RECENT CODING THEOREMS AND CONVERSES FOR MULTI-WAY CHANNELS.


Edward C. van der Meulen

Departement Wiskunde
Katholieke Universiteit Leuven
B-3030 Leuven, Belgium

ABSTRACT. In this survey paper recent results on the multiple-access channel, which have been obtained by various researchers in information theory around the world during the period August 1976 - July 1985, and which are available in the literature, are described. An attempt has been made to include all papers in this area which have come to the author's attention prior to the writing of this article, but no claim of exhaustiveness is made. The word "multiple-access channel" is interpreted here solely from the point of view of multi-user information theory, i.e., in the Shannon-theoretic sense. The recent results are presented through a series of coding theorems and converses, with the emphasis on the statement and interpretation of these results, rather than on proofs. In a previous survey [133] the research advances of several multi-way channels, including the multiple-access channel, were described for the period 1961 - July 1976. The present survey describes the research progress made on the multiple-access channel since then, and thus picks up where that survey left off. Another survey [134], describing the results on the broadcast channel during the period 1976-1980, appeared in the Proceedings of an NATO Advanced Study Institute, and formed Part I of a series on recent coding theorems and converses for multi-way channels. The present paper constitutes Part II of that series.

* The research for this paper was supported in part by Onderzoeksfonds, Katholieke Universiteit Leuven, Project OT/VI/17-OT/IX/20, and by Project GOA 83/88-53, Ministerie van Wetenschapsbeleid, Brussels, Belgium.
1. **INTRODUCTION**

Within information theory, the communication situation in which two or more users send simultaneously information over a common channel to one receiver has been considered first by Shannon [123] in connection with his study of two-way communication channels. Shannon [123] wrote in 1961, "In another paper we will discuss the case of a channel with two or more terminals having inputs only and one terminal with an output only, a case for which a complete and simple solution of the capacity region has been found."

A discrete memoryless (d.m.) multiple-access channel (MAC) with two input users, denoted by \((\mathcal{X}_1 \times \mathcal{X}_2, P(y|x_1,x_2), y)\), consists of two finite input alphabets \(\mathcal{X}_1\) and \(\mathcal{X}_2\), one finite output alphabet \(y\), and a transition probability matrix \(P(y|x_1,x_2)\). Here, \(P(y|x_1,x_2)\) is the probability that the output symbol \(y\) is received, given that the input symbols \(x_1\) and \(x_2\) are transmitted. The problem then is for the two users to communicate simultaneously through the channel as effectively as possible, under various assumptions. The information delivered by the sources to the transmitters may be independent (the classical situation) or correlated. There may be feedback from the receiver to both transmitters, or to only one of them, or no feedback at all. Instead of being discrete, the channel may be continuous, in particular Gaussian. These and other situations have been investigated during the past fifteen years by many workers in information theory.

Consideration of these various situations has led to many new intriguing research problems, and has generated a great research activity around the world in multi-user information theory in general, and with respect to the multiple-access channel in particular. The classical configuration of a d.m. MAC with two senders and two independent sources is depicted in Fig. 1.

One of the main goals in multi-user channel problems is often to characterize
the capacity region of the channel under consideration. Loosely speaking, the capacity region is the set of pairs (or vectors) of signalling rates which can be simultaneously achieved for the various directions with arbitrarily small error probabilities.

Shannon did not publish another paper on multi-way channels after [123], and so it is not known what solution he had in mind when he wrote the above quotation. However, the discussion of the MAC in [123] drew the attention of other researchers in information theory.

Fig. 1. The multiple-access channel in the classical situation.

In 1971, at the 2nd International Symposium on Information Theory in Tsahkadsor, Ahlswede [2] presented a simple and complete characterization of the capacity region of the d.m. MAC in the classical configuration. Since then other characterizations of the same region were obtained, and other communication situations for the MAC investigated. This research resulted in a multitude of papers on the MAC being published from 1973 onwards. While in the early papers on the subject [2,3,13] the communication system under consideration was called "a channel with two (or more) senders and one receiver", the term "multiple access channel" seems to have been used first in this context by Liao [91]. This terminology is now widely adopted.

As mentioned in the abstract, the results on the MAC obtained by various
researchers until August 1976 were already surveyed in [133]. The purpose of this paper is to survey and characterize the research advances on the MAC in the period August 1976 - July 1985. However, in order to be able to describe those results properly, it is necessary to recall most of the earlier results on which the later ones are based or from which these stem. Therefore, this survey paper will give most of the results published since 1971, with the emphasis on those obtained during the period August 1976 - July 1985.

The most important results regarding the MAC during 1971-1976 can be summarized as follows. After Ahlswede's characterization [2] of the capacity region of the d.m. MAC in the classical situation, Liao [91] and Ahlswede [3] gave other, equivalent characterizations. In 1973, Slepian and Wolf [124, 125] considered: (i) the transmission of information from two specially correlated sources over a d.m. MAC, and (ii) the problem of noiseless coding of two arbitrarily correlated information sources. The latter situation can also be viewed as the transmission of information from two arbitrarily correlated sources over a lossless MAC. Ahlswede [2] and Ulrey [130] found simple characterizations of the capacity region of a d.m. MAC with 3 and $r \geq 3$ senders, respectively, in the classical situation. Next, Cover [26] and Wyner [159] proved an achievable rate region for the additive white Gaussian noise (AWGN) MAC. This region was also outlined in [125]. Furthermore, Gaarder and Wolf [61] showed that feedback can increase the capacity region of a d.m. MAC. Subsequently, in 1976, Cover and Leung [30] established an achievable rate region for the d.m. MAC with feedback.

In the period August 1976 - July 1985 research concerning the MAC has developed in the following areas. The d.m. MAC with correlated sources was investigated by Cover, El Gamal, and Salehi [29], Dueck [48], Ahlswede and Han [9], and De Bruyn, Prelov, and van der Meulen [37]. Cover and Leung [30],
Carleial [16], and Ozarow [101] investigated the AWGN MAC with feedback. Ozarow [101,102] established its capacity region. Keilers [85] considered the spectral Gaussian MAC. The d.m. MAC with feedback was further studied by Willems [144,145], and Dueck [46]. Dueck [46] and Willems and van der Meulen [148] investigated the d.m. MAC with partial feedback, Carleial [16] the Gaussian MAC with partial feedback, and King [89], Carleial [17], and Willems [145] the MAC with generalized feedback. Extensions of the d.m. MAC in the Slepian-Wolf configuration to the case of s sources and t input terminals were studied by Han [69]. Dueck [49] and Ahlswede [6] both proved a strong converse for the d.m. MAC in the classical situation. Cover, McEliece, and Posner [31], Grigor'ev [66], Poltyrev [108], Hui [77] and Hui and Humblet [79] considered the asynchronous MAC. The zero-error capacity region of deterministic MAC's in a certain asymmetric communication situation was found by Bassalygo, Pinsker, and Prelov [13]. This so-called asymmetric MAC was further investigated in a series of papers by Prelov [109,110,111], and De Bruyn and van der Meulen [38,39]. Another communication situation, that of cribbing encoders, was investigated by Willems [145], and Willems and van der Meulen [149]. Jahn [80] considered the arbitrarily varying MAC. Following the original investigations by Slepian and Wolf [125], Haroutunian [72], and Dyachkov [51] studied further the subject of error bounds for the d.m. MAC.

The above list provides in a nutshell the bulk of the material presented in this paper. Additional results are mentioned in the various sections. As said earlier, the results themselves will be presented in a systematic way through a series of coding theorems and converses. However, not for all communication situations considered a complete solution (i.e., a characterization of the capacity region) has been obtained. On the contrary, for many problems the research is still in progress. Also, not all of the research has been devoted to merely proving coding theorems and converses, as there are other
related problems of interest, such as determining cardinality constraints for auxiliary random variables, deriving exponential error bounds, and describing constructive coding schemes. Altogether, looking at the variety of the results obtained so far, one may conclude that the research regarding the MAC has progressed considerably in the period August 1976 - July 1985, but that some of the major problems (such as determining necessary and sufficient conditions for transmitting arbitrarily correlated sources over a d.m. MAC and finding the capacity region for a d.m. MAC with feedback) are still unsolved.

As already mentioned, the emphasis in this paper will be on results rather than proofs. In order to prove achievability of the various regions under consideration, several proof techniques have been used over the years; most of them are based on Shannon's random coding method [121]. However, since for each situation new aspects come in, new techniques and tricks had often to be invented. In his first two papers on the MAC, Ahlswede [2,3] uses Shannon's random coding method. Slepian and Wolf [125] give an achievability proof which follows Gallager's random coding proof [65] for the one-way channel. In 1975, two papers by Cover [27,28] appeared in which he put forward a formal development of the idea of joint typicality. It appears that most achievability proofs for the MAC given since 1976 are based on this method. An excellent exposition of this proof technique and its usefulness for proving achievability in multi-user channel problems is given by El Gamal and Cover [56]. In [32] one will find also several sections dealing with the MAC. Whereas most authors used a random coding argument to show achievability, Ozarow [101] used a constructive coding scheme to prove achievability of the region found by him. In [150] a semi-constructive coding scheme is presented to prove achievability in a feedback situation. Recently, in several other papers [137] and [138], achievable rate regions for MAC's with feedback were derived using constructive methods. For proving weak converses, Fano's lemma [57] is still the basic tool.
For proving strong converses, new techniques were developed in \cite{49,6}. For cardinality bounds, a lemma by Ahlswede and Körner \cite{10}, together with its strengthening observed by Wyner and Ziv \cite{161}, forms a basic tool. A paper by Salehi\cite{112} is a useful reference in this regard. Willems \cite{145} obtained new results regarding cardinality bounds and proved convexity of certain regions. In \cite{145} he developed a method to investigate the boundary of certain rate regions for the MAC.

Csiszár and Körner \cite{32} developed a combinatorial approach towards information theory, in which they used types in addition to typical sequences. One chapter of their book is devoted to multiple-access channels. It contains a proof of Theorem 3.1 below and lists some additional results not all of which are mentioned in this paper. Ahlswede \cite{5} discusses the various versions of the capacity region obtained in \cite{2,3} and their similarity with source coding problems. Earlier survey articles dealing with the MAC have been written by Wyner \cite{159}, Cover \cite{26}, and van der Meulen \cite{133}. Schalkwijk\cite{115}, in an extensive article on the quantitative aspects of information theory devotes also a section on the MAC and gives a transparent proof of Theorem 3.1 below. Recently, Gallager \cite{65} reviewed and compared the information theoretic approach and the collision resolution approach to multiaccess channels. He thereby reviewed several aspects of the classical MAC and gave a perspective on the strengths and weaknesses of both approaches.

The division of this paper is as follows. In Section 2 the basic notation is set forth. Section 3 deals with the d.m. MAC in the classical situation, Section 4 with the d.m. MAC in the Slepian-Wolf setting, and Section 5 treats the d.m. MAC with arbitrarily correlated sources. In Section 6 results on the Gaussian MAC are discussed. Sections 7, 8 and 9 deal with the MAC with complete feedback, partial feedback, and generalized feedback, respectively. In Section
10 results pertaining to the d.m. MAC with cribbing encoders are presented. In Section II some recent articles dealing with coding techniques for the d.m. MAC are discussed. The paper concludes with an extensive bibliography.

2. BASIC NOTATION

In this paper, sets are denoted by script letters, such as $\mathcal{X}$. In the discrete case, these sets are assumed to be finite, and the cardinality of a set $\mathcal{X}$ is denoted by $|\mathcal{X}|$. The cartesian product of two sets $\mathcal{X}$ and $\mathcal{Y}$ is indicated by $\mathcal{X} \times \mathcal{Y}$. To a given finite set $\mathcal{X}$ and a probability distribution (PD) $P(x)$ on $\mathcal{X}$, we make always correspond a random variable (RV) $X$ taking values in $\mathcal{X}$ according to this PD, i.e., $Pr\{X=x\} = P(x)$. We assume the reader's familiarity with the standard definitions of entropy $H(X)$, conditional entropy $H(X|Y)$, mutual information $I(X;Y)$, and conditional mutual information $I(X;Y|Z)$ for RV's $X, Y$, and $Z$. (Cf. [64, pp. 20-21].) Given a d.m. MAC $(\mathcal{X}_1 \times \mathcal{X}_2, P(y|x_1,x_2), Y)$, a PD on $\mathcal{X}_1 \times \mathcal{X}_2$ gives rise to the mutual information functions $I(X_1;Y|X_2), I(X_2;Y|X_1)$, and $I(X_1,X_2;Y)$. In the characterizations of the various achievable rate regions often an auxiliary RV $U$ occurs ranging over a finite set $\mathcal{U}$.

The set of sequences of length $n$ of elements of a set $\mathcal{X}$ is denoted by $\mathcal{X}^n$. The members of $\mathcal{X}^n$ are written as $x^n = (x_1, \ldots, x_n)$. A random vector is denoted by underlined, superscripted capitals: $\underline{x}^n = (X_1, \ldots, X_n)$ denotes a random vector of length $n$. A vector of random vectors is denoted by a superscripted capital underscored with a wavy line: $\underline{x}^{kn} = (X_1^n, \ldots, X_k^n)$ represents a vector of $k$ random vectors, each of dimension $n$. This notation for random vectors holds strictly for the written text. (In the figures, a simpler notation is used. There, a single random vector is not underlined nor
superscripted, whereas a vector of random vectors is denoted in the figures
by an underlined capital without a superscript, such as $\underline{X}$ (cf. Fig. 14).) When
the underlying distribution is unspecified, probabilities are denoted by $\Pr$, 
such as $\Pr[U \in A]$. When it is given that one deals with the nth power of a
certain input PD or a certain channel transition probability function, these
probabilities are denoted by $P^n$, as exemplified by the following. The trans-
mission probabilities of a d.m. MAC $\{\mathcal{X}_1 \times \mathcal{X}_2, P(y|x_1,x_2), Y\}$ are defined by

$$P^n(y^n|x^n_1,x^n_2) = \prod_{i=1}^{n} P(y_i|x^n_{1i},x^n_{2i})$$

(2.1)

for all $x^n_1 = (x^n_{11}, \ldots, x^n_{1n}) \in \mathcal{X}^n_1$, $x^n_2 = (x^n_{21}, \ldots, x^n_{2n}) \in \mathcal{X}^n_2$, $y^n = (y_1, \ldots, y_n)$

$\in \mathcal{Y}^n$, and all positive integers $n$. Formula (2.1) expresses the memoryless
feature of the MAC.

As is customary in information theory, logarithms are in the discrete case
taken to the base 2, while in the Gaussian case the natural logarithm is used.
Also, $h(p)$ stands for the usual binary entropy function, i.e., $h(p) = -p \log p - (1-p) \log (1-p)$, for $0 \leq p \leq 1$. The notation $Z \sim \mathcal{N}(0, N)$ means that the RV $Z$
has a normal distribution with mean 0 and variance $N$. Finally, the convex hull
of a set $\mathcal{A}$ is denoted by $\text{co}(\mathcal{A})$.

In this paper there is also, for the convenience of the reader, a classi-
fication of communication situations proposed such as $(K_{21}, I)$, $(K_{21}, II)$, etc.
This classification is by no means standard. Little work has been done sofar


to establish a general classification notation for multi-way communication

situations, apart from the initial classification proposed by Ahlswede [2] and
the classification notation introduced in [133,134]. The classification
notation used here builds forth on the one developed in [133,134].
3. THE MULTIPLE-ACCESS CHANNEL IN THE CLASSICAL SITUATION

A d.m. MAC with two input users, briefly denoted by $K_{21}$, consists of a channel $\mathcal{X}_1 \times \mathcal{X}_2, P(y|x_1, x_2), y$, as defined in Section 1, and with its transition probabilities given by (2.1). This MAC is said to be in the classical situation if it is used to transmit two separate messages from two independent sources. This communication situation is denoted by $(K_{21}, I)$ and shown in Fig. 1. Here, two message sources emit statistically independent messages $m_1 \in \mathcal{M}_1 = \{1, 2, \ldots, M_1\}$ and $m_2 \in \mathcal{M}_2 = \{1, 2, \ldots, M_2\}$, such that each pair $(m_1, m_2)$ occurs with probability $1/(M_1 M_2)$. Message $m_1$ is encoded at terminal 1 into the codeword $x_1^n(m_1) \in \mathcal{X}_1^n$, and message $m_2$ is encoded at terminal 2 into the codeword $x_2^n(m_2) \in \mathcal{X}_2^n$. This is the communication situation first mentioned by Shannon [123].

An $(n, M_1, M_2, \lambda)$-code for the d.m. MAC in situation $(K_{21}, I)$ consists of $M_1$ codewords $x_1^n(m_1) \in \mathcal{X}_1^n$, $M_2$ codewords $x_2^n(m_2) \in \mathcal{X}_2^n$, and $M_1 M_2$ pairwise disjoint decoding sets $B(m_1, m_2)$, such that

$$\frac{1}{M_1 M_2} \sum_{(m_1, m_2) \in \mathcal{M}_1 \times \mathcal{M}_2} P_B^n(B(m_1, m_2) | x_1^n(m_1), x_2^n(m_2)) \geq 1 - \lambda. \quad (3.1)$$

A pair $(R_1, R_2)$ of non-negative real numbers is said to be an achievable rate pair for a d.m. MAC in situation $(K_{21}, I)$, if, for any $\delta > 0$ and any $\lambda, 0 < \lambda < 1$, there exists, for all sufficiently large $n$, an $(n, M_1, M_2, \lambda)$-code for that MAC such that $\log M_i \geq n(R_i - \delta), i = 1, 2$. The capacity region $C(K_{21}, I)$ of a d.m. MAC $K_{21}$ is defined as the set of all achievable rate pairs in situation $(K_{21}, I)$. 
In the period 1961-1976 various results regarding the capacity region $\mathcal{C}(K_{21},I)$ were obtained, which have already been surveyed in [133]. For sake of completeness we will briefly recall some of these results insofar as it is needed.

Although Shannon [123] indicated that he had found the capacity region $\mathcal{C}(K_{21},I)$, he never published this result. A simple characterization of $\mathcal{C}(K_{21},I)$ was first presented by Ahlswede [2] at the 2nd International Symposium on Information Theory, held at Tsahkadsor in September 1971. At the same symposium, van der Meulen [131] put forward a limiting expression for this capacity region, provided computable inner and outer bounds for it, and gave some simple examples of d.m. MAC's, among which the deterministic binary erasure MAC, the two-user binary multiplying channel, and the modulo 2 sum MAC $K_{21}$. For the latter two MAC's he determined the capacity region $\mathcal{C}(K_{21},I)$ by direct methods.

Later, at the IEEE International Symposium on Information Theory held at Asilomar early 1972, Liao [91] announced another simple characterization of $\mathcal{C}(K_{21},I)$, which is slightly different from the one originally found by Ahlswede [2] and has the virtue of being symmetric in its description. Liao [91] also wrote a Ph.D. dissertation on the MAC, and his solution for the characterization of $\mathcal{C}(K_{21},I)$ is quoted in [125]. According to [105], Liao [91] provided a block coding error exponent for the d.m. MAC in situation $(K_{21},I)$ as well. However, since a proof of Liao's result has not yet appeared in print, it is unknown, at least to this author, how he arrived at his characterization. By the end of 1972, Ahlswede [3] had, independently, established this second characterization also, proving both an achievability part and a weak converse. This latter characterization is nowadays preferred, because of its symmetric description. We state it in the following theorem.
Theorem 3.1. ([2], [91], [3]): The capacity region of the d.m. MAC in situation \((K_{21}, I)\) is given by

\[
\mathcal{C}(K_{21}, I) = \co\{(R_1, R_2) : 0 \leq R_1 \leq I(X_1; Y|X_2), \quad 0 \leq R_2 \leq I(X_2; Y|X_1), \quad R_1 + R_2 \leq I(X_1, X_2; Y), \}
\]

for some \(P(x_1, x_2, y) = P(x_1)P(x_2)P(y|x_1, x_2)\).

Observe that in this characterization of \(\mathcal{C}(K_{21}, I)\) the union over PD's \(P(x_1, x_2)\) is with respect to independent PD's only. The characterization given in Theorem 3.1 is in terms of mutual information quantities which involve only a single channel operation and a single-letter PD on inputs, rather than blocks of \(n\) channel operations and a PD on sequences of input letters of length \(n\). Such a characterization is called a single-letter characterization and has the advantage of being in principle computable in the sense that it can be calculated to any desired accuracy in finite time, as opposed to limiting expressions which are generally not computable. (See [32, pp. 259-260] for an instructive discussion on the computability of characterizations.)

Besides situation \((K_{21}, I)\), corresponding to the case of the transmission of two independent messages via two separate encoders over a d.m. MAC \(K_{21}\), we also sometimes consider the situation where both source outputs are made available to both encoders before the codewords are made up. Alternatively, this situation can be viewed as one in which there is a single encoder which is able to observe both messages. This situation is referred to as the d.m. MAC with total cooperation and will here be denoted by \((K_{21}, T)\). The corresponding capacity region will be denoted by \(\mathcal{C}(K_{21}, T)\) and is given by
\[ C(K_{21}, T) = \{(R_1, R_2) : R_1 \geq 0, R_2 \geq 0, R_1 + R_2 \leq C_T\}, \quad (3.3) \]

where \( C_T \) denotes the so-called cooperation capacity defined by

\[ C_T = \max_{P(x_1, x_2)} I(X_1, X_2; Y). \quad (3.4) \]

The region \( C(K_{21}, T) \) is bounded by the total cooperation line \( L_T \), defined by

\[ L_T = \{(R_1, R_2) : R_1 \geq 0, R_2 \geq 0, R_1 + R_2 = C_T\}. \quad (3.5) \]

In this paper, several examples of MAC's will be discussed. It will therefore be convenient to introduce at this point certain classes of d.m. MAC's which are easy to analyze, in analogy to the classes of d.m. one-way channels considered by Ash [11, pp. 50-52]. A d.m. MAC is called
deterministic if \( P(y|x_1, x_2) = 1 \) or 0 for all inputs \( x_1, x_2 \) and all outputs \( y \), or, equivalently, if \( H(Y|X_1, X_2) = 0 \) for all input PD's. Similarly, a d.m. MAC is called lossless if \( H(X_1, X_2|Y) = 0 \) for all input PD's, or, equivalently, if the input pair \( (x_1, x_2) \) is determined by the output \( y \). Finally, a d.m. MAC is called noiseless if it is both deterministic and lossless. Let us denote the capacity region of a deterministic d.m. MAC in situation \( (K_{21}, I) \) by \( C(K_{21}, I, D) \). Then, according to Theorem 3.1, this region is given by

\[ C(K_{21}, I, D) = \text{co}\{(R_1, R_2) : 0 \leq R_1 \leq H(Y|X_2), \quad (3.6a) \]

\[ 0 \leq R_2 \leq H(Y|X_1), \quad (3.6b) \]

\[ R_1 + R_2 \leq H(Y), \quad (3.6c) \]

for some \( P(x_1, x_2, y) = P(x_1)P(x_2)P(y|x_1, x_2) \).
Similarly, if we denote the capacity region of a lossless d.m. MAC in situation \((K_{21}, I)\) by \(C(K_{21}, I, L)\), then, by Theorem 3.1, this region takes the form

\[
C(K_{21}, I, L) = \text{co}\{(R_1, R_2) : 0 \leq R_1 \leq \log |X_1|, 0 \leq R_2 \leq \log |X_2|\}.
\]

(Willems [145] introduced the class of d.m. MAC's \(K_{21}\), which are characterized by the property that \(H(X_1 | Y, X_2) = 0\) for all input PD's or, equivalently, by the property that \(x_1 = f(y, x_2)\) for all triples \((x_1, x_2, y)\). We denote this class of MAC's by \(W_1\). Similarly, one may define the class \(W_2\) to consist of all d.m. MAC's which are characterized by the fact that \(H(X_2 | Y, X_1) = 0\) for all input PD's. We say that a d.m. MAC \(K_{21}\) is conditionally lossless if it belongs to \(W_1\) or \(W_2\). Finally, we define the class \(W\) to consist of all d.m. MAC's \(K_{21}\) which belong to both \(W_1\) and \(W_2\), i.e., \(W = W_1 \cap W_2\). A specific and rather well-known example of a d.m. MAC \(K_{21}\), which occurs many times in this paper, is that of the deterministic binary erasure MAC, abbreviated deterministic BEMAC. In the coding literature, this channel is usually referred to as the two-user noiseless binary adder channel (BAC) (cf. Section 11). This channel, introduced by the author in [131], has

![Fig. 2. The capacity region \((K_{21}, I)\) of the deterministic BEMAC.](image-url)
input alphabets $\mathcal{X}_1 = \mathcal{X}_2 = \{0, 1\}$, output alphabet $\mathcal{Y} = \{0, 1, 2\}$, and is defined by the operation $y = x_1 + x_2$. This MAC is clearly deterministic and belongs to class $\mathcal{W}$ as well. If we denote the capacity region of a deterministic d.m. MAC belonging to class $\mathcal{W}$ in situation $(K_{21}, I)$ by $C(K_{21}, I, \mathcal{D}, \mathcal{W})$, then, according to Theorem 3.1, this region can be characterized as follows:

$$C(K_{21}, I, \mathcal{D}, \mathcal{W}) = \text{co}\{(R_1, R_2) : 0 \leq R_1 \leq H(X_1), 0 \leq R_2 \leq H(X_2), R_1 + R_2 \leq H(Y)\}, \quad (3.8a)$$

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

Characterization (3.8) yields immediately the capacity region of the deterministic BEMAC. For this example, $C(K_{21}, I) = \{(R_1, R_2) : 0 \leq R_1 \leq 1, 0 \leq R_2 \leq 1, R_1 + R_2 \leq 1\}$, as was first computed in [61]. This capacity region is shown in Fig. 2. The total cooperation capacity for this example equals $C_T = \log_2 3 = 1.5850$, as shown in [131].

Having briefly reviewed the main results regarding the d.m. MAC in the classical situation generally known in mid-1976, we are now in a position to describe the results on the d.m. MAC in situation $(K_{21}, I)$ published or obtained in the period August 1976 – July 1985. Those new results are the following.

a. Through an example, Dueck [44] has shown that for the d.m. MAC in situation $(K_{21}, I)$ the capacity region for maximal error is generally smaller than the corresponding average error capacity region given by Theorem 3.1. The determination of the maximal error capacity region in situation $(K_{21}, I)$ remains an open problem.
b. Bierbaum and Wallmeier [14] have shown, also by means of an example, that, in the description of $C(K_2, I)$ as given by Theorem 3.1., it is really necessary to take the convex hull. Thus, some of the points in the region described by (3.2) can only be achieved by time-sharing.

Since the original proofs given in [2] and [3] of the achievability and converse part of Theorem 3.1, several other proofs of both parts have appeared in the literature. In this regard we mention in particular [125],[56],[32], and [115]. Whereas the achievability proofs in [2] and [3] are based on Shannon's random coding method, and the achievability proof in [125] follows Gallager's random coding proof, the achievability proof given by El Gamal and Cover [56] is based on the notion of $\varepsilon$-typical sequences. Csiszár and Körner [32] proved achievability by relating the multiple access network to a multiple source network, an idea which they partly attribute to R. Ahlswede. The achievability proof given in [115] is again in terms of jointly $\varepsilon$-typical sequences. All converse proofs in [2],[3],[125],[56], and [32] are weak converses, and involve Fano's lemma [57]. The proof of the weak converse given in [56] is particularly perspicuous.

c. Dueck [49] was the first to prove the strong converse (in the sense of Wolfowitz [156]) of the coding theorem for the d.m. MAC in situation $(K_2, I)$. Whereas in information theory a weak converse traditionally means that for small error probability $\lambda$ and $n$ sufficiently large no code exists with signaling rate above capacity, a strong converse always means that for any $\lambda$, $0 < \lambda < 1$, and $n$ sufficiently large no code exists with rate above capacity [157]. There are generally several ways to formulate a strong converse. In order to describe the strong converse statement for the d.m. MAC in situation $(K_2, I)$ properly, we introduce another definition, following the approach in [49]. In particular, we will need the concept of $\lambda$-achievability, in addition
to ordinary achievability (as defined after formula (3.1)). For a given \( \lambda, \; 0 < \lambda < 1 \), a pair of non-negative real numbers \((R_1, R_2)\) is said to be \(\lambda\)-achievable for a d.m. MAC in situation \((K_{21}, I)\) if, for any \( \delta > 0 \) and sufficiently large \( n \) there exists an \((n, M_1, M_2, \lambda)\)-code for that MAC such that
\[
\log M_i \geq n(R_i - \delta), \; i = 1, 2.
\]
Clearly, a rate pair \((R_1, R_2)\) is achievable in the ordinary sense if and only if it is \(\lambda\)-achievable for all \( \lambda, \; 0 < \lambda < 1 \). Thus, if we denote for a given \( \lambda \) the set of all \(\lambda\)-achievable rate pairs by \( \mathcal{R}(K_{21}, I, \lambda) \), then we have that for any \( \lambda, \; 0 < \lambda < 1 \), and any MAC \( K_{21} \)
\[
\mathcal{C}(K_{21}, I) = \bigcap_{0 < \lambda < 1} \mathcal{R}(K_{21}, I, \lambda) \subseteq \mathcal{R}(K_{21}, I, \lambda), \quad (3.9)
\]
where the equality on the left-hand side follows from Theorem 3.1. While the weak converse says that
\[
\bigcap_{0 < \lambda < 1} \mathcal{R}(K_{21}, I, \lambda) \subseteq \mathcal{C}(K_{21}, I), \quad (3.10)
\]
the strong converse states that \( \mathcal{R}(K_{21}, I, \lambda) \subseteq \mathcal{C}(K_{21}, I) \) for all \( \lambda, \; 0 < \lambda < 1 \). When we formulate this result as a theorem, thereby taking into account the achievability part of Theorem 3.1, Dueck's strong converse reads as follows.

**Theorem 3.2. ([49]):** For the d.m. MAC in situation \((K_{21}, I)\),
\( \mathcal{R}(K_{21}, I, \lambda) \subseteq \mathcal{C}(K_{21}, I) \) for all \( \lambda, \; 0 < \lambda < 1 \) (and hence by (3.9) \( \mathcal{R}(K_{21}, I, \lambda) = \mathcal{C}(K_{21}, I) \) for all \( \lambda, \; 0 < \lambda < 1 \)).

We remark here that this strong converse theorem involves as performance criterion the average error probability given in (3.1), rather than maximal error probability. For converse theorems average error is a stronger criterion than maximal error. All strong converse theorems established
previously in multi-user information theory used maximal error probability as a criterion. In this connection, Ahlswede, Gács, and Körner [8] developed a general method to prove strong converse theorems in multi-user situations based on maximal error. This method has turned out to be useful in a number of cases, such as e.g. for proving the strong converse for the degraded broadcast channel [10] and the broadcast channel with degraded message sets [90]. (For an account, see [32] and [133].) Therefore, Dueck's strong converse result is rather striking. Dueck [49] proved Theorem 3.2 by making use of the method developed in [8] for maximal error, and Fano's lemma [57], combined with new ideas involving the transformation of "non-single-letter information quantities" into single-letter expressions. This latter method has later been termed a "wringing technique" by Ahlswede [6], who provided another proof of Dueck's result without making use of the method of "blowing up decoding sets" developed in [8]. Rather, Ahlswede [6] followed the approach of [1], in his proof of the strong converse. In doing so, he first derived upper bounds on the length of maximal error codes, and then made this approach applicable to average error codes by introducing a suitable wringing technique, which can be regarded as an improvement of Dueck's wringing technique. At the same time, Ahlswede's proof leads to a somewhat sharper result on coding than Dueck's: it yields a strong converse with $\sqrt{n} \log n$ deviation.

Ohkubo [99] derived a so-called strong outer bound to the capacity region $C(K_{21}, I)$. This outer bound has the property that for rate points contained in it the average probability of decoding error approaches unity as the block length $n$ tends to infinity. Ohkubo [99] assumed maximum likelihood decoding. Since Ohkubo's outer bound does not match the capacity region $C(K_{21}, I)$, his result is obviously superseded by the exact strong converse theorems of Dueck [49] and Ahlswede [6].
d. The so-called asynchronous d.m. MAC in situation \((K_{21}, I)\) has been investigated by a number of authors. Early papers on asynchronous multiplexing and asynchronous multiple access communication include, [25], [126], [127], [129], and [143]. Recent contributions towards establishing the information-theoretic capacity region of the d.m. asynchronous MAC are given in [31], [66], [67], [108], [77] and [79].

An illuminating discussion regarding the problems arising when there is no time synchronization is given in [31]. Cover et al. [31] distinguish between two kinds of cooperation when two users attempt to communicate over the same d.m. MAC \(K_{21}\), the first one being a strategic cooperation (in which both the senders and the receiver agree on the code books that will be used), and the second one occurring when the independent messages are actually sent (thereby deciding e.g. on total cooperation versus independent usage of the channel).

In [31] it is observed that, in the characterization of \(C(K_{21}, I)\) as given by Theorem 3.1, implicit use is made of time-synchronization. In relation to this, we now quote an essential passage of [31] which clearly describes the difficulties which come in the absence of time-synchronization. (The number (3.2) is inserted in the quoted text to make it correspond with the numbering system used here.)

"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 (3.2) when there is a time uncertainty for the users and the receiver? Clearly new code books may have to be constructed. Moreover, the interference of codewords from the users' respective code books 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."
The above quotation adequately describes the problems involved when two users attempt to communicate over a d.m. MAC \((K_{21}, I)\) in the absence of time-synchronization. The difficulties arising in the case of an asynchronous MAC are further compounded by the fact that, as was demonstrated in [14], the region described by (3.2) is in general only convex by virtue of taking the convex hull, thus necessitating time-sharing techniques to obtain all points of the capacity region. In general, such time sharing techniques may not be possible when the delays are arbitrarily large.

In the literature, various notions of asynchronism have been considered.

(i) The d.m. MAC \(K_{21}\) is said to be (fully, totally, or perfectly) synchronous when both senders operate synchronously in the sense that the beginnings of the codewords generated by each sender coincide. This is the classical case \((K_{21}, I)\).

(ii) A d.m. MAC \(K_{21}\) is said to be quasi-synchronous if the two encoders do not maintain block synchronism while the decoder maintains block synchronism with each encoder. Moreover, it is assumed in this case that bit synchronism is maintained amongst the encoders and decoder. This terminology, which stems from the coding literature, is due to Kasami et al. [83]. This is the case considered in all papers discussed in this section, i.e., those concerned with the determination of the information-theoretic capacity region of the so-called asynchronous MAC. In this situation it is assumed that the beginnings of the two codewords may be shifted relative to each other by an arbitrary interval of duration \(\Delta\). The quantity \(\Delta, \Delta = \ldots, -1, 0, 1, \ldots\) is called the synchronization shift. The value of the shift is assumed to be unknown at the MAC inputs, but is known at the output. This condition can be guaranteed by transmitting a special synchronization sequence when each sender commences operation [83], [108], also called a preamble.
sequence [79]. Within the context of a quasi-synchronous MAC it is still possible to distinguish between various situations. Cover et al. [31] considered the case of a quasi-synchronous MAC where the shift is bounded by some constant that is independent of the block length. This case is referred to as "mild asynchronism" in [79]. In general, though, the size of the shift is arbitrary. This situation has been dealt with in [108], [66], [67], [77], and [79]. Hui and Humblet [79] use the terminology "totally asynchronous" in this case.

(iii) Finally, Deatt and Wolf [33] considered the situation where not only the assumption of block synchronism but also that of bit synchronism is dropped altogether, both between the encoders and between the decoder and the encoders. We refer to this situation as the non-synchronized MAC. It will not be further discussed in this section, but rather dealt with in Section 11.

We will now briefly describe the results on the quasi-synchronous d.m. MAC $K_{21}$ obtained in [31], [66], [67], [77], [79], and [108].

Cover et al. [31] defined the maximum relative delay $d$ as the maximum amount by which the two messages are assumed to be out of synchronization relative to a known or prearranged time. They considered two cases. In the first case it is assumed that $d$ is fixed and known to the receiver, whereas in the second case the receiver's knowledge of $d$ is removed and the delay is allowed to grow to infinity at a certain rate. Let us denote the capacity region of the quasi-synchronous d.m. MAC in situation $(K_{21}, I)$ by $C(K_{21}, I, QS)$. Then it is shown in [31] that in the first case ($d$ fixed and known to the receiver) the capacity regions of the quasi-synchronous MAC and of the fully synchronous MAC coincide, i.e., $C_b(K_{21}, I, QS) = C(K_{21}, I)$, where the index $b$ indicates that certain conditions are imposed on $d$, as was done in [31].
The random code construction in [31] is a form of time-sharing that works in the absence of synchronization, provided \( d \) is fixed and known. In the second case (unbounded delays) Cover et al. [31] allow the delay \( d \) and the block length \( n \) to tend to infinity in such a manner that \( d/n \to 0 \). Under this assumption it is shown in [31] that, by using concatenated codes of increasing block lengths, any rate point \( (R_1, R_2) \) in \( C(K_{21}, I) \) is still achievable for the quasi-synchronous d.m. MAC \( K_{21} \). The assumption that the block length is very long compared to the delay may be regarded as somewhat restrictive, though.

In this connection, Grigor'ev [66] considered the problem of quasi-synchronous communication over a specific d.m. MAC when the delay can be an arbitrary function of the block length. The specific MAC for which he investigated quasi-synchronous communication is the deterministic BEMAC described above. Grigor'ev [66] distinguished between the case in which the "users mutually synchronize the boundaries of codewords in such a way that the beginnings and ends of all simultaneously transmitted blocks coincide" (synchronous communication), and the case in which "it is assumed that the position in time of the beginnings and ends of the transmitted words is chosen at random by each user, independently of the others". Moreover, it is assumed in [66] that at the receiving end the positions of the codeword boundaries for both users are known, i.e., that block synchronism for both transmitter/receiver pairs has been established (quasi-synchronous communication).

Grigor'ev [66] defined the length of a suffix \( k \) to be the magnitude of the delay in the symbols between the beginnings of an arbitrary word of user 1 and the first of the subsequent beginnings of a word of user 2. The case \( k = 0 \) corresponds to synchronism, whereas in the absence of synchronism the delay is a random quantity.

Grigor'ev [66] considered any type of delay \( k \) such that \( 0 \leq k \leq \alpha \) or
(1 - \alpha)n \leq k \leq n$, where $n$ is the block length and $0 \leq \alpha \leq 0.5$, $\alpha$ fixed.

The second case considered by Cover et al. [31], i.e., the case where $d/n \rightarrow 0$ as $n \rightarrow \infty$, may be said to correspond to the case $\alpha \rightarrow 0$ in Grigor‘ev’s model.

Grigor‘ev [66] found that for a given $\alpha$, $0 \leq \alpha \leq 0.5$, a rate pair $(R_1,R_2)$ is achievable for the quasi-synchronous deterministic BEMAC in situation $(K_2,I)$ for all delays $k$ (such that $0 \leq k \leq \alpha n$ or $(1 - \alpha)n \leq k \leq n$ for all $n \geq 1$) if $(R_1,R_2)$ belongs to at least one of the following three regions:

(i) $0 \leq R_1 < 0.5$, $0 \leq R_2 < 1$; (ii) $0 \leq R_1 < 1$, $0 \leq R_2 < 0.5$; (iii) $0 \leq R_1 < 1$, $0 \leq R_2 < 1$, $R_1 + R_2 < 1.5 - \alpha$. No converse statement of this result was given in [66]. For a given $\alpha$, $0 \leq \alpha \leq 0.5$, the corresponding achievable rate region is shown in Fig.3, where it is represented by the shaded area bounded by the broken line ABCEFG. The achievable rate region for the extreme case $\alpha = 0.5$ (allowing all delays $k$ such that $0 \leq k \leq n$) is sketched separately in Fig. 4. It is represented there by the shaded region bounded by the broken line ABDHG. Thus, for arbitrary delays $k(0 \leq k \leq n)$ Grigor‘ev’s result [66] guaranteed the achievability of rate pairs in the shaded region of Fig. 4 bounded by the broken line ABDHG, whereas the question of the possible achievability or non-achievability of rate pairs in the triangle BDF was left open in [66]. For $\alpha = 0$ the achievable rate region of Fig. 3 coincides with the classical capacity region shown in Fig. 2.

Following the above results, we mention that Grigor‘ev [67] apparently settled the question of the achievability of the rate pairs in the triangle BDF (Fig.4) for the quasi-synchronous BEMAC in the affirmative, as reported in [108].
Fig. 3. Achievable rate region found in [66] for the quasi-synchronous deterministic BEMAC (situation $(K_{21}, I, QS)$) with delays $k \in [0, an] \cup [(1 - \alpha)n, n]$, $n$ being the block length and $\alpha$ fixed, $0 \leq \alpha \leq 0.5$.

Fig. 4. Achievable rate region found in [66] for the quasi-synchronous deterministic BEMAC (situation $(K_{21}, I, QS)$) with arbitrary delays $k \in [0, n]$. 
Subsequently, Poltyrev [108] determined the capacity region of the d.m. quasi-synchronous MAC (situation $(K_{21}, I, QS)$) when the size of the synchronization shift is arbitrary. Let $\mathcal{R}(K_{21}, I, QS)$ be the region of pairs given by

$$\mathcal{R}(K_{21}, I, QS) = \{ (R_1, R_2) : 0 \leq R_1 \leq I(X_1, Y|X_2), \quad (3.11a)$$

$$0 \leq R_2 \leq I(X_2, Y|X_1), \quad (3.11b)$$

$$R_1 + R_2 \leq I(X_1, X_2; Y), \quad (3.11c)$$

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

Thus, the region $\mathcal{C}(K_{21}, I)$, given by (3.2), equals $co \mathcal{R}(K_{21}, I, QS)$. Poltyrev [108] now showed that for an arbitrary d.m. quasi-synchronous MAC with an arbitrary shift size the capacity region equals $\mathcal{R}(K_{21}, I, QS)$. The capacity region of the d.m. quasi-synchronous MAC $(K_{21}, I)$ with arbitrary delays differs from that of the synchronous d.m. MAC $(K_{21}, I)$ by the fact that in the characterization of $\mathcal{R}(K_{21}, I, QS)$ the convex hull operation is absent. This can be interpreted by saying that time sharing for forming the convex hull is not allowed for two asynchronous encoders. Independently, the same result was established by Hui [77] and Hui and Humblet [79].

It follows that the capacity region of a quasi-synchronous d.m. MAC $(K_{21}, I)$ with arbitrary shift size is strictly less than that of the corresponding synchronous MAC if and only if the region $\mathcal{R}(K_{21}, I, QS)$ is not convex. As it turns out, for the d.m. BEMAC the region $\mathcal{R}(K_{21}, I, QS)$ is convex (cf. [108]), which supports the findings of Grigor'ev [67] that for this MAC the capacity regions for synchronous and quasi-synchronous communication with arbitrary delays coincide. On the other hand, in [14] an example is given of a d.m. MAC $(K_{21}, I)$ for which the region $\mathcal{R}(K_{21}, I, QS)$ is strictly smaller than $\mathcal{C}(K_{21}, I)$. Thus, for that channel, $\mathcal{C}_a(K_{21}, I, QS)$ and $\mathcal{C}(K_{21}, I)$ differ, whereby $\mathcal{C}_a(K_{21}, I, QS)$
denotes the capacity region of the quasi-synchronous d.m. MAC with arbitrary shift size.

Polytrev [108] proved the above result, i.e., $C_a(K_{21},I,\text{QS}) = R(K_{21},I,\text{QS})$, by establishing a weak converse and a direct part based on the use of lattice codes. Hui [77], and Hui and Humblet [79] proved the same result (phrased in terms of total asynchronism), also by giving a weak converse and by establishing a direct part employing random codes. In [108], [77], and [79] it is mentioned that this result can be generalized easily to the case of a quasi-synchronous d.m. MAC with an arbitrary but finite number $t > 2$ of input users.

Recapitulating, we can formulate the main results of this subsection as follows.

Theorem 3.3.a ([31]): When the size of the synchronization shift $\Delta$ is bounded, or small relative to the block length $n$, then the capacity region of the quasi-synchronous d.m. MAC $(K_{21},I,\text{QS})$ is given by

$$C_b(K_{21},I,\text{QS}) = C(K_{21},I). \quad (3.12)$$

Theorem 3.3.b ([108], [77], [79]): When the size of the synchronization shift is arbitrary, $0 \leq \Delta < n$, then the capacity region of the quasi-synchronous d.m. MAC $(K_{21},I,\text{QS})$ is given by

$$C_a(K_{21},I,\text{QS}) = R(K_{21},I,\text{QS}). \quad (3.13)$$

Finally, we mention in this regard the related result by Massey [94], [95], who proved that the asynchronous $t$-user collision channel ($t > 2$) has a total symmetric capacity of $(1-1/t)^{t-1}$, which is smaller than the value of $1$ achievable for the synchronous capacity. (See also [96] and [78].)
Gallager [65] has elaborated upon the connections between the information theoretic approach and the collision resolution approach to multiaccess channels.

e. Ohkubo [100] examined the problem of universal coding and decoding for d.m. MAC's in situation \((K_{21}, I)\), using thereby the method of Fisher [60] based on generalized mutual information functions. In the setup of [100] it is assumed that the actual channel probability matrix of the d.m. MAC \(K_{21}\) governing the transmission of the input symbols \(x_1\) and \(x_2\) is unknown, and furthermore that the (maximum likelihood) decoding algorithm used is based on some hypothetic d.m. MAC \(K'_{21}\) belonging to the set of all d.m. MAC's with the same input and output alphabets as the unknown transmission channel \(K_{21}\). Ohkubo [100] showed that a rate vector \(R = (R_1, R_2)\) is universally achievable, in the sense that there exist codes of rate \(R\) that are asymptotically optimum for all MAC's with fixed input and output alphabets, if \(R\) belongs to a specific subregion of \(C(K_{21}, I)\), the capacity region of the unknown MAC. For rate points \(R\) in the remainder of the capacity region the existence of such universally optimum codes was not demonstrated in [100]. Ohkubo [100] also derived error exponents for the case under consideration.

It should perhaps be mentioned here that various notions of universality are used in the literature. The reader is in this regard referred to e.g. the discussion in [32, p. 172], where the concept of a universally attainable error exponent is defined. Recently, Pokorny and Wallmeier [106] derived such a universally attainable error exponent for the d.m. MAC \((K_{21}, I)\). In particular, using a random coding argument, they derived exponential upper bounds on the average error probability, which are of the informational divergence type [32], and which are universally attainable for every d.m. MAC with given input- and output alphabets. Here, universality means that the choice
of the set of codewords and the decoding rule are independent of the statistics of the MAC. Pokorny and Wallmeier[106] also showed that their random coding exponent is positive for all rate pairs \((R_1, R_2) \in C(K_{21}, I)\).

In this connection it should also be mentioned that in 1974 Ahlswede [3] already established the capacity region of the so-called compound MAC in situation \((K_{21}, I)\). Here, the actual channel probability matrix used for the transmission of a pair of codewords is unknown to the users, but it is known to belong to a certain class of channel probability functions all with the same finite input and output alphabets. In the setup of the compound MAC it is furthermore assumed that the unknown channel probability matrix remains the same during the transmission of a block.

f. Jahn [80] considered the arbitrarily varying d.m. MAC. The mathematical model of an arbitrarily varying one-way channel (AVC) is given by a d.m. channel whose transition probabilities vary arbitrarily from one transmission period to another. The AVC was introduced by Blackwell, Breiman, and Thomasian [15], who proved a coding theorem for AVC's using correlated random codes and average error. In a correlated random code codewords and decoding sets are chosen by a random experiment known to both encoder and decoder. As pointed out in [4] and [80], the concept of a correlated random code is impractical, as it assumes the availability of a device which communicates the outcome of the random experiment to both encoder and decoder. Several authors (cf.[4] for the history and a unifying treatment of the problem) have investigated the AVC using other code concepts, such as deterministic codes with maximal error and randomized encoding with average error. Finally, Ahlswede [4], in the same important paper, determined, among other capacities, the average error capacity of the AVC for deterministic, non-correlated codes under a certain sufficient
condition for the capacity to be positive. Ahlswede [4] called his method to establish this capacity result an "elimination technique".

An arbitrarily varying d.m. MAC (AVMAC) in situation \((K_{21}, I)\) is defined by a collection \(\{P_s(y|x_1, x_2) : s \in S\}\) of d.m. MAC's \(K_{21}\), in which each MAC has the same input alphabets \(\mathcal{X}_1\) and \(\mathcal{X}_2\), and the same output alphabet \(\mathcal{Y}\), and where \(S\) is an arbitrary state set. Operationally, this MAC is allowed to vary from one transmission period to another within a block. Jahn [80] determined the average error capacity region of the d.m. AVMAC \(K_{21}\) under a certain positiveness condition (a sufficient condition for the capacity region to be non-empty), using an adaptation of the elimination technique of [4] to the case of the AVMAC for obtaining a deterministic non-correlated code from a correlated random code.

g. Narayan and Snyder [97] have generalized the concept of a cutoff rate parameter for a single-user channel to that of a "cutoff rate region" for a d.m. MAC \(K_{21}\). The use of the cutoff rate parameter in the study and design of single-user coded communication systems was advocated by Wozencraft and Kennedy [158] and Massey [93], respectively. In [97], the cutoff rate region for a synchronous d.m. MAC \(K_{21}\) is derived by minimizing the exponential error bound found by Slepian and Wolf [125], using Gallager's method [64, p. 142]. This cutoff rate region is a subregion of the capacity region \(C(K_{21}, I)\).

Narayan and Snyder [97] have shown that in situation \((K_{21}, I)\) the cutoff rate region is the same for a synchronous and quasi-synchronous MAC, provided the maximum relative delay by which the messages of the two senders can be out of synchronism is finite and known, as assumed in the model of Cover et al. [31]. In this connection, Gallager [65] has devoted a discussion to the problems involved when one tries to apply sequential decoding to a MAC. He mentions an example by Arikan, which indicates a fundamental problem with sequential decoding when applied to MAC's.
h. Following the work by Ahlswede and Dueck [7], who showed that for a d.m. one-way channel good codes, meeting the random coding bound, can be produced with relatively few (linear in the block length) permutations from a single codeword, Pokorny and Wallmeier [106] investigated the analogous problem for a d.m. MAC (K₂¹,₁). They succeeded in showing that for a d.m. MAC (K₂¹,₁) for which \( C(K₂¹,₁) = R(K₂¹,₁, QS) \), codes can be produced from a single "standard-codeword" with a linearly in the block length growing number of permutations, so that the average error probability associated with this code is upper-bounded by the random coding error bound obtained in [106] (and discussed in Sub-section 3e above), but then applied to the special case of MAC's for which convexification of the rate region is not necessary.

i. Haroutunian [72] has derived an exponential lower bound on the average probability of error of a d.m. MAC (K₂¹,₁), which reflects aspects of the sphere packing bound for the single-user channel. (He actually developed this bound for the more general d.m. MAC situation (K₂¹,II) to be discussed in Section 4.) This lower bound is generally not tight as compared to the exponential upper bound derived in [91] and [125], except for channels with a certain symmetric structure. Gallager [65] has pointed out why error exponents are far more complicated for multiple-access channels than for single-input channels. Dyachkov [51] obtained an upper bound for the average error probability of a d.m. MAC (K₂¹,₁) for constant composition codes, when maximum likelihood and universal decoding are considered. He established the region of rates where this random coding bound is asymptotically tight, and compared his bound with the random coding bound of [91].
4. THE MULTIPLE-ACCESS CHANNEL WITH SPECIALLY CORRELATED SOURCES IN THE 
SENSE OF SLEPIAN AND WOLF

Slepian and Wolf [125] considered a communication situation for the d.m. 
MAC $K_{21}$ in which the information to be transmitted by the two input users 
is correlated in a special way. This communication situation is shown in 
Fig. 5. Here, three message sources emit statistically independent messages 
$m_0 \in \mathcal{M}_0 = \{1, 2, \ldots, M_0\}$, $m_1 \in \mathcal{M}_1 = \{1, 2, \ldots, M_1\}$, and $m_2 \in \mathcal{M}_2 = \{1, 2, \ldots, M_2\}$, 
such that each triple $(m_0, m_1, m_2)$ occurs with probability $1/(M_0 M_1 M_2)$.

Message pair $(m_0, m_1)$ is encoded at terminal 1 by an encoding function $f_1$ into 
the codeword $f_1(m_0, m_1) = x_1^n(m_0, m_1) \in \mathcal{X}_1^n$. Likewise, message pair $(m_0, m_2)$ is 
encoded at terminal 2 by an encoding function $f_2$ into the codeword $f_2(m_0, m_2) = 
x_2^n(m_0, m_2) \in \mathcal{X}_2^n$. Next, the pair $(x_1^n(m_0, m_1), x_2^n(m_0, m_2))$ is transmitted over 
the d.m. MAC $\{\mathcal{X}_1 \times \mathcal{X}_2, p(y|x_1, x_2), Y\}$ using $n$ channel operations, and the 
decoder must estimate the source triple $(m_0, m_1, m_2)$ based on the received 
sequence $y^n$. Thus here, each input terminal attempts to get across one 
private and one common message to the output terminal. We denote this 
communication situation by $(K_{21}, II)$.

![Fig. 5. Slepian-Wolf configuration $(K_{21}, II)$ of a d.m. MAC with specially correlated sources.](image)
The actual inputs to the encoders in situation $(K_{21}, II)$ are of course correlated, since the mappings $f_1$ and $f_2$ are related to each other. This form of correlation is only a special case of the more general situation of two arbitrarily dependent source outputs. The three independent source outputs $m_0$, $m_1$, and $m_2$ can be thought of as being the result of a factorization of two dependent source outputs $u = (m_0, m_1)$ and $v = (m_0, m_2)$ into a common part $w = m_0$ and two private messages $m_1$ and $m_2$. Indeed, if $m_0$, $m_1$, and $m_2$ are independent, then $m_0$ can be regarded as the common part of $u$ and $v$ in the sense of Gács and Körner [62] and Witsenhausen [152]. (See Section 5 for the treatment of the d.m. MAC with arbitrarily correlated sources and the definition of the concept of "common part".)

An $(n, M_0, M_1, M_2, \lambda)$-code for the d.m. MAC $K_{21}$ in situation $(K_{21}, II)$ consists of $M_0 M_1$ codewords $x_1^n(m_0, m_1) \in \mathcal{X}_1^n$, $M_0 M_2$ codewords $x_2^n(m_0, m_2) \in \mathcal{X}_2^n$, and $M_0 M_1 M_2$ pairwise disjoint decoding sets $B(m_0, m_1, m_2) \subset \mathcal{Y}_2^n$, such that

$$\frac{1}{M_0 M_1 M_2} \sum_{(m_0, m_1, m_2)} P^n(B(m_0, m_1, m_2) | x_1^n(m_0, m_1), x_2^n(m_0, m_2)) \geq 1 - \lambda.$$  

(4.1)

A triple $(R_0, R_1, R_2)$ of non-negative real numbers is said to be an achievable rate triple for a d.m. MAC in situation $(K_{21}, II)$, if, for any $\delta > 0$ and any $\lambda$, $0 < \lambda < 1$, there exists, for all sufficiently large $n$, an $(n, M_0, M_1, M_2, \lambda)$-code for this MAC such that $\log M_i \geq n(R_i - \delta)$, $i = 0, 1, 2$. The capacity region $C(K_{21}, II)$ of a d.m. MAC in situation $(K_{21}, II)$ is defined as the set of all achievable rate triples in that situation.

Slepian and Wolf [125] gave the following characterization of the capacity region $C(K_{21}, II)$. 

Theorem 4.1 ([125]): The capacity region of the d.m. MAC in situation (K_{21}, II) is given by

\[ C(K_{21}, II) = \text{co}\{ (R_1, R_2, R_0) : 0 \leq R_1 \leq I(X_1; Y|X_2, U), \]

\[ 0 \leq R_2 \leq I(X_2; Y|X_1, U), \]

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

\[ 0 \leq R_1 + R_2 + R_0 \leq I(X_1, X_2; Y), \]

for some \( P(u, x_1, x_2, y) = P(u)P(x_1|u)P(x_2|u)P(y|x_1, x_2) \).

Here \( U \) is an auxiliary RV taking values in some finite set \( U \) of cardinality \( |U| \). Naturally, the region described in Theorem 4.1 would not be computable if one could not upper-bound \( |U| \). Slepian and Wolf [125] used a random coding argument to prove the achievability part of Theorem 4.1 and Fano's lemma to establish the weak converse. They also derived a random coding error exponent of the Gallager-type [63], upper-bounding the average error probability. In [123] it was furthermore conjectured that \( |U| \leq \exp_2 R_0 \) suffices. Since 1976 the following specific research progress regarding the d.m. MAC in situation (K_{21}, II) was made.

a. Several authors have investigated the problem of upper-bounding \( |U| \), and come up with various cardinality bounds. Salehi [112] proved that it suffices to take \( |U| \leq \min( \| X_1 \| \cdot \| X_2 \| , \| Y \| ) \), while Han [69] proved the weaker bound \( |U| \leq \| X_1 \| \cdot \| X_2 \| + 3 \). Both authors used the basic lemma by Ahlswede and Körner [10] (also called "Support Lemma" in [32]), which provides a general method to obtain cardinality bounds in multi-user information theory. Salehi [112] made also use of the Fenchel-Eggleston strengthening of Caratheodory's theorem [53, p. 35] observed by Wyner and Ziv [161].
Csiszár and Körner [32] and Willems [145] established independently a slightly different formulation of the capacity region $C(K_{21},II)$. They showed that in the description of the capacity region as given by Theorem 4.1, it is not necessary to take the convex hull. Corresponding with this different formulation of the capacity region, new cardinality bounds arise. Csiszár and Körner [32] proved that $|\mathcal{W}| \leq \|\mathcal{Y}\| + 3$ can be assumed, while Willems [145] showed that it suffices to take $|\mathcal{W}| \leq \min\{\|\mathcal{X}_{1}\| \cdot \|\mathcal{X}_{2}\| + 2, \|\mathcal{Y}\| + 3\}$. Because of the different formulations of the capacity region, the cardinality bounds by Salehi [112] and Han [69] on the one hand, and those obtained by Csiszár and Körner [32] and Willems [145] on the other, cannot be compared directly.

b. Dueck [49] has stated that his method, developed for proving a strong converse for the d.m. MAC in situation $(K_{21},I)$ in the case of average error, yields also a strong converse for the d.m. MAC in situation $(K_{21},II)$, again when the average error probability is used as performance criterion.

c. As already mentioned in Subsection 3i, Haroutunian [72] derived an exponential lower bound for the average probability of error of a d.m. MAC in situation $(K_{21},II)$, which reflects aspects of the sphere packing exponent of the single-user channel (cf. [32, p. 166]). He showed that, if the MAC has a certain symmetric structure, this lower bound is tight (i.e., coincides with the exponential upper bound derived in [125]) for all rate triples in a definite subregion of $C(K_{21},II)$, called the critical domain.

d. (The asymmetric MAC) Haroutunian [72], and then Bassalygo, Pinsker, and Prelov [13], considered the communication situation which results from the Slepian-Wolf configuration if one sets one of the private rates, $R_2$ say, equal to zero. This leads to a new communication situation, shown in Fig. 6, which is of independent interest. Here, two message sources emit statistically
independent messages $m_0 \in \mathcal{M}_0 = \{1, 2, \ldots, M_0\}$, and $m_1 \in \mathcal{M}_1 = \{1, 2, \ldots, M_1\}$, such that each pair $(m_0, m_1)$ occurs with probability $1/(M_0 M_1)$. Message pair $(m_0, m_1)$ is encoded at terminal 1 into a codeword $x_1^n(m_0, m_1) \in \mathcal{X}_1^n$, whereas message $m_0$, the output of source 0, is encoded at terminal 2 into a codeword $x_2^n(m_0, m_1) \in \mathcal{X}_2^n$. Thus, input terminal 1 attempts to get across the MAC both one common and one private message, whereas terminal 2 transmits only a common message to the output terminal. This communication situation, denoted by $(K_{21}, III)$, has been termed the "asymmetric MAC" (abbreviated AMAC) in [38], because of its analogy with the configuration of the asymmetric broadcast channel [32]. At this point we will omit the precise definitions of an $(n, M_0, M_1, \lambda)$-code and of a pair $(R_0, R_1)$ of achievable rates for a d.m. MAC in situation $(K_{21}, III)$, since these are easily derived from the corresponding definitions in situation $(K_{21}, II)$. The capacity region $C(K_{21}, III)$ of a d.m. MAC $K_{21}$ is defined as the set of all achievable rate pairs $(R_0, R_1)$ in situation $(K_{21}, III)$.

![Diagram](image-url)

Fig. 6. The d.m. MAC in situation $(K_{21}, III)$. (The asymmetric MAC.)
Haroutunian [72] stated a single-letter characterization of the capacity region \( C(K_{21},III) \) and outlined a proof of his result. Bassalygo, Pinsker, and Prelov [13] established the same characterization and gave a more detailed proof. This characterization is stated in the following theorem.

**Theorem 4.2 ([72],[13]):** The capacity region of the d.m. MAC in situation \( (K_{21},III) \) is given by

\[
C(K_{21},III) = \{ (R_0, R_1) : 0 \leq R_1 \leq I(X_1;Y|X_2), \quad (4.3a)
\]

\[
0 \leq R_0 + R_1 \leq I(X_1,X_2;Y), \quad (4.3b)
\]

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

If we denote the capacity region of a deterministic d.m. MAC in situation \( (K_{21},III) \) by \( C(K_{21},III,D) \), then Theorem 4.2 yields in particular the following characterization of \( C(K_{21},III) \) in the deterministic case, as was first observed in [13]:

\[
C(K_{21},III,D) = \{ (R_0,R_1) : 0 \leq R_1 \leq H(Y|X_2), \quad (4.4a)
\]

\[
0 \leq R_0 + R_1 \leq H(Y), \quad (4.4b)
\]

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

Characterization (4.4) can be regarded as the counterpart of characterization (3.6) for situation \( (K_{21},III) \). Bassalygo et al. [13] evaluated \( C(K_{21},III,D) \) for the deterministic BEMAC and found that in this case the corresponding capacity region is bounded by a curve consisting of three segments, which are respectively specified by the following formulas:

(i) \( R_0 + R_1 = \log_2 3, \ \frac{2}{3} \leq R_0 \leq \log_2 3; \)  
(ii) \( R_0 = 1 - \alpha, \ R_1 = h(\alpha) \) for \( \frac{1}{3} \leq \alpha \leq \frac{1}{2}; \)  
(iii) \( 0 \leq R_0 \leq \frac{1}{2}, \ R_1 = 1. \)  
This region is shown in Fig. 7, which can be regarded as the counterpart of Fig. 2 for situation \( (K_{21},III) \).
Fig. 7. The capacity region $C(K_{21},III)$ of the deterministic BEMAC.

$(h(2/3) = 0.9183; \log_2 3 = 1.5850.)$

Bassalygo et al. [13] particularly concerned themselves with the problem of the transmission of information over a d.m. MAC in situation $(K_{21},III)$ when it is assumed that the upper bound $\lambda$ on the average error probability (cf. $\lambda$ in (4.1)) equals zero. This setup requires some additional definitions. A rate pair $(R_0,R_1)$ is said to be zero-error achievable for a d.m. MAC in situation $(K_{21},III)$, if, for any $\delta > 0$, there exists, for all sufficiently large $n$, an $(n,M_0,M_1,0)$-code for this MAC such that $\log M_i \geq n(R_i - \delta)$, $i = 0,1$. The zero-error capacity region $C_0(K_{21},III)$ is defined as the set of all zero-error achievable rate pairs $(R_0,R_1)$ in situation $(K_{21},III)$. Analogously one may define the zero-error capacity region denoted by $C_0(K_{21},II)$ as the set of all zero-error achievable rate triples $(R_0,R_1,R_2)$.
in situation $(K_{21},II)$. Bassalygo et al. [13] noted that for a general d.m. MAC $K_{21}$ the regions $C_0(K_{21},II)$ and $C_0(K_{21},III)$ are still unknown. However, they [13] were able to characterize $C_0(K_{21},III)$ in the deterministic case. Let us denote the zero-error capacity region of a deterministic d.m. MAC in situation $(K_{21},III)$ by $C_0(K_{21},III,D)$. Then the result by Bassalygo, Pinsker, and Prelov [13] says that $C_0(K_{21},III,D) = C(K_{21},III,D)$, i.e., that for a deterministic d.m. MAC in situation $(K_{21},III)$ the zero-error capacity region and the "ordinary" capacity region (as given by Theorem 4.2) coincide. We state their result as a separate theorem.

**Theorem 4.3 ([13]):** For a deterministic d.m. MAC one has

\[
C_0(K_{21},III) = C(K_{21},III), \quad \text{i.e.,}
\]

\[
C_0(K_{21},III,D) = C(K_{21},III,D).
\]  

(4.5)

Following the methods introduced by Ahlswede and Dueck [7] for the single-user channel, De Bruyn and van der Meulen [38] developed two codeconstructions for the d.m. MAC in situation $(K_{21},III)$, one being an iterative codeconstruction, and the other one being a codegeneration by relatively few (linear in the block length) permutations. In particular, it is shown in [38] that, starting from a single codeword, a code for the d.m. AMAC can be generated by either procedure, thereby keeping the average error probability exponentially upper-bounded, with an error exponent of the informational divergence type, as discussed in [32]. This error exponent can be shown to be equivalent to the error exponent of Slepian and Wolf [125], when $R_2$ is taken to be zero, and thus is positive for all rate pairs $(R_0,R_1) \in C(K_{21},III)$. This result can be regarded as the counterpart of the result by Pokorny and Wallmeier [106] (obtained for the case $(K_{21},I)$, and described in Sub-section 3b) for the d.m. MAC $(K_{21},III)$. However, for this particular situation of a d.m.
AMAC, the result in [38] shows more, as it is proven there not only that
a code for the d.m. AMAC can be produced by permutations, but also that
such code can be generated stepwise.

A separate survey of recent results on the d.m. AMAC appeared in [135].

e. Following the results in [7] and [38], De Bruyn [34] investigated
the problem of finding good codes, generated by few permutations, for the
general d.m. MAC (K_{21},II), i.e., in the Slepian and Wolf [125] setup. She
showed that it is possible to generate a code for the d.m. MAC (K_{21},II)
iteratively with relatively few permutations starting from one single code-
word, so that the associated average probability of error can be exponentially
upper-bounded, with an exponent of the informational divergence type, which
is positive only for rate triples in a definite subregion of C(K_{21},II).
Alternatively, the result by Pokorny and Wallmeier [106] can be extended to the
case (K_{21},II). However, in the case of two private messages (case (K_{21},I))
some basic difficulties arise, so that the problem of code generation
according to the methods of Ahlswede and Dueck [7] is still not completely
solved for the case (K_{21},II).

f. Ahlswede [2] considered the problem of transmitting information
from three independent sources over a d.m. MAC with three input terminals,
when each source output is connected to exactly one input terminal. We denote
a d.m. MAC with three input terminals by K_{31}, the particular communication
situation considered by Ahlswede [2] by (K_{31},I), and the corresponding capacity
region by C(K_{31},I). Ahlswede [2] gave as single-letter characterization
of C(K_{31},I). Ulrey [130] continued these investigations and considered the
problem of transmitting $t \geq 3$ separate messages from $t$ independent sources
over a d.m. MAC with $t$ input terminals. We denote a d.m. MAC with $t \geq 3$
input terminals by K_{t1}, and the situation in which the $i$th encoder
(1 ≤ i ≤ t) observes only the output of the i-th source by \((K_{t1}, I)\). Ulrey [130] characterized the corresponding capacity region, which we denote by \(C(K_{t1}, I)\). Van der Meulen [132] gave a simplified proof of Ulrey's converse theorem.

Slepian and Wolf [125] extended their approach (i.e., the setup of situation \((K_{21}, II)\)) to the case in which information from seven independent sources is to be transmitted over a d.m. MAC \(K_{31}\) and this information, when supplied to the input terminals, is again correlated in a special way. This communication situation may be denoted by \((K_{31}, II)\). Slepian and Wolf [125] provided an expression for the capacity region in this case. They also conjectured a solution for the capacity region in the case of a d.m. MAC with \(t = N\) input terminals and \(S = 2^N - 1\) sources, when the sources are again connected to the input terminals in a special way, similar to the configuration of situation \((K_{21}, II)\).

Han [69] continued these investigations and considered the general communication situation in which information is to be transmitted from \(s ≥ 2\) independent sources over a d.m. MAC \(K_{t1}\), such that the source outputs are connected to the input terminals in an arbitrary manner, thus extending formally the Slepian and Wolf [125] setup of configuration \((K_{21}, II)\) to the case of a d.m. MAC with \(t ≥ 3\) terminals. Here, as in the case \((K_{21}, II)\), the \(s\) sources need not to be into a one-to-one correspondence with the \(t\) input terminals, neither needs \(s\) to be equal to \(t\). In order to describe a specific communication situation (within a whole class of possibilities), Han [69] used the notion of an incidence relation \(δ(s, t) = (δ_1, δ_2, \ldots, δ_t)\) between the \(s\) sources and the \(t\) input terminals, which prescribes for each terminal \(i\) \((1 ≤ i ≤ t)\) the subset \(δ_i\) of sources observed by that terminal. Thus, each incidence relation \(δ(s, t)\) specifies for a given d.m. MAC \(K_{t1}\) and a given
collection of $s$ independent sources a certain communication situation in which
the information to be transmitted is correlated in a special way, similar to
the type of correlation arising in the Slepian and Wolf [125] situation
($K_{21,II}$). The entire communication situation thus defined—consisting of the
d.m. MAC $K_{t1}$, the $s$ independent sources, and the incidence relation $\delta(s,t)$—
is here denoted by $(K_{t1},\delta(s,t))$. Following the terminology of [69] and [70],
we may call this communication situation also a d.m. MAC with cross observation
at the encoders. A typical configuration of this kind is shown in Fig. 8.
It includes the situations considered by Slepian and Wolf [125] and Ulrey
[130] as special case.

---

**Fig. 8.** The d.m. MAC with $s$ sources, $t$ input terminals, and cross
observation at the encoders.
Han [69] established the capacity region $C(K_{t1}, \delta(s,t))$ of a d.m. MAC $K_{t1}$ in situation $(K_{t1}, \delta(s,t))$, for an arbitrary incidence relation $\delta(s,t)$ with $s$ and $t$ finite, by proving a direct part (using a typicality approach) and a weak converse. He showed that his expression for the capacity region includes the expressions obtained by Ahlswede [2] (case $(K_{21}, I)$), Slepian and Wolf [125] (case $(K_{21}, II)$), and Ulrey [130] (case $(K_{t1}, I)$) as special cases. E.g., for the case $(K_{21}, II)$ — in Han's notation the situation $(K_{21}, \delta(3,2))$, with $\delta(3,2) = (\delta_1, \delta_2)$, $\delta_1 = \{1,0\}$, $\delta_2 = \{2,0\}$ — the corresponding expression for the capacity region $C(K_{21}, \delta(3,2))$ obtained by Han [69] takes the following form. Let, for some auxiliary alphabets $U_1, U_2, U_0$ of finite cardinality, and mappings $f_1 : U_1 \times U_0 \rightarrow \mathcal{X}_1$, $f_2 : U_2 \times U_0 \rightarrow \mathcal{X}_2$

\[
R(U_1, U_2, U_0, f_1, f_2) = \{(R_1, R_2, R_0) : 0 \leq R_1 \leq I(U_1; Y|U_0, U_2), (4.6a)
0 \leq R_2 \leq I(U_2; Y|U_0, U_1), (4.6b)
R_1 + R_2 \leq I(U_1, U_2; Y|U_0), (4.6c)
R_0 + R_1 + R_2 \leq I(U_0, U_1, U_2; Y), (4.6d)
0 \leq R_0 \leq I(U_0; Y|U_1, U_2), (4.6e)
R_0 + R_1 \leq I(U_0, U_1; Y|U_2), (4.6f)
R_0 + R_2 \leq I(U_0, U_2; Y|U_1), (4.6g)
\]

for some $P(u_1, u_2, u_0, y) = P(u_1)P(u_2)P(u_0)P(y|f_1(u_1, u_0), f_2(u_2, u_0))$.

Then, Han's expression for $C(K_{21}, II)$ equals

\[
C(K_{21}, \delta(3,2)) = \co \cup R(U_1, U_2, U_0, f_1, f_2), (4.7)
\]

where the union is taken over all choices of finite auxiliary alphabets $U_1$, $U_2, U_0$, and mappings $f_1, f_2$. 
As the region \( C(K_{t1}, \delta(s,t)) \) found by Han [69] is rather complicated to describe in general, its precise formulation will be omitted here. Essential and novel in Han's approach is his use of a test channel. Han [69] provided also cardinality bounds for the auxiliary random variables he had introduced to characterize the region \( C(K_{t1}, \delta(s,t)) \).

Prelov [111], in his investigations of the d.m. MAC \( K_{t1} \) with a special source hierarchy (see Sub-section 4.g below), noted that the expression for the capacity region of the situation \((K_{31},II)\) given by Slepian and Wolf [125, pp. 1058-1059] is incorrect, but that the expression given by Han [69] in this case, i.e., the capacity region \( C(K_{31}, \delta(7,3)) \) with a proper description of the incidence relation \( \delta(7,3) \), is correct. Prelov [111] demonstrated this with an example he credited to S. I. Gel'fand.

g. Prelov[110],[111] considered the transmission of \( t \geq 3 \) independent sources over a d.m. MAC \( K_{t1} \), when the transmission configuration exhibits a special source hierarchy. Specifically, it is assumed that all message sources are accessible to the first encoder, all sources except the first one are accessible to the second encoder, all sources except the first two are accessible to the third encoder, and so forth. This situation can be regarded as the generalization of the d.m. AMAC (case \((K_{21},III)\)) to the case of \( t \geq 3 \) sources and terminals. Alternatively, this communication situation fits also into the framework of the configuration considered by Han [69], for a special choice of the incidence relation \( \delta(t,t), \delta^2 \) say. This situation may be called a d.m. MAC \( K_{t1} \) with a degraded (independent) set of messages; it could be denoted by either \((K_{t1},III)\), or \((K_{t1}, \delta^2(t,t))\), in analogy with the situations considered above.

Prelov[110],[111] obtained for the communication situation under consideration results which are generalizations of those in [13]. First, he gave a
characterization of the capacity region $C(K_{t1},III)$ which does not involve auxiliary random variables $U_1, U_2, \ldots, U_t$, but rather is based on joint probability assignments $P(x_1, x_2, \ldots, x_t)$ on inputs, as in [13] for the case $t = 2$. Secondly, he showed that for a deterministic d.m. MAC $(K_{t1}, \delta^N(t,t))$ the ordinary capacity region and the zero-error capacity region coincide.

h. Wolf [155] also considered the situation in which $t$ independent sources transmit information over a d.m. MAC $K_{t1}$, but he was more concerned with the largest sum rate $R_{\text{SUM}} = R_1 + R_2 + \ldots + R_t$ that can be achieved under various assumptions on cooperation. Wolf [155] distinguished between three different communication situations for a d.m. MAC $K_{t1}$ with $t$ independent sources, and, accordingly, defined three different capacity concepts. These three situations are:

(i) The total cooperation case. This is the generalization of situation $(K_{21}, T)$ to the case of $t$ input terminals, yielding situation $(K_{t1}, T)$. Here, there is a single encoder which is able to observe all $t$ source outputs. The total cooperation capacity will now be denoted by $C_T(K_{t1})$ and is defined as

$$C_T(K_{t1}) = \max_{P(x_1, x_2, \ldots, x_t)} I(X_1, X_2, \ldots, X_t; Y), \quad (4.8)$$

in analogy with (3.4). Here the maximum is taken over all possible joint distributions $P(x_1, x_2, \ldots, x_t)$. If we denote the sum rate in this case by $R_{\text{SUM}}(K_{t1}, T)$, then one has: $R_{\text{SUM}}(K_{t1}, T) \leq C_T(K_{t1})$.

(ii) The partial cooperation case. This is the communication situation $(K_{t1}, I)$ considered by Ulrey [130]. If we denote the sum rate in this case by $R_{\text{SUM}}(K_{t1}, I)$, one has the following upper bound:

$$R_{\text{SUM}}(K_{t1}, I) \leq \max_{Q(x_1, x_2, \ldots, x_t)} I(X_1, X_2, \ldots, X_t; Y). \quad (4.9)$$
Here the maximum is taken over all PD's $Q(x_1, x_2, \ldots, x_L)$ such that $Q$ is a product PD. Following Wolf [155], the right-hand side of (4.9) may be called the "partial cooperation capacity".

(iii) The no cooperation case. This is a new communication situation also considered in [155]. Here it is assumed that the $i$th encoder observes only the $i$th source output, as in situation $(K_{t1}, I)$, but that now there are $t$ different decoders, the $i$th decoder being asked to estimate only the $i$th message, and knowing only the code used by the $i$th encoder. According to [155], this is a situation often encountered in practice. We denote this situation by $(K_{t1}, NC)$, and the corresponding rate sum by $R_{SUM}(K_{t1}, NC)$. Then, as observed by Wolf [155], one has the following upper bound:

$$R_{SUM}(K_{t1}, NC) \leq \sum_{i=1}^{t} \max_{Q(x_1, x_2, \ldots, x_L)} I(X_i; Y). \quad (4.10)$$

Here the maximum is again taken with respect to all product PD's. The right-hand side of (4.10) is called the "no cooperation capacity" in [155]. The latter communication concept is further considered in Section 11.

Wolf [155] also considered the static versus dynamic assignment of codes to a set of $s$ users communicating over a d.m. MAC $K_{t1}$, when it is assumed that no more than $t$ of the users will transmit at any given time and $t << s$. In the static assignment situation each of the $s$ users is given its own set of codewords, which is fixed throughout, thus requiring $s$ distinct codes. In the dynamic assignment situation a code can be reassigned to the next user, once a user becomes inactive. The latter strategy would require only $t$ distinct codes. Several interesting problems arise, also from the viewpoint of capacity. The reader is referred to [155] for further details.
5. THE DISCRETE MEMORYLESS MULTIPLE-ACCESS CHANNEL WITH ARBITRARILY CORRELATED SOURCES

Clearly, the correlation assumed between the messages encoded by the two users in the channel model of Slepian and Wolf [125] is of a special form. In [125] the authors raise the question how to handle more general correlations. Consider thereto the communication situation shown in Fig. 9. On the one hand, there is given a bivariate information source \((\mathcal{U} \times \mathcal{V}, P(u,v))\), putting out independent, identically distributed (i.i.d.) discrete RV's \((U_i, V_i), i = 1, 2, \ldots\), according to an arbitrary but fixed PD \(\Pr\{U_i = u, V_i = v\} = P(u,v), u \in \mathcal{U} \text{ and } v \in \mathcal{V}\), where \(\mathcal{U}\) and \(\mathcal{V}\) are finite sets. On the other hand, there is given a d.m. MAC \(K_{21}\) as defined in Section 3.

Input terminal 1 observes the sequence of outputs of source \(\mathcal{U}\), whereas input terminal 2 observes those of source \(\mathcal{V}\). These source output sequences are subsequently encoded into input sequences of the channel, which are then transmitted. At the receiver end the emitted source sequences are to be decoded. The problem now is to determine at what rates the dependent source outputs can be transmitted over the d.m. MAC with arbitrarily small probability of error.

![Diagram](image)

Fig. 9. The d.m. MAC \(K_{21}\) with two arbitrarily correlated sources.
In a way, this problem traces back to Shannon [123, Section 14], who raised the problem of transmitting a pair of correlated messages over a two-way channel. (Cf. [9] for a discussion of the importance of this problem.) Slepian and Wolf [124],[125] solved the problem of transmitting two correlated messages over a d.m. MAC $K_{21}$ for two special cases: (i) the communication situation ($K_{21}, II$) discussed in Section 4 and (ii) the case where the MAC is lossless. In the latter case the problem coincides with the problem of noiseless coding of two correlated information sources (also called Slepian and Wolf data compression), for which a complete solution was found in [124]. The communication situation exhibited in Fig. 9, will be denoted by $(K_{21}, (U, V), I)$.

The first general treatment of the d.m. MAC $K_{21}$ with arbitrarily correlated sources has been given by Cover, El Gamal, and Salehi [29], who established a sufficient (but not necessary) condition for the reliable transmission of an arbitrarily correlated source over a d.m. MAC $K_{21}$. In the sequel we will follow the setup of [29].

A code of block length $n$ in situation $(K_{21}, (U, V), I)$, denoted by $(\phi_1^n, \phi_2^n, \psi^n)$, consists of two encoding functions

\begin{equation}
\phi_1^n : U^n \rightarrow X_1^n, \tag{5.1a}
\end{equation}

\begin{equation}
\phi_2^n : V^n \rightarrow X_2^n, \tag{5.1b}
\end{equation}

mapping source outputs into codewords, and a decoding function

\begin{equation}
\psi^n : Y^n \rightarrow U^n \times V^n. \tag{5.1c}
\end{equation}

The probability of error when using this code is given by

\begin{equation}
p^n_e = \Pr\{(U^n, V^n) \neq \psi^n(Y^n)\}
= \sum_{(u^n, v^n)} p^n(u^n, v^n) p^n(\psi^n(Y^n) \neq (u^n, v^n) | \phi_1^n(u^n), \phi_2^n(v^n)), \tag{5.2}
\end{equation}
where the joint probability assignment is given by

\[ p_n^{(u^n, v^n, y^n)}(u^n, v^n) = \prod_{i=1}^{n} p(u_i, v_i) p(y_i | \phi_{1i}(u^n), \phi_{2i}(v^n)). \tag{5.3} \]

The correlated source pair \((\mathcal{U} \times \mathcal{V}, P(u,v))\) is said to be reliably transmissible over a d.m. MAC in situation \((K_{21}, (U,V), I)\) if for any \(\lambda, 0 < \lambda < 1\), and sufficiently large \(n\) there exists a code \((\phi_1^n, \phi_2^n, \psi^n)\) for which \(P_n \mathcal{E} < \lambda\). The main problem then is to find a simple characterization of the set of all correlated sources that are reliably transmissible over a given d.m. MAC \(K_{21}\).

There are two versions of the achievability theorem of Cover, El Gamal, and Salehi [29], one involving the common part of \(U\) and \(V\), and a simpler version omitting the common part. Before proceeding, we pause to give the definition of the notion of common part in the sense of Gács and Körner [62] and Witsenhausen [152].

The common part \(W\) of two random variables \(U\) and \(V\) is defined by finding the largest possible integer \(k \geq 2\) such that there exist functions \(f\) and \(g\), \(f : \mathcal{U} \rightarrow \{1, 2, \ldots, k\}\) and \(g : \mathcal{U} \rightarrow \{1, 2, \ldots, k\}\) with \(\Pr\{f(U) = i\} > 0\), \(\Pr\{g(V) = i\} > 0\), \(i = 1, 2, \ldots, k\), such that \(f(U) = g(V)\) with probability one, and then defining \(W = f(U) = g(V)\).

We first state the simpler version of the coding theorem of Cover et al. [29] for the d.m. MAC in situation \((K_{21}, (U,V), I)\), which assumes that the RV's \(U\) and \(V\) contain no common part.

**Theorem 5.1** ([29]): A correlated source \((\mathcal{U} \times \mathcal{V}, P(u,v))\) can be reliably transmitted over a given d.m. MAC \(K_{21}\), in situation \((K_{21}, (U,V), I)\), if there exist probability mass functions \(P(x_i | u)\) and \(P(x_2 | v)\) such that
\begin{align}
H(U|V) &< I(X_1;Y|X_2,V), \quad (5.4a) \\
H(V|U) &< I(X_2;Y|X_1,U), \quad (5.4b) \\
H(U,V) &< I(X_1,X_2;Y), \quad (5.4c)
\end{align}

where \( P(u,v,x_1,x_2,y) = P(u,v)P(x_1|u)P(x_2|v)P(y|x_1,x_2). \)

Cover et al. [29] proved Theorem 5.1 (and its generalization Theorem 5.2 below as well) by employing a random coding argument. As Ahlswede and Han [9] observed, Theorem 5.1 forms the heart of the more general Theorem 5.2. Notice that Theorem 5.1 provides only sufficient (but not necessary) conditions for reliable transmission.

Cover et al. [29] noted that applications of Theorem 5.1 yield several known results as special cases: (i) the direct part of the Slepian and Wolf [124] data compression theorem, (ii) the achievability part of the coding theorem for the d.m. MAC with independent sources (situation \((K_{21},I)\)), as proved by Ahlswede [2], [3] and announced by Liao [91] (cf. Theorem 3.1 above), and (iii) the achievability of the total cooperation capacity \(C_T\) defined by (3.5).

Cover et al. [29] gave an example showing that in general the procedure consisting of factorizing the source-channel transmission problem into two separate problems, i.e., Slepian and Wolf data compression applied to the bivariate source \((\mathcal{U} \times \mathcal{U}, P(u,v))\) followed by the transmission of two independent messages over a d.m. MAC in the sense of situation \((K_{21},I)\), (called the separation principle in [9]), is not optimal, thus motivating the usefulness of Theorem 5.1. The example in [29] involves the deterministic BEMAC defined in Section 3, and the correlated source \((\mathcal{U} \times \mathcal{U}, P(u,v))\), defined by \(\mathcal{U} = \mathcal{U} = \{0,1\}\) and the probability assignment \(P(0,0) = P(1,0) = P(0,1) = 1/3\). For this
example, plain factorization does not yield reliable transmission at all, but it is easy to see that with the simple choices $X_1 = U$ and $X_2 = V$ error-free transmission of the source outputs over the MAC is possible.

Cover, El Gamal, and Salehi [29], in the random code construction underlying the proofs of Theorems 5.1 and 5.2, exploited the dependency structure of the correlated message source. We quote [29]:

"To allow partial cooperation between the two transmitters, we allow our codes to depend statistically on the source outputs. This induces dependence between codewords."

Ahlswede and Han [9] called the coding technique introduced in [29] "a kind of correlation-preserving mapping".

Dueck [48] observed that the construction of Cover, El Gamal, and Salehi [29] is such that the "channel input at time $t$ depends only on the source output at time $t". This observation led him to carefully design an instructive counterexample which shows that the coding strategy of Cover et al. [29] for the d.m. MAC with correlated sources is not optimal in the case $(K_{21}, (U, V), I)$. The counterexample constructed by Dueck [48] involves the definition of a correlated source $(\mathcal{U}, \mathcal{V}, P(u,v))$ and a d.m. MAC $K_{21}$ such that the source does not fulfill the conditions of Theorem 5.1 with respect to the MAC. The correlated source presented by Dueck [48] is chosen to be highly asymmetrical and contains no common part. Dueck [48] showed that for this correlated source a strategy like "a large information source output at time $t$ is encoded in several channel inputs" is definitely better than the coding strategy of Cover, El Gamal, and Salehi [29]. In particular, Dueck [48] demonstrated that, when using such a special strategy, reliable transmission of the chosen source over the given MAC is possible.
We now state the second version of the coding theorem of Cover et al. [29], which is applicable when the source RV's do have a common part.

**Theorem 5.2 ([29]):** A correlated source \((U \times U, P(u,v))\), having a common part \(W = f(U) = g(V)\), can be reliably transmitted over a given d.m. MAC \(K_{21}\) in situation \((K_{21},(U,V),I)\) if there exist probability mass functions \(P(s)\) on a set \(S\), \(P(x_1|u,s)\) on \(\mathcal{X}_1\), and \(P(x_2|v,s)\) on \(\mathcal{X}_2\), such that

\[
\begin{align*}
H(U|V) &< I(X_1;Y|X_2,V,S), \\
H(V|U) &< I(X_2;Y|X_1,U,S), \\
H(U,V|W) &< I(X_1,X_2;Y|W,S), \\
H(U,V) &< I(X_1,X_2;Y),
\end{align*}
\]  

(5.5)

where \(P(s,u,v,x_1,x_2,y) = P(s)P(u,v)P(x_1|u,s)P(x_2|v,s)P(y|x_1,x_2)\).

Cover et al. [29] proved that the region described by Theorem 5.2 is convex and remarked that, if reliable transmission is possible, then in order to generate a random code for error-free transmission, it suffices to consider only those auxiliary RV's \(S\) whose cardinality is bounded above by \(\min\{|\mathcal{X}_1|, |\mathcal{X}_2|, |Y|\}\). Furthermore, it is shown in [29] that Theorem 5.2 includes as a special case the direct part of the coding theorem by Slepian and Wolf [125] for a d.m. MAC in situation \((K_{21},II)\), as given by Theorem 4.1.

Theorem 5.2 is the most general single letter characterization of an achievable rate region for situation \((K_{21},(U,V),I)\) presently known, but Dueck's counter-example [48] shows that this region is not optimal. In addition, Cover et al. [29] characterized the capacity region in situation \((K_{21},(U,V),I)\) by a limiting expression, which however is not computable. They extended the latter characterization to the case of \(t \geq 3\) correlated sources, the outputs of which are to be communicated reliably over a d.m. MAC \(K_{t1}\), whereby each source output is connected
to precisely one channel input.

Ahlswede and Han[9] continued the investigations of [29] and [48] regarding the d.m. MAC $K_{21}$ with arbitrarily correlated sources. To begin with, these authors emphasized the importance of Shannon's paper [123] in connection with the problem of transmitting correlated messages over a noisy channel with two or more senders, thereby quoting some of Shannon's remarks. Next, in Section II of their paper, Ahlswede and Han[9] took a new look at the coding theorem of Cover, El Gamal, and Salehi[29] for situation $(K_{21}(U,V),I)$ (cf. Theorem 5.2 above), viz. from the viewpoint of cross observation at the encoders, as discussed in Section 4 (cf. [69],[70]). In doing so, they[29] provided another proof of Theorem 5.2, by first establishing a simple extension of Theorem 5.1 involving three correlated information sources and a d.m. MAC $K_{21}$ with cross observation at the encoders, and then showing that this result contains Theorem 5.2 as a special case.

Consider in this respect the joint source-channel configuration shown in Fig. 10. Here is given a trivariate information source $(U_1 \times U_2 \times U_3, P(u_1,u_2,u_3))$, producing i.i.d. discrete RV's $(U_1^{u_1},U_2^{u_2},U_3^{u_3}), i = 1,2,\ldots$, according to an arbitrary but fixed PD $Pr(U_1 = u_1, U_2 = u_2, U_3 = u_3) = P(u_1,u_2,u_3)$. Moreover, there is given a d.m. MAC $K_{21}$. The channel is connected to the source through cross observation at the encoders. This source-channel configuration can be regarded as the generalization of the Slepian and Wolf setup (situation $(K_{21},\Pi)$) from the case of three independent sources to that of-three arbitrarily correlated sources. Using the notation introduced in Subsection 4.f, we may describe this source-channel configuration by the incidence relation $\delta(3,2) = (\delta_1,\delta_2)$, with $\delta_1 = \{1,2\}$ and $\delta_2 = \{2,3\}$. Encoder 1 observes the sequences of outputs of sources $U_1$ and $U_2$, and maps these into an input sequence $\Sigma_1$ via an encoding function $\phi_1: U_1^\times U_2^\rightarrow \mathcal{X}_1^n$. 
Fig. 10. Trivariate correlated information source and two-user d.m. MAC
with cross observation at the encoders.

Encoder 2 observes the output sequences of sources \( U_2 \) and \( U_3 \), and maps these into an input sequence \( X_2^n \) via an encoding function \( \phi_2^n : U_2^n \times U_3^n \rightarrow X_2^n \). These input sequences are subsequently transmitted over the channel. The decoder observes a received output sequence \( Y^n \) and must reconstruct the three emitted source sequences using a decoding function \( \psi^n : Y^n \rightarrow U_1^n \times U_2^n \times U_3^n \). The probability of error incurred when using such an encoding and decoding system is defined analogously to (5.2). We denote this communication situation by \( (K_2', (U_1, U_2, U_3), II) \). The correlated source triple \( (U_1 \times U_2 \times U_3, P(u_1, u_2, u_3)) \) is said to be reliably transmissible over a d.m. MAC in situation \( (K_2', (U_1, U_2, U_3), II) \), if for any \( \lambda, 0 < \lambda < 1 \), and sufficiently large \( n \) there exists a code \( (\phi_1^n, \phi_2^n, \psi^n) \) for which the average probability of error \( P_e^n \) satisfies \( P_e^n < \lambda \).

Ahlsweide and Han [9] obtained sufficient conditions in order that an arbitrarily correlated triple source is reliably transmissible over a d.m. MAC in situation \( (K_2', (U_1, U_2, U_3), II) \), thereby generalizing Theorem 5.2 above. We now state those conditions in a theorem, which is a reformulation of Theorem 1 of [9].
Theorem 5.3 ([9]) : A correlated trivariate information source 
\((U_1 \times U_2 \times U_3, P(u_1, u_2, u_3))\) can be transmitted reliably over a given 
d.m. MAC \(K_{21}\) in situation \((K_{21}, (U_1, U_2, U_3), II)\), if there exist random variables 
\(S_1, S_2, S_3\), conditional probability distributions \(P(s_1 | u_1), P(s_2 | u_2), P(s_3 | u_3)\), 
and functions \(f_1 : S_1 \times S_2 \rightarrow X_1, f_2 : S_2 \times S_3 \rightarrow X_2\), such that 
\[
H(U_1 | U_2, U_3) < I(S_1; Y | S_2, S_3, U_2, U_3), \tag{5.6a}
\]
\[
H(U_2 | U_1, U_3) < I(S_2; Y | S_1, S_3, U_1, U_3), \tag{5.6b}
\]
\[
H(U_3 | U_1, U_2) < I(S_3; Y | S_1, S_2, U_1, U_2), \tag{5.6c}
\]
\[
H(U_1, U_2 | U_3) < I(S_1, S_2; Y | S_3, U_3), \tag{5.6d}
\]
\[
H(U_2, U_3 | U_1) < I(S_2, S_3; Y | S_1, U_1), \tag{5.6e}
\]
\[
H(U_1, U_3 | U_2) < I(S_1, S_3; Y | S_2, U_2), \tag{5.6f}
\]
\[
H(U_1, U_2, U_3) < I(S_1, S_2, S_3; Y), \tag{5.6g}
\]

where \(P(u_1, u_2, u_3, s_1, s_2, s_3, y) =
\[
P(u_1, u_2, u_3)P(s_1 | u_1)P(s_2 | u_2)P(s_3 | u_3)P(y | f_1(s_1, s_2), f_2(s_2, s_3)).
\]

In the statement and proof of their result, Ahlswede and Han [9] make use
of an associated test channel. The proof in [9] parallels the random coding
argument used by Cover et al. [29] to prove the simpler version (Theorem 5.1
above) of their coding theorem for situation \((K_{21}, (U, V), I)\). Once the sufficient
conditions for reliable transmission in situation \((K_{21}, (U_1, U_2, U_3), II)\), as given
by Theorem 5.3, are established, Ahlswede and Han [9] derive Theorem 5.2 as
a special case. This way of deriving Theorem 5.2 makes Ahlswede and Han remark
that Theorem 5.1 forms the heart of Theorem 5.2.
Ahlswede and Han [9] also established a coding theorem for a system of source coding with side information via a MAC. Such a system can be regarded as an extension of the noiseless source coding system considered by Ahlswede and Körner [10] and Wyner [160]. Moreover, a new approach is given in [9] to the problem of transmitting correlated messages over a d.m. MAC, based on the concept of a correlated code. Several new coding problems are then formulated. The reader is referred to [9, Section VII] for more details of this new approach.

To complete this section, we now turn to the communication situation of an asymmetric d.m. MAC $K_{21}$ with dependent sources, i.e., to the case of sending information from an arbitrarily correlated bivariate information source over a d.m. MAC $K_{21}$, as in the situation $(K_{21}, \text{III})$ discussed in Subsection 4.d. This communication situation results from the situation $(K_{21}, (U_1, U_2, U_3), \text{II})$ by omitting there the source $U_3$. The communication configuration is shown separately in Fig. 11. This situation was investigated by De Bruyn, Prelov, and van der Meulen [37], who established necessary and sufficient conditions for reliable transmission in this case. In order to describe these results, we first define the concepts which are needed.

![Diagram](image_url)

**Fig. 11.** A d.m. asymmetric MAC $K_{21}$ with arbitrarily correlated sources.
Let be given a bivariate information source \((\mathcal{U} \times \mathcal{Y}, P(u,v))\), emitting i.i.d. discrete RV's \((U_i, V_i), i = 1, 2, \ldots\), according to an arbitrary but fixed PD \(\Pr(U_i = u, V_i = v) = P(u,v), u \in \mathcal{U}, v \in \mathcal{Y}\), where \(\mathcal{U}\) and \(\mathcal{Y}\) are finite sets. Furthermore, let be given a d.m. MAC \(K_{21}\). The encoder at terminal 1 observes the sequences of outputs of both source \(\mathcal{U}\) and source \(\mathcal{Y}\), and maps these into an input sequence \(X_1^n\) via an encoding function \(\phi_1^n : \mathcal{U}^n \times \mathcal{Y}^n \to \mathcal{X}_1^n\). The encoder at terminal 2 observes only the output sequences of source \(\mathcal{Y}\), and maps these into an input sequence \(X_2^n\) via an encoding function \(\phi_2^n : \mathcal{Y}^n \to \mathcal{X}_2^n\). These input sequences are subsequently transmitted over the channel. The decoder observes a received output sequence \(Y^n\) and must reconstruct the two produced source sequences using a decoding function \(\psi^n : Y^n \to \mathcal{U}^n \times \mathcal{Y}^n\). The probability of error incurred when using this encoding and decoding system is defined as in (5.2). We denote this communication situation by \((K_{21}, (U,V), \text{III})\). The correlated source pair \((\mathcal{U} \times \mathcal{Y}, P(u,v))\) is said to be reliably transmissible over a d.m. MAC in situation \((K_{21}, (U,V), \text{III})\), if for any \(\lambda, 0 < \lambda < 1\), and sufficiently large \(n\) there exists a code \((\phi_1^n, \phi_2^n, \psi^n)\) for which the average error probability \(P_e^n\) satisfies \(P_e^n < \lambda\). De Bruyn, Prelov, and van der Meulen [37] now proved the following result.

Theorem 5.4 ([37]): A correlated source pair \((\mathcal{U} \times \mathcal{Y}, P(u,v))\) can be transmitted reliably over a given d.m. MAC \(K_{21}\) in situation \((K_{21}, (U,V), \text{III})\) if there exists a probability distribution \(P(x_1, x_2)\) on \(\mathcal{X}_1 \times \mathcal{X}_2\) such that

\[
H(U|V) < I(X_1;Y|X_2), \tag{5.7a}
\]

\[
H(U,V) < I(X_1, X_2;Y), \tag{5.7b}
\]

where \(P(x_1, x_2, y) = P(x_1, x_2) P(y|x_1, x_2)\).
Conversely, if a correlated source pair \((\mathcal{U} \times \mathcal{V}, P(u,v))\) can be transmitted reliably over a given d.m. MAC \(K_{21}\) in situation \((K_{21}, (U,V), III)\), then the following inequalities must be true for some probability distribution \(P(x_1,x_2)\):

\[
H(U|V) \leq I(X_1; Y|X_2), \tag{5.8a}
\]

\[
H(U,V) \leq I(X_1, X_2; Y). \tag{5.8b}
\]

Thus, Theorem 5.4 provides necessary and sufficient conditions for the reliable transmission of an arbitrarily correlated source pair over a d.m. AMAC, but with no statement when \(H(U|V) = I(X_1; Y|X_2)\) or \(H(U,V) = I(X_1, X_2; Y)\). This boundary situation should be examined for each source-channel combination separately.

The achievability part of Theorem 5.4 is proven in [37] by applying the Slepian-Wolf coding theorem for correlated sources [124, case 0111], and subsequently the coding theorem for the d.m. MAC in situation \((K_{21}, III)\) (Theorem 4.2 above). The converse part is proven by use of Fano's inequality. It is remarkable that for the situation \((K_{21}, (U,V), III)\) the separation principle holds, notwithstanding the fact that in the case \((K_{21}, (U,V), I)\) it is suboptimal.

In [37] the following example is given as an illustration of Theorem 5.4. Consider again the correlated source \((\mathcal{U} \times \mathcal{V}, P(u,v))\) defined by \(\mathcal{U} = \mathcal{V} = \{0,1\}\) and the assignment \(P(0,0) = P(0,1) = P(1,1) = 1/3\), and the deterministic BEMAC defined in Section 3. Suppose we wish to transmit this particular source over this MAC according to situation \((K_{21}, (U,V), III)\). There exists no PD \(P(x_1,x_2)\) such that the inequalities (5.7) are satisfied for this source-channel combination. However, there does exist a PD \(P(x_1,x_2)\) which yields equality in (5.8), viz. the one induced by the encodings \(x_1 \equiv U\).
and $X_2 = V$. The capacity region $C(K_{21},III)$ of this d.m. AMAC is shown in Fig. 12, together with the rate region of the given source ([124, case 0111]). Also is indicated the point where the ideal communication takes place.

![Diagram](image)

Fig. 12. An optimal point for the transmission of a certain correlated source over the asymmetric BEMAC.

6. THE GAUSSIAN MULTIPLE-ACCESS CHANNEL

In Section 3, 4, and 5 we considered only discrete MAC's in the sense that the input and output alphabets were discrete and the information to be transmitted through the MAC consisted of discrete sequences of symbols. In this section we will discuss MAC's with continuous input and output alphabets, in particular the Gaussian MAC. There are two cases to distinguish:
(i) the discrete-time additive white Gaussian noise MAC and (ii) the so-called spectral Gaussian MAC.

The discrete-time additive white Gaussian noise (AWGN) MAC with two input users, input power constraints $P_1$ and $P_2$, and noise variance $N$ is defined by the following conditions: (i) at any particular time the operation of the MAC is given by the equation

$$Y = X_1 + X_2 + Z,$$

where $X_1$ and $X_2$ are the input RV's, $Z$ is a Gaussian noise RV with mean 0 and variance $N$ ($Z \sim \mathcal{N}(0,N)$) independent of $(X_1, X_2)$, and $Y$ is the resulting output RV, all four RV's taking values in the set of real numbers; (ii) when input sequences of length $n > 1$ are transmitted over the MAC the channel operation is given by

$$Y^n = X^n_1 + X^n_2 + Z^n,$$

where $X^n_1 = (X^n_{11}, \ldots, X^n_{1n})$, $X^n_2 = (X^n_{21}, \ldots, X^n_{2n})$, $Z^n = (Z^n_1, \ldots, Z^n_n)$ and $Y^n = (Y^n_1, \ldots, Y^n_n)$, $Z^n$ is independent of $(X^n_1, X^n_2)$, and $Z^n_1, \ldots, Z^n_n$ are i.i.d. RV's with common distribution $\mathcal{N}(0,N)$; and (iii) all input sequences $x^n_1 = (x^n_{11}, \ldots, x^n_{1n})$ and $x^n_2 = (x^n_{21}, \ldots, x^n_{2n})$ which are to be transmitted must satisfy for any block length $n$ the average power constraints

$$\frac{1}{n} \sum_{t=1}^{n} x^2_{it} \leq P_i, \quad i = 1, 2.$$  

The simplest communication situation of a discrete-time AWGN MAC with two senders is the case where two independent messages are to be transmitted, as in situation $(K_{21}, I)$. This communication situation is sketched in Fig. 13. Here the source output $m_1$ is encoded into the codeword $x^n_1(m_1)$ and source output $m_2$ is encoded into the codeword $x^n_2(m_2)$. These codewords are transmitted over
the AWGN MAC, whose operation has just been defined, and an output sequence \( Y^n \) is produced. Applying the decoding rule on this received vector, the source outputs are then to be reconstructed. We denote this communication situation by MAC (AWGN \( K_{21}, N, P_1, P_2, I \)), or, since it is clear that we are dealing in this paper with the MAC only, by (AWGN \( K_{21}, N, P_1, P_2, I \)). The formal definitions of a code, error probability, achievable rate pair, and capacity region are the same as the ones given in Section 3 for situation \( (K_{21}, I) \), except that now discrete probabilities are replaced by probability densities and the constraints (6.3) are imposed on the codewords. We denote the capacity region in this situation by \( C(\text{AWGN } K_{21}, N, P_1, P_2, I) \).

Cover [26] and Wyner [159] established independently an achievable rate region for the MAC (AWGN \( K_{21}, N, P_1, P_2, I \)). This region can be derived from the characterization of \( C(K_{21}, I) \) given in Theorem 3.1 by substituting probability densities for discrete probabilities, taking into account the constraints (6.3), and making use of the Gaussian structure of the channel. However, Cover [26] and Wyner [159] provided direct arguments to prove achievability, based on the well-known properties of the single-input Gaussian channel [118].

![Image of a diagram](image_url)

**Fig. 13.** Discrete-time additive white Gaussian noise MAC with two input users, power constraints \( P_1 \) and \( P_2 \), and independent noise variable \( Z \sim N(0, N) \). (Situation (AWGN \( K_{21}, N, P_1, P_2, I \)).)
Such direct proof is also given in [65]. A perspicuous proof of the achievable rate region based on the notion of joint typicality is provided in [56], where also the extension of the same result is given to the case of a discrete-time AWGN MAC with $t \geq 2$ independent users.

Although the achievable rate region found by Cover [26] and Wyner [159] is indeed the capacity region, no formal proof of the converse appeared until Keilers [89] gave a proof of the weak converse in his Ph. D. dissertation. For sake of completeness, we combine the results of [26], [159], and [85] into one theorem.

**Theorem 6.1 ([26], [159], [85]):** The capacity region of the discrete-time MAC $(\text{AWGN } K_{21}, N, P_1, P_2, I)$ is given by

$$
C(\text{AWGN } K_{21}, N, P_1, P_2, I) = \left\{ (R_1, R_2) : 0 \leq R_1 \leq \frac{1}{2} \ln \left(1 + \frac{P_1}{N}\right), \right. \\
0 \leq R_2 \leq \frac{1}{2} \ln \left(1 + \frac{P_2}{N}\right), \\
R_1 + R_2 \leq \frac{1}{2} \ln \left(1 + \frac{P_1 + P_2}{N}\right) \right\}. 
$$

For the particular choices $P_1 = P_2 = 10$, $N = 1$, the region $C(\text{AWGN } K_{21}, N, P_1, P_2, I)$ is shown in Fig. 19 below (Section 7), together with the region achievable when feedback is allowed, under the name "capacity region without feedback".

Keilers [85] also considered a $k$-dimensional discrete-time additive non-white (arbitrarily correlated) Gaussian noise MAC with $m \geq 2$ users, which he called the spectral Gaussian MAC. This channel is pictured in Fig. 14 for the case of two input users, and can be described as follows.

---

Prof. T. Cover has informed the author that Carroll W. Keilers died before he could publish this result and his other results mentioned in this section in the open, refereed literature.
Source output \( m_i \) (\( i = 1, 2 \)) is encoded into \( X_{1i}^{kn} (m_i) \), consisting of \( k \) \( n \)-dimensional vectors \( (X_{11}^n, X_{12}^n, \ldots, X_{1k}^n) \) of real numbers. At each moment \( t \), \( t = 1, 2, \ldots, n \), the \( k \)-dimensional vector consisting of the \( t \)-th components of the vectors \( X_{1j}^n \), \( j = 1, 2, \ldots, k \), is presented to the channel, i.e., the vector \( X_{1i}^{k} = (X_{11t}, X_{12t}, \ldots, X_{1kt}) \). Thus, the codeword \( X_{1i}^{kn} (m_i) \) is fed into the channel as an \( n \)-sequence of \( k \)-dimensional vectors. Both codewords \( X_{1i}^{kn} (m_1) \) and \( X_{2i}^{kn} (m_2) \) are transmitted, and an output sequence \( \tilde{y}_{kn} \) is received, upon which the decoder applies its decoding function \( \psi_{kn} : y^{kn} \rightarrow M_1 \times M_2 \).

The operation of this spectral MAC is statistically defined by the equation \( \tilde{y}^{k} = X_{1}^{k} + X_{2}^{k} + Z^{k} \), where \( X_1 \) and \( X_2 \) are independent, \( X_{1j}^{k} \) is the \( k \)-th power of \( X_{1j} \) with respect to the independent product distribution (\( i = 1, 2 \)), and \( Z^{k} \) is a \( k \)-dimensional Gaussian random variable (generally not being the product of \( k \) independent one-dimensional Gaussian RV's). In this discussion it is further understood that we focus again on the simplest communication situation, viz. the one where two independent messages are transmitted as in situation (\( K_{21}, I \)), with no cross observation at the encoders. Moreover, we assume that all input sequences \( X_{1i}^{kn} \) and \( X_{2i}^{kn} \) satisfy for any block length \( n \) the average power constraints

\[
\frac{1}{n} \sum_{t=1}^{n} \sum_{j=1}^{k} x_{1jt}^2 \leq P_i, \quad i = 1, 2. \tag{6.5}
\]

In this model of a \( k \)-dimensional discrete-time additive non-white Gaussian noise MAC with 2 users it is assumed that at each time instant \( t \) (\( t = 1, 2, \ldots, n \)) the \( k \)-dimensional noise vector \( Z_{-t}^k = (Z_{1t}, \ldots, Z_{kt}) \) has a \( k \)-variate normal distribution with arbitrary covariance matrix \( A_Z \), and that the \( n \) \( k \)-dimensional noise vectors \( Z_{1}^k, \ldots, Z_{n}^k \) are statistically independent. The definitions of a code, error probability, achievable rate pair, and capacity region are straightforward for this channel, except that it should be emphasized that
Fig. 14. K-dimensional discrete-time additive non-white Gaussian noise MAC with two input users, power constraints $P_1$ and $P_2$, and independent Gaussian noise process $\underline{Z}^k = (Z_1, \ldots, Z_k)$ having an arbitrary covariance matrix $\Lambda_Z$; equivalent to a set of $k$ parallel independent additive white Gaussian noise MAC's.

Each codeword for each user should satisfy the power constraints (6.5).

However, as Keilers [85] observed, "each user can apportion power as desired among the $k$ dimensions of the process as long as the total user power is preserved by the sum".

Keilers [85] showed the equivalence of the above model to that of a collection of $k$ parallel independent AWGN MAC's, whereby the noise vector at time instant $t$, $\underline{Z}^k = (Z_{1t}, \ldots, Z_{kt})$, consists of $k$ independent Gaussian RV's each with mean zero and respective variances $N_1, \ldots, N_k$, i.e., $Z_{jt} \sim \mathcal{N}(0, N_j)$, $j = 1, \ldots, k$; $t = 1, \ldots, n$. Keilers [85] proved this argument by considering a matrix diagonalization procedure whereby the covariance matrix
\( \Lambda_z \) is transformed into a diagonal covariance matrix \( \Lambda_z' \) with as principal diagonal elements the variances \( N_1, \ldots, N_k \). He argued that the collection of \( k \) parallel independent AWGN MAC's thus obtained can serve as an approximation for a continuous-time spectral Gaussian noise MAC, thereby following the approach outlined in Gallager [64, Section 8.5] for reducing the continuous-time AWGN one-way channel to a set of parallel independent AWGN one-way channels. Keilers [85] observed that under the transformation of the original model to that of \( k \) parallel independent AWGN MAC's the power constraints (6.5) are preserved, so that the transformed channel is equivalent to the original channel. Thus, achievability theorems and converses can be proved for either channel model, and the extension to the other model is immediate.

As one of his main results, Keilers [85] established the capacity region of a set of \( k \) parallel independent AWGN MAC's with \( m \geq 2 \) input users. Following the above line of reasoning, this result can then be carried over to the spectral Gaussian MAC. For simplicity we state Keilers' theorem just for the case of two input users.

**Theorem 6.2 ([85]):** The capacity region of a collection of \( k \) parallel independent discrete-time AWGN MAC's with two input users, total input powers \( P_1 \) and \( P_2 \) such that the constraints (6.5) are satisfied, and respective noise variances \( N_1, \ldots, N_k \) is given by

\[
C = \text{co} \left\{ (R_1, R_2) : 0 \leq R_1 \leq \sum_{j=1}^{k} \frac{1}{2} \ln \left( 1 + \frac{P_{j1}}{N_j} \right), \right. \\
0 \leq R_2 \leq \sum_{j=1}^{k} \frac{1}{2} \ln \left( 1 + \frac{P_{j2}}{N_j} \right), \\
0 \leq R_1 + R_2 \leq \sum_{j=1}^{k} \frac{1}{2} \ln \left( 1 + \frac{P_{j1} + P_{j2}}{N_j} \right),
\]

such that \( 0 \leq P_{ji}, j = 1, \ldots, k; i = 1, 2; \) and \( \sum_{j=1}^{k} P_{ji} = P_i, i = 1, 2 \).
Keilers [85] proved the achievability part of Theorem 6.2 by direct application of the results by Cover [29] and Wyner [159]. He also obtained an alternate form of the achievable rate region in terms of the extreme points which generate it. He furthermore showed that these extreme points are described by a sequential water-fill strategy as in Shannon [119]. Moreover, he proved a weak converse.

As a final remark we note that the investigations by Keilers [85] concerning the spectral Gaussian channel can be regarded as dual to the work by Hughes-Hartogs [76] and El Gamal [54] who considered the spectral Gaussian broadcast channel (cf. also [32, Section 11]).

7. THE MULTIPLE-ACCESS CHANNEL WITH COMPLETE FEEDBACK

7.1. Preliminaries

In practical multiple-access situations it is often possible to make use of feedback signals from the receiver to the senders. Consider in this regard the communication situation sketched in Fig. 15. This situation consists of a d.m. MAC with two input users and two independent sources, as in situation $(K_{21}, 1)$, but now with a common noiseless feedback link available from the output to each of the encoders. In this section, and Sections 8 and 9, we will distinguish between different types of feedback for the MAC. With the term "noiseless feedback", or "perfect feedback", or just "feedback" to an encoder we mean that at each time instant $t$ ($t = 1, 2, \ldots, n$) this encoder is allowed to observe an exact copy of the previous output signal $y_{t-1}$ before the transmission of its next input symbol. The feedback may be available to both encoders (as in Fig. 15), or to only one of them (as in Fig. 20 below).
Fig. 15. The d.m. MAC ($K_{21}, I$) with complete feedback.

In the first case we will speak of "complete" feedback, and in the second case of "partial" feedback. (Some authors call the first case "two-sided" feedback and the second case "one-sided" feedback (e.g. [17]), but we prefer to adhere to the terminology defined here so as to allow for possible extensions to MAC's with more than two input terminals.) Complete feedback will be the subject of this section, and partial feedback that of Section 8. In addition, we will discuss the situation where the senders observe a signal that is correlated with, but not necessarily identical to the signal observed by the receiver. This situation, pictured in Fig. 23 below, is called "generalized feedback", and will be the topic of Section 9.

In summary, we will focus on three different types of feedback: (i) complete (noiseless) feedback (this section), (ii) partial (noiseless) feedback (Section 8), and (iii) (complete) generalized feedback (Section 9).

It is a well-known classical result that for the one-way d.m. channel noiseless feedback from the receiver to the single sender cannot increase the capacity ([120][42]). However, as was first demonstrated by Gaarder and Wolf [61], this is generally not the case for a d.m. MAC with complete feedback, or even for a MAC with partial feedback, as we shall see in Section 8.
Indeed, in a MAC with feedback the senders may be able to cooperate with each other, as the feedback allows them to understand each other's transmissions. This observation led Gaarder and Wolf [61] to the striking result that for the deterministic BEMAC complete feedback can enlarge the capacity region. In this section, and in Sections 8 and 9, we shall treat successively the discrete memoryless MAC and the Gaussian MAC in each of the three different feedback situations considered.

7.2 The discrete memoryless MAC with complete feedback

Using an example, Gaarder and Wolf [61] showed that for the d.m. MAC noiseless feedback to both encoders can increase the capacity region. In particular, they gave a two-stage coding scheme for the deterministic BEMAC with which the symmetric rate pair \((R_1, R_2) = (0.76018, 0.76018)\) can be achieved. This rate pair falls outside the classical capacity region \(C(K_2, I)\), which for this channel is sketched in Fig. 2 above.

Subsequently, Cover and Leung [30] proved a generally achievable rate region for the d.m. MAC with complete feedback. Their region is based on a generally applicable two-stage feedback scheme which improves on the specific scheme of Gaarder and Wolf [61] in the case of the deterministic BEMAC with feedback. The proof of the result by Cover and Leung [30] involves the notions of joint typicality, superposition, list codes, and block Markov encoding. Before stating the result of Cover and Leung [30], we first give some definitions. Consider hereto the communication situation of Fig. 15.

In Fig. 15 two message sources produce statistically independent messages \(m_1 \in \{1, 2, ..., M_1\}\) and \(m_2 \in \{1, 2, ..., M_2\}\), each pair \((m_1, m_2)\) occurring with equal probability \(1/(M_1M_2)\). Each encoder is described by a sequence of encoding functions. These functions map the message and the sequence of previous channel outputs into the next channel input. Thus, at time instant \(t, 1 \leq t \leq n,\)
the encoding function $f_1 t$ maps the message $m_1$ and the sequence of channel outputs received up to time $t$, $y^{t-1}$, into the next channel input at terminal 1, i.e., $x_{1t} = f_{1t}(m_1, Y^{t-1})$. Similarly, the encoding operation at time $t$ at terminal 2 is described by $x_{2t} = f_{2t}(m_2, Y^{t-1})$. The decoder gives estimates of $m_1$ and $m_2$ based on his knowledge of the sequence of $n$ channel outputs via a decoding function $g$, i.e., $g(Y^n) = (\hat{m}_1, \hat{m}_2)$. An $(n, M_1, M_2, \lambda)$-code for the d.m. MAC $K_{21}$ with complete feedback consists of two sets of $n$ encoding functions $f_{1t}$ and $f_{2t}$, and a decoding function $g$, such that the probability of error $P_e^n$ defined by $P_e^n = \Pr\{(\hat{m}_1, \hat{m}_2) \neq (m_1, m_2)\}$ satisfies $P_e^n < \lambda$. We denote this communication situation by $(K_{21}, I, CFB)$. A pair $(R_1, R_2)$ of non-negative real numbers is said to be an achievable rate pair for a d.m. MAC in situation $(K_{21}, I, CFB)$, if for any $\delta > 0$ and any $\lambda$, $0 < \lambda < 1$, there exists for all sufficiently large $n$ an $(n, M_1, M_2, \lambda)$-code for this MAC such that $\log M_i \geq n(R_i - \delta)$, $i = 1, 2$. The capacity region, which is as usual defined as the set of all achievable rate pairs, is in this case defined by $C(K_{21}, I, CFB)$.

For a general d.m. MAC $K_{21}$, a computable characterization of the capacity region $(K_{21}, I, CFB)$ is not known at present. However, several partial results towards establishing a characterization are available. We begin with mentioning the following three specific results: (i) the achievable rate region found by Cover and Leung [30], (ii) the outer bound on the capacity region first formulated by Gaarder and Wolf [61], and (iii) the result by Willems [144] that for d.m. MAC's belonging to class $W_1$ or $W_2$ (cf. Section 3) the Cover-Leung rate region is indeed the capacity region. We will first describe these results. Later in this subsection we will mention some other results. We start with stating the achievability theorem of Cover and Leung [30].

Theorem 7.1 ([30]): For a d.m. MAC in situation $(K_{21}, I, CFB)$ an achievable rate region is given by
\[ R_1(X_{21}, I, \text{CFB}) = \text{co} \{ (R_1, R_2) : 0 \leq R_1 \leq I(X_1; Y | X_2, U), \] (7.1a)

\[ 0 \leq R_2 \leq I(X_2; Y | X_1, U), \] (7.1b)

\[ R_1 + R_2 \leq I(X_1, X_2; Y), \] (7.1c)

for some \( P(u, x_1, x_2, y) = P(u)P(x_1|u)P(x_2|u)P(y|x_1, x_2) \)

such that \( \| U \| \leq \min\{ \| X_1 \|, \| X_2 \|, \| Y \| \} \). \]

Cover and Leung [30] attribute the cardinality bound in Theorem 7.1 to Salehi [112]. Willems [144] has shown that, in the description of the achievable rate region as given by Theorem 7.1, it is not needed to take the convex hull, and that corresponding with this new representation the cardinality bound \( \| U \| \leq \min\{ \| X_1 \|, \| X_2 \| + 1, \| Y \| + 2 \} \) suffices. These two different cardinality bounds cannot be compared directly. (See also in this regard the discussion in Subsection 4a concerning cardinality bounds on \( U \) in relation to the auxiliary RV figuring in characterization (4.2).)

Gaarder and Wolf [61] stated, and Ozarow [101], [102] proved, the following outer bound on the capacity region \( C(K_{21}, I, \text{CFB}) \).

**Theorem 7.2** ([61], [101], [102]): The capacity region of a d.m. MAC in situation \((K_{21}, I, \text{CFB})\) is included in the outer bound region given by

\[ R_{OB}(K_{21}, I, \text{CFB}) = \text{co} \{ (R_1, R_2) : 0 \leq R_1 \leq I(X_1; Y | X_2), \] (7.2a)

\[ 0 \leq R_2 \leq I(X_2; Y | X_1), \] (7.2b)

\[ R_1 + R_2 \leq I(X_1, X_2; Y), \] (7.2c)

for some \( P(x_1, x_2, y) = P(x_1, x_2)P(y|x_1, x_2) \).
Moreover, Ozarow [101] proved that in the description of the outer bound region as given by Theorem 7.2 the convexification is unnecessary. In fact, in [102] the convex hull notation is omitted.

Gaarder and Wolf [61] conjectured that the region (7.2) is only an outer bound and that, in general, not all rates in this region are achievable. Willems [144] confirmed this conjecture by proving the remarkable fact that for a particular class of d.m. MAC's, viz. those MAC's which belong to either class \( \mathcal{W}_1 \) or \( \mathcal{W}_2 \) defined in Section 3, the Cover-Leung [30] achievable rate region stated in Theorem 7.1 is indeed the capacity region. The proof of the weak converse given by Willems [144] makes, in one of its steps, use of the special structure of a MAC belonging to \( \mathcal{W}_1 \) or \( \mathcal{W}_2 \). Denoting the capacity region of a d.m. MAC belonging to class \( \mathcal{W}_1 \) in situation \((K_{21}, I, CFB)\) by \( C(K_{21}, I, CFB, \mathcal{W}_1) \), we now state its characterization as obtained by Willems [144] as a separate theorem.

**Theorem 7.3 ([144]):** For a d.m. MAC belonging to class \( \mathcal{W}_1 \) and in situation \((K_{21}, I, CFB)\) the capacity region is given by:

\[
C(K_{21}, I, CFB, \mathcal{W}_1) = \{(R_1, R_2) : 0 \leq R_1 \leq H(X_1|U), \quad (7.3a) \\
0 \leq R_2 \leq I(X_2; Y|X_1, U), \quad (7.3b) \\
R_1 + R_2 \leq I(X_1, X_2; Y), \quad (7.3c)
\]

for some \( P(u, x_1, x_2, y) = P(u)P(x_1|u)P(x_2|u)P(y|x_1, x_2) \)

such that \( \|U\| \leq \min \{|X_1|, |X_2| + 1, |Y| + 2\} \).

A similar characterization of the capacity region can be given for a MAC belonging to class \( \mathcal{W}_2 \). For a deterministic MAC belonging to class \( \mathcal{W} = \mathcal{W}_1 \cap \mathcal{W}_2 \) we denote the capacity region \( C(K_{21}, I, CFB) \) by \( C(K_{21}, I, CFB, D, \mathcal{W}) \).
Specializing Theorem 7.3 to this case we find the following characterization for its capacity region:

\[
C(K_{21}, I, CFb, D, \mathcal{W}) = \{(R_1, R_2) : 0 \leq R_1 \leq H(X_1 | U), \quad (7.4a)
\]

\[
0 \leq R_2 \leq H(X_2 | U), \quad (7.4b)
\]

\[
R_1 + R_2 \leq H(Y), \quad (7.4c)
\]

for some \( P(u, x_1, x_2, y) = P(u)P(x_1 | u)P(x_2 | u)P(y | x_1, x_2) \)

such that \( \| \mathcal{W} \| \leq \min \{ \| X_1 \| \cdot \| X_2 \| + 1, \| Y \| + 2 \} \).

Thus, as observed by Willems [144], [147], the capacity region \( C(K_{21}, I, CFb) \) of the deterministic BEMAC, which belongs to class \( \mathcal{W} \), is given by expression (7.4), with as cardinality constraint \( \| \mathcal{W} \| \leq 5 \).

Using the formulation of Theorem 7.1, van der Meulen [131] showed that for the deterministic BEMAC the rate pair \((R_1, R_2) = (0.791, 0.791)\) can be achieved by putting \( Pr[U = 0] = Pr[U = 1] = 0.5 \), and letting \( X_1 \) and \( X_2 \) be conditionally independent given \( U = u \) with \( Pr[X_i \neq u] = 0.23763 \). Improved calculations by Cover and Leung [30] showed that the rate pair \((R_1, R_2) = (0.79113, 0.79113)\) can be achieved by putting \( Pr[U = 0] = Pr[U = 1] = 0.5 \), and letting \( X_1 \) and \( X_2 \) be conditionally independent given \( U = u \) with \( Pr[X_i \neq u] = 0.237663, i = 1, 2 \). Recently, Willems [147] showed that for the equal rate case (i.e., \( R_1 = R_2 \)) this latter probability assignment is optimal, i.e., that the symmetrical rate pair \((R_1, R_2)\) lies on the boundary of the capacity region \( C(K_{21}, I, CFb) \) of the deterministic BEMAC. This latter result by Willems [147] shows two other interesting facts pertaining to the deterministic BEMAC with complete feedback: (i) the boundary of the capacity region \( C(K_{21}, I, CFb) \) of this MAC does not have a point in common with the total cooperation line, as in this case \( C_T = \log_2 3 = 1.58496 > 1.58226 = 2(0.79113) \), and
(ii) the auxiliary RV U that yields the optimal symmetrical rate pair is binary, which is surprising in view of the fact that the general cardinality constraint derived by Willems [144] gives $|\mathcal{U}| \leq 5$ for this MAC. The capacity region $C(K_{21}, I, CFB)$ of the deterministic BEMAC is shown in Fig. 16, together with the optimal symmetrical rate point $D = (R_1, R_2) = (0.79113, 0.79113)$ and the total cooperation line $L_T$. For comparison, there is also shown in Fig. 16 the contour of the classical capacity region $C(K_{21}, I)$, which has been fully drawn in Fig. 2 above.

Following the investigations of [30] and [147], Vinck [137] described a simple constructive superposition coding scheme for the deterministic BEMAC with complete feedback which achieves the symmetrical rate point $(R_1, R_2) = (0.79095, 0.79095)$, and thus has sum rate $R_{\text{SUM}} = R_1 + R_2 = 1.5819$. Vinck's coding method is related to Schalkwijk's [114] strategy for coding on the unit square.

Vinck, Hoeks, and Post [138] investigated for two specific deterministic d.m. MAC's $K_{21}$ with M-ary inputs $(M \geq 2)$ the Cover-Leung rate region given by Theorem 7.1. These specific MAC's are: (i) the deterministic M-ary erasure MAC (also considered in [20] and called the A channel there), and (ii) the deterministic M-ary adder MAC (referred to as the C channel in [138]). Both MAC's belong to the class $\mathcal{U}$ introduced by Willems [144], and thus for both MAC's the capacity region $C(K_{21}, I, CFB)$ is given by expression (7.4). Vinck et al. [138] examined in particular which symmetrical rate points in this region are optimally achievable. They found that for both types of MAC's the optimally achievable symmetric rate point in the Cover-Leung region lies on the total cooperation line $L_T$ for large enough values of $M$. For the M-ary erasure MAC this happens when $M \geq 6$, and for the M-ary adder MAC when $M \geq 3$. 
Fig. 16. Capacity region $C(K_{21}, I, CFB)$ of the deterministic BEMAC, shown together with optimal symmetrical rate pair $D$ and total cooperation line $L_T$.

In an unpublished manuscript, Paul [103] examined the $T$-user deterministic binary adder MAC with feedback. (In [103] this channel is abbreviated NAMAC (noiseless adder multiple-access channel), whereas elsewhere the name T-user binary adder (BAC) is used [19], [20]. For $T = 2$ this channel reduces to the deterministic BEMAC.) The T-user deterministic binary adder MAC has input alphabets $X_i = \{0, 1\}$, $i = 1, \ldots, T$, output alphabet $Y = \{0, 1, \ldots, T\}$, and the channel operation is defined by $\Pr(Y = X_1 + X_2 + \ldots + X_T) = 1$. Continuing the investigations of Gaarder and Wolf [61], Paul [103] presented automatic repeat request (ARQ) schemes for use on the three-user and four-user BAC and showed that for these MAC's feedback allows information transfer at rates
above the one-way capacity. In the course of his investigations Paul [103] derived the so-called ARQ bound for the T-user BAC. This bound provides an upper bound on the highest sum rate \( R_{\text{SUM}} = R_1 + R_2 + \ldots + R_T \) which can be achieved by any ARQ scheme, and is given by the inequality

\[
R_{\text{SUM}} \leq \frac{T}{1 + \frac{2^{-T}}{\log(T+1)}} \sum_{i=1}^{T-1} \left( \frac{T}{i} \right) \frac{\log(T)}{\log(i)}.
\]  

(7.5)

For \( T = 2 \), (7.5) becomes \( R_{\text{SUM}} \leq 1.52038 \), and for \( T = 3 \) it yields \( R_{\text{SUM}} \leq 1.88163 \). Paul [103] defined the one-way capacity of the T-user BAC by the expression

\[
C_{\text{OW}} = \max_{P(X_1)P(X_2)\ldots P(X_T)} I(X_1,X_2,\ldots,X_T;Y),
\]

(7.6)

which coincides with the right-hand side of expression (4.9). For \( T = 2 \), \( C_{\text{OW}} = 1.5 \), and for \( T = 3 \), \( C_{\text{OW}} = 1.81128 \). Paul [103] noted that, when considering ARQ schemes, profound differences occur between the three-user and two-user BAC, and also between the four-user and the three-user BAC.

Paul [103] applied the Huffman criterion for optimal codeword length to the encoding of the retransmitted information to bound the performance of optimal ARQ strategies. For the case \( T = 3 \), he presented an ARQ scheme meeting the Huffman specification for minimum average codeword length and achieving a sum rate \( R_{\text{SUM}} = 1.83051 > C_{\text{OW}} \). He also presented a non-optimal asymptotic scheme achieving a sum rate equal to 1.84615 bits/\text{use}. For the four-user BAC Paul [103] showed that the Huffman limit cannot be achieved, and he presented an asymptotic ARQ strategy for this MAC achieving a sum rate equal to 2.07240 bits/\text{use} which is larger than \( C_{\text{OW}} = 2.03064 \).
Vinck, Hoeks, and Post [138] investigated also the T-user BAC with complete feedback. For the case $T = 3$, they gave two coding strategies which are not optimal, but are above the ARQ bound. Those strategies use the results of [137]. The first strategy gives $R_{\text{SUM}} = 1.903$ bits/transmission, whereas the second strategy achieves $R_{\text{SUM}} = 1.91$ bits/transmission.

Willems [147] applied his result - i.e., that for the deterministic BENAC with complete feedback the symmetrical rate point $(R_1, R_2) = (0.79113, 0.79113)$ lies on the boundary of the capacity region - to demonstrate the noteworthy fact that the feedback capacity region $C(K_2, I, CFB)$ of the product of two MAC's can be strictly larger than the (Minkowski) sum of the corresponding feedback capacity regions for the separate channels. A similar result was obtained by Poltyrev [107] and El Gamal [54] for the product of two degraded broadcast channels (cf. [32, problem 3.4.8]). These results are in contrast with the theory for single input-single output-channels. Namely, for such channels it is a well-known fact that the capacity of the product of two channels equals the sum of the capacities of the separate channels.

Vinck [139] showed that, as far as two-user binary input MAC's with feedback are concerned, there are only two non-trivial MAC's of that kind.

Hekstra and Willems [74] continued the investigations of Willems [144], and were able to extend the class of channels introduced in [144] (the class $W_1 \cup W_2$ in the notation of this paper) for which the Cover-Leung region (as given by Theorem 7.1) is the feedback capacity region. This class can be formulated as follows. Suppose for the d.m. MAC $(X_1 \times X_2, P(y|x_1, x_2), Y)$ there exists an alphabet $\mathcal{B}$ and mappings $h_i : \mathcal{Y} \times \mathcal{X}_i \rightarrow \mathcal{B}$, $i = 1, 2$, such that $h_1(y, x_1) = h_2(y, x_2) (\Delta b)$ for all $x_1, x_2, y$ for which $P(y|x_1, x_2) \neq 0$. Consider the transition probability matrix $\{P^+(y, b|x_1, x_2)\}$ defined by the
operation \( P^+(y, b|x_1, x_2) = P(y|x_1, x_2) \) if \( b = h_1(y, x_1) = h_2(y, x_2) \), and 0 otherwise. Then the MAC \( K_{21} \) with transition probability matrix \( \{P(y|x_1, x_2)\} \) is said to belong to class \( \mathcal{K} \) if

\[
\frac{P(x_1)P(x_2)P^+(y, b|x_1, x_2)}{P(y, b)} = \frac{\sum_{x_2} P(x_1)P(x_2)P^+(y, b|x_1, x_2)}{P(y, b)} \quad \frac{\sum_{x_1} P(x_1)P(x_2)P^+(y, b|x_1, x_2)}{P(y, b)}
\]

(7.8)

for all \( x_1, x_2, y, b \) for which \( P(y, b) \neq 0 \). Clearly, the class \( \mathcal{K} \) includes the class \( \mathcal{W}_1 \cup \mathcal{W}_2 \). Hekstra and Willems [74] now showed that for a d.m. MAC \( K_{21} \) belonging to class \( \mathcal{K} \) the capacity region in situation (\( K_{21}, I, CFB \)) is, as for MAC's belonging to \( \mathcal{W}_1 \cup \mathcal{W}_2 \), given by

\[
\mathcal{C}(K_{21}, I, CFB, \mathcal{K}) = \{ (R_1, R_2) : 0 \leq R_1 \leq I(X_1; Y|X_2, U), \\
0 \leq R_2 \leq I(X_2; Y|X_1, U), \\
R_1 + R_2 \leq I(X_1, X_2; Y), \}
\]

(7.9a)

(7.9b)

(7.9c)

for some \( P(u, x_1, x_2, y) = P(u)P(x_1|u)P(x_2|u)P(y|x_1, x_2) \)

such that \( \|U\| \leq \min \{ \|X_1\| \cdot \|X_2\| + 1, \|Y\| + 2 \} \)

We now turn for a moment to a discussion of the results obtained with respect to the problem of finding the capacity region of a d.m. MAC with feedback for codes achieving zero-error probability. In Section 3 we defined the concept of a zero-error achievable rate pair, and also of a zero-error capacity region for a d.m. MAC without feedback, in particular the zero-error capacity region \( C_0(K_{21}, III) \), as well as the regions \( C_0(K_{21}, II) \) and
$C_0(K_{21}, I)$. It was noted there that in the deterministic case a simple characterization of $C_0(K_{21}, III)$ is known, but for a general d.m. MAC $K_{21}$, the regions $C_0(K_{21}, I)$, $C_0(K_{21}, II)$, and $C_0(K_{21}, III)$ are still unknown. We are now interested in knowing what happens to these zero-error capacity regions when feedback is allowed, or phrased alternatively, what happens to the feedback capacity regions $C(K_{21}, I, CFB)$, $C(K_{21}, II, CFB)$, $C(K_{21}, III, CFB)$ when we restrict ourselves to zero-error codes. (The latter two capacity regions will be dealt with later in this section.) When complete feedback is available we denote the zero-error capacity regions in situations $(K_{21}, I)$, $(K_{21}, II)$, and $(K_{21}, III)$ by $C_0(K_{21}, I, CFB)$, $C_0(K_{21}, II, CFB)$, and $C_0(K_{21}, III, CFB)$ respectively. As in the non-feedback case, generally these regions are unknown, but some partial results are available.

Prelov [110] observed that for the deterministic BEMAC the zero-error feedback capacity region $C_0(K_{21}, I, CFB)$ is strictly larger than the ordinary capacity region $C(K_{21}, I)$ and thus, a fortiori, is larger than the zero-error non-feedback capacity region $C_0(K_{21}, I)$. He demonstrated this fact by showing that the symmetrical rate pair of Gaarder and Wolf [61] $(R_1, R_2) = (0.76018, 0.76018)$ is achievable with zero-error in the case $(K_{21}, I, CF B)$.

Furthermore, Dueck [50] established the zero-error feedback capacity region $C_0(K_{21}, II, CF B)$ for MAC's in the class $W (W = W_1 \cap W_2$, whereby $W_1$ and $W_2$ were introduced by Willems [144]). The precise formulation of this result is given below, after the situation $(K_{21}, II, CF B)$ has been formally introduced. Similarly, Prelov [110] obtained results regarding $C_0(K_{21}, III, CF B)$, which will be stated in conjunction with other results in situation $(K_{21}, III, CF B)$ dealt with below.

King [89] was the first one to investigate the d.m. MAC in situation $(K_{21}, II, CF B)$, i.e., the situation $(K_{21}, II)$ with feedback to both encoders.
This communication situation is sketched in Fig. 17. Here, three independent message sources emit statistically independent messages $m_1 \in \{1, 2, \ldots, M_1\}$, $m_2 \in \{1, 2, \ldots, M_2\}$, $m_0 \in \{1, 2, \ldots, M_0\}$ such that each triple $(m_1, m_2, m_0)$ occurs with equal probability. In addition, there is a common noiseless feedback link available from the output to each of the encoders. At time instant $t$, $1 \leq t \leq n$, an encoding function $f_{1t}$ maps the message pair $(m_1, m_0)$ and the sequence of channel outputs received up to time $t$, $y_{t-1}$, into the next channel input at terminal 1, i.e., $x_{1t} = f_{1t}(m_1, m_0, y_{t-1})$. Similarly, the encoding operation at time $t$ at terminal 2 is described by $x_{2t} = f_{2t}(m_2, m_0, y_{t-1})$. The decoder must estimate the three source messages $m_1, m_2, m_0$ based on the received sequence $y_n$. We denote this communication situation by $(K_{21}, II, CFB)$. It combines aspects of situations $(K_{21}, I, CFB)$ and $(K_{21}, II)$. The definitions of a code, error probability, achievable rate triple, and capacity region are therefore straightforward and need not to be given here. We denote the capacity region in this case by $C(K_{21}, II, CFB)$. A computable characterization of this capacity region is unknown at present, but King [89] established an achievable rate region which is stated in the following theorem. Theorem 7.4 includes Theorem 7.1 as a special case and extends the capacity region given by Theorem 4.1.

Fig. 17. The d.m. MAC $(K_{21}, II)$ with complete feedback. Situation $(K_{21}, II, CFB)$. 
Theorem 7.4 ([89]) : For a d.m. MAC in situation \((K_{21}, II, CFB)\) an achievable rate region is given by

\[
\mathcal{R}_I(K_{21}, II, CFB) = \text{co} \{(R_1, R_2, R_0) : 0 \leq R_1 \leq I(X_1; Y|X_2, U), \]
\[
0 \leq R_2 \leq I(X_2, Y|X_1, U), \]
\[
0 \leq R_0 + R_1 + R_2 \leq I(X_1, X_2; Y),
\]

for some \(P(u, x_1, x_2, y) = P(u)P(x_1|u)P(x_2|u)P(y|x_1, x_2)\).

King [89] did not state any cardinality bound on \(\mathcal{U}\), but De Bruyn and van der Meulen [39] showed that in the description of the region (7.10) it is not necessary to take the convex hull and that corresponding with this new representation the cardinality bound \(|\mathcal{U}| \leq \min \{\|X_1\| \cdot \|X_2\| + 1, \}
\|Y\| + 2\} \) suffices.

Furthermore, it is shown in [39] that for a d.m. MAC belonging to the class \(\mathcal{W}_1 \cup \mathcal{W}_2\) the region \(\mathcal{R}_I(K_{21}, II, CFB)\) given by Theorem 7.4 (or rather its improved characterization just discussed) is indeed the capacity region \(C(K_{21}, II, CFB)\). Suppose we denote for a deterministic MAC belonging to the class \(\mathcal{W}\) the capacity region \(C(K_{21}, II, CFB)\) by \(C(K_{21}, II, CFB, D, \mathcal{W})\). Then the result just mentioned gives the following characterization of this capacity region, in analogy with expression (7.4) above:

\[
C(K_{21}, II, CFB, D, \mathcal{W}) = \{(R_1, R_2, R_0) : 0 \leq R_1 \leq H(X_1|U), \]
\[
0 \leq R_2 \leq H(X_2|U), \]
\[
0 \leq R_0 + R_1 + R_2 \leq H(Y),
\]

for some \(P(u, x_1, x_2, y) = P(u)P(x_1|u)P(x_2|u)P(y|x_1, x_2)\)

such that \(|\mathcal{U}| \leq \min \{\|X_1\| \cdot \|X_2\| + 1, \|Y\| + 2\}\).
In particular, expression (7.11) gives the capacity region \( C(K_{21}, \Pi, \text{CFB}) \) of the deterministic BEMAC, but now with cardinality constraint \( \|\mathcal{U}\| \leq 5 \). The boundary line \( R_1 = R_2 \) of the region \( C(K_{21}, \Pi, \text{CFB}) \) (and also of the region \( C(K_{21}, \Pi) \)) has been fully characterized in the case of the deterministic BEMAC in [40], with a method similar to the one developed in [147].

We now describe the result by Dueck [50], which has already been mentioned briefly above. Let be given a d.m. MAC \( \mathcal{X}_1 \times \mathcal{X}_2, P(y|x_1, x_2), \mathcal{Y} \) and a finite set \( \mathcal{U} \). Let \( U, X_1, X_2 \) be RV's on the sets \( \mathcal{U}, \mathcal{X}_1, \) and \( \mathcal{X}_2 \) with common distribution \( P_{UX_1X_2} \). Define \( \mathcal{T}(U, X_1, X_2) \) to be the class of triples \( (X_1^x, X_2^x, Y^y) \) of RV's on \( \mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{Y} \) such that the quadruple \( (U, X_1^x, X_2^x, Y^y) \) satisfies \( P_{X_1^x|U} = P_{X_1|U} \), \( P_{X_2^x|U} = P_{X_2|U} \), and \( P_{Y^y|X_1^x, X_2^x}(y|x_1, x_2) = 0 \) whenever \( P(y|x_1, x_2) = 0 \), where \( x_1 \in \mathcal{X}_1, x_2 \in \mathcal{X}_2, y \in \mathcal{Y} \). Then Dueck [50] proved the following theorem.

**Theorem 7.5** ([50]): Let be given a d.m. MAC belonging to class \( \mathcal{U} \). Then if \( C_0(K_{21}, \Pi, \text{CFB}) \neq \{(0, 0, 0)\} \), the zero-error feedback capacity region in situation \((K_{21}, \Pi, \text{CFB})\) is given by

\[
C_0(K_{21}, \Pi, \text{CFB}) = \{(R_1, R_2, R_0) : 0 \leq R_1 \leq H(X_1|U), \quad \text{(7.12a)} \\
0 \leq R_2 \leq H(X_2|U), \quad \text{(7.12b)} \\
0 \leq R_0 \leq I(U; Y^y) - H(X_1^x, X_2^x|U, Y^y), \quad \text{(7.12c)}
\]

for every triple \( (X_1^x, X_2^x, Y^y) \in \mathcal{T}(U, X_1, X_2) \), for some triple \( (U, X_1, X_2) \) such that \( \|\mathcal{U}\| \leq \|\mathcal{X}_1\| \cdot \|\mathcal{X}_2\| + 2 \).

We next turn to the case of the d.m. MAC in situation \((K_{21}, \Pi, \text{CFB})\), i.e., the situation \((K_{21}, \Pi)\) (treated in Subsection 4d) with feedback to
both encoders. This communication situation is shown in Fig. 18. This case was investigated independently by Prelov [110],[111] and De Bruyn and van der Meulen [39]. In these papers the result is derived that complete feedback does not increase the capacity region of the d.m. AMAC $K_{21}$, i.e., that $C(K_{21}, III, CFB) = C(K_{21}, III)$. The latter capacity region is given by Theorem 4.2 and thus need not to be recalled here. However, Prelov [110],[111] proved an additional result, namely that for a deterministic MAC $(K_{21}, III)$ feedback does not increase the zero-error capacity region either. Let us denote the zero-error capacity region and the ordinary (average error) capacity region of a deterministic d.m. MAC in situation $(K_{21}, III, CFB)$ by $C_0(K_{21}, III, CFB, D)$ and $C(K_{21}, III, CFB, D)$ respectively. Then Prelov [110],[111] showed that $C_0(K_{21}, III, CFB, D) = C(K_{21}, III, CFB, D)$. This result, combined with the previous one and with (4.5), yields the following chain of equalities:

$$C_0(K_{21}, III, CFB, D) = C(K_{21}, III, CFB, D) = C(K_{21}, III, D)$$

$$= C_0(K_{21}, III, D),$$

(7.13)

which is Theorem 2 of [111]. Prelov [110],[111] gave also the generalization of the above results to the case of a d.m. MAC $K_{t1}$ ($t \geq 3$) with complete feedback when the transmission configuration exhibits a special source hierarchy, viz. the one described in Subsection 4g.

We conclude this section by mentioning a result by De Bruyn, Prelov, and van der Meulen [37], who considered the problem of reliable transmission of an arbitrarily correlated source pair over a d.m. AMAC $K_{21}$ in the presence of complete feedback. This communication situation is shown in Fig. 19.

In analogy with the notation developed in Section 5 we denote this situation by $(K_{21}, (U,V), III, CFB)$. De Bruyn, Prelov, and van der Meulen [37] showed that the necessary and sufficient conditions for reliable transmission given
in Theorem 5.4 for the case \((K_{21}, (U, V), III)\) continue to hold for the situation \((K_{21}, (U, V), III, CFB)\).

![Diagram](image1)

**Fig. 18.** The d.m. MAC \((K_{21}, III)\) with complete feedback. Situation \((K_{21}, III, CFB)\).

![Diagram](image2)

**Fig. 19.** The d.m. AMAC \(K_{21}\) with arbitrarily correlated sources and complete feedback. Situation \((K_{21}, (U, V), III, CFB)\).
7.3. The Gaussian MAC with complete feedback

The model for a discrete-time additive white Gaussian noise MAC with two input users and complete feedback is shown in Fig. 20. The AWGN MAC under consideration is of the type (AWGN $K_{21},N,P_1,P_2,I$) defined in Section 6, i.e., it has two input users, input power constraints $P_1$ and $P_2$, and independent additive noise with variance $N$, whereas two independent messages are to be transmitted, one separately by each of the two channel encoders. In addition, there is a common noiseless feedback link available from the output to each of the encoders as in Fig. 15. This implies that the channel input at time instant $t$, $1 \leq t \leq n$, at terminal $l$ is a function of the message $m_l$ and the sequence of channel outputs received up to time $t$, i.e., $x_{1t} = f_{1t}(m_l, Y^{t-1}_l)$. Similarly, for the encoding operation at terminal 2. We denote this communication situation by (AWGN $K_{21},N,P_1,P_2,I,CFB$). We will not bother to state the definitions of a code, error probability, achievable rate pair, and capacity region for this case, as those can be easily derived by combining the cases ($K_{21},I,CFB$) and (AWGN $K_{21},N,P_1,P_2,I$). Carleial [16] and Cover and Leung [30] independently established an achievable rate region for this channel using a feedback superposition scheme. We start by stating their result as a separate theorem, which can be regarded as the Gaussian analogue of Theorem 7.1.

![Diagram of discrete-time AWGN MAC](image-url)

Fig. 20. Discrete-time AWGN MAC (AWGN $K_{21},N,P_1,P_2,I$) with complete feedback, yielding situation (AWGN $K_{21},N,P_1,P_2,I,CFB$).
Theorem 7.6 ([16],[30]): For a discrete-time AWGN MAC in situation 
(AWGN K₂₁, N, P₁, P₂, I, CFB) an achievable rate region is given by

\[ R = \left\{ (R₁, R₂) : 0 \leq R₁ \leq \frac{1}{2} \ln \left( 1 + \frac{\alpha₁ P₁}{N} \right), \right. \]
\[ 0 \leq R₂ \leq \frac{1}{2} \ln \left( 1 + \frac{\alpha₂ P₂}{N} \right) \]
\[ R₁ + R₂ \leq \frac{1}{2} \ln \left( 1 + \frac{P₁ + P₂ + 2 \sqrt{\alpha₁ \alpha₂ P₁ P₂}}{N} \right), \]

for \( 0 \leq \alpha₁ \leq 1, \alpha₁ = 1 - \alpha₂, i = 1, 2. \} \tag{7.14a}

Although it is known by now that the achievable rate region given by 
(7.14) is not optimal (this was demonstrated by Ozarow [101],[102] whose 
result is to be discussed below), at the time that Theorem 7.6 was discovered 
(around mid-1976, cf. [133]) the region given by (7.14) (here denoted by 
\( R₁ \) for short) turned out to be quite an improvement over the non-feedback 
capacity region formulated in Theorem 6.1. Properties of region \( R₁ \), together 
with several examples, are discussed in [16],[30], and [133]. Theorem 7.6 has 
been superseded by Theorem 7.8 below, due to Ozarow [101],[102]. However, the 
region \( R₁ \) given by Theorem 7.6 remains of independent interest, because of 
its achievability in the case of partial feedback (see Section 8). Before 
proceeding to describe Ozarow's exact result, we first state the outer bound 
region which he obtained and showed to be the Gaussian counterpart of Theorem 
7.2 above.

Theorem 7.7 ([101],[102]): An outer bound on the capacity region of a 
discrete-time AWGN MAC in situation (AWGN K₂₁, N, P₁, P₂, I, CFB) is given by
\( R_{OB}(\text{AWGN } K_{21}, N, P_1, P_2, I, \text{CFB}) \)

\[
= \bigcup_{0 \leq \rho \leq 1} \left\{ (R_1, R_2) : 0 \leq R_1 \leq \frac{1}{2} \ln \left(1 + \frac{P_1(1-\rho^2)}{N}\right), \quad 0 \leq R_2 \leq \frac{1}{2} \ln \left(1 + \frac{P_2(1-\rho^2)}{N}\right), \quad R_1 + R_2 \leq \frac{1}{2} \ln \left(1 + \frac{P_1 + P_2 + 2\rho \sqrt{P_1 P_2}}{N}\right) \right\}. \quad (7.15) 
\]

Ozarow [101,102] established a significant result by determining the actual capacity region of the AWGN MAC in situation (AWGN \( K_{21}, N, P_1, P_2, I, \text{CFB} \)). In particular he showed that this capacity region, which we denote by \( C(\text{AWGN } K_{21}, N, P_1, P_2, I, \text{CFB}) \), equals the outer bound region given by Theorem 7.7. This result is striking, because in the discrete memoryless case the capacity region \( C(K_{21}, I, \text{CFB}) \) is not known in general; the latter capacity region is only known for MAC's belonging to class \( \mathcal{K} \), for which Hekstra and Willems [74] showed that the inner bound region given by Theorem 7.1 is the capacity region, thereby extending the class \( \mathcal{W}_1 \cup \mathcal{W}_2 \) for which Willems [144] originally found this property to be true. We now proceed to state Ozarow's main result.

**Theorem 7.8 ([101] [102]):** The capacity region of the AWGN MAC in situation (AWGN \( K_{21}, N, P_1, P_2, I, \text{CFB} \)) is given by

\[
C(\text{AWGN } K_{21}, N, P_1, P_2, I, \text{CFB}) = R_{OB}(\text{AWGN } K_{21}, N, P_1, P_2, I, \text{CFB}), \quad (7.16)
\]

where the region \( R_{OB}(\text{AWGN } K_{21}, N, P_1, P_2, I, \text{CFB}) \) is characterized by (7.15).

The parameter \( \rho \) in expression (7.15) corresponds to the correlation coefficient of a bivariate normal probability assignment on the input RV's \( X_1 \) and \( X_2 \), which have zero means and variances bounded by \( N_1 \) and \( N_2 \), respectively. Observe that, when \( \rho = 0 \), the rate constraints (7.15) yield the capacity region of the AWGN MAC without feedback, i.e., the region.
\( C(AWGN K_{21}, N, P_1, P_2, I) \) described in Theorem 6.1.

In proving the forward part of Theorem 7.8, Ozarow \[101],[102]\ used a constructive coding scheme which is clever adaptation towards the case of an AWGN MAC with complete feedback of the well-known constructive coding scheme developed by Schalkwijk and Kailath \[117]\ and Schalkwijk \[113]\ for the single-input AWGN channel with noiseless feedback. In particular, Ozarow \[101],[102]\ showed that, using such a constructive coding scheme, a certain rate point \((R_1^*, R_2^*)\) can be achieved for the AWGN MAC in situation \((AWGN K_{21}, N, P_1, P_2, I, CFB)\) which is defined as follows:

\[
R_1^* = \frac{1}{2} \ln \left( \frac{N + P_1 + P_2 + 2\rho_0 \sqrt{P_1 P_2}}{N + P_2(1 - \rho_0^2)} \right), \tag{7.17a}
\]

\[
R_2^* = \frac{1}{2} \ln \left( \frac{N + P_1 + P_2 + 2\rho_0 \sqrt{P_1 P_2}}{N + P_1(1 - \rho_0^2)} \right), \tag{7.17b}
\]

where \(\rho_0 \in (0, 1)\) is the unique solution of the quartic equation

\[
P(\rho) = 0, \text{ with}
\]

\[
P(\rho) = N(N + P_1 + P_2 + 2\rho \sqrt{P_1 P_2}) - (N + P_1(1 - \rho^2))(N + P_2(1 - \rho^2)). \tag{7.17c}
\]

When one solves \(7.17c\) for \(\rho_0\), the rate point \((R_1^*, R_2^*)\) takes the form

\[
R_1^* = \frac{1}{2} \ln \left( 1 + \frac{P_1(1 - \rho_0^2)}{N} \right), \tag{7.18a}
\]

\[
R_2^* = \frac{1}{2} \ln \left( 1 + \frac{P_2(1 - \rho_0^2)}{N} \right). \tag{7.18b}
\]

Ozarow \[101],[102]\ showed that the rate point \((R_1^*, R_2^*)\) lies on the boundary of the outer bound region \(R_{OB}(AWGN K_{21}, N, P_1, P_2, I, CFB)\) defined in Theorem 7.7, and in fact is that point on the boundary which maximizes \(R_1 + R_2\).
After having established the achievability of the pair \((R_1^*, R_2^*)\), Ozarow [101],[102] extends this result by describing a superposition approach which allows communication at all points along a curve (i) between \((R_1^*, R_2^*)\) and a point \(C = (R_{11}, R_{21})\), and (ii) between \((R_1^*, R_2^*)\) and a point \(D = (R_{12}, R_{22})\), where \(C\) and \(D\) are extreme points of the capacity region 
\[ C(\text{AWGN } K_{21}^N, P_1, P_2, I), \]
and are given by
\[
\begin{align*}
R_{11} &= \frac{1}{2} \ln \left( 1 + \frac{P_1}{N} \right), \\
R_{21} &= \frac{1}{2} \ln \left( 1 + \frac{P_2}{N + P_1} \right), \\
R_{12} &= \frac{1}{2} \ln \left( 1 + \frac{P_1}{N + P_2} \right), \\
R_{22} &= \frac{1}{2} \ln \left( 1 + \frac{P_2}{N} \right).
\end{align*}
\]
(7.19)
(7.20)

Ozarow [101],[102] then shows that the resulting achievable rate region equals the outer bound region \(R_{0B}\), thus completing the proof of Theorem 7.8. The result of Theorem 7.8 is illustrated by the following example due to Ozarow [101]. In Fig. 21 the capacity region 
\[ C(\text{AWGN } K_{21}^N, P_1, P_2, I, \text{CFB}) \]
is shown for the particular choices \(P_1 = P_2 = 10N\), together with the capacity region 
\[ C(\text{AWGN } K_{21}^N, P_1, P_2, I) \]
and the achievable rate region \(R_1^* (\text{AWGN } K_{21}^N, N, P_1, P_2, I, \text{CFB})\). Also shown in Fig. 21 is the symmetrical rate point \((R_1, R_2) = (0.8643, 0.8643)\) which is optimally achievable using Theorem 7.6, and the rate point \((R_1, R_2) = (0.8905, 0.8905)\) which is achievable with Ozarow's constructive coding scheme.

Ozarow [101] extended his results to the case of an AWGN MAC in situation 
\( (\text{AWGN } K_{21}^N, P_1, P_2, \text{II, CFB}) \). This is the Gaussian analogue of situation 
\( (K_{21}^N, \text{II, CFB}) \) considered by King [89], and involves the Slepian and Wolf setup of three independent sources and two channel inputs. Ozarow [101] determined also the capacity region in this case. This capacity region turns out to correspond to the Slepian and Wolf [125] region defined in Theorem 4.1, but now with: (i) joint probability assignments of the form \( P(u, x_1, x_2, y) = \)
\( P(u)P(x_1, x_2 | u)P(y | x_1, x_2) \), (ii) discrete probabilities replaced by probability densities, and (iii) input power constraints at the encoder. Ozarow [101] proved that this modified Slepian and Wolf region is in general an outer bound to the capacity region of a d.m. MAC in situation \((K_{21}, II, CFB)\), and that it is achievable for discrete-time AWGN MAC's in situation \((AWGN K_{12}, N, P_1, P_2, II, CFB)\). No explicit expression for the capacity region \( C(AWGN K_{12}, N, P_1, P_2, II, CFB) \) - similar to expression (7.15) for the region \( C(AWGN K_{12}, N, P_1, P_2, I, CFB) \) - is given in [101], though.

Prelov [110], [111] investigated the AWGN MAC in situation \((AWGN K_{21}, N, P_1, P_2, III, CFB)\), which is the Gaussian analogue of situation \((K_{21}, III, CFB)\) pictured in Fig. 18. This channel may also be called the asymmetric AWGN MAC with complete feedback.

\[ \begin{align*}
R_1 & = (0.8905, 0.8843) \\
R_2 & = (0.8905, 0.8843)
\end{align*} \]

Fig. 21. Capacity region \( C(AWGN K_{21}, N, P_1, P_2, I, CFB) \) of an AWGN MAC with complete feedback, for \( P_1 = P_2 = 10N \), shown together with previously established achievable rate regions.
Prelov [110],[111] showed that in this situation the same fact is true as in the discrete memoryless case, viz. that feedback does not increase the capacity region of an asymmetric AWGN MAC $K_{21}$. More formally, this result can be stated as follows: $C(AWGN K_{21}, N, P_1, P_2, III, CFB) = C(AWGN K_{21}, N, P_1, P_2, III, CFB)$.

Prelov [110],[111] gave also the generalization of this result to the case of an AWGN MAC $K_{t1}$ ($t \geq 3$) with a degraded message set and complete feedback. Formally stated, this result says that $C(AWGN K_{21}, N, P_1, P_2, ..., P_t, III, CFB) = C(AWGN K_{21}, N, P_1, P_2, ..., P_t, III)$. Moreover, he gave an explicit expression for this capacity region for $t \geq 2$.

Başar and Başar [12] investigated the continuous-time white Gaussian noise MAC with complete, noiseless feedback, but from a different point of view than the approach taken sofar. Rather than the derivation of a capacity region, their objective was to obtain real-time implementable "optimal" coding schemes for a class of Gaussian continuous-time MAC's with complete feedback. The communication system considered in [12] involves two independent messages which are realizations of Gaussian RV's, and a white Gaussian noise MAC which is to be used for a period of $T$ seconds. Within this framework, two extreme cases are considered: (i) no cooperation is allowed between the two encoders and their outputs are forced to be statistically independent, and (ii) the encoders are assumed to fully cooperate in the transmission of the messages. Başar and Başar [12] introduced the concept of an "admissible distortion-tradeoff curve" as an alternative to the capacity region, and derived the unique tradeoff curve for each of the two extreme versions of Gaussian continuous-time MAC's with feedback under consideration. They furthermore declared any coding scheme which yields a pair of distortion levels that lies on an admissible distortion-tradeoff curve as an acceptable coding scheme and called it therefore "admissible" or "optimal". For each of the two extreme versions under consideration, Başar and Başar [12] developed coding
schemes which are optimal and real-time implementable. Moreover, they showed that the corresponding encoding structures are linear and exponentially growing, while the decoder in each case has the structure of a Kalman filter.

8. THE MULTIPLE-ACCESS CHANNEL WITH PARTIAL FEEDBACK

8.1. The discrete memoryless MAC with partial feedback

The notion of partial feedback (PFB) for multi-user channels was introduced by Dueck [46] who originally used the term "semi-feedback". (Carleial [17] speaks of "one-sided feedback".) A d.m. MAC $K_{21}$ is said to have partial feedback if there is a noiseless feedback link from the output to only one of the encoders, whereas the other encoder has no information whatsoever about the received signal. A d.m. MAC with partial feedback to encoder 2 in situation $(K_{21}, I)$ is depicted in Fig. 22. We denote this communication situation by $(K_{21}, I, \text{PFB}_2)$. (If the partial feedback link was to be to encoder 1, we would have situation $(K_{21}, I, \text{PFB}_1)$.) Thus, in situation $(K_{21}, I, \text{PFB}_2)$ there are two message sources which produce statistically independent messages $m_1 \in \{1, \ldots, M_1\}$ and $m_2 \in \{1, \ldots, M_2\}$, each pair $(m_1, m_2)$ being equally likely. Encoder 1, the non-feedback encoder, is described by an encoding function $f_1$ that maps $m_1$ into a codeword $x_{1t} = f_1(m_1)$. Encoder 2, the feedback encoder, is described by a set of n encoding functions $f_{2t}$, $1 \leq t \leq n$. At time instant $t$, $1 \leq t \leq n$, the encoding function $f_{2t}$ maps the message $m_2$ and the channel sequence received up to time $t$, $y_{t-1}^t$, into the next channel input at terminal 2, i.e., $x_{2t} = f_{2t}(m_2, y_{t-1}^t)$. The decoder gives estimates of $m_1$ and $m_2$ based on the received sequence $y_n$. The definitions of a code, error probability, achievable rate pair, and capacity region are similar to the ones in situations $(K_{21}, I)$ and $(K_{21}, I, \text{CFB})$, and can be viewed as a combination of those. We denote the capacity region in situation $(K_{21}, I, \text{PFB}_i)$ by $C(K_{21}, I, \text{PFB}_i)$, $i = 1, 2$. 
Dueck [46] showed the interesting fact that the rate point \( (R_1, R_2) = (0.76018, 0.76018) \), found by Gaarder and Wolf [61] for the deterministic BEMAC (using a feedback scheme based on complete feedback), can also be achieved using only partial feedback. Dueck's result indicated that in general not only complete feedback but also partial feedback can increase the capacity region \( C(K_{21}, I) \) of a d.m. MAC. The natural question then arose whether Dueck's result could perhaps be generalized, and whether it could be shown that, with partial feedback only, one could achieve the same rate region as Cover and Leung [30] had shown to be achievable with complete feedback. Carleial [17], Csiszár and Körner [32], and Willems and van der Meulen [148] investigated this problem independently and showed that for a d.m. MAC \( K_{21} \) the rate region given by Theorem 7.1 is indeed achievable just with partial feedback. The various formulations of this result differ slightly among each other in [17], [32] and [148]. We give the formulation as stated and proved by Willems and van der Meulen [148].

**Theorem 8.1** ([148]): For a d.m. MAC either in situation \( (K_{21}, I, \text{PFB}_1) \) or in situation \( (K_{21}, I, \text{PFB}_2) \) an achievable rate region is given by
\[ R_I(K_{21}, I, PFB_1) = \{ (R_1, R_2) : 0 \leq R_1 \leq I(X_1; Y | X_2, U), \]

\[ 0 \leq R_2 \leq I(X_2; Y | X_1, U), \]

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

for some \( P(u, x_1, x_2, y) = P(u)P(x_1 | u)P(x_2 | u)P(y | x_1, x_2) \)

such that \( \| U \| < \min \{ \| X_1 \| \cdot \| X_2 \| + 1, \| Y \| + 2 \} \), \( i = 1, 2 \).

The scheme used in [148] to prove this result is based on the novel ideas of nonrandom partitions and restricted decoding, thus avoiding complicated list coding techniques. Carleial's formulation [17] of the above result appears in a slightly different form (and without cardinality constraints), as a special case of an achievable rate region derived by him for the case of generalized feedback, which is dealt with in Section 9. Csiszár and Körner [32] stated the above result (Theorem 8.1) also, but without going into any convex hull or cardinality considerations. They outlined a proof which involves list codes. Notice that the formulation of the region in Theorem 8.1 reflects the strengthened representation found by Willems [145] of the achievable rate region given in Theorem 7.1, viz. without taking the convex hull, and with other cardinality constraints than the ones found by Cover and Leung [30]. This strengthened representation occurred already once before in this text, namely in expression (7.9). Notice also that interchanging the roles of the two encoders with respect to the feedback reception does not affect the achievable rate region, i.e., \( R_I(K_{21}, I, PFB_1) = R_I(K_{21}, I, PFB_2) \).

As pointed out in Section 7, Willems [144] found that for MAC's belonging to class \( \mathcal{W}_1 \) (characterized by the property that \( x_1 = f(y, x_2) \)) the capacity region in the event of complete feedback is given by the region formulated in Theorem 7.3. It now follows from Theorem 8.1 that the same region is also the capacity region in the case of partial feedback to encoder 2.
More precisely, denoting the capacity region of a d.m. MAC belonging to class \( \mathcal{W}_1 \) in situation \((K_{21}, I, PFB_2)\) by \( C(K_{21}, I, PFB_2, \mathcal{W}_1) \), we have by virtue of Theorems 7.3 and 8.1 that \( C(K_{21}, I, PFB_2, \mathcal{W}_1) = C(K_{21}, I, CFB, \mathcal{W}_1) \), where the latter region is characterized in Theorem 7.3. By the same argument it follows that for channels belonging to class \( \mathcal{W}_1 \) the capacity region in the case of partial feedback to encoder 2 is also equal to the complete feedback capacity region, so that \( C(K_{21}, I, PFB_1, \mathcal{W}_1) = C(K_{21}, I, CFB, \mathcal{W}_1), \quad i = 1, 2. \)

By the same token it follows that for a deterministic MAC belonging to class \( \mathcal{W} \) the capacity region in situation \((K_{21}, I, PFB_1)\), denoted by \( C(K_{21}, I, PFB_1, \mathcal{D}, \mathcal{W}) \), \( i = 1, 2 \), equals the capacity region \( C(K_{21}, I, CFB, \mathcal{D}, \mathcal{W}) \) characterized by (7.4). This holds in particular for the deterministic BEMAC. Thus, the capacity region \( C(K_{21}, I, CFB) \) of the deterministic BEMAC, shown in Fig. 16, is also the capacity region \( C(K_{21}, I, PFB_1) \) of this MAC for either choice of the index \( i \). A similar conclusion can be drawn for the partial feedback capacity region of MAC's belonging to the class \( \mathcal{K} \) considered by Hekstra and Willms [74], [75]. Denoting the capacity region of a d.m. MAC belonging to class \( \mathcal{K} \) in situation \((K_{21}, I, PFB_1)\) by \( C(K_{21}, I, PFB_1, \mathcal{K}), \quad i = 1, 2. \), it follows from Theorem 7.3 and 8.1 that \( C(K_{21}, I, PFB_1, \mathcal{K}) = C(K_{21}, I, CFB, \mathcal{K}), \) where the latter region is characterized by expression (7.9).

Vinck [137] gave an example of a simple coding scheme for the deterministic BEMAC with partial feedback, that achieves the asymmetrical rate point \((R_1, R_2) = (0.986, 0.532)\) with sum rate 1.518 bits per transmission. Vinck's method is based on an application of Schalkwijk's unit square method [116] to a d.m. MAC with partial feedback. The point achieved with this method falls outside the capacity region \( C(K_{21}, I) \).

Using a proof technique similar to the one developed in [148], De Bruyn and van der Meulen [39] showed that for a d.m. MAC in the Slepian and Wolf situation \((K_{21}, II)\), one can achieve with partial feedback to one encoder the
same rate region as King [39] showed to be achievable in situation \((K_{21}, II, CFB)\). Let us denote the communication situation of a d.m. MAC in situation \((K_{21}, II)\) with partial feedback to encoder \(i\) by \((K_{21}, II, PFB_i)\), \(i = 1, 2\), and the corresponding capacity region by \(\mathcal{C}(K_{21}, II, PFB_i)\). The result obtained in [39] is formulated in Theorem 8.2 (cf. also Theorem 7.4 and the discussion following it).

**Theorem 8.2 ([39])**: For a d.m. MAC either in situation \((K_{21}, II, PFB_1)\) or in situation \((K_{21}, II, PFB_2)\) an achievable rate region is given by

\[
\mathcal{K}_{1}(K_{21}, II, PFB_1) = \{(R_1, R_2, R_0) : 0 \leq R_1 \leq I(X_1; Y|X_2, U), \quad (8.2a)
\]

\[
0 \leq R_2 \leq I(X_2; Y|X_1, U), \quad (8.2b)
\]

\[
0 \leq R_0 + R_1 + R_2 \leq I(X_1, X_2; Y), \quad (8.2c)
\]

for some \(P(u,x_1,x_2,y) = P(u)P(x_1|u)P(x_2|u)P(y|x_1,x_2)\)

such that \(\|\mathcal{W}\| \leq \min \{\|\mathcal{I}_1\| \cdot \|\mathcal{I}_2\| + 1, \|\mathcal{Y}\| + 2\}\), \(i = 1, 2\).

As mentioned in Section 7, the region given by expression (8.2) is the capacity region for MAC's belonging to Willems class \(\mathcal{W}_1 \cup \mathcal{W}_2\) in the event of complete feedback. It now follows from Theorem 8.2 that the same region is also the capacity region of a MAC belonging to \(\mathcal{W}_1 \cup \mathcal{W}_2\) in the case of partial feedback. In notation : \(\mathcal{C}(K_{21}, II, PFB_1, \mathcal{W}_1 \cup \mathcal{W}_2) = \mathcal{K}_{1}(K_{21}, II, PFB_1, \mathcal{W}_1 \cup \mathcal{W}_2) = \mathcal{C}(K_{21}, II, CFB, \mathcal{W}_1 \cup \mathcal{W}_2), i = 1, 2\). The same statement is true for MAC's belonging to class \(\mathcal{K}_1\).

**8.2 The Gaussian MAC with partial feedback**

A discrete-time AWGN MAC with two input users and partial feedback to encoder 2 is depicted in Fig. 23. The channel under consideration is similar to the one discussed in Section 7.3, except that now there is only one noise-
Fig. 23. Discrete-time AWGN MAC (AWGN $K_{21}, N, P_1, P_2, I$) with partial feedback to encoder 2, yielding situation (AWGN $K_{21}, N, P_1, P_2, I, PFB_2$).

less feedback link available, viz. to encoder 2. As before, the variance of the noise is $N$, and the transmitter powers for encoder 1 and encoder 2 are $P_1$ and $P_2$ respectively. We denote this communication situation by (AWGN $K_{21}, N, P_1, P_2, I, PFB_2$). Similarly, when there is a noiseless feedback link available from the output to encoder 1 only, we are in situation (AWGN $K_{21}, N, P_1, P_2, I, PFB_1$).

Carleial [17] has shown that for an AWGN MAC in situation (AWGN $K_{21}, N, P_1, P_2, I, PFB_i$), $i = 1, 2$, the same rate region can be achieved as Carleial [16] and Cover and Leung [30] established independently for a discrete-time AWGN MAC in situation (AWGN $K_{21}, N, P_1, P_2, I, CFB$). This is the rate region, which is characterized in Theorem 7.6, and which is, for the particular example discussed in Section 7.3 ($P_1 = P_2 = 10N$), depicted by the dotted line in Fig. 21. We state this result as a separate Theorem.

Theorem 8.3 ([17]): For a discrete-time AWGN MAC, either in situation (AWGN $K_{21}, N, P_1, P_2, I, PFB_1$) or in situation (AWGN $K_{21}, N, P_1, P_2, I, PFB_2$), an achievable rate region is given by
\[ \mathcal{R}_i(\text{AWGN } K_{21}, N, P_1, P_2, I, \text{PFB}_1) = \mathcal{R}_i(\text{AWGN } K_{21}, N, P_1, P_2, I, \text{CFB}), \]  
\[ i = 1, 2, \] where the region \( \mathcal{R}_i(\text{AWGN } K_{21}, N, P_1, P_2, I, \text{CFB}) \) is characterized in Theorem 7.6.

It is unknown at present whether or not the region \( \mathcal{R}_i(\text{AWGN } K_{21}, N, P_1, P_2, I, \text{PFB}_1) \) is also the capacity region in situation \( \text{(AWGN } K_{21}, N, P_1, P_2, I, \text{PFB}_1) \), \( i = 1, 2. \)

Willems, van der Meulen, and Schalkwijk [150] investigated the implications of the Schalkwijk [113] scheme for discrete-time AWGN MAC's in the case of partial feedback. In particular they found that, using a semi-constructive and partial feedback scheme, one can in situation \( \text{(AWGN } K_{21}, N, P_1, P_2, I, \text{PFB}_1) \), \( i = 1, 2, \) in general improve upon the classical capacity region for the AWGN MAC without feedback, thus upon the region \( \mathcal{C}(\text{AWGN } K_{21}, N, P_1, P_2, I) \) formulated in Theorem 6.1. The achievable rate region found in [150] is generated by a set of curves, each curve corresponding to the choice of a certain parameter \( \alpha, 0 \leq \alpha \leq 1. \) For the particular example of the AWGN MAC discussed in Section 7.3 \( (P_1 = P_2 = 10N) \), this achievable rate region is shown in Fig. 24 for the case of partial feedback to encoder 2. This semi-constructively obtained rate region is seen to be strictly larger than the non-feedback capacity region \( \mathcal{C}(\text{AWGN } K_{21}, N, P_1, P_2, I) \), but is dominated by the achievable rate region \( \mathcal{R}_i(\text{AWGN } K_{21}, N, P_1, P_2, I, \text{PFB}_2) \), derived by Carleial [17] using non-constructive methods. For comparison, one particular point is shown in Fig. 24 which lies on the boundary of the latter rate region, viz. the rate point \( (R_1, R_2) = (0.8643, 0.8643) \).
Fig. 24. Achievable rate region for a discrete-time AWGN MAC in situation 
(AWGN $K_{21}, N, P_1, P_2, I, PFB_2$), with $P_1 = P_2 = 10N$, using a semi-constructive coding scheme.

9. THE MULTIPLE-ACCESS CHANNEL WITH GENERALIZED FEEDBACK

9.1. The discrete memoryless MAC with generalized feedback

The notion of a MAC with generalized feedback was introduced by King [89].
A typical example of a d.m. MAC with generalized feedback is shown in Fig. 25.
It consists of a d.m. MAC ($K_{21}, I$) with a common generalized feedback signal
$Y^M$ to both encoders. This generalized feedback signal emanates directly from
the MAC and is not necessarily equal to nor a deterministic function of the
output signal $Y$. The feedback signal $Y^M$ need not to be a degraded version of
the receiver's signal either. The receiver is incapable of affecting the feedback signal. All that can be said about $Y^M$ is that it is correlated with $Y$.  

Fig. 25. The d.m. MAC \((K_{21}, I)\) with common generalized feedback signal.

Also, in this model considered by King [89], the feedback signals observed by the two encoders are assumed to be the same.

A d.m. MAC \(K_{21}\) with two independent sources and a common generalized feedback signal, as shown in Fig. 25, will be denoted here by \((K_{21}, I, CGFB)\). King [89] established an achievable rate region for this channel. The region found by King will not be presented here, partly because it is notationally rather complex to describe, partly because it has been superseded by later results. (Carleial [17] generalized King's region, and Willems, van der Meulen, and Schalkwijk [151] improved it.) King [89] did show, though, that the achievable rate region found by him for the d.m. MAC in situation \((K_{21}, I, CGFB)\) includes the achievable rate region \(R_1(K_{21}, I, CFB)\) found by Cover and Leung [30] for the d.m. MAC \(K_{21}\) with complete feedback (cf. Theorem 7.1). King [89] also considered the special case of a MAC with generalized feedback, whereby the receiver's signal is a degraded version of the feedback signal (also called the "reversely degraded" situation), and derived an achievable rate region for it. (We say that \(Y\) is a degraded version of \(Y^\prime\) with respect to the channel inputs, if \((X_1, X_2), Y^\prime,\) and \(Y\) form a Markov chain in this order, which is denoted by \((X_1, X_2) \rightarrow Y^\prime \rightarrow Y\).)
Moreover, several interesting examples of d.m. MAC's with common generalized feedback are presented and worked out in [89].

Carleial [17] considered the case of a d.m. MAC \( K_{21} \) whereby each sender observes a different generalized feedback signal, thus extending King's model [89]. This situation is depicted in Fig. 26. Here, two different feedback signals, denoted by \( Y_1 \) and \( Y_2 \), emanate directly from the channel, signal \( Y_1 \) being connected to encoder 1, and signal \( Y_2 \) to encoder 2. The two different feedback signals are not necessarily equal to, but only correlated with the decoder's output \( Y \). We denote a d.m. MAC \( K_{21} \) with two independent sources and two different generalized feedback signals, as shown in Fig. 26, by \( (K_{21},I,DGFB) \). Carleial [17] demonstrated an achievable rate region for this channel, using a random coding scheme involving signal superposition. Because of its notational complexity, the region derived by Carleial [17] will not be reproduced here. Its description involves 19 rate constraints, six auxiliary RV's and one mixing RV. Carleial [17] considered two special cases: (i) a d.m. MAC \( K_{21} \) with \( Y \) a degraded version of \( Y_1 \) or \( Y_2 \) or both (the latter case now being called the reversely degraded situation), and (ii) a d.m. MAC \( K_{21} \) with partial generalized feedback. For both situations he derived an achievable rate region by specializing the general rate region found by him for situation \( (K_{21},I,DGFB) \) to these cases. The achievable rate region obtained by Carleial [17] for the d.m. MAC \( K_{21} \) with partial feedback is essentially equivalent to the regions derived by Csiszár and Körner [32] and Willems and van der Meulen [148], apart from the differences in formulation among these regions and the absence of cardinality constraints in [17] and [32]. The general rate region found by Carleial [17] includes as special cases the achievable rate region established by Cover and Leung [30] for the d.m. MAC in situation \( (K_{21},I,CFB) \) and characterized in Theorem 7.1, and also the rate region derived by King [89] for the reversely degraded d.m. MAC \( K_{21} \) with common generalized feedback.
Fig. 26. The d.m. MAC \((K_{21}, I)\) with different generalized feedback signals.

Willems, van der Meulen, and Schalkwijk [151] (see also Willems [145] for a first formulation of these results and complete proofs) recently established another achievable rate region for the d.m. MAC in situation \((K_{21}, I, DGBF)\), which improves on the regions found by King [89] and Carleial [17]. The achievability proof given in [145], [151] involves, besides a three level block Markov superposition scheme, the novel technique of backward decoding. This latter technique was originally developed by Willems in his doctoral dissertation [145] for the case of a d.m. MAC with cribbing encoders (see Section 10), where it proved to be quite successful in establishing optimal rate regions. A direct comparison of the generalized feedback rate region derived in [145], [151] with the corresponding rate region obtained by Carleial [17] is rather difficult to carry out because of the many different rate constraints involved. However, in [145], [151] it is shown that the achievable rate region given there improves on the region obtained by Carleial [17], at least for the case where the generalized feedback situation specializes to the cribbing encoders situations b or e treated in the next section. The achievable rate region derived in [145], [151] is also strictly larger than the region found by King [89] for the case where the generalized feedback specializes to common generalized feedback and King's hypothetical output variable \(\hat{y}\) is suppressed.
Formally, a d.m. MAC $K_{2,1}$ with different generalized feedback outputs $Y_1$ and $Y_2$ is denoted by $\{\mathcal{X}_1 \times \mathcal{X}_2, P(y_1,y_2|x_1,x_2), Y \times Y_1 \times Y_2\}$, where $\mathcal{X}_1$ and $\mathcal{X}_2$ are the input alphabets, $Y$ is the output alphabet, $Y_1$ and $Y_2$ are the (finite) feedback output alphabets, and $\{P(y,y_1,y_2|x_1,x_2)\}$ is the transition probability matrix. The transmission probabilities are defined by

$$p^n(y^n, y_1^n, y_2^n, x_1^n, x_2^n) = \prod_{i=1}^{n} p(y_i, y_{1i}, y_{2i}|x_{1i}, x_{2i})$$

(9.1)

for all $x_t^n = (x_{t1}, x_{t2}, \ldots, x_{tn}) \in \mathcal{X}_t^n$, $t = 1, 2$, $y_t^n = (y_{t1}, y_{t2}, \ldots, y_{tn}) \in \mathcal{Y}_t^n$, $t = 1, 2$, $y^n = (y_1, y_2, \ldots, y_n) \in \mathcal{Y}^n$, and all integers $n \geq 1$. The two message sources shown in Fig. 26 produce statistically independent messages $m_1 \in \{1, 2, \ldots, M_1\}$ and $m_2 \in \{1, 2, \ldots, M_2\}$, each pair $(m_1, m_2)$ occurring with equal probability. Each encoder is completely described by a set of $n$ encoding functions which map the message and the sequence of received feedback outputs into the next channel input. Hence, $x_{ti} = f_t(m_t, y_{ti-1})$, $i = 1, 2, \ldots, n$, $t = 1, 2$. The decoder gives estimates of $m_1$ and $m_2$ based on his knowledge of the received sequence $y^n$. The definitions of a code, error probability, achievable rate pair, and capacity region are now straightforward in the case ($K_{2,1}, I, DGBF$), and therefore will be omitted here. The capacity region in situation ($K_{2,1}, I, DGBF$) is denoted by $\mathcal{C}(K_{2,1}, I, DGBF)$. We are now ready to state the achievable rate region for situation ($K_{2,1}, I, DGBF$) derived in [145], [151].

**Theorem 9.1** ([145], [151]) : For a d.m. MAC $\{\mathcal{X}_1 \times \mathcal{X}_2, P(y,y_1,y_2|x_1,x_2), Y \times Y_1 \times Y_2\}$, with different generalized output signals $Y_1$ and $Y_2$, an achievable rate region in situation ($K_{2,1}, I, DGBF$) is given by

$$R(K_{2,1}, I, DGBF)$$

$$= \{(R_1, R_2) : 0 \leq R_1 \leq I(X_1; Y|X_2, V_1, U) + I(V_1; Y_2|X_2, U),$$

$$0 \leq R_2 \leq I(X_2; Y|X_1, V_2, U) + I(V_2; Y_1|X_1, U)\}$$

(9.2a)  

(9.2b)
\[ R_1 + R_2 \leq \min \{ I(X_1; X_2; Y | V_1, V_2, U) + I(V_1; Y_2 | X_2, U) \\
+ I(V_2; Y_1 | X_1, U), I(X_1; X_2; Y) \} \]

(9.2c)

for some \( P(u, v_1, v_2, x_1, x_2, y, y_1, y_2) \)

\[ = P(u)P(v_1 | u)P(v_2 | u)P(x_1 | v_1, u)P(x_2 | v_2, u)P(y, y_1, y_2 | x_1, x_2). \]

In [145], another formulation is given of the rate region \( R_1(K_{21}, I, DGFB) \). Willems [145] has shown that both formulations are equivalent. It is observed in [145] that, since the cardinalities of the auxiliary RV's \( U, V_1 \), and \( V_2 \) are not explicitly bounded, the region \( R_1(K_{21}, I, DGFB) \) might not be closed. However, it is not difficult to establish that this region is convex. The main arguments of the proof of Theorem 9.1 can be found in [151], whereas a full proof, containing more details, appeared in [145].

Theorem 9.1 is a very general result, in that it contains a number of achievability results in multi-user information theory as a special case. In [145], [151] it is shown that the following rate and rate regions are special cases of \( R_1(K_{21}, I, DGFB) \): (i) the capacity region \( C(K_{21}, I) \) given in Theorem 3.1, (ii) the achievable region \( R_1(K_{21}, I, PPB) \), \( i = 1, 2 \), formulated in Theorem 8.1, (iii) the capacity regions \( C(K_{21}, IV, b) \), \( C(K_{21}, IV, e) \), and \( C(K_{21}, IV, f) \) defined in Section 10, and (iv) the rate \( R_0 \) shown by El Gamal and Aref [55] to be achievable for the d.m. relay channel.

Ideally, one would wish to prove a matching converse theorem for the d.m. MAC \( K_{21} \) with different generalized feedback signals. However, this is recognized to be a very difficult problem, since its solution would imply, among other things, a converse theorem for the d.m. two-way channel, pioneered by Shannon [123]. Even in the case of equal outputs the capacity region of Shannon's two-way channel is still unknown. As observed by Dueck [46], it seems very probable that the problem of determining the capacity region of the
d.m. MAC $K_{21}$ with feedback can only be solved after the capacity region of the d.m. two-way channel has been found. Research on the two-way channel is still in progress, since Shannon's historical paper [123]. Dueck [45] was the first one to give an example showing that Shannon's inner bound region is in general not the capacity region of the two-way channel. Schalkwijk [114],[116] gave specific coding strategies for Blackwell's binary multiplying channel which yield rate pairs beyond Shannon's inner bound. The determination of the capacity region of the binary multiplying channel is still an unresolved problem. Recently, Hekstra and Willems [75] obtained new upper bounds for the capacity region of the d.m. two-way channel with equal outputs and the capacity region of the d.m. MAC with feedback, which improve on previously known bounds.

9.2 The Gaussian MAC with generalized feedback

Both King [89] and Carleial [17] extended their results from the d.m. case to the case of a discrete-time AWGN MAC $K_{21}$ with generalized feedback. King considered thereby equal generalized feedback signals, whereas Carleial was concerned with different generalized feedback signals. In particular, King [89] established an achievable rate region for the AWGN MAC in situation (AWGN $K_{21}, N, P_1, P_2, I$) when there is a common generalized feedback signal to both encoders and the MAC is reversely degraded. Carleial [17] demonstrated an achievable rate region in the more general situation when there are different generalized feedback signals to both encoders and no restrictions on the AWGN MAC $K_{21}$. The Gaussian situation considered by Carleial [17] is formally similar to the situation ($K_{21}, I$, DGFB) considered in the discrete case. Carleial [17] showed that the region obtained by him includes as a special case the region $R_1(AWGN K_{21}, N, P_1, P_2, I, CFB)$ found by Carleial [17] and Cover and Leung [30]. However, the region found by Carleial [17] is not optimal, since it does not contain the capacity region $C(AWGN K_{21}, N, P_1, P_2, I, CFB)$ established by Ozarow
[101],[102] as a special case. On the contrary, the latter capacity region is strictly larger than the achievable region found by Carleial [17] when this one is specialized to the case of complete feedback.

10. THE DISCRETE MEMORYLESS MAC WITH CRIBBING ENCODERS

Shannon [123, Section 14] noted that in the case of the two-way channel the possible existence of dependence between the message sources might be used to attain transmission at rate points outside the inner bound. For the d.m. MAC we have seen examples of the same phenomenon in Sections 4 and 5. There we dealt with the d.m. MAC $K_{21}$ with specially correlated and arbitrarily correlated message sources, respectively, and saw that in those cases generally higher transmission rates can be achieved than in the classical situation. In Sections 7, 8 and 9 we have seen how complete, partial, or generalized feedback can be used to increase the achievable rate region of a d.m. or an AWGN MAC $K_{21}$ beyond the non-feedback capacity region. This latter phenomenon can be partly explained by the fact that, in the case of a MAC $K_{21}$, feedback enables the encoders to cooperate to some extent, in a similar way as dependence between the message sources given rise to cooperation. In this section we will consider still other kinds of cooperation between the encoders of a d.m. MAC $K_{21}$, which also lead to an enlargement of the capacity region $C(K_{21},1)$, and for which situations recently the capacity regions have been established [145],[149]. These new types of cooperation can all be captured under the name of "crbbing encoders".

In the context of a d.m. MAC $K_{21}$ the term "crbbing encoders" means that one of the encoders, encoder 2 say, learns part or all of the codeword put out by the other encoder (encoder 1) before making up his own codeword. This kind of cooperation between the two encoders of a d.m. MAC $K_{21}$ was first
considered by the author [133] for the specific situation where encoder 2 observes the entire codeword put out by encoder 1 before making up his own codeword. This is cribbing situation d below. In [133] an achievable rate region was put forward for this communication situation, which was erroneously claimed to be the capacity region, as was first pointed out to the author by Dueck [43]. The correct capacity region for this communication situation is given by Theorem 10.2 below. In total, five different cribbing situations were considered in [145],[149], for each of which the capacity region was established.

Consider again a d.m. MAC $K_{21}$ with two independent message sources which deliver messages $m_1$ and $m_2$ to two separate encoders 1 and 2, as in the classical situation $(K_{21}, I)$, but now with the following additional features. We first assume that only encoder 2 is cribbing. Encoder 1, the noncribbing encoder, is described by an encoding function $f_1$ which maps $m_1$ into a codeword $X^n_1 = f_1(m_1)$, $n$ being the block length. Encoder 2, the cribbing encoder, is described by a sequence of $n$ encoding functions $f_{2t}$, $1 \leq t \leq n$. The channel input $x_{2t}$ to be emitted by encoder 2 at time instant $t$, $1 \leq t \leq n$, will be based on the message $m_2$ and the knowledge which encoder 2 has gathered about the codeword $X^n_1$ through cribbing from encoder 1, i.e., $x_{2t} = f_{2t}(m_2, X^n_1)$. We can thus distinguish between four essentially different cribbing situations for encoder 2, depending on his knowledge prior to transmission period $t$, $1 \leq t \leq n$.

a. Encoder 2 learns nothing from encoder 1. In that case $x_{2t} = f_{2t}(m_2)$, and we are back in situation $(K_{21}, I)$. This is the classical situation dealt with in Section 3, for which the capacity region is given by Theorem 3.1. In the context of this section this situation is denoted by $(K_{21}, IV,a)$, but it will not be further discussed here.
b. Before emitting $x_{2t}$, encoder 2 has learned the sequence of channel inputs emitted by encoder 1 in all previous transmissions of the block, i.e., $x_{2t} = f_{2t}(m_2, x_{2t-1})$. We denote this communication situation by $(K_{21}, IV, b)$.

c. Before emitting $x_{2t}$, encoder 2 has learned, besides the sequence $x_{1t}^{t-1}$ of previous channel inputs of encoder 1, also the actual channel input that is going to be sent by encoder 1 in transmission $t$, i.e., $x_{2t} = f_{2t}(m_2, x_{1t}^t)$. We denote this communication situation by $(K_{21}, IV, c)$.

d. Before emitting $x_{2t}$, encoder 2 has learned from encoder 1 the entire codeword to be transmitted in the present block, i.e., $x_{2t} = f_{2t}(m_2, x_{1t}^n)$. This is the communication situation considered in [133, Section III C]. Here we denote this situation by $(K_{21}, IV, d)$.

The three communication situations $(K_{21}, IV, b)$, $(K_{21}, IV, c)$, and $(K_{21}, IV, d)$ can all be represented by the diagram of Fig. 27. They have in common that there is a side information link from the output of encoder 1 to the input of encoder 2. (One-way channels with side information at the transmitter were introduced and analyzed by Shannon [122].) This side information may be denoted by $u = t(x_{1t}^n)$.

Sofar we have assumed that only encoder 2 is cribbing, but next we allow also encoder 1 to crib from encoder 2. Encoder 2 must then base his channel input $x_{2t}$ at transmission period $t$, $1 \leq t \leq n$, besides on message $m_1$, on what he has learned through cribbing from encoder 2. This leads to two more causal cribbing situations.

e. Encoder 2 behaves as in situation $(K_{21}, IV, b)$. In addition, before emitting input letter $x_{1t}$, encoder 1 has learned the sequence of channel inputs emitted by encoder 2 in all previous transmissions of the block. The encoding operations for this situation are described by $x_{1t} = f_{1t}(m_1, x_{2t-1}^t)$.
and $x_{2t} = f_{2t}(m_2, x_{1}^{t-1})$, $1 \leq t \leq n$. This communication situation is denoted by $(K_{21}, IV, e)$.

f. Now encoder 2 behaves as in situation $(K_{21}, IV, c)$, while encoder 1 acts as in situation $(K_{21}, IV, e)$. The encoding operations for this cribbing situation are formally described by $x_{1t} = f_{1t}(m_1, x_{2}^{t-1})$ and $x_{2t} = f_{2t}(m_2, x_{1}^{t})$, $1 \leq t \leq n$. We denote this communication situation by $(K_{21}, IV, f)$.

The latter two cribbing situations, $(K_{21}, IV, e)$ and $(K_{21}, IV, f)$, are jointly represented by the schematic of Fig. 28. Both situations have in common that there is a side information link $u = s(x_{2}^{n})$ from the output of encoder 1 to the input of encoder 2, and a side information link $v = s(x_{1}^{n})$ from the output of encoder 2 to the input of encoder 1.

In summary, there are, apart from situation $(K_{21}, I)$, basically five different causal cribbing situations for a d.m. MAC $K_{21}$. As to the definition of a code in each situation, it is assumed throughout that the messages $m_1$ and $m_2$ are selected from finite sets $\{1, 2, \ldots, M_1\}$ and $\{1, 2, \ldots, M_2\}$, respectively, and occur with equal probability $1/(M_1 M_2)$. The encoding functions have been described separately for each of the five situations above.

Fig. 27. The d.m. MAC $(K_{21}, I)$ with one cribbing encoder (encoder 2), yielding situations $(K_{21}, IV, b)$, $(K_{21}, IV, c)$, and $(K_{21}, IV, d)$. 
Fig. 28. The d.m. MAC $(K_{21}, I)$ with both encoders cribbing, yielding communication situations $(K_{21}, IV, e)$ and $(K_{21}, IV, f)$.

The decoder must in each case estimate $m_1$ and $m_2$ based on his knowledge of the received sequence $Y^n$. The definitions of error probability, code, achievable rate pair, and capacity region are straightforward in each situation and hence will not be given here. We denote the capacity regions for each of the five new cribbing situations by respectively $C(K_{21}, IV, b)$, $C(K_{21}, IV, c)$, $C(K_{21}, IV, d)$, $C(K_{21}, IV, e)$, and $C(K_{21}, IV, f)$.

In [145],[149] the capacity region for each of the five communication situations $(K_{21}, IV, b)$, $(K_{21}, IV, c)$, $(K_{21}, IV, d)$, $(K_{21}, IV, e)$, and $(K_{21}, IV, f)$ has been established. The precise statements of these capacity regions are given in Theorems 10.1 through 10.4 below. The general shape of these regions is shown in Fig. 29. In turns out that, as can be seen from Fig. 29, all five capacity regions dominate the classical capacity region $C(K_{21}, I)$, whereas the capacity regions for cribbing situations $(K_{21}, IV, c)$ and $(K_{21}, IV, d)$ coincide. It is also shown in [145],[149] that the capacity regions $C(K_{21}, IV, c)$, $C(K_{21}, IV, d)$, and $C(K_{21}, IV, f)$ have parts in common with the total cooperation line $L_n$, as shown in Fig. 29. We now state the capacity theorems obtained in [145],[149] for the d.m. MAC $K_{21}$ with cribbing encoders.
Theorem 10.1 ([145], [149]): For a d.m. MAC with cribbing encoders in situation \((K_{21}, IV, b)\) the capacity region is given by

\[
C(K_{21}, IV, b) = \{(R_1, R_2) : 0 \leq R_1 \leq H(X_1|U), \quad (10.1a)
\]

\[
0 \leq R_2 \leq I(X_2; Y|X_1, U), \quad (10.1b)
\]

\[
R_1 + R_2 \leq I(X_1, X_2; Y), \quad (10.1c)
\]

for some \(P(u, x_1, x_2, y) = P(u)P(x_1|u)P(x_2|u)P(y|x_1, x_2)\)

such that \(\|U\| \leq \min \{\|X_1\| \cdot \|X_2\| + 1, \|Y\| + 2\}\).

Theorem 10.2 ([145], [149]): For a d.m. MAC with cribbing encoders in situations \((K_{21}, IV, c)\) and \((K_{21}, IV, d)\) the capacity regions are the same and given by

\[
C(K_{21}, IV, c) = C(K_{21}, IV, d)
\]

\[
= \{(R_1, R_2) : 0 \leq R_1 \leq H(X_1), \quad (10.2a)
\]

\[
0 \leq R_2 \leq I(X_2; Y|X_1), \quad (10.2b)
\]

\[
R_1 + R_2 \leq I(X_1, X_2; Y), \quad (10.2c)
\]

for some \(P(x_1, x_2, y) = P(x_1, x_2)P(y|x_1, x_2)\).

Theorem 10.3 ([145], [149]): For a d.m. MAC with cribbing encoders in situation \((K_{21}, IV, e)\) the capacity region is given by

\[
C(K_{21}, IV, e) = \{(R_1, R_2) : 0 \leq R_1 \leq H(X_1|U), \quad (10.3a)
\]

\[
0 \leq R_2 \leq H(X_2|U), \quad (10.3b)
\]

\[
R_1 + R_2 \leq I(X_1, X_2; Y) \quad (10.3c)
\]
for some $P(u,x_1,x_2,y) = P(u)P(x_1|u)P(x_2|u)P(y|x_1,x_2)$

such that $\|u\| \leq \min \{ \|X_1\| \cdot \|X_2\| + 1, \|Y\| + 2 \}$.

Theorem 10.4 ([145],[149]): For a d.m. MAC with cribbing encoders in situation $(K_2,IV,f)$ the capacity region is given by

$$C(K_2,IV,f) = \{ (R_1,R_2) : 0 \leq R_1 \leq H(X_1), \quad \left(10.4a\right)$$

$$0 \leq R_2 \leq H(X_2|X_1), \quad \left(10.4b\right)$$

$$R_1 + R_2 \leq I(X_1,X_2;Y), \quad \left(10.4c\right)$$

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

For each of the above theorems, on achievability part and a (weak) converse is proved in [145],[149]. The achievability proofs of Theorems 10.1 through 10.4 all involve the novel technique of backward decoding. Briefly described, backward decoding is a simultaneous (rather than sequential) decoding technique, whereby several blocks are simultaneously backward decoded. The powerful technique of backward decoding was developed by Willems [145] in his doctoral dissertation. In [145],[149] it was shown that, in the case of a d.m. MAC with cribbing encoders, backward decoding is superior to other techniques, previously used in multi-user information theory to demonstrate achievability in various situations. In the achievability proofs of Theorems 10.2 and 10.4 in particular, the notion of Shannon strategies (as introduced by Shannon [122]) plays an additional role of crucial importance. The proofs of the converses are all based on Fano's inequality [57]. As mentioned earlier, the capacity regions $C(K_{21},VI,b)$ through $C(K_{21},IV,f)$ are all graphically displayed in Fig. 29, together with the classical capacity $C(K_2,I)$ and the total cooperation line $L_T$. 
Fig. 29. Various capacity regions for the d.m. MAC $K_{21}$ in five cribbing situations, together with the classical capacity region $C(K_{21}, I)$ and the total cooperation line.

To conclude this section, we briefly mention another communication situation considered by Willems [145], [146], which in some sense is related to the situation of cribbing encoders. In [145], [146] Willems introduced and analyzed the communication situation whereby the encoders of a d.m. MAC $K_{21}$ are partially cooperating. In this case the encoders are connected with each other by communication links with finite capacities, which permit the encoders to communicate with each other. This enables the encoders to hold simultaneous monologues or a partially simultaneous dialogue, in which
different kinds of information can be communicated. The outcome of such a communication process, or \textit{conference}, is used by the encoders to form their channel inputs. Willems [145],[146] described this way a new communication system for the d.m. MAC $K_{21}$ in which the encoders are partially cooperating. This communication situation contains as special cases the situations $(K_{21}, I)$ and $(K_{21}, II)$, and the total cooperation case. Willems [145],[146] established the capacity region of a d.m. MAC $K_{21}$ with partially cooperating encoders by proving an achievability part and a weak converse.

11. CODING FOR MULTIPLE-ACCESS CHANNELS

Although this survey paper is primarily concerned with the Shannon information-theoretic aspects of the MAC, it would not be complete without mentioning some of the directions in which the coding theory for MAC's has recently developed. In this area, the aim is not so much to find a computable characterization of the capacity region of the MAC under consideration, as well as to describe constructive coding schemes which yield as large as possible rate sums.

Some of the early papers dealing with coding techniques for MAC's include the ones by Cohen, Heller, and Viterbi [25], Gaarder and Wolf [61], Kasami and Lin [81], and Wolf [153]. Cohen, Heller, and Viterbi [25] introduced convolutional codes for the noiseless asynchronous multiple-access OR channel. Gaarder and Wolf [61] gave a constructive coding scheme which demonstrated that if noiseless feedback links are provided from the output of the deterministic BEMAC to the encoders, then a rate pair can be achieved which falls outside the capacity region $C(K_{21}, I)$ (cf. Section 7).
Kasami and Lin [81] investigated the (block) coding problem for both the deterministic BEMAC and the noisy BEMAC. In the coding literature however, the deterministic (resp. noisy) BEMAC is usually referred to as the two-user noiseless (resp. noisy) binary adder channel (BAC). Hence, in this section we will adhere to this terminology. The two-user noisy BAC is identical to the two-user noiseless BAC, except that all 12 transition probabilities are non-zero. Kasami and Lin [81] defined the coding problem for the two-user noiseless BAC to be the construction of a code pair \((C_1, C_2)\) of block length \(n\) that is uniquely decodable (UD), and is such that the corresponding rate pair \((R_1, R_2)\) is as close as possible to the boundary of the capacity region \(\mathcal{C}(K_{21}, I)\). A codepair \((C_1, C_2)\) is said to be uniquely decodable if the sums \(u + v\) of all pairs \((u, v) \in C_1 \times C_2\) are different. Likewise, the coding problem for the two-user noisy BAC is to construct a code pair \((C_1, C_2)\) that is UD and capable of correcting \(t\) or fewer errors, again such that the corresponding sum rate is as high as possible. In the original channel model considered by Kasami and Lin [81], it is assumed that bit (or symbol) and block (or frame) synchronization is maintained amongst the encoders and the decoders, in which case it is said to be synchronized (which terminology is in accordance with the notions defined in Section 3).

In [82] the notion of a \(\delta\)-decodable code pair was introduced, and it was shown that a two-user \(\delta\)-decodable code is capable of correcting \(t = \lfloor((\delta - 1)/2)\rfloor\) or fewer transmission errors over the two-user noisy synchronized BAC, where \(\lfloor q \rfloor\) denotes the largest integer less than or equal to \(q\). It was also proved in [82] that, for a \(\delta\)-decodable two-user code \((C_1, C_2)\), the component codes \(C_1\) and \(C_2\) must have minimum Hamming distances greater than or equal to \(\delta\). A one-decodable two-user code is just a UD code. In the sequel, the size of a code \(C\) is denoted by \(|C|\).
Among other results, Kasami and Lin [81] showed that, for \( n = 2 \), it is possible to construct a UD code pair for the two-user noiseless BAC with \( |C_1| = 2 \) and \( |C_2| = 3 \), yielding the rate pair \((R_1, R_2) = (0.5, 0.7925)\). For a long time this rate point has enjoyed the property that it had the largest sum rate \( R = R_1 + R_2 = 1.2925 \) among the rate pairs for which it was known that they were achievable by constructive methods.

Wolf [153] wrote one of the earliest expository reviews on coding for multi-user communication channels, considering thereby in particular the MAC and the broadcast channel. In [153] several constructive coding schemes were presented for specific d.m. MAC's \( \mathbb{K}_2 \), such as the noiseless and noisy two-user BAC and the two-user modulo 2 or exclusive OR channel.

Since then, a multitude of papers dealing with coding techniques for MAC's have appeared. An excellent and comprehensive survey article describing the results on coding for the d.m. MAC (and other multi-user systems) up to 1980 was written by Farrell [58]. We refer the reader to this survey article for a systematic account of the developments in coding for the MAC which took place prior to 1980, and to the extensive bibliography contained in it. With respect to the d.m. MAC, Farrell [58] covered the following topics: (i) synchronized noiseless adder channel-block codes, (ii) synchronized noisy adder channel-block codes, (iii) synchronized adder channel-convolutional codes, (iv) unsynchronized adder channel, (v) exclusive-OR (modulo-2) channel, and (vi) the T-user M-frequency channel. Because of the complete coverage given by Farrell [58] it is unnecessary to elaborate here further on any of the coding results for MAC's obtained prior to 1980. Rather we will confine ourselves to discussing some recent papers which have appeared after [58] was written. Thereby it will be necessary though to recall some of the results already described in [58], so as to be able to compare new results with older ones.
Kasami, Lin, Wei, and Yamamura [83] related coding for the two-user BAC to a problem in graph theory, and thus were able to improve previously existing lower bounds on the achievable rates of uniquely decodable and δ-decodable codes for the synchronized two-user BAC. For the synchronized noisy two-user BAC, the (asymptotic) lower bound on the achievable rates of δ-decodable codes obtained in [83] improves on the lower bound previously derived by Lin, Kasami, and Yamamura [92]. For the synchronized noiseless two-user BAC, the result in [83] leads to a new (asymptotic) lower bound on the achievable rates of UD codes, which provides an improvement over a previous lower bound derived by Kasami and Lin [82] in the ranges 0.275 < R_1 < 0.5 and 0.7925 < R_1 < 1. This new lower bound is shown in Fig. 30. It is seen to contain the rate point (R_1, R_2) = (0.5, 0.7925) originally found by Kasami and Lin [82] as well as its symmetric counterpart. Katsman [84] derived a similar asymptotic lower bound on the achievable rates of δ-decodable codes for the synchronized noisy two-user BAC.

Kasami et al. [83] also investigated coding for the noiseless two-user BAC without perfect synchronization. In this channel model, which was called the noiseless quasi-synchronized (two-user) BAC in [83], it is assumed that the two encoders do not maintain block synchronism while the decoder maintains block synchronism with each encoder, and moreover that bit synchronism is maintained amongst the encoders and decoder. Earlier, Wolf [154] and Deaett and Wolf [33] had considered both this situation and the case of complete nonsynchronism, whereby the assumption of block synchronism between the decoder and the encoder, as well as the assumption of bit synchronism is dropped. (See in this regard also the discussion in Subsection 3.d concerning the various notions of asynchronism.) Kasami et al. [83] derived lower bounds on the achievable rates of UD codes for the quasi-synchronized two-user noiseless BAC, which fall above the rates of the codes found by Deaett and
Wolf [33] and Lin et al. [92] for the nonsynchronized two-user noiseless BAC. As discussed in Section 3, it has been demonstrated in [31] that the capacity region of the quasi-synchronous MAC is the same as that of the synchronous MAC, i.e., \( C(K_{21}, I, AS) = C(K_{21}, I) \), provided the delay is fixed, or very small compared to the block length.

We now return to the case of the synchronized noiseless two-user BAC. Whereas the lower bound derived in [83] has only an asymptotic meaning, van Tilborg [136] actually described a UD code for the synchronized noiseless two-user BAC which yields a rate point that lies strictly above this lower bound. The code pair \((C_1, C_2)\) found by van Tilborg [136] has parameters \( n = 5, |C_1| = 6, |C_2| = 15 \), leading to the rate pair \((R_1, R_2) = (0.5170, 0.7814)\).

This rate point and its symmetric counterpart are shown in Fig. 31, together with the lower bound obtained in [83]. Coebergh van den Braak and van Tilborg [23], [24] continued these investigations and presented an explicit construction of UD code pairs for the noiseless two-user BAC, yielding many more rate points with sum rate exceeding the non-constructive lower bound of [83]. For example, Coebergh van den Braak and van Tilborg [23] found for \( n = 48 \) a UD code with rate pair \((R_1, R_2) = (0.44514, 0.85820)\) and sum rate \( R = 1.30334 \).

In [24], another code pair is presented, found by Coebergh van den Braak [23] by a different method, yielding a sum rate 1.30565. Khachatrian [86] also constructed a UD code for the synchronized noiseless two-user BAC, with rate pair \((R_1, R_2) = (0.2222, 0.9936)\) and sum rate \( R = 1.2158 \), thus exceeding the lower bound of [83]. In addition, Khachatrian [87] investigated the problem of constructing linear \( \delta \)-decodable code pairs for the synchronized noisy two-user BAC, and presented examples of such codes which, for a given \( \delta \), yield rate pairs that lie above the so-called time sharing bound, which is defined as the maximum possible rate of the binary code of length \( n \) with minimal distance \( \delta \). In [88], these investigations were continued.
Fig. 30. Asymptotic lower bound on achievable rates of uniquely decodable code pairs for the synchronized noiseless two-user BAC, together with two constructively achievable rate pairs dominating this bound.

We finally mention that van Tilborg [136] provided an algebraic and combinatorial upper bound on $|C_2|$, given the code $C_1$, when $(C_1, C_2)$ is a UD code pair for the noiseless two-user BAC.

We next turn to another particular MAC for which coding has been examined. Györfi and Kerekes [68] investigated block coding for the noiseless asynchronous multiple-access OR channel, following earlier work by Cohen, Heller, and Viterbi [25], who introduced this model for multiple access communication. Mathematically, this d.m. MAC can be described as follows. It has $T \geq 2$ binary input alphabets $X_1, X_2, \ldots, X_T$, a binary output alphabet $Y$, and the output $y$ of the channel is zero if and only if $x_1 = x_2 = \ldots = x_T = 0$. In [25] and [127] examples are given of completely asynchronous communication schemes where this channel model is suitable. Wolf [155] described this channel
as a "noiseless model of an on-off pulse channel where the channel produces an output pulse if and only if at least one of the inputs was a pulse".

In the context of classical multiple-access information theory it is usually assumed that the channel is fully synchronized and the decoder is common. If one replaces all variables in this model by their complements one gets the multiple-access binary multiplying channel considered in [131]. In the synchronous case the classical capacity region of this channel equals the total cooperation capacity region, as shown in [131] and [155]. Then the coding problem is fairly trivial: the best rule is time-sharing with zero error. As already mentioned in Subsection 4.h, Wolf [155] referred to the situation \((K_{T_1}, I)\), where there are \(T > 2\) separate encoders, one for each message, and a common decoder, as the case of partial cooperation. In addition to this situation and the total cooperation case, Wolf [155] considered the situation where there are \(T\) separate decoders, the \(i\)th decoder knowing only the code of the \(i\)th encoder, which he called the no cooperation case. It is this latter situation for which Györfi and Kerekes [68] considered the coding problem in the case of the noiseless asynchronous multiple-access OR channel.

In particular, Györfi and Kerekes [68] considered for the noiseless \(T\)-user OR channel the situation where: (i) the \(T\) decoders are separated (the no cooperation case) and (ii) there is no block synchronism between the encoders, but synchronism is maintained between the \(i\)th encoder (the case of quasi-synchronism). For this quasi-synchronous noiseless \(T\)-user OR channel in the no cooperation case, Györfi and Kerekes [68] introduced a block encoding and decoding rule and proved an exponential error bound if the sum of the equal code rates is less than \(\ln 2\). Moreover, Györfi and Kerekes [68] showed that if in the same situation noiseless feedback is allowed zero error may result when again the sum of the equal code rates is less than \(\ln 2\).
As already mentioned by Farrell [58], Chang and Weldon [19] investigated coding schemes for the synchronized T-user (T ≥ 2) BAC, both with and without noise. The T-user BAC has T binary input alphabets, whereas the output alphabet contains T + 1 symbols. If \( x_i \) denotes the binary input of the \( i \)th sender and \( y \) is the output symbol, then the channel operation of the noiseless T-user BAC is defined by \( y = x_1 + x_2 + \ldots + x_T \), where the plus sign denotes real addition. The noisy T-user BAC can be regarded as the cascade of a noiseless T-user BAC with a d.m. one-way channel having non-zero transition probabilities for all possible \( T^2 \) input-output pairs. In the channel model considered by Chang and Weldon [19], it is further assumed that the T encoders maintain bit and word synchronization.

If \( (C_1, C_2, \ldots, C_T) \) denotes a T-user code for a T-user MAC, and \( R_i = \log_2 |C_i| \) is the rate of the \( i \)th constituent code, then the sum rate of this T-user code is defined (cf. [19], [58], [155]) by \( R_{\text{SUM}}(T) = R_1 + R_2 + \ldots + R_T \). In the case of partial cooperation, \( R_{\text{SUM}}(T) \) is upper bounded by the maximum joint mutual information \( C_{\text{SUM}}(T) = \max I(X_1, X_2, \ldots, X_T; Y) \), whereby the maximum is taken over all product distributions on the input RV's \( X_1, X_2, \ldots, X_T \), as discussed in Subsection 4.h (cf. inequality (4.9)). To accentuate that \( C_{\text{SUM}}(T) \) is the maximum sum rate in the partial cooperation case, Wolf [155] used the notation \( C_{\text{PC}}(T) \). Wolf [155] showed that for the synchronized noiseless T-user BAC, \( C_{\text{SUM}}(T) \) is approximately equal to \( (1/2) \log_2 (\pi T/2) \) for large \( T \).

In their article, Chang and Weldon [19] first present the capacity region \( C(K_{21}, T) \) of the synchronized noiseless T-user BAC, which is the characterization found by Liao [91] and Ulrey [130], applied to this MAC. Next, they derive upper and lower bounds on \( C_{\text{SUM}}(T) \) which are asymptotically tight with increasing \( T \). Moreover, Chang and Weldon [19] constructed a class of UD codes for the noiseless T-user BAC with rates which, asymptotically in \( T \),
equal the maximal achievable values. Their method of code construction involves the idea of an iterative construction based on annexing more columns (i.e., bits/word) to the difference matrix of a known UD code, and simultaneously increasing the number of rows (i.e., users) such that the new matrix is a difference matrix for a UD code with larger T. Subsequently, a class of error correcting codes for the noisy T-user BAC is constructed in [19] with the property that the rate vector is above the time-sharing hyperplane, i.e., such that $R_{\text{SUM}}(T) > 1$. Then it is shown that these codes can be used to construct multilevel codes suitable for use on the T-user AWGN MAC.

Continuing the above investigations, Chang and Wolf [20] studied a specific class of noiseless T-user MAC's which includes the noiseless T-user BAC as a special case. Chang and Wolf [20] considered two specific channel models for this class. In each case the goal is to find a T-user code and a decoding rule such that the error probability is arbitrarily small (preferably zero) and the rate sum of the corresponding set of rates is as large as possible. The two specific channel models under consideration in [20] are: (i) the T-user M-frequency MAC with intensity knowledge, and (ii) the T-user M-frequency MAC without intensity knowledge. These channels are referred to as the A channel and B channel, respectively, in [20]. (The A channel also played a role in the investigations of Vinck et al. [138]; cf. Section 7.)

The input alphabet, which is the same for each of the T users in both models, is assumed to consist of the M frequencies $f_1, f_2, \ldots, f_M$. The difference between the two channel models occurs in the output alphabet. For the A channel, "the output at each time instant is a symbol which identifies which subset of frequencies occurred as inputs to the channel at that time instant but not how many of each frequency occurred", whereas for the B channel "the output at each time instant indicates which subset of frequencies was trans-
mitted at that instant and how many of each frequency were transmitted."

For simplicity, in [20] only the noiseless case is pursued for both models. Consequently, the probability of decoding error is required to be zero for the code constructions. Both channel models include the noiseless two-user BAC as a special case. Chang and Wolf [20] observe that although the problem is formulated in terms of frequencies, the results are applicable to any signalling scheme where M orthogonal signals are used in each signalling interval, and thus in particular to pulse position modulation (PPM) where the signalling interval is partitioned into M time slots.

From the information-theoretic characterization of the capacity region \( C(K_T;1,I) \) it follows that for a d.m. noiseless T-user MAC generally \( C_{\text{SUM}}(T) = \max H(Y) \), where the maximum is taken over all product distributions on the input RV's \( X_1, X_2, \ldots, X_T \) (cf. inequality (4.9)). In [20] an attempt is made to calculate \( C_{\text{SUM}}(T) \) for the A and the B channel, for various values of \( M \geq 2 \). In this case, \( C_{\text{SUM}}(T) \) is written as \( C_{\text{SUM}}^{(A)}(T,M) \) and \( C_{\text{SUM}}^{(B)}(T,M) \), respectively, for a fixed value of \( M \). An integral part of the calculation of \( C_{\text{SUM}}(T,M) \) is concerned with the question of finding the input product distribution which maximizes the output entropy. In [20], various results concerning \( C_{\text{SUM}}^{(A)}(T,M) \) and \( C_{\text{SUM}}^{(B)}(T,M) \) have been obtained, some by computer search, others by analytic methods. Among other facts, it was found in [20] that for the T-user M-frequency A channel the maximum output entropy is achieved when all users utilize a common distribution. For \( T < M \), this is the uniform distribution. For \( T > M \), a nonuniform distribution yields the maximum output entropy. For fixed \( M \), \( C_{\text{SUM}}^{(A)}(T,M) \) increases with \( T \) until it reaches its maximum value at a value of \( T \) which is an integer close to \( M \cdot \ln 2 \). As \( T \) further increases, \( C_{\text{SUM}}^{(A)}(T,M) \) decreases until, for very large \( T \), \( C_{\text{SUM}}^{(A)}(T,M) \) asymptotically approaches \( M - 1 \). Fig. 31 gives plots of \( M \cdot C_{\text{SUM}}^{(A)}(T,M) \) versus \( T \) for \( 2 \leq M \leq 10 \) and \( 1 \leq T \leq \min(10, M + 2) \). The maximum \( M \cdot C_{\text{SUM}}^{(A)}(T,M) \)
Fig. 31. $M \cdot C_{\text{SUM}}^{(A)}(T, M)$ for the $T$-user $M$-frequency $A$-channel, $2 \leq M \leq 10$, $1 \leq T \leq \min(10, M + 2)$. Reproduced from [20], [155] with the kind permission of Prof. J.K. Wolf.

for each $M$ is shown by an asterisk.

Furthermore, Chang and Wolf [20] gave constructive coding schemes for both channels which achieve zero error probability and whose rate sum is close to the derived information-theoretic bounds. Two code constructions are described for the $A$ channel, and two for the $B$ channel. One of the code constructions for the $B$ channel generalizes the iterative procedure given by Chang and Weldon [19] from the case of binary input alphabets to that of $M$-ary
input alphabets.

Ferguson [59] generalized the code derived by Chang and Weldon [19] for the noiseless T-user BAC in two ways: (i) by generalizing the iterative construction of [19], and (ii) via the notion of equivalence classes of such codes, thus giving a large class of UD T-user codes for the noiseless T-user BAC.

Following the investigations in [19] and [59], Chang [18] constructed explicitly a class of UD codes for the noiseless T-user BAC, which are of arbitrary length, satisfy a specific recursive relationship, and asymptotically achieve the maximal capacity sum $C_{\text{sum}}(T)$, as $T$ increases.

Dyn'kin and Kurdyukov [52] proposed a method of code construction for the noiseless two-user BAC in situation $(K_{21}, \text{II})$, i.e., where there are three independent sources and two encoders, and the source outputs are connected to the channel inputs according to the Slepian and Wolf [125] setup. In particular, a method is given of constructing binary codes that have zero error probability. Moreover, computational procedures are given that implement the coding and decoding algorithm. In [52], the total sum rate $R = R_0 + R_1 + R_2$ is asymptotically evaluated, as the block length $n$ tends to infinity, and is shown to approach $\log_2 3$ for $0 \leq R_1 + R_2 \leq 1/3$.

De Bruyn [36] investigated the problem of constructing UD codes for the synchronized noiseless two-user BAC in situation $(K_{21}, \text{III})$, i.e., that of the asymmetric MAC discussed in Subsection 4.a. She found several procedures for constructing such codes. One of the proposed constructions yields codes that, for sufficiently large block length, achieve every point within the capacity region, which, for this particular channel, is shown in Fig. 7.
In the area of tree, trellis, and convolutional coding for MAC's, several papers and results should be mentioned, which appeared after Farrell's survey article [58] was written.

Peterson and Costello [104] pointed out that the result in Ohkubo [98] is incorrect in the following sense. First, Ohkubo [98] considered the class of time-varying linear trellis codes, rather than convolutional tree codes for the d.m. MAC $K_{21}$, as the title of [98] suggests. Moreover, whereas Ohkubo [98] attempted to derive an upper bound on decoding error probability for two users communicating over a d.m. MAC $K_{21}$ with linear trellis codes, his result is incorrect since he did not consider all possible error events.

In [105], Peterson and Costello pointed out that previously known results regarding tree, trellis, and convolutional coding for d.m. MAC's included decoding techniques for such codes, and conditions which indicate whether or not a particular binary convolutional (linear) code pair can be used with the two-user BAC, but that up to then no general construction technique for tree and or trellis codes for use with this channel had been found, thus raising the question whether or not these codes do in fact exist. Peterson and Costello [105] considered two-user tree codes for use on an arbitrary two-user d.m. MAC $K_{21}$. The primary result of [105] is an upper bound on the ensemble average maximum likelihood decoding error probability for two-user tree codes. This bound is shown to approach zero exponentially with increasing encoder tail length for all rate pairs in the capacity region $C(K_{21},I)$. This latter result is the correct version of the result erroneously claimed by Ohkubo [98]. Moreover, it entails the existence of good two-user tree codes.
Furthermore, in [105] a two-user tree code exponent is defined and compared with the corresponding block coding error exponent derived by Liao [91]. This comparison is carried out for a specific two-user noisy BAC. It is shown in [105] that for this specific example the tree coding error exponent is larger than the block coding error exponent at all rate pairs in the capacity region \( C(K_{21}, I) \). This corresponds with Viterbi's [141] result for single-user codes. Finally, in [105] the same technique used to derive the error probability bound is used to derive a new lower bound on free distance for two-user tree codes. The derivation of an upper bound on maximum likelihood decoding error probability for random trellis codes is mentioned in [105] as an open problem.

Chevillat [21] investigated trellis coding for a class of \( N \)-user MAC's, in particular the \( N \)-user BAC. In [21] it is shown that \( N \)-user trellis coding, in conjunction with Viterbi decoding, permits the multiple-access function to be combined with forward error correction. Also in [21], the distance and rate properties are analyzed of linear and nonlinear \( N \)-user trellis codes, and a procedure is given for constructing \( N \)-user convolutional-code \( N \)-tuples with large free distance.

Whereas previous investigations regarding the use of trellis coding schemes on d.m. MAC's were concerned with distance and decodability properties of such schemes for linear trellis codes, Sorace [128] has studied the performance of this trellis coding by deriving a new random coding bound in algebraic form. As this bound is not easy to compute, a transfer function bound is derived as a more tractable alternative. This bound demonstrates the existence of code pairs with rate sum exceeding time-sharing for a binary signal set. In [128] the assumption of user synchronism has been maintained.
We conclude this section with a recent paper by Wei and Lin [142]. Whereas all previous studies have focused on homogeneous trellis codes, especially the linear ones called convolutional codes, in [142] non-homogeneous trellis codes are studied and usefully applied to the coding problem for the quasi-synchronous two-user BAC. A trellis code is said to be homogeneous if the number of branches emanating from each node in the trellis diagram is constant. A trellis code is said to be non-homogeneous if the number of branches emanating from the nodes in the trellis diagram are not all identical. Wei and Lin [142] presented some uniquely decodable code pairs \((C_1, C_2)\) that can be used to transmit information reliably over the quasi-synchronous two-user BAC. Such codes are called quasi-synchronous uniquely decodable (QSUD). One of the codes \((C_1)\) of the QSUD code pairs presented in [142] is a non-homogeneous trellis code, the other one \((C_2)\) is a common block code. The code rates found in [142] are better than those of the Deaett-Wolf codes [33], and are close or equal to the asymptotic rates of Kasami et al. [83]. Moreover, a method for calculating the rates of non-homogeneous trellis codes is described, and an algorithm for finding more QSUD code pairs for the quasi-synchronous two-user BAC is formulated in [142].

ACKNOWLEDGEMENT

The author wishes to thank K. De Bruyn for helpful comments and her assistance with the drafting of the figures.
REFERENCES


[50] G. Dueck, "The zero error feedback capacity region of a certain class of multiple-access channels," manuscript, University of Bielefeld, 1984.


