On the Dispersions of the Discrete Memoryless Interference Channel

Sy-Quoc Le[1], Vincent Y. F. Tan[1,2], and Mehul Motani[1]

[1]Department of ECE, National University of Singapore, Singapore
[2]Institute for Infocomm Research (I^{2}R), A*STAR, Singapore

***Abstract*—In this work, achievable dispersions for the discrete memoryless interference channel (DM-IC) are derived. In other words, we characterize the backoff from the Han-Kobayashi (HK) achievable region, the largest inner bound known to date for the DM-IC. In addition, we also characterize the backoff from Sato's region in the strictly very strong interference regime, and the backoff from Costa and El Gamal's region in the strong interference regime. To do so, Feinstein's lemma is first generalized to be applicable to the interference channel. Making use of the generalized Feinstein's lemma, it is found that the dispersions for the DM-IC can be represented by the information variances of eight information densities when HK message splitting is involved, and of six information densities for another encoding strategy. We also derive an outer bound that leverages on a known dispersion result for channels with random state by Ingber-Feder. It is shown that for the strictly very strong interference regime, the inner and outer bound have similar algebraic forms.**

I. INTRODUCTION

In information theory, the capacity is defined as the maximum rate at which reliable communication is admissible for an arbitrarily small probability of error, provided the code blocklength can grow without bound. However, practical codes (such as LDPC, turbo, polar codes) operate at finite blocklengths. Thus, it is important to assess the penalizing deviation from the channel capacity required to maintain the error probability below the desired value at a given finite blocklength. Strassen [1] was the first to address this question for the discrete memoryless (DM) point-to-point channel. In other words, Strassen derived an asymptotic expansion for the logarithm of the maximal size of length-n codes with probability of error $\epsilon \in (0,1)$. Recently, the interest in this topic has been revived by the works [2] and [3]. It has been shown that the maximal codebook size $M^*(n,\epsilon)$, for the DM point-to-point channel, is approximated by the expression

$$\log M^*(n,\epsilon) = nC - \sqrt{nV}Q^{-1}(\epsilon) + O\left(\log n\right), \qquad (1)$$

where C is the *capacity* of the channel, V is the conditional information variance evaluated at the capacity-achieving input distribution (assuming it is unique), and $Q(\cdot)$ is the tail probability of the standard normal distribution. The conditional information variance is also known as the *dispersion* [3]. Thus, $\sqrt{V/n}Q^{-1}(\epsilon)$ is roughly the penalty in terms of rate that one has to pay for operating at a finite blocklength, for the case $\epsilon < \frac{1}{2}$. The quantity $\sqrt{V}Q^{-1}(\epsilon)$ is equivalent to the second-order coding rate in [2]. Recently, the class of DM point-to-point channel for which the third-order rate is known has been enlarged [4].

While the dispersion for the point-to-point channel is fairly well-understood, the understanding of dispersions for multi-user settings is still limited. Achievable dispersions, which are the dispersions for achievable regions, for the DM multiple-access channel have been characterized recently [5]–[7]. The authors in [5] also characterized achievable dispersions for the broadcast channel with degraded message sets and Slepian-Wolf coding.

In this work, we characterize achievable dispersions for the discrete memoryless interference channel (DM-IC). We find that the dispersions can be expressed as information variances of eight information densities when Han-Kobayashi (HK) message splitting is used, and of six such quantities for another encoding strategy. These information densities will be clearly defined below. Furthermore, we define a class of channels called *strictly very strong interference channels*, in which, we show that a general outer bound, using results from [8], has similar algebraic form to our inner bound.

II. SYSTEM MODEL AND PROBLEM FORMULATION

The discrete[1] memoryless interference channel (DM-IC) consists of two finite input alphabets $\mathcal{X}_1$ and $\mathcal{X}_2$, two finite output alphabets $\mathcal{Y}_1$ and $\mathcal{Y}_2$, and a channel transitional probability $p_{Y_1Y_2|X_1X_2}(y_1y_2|x_1x_2)$, whose n-th extension satisfies

$$p_{Y_1^nY_2^n|X_1^nX_2^n}(y_1^ny_2^n|x_1^nx_2^n) = \prod_{k=1}^{n} p_{Y_1Y_2|X_1X_2}(y_{1k}y_{2k}|x_{1k}x_{2k}).$$

Transmitter $j \in \{1,2\}$, wishes to communicate a message $w_j \in \mathcal{W}_j = \{1,2,\ldots,M_j\}$ to receiver $j \in \{1,2\}$, where $\mathcal{W}_j$ are message sources. Messages m_1 and m_2 from each message set $\mathcal{W}_1$ and $\mathcal{W}_2$ respectively are equally probable. It is assumed that W_1 and W_2 are statistically independent. Let $0 < \epsilon_1, \epsilon_2 < 1$ be fixed constants. An $(M_1, M_2, n, \epsilon_1, \epsilon_2)$ *code for the DM-IC* consists of encoding functions

$$f_j : \mathcal{W}_j \to \mathcal{X}_j^n \text{ for } j = 1,2; \qquad (2)$$

and two decoding functions

$$d_j : \mathcal{Y}_j^n \to \mathcal{W}_j \text{ for } j = 1,2; \qquad (3)$$

The work of M. Motani and S.-Q. Le was supported in part by the National University of Singapore under Research Grant WBS R-263-000-579-112. E-mail: {le.sy.quoc, motani}@nus.edu.sg

The work of V. Y. F. Tan was supported by A*STAR, Singapore. Email:vtan@nus.edu.sg

[1]The assumption of discreteness is not restrictive, and is made to ensure that some moments are positive and finite.

such that $p_{e,1} \leq \epsilon_1$ and $p_{e,2} \leq \epsilon_2$, where $p_{e,1}$ and $p_{e,2}$ denote the average error probabilities, $p_{e,j} := \Pr(\hat{W}_j \neq W_j)$.

Definition 1. *A rate pair (R_1, R_2) is $(n, \epsilon_1, \epsilon_2)$*-achievable *for the DM-IC if there exists an $(M_1, M_2, n, \epsilon_1, \epsilon_2)$- code for the channel such that $\frac{1}{n}\log(M_1) \geq R_1$ and $\frac{1}{n}\log(M_2) \geq R_2$. The $(n, \epsilon_1, \epsilon_2)$*-capacity region *of the DM-IC is defined as the set of all $(n, \epsilon_1, \epsilon_2)$-achievable rate pairs.*

First, we make a few important definitions. A well-known achievability technique in the interference channel is HK message splitting in which messages are split into private and common messages. To do so, auxiliary random variables (RVs) U, U_1 and U_2 are introduced, and U is a time-sharing RV.

Definition 2. *In the DM-IC, fix a joint distribution*

$$p(u)p(u_1|u)p(u_2|u)p(x_1|u_1u)p(x_2|u_2u)p(y_1y_2|x_1x_2). \quad (4)$$

Define the information densities

$$i_{11} = i(X_1;Y_1|U_1U_2U) = \log\frac{p(y_1|x_1u_1u_2u)}{p(y_1|u_1u_2u)} \quad (5)$$

$$i_{12} = i(U_2X_1;Y_1|U_1U) = \log\frac{p(y_1|x_1u_1u_2u)}{p(y_1|u_1u)} \quad (6)$$

$$i_{13} = i(X_1;Y_1|U_2U) = \log\frac{p(y_1|x_1u_2u)}{p(y_1|u_2u)} \quad (7)$$

$$i_{14} = i(U_2X_1;Y_1|U) = \log\frac{p(y_1|x_1u_2u)}{p(y_1|u)}. \quad (8)$$

Similarly, we define $i(X_2;Y_2|U_1U_2U)$, $i(U_1X_2;Y_2|U_2U)$, $i(X_2;Y_2|U_1U)$ and $i(U_1X_2;Y_2|U)$ and denote their shortened notations as i_{21}, i_{22}, i_{23} and i_{24}.

Remark: When we want to emphasize the dependence of these quantities on the various RVs, we shall use the second set of notations, e.g., $i(X_1;Y_1|U_1U_2U)$. For the sake of notational convenience and space, we shall often use the first set of notations, e.g., i_{11}.

Denote the mean of i_{11} as I_{11}, and observe that

$$\mathbb{E}[i(X_1;Y_1|U_1U_2U)] = I(X_1;Y_1|U_1U_2U) = I_{11}. \quad (9)$$

Similarly, we can find expression for the expectations of i_{jk} and denote them as I_{jk}, for $j = 1, 2$, and $k = 1, 2, 3, 4$. Correspondingly, we denote the variances of i_{jk} as V_{jk}.

Definition 3. *A DM-IC is said to have* strictly very strong interference *if $I(X_1;Y_1|X_2) < I(X_1;Y_2)$ and $I(X_2;Y_2|X_1) < I(X_2;Y_1)$ for all non-deterministic $p(x_1)p(x_2)$.*

Example 1. Consider a Gaussian IC where $Y_1 = g_{11}X_1 + g_{21}X_2 + Z_1$, $Y_2 = g_{12}X_1 + g_{22}X_2 + Z_2$, where X_1 and X_2 are Gaussian inputs with zero means and unit variances, Z_1 and Z_2 are independent Gaussian noises with zero means and unit variances, and X_j is also independent of Z_j for $j = 1$ and 2. In this channel, treating X_2 as noise, we have $I(X_1;Y_2) = h(Y_2) - h(Y_2|X_1) = 0.5\log(\frac{g_{12}^2}{g_{22}^2+1} + 1)$. $I(X_1;Y_1|X_2) = I(X_1;g_{11}X_1 + Z_1|X_2) = I(X_1;g_{11}X_1 + Z_1) = 0.5\log(g_{11}^2 + 1)$. Similarly, we can find $I(X_2;Y_2|X_1)$ and $I(X_2;Y_1)$. Thus, this Gaussian IC is in the strictly very strong interference regime iff $g_{11}^2 < g_{12}^2/(1 + g_{22}^2)$ and $g_{22}^2 < g_{21}^2/(1 + g_{11}^2)$ and Gaussian inputs are used. This is an extension of [9, Rmk 6.2].

Example 2. Consider a discrete IC where $p(y_1, y_2|x_1, x_2) = p(y_1|x_1, x_2)p(y_2|x_1, x_2)$ and in addition, $p(y_1|x_1, x_2) = p(y_1|x_2)$ and $p(y_2|x_1, x_2) = p(y_2|x_1)$. Thus, Y_1 is independent of X_1 and Y_2 is independent of X_2. Clearly, $I(X_1;Y_1|X_2) = 0$ and $I(X_1;Y_2) > 0$ for every non-deterministic $p(x_1)$. The same is true for the other inequality.

Example 3. Consider a DM-IC where each alphabet is $\mathbb{F}_2$, $Y_1 = (G_{11} \cdot X_1) \oplus (G_{21} \cdot X_2)$ and $Y_2 = (G_{12} \cdot X_1) \oplus (G_{22} \cdot X_2)$. Let $G_{ij} \sim \mathrm{Ber}(q_{ij})$ for some $q_{ij} \in (0, 1)$. Assume that $q_{11} = q_{22}$, $q_{21} = q_{12} = 1 - q_{11}$ and $q_{11} < \delta$ for some $\delta > 0$. $p(x_i) = \mathrm{Ber}(\alpha_i)$, where $\alpha_i \in [\alpha_i^*, 1 - \alpha_i^*]$, for $\alpha_i^* \in (0, 0.5)$ and $i = 1, 2$. We can verify that for arbitrarily small (α_1^*, α_2^*), there is δ, sufficiently small depending on (α_1^*, α_2^*), such that $I(X_1;Y_1|X_2) < I(X_1;Y_2)$. Intuitively, for small q_{ii}, Y_i is almost independent of X_i, for $i = 1, 2$. Also see [10].

Definition 4. *A DM-IC is said to have* strong interference *if $I(X_1;Y_1|X_2) \leq I(X_1;Y_2|X_2)$ and $I(X_2;Y_2|X_1) \leq I(X_2;Y_1|X_1)$ for all $p(x_1)p(x_2)$.*

Alternatively to Definition 2, we can define information densities differently when HK message splitting is not involved. Although doing so will lead to a smaller achievable rate region, such information densities play an important role in the strictly very strong interference regime and the strong interference regime.

Definition 5. *In the DM-IC, fix a joint distribution*

$$p(u)p(x_1|u)p(x_2|u)p(y_1y_2|x_1x_2). \quad (10)$$

Define the information densities

$$i_{11s} = \log\frac{p(y_1|x_1x_2u)}{p(y_1|x_2u)}, \quad (11)$$

$$i_{12s} = \log\frac{p(y_1|x_1x_2u)}{p(y_1|u)} \quad (12)$$

$$i_{13s} = \log\frac{p(y_1|x_1x_2u)}{p(y_1|x_1u)}. \quad (13)$$

Similarly, we define i_{21s}, i_{22s} and i_{23s}.

Denote the expectations and the variances of i_{jks} as I_{jks} and V_{jks} respectively, for $j \in \{1, 2\}$, and $k \in \{1, 2, 3\}$.

We define the following real-valued function, which will be used often. For a given n, define the function f as

$$f(a, b, c) = a - \sqrt{\frac{b}{n}}Q^{-1}(c) + O\left(\frac{\log n}{n}\right). \quad (14)$$

We are going to prove that a $(n, \epsilon_1, \epsilon_2)$-achievable rate region of the DM-IC can be characterized by the first-order and second-order statistics of either the information spectrum quantities defined in Definition 2 or that in Definition 5.

III. Main results for general DM-IC and discussion

A. Main results

The main results of this paper are captured by the following theorems.

Theorem 1. *For any joint distribution satisfying (4), any ϵ_1, ϵ_2 and a sufficiently large blocklength n, there exists a $(2^{n(R_{1c}+R_{1p})}, 2^{n(R_{2c}+R_{2p})}, n, \epsilon_1, \epsilon_2)$ code for the DM-IC satisfying*

$$R_{1p} \le f(I_{11}, V_{11}, \lambda_{11}\epsilon_1) \tag{15}$$
$$R_{1p} + R_{2c} \le f(I_{12}, V_{12}, \lambda_{12}\epsilon_1) \tag{16}$$
$$R_{1p} + R_{1c} \le f(I_{13}, V_{13}, \lambda_{13}\epsilon_1) \tag{17}$$
$$R_{1p} + R_{1c} + R_{2c} \le f(I_{14}, V_{14}, \lambda_{14}\epsilon_1) \tag{18}$$
$$R_{2p} \le f(I_{21}, V_{21}, \lambda_{21}\epsilon_2) \tag{19}$$
$$R_{2p} + R_{1c} \le f(I_{22}, V_{22}, \lambda_{22}\epsilon_2) \tag{20}$$
$$R_{2p} + R_{2c} \le f(I_{23}, V_{23}, \lambda_{23}\epsilon_2) \tag{21}$$
$$R_{2p} + R_{2c} + R_{1c} \le f(I_{24}, V_{24}, \lambda_{24}\epsilon_2), \tag{22}$$

where λ_{jk} are some positive constants satisfying the constraints $\sum_{k=1}^{4} \lambda_{jk} = 1$, and $V_{jk} > 0$, for $j \in \{1, 2\}$, and $k \in \{1, 2, 3, 4\}$.

An outline of the proof for Theorem 1 is presented in Section V.

After applying Fourier-Motzkin elimination process to Theorem 1, we obtain the following theorem.

Theorem 2. *For any joint distribution satisfying (4), the $(n, \epsilon_1, \epsilon_2)$-capacity region of the DM-IC includes the set of all non-negative pairs (R_1, R_2) satisfying*

$$R_1 \le f(I_{13}, V_{13}, \lambda_{13}\epsilon_1) \tag{23}$$
$$R_2 \le f(I_{23}, V_{23}, \lambda_{23}\epsilon_2) \tag{24}$$
$$R_1 + R_2 \le f(I_{14}, V_{14}, \lambda_{14}\epsilon_1) + f(I_{21}, V_{21}, \lambda_{21}\epsilon_2) \tag{25}$$
$$R_1 + R_2 \le f(I_{11}, V_{11}, \lambda_{11}\epsilon_1) + f(I_{24}, V_{24}, \lambda_{24}\epsilon_2) \tag{26}$$
$$R_1 + R_2 \le f(I_{12}, V_{12}, \lambda_{12}\epsilon_1) + f(I_{22}, V_{22}, \lambda_{22}\epsilon_2) \tag{27}$$
$$2R_1 + R_2 \le f(I_{14}, V_{14}, \lambda_{14}\epsilon_1) + f(I_{11}, V_{11}, \lambda_{11}\epsilon_1) + f(I_{22}, V_{22}, \lambda_{22}\epsilon_2) \tag{28}$$
$$R_1 + 2R_2 \le f(I_{24}, V_{24}, \lambda_{24}\epsilon_2) + f(I_{21}, V_{21}, \lambda_{21}\epsilon_2) + f(I_{12}, V_{12}, \lambda_{12}\epsilon_1), \tag{29}$$

where λ_{jk} are positive constants, satisfying the constraints $\sum_{k=1}^{4} \lambda_{jk} = 1$, and $V_{jk} > 0$, for $j \in \{1, 2\}$, and $k \in \{1, 2, 3, 4\}$.

When HK message splitting is involved, a $(n, \epsilon_1, \epsilon_2)$-achievable rate region is characterized in Theorems 1 and 2. Alternatively, a $(n, \epsilon_1, \epsilon_2)$-achievable rate region is characterized in the following theorem.

Theorem 3. *For any joint distribution as in (10), any ϵ_1, ϵ_2 and a sufficiently large blocklength n, there exists a $(2^{nR_1}, 2^{nR_2}, n, \epsilon_1, \epsilon_2)$-code for the DM-IC satifying*

$$R_1 \le f(I_{11s}, V_{11s}, \lambda_{11}\epsilon_1) \tag{30}$$
$$R_1 \le f(I_{23s}, V_{23s}, \lambda_{23}\epsilon_2) \tag{31}$$
$$R_2 \le f(I_{21s}, V_{21s}, \lambda_{21}\epsilon_2) \tag{32}$$
$$R_2 \le f(I_{13s}, V_{13s}, \lambda_{13}\epsilon_1) \tag{33}$$
$$R_1 + R_2 \le f(I_{12s}, V_{12s}, \lambda_{12}\epsilon_1) \tag{34}$$
$$R_1 + R_2 \le f(I_{22s}, V_{22s}, \lambda_{22}\epsilon_2), \tag{35}$$

where λ_{jk} are some positive constants satisfying the constraints $\sum_{j=1}^{3} \lambda_{jk} = 1$, $V_{jk} > 0$, for $j \in \{1, 2\}$, and $k \in \{1, 2, 3\}$.

The proof of Theorem 3 is similar to that of Theorem 1.

B. Discussion

To obtain an inner bound on the $(n, \epsilon_1, \epsilon_2)$-capacity region in Theorems 1 and 2, we made use of HK message splitting [11]. As the blocklength grows unbounded, we recover HK's best achievable rate region [11], [12]. From Theorem 1, we have an interesting observation that there are penalizing deviations from both the common information rates R_{jc} and the private information rates R_{jp}, for $j \in \{1, 2\}$, for operating in the finite blocklength setting with average error probabilities (ϵ_1, ϵ_2). Consequently, in Theorem 2, we observe penalizing deviations from not just the individual rates R_1 and R_2, but also from the weighted sum rates. Another interesting observation is that the penalizing deviations from weighted sum rates depend on both error probabilities ϵ_1 and ϵ_2.

The way we prove Theorem 1 is to make use a non-asymptotic bound presented in Lemma 6. In fact, a stronger version of Lemma 6 is the following bound

$$\begin{aligned}\epsilon_1 \le \inf_{\beta>0} \Big\{ &\Pr[\{i(X_1^n; Y_1^n | U_1^n U_2^n U^n) \le n(R_{1p} + \beta)\} \\ &\cup \{i(U_2^n X_1^n; Y_1^n | U_1^n U^n) \le n(R_{1p} + R_{2c} + \beta)\} \\ &\cup \{i(X_1^n; Y_1^n | U_2^n U^n) \le n(R_{1p} + R_{1c} + \beta)\} \\ &\cup \{i(U_2^n X_1^n; Y_1^n | U^n) \le n(R_{1p} + R_{1c} + R_{2c} + \beta)\}] \\ &+ 2^{-n\beta} \times 4\Big\},\end{aligned} \tag{36}$$

where the distribution of $U^n U_1^n U_2^n X_1^n X_2^n Y_1^n$ is n-th i.i.d. extension of that given in (4). By applying the multi-dimensional Berry-Esséen Theorem [13] to the first term, we can alternatively characterize the achievable dispersions as covariance matrices $\mathbf{V}_1$ and $\mathbf{V}_2$ defined below. The achievable region in terms of private and common information rates, as in Theorem 1, is alternatively given by

$$\mathbf{R}_1 \in \mathbf{I}_1 - \frac{1}{\sqrt{n}} \mathcal{Q}^{-1}(\epsilon_1; \mathbf{V}_1) + O\left(\frac{\log n}{n}\right) \mathbf{1}_4 \tag{37}$$

where

$$\mathbf{R}_1 = [R_{1p}, R_{1p}+R_{2c}, R_{1p}+R_{1c}, R_{1p}+R_{1c}+R_{2c}]^T, \tag{38}$$
$$\mathbf{I}_1 = [I_{11},\ I_{12},\ I_{13},\ I_{14}]^T, \tag{39}$$
$$\mathcal{Q}^{-1}(\epsilon_1; \mathbf{V}_1) = \{\mathbf{z} \in \mathbb{R}^4 | \Pr(\mathcal{N}(\mathbf{0}, \mathbf{V}_1) \le \mathbf{z}) \ge 1 - \epsilon\}, \tag{40}$$
$$\mathbf{V}_1 = \mathrm{Cov}\left([i_{11},\ i_{12},\ i_{13},\ i_{14}]^T\right). \tag{41}$$

We can similarly define $\mathbf{R}_2$, $\mathbf{I}_2$, and $\mathbf{V}_2$ to obtain a vector version of the second half in Theorem 1. It can be shown the $(n, \epsilon_1, \epsilon_2)$-achievable region using the covariance matrices $\mathbf{V}_1$ and $\mathbf{V}_2$, as in (37) will be larger than that in Theorem 1.

In order to obtain the analogue of Theorem 2 and to express the rate region in terms of (R_1, R_2), we have to perform the analogue of Fourier-Motzkin elimination. However, this is not as simple as that in Theorem 2 because the achievable rates as prescribed by (37) are not defined in terms of a polytope. Hence, a linear projection operation of the region given by (37) is required. We omit the details here.

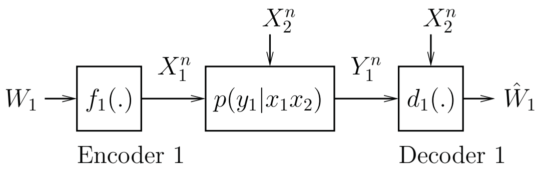

Fig. 1. Channel with X_2^n given to decoder 1.

In point-to-point communication, the second-order statistic defined by only one information density is enough to characterize the dispersion. In the multiple-access channel, the second-order statistics of three information densities are needed to do so. In the DM-IC, we need the second-order statistics of either eight information densities, in the case where HK message splitting is involved, or six such quantities otherwise. This is intuitively acceptable as there are more terminals communicating with one another.

In fact, by not requiring that each receiver correctly decodes the message for the other receiver, (31) and (33) are redundant. The achievable rate region formed by remaining equations in Theorem 3 is reduced to that of the interference channel in the *strong interference* regime as in [14], when $n \to \infty$.

Using a similar argument as in [15], a two-user interference channel (IC) may be converted into a multiple-access channel (MAC) by joining its two outputs into a single product output $Y = Y_1 \times Y_2$. Since this MAC has less restriction in decoding as in the IC, where X_j is decoded with only knowledge of Y_j; thus an outer bound to the new MAC is an outer bound to the IC. Therefore, the outer bounds in [7] are applicable to our problem as well. In fact we present a tighter outer bound that captures the independence of the inputs in Theorem 5.

IV. The strictly very strong interference regime

In the *strictly very strong interference* regime, (31), (33), (34) and (35) can be shown to be redundant when n is sufficiently large. We obtain the following corollary (to Theorem 3) which yields an inner bound to the DM-IC.

Corollary 4. *The $(n, \epsilon_1, \epsilon_2)$-achievable rate region of the DM-IC in the strictly very strong interference regime includes the set of non-negative pairs (R_1, R_2) satisfying*

$$R_1 \le f(I_{11s}, V_{11s}, \epsilon_1) \tag{42}$$
$$R_2 \le f(I_{21s}, V_{21s}, \epsilon_2). \tag{43}$$

for some joint distribution as in (10).

As $n \to \infty$, this region reduces to the asymptotic achievable region of the very strong interference regime as in [16]. In fact, (42) and (43) are algebraically similar to an outer bound, derived in the following theorem, to the $(n, \epsilon_1, \epsilon_2)$-capacity region of any DM-IC.

Theorem 5. *The $(n, \epsilon_1, \epsilon_2)$-capacity region of any DM-IC is included in the set of non-negative pairs (R_1, R_2) satisfying*

$$R_1 \le \max_{p(x_1), p(x_2)} f(\tilde{I}_{11s}, \tilde{V}_{11s}, \epsilon_1) \tag{44}$$
$$R_2 \le \max_{p(x_1), p(x_2)} f(\tilde{I}_{21s}, \tilde{V}_{21s}, \epsilon_2), \tag{45}$$

where $\tilde{I}_{11s}, \tilde{I}_{21s}, \tilde{V}_{11s}, \tilde{V}_{21s}$ are the mutual information and dispersions with respect to the information densities defined in Definition 5 with the time-sharing variable $U = \emptyset$. More precisely,

$$\tilde{V}_{11s} = \mathbb{E}\left[V(W_{X_2})\right] + \mathrm{Var}\left[C(W_{X_2})\right] \tag{46}$$

where $C(W)$ and $V(W)$ are the capacity and the dispersion of channel W and $W_{x_2}(y_1|x_1) := p_{Y_1|X_1X_2}(y_1|x_1x_2)$.

Remark: A few comments are in order with regard to Corollary 4 and Theorem 5. Note that the mutual information and dispersions for the inner bound in (42) and (43) are computed with respect to a *common* input distributions $p(u)p(x_1|u)p(x_2|u)$. However, for the outer bound in (44) and (45), the mutual information and dispersions are, in general, computed with respect to *different* input distributions $p(x_1), p(x_2)$. This is the reason why the outer bound includes the inner bound. Note that unlike in [7] for the multiple-access channel, the maximizations in (44) and (45) are over product distributions, capturing the intrinsic nature of the DM-IC where the channel inputs X_1^n and X_2^n must be independent. Even though the inner and outer bounds prescribed by the previous two results are not matching, the algebraic forms of the bounds are very similar.

Proof: Any decoder at receiver 1 cannot do better than a genie-aided system where X_2^n (in addition to Y_1^n) is given to receiver 1. See Fig. 1. Thus, any outer bound on R_1 applicable to the system in Fig. 1 is also applicable to the standard DM-IC. This system can be interpreted as a DM channel with state, with input X_1^n, output (Y_1^n, X_2^n), where X_2^n is interpreted as the random channel state. The dispersion of this channel model was derived by Ingber-Feder [8] and is given by (46). Note that as in the channel with state problem where the state is only known to the receiver [9, Eq. (7.2)], the state is independent of the codeword. This corresponds to the situation here. The other bound is proved similarly. ■

V. Outline of proof for Theorem 1

Before proving Theorem 1, we generalize Feinstein's lemma [17, Lemma 3.4.1] for the point-to-point channel to be applicable to the DM-IC.

Lemma 6. *Consider the DM-IC defined in Theorem 1. Fix a joint distribution satisfying (4). For any $n \in \mathbb{N}$ and an arbitrary positive constant β, there exists a $(2^{n(R_{1c}+R_{1p})}, 2^{n(R_{2c}+R_{2p})}, n, \epsilon_1, \epsilon_2)$-code such that*

$$\begin{aligned}\epsilon_1 \le &\Pr\left[\frac{1}{n} i(X_1^n; Y_1^n | U_1^n U_2^n U^n) \le R_{1p} + \beta\right] \\ &+ \Pr\left[\frac{1}{n} i(U_2^n X_1^n; Y_1^n | U_1^n U^n) \le R_{1p} + R_{2c} + \beta\right] \\ &+ \Pr\left[\frac{1}{n} i(X_1^n; Y_1^n | U_2^n U^n) \le R_{1p} + R_{1c} + \beta\right] \\ &+ \Pr\left[\frac{1}{n} i(U_2^n X_1^n; Y_1^n | U^n) \le R_{1p} + R_{1c} + R_{2c} + \beta\right] \\ &+ 2^{-n\beta} \times 4,\end{aligned} \tag{47}$$

and the average probability of error ϵ_2 at receiver 2 is similarly upper-bounded.

The proof of Lemma 6 is similar to the proof of Lemma 7.10.1 in Han's book [17].

Remark: Lemma 6 holds for any blocklength n, even when $n = 1$. There is no restriction on the structures of $(U^n U_1^n U_2^n X_1^n X_2^n)$ and the channel $p_{Y_1^n Y_2^n | X_1^n X_2^n}$. Thus, if the structures are chosen carefully (e.g., stationary, memoryless), we can potentially obtain useful results.

With Lemma 6, we are ready to prove Theorem 1. First we choose distributions for various RVs in Lemma 6. For $j \in \{1, 2\}$, define the product distributions

$$p(u^n) = \prod_{k=1}^{n} p(u_k), \tag{48}$$

$$p(u_j^n | u^n) = \prod_{k=1}^{n} p(u_{jk} | u_k), \tag{49}$$

$$p(x_j^n | u_j^n, u^n) = \prod_{k=1}^{n} p(x_{jk} | u_{jk}, u_k). \tag{50}$$

Due to DM property of the channel and the construction of the codebooks in (48)–(50), each of the information densities $i(X_1^n; Y_1^n | U_1^n U_2^n U^n)$, $i(U_2^n X_1^n; Y_1^n | U_1^n U^n)$, $i(X_1^n; Y_1^n | U_2^n U^n)$ and $i(U_2^n X_1^n; Y_1^n | U^n)$ are sums of n i.i.d. RVs. Applying Berry-Esséen Theorem [18, Theorem 2, Chapter XVI] to the first term in (47), we have

$$\begin{aligned} &\Pr\left[\frac{1}{n} i(X_1^n; Y_1^n | U_1^n U_2^n U^n) \le R_{1p} + \beta\right] \\ &\le Q\left(\frac{nI(X_1; Y_1 | U_1 U_2 U) - nR_{1p} - n\beta}{\sqrt{nV_{11}}}\right) + \frac{B_{11a}}{\sqrt{n}}, \end{aligned} \tag{51}$$

where $B_{11a} = \frac{6T_{11}}{V_{11}^{3/2}}$ with T_{11} is the third moment of $i(X_1; Y_1 | U_1 U_2 U)$.

Applying similar bounding techniques to other terms in (47), we obtain

$$\begin{aligned} \epsilon_1 \le\ & 4 \times 2^{-n\beta} + Q\left(\frac{nI_{11} - nR_{1p} - n\beta}{\sqrt{nV_{11}}}\right) + \frac{B_{11a}}{\sqrt{n}} \\ & + Q\left(\frac{nI_{12} - n(R_{1p} + R_{2c}) - n\beta}{\sqrt{nV_{12}}}\right) + \frac{B_{12a}}{\sqrt{n}} \\ & + Q\left(\frac{nI_{13} - n(R_{1p} + R_{1c}) - n\beta}{\sqrt{nV_{13}}}\right) + \frac{B_{13a}}{\sqrt{n}} \\ & + Q\left(\frac{nI_{14} - n(R_{1p} + R_{1c} + R_{2c}) - n\beta}{\sqrt{nV_{14}}}\right) + \frac{B_{14a}}{\sqrt{n}}, \end{aligned} \tag{52}$$

where $B_{1ka} := 6T_{1k}/V_{1k}^{3/2}$ with T_{1k} is the third absolute moment of i_{1k}, for $k = 1, 2, 3$ and 4. Note that the third moments are finite and the variances are positive because all alphabets are discrete.

By splitting the error ϵ_1 and choose $\beta = \frac{\log n}{n}$, we obtain the first part of Theorem 1. The second half of the theorem is proved similarly to the above.

Remark: The generation of the time-sharing variable u^n as in (48) is worse than fixing u^n to be a non-random sequence whose type is $O(\frac{1}{n})$-close to $p(u)$; cf. [7]. In [7], an achievable dispersion for the MAC is obtained and it was shown that the strategy contained therein is better than the generation of u^n in (48) because the conditional variance is no larger than the unconditional variance. However, we here generate u^n according to (48) for simplicity.

VI. Conclusion

In this work, we characterized achievable dispersions for the 2-user DM-IC. To do so, we first generalized Feinstein's lemma to be applicable to the interference channel. Next, using the generalized Feinstein's lemma, we showed that the dispersions are given by information variances of a set of information densities. The first- and second-order statistics of the information densities are shown to characterize an inner bound on the $(n, \epsilon_1, \epsilon_2)$-capacity region of the DM-IC, where n is a given finite (but large) blocklength, and ϵ_1 and ϵ_2 are desired (or tolerable) average probabilities of error at receiver 1 and 2 respectively. We also derived a general converse bound, applicable for any DM-IC, which has a similar algebraic form to our inner bound in the strictly very strong interference regime.

References

[1] V. Strassen. Asymptotische abschatzungen in shannon's informationstheorie. *Trans. Third Prague Conf. Information Theory*, pages 689–723, 1962.

[2] M. Hayashi. Information spectrum approach to second-order coding rate in channel coding. *IEEE Transactions on Information Theory*, 55(11):4947–4966, Nov. 2009.

[3] Y. Polyanskiy, H. V. Poor, and S. Verdu. Channel coding rate in the finite blocklength regime. *IEEE Transactions on Information Theory*, 56(5):2307–2359, May. 2010.

[4] M. Tomamichel and V. Y. F. Tan. A tight upper bound for the third order asymptotics of discrete memoryless channel. *Arxiv:1212.3689v1*, Dec. 2012.

[5] V. Y. F. Tan and O. Kosut. On the dispersions of three network information theory problems. *Arxiv:1201.3901v2*, Feb. 2012.

[6] E. MolavianJazi and J. N. Laneman. Discrete memoryless multiple access channel in the finite blocklength regime. In *Proc. IEEE International Symposium on Information Theory*, pages 36–40, Cambridge, MA, 2012.

[7] Y.-W. Huang and P. Moulin. Finite blocklength coding for multiple access channels. In *Proc. IEEE International Symposium on Information Theory*, pages 836–840, Cambridge, MA, 2012.

[8] A. Ingber and M. Feder. Finite blocklength coding for channels with side information at the receiver. In *Convention of Electrical and Electronics Engineers in Israel (IEEEI)*, 2010.

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

[10] H. Sato. On the capacity region of a discrete two-user channel under strong interference. *IEEE Transactions on Information Theory*, 24(3):377–379, May 1978.

[11] T. S. Han and K. Kobayashi. A new achievable rate region for the interference channel. *IEEE Transactions on Information Theory*, 27(1):49–60, Jan. 1981.

[12] H. F. Chong, M. Motani, H. K. Garg, and H. El Gamal. On the Han-Kobayashi region for interference channel. *IEEE Transactions on Information Theory*, 54(7):3188–3195, Jul. 2008.

[13] V. Bentkus. On the dependence of the berry esseen bound on dimension. *Journal of Statistical Planning and Inference*, 113:385–402, 2003.

[14] M. H. M. Costa and A. El Gamal. The capacity region of the discrete memoryless interference channel with strong interference. *IEEE Transactions on Information Theory*, 28:343–346, Mar. 1982.

[15] H. Sato. Two-user communication channels. *IEEE Transactions on Information Theory*, 23(3):295–304, 1977.

[16] H. Sato. The capacity of the Gausian interference channel under strong interference. *IEEE Transactions on Information Theory*, 27(6):786–788, Nov. 1981.

[17] T. S. Han. *Information-Spectrum Methods in Information Theory*. Springer Berlin Heidelberg, 1^{st} edition, 2003.

[18] W. Feller. *An Introduction to Probability Theory and Its application*, volume II. John Wiley and Sons, 2nd edition, 1971.