ASYNCHRONOUS MULTIPLE ACCESS CHANNEL CAPACITY

BY

THOMAS M. COVER
ROBERT J. McELIECE
EDWARD C. POSNER

TECHNICAL REPORT NO. 35
DECEMBER 1978

PREPARED UNDER THE AUSPICES
OF
NATIONAL SCIENCE FOUNDATION
GRANT ENG 76-03684

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
ASYNCHRONOUS MULTIPLE ACCESS CHANNEL CAPACITY

Thomas M. Cover†
Robert J. McEliece††* Edward C. Posner†††*

TECHNICAL REPORT NO. 35
DECEMBER 1978

PREPARED UNDER THE AUSPICES
OF
NATIONAL SCIENCE FOUNDATION
GRANT ENG 76-03684

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

†Stanford University. This work was supported by NSF Grant ENG 76-03684.
††Jet Propulsion Laboratory, California Institute of Technology.
†††Dept. of Electrical Engineering and Jet Propulsion Laboratory, California Institute of Technology.
*The work of the second two authors represents, in part, one phase of research carried out at the Jet Propulsion Laboratory of the California Institute of Technology for the National Aeronautics and Space Administration under Contract N157-100.
Asynchronous Multiple Access Channel Capacity

ABSTRACT

The capacity region for the multiple access channel without time synchronization at the transmitters and receivers is shown to be the same as the known capacity region for the ordinary multiple access channel. An optimal code for the unsynchronized MAC uses time-sharing, at each transmitter, of two optimal codes for the ordinary MAC. Optimal decoding uses maximum likelihood decoding over shifts of the hypothesized transmitter words.
0. Introduction

A two-user multiple access channel \( p(y|x_1,x_2) \) has two senders \( x_1 \) and \( x_2 \), and a receiver \( y \). When two users are attempting to use the same channel, there are two kinds of cooperation that make physical sense.

The first is a strategic cooperation—both the senders and the receiver agree on the codebooks that will be used and the decoding mapping for the receiver. This is the usual assumption for the Shannon channel.

The second possible cooperation occurs when the independent messages are actually sent. If both senders somehow know the two messages \( W_1 \in 2^{nR_1} \), \( W_2 \in 2^{nR_2} \) to be sent, then they can use the channel cooperatively as an ordinary 1-sender Shannon channel with capacity

\[
R_1 + R_2 \leq C = \max_{p(x_1,x_2)} I(X_1,X_2;Y). \tag{1}
\]

Much more commonly, however, it is the case that \( W_1 \) is known only to \( x_1 \) and \( W_2 \) only to \( x_2 \), thus allowing only rates \((R_1,R_2)\) satisfying

\[
R_1 \leq I(X_1;Y|X_2)
\]

\[
R_2 \leq I(X_2;Y|X_1)
\]

\[
R_1 + R_2 \leq I(X_1,X_2;Y)
\]

for \( p(x_1,x_2) = p(x_1)p(x_2) \). This independent-user region is the multiple access channel capacity found by Ahlswede [1] and Liao [2]. We shall only be concerned with the independent-user capacity of Eq. (2).

Returning now to the strategic cooperation used to derive Eq. (1) and
Eq. (2), we see that implicit use is made of time-synchronization. Even simple time-sharing, in which each sender is quiet while the other sends, requires a common time base. What happens to the capacity region in Eq. (2) when there is a time uncertainty for the users and the receiver? Clearly new codebooks may have to be constructed. Moreover, the interference of codewords from the users' respective codebooks cannot now be cooperatively allowed for. For example, the strategic cooperation by time-sharing may be ruined by unavoidable overlapping of the transmission periods. Finally, even the receiver must revise his decoding strategy in order to look for the joint transmissions with arbitrary time shifts.

We shall show that the capacity region is unaffected by lack of synchronization.
Section 1: Definitions and Review of Multiple Access Capacity

An \( \{ (M_1, M_2), n, P_n \} \) code for the (discrete memoryless) multiple access channel \( \{ X_1, X_2, Y, p(y|x_1, x_2) \} \) is a pair of maps \( x_1 : \{1, 2, \ldots, M_1 \} \rightarrow X_1^n \), \( x_2 : \{1, 2, \ldots, M_2 \} \rightarrow X_2^n \), and a map \( g : Y^n \rightarrow \{1, 2, \ldots, M_1 \} \times \{1, 2, \ldots, M_2 \} \).

The probability of error \( P_n \) of this code is defined by

\[
P_n = P\{ g(Y^n) \neq (i, j) \} = \frac{1}{M_1 M_2} \sum_{i,j} P\{ g(Y^n) \neq (i, j) \}, \tag{1}
\]

where

\[
p(y^n|i, j) = \prod_{k=1}^{n} p(y_k|x_{1k}(i), x_{2k}(j)). \tag{2}
\]

A pair of rates \( (R_1, R_2) \) is said to be achievable if there exists a sequence of \( \{ (2^{-n R_1}, 2^{-n R_2}), n, P_n \} \) codes with \( P_n \rightarrow 0 \). The capacity region \( C^* \) is the closure of the set of all achievable rates \( (R_1, R_2) \).

Theorem 1 establishes the capacity region \( C^* \). An alternative proof to those in [1, 2] will be given as a model for the subsequent proof that \( C^* \) remains unchanged when there is no time reference, i.e., no synchronization.

For the multiple access channel without synchronization, the error criterion is more stringent—the decoding must be correct for arbitrary shifts of the \( X_1(i) \) and \( X_2(j) \), where the shifts are imbedded in arbitrary sequences from \( X_1 \) and \( X_2 \), respectively. This is the appropriate error criterion. To average over all shifts is not appropriate because each user will want to be sure that most of his blocks are in fact received correctly. The rejected relaxed definition would permit a positive probability that almost none of his blocks are received correctly. We shall first define a code for a channel in which the shifts are known not to exceed \( d \). Then a rate region independent of \( d \) will be defined.
An \((M_1,M_2,n,d,P_n)\) code for a multiple access channel with maximum relative delay \(d\) is a pair of maps

\[
x_1 : \{1,2,\ldots,M_1\} \rightarrow X_1^n,
\]

\[
x_2 : \{1,2,\ldots,M_2\} \rightarrow X_2^n,
\]

and a map

\[
g : Y^{n+d} \rightarrow \{1,2,\ldots,M_1\} \times \{1,2,\ldots,M_2\}.
\]

The probability of error of this code is

\[
P_n = \frac{1}{M_1 M_2} \sum_{i,j} \max_{1 \leq d_1 \leq d} \max_{1 \leq d_2 \leq d} P\{g(Y^{n+d}) = (i,j) \mid d_1, d_2\},
\]

\[
\sum_{X_1 \in X_1, \tilde{X}_1 \in \tilde{X}_1} \sum_{X_2 \in X_2, \tilde{X}_2 \in \tilde{X}_2}
\]

where

\[
p(y^{n+d} \mid i,j,d_1,d_2) = \prod_{k=1}^{n+d} p(y_k \mid x_1^{k-d_1}(i), x_2^{k-d_2}(j))
\]

and the first \(d_1\) symbols and last \(d-d_1\) symbols of \(x_1(i)\) are arbitrary sequences \(\tilde{x}_1 \in X_1^{d_1}, \tilde{x}_1 \in \tilde{X}_1^{d-d_1}\) (with a parallel condition for \(d_2\) and \(x_2(j)\)).

We shall explain this definition. Here \(X_1\) and \(X_2\) are the channel input alphabets for the two users. The integers \(M_1\) and \(M_2\) are equal to \(2^{nR_1}\) and \(2^{nR_2}\), where \(R_1\) and \(R_2\) are the rates of the two users. The mappings \(x_1\) and \(x_2\) are the encoding mappings which produce codewords of length \(n\) from the index message. The probability \(p(y \mid x_1,x_2)\) is the transition probability for the channel, which gives the probability of output \(y\) when \(x_1\) and \(x_2\) are the inputs of the two users.
The set $Y$ is the output alphabet. We assume that there is one receiver trying to reconstruct the two inputs.

The relative delay $d$ is the maximum amount by which the two messages are assumed to be out of synchronism relative to a known or prearranged time. In reasonable classes of applications, we can assume such a $d$ exists. The map $g$ is the decoding operation which can commence when $n + d$ symbols have been received.

The error probability $P_n$ is a two-dimensional average word error probability. It assumes all $M_1 M_2$ pairs $(i,j)$ of inputs are equally likely, and takes an average over these pairs. For each pair, the term to be averaged is the maximum over $d$ of the probability of decoding at least one of $i$ and $j$ incorrectly, given that the delay of the first user relative to the prearranged start of block codeword is $d_1$, and of the second user is $d_2$. Here $d_1$ and $d_2$ are constrained to be at most the maximum relative delay $d$. The maximum is also taken over all possible "head" and "tail" sequences from prior or subsequent codewords. These intrude into the $n + d$ symbols input to the decoder to pad out the length from $n$, the code block length, to $n + d$, the block length plus the maximum relative delay.

A pair of rates $(R_1, R_2)$ will be said to be achievable if there exists a sequence $d_n \to \infty$ and a sequence of \(((2^n R_1, 2^n R_2, n, d_n, P_n)\) codes with $P_n \to 0$ as $n \to \infty$. This means that we can guarantee arbitrarily low word error probabilities at these rates, no matter how large the relative delay bound maybe as long as we know a bound for the relative delay.
A stronger sense of achievability independent of knowledge of a relative delay bound will be demonstrated in Section 3.

Finally the capacity region $C$ is as usual the closure of the set of achievable rates.

**Example:** (The binary erasure multiple access channel.)

Let $\overline{X}_1 = \overline{X}_2 = \{0,1\}$, $Y = \{0,1,2\}$. Consider the deterministic channel $y = x_1 + x_2$. For obvious reasons, $y = 1$ is called an erasure. The capacity region (see Theorem 1) is given by $R_1 \leq 1$, $R_2 \leq 1$, $R_1 + R_2 \leq 1.5$. See Gaarder and Wolf [5].

A new proof of the direct part of the following theorem will be given and used as the model for the proof in the next section. This is the known result for the synchronized multiple access channel.

**Theorem:** (Ahlswede, Liao).

The capacity region of the multiple access channel is given by the set of all rates in the convex closure of the set of rates $(R_1, R_2)$ satisfying

$$R_1 < I(X_1;Y|X_2)$$

$$R_2 < I(X_2;Y|X_1)$$

$$R_1 + R_2 < I(X_1,X_2;Y),$$

for some $p(x_1,x_2,y) = p(x_1)p(x_2)p(y|x_1,x_2)$.

**Proof:** Fix $p(x_1), p(x_2)$. Let $p(x_1,x_2) = p(x_1)p(x_2)$. Choose a random code of $2^{nR_1} x_1$ 's $\in \overline{X}_1^n$ i.i.d. $\sim \prod_{i=1}^{n} p(x_{1i})$, and independently
choose $2^nR_2$ $x_2$'s $\in X_2^n$ i.i.d. $\sim \prod_{i=1}^{n} p(x_{2i})$.

Define the set $A^n_\varepsilon$ of $\varepsilon$-typical $(x_1, x_2, y)$ triples by

$$A^n_\varepsilon = \{(x_1, x_2, y) \in X_1^n \times X_2^n \times Y^n :$$

$$\left| -\frac{1}{n} \log p(x_1) - H(x_1) \right| < \varepsilon ,$$

$$\left| -\frac{1}{n} \log p(x_2) - H(x_2) \right| < \varepsilon ,$$

$$\left| -\frac{1}{n} \log p(y) - H(y) \right| < \varepsilon ,$$

$$\left| -\frac{1}{n} \log p(x_1, x_2) - H(x_1, x_2) \right| < \varepsilon ,$$

$$\left| -\frac{1}{n} \log p(x_1, y) - H(x_1, y) \right| < \varepsilon ,$$

$$\left| -\frac{1}{n} \log p(x_2, y) - H(x_2, y) \right| < \varepsilon ,$$

$$\left| -\frac{1}{n} \log p(x_1, x_2, y) - H(x_1, x_2, y) \right| < \varepsilon \}$$

(6)

We note that $(x_1, x_2, y)$ are i.i.d. $\sim p(x_1)p(x_2)p(y|x_1, x_2)$. Thus, by $\log \prod = \sum \log$ and the law of large numbers, $-\frac{1}{n} \log p(\cdot) \to H$, with probability one, for each of the 7 constraints in (6). Hence there exists an $n_0$ such that, for $n \geq n_0$, $P(A^n_\varepsilon) \geq 1 - \varepsilon$. Also, it can be seen from (2) that

$$|A^n_\varepsilon| \leq 2^{-n(H(x_1, x_2, y) + \varepsilon)}$$

(7)

For decoding, given $y$, simply choose the pair $(i, j)$ such that

$$(x_1(i), x_2(j), y) \in A^n_\varepsilon,$$

(8)

if such an $(i, j) \in \{1, 2, \ldots, 2^nR_1\} \times \{1, 2, \ldots, 2^nR_2\}$ exists and is unique—otherwise declare an error.
By symmetry of the random code construction, the probability of error
(averaged over the random code) is independent of the index \((i,j)\) sent.
Thus without loss of generality assume that \((i,j) = (1,1)\) is sent. Con-
sider the events

\[
E_{ij} = \{(X_1(i), X_2(j), y) \in A^n_\varepsilon \}. \tag{9}
\]

Then by the union of events bound

\[
P_n = P\left(\bigcup_{(i,j) \neq (1,1)} E_{ij}\right) \leq P(E_{11}^C) + \sum_{i \neq 1} P(E_{i1}) + \sum_{j \neq 1} P(E_{1j}) + \sum_{i \neq 1, j \neq 1} P(E_{ij}). \tag{10}
\]

Assume henceforth that \(n \geq n_0\). Thus \(P(E_{11}^C) \leq \varepsilon\).

Next, for \(i \neq 1\),

\[
P(E_{i1}) = \sum_{(x_1, x_2, y) \in A} P(x_1, x_2, y) \]

\[
= \sum_A p(x_1)p(x_2, y) \]

\[
\leq |A| \cdot \left(2^{-n(H(X_1) - \varepsilon)} \cdot 2^{-n(H(X_2, Y) - \varepsilon)} \cdot 2^{-n(I(X_1; X_2, Y) - 3\varepsilon)} \right) \leq 2^{-n(I(X_1; Y|X_2) - 3\varepsilon)}.
\]

where the 1st equality is by definition of \(E_{i1}\), the 2nd from \(i \neq 1\)
which implies the independence of \(X_1\) from \((X_2, Y)\), the 3rd from the
definition of \(A = A^n_\varepsilon\), the 4th from (7), and the last equality follows
from the independence of $X_1$ and $X_2$ and the identity $I(X_1; Y|X_2) = I(X_1; Y, X_2) - I(X_1; X_2)$.

Similarly, for $j \neq 1$,

$$P(E_{1j}) \leq 2^{-n(I(X_1; Y|X_2) - 3\varepsilon)}$$

(12)

and, for $i \neq 1, j \neq 1$,

$$P(E_{ij}) \leq 2^{-n(I(X_1, X_2; Y) - 3\varepsilon)}$$

(13)

Hence, returning to (10) we have

$$P_n \leq \varepsilon + 2^{nR_1} -n(I(X_1; Y|X_2) - 3\varepsilon) + 2^{nR_2} -n(I(X_2; Y|X_1) - 3\varepsilon) + 2^{n(R_1 + R_2)} -n(I(X_1, X_2; Y) - 3\varepsilon).$$

(14)

Thus the conditions of the theorem cause each term to tend to zero as $n \to \infty$.

Time sharing (allowable because of time synchronization) achieves any $(R_1, R_2)$ in the convex hull, and the direct part of the theorem is proved.

The converse is well known and will not be repeated.
Section 2: Capacity without Synchronization.

We shall show that the same sequence of random codes causing \( P_n \to 0 \) in the previous section will also cause \( P_n \to 0 \) if the words are not synchronized. The construction is a form of time-sharing that works in the absence of synchronization. We thus obtain the same capacity region as if we had time synchronization between the two users. There is no decrease in the capacity region due to lack of time synchronization.

Let \( d_1 \) and \( d_2 \) be fixed nonnegative integers unknown to the receiver. Sender \( k, k=1,2 \), sends an arbitrary sequence of \( d_k \) symbols from alphabet \( X_k \) followed by codeword \( x_k(i_k) \) of blocklength \( n \), followed by more arbitrary symbols from \( X_k \).

We shall first assume that the receiver knows a bound \( d \) on the delays, i.e., \( d_1, d_2 \leq d \). Hence the receiver inspects \( y \in Y^{n+d} \) for the presence of \( x_1(i_1), x_2(i_2) \) imbedded with arbitrary shifts in arbitrary transmitter sequences. Later we shall remove the receiver's knowledge of \( d \).

In the general case for a multiple access channel without synchronization, it will be necessary to obtain the convex combination of rate points \( (R_1, R_2) \) and \( (R'_1, R'_2) \) to achieve the point \( (R_1^0, R_2^0) = \alpha(R_1, R_2) + \bar{\alpha}(R'_1, R'_2), \ 0 < \alpha < 1, \bar{\alpha} = 1 - \alpha \). This time sharing is necessitated by the possible lack of convexity of the union of the set of \( (R_1, R_2) \) satisfying Eq. (1).

Let \( p_1(x_1, x_2) = p_1(x_1)p_1(x_2) \) induce a region given in (1) that has \( (R_1, R_2) \) as an extreme point, and let \( p_2(x_1, x_2) = p_2(x_1)p_2(x_2) \) induce a region that has \( (R'_1, R'_2) \) as an extreme point. Using the random coding procedure of Section 1, generate a random \( (2^{\alpha nR_1}, 2^{\alpha nR_2}, \alpha n) \) code according to \( p_1 \) and a random \( (2^{\alpha nR'_1}, 2^{\alpha nR'_2}, \alpha n) \) code according to \( p_2 \).

The sent and received sequences will then look like this for some \( d_1, d_2 \) :
The crucial point is that \( x_1(i) \) and \( x_2(j) \) will have substantial overlap, and the region of overlap can be prespecified. This overlap will be sufficient to detect typicality and reject atypicality. In fact inspection of Figure 2 shows the overlap regions to be of lengths at least \( \alpha n - d \) and \( \bar{\alpha} n - d \), independent of \( d_1, d_2 \) for \( 0 \leq d_1, d_2 \leq d \).

The decoding is as follows. We must look for codewords in all their possible shifts, i.e., up to maximum delay \( d \). Let the maximal delay \( d \) be fixed and known. Let \( \tau^k \) denote a cyclic shift \( k \) units to the right.
of a given \((n+d)\)-tuple. Let \(W_{\alpha,n,d} = W\) denote the window function that
inspects only the values of the vector in the first window specified in
Figure 2. Note that no dummy symbols could be in the window. Define the
set of \(p_1(x_1, x_2, y)\)-typical sequences \(A_{\varepsilon}^1\) only over the \((an-d)\) coordinates
specified in the first window in Fig. 2. Thus, for example, there are at most
\((an-d)(H_1(x_1, x_2, y) + \varepsilon)\)
jointly typical triples in the first window, and each
triple in \(A_{\varepsilon}^1\) has probability \(\leq 2^{-(an-d)(H_1(x_1, x_2, y) - \varepsilon)}\). The second
window will be treated by similar techniques.

Again, without loss of generality assume that \((1,1)\) was sent and that
the delays were \(d_1, d_2\), where \(1 \leq d_1, d_2 \leq d\). To place an upper bound on
the probability of error \(p_n^1\) in the first code, define the events, for
\(\begin{align*}
1 \leq k_1, k_2 \leq d, \ i \in \{1, 2, \ldots, 2^{nR_1}\}, \ j \in \{1, 2, \ldots, 2^{nR_2}\}, \ E_{k_1, k_2, i, j} = \\
&\text{the event that } \left(\tau_1^{k_1} x_1(i), \tau_2^{k_2} x_2(j), \tau (y)\right) \in A_{\varepsilon} .
\end{align*}\)
That is, the
event \(E_{k_1, k_2, i, j}\) occurs if the \(k_1\)-shift of transmitter 1's \(i\)th codeword
\(x_1(i)\) and the \(k_2\)-shift of transmitter 2's \(j\)th codeword are seen to be
jointly typical with \(y\) in the first window.

Since we have assumed \((1,1)\) was sent with delay \((d_1, d_2)\), an error
will occur if \(U_{k_1, k_2} E_{k_1, k_2, 1, 1}\) does not occur (i.e., \((1,1)\) is not a
candidate) or if for some \(1 \leq k_1, k_2 \leq d\) and some \((i, j) \neq (1,1)\),
\(E_{k_1, k_2, i, j}\) does occur (an incorrect candidate). Observe that
\[
\left( \bigcup_{k_1, k_2} E_{k_1, k_2, 1, 1} \right)^c = \bigcap_{k_1, k_2} E_{k_1, k_2, 1, 1}^c \subseteq E_{d_1, d_2, 1, 1}^c .
\]
Thus
\[
p_n^1 \leq p \left( E_{d_1, d_2, 1, 1}^c \right) + \\
p \left( \bigcup_{1 \leq k_1, k_2 \leq d} E_{k_1, k_2, i, j} \right) .
\]
The first term can be made \(\leq \varepsilon\) for \(n\)
\((i, j) \neq (1,1)\)
sufficiently large by the asymptotic equipartition property. Expanding and bounding the second term we have

\[
p_n^{1} \leq \varepsilon + \sum_{1 \leq k_1, k_2 \leq d} \sum_{i=1, j \neq 1} p_{E_{k_1, k_2}, i, j} + \sum_{k_1, k_2} p(\cdot) + \sum_{k_1, k_2} p(\cdot)
\]

(15)

Treating the last term first, we note that

1) there are \(d^2 \binom{2}{2} (\alpha, d R_1 - 1) \binom{2}{2} (\alpha, d R_2 - 1) \leq d^2 (\alpha, d) (R_1 + R_2)

(16)

terms; and

ii) each term is upper bounded by

\[
p_{E_{k_1, k_2}, i, j} \leq |A_{\varepsilon}|^2 \left( \frac{-(\alpha, d)(H_1(X_1) - \varepsilon)}{2} - \frac{(\alpha, d)(H_1(X_2) - \varepsilon)}{2} - (\alpha, d)(H_1(Y) - \varepsilon) \right.
\]

\[
- (\alpha, d)(I_1(X_1, X_2; Y) - 4 \varepsilon)
\]

\[\leq 2^{-(\alpha, d)(H_1(X_1, X_2, Y) - \varepsilon)}\]

(17)

where we have used

\[
|A_{\varepsilon}|^2 \leq 2^{-(\alpha, d)(H_1(X_1, X_2, Y) - \varepsilon)}
\]

(18)

Thus the last sum in Eq. (15) tends to zero if

\[
\frac{1}{n} \left( \log d^2 + (\alpha, d)(R_1 + R_2) - (\alpha, d)(I_1(X_1, X_2; Y) - 4 \varepsilon) \right) < 0
\]

(19)

or equivalently,

\[
R_1 + R_2 < I_1(X_1, X_2; Y) - \frac{2 \log d}{\alpha, d} - 4 \varepsilon
\]

(20)
Similarly treating the first two terms, we see that these terms → 0
if
\[ R_1 < I_1(X_1; Y|X_2) - \frac{2 \log d}{n} - 4 \varepsilon \]  
\[ R_2 < I_1(X_2; Y|X_1) - \frac{2 \log d}{n} - 4 \varepsilon . \]  
(21)

A similar calculation is made for the second window, at rates \( R^1_n = (R^1, R^2) \), probability of error \( p^2_n \), and typical set \( A^2_\varepsilon \) defined under \( p_2(\cdot, \cdot, \cdot) \).

Finally, \( p_n \leq p^1_n + p^2_n \), and, for every \( d \) and \( \varepsilon > 0 \), \( n \) can be chosen so that \( p_n \leq \varepsilon \). The rate pair for such a code is

\[ R^0 = \left( \alpha n R + \overline{\alpha} n R' \right) / (n + d) \]
\[ = \left( n / (n + d) \right) \left( \alpha R + \overline{\alpha} R' \right) \]
\[ \rightarrow \alpha R + \overline{\alpha} R' , \quad \text{as} \quad n \rightarrow \infty , \]  
(22)
since \( d \) is fixed.

Thus we have a proof that any rate point \( R^0 \) in \( C^* \) can be achieved with probability of error \( p_n \rightarrow 0 \).
Section 3: Elimination of Knowledge of Delay \( d \).

For known maximal synchronization delay \( d \) and desired probability of error \( \epsilon \), there exist block codes \( C_1, C_2 \) of block lengths \( n(d, \epsilon) \) achieving any rate \((R_1, R_2)\) in the capacity region and achieving average probability of error \( \epsilon \). However, if the true delay is greater than \( d \), the probability of error may be high.

We overcome this problem by concatenating codes of increasing block lengths \( n_1, n_2, \ldots \). The \( i \)th block code is designed to have rate \((R_1, R_2)\) and probability of error \( \epsilon_i \) for all delays \( \leq d_i \).

For a given \((R_1, R_2)\) in the capacity region, choose \( d_i \to \infty \), and let \( n_i \to \infty \) in such a manner that

\[
\epsilon_i \to 0
\]

\[d_i/n_i \to 0 \quad (24)\]

and

\[
n_i/(\sum_{j=1}^{i} n_j) \to 1 \quad (25)\]

Moreover, in the \( i \)th block, with block length \( n_i \), retransmit all of the bits from the previously sent blocks.

For any \( d \), there exists an \( i_0 \) such that \( d_i \geq d \) for \( i \geq i_0 \). Now \((n_i)(R_1, R_2)\) bits are received in \((n_1 + n_2 + \ldots + n_i + d_i)\) transmissions for an overall rate vector of

\[
(R_1^*, R_2^*) = (R_1, R_2) \frac{n_i}{\sum_{j=1}^{i} n_j + d_i} \quad (26)
\]

as \( i \to \infty \), by conditions (24) and (25). Thus no bits are lost and the achievable rates are not affected.
Finally, if we add the condition
\[ \sum_{i=1}^{\infty} e_i < \infty \] (27)
to (1), (2), and (3), it follows from the Borel Cantelli lemma that with probability one only a finite number of block decoding errors will be made. At that time all previous errors will have been corrected and no future errors will be made.

The choice of block lengths \( n_i \) can be made in two interesting ways:

a) \( n_i / (\sum_{j=1}^{i} n_j + d_i) + 1 \), with resulting overprints on the bits already received. The result is that any given bit will eventually be correct with probability one after a finite number of changes (overprints). The problem is that the decoding delays increase very fast, resulting in \( \overline{\text{Tim}} R_n = C \) but \( \lim R_n = 0 \).

b) \( n_i \to \infty, \ n_{i+1}/(n_i + d_i) + 1 \). Now \( \overline{\text{Tim}} R_n = \lim R_n = C \). However, bits are no longer eventually correct with probability one. On the other hand, the expected proportion of bit errors in the first \( n \) transmissions tends to zero as \( n \to \infty \).

This increasing block length construction may not be completely satisfactory, however. For no time \( T \) do we know that any bits will be correctly decoded at time \( T \). The decoding delay has been allowed to grow to infinity. Most people would not accept such a communication system. There is thus a minor gap between the results of [3] and [4] and these results which can only be closed by further research.

We do, however, have a very precise result when the delay can be bounded in advance. This is a reasonable assumption for actual channels.
Thus this paper has shown that the benefits of time sharing to achieve convexity of the capacity region can be obtained without synchronization.
References


