Simpler Achievable Rate Regions for Multiaccess with Finite Blocklength

Ebrahim MolavianJazi and J. Nicholas Laneman
Department of Electrical Enginerring
University of Notre Dame
Notre Dame, IN 46556, USA
Email: {emolavia,jnl}@nd.edu

***Abstract*—Although practical communication networks employ coding schemes with blocklengths as low as several hundred symbols, classical information theoretic setups consider blocklengths approaching infinity. Building upon information spectrum concepts and recent work on channel dispersion, we develop a non-asymptotic inner bound on as well as a low-complexity, second-order achievable rate region for a discrete memoryless multiple access channel with a given finite blocklength and positive average error probability. Our bounds appear to capture essentially the same region as those of Tan and Kosut, but are less computationally complex because they require only the means and variances of the relevant mutual information random variables instead of their full covariance matrix.**

I. Introduction

Traditional channel coding theorems of information theory study the fundamental limits of communication in the presence of noise and interference using coding schemes of asymptotically large blocklengths. In such extremes, information can be encoded at a rate approaching a first order statistic (the *channel average mutual information*). Delay and complexity limitations of many practical applications, however, require coding with finite blocklength, even on the order of several hundred symbols, for which classical results do not hold. Following Strassen [1], it has recently been shown [2], [3] that a second order statistic (the *channel dispersion*) plays an important role in the fundamental limits with finite blocklength.

From a high-level perspective, both analyses stem from the common framework of information spectrum approach [4], i.e., treating mutual information as a random variable (RV); its *limiting* version for asymptotically large blocklength [4], and its n-letter form for finite blocklength [1], [2], [3]. In either case, the cumulative distribution function (CDF) of this RV characterizes performance in terms of the probability that channel cannot support the communication rate and causes an "outage" for the actual codeword to be correctly detected at the receiver. High coding rates turn out to arise when error probability is approximated by the outage probability, and the probability of "confusion", i.e., the observation is wrongly decoded to any incorrect codeword, decays to zero.

For the important case of stationary memoryless channels, the limiting mutual information rate RV concentrates with probability one at the average mutual information, so we are dealing with a zero or one outage probability, depending upon whether the communication rate is less than or greater than the average mutual information, or first-order characterization. Similarly, in the regime of finite blocklength, according to the Central Limit Theorem, the CDF of the n-letter mutual information rate RV approaches that of a Gaussian, so we can estimate the outage probability and approximate achievable rates by using first and second moments of the mutual information rate RV, or the second-order characterization.

In this paper, we show how similar ideas can be extended to a multi-user setting in which multiple users are communicating several independent messages to a single receiver over a multiple access channel. In particular, we explore the increase in coding rate, especially its second-order, as a function of the finite blocklength for a fixed average error probability. A key element of our work is to use an *outage-splitting* approach for the problem of assigning a single average error probability to several outage events arising in a DM-MAC. We demonstrate that this approach leads to simple, but rather tight achievable regions in the finite blocklength regime.

II. Problem Statement and Background

A 2-user discrete memoryless multiple access channel (DM-MAC) without feedback consists of two finite input alphabets $\mathcal{X}_1$ and $\mathcal{X}_2$, a finite output alphabet $\mathcal{Y}$, and a channel transition probability matrix $P_{Y|X_1X_2}(y|x_1,x_2)\colon \mathcal{X}_1 \times \mathcal{X}_2 \to \mathcal{Y}$ whose n-th extension follows

$$P_{Y^n|X_1^nX_2^n}(y^n|x_1^n,x_2^n) = \prod_{l=1}^{n} P_{Y|X_1X_2}(y_l|x_{1l},x_{2l}).$$

For such a DM-MAC $(\mathcal{X}_1,\mathcal{X}_2,P_{Y|X_1X_2}(y|x_1,x_2),\mathcal{Y})$, an (n,M_1,M_2,ϵ) code is composed of two message sets $\mathcal{M}_1 = \{1,...,M_1\}$ and $\mathcal{M}_2 = \{1,...,M_2\}$, and a corresponding set of codeword pairs and mutually exclusive decoding regions $\{(x_1^n(j),x_2^n(k),D_{j,k})\}$, with $j\in\mathcal{M}_1$ and $k\in\mathcal{M}_2$, such that the average error probability satisfies

$$P_e^{(n)} \triangleq \frac{1}{M_1M_2}\sum_{j=1}^{M_1}\sum_{k=1}^{M_2}\Pr[Y^n \notin D_{j,k}|X_1^n(j),X_2^n(k)\text{ sent}] \le \epsilon.$$

Accordingly, a $(\log M_1(n,\epsilon)/n, \log M_2(n,\epsilon)/n)$ pair is *achievable* for a DM-MAC $(\mathcal{X}_1,\mathcal{X}_2,P_{Y|X_1X_2}(y|x_1,x_2),\mathcal{Y})$ with finite blocklength n, and average error probability ϵ if such an (n,M_1,M_2,ϵ) code exists.

978-1-4673-2579-0/12/$31.00 ©2012 IEEE

As mentioned in Section I, channel coding rates depend upon the behavior of the relevant mutual information RVs. Specifically, a 2-user DM-MAC involves the following three mutual information RVs:

$$i(X_1;Y|X_2T) \triangleq \log \frac{P_{Y|X_1X_2}(Y|X_1X_2T)}{P_{Y|X_2T}(Y|X_2T)},$$
$$i(X_2;Y|X_1T) \triangleq \log \frac{P_{Y|X_1X_2}(Y|X_1X_2T)}{P_{Y|X_1T}(Y|X_1T)},$$
$$i(X_1X_2;Y|T) \triangleq \log \frac{P_{Y|X_1X_2}(Y|X_1X_2T)}{P_{Y|T}(Y|T)},$$

where T is an auxiliary "time sharing" RV satisfying the Markov Chain $T \to X_1X_2 \to Y$. In the regime of asymptotically large blocklength, achievable rates will depend on the first order statistics of these RVs:

$$\mathbb{I}(X_1;Y|X_2T) \triangleq \mathbb{E}[i(X_1;Y|X_2T)],$$
$$\mathbb{I}(X_2;Y|X_1T) \triangleq \mathbb{E}[i(X_2;Y|X_1T)],$$
$$\mathbb{I}(X_1X_2;Y|T) \triangleq \mathbb{E}[i(X_1X_2;Y|T)],$$

where expectation is taken with respect to the distribution $p(t)p(x_1|t)p(x_2|t)P_{Y|X_1X_2}(y|x_1,x_2)$. Using these quantities, Ahlswede and Liao [5] established the capacity region of a 2-user DM-MAC. Subsequently, Dueck [6] and Ahlswede [7] proved the strong converse, concluding that, even for a non-vanishing average error probability $0 < \epsilon \le 1$, the first-order characterization of the capacity region of a DM-MAC $(\mathcal{X}_1, \mathcal{X}_2, P_{Y|X_1X_2}(y|x_1,x_2), \mathcal{Y})$ is given by the closure as $n \to \infty$ of all $(\log M_1(n,\epsilon)/n, \log M_2(n,\epsilon)/n)$ pairs satisfying

$$\log M_1(n,\epsilon) < n\mathbb{I}(X_1;Y|X_2T) + o(n)$$
$$\log M_2(n,\epsilon) < n\mathbb{I}(X_2;Y|X_1T) + o(n)$$
$$\log M_1(n,\epsilon) + \log M_2(n,\epsilon) < n\mathbb{I}(X_1X_2;Y|T) + o(n)$$

for some choice of the joint distribution $p(t)p(x_1|t)p(x_2|t)P_{Y|X_1X_2}(y|x_1,x_2)$ with the auxiliary random variable T defined on a set $|\mathcal{T}| \le 2$.

In the following, we sharpen these classical results for the finite blocklength regime using the second order statistics or *dispersions* of the relevant mutual information RVs:

$$\mathbb{V}(X_1;Y|X_2T) \triangleq \mathrm{Var}[i(X_1;Y|X_2T)],$$
$$\mathbb{V}(X_2;Y|X_1T) \triangleq \mathrm{Var}[i(X_2;Y|X_1T)],$$
$$\mathbb{V}(X_1X_2;Y|T) \triangleq \mathrm{Var}[i(X_1X_2;Y|T)],$$

where the variances are again calculated with respect to the distribution $p(t)p(x_1|t)p(x_2|t)P_{Y|X_1X_2}(y|x_1,x_2)$.

III. Main Results

This section summarizes our main results in this paper. We first state the following Dependence Testing (DT) bound for a DM-MAC, which provides a non-asymptotic achievable region valid for any blocklength. It basically describes the error probability in terms of the outage and confusion probabilities, and is based on the DT bound of [2] and ideas from the information-spectrum approach for a general MAC [4].

Theorem 1. *For a DM-MAC $(\mathcal{X}_1, \mathcal{X}_2, P_{Y|X_1X_2}(y|x_1,x_2), \mathcal{Y})$ and for any joint distribution $p(t)p(x_1|t)p(x_2|t)$, there exists a (n, M_1, M_2, ϵ) code such that*

$$\begin{aligned}
\epsilon \le\ & \Pr\left[i(X_1^n;Y^n|X_2^nT^n) \le \log\gamma_1(X_1^n,X_2^n)\right] \\
&+ \frac{M_1-1}{2}\Pr\left[i(X_1^n;\bar{Y}_2^n|X_2^nT^n) > \log\gamma_1(X_1^n,X_2^n)\right] \\
&+ \Pr\left[i(X_2^n;Y^n|X_1^nT^n) \le \log\gamma_2(X_1^n,X_2^n)\right] \\
&+ \frac{M_2-1}{2}\Pr\left[i(X_2^n;\bar{Y}_1^n|X_1^nT^n) > \log\gamma_2(X_1^n,X_2^n)\right] \\
&+ \Pr\left[i(X_1^nX_2^n;Y^n|T^n) \le \log\gamma_3(X_1^n,X_2^n)\right] \\
&+ \frac{(M_1-1)(M_2-1)}{2}\Pr\left[i(X_1^nX_2^n;\bar{Y}_0^n|T^n) > \log\gamma_3(X_1^n,X_2^n)\right]
\end{aligned} \tag{1}$$

where $Y^n, \bar{Y}_0^n, \bar{Y}_1^n, \bar{Y}_2^n$ are n-fold distributions according to $P_{Y\bar{Y}_0\bar{Y}_1\bar{Y}_2|X_1X_2T}(y,a,b,c|x_1,x_2,t) = P_{Y|X_1X_2}(y|x_1,x_2)P_{Y|T}(a|t)P_{Y|X_1T}(b|x_1,t)P_{Y|X_2T}(c|x_2,t)$, and where $\gamma_1,\gamma_2,\gamma_3 : \mathcal{X}_1^n \times \mathcal{X}_2^n \to [0,\infty)$ are arbitrary measurable functions whose optimal choices to give highest rates are as follows:

$$\gamma_1(X_1^n,X_2^n) \equiv \frac{M_1-1}{2}, \qquad \gamma_2(X_1^n,X_2^n) \equiv \frac{M_2-1}{2},$$
$$\gamma_3(X_1^n,X_2^n) \equiv \frac{(M_1-1)(M_2-1)}{2}.$$

The above expression for DT bound is stated to match our outage-splitting approach later in Theorem 3. It is, however, possible to strengthen this bound by focusing on the three outages *jointly*.

Theorem 2. *For a DM-MAC $(\mathcal{X}_1, \mathcal{X}_2, P_{Y|X_1X_2}(y|x_1,x_2), \mathcal{Y})$ and for any joint distribution $p(t)p(x_1|t)p(x_2|t)$, there exists a (n, M_1, M_2, ϵ) code such that*

$$\begin{aligned}
\epsilon \le\ & \Pr\big[i(X_1^n;Y^n|X_2^nT^n) \le \log\gamma_1(X_1^n,X_2^n) \\
&\qquad \cup\ i(X_2^n;Y^n|X_1^nT^n) \le \log\gamma_2(X_1^n,X_2^n) \\
&\qquad \cup\ i(X_1^nX_2^n;Y^n|T^n) \le \log\gamma_3(X_1^n,X_2^n)\big] \\
&+ \frac{M_1-1}{2}\Pr\left[i(X_1^n;\bar{Y}_2^n|X_2^nT^n) > \log\gamma_1(X_1^n,X_2^n)\right] \\
&+ \frac{M_2-1}{2}\Pr\left[i(X_2^n;\bar{Y}_1^n|X_1^nT^n) > \log\gamma_2(X_1^n,X_2^n)\right] \\
&+ \frac{(M_1-1)(M_2-1)}{2}\Pr\left[i(X_1^nX_2^n;\bar{Y}_0^n|T^n) > \log\gamma_3(X_1^n,X_2^n)\right].
\end{aligned} \tag{2}$$

where $Y^n, \bar{Y}_0^n, \bar{Y}_1^n, \bar{Y}_2^n$ are n-fold distributions according to $P_{Y\bar{Y}_0\bar{Y}_1\bar{Y}_2|X_1X_2T}(y,a,b,c|x_1,x_2,t) = P_{Y|X_1X_2}(y|x_1,x_2)P_{Y|T}(a|t)P_{Y|X_1T}(b|x_1,t)P_{Y|X_2T}(c|x_2,t)$,

Next, we give an achievable region for a DM-MAC with (sufficiently large) finite blocklength, which is a consequence of the DT bound for DM-MAC in Theorem 1 by appealing to the Central Limit Theorem to utilize a Gaussian approximation for the relevant mutual information RVs, thus estimating the outage and confusion probabilities. In the following, $Q^{-1}(\cdot)$ is the well-known inverse of the complementary-CDF function of a standard Gaussian distribution $Q(x) \triangleq \frac{1}{2\pi}\int_x^\infty e^{-t^2/2}dt$.

Theorem 3. *An achievable region for the DM-MAC $(\mathcal{X}_1, \mathcal{X}_2, p(y|x_1, x_2), \mathcal{Y})$ is given by the union of all $(\log M_1(n,\epsilon)/n, \log M_2(n,\epsilon)/n)$ pairs satisfying*

$$\begin{aligned} \log M_1(n,\epsilon) &< n\mathbb{I}(X_1;Y|X_2T) \\ &\quad - Q^{-1}(\lambda_1\epsilon)\sqrt{n\mathbb{V}(X_1,Y|X_2T)}+O(1), \\ \log M_2(n,\epsilon) &< n\mathbb{I}(X_2;Y|X_1T) \\ &\quad - Q^{-1}(\lambda_2\epsilon)\sqrt{n\mathbb{V}(X_2,Y|X_1T)}+O(1), \\ \log M_1(n,\epsilon)+\log M_2(n,\epsilon) &< n\mathbb{I}(X_1X_2;Y|T) \\ &\quad - Q^{-1}(\lambda_3\epsilon)\sqrt{n\mathbb{V}(X_1X_2,Y|T)}+O(1), \end{aligned} \tag{3}$$

for some choice of the joint distribution $p(t)p(x_1|t)p(x_2|t)P_{Y|X_1X_2}(y|x_1,x_2)$ with the auxiliary random variable T defined on a set $|\mathcal{T}| \leq 6$, and for some positive constants $\lambda_1, \lambda_2, \lambda_3$ satisfying $\lambda_1+\lambda_2+\lambda_3 = 1$.

In light of our discussions in Section I, this theorem suggests that high rates arise from coding schemes in which outages dominate confusions, such that the average error probability ϵ is split among the three outage events of a 2-user DM-MAC according to some $\lambda_1, \lambda_2, \lambda_3$ partitioning. In comparison with Ahlswede and Liao's result [5], Theorem 3 suggests that taking finite blocklength into account introduces rate penalties (for the interesting case of $\epsilon < \frac{1}{2}$) that depend on blocklength, error probability and DM-MAC dispersions.

Our main result depends only on the mean and variance of the relevant mutual information RVs, each of which is approximated with a scalar Gaussian distribution. By contrast, in a concurrent work on this problem, Tan and Kosut [8] treat the outage events jointly, without using a union bound to split the outages. In fact, although our Theorem 3 follows from the DT bound of Theorem 1, the Tan and Kosut result [8] can be obtained as a second-order approximation of the generalized DT bound of Theorem 2. Although the approach of [8] leads to an inner bound for the DM-MAC that is larger in principle, it requires dealing with a full covariance matrix and the inverse CDF of a multi-dimensional Gaussian distribution, which we expect to be more computationally complex than our result particularly as the number of users grows. It is worth mentioning that choosing $\lambda_1 \to 1$ or $\lambda_2 \to 1$ in our Theorem 3 recovers the point-to-point results of [2] along the two axes, and selecting $\lambda_3 \to 1$ recovers (a significant part of) the dominant face of the achievable region of [8].

IV. Numerical Example

In this section, we illustrates our results through the example of the "real adder" DM-MAC $Y = X_1 + X_2 + Z$ that takes the real addition of binary inputs X_1, X_2 and Bernoulli($\frac{1}{2}$) noise Z, leading to a quaternary output Y.

Figure 1 depicts, in addition to the classical capacity region, the achievable region of Theorem 3 for $n = 200$ and $\epsilon = 10^{-3}$. For each valid selection of the parameters $\lambda_1, \lambda_2, \lambda_3$, a pentagon is obtained and taking the union over all such choices gives rise to a convex hull, that is, a curved shape.

Figure 2 compares our achievable rate region in Theorem 3 for blocklengths $n = 200, 300, 500, 1000, 5000$ and $\epsilon = 10^{-3}$ with that of Tan and Kosut [8], the genie-aided single-user outer bound implied by [2], and the classical capacity region. Both achievable regions in the finite blocklength regime operate quite close to the outer bound and both have smooth shapes with no sharp corners, but as blocklength grows, they approach the well-known pentagon shape of the capacity region. Furthermore, for this example, our region appears to achieve much of the region of [8], except for a very slight gap at the blunt "corners" of the region which shrinks as blocklength grows. In terms of numerical evaluation, our result takes at least one order of magnitude less time than that of [8] for the same resolution.

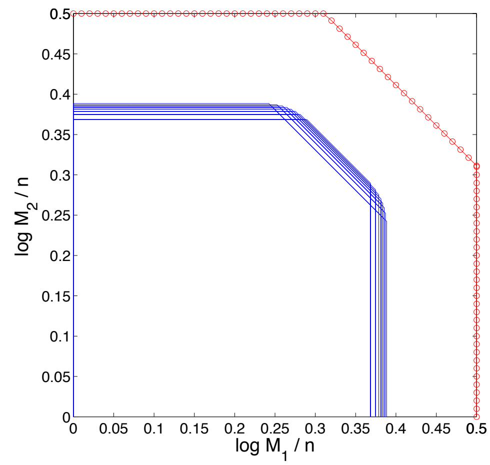

Fig. 1. Achievable region of Theorem 3 for the real adder DM-MAC $Y = X_1 + X_2 + Z$ with $n = 200$ and $\epsilon = 10^{-3}$. The full union is approximated by taking several values of $\lambda_1, \lambda_2, \lambda_3$ with steps of 0.1. The capacity region $(n \to \infty)$ is also shown.

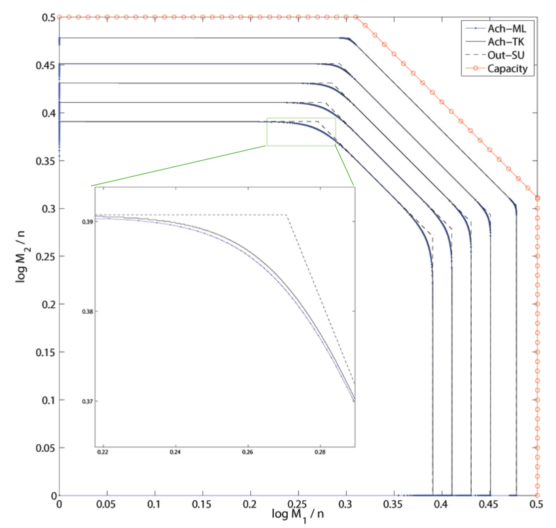

Fig. 2. Achievable region of Theorem 3 for the real adder DM-MAC $Y = X_1 + X_2 + Z$ with $n = 200, 300, 500, 1000, 5000$ and $\epsilon = 10^{-3}$. Also shown for comparison are: the result of Tan-Kosut [8], the single-user outer bound of [2], and the capacity region $(n \to \infty)$.

V. Proof of Theorems 1 and 2

Here, we summarize the proof of the DT bound for DM-MAC. The proof uses the coded time sharing method of [5] with the usual random coding technique, i.e., proving that the average error probability over the ensemble of all codes generated at random satisfies the DT bound, thus concluding the existence of a code with the described rate and error probability performance. The decoding rule, however, is likelihood ratio test (LRT) decoding [2], which can be considered as a generalization of joint typicality decoding that evaluates the partial (conditional) and full (unconditional) dependence of the output on the input codeword pair under test.

A time sharing realization t^n is generated according to $\prod_{l=1}^n p(t_l)$ and revealed to both the receiver and the two transmitters. Then, M_1 codewords $X_1^n(j)$, $j\in\mathcal{M}_1$, and M_2 codewords $X_2^n(k)$, $k\in\mathcal{M}_2$, all of length n are generated independently according to $\prod_{l=1}^n p(x_{1l}|t_l)$ and $\prod_{l=1}^n p(x_{2l}|t_l)$, respectively. These two codebooks are also revealed to the receiver and both transmitters. Given the time sharing realization t^n, the codebook pair $\{x_1^n(j)\}_{j=1}^{M_1}\times\{x_2^n(k)\}_{k=1}^{M_2}$, and the channel output y^n, the decoder runs for all M_1M_2 codeword pairs the following three LRTs

$$Z_{j,k}^{(1)}(y^n)=1\{i\left(x_1^n(j);y^n|x_2^n(k)t^n\right)>\log\gamma_1\left(x_1^n(j),x_2^n(k)\right)\},$$
$$Z_{j,k}^{(2)}(y^n)=1\{i\left(x_2^n(j);y^n|x_1^n(k)t^n\right)>\log\gamma_2\left(x_1^n(j),x_2^n(k)\right)\},$$
$$Z_{j,k}^{(3)}(y^n)=1\{i\left(x_1^n(j)x_2^n(k);y^n|t^n\right)>\log\gamma_3\left(x_1^n(j),x_2^n(k)\right)\},$$

choosing the first pair (j,k) for which $Z_{j,k}^{(1)}(y^n)=Z_{j,k}^{(2)}(y^n)=Z_{j,k}^{(3)}(y^n)=1$, where "first" is defined as a row-by-row search, so that $(m,p)<(j,k)$ iff either $m<j$ or $m=j,p<k$. The error probability for message pair (j,k) is thus given by

$$\begin{aligned}\epsilon_{j,k}=&\Pr\Big[\Big\{Z_{j,k}^{(1)}(Y^n)=0\cup Z_{j,k}^{(2)}(Y^n)=0\cup Z_{j,k}^{(3)}(Y^n)=0\Big\}\\&\bigcup_{(m,p)<(j,k)}\Big\{Z_{m,p}^{(1)}(Y^n)=Z_{m,p}^{(2)}(Y^n)=Z_{m,p}^{(3)}(Y^n)=1\Big\}\Big|A_{j,k,t^n}\Big]\\\le&\sum_{s=1}^{3}\Pr\Big[Z_{j,k}^{(s)}(Y^n)=0\Big|A_{j,k,t^n}\Big]+\sum_{m<j,p=k}\Pr\Big[Z_{m,k}^{(1)}(Y^n)=1\Big|A_{j,k,t^n}\Big]\\&+\sum_{p<k,m=j}\Pr\Big[Z_{j,p}^{(2)}(Y^n)=1\Big|A_{j,k,t^n}\Big]+\sum_{m<j,p\neq k}\Pr\Big[Z_{m,p}^{(3)}(Y^n)=1\Big|A_{j,k,t^n}\Big]\end{aligned}$$

where the notation A_{j,k,t^n} in the conditioning is a shorthand for the event $\{X_1^n=x_1^n(j),X_2^n=x_2^n(k),T^n=t^n\}$, and we have used the definition of the LRT's and the union bound. Not using the union bound for the first three summands and leaving them as a joint outage event, while keeping the rest of proof unchanged, will result in the generalized bound (2).

Now, since the time sharing sequence is generated i.i.d according to distribution $p(t)$ and all the codewords in the first, resp. second, codebook are generated independently according to the distribution $p(x_1|t)$, resp. $p(x_2|t)$, we can take the average of the above inequality over all possible time sharing realizations and the ensemble of all codebook pairs.

$$\begin{aligned}\mathbb{E}[\epsilon_{j,k}]\le&\Pr\left[i\left(X_1^n;Y^n|X_2^nT^n\right)\le\log\gamma_1\left(X_1^n,X_2^n\right)\right]\\&+(j-1)\Pr\left[i\left(X_1^n;\bar{Y}_2^n|X_2^nT^n\right)>\log\gamma_1\left(X_1^n,X_2^n\right)\right]\\&+\Pr\left[i\left(X_2^n;Y^n|X_1^nT^n\right)\le\log\gamma_2\left(X_1^n,X_2^n\right)\right]\\&+(k-1)\Pr\left[i\left(X_2^n;\bar{Y}_1^n|X_1^nT^n\right)>\log\gamma_2\left(X_1^n,X_2^n\right)\right]\\&+\Pr\left[i\left(X_1^nX_2^n;Y^n|T^n\right)\le\log\gamma_3\left(X_1^n,X_2^n\right)\right]\\&+(j-1)(M_2-1)\Pr\left[i\left(X_1^nX_2^n;\bar{Y}_0^n|T^n\right)>\log\gamma_3\left(X_1^n,X_2^n\right)\right].\end{aligned}$$

Averaging this over all message pairs (j,k) gives the DT bound (1). To conclude the proof of Theorem 1, it is sufficient to observe that each line on its RHS is a weighted sum of two types of error in a Bayesian binary hypothesis test, and therefore corresponds to average error probability of the test. Then, it is known that the optimal test is an LRT (as we have used) with the optimal threshold equal to the ratio of priors or simply the ratio of the coefficients of the two error probabilities in each test. ■

VI. Proof of Theorem 3

Here, we sketch how Theorem 1 is used for proving Theorem 3. We basically expand each of the three mutual information RVs in the DT bound (1) as sums of i.i.d. RVs and use the Central Limit Theorem, or more specifically the Berry-Esseen Theorem, to calculate the associated outage and confusion probabilities, analogous to [2].

Upon fixing the distribution $p(t)p(x_1|t)p(x_2|t)$, the output Y^n of the DM-MAC $p(y|x_1,x_2)$ corresponding to the inputs X_1^n and X_2^n and the time sharing variable T^n described in the DT bound above satisfies

$$\begin{aligned}i\left(X_1^n;Y^n|X_2^nT^n\right)&=\sum_{l=1}^{n}\log\frac{p(Y_l|X_{1l}X_{2l}T_l)}{p(Y_l|X_{2l}T_l)}\triangleq\sum_{l=1}^{n}i_{1l},\\i\left(X_2^n;Y^n|X_1^nT^n\right)&=\sum_{l=1}^{n}\log\frac{p(Y_l|X_{1l}X_{2l}T_l)}{p(Y_l|X_{1l}T_l)}\triangleq\sum_{l=1}^{n}i_{2l},\\i\left(X_1^nX_2^n;Y^n|T^n\right)&=\sum_{l=1}^{n}\log\frac{p(Y_l|X_{1l}X_{2l}T_l)}{p(Y_l|T_l)}\triangleq\sum_{l=1}^{n}i_{3l},\end{aligned}$$

where for all $l=1,...,n$,

$$i_{1l}\sim i\left(X_1;Y|X_2T\right),\qquad i_{2l}\sim i\left(X_2;Y|X_1T\right),$$
$$i_{3l}\sim i\left(X_1X_2;Y|T\right),$$

and their mean and variance can be described in terms of the corresponding average mutual information and dispersion terms as

$$\begin{aligned}\mathbb{E}[i_{1l}]&=\mathbb{I}(X_1;Y|X_2T),&\mathrm{Var}[i_{1l}]&=\mathbb{V}(X_1;Y|X_2T),\\\mathbb{E}[i_{2l}]&=\mathbb{I}(X_2;Y|X_1T),&\mathrm{Var}[i_{2l}]&=\mathbb{V}(X_2;Y|X_1T),\\\mathbb{E}[i_{3l}]&=\mathbb{I}(X_1X_2;Y|T),&\mathrm{Var}[i_{3l}]&=\mathbb{V}(X_1X_2;Y|T).\end{aligned}$$

Assume that all three dispersions are strictly positive. In such a case, we obtain using the Berry-Esseen Theorem that

$$\Pr\left[i\left(X_1^n;Y^n|X_2^nT^n\right)\le\log\gamma_1\right]=\Pr\left[\sum_{l=1}^{n}i_{1l}\le\log\gamma_1\right]$$
$$\le Q\left(\frac{n\mathbb{I}(X_1;Y|X_2T)-\log\gamma_1}{\sqrt{n\mathbb{V}(X_1;Y|X_2T)}}\right)+\frac{B_{11}}{\sqrt{n}}, \tag{4}$$

where the constant $B_{11}\triangleq\frac{6S[i(X_1;Y|X_2T)]}{\mathbb{V}(X_1;Y|X_2T)^{3/2}}$, with $S[\cdot]$ being the third moment operator, represents the Berry-Esseen gap to the Gaussian distribution. On the other hand, we can use a change of measure technique as in [2]

$$Q\left[\frac{dP}{dQ}>\gamma\right]=\int 1\left\{\frac{dP}{dQ}>\gamma\right\}dQ=\int\left(\frac{dP}{dQ}\right)^{-1}1\left\{\frac{dP}{dQ}>\gamma\right\}dP \tag{5}$$

to obtain

$$\Pr\left[i\left(X_1^n;\bar{Y}_2^n|X_2^nT^n\right)>\log\gamma_1\right]$$
$$=\mathbb{E}\left[\exp\left\{-\sum_{l=1}^{n}i_{1l}\right\}1\left\{\sum_{l=1}^{n}i_{1l}>\log\gamma_1\right\}\right]\le\frac{B_{12}}{\sqrt{n}}\gamma_1^{-1}, \tag{6}$$

where (6) is according to [2, Lemma 47].

By substituting (4) and (6) and the analogous bounds for the other two mutual information RVs into the DT bound (1) with the optimal selection for thresholds $\gamma_1,\gamma_2,\gamma_3$, we obtain

$$\epsilon\le Q\left(\frac{n\mathbb{I}(X_1;Y|X_2T)-\log\frac{M_1-1}{2}}{\sqrt{n\mathbb{V}(X_1;Y|X_2T)}}\right)+\frac{B_1}{\sqrt{n}}$$
$$+Q\left(\frac{n\mathbb{I}(X_2;Y|X_1T)-\log\frac{M_2-1}{2}}{\sqrt{n\mathbb{V}(X_2;Y|X_1T)}}\right)+\frac{B_2}{\sqrt{n}}$$
$$+Q\left(\frac{n\mathbb{I}(X_1X_2;Y|T)-\log\frac{(M_1-1)(M_2-1)}{2}}{\sqrt{n\mathbb{V}(X_1X_2;Y|T)}}\right)+\frac{B_3}{\sqrt{n}},$$

where $B_1=B_{11}+B_{12}$ and analogously for B_2 and B_3. Now, splitting ϵ among the three first terms of each line gives

$$\log M_1\le n\mathbb{I}(X_1;Y|X_2T)-\sqrt{n\mathbb{V}(X_1;Y|X_2T)}Q^{-1}\left(\lambda_1\epsilon-\frac{B_1}{\sqrt{n}}\right)$$
$$\log M_2\le n\mathbb{I}(X_2;Y|X_1T)-\sqrt{n\mathbb{V}(X_2;Y|X_1T)}Q^{-1}\left(\lambda_2\epsilon-\frac{B_2}{\sqrt{n}}\right)$$
$$\log M_1+\log M_2\le n\mathbb{I}(X_1X_2;Y|T)$$
$$-\sqrt{n\mathbb{V}(X_1X_2;Y|T)}Q^{-1}\left(\lambda_3\epsilon-\frac{B_3}{\sqrt{n}}\right) \tag{7}$$

where positive constants $\lambda_1,\lambda_2,\lambda_3$ that sum up to 1 can be arbitrarily chosen to represent the weight of each of the three types of outage for communication over a DM-MAC with average error probability ϵ. We can further simplify the bounds in (7) using Taylor's expansion $Q^{-1}(\lambda\epsilon-\frac{B}{\sqrt{n}})\ge Q^{-1}(\lambda\epsilon)+\tilde{B}/\sqrt{n}$ and the fact that dispersions are finite over the set of distributions $p(t)p(x_1|t)p(x_2|t)$, so $\mathbb{V}\le V_{\max}$ and $\tilde{B}\le\tilde{B}_{\max}$, to obtain the bounds of Theorem 3.

In the case that one or more of the dispersion terms are zero, we directly evaluate the corresponding probabilities, using the fact that a mutual information RV with zero dispersion (variance) is concentrated almost surely at the average mutual information (mean). For example, if the first dispersion is zero, $\mathbb{V}(X_1;Y|X_2T)=0$, then with the optimal choice of threshold

$$\log\gamma_1=\log\frac{M_1-1}{2}=n\mathbb{I}(X_1;Y|X_2T)+\underbrace{\log(\lambda_1\epsilon)}_{\le 0},$$

the first two summands on the RHS of the DT bound (1) can be evaluated using a change of measure technique as in (5).

$$\Pr\left[i\left(X_1^n;Y^n|X_2^nT^n\right)\le\log\gamma_1\right]$$
$$+\frac{M_1-1}{2}\Pr\left[i\left(X_1^n;\bar{Y}_2^n|X_2^nT^n\right)>\log\gamma_1\right]$$
$$=0+\frac{M_1-1}{2}\exp\{-n\mathbb{I}(X_1;Y|X_2T)\}\cdot 1=\lambda_1\epsilon.$$

Now, notice that the achievable rate is simply

$$\log M_1<n\mathbb{I}(X_1;Y|X_2T)-Q^{-1}(\lambda_1\epsilon)\underbrace{\sqrt{n\mathbb{V}(X_1,Y|X_2T)}}_{=0}+O(1),$$

as in (3). This concludes the proof of Theorem 3. ■

VII. Conclusion

We have proved a simple achievable rate region for DM-MAC in the regime of finite blocklength by splitting the allowed average error probability among several "outage" events, in which the channel cannot support the target rates of a subset of the users. This region appears to have a curved blunt shape in general and implies rate penalties with respect to the infinite blocklength regime that depend on the allowed error probability, the chosen finite blocklength, and the dispersions of the DM-MAC, i.e., the variances of the relevant mutual information RVs. We have observed that our achievable rate region covers a significant portion of the concurrent result in [8], while its numerical computation is much easier. We aim as a future direction to prove the tightness of an outer bound we have developed and compare it with the existing inner bounds.

References

[1] V. Strassen, "Asymptotische Abschatzungen in Shannon's Informationstheorie," in *Transaction of the 3rd Prague Conference on Information Theory*, Prague, pp. 689–723, 1962.

[2] Y. Polyanskiy, H. V. Poor and S. Verdú, "Channel Coding Rate in the Finite Blocklength Regime," *IEEE Transactions on Information Theory*, vol. 56, no. 5, pp. 2307–2359, May 2010.

[3] M. Hayashi, "Information Spectrum Approach to Second-Order Coding Rate in Channel Coding," *IEEE Transactions on Information Theory*, vol. 55, no. 11, pp. 4947–4966, Nov. 2009.

[4] T.-S. Han, *Information-Spectrum Methods in Information Theory*, Springer-Verlag, Berlin, Germany, 2003.

[5] A. El Gamal and Y.-H. Kim, *Network Information Theory*, Cambridge University Press, UK, 2012.

[6] G. Dueck, "The Strong Converse to the Coding Theorem for the Multiple–Access Channel," *J. Comb. Inform. Syst. Sci.*, vol. 6, pp. 187–196, 1981.

[7] R. Ahlswede, "An Elementary Proof of the Strong Converse Theorem for the Multiple Access Channel," *J. Comb. Inform. Syst. Sci.*, vol. 7, no. 3, pp. 216–230, 1982.

[8] V. Y. F. Tan and O. Kosut, "On the Dispersion of Three Network Information Theory Problems," submitted to *IEEE Transactions on Information Theory*, Available at http://arxiv.org/abs/1201.3901, Jan 2012.