Second-Order Asymptotics for the Gaussian Interference Channel with Strictly Very Strong Interference

Sy-Quoc Le, Vincent Y. F. Tan, and Mehul Motani

Department of Electrical and Computer Engineering, National University of Singapore

***Abstract*—The second-order asymptotics of the Gaussian interference channel in the strictly very strong interference regime are considered. The rates of convergence to a given point on the boundary of the (first-order) capacity region are determined. These rates are expressed in terms of the average probability of error and variances of selected modified information densities which coincide with the dispersion of the (single-user) Gaussian channel. Interestingly, under the strictly very strong interference assumption, the intuition that receivers can decode messages from non-intended transmitters carries over to the second-order analysis.**

I. Introduction

Recently, the issue of second-order coding rates has become increasingly prominent in network information theory. Strassen [1], Hayashi [2], and Polyanskiy, Poor and Verdú [3] characterized the second-order capacity of the discrete memoryless (DM) channel. However, it is not trivial to generalize this result from single-user to multi-user settings. Initial efforts focused on global achievable dispersions for the DM multiple-access channel (MAC), for the DM asymmetric broadcast channel [4], and for the DM interference channel (IC) [5]. However, as pointed out by Haim et al. [6], global dispersion analysis has certain drawbacks such as the failure to precisely capture the nature of convergence to the boundary of the capacity region, the inability in characterizing the deviation from a specific point on the boundary and the difficulty in getting tight second-order results. To overcome these weaknesses, the same authors [6] proposed "local" dispersion analysis. Tan-Kosut [4] and Nomura-Han [7] characterized the second-order capacity region for distributed source coding, i.e. the Slepian-Wolf problem [8]. While it is possible to obtain tight second-order converse bounds for distributed source coding, it is challenging to do so for channel coding problems such as MAC (due in part to the union over input distributions). Scarlett-Tan [9] obtained the second-order capacity region for the Gaussian MAC with degraded message sets. The degradedness of the message sets makes it possible to avoid certain difficulties in getting a tight converse. It is an open problem to characterize the local second-order capacity region for the Gaussian MAC with non-degraded message sets.

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 and an NUS startup grant WBS R-263-000-A98-750. Email:vtan@nus.edu.sg

In this paper, we study the local dispersions of the Gaussian IC in the strictly very strong (VS) interference regime. More precisely, we characterize the $(\kappa_1, \kappa_2, \epsilon)$-second-order capacity region (which will be defined shortly) for the channel in terms of the average probability of error ϵ and the variances of selected modified information densities. We focus on the interesting case where (κ_1, κ_2) lies on the boundary of the capacity region. The converse is proved with the help of a generalized version of Verdú-Han lemma [10], which involves only two error events. The direct part is proved with the help of a generalized Feinstein lemma [11], which involves four error events. The condition of being in the strictly VS interference regime reduces the number of error events involved in the direct part, thus allowing the converse to match the direct part.

We emphasize that apart from Scarlett-Tan's work [9], this is the only work that completely characterizes the local dispersions for a channel-type network information theory problem. In addition, this is the only work that completely characterizes the local dispersions for a channel-type network information theory problem where input distributions are independent.

II. System model and problem formulation

The two-user Gaussian interference channel (IC) is defined by the following input-output relationships

$$Y_{1i} = h_{11}X_{1i} + h_{21}X_{2i} + Z_{1i}, \tag{1}$$

$$Y_{2i} = h_{12}X_{1i} + h_{22}X_{2i} + Z_{2i}, \tag{2}$$

where X_{ji} denotes the signal sent by transmitter j (Tx_j), Y_{ji} denotes the output at receiver j, for $j = 1, 2$, at time i, for $i \in \{1, 2, ..., n\}$, and $\{Z_{ji}\}_{i=1}^n$ are independent, additive white Gaussian noise processes with zero means and unit variances. Denote the transitional probability $P_{Y_1^n Y_2^n | X_1^n X_2^n}(y_1^n y_2^n | x_1^n x_2^n)$ as $W^n(y_1^n y_2^n | x_1^n x_2^n)$. Denote the Y_1- and Y_2-marginals of W as W_1 and W_2 respectively. The forward channel gains $\{h_{11}, h_{21}, h_{12}, h_{22}\}$ are assumed to be positive constants and known at all terminals. Tx_j, for $j = 1, 2$, wishes to communicate a message $S_j \in \{1, 2, ..., M_j\} = \mathcal{W}_j$ to receiver Rx_j. It is assumed that S_1 and S_2 are independent, and uniformly distributed. We use nats as the units of information.

Define $\mathcal{F}_{jn} \triangleq \{x_j^n \in \mathcal{X}_j^n | \sum_{k=1}^n x_{jk}^2 \leq nP_j\}$ for positive powers $P_j, j = 1, 2$. An $(M_{1n}, M_{2n}, n, \epsilon_n, P_1, P_2)$-code for the IC consists of two encoding functions $f_{jn} : \mathcal{W}_j \to \mathcal{F}_{jn}$ and two decoding functions $g_{jn} : \mathcal{Y}_j^n \to \mathcal{W}_j$ for $j = 1, 2$, where the average probability of error is given as $\epsilon_n \triangleq \Pr(\hat{S}_1 \neq S_1 \text{ or } \hat{S}_2 \neq S_2)$, and P_1 and P_2 are the average power constraints at the transmitters.

In the spirit in the works [2], [4], [7], [9], we define the second-order capacity region as follows.

Definition 1. Fix any two non-negative numbers κ_1 and κ_2. A pair (L_1, L_2) is said to be $(\kappa_1, \kappa_2, \epsilon)$-*achievable* [1] if there exists a sequence of $(M_{1n}, M_{2n}, n, \epsilon_n, P_1, P_2)$-codes such that

$$\limsup_{n\to\infty} \epsilon_n \le \epsilon, \text{ and } \liminf_{n\to\infty} \frac{1}{\sqrt{n}}(\log M_{jn} - n\kappa_j) \ge L_j \quad (3)$$

for $j = 1, 2$. The $(\kappa_1, \kappa_2, \epsilon)$-*second-order capacity region* of the IC $\mathcal{L}(\kappa_1, \kappa_2, \epsilon) \subset \mathbb{R}^2$ is defined as the closure of the set of all $(\kappa_1, \kappa_2, \epsilon)$-achievable rate pairs (L_1, L_2).

Definition 2. The IC is said to have *VS interference* if

$$h_{22}^2 \le \frac{h_{21}^2}{1 + h_{11}^2 P_1} \text{ and } h_{11}^2 \le \frac{h_{12}^2}{1 + h_{22}^2 P_2}. \quad (4)$$

The IC is said to have *strictly VS interference* if both inequalities in (4) are strict.

Definition 3. Define the *Gaussian capacity function* $\mathsf{C}(z) \triangleq \frac{1}{2}\log(1+z)$ and the following first-order quantities:

$$I_{11} \triangleq \mathsf{C}(h_{11}^2 P_1), \ I_{12} \triangleq \mathsf{C}(h_{11}^2 P_1 + h_{21}^2 P_2), \quad (5)$$
$$I_{21} \triangleq \mathsf{C}(h_{22}^2 P_2), \ I_{22} \triangleq \mathsf{C}(h_{22}^2 P_2 + h_{12}^2 P_1), \quad (6)$$
$$\mathbf{I}_{\mathrm{c}} \triangleq [I_{11}\, I_{21}]^T, \quad \mathbf{I}_{\mathrm{d}} \triangleq [I_{11}\, I_{12}\, I_{22}\, I_{21}]^T. \quad (7)$$

The vectors $\mathbf{I}_{\mathrm{c}}$ and $\mathbf{I}_{\mathrm{d}}$ characterize the first-order regions obtained in the converse and direct parts respectively.

Carleial [12] proved that the capacity region $\mathcal{C}$ of the Gaussian IC with VS interference is given by

$$\mathcal{C} = \{(R_1, R_2) \in \mathbb{R}_+^2 \mid R_1 \le I_{11}, \ R_2 \le I_{21}\}. \quad (8)$$

A certain set of information densities plays an important role in the IC [5], [13], [14]. However, in dealing with channels with cost constraints, modified information densities [2], [15] offer certain advantages. In that spirit, we define the following modified information densities.

Definition 4. Fix a joint distribution

$$P_{X_1^n}(x_1^n) P_{X_2^n}(x_2^n) W_1^n(y_1^n|x_1^n x_2^n) W_2^n(y_2^n|x_1^n x_2^n). \quad (9)$$

Given two auxiliary (conditional) output distributions $Q_{Y_1^n|X_2^n}$ and $Q_{Y_1^n}$, define the modified information densities

$$\tilde{i}_{11}^n(X_1^n X_2^n Y_1^n) \triangleq \log \frac{W_1^n(Y_1^n|X_1^n X_2^n)}{Q_{Y_1^n|X_2^n}(Y_1^n|X_2^n)}, \quad (10)$$

$$\tilde{i}_{12}^n(X_1^n X_2^n Y_1^n) \triangleq \log \frac{W_1^n(Y_1^n|X_1^n X_2^n)}{Q_{Y_1^n}(Y_1^n)}. \quad (11)$$

We will often use the shorthands $\tilde{i}_{11}^n$ and $\tilde{i}_{12}^n$.

Similarly, given two auxiliary (conditional) output distributions $Q_{Y_2^n|X_1^n}$ and $Q_{Y_2^n}$, we define $\tilde{i}_{21}^n(X_1^n X_2^n Y_2^n)$ and $\tilde{i}_{22}^n(X_1^n X_2^n Y_2^n)$.

[1]We note that it is more precise to define a pair being $(P_1, P_2, \kappa_1, \kappa_2, \epsilon)$-achievable. However, we omit the dependence on (P_1, P_2) as (P_1, P_2) are fixed throughout the paper.

In addition, we define

$$\tilde{\mathbf{i}}_{\mathrm{c}}^n(X_1^n X_2^n Y_1^n Y_2^n) \triangleq [\tilde{i}_{11}^n \quad \tilde{i}_{21}^n]^T \quad (12)$$
$$\tilde{\mathbf{i}}_{\mathrm{d}}^n(X_1^n X_2^n Y_1^n Y_2^n) \triangleq [\tilde{i}_{11}^n \quad \tilde{i}_{12}^n \quad \tilde{i}_{22}^n \quad \tilde{i}_{21}^n]^T. \quad (13)$$

Definition 5. Define the *Gaussian cross-dispersion function* $\mathsf{V}(u,v) \triangleq \frac{u(v+2)}{2(u+1)(v+1)}$ and the *Gaussian dispersion function* $\mathsf{V}(u) \triangleq \mathsf{V}(u,u)$. Define the second-order quantities

$$V_1 \triangleq \mathsf{V}(h_{11}^2 P_1), \quad \text{and} \quad V_2 \triangleq \mathsf{V}(h_{22}^2 P_2). \quad (14)$$

Definition 6. Let $\Phi(z) \triangleq \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-t^2/2}\, dt$ be the Gaussian cumulative distribution function (cdf).

III. Main Result

Theorem 1. *For any $0 < \epsilon < 1$, the $(\kappa_1, \kappa_2, \epsilon)$-second-order capacity region in the following special cases is given by:*

i) When $\kappa_1 = I_{11}$ and $\kappa_2 < I_{21}$ (vertical boundary),

$$\mathcal{L}(\kappa_1, \kappa_2, \epsilon) = \left\{ (L_1, L_2) \middle| \Phi\Big(\frac{L_1}{\sqrt{V_1}}\Big) \le \epsilon \right\}. \quad (15)$$

ii) When $\kappa_1 = I_{11}$ and $\kappa_2 = I_{21}$ (corner point),

$$\mathcal{L}(\kappa_1, \kappa_2, \epsilon) = \left\{ (L_1, L_2) \middle| 1 - \epsilon \le \Phi\Big(-\frac{L_1}{\sqrt{V_1}}\Big)\Phi\Big(-\frac{L_2}{\sqrt{V_2}}\Big) \right\}. \quad (16)$$

iii) When $\kappa_1 < I_{11}$ and $\kappa_2 = I_{21}$ (horizontal boundary),

$$\mathcal{L}(\kappa_1, \kappa_2, \epsilon) = \left\{ (L_1, L_2) \middle| \Phi\Big(\frac{L_2}{\sqrt{V_2}}\Big) \le \epsilon \right\}. \quad (17)$$

Proof: This theorem is proved in the appendix. ■

The result is applicable to any $(\kappa_1, \kappa_2) \in \mathbb{R}_+^2$. If (κ_1, κ_2) is inside the $\mathcal{C}$, then $\mathcal{L}(\kappa_1, \kappa_2, \epsilon) = \mathbb{R}^2$. If (κ_1, κ_2) is outside the $\mathcal{C}$, then $\mathcal{L}(\kappa_1, \kappa_2, \epsilon) = \emptyset$. This implies the strong converse. The only interesting cases are presented in Thm. 1.

In case (i), the $(\kappa_1, \kappa_2, \epsilon)$-capacity region depends on ϵ and V_1 only. This is because $\kappa_2 < I_{21}$ which implies that Tx_2 operates at a rate strictly below the optimal first-order rate I_{21}. Thus, the second channel operates in large-deviations (error exponents) regime so the second constraint is not featured in dispersion analysis as the error probability is exponentially small. See [4], [6], [7], [9]. By symmetry, case (iii) is similar to case (i). In case (ii), the $(\kappa_1, \kappa_2, \epsilon)$-second-order capacity region is a function of ϵ and *both* V_1 and V_2 because we are operating near the corner point of $\mathcal{C}$ and the two constraints on the rates come into play in the characterization of $\mathcal{L}(\kappa_1, \kappa_2, \epsilon)$.

One of the input distributions that achieves the capacity, dispersion and even the third-order coding rate of the Gaussian point-to-point channel [3], [16] is the uniform distribution on the power sphere. MolavianJazi-Laneman [15] obtained global achievable dispersions for the Gaussian MAC using uniform distributions on power spheres. In this work, we also use the *uniform input distributions on power spheres*. It is not easy to use the *cost constrained ensemble* in [9] as that input distribution is more suited to superposition coding.

The converse makes use of a generalized version of the Verdú-Han lemma [10], which involves only two error events. The proof of the direct part makes use of a generalized version of Feinstein's lemma [11], which involves four error events.

The strictly VS interference condition is needed to reduce the number of error events in the direct part, so that it matches the converse. For ICs in the VS interference regime [12], the intuition is that each receiver can reliably decode information from the non-intended transmitter. Interestingly, this intuition carries over to second-order analysis with the caveat that the interference must be *strictly* very strong. Indeed, in the proof of the generalized Verdú-Han lemma which we use in our converse, we assume a genie provides the codeword of Tx_2 to Rx_1 and vice versa. This is partly reflected in case (ii) in (16), where the second-order rate L_1 is constrained by V_1 but not V_2, and vice versa. It is also remarkable that the form of (16) suggests that the two error events $\{\hat{S}_j \neq S_j\}$ for $j = 1, 2$ are independent when clearly, interference still exists.

Finally, it is somewhat surprising that in the converse, even though we must ensure that the transmitter outputs are independent, we do not need to use the wringing technique, invented by Ahlswede [17] and used originally to prove that the DM-MAC admits a strong converse. This is due to Gaussianity which allows us to show that the first- and second-order statistics of a certain set of information densities are independent of x_1^n and x_2^n on power spheres. See (28)-(29).

IV. Appendix

A. Supporting lemmas

The following lemma is a generalized version of Feinstein's lemma [11], which is used in the direct proof of Theorem 1.

Lemma 1. *Fix a joint distribution satisfying (9). For any $n \in \mathbb{N}$, any $\gamma > 0$, and any auxiliary (conditional) output distributions $Q_{Y_1^n|X_2^n}$, $Q_{Y_1^n}$, $Q_{Y_2^n|X_1^n}$ and $Q_{Y_2^n}$, there exists an $(M_{1n}, M_{2n}, n, \epsilon_n, P_1, P_2)$-code for the Gaussian IC, such that*

$$\epsilon_n \leq \Pr(\mathcal{E}_{11} \cup \mathcal{E}_{12} \cup \mathcal{E}_{21} \cup \mathcal{E}_{22}) + K e^{-n\gamma} + P_{X_1^n}(\mathcal{F}_{1n}^c) + P_{X_2^n}(\mathcal{F}_{2n}^c) \quad (18)$$

where $\mathcal{E}_{j1} \triangleq \{\tilde{i}_{j1}^n \leq \log M_{jn} + n\gamma\}$ and $\mathcal{E}_{j2} \triangleq \{\tilde{i}_{j2}^n \leq \log M_{1n}M_{2n} + n\gamma\}$ for $j = 1, 2$, and

$$K \triangleq K_{11} + K_{12} + K_{21} + K_{22}, \quad (19)$$

$$K_{11} \triangleq \sup_{x_2^n, y_1^n} \frac{P_{Y_1^n|X_2^n}(y_1^n|x_2^n)}{Q_{Y_1^n|X_2^n}(y_1^n|x_2^n)},\ K_{12} \triangleq \sup_{y_1^n} \frac{P_{Y_1^n}(y_1^n)}{Q_{Y_1^n}(y_1^n)}, \quad (20)$$

$$K_{21} \triangleq \sup_{x_1^n, y_2^n} \frac{P_{Y_2^n|X_1^n}(y_2^n|x_1^n)}{Q_{Y_2^n|X_1^n}(y_2^n|x_1^n)},\ K_{22} \triangleq \sup_{y_2^n} \frac{P_{Y_2^n}(y_2^n)}{Q_{Y_2^n}(y_2^n)}. \quad (21)$$

The following lemma is a generalized version of Verdú-Han lemma [10]. It is useful in the converse proof of Theorem 1.

Lemma 2. *For every $n \in \mathbb{N}$, for every $\gamma > 0$, and for any auxiliary (conditional) output distributions $Q_{Y_1^n|X_2^n}$ and $Q_{Y_2^n|X_1^n}$, every $(M_{1n}, M_{2n}, n, \epsilon_n, P_1, P_2)$-code for the Gaussian IC satisfies*

$$\epsilon_n \geq \Pr(\tilde{i}_{11}^n \leq \log M_{1n} - n\gamma \text{ or } \tilde{i}_{21}^n \leq \log M_{2n} - n\gamma) - 2e^{-n\gamma}, \quad (22)$$

where X_j^n is uniformly distributed over the j-th codebook.

For proofs of Lemma 1 and Lemma 2, refer to [18].

B. Proof of Theorem 1: Converse Part

Define the auxiliary (conditional) output distributions

$$\hat{Q}_{Y_1}(y_1) \triangleq N(0, h_{11}^2 P_1 + h_{12}^2 P_2 + 1) \quad (23)$$

$$\hat{Q}_{Y_2}(y_2) \triangleq N(0, h_{12}^2 P_1 + h_{22}^2 P_2 + 1) \quad (24)$$

$$\hat{Q}_{Y_1|X_2}(y_1|x_2) \triangleq N(h_{21}x_2, h_{11}^2 P_1 + 1) \quad (25)$$

$$\hat{Q}_{Y_2|X_1}(y_2|x_1) \triangleq N(h_{12}x_1, h_{22}^2 P_2 + 1), \quad (26)$$

which are the distributions at the output of the IC when the inputs are $X_1 \sim N(0, P_1)$ and $X_2 \sim N(0, P_2)$.

Fix any pair of rates (κ_1, κ_2) on the boundary of $\mathcal{C}$ in (8). Consider any second-order pair (L_1, L_2) that is $(\kappa_1, \kappa_2, \epsilon)$-achievable for the Gaussian IC. This implies that there exists a sequence of $(M_{1n}, M_{2n}, n, \epsilon_n, P_1, P_2)$-codes satisfying (3).

By the definition of $\liminf$, for any $\beta > 0$, there exists an integer N_β such that for all $n > N_\beta$

$$\log M_{jn} - n\kappa_j \geq \sqrt{n}(L_j - \beta). \quad (27)$$

Let $\mathcal{L}_{\mathrm{eq}}(\kappa_1, \kappa_2, \epsilon)$ be the $(\kappa_1, \kappa_2, \epsilon)$-second-order capacity region of the IC with equal power constraints, i.e. each codeword x_j^n satisfies $\sum_{k=1}^n x_{jk}^2 = nP_j$ for $j = 1, 2$. It can be shown that (cf. [3, Lem. 39]) $\mathcal{L}_{\mathrm{eq}}(\kappa_1, \kappa_2, \epsilon) = \mathcal{L}(\kappa_1, \kappa_2, \epsilon)$. Therefore, in this converse proof, it is sufficient to assume equal power constraints.

Choose the output distribution $Q_{Y_1^n}$ in Lemma 2 as the n-fold product of $\hat{Q}_{Y_1}(y_1)$, defined in (23), i.e. $Q_{Y_{1k}} = \hat{Q}_{Y_1}$ for all $k = 1, \ldots, n$. Similarly, choose the (conditional) output distributions $Q_{Y_2^n}$, $Q_{Y_1^n|X_2^n}$ and $Q_{Y_2^n|X_1^n}$ in Lemma 2, as the n-fold products of $\hat{Q}_{Y_2}(y_2)$, $\hat{Q}_{Y_1|X_2}(y_1|x_2)$ and $\hat{Q}_{Y_2|X_1}(y_2|x_1)$ respectively. Next, choose $\gamma = \frac{\log n}{2n}$. Let V_{c} be the 2×2 diagonal matrix with V_1 and V_2 along its diagonals. It can be shown (cf. [9]) that for all x_1^n and x_2^n (satisfying $\|x_j^n\|^2 = nP_j$),

$$\mathbb{E}\left[\frac{1}{\sqrt{n}} \sum_{k=1}^n \tilde{\mathbf{i}}_{ck}(x_{1k} x_{2k} Y_{1k} Y_{2k})\right] = \sqrt{n}\mathbf{I}_{\mathrm{c}} \quad (28)$$

$$\mathrm{cov}\left[\frac{1}{\sqrt{n}} \sum_{k=1}^n \tilde{\mathbf{i}}_{ck}(x_{1k} x_{2k} Y_{1k} Y_{2k})\right] = V_{\mathrm{c}}. \quad (29)$$

Let $t_{\mathrm{c}} \triangleq \frac{1}{n}\sum_{k=1}^n \mathbb{E}[\|\tilde{\mathbf{i}}_{\mathrm{c}k}(x_{1k}x_{2k}Y_{1k}Y_{2k})\|^3]$ be the third absolute moment and $\phi_{\mathrm{c}} \triangleq \frac{254\sqrt{2}t_{\mathrm{c}}}{\lambda_{\min}(V_{\mathrm{c}})^{3/2}}$. Define the rate pair $\mathbf{R}_{\mathrm{c}} \triangleq [\frac{\log M_{1n}}{n}, \frac{\log M_{2n}}{n}]^T$. Note that $V_{\mathrm{c}} \succ 0$ because the channel gains and powers are all positive. Also $t_{\mathrm{c}} < \infty$ from [9, App. A]. Thus, ϕ_{c} is finite. Now we define $\Psi(\mathbf{t}; \mathbf{m}, \Sigma) \triangleq \int_{-\infty}^{t_1} \ldots \int_{-\infty}^{t_k} N(\mathbf{z}; \mathbf{m}, \Sigma)\, d\mathbf{z}$ to be the generalization of the Gaussian cdf to the multivariate setting. Then we have

$$\begin{aligned} A(x_1^n, x_2^n) &\triangleq \Pr\left(\frac{1}{n}\sum_{k=1}^n \tilde{\mathbf{i}}_{\mathrm{c}k}(x_{1k}x_{2k}Y_{1k}Y_{2k}) > \mathbf{R}_{\mathrm{c}} - \gamma\mathbf{1}\right) \\ &= \Pr\left(\frac{1}{\sqrt{n}}\sum_{k=1}^n \tilde{\mathbf{i}}_{\mathrm{c}k} - \sqrt{n}\mathbf{I}_{\mathrm{c}} > \sqrt{n}(\mathbf{R}_{\mathrm{c}} - \mathbf{I}_{\mathrm{c}} - \gamma\mathbf{1})\right) \\ &\overset{(a)}{\leq} \Psi(-\sqrt{n}(\mathbf{R}_{\mathrm{c}} - \mathbf{I}_{\mathrm{c}} - \gamma\mathbf{1}); \mathbf{0}, V_{\mathrm{c}}) + \frac{\phi_{\mathrm{c}}}{\sqrt{n}} \\ &\overset{(b)}{\leq} \Psi(-\sqrt{n}(\mathbf{R}_{\mathrm{c}} - \mathbf{I}_{\mathrm{c}}); \mathbf{0}, V_{\mathrm{c}}) + O\left(\frac{\log n}{\sqrt{n}}\right). \end{aligned} \quad (30)$$

where (a) follows from the multivariate Berry-Esseen Theorem (see [19] or Corollary 38 in [20]); and (b) follows from Taylor expansion of the function $\Psi(\mathbf{t};\mathbf{0},V_{\rm c})$, which is differentiable wrt $\mathbf{t}$.

From Lemma 2, we have

$$\begin{aligned}\epsilon_n &\geq 1-\Pr\left(\frac{1}{n}\tilde{\mathbf{i}}_{\rm c}^n(X_1^nX_2^nY_1^nY_2^n)>\mathbf{R}_{\rm c}-\gamma\mathbf{1}\right)-2e^{-n\gamma}\\ &=1-\mathbb{E}\left[A(X_1^n,X_2^n)\right]-2e^{-n\gamma}. \end{aligned}\tag{31}$$

Note that $e^{-n\gamma}=\frac{1}{\sqrt{n}}$. Combining (30) and (31), we have

$$\begin{aligned}\epsilon_n &\geq 1-\Psi(-\sqrt{n}(\mathbf{R}_{\rm c}-\mathbf{I}_{\rm c});\mathbf{0},V_{\rm c})-O\left(\frac{\log n}{\sqrt{n}}\right)-\frac{2}{\sqrt{n}}\\ &\overset{(a)}{\geq} 1-\Psi\left(\begin{bmatrix}\sqrt{n}(I_{11}-\kappa_1)-L_1+\beta\\ \sqrt{n}(I_{21}-\kappa_2)-L_2+\beta\end{bmatrix};\mathbf{0},V_{\rm c}\right)\\ &\qquad -O\left(\frac{\log n}{\sqrt{n}}\right)-\frac{2}{\sqrt{n}} \end{aligned}\tag{32}$$

where (a) holds for all $n>N_\beta$ and follows because $\mathbf{t}\mapsto\Psi(\mathbf{t};\mathbf{0},V_{\rm c})$ is monotonically increasing in $\mathbf{t}$ and (27).

We now consider three different cases.

Case 1: When $\kappa_1=I_{11}$ and $\kappa_2<I_{21}$

For any fixed L_2, if $\kappa_2<I_{21}$, we have $\sqrt{n}(I_{21}-\kappa_2)-L_2+\beta\to+\infty$. Thus, the second term on the RHS of (32) converges to $\Psi\left(-L_1+\beta;0,V_1\right)=\Phi\big(\frac{-L_1+\beta}{\sqrt{V_1}}\big)$. Taking $\limsup$ on both sides of (32), and using (27), we have

$$\epsilon\geq\limsup_{n\to\infty}\epsilon_n\geq 1-\Phi\Big(\frac{-L_1+\beta}{\sqrt{V_1}}\Big).\tag{33}$$

Since this is true for any $\beta>0$, we may let $\beta\downarrow 0$ and deduce that $\Phi\big(\frac{L_1}{\sqrt{V_1}}\big)\leq\epsilon$. This case is proved.

Case 2: When $\kappa_1=I_{11}$ and $\kappa_2=I_{21}$

In this case, the second term on the RHS of (32) converges to $\Psi\left([-L_1+\beta,-L_2+\beta]^T;0,V_{\rm c}\right)$. The rest of the arguments are similar to that in case 1. Note that because $V_{\rm c}$ is diagonal,

$$\Psi\left([-L_1,-L_2]^T;0,V_{\rm c}\right)=\Phi\Big(-\frac{L_1}{\sqrt{V_1}}\Big)\Phi\Big(-\frac{L_2}{\sqrt{V_2}}\Big).\tag{34}$$

Case 3: When $\kappa_1<I_{11}$ and $\kappa_2=I_{21}$

By symmetry, case 3 is proved similarly to case 1.

C. Proof of Theorem 1: Direct Part

Case 1: When $\kappa_1=I_{11}$ and $\kappa_2<I_{21}$

Fix any pair (L_1,L_2) satisfying

$$\Phi\Big(\frac{L_1}{\sqrt{V_1}}\Big)\leq\epsilon.\tag{35}$$

Let the number of codewords in the j-th codebook be

$$M_{nj}=\lfloor e^{n\kappa_j+\sqrt{n}L_j+n^{1/4}\beta}\rfloor\tag{36}$$

for $j=1,2$, and a fixed $\beta>0$. It is clear that

$$\liminf_{n\to\infty}\frac{1}{\sqrt{n}}(\log M_{jn}-n\kappa_j)\geq L_j.\tag{37}$$

Therefore, it suffices to show the existence of a sequence of $(M_{1n},M_{2n},n,\epsilon_n,P_1,P_2)$-codes such that $\limsup_{n\to\infty}\epsilon_n\leq\epsilon$, in order to show that (L_1,L_2) is $(\kappa_1,\kappa_2,\epsilon)$-achievable. For this, we define an appropriate input distribution to be used in Lemma 1. Inspired by [15], [16], we define the input distributions to be uniform on the respective power shells, i.e.

$$P_{X_j^n}(x_j^n)\triangleq\frac{\delta(\|x_j^n\|-\sqrt{nP_j})}{S_n(\sqrt{nP_j})},\tag{38}$$

for $j=1,2$ and where $\delta(\cdot)$ is the Dirac delta and $S_n(r)=\frac{2\pi^{n/2}}{\Gamma(n/2)}r^{n-1}$ is the surface area of a sphere in $\mathbb{R}^n$ with radius r. With this choice, we have $P_{X_1^n}(\mathcal{F}_{1n}^c)+P_{X_2^n}(\mathcal{F}_{2n}^c)=0$, i.e. the power constraints are satisfied with probability 1.

Choose the auxiliary (conditional) output distributions $Q_{Y_1^n}(y_1^n)$, $Q_{Y_2^n}(y_2^n)$, $Q_{Y_1^n|X_2^n}(y_1^n|x_2^n)$ and $Q_{Y_2^n|X_1^n}(y_2^n|x_1^n)$ in Lemma 1 to be the n-fold memoryless extensions of $\hat{Q}_{Y_1}(y_1)$, $\hat{Q}_{Y_2}(y_2)$, $\hat{Q}_{Y_1|X_2}(y_1|x_2)$ and $\hat{Q}_{Y_2|X_1}(y_2|x_1)$ respectively, where these distributions are defined in (23-26).

For n sufficiently large, it can be proved that K_{11}, K_{21}, K_{12} and K_{22} are finite (see [15, Prop. 3]). Thus, K in (19) is also finite. Define

$$\alpha_{11}\triangleq 1+h_{11}^2P_1,\quad \alpha_{12}\triangleq 1+h_{11}^2P_1+h_{21}^2P_2,\tag{39}$$

$$\alpha_{21}\triangleq 1+h_{22}^2P_2,\quad \alpha_{22}\triangleq 1+h_{12}^2P_1+h_{22}^2P_2.\tag{40}$$

It can be shown that the four modified information densities take the form

$$\begin{aligned}\tilde{i}_{11}^n&=nI_{11}+\frac{1}{2\alpha_{11}}[(\alpha_{11}-1)(n-\|Z_1^n\|^2)+2h_{11}\langle X_1^n,Z_1^n\rangle]\\ \tilde{i}_{21}^n&=nI_{21}+\frac{1}{2\alpha_{21}}[(\alpha_{21}-1)(n-\|Z_2^n\|^2)+2h_{22}\langle X_2^n,Z_2^n\rangle]\\ \tilde{i}_{12}^n&=nI_{12}+\frac{1}{2\alpha_{12}}[(\alpha_{12}-1)(n-\|Z_1^n\|^2)\\ &\qquad+2h_{11}h_{21}\langle X_2^n,X_1^n\rangle+2h_{11}\langle X_1^n,Z_1^n\rangle+2h_{21}\langle X_2^n,Z_1^n\rangle]\\ \tilde{i}_{22}^n&=nI_{22}+\frac{1}{2\alpha_{22}}[(\alpha_{22}-1)(n-\|Z_2^n\|^2)\\ &\qquad+2h_{22}h_{12}\langle X_2^n,X_1^n\rangle+2h_{22}\langle X_2^n,Z_2^n\rangle+2h_{12}\langle X_1^n,Z_2^n\rangle].\end{aligned}$$

Next, we use the *central limit theorem for functions* technique proposed by MolavianJazi-Laneman [15] to transform these modified information densities into functions of sums of independent random vectors. Let $T_j^n\sim N(\mathbf{0},\mathbf{I}_{n\times n})$, for $j=1,2$, be standard Gaussian random vectors that are independent of each other and of the noises Z_j^n. Note that the input distribution in (38) results in $X_{jk}=\sqrt{nP_j}\frac{T_{jk}}{\|T_j^n\|}$, for $k\in\{1,\dots,n\}$. Indeed, $\|X_j^n\|^2=nP_j$ with probability one. Now consider the length-10 random vector $\mathbf{U}_k\triangleq(\{U_{j1k}\}_{j=1}^4,\{U_{j2k}\}_{j=1}^4,U_{9k},U_{10k})$, where

$$\begin{aligned}&U_{11k}\triangleq 1-Z_{1k}^2, && U_{21k}\triangleq h_{11}\sqrt{P_1}T_{1k}Z_{1k},\\ &U_{31k}\triangleq h_{21}\sqrt{P_2}T_{2k}Z_{1k}, && U_{41k}\triangleq h_{11}h_{21}\sqrt{P_1P_2}T_{1k}T_{2k},\\ &U_{12k}\triangleq 1-Z_{2k}^2, && U_{22k}\triangleq h_{22}\sqrt{P_2}T_{2k}Z_{2k},\\ &U_{32k}\triangleq h_{12}\sqrt{P_1}T_{1k}Z_{2k}, && U_{42k}\triangleq h_{12}h_{22}\sqrt{P_1P_2}T_{1k}T_{2k},\\ &U_{9k}\triangleq T_{1k}^2-1, && U_{10k}\triangleq T_{2k}^2-1.\end{aligned}\tag{41}$$

It is easy to verify that $\mathbf{U}_k$ is i.i.d. across all channel uses $k\in\{1,\dots,n\}$, and $\mathbb{E}(\mathbf{U}_k)=0$ and $\mathbb{E}(\|\mathbf{U}_k\|^3)$ is finite.

Define the functions $\tau_{11}, \tau_{12} : \mathbb{R}^{10} \to \mathbb{R}$ as follows

$$\tau_{11}(\mathbf{u}) \triangleq (\alpha_{11}-1)u_{11} + \frac{2u_{21}}{\sqrt{1+u_9}} \tag{42}$$

$$\tau_{12}(\mathbf{u}) \triangleq (\alpha_{12}-1)u_{11} + \frac{2u_{21}}{\sqrt{1+u_9}} + \frac{2u_{31}}{\sqrt{1+u_{10}}} + \frac{2u_{41}}{\sqrt{1+u_9}\sqrt{1+u_{10}}}, \tag{43}$$

for receiver 1. Similarly, define $\tau_{21}(\mathbf{u})$ and $\tau_{22}(\mathbf{u})$ for receiver 2. It can be shown that, for $l \in \{11, 12, 21, 22\}$,

$$\tilde{i}_l^n = nI_l + \frac{n}{2\alpha_l}\tau_l\left(\frac{1}{n}\sum_{k=1}^{n}\mathbf{U}_k\right). \tag{44}$$

Next, we can show the random vector $\frac{1}{\sqrt{n}}\tilde{\mathbf{i}}_{\rm d}^n - \sqrt{n}\,\mathbf{I}_{\rm d}$ converges in distribution to a zero-mean Gaussian with covariance matrix of the form

$$V_{\rm d} \triangleq \begin{bmatrix} V_1 & * & * & 0 \\ * & * & * & * \\ * & * & * & * \\ 0 & * & * & V_2 \end{bmatrix}. \tag{45}$$

In the above, the $*$'s represent entries that are inconsequential for the purposes of subsequent analyses.

Define the length-4 rate vector $\mathbf{R}_{\rm d} \triangleq [\frac{\log M_{1n}}{n}, \frac{\log(M_{1n}M_{2n})}{n}, \frac{\log(M_{1n}M_{2n})}{n}, \frac{\log M_{2n}}{n}]^T$. Appealing to Lemma 1, with $\gamma = \frac{\log n}{2n}$, we have

$$\begin{aligned} \epsilon_n &\le 1 - \Pr\left(\frac{1}{\sqrt{n}}\tilde{\mathbf{i}}_{\rm d}^n(X_1^nX_2^nY_1^nY_2^n) > \sqrt{n}(\mathbf{R}_{\rm d}+\gamma\mathbf{1})\right) - Ke^{-n\gamma} \\ &\le 1 - \Pr\left(\frac{1}{\sqrt{n}}\sum_{k=1}^{n}(\tilde{\mathbf{i}}_{{\rm d}k} - \mathbf{I}_{\rm d}) > \sqrt{n}\left(\mathbf{R}_{\rm d} - \mathbf{I}_{\rm d} + \gamma\mathbf{1}\right)\right) - \frac{K}{\sqrt{n}} \\ &\overset{(a)}{\le} 1 - \Psi\left(-\sqrt{n}\left(\mathbf{R}_{\rm d} - \mathbf{I}_{\rm d} + \gamma\mathbf{1}\right); \mathbf{0}, V_{\rm d}\right) - O\left(\frac{1}{\sqrt{n}}\right) \\ &\overset{(b)}{\le} 1 - \Psi(-\sqrt{n}(\mathbf{R}_{\rm d} - \mathbf{I}_{\rm d}); \mathbf{0}, V_{\rm d}) + O\left(\frac{\log n}{\sqrt{n}}\right), \end{aligned} \tag{46}$$

where (a) follows from a variant of the multivariate Berry-Esseen theorem (see [15, Prop. 1]); and (b) follows from Taylor expanding $\mathbf{t} \mapsto \Psi(\mathbf{t}; \mathbf{0}, V_{\rm d})$.

Due to the strictly VS interference assumption (Definition 2),

$$h_{22}^2P_2 + 1 < \frac{h_{21}^2P_2 + h_{11}^2P_1 + 1}{h_{11}^2P_1 + 1}. \tag{47}$$

Thus, $I_{11} + I_{21} < I_{12}$. Similarly, we have $I_{11} + I_{21} < I_{22}$. Therefore, as $n \to \infty$, we have

$$\begin{aligned} &-\sqrt{n}(\mathbf{R}_{\rm d} - \mathbf{I}_{\rm d}) \\ &= -\sqrt{n}\begin{bmatrix} \kappa_1 + \frac{L_1}{\sqrt{n}} + \frac{\beta}{n^{3/4}} - I_{11} \\ \kappa_1 + \kappa_2 + \frac{L_1}{\sqrt{n}} + \frac{L_2}{\sqrt{n}} + 2\frac{\beta}{n^{3/4}} - I_{12} \\ \kappa_1 + \kappa_2 + \frac{L_1}{\sqrt{n}} + \frac{L_2}{\sqrt{n}} + 2\frac{\beta}{n^{3/4}} - I_{22} \\ \kappa_2 + \frac{L_2}{\sqrt{n}} + \frac{\beta}{n^{3/4}} - I_{21} \end{bmatrix} \to \begin{bmatrix} -L_1 \\ +\infty \\ +\infty \\ +\infty \end{bmatrix}. \end{aligned} \tag{48}$$

Thus, $\Psi(-\sqrt{n}(\mathbf{R}_{\rm d} - \mathbf{I}_{\rm d}); \mathbf{0}, V_{\rm d}) \to \Psi(-L_1; 0, V_1) = \Phi\big(-\frac{L_1}{\sqrt{V_1}}\big)$. Taking $\limsup$ on both sides of (46), we have

$$\limsup_{n\to\infty} \epsilon_n \le 1 - \Phi\Big(-\frac{L_1}{\sqrt{V_1}}\Big) = \Phi\Big(\frac{L_1}{\sqrt{V_1}}\Big) \le \epsilon, \tag{49}$$

where the final inequality follows the choice of L_1 in (35). This completes the proof of the direct part for Case 1.

Case 2: When $\kappa_1 = I_{11}$ and $\kappa_2 = I_{21}$.

In this case, we have $\Psi(-\sqrt{n}(\mathbf{R}_{\rm d} - \mathbf{I}_{\rm d}); \mathbf{0}, V_{\rm d}) \to \Psi([-L_1 - L_2]^T; 0, V_{\rm c})$ because the second and third entries in (48) tend to $+\infty$ (by the strictly VS interference assumption) while the first and fourth entries tend to L_1 and L_2 respectively. Thus, as mentioned previously, only the $(1,1), (1,4), (4,1)$ and $(4,4)$ entries in $V_{\rm d}$ are required. Note that $V_{\rm d}(\{1,4\},\{1,4\}) = V_{\rm c}$. Furthermore, the relation in (34) holds. The rest of the proof is similar to case 1.

Case 3: When $\kappa_1 < I_{11}$ and $\kappa_2 = I_{21}$.

By symmetry, case 3 is proved similarly to case 1.

Acknowledgments: The authors would like to acknowledge several helpful discussions with Jonathan Scarlett and Masahito Hayashi.

References

[1] V. Strassen. Asymptotische abschatzungen in Shannon's informationstheorie. *Trans. 3rd Prague Conf. Inform. Th.*, 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. Verdú. Channel coding rate in the finite blocklength regime. *IEEE Transactions on Information Theory*, 56(5):2307–2359, May. 2010.

[4] V. Y. F. Tan and O. Kosut. On the dispersions of three network information theory problems. *IEEE Transactions on Information Theory*, 2013, DOI 10.1109/TIT.2013.2291231.

[5] S.-Q. Le, V. Y. F. Tan, and M. Motani. On the dispersions of the discrete memoryless interference channel. In *Proc. IEEE Int'l Symposium on Information Theory*, pages 1859–1864, Istanbul, Turkey, 2013.

[6] E. Haim, Y. Kochman, and U. Urez. A note on the dispersion of network problems. In *Proc. IEEE Convention of Electrical and Electronics Engineerings in Israel*, pages 1–9, Israel, 2012.

[7] R. Nomura and T. S. Han. Second-order Slepian-Wolf coding theorems for non-mixed and mixed sources. In *Proc. IEEE Int'l Symposium on Information Theory*, pages 1974–1979, Istanbul, Turkey, 2013.

[8] D. Slepian and J. K. Wolf. Noiseless coding of correlated information sources. *IEEE Transactions on Information Theory*, 19:471–80, 1973.

[9] J. Scarlett and V. Y. F. Tan. Second-order asymptotics for the Gaussian MAC with degraded message sets. Oct 2013. arXiv:1310.1197v2.

[10] S. Verdú and T. S. Han. A general formula for channel capacity. *IEEE Transactions on Information Theory*, 40(4):1147–57, Apr 1994.

[11] A. Feinstein. A new basic theorem of information theory. *IRE Transactions on Information Theory*, 4(4):2–22, 1954.

[12] A. B. Carleial. A case where interference does not reduce capacity. *IEEE Transactions on Information Theory*, 21:569–570, Sep. 1975.

[13] 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.

[14] 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.

[15] E. MolavianJazi and J. N. Laneman. A finite-blocklength perspective on Gaussian multi-access channels. 2013. arXiv:1309.2343v1.

[16] V. Y. F. Tan and M. Tomamichel. The third-order term in the normal approximation for the AWGN channel. 2013. arXiv:1311.2237v2.

[17] R. Ahlswede. An elementary proof of the strong converse theorem for the multiple access channel. *J. of Combinatorics, Information & System Sciences*, pages 216–230, 1982.

[18] S.-Q. Le, V. Y. F. Tan, and M. Motani. A case where interference does not affect the channel dispersion. 2014. Submitted to IEEE Transactions on Information Theory [http://arxiv.org/abs/1404.0255].

[19] F. Gotze. On the rate of convergence in the multivariate CLT. *The Annals of Probability*, 19(2):721–739, 1991.

[20] S. Watanabe, S. Kuzuoka, and V. Y. F. Tan. Non-asymptotic and second-order achievability bounds for coding with side-information. 2013. arXiv:1301.6467.