Joint Source-Channel Coding for the Multiple-Access Channel with Correlated Sources

Arezou Rezazadeh[1], Josep Font-Segura[1], Alfonso Martinez[1], Albert Guillén i Fàbregas[123]
[1]Universitat Pompeu Fabra, [2]ICREA and [3]University of Cambridge
arezou.rezazadeh@upf.edu, {josep.font,alfonso.martinez,guillen}@ieee.org

***Abstract*—This paper studies the random-coding exponent of joint source-channel coding for the multiple-access channel with correlated sources. For each user, by defining a threshold, the messages of each source are partitioned into two classes. The achievable exponent for correlated sources with two message-dependent input distributions for each user is determined and shown to be larger than that achieved using only one input distribution for each user. A system of equations is presented to determine the optimal thresholds maximizing the achievable exponent. The obtained exponent is compared with the one derived for the MAC with independent sources.**

I. Introduction

Some studies show that for point-to-point communication, using a partition of the message set into source-type classes and assigning one input distribution for each class leads to a larger exponent than having codewords drawn from a single product distribution [1], [2]. Recent studies generalize this result to the multiple-access channel (MAC) [3]. In [4], the exponent with message-dependent random-coding across two classes is found to beat independent identically distributed (iid) random-coding for a two-user MAC with independent sources.

For a two-user MAC with correlated sources, [5] studied a message-dependent ensemble where codewords are generated by a symbol-wise conditional probability distribution that depends on the instantaneous source symbol and on the empirical distribution of the source sequence. The derived exponent were given as a multidimensional optimization problem over distributions i.e., primal domain [5]. In this paper, we apply Lagrange duality theory to the results in [5] and find the exponent in the dual domain, i.e. as a lower dimensional problem over parameters in terms of Gallager functions. We show that the obtained exponent is larger than that achieved using only one input distribution for each user.

A. System Model, Definitions and Notations

Using the convention that scalar random variables are denoted by capital letters, we consider two correlated sources characterized by $P_{U_1U_2} \in \mathcal{P}_{\mathcal{U}_1\mathcal{U}_2}$ on the alphabet $\mathcal{U}_1 \times \mathcal{U}_2$, where $\mathcal{U}_1$ and $\mathcal{U}_2$ are the respective source alphabets, and $\mathcal{P}_{\mathcal{U}_1\mathcal{U}_2}$ is the set of all possible distributions of (U_1, U_2). In addition, the set of all empirical distributions on a joint vector in $\mathcal{U}_1^n \times \mathcal{U}_2^n$ (i.e. types) is denoted by $\mathcal{P}^n_{\mathcal{U}_1\mathcal{U}_2}$.

This work has been funded in part by the European Research Council under grant 725411, and by the Spanish Ministry of Economy and Competitiveness under grant TEC2016-78434-C3-1-R.

Encoder $\nu = 1, 2$ maps a length-n source message $\boldsymbol{u}_\nu$ to the length-n codeword $\boldsymbol{x}_\nu(\boldsymbol{u}_\nu)$ drawn from the codebook $\mathcal{C}^\nu = \{\boldsymbol{x}_\nu(\boldsymbol{u}_\nu); \boldsymbol{u}_\nu \in \mathcal{U}_\nu^n\}$. To simplify some expressions, we use underline to represent a pair of quantities for users 1 and 2, such as $\underline{u} = (u_1, u_2)$, $\underline{\boldsymbol{u}} = (\boldsymbol{u}_1, \boldsymbol{u}_2)$ or $\underline{\mathcal{U}} = \mathcal{U}_1 \times \mathcal{U}_2$. Both users send their respective codewords over discrete memoryless MAC with transition probability $W(y|\underline{x})$, input alphabets $\mathcal{X}_1$ and $\mathcal{X}_2$, and output alphabet $\mathcal{Y}$. By receiving the sequence $\boldsymbol{y}$, the decoder estimates the transmitted pair messages $\underline{\boldsymbol{u}}$ based on the maximum a posteriori criterion:

$$\hat{\underline{\boldsymbol{u}}} = \arg\max_{\underline{\boldsymbol{u}} \in \underline{\mathcal{U}}^n} P^n_{\underline{U}}(\underline{\boldsymbol{u}}) W^n\big(\boldsymbol{y} \,|\, \boldsymbol{x}_1(\boldsymbol{u}_1), \boldsymbol{x}_2(\boldsymbol{u}_2)\big). \tag{1}$$

An error occurs if the decoded messages $\hat{\underline{\boldsymbol{u}}}$ differ from the transmitted $\underline{\boldsymbol{u}}$; the error probability for a given pair of codebooks is thus given by

$$\epsilon^n(\mathcal{C}^1, \mathcal{C}^2) \triangleq \mathbb{P}\left[\hat{\underline{\boldsymbol{U}}} \neq \underline{\boldsymbol{U}}\right]. \tag{2}$$

The error event $\hat{\underline{\boldsymbol{U}}} \neq \underline{\boldsymbol{U}}$ can be split into three disjoint types of error events, namely $(\hat{\boldsymbol{U}}_1, \boldsymbol{U}_2) \neq (\boldsymbol{U}_1, \boldsymbol{U}_2)$, $(\boldsymbol{U}_1, \hat{\boldsymbol{U}}_2) \neq (\boldsymbol{U}_1, \boldsymbol{U}_2)$ and $(\hat{\boldsymbol{U}}_1, \hat{\boldsymbol{U}}_2) \neq (\boldsymbol{U}_1, \boldsymbol{U}_2)$. These events are respectively labeled by τ, with $\tau \in \{\{1\}, \{2\}, \{1, 2\}\}$. To further simplify some expressions, we adopt the following convention,

$$u_\tau = \begin{cases} \emptyset & \tau = \emptyset \\ u_1 & \tau = \{1\} \\ u_2 & \tau = \{2\} \\ \underline{u} & \tau = \{1, 2\} \end{cases}, \tag{3}$$

for the variable u_ν, and similarly for the probability distribution Q_ν and the set $\mathcal{X}_\nu$. We denote the complement of ν (or τ) in the set $\{1, 2\}$ (or the subsets of $\{1, 2\}$) by ν^c (or τ^c), e.g. $\tau^c = \{2\}$ for $\tau = \{1\}$ and $\tau^c = \emptyset$ for $\tau = \{1, 2\}$.

The pair of sources (U_1, U_2) is transmissible over the channel if there exists a sequence of codebooks $(\mathcal{C}^1_n, \mathcal{C}^2_n)$ such that $\lim_{n\to\infty} \epsilon^n(\mathcal{C}^1_n, \mathcal{C}^2_n) = 0$. An exponent $E(P_{\underline{U}}, W)$ is achievable if there exists a sequence of codebooks such that

$$\liminf_{n\to\infty} -\frac{1}{n} \log\left(\epsilon^n(\mathcal{C}^1_n, \mathcal{C}^2_n)\right) \geq E(P_{\underline{U}}, W). \tag{4}$$

II. Joint Source-Channel Random-Coding

For point-to-point transmission of a discrete memoryless source P_U over a discrete memoryless channel W, the joint source-channel iid random-coding with input distribution Q is expressed in terms of Gallager source and channel functions, respectively given by [6]

$$E_s(\rho, P_U) = (1+\rho)\log\left(\sum_u P_U(u)^{\frac{1}{1+\rho}}\right), \tag{5}$$

$$E_0(\rho, Q, W) = -\log\sum_y\left(\sum_x Q(x)W(y|x)^{\frac{1}{1+\rho}}\right)^{1+\rho}. \tag{6}$$

For a two-user MAC with two correlated sources $P_{\underline{U}}$, transition probability W and given input distributions Q_1 and Q_2, the i.i.d random-coding exponent is given by

$$E^{\text{i.i.d}}(P_{\underline{U}}, W) = \min_{\tau\in\{\{1\},\{2\},\{1,2\}\}} \max_{\rho\in[0,1]} E_0(\rho, Q_\tau, WQ_{\tau^c}) - E_{s,\tau}(\rho, P_{\underline{U}}), \tag{7}$$

where $E_0(\cdot)$ is given by (6), and $E_{s,\tau}(\cdot)$ is the generalized Gallager's source functions for error type τ where

$$E_{s,\tau}(\rho, P_{\underline{U}}) = \log\sum_{u_{\tau^c}}\left(\sum_{u_\tau} P_{\underline{U}}(\underline{u})^{\frac{1}{1+\rho}}\right)^{1+\rho}. \tag{8}$$

Another possible strategy, known as message-dependent random-coding [4], is to assign source messages to disjoint classes, and to use codewords generated according to a distribution that depends on the class index. The primal form of the message-dependent exponent for the MAC with two correlated sources has been given by [5, Eq. (9)] where codewords are generated according to conditional input distributions that depend on the composition of the source message.

For user $\nu = 1, 2$, let the set of input distributions $\{Q_{\nu,\hat{P}_{U_\nu}}\}$ be given. By applying the same approach in [5], the achievable exponent of [5, Eq. (9)] for statistically independent messages and codewords is simplified to (9) at the bottom of the page, where $[x]^+ = \max\{0, x\}$. In order to find the dual-domain version of (9), we firstly analyze the source-exponent terms.

A. Source exponent function

In [4], for each user, a fixed threshold was considered to partition the source-message set into two classes, i.e.,

$$\mathcal{A}_\nu^1(\gamma_\nu) = \{\boldsymbol{u}_\nu \in \mathcal{U}_\nu^n \ : \ P^n_{\boldsymbol{U}_\nu}(\boldsymbol{u}_\nu) \geq \gamma_\nu^n\}, \tag{10}$$
$$\mathcal{A}_\nu^2(\gamma_\nu) = \{\boldsymbol{u}_\nu \in \mathcal{U}_\nu^n \ : \ P^n_{\boldsymbol{U}_\nu}(\boldsymbol{u}_\nu) < \gamma_\nu^n\}. \tag{11}$$

Here, we use the same idea in the primal domain. Exploiting that the source messages are encoded independently for each user in distributed source coding [7], the following Lemma gives the asymptotic form of (10) and (11) for correlated sources.

Lemma 1: Let $P_{\underline{U}}$ be a probability distribution of two correlated sources and P_{U_ν} be the marginal distribution for source $\nu = 1, 2$. Given partitioning thresholds $\gamma_\nu \in [0, 1]$, the set of probability distributions $\mathcal{P}_{\underline{\mathcal{U}}}$ can be partitioned into disjoint classes $\mathcal{B}_\nu^1(\gamma_\nu)$ and $\mathcal{B}_\nu^2(\gamma_\nu)$ where

$$\mathcal{B}_\nu^1(\gamma_\nu) = \left\{\hat{P}_{\underline{U}} \in \mathcal{P}_{\underline{\mathcal{U}}} : \sum_{\underline{u}} \hat{P}_{\underline{U}}(\underline{u})\log P_{U_\nu}(u_\nu) \geq \log(\gamma_\nu)\right\}, \tag{12}$$

$$\mathcal{B}_\nu^2(\gamma_\nu) = \left\{\hat{P}_{\underline{U}} \in \mathcal{P}_{\underline{\mathcal{U}}} : \sum_{\underline{u}} \hat{P}_{\underline{U}}(\underline{u})\log P_{U_\nu}(u_\nu) < \log(\gamma_\nu)\right\}. \tag{13}$$

Proof: See Appendix A in [8]. ■

Roughly speaking, $\mathcal{B}_\nu^1(\gamma_\nu)$ in (12), can be interpreted as the asymptotic limit of the union of sequences $(\boldsymbol{u}_1, \boldsymbol{u}_2)$ with type $\hat{P}^n_{\underline{U}}$, where as long as the marginal probability $P^n_{U_\nu}(\boldsymbol{u}_\nu)$ is not less than the threshold γ_ν^n, the empirical distribution of $\boldsymbol{u}_{\nu^c}$ can be arbitrary (similarly for $\mathcal{B}_\nu^2(\gamma_\nu)$ in (13)). The following Proposition finds the Gallager source exponent function for the messages corresponding to $\mathcal{B}_\nu^1(\gamma_\nu)$ and $\mathcal{B}_\nu^2(\gamma_\nu)$.

Proposition 1: For given $\gamma_\nu \in [0, 1]$ and $i_\nu \in \{1, 2\}$, in view of $\mathcal{B}_\nu^{i_\nu}(\gamma_\nu)$ given by (12) and (13), we have

$$\min_{\hat{P}_{\underline{U}}\in\mathcal{P}_{\underline{\mathcal{U}}}:\hat{P}_{\underline{U}}\in\mathcal{B}_1^{i_1}(\gamma_1),\hat{P}_{\underline{U}}\in\mathcal{B}_2^{i_2}(\gamma_2)} D(\hat{P}_{\underline{U}}||P_{\underline{U}}) - \rho H(\hat{P}_{U_\tau|U_{\tau^c}}) = -E_{s,\tau,i_1,i_2}(\rho, P_{\underline{U}}, \underline{\gamma}), \tag{14}$$

where

$$E_{s,\tau,i_1,i_2}(\rho, P_{\underline{U}}, \underline{\gamma}) = \min_{\lambda_1\geq 0,\lambda_2\geq 0} \log\sum_{u_{\tau^c}}\left(\sum_{u_\tau} P_{\underline{U}}(\underline{u})^{\frac{1}{1+\rho}} \times \left(\frac{P_{U_1}(u_1)}{\gamma_1}\right)^{-\frac{(-1)^{i_1}\lambda_1}{1+\rho}}\left(\frac{P_{U_2}(u_2)}{\gamma_2}\right)^{-\frac{(-1)^{i_2}\lambda_2}{1+\rho}}\right)^{1+\rho}. \tag{15}$$

Proof: See Appendix B in [8]. ■

In fact, in (15), the objective function is a convex function with respect to λ_ν for $\nu = 1, 2$, and the optimal λ_ν minimizing (15) are the solution of an implicit equation obtained by setting the partial derivative of the objective function of (15) with respect to λ_ν equal to zero. To be precise, for the cases where both constraints $\hat{P}_{\underline{U}} \in \mathcal{B}_1^{i_1}(\gamma_1)$ and $\hat{P}_{\underline{U}} \in \mathcal{B}_2^{i_2}(\gamma_2)$ are active, λ_1 and λ_2 derived as the solution of the implicit equation, are greater than zero. Otherwise, the solution of the implicit equation is negative and the optimal λ_ν is zero.

$$E_1(P_{\underline{U}}, W) = \min_{\tau\in\{\{1\},\{2\},\{1,2\}\}} \min_{\hat{P}_{\underline{U}}\in\mathcal{P}_{\underline{\mathcal{U}}}} \min_{\hat{P}_{\underline{X}Y}\in\mathcal{P}_{\underline{\mathcal{X}}\times\mathcal{Y}}} D(\hat{P}_{\underline{U}}||P_{\underline{U}}) + D(\hat{P}_{\underline{X}Y}||Q_{1,\hat{P}_{U_1}}Q_{2,\hat{P}_{U_2}}W)$$
$$+\left[\min_{\substack{\tilde{P}_{\underline{U}}\in\mathcal{P}_{\underline{\mathcal{U}}}:\tilde{P}_{U_{\tau^c}}=\hat{P}_{U_{\tau^c}},\\ \mathbb{E}_{\tilde{P}}\log P_{\underline{U}}\geq\mathbb{E}_{\hat{P}}\log P_{\underline{U}}}} \min_{\substack{\tilde{P}_{\underline{X}Y}\in\mathcal{P}_{\underline{\mathcal{X}}\times\mathcal{Y}}:\tilde{P}_{X_{\tau^c}Y}=\hat{P}_{X_{\tau^c}Y},\\ \mathbb{E}_{\tilde{P}}\log W\geq\mathbb{E}_{\hat{P}}\log W}} D(\tilde{P}_{\underline{X}Y}||Q_{\tau,\tilde{P}_{U_\tau}}\hat{P}_{X_{\tau^c}Y}) - H(\tilde{P}_{U_\tau|U_{\tau^c}})\right]^+ \tag{9}$$

Here, we compare the result given by (15) with that for independent sources. In [2], [4], it was shown that the exponent is expressed in terms of two $E_{s,i_\nu}(\cdot)$ functions, namely

$$E_{s,1}(\rho, P_{U_\nu}, \gamma_\nu) = \begin{cases} E_s(\rho, P_{U_\nu}) & \frac{1}{1+\rho} \geq \frac{1}{1+\rho_{\gamma_\nu}}, \\ E_s(\rho_{\gamma_\nu}, P_{U_\nu}) + E_s'(\rho_{\gamma_\nu})(\rho - \rho_{\gamma_\nu}) & \frac{1}{1+\rho} < \frac{1}{1+\rho_{\gamma_\nu}}, \end{cases} \tag{16}$$

and a similar definition for $E_{s,2}(\rho, P_{U_\nu}, \gamma_\nu)$, with the two conditions swapped. In the definition of the $E_{s,i_\nu}(\cdot)$ functions, the parameter ρ_{γ_ν} is the solution of the implicit equation

$$\frac{\sum_u P_{U_\nu}(u_\nu)^{\frac{1}{1+\rho}} \log P_{U_\nu}(u_\nu)}{\sum_{u_\nu} P_{U_\nu}(u_\nu)^{\frac{1}{1+\rho}}} = \log(\gamma_\nu), \tag{17}$$

as long as $\min_{u_\nu} P_{U_\nu}(u_\nu) \leq \gamma_\nu \leq \max_{u_\nu} P_{U_\nu}(u_\nu)$. If $\gamma_\nu \in [0, \min_{u_\nu} P_{U_\nu}(u_\nu))$, we have $\rho_{\gamma_\nu} = -1_-$ and if $\gamma_\nu \in (\max_{u_\nu} P_{U_\nu}(u_\nu), 1]$, we have $\rho_{\gamma_\nu} = -1_+$.

Additionally, from [9, Lemma 3], for each source $\nu = 1, 2$ with distribution P_{U_ν}, threshold γ_ν, and $i_\nu = 1, 2$, we have

$$E_{s,i_\nu}(\rho, P_{U_\nu}, \gamma_\nu) = \min_{\lambda_\nu \geq 0} \log \sum_{u_\nu} P_{U_\nu}(u_\nu)^{\frac{1}{1+\rho}} \left(\frac{P_{U_\nu}(u_\nu)}{\gamma_\nu}\right)^{-\frac{(-1)^{i_\nu}\lambda_\nu}{1+\rho}}. \tag{18}$$

For independent sources, by applying $P_{\underline{U}}(\underline{u}) = P_{U_1}(u_1)P_{U_2}(u_2)$ in (15), and in view of (18), the function $E_{s,\tau,i_1,i_2}(\rho, P_{\underline{U}}, \underline{\gamma})$ is simplified as

$$E_{s,\tau,i_1,i_2}(\rho, P_{U_1}(u_1)P_{U_2}(u_2), \underline{\gamma}) = E_{s,i_\tau}(\rho, P_{U_\tau}, \gamma_\tau) + E_{s,i_{\tau^c}}(0, P_{U_{\tau^c}}, \gamma_{\tau^c}), \tag{19}$$

where as discussed in [4, Eq. (15)], for $\tau = \{1, 2\}$, $E_{s,i_{\{1,2\}}}(\rho, P_{\underline{U}}, \underline{\gamma}) = E_{s,i_1}(\rho, P_{U_1}, \gamma_1) + E_{s,i_2}(\rho, P_{U_2}, \gamma_2)$. In fact, depending on the tangent points in (17), $E_{s,\{1,2\},i_1,i_2}(\cdot)$ as a function of ρ is either $E_s(\rho, P_{U_\nu}) + E_s(\rho, P_{U_{\nu^c}})$ or $E_s(\rho, P_{U_\nu}) + E_{s,i_{\nu^c}}(\rho, P_{U_{\nu^c}}, \gamma_{\nu^c})$ where ν can be 1 or 2, and ν^c denotes the complement index of ν among the set $\{1, 2\}$.

For error type $\tau \in \{\{1\}, \{2\}\}$ and for the four combinations of $i_1, i_2 \in \{1, 2\}$, Fig. 1 shows (19) for two independent sources with given γ_1, γ_2. As shown in (19) and for Fig. 1, the functions $E_{s,\tau,1,1}(\cdot)$ and $E_{s,\tau,2,1}(\cdot)$ follow $E_s(\rho, P_{U_\tau})$ given by (5), for an interval of ρ, while they are the straight line tangent to Gallager's source function beyond that interval. However, the functions $E_{s,\tau,1,2}(\cdot)$ and $E_{s,\tau,2,2}(\cdot)$ are either the Gallager's source function shifted by $E_{s,i_{\tau^c}}(0, P_{U_{\tau^c}}, \gamma_{\tau^c})$ or the straight line tangent to it.

On the other hand, for correlated sources with four combinations of $i_1, i_2 \in \{1, 2\}$, Fig. 2 shows (15) for two correlated sources with given γ_1, γ_2 and error type τ. It can be seen that for the example of Fig. 2, the functions $E_{s,\tau,1,1}(\cdot)$ and $E_{s,\tau,2,1}(\cdot)$ are the generalized Gallager's source function (8) for an interval of ρ, while they are a curve tangent to $E_{s,\tau}(\cdot)$ beyond that interval. Thus, unlike the independent sources, instead of a straight line tangent to Gallager's source function, for correlated sources, a curve is tangent to $E_{s,\tau}(\cdot)$. The reason for this is explained in the following.

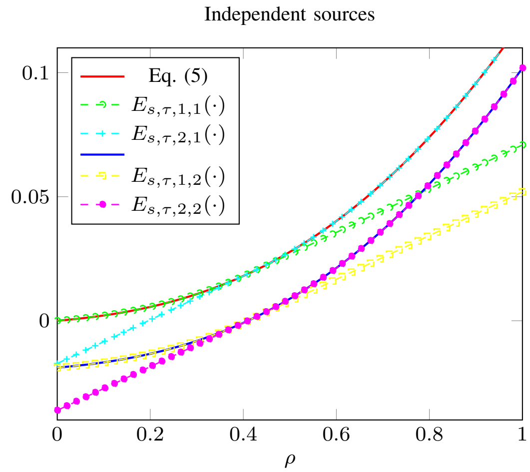

Fig. 1. $E_{s,\tau,i_1,i_2}(\cdot)$ in (19) for two independent sources versus ρ, for fixed γ_1 and γ_2 where $i_1, i_2 = 1, 2$. For error type $\tau \in \{\{1\}, \{2\}\}$, the solid red and blue curves are respectively $E_s(\rho, P_{U_\tau})$ and $E_s(\rho, P_{U_\tau}) + E_{s,i_{\tau^c}}(0, P_{U_{\tau^c}}, \gamma_{\tau^c})$.

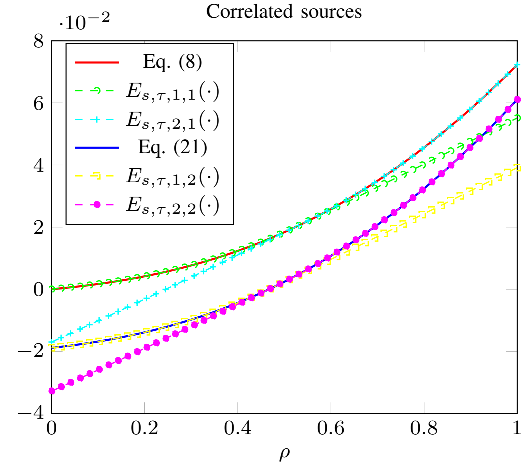

Fig. 2. $E_{s,\tau,i_1,i_2}(\cdot)$ in (15) for two correlated sources versus ρ, for fixed γ_1 and γ_2 where $i_1, i_2 = 1, 2$. The solid red and blue curves are respectively given by (8) and (21).

In Fig. 2, consider $E_{s,\tau,2,1}(\cdot)$ where $i_1 = 2$ and $i_2 = 1$. For the region of ρ where $E_{s,\tau,2,1}(\cdot)$ equals to $E_{s,\tau}(\cdot)$, both constraints $\hat{P}_{\underline{U}} \in \mathcal{B}_1^2(\gamma_1)$ and $\hat{P}_{\underline{U}} \in \mathcal{B}_2^1(\gamma_2)$ are inactive, while for the region of ρ where $E_{s,\tau,2,1}(\cdot)$ equals to the curve tangent to $E_{s,\tau}(\cdot)$, only one of the constraints $\hat{P}_{\underline{U}} \in \mathcal{B}_1^2(\gamma_1)$ or $\hat{P}_{\underline{U}} \in \mathcal{B}_2^1(\gamma_2)$ is active (similarly for $E_{s,\tau,1,1}(\cdot)$). For given i_1, i_2, let $\nu \in \{1, 2\}$ correspond to the active constraint. For example, in Fig. 2, for the region of ρ where $E_{s,\tau,2,1}(\cdot)$ equals the tangent curve, only the constraint $\hat{P}_{\underline{U}} \in \mathcal{B}_\nu^{i_\nu}(\gamma_\nu)$ is active. Then, the primal form of the curve is

$$-\min_{\substack{\hat{P}_{\underline{U}} \in \mathcal{P}_{\underline{\mathcal{U}}}: \\ \sum_{\underline{u}} \hat{P}_{\underline{U}}(\underline{u}) \log P_{U_\nu}(u_\nu) = \log(\gamma_\nu)}} D(\hat{P}_{\underline{U}} \| P_{\underline{U}}) - \rho H(\hat{P}_{U_\tau | U_{\tau^c}}), \tag{20}$$

as corresponds to the Gallager's source exponent function of messages source ν whose empirical distributions are fixed, i.e., $\{\hat{P}_{\underline{U}} \in \mathcal{P}_{\underline{\mathcal{U}}} : \sum_{\underline{u}} \hat{P}_{\underline{U}}(\underline{u}) \log P_{U_\nu}(u_\nu) = \log(\gamma_\nu)\}$.

We note that (20) describes the situation where only the type class of one source is fixed. Thus, we have more freedom in the source type class of the other source. This implies that for correlated sources the joint type class is not fixed, but rather contains the union of joint type classes whose type class of one of the sources is fixed. Unlike for independent sources, for correlated sources (20) is a curve rather than a straight line.

Coming back to Fig. 2, for an interval of ρ, the function $E_{s,\tau,1,2}(\cdot)$ ($E_{s,\tau,2,2}(\cdot)$) is

$$\min_{\lambda_\nu\geq 0}\log\sum_{u_{\tau^c}}\left(\sum_{u_\tau}P_{\underline{U}}(\underline{u})^{\frac{1}{1+\rho}}\left(\frac{P_{U_\nu}(u_\nu)}{\gamma_\nu}\right)^{-\frac{(-1)^{i_\nu}\lambda_\nu}{1+\rho}}\right)^{1+\rho}, \tag{21}$$

where $\nu\in\{1,2\}$ indicates that only the constraint $\hat{P}_{\underline{U}}\in\mathcal{B}_\nu^{i_\nu}(\gamma_\nu)$ is active. In addition, beyond that interval of ρ, the functions $E_{s,\tau,1,2}(\cdot)$ ($E_{s,\tau,2,2}(\cdot)$) is (15) where both constraints $\hat{P}_{\underline{U}}\in\mathcal{B}_1^{i_1}(\gamma_1)$ and $\hat{P}_{\underline{U}}\in\mathcal{B}_2^{i_2}(\gamma_2)$ are active.

B. Error Exponent Analysis

The primal form of the message-dependent exponent for the MAC with two correlated sources is given by (9). To find the dual-domain of (9), we use the following Lemma.

Lemma 2: $E_1(P_{\underline{U}},W)$ given by (9) is bounded as

$$\begin{aligned}E_1(P_{\underline{U}},W)\geq\min_\tau\min_{\hat{P}_{\underline{U}}\in\mathcal{P}_{\underline{\mathcal{U}}}}\min_{\hat{P}_{\underline{X}Y}\in\mathcal{P}_{\underline{\mathcal{X}}\times\mathcal{Y}}}&D(\hat{P}_{\underline{U}}\|P_{\underline{U}})\\&+D(\hat{P}_{\underline{X}Y}\|Q_{1,\hat{P}_{U_1}}Q_{2,\hat{P}_{U_2}}W)\\+\max_{\rho\in[0,1]}\rho D(\hat{P}_{\underline{X}Y}\|Q_{\tau,\hat{P}_{U_\tau}}\hat{P}_{X_{\tau^c}Y})&-\rho H(\hat{P}_{U_\tau|U_{\tau^c}}).\end{aligned}\tag{22}$$

Proof: See Appendix C in [8]. ■

The optimization problem over $\hat{P}_{\underline{X},Y}$ in (22) is coupled with the minimization problem over $\hat{P}_{\underline{U}}$ through $Q_{\nu,\hat{P}_{U_\nu}}$ for $\nu=1,2$. In view of classes defined by (12) and (13), we express the dependency of the input distribution $Q_{\nu,\hat{P}_{U_\nu}}$ on $\hat{P}_{U_\nu}$, through the class index. In other words, for $\hat{P}_{U_\nu}\in\mathcal{B}_\nu^1(\gamma_\nu)$, we let $Q_{\nu,\hat{P}_{U_\nu}}=Q_{\nu,1}$ and similarly for $\hat{P}_{U_\nu}\in\mathcal{B}_\nu^2(\gamma_\nu)$, we let $Q_{\nu,\hat{P}_{U_\nu}}=Q_{\nu,2}$. Applying this to (22), and splitting the minimization over $\hat{P}_{\underline{U}}$ into minimization over disjoint classes as $\min_{i_1,i_2=1,2}\min_{\hat{P}_{\underline{U}}\in\mathcal{P}_{\underline{\mathcal{U}}}:\hat{P}_{\underline{U}}\in\mathcal{B}_1^{i_1}(\gamma_1),\hat{P}_{\underline{U}}\in\mathcal{B}_2^{i_2}(\gamma_2)}$, we find that

$$\begin{aligned}&E_1(P_{\underline{U}},W)\geq\\&\min_\tau\min_{i_1,i_2=1,2}\min_{\hat{P}_{\underline{U}}\in\mathcal{P}_{\underline{\mathcal{U}}}:\hat{P}_{\underline{U}}\in\mathcal{B}_1^{i_1}(\gamma_1),\hat{P}_{\underline{U}}\in\mathcal{B}_2^{i_2}(\gamma_2)}D(\hat{P}_{\underline{U}}\|P_{\underline{U}})\\&\qquad+\min_{\hat{P}_{\underline{X}Y}\in\mathcal{P}_{\underline{\mathcal{X}}\times\mathcal{Y}}}D(\hat{P}_{\underline{X}Y}\|Q_{1,i_1}Q_{2,i_2}W)\\&\qquad+\max_{\rho\in[0,1]}\rho D(\hat{P}_{\underline{X}Y}\|Q_{\tau,i_\tau}\hat{P}_{X_{\tau^c}Y})-\rho H(\hat{P}_{U_\tau|U_{\tau^c}}).\end{aligned}\tag{23}$$

By using the min-max inequality, we swap the maximization over ρ with the minimizations over $\hat{P}_{\underline{X}Y}\in\mathcal{P}_{\underline{\mathcal{X}}\times\mathcal{Y}}$ and $\hat{P}_{\underline{U}}$ in (23), i.e., $E_1(P_{\underline{U}},W)\geq E(P_{\underline{U}},W)$ where $E(P_{\underline{U}},W)$ is given by (24) at the bottom of the page. In (24), the inner minimization problems over $\hat{P}_{\underline{X}Y}\in\mathcal{P}_{\underline{\mathcal{X}}\times\mathcal{Y}}$ and $\hat{P}_{\underline{U}}\in\mathcal{P}_{\underline{\mathcal{U}}}$, respectively lead to the channel and source exponent functions. The minimization over $\hat{P}_{\underline{U}}$ is discussed in Proposition 1, while to find channel exponent function, we use [8, Lemma 5]. By setting $\hat{P}_{XY}=\hat{P}_{\underline{X}Y}$ and $Q=Q_{\tau,i_\tau}$ in [8, Lemma 5], the minimization over $\hat{P}_{\underline{X}Y}$ in (24), is solved as

$$\begin{aligned}\min_{\hat{P}_{\underline{X}Y}\in\mathcal{P}_{\underline{\mathcal{X}}\times\mathcal{Y}}}&D(\hat{P}_{\underline{X}Y}\|Q_{1,i_1}Q_{2,i_2}W)\\&+\rho D(\hat{P}_{\underline{X}Y}\|Q_{\tau,i_\tau}\hat{P}_{X_{\tau^c}Y})=E_0(\rho,Q_{\tau,i_\tau},WQ_{\tau^c,i^c_\tau}),\end{aligned}\tag{25}$$

where $E_0(\cdot)$ is given by (6).

Now, putting back the results obtained in equations (25) and (14) into the respective minimization problems over $\hat{P}_{\underline{X}Y}$ and $\hat{P}_{\underline{U}}$ of (24), and defining

$$\begin{aligned}f_{i_1,i_2}(\gamma_1,\gamma_2)=\min_{\tau\in\{\{1\},\{2\},\{1,2\}\}}\max_{\rho\in[0,1]}&E_0(\rho,Q_{\tau,i_\tau},WQ_{\tau^c,i^c_\tau})\\&-E_{s,\tau,i_1,i_2}(\rho,P_{\underline{U}},\underline{\gamma}),\end{aligned}\tag{26}$$

an alternative expression for (24) is derived as

$$E(P_{\underline{U}},W)=\max_{\gamma_1,\gamma_2\in[0,1]}\min_{i_1,i_2=1,2}f_{i_1,i_2}(\gamma_1,\gamma_2),\tag{27}$$

where in (27), we optimized the exponent over γ_ν for $\nu=1,2$. We recall that since two source-message classes namely $\mathcal{B}_\nu^1(\gamma_\nu)$, $\mathcal{B}_\nu^2(\gamma_\nu)$ and two input distributions $Q_{\nu,1},Q_{\nu,2}$ are considered for each user $\nu=1,2$, there are four possible assignments where in (27) the optimal assignment of input distributions is considered.

In [8, Appendix D], we show that for $\nu=1,2$, the function $\max_{\rho\in[0,1]}E_0(\rho,Q_{\tau,i_\tau},WQ_{\tau^c,i^c_\tau})-E_{s,\tau,i_1,i_2}(\rho,P_{\underline{U}},\underline{\gamma})$ is non-decreasing with respect to γ_ν when $i_\nu=1$ and is non-increasing with respect to γ_ν when $i_\nu=2$. Considering this fact, to find the optimal $\underline{\gamma}$ maximizing (27), we can use the same approach proposed in [4, Proposition 2]. In other words, the optimal γ_1 and γ_2 are the points where the minimum of all non-decreasing functions with respect to γ_ν are equal with the minimum of all non-increasing functions.

Proposition 2: The optimal $\gamma_1^\star$ and $\gamma_2^\star$ maximizing (27) satisfy

$$\begin{cases}\min_{i_2=1,2}f_{1,i_2}(\gamma_1^\star,\gamma_2^\star)=\min_{i_2=1,2}f_{2,i_2}(\gamma_1^\star,\gamma_2^\star),\\\min_{i_1=1,2}f_{i_1,1}(\gamma_1^\star,\gamma_2^\star)=\min_{i_1=1,2}f_{i_1,2}(\gamma_1^\star,\gamma_2^\star).\end{cases}\tag{28}$$

If (28) has no solutions, $\gamma_\nu^\star\in\{0,1\}$: if $f_{1,i_2}(0,\gamma_2)>f_{2,i_2}(0,\gamma_2)$ then $\gamma_1^\star=0$, otherwise $\gamma_1^\star=1$; and if $f_{i_1,1}(\gamma_1,0)>f_{i_1,2}(\gamma_1,0)$, we have $\gamma_2^\star=0$, otherwise $\gamma_2^\star=1$.

Proof: See Appendix D in [8]. ■

$$\begin{aligned}E(P_{\underline{U}},W)=\min_{i_1,i_2=1,2}\min_{\tau\in\{\{1\},\{2\},\{1,2\}\}}\max_{\rho\in[0,1]}\min_{\hat{P}_{\underline{X}Y}\in\mathcal{P}_{\underline{\mathcal{X}}\times\mathcal{Y}}}&D(\hat{P}_{\underline{X}Y}\|Q_{1,i_1}Q_{2,i_2}W)+\rho D(\hat{P}_{\underline{X}Y}\|Q_{\tau,i_\tau}\hat{P}_{X_{\tau^c}Y})\\+\min_{\hat{P}_{\underline{U}}\in\mathcal{P}_{\underline{\mathcal{U}}}:\hat{P}_{\underline{U}}\in\mathcal{B}_1^{i_1}(\gamma_1),\hat{P}_{\underline{U}}\in\mathcal{B}_2^{i_2}(\gamma_2)}&D(\hat{P}_{\underline{U}}\|P_{\underline{U}})-\rho H(\hat{P}_{U_\tau|U_{\tau^c}})\end{aligned}\tag{24}$$

III. Numerical Example

In this section, we present an example showing that using two input distributions for each user attains larger achievable exponent than the case where each user uses one input distribution, a case whose exponent is given by

$$\max_{i_1\in\{1,2\}} \max_{i_2\in\{1,2\}} \min_{\tau} F^{\mathrm{L}}_{\tau,i_\tau,i_{\tau^c}}, \tag{29}$$

where

$$F^{\mathrm{L}}_{\tau,i_\tau,i_{\tau^c}} = \max_{\rho\in[0,1]} E_0(\rho, Q_{\tau,i_\tau}, WQ_{\tau^c,i_{\tau^c}}) - E_{s,\tau}(\rho, P_{\underline{U}}). \tag{30}$$

We consider two correlated discrete memoryless sources with alphabet $\mathcal{U}_\nu = \{1,2\}$ for $\nu = 1,2$ where

$$P_{\underline{U}} = \begin{pmatrix} 0.0005 & 0.0095 \\ 0.0005 & 0.9895 \end{pmatrix}. \tag{31}$$

We also consider a discrete memoryless MAC, similar to the one given by [4, Eq. (31)], with $\mathcal{X}_1 = \mathcal{X}_2 = \{1,\dots,6\}$ and $|\mathcal{Y}| = 4$. Let W be the transition probability of this channel,

$$W = \left(W_1^T, W_2^T, W_3^T, W_4^T, W_5^T, W_6^T\right)^T, \tag{32}$$

where

$$W_1 = \begin{pmatrix} 1-3k_1 & k_1 & k_1 & k_1 \\ k_1 & 1-3k_1 & k_1 & k_1 \\ k_1 & k_1 & 1-3k_1 & k_1 \\ k_1 & k_1 & k_1 & 1-3k_1 \\ 0.5-k_2 & 0.5-k_2 & k_2 & k_2 \\ k_2 & k_2 & 0.5-k_2 & 0.5-k_2 \end{pmatrix}, \tag{33}$$

for $k_1 = 0.045$ and $k_2 = 0.01$. W_2 and W_3 are 6×4 matrices whose rows are all the copy of 5^{th} and 6^{th} row of matrix W_1, respectively. W_4 is a 6×4 matrix with rows numbers 2, 3, 4, 1, 6, and 5 of W_1. Similarly, W_5 is a 6×4 matrix with rows numbers 3, 4, 1, 2, 5, and 6 of W_1 and W_6 is a 6×4 matrices with rows numbers 4, 1, 2, 3, 6, and 5 of W_1.

We observe that W is a 36×4 matrix where the transition probability $W(y|x_1,x_2)$ is located at row $x_1 + 6(x_2-1)$ of matrix W, for $(x_1,x_2) \in \{1,2,...,6\}\times\{1,2,...,6\}$. Recalling that each source has two classes and that four input distributions generate codewords, there are four possible assignments of input distributions to classes. Among all possible permutations, we select the one that gives the highest exponent. Here, for user $\nu = 1,2$, we consider the set of input distributions $\big\{[0\ 0\ 0\ 0\ 0.5\ 0.5],\ [0.25\ 0.25\ 0.25\ 0.25\ 0\ 0]\big\}$. For the channel given in (32), the optimal assignment is

$$Q_{\nu,1} = [0\ 0\ 0\ 0\ 0.5\ 0.5], \tag{34}$$
$$Q_{\nu,2} = [0.25\ 0.25\ 0.25\ 0.25\ 0\ 0], \tag{35}$$

for both $\nu = 1,2$.

For this example, from (28), we numerically compute the optimal $\gamma_1^\star$ and $\gamma_2^\star$ maximizing (27) leading to $\gamma_1^\star = 0.8469$ and $\gamma_2^\star = 0.6581$. The message-dependent exponent is derived as $E(P_{\underline{U}}, W) = 0.2611$, while i.i.d. exponent for the best assignment is derived as 0.2503. Fig. 3 shows $\min_{i_1,i_2} f_{i_1,i_2}(\underline{\gamma})$ with respect to γ_1 and γ_2. It can be seen that the maximum of $\min_{i_1,i_2} f_{i_1,i_2}(\underline{\gamma})$ is derived at $(0.8469, 0.6581)$; however, the lower bound is obtained at $(1,0)$.

Table I: Values of $\max_{\rho\in[0,1]} E_0(\rho, Q_{\tau,i_\tau}, WQ_{\tau^c,i^c_\tau}) - E_{s,\tau,i_1,i_2}(\rho, P_{\underline{U}}, \underline{\gamma})$ with optimal thresholds $\gamma_1^\star = 0.8469$ $\gamma_2^\star = 0.6581$, for types of error τ, and user classes i_τ, i_{τ^c}.

	(i_1,i_2)			
	(1,1)	(1,2)	(2,1)	(2,2)
$\tau=\{1\}$	0.3172	0.2735	0.3120	0.2611
$\tau=\{2\}$	0.3986	0.4372	0.2611	0.4119
$\tau=\{1,2\}$	0.2611	0.2972	0.2630	0.2883

Table II: Values of $F^{\mathrm{L}}_{\tau,i_\tau,i_{\tau^c}}$ in (30) for types of error τ, and input distribution Q_{1,i_1}, Q_{2,i_2}.

	$Q_{1,1}$,$Q_{2,1}$	$Q_{1,1}$,$Q_{2,2}$	$Q_{1,2}$,$Q_{2,1}$	$Q_{1,2}$,$Q_{2,2}$
$\tau=\{1\}$	0.2682	0.0642	0.3120	0.0879
$\tau=\{2\}$	0.3986	0.3986	0.2503	0.3696
$\tau=\{1,2\}$	0.2097	0.2097	0.2630	0.2360

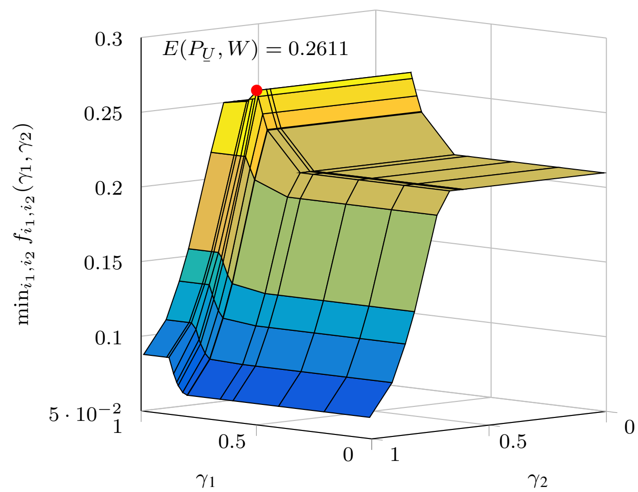

Fig. 3. $\min_{i_1,i_2} f_{i_1,i_2}(\gamma_1,\gamma_2)$ with respect to γ_1 and γ_2.

References

[1] I. Csiszár, "Joint source-channel error exponent," *Probl. Control Inf. Theory*, vol. 9, no. 1, pp. 315–328, 1980.

[2] A. Tauste Campo, G. Vazquez-Vilar, A. Guillén i Fàbregas, A. Martinez, and T. Koch, "A derivation of the source-channel error exponent using nonidentical product distributions," *IEEE Trans. Inf. Theory*, vol. 60, no. 6, pp. 3209–3217, Jun. 2014.

[3] L. Farkas and T. Kói, "Random access and source-channel coding error exponents for multiple access channels," *IEEE Trans. Inf. Theory*, vol. 61, no. 6, pp. 3029–3040, June 2015.

[4] A. Rezazadeh, J. Font-Segura, A. Martinez, and A. Guillén i Fàbregas, "Multiple-access channel with independent sources: Error exponent analysis," in *2018 Information Theory Workshop*, Guangzhou, China, Nov. 2018.

[5] ——, "An achievable error exponent for the multiple access channel with correlated sources," in *IEEE Int. Symp. Inf. Theory (ISIT)*, Aachen, Germany, Jun. 2017.

[6] R. Gallager, *Information Theory and Reliable Communication*. John Wiley & Sons, 1968.

[7] D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inf. Theory*, vol. 19, no. 4, pp. 471–480, July 1973.

[8] A. Rezazadeh, J. Font-Segura, A. Martinez, and A. Guillén i Fàbregas, "Joint Source-Channel Coding for the Multiple-Access Channel with Correlated Sources." [Online]. Available: https://arxiv.org/abs/1901.05778v2

[9] ——, "Multiple-Access Channel with Independent Sources: Error Exponent Analysis." [Online]. Available: https://arxiv.org/abs/1805.11315v2