitted scalar field is described by

\[
\nu(t) = \sum_{n=0}^{\infty} a_n A \cos(\omega t + \theta_n) = V_2 \Re \left\{ \sum_{n=0}^{\infty} \frac{a_n A}{\sqrt{2}} e^{i n \theta_n} \right\},
\]

where \( V_2 \) is the carrier frequency. The decoding from \( \{a_n\} \) to \( \{b_n\} \) is performed in the same operation that compares the received signal in \( nT, (n + 1)T \) to the reference signal. This suggests the demodulation scheme shown in Fig. 1. The optical signal described in (1) is divided by a half-silvered mirror into two signals of equal power, one of which is delayed by the symbol period \( T \), and then added to the other. The resultant optical signal is incident on the photodetector in Fig. 1 and is given by

\[
r(t) = \frac{1}{\sqrt{2}} \left\{ a_n A \cos(\nu t + \theta_n) + a_{n-1} A \cos((\nu t + \theta_n) - [\nu T + \Delta \theta_n]) \right\}
\]

where the Brownian motion \( \{\theta_n, t \in \mathbb{R}\} \) models the transmitted laser phase noise and \( \nu \) is the carrier frequency. Note that in the absence of laser phase noise the transmitted optical energy is \( \frac{1}{2} A^2 T/2 \) photons per bit.

The decoding from \( \{a_n\} \) to \( \{b_n\} \) is performed in the same operation that compares the received signal in \( nT, (n + 1)T \) to the reference signal. This suggests the demodulation scheme shown in Fig. 1. The optical signal described in (1) is divided by a half-silvered mirror into two signals of equal power, one of which is delayed by the symbol period \( T \), and then added to the other. The resultant optical signal is incident on the photodetector in Fig. 1 and is given by
where $E = \lambda_0/A2e^{\nu t}(a_0 + a_1e^{-j(\nu T + c)})$ is the complex envelope of the electric field incident on the photodetector, and $\Delta \theta$ is an increment of Brownian motion given by

$$\Delta \theta \triangleq \theta_1 - \theta_2 - \tau$$

The remaining receiver structure will be based on the photon point process observations at the output of the photodetector. This process may be modeled as a doubly stochastic point process with count $N$, occurrence times $\{W_1, W_2, \ldots, W_N\}$, and rate $\lambda = \alpha E[\cos^2 s] + \lambda_0$ where $\alpha$ and $\lambda_0$ are parameters of the photodetector [4]. It should be pointed out that it is equivalent to express the rate as $\lambda_0 = \alpha E[\cos^2 s] + \lambda_0$ and ignore the double frequency terms. Although the following results may easily be extended for arbitrary values of $\alpha$ and $\lambda_0$, we assume in the following that the photoelectric conversion is ideal, i.e., $\alpha = 1$ and $\lambda_0 = 0$. Under these assumptions the rate at the electron departure process for $0 \leq t < T$ is

$$\nu = \frac{1}{2}[1 + c_0 \cos \Delta \theta_1], \quad 0 \leq t < T$$

(4)

where we have assumed $\nu T = 0 \mod 2\pi$. The hypothesis pair or the interval $(0, T)$ may be described as

$\mathcal{H}_0$: $\lambda_0 = \lambda_0^{(0)} \triangleq \frac{1}{T} \left(1 - \cos \Delta \theta_1\right), \quad 0 \leq t < T$

$\mathcal{H}_1$: $\lambda_0 = \lambda_0^{(1)} \triangleq \frac{1}{T} \left(1 + \cos \Delta \theta_1\right), \quad 0 \leq t < T$

(5)

where $\nu$ is the transmitted optical energy in photons per bit as defined earlier. If we define $X(t) = E[\lambda(t)N(t); 0 \leq s < t]$, then the test which minimizes the probability of error is [5]

$$N_T \geq \sum_{i=1}^{\lfloor N_T/T \rfloor} \log \frac{2\nu e^{-\nu T} - \lfloor \nu T \rfloor - \sum_{i=1}^{\lfloor N_T/T \rfloor} \sum_{j=1}^{\lfloor N_T/T \rfloor} \lfloor \nu T \rfloor - \sum_{i=1}^{\lfloor N_T/T \rfloor} \lfloor \nu T \rfloor \geq 0.\]$$

(6)

The formulation of this log-likelihood ratio has an interesting, two part structure. First derive the minimum mean-squared error causal estimates of the intensities given the observation under each hypothesis, and then solve a binary hypothesis testing problem with observations from a nonhomogeneous Poisson process whose conditional rates are these estimates. This fact has commonly been referred to as the separation theorem of detection, and motivates the use of suboptimal estimates in hypothesis testing with doubly stochastic point process observations [4], [5].

Unfortunately, the explicit structure of (6) is unknown due to the difficulty in evaluating the conditional estimates $\lambda_0^{(i)}$. One approach to this problem is to replace the optimum estimates $\hat{\lambda}_0^{(i)}$ with suboptimum estimates. In this paper, we propose the suboptimum estimates $E[\lambda_0^{(i)}]$. A justification of this approach is the following. Denoting $B_i$ as the laser bandwidth in Hertz, $\gamma = 1/B_i T$, and choosing $x_0$ such that $1 - e^{-\nu t} < x_0 < e^{-\nu t}$, we have

$$[\exists t \in [0, T] : |\cos \Delta \theta_1 - E[\cos \Delta \theta_1]| > x_0].$$

(7)

This upper bound follows from Lemma 1 in Section III. For $\gamma = 1000$, and $x_0 = 20$ percent of $E[\cos \Delta \theta_1]$, the left-hand side of (7) is less than 0.03. For $\gamma = 1500$ and $x_0 = 20$ percent of $E[\cos \Delta \theta_1]$, the probability is less than 0.003. So the probability that a sample path of the intensity deviates from the mean in a symbol interval by more than 20 percent is bounded above by a small number for reasonable values of $\gamma$. Based on this argument, we employ the constant $E[\cos \Delta \theta_1]$ to estimate the process $\{\cos \Delta \theta_1, t \in [0, T]\}$, and the suboptimal estimates of the conditional rates follow directly. The advantage to assuming a constant estimate is that the test assumes homogeneous Poisson point process observations, and the decision strategy is very simple: compare the photon count to a threshold. By (7) we can expect this strategy to be optimum for large $\gamma$ although it may be far from optimum for low $\gamma$.

A suboptimal receiver design is specified below. Since $\{\theta_i, t \in [0]\}$ is a Brownian motion with zero drift and diffusion coefficient $\sqrt{2\pi}$, $\{\Delta \theta_i, t \in [0]\}$ is a stationary Gaussian random process with zero mean and autocorrelation function

$$R(r) = \begin{cases} \frac{2\pi}{\gamma} & |r| < T \\ 0 & \text{otherwise} \end{cases}$$

(8)

The means of $\lambda_0^{(i)}$ are easily computed

$$E[\lambda_0^{(0)}] = \frac{\gamma}{T} (1 - e^{-\gamma T}),$$

(9)

$$E[\lambda_0^{(1)}] = \frac{\gamma}{T} (1 + e^{-\gamma T}).$$

Employing these suboptimal estimates in (6), a suboptimal hypothesis test is

$$N_T \geq \sum_{i=1}^{\lfloor N_T/T \rfloor} \log \frac{2\nu e^{-\nu T} - \lfloor \nu T \rfloor - \sum_{i=1}^{\lfloor N_T/T \rfloor} \sum_{j=1}^{\lfloor N_T/T \rfloor} \lfloor \nu T \rfloor - \sum_{i=1}^{\lfloor N_T/T \rfloor} \lfloor \nu T \rfloor \geq 0.\]$$

where $\lfloor \cdot \rfloor$ indicates the greatest integer function. Due to the stationarity of the laser phase noise increment, the suboptimal test is not a function of the photon arrival times, and the remaining portion of the DPSK receiver need only count the number of received photons in the interval $[0, T]$. The entire suboptimal receiver structure is shown in Fig. 1. Note that the threshold obtained in (10) is not the threshold that minimizes the error probability of a test which uses the statistic $N_T$. The optimum threshold is also a function of the optical power and $\gamma$, and is closely approximated as a by-product of the error probability analysis of Section III.

It is worthwhile to compare this DPSK receiver to the balanced receiver analyzed in [2]. The balanced receiver divides the received optical signal into four streams of equal intensity, two of which are delayed by the symbol period $T$. A delayed stream is added to a nondelayed stream in the same way as in the proposed receiver, except at one-half the power. The remaining pair of optical streams are subtracted. The outputs of the summer and subtractor are each input to photodetectors, and the resulting counts are subtracted. The resulting hypothesis test requires no threshold setting. That is, the difference of the electron counts is compared to zero, and the receiver is independent of $\gamma$. 

---

\[fig:1\] Proposed optical DPSK receiver.
However, in using this information about the laser phase noise, the proposed DPSK receiver requires one-half the transmitted optical power of the balanced DPSK receiver to achieve the same error rate. This claim will be verified in Section III. This comparison assumes that the three-port beam combiners used in both receivers are lossless. It should be noted that if a Mach-Zehnder interferometer is used as the beam combiner [6], then the proposed DPSK detector uses only one of the two output port beams, and the energy of the signal incident on the photodetector is one-half of that assumed in the analysis of this paper. In this case, the performance of our receiver is asymptotically equivalent to that of the balanced detector in [2], implemented using both output ports of a Mach-Zehnder interferometer [7].

As suggested above, the parameter \( \gamma \) is central in the analysis of optical communication systems employing coherent light with nonzero linewidth. It characterizes the performance degradation due to the transmitting laser phase jitter relative to the symbol rate. For fixed laser linewidth, the effect of the phase noise on system performance is less pronounced as the symbol rate increases, as reflected by an increase of \( \gamma \).

Typically, \( \gamma \in [50, 1600] \), which follows from \( B_2 \in [6 \, \text{MHz}, 20 \, \text{MHz}] \) for semiconductor injection lasers [8], and symbol rates from 1 to 10 Gbits/s.

### III. PERFORMANCE ANALYSIS

In this section, we characterize the performance of the proposed DPSK receiver. We show that as \( \gamma \to \infty \) the probability of error is the quantum limit of optical communications. We then derive upper and lower bounds on the probability of error for arbitrary \( \gamma \) and show that these bounds converge to the true value as \( \gamma \to \infty \). Finally, we present Monte Carlo simulation results of the receiver performance and compare them to the error probability bounds.

We begin by showing that the performance of the proposed DPSK receiver is quantum limited as \( \gamma \to \infty \). In this case, the transmitting laser is ideal and it is easy to see that (10) becomes

\[
N_f \leq \frac{N_0}{\gamma_0}. \tag{11}
\]

As \( \gamma \to \infty \), the rates under each hypothesis are deterministic and \( N_f \) is an unconditionally Poisson random variable under each hypothesis. If we define \( \Lambda_i \triangleq \frac{1}{\gamma} N_i, i \in \{0, 1\} \) then the probability of error is

\[
P[\text{error}] = P[N_f = 0, H_1] = \frac{1}{2} e^{-\Lambda_1}. \tag{12}
\]

Since \( \Lambda_1 = 2T \) for an ideal transmitting laser, (12) becomes

\[
P[\text{error}] = \frac{1}{2} e^{-2T} \tag{13}
\]

which is known as the quantum limit. Thus, the receiver performance is quantum limited as the transmitting laser linewidth goes to zero.

Next, we consider the error probability for finite values of \( \gamma \). It is convenient to define the moment generating function \( M_i(v) \triangleq E[e^{i \lambda}] \) and to let \( N(\Lambda) \) denote a Poisson random variable with mean \( \Lambda \). Conditioned on the rate and the hypothesis, the observation process is a Poisson point process. Therefore, the probability of error under \( H_i \) may be found by first conditioning on \( \{\Delta \theta_i, 0 < t < T\} \)

\[
P[\text{error}|H_1] = E[P[N(\Lambda_1) \leq \lfloor \Delta \theta_i, 0 < t < T \rfloor)]
= E \left[ \sum_{k=0}^{\lfloor \Delta \theta_i \rfloor} \frac{(\Lambda_1)^k}{k!} e^{-\Lambda_1} \right]
= \sum_{k=0}^{\lfloor \Delta \theta_i \rfloor} \frac{1}{k!} \int_0^T M_k(v) \, dv \tag{14}
\]

where \( f \) is an arbitrary nonnegative integer threshold, and the last line follows from an application of the bounded convergence theorem. By a similar argument we have under \( H_0 \)

\[
P[\text{error}|H_0] = 1 - \sum_{k=0}^{\lfloor \Delta \theta_i \rfloor-1} \frac{1}{k!} \int_0^T M_k(v) \, dv. \tag{15}
\]

It appears that there is no closed-form expression for the foregoing moment generating functions when \( \{\Delta \theta_i, 0 < t < T\} \) is a \( T \) increment of Brownian motion. In the remainder of this section, we consider upper and lower bounds to (14) and (15). Based on the fact that the mgf of \( 1/2T \int_0^T (\Delta \theta_i)^2 \, dt \) is computable (see the Appendix), our approach is to find upper and lower bounds on the error probability based on quadratic bounds of \( \cos x \)

\[
1 - \frac{x^2}{2} \leq \cos x \leq \frac{1}{2} \left( 1 - \frac{ax^2}{2} \right) \text{ if } |x| < x_o
\]

\[
1 \text{ otherwise}
\]

where \( 0 < a < 1 \) and \( x_o \) is the smallest positive real number such that \( 1 - ax^2/2 = \cos x \). If each cosine in the expressions for \( \Lambda_i \) is replaced by the upper bounding function, the corresponding rates are further apart, hence it is easier to discriminate between them and a lower bound on the error probability is obtained. Analogously, an upper bound is obtained if each cosine is replaced by the lower bounding function. More precisely, the cumulative distribution function (cdf) of a Poisson random variable is a strictly decreasing function of its mean, that is, \( P[N(\Lambda) \leq \lfloor \Lambda \rfloor] \) is decreasing in \( \Lambda \).

If the error probability under each hypothesis is found by first conditioning on \( \{\Delta \theta_i, 0 < t < T\} \), then the conditional probabilities are Poisson cdf's, and substituting the bounds from (16) in the expressions for \( \Lambda_i \) will yield upper and lower bounds for the conditional probabilities as well as the unconditioned probabilities. Denoting \( \Lambda_i^U \) as the bound on \( \Lambda_i \) which yields an upper bound to the error probability on \( H_i \), and \( \Lambda_i^L \) as the bound on \( \Lambda_i \) which yields a lower bound to the error probability under \( H_i \), we have

\[
\Lambda_i^U = \frac{\lambda}{2T} \int_0^T (\Delta \theta_i)^2 \, dt
\]

\[
\Lambda_i^U = 2T - \Lambda_i^U
\]

\[
\Lambda_i^L = \frac{\lambda}{2T} \int_0^T 1 - f_\Lambda(\Delta \theta_i) \, dt
\]

\[
\Lambda_i^L = 2T - \Lambda_i^L. \tag{17}
\]

As seen by (14), (15), and (17) computation of these bounds requires the moment generating functions of \( \Lambda_i^U \) and \( \Lambda_i^L \). The following two lemmas loosen two of the bounds so that we require only \( \Theta(v) \triangleq E[e^{\lambda B_i}] \), which is derived in the Appendix. Lemma 1 quantifies the fact that the probability that a \( T \) length sample path of \( \Delta \theta \) deviates far from the origin decreases with increasing \( \gamma \). Lemma 2 uses this fact in deriving a lower bound on the error probability under \( H_0 \), which does not depend on the mgf of \( \Lambda_i^L \).

Lemma 1: Let \( x \in \mathbb{R} \). Then

\[
P[3t \in [0, T]: |\Delta \theta| > x] \leq 4Q \left[ \frac{x}{4\gamma} \right]
\]

where \( Q(x) \triangleq 1/\sqrt{2\pi} \int_x^{\infty} e^{-\lambda^2} \, d\lambda \).
For \( n = 0, + \), we define the stopping times \( T_n = \inf \{ t : W_t > n \} \) and \( T_n = \inf \{ t : W_t < -n \} \). Then,

\[
P(3t \in [0, T) : |A_t| > n) = P(3t \in [0, T) : |\theta_{T_n} - \theta_T| > n)
\]

\[
\leq P(3t \in [0, 2T) : |\theta_{T_n} - \theta_T| > n)
\]

\[
P(3t \in [0, 2T) : |\theta_T| > n)
\]

\[
P(\min(T_n, T_0) < 2T)
\]

\[
P(T_n < 2T) + P(T_0 < 2T)
\]

\[
= 2P(T_n < 2T) + P(T_0 < 2T)
\]

\[
= 4P(T_n < 2T, W_T > n)
\]

\[
= 4P(W_T > n)
\]

Equation (18) follows from the Markov property of the Brownian motion, (19) from the union bound, (20) from the fact that \(- W_t\) has the same probability law as \( W_t \), and (21) from the reflection principle of the Wiener process.

It should be pointed out that a tighter upper bound is possible by using the first passage times for the process \( \{A0, A1\} \) from the reflection principle of the Wiener process. However, the easy upper bound used in the proof suffices for our needs.

**Lemma 2:** Let \( k \) be an arbitrary, nonnegative integer. Lower bounds on the conditional error probabilities are

\[
1 - 4Q \left[ x \frac{\gamma}{4} + \frac{\gamma}{\pi} \right] - E[P(N(\Delta t) > 1) | \{\Delta t, 0 \leq t \leq T\}]
\]

where \( 1 - a x^2 / 2 \geq \cos x \forall x \in [0, x_0] \), and

\[
\sum_{k=0}^{\infty} \frac{1}{k!} (2j)^k e^{-2j} \leq E[P(N(\Delta t) > 1) | \{\Delta t, 0 \leq t \leq T\}]
\]

**Proof:** Let \( K \) be the indicator function of event \( A \). Then

\[
E[P(N(\Delta t) > 1) | \{\Delta t, 0 \leq t \leq T\}]
\]

\[
= 1 - E[P(N(\Delta t) > 1) | \{\Delta t, 0 \leq t \leq T\}]
\]

\[
= 1 - E[P(N(\Delta t) > 1) | \{\Delta t, 0 \leq t \leq T\}]
\]

\[
\cdot \left[ \{\Delta t \leq x_0, \forall x \in [0, x_0]\} + 1 \{x \in (0, x_0) | \Delta t > x_0\} \right]
\]

\[
\geq 1 - E[P(N(\Delta t) > 1) | \{\Delta t, 0 \leq t \leq T\}]
\]

\[
\cdot \left[ \{\Delta t \leq x_0, \forall x \in [0, x_0]\} - P(\exists t \in [0, T) : |\Delta t| > x_0) \right]
\]

\[
\cdot E[P(N(\Delta t) > 1) | \{\Delta t, 0 \leq t \leq T\}]
\]

\[
\cdot \left[ \{\Delta t > x_0, \forall x \in [0, x_0]\} \right]
\]

where \( 1 - a x^2 / 2 - \cos x \geq 0 \forall x \in [0, x_0] \).

Applying Lemma 1, we have proved the first bound of this claim. To prove the second bound, we recognize the fact that \( A_i < 2 \), and that the cdf of \( N(A) \) is a monotonically decreasing function of \( A \).

As a direct result of Lemma 2 a lower bound on the total probability of error is

\[
1 - 4Q \left[ x \frac{\gamma}{4} + \frac{\gamma}{\pi} \right] + \sum_{k=0}^{\infty} \frac{1}{k!} (2j)^k e^{-2j}
\]

\[
\sum_{k=0}^{\infty} \frac{a_k}{k!} d_k^k e^{-2j} \Phi(\mu) \] is defined earlier, and \( l \) is an arbitrary nonnegative integer. From (17) as well the fact that the Poisson cdf is a decreasing function of the mean, an upper bound on the total error probability is

\[
2P[error] \leq 1 + \sum_{k=0}^{\infty} \frac{1}{k!} d_k^k e^{-2j} \Phi(\mu) \]

Both (22) and (23) depend on \( \Phi(\mu) \), which is computed in the Appendix.

In the next lemma, we show that the bounds in (22) and (23) converge as \( x \to -\infty \).

**Lemma 3:** Let \( k \) be an arbitrary, nonnegative integer.

\[
\lim_{x \to -\infty} 1 + \sum_{k=0}^{\infty} \frac{1}{k!} d_k^k e^{-2j} \Phi(\mu) \]

\[
= \lim_{x \to -\infty} 1 - 4Q \left[ x \frac{\gamma}{4} + \frac{\gamma}{\pi} \right] + \sum_{k=0}^{\infty} \frac{1}{k!} (2j)^k e^{-2j}
\]

\[
\sum_{k=0}^{\infty} \frac{a_k}{k!} d_k^k \Phi(\mu) \]

**Proof:** We rewrite (22) as

\[
1 - 4Q \left[ x \frac{\gamma}{4} + \frac{\gamma}{\pi} \right] + \sum_{k=0}^{\infty} \frac{1}{k!} (2j)^k e^{-2j}
\]

\[
- E[P(N(\Delta t) > 1) | \{\Delta t, 0 \leq t \leq T\}] \leq 2P[error].
\]

Taking limits of both sides as \( x \to -\infty \), we have

\[
P(N(2T) \leq 1) + \lim_{x \to -\infty} E[P(N(\Delta t) > 1) | \{\Delta t, 0 \leq t \leq T\}] \leq 2P[error].
\]

The last two terms cancel by an application of the bounded convergence theorem, and the continuity of the Poisson cdf with the mean. The same result follows from the limit of (23) by similar arguments.

Since the upper and lower bounds result by replacing the mgf of \( 1/T \) with \( \cos \Delta t, dt \) by that of \( 1/T \) for \( \Delta t \) between the range of values of interest in the analysis of the proposed detector. The theoretical mgf of \( 1/T \) begins to differ from the mgf of \( 1/T \) obtained by Monte Carlo simulation at \( y = 30 \), which is conservative well below the range of values of interest in the analysis of the proposed detector. The theoretical mgf of \( 1/T \) for \( \Delta t \) begins to differ from the mgf of \( 1/T \) obtained by Monte Carlo simulation at \( y = 26 \). As seen in the Appendix, this sharp rise of the mgf is due to a branch point in
Fig. 2. Moment generating functions ($\gamma = 30$).

Fig. 3. Error rate bounds and simulation results (6 photons/bit).

the infinite product expression of the mgf. As $\gamma$ increases, this branch point occurs for smaller values of $v$. Fortunately, we are concerned with the region $v \in [-1, 1]$ where we evaluate the mgf for the error probability bounds.

Fig. 3 presents the upper and lower bounds on the error probability of the test in (10) for an optical power of 6 photons per bit. Also shown are results of Monte-Carlo simulations of the hypothesis test. It appears from the simulation results that the upper bound is tighter than the lower bound. This is because the lower bound on $E[P(N(t)) \leq I] \{\Delta t \leq t \leq \tau \}$ obtained in Lemma 2 is derived by trivially upper bounding $\cos(\chi)$ by unity for $|\chi| > x_\gamma$. In Fig. 3, it can also be seen that the error probability bounds are discontinuous functions of $v$. These discontinuities occur for values of $(\zeta, \gamma)$ where the suboptimal threshold, given by the RHS of (10), changes value. Indeed, these discontinuities result from the use of a suboptimal threshold. As the LHS of (10) is integer-valued, it is straightforward to optimize the threshold so as to minimize either the upper or lower bounds. Because of the tightness exhibited by the upper bound, we choose the threshold function which minimizes (23). Additionally, since the process $\{X(t), t \in [0, T]\}$ is close to the mean $E[X(t)]$ for moderate $\gamma$, the suboptimum threshold function in (10) is equal to the optimized threshold function except in very small intervals in the range of interest of $\gamma$. In Fig. 4, we have displayed the lower envelope of the upper bounds corresponding to all integer thresholds, (which is, obviously, an upper bound to the error probability of the test obtained with the optimum threshold) together with the lower bound computed at the threshold that minimizes the upper bound for several values of $\gamma$. That is, Fig. 4 displays (22) and (23) replacing $I$ by the optimized threshold function.

Fig. 4. (a) Error probability bounds with optimized threshold. (b) Error probability bounds with optimized threshold.

Fig. 5. Bounds on the power penalty of proposed DPSK receiver.

The power penalty is an alternative way to characterize the receiver performance. Fig. 5 shows bounds on the power penalty of the proposed DPSK receiver at $10^{-4}$ BER. These curves were obtained by recording the values of $\gamma$, for fixed optical energy, at which the lower bound (22) and upper bound (23) were equal to $10^{-4}$. The optimized threshold was employed for these curves as well. It should be noted that the lower bounding curve in Fig. 5 is a smooth lower bounding envelope to the power penalty data. By comparison, the power penalty for the balanced DPSK receiver as described in [2] is always greater than 3 dB, and attains this value only as $\gamma \rightarrow \infty$. As Fig. 5 illustrates, the power penalty for the quantum-limited DPSK receiver is below 3 dB for $\gamma > 700$.

IV. SUMMARY

In this paper, we have analyzed the error probability of an asymptotically quantum-limited direct-detection DPSK receiver. The receiver consists of a delay-and-sum optical preprocessor in tandem with a photoelectric converter and an integrate-and-dump circuit. The output is initially compared to a suboptimal threshold that was derived under the assumption that the conditional rates are constants. We tightly bounded the error probability for arbitrary thresholds by developing upper and lower bounds on the conditional intensities of the photon point process at the photodetector. Prompted by the tightness of the upper bound, we then improved the receiver performance by minimizing the upper bound over all integer threshold.
and as in (8). By substituting these equations into (A.2), we find that \{X_{i}, i = 1, 2 \cdots \} are unknowns.

Equation (A.5) suggests the general solution

\[ \phi(i) = A_i \cos \omega_0 t + B_i \sin \omega_0 t \quad 0 < \sigma < T. \]  

If we substitute (A.6) into (A.3) and (A.4) to solve for the unknowns \(A_i, \omega_0,\) and \(B_i,\) we find that \(\{\lambda_i, i = 1, 2 \cdots \}\) are the solutions to

\[ \frac{1}{2} \sqrt{\frac{2 \pi}{\lambda_i}} \sin \frac{2 \pi x_0}{\sqrt{\lambda_i}} = 1 + \cos \frac{2 \pi x_0}{\sqrt{\lambda_i}}. \]

Included among these eigenvalues are \(\{2\pi/\sqrt{(2n+1)^2}, n = 0, 1, 2 \cdots \}.\) The remaining portion of the eigenvalues are close to, but not exactly \(\{2\pi/\sqrt{2n}, n = 1, 2 \cdots \}.\) Now that the eigenvalues are known, the moment generating function of \(1/2T \int_0^T \Delta \theta_i dt\) may be found by (A.1).

REFERENCES


[11] D. T. P. M. L. Philipps, "Generating function of the random variable \(1/2\sqrt{2 \pi},\) the moment generating function of the random variable \(1/2 \int_0^T \Delta \theta_i dt.\) A more general problem has been solved previously [12]. Consider the random process \(\{x_t = \int_0^t \tau(t, \tau) d\tau, t \in \mathbb{R},\}\) where \{\(x_t, \tau \in \mathbb{R},\}\) is a zero mean, wide-sense stationary random process with autocorrelation function \(R(\tau).\) Then the moment generating function \(M_\theta(t)\) is given in infinite product form as

\[ M_\theta(t) = \prod_{i=1}^\infty \frac{1}{1 - 2 \pi i^2 \Delta \theta_i} \]  

where \{\(\lambda_i, i = 1, 2 \cdots \}\} are obtained by solving the homogeneous integral equation

\[ \lambda_i \phi(i) = \int_0^T R(\tau - \sigma) \phi(i) d\tau. \]  

For our particular case, we have

\[ \phi(i) = \frac{2 \pi}{\lambda_i} \phi(i) \quad 0 < \sigma < T. \]

Similar equations result for other values of \(\sigma.\) Taking the first- and second-partial derivatives of (A.3) with respect to \(\sigma\) we get

\[ \lambda_i \phi(i) = \frac{\pi \int^T \sigma T - |\sigma - \tau| \phi(i) d\tau}{\gamma} \]  

\[ \lambda_i \phi(i) = \frac{\pi \int^T \sigma T - |\sigma - \tau| \phi(i) d\tau}{\gamma} \quad 0 < \sigma < T. \]

and

\[ \phi(i) = 2 \pi \phi(i) \quad 0 < \sigma < T. \]  

Equation (A.5) suggests the general solution

\[ \phi(i) = A_i \cos \omega_0 t + B_i \sin \omega_0 t \quad 0 < \sigma < T. \]  

If we substitute (A.6) into (A.3) and (A.4) to solve for the unknowns \(A_i, \omega_0,\) and \(B_i,\) we find that \(\{\lambda_i, i = 1, 2 \cdots \}\) are the solutions to

\[ \frac{1}{2} \sqrt{\frac{2 \pi}{\lambda_i}} \sin \frac{2 \pi x_0}{\sqrt{\lambda_i}} = 1 + \cos \frac{2 \pi x_0}{\sqrt{\lambda_i}}. \]
Multiple-Access Channels with Memory with and without Frame Synchronism

Sergio Verdu

Reprinted from
IEEE TRANSACTIONS ON INFORMATION THEORY
vol. 35, No. 3, May 1989
Multiple-Access Channels with Memory with and without Frame Synchronism

SERGIO VERDÚ, SENIOR MEMBER, IEEE

Abstract—The capacity region of frame-synchronous and asynchronous discrete two-user multiple-access channels with finite memory is obtained. Frame synchronism refers to the ability of the transmitters to send their codewords in unison. The absence of frame synchronism in memoryless multiple-access channels is known to result in the removal of the convex hull operation from the expression of the capacity region. We show that when the channel has memory, frame synchronism rules out nonstationary inputs to achieve any point in the capacity region, thereby allowing only coding strategies that involve cooperation in the frequency domain but not in the time domain. This restriction drastically reduces the capacity region of some multiple-access channels with memory, and in particular the total capacity of the channel, which is invariant to frame synchronism for memoryless channels.

I. INTRODUCTION

THE CENTRAL result in multiuser information theory states that the capacity region of a two-user discrete multiple-access channel is equal to the convex closure of the set of rate pairs \( (R_1, R_2) \) satisfying

\[
0 \leq R_1 \leq I(X;Z|Y),
\]

\[
0 \leq R_2 \leq I(Y;Z|X),
\]

\[
R_1 + R_2 \leq I(X,Y;Z)
\]

for some independent input distributions \( X \) and \( Y \) and output distribution \( Z \). This result was obtained by Ahlswede [1] in 1971 under the key assumptions that the channel is memoryless and frame (or block)-synchronous, i.e., the beginnings of the codewords sent by the transmitters coincide. In the absence of frame synchronism, an unpredictable offset exists between the epochs at which the codewords of each user are received at the decoder. Even though the receiver can easily acquire timing synchronism with each user and hence know the value of the offset prior to decoding, the transmitters must encode their messages without knowing the offset between their codewords. Assuming that the offset can be guaranteed to be negligible with respect to the codeword length (e.g., if an upper bound on the offset is known by the transmitters), Cover et al. [5] and Narayan and Snyder [15] proved that the capacity region and the cutoff rate region, respectively, of the discrete memoryless multiple-access channel are the same as in the frame-synchronous case. Polytyrev [18] and, independently, Hui and Humblet [12] have shown that if no information on the actual value of the offset is available to the transmitters (i.e., if the channel is "completely" frame-synchronous), then the capacity region of the discrete memoryless multiple-access channel is as stated above but without the convex hull operation. This implies that in most memoryless channels of interest, frame asynchronism does not change the capacity region, the most well-known exceptions being, perhaps, the Massey-Mathys collision channel without feedback [13] and the counterexamples in [6, p. 287] and [3]. Furthermore, the Polytyrev-Hui-Humblet result implies that the maximum achievable rate sum \((R_1 + R_2)\) (or total capacity) of the memoryless multiple-access channel is never decreased by the lack of frame synchronism, because the rate sum of any convex combination of rate pairs is equal to the convex combination of the respective rate sums. As we shall see, these conclusions are no longer true when the multiple-access channel has memory.

Even though the study of the capacity of single-user channels with memory has occupied a prominent position in the development of the Shannon theory, multiple-access channels with memory have received scant attention in the literature (see van der Meulen [14, open problem 12]). Aside from their inherent conceptual and practical interest, multiple-access channels with memory play a key role in the modeling of symbol-asynchronous channels. These are continuous-time channels where each codeword symbol modulates a finite-duration signal waveform and the transmitters do not cooperate so that the symbol epochs are aligned at the receiver. Since each symbol overlaps with two consecutive symbols transmitted by the other user, the equivalent discrete-time multiple-access channels required to model symbol-asynchronous channels have memory [21]. Therefore, the study of multiple-access channels where the transmitters are completely asynchronous leads to frame-asynchronous discrete-time multiple-access channels with memory.

One more type of channel "asynchronism" is that which allows deletions and insertions of symbols at locations unknown to the decoder. This has been studied by Dobrushin [7] and by Ahlswede and Csis [2] in the context of single-user and multiple-access memoryless channels, respectively.

Manuscript received November 23, 1987; revised August 25, 1988. This work was supported in part by the Office of Naval Research under contract N00014-87-K-0054. The material in this paper was partially presented at the 1988 IEEE International Symposium on Information Theory, Kobe, Japan, June 1988.

018-9448/89/0500-0605$01.00 ©1989 IEEE
The multiple-access channel with memory studied in this paper has finite input alphabets $A_1$ and $A_2$ and finite output alphabet $B$. Except for the general converse theorem for synchronous channels proved in Section II, our results are obtained under the assumption that the multiple-access channel is stationary and has finite (input) memory, in the sense that each channel output depends on up to $m$ consecutive inputs of each user, and the outputs are conditionally independent given the inputs. The multiple-access channel with finite memory encompasses many cases of practical interest such as the symbol-asynchronous channel and channels with finite-length intersymbol interference; its capacity has been solved in the single-user case in the works of Tsaregradsky [20], Feinstein [8], and Wolfowitz [22].

As usual when dealing with sources or channels with memory, the capacity region of the multiple-access channel with memory does not admit single-letter characterizations and, rather, is given in terms of a limit of regions. This fact does not curtail the applicability or interest of these results because those limits are computable, as we show in several examples where they result in explicit closed-form expressions. Moreover, we provide a theorem (which generalizes Wolfowitz's result [22, theorem 5.5.1]) on the speed of convergence of single-user capacity) that allows the computation of the capacity region of the channel with memory up to any desired degree of approximation via the computation of achievable regions for memoryless channels.

As in the case of the memoryless multiple-access channel, the frame-synchronous channel with memory is shown to satisfy the time-sharing principle, i.e., its capacity region is convex. As a form of cooperation in the time domain, time-sharing requires nonstationary input distributions. Note that, while stationary inputs always achieve capacity in time-invariant single-user channels, there are time-invariant multiple-access channels (e.g., the aforementioned channels whose capacity region is decreased without frame-synchronism) that require nonstationary inputs to achieve all points in the capacity region. In this paper we show that only stationary inputs are allowed for frame-asynchronous multiple-access channels with memory. Hence cooperation between the users is beneficial in the frequency domain (dependent inputs are necessary to achieve capacity because the channel has memory) but not in the time domain due to the lack of a common time reference. The opposite situation is encountered in the frame-synchronous memoryless multiple-access channel, where it is enough to restrict attention to independent input sequences and time-sharing (hence nonstationary) inputs may be required to achieve capacity. In the light of our results, the Poltyrev--Hui--Humblet result for memoryless channels is a consequence of the nonstationarity of time-sharing strategies.

The actual impact that the lack of frame synchronism (i.e., the restriction to stationary inputs) has on the capacity region of the multiple-access channel with memory is quite diverse. On one hand, there are many frame-synchronous channels (e.g., the symbol-asynchronous multiple-access channel considered in [21]) whose capacity regions are achieved by stationary inputs, and therefore, they do not decrease if the users are not guaranteed to transmit their codewords in unison. On the other hand, we show in this paper the existence of channels with memory where not only the capacity region but the total capacity is drastically reduced by the lack of frame synchronism.

II. FRAME-SYNCHRONOUS CAPACITY REGION

We give first a general converse coding theorem for the discrete frame-synchronous multiple-access channel that puts no restrictions on its transition probabilities.

Theorem 1: The capacity region of the discrete frame-synchronous multiple-access channel satisfies

$$ C \subseteq \text{closure} \left( \liminf_{n \to \infty} \frac{1}{n} C_n \right) $$

where

$$ C_n = \bigcup_{X,Y,Z} \{ (R_1, R_2) : 0 \leq R_i \leq I(X^n; Z^m | Y^n) $$

$$ 0 \leq R_1 + R_2 \leq I(X^n, Y^n; Z^m) $$

and the union is over independent $n$-dimensional input distributions. Note that the convex closure of $C_n$ is the capacity region of the discrete memoryless multiple-access channel whose input and output alphabets are $A_1^n$, $A_2^n$, and $B^n$, respectively, and whose transition probabilities are $P_{1|X,Y}$. The capacity region of the discrete memoryless multiple-access channel is denoted by $C_{\text{convex}}$.

Proof: We need to show that, for all $0 < \epsilon < 1$, every $\epsilon$-achievable rate pair $(R_1, R_2)$ belongs to the right side of (2). If $(R_1, R_2)$ is $\epsilon$-achievable, then for all $\gamma > 0$ and for all sufficiently large $n$ there exists an $(n, M_1, M_2, \epsilon)$ code (i.e., a code with block length $n$, $M_1$, codewords for user $i$, and average probability that both messages are correctly decoded greater or equal than $1 - \epsilon$) such that

$$ \begin{align*}
\log M_i &= \frac{1}{n} R_i - \gamma, \\
&= 1, 2,
\end{align*} $$

Fix one such code and let $S_1$ and $S_2$ denote independent random variables uniformly distributed on $\{1, \ldots, M_1\}$, and $\{1, \ldots, M_2\}$. The message transmitted by user $i$ is a realization of $S_i$. Let $Z^n$ denote the output of the channel, when $S_1$ and $S_2$ are transmitted using the above $(n, M_1, M_2, \epsilon)$ code, and let $(S_1, S_2)$ be the messages se-
lected by the decoder. The Fano inequality states that
\[ H(S_1, S_2 | Z^*) \leq \log 2 + P \left( (S_1, S_2) \neq (\hat{S}_1, \hat{S}_2) \right) \]
\[ \log M_1 M_2 \]  
(5a)
\[ H(S_1 | Z^*) \leq \log 2 + P \left( \hat{S}_1 \neq S_1 \right) \log M_1 \]  
(5b)
\[ H(S_2 | Z^*) \leq \log 2 + P \left( \hat{S}_2 \neq S_2 \right) \log M_2. \]  
(5c)
Since the average probability of error of the code does not exceed \( \epsilon \), the probabilities in the upper bounds of (5) can be replaced by \( \epsilon \), and because \( S_1 \) and \( S_2 \) are uniformly distributed, we can write
\[ I(S_1; Z^*) \geq (1 - \epsilon) \log M_1 - \log 2 \]  
(6a)
\[ I(S_2; Z^*) \geq (1 - \epsilon) \log M_2 - \log 2 \]  
(6b)
\[ I(S_1, S_2; Z^*) \geq (1 - \epsilon) \log M_1 M_2 - \log 2. \]  
(6c)
If \( f_i; \{1, \ldots, M_i\} \to A_i^* \) denotes the code book of user \( i \), then since \( S_1 \) and \( S_2 \) are independent,
\[ I(f_i(S_i); Z^n | f_i(S_i)) = I(f_i(S_i); f_i(S_i)) \]
\[ + I(f_i(S_i); Z^n | f_i(S_i)) \]
\[ = I(f_i(S_i); Z^n, f_i(S_i)) \]
\[ \geq I(f_i(S_i); Z^n) \]
\[ \geq I(S_i; Z^*) \]  
(7)
where the last inequality follows from the data-processing lemma. In a similar way, we obtain
\[ I(f_i(S_i); Z^n | f_i(S_i)) \geq I(S_i; Z^*) \]  
(8a)
\[ I(f_1(S_1), f_2(S_2); Z^n) \geq I(S_1, S_2; Z^*). \]  
(8b)
Now, putting (4) together with (6)–(8), we get
\[ (1 - \epsilon)(R_1 - \gamma) - \frac{\log 2}{n} \leq - \frac{1}{n} I(f_i(S_i); Z^n | f_i(S_i)) \]
\[ (1 - \epsilon)(R_2 - \gamma) - \frac{\log 2}{n} \leq - \frac{1}{n} I(f_i(S_i); Z^n | f_i(S_i)) \]
\[ (1 - \epsilon)(R_1 + R_2 - 2\gamma) - \frac{\log 2}{n} \leq - \frac{1}{n} I(f_i(S_i); f_i(S_i); Z^n) \]
which implies that
\[ (1 - \epsilon)(R_1 - \gamma, R_2 - \gamma) = 1 - C_n \]  
for all sufficiently large \( n \), and consequently,
\[ (1 - \epsilon)(R_1 - 2\gamma, R_2 - 2\gamma) = 1 - C_n \]  
for all sufficiently large \( n \), or in other words
\[ 1 - \epsilon)(R_1 - 2\gamma, R_2 - 2\gamma) \leq \liminf_{n \to \infty} 1 - C_n. \]  
(9)
However, since \( \epsilon \) and \( \gamma \) are arbitrarily small, (9) implies that \( (R_1, R_2) \) must be the limit of a sequence of points converging to \( \liminf_{n \to \infty} (1/n) C_n \), and therefore it belongs on the right side of (2) (as was to be shown).

In the case of the single-user channel with memory, no universal direct coding theorem is known to hold with the same generality as Theorem 1. We will henceforth focus our attention on the following class of multiple-access channels with memory.

**Definition:** A multiple-access channel with finite memory \( m \) is one whose channel transition probability satisfies
\[ p(z_i | x_i, \ldots, x_{i-m}) = p(z_i | x_i, \ldots, x_{i-m}, y_{i-1}, \ldots, y_{i-1}) \]
\[ = \prod_{i=m}^n p_r(w_i | a_{i-m+1}, \ldots, a_i, b_{i-m+1}, \ldots, b_i) \]  
(10)
for all \( n > 0 \).

This implies that the outputs \( Z_m, \ldots, Z_n \) are conditionally independent given the inputs, and each of them depends on \( m \) consecutive inputs of each user, thus encompassing intersymbol interference of finite duration. This definition allows us to handle the boundary outputs \( Z_1, \ldots, Z_{m-1} \) (which depend on fewer than \( m \) input symbols from each user) in any arbitrary way, and it is therefore preferable to the single-user definition of [8] and [22] where the boundary outputs are not available to the decoder. As we shall see, the capacity region of the multiple-access channel with memory depends only on the transition probability \( p_r \) and not on the conditional distribution of the first \( m - 1 \) outputs.

In what follows it is convenient to refer to a memoryless multiple-access channel derived from the channel with memory in the following way.

**Definition:** Let \( l \geq m \). The \( l \)-block multiple-access channel derived from a multiple-access channel with finite memory \( m \) is a memoryless channel characterized by input alphabets \( A_1^l \) and \( A_2^l \), output alphabet \( B^{l-m+1} \), and transition probability
\[ p_r(w_m, \ldots, w_l) | (a_1, \ldots, a_l), (b_1, \ldots, b_l) \]
\[ = \prod_{l=m}^n p_r(w_i | a_{i-m+1}, \ldots, a_i, b_{i-m+1}, \ldots, b_i). \]  
(11)

It follows from (1) that the capacity region of the \( l \)-block memoryless multiple-access channel is equal to the convex closure of
\[ Q_l = \bigcup_{X', Y'} \{ (R_1, R_2) \in [X', Y'] : 0 \leq R_1 \leq I(X'; Z_{m-1}' | Y') \}
\[ 0 \leq R_2 \leq I(Y'; Z_m' | X') \}
\[ X_1 + R_2 \leq I(X'_l, Y'_l; Z_{m-1}' | X') \} \]  
(12)

where the union is over independent distributions on the sets \( A_1^l \) and \( A_2^l \), respectively, and \( Z_m'^{l-m} = (Z_m, \ldots, Z_l) \). The direct coding theorem for the multiple-access channel with finite memory gives the following achievable region as a function of the achievable region in (12) for the \( l \)-block multiple-access channel.

**Theorem 2:** The capacity region of the frame-synchronous multiple-access channel with finite memory \( m \) satis-
For sufficiently large \( tR/((k + t) \gamma / 2) \), however, in each case the right side of (16) can be further lower bounded by \( R_i - \gamma \). Thus even though the new code has lower rates than the original code with block length \( kl \), the decrease is inappreciable for large \( k \).

The following result proves that the inner and outer bounds shown in Theorems 1 and 2 coincide.

**Theorem 3:** The capacity region of the frame-synchronous multiple-access channel with finite memory \( m \) is given by

\[
C = \text{closure} \left( \bigcup_{n \geq m} \frac{1}{n} Q_* \right) = \text{closure} \left( \liminf \frac{1}{n} Q_* \right)
\]

Proof: The essence of the proof is the following inequlity which holds for all \( l \geq m \):

\[
\begin{align*}
\sum_{t=0}^{l-1} I(Y^t; Z^t) &= I(X^t, Y^t; Z^t) + I(X^t, Y^t; Z^t | Z^t) \\
&\leq l(X^t, Y^t; Z^t) + (m-1) \log |B| \quad (19a)
\end{align*}
\]

and similarly,

\[
\begin{align*}
I(X^t; Z^t) &\leq l(X^t; Z^t) + (m-1) \log |B| \quad (19b) \\
I(Y^t; X^t) &\leq l(Y^t; X^t) + (m-1) \log |B| \quad (19c)
\end{align*}
\]

which imply that

\[
C \subseteq \bigcup_{n \geq m} \frac{1}{n} Q_* + (m-1) \log |B| U
\]

where \( U \) is the unit square \( \{(x_1, x_2): 0 \leq x_1, x_2 \leq 1\} \).

We now have the following chain of inclusions:

\[
\text{closure} \left( \limsup \frac{1}{n} C_\gamma \right) \subseteq \text{closure} \left( \liminf \frac{1}{n} Q_* \right) \subseteq C \subseteq \text{closure} \left( \liminf \frac{1}{n} C_\gamma \right)
\]

(20)

where the first and last inclusions follow easily from (19) and the third and fourth inclusions are Theorems 2 and 1, respectively. Finally, since the liminf is a subset of the limsup all the inclusions in (20) are in fact equalities.

The closure operation in (17) is indeed necessary because even if \( \lim_{n \to \infty} (Q_* / n) \) exists, it may not be a closed set (e.g., if the first \( m-1 \) boundary outputs are independent of the inputs). At first sight it may seem surprising that the capacity region of Theorem 3 does not involve an explicit convex hull operation, especially in light of the fact that the particular case of the frame-synchronous memoryless multiple-access channel is known to require the convex hull operation. In fact the capacity region of Theorem 3 is already convex because it is given as a limit of achievable regions for \( n \)-block channels whose input distributions are allowed to time-share among several distributions as a...
result of the assumption that the users are frame-synchronous. This is formalized in the following result.

Corollary (Time-Sharing Principle): The capacity region of the frame-synchronous multiple-access channel with finite memory is a convex set.

Proof: It follows from Theorem 3 that the capacity region is independent of the conditional probability of the \((m-1)\) boundary outputs; therefore, we can prove the corollary for any arbitrary choice of this probability; in particular, we shall assume that \(Z_1, \ldots, Z_{m-1}\) are independent of the inputs. Then we can follow the same approach as in the proof of the time-sharing principle for memoryless channels [6] which juxtaposes two codes. If we impose the restriction that the decoder must discard the leading \((m-1)\) output symbols of each of the two blocks, then the decoding of the new code is decoupled and equivalent to the case when the codewords are sent sequentially. Therefore, the error probability of the new code is better than the sum of the probabilities of error of the two component codes, the rate pair is a convex combination of the rate pairs of both codes, and the proof proceeds as in the memoryless case [6, p. 272].

Theorem 3 can easily be generalized in several directions. For example, the proof of both the converse and the direct theorems remain essentially unchanged for continuous-alphabet channels with input constraints. Another generalization which is of interest in the symbol-asynchronous channel [21] is that of a compound multiple-access channel where the transmitters only know that the channel belongs to an uncertainty set \(\Gamma\) (cf. [6, p. 288] for the corresponding memoryless result). In that problem, the proof of the direct theorem requires very little modification since the construction of codebooks therein is independent of the channel, and the proof of the converse only needs to take care of the fact that a good code must be so for any possible channel in the uncertainty set. Then Theorem 3 holds by replacing \(C_0\) by

\[
C'_n = \bigcup_{X^m \in \Gamma} \bigcap \{ (R_1, R_2); 0 \leq R_1 \leq I(X^n; Z^n(\omega))|X^n) \\
0 \leq R_2 \leq I(Y^n; Z^n(\omega)|X^n) \\
R_1 + R_2 \leq I(X^n, Y^n; Z^n(\omega)) \}
\]

where \(Z^n(\omega)\) is connected to \(X^n\) and \(Y^n\) through channel \(\omega \in \Gamma\).

For the purposes of illustration we will show several examples where the limits of Theorem 3 are explicitly computable. However, in cases without much structure an alternative to the analytical computation of those limits is their numerical approximation. This can be done using the following theorem, which allows the computation of the capacity region as accurately as desired via the computation of achievable regions for memoryless channels. Theorem 4 is a generalization of the single-user result obtained by Wolfowitz [22, theorem 5.5.1].

Theorem 4: The capacity region of the frame-synchronous multiple-access channel with finite memory satisfies for every \(l \geq m\)

\[
\text{closure} \left( \frac{1}{l} Q_l \right) \subseteq C \subseteq \text{closure} \left( \frac{1}{l} Q_l \right) + \frac{m-1}{l} \log |B| |U| \tag{21}\]

where \(U\) is the unit square \((x_1, x_2): 0 \leq x_1 \leq 1, 0 \leq x_2 \leq 1\).

Proof: The inner bound is a consequence of Theorem 2 and the corollary to Theorem 3. To show the upper bound, fix \(l \geq m\) and notice that for any \(n = kl, X^n\), and \(Y^n\),

\[
I(X^n, Y^n; Z^n) = I(X^n, Y^n; Z_{m+l}^{2l} \cdots Z_{m-(k-1)l}^{kl}) \\
+ I(X^n, Y^n; Z_{m+l}^{2l} \cdots Z_{m-(k-1)l}^{kl}) \\
\times \cdots \times Z_{m+(k-1)l}^{kl} \cdots Z_{m-(k-1)l}^{kl}) \\
\leq k(m-1) \log |B| \\
+ I(X^n, Y^n; Z_{m+l}^{2l} \cdots Z_{m+(k-1)l}^{kl}) \\
= k(m-1) \log |B| + H(Z_{m+l}^{2l} \cdots Z_{m+(k-1)l}^{kl}) \\
- H(Z_{m+l}^{2l} \cdots Z_{m+(k-1)l}^{kl}) \\
\leq k(m-1) \log |B| + \sum_{j=0}^{k-1} H(Z_{m+l}^{2l}Y_{m+j+1}^{j+1}|X^n, Y^n) \\
- H(Z_{m+l}^{2l} \cdots Z_{m+(k-1)l}^{kl}) \\
= k(m-1) \log |B| \\
+ \sum_{j=0}^{k-1} I(X_{m+j+1}^{j+1}Y_{m+j+1}^{j+1}; Z_{m+j+1}^{j+1}) \\
= k(m-1) \log |B| \tag{22}
\]

where the next-to-last identity follows from the definition of the channel with finite memory. Similarly, we can upper-bound

\[
I(X^n; Z^n|Y^n) \leq k(m-1) \log |B| \\
+ \sum_{j=0}^{k-1} I(X_{m+j+1}^{j+1}Z_{m+j+1}^{j+1}|Y_{m+j+1}^{j+1}) \tag{23a}
\]

\[
I(Y^n; Z^n|X^n) \leq k(m-1) \log |B| \\
+ \sum_{j=0}^{k-1} I(Y_{m+j+1}^{j+1}Z_{m+j+1}^{j+1}|X_{m+j+1}^{j+1}) \tag{23b}
\]

However, we saw in the proof of Theorem 1 that if \((R_1, R_2)\) is \(\gamma\)-achievable, then for all \(\gamma > 0\) and for all sufficiently large \(n\), there exist \((n, M_1, M_2, \epsilon)\) codes such that

\[
\frac{\log M_i}{n} \geq R_i - \gamma, \quad i = 1, 2 \tag{24}
\]
and for some input distributions $X^*$ and $Y^*$,
\[(1 - \epsilon) \log M_1 - \log 2 \leq I(X^*; Y^*)\]  
\[(1 - \epsilon) \log M_2 - \log 2 \leq I(Y^*; Z^*)\]  
\[(1 - \epsilon) \log M_1 M_2 - \log 2 \leq I(X^*, Y^*; Z^*)\]

Now, combining (22)-(25), we obtain for all sufficiently large $k$
\[
\frac{1}{k} \left[ (1 - \epsilon) \log M_1 - \log 2 \right] - (m - 1) \log |B| \\
\frac{1}{k} \left[ (1 - \epsilon) \log M_2 - \log 2 \right] - (m - 1) \log |B| \\
\frac{1}{k} \left[ (1 - \epsilon) \log M_1 M_2 - \log 2 \right] - (m - 1) \log |B| \\
\leq \frac{1}{k} \sum_{j=0}^{k-1} I(X_j^*; Y_j^*; Z_j^*) \\
\leq \frac{1}{k} \sum_{j=0}^{k-1} I(Y_j^*; Z_j^*; X_j^*) \\
\leq \frac{1}{k} \sum_{j=0}^{k-1} I(X_j^*; Y_j^*; Z_j^*) \\
\leq \frac{1}{k} \sum_{j=0}^{k-1} I(Y_j^*; X_j^*; Z_j^*)
\]

which implies that
\[
\frac{1}{k} (1 - \epsilon) (\log M_1, \log M_2) \\
\leq \text{convex} \left\{ Q_j + \frac{\log 2}{k} + \frac{(m - 1) \log |B|}{k} \right\} U
\]

for all sufficiently large $k$. This together with (24) implies that any $\epsilon$-achievable pair $(R_1, R_2)$ satisfies
\[
(1 - \epsilon)(R_1, R_2) \\
\leq \text{closure} \left( \text{convex} + Q_j + \frac{m - 1}{k} \log |B| U \right).
\]

Thus if $(R_1, R_2)$ is an achievable pair, then it must belong to the closed set in the right side of (26).

The following examples serve to illustrate the analytical evaluation of the capacity region of the frame-synchronous multiple-access channel with memory. In Section III, we derive the capacity of these channels in the absence of synchronism.

Example 1: Consider the following multiple-access channel with finite $(m = 2)$ memory which is a simple discrete-time noiseless model of two-user duobinary transmission: $A_1 = A_2 = (0, 1), B = (0, 1, 2, 3, 4)$
\[
z_i = x_i + x_{i-1} + y_i + y_{i-1}
\]

where, according to Theorem 3, it is not necessary to specify the initial conditions as far as computing the capacity region is concerned. To evaluate $C$, first we compute the mutual informations in the definition of $Q_*$ (12). Since the outputs are deterministic given the inputs,

we have
\[
I(X_1, \cdots, X_n; Y_1, \cdots, Y_n; Z_1, \cdots, Z_n) \\
= H(Z_1, \cdots, Z_n) \\
\leq H(Z_1, \cdots, Z_n|X_1 + Y_1) + H(X_1 + Y_1)
\]

and
\[
H(Z_1, \cdots, Z_n|X_1 + Y_1) \\
= H(Z_1|X_1 + Y_1) + H(Z_2|X_1 + Y_1|Z_1) \\
+ \cdots + H(Z_n|X_1 + Y_1, \cdots, Z_{n-1})
\]

Moreover, the properties of conditional entropy result in
\[
H(Z_1, \cdots, Z_n|X_1 + Y_1) \\
= H(X_1 + Y_1, \cdots, X_n + Y_n) - H(X_1 + Y_1).
\]

Also, using the definition of the channel and the fact that $(X_1, \cdots, X_n)$ and $(Y_1, \cdots, Y_n)$ are mutually independent, we can write
\[
H(Z_1, \cdots, Z_n|X_1, \cdots, X_n) \\
= H(X_1 + Y_1, \cdots, X_{n-1} + Y_{n-1})
\]

and, similarly,
\[
H(Z_1, \cdots, Z_n|Y_1, \cdots, Y_n) \\
= H(Y_1 + Y_2, \cdots, Y_{n-1} + Y_n).
\]

Now, putting together (28)-(31), we obtain
\[
C = \text{closure} \left( \liminf_{n \to \infty} \frac{1}{n} Q_n \right) \\
= \text{closure} \left( \liminf_{n \to \infty} \bigcup_{x^n y^n} \left( (R_1, R_2) : \right. \right.
\]
\[
0 \leq R_1 \leq \frac{1}{n} H(X_1 + X_2, \cdots, X_{n-1} + X_n) \\
0 \leq R_2 \leq \frac{1}{n} H(Y_1 + Y_2, \cdots, Y_{n-1} + Y_n) \\
R_1 + R_2 \leq \frac{1}{n} H(X_1 + Y_1, \cdots, X_n + Y_n) \right) \bigg)
\]

\[
= \text{closure} \left( \liminf_{n \to \infty} \bigcup_{x^n y^n} \left( (R_1, R_2) : \right. \right.
\]
\[
0 \leq R_1 \leq \frac{1}{n} H(X_1, \cdots, X_n) \\
0 \leq R_2 \leq \frac{1}{n} H(Y_1, \cdots, Y_n) \\
R_1 + R_2 \leq \frac{1}{n} H(X_1 + Y_1, \cdots, X_n + Y_n) \right) \bigg) \right) \bigg)
\]

(27)
and since each of the three entropies in (32) is maximized by independent equiprobable inputs \((\max_{x,y} H(X+Y)}\) over independent binary \(X\) and \(Y\) is equal to 1.5 bit and is achieved by equiprobable distributions), the right side of (32) is equal to the pentagon \(C = \{0 \leq R_1 \leq 1, 0 \leq R_2 \leq 1, R_1 + R_2 \leq 1.5\}\).

**Example 2:** Let \(A_1 = A_2 = (0,1,2)\) and \(B = (1,2)\)

\[
\begin{align*}
z_1 &= \begin{cases} 
  x_1, & \text{if } y_1 = 0 \\
  (1/2, 1/2), & \text{if } y_1 \neq 0
\end{cases} \quad (33)
\end{align*}
\]

where \((1/2, 1/2)\) indicates that \(z_1\) is equally likely to be 1 or 2.

In this channel it is necessary for the encoders to use some sort of time-sharing to achieve optimum rates because simultaneous zeros or nonzeros and consecutive zeros result in equally likely outputs. We take the following initial conditional distribution (this choice does not affect the capacity region but simplifies the proof):

\[
z_1 = \begin{cases} 
  x_1, & \text{if } y_1 = 0 \\
  (1/2, 1/2), & \text{if } y_1 \neq 0
\end{cases} \quad (34)
\]

We will now investigate the maximum achievable rates when transmitter 1 (respectively, 2) sends nonzeros at odd-numbered (respectively, even-numbered) times (with no restrictions otherwise). Then, it follows from (33) and (34) that

\[
z_{2k+1} = \begin{cases} 
  x_{2k+1}, & \text{if } y_{2k+1} = 0 \\
  (1/2, 1/2), & \text{if } y_{2k+1} \neq 0
\end{cases} \quad (35a)
\]

\[
z_{2k} = \begin{cases} 
  y_{2k}, & \text{if } x_{2k} = 0 \\
  (1/2, 1/2), & \text{if } x_{2k} \neq 0
\end{cases} \quad (35b)
\]

which means that the channel is actually decoupled into two identical memoryless channels whose capacity region is obtained as follows.

If \((X_{2k}, Y_{2k}, Z_{2k})\) are connected by (35b), then their mutual informations are easily shown to be given (in bits) by

\[
\begin{align*}
(X_{2k}; Z_{2k} | Y_{2k}) &= h_b(P[X_{2k} = 0]/2) - P[X_{2k} = 0] \\
(Y_{2k}; Z_{2k} | X_{2k}) &= h_b(P[Y_{2k} = 1])P[X_{2k} = 0] \\
(X_{2k}, Y_{2k}; Z_{2k}) &= h_b(1/2 + P[X_{2k} = 0]) - P[X_{2k} = 0]
\end{align*}
\]

where \(h_b(x) = -x \log x - (1-x) \log (1-x)\). All these mutual informations are maximized simultaneously by \(\{Y_{2k} = 1\} = P[Y_{2k} = 2] = 1/2\), and so the capacity region of the channel in (35b) is

\[
J = \left\{0 \leq R_1 \leq h_b\left(\frac{P}{2}\right), 0 \leq R_2 \leq h_b\left(\frac{P}{2}\right) \mid 0 \leq R_1 + R_2 \leq 1 - p \right\}
\]

and the capacity region of (35a) is obtained by interchanging \(R_1\) and \(R_2\) in (36). The sum of these two regions divided by 2 (since each channel is only used half the time) is found in Fig. 1. Another example where the capacity region of the multiple-access channel is explicitly computed is the symbol-asynchronous energy-constrained Gaussian channel ([21] is devoted to the evaluation of the limit characterizing the capacity region).

**III. FRAME-ASYNCHRONOUS CAPACITY REGION**

Unlike frame-synchronous channels where it is enough to consider "one-shot" models in which each user transmits only one codeword, the (completely) frame-asynchronous multiple-access channel cannot be decoupled into independent blocks due to the overlap between consecutive codewords, and the optimum decoder needs to decode all messages simultaneously, i.e., all outputs are useful in making decisions about any particular codeword. Ideally, the goal would be to analyze a model with doubly infinite streams of codewords subject to an arbitrary shift. However, to formulate a well-posed problem, it is necessary (at least within the realm of channel block coding) to work with a finite number \(N\) of transmitted codewords per user and then analyze the limiting behavior of the capacity as \(N \to \infty\). Since the offset between both strings is arbitrary, the approach we take is to arrange the \(N\) codewords of each user in a ring (codeword \(N\) is followed by codeword 1) and to model the offset by an arbitrary relative rotation of both rings (Fig. 2). As \(N \to \infty\), the radius of the ring becomes infinite, and the ring models the desired infinite codeword streams offset by an arbitrary shift because, for each output symbol, the boundary condition at infinity is irrelevant. As we will see and should expect (because of the finite memory the codeword boundaries become irrelevant.
as the codeword length goes to infinity), the capacity region (per channel use) of the $N$-ring does not depend on $N$, and therefore, it is not necessary to investigate its limiting behavior. The main disadvantage of an alternative linear arrangement of the $N$ codewords is that, due to the lack of synchronism, not all the codewords overlap with the other user's stream, and those that do overlap have different decoding error probabilities depending on the offset and their relative location to the boundaries. Note that this problem can be partially avoided by restricting the shift not to exceed the length of a codeword, but since the total number of codewords is assumed finite, it can be argued that such an approach would assume a certain degree of cooperation between the transmitters.

Each transmitter encodes its $N$ messages independently (each message is drawn independently from $\{1, \cdots, M_i\}$) and is not restricted to use the same code book for each message. While the receiver acquires the location of each codeword prior to decoding (this can be easily accomplished using synchronization prefixes), the messages are encoded without knowledge of the relative rotation. Therefore, the channel is a decoder-informed compound channel, which is equivalent, from the viewpoint of finding the capacity region, to a bank of parallel multiple-access channels (one per rotation value) sharing the same inputs.

**Theorem 5:** If $(R_1, R_2)$ is an achievable rate pair for the $N$-ring, then

$$ (R_1, R_2) \in \text{closure} \left( \liminf_{n \to \infty} \frac{1}{nN} Q_n^S \right) $$

$$ \subset \text{closure} \left( \limsup_{n \to \infty} \frac{1}{n} Q_n^S \right) $$

$$ = \text{closure} \left( \limsup_{n \to \infty} \frac{1}{n} C_n^S \right) $$

(37)

where $Q_n^S$ and $C_n^S$ are defined as in (12) and (3), except that the union therein is taken only over $n$-dimensional distributions induced by stationary probability measures.

**Proof:** There are $N (n, M_i)$ code books for the $i$th user and each of the $N$ messages are encoded independently. Thus the $(nN, M_i^N)$ juxtaposition code book of the $i$th user for the $N$-ring consists of the Cartesian product of the $N (n, M_i)$ code books. If $(R_1, R_2)$ is achievable, then for all $\epsilon > 0$, $\delta > 0$ and all $n$ sufficiently large, there exists a $(nN, M_i^N, M_j^N, \epsilon)$ juxtaposition code for the $N$-ring such that

$$ R_1 \leq \frac{\log M_i}{n} + \delta, $$

i.e., there exist $N (n, M_i)$ code books for user 1 and $N (n, M_j)$ code books for user 2 (which are independent of the offset) and a decoding strategy (which depends on the offset) such that the average (over the set of equiprobable messages) probability of error does not exceed $\epsilon$ regardless of the offset. Select one such code and denote the independent messages of both users by $(S_i, \cdots, S_N)$ and $(T_1, \cdots, T_N)$, respectively. Then, the Fano inequality implies that

$$ H(S_i, \cdots, S_N) \leq \epsilon \log M_i^N + \log 2 $$

(38a)

$$ H(T_1, \cdots, T_N) \leq \epsilon \log M_j^N + \log 2 $$

(38b)

$$ H(S_i, \cdots, S_N, T_1, \cdots, T_N) \leq \epsilon \log M_i^N M_j^N + \log 2 $$

(38c)

where $Z$ is the distribution of the totality of the outputs of the $N$-ring. If $X$ and $Y$ are the distributions of the $A_i^n$ and $A_j^n$ valued random variables resulting from the encoding of the messages by the selected code books, the lack of frame synchronism is modeled by assuming that the inputs to the $N$-ring are rotated versions of $X$ and $Y$. If $x$ is an $M$-vector, then $r_i(x)$ denotes an $M$-vector whose components coincide with those of $x$ rotated by $\tau$ positions, where $\tau \in \{0, \cdots, M-1\}$, i.e.,

$$ r_i(a_1, \cdots, a_M) = (a_{M-\tau+1}, \cdots, a_{\tau+1}). $$

Even though it is enough to consider a relative rotation of both rings, it is more convenient in the proof of the converse to allow a rotation of both input rings with respect to an arbitrary reference. Denoting the rotations by $\tau_1$ and $\tau_2$, the data-processing lemma implies that

$$ I(r_\tau(X); Z, r_\tau(Y)) \geq I(S_i, \cdots, S_N; Z) $$

$$ = H(S_i, \cdots, S_N) - H(S_i, \cdots, S_N | Z) $$

$$ \geq (1-\epsilon) \log M_i^N - \log 2 $$

(39)

where the last inequality follows from (38a) and with a slight abuse of notation we have denoted by $r_\tau(X)$ the probability measure that assigns the same mass to $r_i(a_1, \cdots, a_M)$ as $X$ assigns to $(a_1, \cdots, a_M)$. From the independence of $X$ and $Y$ and (39), we have that

$$ I(r_\tau(X); Z | r_\tau(Y)) \geq (1-\epsilon) \log M_i^N - \log 2. $$

However, since this is true regardless of the actual value of the offsets $\tau_1$ and $\tau_2$ we can write

$$ (1-\epsilon) \log M_i^N - \log 2 \leq \frac{1}{nN} \sum_{\tau_1=0}^{nN-1} I(r_\tau(X); Z | r_\tau(Y)) $$

$$ = I(r_\tau(X); Z | c(Y)) $$

(40)

where the probability measure $c(Y)$ is equal to the follow-
ing mixture of the probability measures \( r_\tau(Y) \), \( \tau = 0, \ldots, nN - 1 \):

\[
c(Y) = \frac{1}{nN} \sum_{\tau=0}^{nN-1} r_\tau(Y),
\]

and the second equation in (40) follows from the fact that the distribution of the conditioning random variable enters linearly in the definition of conditional mutual information.

An \( M \)-dimensional probability measure \( p \) will be referred to as circulant if \( r_\tau(p) = p \) for all \( \tau \), and we will say that \( p \) is stationary if, for any subset \( \{i_0, \ldots, i_M\} \subset \{1, \ldots, M\} \) and shift \( s > 0 \) such that \( \{i_0 + s, \ldots, i_M + s\} \subset \{1, \ldots, M\} \),

\[
p(a_{i_0}, \ldots, a_{i_M}) = p(a_{i_0 + s}, \ldots, a_{i_M + s})
\]

For any probability measure \( p \), \( c(p) \) (defined in (41)) is a circulant probability measure. To see this note that, for any \( \lambda \subset \{0, \ldots, M\} \),

\[
r_\lambda(c(p)) = \frac{1}{M} \sum_{\tau=0}^{M-1} r_{\tau+\lambda}(p) = \frac{1}{M} \sum_{\tau=0}^{M-1} r_\tau(p) = c(p).
\]

Furthermore, it is easy to check that an \( M \)-dimensional circulant probability measure is an \( M \)-dimensional stationary probability measure. Now, since (40) holds for all \( \tau \in \{0, \ldots, nN - 1\} \),

\[
(1 - \epsilon) \log M^N - \log 2 \leq \frac{1}{nN} \sum_{\eta=0}^{nN-1} I(r_\eta(X); Z|c(Y))
\]

\[
\leq I(c(X); Z|c(Y))
\]

(42)

where the second inequality follows from the concavity of mutual information. Proceeding in a similar way we obtain from (38)

\[
(1 - \epsilon) \log M^N - \log 2 \leq I(c(X); Z|c(Y))
\]

(43)

and

\[
(1 - \epsilon) \log M^N M^2 - \log 2 \leq I(c(X), c(Y); Z). \tag{44}
\]

If \( m - 1 \) consecutive components of \( Z \) are discarded, then we have a channel analogous to an \( nN \)-block channel whose \( B^{nN-m+1} \) valued output random variable is denoted by \( Z_{nN}^m \). Then, the following upper bounds follow in a way similar to (18)

\[
I(c(X); Z|c(Y))
\]

\[
\leq I(c(X); Z_{nN}^m|c(Y)) + (m - 1) \log |B|
\]

(45a)

\[
I(c(Y); Z|c(X))
\]

\[
\leq I(c(Y); Z_{nN}^m|c(X)) + (m - 1) \log |B|
\]

(45b)

\[
I(c(X), c(Y); Z)
\]

\[
\leq I(c(X), c(Y); Z_{nN}^m) + (m - 1) \log |B|
\]

(45c)

Finally, it follows from (38), (42)–(43) and the stationarity of \( c(X) \) and \( c(Y) \) that

\[
[1 - \epsilon](R_1 - \delta, R_2 - \delta) - \log \frac{2}{nN} + \frac{m - 1}{nN} \log |B|)
\]

\[
(1, 1) \leq \frac{1}{nN} Q^R_{nN}.
\]

Thus if \( n > \log 2 + (m - 1) \log |B|/\delta(1 - \epsilon) \), then

\[
[1 - \epsilon](R_1 - 2\delta, R_2 - 2\delta) \leq \frac{1}{nN} Q^R_{nN}.
\]

which implies that

\[
[1 - \epsilon](R_1 - 2\delta, R_2 - 2\delta) \leq \liminf_{n \to \infty} \frac{1}{nN} Q^R_{nN}.
\]

However, since \( \epsilon \) and \( \delta \) are arbitrarily small, \( (R_1, R_2) \) has to be a limit point of a sequence of points belonging to \( \liminf_{n \to \infty}(1/nN)Q^R_{nN} \), and (37) follows.

**Theorem 6:** The following set is an achievable region for the \( N \)-ring:

\[
\text{closure } \bigcup_{\mu_X, \mu_Y \text{ stationary}} \{(R_1, R_2) : 0 \leq R_1 \leq I(\mu_X; \mu_Z|\mu_Y)
\]

\[
0 \leq R_2 \leq I(\mu_Y; \mu_Z|\mu_X)
\]

\[
R_1 + R_2 \leq I(\mu_X, \mu_Y; \mu_Z)
\]

(46)

where

\[
I(\mu_X, \mu_Y; \mu_Z) = \lim_{n \to \infty} \frac{1}{n} I(X^n, Y^n; Z^n)
\]

(47a)

\[
I(\mu_X; \mu_Z|\mu_Y) = \lim_{n \to \infty} \frac{1}{n} I(X^n; Z^n|Y^n)
\]

(47b)

\[
I(\mu_Y; \mu_Z|\mu_X) = \lim_{n \to \infty} \frac{1}{n} I(Y^n; Z^n|X^n)
\]

(47c)

and \( X^n, Y^n \) are the \( n \)-dimensional distributions induced by the stationary probability measures \( \mu_X, \mu_Y, \mu_Z \) and \( Z^n \) is the output of the frame-synchronous multiple-access channel with memory when the inputs are independent with distributions \( X^n \) and \( Y^n \).

**Proof:** The existence of the limits in (47) is an easy consequence of the stationarity of the inputs, the time-invariance of the channel, and the existence of entropy rate for any discrete stationary process (e.g., [9]). The symbols transmitted by each user will be denoted by

\[
x = \{x_k(i), k = 1, \ldots, N, i = 1, \ldots, n\}
\]

\[
y = \{y_k(i), k = 1, \ldots, N, i = 1, \ldots, n\}
\]

where \( n \) is the codeword length and \( N \) is the number of codewords in the ring. Similarly, the output symbols are labeled by

\[
z = \{z_k(i), k = 1, \ldots, N, i = 1, \ldots, n\}
\]

If the users were frame-synchronous, then \( z_k(i) \) would depend on \( \{x_k(i-j) \text{ or } x_{k-j}(i-j+n) \text{ if } i \leq j\}^{n-1}_{j=0} \) and \( \{y_k(i-j) \text{ or } y_{k-j}(i-j+n) \text{ if } i \leq j\}^{n-1}_{j=0} \). The lack of frame synchronism introduces a relative rotation of
the rings which can be quantified by \( s \in \{0, \ldots, N - 1\} \), the number of codewords shifted, and \( r \in \{0, \ldots, n - 1\} \), the rotation modulo the codeword length. More precisely, the input \( x_k(i) \) is aligned with \( y_k(i) \), defined by

\[
y_k(j) = \begin{cases} j_{(i-1)n + (j-r)}, & 1 \leq j \leq r \\ j_{(i-1)n + (j-r+n)}, & r < j \leq n \end{cases}
\]

or, equivalently,

\[
y_k(i) = \begin{cases} \bar{y}_{k+(i)}, & i + r \leq n \\ \bar{y}_{k+(i+n)}, & i + r > n. \end{cases}
\]

We will now fix an integer \( l \geq 0 \) independent of all other parameters and force the decoder to discard the following output values:

\[ z_k(i), i \in I; I = \{l + m, \ldots\} \cup \{l + m + r, \ldots\} \]

which corresponds to discarding the \( l + m - 1 \) symbols following the beginning of each received codeword. Note that if \( l \) is large enough, \( I = \varnothing \); however, we will eventually be interested only in the asymptotic behavior as \( n \to \infty \), in which case \( I \) is practically identical to \( \{l, \ldots, n\} \). Note further that the relative shift \( r \) is allowed to be any integer \( \{0, \ldots, n - 1\} \), and so the cardinality of each of the two components of \( I \) may grow linearly in \( n \). We may rearrange the codewords of Fig. 2 in the matrix form shown in Fig. 3. In this figure, each codeword of user 1 occupies a single row, whereas each codeword of user 2 occupies two consecutive rows. The blacked-out outputs correspond to the \( l + m - 1 \) symbols following the beginning of each codeword which are discarded by the decoder. In connection with this figure, it is useful to introduce the following notation:

\[
x_k^i = \{x_k(i), i = l + 1, \ldots, r\} \\
x_k^e = \{x_k(i), i = l + r + 1, \ldots, n\} \\
y_k^i = \{y_k(i), i = l + 1, \ldots, r\} \\
y_k^e = \{y_k(i), i = l + r + 1, \ldots, n\} \\
z_k^i = \{z_k(i), i = l + m, \ldots, r\} \\
z_k^e = \{z_k(i), i = l + r + m, \ldots, n\} \\
e = \{z_k(i), k = 1, \ldots, N, i \in I\} = \{(z_k^i, z_k^e), k = 1, \ldots, N\}.
\]

Then, the definition of the multiple-access channel with finite memory (10) implies that

\[
P_{\bar{x} | x}(\bar{z} | x, \bar{y}) = \prod_{k=1}^N \prod_{i=1}^r p_e(z_k(i) | x_k(i-m+1), \ldots, x_k(i)), \bar{y}_k(i)) = \prod_{k=1}^N p_{\bar{x} | x}(z_k^i | x_k^i, \bar{y}_k^i)
\]

which implies that the output subblocks \( \{z_k^i, z_k^e, k = 1, \ldots, N\} \) are conditionally independent given the inputs and only depend on their corresponding input subblocks. Notice also that the inputs with indices \( i \in \{1, \ldots, l\} \cup \{r + 1, \ldots, r + l\} \) do not affect any outputs used by the decoder.

We now proceed to show that for any pair \((\mu_x, \mu_y)\) of stationary \( l \)-dependent measures defined on the infinite sequences drawn from \( A_1 \) and \( A_2 \), respectively, the following pentagon is achievable:

\[
C(\mu_x, \mu_y) = \{(R_1, R_2) : 0 \leq R_1 \leq I(\mu_x; \mu_y) \quad 0 \leq R_2 \leq I(\mu_y; \mu_z) \quad R_1 + R_2 \leq I(\mu_x, \mu_y; \mu_z)\}.
\]

To this end, we must show that for any fixed \((R_1, R_2) \in C(\mu_x, \mu_y), \epsilon > 0, \) and \( \gamma > 0, \) and all sufficiently large \( n \) there exist \((nN, M_1^n, M_2^n, \epsilon)\) juxtaposition codes for the \( N \)-ring such that

\[
\log \frac{M_j}{n} \geq R_i - \gamma, \quad i = 1, 2
\]

i.e., there exist \( N (n, M_i) \) code books for user \( i \) (independent of the relative rotation), and a decoding rule (possibly dependent of the relative rotation) such that the probability that any of the \( N \) messages transmitted by each user are decoded incorrectly is not higher than \( \epsilon \). The codes are chosen as follows.

**Random Coding:** The \( N \) code books of user \( i \) are denoted by \((f_i : \{1, \ldots, M_i\} \to A_1^n)_{i=1}^M \) and are the outcomes of random selection where each codeword in each code book is independently selected with probability

\[
p_{R_i(m)}(a_1, \ldots, a_s) = p_X(a_1, \ldots, a_s) \\
p_{R_2(m)}(b_1, \ldots, b_s) = p_Y(b_1, \ldots, b_s)
\]

where \( X^n \) and \( Y^n \) are the independent \( n \)-dimensional distributions induced by \((\mu_x, \mu_y)\). The overall code book
of user $i$ resulting from the juxtaposition of the foregoing $N$ code books is denoted by $f_i \colon (1, \ldots, M_i)^N \to A_i^N$.

Decoding: The decoder performs simultaneous decoding of the $N$ messages transmitted by both users, upon observing $\bar{z}$ and the rotation $(r,s)$, in the knowledge of the code books $f_1$ and $f_2$. The decoder selects the messages $(m_1, m_2) \in (1, \ldots, M_1)^N \times (1, \ldots, M_2)^N$ if $(m_1, m_2)$ is the unique pair that satisfies

$$(f_1(m_1), f_2(m_2), \bar{z}) \in J(n, \delta)$$

where $0 < \delta < \gamma$, and it outputs a decoding error if there are zero or more than one such pairs. The set $J(n, \delta)$ in (49) is the set of jointly typical sequences according to three criteria:

$$J(n, \delta) = J_1(n, \delta) \cap J_2(n, \delta) \cap J_3(n, \delta)$$

where

$$J_1(n, \delta) = \left\{ (x, \bar{y}, \bar{z}) : \left| \frac{1}{n} i_1(x, \bar{y}, \bar{z}) - \frac{1}{n} I(X; \bar{Z}|\bar{Y}) \right| \leq \delta \right\}$$

(50a)

$$J_2(n, \delta) = \left\{ (x, \bar{y}, \bar{z}) : \left| \frac{1}{n} i_2(x, \bar{y}, \bar{z}) - \frac{1}{n} I(\bar{Y}; \bar{Z}|X) \right| \leq \delta \right\}$$

(50b)

$$J_3(n, \delta) = \left\{ (x, \bar{y}, \bar{z}) : \left| \frac{1}{n} i_3(x, \bar{y}, \bar{z}) - \frac{1}{n} I(X, \bar{Y}; \bar{Z}) \right| \leq \delta \right\}$$

(50c)

where

$$i_1(x, \bar{y}, \bar{z}) = \log \frac{P_{ZX}(\bar{z} | x, \bar{y})}{P_{Z}(\bar{z} | \bar{y})}$$

(51a)

$$i_2(x, \bar{y}, \bar{z}) = \log \frac{P_{ZX}(\bar{z} | x, \bar{y})}{P_{XZ}(\bar{z} | x)}$$

(51b)

$$i_3(x, \bar{y}, \bar{z}) = \log \frac{P_{ZX}(\bar{z} | x, \bar{y})}{P_{Z}(\bar{z})}$$

(51c)

Note that the expected values of the functions in (51) evaluated with the distribution $(X, \bar{Y}, \bar{Z})$ are equal to the mutual informations appearing in (50). We can decompose the functions in (51) taking advantage of the assumed $l$-dependence of the inputs, which implies that the random variables $(X_k, X_k^c, k = 1, \ldots, N)$ (and $(\bar{Y}_k, \bar{Y}_k^c, k = 1, \ldots, N)$) are independent. Then we obtain via (48) that

$$i_1(x, \bar{y}, \bar{z}) = \sum_{k=1}^{N} \log \frac{P_{Z|X}(\bar{z}|x, \bar{y})}{P_{Z|Y}(\bar{z}|\bar{y})} + \frac{P_{Z|X}(\bar{z}|x, \bar{y})}{P_{Z|Y}(\bar{z}|\bar{y})}$$

$$i_2(x, \bar{y}, \bar{z}) = \sum_{k=1}^{N} \log \frac{P_{Z|X}(\bar{z}|x, \bar{y})}{P_{Z|X^c}(\bar{z}|x, \bar{y})} + \frac{P_{Z|X}(\bar{z}|x, \bar{y})}{P_{Z|X^c}(\bar{z}|x, \bar{y})}$$

(52a)

$$i_3(x, \bar{y}, \bar{z}) = \sum_{k=1}^{N} \log \frac{P_{Z|X}(\bar{z}|x, \bar{y})}{P_{Z}(\bar{z})}$$

(52b)

Taking expected values of (52) with respect to $(X, \bar{Y}, \bar{Z})$ and recalling that the inputs are stationary we obtain that

$$I(X; \bar{Z}|\bar{Y}) = N I(X; \bar{Z}|\bar{Y})$$

(53a)

$$I(\bar{Y}; \bar{Z}|X) = N I(\bar{Y}; \bar{Z}|X)$$

(53b)

$$I(X, \bar{Y}; \bar{Z}) = N I(X, \bar{Y}; \bar{Z})$$

(53c)

where $\bar{Z}^n$ and $\bar{Y}^n$ denote $(Z_i(i), i \in I)$ and $(\bar{Y}_i(i), i = 1, \ldots, n)$, respectively. Furthermore, since $(Z_i(i), i \in I)$ depends on the inputs only through $X_i^n$, $X_i^c$, $Y_i^n$, and $\bar{Y}_i^n$, and since $X^n$ is stationary and $l$-dependent, $(\bar{Y}_i^n, \bar{Y}_i^c)$ has the same distribution as $(Y_i^n, Y_i^c)$ and therefore (53) can be written as

$$I(X; \bar{Z}|\bar{Y}) = N I(X^n; \bar{Z}^n|Y^n)$$

(54a)

$$I(\bar{Y}; \bar{Z}|X) = N I(Y^n; \bar{Z}^n|X^n)$$

(54b)

$$I(X, \bar{Y}; \bar{Z}) = N I(X^n, Y^n; \bar{Z}^n)$$

(54c)

The probability that the transmitted messages $(S_1, S_2) = (m_1, m_2)$ are not decoded correctly given that $(f_1, f_2)$ are the chosen code books is

$$e_{m,m}(f_1, f_2) = P \left[ (F_1(m_1), F_2(m_2), \bar{Z}) \in J(n, \delta) \right]$$

where

$$e_{m,m}(f_1, f_2) = P \left[ (F_1(m_1), F_2(m_2), \bar{Z}) \in J(n, \delta) \right]$$

(55)
The first term in the right side of (55) is smaller than $\epsilon/2$ for sufficiently large $n$ because

$$
\lim_{n \to \infty} P\left[(F_i(m_1), F_i(m_2), \bar{Z}) \in J_k(n, \delta)(S_1, S_2) = (m_1, m_2)\right] = 1, \quad k = 1, 2, 3. (56)
$$

This holds because the inputs and output are jointly stationary and ergodic [4] (the output is $(I + m + 1)$-dependent) and therefore the Shannon–McMillan theorem (see e.g., [9]) implies that each of the $2N$ terms in the right sides of (52) converges in probability (when scaled by $n$) to its expected value (which may be zero if, for example, $r$ remains finite as $n \to \infty$). (Notice that this holds even though each output subblock $z_{ij}$ (or $z_{ij}^*$) has $m-1$ fewer elements than the corresponding input subblocks $x_{ij}$ and $y_{ij}$ (or $x_{ij}^*$ and $y_{ij}^*$)), since convergence is not affected by any fixed number of elements.)

To investigate the behavior of the second term on the right side of (55), we will introduce the independent random vectors $U$ and $V$ defined on $A^n$ and $A^n$, respectively, whose distributions are $p_x$ and $p_y$, but which, unlike $X$ and $Y$, are independent of $Z$, i.e.,

$$
\rho_{UVZ}(x, y, z) = p_x(x)p_y(y)p_z(z) \exp(-i(x, y, z)). (57)
$$

If $(x, y, z) \in J_k(n, \delta)$, however, then (50c) implies that

$$
\exp(-i(x, y, z)) \leq \exp(-I(X, \tilde{Y}; Z) + n\delta)
$$

and, consequently,

$$
P\left[(F_i(m_1), F_i(m_2), \bar{Z}) \in J_k(n, \delta)(S_1, S_2) = (m_1, m_2)\right] \leq \exp(-I(X, \tilde{Y}; Z) + n\delta). (58)
$$

Proceeding similarly with the third and fourth terms on the right side of (55) we obtain that, for sufficiently large $n$,

$$
E\left[e_{m,m}(F_i, F_j)\right] \leq M_1^n \exp(-(I(X; \bar{Z}) + n\delta)
$$

$$
+ M_2^n \exp(-(I(\bar{Y}; Z|X) + n\delta)
$$

$$
+ M_1^n M_2^n \exp(-(I(X, \bar{Y}; Z) + n\delta) + \epsilon/2. (59)
$$

Thus if $M_1$ and $M_2$ grow sufficiently slowly with $n$ we will be able to show that for large $n$ the right side of (59) does not exceed $\epsilon$. Specifically, we choose $M_1$ and $M_2$ to satisfy

$$
R_i - \frac{\gamma}{N} < \log M_i \leq R_i - \frac{\delta + \gamma}{2N}, \quad i = 1, 2. (60)
$$

(This choice is possible for all sufficiently large $n$.) Then, (59) is further upper-bounded by

$$
E\left[e_{m,m}(F_i, F_j)\right] \leq \exp\left(nN\left[R_i - \frac{1}{nN}I(X; \bar{Z}|\bar{Y})\right] - n\frac{\gamma - \delta}{2}\right)
$$

$$
+ \exp\left(nN\left[R_i - \frac{1}{nN}I(\bar{Y}; Z|X)\right] - n\frac{\gamma - \delta}{2}\right)
$$

$$
+ \exp\left(nN\left[R_i + R_j - \frac{1}{nN}I(X, \bar{Y}; Z)\right] - n\gamma\right) + \frac{\epsilon}{2}. (61)
$$

Using (54) and recalling that $\delta < \gamma$, it is seen that if

$$
0 \leq R_i \leq \liminf_{n \to \infty} \frac{1}{n} I(X^n; Z^n|Y^n) \quad (62a)
$$

$$
0 \leq R_j \leq \liminf_{n \to \infty} \frac{1}{n} I(Y^n; Z^n|X^n) \quad (62b)
$$

$$
R_i + R_j \leq \liminf_{n \to \infty} \frac{1}{n} I(X^n, Y^n; Z^n), \quad (62c)
$$

then the right side of (61) does not exceed $\epsilon$ for sufficiently large $n$. Therefore, at least one realization of the code books must exist that results in probability of error better than $\epsilon$, and so the pentagon in (62) is achievable. Actually, that region coincides with $C(\mu_x, \mu_y)$ because $Z^n$ was obtained by discarding $2(1+m-1)$ elements from $Z^n = (Z_i(i), i = 1, \cdots, n)$ (which is the output of the frame-synchronous channel with inputs $X^n$ and $Y^n$) and therefore (cf. (18))

$$
I(X^n; Z^n|Y^n) \leq I(X^n; Z^n|Y^n) + 2(I + m - 1) \log |B| \quad (63a)
$$

$$
I(Y^n; Z^n|X^n) \leq I(Y^n; Z^n|X^n) + 2(I + m - 1) \log |B| \quad (63b)
$$

$$
I(X^n, Y^n; Z^n) \leq I(X^n, Y^n; Z^n) + 2(I + m - 1) \log |B| \quad (63c)
$$

It remains to show that the restriction to $l$-dependent input distributions can be dropped without changing the region in (64). We will do so in three steps where we show that the union can be written over 1) $B$-processes, 2) ergodic stationary processes, and finally, 3) stationary processes.
Step 1: The B-processes are an important class of stationary ergodic discrete-time random processes (introduced by Ornstein [16]) that can be defined as the outputs of the time-invariant systems driven by independent identically distributed (i.i.d.) inputs. This is, in effect, a mixing condition requiring that the influence of the sufficiently distant past becomes negligible. It was shown by Ornstein [16] that the closure in the J-metric of the stationary l-dependent processes is equal to the set of B-processes. The J-metric between two stationary ergodic measures μ and μ is equal to the minimum percentage of time samples we need to change a representative realization of μ to make it look like a representative realization of μ. Due to the finite memory of the channel, it is easy to show that if ξ sequence of stationary ergodic input measures yzar converges in the J-metric to μ and μ, then the corresponding output measures also converge in the J-metric, because one way to generate a representative sequence of stationary ergodic measures is by modifying representative strings of μ and μ to get representative strings of μ and μ without changing the output samples unaffected by those modifications. Therefore, d(μ, μ) ≤ m[d(μ, μ) + d(μ, μ)] since each input value affects at most m output values. Now, since the entropy rate is a continuous function of the stationary measure under the J-metric [19] and the three constraints in (64) can be written as

\begin{align}
I(μx; μz | μγ) &= H(μγ, μz) + H(μx, μy, μz) \\
I(μy; μz | μx) &= H(μx, μz) + H(μy, μx, μz) \\
I(μx, μy; μz) &= H(μx) + H(μy) - H(μx, μy, μz)
\end{align}

(65a) (65b) (65c)

where

\begin{align}
G_α(μx, μy) &= \max \limits_{0 ≤ R_1 ≤ 1} \alpha R_1 + (1 - \alpha) R_2 \\
&= \begin{cases} 
(2α - 1)I(μx; μz | μγ) + (1 - α)I(μx, μy; μz), & 1/2 ≤ α ≤ 1 \\
(1 - 2α)I(μy; μz | μx) + αI(μx, μy; μz), & 0 ≤ α ≤ 1/2
\end{cases}
\end{align}

(67)

the region (64) is unchanged if we enlarge the set of stationary l-dependent processes to its closure, the set of B-processes.

Step 2: The stationary mixing multistep Markov processes are B-processes (the mixing condition essentially rules out processes with periodicities), whose closure in the weak topology is the set of stationary ergodic processes [10, p. 360], where we say that μ converges weakly to μ if, for all n > 0, the n-dimensional distribution induced by μ converges to that of μ. Again, due to the finite memory of the channel it is easy to show that if μ converges weakly to μ and μ, then the corresponding output measures also converge weakly, because the n-dimensional distribution induced by μ depends linearly on the n + m - 1-dimensional distributions of μ and μ. Furthermore, since the entropy rate is a lower semi-continuous function in the weak topology, we can write

\begin{align}
\liminf_{k \to \infty} H(μ_z^{(k)}) &= H(μ_z) \\
\liminf_{k \to \infty} H(μ_z^{(k)} | μ_y^{(k)}) &= H(μ_z | μ_y)
\end{align}

(66)

and (cf. [17])

\begin{align}
\lim_{k \to \infty} H(μ_z^{(k)} | μ_y^{(k)}) &= H(μ_z | μ_y)
\end{align}

since the latter expression is linear in the conditioning measures. This implies that the union appearing in the achievable region can indeed be extended to the stationary ergodic measures.

Step 3: This step has a well-known counterpart in the solution of the capacity of single-user channels with memory (cf. [11, sec. III]). There, the ergodic assumption is needed to invoke the Shannon-McMillan theorem in the proof of the direct theorem, whereas the usual converse techniques upper-bound capacity by the minimum of mutual information rates over all stationary inputs. A proof that the lower and upper bounds thus obtained coincide was given by Parthasarathy [17] using the ergodic decomposition theorem. Even though in the multiuser case capacity is not given as the maximization of a scalar function, we can use Parthasarathy’s result by noticing that all we need to show is that for every 0 ≤ α ≤ 1 (cf. [21])

\begin{align}
\sup \limits_{μx, μy \text{ stationary ergodic}} G_α(μx, μy) = \sup \limits_{μx, μy \text{ stationary}} G_α(μx, μy)
\end{align}

(66)

where the second equality holds because the maximization on the left side is attained at one of the two (Pareto) optimum vertices of the feasible pentagon (note that I(μ; μγ, μz) ≤ I(μ; μγ, μz) + I(μγ; μ; μz)). We may now fix α ∈ [0, 1/2], the other case being entirely parallel. We then obtain that for all stationary pairs μ, μγ

\begin{align}
G_α(μx, μy) &= (1 - α)I(μx, μy; μz) - (2α - 1)I(μx; μz | μγ) - (1 - 2α)I(μy; μz | μx) \\
&= (1 - α) \int I(μx, μy; μz) dP_μ(x) dP_μ(y) \\
&- (2α - 1) \int I(μx; μz | μγ) dP_μ(x) \\
&- (1 - 2α) \int I(μy; μz | μx) dP_μ(x) \\
&= \int G_α(μx, μy) dP_μ(x) dP_μ(y)
\end{align}

(68)
where the second theorem follows from Parthasarathy's representation theorem [17], and \( \{\mu_x, x \in \Lambda_1\} \) and \( \{\mu_y, y \in \Lambda_2\} \) are the stationary ergodic measures in the ergodic decompositions of \( \mu_x \) and \( \mu_y \):

\[
\mu_x(E) = \int \mu_x(E) dP_x(x) \tag{69a}
\]

\[
\mu_y(F) = \int \mu_y(F) dP_y(y) \tag{69b}
\]

for all measurable sets \( E \) and \( F \). Notice that the only restriction of Parthasarathy's result is that the channel connecting input and output be stationary, and this is the case for the channel that connects \( \{\mu_x, \mu_y\} \) with \( \mu_z \), as well as the channel seen by the first user, which connects \( \mu_x \) and \( \mu_y \) because both \( \mu_x \) and the multiple-access channel are stationary.

Finally, for a stationary pair \( \mu_x, \mu_y \) to achieve a value of \( G(x,y) \) close to the supremum, there must exist \((x,y) \in \Lambda_1 \times \Lambda_2 \) such that \( G(x,y) \) is close to the supremum, because the average with respect to \( P_x \times P_y \) in (68) cannot be larger than each of its sample values. This fact shows (66) and completes the proof of the theorem.

We will now use Theorems 5 and 6 to find the frame-asynchronous capacity region of the examples we studied at the end of Section II.

**Example 1:** Since the frame-synchronous capacity region is achieved by stationary (i.i.d.) inputs, it remains the same if the users are frame-asynchronous. The same is true for the symbol-asynchronous Gaussian multiple-access channel [21] where the capacity region is achieved by stationary colored Gaussian processes. (In that case, the capacity region does depend, in general, on whether the transmitters are symbol-synchronous.)

**Example 2:** We will show first that the capacity region is achieved by stationary (i.i.d.) inputs, and thus the following region is achievable:

\[
0 \leq R_1 \leq 1/2 \quad \text{and} \quad \max_{\gamma_1, \gamma_2} \gamma_1 \beta_2 + \gamma_2 \beta_1 = 1/2
\]

which, together with (71), implies that the total capacity of the frame-asynchronous channel is bounded from above by

\[
\lim_{n \to \infty} \frac{1}{n} \max_{\text{stationary}} I(X^n, Y^n; Z^n) \leq 1/2,
\]

in contrast to the total capacity of the frame-synchronous channel, which is equal to 1 bit.

To show achievability of the triangle, we choose both input processes to be stationary and Markov with

\[
P[X_i \neq 0 | X_{i-1} = 0] = 1
\]

\[
P[Y_i \neq 0 | Y_{i-1} = 0] = 1
\]

(i.e., \( \gamma_k = 1 - \beta_k \), \( k = 1,2 \)). Then, as there are no consecutive input zeros, the channel is equivalent to the memoryless channel

\[
z_i = \begin{cases} x_i, & \text{if } x_i \neq 0 \text{ and } y_i = 0 \\ y_i, & \text{if } y_i \neq 0 \text{ and } x_i = 0 \\ (1/2,1/2), & \text{otherwise} \end{cases}
\]

Furthermore, we will only consider inputs whose nonzero values are independent and equally likely to be 1 or 2. Then the outputs are independent both unconditionally and conditioned on either input sequence:

\[
H(Z^n_i | X^n, Y^n) = n - m - 1
\]

\[
H(Z^n_i | Y^n) = \sum_{i=m}^{n} H(Z_i | Y^n) = (n - m + 1) \left[ 1 - \gamma_1 + \gamma_2 h_{\gamma_1/2} \right]
\]

where the second equation follows from (72) and the independence of \( X^n \) and \( Y^n \). Then (70), (73), and (74) imply that

\[
I(X^n; Z^n_i | Y^n) = (n - m + 1)
\]

\[
I(Y^n; Z^n_i | X^n) = (n - m + 1)
\]

\[
I(X^n, Y^n; Z^n_i) = (n - m + 1)
\]

However, \( h_{\gamma_k/2} \geq \lambda_k \), and thus the following region is achievable:

\[
\bigcup_{\gamma_k \in \{1/2,1\}} \{(R_1, R_2) : 0 \leq R_1 \leq 1/2, 0 \leq R_2 \leq \gamma_k (1 - \gamma_k) \}
\]
which can be shown to coincide with the triangle \((R_1, R_2): 1 \leq R_1, 0 \leq R_2, R_1 + R_2 \leq 1/2\).

It can be seen that in this example full frame-synchronous capacity would be achieved if the encoders were informed of the relative shift modulo 2, and without this information, they cannot do better than the frame-asynchronous region. This points out that, in contrast to the memoryless channel, even a mild form of asynchronism where the shift may be only 0 or 1 reduces the capacity region. The reason is that in the memoryless channel, a large time-scale type of cooperation (time-sharing) is enough to achieve capacity, whereas in a channel with memory, the encoders may need to cooperate in a small time-scale. Mild frame-asynchronism only precludes cooperation in the small time-scale.

REFERENCES

The Capacity Region of the Symbol-Asynchronous Gaussian Multiple-Access Channel

Sergio Verdú

Reprinted from IEEE TRANSACTIONS ON INFORMATION THEORY Vol. 35, No. 4, July 1989
The Capacity Region of the Symbol-Asynchronous Gaussian Multiple-Access Channel

SERGIO VERDÚ, SENIOR MEMBER, IEEE

Abstract — In the information theory of the multiple-access channel, two types of synchronism are usually assumed among the transmitters, namely, frame and symbol synchronism. Frame synchronism refers to the ability of the users to start the transmission of their codewords in unison. The issue of symbol synchronism arises in continuous-time channels in which each codeword symbol modulates a fixed assigned waveform; the channel is symbol synchronous if the users cooperate so that their symbol epochs coincide at the receiver. In practice, symbol synchronism is harder to achieve, and the only reported progress so far has been in the removal of the assumption of frame synchronism. It is shown that if the transmitters are assigned the same waveform, symbol asynchronism has no effect on the two-user capacity region of the white Gaussian channel which is equal to the Cover-Wyner pentagon, whereas if the assigned waveforms are different (e.g., code division multiple access), the symbol-asynchronous capacity region is no longer a pentagon.

I. INTRODUCTION

THE MAIN GOAL of the information-theoretic study of the multiple-access channel is to find its capacity region, i.e., the set of information rates at which simultaneous reliable communication of the messages of each user is possible. This problem was solved in the pioneering work of Ahlswede [1], [2] on the two-user discrete memoryless channel; later, an explicit expression for the capacity region of the Gaussian memoryless discrete-time multiple-access channel was given by Cover [3] and Wyner [4]. These and most of the subsequent results on the subject assumed so-called frame (or block) synchronism among the users in the sense that the beginnings of the codewords of each user were guaranteed to coincide at the receiver. It has been shown by Polyá and Ráth [5] and, independently, by Hui and Humblet [6] that the only effect of frame synchronism on the discrete memoryless multiple-access channel is the removal of the convex hull operation from the expression for the capacity region. It was recently shown [7] that if the multiple-access channel has memory, frame synchronism may drastically reduce the capacity region and, in particular, the maximum achievable rate sum. At any rate, in many practical situations it is perfectly reasonable to assume that this type of synchronism is achievable with a modicum of channel feedback or cooperation among transmitters.

The type of synchronism that is difficult to achieve in many practical situations (owing to the much smaller time scale involved) is symbol synchronism. This issue arises in continuous-time channels where each codeword symbol modulates a signal waveform of finite duration, as is the case in most conventional digital communication systems. In these systems, user $k$ transmits a codeword $(M_k(1), \cdots, M_k(n)) \in A_k^n$ by sending the signal

$$\sum_{i=1}^{n} s_k(t - iT; b_k(i))$$

where the waveforms $\{s_k(t; b), b \in A_k\}$ vanish outside the interval $[0,T]$ and constitute the fixed signaling alphabet of user $k$, which is known to all transmitters and to the receiver. If the symbol epochs of the signals transmitted by the users are not aligned at the receiver, then the channel is symbol asynchronous (Fig. 1). For a channel with two senders and one receiver, assuming frame synchronism and an additive white Gaussian noise channel model, we can write the channel output as

$$y(t) = \sum_{i=1}^{n} s_1(t - iT - \tau_1; b_1(i))$$

$$+ \sum_{i=1}^{n} s_2(t - iT - \tau_2; b_2(i)) + z(t)$$

where $\tau_1$ and $\tau_2$ are the symbol durations of senders 1 and 2, respectively.
where the delays or offsets $\tau_1 \in [0, T)$, $\tau_2 \in [0, T)$ account for the symbol asynchronism between the users and are known to the receiver (because it acquires the timing of each of the received signals to decode reliably each of the transmitted messages) and unknown to the transmitters.

While the derivation of coding strategies for symbol-asynchronous channels has been addressed before [8], it appears that no results on the capacity region of the multiple-access channel are available when symbol synchronism is not assumed. In this paper we find the capacity region of a fairly general symbol-asynchronous Gaussian multiple-access channel in which user $k$ modulates linearly a fixed signature waveform $s_k(t)$, i.e., $s_k(t; b) = b s_k(t)$. This encompasses many interesting channels in applications, such as direct-sequence spread-spectrum code-division multiple-access channels (CDMA) wherein each transmitter is assigned a distinct signature waveform which is used to modulate information simultaneously and independently of the other transmitters. We focus our attention on energy-limited channels where $A_k = \mathbb{R}$, and each codeword of user $k$ is constrained to satisfy

$$\frac{1}{n} \sum_{i=1}^{n} b_k^2(i) \leq w_k, \quad k=1,2. \quad (1.2)$$

The methods employed in this paper can be used to solve the case where the $A_k$ are finite alphabets; however, in this case, as in the single-user discrete-time Gaussian channel with finite alphabets or amplitude constraints, no explicit expressions for capacity can be obtained.

If the transmitters are assigned identical signature waveforms and are symbol synchronous, i.e., $\tau_1 = \tau_2$, then it is easy to see that the channel is equivalent to the standard one-dimensional discrete-time Gaussian multiple-access channel, and therefore, its capacity region is given by the Cover-Wyner pentagon: each individual rate is constrained not to exceed single-user capacity and the sum of the rates cannot exceed the capacity of a single-user channel whose signal-to-noise ratio is the sum of the signal-to-noise ratios of both users. In this paper it is shown that the same result holds even if the users are not symbol synchronous. However, that is no longer true when the transmitters are assigned different signature waveforms. Then the symbol-asynchronous capacity region is no longer a pentagon and depends not only on the respective signal-to-noise ratios, but also on the similarity between the signature waveforms quantified by their cross correlations. In some applications it may be of interest to use those results to find the sought-after capacity region of the compound channel. Finally, in Section IV we consider an alternative representation of the capacity region which results in a particularly compact characterization of the fundamental limits of the multiple-access channel in the region of high signal-to-noise ratios.

II. CHANNEL MODEL

The goal of this section is to obtain a channel with discrete-time outputs whose capacity is the same as that of the channel with continuous-time output

$$y(t) = \sum_{i=1}^{n} b_i(i) s_1(t-iT-\tau_1) + \sum_{i=1}^{n} b_i(i) s_2(t-iT-\tau_2) + n(t) \quad (2.1)$$

where $n(t)$ is white Gaussian noise with power spectral density equal to $\sigma^2$. This goal is achieved by considering the projection of the observation process ($y(t)$) along the direction of the unit energy signals ($s_1(i)$) and ($s_2(i)$) and their $T$-shifts:

$$y_k(t) = \int_{T+\tau_k}^{T} y(t) s_k(t-iT-\tau_k) dt, \quad k=1,2. \quad (2.2)$$

1 Most capacity analyses of the CDMA channel have focused on single-user receivers and approximated the multiple-access interference by a white Gaussian process [9]-[11], thereby providing limited insight into the fundamental limits of that channel.
THE CAPACITY REGION OF THE SYMBOL-ASYNCHRONOUS GAUSSIAN MULTIPLE-ACCESS CHANNEL

VEXDÜ: \( v(t) \) respectively. The key observation is that this operation does not destroy any information that is valuable in deciding which messages were transmitted. This is because the likelihood function (i.e., the conditional expectation of the Radon-Nikodym density) at the receiver

\[
\rho_{12} = \int_0^T s_1(t) s_2(t) (t + \tau_1 - \tau_2) \, dt 
\]
\[
\rho_{21} = \int_0^T s_1(t) s_2(t) (t + T + \tau_1 - \tau_2) \, dt. 
\]

It follows easily that

\[
\begin{bmatrix} y_1(i) \\ y_2(i) \end{bmatrix} = \begin{bmatrix} 0 & \rho_{21} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} b_1(i-1) \\ b_2(i-1) \end{bmatrix} + \begin{bmatrix} 1 & \rho_{12} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} b_1(i) \\ b_2(i) \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ \rho_{21} & 0 \end{bmatrix} \begin{bmatrix} b_1(i+1) \\ b_2(i+1) \end{bmatrix} + \begin{bmatrix} \rho_1(1) \\ \rho_2(1) \end{bmatrix} + \begin{bmatrix} \rho_1(n) \\ \rho_2(n) \end{bmatrix} 
\]

for \( 1 \leq i \leq n \) (with \( b_1(0) = b_2(n+1) = 0, k = 1, 2 \)); the discrete-time random process \( \{s(i), n_2(i)\}^T \) is Gaussian with zero mean and covariance matrix:

\[
E \left[ \begin{bmatrix} n_1(i) \\ n_2(i) \end{bmatrix} \begin{bmatrix} n_1(j) \\ n_2(j) \end{bmatrix} \right] = \sigma^2 H(i-j) 
\]

where \( H(i) = 0 \) if \(|i| > 1\), and \( H(1), H(0), \) and \( H(-1) \) are the matrices appearing in (2.4), i.e.,

\[
I(0) = \begin{bmatrix} 1 \\ \rho_{12} \end{bmatrix}, \quad H(1) = H(-1) = \begin{bmatrix} \rho_{12} \\ 0 \end{bmatrix}, 
\]

since the receiver knows the assigned waveforms \( \{s_1(t)\} \) and \( \{s_2(t)\} \) as well as the symbol epochs \( \{R_T + \tau_1\} \) and \( \{R_T + \tau_2\} \), it can compute \( \{y_1(i)\}_{i=1}^{\infty} \) and \( \{y_2(i)\}_{i=1}^{\infty} \) by passing the observations through two matched filters for signals \( \{s_1(t)\} \) and \( \{s_2(t)\} \), respectively. The key observation is that this operation does not destroy any information that is valuable in deciding which messages were transmitted. This is because the likelihood function (i.e., the conditional expectation of the Radon-Nikodym derivative between the measure induced by the observations and Wiener measure given that \( \{b_1(i)\}_{i=1}^{\infty} \) and \( \{b_2(i)\}_{i=1}^{\infty} \) are the transmitted codewords) is equal to a constant times (e.g., [12])

\[
\mathbb{E} \left[ \sum_{k=1}^{n} \sum_{i=1}^{n} b_k(i) s_k(t-iT-\tau_k) \right]^2 dt 
\]

which can be factored into

\[
h_1(\{y(i)\}) h_2(\{b_i(i)\}, \{y_1(i)\}, \{y_2(i)\}) \]

hence because of the factorization theorem [13], \( \{y_1(i)\}_{i=1}^{\infty} \) and \( \{y_2(i)\}_{i=1}^{\infty} \) are sufficient statistics for the transmitted messages. This implies that the channel output \( \{y(t)\} \) enters in the computation of the posterior probability of each message only through \( \{y_1(i)\}_{i=1}^{\infty} \) and \( \{y_2(i)\}_{i=1}^{\infty} \). Thus, no matter which codebooks are chosen by the transmitters, the probability that the maximum \textit{a posteriori} decoder selects the true transmitted message remains the same if instead of working with the original continuous-time observations \( \{y(t)\} \) the decoder is constrained to work with the discrete-time sequences \( \{y_1(i)\}_{i=1}^{\infty} \) and \( \{y_2(i)\}_{i=1}^{\infty} \). Therefore, if a rate pair is \( \epsilon \)-achievable for the multiple-access channel (2.1), it is also \( \epsilon \)-achievable for the multiple-access channel (2.4), and hence the capacity regions of both channels coincide. In this respect, notice for future use that if \( t(W) \) is a sufficient statistic for \( Z \), then the data-processing inequality is satisfied with equality because \( Z \) and \( W \) are conditionally independent given \( t(W) \). Therefore,

\[
I(Z; t(W)) = I(Z; W, t(W)) - I(Z; W|t(W)) 
\]
\[
= I(Z; W) + I(Z; t(W)|W) 
\]
\[
= I(Z; W). 
\]

Note that even though channel (2.4) has two output sequences, it is a multiple-access channel rather than an interference channel because both outputs are available to the multiuser receiver. Channel (2.4) is characterized by the cross correlations \( \rho_{12} \) and \( \rho_{21} \), which depend on the relative offset \( \tau_2 - \tau_1 \) and, therefore, in general, are unknown to the transmitters. Consequently, it is necessary to analyze a \textit{compound} multiple-access channel where the transmitters only know that \( (\rho_{12}, \rho_{21}) \) belongs to an uncertainty set determined by \( (s_1(t)) \) and \( (s_2(t)) \).

The main characteristic of the discrete-time multiple-access channel in (2.4) is that it has memory because the noise sequence is correlated and each output value depends on three input symbols, while each of these symbols affects two consecutive output vectors (cf. Fig. 2). It is possible to obtain an equivalent multiple-access channel (Appendix I) whose noise process is independent at the expense of an enlarged set of observables. The advantage of the latter discrete-time model is that it is possible to invoke coding theorems for channels where the outputs are conditionally independent given the inputs [7].

If either \( \rho_{12} = 0 \) or \( \rho_{21} = 0 \), then the channel becomes memoryless because in that case the users are in effect symbol synchronous. For example, if the users are assigned the same signal and the channel is symbol synchronous, both outputs in (2.4) coincide and are equal to

\[
y(i) = b_1(i) + b_2(i) + n(i) 
\]

where \( (n(i)) \) is a Gaussian independent sequence. Then the channel is the conventional scalar discrete-time Gaussian channel, whose capacity region, subject to the energy
constraints in (1.2), is the Cover-Wyner region:

\[ r = \log \left[ \frac{1}{2} \log \left( 1 + \frac{w_1}{\sigma^2} \right) \right] \]

\[ 0 \leq R_1 \leq \frac{1}{2} \log \left( 1 + \frac{w_1}{\sigma^2} \right) \]

\[ R_1 + R_2 \leq \frac{1}{2} \log \left( 1 + \frac{w_1 + w_2}{\sigma^2} \right) \]  

(2.7)

If the assigned signals are not equal but the users remain symbol synchronous, then (2.4) reduces to the memoryless multiple-access channel

\[ \left[ \begin{array}{c} y_1(i) \\ y_2(i) \end{array} \right] = \left[ \begin{array}{cc} 1 & \rho \\ \rho & 1 \end{array} \right] \left[ \begin{array}{c} b_1(i) \\ b_2(i) \end{array} \right] + \left[ \begin{array}{c} n_1(i) \\ n_2(i) \end{array} \right] \]  

(2.8)

where \( [(n_1(i)n_2(i))^T] \) is an independent Gaussian process with \( E[n_i^2(i)] = \sigma^2 \) and \( E[n_i(i)n_2(i)] = \rho \sigma^2 \), and

\[ \rho = \int_0^T s_1(t)s_2(t) \, dt. \]

In this case, the Cover-Wyner region can be easily generalized (Section III) thanks to the lack of memory when the users are symbol synchronous.

### III. Capacity Region

Before we obtain the capacity region of the symbol-asynchronous Gaussian multiple-access channel, we will generalize the Cover-Wyner region (2.7) to the symbol-synchronous channel where both users are not necessarily assigned the same waveform. To this end, according to (2.8) we need to find the convex closure over independent random variables \( X_1 \) and \( X_2 \) such that \( E[X_1^2] \leq w_1 \) and \( E[X_2^2] \leq w_2 \) of the union of the pentagons \( 0 \leq R_1 \leq I(X_1; Y|X_2), 0 \leq R_2 \leq I(X_2; Y|X_1), R_1 + R_2 \leq I(X_1, X_2; Y) \), with the output \( Y \) given by

\[ Y = \left[ \begin{array}{c} X_1 \\ X_2 \end{array} \right] + \left[ \begin{array}{c} N_1 \\ N_2 \end{array} \right] \]  

(3.1)

where \( N_1 \) and \( N_2 \) are jointly Gaussian with zero mean, \( E[N_1^2] = \sigma_1^2 \), and \( E[N_1N_2] = \rho \sigma_1^2 \). The case \( |\rho| = 1 \) results in the region (2.7); we will therefore assume \( |\rho| < 1 \). Since \( X_1 \) and \( X_2 \) are independent random variables, the covariance of \( Y \) is equal to

\[ \text{cov}(Y) = \left[ \begin{array}{cc} 1 & \rho \\ \rho & 1 \end{array} \right] \sigma_1^2 I_2 + \left[ \begin{array}{cc} \text{var}(X_1) & 0 \\ 0 & \text{var}(X_2) \end{array} \right] \left[ \begin{array}{cc} 1 & \rho \\ \rho & 1 \end{array} \right]. \]  

(3.2)

\( I \), denotes the \( n \times n \) identity matrix.

and we can upper-bound the mutual informations

\[ I(X_1, X_2; Y) \leq \frac{1}{2} \log \left( 1 + \frac{w_1}{\sigma^2} \right) \]  

\[ = \frac{1}{2} \log \left( 1 + \frac{|\rho|}{\sigma^2} \left( \text{var}(X_1) + \text{var}(X_2) \right) \right) \]  

(3.3)

\[ I(X_1; Y|X_2) \leq \frac{1}{2} \log \left( 1 + \frac{w_1}{\sigma^2} \right) \]  

\[ = \frac{1}{2} \log \left( 1 + \frac{w_1}{\sigma^2} \right) \]  

(3.4)

\[ I(X_2; Y|X_1) \leq \frac{1}{2} \log \left( 1 + \frac{w_2}{\sigma^2} \right) \]  

(3.5)

\[ I(X_1, X_2; Y) \leq \frac{1}{2} \log \left( 1 + \frac{w_1}{\sigma^2} \right) \]  

\[ = \frac{1}{2} \log \left( 1 + \frac{w_1}{\sigma^2} \right) \]  

(3.6)

which differs from (2.7) in that the maximum rate sum is no longer the capacity of a single-user channel whose signal-to-noise ratio is equal to \( (w_1 + w_2)/\sigma^2 \). Notice that when \( s_1(t) \) and \( s_2(t) \) are orthogonal \( (\rho = 0) \), then, effectively, both users transmit in separate noninterfering channels and can send information at single-user rates. The \( K \)-user capacity region of the symbol-synchronous Gaussian channel can be found in [14].

Before we state and prove the formula for the capacity of the symbol-asynchronous Gaussian multiple-access channel, we will motivate the expression of the capacity region by finding the mutual information rates in channel (2.4) when the inputs are stationary Gaussian processes with power spectral densities \( S_1(\omega), \omega \in [-\pi, \pi] \) and \( S_2(\omega), \omega \in [-\pi, \pi] \). Channel (2.4) is a two-input two-
output dynamic linear time-invariant system whose outputs are embedded in colored stationary Gaussian noise. If the inputs are stationary Gaussian processes, then the mutual information rates can be written as the difference between the differential entropy rates of the output with and without each of the input processes. Consequently, all that's needed is an expression for the differential entropy rate of a stationary vector Gaussian discrete-time process. In the scalar case the differential entropy rate of a Gaussian process whose power spectral density is $S(\omega)$ is equal to [15, p. 542]

$$L(S) = \exp \frac{1}{2\pi} \int_{-\pi}^{\pi} \log S(\omega) \, d\omega. \quad (3.8)$$

This follows because the differential entropy of a Gaussian vector with covariance matrix $\Sigma$, is $n/2 \log(2\pi e \det \Sigma^{1/2})$ and according to the Toeplitz distribution theorem [16], $\lim_{n \to \infty} \det \Sigma^{1/2}$ coincides with the geometric mean of the Fourier transform of the covariance sequence. What we need for our purposes is a generalization of this result to vector random processes, i.e., we need to find $\lim_{n \to \infty} \det \Sigma^{1/2}$ when $\Sigma$ is an $n$-block Toeplitz matrix whose elements are 2x2 covariance matrices $R(i-j) \otimes \Sigma_k(i,j)$. A solution to this problem can be found in [17] where it is shown that if the power spectral density matrix $M(\omega) = \Sigma_n(\omega) e^{-i\omega n} R(n)$ is continuous and positive definite in $[-\pi, \pi]^2$ then the foregoing limit is equal to the geometric mean of the determinant of $M(\omega)$.

Now the output of channel (2.4) is a zero-mean vector Gaussian process with power spectral density matrix given by

$$I(\omega) = \begin{bmatrix} S_1(\omega) & 0 \\ 0 & S_2(\omega) \end{bmatrix} T(\omega) + \sigma^2 T(\omega),$$

$$\sigma^2 = \begin{bmatrix} \rho_{12} + \rho_{21} e^{-2\omega} \\ \rho_{12} + \rho_{21} e^{-\omega} \end{bmatrix}.$$

Therefore, the mutual information rate between the output and the inputs is equal to

$$\lim_{n \to \infty} \frac{1}{n} I(X_1^n, X_2^n; Y^n) = h(\Xi) - h(\sigma^2 T) \quad \text{where}$$

$$\rho(\omega) = |\rho_{12} + \rho_{21} e^{i\omega}|^2 = \rho_{12}^2 + \rho_{21}^2 + 2 \rho_{12} \rho_{21} \cos \omega. \quad (3.10)$$

Similarly, setting $S_2(\omega) = 0$ and $S_1(\omega) = 0$, respectively, in (3.9) we get

$$\lim_{n \to \infty} \frac{1}{n} I(X_1^n; Y^n | X_1^n) = \frac{1}{4\pi} \int_{-\pi}^{\pi} \log \left[ 1 + \frac{S_l(\omega)}{\sigma^2} \right] d\omega. \quad (3.11)$$

$$\lim_{n \to \infty} \frac{1}{n} I(X_1^n; Y^n | X_1^n) = \frac{1}{4\pi} \int_{-\pi}^{\pi} \log \left[ 1 + \frac{S_2(\omega)}{\sigma^2} \right] d\omega. \quad (3.12)$$

As mentioned in Section I, we will find the capacity region of the asynchronous channel first in the case where the transmitters know the offset, and hence the cross-correlations between their signals, and then in the case where they do not.

**Theorem 1:** The capacity region of the energy-constrained asynchronous Gaussian multiple-access channel when the transmitters know their mutual offset is given by

$$C = \text{closure} \left( \liminf_{n \to \infty} \frac{1}{n} C_n \right)$$

where $C_n$ is the following achievable region for the $n$-block memoryless multiple-access channel whose input symbols correspond to $n$ consecutive channel uses of the original channel.

Proof: It is shown in [7, Theorem 3] that the capacity region of the frame-synchronous discrete-time multiple-access channel with finite memory where the outputs depend on several consecutive input symbols and are conditionally independent given the inputs is equal to

$$C = \text{closure} \left( \liminf_{n \to \infty} \frac{1}{n} C_n \right)$$

where $C_n$ is the following achievable region for the $n$-block memoryless multiple-access channel whose input symbols correspond to $n$ consecutive channel uses of the original channel.
channel with memory:

\[ C_\text{m} = \bigcup_{X_1^7, X_2^7} \{ (R_1^7, R_2^7) : 0 \leq R_1^7 \leq I(X_1^7; Y^n | X_2^7) \}
\]

\[ 0 \leq R_2^7 \leq I(X_2^7; Y^n | X_1^7), \]

\[ R_1^7 + R_2^7 \leq I(X_1^7, X_2^7; Y^n) \} \]  \quad (3.15)

where the union is over the independent random variables \( X_1^7 \) and \( X_2^7 \) satisfying in the present case \( E[X_k^7 X^7_k] \leq \omega_k, \quad k = 1, 2 \).

The aforementioned class of multiple-access channels with finite memory includes as a special case the discrete-time channel in (1.3) whose noise sequence is independent. It does not encompass channel (2.4) directly because the noise sequence therein is dependent. However, since the observables of (2.4) are sufficient statistics for the inputs and are deterministic transformations of the redundant observables in (1.3), not only their capacities coincide but, according to (2.5), the mutual informations arising in the achievable regions, \( C_\text{m} \), of the respective induced \( n \)-block memoryless channels are also equal. Therefore, it is enough to show that the closed set in (3.13) is equal to

\[ \lim_{n \to \infty} (1/n) C_\text{m} \]  \quad (3.16)

As in (3.3), (3.4), and (3.5) we can upper-bound the mutual informations by

\[ I(X_1^7, X_2^7; Y^n) \leq \frac{1}{2} \log \det \left[ \sigma^2 \mathbf{R} \right] \]  \quad (3.19a)

\[ I(X_1^7; Y^n | X_2^7) \leq \frac{1}{2} \log \det \left[ \mathbf{I} + \sigma^{-2} \mathbf{S}_1 \right] \]  \quad (3.20a)

\[ I(X_2^7; Y^n | X_1^7) \leq \frac{1}{2} \log \det \left[ \mathbf{I} + \sigma^{-2} \mathbf{S}_2 \right] \]  \quad (3.20b)

and

\[ \det \left[ \mathbf{I}_n + \sigma^{-2} \mathbf{S} \right] \]  \quad (3.21)

where \( \mathbf{S}_k = \text{cov}(X_k^n), \quad k = 1, 2 \), and equality holds in (3.19) and (3.20) if \( X_1^7 \) and \( X_2^7 \) are Gaussian. The following identity whose proof is in Appendix II gives an explicit expression of (3.19) as a function of \( \Sigma_1, \Sigma_2, \rho_{12}, \) and \( \rho_{21} \).

**Lemma 1:** The following identity holds

\[ \det \left[ \mathbf{I}_n + \sigma^{-2} \mathbf{S} \right] \]  \quad (3.22)
Therefore, (3.15) reduces to the following union over all trace-constrained nonnegative-definite $n \times n$ matrices:

$$C_n = \bigcup_{r_1 \geq 0} \left\{ (r_1, r_2), 0 \leq r_1 \leq \frac{1}{2} \log \det \left[ I_n + \sigma^{-2} \Sigma_1 \right] \right\} \bigcup_{r_2 \geq 0} \left\{ (r_1, r_2), 0 \leq r_2 \leq \frac{1}{2} \log \det \left[ I_n + \sigma^{-2} \Sigma_2 \right] \right\} \bigcup_{k=1}^{n} \left\{ (r_1, r_2), 0 \leq r_1 \leq \frac{1}{2} \log \det \left[ I_n + \sigma^{-2} \Sigma_k \right] \right\}.$$  

Region $C_n$ is a convex set because each of the three rate constraints in (3.21) is a concave function of $(\Sigma_1, \Sigma_2)$. This is a consequence of the fact that $\log \det (A)$ is concave in $A$ if $A$ is a positive-definite matrix [18, p. 125]. Note that even though the determinant appearing in the rate-sum constraint is not that of a symmetric matrix, it is equal to the determinant of the positive-definite matrix

$$I_{2n} + \frac{1}{\sigma^2} \left[ \begin{array}{cc} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{array} \right] Q^T Q,$$

where $Q^T Q = \left[ I_n \quad S^T \right]$. There is no covariance pair $(\Sigma_1, \Sigma_2)$ that maximizes all three rate constraints in (3.23) simultaneously. This is in contrast to the symbol-synchronous channel where we saw that the mutual informations in (3.3)–(3.5) are maximized simultaneously by a pair of input distributions, thus resulting in a capacity region which is equal to a pentagon. Nevertheless, we will be able to show that there is a set of optimum eigenvectors for each user in the sense that it is enough to take the union in (3.23) only over the subset of covariance matrices having those eigenvectors, thereby effectively reducing the union to one over diagonal matrices.

To prove this, the first step is to apply the singular-value decomposition theorem to the matrix $S$ defined in (3.22). According to this result [19, p. 192], we can write

$$S = UDV^* \quad (3.24)$$

where $U$ and $V$ are orthogonal matrices (of the eigenvectors of $SS^T$ and $S^T S$, respectively) and $D$ is a diagonal matrix of the singular values $\{d_k\}_{k=1}^n$ of $S$, i.e., the nonnegative square roots of the eigenvalues of the nonnegative-definite Jacobi matrix

$$S^T S = \left[ \begin{array}{cccc} \rho_1^2 & \rho_{12} \rho_{21} & \rho_{13} \rho_{21} & \rho_{12} \rho_{31} \\ \rho_{12} \rho_{21} & \rho_1^2 + \rho_{21}^2 & \rho_{12} \rho_{21} & \rho_{12} \rho_{31} \\ \rho_{13} \rho_{21} & \rho_{12} \rho_{21} & \rho_2^2 + \rho_{21}^2 & \rho_{12} \rho_{21} \\ \rho_{12} \rho_{31} & \rho_{12} \rho_{21} & \rho_{12} \rho_{21} & \rho_3^2 + \rho_{21}^2 \end{array} \right].$$  

(3.25)

Now, using the orthogonality of $U$ and $V$, we can express the determinant in the rate-sum constraint in (3.23) as

$$\det \left[ \begin{array}{cc} V^* & 0 \\ 0 & U^* \end{array} \right] \det \left[ I_{2n} + \frac{1}{\sigma^2} \left[ \begin{array}{cc} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{array} \right] \right] \det \left[ \begin{array}{cc} I_n & S^T \\ S & I_n \end{array} \right] \det \left[ \begin{array}{cc} V^* \Sigma V & V^* \Sigma U \\ U^* \Sigma V & U^* \Sigma U \end{array} \right]$$

$$= \det \left[ I_{2n} + \frac{1}{\sigma^2} \left[ \begin{array}{cc} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{array} \right] \right] \det \left[ \begin{array}{cc} I_n & D \\ D & I_n \end{array} \right]$$

(3.26)

where we have set $\Lambda_1 = V^* \Sigma V$ and $\Lambda_2 = U^* \Sigma U$. Since $tr(\Lambda_k) = tr(\Sigma_k), \Sigma_k \geq 0$ if and only if $\Lambda_k \geq 0$, and

$$\det \left[ I_n + \sigma^{-2} \Lambda_k \right] = \det \left[ I_n + \sigma^{-2} \Sigma_k \right],$$

(3.27)

the region in (3.23) is equal to

$$C_n = \bigcup_{\Lambda_1 \geq 0 \atop \Lambda_2 \geq 0 \atop \Sigma_k \geq 0 \atop k=1,2} \left\{ (r_1, r_2), 0 \leq r_1 \leq \frac{1}{2} \log \det \left[ I_n + \sigma^{-2} \Lambda_1 \right] \right\} \bigcup_{\Lambda_2 \geq 0 \atop \Sigma_k \geq 0 \atop k=1,2} \left\{ (r_1, r_2), 0 \leq r_2 \leq \frac{1}{2} \log \det \left[ I_n + \sigma^{-2} \Lambda_2 \right] \right\}$$

(3.28)

$$\cdot R_1 + R_2 \leq \frac{1}{2} \log \det \left[ I_{2n} + \frac{1}{\sigma^2} \left[ \begin{array}{cc} \Lambda_1 & 0 \\ 0 & \Lambda_2 \end{array} \right] \right]$$

(3.29)
Thus in effect the singular-value decomposition theorem has allowed us to substitute the matrix $S$ in (3.23) by the diagonal matrix $D$. This is advantageous because the set in (3.28) is actually the capacity region of the two-user Gaussian memoryless channel shown in Fig. 4. This channel differs from the one in Fig. 3 in that the inputs corresponding to different coordinates do not interfere, and the noise covariance matrix is

$$
(3.29)
$$

Therefore, the singular-value decomposition of $S$ effectively decouples the original channel in (3.16) into independent $2 \times 2$ multiple-access subchannels. The capacity region of this decoupled channel is achieved by input distributions whose coordinates are independent. To prove this, we will show that the rate constraints in (3.28) are maximized by diagonal matrices $A_1$ and $A_2$, and therefore, the matrices of optimum eigenvectors for $\Sigma_1$ and $\Sigma_2$ are $V$ and $U$, respectively. First we apply the Hadamard inequality (the determinant of a nonnegative-definite matrix is upper-bounded by the product of its diagonal elements) to the individual rate constraints in (3.28):

$$
1/2 \log \det \left[ I_n + \sigma^2 \Lambda_k \right] \leq \frac{1}{2} \sum_{i=1}^{n} \log \left( 1 + \frac{\lambda_i}{\sigma^2} \right), \quad k=1,2
$$

(3.30)

where $\lambda_i$ is the $i$th diagonal entry of $\Lambda_k$, and equality holds in (3.30) when $\Lambda_k$ is diagonal. Second, to upper-bound the rate-sum constraint in (3.28) in terms of the diagonal elements of $\Lambda_1$ and $\Lambda_2$, we will invoke the following result proved in Appendix III.

**Lemma 2.** Let $A$ and $B$ be $n \times n$ nonnegative-definite matrices, and let $\Delta = \text{diag} (\delta_1, \ldots, \delta_n)$, where $\delta_i$ is a complex scalar such that $|\delta_i| \leq 1$ for $i=1, \ldots, n$. Then

$$
\det \left[ I_n + \frac{A}{A^*} \right] \leq \prod_{i=1}^{n} \left( 1 + \delta_i \right)
$$

(3.31)

with equality if $A$ and $B$ are diagonal.

We apply Lemma 2 to the case $A = \sigma^{-2} \Lambda_1$, $B = \sigma^{-2} \Lambda_2$ and $\Delta = D$, where the singular values of $S$ (i.e., the diagonal elements of $D$) are real numbers belonging to the interval $(0,1)$ since $R$ is positive-definite. (See [19, p. 382], and (II.7).) It then follows from (3.30) and Lemma 2 that the three constraints in (3.28) are maximized by diagonal matrices; hence we can now write

$$
\frac{1}{n} C_n = \bigcup_{\lambda_i \geq 0, i=1, \ldots, n} \left\{ \frac{1}{n} \sum_{i=1}^{n} \log \left( 1 + \frac{\lambda_i}{\sigma^2} \right) \right\}
$$

for some $0 \leq \alpha \leq 1$ (see [20]).

For each spectral pair, denote the rate-sum constraint in (3.13) by

$$
F(S_1, S_2) = \frac{1}{4\pi} \int_{-\pi}^{\pi} \log \left( 1 + \frac{S_1(\omega)}{\sigma^2} + \frac{S_2(\omega)}{\sigma^2} \right) d\omega.
$$

(3.34)

Notice that the individual rate constraints in (3.13) are $F(S_1,0)$ and $F(0,S_2)$, respectively. Furthermore, to simplify the notation, the $L_1([-\pi,\pi])$ subset of power spectra satisfying (1/2$\pi$) $\int_{-\pi}^{\pi} S(\omega) d\omega \leq w$. will be denoted by

$$
\text{C}(w).
$$

(3.35)

It remains to show that the limit as $n \to \infty$ of the set (3.32) is equal to (3.13). The approach we follow is to show that the Pareto-optimal rate pairs of (3.13) coincide with those of $\lim_{n \to \infty} (1/n) C_n$.

The integrand in the rate-sum constraint of the region $C$ in (3.13) is equal to (cf.3.9)

$$
\log \det \left[ I_n + \frac{S_1(\omega)}{\sigma^2} \frac{S_2(\omega)}{\sigma^2} \right] \leq \frac{1}{\alpha} \cdot \frac{\lambda_1}{\sigma^2} \left( 1 - \rho^2(\omega) \right)
$$

(3.32)

with equality if $A$ and $B$ are diagonal.
Then for every $0 \leq \alpha \leq 1$, (3.33) is equal to

$$\max_{S_1 \in \mathcal{P}(\mathcal{W})} \max_{S_2 \in \mathcal{P}(\mathcal{W})} \alpha R_1 + (1-\alpha) R_2$$

$$= \max_{S_1 \in \mathcal{P}(\mathcal{W})} \max_{S_2 \in \mathcal{P}(\mathcal{W})} \left\{ \max \left\{ \frac{1}{2} \sum_{i=1}^{n} (2\alpha - 1) \log \left( 1 + \frac{\lambda_{S_1 S_2}}{\sigma^2} \right) + (1-\alpha) \log \left( 1 + \frac{\lambda_{S_2}}{\sigma^2} - \frac{\lambda_{S_1 S_2}}{\sigma^2} \right) \right\}, \quad \text{if } \frac{1}{2} \leq \alpha \leq 1 \right. \right.$$ 

$$= \max_{S_1 \in \mathcal{P}(\mathcal{W})} \max_{S_2 \in \mathcal{P}(\mathcal{W})} \left\{ \max \left\{ \frac{1}{2} \sum_{i=1}^{n} (1-2\alpha) \log \left( 1 + \frac{\lambda_{S_2}}{\sigma^2} \right) + \alpha \log \left( 1 + \frac{\lambda_{S_1 S_2}}{\sigma^2} + \frac{\lambda_{S_2}}{\sigma^2} \right) \right\}, \quad \text{if } 0 \leq \alpha \leq \frac{1}{2} \right. \right.$$ 

(3.35)

where (3.35) follows from

$$F(S_1, S_2) \leq F(S_1, 0) + F(0, S_2). \quad (3.36)$$

Following the same approach with the convex set (3.32), we obtain that every Pareto-optimal pair in $(1/n)C_1$ attains

$$\max_{(R_1, R_2) \in \mathcal{C}_1} \alpha R_1 + (1-\alpha) R_2$$

for some $0 \leq \alpha \leq 1$. To show that, for every $0 \leq \alpha \leq 1$, the limit as $n \to \infty$ of the right side of (3.37) coincides with (3.35), we will fix $1/2 \leq \alpha \leq 1$ (the proof for $0 \leq \alpha < 1/2$ is identical) and we will prove that

$$\max_{\phi_1, \phi_2 \in \mathcal{P}(\mathcal{N})} \frac{1}{4\pi} \int_{-\infty}^{\infty} f(\Phi_1(\omega), \Phi_2(\omega), \rho^2(\omega)) d\omega$$

$$= \lim_{n \to \infty} \frac{1}{4\pi} \int_{-\infty}^{\infty} f(\Phi_1(\omega), \Phi_2(\omega), \rho^2(\omega)) d\omega$$

(3.38)

where

$$(z_1, z_2, \rho^2) = (2\alpha - 1) \log(1 + z_1)$$

$$\left(1 - \alpha\right) \log\left(1 + z_1 + z_2 + z_1z_2(1 - \rho^2)\right)$$

(3.39)

Lemma 3: If $1/2 \leq \alpha < 1$, then

$$\max_{\phi_1, \phi_2 \in \mathcal{P}(\mathcal{N})} \frac{1}{4\pi} \int_{-\infty}^{\infty} f(\Phi_1(\omega), \Phi_2(\omega), \rho^2(\omega)) d\omega$$

$$= \lim_{n \to \infty} \frac{1}{4\pi} \int_{-\infty}^{\infty} g(\rho^2(\omega), \theta_1, \theta_2) d\omega$$

(3.40)

where $\theta_1, \theta_2$ are positive scalars such that

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} \gamma_k(\rho^2(\omega), \theta_1, \theta_2) d\omega = \frac{w_k}{\sigma^2}, \quad k = 1, 2 \quad (3.41)$$

and $\gamma_k(\cdot, \theta_1, \theta_2), k = 1, 2$ are continuous functions.

Proceeding as in the proof of Lemma 3, it follows that the same result holds for the finite-dimensional optimiza-
tion problem in the right side of (3.38), i.e.,

$$\max_{\phi_m, \phi_n \geq 0, m=1, \ldots, n} \frac{1}{2n} \sum_{i=1}^{n} f(\phi_m, \phi_n, d_i^2) = \frac{1}{2n} \sum_{i=1}^{n} g(d_i^2, \theta_1, \theta_2)$$

(3.43)

with \( \theta_1, \theta_2 \) such that

$$\frac{1}{n} \sum_{i=1}^{n} \gamma_k(d_i^2, \theta_1, \theta_2) = \frac{w_k}{\sigma^2}, \quad k=1,2.$$ (3.44)

Since for any pair of signal-to-noise ratios \((w_1/\sigma^2, w_2/\sigma^2)\) there exist solutions \( \theta_1 \) and \( \theta_2 \) to (3.41) and to (3.44), the identity in (3.38) will follow if we can show that for every fixed positive pair \((\theta_1, \theta_2)\)

$$\lim_{n \to \infty} \frac{1}{2n} \sum_{i=1}^{n} g(d_i^2, \theta_1, \theta_2) = \frac{1}{4\pi} \int_{-\pi}^{\pi} g(\rho^2(\omega), \theta_1, \theta_2) \, d\omega$$

(3.45)

and

$$\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \gamma_k(d_i^2, \theta_1, \theta_2) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \gamma_k(\rho^2(\omega), \theta_1, \theta_2) \, d\omega.$$ (3.46)

To prove this, we need to examine the behavior as \( n \to \infty \) of \((d_1^2, \ldots, d_n^2)\), the eigenvalues of the Jacobi matrix \( S^T S \) in (3.25). It can be shown that

$$d_i^2 = \rho_{12}^2 + \rho_{11}^2 + 2\rho_{12}\rho_{21} \beta_i, \quad i=1, \ldots, n$$ (3.47)

where \((\beta_i, i=1, \ldots, n)\) are the roots of the \( n \)th degree polynomial \( T_{n+1}(x) \) obtained through the recursion

$$T_{k+1}(x) = 2xT_k(x) - T_{k-1}(x) \quad \text{with} \quad T_0(x) = 1; \quad T_1(x) = -\frac{\rho_{21}}{\rho_{12}}.$$

In special cases it is indeed possible to obtain closed-form expressions for the eigenvalues of \( S^T S \), for example, if \( \rho_{12} = \rho_{21} \), then [22]

$$d_i^2 = \rho_{12}^2 + \rho_{11}^2 + 2\rho_{12}\rho_{21}, \quad i=1, \ldots, n.$$ (3.48)

At any rate, it is easy to show that the eigenvalues of the Toeplitz matrix \( T \) obtained by substituting the entry \((S^T S)_{i,j} = \rho_{12}^2 \) by \( \rho_{12}^2 + \rho_{21}^2 \) are equal to [22]

$$d_i^2 = \rho_{12}^2 + \rho_{11}^2 + 2\rho_{12}\rho_{21} \cos\left(\frac{2\pi i}{2n+1}\right), \quad i=1, \ldots, n.$$ (3.49)

Thus if \( d_i^2 \) were replaced by \( d_i^2 \) in the left sides of (3.45) and (3.46), these equations would follow immediately because of the continuity of the Riemann integrands therein. It is indeed valid to carry out this replacement because \( T \)

*Except for the initial conditions, this is the recursive definition of the Chebyshev polynomials, whose zeros are trigonometric [21].
These rates are achieved (by one user at a time) at the special pair \( S^*_1, S^*_2 \), where \( F(S^*_1, S^*_2) \) is achieved by the points \( \{ (R_1, R_2): 0 \leq R_1 \leq F(S^*_1, 0), 0 \leq R_2 \leq F(0, S^*_2), R_1 + R_2 \leq F(S^*_1, S^*_2) \} \)

achievable with the unique spectral pair \( (S^*_1, S^*_2) \) that maximizes the rate-sum constraint, i.e.,

\[
F(S^*_1, S^*_2) = \max_{S^*_1 \in P(\omega)} \max_{S^*_2 \in P(\omega)} F(S^*_1, S^*_2)
\]

(3.50)

and the rate-pairs \( B \) and \( B' \) correspond to \( (R_1, R_2) = (F(S^*_1, 0), F(S^*_1, S^*_2) - F(S^*_1, 0)) \) and \( (R_1, R_2) = (F(S^*_1, S^*_2) - F(0, S^*_2), F(0, S^*_2)) \), respectively.

Note that according to the optimality conditions in (IV.3), (IV.4) (particularized to \( \alpha = 1/2 \)), the spectral pair \( (S^*_1, S^*_2) \) is the solution to

\[
\frac{S^*_k(\omega)}{\sigma^2} = \phi^*_k(\omega)
\]

\[
= \max \left\{ \frac{1}{2} \theta_k, 1 + \frac{1 + \phi^*_k(\omega)}{1 + \phi^*_k(\omega)(1 - \rho^2(\omega))} \right\},
\]

(3.51)

where \( \theta_k \) is chosen so that (IV.2) is satisfied. Since \( \rho_{12} \) and \( \rho_{21} \) are nonzero, \( \rho(\omega) \) is not a constant function of \( \omega \), and hence neither are \( S^*_1(\omega) \) and \( S^*_2(\omega) \). However, the individual rate constraints

\[
\frac{1}{4\pi} \int_{-\infty}^{\infty} \log \left( 1 + \frac{S^*_k(\omega)}{\sigma^2} \right) d\omega
\]

are maximized over the set \( P(\omega) \) by the constant functions \( S^*_k(\omega) = w_k \) and thus \( (S^*_1, S^*_2) \) fails to achieve the largest possible individual rates:

\[
C_k = \frac{1}{2} \log \left( 1 + \frac{w_k}{\sigma^2} \right), \quad k = 1, 2.
\]

These rates are achieved (by one user at a time) at the points \( A \) and \( A' \) in Fig. 5. Point \( A \) is achieved by the spectral pair \( (w_1, S^*_2) \), where

\[
F(w_1, S^*_2) = \max_{S^*_2 \in P(\omega)} F(w_1, S^*_2)
\]

\[
= \max_{S^*_2 \in P(\omega)} \frac{1}{4\pi} \int_{-\infty}^{\infty} \log \left( 1 + \frac{w_1}{\sigma^2} + \frac{S^*_2(\omega)}{\sigma^2} \right) d\omega
\]

\[
\cdot \left[ 1 + \frac{w_1}{\sigma^2} (1 - \rho^2(\omega)) \right] d\omega
\]

(3.51)

where \( \beta \) is chosen so that \( 1/2\pi \int S^*_k(\omega) d\omega = w_k \). Note that (3.52) admits the classical water-filling interpretation [24], [25] arising in the study of colored Gaussian single-user channel capacity.

The segment uniting \( A \) and \( B \) does not belong to the boundary of the capacity region, and therefore, \( C \) is not a heptagon. This property which is illustrated by the capacity region in Fig. 5 can be proved as follows. Choose \( 1/2 < \alpha < 1 \) such that the rate pairs \( A \) and \( B \) (and their convex combinations) achieve the same value of the function \( \alpha^2 R_1 + (1 - \alpha^2) R_2 \), i.e. (cf. (3.35)),

\[
(2\alpha^2 - 1) F(w_1, 0) + (1 - \alpha^2) F(w_1, S^*_2) = (2\alpha^2 - 1) F(S^*_1, 0) + (1 - \alpha^2) F(S^*_1, S^*_2).
\]

If the segment between \( A \) and \( B \) belonged to the boundary of the capacity region, then both \( A \) and \( B \) would attain \( \max_{(R_1, R_2) \in \mathbb{C}} \alpha^2 R_1 + (1 - \alpha^2) R_2 \). However, this is not possible due to the strict concavity of the function \( (2\alpha - 1) F(S^*_1, 0) + (1 - \alpha) F(S^*_1, S^*_2) \); any convex combination of the spectral pairs \( (w_1, S^*_2) \) and \( (S^*_1, S^*_2) \) will achieve strictly higher values of \( \alpha^2 R_1 + (1 - \alpha^2) R_2 \) than \( A \) and \( B \). In fact, the same argument can be employed to show that the transition from \( A \) to \( B \) contains no straight lines.

We are now ready to state and prove our main result concerning the capacity region of the asynchronous Gaussian multiple-access channel wherein the transmitters ignore their mutual offset \( \tau_2 = \tau_1 \). The transmitters only know that the crosscorrelations \( (\rho_{11}, \rho_{21}) \) that parametrize the channel belong to an uncertainty set \( \Gamma \), which is determined by the choice of the signature waveforms. For example, if both users are assigned a rectangular waveform then the uncertainty set is equal to the segment \( \Gamma = \{ 0 \leq \rho_{12} \leq 1, 0 \leq \rho_{21} \leq 1, \rho_{12} + \rho_{21} = 1 \} \). Note that in practical applications it may be of interest to model channels where the offset is not the only source of uncertainty for the crosscorrelations; for example, if the signature waveforms are sinusoidally modulated, the crosscorrelations depend on the relative phase between the carriers (e.g., for rectangular signals modulated by sinusoids whose frequency is a large multiple of the inverse of the symbol period \( T \), we get \( \Gamma = \{ 0 \leq \rho_{12} \leq 1, 0 \leq \rho_{21} \leq 1, \rho_{12} + \rho_{21} \leq 1 \} \) ). The following result puts no restrictions on the source of the uncertainty of the set of cross correlations \( \Gamma \).

**Theorem 2:** The capacity region of the energy-constrained asynchronous Gaussian multiple-access channel where the transmitters do not know their mutual offset is
true, and we need to follow an alternative route, suggested by the following result.

\[
S_1(\omega) \frac{1}{\sigma^2} d\omega, \quad 0 \leq R_1 \leq \frac{1}{4\pi} \int_{-\pi}^{\pi} \log \left( 1 + \frac{S_1(\omega)}{\sigma^2} \right) d\omega
\]

\[
S_2(\omega) \frac{1}{\sigma^2} d\omega, \quad 0 \leq R_2 \leq \frac{1}{4\pi} \int_{-\pi}^{\pi} \log \left( 1 + \frac{S_2(\omega)}{\sigma^2} \right) d\omega
\]

Proof: Having shown the result for the special case of a singleton uncertainty set \( \Gamma = \{(\rho_{12}, \rho_{21})\} \), we will be able to proceed at a faster tempo by invoking several lemmas used in the proof of Theorem 1. The capacity of the compound decoder-informed multiple-access channel with memory can be shown to satisfy ([7], see also [26, p. 288])

\[
C^* = \text{closure} \left( \lim_{n \to \infty} \frac{1}{n} C_n^* \right)
\]

with

\[
C_n^* = \bigcup_{\mathcal{P}^* \neq 0} \bigcap_{\mathcal{X}^* \neq 0} \left\{ (R_1^*, R_2^*) : \begin{eqnarray}
0 \leq R_1 \leq \frac{1}{2} \log \det \left[ I_n + \rho_{12} \rho_{21} \right] \\
0 \leq R_2 \leq \frac{1}{2} \log \det \left[ I_n + \rho_{12} \rho_{21} \right]
\end{eqnarray} \right\}
\]

where \( Y^*(\rho_{12}, \rho_{21}) \) denotes the output of channel (2.4) with crosscorrelations \( (\rho_{12}, \rho_{21}) \), and \( X_1^* \) and \( X_2^* \) are independent random variables satisfying the same input constraints as in (3.15). This follows simply because the direct coding theorem can be proved using codebooks that do not depend on the actual channel (via random selection) while the fact that for reliable communication a code has to be good no matter which actual channel is in effect establishes the converse theorem. Using Lemma 1 and proceeding as in Theorem 1, we obtain that

\[
C_n^* = \bigcup_{\mathcal{P}^* \neq 0} \bigcap_{\mathcal{X}^* \neq 0} \left\{ (R_1, R_2) : 0 \leq R_1 \leq \frac{1}{2} \log \det \left[ I_n + \rho_{12} \rho_{21} \right] \right\}
\]

\[
R_1 + R_2 \leq \inf_{(\rho_{12}, \rho_{21}) \neq 0} \frac{1}{n} \log \det \left[ I_n + \frac{\rho_{12} \rho_{21}}{\sigma^2} \right] 
\]

where the only difference with respect to (3.23) is the minimization of the rate-sum constraint with respect to the cross-correlation-dependent matrix \( S \).

In Theorem 1, we showed using the singular-value decomposition theorem that a set of eigenvectors exists which maximizes the three constraints in (3.23) no matter which eigenvalues are used, thus reducing the union therein to one over diagonal matrices. Here this property is no longer
Using the decomposition (3.58) in lieu of (3.24) and Lemma 2 with $\Delta = D$, and proceeding in a way similar to the proof of Theorem 1, we obtain

$$
\bigcup_{\Sigma_i \geq 0} \left\{ (R_1, R_2), 0 \leq R_1 \leq \frac{1}{2n} \log \det \left[ I_n + \sigma^{-2} \Sigma_1 \right], 0 \leq R_2 \leq \frac{1}{2n} \log \det \left[ I_n + \sigma^{-2} \Sigma_2 \right] \right\}
$$

$$
\bigcup_{\sum_{k=1}^2 \Sigma_k \leq \sigma^2} \left\{ (R_1, R_2), 0 \leq R_1 \leq \frac{1}{2n} \sum_{i=1}^n \log \left( 1 + \frac{\lambda_i^1}{\sigma^2} \right), 0 \leq R_2 \leq \frac{1}{2n} \sum_{i=1}^n \log \left( 1 + \frac{\lambda_i^2}{\sigma^2} \right) \right\}
$$

where $d_i = d_i^1 + d_i^2 + 2d_i^1d_i^2 \cos \frac{i\pi}{n} - \frac{2\pi(i-1)}{n}$. However, note that the right side of (3.63) can be written as

$$
\lim_{n \to \infty} \max_{\phi_1, \phi_2} \inf_{\rho_1, \rho_2} \frac{1}{\rho_1^2 + \rho_2^2 + 2\rho_1\rho_2 \cos \frac{i\pi}{n} - \frac{2\pi(i-1)}{n}}
$$

where

$$
\rho_1^2 = \rho_1^2 + \rho_2^2 + 2\rho_1\rho_2 \cos \frac{i\pi}{n} - \frac{2\pi(i-1)}{n}
$$

Again, notice that since the capacity region is achieved with stationary inputs, Theorem 2 holds regardless of whether or not the transmitters are frame synchronous.

**Corollary:** If both users are assigned identical waveforms (and they do not know their mutual offset), then the capacity region is invariant to symbol (and frame) asynchronism.
Fig. 7. Symbol-asynchronous Gaussian capacity region when transmitters do not know offset between their signals.

Proof: Because $1 - \rho_1^2 - \rho_2^2 - 2\rho_1\rho_2\cos \omega \geq 0$, it is easy to see that the asynchronous capacity region (3.53) reduces to the Cover-Wyner region (2.7) if the uncertainty set $\Gamma$ includes either $(0,1)$ or $(1,0)$. This occurs when both users are assigned the same waveform.

In Fig. 7 we can see the asynchronous capacity regions corresponding to two different assignments of the signature waveform: a) identical signals, resulting in the Cover-Wyner pentagon, and b) signals that are orthogonal when symbol synchronous, resulting in a pentagon with smooth corners.

IV. EFFICIENCY REGION

A fruitful way to represent the multiple-access capacity region is to consider the effective signal-to-noise ratio of a user who transmits at rate $R$, which we define as the signal-to-noise ratio, $\gamma$, required to achieve capacity $R$ in a single-user channel, i.e.,

$$\gamma = \exp\{2R\} - 1.$$ (4.1)

Since the mapping in (4.1) is one-to-one, the rate and the effective signal-to-noise ratio give the same information. It is convenient to normalize the effective signal-to-noise ratio with respect to the actual signal-to-noise ratio. This results in the performance measure we refer to as efficiency $\eta$, which is a parameter ranging from 0 to 1 that quantifies the performance degradation suffered by each user because of the presence of other users in the channel. Once the capacity region of a multiple-access channel is known it is immediate to obtain the efficiency region, by substituting each of the individual rates in terms of the respective efficiencies, i.e.,

$$R_i = \frac{1}{2} \log\left(1 + \eta_i \frac{\sigma_i^2}{\sigma_0^2}\right).$$ (4.2)

For example, it follows from the capacity region in (3.6) that the efficiency region of the symbol-synchronous channel is equal to

$$\left\{(\eta_1, \eta_2) : 0 \leq \eta_1 \leq 1, 0 \leq \eta_2 \leq 1, \eta_1\eta_2 - (1 - \eta_1)\frac{\sigma_2^2}{\sigma_1^2} - (1 - \eta_2)\frac{\sigma_1^2}{\sigma_2^2} \leq 1 - \rho^2\right\}.$$ (4.3)

where recall that $\rho$ is the crosscorrelation between the assigned waveforms.

This efficiency region is illustrated in Fig. 8, as a function of the background noise level. For low signal-to-noise ratios the efficiency region occupies nearly all of the unit square because the main mechanism limiting performance is the background Gaussian noise, rather than the multiple-access interference. Conversely, it is apparent from Fig. 8 that for moderate-to-large signal-to-noise ratios the efficiency region converges to an asymptotic region which quantifies the underlying limitation of the multiple-access channel due to the cross correlation between the assigned signal waveforms. The region in (4.3) admits a particularly simple asymptotic expression as the noise spectral density goes to zero:

$$E = \left\{(\eta_1, \eta_2) : 0 \leq \eta_1 \leq 1, 0 \leq \eta_2 \leq 1, \eta_1\eta_2 \leq 1 - \rho^2\right\}.$$ (4.4)

The usefulness of the asymptotic efficiency region is threefold: it provides a simple way to characterize multiple-access capacity in high signal-to-noise ratio situations; it gives a lower bound to the efficiency region achievable at any background noise level, which depends only on the assigned signal waveforms and not on the signal-to-noise ratios, and it gives an intuitive characterization of the performance degradation in a multiuser channel in terms

---

"An analogous performance measure was defined in the analysis of the minimum uncoded error probability of Gaussian multiple-access channels [27]."

"It is shown in [14] that the efficiency region is monotonically increasing in $\sigma^2$ (cf. Fig. 8)."
The capacity region of the symbol-asynchronous Gaussian multiple-access channel is (via (3.53) and (4.2)) equal to

\[
\left\{ (\eta_1, \eta_2) : \eta_1 \leq \exp \frac{1}{2\pi} \int_{-\pi}^{\pi} \log S_1(\omega) \, d\omega, \eta_2 \leq \exp \frac{1}{2\pi} \int_{-\pi}^{\pi} \log S_2(\omega) \, d\omega \right\}
\]

of the additional power required to achieve single-user capacity.

For example, suppose that the users' objective is to transmit at rates \( \bar{R}_1 \) and \( \bar{R}_2 \), respectively. If they were operating in a single-user channel, these rates could be achieved with powers: \( w_k = \sigma^2 (\exp (2\bar{R}_k) - 1) \), \( k = 1, 2 \). However, when they share the same channel, these powers are no longer sufficient to guarantee reliable communication at rates \( \bar{R}_1 \) and \( \bar{R}_2 \). The asymptotic efficiency region (4.4) indicates that the sum of their powers in dB has to increase by \(-10 \log (1 - \rho^2)\) dB and that the way the users split the burden of increasing their powers is immaterial as long as the total power increases by the prescribed amount.

In the conventional scalar multiple-access channel, which corresponds to the users being assigned the same waveforms, i.e., \( \rho = 1 \), the asymptotic efficiency region is (via (4.4)):

\[
\{ 0 \leq \eta_1 \leq 1, \eta_2 = 0 \} \cup \{ \eta_1 = 0, 0 \leq \eta_2 \leq 1 \}.
\]

Thus when the signal-to-noise ratio is high, the best strategy is to let one of the users transmit at practically full single-user speed, while the other user's rate is kept at a very low level. This is considerably more efficient than time-division multiple-access (TDMA) signaling whose asymptotic efficiency is equal to zero for both users—although if both rates are required to be the same, then TDMA is indeed almost as good as the best coding for low background noise (see [28]).

These conclusions do not hold in the case where the assigned waveforms are different \((|\rho| < 1)\). For example, suppose that \( \rho = 0.1 \) and two equal-rate equal-energy users with signal-to-noise ratio equal to 20 dB transmit at the maximum possible rates. Had the users employed TDMA, each of them would have required approximately 40 dB to attain the same rate. Even in the case where there is heavy cross correlation between the signals, TDMA is not near-optimum, e.g., if \( \rho = 0.9 \), then TDMA would still require 33 dB to attain the same rate.

The efficiency region of the asynchronous Gaussian multiple-access channel is (via (3.53) and (4.2)) equal to

\[
\left\{ (\eta_1, \eta_2) : \eta_1 \leq \exp \frac{1}{2\pi} \int_{-\pi}^{\pi} \log \left( \sigma^2 + S_1(\omega) \right) \, d\omega, \eta_2 \leq \exp \frac{1}{2\pi} \int_{-\pi}^{\pi} \log \left( \sigma^2 + S_2(\omega) \right) \, d\omega \right\}
\]

Letting \( \sigma^2 \to 0 \), the asymptotic efficiency region results:

\[
\begin{align*}
\eta_1 \eta_2 & \leq \inf_{(p_{11}, p_{21}) \in \Gamma} \exp \frac{1}{2\pi} \int_{-\pi}^{\pi} \log \left( 1 - (\rho_{11}^2 + \rho_{21}^2) - 2 \rho_{12} \rho_{21} \cos \omega \right) \, d\omega \\
& \leq \inf_{(p_{11}, p_{21}) \in \Gamma} \left[ \frac{\sqrt{1 - (\rho_{12} + \rho_{21})^2} + \sqrt{1 - (\rho_{12} - \rho_{21})^2}}{2} \right]^2
\end{align*}
\]

The constraints in (4.6) depend on the spectral densities only through their geometric mean; therefore, all three constraints are maximized simultaneously by a single pair of spectral densities because the function that maximizes the geometric mean subject to a constraint on its arithmetic mean is constant. Therefore, (4.6) is equal to the efficiency region achievable with white spectral densities \( S_1(\omega) = 1, \omega \in [-\pi, \pi] \), which implies that white inputs, while not optimum in general, achieve capacity asymptotically as the background noise level goes to zero. Then the asymptotic efficiency region is the intersection of the unit square with the hyperbolic region

\[
\eta_1 \eta_2 \leq \inf_{(p_{11}, p_{21}) \in \Gamma} \exp \frac{1}{2\pi} \int_{-\pi}^{\pi} \log \left( \sigma^2 + S_1(\omega) \right) \, d\omega
\]

where the definite integral is found in [29, p. 560] (see also [30, p. 384]). Note that this result generalizes the constraint \( 1 - \rho^2 \) obtained for the product of the asymptotic efficiencies in (4.4) when the users are synchronous. Equation (4.7) indicates that contrary to what is sometimes assumed in pseudonoise sequence design, it is as important to minimize the\textit{difference} between the cross correlations as to minimize their\textit{sum} (the so-called periodic cross correlation). The function on the right side of (4.7) is tightly approximated by \( 1 - \rho_{12}^2 - \rho_{21}^2 \) for low cross-correlation values such as those in Fig. 9, where we can see the uncertainty set of cross correlations between two carrier-modulated spread spectrum waveforms used in CDMA [31]. In this case, the minimization in (4.7) is attained by the rightmost point in Fig. 9.
IEEE TRANSACTIONS ON INFORMATION THEORY, VOL 35, NO. 4. JULY 1989

sufficiency of the respective outputs, the channels in (1.3) have the same capacity region.

APPENDIX I

To simplify the coding theorems for multiple-access channels with memory invoked to find the capacity region of the asynchronous Gaussian channel, it is convenient to obtain a discrete-time channel equivalent to (2.4) and whose noise process is independent. The idea is to obtain a set of sufficient statistics that are independent given the transmitted symbols, but that unlike those in (2.4) are not a minimal set. We define (cf. Fig. 2):

\[ y_i(t) = \int_{(t-iT) + n}^{(t(i+1)T + n)} x_j(t) s_j(t-iT - \tau_j) dt \]  

\[ y_i^*(t) = \int_{T + \tau_j}^{(t-iT + \tau_j)} x_j(t) s_j(t-iT - \tau_j) dt \]  

It is clear from (2.2) and (1.1) that

\[ y_i(i) = y_i^*(i) + y_i^*(i+1) \]  

Thus the set of quantities \( \{ y(i) \}_{i=1}^n \) and \( \{ y^*(i) \}_{i=1}^n \) are sufficient statistics for the transmitted messages. To obtain the explicit dependence of \( y_i(i) \) and \( y_i^*(i) \) on the transmitted symbols, it is convenient to define the partial energies of \( \{ s_j(t) \} \):

\[ E_s(i) = \int_{(i-1)T + \tau_j}^{iT + \tau_j} s_j(t)^2 dt \]  

\[ E_{s^*}(i) = \int_{T + \tau_j}^{iT + \tau_j} s_j(t)^2 dt \]  

It follows from (2.1), (2.4), (1.1), and (1.2) that we can write

\[ y_i(i) = \sum_{j=1}^n \rho_{j2} \rho_{j1} \left[ b_j(i) \right] \]  

\[ y_i^*(i) = \sum_{j=1}^n \rho_{j2} e_j \left[ b_j(i) \right] \]  

\[ y_i^*(i-1) = \sum_{j=1}^n \rho_{j2} e_j \left[ b_j(i-1) \right] \]  

where

\[ E = \left[ L_1 \quad S \quad I_s \right] \]  

The channel in (1.3) has memory because of the dependence on previous inputs; however, since the random process \( \{ n(i) \} \) is white, the noise sequence in (1.3) is independent. Because of the sufficiency of the respective outputs, the channels in (2.4) and (1.3) have the same capacity region.

APPENDIX II

Proof of Lemma 1: To prove the identity

\[ \det \left[ I_{2n} + \sigma^{-2} E[X^TX^T] R \right] = \det \left[ I_{2n} + \frac{1}{\sigma^2} \left[ \begin{array}{cc} \Sigma_j & 0 \\ 0 & I_s \end{array} \right] \right], \]  

(II.1)

we will first decompose the crosscorrelation matrices \( E[X^TX^T] \) and \( R \) using the Kronecker product:

\[ E[X^TX^T] = E \left[ \begin{array}{cc} X_1^T & X_2^T \\ X_2^T & X_1^T \end{array} \right] \]  

\[ = \sigma^{-2} \left[ \begin{array}{cc} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{array} \right] \]  

(II.2)

and

\[ R = \left[ \begin{array}{cc} I_1 & 0 \\ 0 & I_2 \end{array} \right] \]  

(II.3)

It is straightforward to check that if \( A, B, C, D \) are \( n \times n \) matrices, then

\[ P = \left[ \begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array} \right] + \Sigma_1 \otimes \left[ \begin{array}{cc} 0 & 0 \\ 0 & 1 \end{array} \right] + \Sigma_2 \otimes \left[ \begin{array}{cc} 0 & 0 \\ 0 & 1 \end{array} \right] \]  

(II.4)

where \( P \) is the permutation (orthogonal) matrix whose only nonzero entries are

\[ P_{j-1,j} = 1, \quad j=1, \ldots, n \]  

(II.5a)

\[ P_{2j-2,j} = 1, \quad j=1, \ldots, 2n \]  

(II.5b)

Therefore, (II.2) and (II.3) can be written, respectively, as

\[ \gamma^T \]  

(II.6)

where \( \gamma \) is the stochastic vector that enters the problem. It follows from (II.5), (II.7), and the orthogonality of \( P \) upon using the identity \( \det(I + AB) = \det(I + BA) \).

APPENDIX III

Proof of Lemma 2: Define the following diagonal matrices:

\[ R = \text{diag}(\cos \theta_1, \ldots, \cos \theta_n) \]  

\[ G = \text{diag}(e^{i\theta_1}, \ldots, e^{i\theta_n}) \]  

where \( \theta_i = (1/2) \sin \theta_i \in [0, \pi/4] \) and \( \delta_i = \sin \theta_i \). It is straightforward to check that

\[ F^* G^* \]  

(III.1)
therefore, we can write
\[\begin{align*}
&\text{set } I_{2n} + \frac{1}{\sigma^2} \left[ \begin{array}{cc}
A & 0 \\
0 & B
\end{array} \right] \left[ \begin{array}{cc}
I_n & \Delta^* \\
\Delta & I_n
\end{array} \right], \\
&= \text{set } I_{2n} + \frac{1}{\sigma^2} P \left[ \begin{array}{cc}
F & G^* \\
G & F
\end{array} \right] P^T \left[ \begin{array}{cc}
A & 0 \\
0 & B
\end{array} \right] P \left[ \begin{array}{cc}
F & G^* \\
G & F
\end{array} \right] P^T.
\end{align*}\]

where \( P \) is the orthogonal matrix introduced in Appendix II. It follows from (II.4) that
\[P \left[ \begin{array}{cc}
F & G^* \\
G & F
\end{array} \right] P^T\]
is a block diagonal matrix whose \( i \)-th diagonal block is the \( 2 \times 2 \)
matrix
\[\begin{bmatrix}
\cos \theta_i & e^{-i\phi} \sin \theta_i \\
e^{i\phi} \sin \theta_i & \cos \theta_i
\end{bmatrix},\]

whereas
\[P \left[ \begin{array}{cc}
A & 0 \\
0 & B
\end{array} \right] P^T\]
is a block matrix whose \( i \)-th diagonal block is
\[\begin{bmatrix}
a_i & 0 \\
0 & b_i
\end{bmatrix}.\]

Because of the assumption on the magnitude of the elements of \( \Delta \), the hermitian matrix
\[\begin{bmatrix}
I_n & \Delta^* \\
\Delta & I_n
\end{bmatrix}\]
is nonnegative definite and, therefore, so is the matrix in the right side of (III.2). This allows us to apply the generalized Hadamard inequality [18], which when substituted in (IV.3) results in
\[\begin{align*}
L(\Phi_1, \Phi_2) &= -\int_{-\pi}^{\pi} \frac{d\omega}{\sigma^2} \left[ \int_{-\pi}^{\pi} \Phi_1(\omega) d\omega - \frac{w_i}{\sigma^2} \right] - \frac{1}{2} \int_{-\pi}^{\pi} \Phi_2(\omega) d\omega - \frac{w_i}{\sigma^2} \\
&= \frac{1}{2} \int_{-\pi}^{\pi} \Phi_2(\omega) d\omega - \frac{w_i}{\sigma^2}, \quad k=1,2.
\end{align*}\]

The optimum pair \((\Phi_1^*, \Phi_2^*)\) is the unique element in the cone of nonnegative function pairs that satisfies \(^{19}\) (c.g. [33, p. 227])
\[\begin{align*}
\theta_1 &\geq f_1(\Phi_1^*(\omega), \Phi_2^*(\omega), \rho^2(\omega)) \\
&= \frac{2a-1}{1 + \Phi_1^*(\omega)} \\
&+ \frac{(1-a)[1 + \Phi_1^*(\omega)(1 - \rho^2(\omega))]}{1 + \Phi_1^*(\omega) + \Phi_2^*(\omega) + \Phi_1^*(\omega) \Phi_2^*(\omega)(1 - \rho^2(\omega))},
\end{align*}\]

and the above relations hold for otherwise \( \Phi_2^*(\omega) = 0\) for all \( \omega \in \{\pi, \pi\}\), which does not satisfy (IV.2). Let us now see what conditions force each of the solutions to be zero at a particular frequency.

On the one hand, if \( \Phi_1^*(\omega) = 0\), then (IV.5) and (IV.6) imply that
\[\begin{align*}
\psi_1(\omega) &= \frac{1 - \alpha}{\theta_1} \\
&= \frac{1}{\theta_1},
\end{align*}\]

which implies that
\[\begin{align*}
\theta_1 \geq \theta_2, \quad i \geq 0. \quad \text{(IV.6)}
\end{align*}\]

for otherwise \( \Phi_2^*(\omega) = 0\) for all \( \omega \in \{\pi, \pi\}\), which does not satisfy (IV.2).

Let us now see what conditions force each of the solutions to be zero at a particular frequency.

On the one hand, if \( \Phi_1^*(\omega) = 0\), then (IV.3) and (IV.5) imply that
\[\begin{align*}
\psi_1(\omega) &= \frac{1 - \alpha}{\theta_1} \\
&= \frac{1}{\theta_1},
\end{align*}\]

which when substituted in (IV.3) results in
\[\begin{align*}
\rho^2(\omega) &\geq \frac{\alpha - \delta_i}{(1-a) - \theta_2}.
\end{align*}\]

One conclusion that can be drawn from condition (IV 8) is that
\[\begin{align*}
\alpha &> \theta_2, \quad \text{(IV.9)}
\end{align*}\]

because otherwise, \( \Phi_2^*(\omega) = 0\) for all \( \omega \in \{\pi, \pi\}\).

On the other hand, if \( \Phi_1^*(\omega) = 0\), then (IV.3) and (IV.9) imply that
\[\begin{align*}
\psi_1(\omega) &= \frac{\alpha}{\theta_1} \\
&= \frac{1}{\theta_1}.
\end{align*}\]
which upon substitution in (IV.5) results in the condition
\[ p^2(\omega_b) \geq \frac{\alpha}{1-\alpha} \frac{(1-\alpha) - \theta_1}{\alpha - \theta_1}. \] (IV.11)

Note that since \(1/2 \leq \alpha < 1\), (IV.8) and (IV.10) cannot be true simultaneously if \(p^2(\omega_b) < 1\). However, they can indeed be false simultaneously, in which case (IV.3) and (IV.4) are satisfied with equality and \((\Phi^*\omega), \Phi^*\omega)\) are the positive solutions to the system of the following equations:
\[ \Phi^*\omega = \frac{1-\alpha}{\theta_1} - \frac{1 + \Phi^*\omega}{1 + \Phi^*\omega}(1 - p^2(\omega)) \] (IV.12)
\[ 2\alpha - 1 \frac{(1-\alpha)[1 + \Phi^*\omega(1 - p^2(\omega))]}{1 + \Phi^*\omega + \Phi^*\omega(\Phi^*\omega + \Phi^*\omega)(1 - p^2(\omega))} = \theta_1. \] (IV.13)

It follows that for each fixed pair of Lagrange multipliers, the maximizing spectra \((\Phi^*\omega), \Phi^*\omega)\) depend on \(\omega\) only through \(p^2(\omega)\), i.e., \(\Phi^*\omega = \gamma_1(p^2(\omega), \theta_1, \theta_2)\), which is a continuous function of \(p^2(\omega)\).

**APPENDIX V**

**Proof of Lemma 4:** Let \(E\) be the \(2n \times 2n\) matrix whose only nonzero elements are \(E_{2i} = E_{2i-1} = 1\), and let \(Z\) be the nonnegative-definite square root of \(\Sigma^*\). Then we can write
\[ \log \det \left[ I_{2n} + \frac{1}{\sigma^2} \left[ \begin{array}{c} \Sigma_0 \ 0 \\ 0 \ \Sigma_1 \end{array} \right] \left[ \begin{array}{c} I_s \ S^T \\ S \ I_s \end{array} \right] \right] \]
\[ - \log \det \left[ I_{2n} + \frac{1}{\sigma^2} \left[ \begin{array}{c} \Sigma_0 \ 0 \\ 0 \ \Sigma_1 \end{array} \right] \left[ \begin{array}{c} I_s \ S^T \\ S \ I_s \end{array} \right] \right] \]
\[ = \log \det \left[ I_{2n} + \rho_n \text{ME} \right] \] (V.1)
where
\[ M = \sigma^{-2} \left[ \begin{array}{c} Z_1 \ 0 \\ 0 \ Z_2 \end{array} \right] \left[ I_{2n} + \frac{1}{\sigma^2} \left[ \begin{array}{c} \Sigma_0 \ 0 \\ 0 \ \Sigma_1 \end{array} \right] \left[ \begin{array}{c} I_s \ S^T \\ S \ I_s \end{array} \right] \right]^{-1} \left[ \begin{array}{c} Z_1 \ 0 \\ 0 \ Z_2 \end{array} \right]. \] (V.2)

The determinant in the right side of (V.1) is easily computed:
\[ \det \left[ I_{2n} + \rho_n \text{ME} \right] = 1 + p^2 M_{12}^2 + 2p^2 M_{12} M_{21} - p^2 M_{11} M_{21} \]
\[ \leq 1 + 2p^2 M_{12} \]
\[ \leq 1 + 2[p^2 M_{11} M_{21}]^{1/2} \]
\[ + \frac{1}{\gamma_1 + w_1} \frac{1}{\gamma_1 + w_1} \frac{1}{\gamma_1 + w_1} \frac{1}{\gamma_1 + w_1} \frac{1}{\gamma_1 + w_1} \]
\[ \gamma = \frac{1}{\gamma_1 + w_1} \frac{1}{\gamma_1 + w_1} \frac{1}{\gamma_1 + w_1} \frac{1}{\gamma_1 + w_1} \frac{1}{\gamma_1 + w_1} \] (V.3)

where the first two inequalities follow from the nonnegative definiteness of \(M\), and the third inequality is a consequence of the fact that, according to (V.2), the largest eigenvalue of \(M\) is upper-bounded by that of
\[ \left[ \begin{array}{c} \Sigma_0 \ 0 \\ 0 \ \Sigma_1 \end{array} \right] \]
which is in turn upper-bounded by \((\text{tr} \Sigma_0 + \text{tr} \Sigma_1)/\alpha^2 \leq (w_1 + n)/\alpha^4\). Interchanging the roles of \(S\) and \(S^*\), the same bound can be shown for the reverse difference in (V.1), i.e.,
\[ \frac{1}{n} \log \det \left[ I_{2n} + \frac{1}{\sigma^2} \left[ \begin{array}{c} \Sigma_0 \ 0 \\ 0 \ \Sigma_1 \end{array} \right] \left[ \begin{array}{c} I_s \ S^T \\ S \ I_s \end{array} \right] \right] \]
\[ - \frac{1}{n} \log \det \left[ I_{2n} + \frac{1}{\sigma^2} \left[ \begin{array}{c} \Sigma_0 \ 0 \\ 0 \ \Sigma_1 \end{array} \right] \left[ \begin{array}{c} I_s \ S^T \\ S \ I_s \end{array} \right] \right] \]
\[ \leq \frac{1}{n} \log \left( 1 + \frac{w_1 + n}{\sigma^2} \right), \] (V.4)

completing the proof of Lemma 4.

**REFERENCES**

Control and Optimization Methods in Communication Network Problems

Anthony Ephremides
Sergio Verdú

Reprinted from
IEEE TRANSACTIONS ON AUTOMATIC CONTROL
Vol. 34, No. 9, September 1989
Control and Optimization Methods in Communication Network Problems

ANTHONY EPHERMIDES, FELLOW, IEEE, AND SERGIO VERDÚ, SENIOR MEMBER, IEEE

Abstract—In this paper we focus on two areas of communication network design in which methods of control and optimization theory have proven useful. These are the area of multiple access communication (for networks with shared links such as radio networks and local area networks) and the area of network routing (for networks with point-to-point interconnections). We review a few selected problems in each area to show the role of the control concepts involved and we then proceed to identify other areas of communication network design in which the same control theoretic and optimization methodology may be applicable and useful. We do not survey the work done in this area, nor do we review work in control areas whose methods are applicable in other communication network problems. Instead, we attempt to bring to the attention of the control systems community the numerous instances of problems arising in the pure communication network design process that can benefit from the attention and the capabilities of this community.

I. INTRODUCTION

COMMUNICATION networks are designed and built in order to share resources. If interconnecting systems and bandwidths were available at no cost, then the solution to the problem of communication would be to assign dedicated communication links (channels) of sufficient capacity to every pair of conceivable users to meet their needs. This not being the case, it is necessary to multiplex the sources of communication traffic in order to optimize various cost criteria. Frequently, this optimization is dynamic and done on the basis of feedback that monitors the evolution of the degree of utilization of the network resources. Thus, we should expect a number of problems arising in communication network design to fit naturally in the framework of control systems design. In this paper we wish to demonstrate that indeed this is the case and to show how various control and optimization methodologies have been used in the study of communication networks.

In the beginning there was a single communication network, the telephone network. It represented a multibillion dollar investment and seemed to serve reasonably adequately the voice communication needs. The explosive growth in data communication needs during the last 30 years built up the pressure for additional and alternative networking options. As a result, the notion of store-and-forward switching (known also as message switching) was introduced in the early 1960's. This notion represented a breakthrough since it constituted a radical reversal of thinking with respect to the circuit-switching process, namely, instead of securing an open, dedicated "pipe" for the transmission of messages by means of hardware switches, it allowed a step-by-step (node-by-node) forwarding of messages, thereby permitting each node to switch messages by deciding when and where to deliver the messages in its buffer. In the last 20 years we have seen an avalanche of technologies (fast switching, time division multiplexing, local area networks, fiber optical networks, integrated services digital networks, etc.) and a proliferation of operational public and private networks that put these technologies to test and challenged communication engineers. In addition, they should challenge control engineers as well.

Without attempting a survey of this vast application area we wish to promulgate the viewpoint that many (if not most) specific sub-problems in the network design process are natural control problems. In support of this thesis, we choose, first, to illustrate how two major areas in communication networks (routing and multiple access) have benefitted from the use of techniques borrowed from what is traditionally perceived as control systems methodology and, second, to mention additional areas that are likely to benefit from the control systems community. As illustrated in this paper, the techniques that have proved useful in communication networks include: dynamic programming (e.g., [2], [6], [8]-[10], [22], [29], [38], [59], [47], [49], [54]); linear programming (e.g., [30], [51]); constrained and iterative optimization (e.g., [5], [14], [16], [42]); Markov decision theory tools (e.g., [2], [26], [29], [38]); control of Markov chains (e.g., [11], [17], [18], [20], [40], [45]); stability analysis of stochastic systems via Lyapunov methods (e.g., [31], [43]); sample path dominance (e.g., [2], [52]); and convergence of distributed and asynchronous algorithms (e.g., [6], [16], [42]).

The problem of routing is encountered in all and every network that does not permit the source to reach the destination in a single transmission hop, but instead it must traverse a path of intermediate links. By contrast, the problem of multiple access is encountered primarily in those networks that permit the nodes to reach their destination directly in one hop by having to share the same link with other transmitting nodes. In addition, the two problems are fundamentally different in nature and, jointly, cover considerable ground in the networking area. Finally, together they facilitate the identification of additional design issues and the extension of the applicability of suitable control methods. Thus, they represent "cornerstone" areas of network design.

Routing can be studied either macroscopically or microscopically. The macroscopic viewpoint considers basically a flow model and determines the splitting of the flow in order to reach the destination in minimum time with efficient use of the network resources. It is traditionally referred to as static routing. The microscopic viewpoint dissects the flow process down to the atomic level of the individual transmission unit, the message (a string of bits commonly referred to as packet), and determines the path each message must follow at each of its hops through the network. It is traditionally referred to as dynamic routing. Both viewpoints are explored in Section II.

Multiple access is a collective term that refers to numerous problems that deal with the dynamic allocation of a single resource among users who can coordinate their use of that resource only by making use of that resource. These problems arise primarily in the context of radio channels but also in the
context of shared cable resources in local area networks. In Section III, we explore the main multiple access problems where control methods have been successfully applied.

Both in the case of routing as well as in the case of multiple access we place the emphasis on the control techniques that have been used. We then show how these techniques, sometimes with slight modification, can be naturally transported to other problem areas such as voice-data integration, flow control, and the scheduling of messages and links. This is done in Section IV.

II. NETWORK ROUTING

The problem of routing in communication networks is one that has received early attention and has experienced significant breakthroughs in the brief history of the field of communication networks. It is one of the first problems that gained prominence as a result of the emergence of store-and-forward switching. It is also one in which analytical tools and available theories applied nicely from the beginning.

A. Static Routing

Given a network (a set of nodes connected by directed links) a path connecting the source node to the destination node has to be selected from the set of all possible such paths. In the simplest formulation, the problem is one of finding the shortest path, i.e., a path whose length is assigned to each link and the optimization criterion is the total path length. This problem is one of the archetypical combinatorial optimization problems (the solution can be found by exhaustive enumeration of a finite set of possibilities—all possible paths from source to destination). Among the many existing shortest path algorithms (see, e.g., [41]), the Bellman-Ford algorithm (1956) is of particular interest to our exposition, both because it is based on dynamic programming and because, as we will see below, it easily lends itself to distributed asynchronous implementation. A natural choice to find the shortest path from source to destination in a layered network (i.e., one in which the nodes can be grouped in subsets $U_1, \cdots, U_N$ such that the source and destination nodes belong to $U_1$ and $U_N$, respectively, and there are links only between nodes in adjacent layers $U_{k-1}$ and $U_k$) such as the one in Fig. 1, is the dynamic programming algorithm, where the shortest paths and distances (costs-to-go) of the nodes in layer $U_k$ to the destination are computed based on the shortest paths and distances of the nodes in layer $U_{k+1}$. If the network is not layered (such as that in Fig. 2), its shortest path can be obtained by finding the shortest path in a layered network derived from the original one as specified in the Bellman-Ford algorithm: the number of layers is equal to the number of nodes in the original network, say $N$, each layer contains a copy of each of the $N$ nodes, and there is a link connecting two nodes in consecutive layers if such a link exists in the original network, in addition, copies of the same node in consecutive stages are connected by a zero-length link. (Fig. 1 was actually derived from Fig. 2 using this rule.) It is easy to see by induction that $D_k(i)$, the cost-to-go of node $i$ in layer $N - k$, is the minimum length of any path from $i$ to the destination that uses at most $k$ links (in the original network). Since no shortest path uses more than $N - 1$ links (link lengths are assumed nonnegative and, therefore, no path containing loops need be considered), the cost-to-go of node $i$ at layer $1$, $D_{N-1}(i)$, will indeed be the length of the shortest path from node $i$ to the destination. Thus, the Bellman-Ford algorithm can be formulated as the iteration

$$D_k(i) = \min_{j \in N(i)} \{D_{k-1}(j) + d_{ij}\} \quad \text{for } k = 1, \cdots, N-1 \quad (2.1)$$

where $d_{ij}$ is the length of the link from $i$ to $j$, $N(i)$ is the set of nodes for which such a link exists and it is assumed that $D_k(i) = \infty$ if $i$ is not the destination node, which corresponds to the removal of all the nodes but the destination in the final layer (Fig. 1).

Contrary to what may appear at first glance there is a lot more to network routing than finding shortest paths. After all, the shortest path may not be the best path. The reason is that the real goal is to minimize the delay experienced in going from source to destination, and the delay encountered in each link is usually a function of the amount of traffic carried by the link (as the link becomes congested, it takes longer to go through it), which is referred to as the link flow and is quantified in packets (or messages) per second. Then, assuming a given desired flow level from source to destination, the problem is how to distribute it among all the possible paths so as to minimize the total delay. In contrast to the previous more elementary formulation of the routing problem which led to the shortest path combinatorial optimization problem and which corresponds to the special case in which the link delays are independent of the flows, we now face a continuous optimization problem which can be written as

$$\begin{align*}
\text{minimize} & \quad F(x) = \sum_{(u_j)} D_0 \left( \sum_{x(n) \in P(u_j)} x(n) \right) \\
\text{subject to} & \quad x \in X = \left\{ (x(1), \cdots, x(J)) \in R^J : \sum_{n=1}^J x(n) = \lambda, x(n) \geq 0 \right\} \quad (2.2)
\end{align*}$$

1 All the algorithms and results discussed in this section can be extended to the case where there are several source-destination pairs in the network.
Fig 3 Characterization of the solution to the minimum-delay routing problem.

where the set of all paths from source to destination is labeled \{1, \ldots, J\}; \(x = (x(1), \ldots, x(J))\) is the vector of unknown nonnegative path flows which sum up to \(\lambda\), the desired flow from source to destination; \(P(i,j) \subseteq \{1, \ldots, J\}\) is the subset of paths that traverse link \((i, j)\); and \(D_y(x)\) is the portion of the overall delay contributed by the link from node \(i\) to node \(j\) when the flow it carries is equal to \(x\). In order to characterize a global solution to the optimization over a convex set in (2.2), it is natural to restrict attention to convex penalty functions. In practice, it is common that the incremental delay in a link grows with the amount of traffic it carries and, therefore, it can be assumed that the functions \(D_y\) are convex without affecting significantly the practical applicability of the results.

Now, the characterization of the solution to (2.2), \(x^*\), is straightforward. Since the feasible set \(X\) and the penalty function \(F\) are convex, it is necessary and sufficient that the directional derivative of the penalty function be nonnegative when evaluated at \(x^*\) in the direction of any of the elements of \(X\) (e.g., [37])

\[
0 \leq \lim_{\alpha \to 0} \frac{1}{\alpha} [F((1-\alpha)x^*+\alpha x) - F(x^*)] \quad \forall x \in X \tag{2.3}
\]

which translates into

\[
0 \leq \sum_{m \in P(i,j)} d_y'(m) \sum_{m \in P(i,j)} \sum_{n \in P(i,j)} [x(n) - x^*(n)]
\]

\[
= \sum_{n=1}^J [x(n) - x^*(n)] d_x'(n) \quad \forall x \in X \tag{2.4}
\]

where \(d_y'(m) = \sum _{\nu \in \mathbb{P}(m) \cap P(i,j)} d_x'(\nu) \sum \mathbb{P}(m) x^*(m)\) is the length of path \(n\) when the length of each link is equal to the derivative of its delay evaluated at the set of flows \(x\), and \(L(n)\) is the set of links used by path \(n\). The solution to (2.4), \(x^*\), is the vector in \(X\) that minimizes its inner product with the vector of distances \(d_x'(n)\). Thus, \(x^*\) puts all its weight on the smallest component(s); of \(d_x\). The conclusion is that the optimum flow uses only shortest paths computed according to the derivative of the link delays.

This solution to the minimum-delay routing problem allows us to check whether a given set of flows is optimum. Unfortunately, it does not tell us how to find the optimum flows. Indeed, we face the chicken-and-egg situation depicted in Fig. 3. The optimum flows are obtained by solving a shortest path problem; but in order to compute the link lengths it is necessary to know the optimum flows. Nevertheless, the foregoing characterization of the optimal solution does suggest a possible iterative procedure to find the optimum set of flows. Starting with a given set of flows \(x\) one can compute the minimum derivative shortest paths for that flow and, hence, a new flow, \(x^*(x)\) that is positive only along those shortest paths. The process can then be repeated, until there is no appreciable cost decrease. The region of convergence of such a procedure can be improved by letting the new flow be a convex combination of \(x\) and \(x^*(x)\), i.e.,

\[
x_{k+1} = (1-\alpha_k)x_k + \alpha_k x^*(x_k).
\]

This is the so-called flow deviation method of Fratta, Gerla, and Kleinrock [14], where \(0 \leq \alpha_k \leq 1\) is chosen to minimize

\[
F((1-\alpha_k)x_k + \alpha_k x^*(x_k))
\]

which is a special case of the feasible-direction nonlinear programming algorithm due to Frank and Wolfe [13]. The convergence of the flow deviation method to the optimum routing is rather slow because unfavorable paths tend to carry considerable flow during many iterations unless the initial routing guess is particularly fortuitous. Such a behavior can be improved by reducing the flow along each nonminimum derivative path in accordance to the delay experienced in that path. This is the idea of iterative routing algorithms based on gradient projection nonlinear optimization methods (e.g., [4]) in which the flow decrease along a nonminimum derivative path is proportional to the difference between its length and that of the shortest path (according to the first derivative of the delay function). If such a decrease would result in a negative flow, then the flow along that path is set to zero (hence, the projection to the set of feasible flows).

We have seen that the problem of static network routing can be formulated as a conceptually straightforward optimization problem that admits well-known solutions in nonlinear programming. What sets optimum routing in communication networks apart from other multicommodity flow problems arising in operations research is the fact that the optimization is carried out in real time, and often, in distributed fashion, where each node makes its own routing decisions based on local information. The review of centralized routing has revealed that the shortest path problem plays a central role in solving for the optimum routing regardless of whether the link congestion measures depend on the link flow or not. Hence, we will start the exposition of distributed routing algorithms by discussing the distributed version of the Bellman-Ford shortest path algorithm.

The Bellman-Ford updating equation in (2.1) suggests that the algorithm is suited for decentralized operation because each node can update its own estimate of distance to the destination (cost-to-go) provided it receives from its neighbors their own estimates (appearing on the right-hand side of (2.1)). The feature that makes the study of the distributed Bellman-Ford algorithm interesting is that it can run completely asynchronously, in the sense that the updating and communication times need not be coordinated and convergence can be guaranteed by simply assuming that updating and communication between nodes never cease, without any requirements whatsoever on the rate of communication. The proof of convergence is a nice illustration of the analysis of decentralized algorithms where the processors are allowed to perform their computations and to communicate the corresponding results completely independently of one another [5], [6]. The idea is to show that the estimates computed in the distributed asynchronous algorithm are always sandwiched by the estimates computed by the centralized version of the algorithm when started at two different initial conditions, and that both centralized estimates converge to the true distances to the destination node.

Those centralized estimates are denoted by \(D_k = (D_k(1), \ldots, D_k(N))\) and \(D_k = (D_k(1), \ldots, D_k(N))\), and are the result of the centralized Bellman-Ford iteration (2.1) when \(k\) is started with initial conditions \(D_0 = (\infty, \ldots, \infty, 0)\) and \(D_0 = (0, \ldots, 0)\), respectively. (The destination node is assumed to be the \(N\)th node.) Define the operator [see (2.1)]

\[
B[D_k] = \min_{D_k \in \mathbb{D}} [D_k(j) + d_j] \tag{2.6}
\]

if \(1 \leq i \leq N\), and \(B[D_k] = D_k(N)\). This operator is monotone in the sense that if \(D \leq D^*\) (i.e., if \(D(i) \leq D^*(i), i = 1, \ldots, N\), then

\[
B[D] \leq B[D^*]. \tag{2.7}
\]

The monotonicity of \(B_k\) implies that

\[
D_k \leq D_{k+1} \leq D_{k+1} \leq D_k \tag{2.8}
\]
and, moreover, it is easy to show that for sufficiently large $k$

$$D_k = D_{N-1} = D_k$$

(2.9)

which is the vector of distances from each node to destination as we saw in the discussion of the centralized algorithm.

In the asynchronous distributed version of the algorithm, it is assumed that each node $i$ keeps at time $t \geq 0$ an estimate of its distance to destination $A_i(t)$, and an estimate of the distance from each of its neighbors $j \in N(i)$ to destination $A'_j(t)$, which is simply the latest estimate received from node $j$. In view of (2.8) and (2.9), convergence of the algorithm will follow if we show that for every index $k$, there exists a time $t_k > 0$ such that for all $t \geq t_k$

$$D_k \leq A_i \leq D_k$$

(2.10)

and for $i = 1, \ldots, N - 1$

$$D_k(\cdot) \leq A'_i(\cdot) \leq D_k(\cdot) \quad j \in N(i).$$

(2.11)

This is shown by induction. If $k = 0$, then (2.10) and (2.11) hold as long as the initial estimates of the decentralized algorithm are nonnegative. Assuming that the induction hypothesis is true for $k$, the monotonicity of $B_i$ implies that if $t \geq t_k$, then

$$D_{k+1}(i) = B_k[D_k(\cdot)] \leq B_k[D_k(\cdot)] = D_{k+1}(i).$$

(2.12)

But $A_i(i)$ is a piecewise constant function of time which only jumps at the updating times of node $i$, at which times it takes the value

$$A_i(i) = B_i[A'_i].$$

Therefore, we can write

$$D_{k+1}(i) \leq A_i(i) \leq D_{k+1}(i) \quad \text{for } t \geq t_k$$

(2.13)

where $t_k$ is the smallest updating time of node $i$ which is greater than $t_k$. Moreover, if we wait long enough after $t_k$, not only all the nodes will have carried out their first updates after $t_k$, but the result of those computations will have been communicated to their neighbors because of the assumption that updating and communication occur infinitely often. Hence, there exists $t_{k+1} \geq t_k$ such that for all $t \geq t_{k+1}$ and for all $i$ and $j$

$$A'_i(j) = A_i(j)$$

for some $s \geq t_k$ (which depends on $t, i,$ and $j$). Thus, it follows from (2.13) that

$$D_{k+1}(j) \leq A'_i(j) \leq D_{k+1}(j) \quad j \in N(i) \quad i = 1, \ldots, N-1$$

completing the induction proof and, therefore, the proof of convergence of the distributed asynchronous Bellman-Ford algorithm.

When the link delays depend on the traffic flows, it is also possible to obtain the optimal routing that solves (2.2) in a distributed asynchronous fashion. Gradient projection algorithms are better suited for this task than the flow deviation method because in the latter method a higher degree of synchronization is required in order for the nodes to use the same step size at each iteration. In the distributed asynchronous implementation of gradient projection optimum routing algorithms, each node broadcasts from time to time the values of its outgoing flows to its upstream neighbors, who in turn pass that information on to their upstream neighbors. In this way, the source keeps estimates at all times of the link flows and can carry out the gradient projection iteration autonomously based on those estimates. The first algorithm based on this idea was due to Gallager [16], who posed an alternative formulation to (2.2), where the unknowns are the fractions of flow routed to each outgoing link at each node, rather than the path flows. Tsitsiklis and Bertsekas [42] showed the convergence of the distributed asynchronous implementation of gradient projection optimal routing algorithms provided the time between consecutive broadcasts is small enough relative to the speed at which the flows generated by the algorithm change. The approach for showing the stability of this algorithm is very different from the proof of convergence of the distributed Bellman-Ford algorithm where the monotonicity of the dynamic programming mapping implied that the estimates are closer and closer to the solution regardless of the actual sequence of communication and computation times. The idea here is that if the step size of the algorithm is small enough, then the flows change so slowly with respect to the periods between communication times that their evolution is very close to that of the centralized algorithm which uses the unique, true value of each link flow.

B. Dynamic Routing

As mentioned earlier, there are two fundamentally different philosophies to network routing: either viewing it as a "flow" problem in which the traffic of messages is modeled as a "macro"-commodity entering the network as a single entity (static or quasi-static routing), or as an individualized-message path-finding problem in which the traffic is broken down to its constituent elementary units (dynamic routing)—a dichotomy akin to that of statistical/quantum mechanics in physics. Whereas the first approach leads to optimization problems where time plays no role, the essential ingredient of the second approach is the randomness of the time-evolution of the buffers in the network, thus placing dynamic routing within the sphere of stochastic control.

The most elementary instance of dynamic routing is the simple queueing system shown in Fig. 4 which models a node with one incoming link and two outgoing links. It simplifies considerably the dynamics of the message arrival process and of the service time characteristics and ignores processing delay. Thus, the arrival instants of messages over the incoming link are assumed to constitute a Poisson process of constant rate $\lambda$. Upon arrival each message is put in the buffer of one of the two outgoing links. This action represents the "control." The buffers are assumed to have unlimited (infinite) capacity and the message lengths are assumed to be random with exponential distribution (an obvious additional simplification) with parameter $\mu$. The two outgoing links have equal capacity of $C$ bits/s. Thus, each link is modeled as a queueing system with exponential service time distribution with parameter $\mu C$. It is desired to characterize the optimal control policy that minimizes the average total delay per message based on the observations of the "state" of the system, namely the number of messages $q_1$ and $q_2$ in the two buffers. The model, of course, assumes that the head-of-the-line message is dropped from the buffer as soon as the transmission of its last bit is completed.

This model, despite its simplicity, proved to be rather difficult to analyze. For details, see [10], it is not important to repeat them here. It should suffice to state that the main result, which simply requires that upon arrival a message should join the shortest queue (with arbitrary decision in case the two queues have equal numbers of messages), was hardly surprising. Yet an intricate
argument on the dynamic programming equation (DPE) was needed and there were some counter-intuitive side-results including the relaxation of the Poisson assumption on the arrivals, and the fact that in the incomplete state information case, the certainty-equivalent control (i.e., send the message to the expected shortest queue) need not be optimum unless both queues save the same number of customers initially.

The optimality of the send-to-shortest-queue (SS) policy in the state information case can be proved in a rather strong sense. At all times, the sum \( q_1 + q_2 \) and maximum \( \max \{ q_1, q_2 \} \) of the number of messages in both buffers are stochastically minimized by the SS policy in the sense of the partial order between random variables according to which the random variable \( q \) is stochastically smaller than \( q \). If \( P[X \leq a] \geq P[Y \leq a] \) for all \( a \). The proof of optimality can be obtained by the method of forward induction [53], whereby the desired stochastic ordering between the queue sizes under the optimum and an arbitrary policy is shown to be preserved at each transition.

The problem formulation of [10] is one of many related ones (see [8], [9], [22], [24], [33], [38], [54], [55]) which are slightly more complicated but share some fundamental characteristics. In fact, extend beyond the confines of the routing problem into the areas of priority assignment, resource allocation, and flow control. They are all Markovian decision process (MDP) problems. In the sequel we will describe a fairly general MDP that includes the dynamic routing problem as a special case. In fact, it includes almost all of the queueing control problems that have been studied in connection with communication network issues. We will then outline the solution methodologies that have been used. These include basically: 1) the derivation of optimality conditions from the DPE associated with the corresponding MDP; 2) the use of simple path stochastic dominance arguments, and finally, 3) the reformulation of the MDP as a linear program. We would emphasize, lest the reader be unduly encouraged, that the problems in this area are sufficiently complex so that only modest results can be generally obtained despite involved arguments and contrivance machinery. Typically, these results characterize some structural properties of the optimal policy. However, knowledge of such structure is often sufficient tosnber close approximation to the actual optimal policy by well-founded heuristics.

Let us recall briefly what an MDP is (for details, see [30]). We need a state description of the process to be controlled. Let \( S \) be its state space. When in state \( s \in S \), a set \( A_s \) of admissible control actions is specified. When action \( a \in A_s \) is applied, there is a transition from state \( s \) to \( s' \) that is governed by the probability distribution \( p(s'|s,a) \), which occurs after a random time \( t \) which is exponentially distributed with distribution denoted by \( (t|s,a,s') \). Clearly, \( p \) and \( t \) together describe the stochastic dynamics of the process to be controlled. Finally, each transition is accompanied by a cost penalty that we denote by \( c(s,a,s') \).

The dynamic routing problem we considered before fits in this formulation easily. In that case, the state space is \( S = \{ 0, 1, 2, 3, \ldots \} \). An element \( s = (q_1, q_2) \in S \) is simply the pair of values of the respective queue sizes. The set of actions \( A_s \) is the same for any state and consists of \( a_1 \) and \( a_2 \) where \( a_i \) is the action that permits an arriving message to the buffer of link \( i \). The probability \( p \) is of trivial form, in that the transitions are deterministic. Assignment of an arrival to queue \( i \) augments \( q_i \) by one. Note, now, that in addition to the arrival instants, the departure (or service completion) instants are important because two induce state transitions as well. A departure from queue \( i \) causes \( a_i \) by one. When a departure occurs there is no meaningful control action that can be applied in this particular problem. The exponential distribution \( t \) corresponds to times between arrivals and/or departures. Finally, the cost rate \( c \) must reflect the delay. By Little's result in queueing theory, we know that the average delay is proportional to the average number of customers in the queue. Thus, \( c(t, a, s, s') \) can be taken to be simply equal to \( (q_1 + q_2) \). This MDP formulation can be extended to encompass more complicated queueing control problems.

Let us return now to the general MDP. We need to specify the notion of a control policy and the optimization criterion. Let us denote by \( \xi_1, \xi_2, \ldots \), the state transitions that occur at instants \( t_1, t_2, \ldots \). A policy \( \pi \) is a sequence of decision rules \( \pi_1, \pi_2, \ldots \), where \( \pi_x \) determines the choice of action at the transition time \( t_x \). It can be viewed as a conditional distribution on the set of actions parameterized by the past history of the process.

The optimization criterion that corresponds to the practical case of expected total delay is the long-run average expected cost; namely, if we denote by \( V(\pi, i, t) \) the expected cost incurred under policy \( \pi \), with initial state \( i \), until time \( t \) we consider as the optimization criterion the value function

\[
V(\pi, i, t) = \lim_{T \to \infty} \inf_{\pi} \frac{1}{t} \sum_{t=1}^{T} c(i, a, i') + \beta(i, a, i')V(\pi(i'))p(i'|a, i)
\]

For technical reasons, however, that are well known to optimization specialists, it is easier to establish optimality conditions if we consider, instead, the so-called \( \alpha \)-discounted cost, i.e.,

\[
V_\alpha(\pi, i, t) = \sum_{t=0}^{\infty} e^{-\alpha t} c(i, a, i') + \beta(i, a, i')V_\alpha(\pi(i'))p(i'|a, i)
\]

The latter converges to the former as \( \alpha \to 1 \) under a variety of stationarity conditions. For technical reasons that will become apparent in the sequel, we will also consider the finite-horizon costs. These are defined in a similar fashion except that we let time extend only to \( t \), the instant of the \( n \)th transition. If we denote by \( V_\alpha(i) \) and \( V_\alpha(i) \) for the finite horizon cases the values of these cost functions when \( \pi \) is chosen optimally, we are led to the following DPE:

\[
V_\alpha(i) = \inf_{\pi \in A} \sum_{t} c(i, a, i') + \beta(i, a, i')V_\alpha(i')p(i'|a, i)
\]

or

\[
V_\alpha(i) = \inf_{\pi \in A} \sum_{t} c(i, a, i') + \beta(i, a, i')V_\alpha(i')p(i'|a, i)
\]

where

\[
\beta(s, a, s') = \int_{0}^{\infty} e^{-\alpha t} dt(s, a, s')
\]

and

\[
c(s, a, s') = \int_{0}^{\infty} c(r(s, a, s') dt(r(s, a, s'))
\]

are the discount factor and cost functions per transition, respectively.

The DPE is of fundamental importance in the study of MDP's because the value function \( V_\alpha \) has the usually convenient properties of convexity, supermodularity, and other forms of monotonicity that lead readily to sufficient conditions for optimality. The difficulty with the analysis of the DPE is that the optimality conditions are heavily problem-dependent and often lead to explosively large numbers of cases to be verified separately. This is especially true for MDP's that arise from queueing models. For this reason, and because of additional difficulties that arise when the state is on the boundaries (see [22]), it became evident that alternative methods of solution were needed.

A slight modification of the model of transitions, called uniformization, is useful in that it introduces dummy transitions from a state into itself, thus, some transitions which introduce conceptual complications can be handled variously from this discrete transition time formulation.
One alternative method that has received attention recently and which produced successful results in problems of queueing control (akin to the routing problem) is a probabilistic method called sample-path or stochastic dominance. This method bypasses completely dealing with the value function. Instead, it focuses directly on seeking the optimal policy. Let \( G \) be the class of admissible policies. If we suspect that the optimal policy \( \pi \) has a property \( p \), then we can proceed as follows in order to prove that it actually does have that property. Let \( S \) be a subset of \( G \), to which we know the optimal policy belongs. We consider a subset of polices \( S \subset S \), all elements of which have the property \( p \). For every \( \pi \in S \), we attempt to construct a policy \( \tilde{\pi} \) which outperforms \( \pi \). If we succeed, we must conclude that the optimal policy belongs to \( S \). In constructing \( \tilde{\pi} \) we often need to engage in a careful reorganization of the underlying probability space in order to align the sample paths properly, so that the comparison of the two policies can be made for every sample path. This procedure is full of risks and extreme care is required to avoid faulty arguments. Note, also, that to apply this method usefully, we must have "guessed" the properties of the optimal policy correctly. Thus, at best, it is a method to verify the validity of our conclusions, rather than a method that leads us to the right conclusions.

Successful use of the stochastic dominance approach was made in [52] and [50] where a problem that is dual to the problem of dynamic routing was studied. Specifically, in a two-server queueing system in which the two servers have unequal service rates, we wish to determine whether and when the slower server needs to be activated if we are interested in minimizing the usual total expected delay function. That the optimal policy has a threshold form (namely that the slower server must be activated when the queue size exceeds a crucial value) was proven in [29] via the DPE method. However, the alternative proof via arguments of stochastic dominance was much simpler and led to a generalization of the result to cases of nonexponential arrivals and/or service, that could not have been easily accomplished by means of the DPE method.

Another successful use of the stochastic dominance method has been noted in [2]. In this case the problem of optimally choosing which customer to serve next in a single queueing system was considered under the constraint that each customer must begin (or terminate) service by an individually assigned random deadline or else it is dropped from the system. The cost criterion is then to minimize the expected number of lost customers. It was proven that scheduling the customer with shortest time to extinction and/or service, that could not have been easily accomplished by means of the DPE method.

Although these problems differ from routing, the model structures are quite similar, and it has been observed that, usually, queueing control problems with such structural similarities can be studied equally successfully.

The third method, which was first used in [38] in the study of a specific queueing control problem, and which has been broadly extended recently in [51], is the linear programming approach. Almost any queueing control problem that can be formulated as an MDP (therefore the problem of dynamic routing, as well) can be converted to an equivalent linear program (LP). The advantages of this conversion are that it is problem-independent and it leads occasionally to successful study of semi-Markov decision problems as well. Furthermore, it facilitates considerably the characterization of optimal solution properties. Here is how this equivalence can be demonstrated.

Let us concentrate on an MDP under a finite-horizon, discounted cost formulation. We shall consider a queueing model with state dynamics given by

\[
x_{k+1} = x_k + z_k, \quad k = 0, 1, \ldots, n.
\]

Here, \( x_k \) denotes the state at \( t_k \), the instant of the \( k \)th transition. \( z_k \) represents that transition, and \( x_k \) represents the control action at that transition. The transition \( z_k \) can represent an arrival or a departure as an increment of the state. The control \( x_k \) is conveniently defined to enable \( z_k = 1 \) or disable \( z_k = 0 \) a transition. For example, in the routing model discussed at the beginning of the section, the state is equal to a two-dimensional vector of queue sizes, and the transition corresponding to sending an arriving message to the first queue would be represented by \( z_k = 1 \). Indeed, a variety of queueing control problems (in fact, the vast majority of those that have been considered in connection with communication network problems) can be so represented.

Note that the crucial aspect of this state equation is the linear dependence on the controls. Note also that usually the cost function is linear in the state (since the usual cost criterion is the expected delay which is coupled to the queue sizes, and hence the state, by Little's result). Consequently, the cost is linear in the controls. The minimization of the cost over the set of control trajectories is constrained since the state equation must be satisfied and the state must always belong to an admissible set (typically, a set of vectors with integer-valued coordinates belonging to given ranges). Thus, the constraints are also linear in the controls, and the problem is easily formulated as an LP. There are, however, two points that require attention. First, the controls are integer valued, i.e., \( z_k \in \{0, 1\} \). Second, in the MDP the vectors \( z_k \) are random and depend on past history.

The first problem is taken care of in one of two ways: by construction or by use of a property of the constraint matrix of the linear program, called unimodularity. The construction method involves using a noninteger optimum control whose quantized version satisfies the MDP optimality conditions (see [38], [51] for details). The use of unimodularity involves a well-known result in the theory of integer linear programming (e.g., [34]); if the constraint matrix of an LP is integer valued and totally unimodular (i.e., each of its sub-determinants is \( +1, -1, \) or \( 0 \)), then all the vertices of the feasible polytope are integer valued. Therefore, no further restrictions are needed to guarantee that the solution of a conventional LP will result in the integer-valued optimal control. Fortunately, in many queueing problems of interest (including the dynamic routing problem), the constraint matrix is indeed totally unimodular.

The second problem is easily taken care of by thinking of the \( z_k \)'s as functions from the sample space \( \Omega \) to the action space. Thus, the cost criterion can be written as a functional on the underlying probability space.

Let \( z_k(\omega) \) represent the control action at the \( k \)th transition, where \( \omega \) denotes the random "history" until the \( k \)th transition. We have

\[
x_{k+1}(\omega_{k+1}) = x_k(\omega_{k+1}) + z_k(\omega_{k+1}), \quad k = 0, 1, \ldots, n - 1.
\]

Let \( S \) and \( Z \) be the set of admissible states and controls, respectively. The \( \beta \)-discounted, \( n \)-step, expected cost under policy \( z \) and initial condition \( x \) is given by

\[
J_S^n(x, Z) = E_z \left( \sum_{k=0}^{n-1} \beta^k L(z_k) \right)
\]

where

\[
L(z_k) = c^T x_k + d^T z_k
\]

(\( c \) and \( d \) denote constant column vectors). This is a cost function that is adequately general. For example, in a pure resource allocation problem without blocking or rejection of messages we have \( d = 0 \), while in pure blocking problems we take \( c = 0 \). The state equation, after repeated iterations, yields

\[
x_k(\omega_k) = x + \sum_{j=1}^k z_j(\omega_j) \xi_j(\omega_j), \quad k > 0.
\]
Therefore,

\[ J_Z^T(x, z) = E_x \sum_{k=1}^{\infty} \beta^k \left\{ c^T x + c^T \sum_{j=1}^{z} \xi_j + d^T Z_k \right\} \]

\[ = \frac{1 - \beta^x}{1 - \beta} c^T x + E_x \sum_{k=1}^{\infty} \beta^k \left\{ c^T \xi_k + d^T Z_k \right\} . \]

But

\[ E_x(x_k) = \sum_{\omega_k} z_k(\omega_k) Pr(\omega_k). \]

Hence

\[ J_Z^T(x, z) = \frac{1 - \beta^x}{1 - \beta} c^T x + \sum_{k=1}^{\infty} \gamma_k(\omega_k) z_k(\omega_k) \]

where \( \gamma_k(\omega_k) \) is a known function that depends on \( Pr(\omega_k), c, \xi_k, \) and \( \beta^x \). Consequently, the MDP is equivalent to

\[ \min \sum_{k=1}^{\infty} \sum_{\omega_k} \gamma_k(\omega_k) z_k(\omega_k) \]

subject to

\[ (x + \sum_{j=1}^{z_k} \xi_j(\omega_j)) \in S \]

which is a conventional LP where the initial condition plays the role of a parameter, the sensitivity with respect to which can be studied by the well-developed theory of sensitivity analysis of linear programming [15].

In conclusion, we see that the MDP is converted to an equivalent LP under very mild conditions that are usually satisfied by dynamic routing and other queueing control problems. Thus, a third alternative methodology becomes generally available for the study of these problems. Whether to choose from the arsenal the DPE approach, or the LP method, or stochastic dominance tools, depends on the problem and on the, as yet undeveloped, intuition that the investigator should possess.

III. MULTIPLE-ACCESS COMMUNICATIONS

The communication networks considered in the discussion of outing problems in Section II consist typically of a set of nodes connected by point to point communication links. Each of these links viewed in isolation can be modeled as a classical communication channel with one sender and one receiver. In this section, we consider multipoint to point communication links where several transmitters share a common channel. Multiple access channels are the basic building blocks of radio networks, satellite communication, and local area networks, and during the last 15 years have attracted the attention of many communication, information, and control theorists.

There is a wide variety of strategies to divide the "resource" of a communication channel among several geographically separated transmitters. The simplest methods are those that assign a permanent independent sub-channel to each transmitter (e.g., in frequency division multiple access and time division multiple access); these strategies are easy to analyze and are widely used in practice in situations where the users need to transmit at fairly low rates. If the transmitters are bursty (i.e., the radio of peak average rate at which the need to transmit is high) those static methods are inefficient since most of the time the channel is unutilized while demand (and induced delay) accumulates at busy terminal locations. Dynamic channel sharing strategies overcome this problem by allocating channel resources on an on-demand basis. Consistent with the overall spirit of this paper, our goal here is not to review this vast topic, but rather to demonstrate how control theory can play a useful role in its study. Here we wish to single out two multiple access strategies: random access and simultaneous transmission, which are broadly representative of dynamic channel sharing systems and in which control theoretic concepts have played a pivotal role.

In random access communication, the conceptual allocation model is addressed without an effort to exploit the signaling degrees of freedom and the micro-structure of the transmitted messages. For this purpose, a crude channel model is considered, that achieves this separation of the "macro" from the "micro" problem. In simultaneous transmission systems, however, a more refined viewpoint is adopted, by taking the realities of the medium into account, modeling them, and exploiting them.

A. Random-Access

The object of interest here is the so-called collision channel model, in which messages (called packets) require one time unit (called slot) for transmission and are sent by a population of users who are synchronized so that their slots coincide at the receiver, but are otherwise uncoordinated and unaware of which and how many users have packets to transmit. If two or more packets are simultaneously transmitted, it is assumed that the receiver is unable to recover any of the messages, and they have to be retransmitted in a future slot. In the ALOHA algorithm, which was developed in the early 1970's [1] at the University of Hawaii and marked the beginning of the area of random-access communication, each packet that has been unsuccessfully transmitted before is transmitted with probability \( p \) in the next slot. New packets which have not attempted transmission before are transmitted with probability either \( 1 - p \) or \( p \) depending on which version of the ALOHA algorithm is used. In our discussion, we will assume the latter choice.

Under these conditions, and assuming that the number of newly generated packets in each slot is a random variable \( \lambda \) independent from slot to slot, the number of packets awaiting transmission (called backlog) is a Markov chain taking values in \( \{ 0, 1, 2, \ldots \} \). The central problem is to investigate under what conditions the backlog Markov chain is ergodic, i.e., it is stable in the sense that it reaches a steady state in which the periods between the times when there are no packets to transmit are not too infrequent (they have finite expected value). The transition probabilities of the Markov chain are parametrized by the rate of arrival of new packets \( \lambda \) and the retransmission probability \( p \).

Whereas \( \lambda \) is fixed and given, \( p \) is chosen by the transmitters. Hence, we are dealing with a fairly simple controlled Markov chain whose control space is the interval \( (0, 1] \). In the original ALOHA algorithm, the control \( p \) remained constant and common to all transmitters regardless of the information acquired by listening to the channel, thereby resulting in the open-loop control of the Markov chain. Despite several "proofs" of the stability of ALOHA published during the 1970's, neither the actual system built in Hawaii nor the ideal Markov chain model were stable. The reason why the open loop system is unstable can be easily understood by considering the backlog drift, \( d(n) \), which is defined as the expected increase in the backlog over the next slot when the current value of the backlog is equal to \( n \). It is easy to see that the backlog drift is given simply by the expected number of new packets per slot minus the expected number of successfully transmitted packets in the next slot, i.e.,

\[ d(n) = \lambda - n(p(1-p))^{n-1}. \]  

The drift quantifies the expected evolution of the Markov chain from each state, and therefore it is a valuable tool in analyzing the stability of the chain. For any \( p \in (0, 1) \) the term in brackets in
(3.1) goes to 0 as \( n \to \infty \), and hence, the drift is positive and close to \( \lambda \) for sufficiently large backlogs. This implies that when the backlog is large it tends to grow, thereby eliminating any hope for stability. Using standard results, this reasoning can be formalized straightforwardly to prove not only the instability of the open-loop system \([11]\) for all values of \( \lambda \) and \( \rho \), but the fact that the backlog goes to infinity with probability one \([25, 35, 40]\).

Fortunately, the system can be stabilized by closed-loop control. Let us examine first the case of complete-state information, i.e., each station is informed at the end of each slot of the current value of the backlog and chooses the retransmission probability on the basis of that information. As far as stability is concerned, the best choice of the retransmission probability \( p \) is the value that minimizes the drift because that results in the maximum possible arrival rate that guarantees stability (called the throughput). It follows from (3.1) that the optimum value of \( p \) is

\[
\rho^* (n) = \frac{1}{n}, \quad n = 1, 2, \ldots
\]  

(3.2)

and the resulting drift is

\[
d^*(n) = \lambda - \left[ 1 - \frac{1}{n} \right] n - 1
\]  

(3.3)

which is negative for \( n > 1 \) when \( \lambda < e^{-1} \), and is positive for large backlogs when \( \lambda > e^{-1} \). Therefore, the throughput of the closed-loop system with complete state information is \( e^{-1} = 0.368 \). However, the relevance of complete state information feedback is rather limited in practice. This is because the instantaneous value of the backlog is available to each station only if there exists so large a degree of communication among the transmitters that much more efficient algorithms than ALOHA can be used.

The case of partial state information is the problem of interest in practice, since the only feedback available to each station is the outcome (collision, success, empty) of the transmission in each slot. The analysis of the controlled system with partial state information was pioneered by Hajek and Van Loon \([20]\) who proposed a recursive updating law of the retransmission probabilities as a function of the channel outcomes. This feedback policy was shown in \([21]\) to attain the throughput achievable with complete-state information, namely \( e^{-1} \). Those papers and subsequent works have referred to the problem as decentralized control of ALOHA, motivated by the fact that each station chooses the retransmission probability autonomously based on the channel feedback. However, it is useful to recognize that the problem boils down to (centralized) stochastic control with one decision maker and incomplete state information because all stations are constrained to use the same retransmission probabilities.

We will review here the proof of stability of the following certainty-equivalence closed-loop control:

\[
\rho (\hat{n}) = \frac{1}{\hat{n}}
\]  

(3.4)

where \( \hat{n} \) is an estimate of the backlog updated according to

\[
\hat{n}_{k+1} = \begin{cases} 
1, & \text{kth slot is idle} \\
\hat{n}_k + \beta, & \text{kth slot is success or collision} 
\end{cases}
\]  

(3.5)

The throughput attainable with this feedback law depends on the constants \( \alpha < 0 \) and \( \beta > 0 \). As we will see, there exists a set of choices for those constants that results in throughput equal to \( e^{-1} \).

Unlike the case of complete-state information, the proof of stability is not straightforward because now it is the two-dimensional process formed by the backlog and its estimate \( \{(n_k, \hat{n}_k)\} \) (rather than the backlog itself) which is a Markov process. According to (3.4) and (3.5) the drift of this Markov process is given by

\[
E ((n_{k+1}, \hat{n}_{k+1}) - (n_k, \hat{n}_k)) = (n, \hat{n})
\]  

(3.6)

Conversely, to what we saw in the case when the state is known, it is not true that the backlog drift is negative for sufficiently large backlogs. As we can see in Fig. 5, if the estimate is far from the true value, then the backlog may actually increase.

However, at every point in the state space the tendency of the process is to approach the diagonal where the estimate is equal to the true value of the backlog. Furthermore, as Fig. 5 or the analysis of the perfect-state information case shows, the drift along the diagonal is negative. Such a behavior is a strong indication of the stability of the controlled Markov process. This can be proved using a powerful sufficient condition found by Mikhailov \([31]\) for the stability of a Markov process taking values in \( \mathbb{R}^+ \times \mathbb{R}^+ \). In essence, Mikhailov's condition states that it is enough to restrict attention to those points of the state space where either the backlog or its estimate are large and at which the drift is radial, i.e.,

\[
\frac{d(n, s)}{c(n, s)} = \frac{n}{s};
\]

then, it is sufficient for stability that the drift point towards the origin at those states. To see that this condition is indeed satisfied for our system, we compute first the asymptotic drifts along the radius \( \{(n, s); n/s = \psi \} \) for \( \psi \in [0, \infty) \):

\[
d(\psi) = \lim_{s \to \infty} d(\psi, s) = \lambda - \psi e^{-\psi}
\]  

(3.7a)

\[
c(\psi) = \lim_{s \to \infty} c(\psi, s) = \beta + (\alpha - \beta) e^{-\psi}
\]  

(3.7b)

It can be checked using (3.7) that if the constants \( \alpha \) and \( \beta \) in (3.5) are chosen such that \( \beta > 0.23 \) and \( \lambda - e^{-1} = \beta + (\alpha - \beta) e^{-1} \), then the drift is radial only at \( \psi = 1 \) (cf. Fig. 5), where it points towards the origin as long as \( d(1) = \lambda - e^{-1} < 0 \). Mikhailov's sufficient condition can be justified constructing a
stochastic Lyapunov function to prove the stability of a Markov process \( \{ x_t \} \) with state space \( \mathbb{Z}^d \times \mathbb{Z}^d \). To that end, it is advantageous to switch to polar coordinates \((r, \phi)\) and to define the radial drift \( \delta(r, \phi) \) as the projection of the drift along the direction of the point \((r, \phi)\) and the tangential drift \( \mu(r, \phi) \) as the projection of the drift along the direction perpendicular to \((r, \phi)\). Denote the asymptotic drifts \( \delta(\phi) = \lim_{r \to \infty} \delta(r, \phi) \) and \( \mu(\phi) = \lim_{r \to \infty} \mu(r, \phi) \) and define the function

\[
V(r, \phi) = r \delta(\phi)
\]

where

\[
\Phi(\phi) = \exp \left( -C \int_0^\phi \mu(u) \, du \right) \quad \phi \in \left[ 0, \frac{\pi}{2} \right].
\]

Note that \( V(r, \phi) \) is a candidate Lyapunov function because it is positive outside the origin and \( V(r, \phi) \to \infty \) as \( r \to \infty \). Furthermore, it can be shown [31] that the asymptotic drift of the candidate Lyapunov function is equal to

\[
\lim_{r \to \infty} E[V(x_{t+1}) - V(x_t)] = \Phi(\delta(0) - C \mu^2(0)). \quad (3.8)
\]

Now, under Mikhailov's condition, the asymptotic drifts are assumed continuous on \([0, \pi/2]\) and \( \delta(\phi) < \epsilon \) for any phase such that \( \mu(\phi) = 0 \) (i.e., whenever the drift is radial it points towards the origin), therefore, the constant \( C \) can be chosen large enough so that the left side of (3.8) is upper bounded by a negative constant. This implies that \( V(r, \phi) \) is indeed a stochastic Lyapunov function and therefore standard results on the stability of stochastic systems [27], [45] can be applied to show the stability of the system.  

In some multiaccess environments, the receiver can indeed demodulate reliably one or more packets even in the presence of other interfering packets and the collision channel model no longer applies to those cases. The results reviewed in this section can be generalized to a general channel with multipacket reception capability, to show that: 1) the throughput of open-loop ALOHA is equal to the limit of the expected number of successfully received packets per slot as the backoff goes to infinity [17]; and 2) the throughput of closed-loop ALOHA (with either complete or partial state information) is equal to the maximum over \( \epsilon \) of the expected number of successfully received packets per slot when the number of attempted transmissions is a Poisson random variable with mean \( \mu \) [18].

Returning to the case of the collision channel, the next natural step is to drop the main restriction in the ALOHA algorithm, namely, that all stations use the same retransmission probability. This is done in a class of random-access algorithms referred to as collision resolution algorithms which are characterized by the fact that not only are all blocked packets eventually retransmitted successfully, but all users eventually become aware that these packets have been successfully retransmitted. Contrary to the ALOHA algorithm, the decision whether or not to transmit a packet takes into account the previous history of attempted retransmissions of that particular packet. The introduction of this new dimension into the problem renders Markov chain tools considerably less useful than in the foregoing analysis and converts it into a very difficult decentralized stochastic control problem, for which the optimum throughput remains unknown despite many efforts.

4 Another choice of stochastic Lyapunov function for the specific case of decentralized control of ALOHA can be found in [43].

5 The best known algorithm has been shown to achieve a throughput of 0.488 using Howard’s policy iteration for sequential infinite-horizon problems [32] or by reduction to a simple optimization problem [48]. On the other hand, it is known that the optimum throughput is upper bounded by 0.568 [44].

B. Simultaneous Transmission

In contrast to random-access communication systems, in simultaneous transmission multiple-access systems, the transmitters send their messages simultaneously, independently, and without monitoring the channel in any way. The most common type of simultaneous transmission system is code-division multiplexing, where each user modulates a preassigned signature waveform known by the receiver.

Specifically, we will assume that in order to send the message \( \{ b_i(t) \} \in A \) (i.e., a string of \( M \) symbols drawn from a finite set \( A \)), the \( k \)th user transmits

\[
\sum_{i=0}^{M-1} b_k(i)s_k(t-iT)
\]

where \( s_k(t), 0 \leq t \leq T \) is the waveform assigned to the \( k \)th user, and \( T \) is the symbol period. Then the demodulator receives the sum of the signals transmitted by the \( K \) active users embedded in noise

\[
r(t) = \sum_{k=1}^{K} \sum_{i=0}^{M-1} b_k(i)s_k(t-iT) + n(t)
\]

and the multiuser signal in (3.8)

\[
S(t, d) = \sum_{k=1}^{K} \sum_{i=0}^{M-1} b_k(i)s_k(t-iT) = \sum_{i=1}^{ME} d_i z_i(t)
\]

where \( z_{k, d}(t) = s_k(t-iT-\tau_k) \).

A reasonable criterion for demodulating the information carried in \( S(t, d) \) upon observation of \( r(t) \) is to select the optimum vector \( d \) that best explains the received waveform in the sense of minimizing the energy of the corresponding noise realization, i.e.,

\[
\min_{d \in A^K} \| S(t, d) - r(t) \|^2. \quad (3.10)
\]

If the noise \( n(t) \) is white and Gaussian, then this criterion results in maximum likelihood decisions. Equivalently, the objective is to find the vector that solves

\[
\max_{d \in A^K} Q(d) \quad (3.11)
\]

where

\[
Q(d) = 2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} S(t, d) r(t) \, dt - \int_{-\infty}^{\infty} S(t, d) \, dt. \quad (3.12)
\]

Since the maximization in (3.11) is over a finite set, we could solve it by brute-force method of evaluating \( Q(d) \) for each possible argument. However, it is possible to decompose \( Q(d) \) in a sequential fashion that lends itself to efficient optimization. From
(3.9) it is immediate to write the first integral in (3.12) sequentially

\[ \int_{-\infty}^{\infty} S(t, d) r(t) dt = \sum_{i=1}^{2K} d_i \gamma_i, \quad (3.13) \]

where

\[ \gamma_i = \int_{-\infty}^{\infty} z(t) r(t) dt. \quad (3.14) \]

This implies that the objective function (3.12) depends on \( r(t) \) only through the quantities \( \{ \gamma_i \}_{i=1}^{2K} \), which are obtained by correlating \( r(t) \) with each of the signature waveforms during each symbol epoch. In order to find an explicit expression for the second integral on the right-hand side of (3.12), which is the first integral in (3.15), we have

\[ R(j, l) = \int_{-\infty}^{\infty} z(t) r(t) dt. \quad (3.15) \]

It follows immediately from the definition that these coefficients satisfy the following properties.

1. \( R(k + iK, k + iK) = \{ \delta(t) \} \equiv \delta_k. \)
2. \( R(k + iK, n + iK) = R(k, n) \) for all \( i \).
3. \( R(j, l) = 0 \) unless \( j - l < 0 \).

The first property indicates that each of the diagonal elements of \( R(i, j) \) is equal to the energy of one of the \( K \) assigned waveforms. The second and third properties can be illustrated by referring to Fig. 6 which represents the symbol epochs of three asynchronous users sending strings of \( M = 5 \) symbols. Each symbol period in Fig. 6 is labeled with the index of the corresponding component of the vector \( d \). The second property indicates that the cross-correlations between two signals depend on their relative location (e.g., \( R(4, 6) = R(13, 15) \) in Fig 6) and the third property states that each symbol only interferes with \( 2K - 2 \) symbols of the other users (e.g., in Fig. 6, \( d_5 = b_5(2) \) only overlaps with \( d_1 = b_1(2), d_2 = b_2(2), d_0 = b_0(3), \) and \( d_4 = b_4(3) \)). It follows from these properties that the coefficients in (3.15) can be obtained from the \( K \times K \) matrix \( R(k, n) \) whose diagonal elements correspond to the energy per symbol of each user and whose off-diagonal elements correspond to the cross-correlations between the signature waveforms of each pair of users. Using (3.15), the foregoing properties, and letting \( \kappa(j) \in \{ 1, \ldots, K \} \) be the modulo-\( K \) remainder of \( j \) (i.e., for some \( i, j = \kappa(j) + iK \)), we can write

\[ \int_{-\infty}^{\infty} S(t, d) r(t) dt = \sum_{j=1}^{MK} \sum_{l=1}^{MK} d_j d_l R(j, l), \quad (3.16) \]

where

\[ \Omega(d) = \sum_{j=1}^{MK} \lambda_j \sum_{\gamma_j = 1}^{d_j} R(j, l) \]
A. Flow Control

A stark reality in the design of networks is that despite the reduction of the cost of memory, storage at each node is going to be finite. Coupled with another reality, namely that data transmissions on the whole continue to be bursty, it implies that buffer overflow may occur and, along with it, congestion and deadlocks. Flow control is the name we use to describe the collection of measures taken to avoid buffer overflow and highly congested nodes in the network. Congestion and saturation are often the consequences of diverging, unstable behavior. Thus, it is of interest not only to optimize over possible flow control strategies, but to determine their robustness against disturbances or modeling inaccuracies that may lead to unstable behavior.

The control variables in flow control problems are admission (or blocking) probabilities for messages or sessions at the source node. In practice these are often implemented in terms of a bang-bang control strategy known as window flow control whereby input ports are allowed to continuously inject messages into the network at the full desired input rate until the number of unacknowledged messages exceeds the value of the "window size" W. A simple, yet unanswered question is, what should the value of W be?

Previous efforts to use control theory tools to analyze optimal flow control problems include [28] and [46] where the optimality of window flow control is proved within the domain of a simplified model, and [39] where dynamic programming value iteration techniques are used to characterize optimal flow control performance. An alternative approach to the flow control problem is to subsume it into the static routing problem considered in Section II-A [19]: suppose that for every source-destination pair a fictitious direct link is added between them. We can then interpret the blocking action of a flow control procedure as a diversion of the blocked portion of the traffic through this fictitious link to the destination. Thus, we can consider that no traffic is blocked. Of course, in order to discourage the use of this fictitious link we must augment the overall delay cost function with a term that penalizes appropriately the use of this link.

B. Integrated Switching

A revolutionary development in the field of networks whose implementation is currently under way is the combination of the capabilities of what have been separately developed in the past and called voice networks and data networks. Voice is a commodity that must meet different requirements than data. For example, speech signals have inherent redundancy that make them quite robust with respect to occasional errors or deliberate compression. At the same time, except in applications of voice messaging, speech signals occur in the context of real-time conversations and, as such, must encounter short and, more importantly, constant delay. On the other hand, data must preserve their integrity and cannot tolerate errors; however, long and variable delays can be often tolerated.

How does one design a single network that can handle such dissimilar commodities with automated procedures? The natural course of events in the last decade or two was to attempt to force data on primarily voice networks or to let voice ride on what were mainly data networks. The literature is full of ideas for baseline integration that are mostly heuristic and difficult to analyze. An attempt to formulate the problem of integrated switching as an optimization problem was presented in [50]. In its simplest form, the model is as follows: consider a single node in the network with a single outgoing link on which incoming voice calls and data packets must be multiplexed. Let W be the bandwidth of the outgoing link. Let V be the bandwidth required for the continuous, uninterrupted accommodation of a single voice call. Let, therefore, N = W/V be the maximum number of calls that can be assigned dedicated circuits simultaneously if no data packets are transmitted. A voice call can either be accepted (and assigned the necessary bandwidth V) or blocked. Data packets can be stored in a buffer if available. If, at a given time, there is an i calls in the system, the data packets can be served at the full rate corresponding to the remaining bandwidth W - iV. Such a switching architecture represents what has been called the movable boundary idea in integration. A natural MDP can be simply formulated as follows: choose the control action of blocking or accepting a call upon arrival in order to minimize the weighted sum of the average data packet delay and the call-blocking probability. If we assume that both arrival streams (voice calls and data) are independent Poisson processes, that the call holding time is exponentially distributed, and that the message lengths are likewise exponential, we can apply the technique described in Section II of converting the MDP to an LP and show that the optimal policy has the useful switching-type form. Namely, if I is the number of ongoing calls and J the total number of data messages at the node, the optimal control action should be to block the call in region B of the state space as shown in Fig. 7 and to accept it in region A.

C. Link Scheduling

Let us now turn our attention back to the radio network environment. In Section III the multiple access channel was considered and a number of difficult but interesting control problems were identified. Throughout that discussion, it was assumed that all terminals are within a single transmission hop from the destination. In many radio networks, however, this is not the case. Messages need to be relayed via intermediate nodes to their final destinations. Thus, the familiar problem of routing arises again, except that this time there is a new twist to it. In point-to-point networks, transmissions between different node pairs can take place simultaneously because there are dedicated, "hard-wired" links between the corresponding nodes. In a radio (or, more generally, a multaccess/broadcast) environment, if the nodes are densely connected, not all transmissions can take place simultaneously (unless separate dedicated channels or simultaneous transmission signaling techniques (Section III-B) are used). They must be scheduled in time to avoid the interference that would occur otherwise.

It becomes evident that the mere fact that the transmission among a group of nodes must take place at a time raises the question whether the intended transmissions are routing-wise optimal any more. Several versions of this problem have been studied [3], [23], [36]. In every case and even if the routing problem is sidestepped, we are led to hard combinatorial optimization problems where questions of computational complexity and distributed implementation are of primary importance.

V. Conclusion

It should be clear by now that the theory of linear and nonlinear optimization, dynamic programming, stochastic control, stability
analysis, and distributed control have found interesting applications arising in the analysis and design of communication networks. Unlike other complex systems that have been successfully studied by control system theorists in the past (such as chemical plants, flexible aircraft, robot systems, etc.) communication networks stand out in that the commodity to be controlled is information (including its transmission, storage, processing, etc.). This feature, perhaps, misleads and intimidates those who do not feel sufficiently inter-disciplinarian to tackle these problems. We hope that by having selected to present a few examples in which concrete, purely control-theoretic problems can be formulated and have been (or can be) studied successfully, we may encourage attention by the control community to this application area that is especially rich in new challenges.

As stated from the outset, we did not attempt to survey or completely cover the multiple control facets of communication networks. The collection in this paper merely represents an effort to illuminate a few selected problem areas and to show how control techniques apply to them.

REFERENCES


Anthony Ephremides (S’68–M’71–SM’77–F’84) was born in Athens, Greece, in 1943. He received the Ph.D. degree in electrical engineering from Princeton University, Princeton, NJ, in 1971.

He has been with the University of Maryland, College Park, since 1971. He has also spent semesters on leave at M.I.T., the University of California, Berkeley, and ETH, Zurich. He is active in professional consulting as President of Pontos, Inc. Currently, his research interests lie in the areas of communication systems, performance analysis, modeling, optimization, simulation, and design.

Dr. Ephremides is Director of Division X of the IEEE and has served as President of the Information Theory Society and on the Board of Governors of the Control Systems Society. He has been Associate Editor of the IEEE TRANSACTIONS ON AUTOMATIC CONTROL and General Chairman of major IEEE Conferences.

Sergio Verdú (S’80–M’84–SM’88) was born in Barcelona, Catalonia, Spain, in 1958. He received the Telecommunication Eng. degree from the Technical University of Barcelona in 1980 and the Ph.D. degree in electrical engineering from the University of Illinois at Urbana-Champaign in 1984.

Upon completion of his doctorate he joined the faculty of Princeton University, Princeton, NJ, where he is currently an Associate Professor of Electrical Engineering. His current research interests are in the areas of multiuser communication and information theory.

Dr. Verdú is a recipient of the National University Prize of Spain, the Rheinstein Outstanding Junior Faculty Award of the School of Engineering and Applied Science at Princeton University, and the NSF Presidential Young Investigator Award. He is currently serving as Associate Editor of the IEEE TRANSACTIONS ON AUTOMATIC CONTROL, and as a member of the Board of Governors of the IEEE Information Theory Society.
Optimal Decentralized Control in the Random Access Multipacket Channel

Sylvie Ghez
Sergio Verdu
Stuart C. Schwartz

Reprinted from
IEEE TRANSACTIONS ON AUTOMATIC CONTROL
Vol. 34, No. 11, November 1989
Optimal Decentralized Control in the Random Access Multipacket Channel

SYLVIE GHEZ, STUDENT MEMBER, IEEE, SERGIO VERDÚ, SENIOR MEMBER, IEEE, AND STUART C. SCHWARTZ, SENIOR MEMBER, IEEE

Abstract—A decentralized control algorithm is sought that maximizes the stability region of the infinite-user slotted multipacket channel and is easily implementable. To this end, the perfect state information case where the stations can use the instantaneous value of the backlog to compute the retransmission probability is studied first. The best throughput possible for a decentralized control protocol is obtained, as well as an algorithm that achieves it. Those results are then applied to derive a control scheme when the backlog is unknown, which is the case of practical relevance. This scheme, based on a binary feedback, is shown to be optimal given some restrictions on the channel multipacket reception capability.

I. INTRODUCTION

Most studies on random access communications rely on the assumption that when two or more packets overlap, all the information that was sent is irretrievably lost, hence the need to repeat all transmissions at some later time. This is actually a pessimistic point of view, since there are many examples of random access systems where one or more packets may be successful in the presence of other simultaneous transmissions. In order to represent such random access systems, a model for a channel with multipacket reception capability has been developed in [6]-[8]. We consider a slotted channel with an infinite population of users, and we assume that the probability of having $k$ successes in a slot where there are $n$ transmissions depends only on the collision size $n$

$$e_{nk} = P[k \text{ packets are correctly received|} n \text{ transmissions}]$$

$n = 1, 0 \leq k \leq n$.

We define the reception matrix as

$$E = \begin{bmatrix}
e_{10} & e_{11} & 0 & \cdots \\
e_{20} & e_{21} & e_{22} & \cdots \\
e_{30} & e_{31} & e_{32} & \cdots \\
\vdots & \vdots & \vdots & \ddots
\end{bmatrix}$$

This model can be applied to channels with capture [1]-[3], [10], [16], [18], [20], [23], [26], [28], [34] and to systems using CDMA [22], [24], [29]. It is also relevant for many other applications, such as systems with multuser detectors [33] or, for instance, the channel studied in [17], [31]. For more details about this model, the reader is referred to [6] and [8]. Denoting by $C_n = \sum_{k=1}^{n} e_{nk}$ the average number of packets correctly received in collisions of size $n$, we assume that the limit $C = \lim_{n \to \infty} C_n$ exists, as is usually the case with models of practical interest. It has been proved in [8] that the Aloha random access algorithm has a maximum stable throughput $\eta_0 = C$ in the multipacket channel. Decentralized control strategies have been shown [11], [12], [19], [25], [30] to stabilize the slotted Aloha algorithm in the case of the usual collision channel, hence, it is reasonable to expect that when those strategies are used in the multipacket channel, the resulting throughput will be higher than $\eta_0$. We consider schemes of the form

$$\eta_n = F(S_n)$$

$$S_{n+1} = G(S_n, Z_n)$$

(1)

where $\eta_n$ is the retransmission probability in slot $n$, $S_n$ is an estimate of the backlog $X_n$ at the beginning of slot $n$, and $Z_n$ is the feedback at the end of slot $n$. The number of new packets arriving during slot $n$, $A_n$, is assumed to form a sequence of i.i.d. random variables with probability distribution $P[A_n = k] = \lambda_n(k \geq 0)$, such that the mean arrival rate $\Lambda = \sum_{n=1}^{\infty} n \lambda_n$ is finite. Each of the $A_{n-1}$ new packets that arrived during slot $n-1$ is transmitted in slot $n$ with probability $\eta_n$.

As in the case of conventional channels, it is useful to study first the case of control with perfect state information where the value of the backlog is given to the users prior to the selection of the retransmission probability. To keep track of the exact value of the backlog, a central controller is usually necessary, which is an unreasonable requirement for most practical random access channels. However, the study of the perfect state information case allows us to determine an upper bound to the best throughput $\eta$, achievable by any decentralized control of the form (1), and suggests a simple implementation. Those results are in turn helpful to derive control protocols in the case where the backlog is unknown. This is done in Section III where we consider a backlog estimate which is recursively updated using the binary feedback. In addition, it is assumed throughout the paper that each station is informed when its packet is successfully received. It is proved that provided a certain condition on the reception matrix holds, the throughput achievable with this type of feedback is the same as the perfect state information throughput.

In a paper whose translation appeared only very recently [19] (after our work [7]), Mikhailov has derived sufficient conditions for stability and instability of two-dimensional Markov chains. Although this was meant to be used for decentralized control schemes in the usual collision channel, this approach is powerful enough to be applied to the multipacket channel. In Section IV we show by using Mikhailov's result that the scheme presented in Section III is stable under weaker assumptions. However, only a weaker form of stability can be proved in this way.

0018-9286/89/1100-1153$1.00 © 1989 IEEE
II. CONTROL OF THE MULTIPACKET CHANNEL WITH PERFECT STATE INFORMATION

In this section we assume that all the users know the value of \( X_s \) at the beginning of slot \( n \), and we let the retransmission probability be a function of the exact value of the backlog, i.e., \( p_x = F(X_s) \). In this ideal case, the system is much simpler to analyze than in the general case (1) since \( (X_s)_{s>0} \) is a homogeneous Markov chain. Our goal is to determine the optimal control function \( F^* \) that yields the largest ergodicity region, and the corresponding throughput, denoted by \( \eta \). For instance, it is well known [4] that for the usual collision channel with the access rule in effect here, \( F^*(X_s) = 1/X_s \) is the retransmission probability that minimizes the drift at each step, resulting in an ideal throughput of \( \eta = e^{-1} \).

First note that all the results herein are valid provided that the backlog Markov chain \( (X_s, S_{s>0}) \) corresponding to a control (1) is irreducible and aperiodic. It can be easily checked that for both access rules considered in this paper (see below), as well as all the algorithms, a simple set of sufficient conditions for irreducibility and aperiodicity is

a) \( \lambda_0 \neq 0 \)

b) \( \lambda_0 + \sum_{n=1}^{\infty} \lambda_n \epsilon_n < 1 \)

c) \( \epsilon_{10} \neq 0 \)

which are analogous to the conditions for the open-loop system studied in [6]. The theorem below gives the best throughput possible for a control protocol (1).

Theorem 1: There exists a retransmission probability \( p^* \) that minimizes the expected backlog increase when the backlog is equal to \( n \).

With such a retransmission probability, the system is stable for \( \lambda < \eta \) and unstable for \( \lambda > \eta \), with

\[
\eta = \sup_{x>0} e^{-x} \sum_{n=1}^{\infty} C_n \frac{x^n}{n!}.
\]

Proof of Theorem 1: The proof is based on standard drift analysis techniques. \( (X_s)_{s>0} \) is a homogeneous Markov chain which evolves according to

\[
X_{t+1} = X_t + A_t - \Sigma_t
\]

where \( \Sigma_t \) is the number of packets successfully transmitted in slot \( t \). The system is defined to be stable if \( (X_t)_{t>0} \) is ergodic and un

b\) otherwise. Let \( d_e \) be the drift of \( X_t \) at state \( n \): \( d_e = E[X_{t+1} - X_t | X_t = n] \). We have \( 0 \leq \Sigma_t \leq X_t \), and if we denote by \( p \) the retransmission probability used in slot \( t \), then for \( n \geq 1 \), the probability of having \( k \) successes is given by

\[
d_e = \lambda - \sum_{k=1}^{n} \sum_{j=1}^{n} \binom{n}{j} p^j(1-p)^{n-j} e^{-(j-1)n} (1 \leq k \leq n). \tag{3}
\]

t then follows from (2) that the backlog drift at state \( n \geq 1 \) is given by

\[
t_e = \lambda - \sum_{k=1}^{n} \binom{n}{j} p^j(1-p)^{n-j} e^{-(j-1)n} \frac{1}{k} \tag{4}
\]

which becomes \( d_e(p) = \lambda - t_e(p) \) if we define \( t_e(p) \) to be the average number of successes given the backlog \( n \) and the retransmission probability \( p \)

\[
t_e(p) = \sum_{j=1}^{n} \binom{n}{j} p^j(1-p)^{n-j} C_j. \tag{5}
\]

Since \( t_e(p) \) is a polynomial on the compact \([0,1]\), it achieves its maximum and we can define

\[
\rho^*_n = \arg \max_{p \in [0,1]} t_e(p) = \arg \min_{p \in [0,1]} d_e(p).
\]

We now proceed to compute the limit of the drift when the retransmission probability \( p^*_n \) is used. We show that

\[
\lim_{n \to \infty} t_e(p^*_n) = \sup_{x>0} e^{-x} \sum_{n=1}^{\infty} C_n \frac{x^n}{n!} = \sup_x (t(x)). \tag{6}
\]

Let us first assume that \( \lambda < +\infty \).

Property 1:

We have for \( n > M \)

\[
|t(x) - C| \leq e^{-x} C + e^{-x} \sum_{n=1}^{M} \frac{x^n}{n!} |C_n - C| + \sum_{n=M+1}^{\infty} \frac{x^n}{n!} |C_n - C|.
\]

Choosing \( M \) as for Property 1 we get

\[
|t_e(p) - C| \leq \sum_{j=0}^{n} \binom{n}{j} p^j(1-p)^{n-j} |C_j - C| + (1-p)^n C.
\]

Choosing \( M \) as for Property 1 we get

\[
|t_e(p) - C| \leq 2B_1 \sum_{j=0}^{n} \binom{n}{j} p^j(1-p)^{n-j} + \epsilon.
\]

Properties 2 and 3 follow.
a random access multipacket channel

Proof of Theorem 2: To prove the first part of the theorem we use a result of [27] which is a generalization of Kaplan's theorem. If \( p_i = F(S_i) \) and \( S_{i+1} = G(S_i, Z_i) \), consider the Markov chain \((X_i, S_i)\) and the Lyapunov function \( V(n, s) = n \).

\[
\lambda - \sum_{i=1}^{n} p_i F(s)/(1 - F(s)) \geq C_j
\]

for all \( n \) large enough and all \( s \). Therefore, the drift of \( V \) is strictly positive outside a finite subset of the state space. Since it is shown in the Appendix that the generalized Kaplan's condition is verified, it is enough to conclude that \((X_i, S_i)\) is nonergodic. Hence, \( \eta_2 \) is indeed the best throughput achievable by any decentralized control algorithm of the form (1).

To prove the second part of the theorem, we need the following property.

**Property 5:** If for all \( x \geq 0 \), \( t(x) < \sup_{x \geq 0} t(x) \), then \( \eta_2 = C \).

If \( \sup_{x \geq 0} t(x) = +\infty \), it is easily seen that \( C = +\infty \). If \( \sup_{x \geq 0} t(x) = +\infty \), consider a sequence \((x_k)_{k \geq 1}\) of nonnegative reals such that \( \lim_{k \to \infty} t(x_k) = \sup_{x \geq 0} t(x) \). If \((x_k)_{k \geq 1}\) was bounded above by \( K < +\infty \), we would have for all \( n \geq 1 \), \( t(x_k) \leq \sup_{x \geq 0} t(x) \), and in the limit \( \sup_{x \geq 0} t(x) = \sup_{x \geq 0} t(x) \). Then there would exist \( x_0 \in [0, K] \) such that \( t(x_0) = \sup_{x \geq 0} t(x) \), which is a contradiction. Therefore, \((x_k)_{k \geq 1}\) is unbounded, and one can build a subsequence \((x_k)_{k \geq 1}\) such that \( t(x_k) \to +\infty \). We still have, of course, \( \lim_{x \to 0} t(x) = \sup_{x \geq 0} t(x) \), but on the other hand, we have \( \lim_{x \to 0} t(x) \neq \lim_{x \to 0} t(x) \). From Property 1 in the proof of Theorem 1, \( \lim_{x \to 0} t(x) = C \) and Property 5 follows.

Thus, if \( \eta_2 > C \), then \( t(x) \) achieves its supremum at some finite positive real \( A \). Let us consider the control \( \rho_i = A/X_i \), for \( X_i \geq A \). (Note that the value of the retransmission probability is unspecified for \( X_i < A \) because it does not affect the throughput.) Then from (4) \( d_\rho = -t(A) \), and from Property 3 in the proof of Theorem 1 \( \lim_{n \to \infty} d_\rho = t(A) \). It follows then from (21) that \( (X_i)_{i \geq 0} \) is ergodic if \( \eta_2 > C \), and I assume that the number of transmissions in each slot, \( N \), is Poisson distributed, and if we could choose any positive number as the mean of \( N \) by regulating the retransmission probability, the throughput would equal the average number of successes per slot, \( E(C_j) \), maximized over the mean of \( N \). As in the usual collision channel, a wrong analysis leads to a correct conclusion. Several examples are gathered in Table I (see [8] for details).

Probably the most important conclusion of this section is that in general it is not necessary to compute the exact value of \( p^* \), which would require a large amount of on-line computations, and seriously hinder any application of Theorem 1 to the case where the backlog is unknown. Two cases may occur. If \( t(x) \) does not attain its supremum, from Property 5 in the proof of Theorem 2, we have \( \eta_2 = C \) (e.g., this happens in the model developed in [6] for mobile users with pairwise transmissions). In this case no throughput improvement can be achieved by varying the retransmission probability, and therefore it is enough to restrict attention to the open-loop strategy studied in [8]. On the other
for the usual collision channel in the IFT case, the retransmission probability is

\[ p^* = \frac{X_0}{X_0 - X_1}, \]

and it can be shown that there exists a retransmission probability that minimizes the drift of the Markov chain with partial state information provided by the reception matrix is verified.

II. Property 1

Proof of Theorem 3: If \( \lim_{\lambda \to \infty} C_\lambda = \infty \), then the best throughput achievable by the class of algorithms in (1) is \( \neq \sup \{ \lambda : \lambda < \sup_{x>0} T(x) \} \).

Thus, in this case, \( T(x) \) depends only on \( \lambda \), and to clarify the proof below, we denote it by \( T_\lambda(x) \).

\[ T_\lambda(x) = t(x) + \lambda. \]

Assume that \( t(x) \) does not achieve its supremum. Then from Property 3 in the proof of Theorem 2, we have \( \eta_\lambda \equiv C \equiv \lim_{\lambda \to \infty} T_\lambda(x) \). It follows from (14) that for all \( \lambda > 0 \), \( \lim_{\lambda \to \infty} T_\lambda(x) = C \).

Therefore, for all \( \lambda > 0 \), \( \sup_{x>0} T_\lambda(x) \) is finite. Hence, for all \( \lambda > 0 \), \( \sup_{x>0} T_\lambda(x) = \sup_{x>0} T(x) \), and by definition of \( \eta_\lambda \), we finally get \( \eta_\lambda = \sup_{x>0} T(x) \).

Note that \( T(x) \) does not achieve its supremum, in the sense that if there existed \( \lambda \in (0, \infty) \) and \( x_0 \geq 0 \) such that \( \sup_{x>0} T(x) = (t(x) + \lambda) \), then for all \( \lambda \in [0, x_0] \),

\[ T(x) = \sup_{x>0} T(x) = t(x) + \lambda. \]

We have for all \( x \geq 0 \) (15), therefore \( \sup_{x>0} T(x) \leq x_0 \).

Together with (15), it follows that for all \( \lambda \in (0, \sup_{x>0} T(x)) \), \( \lambda < \sup_{x>0} T_\lambda(x) \), and therefore \( \eta_\lambda = \sup_{x>0} T(x) = \eta_\lambda \).

Since from (14) \( \sup_{x>0} T(x) \leq \eta_\lambda \) for all \( \lambda \), we get \( \eta_\lambda \leq \eta_\lambda \) and finally \( \eta_\lambda = \sup_{x>0} T(x) \).

Note that from (14), \( T(x) \) reaches its supremum too, since for all \( \lambda < \infty \), there exists \( x_0 \geq 0 \) such that \( T_\lambda(x_0) = \eta_\lambda \).

Claim that we have also shown in this proof that \( T(x) \) reaches its supremum iff \( t(x) \) does, which means that \( \eta_\lambda \) can be achieved with a control of the form \( p_\lambda = A/X_\lambda \), if \( \eta_\lambda \) can.

III. OPTIMAL CONTROL FOR THE MULTIPACKET CHANNEL

It is assumed from now on that the users do not have access to the value of the backlog, so the problem becomes one of control of the Markov chain with partial state information provided by the channel feedback. We build a backlog estimate \( S \), with feedback which is such that \( Z = 0 \) if slot \( t \) was empty, and \( Z = 1 \) otherwise. The results of the previous section strongly suggest that we should use as a retransmission probability \( p_\lambda = A/S \), where \( A \) is a point at which \( t(x) \) achieves its supremum (according to Property 3, \( A \) is assumed to be finite). We show that the resulting control algorithm achieves the optimal maximum stable throughput \( \eta_\lambda \). This holds provided that the following assumption on the reception matrix is verified.
Theorem 4: Assume that there exists $A \in (0, +\infty)$ such that $(A) = \sup_{z \geq 0} e^{z(A)}$, that the new packet arrivals $(A)_{t \geq 0}$ are exponential type, and that condition $C$ holds if $\alpha < 0$ and $\beta < 1$ verify the following two conditions:

1. $\beta > \lambda$
2. $\beta(1 - e^{-\beta}) + \eta - \lambda + \alpha e^{-\beta} = 0$

then the control algorithm (cf. the control laws proposed in [15], [19], and [25])

\[ d(n, s) = \max \{ A, |s| + \alpha I(Z = 0) + \beta I(Z = 0) \} \]

\[ \eta = \frac{4}{\lambda} \]

\[ \eta = \max \{ n, s \} \]

\[ \eta = \max \{ n, s \} \]

Proposition 1: There exist $\gamma \in (0, 1/5), \delta > 0, \text{ and an integer } n$ such that $\tilde{c}(n, s) < \eta$ for all $(n, s) \in C(-\gamma, \gamma) \cap U_M, c(n, s) \leq -\delta$ and $c(n, s, J) \leq -\delta + \upsilon(J)$.

The detailed proof of Proposition 1 can be found in [9]. After computing the value of the drifts

\[ d(0, s) = \lambda - \sum_{j=1}^{n} \left( \frac{n}{A} \right)^{\frac{1}{A}} \left( 1 - \frac{A}{s} \right)^{\frac{n}{A}} C_j \quad (n \geq 1) \]  

\[ d(n, s) = \beta - \max \{ \lambda - (A, s) - \lambda \} \]

\[ d(n, s) = \beta - \lambda + \max \{ A - s, \alpha \} - \alpha \left( 1 - \frac{A}{s} \right) \]

we work out upper and lower bounds by truncating the sums (16) and (17) to a fixed number of terms, and then we approximate those bounds as a function of the sole variable $n$. The main idea is that the dynamic behavior of the Markov vector $M = (X, S)$ depends essentially on the ratio $X/S$. For instance, if $x$ is nearly equal to $1$, the backlog estimate is close to its ideal value, and we should have $c(n, s) < 0$ since the backlog drift is negative in the perfect state information case. Also, a well-behaved estimate should be such that if $x < 1$, then the error $s - n$ is positive, and therefore should have a negative drift $d(n, s) < 0$ (see [15]).

In the same way, we expect to have $d(n, s) > 0$ for $x > 1$.

We define the following Lyapunov function:

\[ V(n, s) = \max \left\{ \frac{n}{\gamma} - 3\gamma (s-n), \frac{1-3\gamma}{\gamma} (s-n) \right\} \]

where the constants have been chosen so that $V$ is continuous. $V(n, s)$ is equal to the first, second, and third term inside the bracket when $(n, s)$ is in $C(-\gamma, \gamma), C(\gamma, +\infty)$, and $C(-\gamma, -\gamma)$, respectively. Notice that $V$ is defined so as to take the best advantage of the drift properties listed in Proposition 1. For
instance, when $V(n, s)$ is equal to $n$, then the Markov chain $M_t$ belongs to $C(-3, 3)$ which is included in $C(-3, 5)$ where the backlog drift is negative provided that either $n$ or $s$ is sufficiently large. Similar comments can be made about the other two regions. Unfortunately, this does not enable us to conclude that the drift of the Lyapunov function is negative in $U_M$ because $M_t$ may well be in a different region than $M_t$. However, this change of region becomes unlikely if we exclude a small zone around the lines $x = 1 \pm 3y$ where $V$ changes definition and indeed the second part of this proof consists of showing that the Lyapunov function has a negative drift in the remainder of the state space.

Proposition 2: There exist $M_0 \geq 0$ and $\delta_0 > 0$ such that for all $N \geq M_0$ and for all $(n, s) \in U_N \cap \{C(-\infty, -4y) \cup C(-2y, 2y) \cup C(4y, \infty)\}$, we have

$$E[V(M_{n+1}) - V(M_n)] \leq (t, s) < -\delta_0.$$  

Proof of Proposition 2: We consider separately likely and unlikely events

$$E[V(M_{n+1}) - V(M_n)]M_t = (n, s)$$

$$E[V(M_{n+1}) - V(M_n)]I(\{|A_1 - \Sigma_1| \leq J\})M_t = (n, s).$$

(18)

We start by showing that the first term, which corresponds to likely events, is negative when $J$ is large by using the properties of the truncated drifts from Proposition 1 and a simple geometric result. The lemma below, whose proof is in [9], gives a measure of how much a cone $C(\lambda_0, \lambda_1)$ expands if each of its points is allowed to move of some distance that cannot exceed an absolute value along each axis.

Lemma: Consider $\gamma > 0$, $B > 0$, and $1 < \lambda_0 < \lambda_1 < +\infty$ and assume that $n - n' \leq B$, $|s - s'| \leq B$, and $Q \geq B/\gamma (1 + |\lambda_0|/|\lambda_1 + 2 + \gamma$). Then:

1) $(n, s) \in C(\lambda_0, \infty) \cap U_0$  
   
2) $(n, s) \in C(-\infty, \lambda_1) \cap U_0$  
   
3) $(n, s) \in C(\lambda_0, \lambda_1) \cap U_0$  

Let us denote by $T_1(n, s, J)$ and $T_2(n, s, J)$ the two terms on the right-hand side of (18). The first term $T_1(n, s, J)$ corresponds to a case where the variation in the backlog is bounded below, and can be shown to vanish as $J$ increases by using the sole fact that the mean arrival rate $\lambda$ is finite. Consider now $T_2(n, s, J)$. If $M_0 = (n, s)$ belongs to a region such that $x = n/s > 3\rho$, then $x_0$ can be chosen large enough so that if $M_{n+1}$ belongs to $C(-\infty, -3y)$, then the error in the backlog estimate will result from the large number of successes just compensates the initial error $n - s > 0$. On the other hand, when $M_0$ belongs to any region such that $x$ is bounded above, then $E\Sigma(C, J)]M_0 = (n, s)$ goes to zero uniformly in $(n, s)$ and $T_2(n, s, J)$ can be dealt with by using the following rather crude bound for the variation of $V$:

$$|V(M_{n+1}) - V(M_n)| \leq \max \{1, 1 + 3y - 3y\}.$$  

(24)

where $R$ is some positive constant. It is shown in [9] that

$$E[V(M_{n+1}) - V(M_n)](\{A_1 - \Sigma_1 > J\})M_0 = (n, s)$$

$$\leq \delta_t(n, s).$$  

(25)

where $\lim_{n \to \infty} \delta_t(n, s) = 0$ and $\delta_t(n, s)$ is a nonnegative function that depends on $J$ and goes to zero as either $n$ or $s$ goes to infinity. By using (22), (25), and the decomposition (18), we get the desired result that the drift of $V$ is negative in this part of the state space: fix an integer $J_{\text{max}}$ such that $J_{\text{max}} > J_0$ and then for all $J \geq J_{\text{max}}$, $\delta_t(n, s) \leq \delta_t/3$. Then from (22) and (25), we have for all $(n, s) \in U_{Q(\text{max})} \cap \{C(-\infty, -4y) \cup C(-2y, 2y) \cup C(4y, \infty)\}$

$$E[V(M_{n+1}) - V(M_n)]M_0 = (n, s) \leq \frac{2}{3} \delta_t(n, s),$$

(26)

Then we can choose an $M_0 > Q(\text{max})$ which is large enough so that $\delta_t(n, s) < \delta_t/3$ for all $(n, s) in U_{M_0}$. This concludes the second part of the proof. Unfortunately, it is not always true that the drift of $V$ is negative outside a finite subset of the state space. For instance, we have proved that in the case of the usual collision channel with Poisson new packet arrivals, there exist constants $B_{\text{MAX}} > 0$ and $M_0$ such that for all $(n, s) \in U_{M_0}$ for which $x = 1 \pm 3y$, and for all $\alpha$ and $\beta$ verifying $C1$ and $C2$, $E[V(M_{n+1}) - V(M_n)]M_0 = (n, s) > B_{\text{MAX}}$. However, discontinuities around the lines $x = 1 \pm 3y$ can occur when the backlog grows long enough, and in the last part of this proof we show that the $y$-step drift of $V$, $E[V(M_{n+1}) - V(M_n)]M_0 = (n, s)$ is negative for some integer $J$.

Proposition 3: There exist $J_0 > 0$, $\rho > 0$, and $M_0 > 0$ such...
that for all \((n, s) \in U_{M_t}\)
\[
E[V(M_{t+j}) - V(M_t)]_{M_t = (n, s)} \leq -\rho.
\]

**Proof of Proposition:** One of the main problems in dealing with the \(J\)-step drift of \(V\) is to control the changes of regions between \(M_t\) and \(M_{t+j}\). To this end, we define the stopping time
\[
\tau_j = \min \left\{ s \geq 0, \sum_{k=0}^j \left( A_{n+s} - S_{n+s} \right) > J \right\}.
\]
If \(\tau_j \geq J\), then for \(1 \leq k \leq J\), \(|X_{n+k} - X_n| \leq J\) and \(|S_{n+k} - S_n| \leq J(|\alpha| + \beta)\). Thus, if we define \(B'(J) = \max \left\{ J(|\alpha| + \beta), J^j \right\}\) and \(Q'(J)\) to be any integer such that \(Q'(J) \geq B'(J) + \max \left\{ M_0, M' \right\}\), and \(Q'(J) \geq 2B'(J)/(1 + 9/2\gamma)(5\gamma + 2)\), then, still assuming that \(\tau_j \geq J\), we get from the lemma for \(0 \leq k \leq J\),

\[
M_t \in C(-\infty, -4\gamma - J/2) \cap U_{Q'(J)}
\]

\[= M_{n+j} \in C(-\infty, -4\gamma) \cap U_{M_0}\] (26)

\[
M_t \in C(-2\gamma + J/2, 2\gamma - J/2) \cap U_{Q'(J)}
\]

\[= M_{n+j} \in C(-2\gamma, 2\gamma) \cap U_{M_0}\] (27)

\[
M_t \in C(4\gamma + J/2, \infty) \cap U_{Q'(J)}
\]

\[= M_{n+j} \in C(4\gamma, \infty) \cap U_{M_0}\] (28)

\[
M_t \in C(-4\gamma - J/2, -2\gamma + J/2) \cap U_{Q'(J)}
\]

\[= M_{n+j} \in C(-4\gamma, -2\gamma) \cap U_{M}\] (29)

\[
M_t \in C(2\gamma - J/2, 4\gamma + J/2) \cap U_{Q'(J)}
\]

\[= M_{n+j} \in C(2\gamma, 4\gamma) \cap U_{M}\] (30)

In other words, we have partitioned the plane into two zones

\[
Z_N = C(-\infty, -4\gamma - J/2) \cup C(-2\gamma + J/2, 2\gamma - J/2)
\]

\[\cup C(4\gamma + J/2, \infty),\]

and

\[
Z_P = C(-4\gamma - J/2, -2\gamma + J/2) \cup C(2\gamma - J/2, 4\gamma + J/2).
\]

Then we have chosen \(Q'(J)\) such that if \(M_t\) belongs to \(Z_N\), which is slightly smaller than the region in which the drift of the Lyapunov function is negative, and if \(\tau_j \geq J\), then the Markov chain remains in the region in which Proposition 2 applies up to time \(t + J\) (see (26)-(28) and Fig. 2). \(Q'(J)\) is also such that if \(M_t\) is in \(Z_P\) and if \(\tau_j \geq J\), then up to time \(t + J\) the chain stays in a region such that two out of the three properties of Proposition 1 hold at each step (see (29), (30), and Fig. 3).

We start by showing that the \(J\)-step drift of \(V\) is negative at \((n, s)\) when \((n, s)\) belongs to \(Z_N\). We decompose the \(J\)-step drift of \(V\) as follows:

\[
E[V(M_{t+j}) - V(M_t)]_{M_t = (n, s)} = \sum_{k=0}^{J-1} E[E[V(M_{t+k+1})] - V(M_{t+k})]
\]

\[\cdot I(\tau_j < J)|M_t = (n, s)] + \sum_{k=0}^{J-1} E[V(M_{t+k+1})] - V(M_{t+k})I(\tau_j < J)|M_t = (n, s)].\] (31)

Denote by \(U_t(J, n, s)\) and \(U_t(J, n, s)\) the two sums on the right-hand side of (31). If \(\tau_j \geq J\), then (26)-(28) hold, and therefore we can apply Proposition 2

\[
U_t(J, n, s) = -J\delta P(\tau_j \geq J)|M_t = (n, s)].\] (32)

Let us now show that \(\tau_j < J\) is indeed an unlikely event, the
Lemma 2. Assume that 

$$E[\Sigma_{t+n} | M_n = (n, s)]$$

is bounded above independent of \(n\).

From Markov’s inequality we have

$$\Pr \{ | \Sigma_{t+n} | > M_n \} \leq \frac{1}{2^n} \sum_{k=0}^{2^n} E[\Sigma_{t+n} | M_n = (n, s)].$$

Denoting by \(B\) an upper bound on the sequence \(t_r(p^*_r)\), it follows from Section II that

$$E[\Sigma_{t+n} | M_n = (n, s)] = E[E[\Sigma_{t+n} | M_{t+n}] | M_n = (n, s)] \leq B,$$

so we get

$$\Pr \{ | \Sigma_{t+n} | > M_n \} \leq \frac{1}{2^n} \sum_{k=0}^{2^n} E[\Sigma_{t+n} | M_n = (n, s)].$$

where \(B\) is some positive constant. From (24), it is easy to check that the drift of \(V\) is bounded by some positive constant \(B\), so that

$$\Pr \{ \Sigma_{t+n} \leq B \} \geq 1 - \frac{B}{2^n}.$$

By (33), the J-step drift of \(V\) restricted to likely events \(\tau_J \geq J\) goes to \(-\infty\) as \(J\) increases, and then we prove that the J-step drift of \(V\) restricted to unlikely events \(\tau_J < J\) is bounded. Using (30) and (38c), it follows that there exists a constant \(\delta_1\) independent of \(J\) such that \(\delta_1 \leq \min \{ 1, (1-3\gamma)/3\gamma \}\) has been defined in (22). We show that the first two terms on the right-hand side of (36) are bounded. Since (33) \(\lim_{J \to \infty} \delta_1 J/2P(\tau_J \geq J) = -\infty\), this will be sufficient to prove that \(\lim_{J \to \infty} E[V(M_{t+n}) - V(M_n)]I(\tau_J \geq J) = -\infty\). Define \(W_j = X_j - X_{j+1} + kT_{j+1}/2\) and \(\delta_j = F_{j+1}\), where \(F_i\) is the sigma-field generated by \(\{A_t, s \leq t < j, X_t, 5 \leq j\}\), representing the history of the process \((M_{t+n})_+\) up to time \(t\). To prove that the first term in (36) is bounded, we show that there exists \(\delta_2 > 0\) such that \((Y_n, F_{n+1})\) is a supermartingale, with \(Y_n = e^{kT_{n+1}}(I(\tau_n \geq k)).\) We need to show that \(E[Y_{n+1} | F_n] \leq Y_n\), which is equivalent to

$$\Pr \{ e^{kT_{n+1} + k + 1 - X_n} \leq k + 1 \} I(\tau_n \geq k + 1) = 0.$$

Now if \(\tau_n \geq k + 1\), then from (30), \(M_{t+n} \in C(\gamma \cdot 5, \gamma \cdot 5) \cap U_{n+1}\), and there exist two constants \(\gamma > 0, \delta < 1\) such that \(E[e^{kS}] < \delta < 1\). Hence, there exists \(\delta > 0\) such that

$$E[Y_{n+1} | F_n] = E[e^{kS} I(\tau_n \geq k)] F_n \leq E[Y_n | F_n] = 1.$$
Theorem 4: Suppose that:

i) the number of new packet arrivals per slot has finite second moment $E[A^2] < +\infty$;

ii) there exists $A \in (0, +\infty)$ such that $t(A) = sup \tau_p x (x)$;

iii) $C^0$: there exists $B < +\infty$ such that for all $n \geq 1$,

$$\Sigma_{x \in \mathbb{Z}} x^2 \eta_x \leq B.$$ 

Fix $\lambda < \eta_n$ and $\xi > 0$ such that $\lambda < \tau(A \xi)$. Choose $\alpha < 0$ and $\beta > 0$ such that

$$\omega^1 : \beta(e^{\xi - t}) = -\frac{\lambda - \tau(A \xi)}{\xi} e^{\xi - \alpha}.$$ 

Then the control algorithm

$$c(n, s) = \begin{cases} \lambda - \tau \left( A^2 \right) & \text{if } \eta \in (0, 0) \\ \sup_{\lambda > \eta} \frac{\lambda - \tau(A \xi)}{\xi} e^{\xi - \alpha} & \text{if } \eta \notin (0, 0) \end{cases}$$

is stable.

Proof of Theorem 5: Let us state first Mikhailov's Theorem (cf. [35] for an exposition of this result and its application in the decentralized control of the conventional collision channel).

Theorem (Mikhailov [19]): Let $M_t = (X_t, S_t)$ be a homogeneous Markov process on $R^* \times R^*$ with drifts

$$c(n, s) = E[M_{t+1} - M_t | M_t = (n, s)].$$

Suppose that:

i) there exists $B < +\infty$ such that for all $(n, s) \in R^* \times R^*$,

$$E[|M_{t+1} - M_t|^2 | M_t = (n, s)] \leq B;$$

ii) for all $\psi \in (0, +\infty)$, the drifts $(c(n, n/\psi), c(n, n/\psi))$ converge uniformly in $\psi$ as $n$ goes to infinity to $(c(\psi), c(\psi))$;

iii) the limit drifts $(c(\psi), c(\psi))$ are differentiable on $(0, +\infty)$, with $(c(0), c(0)) = \lim_{\psi \to 0} (c(0, s), c(0, s)), c(0, s)$; and

iv) there exists $\epsilon > 0$ such that if $c(\psi_0) = \psi_0 c(\psi_0)$, then $c(\psi_0) < -\epsilon$.

Then $M_t$ is stable.

Since both the new packet arrivals and the rows of the reception matrix have finite variance, it is easy to check that condition i) in Mikhailov's Theorem holds

$$\epsilon^2 = E[|M_{t+1} - M_t|^2 | M_t = (n, s)] \leq E[|A|^2] E[|B|^2] = 1.$$ 

Now

$$E[(S_{t+1} - S_t)^2 | M_t = (n, s)] \leq E[|A|^2] + E[|S|^2] M_t = (n, s).$$

From C0' the variance of the number of successes is also bounded

$$\Sigma_{n \in \mathbb{Z}} |M_t = (n, s)| = \frac{\sum_{k=1}^{\infty} k^2 \sum_{j=k}^{\infty} \left( \frac{n}{A} \right)^j \left( 1 - \frac{A}{s} \right)^{j-k}}{e^t - B}.$$

It follows directly from (16) and (17) that the limit drifts are given by

$$c(\psi) = \lambda - \tau(A \psi)$$

$$c(\psi) = \beta + (\alpha - \beta) e^{-\psi \phi},$$

respectively, for $\psi \in (0, +\infty)$. Uniform convergence to the limit drifts follows immediately from the results given for the perfect state information case (Property 4). Also it is clear the $t(x)$ is
differentiable (see (6), where \(0 \leq C_s \leq n\)). Therefore, properties ii) and iii) in Mikhailov's Theorem are satisfied.

In order to check property iv) note that if \(\psi_0 = \xi\), then it follows from C1' that

\[
c(\psi_0) = \psi_0^e(\psi_0).
\]

But, at that point, \(c(\psi_0) < 0\) because of the choice of \(\xi\). There is no other root of the equation \(c(\psi) = \psi^e(\psi)\), and, therefore, property v) follows. To see this, note that because of C1', \(c(\psi) = \psi^e(\psi)\) for \(\psi \neq \xi\) is equivalent to

\[
\frac{\lambda - t(A \psi)}{\psi} - e^{t(A \xi) - t(A \psi)} = 0
\]

which is impossible if \(\psi \neq \xi\) because of C2'.

It can be shown [9] that \(m_l(\lambda)\) is finite for all nonnegative \(\lambda\) and \(\xi\), and therefore the set of control laws defined by C1' and C2' is nonempty. Actually, the set of control laws in Theorem 4 is a subset of those in Theorem 5 because in Theorem 5 we can choose \(\lambda = 1\), in which case C2' is equivalent to C1' and C1 is more restrictive than C2 because \(\lambda \geq m_l(\lambda)\) [9]. □

V. Conclusion

In this paper we have investigated the properties of decentralized control algorithms for a random access channel with multipacket reception capability. By using the working hypothesis that the users are aware of the value of the backlog, we have determined the best throughput achievable by any such protocol, as well as a simple way to achieve it. The optimum throughput has been shown to be given by the maximum average number of successes per slot when the number of transmissions, per slot is Poisson distributed. In the imperfect state information case, we have shown that the same throughput achieved in the perfect state information case can be achieved by using in lieu of the true backlog, an estimate of the backlog computed at each station using binary feedback, and we have used this estimate to derive a control scheme which is optimal in the sense that it achieves the optimal throughput determined earlier. This is true provided the reception matrix verifies condition C0, which puts some restrictions on the number of successes per slot. By using Mikhailov's result, C0 can be replaced by the weaker condition C0'. In this case however, geometric ergodicity was not ensured. Note that the feedback empty/nonempty used in Sections III and IV may be less than the available feedback in many practical situations, but no further information is needed: a ternary feedback would not shorten the proof or achieve better throughput.

Finally, let us mention that one can easily modify the proof of Theorem 4 to show that a similar result holds with the IFT access rule. More precisely, under a hypothesis paralleling those of Theorem 4, one can build a control scheme based on a binary feedback empty/nonempty such that the Markov vector \((X, S)\) is geometrically ergodic for \(\lambda < \sup_{x \in D}(T(x))\). Using Theorem 3, it can be seen that the maximum stable throughput is the same for both access rules when the new packet arrivals are Poisson distributed.

APPENDIX

KAPLAN'S CONDITION

Consider a Markov chain with denumerable state-space \(D\), and one-step transition probability matrix \(P_{xy}\) \(x, y \in D\). Let \(V(x)\) by a Lyapunov function on \(D\). Then the generalized Kaplan's condition holds if there exists a positive constant \(B\) such that for all \(x \in [0, 1]\) and all \(x \in D\)

\[
z^{\psi(x)} - \sum_{y \in D} P_{xy} z^{\psi(y)} \geq -B(1 - z).
\]

1) One-Dimensional Kaplan's Condition: Consider the model of Section II with a control scheme \(P_n = F(X_n)\), and the Lyapunov function \(V(x) = x\). To check Kaplan's condition, it is enough from (27) to show that the downward part of the drift \(-D(x) = \Sigma_{j=0}^{L} k P_{j+1}-1\) is bounded below. For \(i \geq k + 1\) and \(1 \leq k \leq i\) we have

\[
P_{k+1;i-k} = \sum_{n=0}^{i-k} \lambda_n \sum_{k+1}^{i} \sum_{j=n+1}^{i} (j) F(i)-(1-F(i))^{j} \epsilon_{j,k+n}.
\]

After a change of variable, it follows that

\[
D(i) = \sum_{j=1}^{i} (j) F(i)-(1-F(i))^{j} \sum_{n=0}^{i-k} \lambda_n \sum_{k+1}^{i} \sum_{j=n+1}^{i} (j) F(i) \epsilon_{j,k+n}.
\]

If \((C_n)_{n=1}^\infty\) is bounded, then Kaplan's condition holds independent of the retransmission policy. Denoting by \(B_2\) an upper bound for \((C_n)_{n=1}^\infty\)

\[
D(i) \geq \sum_{j=1}^{i} (j) F(i)-(1-F(i))^{j} \lambda_n \sum_{k+1}^{i} \sum_{j=n+1}^{i} (j) F(i) \epsilon_{j,k+n}.
\]

\[
\geq \sum_{j=1}^{i} (j) F(i)-(1-F(i))^{j} C_2 \geq -B_2.
\]

2) Two-Dimensional Kaplan's Condition: Consider now the multipacket channel with a general control algorithm (1). Then \((X, S)\) is the Markov chain of interest, and the relevant Lyapunov function is \(V(n, s) = n\). We prove again that Kaplan's condition holds provided that \((C_n)_{n=1}^\infty\) is bounded. From (27), it is enough also in this case to show that the downward part \(T(x)\) of the generalized drift is bounded below, with \(T(x) = \Sigma_{P_{n+1}2 \leq \Sigma_{P_{n+1}2} P_n} (V(y) - V(x))\). Given a state \(x = (i, s)\), we have

\[
T(x) = -\sum_{r=1}^{i} \sum_{k} P_{X_S+1;X_S+1} = \Sigma_{r=1}^{i} \sum_{k} P_{X_S+1;X_S+1} = \Sigma_{r=1}^{i} \sum_{k} P_{X_S+1;X_S+1} = \Sigma_{r=1}^{i} \sum_{k} P_{X_S+1;X_S+1}.
\]

which is, in the same way as before

\[
T(x) = -\sum_{r=1}^{i} \sum_{k} \lambda_i \sum_{j=n}^{i} (j) F(i)-(1-F(i))^{j} \epsilon_{j,k+n}.
\]

this expression is similar to (A-1), and the end of the proof is the same as in (A-2).

REFERENCES


Sylvie Ghez (S'80) was born in Paris, France, on June 3, 1961. She received the engineering diploma in 1984 from the Ecole Nationale Superieure des Telecommunications, Paris, and the M.S. degree in electrical engineering from Princeton University, Princeton, NJ, in 1987.

At present she is a doctoral candidate at Princeton University. Her research interests include multiple access protocols, network routing problems, queueing theory, and applied probability theory.

Ms. Ghez was the recipient of the Princeton Wallace Memorial Fellowship for the year 1987-1988.

Sergio Verdú (S'80-M'84-SM'88), for a photograph and biography, see p 942 of the September 1989 issue of this TRANSACTIONS.

Stuart C. Schwartz (S'64-M'66-SM'83) was born in Brooklyn, NY, on July 12, 1939. He received the B.S. and M.S. degrees in aeronautical engineering from the Massachusetts Institute of Technology, Cambridge, in 1961, and the Ph.D. degree from the Information and Control Engineering Program, University of Michigan, Ann Arbor, in 1966.

While at M.I.T. he was associated with the Naval Supersonic Laboratory and the Instrumentation Laboratory. During the year 1961-1962 he was at the Jet Propulsion Laboratory, Pasadena, CA, working on problems in orbit estimation and telemetry. He is presently a Professor of Electrical Engineering and Chairman of the department at Princeton University, Princeton, NJ. He served as Associate Dean of the School of Engineering during the period July 1977-June 1980. During the academic year 1972-1973 he was a John S. Guggenheim Fellow and a Visiting Associate Professor in the Department of Electrical Engineering, Technion-Israel Institute of Technology, Haifa, Israel. During the academic year 1980-1981, he was a member of the Technical Staff at the Radio Research Laboratory, Bell Telephone Laboratories, Crawford Hill, NJ. His principal research interests are in the application of probability and stochastic processes to problems in statistical communication and systems theory.

Dr. Schwartz is a member of Sigma Gamma Tau,Eta Kappa Nu, and Sigma Xi. He has served as an Editor for SIAM Journal on Applied Mathematics and as Program Chairman for the 1989 ISIT.
Abstract

This work studies optical Code Division Multiple Access (CDMA) systems, and presents the exact error expression for the noncoherent, single-user matched-filter receiver based on the electron count in a symbol period. This analysis is valid for arbitrary photodetectors, adheres fully to the semi-classical model of light, and does not depend on approximations for large user groups or strong received optical fields.

The general error rate expression is specialized to the case of unity gain photodetectors and prime sequences, and the exact minimum probability of error and optimal threshold are compared to those obtained with simplifying assumptions on user transmission coordination or multiple-access-interference (MAI) distribution. We find that the approximation of chip synchronization yields a weak upper bound on the true error rate, and we demonstrate that the approximations of perfect optical-to-electrical conversion and Gaussian MAI yield an optimal hypothesis test whose error rate overestimates the true minimum error rate and underestimates the optimal threshold for moderate and large received optical energies.

Optical CDMA Model

The digital modulation format studied in this paper is optical Direct Sequence Spread Spectrum, i.e., during each symbol interval of duration T, the jth transmitting laser is amplitude-modulated by the product of the data symbol, v(t), and an assigned, signature sequence of relatively short rectangular pulses. This scheme divides the symbol interval into N equal length subintervals, called chips, on which the signature sequence is constant and takes on values in {0,1}. Further, we define Pj as the number of non-zero chips in each signature sequence, b1j as the transmitted symbol of the jth user in the interval [nT, (n + 1)T], and cj(t) as a periodic replication of the signature sequence of the jth user such that c_j(t), t ∈ [nT, (n + 1)T] is the jth signature sequence for any fixed integer n. Then the transmitted complex scalar field from the jth laser may be expressed as

\[ r_j(t) = \sqrt{\frac{s}{T}} c_j(t) b_{1j} e^{i(2\pi f_0 t + \theta_j + \phi_j(t))}, \quad nT \leq t < (n + 1)T \]

where s is proportional to the optical energy per bit of the transmitting laser, \( \nu \) denotes the optical carrier frequency (assumed to be identical for all users), and \( \theta_j \) is the phase offset of the jth laser from the first laser. In this expression, \( W_j(t) \) is a standard Brownian motion, and \( a_j \) is related to the jth transmitting laser linewidth, \( B_j \), by \( a_j = \sqrt{2\pi B_j} \). The relative delays \( \tau_j \) are defined on \([0,T]\) with reference to the receiver of the first user. With dispersion-free transmission (1) also represents the complex scalar field at the first receiver due to user j.

We shall assume that the symbol rate of each user is the same, the optical fields of the K users add in a noncoherent fashion, and that each single-user receiver acquires the timing of its transmitter's symbol epochs. As there is no cooperation between the users, it is appropriate to model the remaining relative delays, \( \{\tau_j \}_{j=2}^K \), as independent, identically distributed, random variables that are uniformly distributed on the interval \([0,T]\). It follows that the intensity of the optical field at the receiver of the first user is

\[ \langle r(0) \rangle^2 = \frac{4s}{T} \sum_{j=1}^{K} b_{1j-1} c_j(t - \tau_j) p_t(0, \tau_j) + \sum_{j=1}^{K} a_j b_{1j} c_j(t - \tau_j) p_t(\tau_j, T) \]

Where \( p_t(\alpha, \beta) \) a rectangular pulse of unit height with support \([\alpha, \beta]\). Due to the modulation shown in (1), the resulting photon point process depends on the data \( b_{1,0} \) only on the set \( \{c_1(t) = 1, 0 \leq t < T\} \). A commonly used receiver for this channel is the noncoherent matched-filter, which sums the photon counts in each of the nonzero chip subintervals of the user of interest. Given that the function \( c_1(t) \) takes values on \([0,1]\), the correlation operation would be easily achieved at extremely low chip rates by an electro-optic modulator, which would allow received light to pass only when \( c_1(t) = 1 \). A more effective device to achieve the matched-filtering operation at higher chip rates is the fiber optic tap delay line, which uses the finite offset of the laser from the first laser. In this expression, \( f_0 \) is the phase offset of the jth laser from the first laser. In this expression, \( f_0 \) is the phase offset of the jth laser from the first laser. In this expression, \( f_0 \) is the phase offset of the jth laser from the first laser. In this expression, \( f_0 \) is the phase offset of the jth laser from the first laser. In this expression, \( f_0 \) is the phase offset of the jth laser from the first laser. In this expression, \( f_0 \) is the phase offset of the jth laser from the first laser.
propagation velocity of light to achieve the proper relative delay of two optical signals by passing them through fibers of different lengths. The matched-filter direct-detection receiver has been studied in several experiments 1,2] and will be the CDMA receiver analyzed in this work.

![Optical Noncoherent Matched-Filter CDMA Receiver](image)

\[ R_{ij}(r) \triangleq \frac{N}{\pi} \int_{-T}^{T} c_{j}(t-r) c_{i}(t) dt \]

\[ \hat{R}_{ij}(r) \triangleq \frac{N}{\pi} \int_{-T}^{T} c_{j}(t-r) c_{i}(t) dt \]

These signals are input to the photodetector. To decide on the value of \( t_{j} \), we use the secondary electron count during the last non-zero non-zero chip of the same bit interval in the involved signal. Thus, the first tap requires more filter xule than the second. The tapped signals are noncoherently recombined, and the output optical signal is incident on the photodetector. To decide on the value of \( b_{j} \), we use the secondary electron count during the last non-zero non-zero interval of the first signature sequence. For the remainder of this work we denote this secondary electron count by \( \Lambda \). We shall employ a common photomultiplier tube, in which the intensity of primary electrons is given by \( \alpha r(t) + \beta \), where \( \alpha \) is proportional to the quantum efficiency of the photodetector, and \( \beta \) denotes the rate of primary electrons due to an independent dark current. The primary electron count mean due to thermoelectrons. In the remainder of this work we set the quantum efficiency of the photodetector to unity, as this effects the distribution of \( N \) only through an attenuation of intensity. Further, we set \( z = \infty \) and \( x = x \) under hypothesis \( H_{1} \).

\[ P[\mathcal{N} = n | x] \]

In this section we obtain the general expression for the pdf of the secondary electron count \( \mathcal{N} \), at the integrator output for an arbitrary photomultiplier and for synchronous or asynchronous transmission. We will use this result in a later section to compare the error rates under various simplifying approximations to the exact error rate. Also, the form of the general expression will be used in the next section to develop arbitrarily tight, computationally efficient bounds on the cumulative distribution function of \( \mathcal{N} \).

In the following, we define \( M \) as the upper bound on the set of total cross-correlations \( R_{ij} + \hat{R}_{ij} \), and as the signature sequences are from \( \{0,1\}^{N} \), these bounds hold for the partial cross-correlations as well. Since the relative delays are uniformly distributed and the chip waveform is rectangular, it is straightforward to show that each cross-correlation is a mixed random variable whose measures have point masses on the integers \( \{0,1,\ldots,M\} \) and continuous portions that are constant between these integers. We shall employ the following notation

\[ P[R_{ij} = i] = d_{j}(i) \quad i \in \{0,1,\ldots,M\} \]

\[ P[R_{ij} \in [v, v + dv]] = c_{j}(v) \quad [v, v + dv] \in (s, s + 1) \]

and we denote the distribution of \( R_{ij} \) as \( (d_{j}(0), d_{j}(1), \ldots, d_{j}(M), c_{j}(0), \ldots, c_{j}(M-1)) \). Thus the marginal distribution of each cross-correlation is completely specified by 2M parameters. Further, the superscript-T notation will be used to distinguish the distribution of the total cross-correlation \( R_{ij} + \hat{R}_{ij} \) from that of \( R_{ij} \), and the hat notation will be used for the distribution of \( \hat{R}_{ij} \).
Our approach to finding the PMF of \( N \) is the following: we will derive the z-transform of \( N \) from its conditional compound Poisson nature, and then show that this z-transform has a particularly straightforward and explicit Maclaurin series expansion. The PMF is the collection of coefficients of this series, and may be explicitly represented.

By conditioning on \( \{ x, \{ R_{1}, \ldots, R_{K} \} \} \), \( j = 2, \ldots, K \), the count \( N \) has a compound Poisson distribution, whose z-transform is given by

\[
E \left[ z^{N} \right] \mid \{ x, \{ R_{1}, \ldots, R_{K} \} \} = \sum_{j=2}^{K} E \left[ z^{N} \right] \mid \{ x, \{ R_{1}, \ldots, R_{j-1}, R_{j}, \ldots, R_{K} \} \} \]

Due to the mutual independence of the pairs \( \{ R_{j}, \tilde{R}_{j} \} \) we need to determine only the expectation of each factor in (3), as the \( j^{th} \) factor depends only on the random mixture \( b_{j-1} R_{j} + b_{j} \tilde{R}_{j} \). It is clear that the random mixture has the same kind of distribution as \( R_{j-1} \), as the \( j^{th} \) factor depends only on the random mixture \( R_{j-1} \), and we denote this mixed distribution by \( D_{j}(0), D_{j}(1), \ldots, D_{j}(M) \). With this notation, the closed form expression of the power series of interest is

\[
E \left[ z^{N} \right] \mid \{ x, \{ R_{1}, \ldots, R_{K} \} \} = \prod_{j=2}^{K} \left( \sum_{i=0}^{M} D_{j}(i) \exp(rG(i-1)z) \right)
\]

We are interested in finding \( P \{ N = n \mid x \} \), which is the coefficient of \( z^n \) in the power series of (4) about the origin. This power series is straightforward but unwieldy in general for most signature sequence sets of interest. For example, the number of parameters in the power series is reduced by a factor of \( K-1 \) by assuming that the marginals of \( R_{1}, R_{2}, \ldots, R_{K} \) are independent of \( j \), i.e., the contribution of user \( j \) to the MAI is statistically indistinguishable from the other interferers. We have verified that this is an excellent approximation when the signature sequences come from the prime codes, and will drop the subscript from the distribution of the random mixtures in the sequel. Also, the power series of this expression is

\[
\sum_{n=0}^{\infty} \frac{1}{(1-G(z))^{b+1}} = G(z) \frac{e^{aG(z)-1}}{(1-G(z))^{b+1}} + \frac{e^{aG(z)-1}}{(1-G(z))^b}
\]

where \( \beta \in (0,1,2,\ldots) \), into the definition for \( \text{Res} \) (7). This yields

\[
(1 - p_{0}) \text{Res}(n + 1, a, b + 1) = \sum_{l=0}^{n} p_{l} \text{Res}(n + l, a, \beta + 1)
\]

where \( G(z) = e^{aG(z)-1} \). For most photomultiplier models \( p_{0} = 0 \), which we will assume in the sequel. The initial conditions of this recursion are also easily extracted from the definition of \( \text{Res} \),

\[
\text{Res}(0, a, b) = e^{-a}, \quad b \in (0,1,2,\ldots)
\]

The linear recursion for \( \text{Res} \) on \( n \) and \( \beta \) permits fast, efficient computation for any arguments \( n, \beta \geq 1 \). Note
that the second initial condition for this recursion depends on \( P(\sum_{k=1}^{k} g_{l} = n) \), which must be known for \( n, k \in \{0,1,2,\ldots\} \). These probabilities require iterated convolutions of the PMF of the random gain \( g_{l} \), may be precomputed and stored for small \( n \) and \( k \), and may be accurately approximated online for large \( n, k \). We are naturally interested in special cases where \( P(A^2) = 1 \), which must be known for \( n, k \in \{0,1,2,\ldots\} \). These probabilities require iterated convolutions of the PMF of the random gain \( g_{l} \), may be precomputed and stored for small \( n \) and \( k \), and may be accurately approximated online for large \( n, k \).

Arbitrarily Tight Bounds on \( P[N \leq n] \)

Computationally efficient bounds must reduce the complexity of (6) in both the multinomial summation and the computation of \( \Pi_{s} \), while controlling the loss of accuracy by a parameter of our selection. In this section we show that by quantizing the random mixtures, we achieve all three objectives.

The complexity of the PMF is due to the smoothing over the joint distribution of the random mixtures we originally conditioned on these random variables to take advantage of the conditional compound Poisson nature of \( N \). We could have also conditioned according to the conditional mean \( \lambda \), for which \( N \) is also compound Poisson. However, the exact distribution of the conditional mean \( \lambda \) is not easily obtained, as it is formed by the convolution of \( K-1 \) mixed distributions. It is obvious that if the convolved distributions were discrete, say, with \( QM + 1 \) points, then the exact distribution of \( N \) would be straightforward to compute. More importantly, the distribution of \( \lambda \) would take on \((K-1)QM + 1 \) points, rather than a number that is exponential in the number of interferers.

But how do we obtain bounds on \( P[N \leq n] \) that use a discrete distribution on \( \lambda \), and are arbitrarily tight? Suppose we quantize the random mixtures \( b_{j-1}R_{j} + b_{j}R_{j+1} \) with a \( Q \) quantization step size, \( Q \in \{1,2,\ldots\} \), and round-up or round-down to form bounds on the random variables. That is, we form \( A_{l}, A_{u} \) given by

\[
A_{l} = z + d + \frac{d}{Q} \sum_{j=2}^{K} b_{j-1} \left[ \frac{1}{Q} R_{j} \right] + b_{j} \frac{1}{Q} \left[ \frac{1}{Q} R_{j} \right]
\]

\[
A_{u} = z + d + \frac{d}{Q} \sum_{j=2}^{K} b_{j-1} \left[ \frac{1}{Q} R_{j} \right] + b_{j} \frac{1}{Q} \left[ \frac{1}{Q} R_{j} \right]
\]

where \([R] \) \( ([R]) \) is the greatest (least) integer function of \( R \). Then it is obvious that \( A_{l} \leq \lambda \leq A_{u} \), but we use \( A_{l}, A_{u} \) to form bounds on the secondary electron count CDF.

A subtle point is raised by considering the form of \( N \)

\[
N(\lambda) = \sum_{p=1}^{\Pi(\lambda)}
\]

where \( \Pi(\lambda) \) is the conditionally Poisson number of primary electrons with conditional mean \( \lambda \). Since \( g_{l} \) are non-negative, we have that \( N \) is an increasing function of the primary electron count, \( \Pi \). It is not clear that (a.s.) bounds on \( \lambda \) produce similar bounds on \( \Pi(\lambda) \), as \( P(\Pi(\lambda) > \Pi(\lambda) | z) > 0 \), and this representation of \( N \) does not guarantee bounds on \( P[N \leq n] \).

In the lemma below we use a statistically equivalent representation of \( N \) to show that we may achieve bounds on \( P[N \leq n] \) by using the distributions of \( A_{l}, A_{u} \).

Lemma. Let \( \Pi(\lambda) \) be a conditional Poisson random variable with mean \( \lambda \) given \( \lambda \), and let \( N(\lambda) = \sum_{p=1}^{\Pi(\lambda)} g_{l} \), where \( \{g_{l}\} \) are independent, identically distributed, non-negative integer-valued random variables. Let \( \lambda < \Lambda \), a.s. Then

\[
P[N(\lambda) \leq n] \leq P[N(\lambda') \leq n], \quad n \geq 0.
\]

Proof. We recall that \( p_{k} = P(g_{l} = k) \), and define \( \{\Pi(\lambda'A_{k}) \} \) to be a set of conditionally mutually-independent, Poisson random variables with the indicated means given \( \lambda \) so that \( \Pi(\lambda) = \sum_{k=1}^{\Pi(\lambda)} g_{l} \), where \( \{g_{l}\} \) are independent, identically distributed, non-negative integer-valued random variables. Let \( \lambda < \Lambda \), a.s. Then

\[
P[N(\lambda) \leq n] \leq P[N(\lambda') \leq n], \quad n \geq 0.
\]

It is straightforward to show that if \( \{X_{1}, X_{2}, Y_{1}, Y_{2}\} \) are conditionally mutually-independent random variables given \( \Lambda \), \( \Lambda' \) and

\[
P[X_{i} \leq n \mid \Lambda, \Lambda'] \leq P[Y_{i} \leq n \mid \Lambda, \Lambda'], \quad i = 1,2
\]

then the same is true for the sum

\[
P[X_{1} + X_{2} \leq n \mid \Lambda, \Lambda'] \leq P[Y_{1} + Y_{2} \leq n \mid \Lambda, \Lambda']
\]

Since the Poisson CDF is a decreasing function of the mean, we have for \( I = 1 \)

\[
P\left[ \sum_{k=1}^{k} \Pi(\lambda' \mid A_{k}) \leq n \mid \Lambda, \Lambda' \right] \leq P\left[ \sum_{k=1}^{k} \Pi(\lambda' \mid A_{k}) \leq n \mid \Lambda, \Lambda' \right]
\]

The same is true for the unconditioned CDFs by smoothing. The same holds for finite \( I \) by induction on the above fact, and for \( I \rightarrow \infty \) by monotone sequential continuity of the probability measures.
Examples: Prime Sequences and PIN Photodiodes

A necessary prerequisite to the comparison between error rates of the CDMA matched-filter receiver is the computation of the random mixture distribution \((D(0), \ldots, D(M), C(0), \ldots, C(M-1))\), as seen in (8). These are computed by a knowledge of the signature sequences, as well as the distribution of the relative delay. Since the cross-correlations of prime sequences are bounded above by \([4]\) \(M = 2\), we must compute \((D(0), D(1), D(2), C(0), C(1))\) for the chip-synchronous, and asynchronous cases. For the prime sequences from \(GF(17)\), we have found that the average distributions for the random mixtures are

\[
\begin{align*}
\text{chip synchronous} &\Rightarrow (0.57, 0.36, 0.07, 0.00, 0.00) \\
\text{asynchronous} &\Rightarrow (0.44, 0.22, 0.01, 0.24, 0.09)
\end{align*}
\]

As noted earlier, we have verified that the MAI for prime sequences is well-modeled by a sum of independent, identically distributed (IID) random variables in the \(\mu\) sense that the mean, variance, and third central moment of the MAI are identical to the exact MAI moments, while the fourth central moments differed by less than \(0.004\%\) for 29 interferers. Further, these distributions did not differ significantly for the prime sequences from \(GF(11)\) and \(GF(17)\), and we use the IID assumption and the average distribution for all calculations.

In Figure 2 we have plotted the minimum error probability of the matched-filter CDMA receiver for the chip-synchronous assumption and for completely asynchronous transmission. We have used the weight 17 and length 290 prime sequences from \(GF(17)\), a received optical energy of \(1000\) photons per bit, and a dark current contribution of \(50\) thermo-electrons per bit. For a bit rate of \(100\) kbps, these numbers correspond to a peak received power of \(4 \times 10^{-7}\) W and a photodetector dark current of approximately \(4 \times 10^{-9}\) A. From Figure 2 we see that over this particular case the chip asynchronous approximation upper bounds the error in the asynchronous case by at least one order of magnitude.

The error rates are ordered in this way due exclusively to the difference of the distributions of the random mixtures shown above. Note that the means of the random mixtures are identical in both cases, while the ordering of the variances coincides with that of the error rates. Thus, MAI has identical means under these distributions, and exact moments whose ordering coincides with that of the error rates. It is easy to show that \(\mathbb{E}[N^2] = \lambda\), and \(\text{var}(N^2) = \lambda (\lambda + 1)\), which implies that under each hypothesis on \(N\) the mean of \(N\) is unchanged by the approximation of chip synchronism, yet the variance of \(N\) given \(x\) increases as we proceed from complete asynchronism to chip-synchronism. From the ordering of the minimum error rate curves in Figure 2, we see that an increase in the variance of \(N\) under each hypothesis results in an increased error rate as the conditional means of \(N\) are fixed.

Direct detection systems often require large received optical energies to achieve an acceptable error rate when a PIN photodiode is used, so we are interested in the asymptotic distribution of \(N\), which is defined as \(\lim_{\lambda \to \infty} \mathbb{E}[N^2]/\lambda\). If \(\rho = 0\), then the normalized count converges in distribution to a standard Gaussian random variable. Is this the case in general?

The answer was solved independently by Serfozo [6] and Grandell [7] for the special case when \(\lambda \to \infty\), and depends on the limit \(\rho\) defined as \(\lim_{\lambda \to \infty} \mathbb{E}[N^2]/\lambda\). If \(\rho = 0\), then the normalized count converges in distribution to a standard Gaussian. If \(\rho = \infty\), then the normalized count converges in distribution to \(\phi\). Finally, if \(0 < \rho < \infty\), then the normalized count converges in distribution to an independent mixture of a standard Gaussian and \(\phi\).

In our case, the parameter is the received signal energy per bit, \(s\), and the condition \(\lambda \to \infty\) is satisfied as \(\lambda \) is proportional to \(s\). It is this fact that also sets \(\rho\) to \(\infty\), and we have from the result above that for large signal energies the normalized count converges in distribution to the scaled conditional mean \(\phi\). This asymptotic result is a weaker form of what is more commonly known as "perfect optical-to-electrical conversion", in which the integrated photocurrent is equal (a.s.) to the integrated optical intensity. It will be seen in the numerical results presented...
next that the asymptotic statistic is far from being a
deterministic signal in Gaussian noise, as the MAI is far from
Gaussian even for a moderate number of users.

In Figure 3 we have compared the minimum error
rates of the CDMA matched-filter receiver based on
perfect optical-to-electrical conversion (the high energy limit)
to those for the true distribution of $N$ at various finite opti-
cal energies. In this example we have used the prime se-
quence from GF(11). Also, we have plotted the minimum error
rate under the additional assumption of Gaussian-distributed MAI. We note that even for modest received
optical energies of 10,000 photons per bit the error rate
exceeds that predicted by the asymptotic distribution by
at least an order of magnitude. Figure 3 shows that the
minimum error rate is a decreasing function of the received
optical energy, as expected. Further, we note that a Gauss-
ian assumption on the MAI, together with the perfect
optical-to-electrical assumptions is a poor estimate of the
true minimum error rate curve, except for user group sizes
exceeding, say, 10 users. In particular, this assumption
overestimates the error rate for moderate to large incident
optical energies.

As a result of the perfect optical-to-electrical conversion
approximation, the boundedness of the MAI leads to an
"error-free" condition for sufficiently small numbers of in-
erferers. This occurs since the supports of the conditional
* distributions of the test statistic are disjoint under these as-
sumptions. Since prime sequences have cross corre-
lations that are bounded above by 2, the necessary condition
for prime sequences is $K - 1 < P/2$. This assumption pre-
dicts zero error rate for $K \leq 6$ in Figure 3, which indicates
that the perfect optical-to-electrical assumption accurately
predicts the "error-free condition" only for incident optical
energies exceeding 10,000 photons per bit - the error rate
or $K=6$ at this energy is roughly $10^{-14}$.

In Figure 4 we have plotted the optimal thresholds,
normalized by the signal energy, $s$, for those error rate-

References
Fiber-Optic Local Area Network Using Optical Processing",
J. Lightwave Tech., Vol. LT-4, No. 5, pp. 547-554,
May,1986
Fiber Optic Code Division Multiple Access (FO-CDMA)",
International Conference on Communications, 1987
Son,1976
Sequences For OR Channel Code Division Multiplexing",
Electronic Letters Vol. 10, No.21,pp. 888-889, 13th,
October, 1983
Levy, J. Wiley, New York, 1972
Abstract

The optimal signal design problem for a band-limited PAM symbol-synchronous Gaussian two-user multiple-access channel is investigated. Using the root-mean-square and the fractional out-of-band energy bandwidth definitions, we find the capacity region of the channel and the signature waveforms to achieve each point inside the capacity region. The optimal pair of signature waveforms are mirror images of each other, and are obtained by minimizing their cross-correlation subject to a fixed finite duration and the bandwidth constraint. The two-user capacity region, in the rms case, is found to contain the capacity region of the two-user strictly band-limited Gaussian channel. This demonstrates the fact that by relaxing constraints in the frequency domain, we can introduce structure (PAM) in the time domain and obtain a larger capacity region.

1. Introduction

The capacity region of the two-user discrete-time Gaussian multiple-access channel

\[ y_i = x_{i1} + x_{i2} + n_i \]

where \( n_i \) is an i.i.d. Gaussian sequence with variance equal to \( \sigma^2 \) and the energy of each codeword is constrained to satisfy

\[ \frac{1}{n} \sum_{i=1}^{n} x_i^2 \leq W_k \quad k = 1, 2 \]

is equal to the Cover-Wyner pentagon [1], [2]:

\[ C_D = \left\{ (R_1, R_2) : 0 \leq R_1 \leq \frac{1}{2} \log(1 + \frac{\sigma^2}{B^2}) \right\} \]

in information units per channel use. Analogously, the capacity region of the continuous-time band-limited channel with noise power spectral density \( B \) and \( k \)-th user signal power equal to \( \sigma^2 \), is given by [2], [3] (in units per second)

\[ C_C = \left\{ (R_1, R_2) : 0 \leq R_1 \leq B \log(1 + \frac{\sigma^2}{B^2}) \right\} \quad \right\} \]

Thus capacity region is achieved by approximately band-limited and approximately time limited waveforms which have no particular structure. In order to deal with modulation and demodulation schemes with manageable complexity, it is customary in digital communications to introduce structure on the transmitted waveforms by slotting the time domain into intervals of length \( T \) and sending a symbol in each slot by means of a digital modulation format such as PAM, PSK, FSK, etc. In the case of PAM (Pulse Amplitude Modulation), the \( k \)-th user is assigned a fixed deterministic waveform, \( s_k(t) \), which is time-limited to \([0, T]\) and is modulated by the information stream. Then, assuming that the transmitters are symbol-synchronous, the PAM two-user multiple-access channel becomes

\[ y(t) = \sum_{i=1}^{n} b_i(t) s_i(t - iT) + b_i(t) s_i(t - iT) + n(t) \quad (3) \]

where \( n(t) \) is white Gaussian noise with spectral density \( \sigma^2 \) and \( b_i(t) \) is the symbol stream transmitted by the \( k \)-th user. Assuming that, without loss of generality, the signature waveforms have unit energy, the energy constraints on the transmitted waveforms become

\[ \frac{1}{n} \sum_{i=1}^{n} s_i^2(t) \leq W_k \quad k = 1, 2 \quad (4) \]

It is easy to show that if \( s_1(t) = s_2(t) \), then the capacity of (3) under constraints (4) is equal to the Cover-Wyner pentagon (1) (this result remains true even if the users are completely asynchronous [3]), if the signature waveforms are not necessarily identical, then the Cover-Wyner pentagon generalizes to [4]

\[ C_V = \left\{ (R_1, R_2) : 0 \leq R_1 \leq \frac{1}{2} \log(1 + \frac{\sigma^2}{B^2}) \right\} \]

in information units per channel use or

\[ C_V = \left\{ (R_1, R_2) : 0 \leq R_1 \leq \frac{1}{2} \log(1 + \frac{\sigma^2}{B^2}) \right\} \]

in information units per second, where \( \rho = \int s_1(t) s_2(t) \, dt \) is the cross-correlation between the signature waveforms.

A natural question to address is the choice of the unit-energy waveforms \( s_1(t) \) and \( s_2(t) \) to maximize the capacity region \( C_V \). It is clear that the unconstrained solution is to choose orthogonal signature waveforms. Then, \( \rho = 0 \), and the multiple-access channel is decoupled into independent single-user channels, and each transmitter can transmit at single-user capacity. However, in practice, there are constraints on the choice of the signals (e.g. in Spread Spectrum CDMA systems, the waveforms may be constrained to be Pseudo Noise shift register sequences of given period, and it is not always possible to assign orthogonal waveforms for all users. In this paper, we will address the optimization of the signature waveforms and their duration \( T \) under bandwidth constraints. Since the signature waveforms are strictly time-limited, they cannot be strictly band-limited, and the need arises to quantify the bandwidth of these signals. There are several established ways to accomplish this [5]. In this paper, we will consider the two bandwidth measures of baseband signals that have received most attention from the information theoretic community: the
root mean square (rms) bandwidth and the fractional out-of-band energy (fobe) bandwidth. The rms bandwidth was popularized by Gabor [6] (it is sometimes referred to as Gabor bandwidth) and studied subsequently in [5], [7], and [8]. A finite-energy signal $s(t)$ has rms bandwidth $B$ if its Fourier transform $S(f)$ satisfies

$$\int_{-\infty}^{\infty} S(f)^2 df = B^2$$

i.e., the rms bandwidth is the square root of the "second moment" of the energy spectral density $|S(f)|^2$ of the normalized signal $s(t)$, proportional to the square root of the energy of its derivative,

$$\int_{-\infty}^{\infty} S(f)^2 df = B^2$$

The fobe bandwidth has been used in e.g. [5], [8] and is defined as the bandwidth necessary to encompass a given fraction (say $\alpha$) of the signal energy, i.e., the $\alpha$-fobe bandwidth is $B$ if

$$\int_{-B}^{B} |S(f)|^2 df = \alpha \int_{-\infty}^{\infty} |S(f)|^2 df$$

Notice that the bandwidth constraints imposed on the signature waveforms will be inherited by the transmitted signals because, as is well known [9], the power spectral density of $\sum b_k(t) s(t - iT - \tau)$ where $\tau$ is uniformly distributed in $[0,T]$ and $(b_k(t))$ is an i.i.d. sequence, is a scaled version of the energy spectral density $|S_\alpha(f)|^2$.

2. Single-user Channel

Before solving for the capacity region of the PAM multiple-access channel under bandwidth constraints, it is enlightening to examine the PAM single-user channel with constrained rms bandwidth. This channel differs from the classical band-limited Gaussian channel in that the allowable transmitted signals have much more structure (PAM) and 2) are rms band-limited but not strictly band-limited. It turns out that the effect of the laser bandwidth measure cancels the effect of the additional structure imposed on the transmitted signals in the time domain, and the capacity of the channel is given by the celebrated Shannon formula [10].

**Theorem 2.1.**

The capacity of the single-user PAM white Gaussian channel with noise power spectral density $N$, rms bandwidth and signal power equal to $\alpha^2$, $B$, and $S$ respectively is given by (in units per second)

$$C_S = B \log_2[1 + \frac{S}{2\alpha^2 B}]$$

Proof.

The single-user PAM white Gaussian channel is a special case of (8):

$$y(t) = \sum_{i=1}^{n} b(i)s(t - iT) + n(t)$$

Assuming that, without loss of generality, $s(t)$ has unit energy, the power constraint becomes

$$\frac{1}{n} \sum_{i=1}^{n} b^2(i) \leq TS$$

and the $T$-shifts of $s(t)$, $(s(t - iT))_{i=1}^{n}$ form an orthonormal set. The projections of $y(t)$ on this orthonormal set are equal to

$$y(i) = \int_{-T}^{T} y(t)s(t - iT) dt \quad i = 1, \ldots, n$$

or, substituting $y(t)$ from (10),

$$y(i) = b(i) + n(i)$$

where $(n(i))$ is an i.i.d. Gaussian sequence with variance equal to $\sigma^2$.

The important point to note is that $(y(i))_{i=1}^{n}$ are sufficient statistics for the transmitted messages; therefore, the capacity of the PAM channel (10) for a fix $T$ coincides with the capacity of the discrete time memoryless channel (13) with constraint (11), which is given by (e.g., [11]) (in units per second)

$$C_S(T) = \frac{1}{2T} \log_2[1 + \frac{ST}{\alpha^2}]$$

Since $C_S(T)$ is monotonically decreasing in $T$, the capacity is maximized by minimizing $T$. However, due to the rms bandwidth constraint, the value of $T$ cannot be arbitrarily small. Using the fact that the set $(\sqrt{\frac{2}{\pi}} \sin(\frac{\pi fT}{T}))_{i=0}^{\infty}$ is a complete orthonormal set in the space of all rms band-limited signals in $[0,T]$, we can express $s(t)$ as

$$s(t) = \sum_{i=1}^{\infty} d_i \sqrt{\frac{T}{2\pi}} \sin(\frac{\pi fT}{T})$$

Then, the unit energy assumption on the constraint in the rms bandwidth (7) translate into

$$\sum_{i=1}^{\infty} d_i^2 = 1$$

and

$$\sum_{i=1}^{\infty} \frac{d_i^2}{T^2} \leq (2BT)^2$$

respectively.

The minimum $T$ consistent with (16) and (17) is chosen by taking equality in (17) and minimizing the left hand side of (17) subject to (16). Since

$$1 = \sum_{i=1}^{\infty} d_i^2 \leq \sum_{i=1}^{\infty} \frac{d_i^2}{T^2}$$

with equality if and only if $d_i = 1$ and $d_j = 0$ if $i \neq j$, it follows that the optimum $T$ is equal to $\frac{1}{2B}$ which upon substitution in (14) results in the desired result.

3. Two-user Channel

We turn our attention to the main results of the paper, namely the optimization of the capacity region of the synchronous PAM channel (5b) with respect to the choice of the signature waveforms, including their duration $T$. In both the rms and the fobe bandwidth constrained problems, we will solve the problem in two stages:
1. Fix $T$, and find $\rho(TB)$, the minimum absolute cross-correlation, $|\rho|$, achievable under the time-bandwidth constraint (and the optimal waveforms which achieve that $\rho$). Then, the capacity region for fixed $T$ is given by $C(T)$ in (5b) evaluated at $\rho = \rho(TB)$. This is because $C(T)$ depends on the signature waveforms only through the rate-sum constraint which is monotonic decreasing in $\rho$.

2. Take the union of the capacity regions found in the first stage over all $T$. Note that there is a minimum value of $T$ below which the time-bandwidth product is so small that no waveform can be found to satisfy the bandwidth constraint, and therefore, the capacity region is an empty set. Also, there is a maximum value of $T$ above which the allowed time-bandwidth product is so large that orthogonal signals can be assigned to both users, and therefore, the capacity region decreases with $T$ beyond that maximum value of $T$.

Theorem 3.2.

The capacity region of the two-user PAM white Gaussian multiple-access channel with noise power spectral density, $B$, $S_1$ and $S_2$, re-
specially, is given by

\[
C_G = \bigcup_{1 \leq R_1 \leq 2} \left\{ (R_1, R_2) : R_1 + R_2 \leq \frac{B}{\log \left( 1 + \frac{\pi^2}{\gamma^2} \right)} \right\}
\]

\[
= \left\{ (R_1, R_2) : R_1 + R_2 \leq \frac{B}{\log \left( 1 + \frac{\pi^2}{\gamma^2} \right)} \right\}
\]

The set of values of \( \gamma \) that are minor images of each other and as \( \gamma \) increases, they become more asymmetric so as to decrease the cross-correlation while maintaining the same rms bandwidth.

Finally, although the union in Theorem 3.2 is taken over \( \gamma \) in the interval \([1, \sqrt{3}]\), not every \( \gamma \) in that interval achieves some boundary points of \( C_G \). The set of values of \( \gamma \) that achieve some boundary points of \( C_G \) is a function of the signal-to-noise ratio, \( \frac{\pi^2}{\gamma^2} \), \( k = 1, 2 \). According to Figure 1, the boundary points in the segments AB and EF are achieved by \( \gamma = 1 \), while those in the segment CD are achieved by some \( \gamma_{\text{max}} \) in \([1, \sqrt{3}]\) depending on the signal-to-noise ratio. The boundary points in BC and DE are achieved by \( 1 \leq \gamma \leq \gamma_{\text{max}} \).

We now proceed to the optimal signal design problem under \( \alpha\)-lobe bandwidth constraint. Denote the prolate spheroidal wave functions (12], [13], and [14]) as \( \psi_i(TB, t) \) and the associated eigenvalues as \( \lambda_i(TB) \), i.e.,

\[
\lambda_i(TB) \psi_i(TB, t) = \int_{-\frac{T}{2}}^{\frac{T}{2}} \psi_i(TB, r) \frac{\sin[2\pi TB(t-r)]}{\pi(t-r)} dr
\]

for \( i = 0, 1, 2, \ldots \) and \( \lambda_0(TB) > \lambda_1(TB) > \lambda_2(TB) > \ldots \). It is known [8] that \( \psi_0(TB, \frac{T}{2} - \frac{t}{2}) \) and \( \psi_1(TB, \frac{T}{2} - \frac{t}{2}) \) are even and odd about \( \frac{T}{2} \) respectively and the set \( \left\{ \frac{1}{\sqrt{\pi TB}} \psi_i(TB, \frac{T}{2}) \right\} \) forms a complete orthonormal set in \([0, T] \). Also, \( \lambda_0(TB) \) and \( \lambda_1(TB) \) are continuous and monotonic increasing in \( TB \) (Figure 4).

Theorem 3.3.

For any \( 0 < \alpha < 1 \),

\[
|TB - \lambda_0^{-1}(\alpha)|, \text{ then the minimum cross-correlation, } \rho^*(TB), \text{ between any two unit-energy signals of duration } T, \text{ and } \alpha\text{-lobe bandwidth less than or equal to } B \text{ is}
\]

\[
\rho^*(TB) = \max \left\{ 0, \frac{2\alpha - \lambda_0(TB) - \lambda_1(TB)}{\lambda_0(TB) - \lambda_1(TB)} \right\}
\]

and is achieved by the signature waveforms

\[
\psi_1(TB) = \sqrt{\frac{2\alpha - \lambda_0(TB) - \lambda_1(TB)}{\lambda_0(TB) - \lambda_1(TB)}} \psi_0(TB, \frac{T}{2} - \frac{1}{2})
\]

\[
\psi_2(TB) = \sqrt{\frac{2\alpha - \lambda_0(TB) - \lambda_1(TB)}{\lambda_0(TB) - \lambda_1(TB)}} \psi_1(TB, \frac{T}{2} - \frac{1}{2})
\]

\[
TB < \lambda_0^{-1}(\alpha), \text{ then there exists no signal of duration } T \text{ and } \alpha\text{-lobe bandwidth less than or equal to } B \text{.}
\]

Proof.

As in Theorem 3.1, we would like to find a suitable complete orthonormal set in \([0, T] \). To that end, we rewrite the definition of \( \alpha\)-lobe bandwidth as

\[
\alpha \leq \int_{-B}^{B} |S(f)|^2 df
\]

\[
= \int_{-T}^{T} \int_{-T}^{T} S(t) s(t') dt' dt
\]

\[
= \int_{-T}^{T} \int_{-T}^{T} s(t) s(t') dt' dt
\]

\[
= \int_{-T}^{T} \int_{-T}^{T} \sin[2\pi TB(t-t')]^2 dt' dt
\]

\[
= \int_{-T}^{T} \int_{-T}^{T} \sin[2\pi TB(t-t')]^2 dt' dt
\]

\[
= \int_{-T}^{T} \int_{-T}^{T} \sin[2\pi TB(t-t')]^2 dt' dt
\]

Since the prolate spheroidal wave functions are eigenfunctions of the kernel \( TB \sin(t-t') \), a good choice for the complete orthonormal set will be the set of all prolate spheroidal wave functions.

For notational convenience, we will drop the explicit dependence on \( TB \) of the eigenvalues of the prolate spheroidal wave functions. If \( TB \geq \lambda_0^{-1}(\alpha) \), we can express any \( s(t) \) and \( s_1(t) \) in terms of \( \psi_i(t) \) as \( s(t) = \sum_{i=0}^{\infty} \tilde{a}_i \psi_i(t) \) and \( s_1(t) = \sum_{i=0}^{\infty} \tilde{a}_i \psi_i(t) \).

Using (27) and (28), we have

\[
\alpha \leq \int_{-B}^{B} |S(f)|^2 df
\]

\[
= \sum_{i=0}^{\infty} \tilde{a}_i^2 \lambda_0(TB) = \| \tilde{a}_i \|_2^2 \lambda_0(TB)
\]

\[
= \sum_{i=0}^{\infty} \tilde{a}_i^2 \lambda_0(TB) = \| \tilde{a}_i \|_2^2 \lambda_0(TB)
\]

where \( \Lambda = \text{diag}[\lambda_0, \lambda_1, \lambda_2, \ldots] \). Also, the cross-correlation matrix, \( H \), is

\[
H = A \Lambda A^T = \begin{bmatrix} a_1^T & a_2^T & \cdots \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}
\]

(30)
Similar to the rms case, we find the lower bound by maximizing the average over $k = 1, 2$ of the right hand side of (29).

Rewriting the average, we have

$$\frac{1}{2}\sum_{k=1}^{2} \xi_k \Lambda^T \Xi_k \Lambda = \frac{1}{2} \text{tr}(\Lambda \Lambda^T \Lambda)$$

Then, only the first two rows of $P^T$, and $\Xi \Lambda$ are distinct and decreasing down the diagonal axis, since the eigenvalues of $\Lambda^T \Lambda$ and $P^T \Lambda$ are the same, we have $\xi_1 = 1 + \rho$ and $\xi_2 = 1 - \rho$.

Now, let's denote $P$ as the $2 \times oo$ matrix formed by taking the maximum of the average is

$$\text{tr}(\Xi \Lambda^T \Xi \Lambda)$$

We will solve the maximization problem using the Lagrange multiplier method. We form the Lagrangian,

$$\sum_{k=1}^{2} \xi_k p_k^T \Lambda p_k + \sum_{k=1}^{2} \sum_{i=1}^{\infty} \lambda_m (p_k^T \Lambda p_k - \epsilon_m)$$

where $p_k$ is the $k$th row of $P$. Taking derivative with respect to $p_k$, we have

$$\xi_1 \Lambda p_1 + z_{11} p_1 + z_{12} p_2 = 0$$

and

$$\xi_2 \Lambda p_2 + z_{22} p_2 + z_{12} p_1 = 0$$

If we pre-multiply (34) by $p_1^T$ and (35) by $p_2^T$, we have $z_{12} = 0$ since $\xi_1 \neq \xi_2$; therefore, from (34) and (35), $p_1$ and $p_2$ are eigenvectors of $\Lambda$. Since $\Lambda$ is diagonal and the diagonal elements are distinct and decreasing down the diagonal axis, we have

$$P = \begin{bmatrix} 1 & \xi_2 \\ 0 & 0 \end{bmatrix}$$

Substituting back into (31), we show that the maximum value of (31) is $\lambda_0 - \frac{\rho}{2}^2 + \lambda_1 - \frac{\rho}{2}^2$. Comparing to (29), we have

$$\rho \leq \frac{1}{2} \left( (\lambda_0 + \lambda_1) + \rho (\lambda_0 - \lambda_1) \right)$$

or, together with $0 \leq \rho \leq 1$,

$$\rho \geq \max \left\{ 0, \frac{2a - \lambda_0 - \lambda_1}{\lambda_0 - \lambda_1} \right\}$$

The achievability of the lower bound can be verified, as in the rms case, by letting

$$A = \begin{bmatrix} \sqrt{\frac{\rho}{2}} & \sqrt{\frac{\rho}{2}} & 0 & \cdots \\ \sqrt{\frac{\rho}{2}} & -\sqrt{\frac{\rho}{2}} & 0 & \cdots \\ \sqrt{\frac{\rho}{2}} & \sqrt{\frac{\rho}{2}} & 0 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

which corresponds to the optimal set of signals stated in the theorem.

The proof of the second part of the theorem ($TB < \lambda_0^{-1} (\alpha)$) can be found in [13, p.54].

Theorem 3.4.

The capacity region of the two-user PAM white Gaussian multiple-access channel with noise power spectral density, $a$-fobe bandwidth and signal powers equal to $\sigma^2$, $B$, $S_1$, and $S_2$, respectively, is given by

$$C_F = \left\{ (R_1, R_2) : R_1 + R_2 \leq \frac{B}{T} \log \left[ 1 + \frac{\rho}{2}(\frac{S_1}{S_2})^2 \right], \frac{S_1}{S_2} \leq \frac{1}{2} \left( \frac{\lambda_0 (2TB)}{\lambda_1 (2TB)} + \lambda_1 (2TB) \right) \right\}$$

where $\alpha = \lambda_0 (2TB) = \frac{1}{2} \left( \lambda_0 (2TB) + \lambda_1 (2TB) \right)$.

Proof.

The proof is very similar to that in Theorem 3.2 where $3T = 2TB$. The lower limit of $\gamma$ is carried over from Theorem 3.3, while $\gamma_{\text{max}}$ is the smallest $\gamma$ such that $\rho(\frac{1}{2}) = 0$.

Notice that the range of $\gamma$ in taking the union is only a function of $\alpha$. In Figure 4, we show $\lambda_0 (2TB)$ and $\frac{1}{2} \left( \lambda_0 (2TB) + \lambda_1 (2TB) \right)$ vs the time-bandwidth product, and $\gamma_{\text{min}}$ and $\gamma_{\text{max}}$ can be obtained directly from the figure. Also, Figure 5 shows the capacity region, $C_F$, with the capacity region of the strictly band-limited channel, $C_C$. Similar comments to those we made in the rms case apply to the values of $\gamma$ that achieve the boundary points in the capacity region. However, we see that for sufficiently high $\alpha$, $C_F$ does not contain $C_C$ in contrast to the rms case.

Finally, in Figure 6, we show the signature waveforms which are, as expected, mirror images of each other. However, in contrast to the rms case where the signature waveforms must be zero at the end points to have finite rms bandwidth, the transmitted signal waveform in the $a$-fobe case may have jumps at $t = \pm T$.

References


Fig. 1. Capacity Regions in the rms case for SNR1=SNR2=20dB, B=1kHz

Fig. 2. Signature waveforms for two-user rms band-limited channel, TB=0.6

Fig. 3. Signature waveforms for two-user rms band-limited channel, TB=0.525

Fig. 4. Eigenvalues vs. time-bandwidth product

Fig. 5. Capacity Regions in the lobe case for SNR1=SNR2=20dB, B=1kHz

Fig. 6. Signature waveforms for two-user 0.8-lobe band-limited channel, TB=0.955
TOTAL CAPACITY OF THE RMS BANDLIMITED K-USER PAM SYNCHRONOUS CHANNEL

ROGER S. CHENG & SERGIO VERDÚ
Department of Electrical Engineering
Princeton University, Princeton, NJ 08544

ABSTRACT

Continuous-time additive white Gaussian noise channels with strictly time-limited and root mean square (RMS) bandlimited inputs are studied. RMS bandwidth is equal to the normalized second moment of the spectrum, which has proved to be a useful and analytically tractable measure of the bandwidth of strictly time-limited waveforms.

We find the Total Capacity (TC) of the K-user channel under total power and power-weighted average RMS bandwidth constraints. A lower bound to the TC under equal-power constraint is obtained. Total Capacity Ratio (TCR) is defined as the ratio of the K-user TC to K times the single-user capacity. Power (Bandwidth) efficiency is defined as the ratio of the effective power (bandwidth) to the actual power (bandwidth). The effective power (bandwidth) is the corresponding power (bandwidth) needed for a single user channel to achieve the same capacity. We find lower bounds to the TCR and efficiencies which indicate that savings in bandwidth compared to the FDMA scheme can be achieved by the CDMA scheme at the expense of more complicated decoding hardware.

1. INTRODUCTION

In this paper, we deal with the continuous-time Pulse Amplitude Modulation (PAM) Gaussian multiple-access channel (MAC). Each user is assigned a fixed deterministic continuous-time signature waveform, $s_k(t)$, which is time-limited to $[0, T]$ and is modulated linearly by the information stream. Assuming that the transmitters are symbol-synchronous, the channel can be expressed as

$$y(t) = \sum_{i=1}^{n} \sum_{k=1}^{K} b_k(i)s_k(t - iT) + n(t)$$

where $n(t)$ is white Gaussian noise with spectral density $\frac{N_0}{2}$ and $\{b_k(i)\}$ is the symbol stream transmitted by the $k^{th}$ user.

The capacity region of this channel has been found by Verdú [1] [2]. Denoting $W$ and $H$ as the diagonal matrix with the users' powers as its diagonal entries, and the cross-correlation matrix of the normalized signature waveforms, respectively, the capacity region is expressed as

$$C_V = \left\{ (R_1, R_2, \ldots, R_K) : \sum_{j \in J} R_j \leq \frac{1}{2T} \log(\det(\mathbf{I}_J + \frac{2T}{N_0} W_J H_J)) \quad \forall J \subset \{1, \ldots, K\} \right\}$$

This work was supported by the office of Naval Research under Contract N00014-87-K-0054 and by the National Science Foundation under PYI Grant ECSE-8857689.
The capacity region is maximized by orthogonal signature waveforms. However, under bandwidth constraints, orthogonal signature waveforms are not necessarily optimal since orthogonality can only be achieved by lowering the symbol rate, $1/T$.

There are many different bandwidth definitions [3]. In this paper, we concentrate on the root mean square (RMS) bandwidth because it is analytically tractable and can be applied to strictly time-limited signals. The RMS bandwidth was introduced by Gabor [4] and studied subsequently in [3], [5] and [6]. It is the square root of the second moment of the energy spectral density $(|S_{kk}(f)|^2)$ of the normalized signal which is proportional to the square root of the energy of its derivative.

In the two-user case, the capacity region of the RMS bandlimited PAM channel has been found in [7] and the total capacity (the maximum rate sum over the capacity region) is larger than the single-user capacity with the power equal to the sum of the users' powers. The gain in the total capacity from the single-user to the two-user case can be explained by the increase in the dimensionality of the signal set. We can consider the transmitted signal in a symbol interval as a signal drawn from a signal set. Then, the signal set in the single-user and the two-user case are one-dimensional and two-dimensional, respectively. From this viewpoint, it is easy to see that the total capacity increases as the number of users increases while the total power remains constant.

In this paper, we find the total capacity (TC) of the $K$-user channel under the total power constraint

$$\text{tr}(W) \leq W$$

and the power-weighted averaged RMS bandwidth constraint

$$\frac{1}{\text{tr}(W)} \sum_{k=1}^{K} W_{kk} \int_{-\infty}^{\infty} f^2 |S_{kk}(f)|^2 df \leq B^2$$

The power constraint is placed on the total power instead of the individual power since the later requires finding all possible sets of eigenvalues of a positive definite matrix with fixed diagonal entries which is, in general, intractable. The bandwidth constraint is justified because the power-weighted average RMS bandwidth is the RMS bandwidth of the power spectral density of the transmitted signal.

Several performance measures, Total Capacity Ratio, Power efficiency and Bandwidth efficiency, are defined and analyzed. Bounds and limiting values of these measures are also obtained.

**2. TOTAL CAPACITY**

Theorem 2.1.

The Total Capacity of the $K$-user RMS bandlimited PAM Gaussian MAC with total power and a power-weighted average RMS bandwidth constraints is

$$C(B, K, \Lambda) = \max \left\{ \sum_{n=1}^{N} \log[1 + h_n(\lambda)] \right\} \log e$$
where the maximization is over \(1 \leq N \leq K\), \(\frac{\gamma KA}{N} \leq \lambda\), and \(1 \leq \gamma \leq \sqrt{f_N}\) such that \(\gamma = \sqrt{f_N}\) iff \(\lambda = \frac{\gamma KA}{N}\), and

\[
\sum_{n=1}^{N} h_n(\lambda) = \gamma KA
\]

where

\[
h_n(\lambda) \triangleq \frac{N \lambda f_N + \gamma KA[\lambda(\gamma^2 - 1) - 1] - n^2(N\lambda - \gamma KA)}{N(f_N - 1 - \lambda) + \gamma^3 KA + n^2(N\lambda - \gamma KA)} > 0
\]

\[
f_N \triangleq \frac{1}{N} \sum_{n=1}^{N} n^2 = \frac{1}{6}(N + 1)(2N + 1)
\]

and the average signal-to-noise ratio is denoted by

\[
\Lambda \triangleq \frac{W}{KN_0B}
\]

Proof.

Since the signature waveforms are RMS bandlimited, and the set \(\{\phi_i(t, T)\}_{i=1}^{\infty}\) where

\[
\phi_i(t, T) \triangleq \begin{cases} \sqrt{2} \sin\left(\frac{\pi i t}{T}\right) & \text{if } t \in [0, T]; \\ 0 & \text{otherwise.} \end{cases}
\]

forms a complete orthonormal basis for all RMS bandlimited signals, we can write

\[
\begin{bmatrix} s_1(t) \\ s_2(t) \\ \vdots \\ s_K(t) \end{bmatrix} = A^T \Phi(t, T)
\]

where \(\Phi(t, T) \triangleq [\phi_1(t, T) \phi_2(t, T) \ldots]^T\). Then, the power-weighted average RMS bandwidth constraint can be written, via the Parseval's theorem, as

\[
\frac{1}{\text{tr}(W)} \sum_{k=1}^{K} W_{kk} \int_0^T \left[ \frac{d}{dt}s_k(t) \right]^2 dt = \frac{1}{(2T)^2 \text{tr}(W)} \text{tr}(WA^T \Pi A)
\]

\[
= \frac{1}{(2T)^2 \text{tr}(W)} \text{tr}(\Pi AW A^T)
\]

\[
\leq B^2
\]

From the capacity region in (2), it is clear that the total capacity is maximized when \(\text{tr}(W) = W\). We denote the time-bandwidth product by \(\gamma \triangleq 2BT\), the average signal-to-noise ratio by \(\Lambda \triangleq \frac{W}{KN_0B}\), and the eigenvalues of \(\frac{1}{2T}W H\) by \(\lambda_k\) such that \(\lambda_j \leq \lambda_i, \forall i < j \leq K\). Then, the total capacity becomes

\[
\text{TC}_V = \frac{B}{\gamma} \sum_{k=1}^{K} \log[1 + \lambda_k]
\]

and the power constraint becomes

\[
\sum_{k=1}^{K} \lambda_k = \gamma KA
\]
Since the eigenvalues of $\frac{2\pi}{N_0} WH$ are also the eigenvalues of $\frac{2\pi}{N_0} A W A^T$, and once $\{\lambda_k\}_{k=1}^K$ are fixed, the left hand side of (12) is minimized when $AWA^T$ is diagonal with decreasing diagonal entries, we can rewrite (12) as

$$\sum_{k=1}^{K} k^2 \lambda_k \leq \gamma^3 K A$$

(15)

For fixed $T$, the total capacity is found by maximizing (13) over all $\lambda_k \geq 0$, $k = 1,\ldots,K$ under the constraints (14) and (15). Using the Kuhn-Tucker Theorem, we form the Lagrangian

$$-\sum_{k=1}^{K} \log[1 + \lambda_k] + \mu(\sum_{k=1}^{K} \lambda_k - \gamma K A) + \gamma(\sum_{k=1}^{K} k^2 \lambda_k - \gamma^3 K A)$$

(16)

and obtain the necessary conditions:

$$\lambda_n = \frac{1}{(1 + \lambda_n)^2} - 1 > 0 \quad n = 1,\ldots,N.$$  

(17)

and $\lambda_n = 0$ for all $n > N$,

$$\gamma(\sum_{n=1}^{N} n^2 \lambda_n - \gamma^3 K A) = 0$$

(18)

and $0 \leq \gamma$.

Rewriting (17) as $(x + y) (1 + \lambda_n) = 1$, and summing over all $n$, we have, from (14) and (18),

$$(N + \gamma K A) x + (N f_N + \gamma^3 K A) y = N$$

(19)

Particularizing (17) to $n = 1$, and substituting in (19), we have

$$y = \frac{N \lambda_1 - \gamma K A}{(1 + \lambda_1)(N (f_N - 1) + \gamma K A (\gamma^2 - 1))}$$

(20)

and

$$z = \frac{N(f_N - 1 - \lambda_1) + \gamma^3 K A}{(1 + \lambda_1)(N (f_N - 1) + \gamma K A (\gamma^2 - 1))}$$

(21)

Substituting (20) and (21) into (17), and denoting $\lambda_1$ by $\lambda$ and $\lambda_n$ by $h_n(\lambda)$, we have (7), and the power constraint in (14) becomes (6).

When $\gamma = 0$, $\lambda = h_n(\lambda) = \frac{N \lambda_1}{N}$ for all $n = 1,\ldots,N$. Upon substituting into (15), we have $\sqrt{f_N} \leq \gamma$. Since the total capacity becomes $\frac{BN}{\gamma} \log[1 + \frac{N \lambda_1}{N}]$ which is monotonically decreasing in $\gamma$, the optimal $\gamma$ is equal to $\sqrt{f_N}$ and (15) is satisfied with equality. If we rewrite (7) and sum up over all $n$, we have

$$(N \lambda_1 - \gamma K A)(\sum_{n=1}^{N} n^2 h_n(\lambda) - \gamma^3 K A) = 0$$

(22)

When $0 < \gamma$, $\frac{N \lambda_1}{N} < \lambda$, and from (22), (15) is again satisfied with equality. Therefore, if we require $\gamma = \sqrt{f_N}$ iff $\lambda = \frac{N \lambda_1}{N}$, (15) is superfluous. Finally, specifying the range of $\gamma$ and $\lambda$ and the condition that $\gamma = \sqrt{f_N}$ iff $\lambda = \frac{N \lambda_1}{N}$, we have the desired result.

This theorem gives the exact calculation needed for the TC. The main reason why we cannot obtain a simpler solution is the lack of a closed form expression of $\sum_{n=1}^{N} \frac{1}{a + bn \gamma}$. However,
despite the complicated expression, the TC can be computed once the average signal-to-noise ratio, \( \Lambda \), and \( K \) is given. In Figure 1, we show the TC with different values of \( K \) and \( \Lambda \).

For a given \( W \), we show that any set of signature waveforms, with \( A \) such that \( AWAT \) is a diagonal matrix with the \( n \)th diagonal entry equal to \( h_n(\lambda) \), is optimal. However, such an \( A \) does not always exist for any arbitrarily given \( W \). For fixed total power, \( W \), finding the set of \( W \) where \( A \) exists is equivalent to finding the possible set of diagonal entries of a positive definite matrix with fixed eigenvalues, which seems intractable. Reversing the problem, one may want to fix the \( W \) and find the total capacity. In general, this is equivalent to finding the possible set of eigenvalues of a positive definite matrix with fixed diagonal entries, which is again intractable.

In the following theorem, we give a lower bound to the TC in the equal-power case where \( W = \frac{W}{K} I \). Clearly, this is also a lower bound to the capacity of the original channel with the total power constraint in Theorem 2.1.

**Theorem 2.2.**

The lower bound to the Total Capacity when the users' powers are the same is

\[
TC_{BP}(W) \geq \max_{1 \leq \gamma \leq \sqrt{K}} \frac{B}{\gamma} \log \left\{ \frac{1 + \gamma \Lambda^2}{f_K - 1} \right\} \left\{ \frac{1 + \gamma \Lambda^2}{f_K - 1} \right\}^{K-1}
\]

(23)

**Proof.**

The lower bound is found by exhibiting a symmetric positive definite matrix \( H \), such that the total capacity for that particular signature waveform set is easy to find. We let \( H \) be

\[
H = \begin{bmatrix}
1 & \rho & \cdots & \rho \\
\rho & 1 & \cdots & \vdots \\
\vdots & \ddots & \ddots & \rho \\
\rho & \cdots & \rho & 1
\end{bmatrix}
\]

(24)

the eigenvalues of \( H \) with \( 0 \leq \rho \leq 1 \) to be be specified in the sequel are \( 1 + (K - 1)\rho \) and \( 1 - \rho \) with multiplicity \( K - 1 \). Then, the total capacity under the equal-power constraint becomes

\[
TC_V = \frac{B}{\gamma} \log \left\{ \det(I_K + \gamma \Lambda H) \right\}
\]

\[
= \frac{B}{\gamma} \log \left\{ \left[ 1 + \gamma \Lambda(1 + (K - 1)\rho) \right] \left[ 1 + \gamma \Lambda(1 - \rho) \right]^{K-1} \right\}
\]

(25)

while the bandwidth constraint (15) becomes

\[
\rho \geq \frac{f_K - \gamma^2}{f_K - 1}
\]

(26)

Since (25) is monotonically decreasing in \( \rho \) when \( 0 \leq \rho \leq 1 \), the TC is maximized when \( \rho \) achieves equality in (26). Substituting \( \rho \) from (26) with equality into (25), and maximizing over all \( \gamma \), we have (23).
In Figure 1, we plot the lower bound to the TCep for different values of $K$ and $\Lambda$. Since the TC under the total power constraint serves as an upper bound to the TCep, Figure 1 gives a tight upper and lower bound to the Total Capacity of an equal-power constrained channel, for moderate number of users.

As a performance measure, we define the Total Capacity Ratio (TCR) as the ratio of the $K$-user TC to $K$ times the single-user capacity with the same RMS bandwidth and average signal-to-noise ratio constraints. Since the single-user capacity of a RMS bandlimited PAM channel is equal to $B \log(1 + \Lambda)$ (see [7]), the TCR can be written as

$$\text{TCR}(K, \Lambda) \triangleq \frac{\text{TC}(B, K, \Lambda)}{KB \log(1 + \Lambda)}$$

(27)

The TCR gives the ratio of the capacity available to an average user (when the channel is shared by $K$ users) to the single-user capacity. In other words, it measures, from the user's viewpoint, the ratio of the average user capacity in a multi-user channel to the capacity in a single-user channel. Notice that the TCR depends only on $K$ and $\Lambda$, and is independent of $B$.

Using the lower bound in Theorem 2.2, we obtain a lower bound to the TCR under the equal-power constraint for all signal-to-noise ratios.

**Corollary 2.1.**

A lower bound to the TCR under the equal-power constraint for all signal-to-noise ratio is

$$\text{TCR}(K, \Lambda) \geq \frac{1}{\gamma}$$

(28)

where $\gamma$ is the positive real root of the equation

$$\gamma(\gamma^2 - 1) = fK - 1$$

(29)

**Proof.**

In order to obtain (28), we simply substitute $\gamma$ from (29) into (23) and (27). Since there is one and only one real positive solution in (29), there is no ambiguity in the value of $\gamma$.

In Figure 2, we show the TCR under the total power constraint and the lower bound to the TCR under the equal-power constraint for different number of users and different average signal-to-noise ratios.

### 3. EFFICIENCIES

The TCR gives the performance degradation, from the user's viewpoint, when a bandlimited channel is shared by $K$ users instead of a single user. A natural question to be asked is "How to maintain the same rate in the presence of other users?" If we want to maintain the same information rate, we have to modify some of the parameters. In the following, we will analyze two alternatives. First, we increase the signal-to-noise ratio by increasing the power while the bandwidth remains constant. Second, we increase the bandwidth of the channel while the power of each user remains the same.
The power efficiency, denoted by \( \eta_p(K, \Lambda) \), is defined as

\[
\eta_p(K, \Lambda) \triangleq e^{\frac{TC(B,K,\Lambda)}{BK} - 1}
\]  

(30)

or, equivalently, implicitly as

\[
TC(B,K,\Lambda) = BK \log(1 + \eta_p(K, \Lambda) \Lambda)
\]

(31)

The bandwidth efficiency, denoted by \( \eta_B(K, \Lambda) \), is defined implicitly as

\[
TC(B,K,\Lambda) = \eta_B(K, \Lambda) BK \log(1 + \frac{\Lambda}{\eta_B(K, \Lambda)})
\]

(32)

The power efficiency, \( \eta_p(K, \Lambda) \) (bandwidth efficiency, \( \eta_B(K, \Lambda) \)) gives the ratio of the effective power (bandwidth) to the actual power (bandwidth) when the actual signal-to-noise ratio is \( \Lambda \). The actual power (bandwidth) is the power (bandwidth) used in transmission while the effective power (bandwidth) is the corresponding power (bandwidth) needed for a single user channel to achieve the same capacity. In other words, \(-10 \log[\eta_p(K, \Lambda)]\) gives the power in db that we have to add to each user in order to maintain the single-user capacity. Similarly, \(1/\eta_B(K, \Lambda)\) gives the ratio that we have to increase the bandwidth in order to maintain the same information rate.

**Theorem 3.1.**

The power efficiency satisfies,

\[
\lim_{\Lambda \to \infty} \eta_p(K, \Lambda) = 0
\]

(33)

A lower bound to the bandwidth efficiency, \( \eta_B(K, \Lambda) \), for all signal-to-noise ratio under the equal-power constraint is

\[
\eta_B(K, \Lambda) \geq \frac{1}{\sqrt{f_K}}
\]

(34)

where \( f_K \) is defined in (8).

**Proof.**

From the definition of \( \eta_p(K, \Lambda) \), we have,

\[
\prod_{n=1}^{N} (1 + h_n(\Lambda)) = [1 + \eta_p(K, \Lambda) \Lambda]^{K\gamma}
\]

(35)

where \( N, h_n(\Lambda) \) and \( \gamma \) are all optimally selected for that \( \Lambda \). Subtracting (14) from (15), it is easy to get

\[
h_n(\Lambda) \leq \gamma K \Lambda (\gamma^2 - 1) \quad n = 1, \ldots, N.
\]

(36)

Substituting (36) into (35), and dividing both side by \( \Lambda^{K\gamma} \), we have, in the limit as \( \Lambda \to \infty \),

\[
\lim_{\Lambda \to \infty} \Lambda^{-K\gamma} \prod_{n=1}^{N} \left( \frac{1}{\Lambda} + \gamma K (\gamma^2 - 1) \right) \geq \eta_p(K, \Lambda)^{K\gamma}
\]

(37)
If \( \gamma \to 1 \) as \( \Lambda \to \infty \), the second factor on the left hand side of (37) tends to 0, while the first factor tends either to 0 (\( N < K \)) or to 1 (\( N = K \)). On the other hand, if \( \gamma \to \alpha > 1 \) as \( \Lambda \to \infty \), the first factor on the left hand side of (37) tends to 0 for any \( N \leq K \), while the second factor is bounded. Therefore, in both cases, the left hand side tends to 0 and since \( 1 \leq \gamma \leq \sqrt{\frac{1}{K}} \), we have \( \eta_p(K, \Lambda) \to 0 \) as \( \Lambda \to \infty \).

Substituting \( \gamma = \sqrt{\frac{1}{\Lambda}} \) in Theorem 2.2, we have

\[
TC_{EP}(B, K, \Lambda) \geq \frac{1}{\sqrt{\frac{1}{\Lambda}}} BK \log[1 + \sqrt{\frac{1}{\Lambda}}] \tag{38}
\]

Since the right hand side of (32) is monotonically increasing in \( \eta_B(K, \Lambda) \), we have (34) when compared to (38).

The TC is obtained by optimizing the balance between the “symbol rate” factor, \( B/\gamma \), and the “information sent per symbol” factor, \( \log[\cdot] \). As the average signal-to-noise ratio tends to infinity, the “symbol rate” factor dominates and the optimal users' signature waveforms are asymptotically identical. Then, the product term of the signal-to-noise ratios inside the log function in the TC becomes relatively small, and the asymptotic power efficiency is equal to zero. The bandwidth efficiency indicates the increase in bandwidth needed to maintain the same user rate when a single-user channel is shared by \( K \) users.

In Figure 3 and 4, we plot the Power and Bandwidth efficiency for different values of \( K \) and \( \Lambda \). Also, in the same graphs, we show the lower bound to the efficiencies for the equal-power constrained channel. It shows that regardless of the signal-to-noise ratios, increasing the bandwidth by a factor of 10, we can accommodate about 50 users on the multi-user channel. This indicates a 80 percent reduction in the bandwidth required by Frequency Division Multiple Access (FDMA). Clearly, the tradeoff is a more complicated demodulating and decoding process in the Synchronous Code Division Multiple Access (CDMA) channel, which is a special case of the current model.

REFERENCES


Figure 3. Bandwidth efficiency for different number of users and average signal-to-noise ratios.

Figure 4. Power efficiency for different number of users and average signal-to-noise ratios.