LIMITATIONS ON THE CAPACITY OF THE M-USER BINARY ADDER CHANNEL DUE TO PHYSICAL CONSIDERATIONS*

I. Bar-David†, E. Plotnik, R. Rom‡

Department of Electrical Engineering
Technion - Israel Institute of Technology
Haifa 32000, Israel

April 1992
Revised: June 1993

ABSTRACT

The capacity of the M-User Binary Adder Channel, subjected to various restrictions of physical nature, is investigated. The underlying propagation media considered are (i) fiber-optical, with lossless coupling and Poisson statistics (ii) radio, under Rayleigh fading and (iii) radio with constant amplitudes and random phases. Whereas the capacity of the unrestricted (ideal) model for the Binary Adder Channel is known to increase without limit with the number of users, it is shown here that, for each of these cases, the total capacity is upper bounded by a constant independent of of the number of users: in case (i) by $1.7Q_T$ bits per channel use, where Q_T is the parameter of the Poisson process, in case (ii) by 4.33 bits per channel use and in case (iii) by 4.27 bits per channel use.

Keywords:

Multiuser channel, Capacity, Radio channel multiplexing, Optical fiber communications.

---

* An early version of this paper was presented at the International Symposium on Information Theory, Japan, 1988

† Supported in part by the Technion fund for Research and Development

‡ Supported in part by Sun Microsystems, Inc. USA.

I. INTRODUCTION

Multiuser communications has been the subject of extensive research and development in the past few decades. Essentially, multiuser communication is a method of communication in which all or some of the users transmit simultaneously over a common channel in a way that allows the proper receiving of the transmitted messages [1,2]. A major consideration in devising codes and designing access scheme is clearly the nature of the channel over which this communication takes place. Two types of channels have been most popular: the Collision Channel [3] and the Binary Adder Channel [4]. The latter is the subject of this paper.

The M-user Binary Adder Channel (BAC) is defined such that at any time its output is the real sum of its binary inputs (Figure 1). Formally, let $X_i \in \{0, 1\}$ be the i-th input to the channel. Then the output is defined as

$$Z \stackrel{\Delta}{=} \sum_{i=1}^{M} X_i$$

where the summation is over the real numbers. Bit and block synchronization are assumed. We qualify this model as the *ideal* BAC.

Consider now M users communicating through an ideal BAC at a transmission rate of 1 symbol per unit time for each user. The i-th user $(i = 1, 2, \ldots, M)$ has a codebook consisting of $2^{R_i N}$ codewords each N bits long, meaning that the *information rate* of user i is R_i. The *sum rate* $R_{\text{SUM}}(M)$ is defined as $R_{\text{SUM}}(M) = \sum_{i=1}^{M} R_i$, and the *total capacity* $C_{\text{SUM}}(M)$ is the maximum value of $R_{\text{SUM}}(M)$ such that

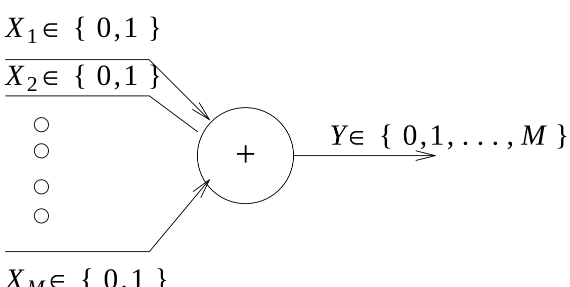

Figure 1: The M-User Binary Adder Channel (BAC)

reliable communication can be assured. The users are independent of one another, and P_i denotes the probability that the i-th user transmits a $"1"$ at any given time unit, i.e., $P_i \stackrel{\Delta}{=} Prob[X_i = "1"]$. Ahlswede and Liao [2,5], proved that for any discrete time and memoryless multiple access channel with independent inputs (of which our BAC is a special case):

$$C_{\text{SUM}}(M) \leq \max I(X_1, X_2, \ldots, X_M ; Y) \tag{1}$$

where the maximum is taken over all possible probability distributions $P(X_1, X_2, \ldots, X_M)$ of the form $P(X_1) \cdot P(X_2) \cdot \ldots \cdot P(X_M)$, and where $I(X_1, X_2, \ldots, X_M ; Y)$ is the mutual information between $(X_1, X_2, \ldots, X_M)$, and Y, [13, pp. 18]. In fact, since we consider only the symmetrical case in which $R_i = R$ for all i, and due of the symmetry of the problem, a result by Shamai (Shitz) [12] leads to the conclusion that equation (1) can be satisfied with equality.

For the ideal BAC Wolf [1] has shown that

$$\frac{1}{2} \log \frac{\pi M}{2} \leq C_{\text{SUM}}(M) \leq \frac{1}{2} \log (\pi e \left\lceil \frac{M}{2} \right\rceil) \tag{2}$$

implying that the capacity increases logarithmically with the number of users. This is a remarkable and a rather unexpected result which raises the question whether this increase in capacity with the number of users, who are also mutual interferers, is due to laxity in the ideal model. We therefore subjected the M-user BAC to various restrictions which reflect physical characteristics of the underlying channels. To that end we examined: (a) The mechanism by which the signals are added by the channel, and (b) The statistical properties of the information carrying signals. Indeed, in three variants of practical situations, the logarithmic increase in capacity with the number of users, as promised by the ideal BAC model, cannot be maintained; rather $C_{\text{SUM}}(M)$ is bounded by constants independent of M.

In Section II direct detection fiber-optical communication is considered. The light-wave signals are added in a lossless coupler and ideal photon counting is assumed. The photon counts obey Poisson statistics. For this channel $C_{\text{SUM}}(M) < 1.627 Q_T$ bits per channel use (b/c.u.), where Q_T is the expected number of photons emitted by a single user when sending $"1"$ and lossless fibers are assumed. Section III addresses radio transmissions with the summation being carried out by the receiving antenna. In subsection III.A a Rayleigh-fading channel is considered and it is shown that for any M, $C_{\text{SUM}}(M) < 4.328$ b/c.u.

A tighter bound obtained by a numerical technique, practical for small M values only, indicates that the bound is in the neighborhood of 2.5. An exact calculation for $M = 2$ indicates that the bounds might be quite loose, since $C_{\mathrm{SUM}}(2) = 1.027$b/c.u. Since an aggregated transmission rate of a 1 b/c.u. can be achieved by the simple time-sharing of the M users, we conclude that in this case no significant gain in performance is possible by *any* coding technique. In subsection III.B radio channels, in which information carrying signals have constant amplitude and random phases, are considered. Then, the total capacity $C_{\mathrm{SUM}}(M)$ is found to be upper-bounded by 4.261 b/c.u. There is no indication about the tightness of this bound.

II. M-USER FIBER-OPTICAL BINARY ADDER CHANNEL

Consider the communication system depicted in Figure 2 in which each user has an independent (On-Off-Keyed) laser modulating his emited bits. When the bit sent by the i-th user is $''1''$ ($X_i = ''1''$), a packet containing a random number of photons is emitted by its modulator for a period of T seconds. When a $''0''$ is transmitted ($X_i = ''0''$) no photons are transmitted. Denote by Q_T the expected number of photons emitted during a single $''1''$ transmission. A lossless optical device such as a Symmetrical Coupler is used to sum the incoming signals at the receiver's input, the power at the output port of the adder S_{out} is given by

$$S_{out} = \frac{1}{M} \sum_{i=1}^{M} S_i \quad , \tag{3}$$

where S_i is the power incoming from user i (see [14] page 22). We assume Direct Detection with an ideal photon counter detector, in which case the photon count obeys Poisson statistics. It follows from (3) that the expected photon count, Q, due to one user is $Q \overset{\Delta}{=} \frac{Q_T}{M}$. Then, the conditional probability distribution of Y, the photon counter output, (conditioned on the number of users that sent $''1''$) is given by

$$P(Y = k \mid j \text{ users sent } ''1'') = \frac{e^{-jQ} \cdot (jQ)^k}{k!} \qquad \begin{matrix} (k = 0, 1, \ldots, \infty) \\ (j = 0, 1, \ldots, M) \end{matrix} \tag{4}$$

where for $j = k = 0$, 0^0 is taken as 1.

To calculate the capacity we use equation (1) and get:

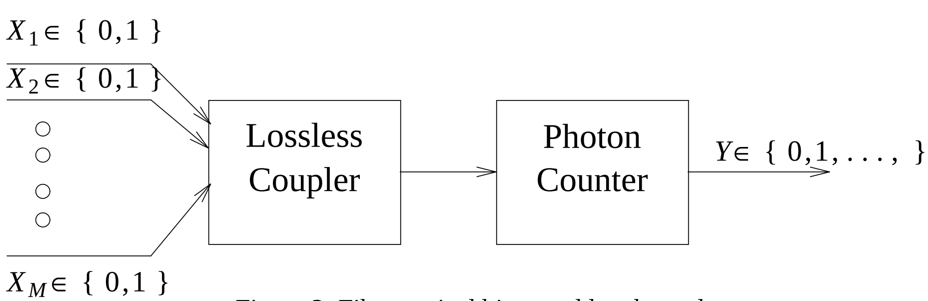

Figure 2: Fiber-optical binary adder channel

$$C_{\text{SUM}}(M) = \max_{P_1, P_2, \ldots, P_M \in [0,1]^M} I(X_1, X_2, \ldots, X_M ; Y) =$$

$$= \max_{P_1, P_2, \ldots, P_M \in [0,1]^M} E\left[\log \frac{P(Y | X_1, \ldots, X_M)}{P(Y)}\right] .$$

From the symmetry of the problem and the convexity of $I(X_1, X_2, \ldots, X_M ; Y)$ in $P_1, P_2, \cdots, P_M$ [6, p. 28], the maximum is achieved for $P_1 = P_2 = \ldots = P_M \stackrel{\Delta}{=} P$, where P is the probability that a certain user transmits a $"1"$. To find $P(Y)$, we note that $Prob(j$ users sent $"1")$ is given by $\begin{bmatrix} M \\ j \end{bmatrix} P^j (1-P)^{M-j}$. Combined with (4) we have

$$P(Y=k) = \sum_{j=0}^{M} \frac{\begin{bmatrix} M \\ j \end{bmatrix} P^j (1-P)^{M-j} e^{-Qj} (jQ)^k}{k!} \qquad (k = 0, 1, \ldots) \tag{5}$$

Therefore the explicit expression for the mutual information is

$$I(X_1, X_2, \ldots, X_M ; Y) = \sum_{k=0}^{\infty} \sum_{j=0}^{M} \begin{bmatrix} M \\ j \end{bmatrix} P^j (1-P)^{M-j} \frac{e^{-Qj} (Qj)^k}{k!} \cdot \log \frac{e^{-Qj} (Qj)^k}{\sum_{i=0}^{M} \begin{bmatrix} M \\ i \end{bmatrix} P^i (1-P)^{M-i} \cdot e^{-Qi} \cdot (Qi)^k} \tag{6}$$

It is shown in Appendix A that the right handside of the above expression can be upper bounded, that is

$$I(X_1, X_2, \ldots, X_M ; Y) \le Q_T (1-P) \log e - Q_T P \log P .$$

Maximizing the right hand side of the above expression with respect to P leads to

$$I(X_1, X_2, \ldots, X_M ; Y) \le 1.627 Q_T$$

meaning that for all $M \ge 2$, $C_{\text{SUM}}(M) \le 1.627 Q_T$, independent of M.

III. THE M-USER RADIO BINARY ADDER CHANNEL

Consider the communication system depicted in Figure 3. When $X_i = "0"$, $X_i(t) \equiv 0$ $0 \le t \le T$, and when $X_i = "1"$, $S_i(t) = \sin \omega t \quad 0 \le t \le T, \quad (i = 1, 2, \ldots, M)$. The signals are sent over independent channels, and summed by the receiving antenna. The received signal is therefore

$$Y(t) = \sum_{i=1}^{M} V_i \sin[\omega t + \theta_i] \qquad 0 \le t \le T \quad , \tag{7}$$

where $V_i = 0$ if $X_i = "0"$ and if $X_i = "1"$, V_i is the amplitude of the received signal. We consider below two cases: one in which the V_i are i.i.d Rayleigh distributed and another in which $V_i = const \equiv 1$. In both cases the θ_i are i.i.d random variable uniformly distributed in $[0, 2\pi]$. The piecewise constant model assumption, allows us to consider a discrete time version of the channel. We assume the channel is memoryless, that is, the random variables of the lth bit transmission period $(lT \le t \le (l+1)T)$ are independent of the random variables of any other bit transmission period.

The channel output $Y(t)$ is passed through two matched filters, one for each quadrature component; the filters outputs, W and Z are sampled every T seconds and are a sufficient statistic. Using equation (1) for this equivalent time-discrete version of the channel, $C_{\text{SUM}}(M)$ is given by:

$$C_{\text{SUM}}(M) = \max_{P_1, P_2, \ldots, P_M \in [0,1]^M} I(X_1, X_2, \ldots, X_M ; W, Z) \tag{8}$$

where $P_i \overset{\Delta}{=} P(X_i = "1")$.

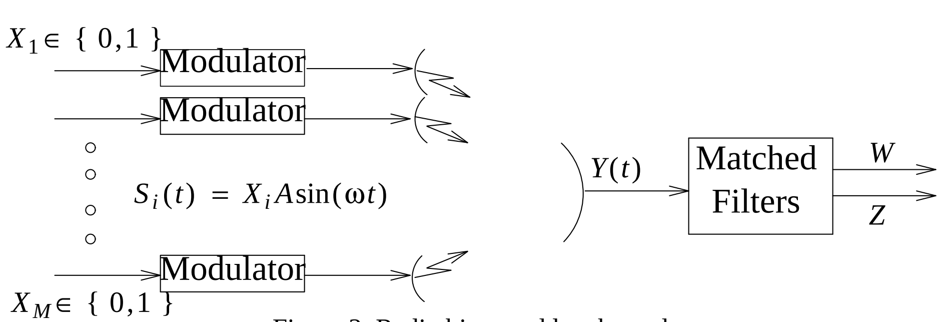

Figure 3: Radio binary adder channel

A. The M-User Radio BAC Under Rayleigh Fading

In this subsection we consider the first variant of the radio channel in which V_i (see equation (7)) is i.i.d Raleigh distributed. Equation (8), applied to this case becomes

$$C_{\text{SUM}}(M) = \max_{P_1, P_2, \ldots, P_M \in [0,1]^M} E\left[\log \frac{f(w, z \mid X_1, \ldots, X_M)}{f(w, z)}\right] \quad . \tag{9}$$

Due to the symmetry of the problem and the convexity of $I(X_1, X_2, \ldots, X_M ; W, Z)$, it suffices to consider one maximization parameter P in (9). To derive the explicit expression for $I(X_1, X_2, \ldots, X_M ; W, Z)$, notice that when $j \geq 1$ users transmit $"1"$, each filter's output is a sum of j independent Gaussian random variables with zero mean and variance σ^2. (The quadrature components of a fading signal are independent Gaussian random variables [7, pp. 527-532]). Hence, W and Z are independent and their conditional joint density function is given by

$$f_{W,Z}(w, z \mid j \quad \text{users transmit} \quad "1") = \begin{cases} \dfrac{1}{2\pi j\sigma^2} e^{-\frac{w^2 + z^2}{2j\sigma^2}} & j \geq 1 \\ \delta(w)\delta(z) & j = 0 \end{cases}$$

where $\delta(\cdot)$ is the Dirac delta function. The probability that j users transmit $"1"$ is $\begin{bmatrix} M \\ j \end{bmatrix} P^j (1-P)^{M-j}$, thus the unconditional density function $f_{W,Z}(w, z)$ is given by

$$f_{W,Z}(w, z) = (1-P)^M \delta(w) \cdot \delta(z) + \sum_{i=1}^{M} \begin{bmatrix} M \\ i \end{bmatrix} P^i (1-P)^{M-i} \cdot \frac{1}{2\pi i\sigma^2} e^{-\frac{w^2 + z^2}{2i\sigma^2}} \quad .$$

Because this density function is a mixture of a discrete and continuous function, the mutual information can be calculated by considering separately the discrete and continuous portions, leading to

$$I(X_1, X_2, \ldots, X_M; W, Z) = (1-P)^M \log \frac{1}{(1-P)^M} + \sum_{j=1}^{M} \binom{M}{j} P^j (1-P)^{M-j} \cdot$$

$$\cdot \iint_{w,z=-\infty}^{\infty} \log \left[\frac{\frac{1}{2\pi j\sigma^2} e^{-\frac{w^2+z^2}{2j\sigma^2}}}{\sum_{i=1}^{M} \binom{M}{i} P^i (1-P)^{M-i} \cdot \frac{1}{2\pi i\sigma^2} e^{-\frac{w^2+z^2}{2i\sigma^2}}} \right] \cdot \frac{1}{2\pi j\sigma^2} e^{-\frac{w^2+z^2}{2j\sigma^2}} dwdz \quad (10)$$

which, after some algebra reduces to

$$I(X_1, X_2, \ldots, X_M; W, Z) =$$

$$= -(1-P)^M \log [(1-P)^M] - \sum_{j=1}^{M} \binom{M}{j} P^j (1-P)^{M-j} \log j + [1-(1-P)^M] \cdot \log \frac{1}{2\pi\sigma^2} +$$

$$+ [1-(1-P)^M] \log e - \iint_{w,z=-\infty}^{\infty} \left[\sum_{j=1}^{M} \binom{M}{j} P^j (1-P)^{M-j} \frac{1}{2\pi j\sigma^2} e^{-\frac{w^2+z^2}{2j\sigma^2}} \right] \cdot \quad (11)$$

$$\log \left[\sum_{i=1}^{M} \binom{M}{i} P^i (1-P)^{M-i} \frac{1}{2\pi i\sigma^2} e^{-\frac{w^2+z^2}{2i\sigma^2}} \right] dwdz$$

For $M=2$, $I(X_1, X_2; W, Z)$ is directly calculated (see Appendix B), as a function of P, and is maximized with respect to P to yield $C_{\text{SUM}}(2) = 1.02653$ bits per channel use, when $P = 0.30568$. In Appendix C it is shown that for $M > 2$, the total capacity $C_{\text{SUM}}(M)$ in this case is upper-bounded by 4.328 b/c.u. independent of M.

In table I a tighter bound on $C_{\text{SUM}}(M)$ for several values of M is given. The results in the table were obtained by numerical maximization of equation C-5 (of Appendix C) with respect to P, for the given values of M. From the table we speculate that $C_{\text{SUM}}(M) < 2.5$ bits per channel use $\forall$ M. Since in the straight forward time-sharing strategy for M users, transmission rate of 1 bit per channel use is

M	Upper Bound (b/c.u)
3	2.2769
4	2.3182
5	2.3414
10	2.3843
20	2.4039
50	2.4152

Table I: Upper bounds on $C_{\text{SUM}}(M)$

achieved, it is evident that no significant gain in performance is possible by any coding technique in this channel.

B. The M-User Radio BAC with Constant Amplitudes and Random Phases

In this section we consider the second variant of the M-user radio channel. We consider again the communication system depicted in Figure 3 without fading but with random phase drifts. In this case the output signal $Y(t)$ of equation (7) is such that $V_i \equiv 0$ when $X_i = "0"$ and $V_i \equiv 1$ when $X_i = "1"$. A simplified model is obtained by defining the random variable $r \triangleq \sqrt{W^2 + Z^2}$. In this special case r takes values between 0 and M and is a sufficient statistic. Equation (8) therefore becomes

$$C_{\text{SUM}}(M) = \max_{0 \le P \le 1} I(X_1, X_2, \ldots, X_M; r) \quad , \tag{12}$$

where $P \triangleq P(X_i = "1")$, $\forall \ i = 1, 2, \ldots, M$.

The explicit expression for the mutual information between the inputs X_i and the output r is

$$I(X_1, X_2, \ldots, X_M; r) = \sum_{X_1, X_2, \ldots, X_M \in \{0,1\}^M} P(X_1) \cdot P(X_2) \cdot \ldots \cdot P(X_M) \cdot \int_{r=0}^{M} f_r(r | X_1, X_2, \ldots, X_M) \log \frac{f_r(r | X_1, \ldots, X_M)}{f_r(r)} \, dr \tag{13}$$

where the probability density function $f_r(r)$ can be written in terms of the conditional density function $f_r(r | X_1, X_2, \ldots, X_M)$ as

$$f_r(r) = \sum_{X_1, X_2, \ldots, X_M \in \{0,1\}^M} P(X_1) \cdot P(X_2) \cdot \ldots \cdot P(X_M) \cdot f_r(r | X_1, X_2, \cdots, X_M) \tag{14}$$

The symmetry of the problem allows us again to consider only the case $P(X_1) = P(X_2) \ldots P(X_M) \triangleq P$.

In upper-bounding $I(X_1, X_2, \ldots, X_M; r)$, we make use of the notion of *typical sequences*. We start with a brief overview of these sequences and some of their properties.

Let P be a discrete probability measure defined over an alphabet χ. A sequence $\underline{X} \in \chi^M$ is called δ-*typical sequence* for an information source emitting letters according to the measure P, if

$$\left| \frac{1}{M} \cdot N(a | \underline{X}) - P(a) \right| \le \delta \qquad \forall \quad a \in \chi \tag{15}$$

In equation (15) $N(a|\underline{X})$ is the number of appearances of the letter a in $\underline{X}$, and M is the length of the sequence $\underline{X}$. The set of all δ-typical sequences for a source with measure P is denoted by $(ZMTP$, *and the set of non-typical sequences by* $\overline{(ZMTP}$. *If* k *is the number of* $"1"$*'s in a sequence* $\underline{X}$ *of length* M *drawn from a memoryless binary source with* $P("1") = P$, *then from equation (15)* $\underline{X}$ *is* δ-*typical sequence if*

$$M(P-\delta) \leq k \leq M(P+\delta) \quad . \tag{16}$$

The following Lemma gives an upper-bound on the probability of the set $(ZMTP$.

Lemma 1: $P_r(\overline{(ZMTP}) \leq 2e^{-M \cdot f(P,\delta)}$ *where*

$$f(P,\delta) \stackrel{\Delta}{=} (\delta + P) \ln \frac{(1-P)(\delta+P)}{(1-P-\delta) \cdot P} + \ln \frac{(1-P-\delta)}{(1-P)} . \tag{17}$$

and $f(P,\delta)$ *is defined for* $0 < P < 1, \ \ \delta > 0$ *and* $P+\delta < 1$.

Proof: The proof is based on the Chernoff bound and appears in Appendix D.

To tie the notion of typical sequences with our case, we note that every vector $\underline{X} = X_1, \ldots, X_M$ that appears on the right handside of equation (13) with the associated probability $P(X_1) \cdot \cdots \cdot P(X_M)$ can be viewed as a sequence $\underline{X}$ of M letters emitted from a binary memoryless source with $P(X_i="1") = P$, $(i=1,\ldots,M)$.

Finally, we note that for large values of M and $\delta = C' \cdot P$ where C' is a constant such that $0 < C' < 1$, the maximum of $I(\underline{X}\ ; r)$ with respect to P, $(0 \leq P \leq 1)$ is obtained for $P > \frac{1}{M}$. If this were not so, it would have meant that the typical sequences $\underline{X}$ of length M are those containing a single $"1"$ (or the all zero sequence). But this means, in turn, that time-sharing is the optimal access strategy to the channel, meaning that $C_{\mathrm{SUM}}(M) = 1$ b/c.u., which we can clearly defeat.

On the other hand the maximum of $I(X_1, \ldots, X_M\ ; r)$ is achieved for $P < \frac{1}{2}$, since in such a system transmission of zeros is preferred in order to reduce the uncertainty at the channel's output associated with the transmission of $"1"$.

We present now the major steps in bounding $I(X_1, X_2, \ldots, X_M ; r)$. As can be seen from equation (13), expressions for $f_r(r \mid \underline{X})$ and $f_r(r)$ are needed. Since for a transmission of a $"0"$ r does not change, for a sequence $\underline{X}$ with k $"1"$'s and $M-k$ $"0"$'s, the problem of finding $f_r(r \mid \underline{X})$ is identical to finding the probability density function of the distance from the origin after k steps, in a two-dimensional random walk [9]. This density function is derived by differentiating the distribution $P(r, k) \triangleq r \cdot \int_0^\infty [J_o(x)]^k \cdot J_1(rx)\, dx$ with respect to r [9, equation (1.3)]. Unfortunately, tight bounds on $f(r \mid \underline{X})$ can be derived only for very large values of k. To overcome this difficulty, recall that $I(X_1, X_2, \ldots, X_M ; r)$ is nondecreasing function of M [6, p. 26], meaning that

$$I(X_1, X_2, \ldots, X_L ; r) \le \lim_{M \to \infty} I(X_1, X_2, \ldots, X_M ; r) \quad \forall\, L = 1, 2, \cdots \tag{18}$$

and thus it suffices to upper-bound $\lim_{M \to \infty} I(X_1, X_2, \ldots, X_M ; r)$ which we do next.

We start by breaking down the mutual information into two components based on the typical sequences. That is,

$$I(X_1, X_2, \ldots, X_M ; r) \triangleq I_{(ZMTP}(\underline{X} ; r) + I_{\overline{(ZMTP}}(\underline{X} ; r) \tag{19}$$

where

$$I_{(ZMTP}(\underline{X} ; r) = \sum_{\underline{X} \in (ZMTP} P(\underline{X}) \int_{r=0}^{M} f_r(r \mid \underline{X}) \log \left[\frac{f_r(r \mid \underline{X})}{\sum_{\underline{X}} f_r(r \mid \underline{X}) P(\underline{X})} \right] dr$$

$$I_{\overline{(ZMTP}}(\underline{X} ; r) = \sum_{\underline{X} \in \overline{(ZMTP}} P(\underline{X}) \int_{r=0}^{M} f_r(r \mid \underline{X}) \log \left[\frac{f_r(r \mid \underline{X})}{\sum_{\underline{X}} f_r(r \mid \underline{X}) P(\underline{X})} \right] dr \tag{20}$$

For a fixed value of $P (0 < P < 1)$ and large enough M, any sequence $\underline{X} \in (ZMTP$ *contains approximately* MP $"1"$*'s (see equation (16)), and when* M *tends to infinity bounds on* $f_r(r \mid \underline{X})$ *can be found. For sequences in* $\overline{(ZMTP}$ *the density function* $f_r(r \mid \underline{X})$ *is harder to evaluate, but the contribution of these sequences to* $I(X_1, X_2, \ldots, X_M ; r)$ *is negligible, as will be shown with the aid of Lemma 1.*

A detailed rigorous analysis along these lines is presented in appendix E. It is shown that $C_{\text{SUM}}(M)$ is upper bounded by 4.261b/c.u., independently of the number of users.

IV. CONCLUSION

We have shown that the ideal binary adder channel is not a realistic model for several practical cases such as the fiber optical and radio channels which are considered in this paper. Whereas the ideal BAC promises logarithmic increase in aggregate capacity with the number of users, in the cases considered the capacity is bounded by rather small values independent of the number of users.

The three cases considered are only representations of physical situations and by no means exhaustive. The conclusion, therefore, is only a caveat to remind the designer to look very suspiciously into the physical nature of the transmission medium before applying to it the ideal BAC model.

APPENDIX A

In this appendix we derive an upper bound on the mutual information for the optical channel. The derivation of the bound is based on the monotonicity of the functions $\log x$ and e^x, and the convexity of the function x^k that enable the use of Jensen's inequality.

Separating equation (6) into two terms we get

$$I(X_1, X_2, \ldots, X_M; Y) = \sum_{j=0}^{M} \binom{M}{j} P^j (1-P)^{M-j} \cdot \sum_{k=0}^{\infty} \frac{e^{-jQ} (Qj)^k}{k!} \log \left[e^{-jQ} \cdot j^k \right] - \\ - \sum_{j=0}^{M} \binom{M}{j} P^j (1-P)^{M-j} \cdot \sum_{k=0}^{\infty} \frac{e^{-jQ} (Qj)^k}{k!} \log \sum_{i=0}^{M} \binom{M}{i} P^i (1-P)^{M-i} e^{-Qi} i^k \qquad \text{(A-1)}$$

Expanding the first term we get

$$\sum_{j=0}^{M} \binom{M}{j} P^j (1-P)^{M-j} \cdot \sum_{k=0}^{\infty} \frac{e^{-jQ} (Qj)^k}{k!} \log (e^{-Qj} \cdot j^k) = \\ = -Q \log e \sum_{j=0}^{M} \binom{M}{j} P^j (1-P)^{M-j} \cdot j + \sum_{j=0}^{M} \binom{M}{j} P^j (1-P)^{M-j} \log(j) \cdot Qj \qquad \text{(A-2)}$$

Recalling that

$$\sum_{j=0}^{M} \binom{M}{j} P^j (1-P)^{M-j} \cdot j = MP \qquad \text{(A-3)}$$

we find that equation (A-2) is equal to

$$- QMP \log e + Q \sum_{j=0}^{M} \binom{M}{j} P^j (1-P)^{M-j} j \log j \qquad \text{(A-4)}$$

where $0 \log 0$ is considered as 0. Since $j \log j \le j \log M$ for $j = 1, 2, \ldots, M$ it follows that $E[j \log j] \le E[j] \log M$ and an upper bound on (A-4) (using (A-3)) is

$$-QMP \log e + QMP \log M \qquad \text{(A-5)}$$

To upper-bound $I(X_1, X_2, \ldots, X_M; Y)$, a lower bound on the second term of (A-1) is needed. For all $i \le M$, $e^{-Qi} \ge e^{-QM}$ thus, a lower bound is

$$\sum_{j=0}^{M} \binom{M}{j} P^j (1-P)^{M-j} \cdot \sum_{k=0}^{\infty} \frac{e^{-jQ} (Qj)^k}{k!} \log \left[e^{-QM} \cdot \sum_{i=0}^{M} \binom{M}{i} P^i (1-P)^{M-i} \cdot i^k \right] . \quad \text{(A-6)}$$

Using Jensen's inequality (since the function $f(x) = x^k$ is convex *cup*) we get

$$\sum_{i=0}^{M} \binom{M}{i} P^i (1-P)^{M-i} \cdot i^k \geq (M \cdot P)^k \quad \text{(A-7)}$$

and therefore equation (A-6) is lower bounded by

$$\sum_{j=0}^{M} \binom{M}{j} P^j (1-P)^{M-j} \cdot \sum_{k=0}^{\infty} \frac{e^{-jQ} (Qj)^k}{k!} \log [e^{-QM} \cdot (MP)^k] = -QM \log e + QMP \log MP. \quad \text{(A-8)}$$

Combining equations (A-5) and (A-8) yields

$$I(X_1, X_2, \ldots, X_M; Y) \leq QM(1-P) \log e - QMP \log P \quad \text{(A-9)}$$

and substituting QM by Q_T yields the desired result.

APPENDIX B

In this appendix we evaluate the mutual information for the 2-user fading channel. We derive an explicit expression for $I(X_1, X_2 ; W, Z)$ as a function of the probability that user transmits a $''0''$ (denoted P), and maximize this expression with respect to P.

Setting $M=2$ in (10) gives

$$I(X_1, X_2 ; W, Z) = (1-P)^2 \log \frac{1}{(1-P)^2} +$$

$$+ 2P(1-P) \iint_{w,z=-\infty}^{\infty} \log \left[\frac{e^{-\frac{w^2+z^2}{2\sigma^2}}}{2P(1-P)e^{-\frac{w^2+z^2}{2\sigma^2}} + \frac{P^2}{2} e^{-\frac{w^2+z^2}{4\sigma^2}}} \right] \cdot \frac{e^{-\frac{w^2+z^2}{2\sigma^2}}}{2\pi\sigma^2}\, dw\, dz + \tag{B-1}$$

$$3c' + P^2 \iint_{w,z=-\infty}^{\infty} \log \left[\frac{e^{-\frac{w^2+z^2}{4\sigma^2}}}{4P(1-P)\, e^{-\frac{w^2+z^2}{2\sigma^2}} + P^2 e^{-\frac{w^2+z^2}{4\sigma^2}}} \right] \frac{e^{-\frac{w^2+z^2}{4\sigma^2}}}{4\pi\sigma^2}\, dw\, dz$$

By the transformation $W = r \cos \theta$, $Z = r \sin \theta$, we get

$$I(X_1, X_2 ; W, Z) =$$

$$= (1-P)^2 \log \frac{1}{(1-P)^2} + \frac{2P(1-P)}{2\pi\sigma^2} \int_0^{2\pi} d\theta \int_0^{\infty} e^{\frac{-r^2}{2\sigma^2}} \cdot r \cdot \log \frac{e^{\frac{-r^2}{4\sigma^2}}}{\frac{P^2}{2} + 2P(1-P)\, e^{\frac{-r^2}{4\sigma^2}}}\, dr \tag{B-2}$$

$$+ \frac{P^2}{4\pi\sigma^2} \int_0^{2\pi} d\theta \int_0^{\infty} e^{\frac{-r^2}{4\sigma^2}} \cdot r \cdot \log \frac{\frac{1}{2}}{\frac{P^2}{2} + 2P(1-P)\, e^{\frac{-r^2}{4\sigma^2}}}\, dr \quad .$$

Solving these integrals [8] yields

$$I(X_1, X_2 ; W, Z) = -(1-P)^2 \log (1-P)^2 -$$

$$- \frac{p}{8(1-p)} \Big[4p(1-p)\log e + (1-p)^2 \log 2 - (16 - 24p + 9p^2)\log(4-3p) \tag{B-3}$$

$$+ (16-24p-7p^2)\log p \Big]$$

Maximizing (B-3) over $0 \leq P \leq 1$ leads to $C_{\mathrm{SUM}}(2) = 1.02653$b/c.u., for $P = 0.30568$.

APPENDIX C

In this appendix we upper bound the mutual information for the fading channel. Before we start with the main derivation we prove two technical propositions of which we shall make use later on.

The first one provides an upper bound on the differential entropy of a weighted sum of Gaussian random variables. We start by proving a lemma on mixing of distributions.

Lemma: Given two sets of nonnegative numbers $\{ \sigma_i \}$ and $\{ a_i \}$ $(i = 1, 2, \ldots, M)$ such that $\sum_{i=1}^{M} a_i = 1$, define a family of functions $\phi_i(x,y) \stackrel{\Delta}{=} \frac{1}{2\pi\sigma_i^2} e^{-(x^2+y^2)/2\sigma_i^2}$, and $\phi(x,y) \stackrel{\Delta}{=} \sum_{i=1}^{M} a_i \, \phi_i(x,y)$. Let the operator H (the differential entropy) be defined as $H\phi \stackrel{\Delta}{=} \iint_{-\infty}^{\infty} \phi(x,y) \log \frac{1}{\phi(x,y)} \, dxdy$. Then:

$$
\begin{aligned}
H\phi &\le \sum_{i=1}^{M} a_i \log(2\pi\sigma_i^2) + \log e \cdot E[\sigma^2] \cdot E\left[\frac{1}{\sigma^2}\right] = \\
&= \sum_{i=1}^{M} a_i H\phi_i + \left[E[\sigma^2] \cdot E\left[\frac{1}{\sigma^2}\right] - 1 \right] \log e
\end{aligned}
\qquad \text{(C-1)}
$$

where $E[\sigma^2] \stackrel{\Delta}{=} \sum_{i=1}^{M} a_i \, \sigma_i^2$, and $E\left[\frac{1}{\sigma^2}\right] = \sum_{i=1}^{M} \frac{a_i}{\sigma_i^2}$

Since $E[\sigma^2] \cdot E\left[\frac{1}{\sigma^2}\right] > 1$ by Jensen's inequality, the interpretation is that when Gaussian distributions of differing variances are mixed by linear weights, the resulting entropy is larger than the same linear weighting of the individual entropies. Equality is obtained above only if the original distributions are identical, in which case the mixing maintains the original distribution.

Proof: Directly from the definitions above it follows that

$$\sum_{i=1}^{M} a_i \log \phi_i(x,y) = \sum_{i=1}^{M} a_i \log \frac{1}{2\pi \sigma_i^2} - \frac{x^2 + y^2}{2} \log e \cdot \sum_{i=1}^{M} a_i \frac{1}{\sigma_i^2} =$$

$$= -\sum_{i=1}^{M} a_i \log(2\pi \sigma_i^2) - \frac{x^2 + y^2}{2} \log e \cdot E\left[\frac{1}{\sigma^2}\right]. \qquad \text{(C-2)}$$

By Jensen's inequality $\log (\sum_{i=1}^{M} a_i \phi_i(x)) \geq \sum_{i=1}^{M} a_i \log \phi_i(x)$. Therefore

$$H\phi = -\int_{-\infty}^{\infty}\int (\sum_{i=1}^{M} a_i\phi_i(x,y)) \log(\sum_{i=1}^{M} a_i\phi_i(x,y))\,dxdy \leq -\int_{-\infty}^{\infty}\int (\sum_{i=1}^{M} a_i\phi_i(x,y))(\sum_{i=1}^{M} a_i\log(\phi_i(x,y))\,dxdy =$$

$$= -\int_{-\infty}^{\infty}\int (\sum_{i=1}^{M} a_i \phi_i(x,y)) \left[\sum_{i=1}^{M} a_i \log\frac{1}{2\pi \sigma_i^2} - \frac{x^2 + y^2}{2} \log e \cdot \sum_{i=1}^{M} a_i \frac{1}{\sigma_i^2}\right] dxdy =$$

$$= -\sum_{i=1}^{M} a_i \log \frac{1}{2\pi \sigma_i^2} + \frac{1}{2}\log e \cdot E\left[\frac{1}{\sigma^2}\right] \cdot \int_{-\infty}^{\infty}\int \left[\sum_{i=1}^{M} a_i (x^2 + y^2) \cdot \phi_i(x,y)\right] dxdy = \qquad \text{(C-3)}$$

$$= \sum_{i=1}^{M} a_i \log(2\pi \sigma_i^2) + \log e \cdot E\left[\frac{1}{\sigma^2}\right] \cdot E[\sigma^2].$$

Equality in (C-1) holds iff $\sigma_i^2 = \sigma^2 \;\forall\, i = 1, \ldots, M$. □

We next prove a proposition that upper bounds $\sum_{i=1}^{M} \frac{\binom{M}{i} P^i (1-P)^{M-i}}{i}$ by $\frac{2}{M}$ and that will be used to upper bound the mutual information $I(X_1, X_2, \ldots, X_M; W, Z)$.

Proposition C-2: $\sum_{i=1}^{M} \frac{\binom{M}{i} P^i (1-P)^{M-i}}{i} < \frac{2}{M}$

Proof: Define $X(P) \triangleq \sum_{i=1}^{M} \frac{\binom{M}{i} P^i (1-P)^{M-i}}{i}$. To prove the proposition we need to show that $X(P) < \frac{2}{M}$. We do this by first calculating a closed form for $X(P)$ and then bounding it. Differentiating $X(P)$ with respect to P yields

$$X'(P) = \sum_{i=1}^{M} \binom{M}{i} P^{i-1} (1-P)^{M-i} - \sum_{i=1}^{M} \frac{(M-i)}{i} \binom{M}{i} P^i (1-P)^{M-i-1}$$

Rearranging the above leads to the following differential equation:

$$X'(P) + \frac{M}{1-P} X(P) = \frac{1-(1-P)^M}{P(1-P)}$$

the solution of which is given by

$$X(P) = \sum_{i=0}^{M-1} \frac{(1-P)^i - (1-P)^M}{M-i}$$

where we have made use of the boundary condition $X(0) = 0$.

By inductive proof it can be established that $X(P)$ is bounded from above by $2/M$.

□

We now compute the bound on the mutual information. Our point of departure is equation (11). We use the notation $a'_i = \begin{bmatrix} M \\ i \end{bmatrix} P^i (1-P)^{M-i} \quad (i = 1, \ldots, M)$ and define $\sigma_j^2 \stackrel{\Delta}{=} j\sigma^2$ and $a_i \stackrel{\Delta}{=} \frac{a'_i}{1 - (1-P)^M}$ (so that $\sum_{i=1}^{M} a_i = 1$). Equation (11) can now be rewritten as

$$I(X_1, X_2, \ldots, X_M ; W, Z) =$$
$$= (1-P)^M \log \frac{1}{(1-P)^M} - [1-(1-P)^M] \sum_{j=1}^{M} a_i \log (j) + [1-(1-P)^M] \cdot \log \frac{1}{2\pi\sigma^2} +$$
$$+ [1-(1-P)^M] \log e - \qquad \text{(C-4)}$$
$$- [1-(1-P)^M] \iint_{w,z=-\infty}^{\infty} \left[\sum_{j=1}^{M} a_i \phi_i(w,z) \right] \cdot \log \left[[1-(1-P)^M] \sum_{i=1}^{M} a_i \phi_i(w,z) \right] dwdz$$

and using the result of Lemma we get

$$I(X_1, X_2, \ldots, X_M ; W, Z) \leq$$
$$\leq (1-P)^M \log \frac{1}{(1-P)^M} - [1-(1-P)^M] \sum_{j=1}^{M} a_i \log (j) + [1-(1-P)^M] \cdot \log \frac{1}{2\pi\sigma^2} +$$
$$+ [1-(1-P)^M] \log e + [1-(1-P)^M] \left[\sum_{i=1}^{M} a_i \log(2\pi \sigma_i^2) + \log e \cdot E\left[\frac{1}{\sigma^2}\right] \cdot E[\sigma^2] \right] = \qquad \text{(C-5)}$$
$$= (1-P)^M \log \frac{1}{(1-P)^M} + [1-(1-P)^M] \log \frac{1}{1-(1-P)^M} +$$
$$+ [1-(1-P)^M] \left[\log e \cdot E\left[\frac{1}{\sigma^2}\right] \cdot E[\sigma^2] + \log e \right]$$

The term $E\left[\sigma^2\right]$ is an expectation of a binomial random variable:

$$E\left[\sigma^2\right] = \sum_{i=1}^{M} a_i i\sigma^2 = \frac{MP\sigma^2}{[1-(1-P)^M]} \ . \tag{C-6}$$

To bound the value of $E\left[\frac{1}{\sigma^2}\right]$ we use the result of the Proposition and get

$$E\left[\frac{1}{\sigma^2}\right] = \sum_{i=1}^{M} \frac{a_i}{i\sigma^2} = \sum_{i=1}^{M} \frac{\begin{bmatrix} M \\ i \end{bmatrix} P^i (1-P)^{M-i}}{[1-(1-P)^M] \cdot i \cdot \sigma^2} \leq \frac{2}{[1-(1-P)^M] M\sigma^2} \tag{C-7}$$

Finally, substituting equations (C-6) and (C-7) into (C-5) yields

$$\begin{aligned} I(X_1, X_2, \ldots, X_M; W, Z) \leq \\ \leq (1-P)^M \log\frac{1}{(1-P)^M} + [1-(1-P)^M] \log\frac{1}{[1-(1-P)^M]} + \\ + [1-(1-P)^M] \log e + \frac{2P\log e}{[1-(1-P)^M]} \end{aligned} \tag{C-8}$$

The maximum of (C-8) is obtained for $P = 1$ ($\forall M$) thus $C_{\mathrm{SUM}}(M) < 4.328$b/c.u.

APPENDIX D

Proof of Lemma 1.

From equation (15) it is clear that

$$P_r(\overline{T}_{[\delta]}) = P_r\left[\left| \frac{k}{M} - P \right| > \delta \right] \tag{D-1}$$

The Chernoff bound states that for any random variable x and every $S > 0$

$$P_r[X \geq A] \leq e^{-SA} \cdot E[e^{SX}] \quad . \tag{D-2}$$

Define $Y \overset{\Delta}{=} \frac{k}{M} - P$. From (D-2) we get

$$P_r\left[\left| \frac{k}{M} - P \right| > \delta \right] \leq e^{S_1 \cdot \delta} \cdot E[e^{S_1 \cdot Y}] + e^{S_2 \cdot \delta} \cdot E[e^{S_2 \cdot Y}] \qquad (\forall\, S_1 > 0\,,\, S_2 < 0) \tag{D-3}$$

But

$$E[e^{S_1 Y}] = e^{-S_1 \cdot P} \cdot E[e^{S_1 \cdot \frac{k}{M}}] \tag{D-4}$$

and

$$E[e^{S_1 \frac{k}{M}}] = \sum_{i=0}^{M} e^{S_1 \frac{i}{M}} \cdot P^i (1-P)^{M-i} \cdot \begin{bmatrix} M \\ i \end{bmatrix} = [\frac{e^{S_1}}{M} P + (1-P)]^M \tag{D-5}$$

Therefore

$$e^{S_1 \cdot \delta} \cdot E[e^{S_1 \cdot Y}] = e^{-S_1 (p+\delta)} \cdot [e^{\frac{S_1}{M}} \cdot P + (1-P)]^M \tag{D-6}$$

The right hand side of (D-6) is minimized for

$$S_1 = M \cdot \ln \frac{(1-P)(p+\delta)}{P(1-p-\delta)} \quad . \tag{D-7}$$

Similar arguments hold for S_2 as well.

Replacing these values for S_1 and S_2 in (D-4) completes the proof.

APPENDIX E

In this appendix an upper bound on the mutual information $I(X_1, \dots, X_M; r)$ is found where r is the distance from the origin to the sum of M randomly phased phasors. We briefly describe now the major stages in the derivation of the bound. First we show that the contribution of nontypical sequences $I_{\overline{(ZMTP}}(\underline{X}\ ;\ r)$ *to* $I(X_1, \dots, X_M; r)$ *tends to zero as* M *grows to infinity. Then we develop lower and upper bounds on the conditional density function* $f(r \mid \underline{X})$ *when* $\underline{X}$ *is a typical sequence. Finally, these bounds are used to upper bound the contribution of of typical sequences,* $I_{(ZMTP}(\underline{X}\ ;\ r)$, *to* $I(X_1, \dots, X_M; r)$.

Our point of departure is equation (19). We bound separately each of the terms and, as mentioned before, we must consider only the case $\frac{1}{M} < P < \frac{1}{2}$.

Starting with the first term of equation (19) we get

$$I_{\overline{(ZMTP}}(\underline{X}\ ;\ r) =$$

$$\sum_{\underline{X} \in \overline{(ZMTP}} P(\underline{X}) \cdot \int_{r=0}^{M} f_r(r \mid \underline{X}) \cdot \log \left[\frac{f_r(r \mid \underline{X})}{\sum_{\underline{X} \in (ZMTP} f_r(r \mid \underline{X}) P(\underline{X}) + \sum_{\underline{X} \in \overline{(ZMTP}} f_r(r \mid \underline{X}) P(\underline{X})} \right] dr \leq \tag{E-1}$$

$$\leq \sum_{\underline{X} \in \overline{(ZMTP}} P(\underline{X}) \cdot \int_{r=0}^{M} f_r(r \mid \underline{X}) \log \left[\frac{f_r(r \mid \underline{X})}{\sum_{\underline{X} \in \overline{(ZMTP}} f_r(r \mid \underline{X}) P(\underline{X})} \right] dr \leq \tag{E-2}$$

$$\leq \sum_{\underline{X} \in \overline{(ZMTP}} P(\underline{X}) \cdot \int_{r=0}^{M} f_r(r \mid \underline{X}) \log \left[\frac{f_r(r \mid \underline{X})}{P^M \cdot f_r(r \mid \underline{X})} \right] dr = \tag{E-3}$$

$$= \sum_{\underline{X} \in \overline{(ZMTP}} P(\underline{X}) M \log \frac{1}{P} \leq M \log \frac{1}{P} \cdot 2\, e^{-Mf(P,\,\delta)} \leq \tag{E-4}$$

$$\leq 2M \cdot \log M \cdot e^{-M \cdot f(P,\, \delta)} \tag{E-5}$$

Here the transition form equation (E-1) to (E-2) follows from the monotonicity of the logarithm function. The one from (E-2) to (E-3) is by taking into account only the nontypical sequence $\underline{X} = \underline{1}$. The transition from equation (E-3) to (E-4) is due to the fact that $\int_{r=0}^{M} f(r \mid \underline{X})\, dr = 1$. The inequality (E-4) is due to Lemma 1 and the final transition to (E-5) follows from $P > \frac{1}{M}$.

It is evident from equation (E-5) that $I_{\overline{(ZMTP}}(\underline{X}\,;\, r)$, *the contribution of the nontypical sequences, to* $I(X_1, X_2, \ldots, X_M\,;\, r)$ *tends to zero as M grows to infinity.*

To proceed with the bounding of the mutual information we first develop some bounds on the density function $f(r \mid \underline{X})$. We define

$$f(r \mid n) \overset{\Delta}{=} f_r(r \mid \underline{X} \text{ contains exactly } n \text{ "1"}) \tag{E-6}$$

and note that $f(r \mid n) = 0$ for $r > n \;\; \forall\, n = 1, \ldots, M$. Greenwood and Durand [9] have shown that this density function fulfills the equation

$$G(r,\, n) \overset{\Delta}{=} 1 - \int_0^r f(s \mid n)\, ds$$

and show that the function $G(\cdot\,,\,\cdot)$ is given by[7, (6.4)]

$$G(r,n) = e^{-z}\left[1 + \frac{1}{2n}\left[z - \frac{z^2}{2!}\right] + \frac{1}{12n^2}\left[-z + \frac{11z^2}{2!} - \frac{19z^3}{3!} + \frac{9z^4}{4!}\right] + \right.$$
$$\left. + \frac{1}{24n^3}\left[-2z - \frac{4z^2}{2!} + \frac{69z^3}{3!} - \frac{163z^4}{4!} + \frac{145z^5}{5!} - \frac{45^6}{6!}\right] + \cdots\right] \tag{E-7}$$

where $z = \frac{r^2}{n}$ and the higher terms in (E-7) are contributions of order lower than n^{-4}.

The density function $f(r \mid n)$ can be derived from (E-7) by differentiating $1 - G(r,\, n)$ with respect to r, and is given by

$$f(r|n) = \frac{2r}{n} \cdot e^{\frac{-r^2}{n}} \left[1 - \frac{1}{2n} + \frac{1}{12n^2} - \frac{1}{12n^3} + \frac{r^2}{n}\left[\frac{1}{n} + \frac{1}{12n^2} - \frac{1}{12n^3}\right] + \ldots\right] \quad \text{(E-8)}$$

We are looking for ways to bound $f(r|n)$ and as equation (E-8) suggests, we must look for an upper bound of the type $\gamma\frac{r}{n}e^{\frac{-\xi r^2}{n}}$ and a lower bound of the type $\beta\frac{r}{n}e^{\frac{-\alpha r^2}{n}}$ where α , β , γ , ξ are constants to be determined later. Note that since we shall evaluate $I_{(ZMTP)}(\underline{X}\ ;\ r)$ *for large values of* n *it is only necessary to find a bound that holds for* $n >> 1$.

By direct computation it can be shown that $\gamma = 2.1$ and $\xi = 0.7$ can serve as our upper bound i.e., there exist N_0 such that $\forall \ \ n > N_0$

$$f(r|n) < 2.1 \cdot \frac{r}{n} e^{\frac{-0.7r^2}{n}} \qquad r \in [0, n]. \quad \text{(E-9)}$$

Lower bounding $f(r|n)$ is a bit more complicated since for $r > n$, $f(r|n) = 0$ whereas $\beta\frac{r}{n}e^{\frac{-\alpha r^2}{n}} > 0$. We therefore divide $[0, n]$ into two segments, $[0, n-2)$ and $[n-2, n]$ and treat them separately. For the range $r \in [0, n-2]$ it can be shown by direct computation that $\beta = 2.0$ and $\alpha = 1.5$ can serve as our bound, i.e., there exist N_1 such that $\forall \ \ n > N_1$

$$f(r|n) > \frac{2r}{n} e^{\frac{-1.5r^2}{n}} \qquad r \in [0, n-2]. \quad \text{(E-10)}$$

The last case to consider is the range $r \in (n-2, n]$ for which Rice provides the following explicit formula [10, equation (4.1)]:

$$f(r|n) = \frac{n^{1/2}}{2\pi\,\Gamma\left(\frac{n-1}{2}\right)} \left[\frac{n-r}{2\pi}\right]^{(n-3)/2} \left[1 + \frac{(n-1)(n-4)}{4n} + \frac{(n^2 + 4n - 9)(n-2)^2}{32n^2} + \ldots\right] \quad \text{(E-11)}$$

Having derived these bounds we next develop some bounds on certain integrals of $f(r|n)$.

Proposition E-1: For $M(P - \delta) \le n \le M(P + \delta)$ the integral $\int_{M(P-\delta)}^{n} f(r|n)\,dr$ decreases exponentially with increasing M.

Proof: We prove the proposition by bounding the integral both from and above and below using the bounds on $f(r|n)$. We start with the bounds developed in equations (E-10) and (E-11).

$$\int_{M(P-\delta)}^{n} f(r|n)\,dr = \int_{M(P-\delta)}^{n-2} f(r|n)\,dr + \int_{n-2}^{n} f(r|n)\,dr \tag{E-12}$$

and address each of the above components separately. For the first component we get

$$\int_{M(P-\delta)}^{n-2} f(r|n)\,dr \geq \int_{M(P-\delta)}^{n-2} \frac{2r}{n} e^{\frac{-1.5r^2}{n}}\,dr = \left[\frac{2}{3} e^{\frac{-1.5r^2}{n}}\right]_{M(P-\delta)}^{n-2} =$$

$$= \frac{2}{3}\left[e^{-\frac{1.5[M(P-\delta)]^2}{n}} - e^{-\frac{1.5(n-2)^2}{n}}\right] \geq \frac{2}{3}\left[e^{-\frac{1.5[M(P-\delta)]^2}{M(P+\delta)}} - e^{-\frac{1.5(n-2)^2}{n}}\right] \tag{E-13}$$

Focusing on the second term of equation (E-12) we consider equation (E-11) and notice that in the range under consideration, $\frac{n-r}{2\pi}$ is at most $\frac{1}{\pi}$ which means that in this range $f(r|n)$ decreases exponentially with n. (The Γ function in the denominator, ensures an even faster decrease).

Using equation (E-9), we derive in a similar manner

$$\int_{M(P-\delta)}^{n} f(r|n)\,dr \leq \int_{M(P-\delta)}^{n} \frac{2.1r}{n} e^{\frac{-0.7r^2}{n}} \leq \frac{3}{2}\left[e^{-0.7M(P-\delta)} - e^{-0.7n}\right]$$

which concludes the proof. □

Corollary E-2: From the above proposition we deduce immediately:

$$\lim_{M\to\infty} M \log(M) \int_{M(P-\delta)}^{n} f(r|n)\,dr = 0\$$$

Finally, we introduce the notation

$$Q_i \triangleq \begin{bmatrix} M \\ i \end{bmatrix} P^i (1-P)^{M-i} \quad . \tag{E-14}$$

We now proceed to bound $I_{(ZMTP}(\underline{X}\,;\, r)$, *the second term of equation (19). Recalling that* $f(r|n) = 0$ *for* $r > n \;\; \forall\, n = 1, \ldots, M$ *this term becomes*

$I_{(ZMTP}(\underline{X}\,;\,r) \; =$

$$\sum_{\underline{X}\,\in\,(ZMTP} P(\underline{X}) \int_{r=0}^{M} f_r(r|\underline{X}) \, \log \left[\frac{f_r(r|\underline{X})}{\sum_{\underline{X}\,\in\,(ZMTP} f_r(r|\underline{X}) P(\underline{X}) \; + \; \sum_{\underline{X}\,\in\,\overline{(ZMTP}} f_r(r|\underline{X}) P(\underline{X})} \right] dr \; \le$$

$$\le \sum_{\underline{X}\,\in\,(ZMTP} P(\underline{X}) \int_{r=0}^{M} f_r(r|\underline{X}) \, \log \left[\frac{f_r(r|\underline{X})}{\sum_{\underline{X}\,\in\,(ZMTP} f_r(r|\underline{X}) \; P(\underline{X})} \right] dr \; = \qquad \text{(E-15)}$$

$$= \sum_{n=M(P-\delta)}^{M(P+\delta)} Q_n \int_{r=0}^{n} f(r|n) \, \log \frac{f(r|n)}{\sum_{j=M(P-\delta)}^{M(P+\delta)} Q_j \cdot f(r|j)} \; dr$$

Since the lower-bound on $f(r|n)$ given in equation (E-10), holds only for $0 < r < n-2$ we rewrite (E-15) as

$$I_{(ZMTP}(\underline{X}\,;\,r) \le \sum_{n=M(P-\delta)}^{M(P+\delta)} Q_n \left[\int_{0}^{M(P-\delta)-2} f(r|n) \log \frac{f(r|n)}{\sum_j Q_j \cdot f(r|j)} \; dr \; + \right. \qquad \text{(E-16)}$$

$$\left. + \int_{M(P-\delta)-2}^{n} f(r|n) \log \frac{f(r|n)}{\sum_j Q_j \cdot f(r|j)} \; dr \right]$$

and address the two terms separately. Considering first the second term of equation (E-16) we get

$$\sum_{n=M(P-\delta)}^{M(P+\delta)} Q_n \int_{M(P-\delta)-2}^{n} f(r|n) \log \frac{f(r|n)}{\sum_j Q_j \cdot f(r|j)} \; dr \; \le$$

$$\le \sum_{n=M(P-\delta)}^{M(P+\delta)} Q_n \int_{M(P-\delta)-2}^{n} f(r|n) \log \frac{f(r|n)}{Q_n \cdot f(r|n)} \; dr \; =$$

$$= \; - \sum_{n=M(P-\delta)}^{M(P+\delta)} Q_n \int_{M(P-\delta)-2}^{n} f(r|n) \log Q_n dr \; \le \; - \sum_{n=M(P-\delta)}^{M(P+\delta)} Q_n \int_{M(P-\delta)-2}^{n} f(r|n) \log P^M dr \; \le \qquad \text{(E-17)}$$

$$\le \sum_{n=M(P-\delta)}^{M(P+\delta)} Q_n M \log(M) \int_{M(P-\delta)-2}^{n} f(r|n) \, dr$$

where in the last transition we substituted P by its lowest possible value $\frac{1}{M}$. By Corollary E-2 this term vanishes as M increases to infinity, which means that in the limit

$$I_{(ZMTP}(\underline{X}\,;\,r) \le \sum_{n=M(P-\delta)}^{M(P+\delta)} Q_n \int_{0}^{M(P-\delta)-2} f(r|n) \log \frac{f(r|n)}{\sum_j Q_j \cdot f(r|j)} \; dr$$

Here the bounds in (E-9) and (E-10) are valid, since for any $M(p-\delta) \le n \le M(p+\delta)$ the value of r is less than $n-2$. By Jensen's inequality we have

$$\log \Big(\sum_j Q_j f(r|j)\Big) \geq \sum_j Q'_j \log [(1-\varepsilon) f(r|j)] \tag{E-18}$$

where $1 - \varepsilon \triangleq \sum_{j=M(P-\delta)}^{M(P+\delta)} Q_j$ and $Q'_j \triangleq \dfrac{Q_j}{1 - \varepsilon}$ (note that ε approaches 0 as M approaches infinity).

Hence,

$$I_{(ZMTP}(\underline{X}\,;\,r) \leq \sum_{n=M(P-\delta)}^{M(P+\delta)} Q_n \int_0^{M(P-\delta)-2} f(r|n) \log \frac{f(r|n)}{\sum_j Q_j \cdot f(r|j)}\, dr = \tag{E-19}$$

$$= \sum_{n=M(P-\delta)}^{M(P+\delta)} Q_n \int_0^{M(P-\delta)-2} f(r|n) \left[\log f(r|n) - \log[\sum_j Q_j \cdot f(r|j)]\right] dr \leq$$

$$\leq \sum_{n=M(P-\delta)}^{M(P+\delta)} Q_n \int_0^{M(P-\delta)-2} f(r|n) \left[\log f(r|n) - \sum_j Q'_j \log [(1-\varepsilon)\ f(r|j)]\right] dr =$$

$$= \sum_{n=M(P-\delta)}^{M(P+\delta)} Q_n \int_0^{M(P-\delta)-2} f(r|n) \left[\log f(r|n) - \sum_j [Q'_j \log f(r|j)]\right] dr -$$

$$- \log (1-\varepsilon) \cdot \sum_n Q_n \int_0^{M(P-\delta)-2} f(r|n)\ dr$$

The last term of the above expression is negligible. This is due to the fact that the integral is bounded and therefore so is the summation, while the logarithm approaches 0 as M (and n) increase. We can thus focus on the first two terms. We first replace $f(r|n)$ with the appropriate bounds from equations (E-9) through (E-11) and get

$$I_{(ZMTP}(\underline{X}\,;\,r) \le \sum_{n=M(P-\delta)}^{M(P+\delta)} Q_n \int_0^{M(P-\delta)-2} f(r|n)\left[\log f(r|n) - \sum_j [Q'_j \log f(r|j)]\right]dr =$$

$$= \sum_{n=M(P-\delta)}^{M(P+\delta)} Q_n \int_0^{M(P-\delta)-2} f(r|n)\left[\log\left[\frac{2.1r}{n}e^{-0.7\frac{r^2}{n}}\right] - \sum_j \left[Q'_j \log \frac{2r}{j}e^{-\frac{1.5r^2}{j}}\right]\right]dr =$$

$$= \sum_{n=M(P-\delta)}^{M(P+\delta)} Q_n \int_0^{M(P-\delta)-2} f(r|n)\left[\log\frac{2.1}{2} - \log n + \sum_j [Q'_j \log(j)] + \right.$$

$$\left. 4c' + \log e^{-0.7\frac{r^2}{n}} - \sum_j Q'_j \log e^{-\frac{1.5r^2}{j}}\right]dr = \qquad \text{(E-20)}$$

$$= \log\frac{2.1}{2}\sum_{n=M(P-\delta)}^{M(P+\delta)} Q_n \int_0^{M(P-\delta)-2} f(r|n)\,dr - \sum_{n=M(P-\delta)}^{M(P+\delta)} Q_n \log n \int_0^{M(P-\delta)-2} f(r|n)\,dr +$$

$$+ \left[\sum_j Q'_j \log(j)\right]\sum_{n=M(P-\delta)}^{M(P+\delta)} Q_n \int_0^{M(P-\delta)-2} f(r|n)\,dr +$$

$$+ \sum_{n=M(P-\delta)}^{M(P+\delta)} Q_n \int_0^{M(P-\delta)-2} f(r|n)\left[\log e^{-0.7\frac{r^2}{n}} - \sum_j Q'_j \log e^{-\frac{1.5r^2}{j}}\right]dr =$$

Because Q_n and $f(r|n)$ are probability measures we have that

$$\sum_{n=M(P-\delta)}^{M(P+\delta)} Q_n \int_0^{M(P-\delta)-2} f(r|n)\,dr \le 1 \quad .$$

In addition,

$$\int_0^{M(P-\delta)-2} f(r|n)\,dr = 1 - \int_{M(P-\delta)-2}^{n} f(r|n)\,dr \quad .$$

Substituting these into equation (E-20) yields

$$I_{(ZMTP}(\underline{X}\,;\,r) \le \log\frac{2.1}{2} + \sum_{n=M(P-\delta)}^{M(P+\delta)} Q_n \log n \int_{M(P-\delta)-2}^{n} f(r|n)\,dr -$$

$$- \left[\sum_j [Q'_j \log(j)\right] \cdot \sum_{n=M(P-\delta)}^{M(P+\delta)} Q_n \int_{M(P-\delta)-2}^{n} f(r|n)\,dr + \qquad \text{(E-21)}$$

$$+ \sum_{n=M(P-\delta)}^{M(P+\delta)} Q_n \int_0^{M(P-\delta)-2} f(r|n)\left[\log e^{-0.7\frac{r^2}{n}} - \sum_j Q'_j \log e^{-\frac{1.5r^2}{j}}\right]dr =$$

By corollary E-2 the second and third terms above approach 0 as M is increasing. To calculate the last term we replace our bounds from equations (E-9) and (E-10) and get

$$I_{(ZMTP}(\underline{X}\,;\,r) \le \log\frac{2.1}{2} + \sum_{n=M(P-\delta)}^{M(P+\delta)} Q_n \int_0^{M(P-\delta)-2} f(r\,|\,n)\left[\log e^{-0.7\frac{r^2}{n}} - \sum_j Q'_j \log e^{-\frac{1.5r^2}{j}}\right]dr$$

$$\le 0.071 + \sum_{n=M(P-\delta)}^{M(P+\delta)} Q_n \int_{M(P-\delta)-2}^{n} \frac{2r}{n} e^{-\frac{1.5r^2}{n}} \log e^{-0.7\frac{r^2}{n}}\, dr - \qquad \text{(E-22)}$$

$$- \sum_{n=M(P-\delta)}^{M(P+\delta)} Q_n \int_0^{M(P-\delta)-2} \frac{2.1r}{n} e^{-0.7\frac{r^2}{n}} \left[\sum_j Q'_j \log e^{-\frac{1.5r^2}{j}}\right]dr$$

The first integral above can be solved directly, and letting M increase to infinity we get the value $\frac{-\log e \cdot 0.7 \cdot 7.2}{2 \cdot (1.5)^2} = -0.448$. To bound the second integral, we observe that for $M(p-\delta) \le n,\ j \le M(p+\delta)$, holds $j \ge M(p-\delta) \ge \frac{n(P-\delta)}{(P+\delta)}$, thus

$$e^{-\frac{1.5r^2}{j}} > e^{-\frac{1.5r^2(p+\delta)}{n(p-\delta)}} \quad . \qquad \text{(E-23)}$$

A computation similar to the previous one bounds this term by 4.638.

Substituting all these results we find that

$$I_{(ZMTP}(\underline{X}\,;\,r) \le 0.071 - 0.448 + 4.638 = 4.261$$

which finally leads to

$$I(\underline{X}\,;\,r) \le 4.261 \text{ b/c.u.}$$

References

[1] J.K. Wolf, Multi-User Communication Networks in Communication Systems and Random Process Theory, J.K. Skwirgynski, Ed., Alphen aan den Rijn: The Netherlands, 1978, pp. 37-53.

[2] R. Ahlswede, Multi-Way Communication Channels, in Proc. 2nd Int. Symp. Information Theory, Tsahkadsor, Armenian S.S.R., Hungarian Acad. Sc., pp. 23-52, 1973.

[4] J.L. Massey and P. Mathys The Collision Channel without Feedback, IEEE Transactions on Information Theory, Vol IT-31, pp. 192-204, March 1985.

[4] E.C. Van der Muelen, A Survey of Multi-Way Channels in Information Theory: 1961-1976, IEEE Transactions on Information Theory, Vol IT-23, pp. 1-37, January 1977.

[5] H. Liao, A coding theorem for multiple access communications, in Proc. Int. Symp. Information Theory, Asilomar, Ca., 1972; also Multiple Access Channels, Ph.D. dissertation, Dept. Elec. Eng., Univ. of Hawaii, 1972.

[6] R.J. McEliece, *The Theory of Information and Coding*, Vol. 3 of the Encyclopedia of Mathematics and its Applications, Addison-Wesley Publishing Company, 1977.

[7] J.M. Wozencraft and I.M. Jacobs, *Principles of Communication Engineering*, John Wiley & Sons, Inc., 1967.

[8] I.S. Gradshteyn and I.M. Ryzhik, *Tables of Integrals Series and Products*, Academic Press, 1965.

[9] J.R. Greenwood and D. Durand, The Distribution of Length and Components of the Sum of n Random Unit Vectors, Ann. Math. Statist. Vol.26, pp. 233-246, 1955.

[10] S.O. Rice, Distribution of the Extreme Values of the Sum of n Sine Waves Phased at Random, Quart. Appl. Math., Vol.12, pp. 375-381, 1955.

[11] I.F. Blake and H. Darabian, Approximation for the Probability in the Tails of the Binomial Distribution, IEEE Trans. on Information Theory, Vol. IT-33, pp. 426-428, May 1987.

[12] S. Shitz, On Permutation Invariant Symmetric Multiple Access Channels with Application to Multiple Access Spread Spectrum, AEU, Vol.41 No. 6, 1987 pp. 347-355

[13] R.G. Gallager, *Information Theory and Reliable Communication*, New York, John Wiley, 1968.

[14] P.S. Henry, High-Capacity Lightwave Local Area Networks, IEEE Communications Magazine, Vol. 27, No. 10, 1989, p. 22.