Transmission of Correlated Sources over a MAC: A Gaussian Approximation-Based Analysis

Vincent Y. F. Tan

Institute for Infocomm Research (I2R), A*STAR, Singapore
National University of Singapore (NUS)
Emails: tanyfv@i2r.a-star.edu.sg, eletyfv@nus.edu.sg

***Abstract*—The problem of separately encoding and subsequently transmitting correlated sources over a discrete memoryless multiple-access channel is revisited. In particular, we examine the sufficient conditions on the source and the channel under which there exists a n-length block code satisfying the condition that the error probability of reconstructing the discrete memoryless multiple source is no larger than some fixed constant $\epsilon > 0$, or in short, ϵ-lossless transmission. We modify the decoding rule of Cover-El Gamal-Salehi and analyze the error probability using Gaussian approximations to derive a second-order generalization their sufficient condition for the ϵ-lossless transmission of the correlated sources.**

***Index Terms*—Correlated sources, Multiple-access channel, Gaussian approximations, Second-order coding rates, Dispersion**

I. Introduction

In this paper, we revisit and present a Gaussian approximation-based analysis of one of the most fundamental problems in network information theory—that of transmitting correlated (discrete memoryless multiple) sources over a discrete memoryless multiple-access channel (MAC) [1], [2]. The correlated information source is modeled as the realization of two (or more) random variables S and T jointly distributed independently and identically as $p_{S,T}$. Using a length-n block code, the sources are separately encoded as X_1^n and X_2^n respectively and these representations are then transmitted across n uses of a discrete memoryless multiple access channel $W(y|x_1,x_2)$. By using a separation strategy involving Slepian-Wolf compression [3] followed by encoding the compressed bits using a multiple-access channel code [4], [5], we observe that if there exists a $p_Q(q)$, $p_{X_1|Q}(x_1|q)$ and $p_{X_2|Q}(x_2|q)$ such that the following conditions hold:

$$\begin{aligned} H(S|T) &< I(X_1;Y|X_2,Q) \\ H(T|S) &< I(X_2;Y|X_1,Q) \\ H(S,T) &< I(X_1,X_2;Y|Q), \end{aligned} \tag{1}$$

then the probability of decoding error tends to zero as the blocklength tends to infinity, i.e.,

$$\lim_{n\to\infty} \mathbb{P}\big[(\hat{S}^n,\hat{T}^n) \ne (S^n,T^n)\big] = 0. \tag{2}$$

Cover, El Gamal and Salehi [1] proved that, unlike the point-to-point setting, such a separation strategy is suboptimal. They proceeded to exploit the *common information* between S and T to obtain a better achievable condition. By defining K as the common part of S and T (in the sense of Gács and Körner [6]), they showed that if there exists distributions $p_C(c)$, $p_{X_1|S,C}(x_1|s,c)$ and $p_{X_2|T,C}(x_2|t,c)$ such that

$$\begin{aligned} H(S|T) &< I(X_1;Y|X_2,T,C) \\ H(T|S) &< I(X_2;Y|X_1,S,C) \\ H(T,S|K) &< I(X_1,X_2;Y|K,C) \\ H(S,T) &< I(X_1,X_2;Y) \end{aligned} \tag{3}$$

then the probability of error tends to zero as the blocklength tends to infinity. The independent auxiliary random variable C represents the common part of the sources K. Ahlswede and Han [2] simplified the proof of the achievability of (3), and we will adopt their proof strategy. It is known (see [7]) that the condition in (3) is suboptimal. The best known outer bound is by Kang and Ulukus [8] but this paper only focuses on generalizing the achievability result in (3).

In this paper, we are interested in deriving a sufficient condition on the source (S,T) and the channel W (analogous to (3)) such that for sufficiently large blocklengths n, the error probability in reconstructing the source is no larger than some $\epsilon > 0$, i.e., instead of (2), we have the constraint

$$\mathbb{P}\big[(\hat{S}^n,\hat{T}^n) \ne (S^n,T^n)\big] \le \epsilon. \tag{4}$$

This analysis brings us a step closer to understanding the non-asymptotic fundamental limits for transmitting correlated sources over a MAC. The sufficient condition (see (5) below) is termed a *second-order* (or *dispersion-type*) condition which is in line with the recent works on second-order coding rates [9] and finite blocklength analysis [10]. In our context, a sufficiently large n and an $\epsilon \in (0,1)$ are fixed and we ask what are the conditions on the source (S,T) and the channel W such that there exists a code for which (4) holds.

A. Our Main Contribution

Our main result in this paper can be roughly summarized as follows: Let the set $\mathscr{S}_{\mathbf{V}}(\epsilon) \subset \mathbb{R}^6$ (to be defined later) play the role of the $Q^{-1}(\,\cdot\,)$ function for a zero-mean Gaussian with covariance matrix $\mathbf{V}$ [11]. We show, by using the proof techniques in [1], [2] and [12] but with a different decoder, that if there exists distributions $p_C(c)$, $p_{X_1|S,C}(x_1|s,c)$ and

$p_{X_2|S,T}(x_2|t,c)$ such that

$$\mathbf{I} - \mathbf{H} \in \frac{\mathscr{S}_{\mathbf{V}}(\epsilon)}{\sqrt{n}} + \frac{\log n}{n}\mathbf{1}_6 \tag{5}$$

as n grows, then (4) holds. Here, $\mathbf{1}_d$ is the vector of length d consisting of ones. We briefly describe the terms in (5): The vectors $\mathbf{I}, \mathbf{H} \in \mathbb{R}_+^6$ are the mutual information and entropy vectors (analogous to those in (3) but there are two extra elements). These are functions of the distributions $p_C(c)$, $p_{X_1|S,C}(x_1|s,c)$ and $p_{X_2|S,T}(x_2|t,c)$. The matrix $\mathbf{V}$ (to be defined later) is termed the *information dispersion matrix* of the source and channel and is a generalization of the dispersion matrices derived in [11]. Note that the result in (5) generalizes the asymptotic results in [1], [2], [12] because as $n \to \infty$, we recover the first-order condition in (3) for every $\epsilon > 0$.

An interesting interpretation of (5) is as follows: If n is small, then the mutual information quantities in (3) have to be significantly larger than the corresponding entropy quantities. Conversely, for large n, the mutual information quantities can be rather close to the corresponding entropy quantities. Our result differs from the first-order result in (3) in that there are two extra constraints on the mutual informations and entropies, namely that $I(X_1, C; Y|X_2, T) - H(S|T)$ and $I(X_2, C; Y|X_1, S) - H(T|S)$ have to be sufficiently positive at blocklength n (also relative to the other constraints). It is not clear that these constraints can be subsumed by other ones because the second-order terms governed by $\mathscr{S}_{\mathbf{V}}(\epsilon)/\sqrt{n}$ come into play at a finite blocklength n.

B. Related Work

This work is an extension of the work by Tan and Kosut [11] where second-order coding rates for Slepian-Wolf coding [3] and the discrete memoryless MAC [4], [5] were examined in detail. Clearly, if there is a non-empty intersection between the $(n, \epsilon_{\mathrm{s}})$-Slepian-Wolf and $(n, \epsilon_{\mathrm{c}})$-MAC regions and the error probabilities satisfy $\epsilon_{\mathrm{s}} + \epsilon_{\mathrm{c}} - \epsilon_{\mathrm{s}}\epsilon_{\mathrm{c}} \le \epsilon$, then (4) is satisfied. Instead of separation, we study a joint source-channel coding scheme (JSCC) in this paper. The work by Tan and Kosut [11] builds on a line of work on dispersion analysis for channel coding that was studied extensively by Hayashi [9] and Polyanskiy et al. [10]. In the latter, the authors introduced new channel coding rate bounds and used these bounds to strengthen the results in Strassen's seminal work [13]. In addition, finite blocklength analysis has also been applied to lossy source coding [14], [15] and point-to-point JSCC [16], [17] just to name a few. The general version of the point-to-point-JSCC problem was studied by Han [18, Chapter 5] and extended by Campo et al. [19]. Iwata and Oohama [12] studied the multi-sender-single-receiver version of the JSCC problem using information spectrum methods [18]. We will also use some information spectrum ideas in our proofs.

C. Structure of Paper

This paper is structured as follows: In Section II, we state the definition of the problem. In Section III, we state our main result by first generalizing the main result of Ahlswede and Han [2] then specializing it to S, T and their common part K to obtain the generalization of (3). In Section IV, we present the proofs of our main theorems. In Section V, we conclude our discussion and suggest avenues for further research.

II. Definitions

In this section, we present some definitions for the problem of transmitting a correlated source over a MAC. Recall that we wish to send a discrete memoryless multiple source $(S, T) \sim p_{S,T}(s,t)$ over a discrete memoryless MAC $(\mathcal{X}_1, \mathcal{X}_2, W(y|x_1, x_2), \mathcal{Y})$.

Definition 1. *A* $(|\mathcal{S}|^n, |\mathcal{T}|^n, n, \epsilon)$ joint source-channel code *for transmitting the correlated source* (S, T) *over the discrete memoryless MAC* $(\mathcal{X}_1, \mathcal{X}_2, W(y|x_1, x_2), \mathcal{Y})$ *consists of*

1) *Two encoders: Encoder* 1 *assigns a sequence* $\phi_n^1(s^n) \in \mathcal{X}_1^n$ *to every source sequence* $s^n \in \mathcal{S}^n$*. Encoder* 2 *assigns a sequence* $\phi_n^2(t^n) \in \mathcal{X}_2^n$ *to every source sequence* $t^n \in \mathcal{T}^n$*.*
2) *A decoder* ψ_n *that assigns an estimate* $(\hat{s}^n, \hat{t}^n) \in \mathcal{S}^n \times \mathcal{T}^n$ *to each sequence (channel output)* $y^n \in \mathcal{Y}^n$

such that the error probability satisfies (4)*.*

Definition 2. *We say that the source* (S, T) *can be* (n, ϵ)-transmissible *over the discrete memoryless MAC* $(\mathcal{X}_1, \mathcal{X}_2, W(y|x_1, x_2), \mathcal{Y})$ *if there exists a* $(|\mathcal{S}|^n, |\mathcal{T}|^n, n, \epsilon)$ *joint source-channel code.*

Given a vector $\boldsymbol{\mu} \in \mathbb{R}^d$ and a positive definite matrix $\mathbf{\Sigma} \in \mathbb{S}_+^d$, we use the notation $\mathcal{N}(\boldsymbol{\mu}, \mathbf{\Sigma})$ to denote a multivariate Gaussian probability density function with mean $\boldsymbol{\mu}$ and covariance matrix $\mathbf{\Sigma}$.

Definition 3. *Define the set [11]*

$$\mathscr{S}_{\mathbf{V}}(\epsilon) := \left\{ \mathbf{z} \in \mathbb{R}^d : \mathbb{P}[\mathbf{Z} \le \mathbf{z}] \ge 1 - \epsilon \right\} \tag{6}$$

where $\mathbf{Z} \sim \mathcal{N}(\mathbf{0}, \mathbf{V})$ *and for two vectors* $\mathbf{a}, \mathbf{b} \in \mathbb{R}^d$*,* $\mathbf{a} \le \mathbf{b}$ *means that* $a_j \le b_j$ *for all* $1 \le j \le d$*.*

The set $\mathscr{S}_{\mathbf{V}}(\epsilon)$ is analogous to the inverse Q-function for univariate Gaussians. Indeed, observe that if $d = 1$, $\mathscr{S}_{\sigma^2}(\epsilon) = [\sigma Q^{-1}(\epsilon), \infty)$. The set $\mathscr{S}_{\mathbf{V}}(\epsilon)$ characterizes the family of $\mathbf{z}$'s that are points of sufficiently "high probability" in $\mathbb{R}^d$. Our characterization of (n, ϵ)-transmissibility will be in terms of the set $\mathscr{S}_{\mathbf{V}}(\epsilon)$ for an appropriate (information dispersion) matrix $\mathbf{V}$.

III. Main Results

A. (n, ϵ)-Transmissibility for an Auxiliary System

While we would like to directly characterize the (n, ϵ)-transmissibility for the joint source (S, T), it is perhaps more straightforward and intuitive to adopt the Ahlswede-Han approach [2] to solve a different but related and simpler problem first. Loosely speaking, this problem decouples the effect of the common part from the actual sources, thus simplifying the proof of the original problem. Here, we have a discrete memoryless multiple source $(S_1, S_2, S_3) \sim p_{S_1,S_2,S_3}$ and a MAC $W(y|x_1, x_2)$. We would like to design two encoders observing

(S_1^n, S_2^n) and (S_2^n, S_3^n) and a decoder which observes the output of the MAC and estimates the three sources. That is the first encoder is $\phi_n^1 : \mathcal{S}_1^n \times \mathcal{S}_2^n \to \mathcal{X}_1^n$, the second encoder is $\phi_n^2 : \mathcal{S}_2^n \times \mathcal{S}_3^n \to \mathcal{X}_2^n$ and the decoder is $\psi_n : \mathcal{Y}^n \to \mathcal{S}_1^n \times \mathcal{S}_2^n \times \mathcal{S}_3^n$. The error probability $\mathbb{P}[(\hat{S}_1^n, \hat{S}_2^n, \hat{S}_3^n) \neq (S_1^n, S_2^n, S_3^n)]$ is to be no larger than ϵ for sufficiently large blocklength n. We define (n, ϵ)-transmissibility of the joint source (S_1, S_2, S_3) analogously to Definitions 1 and 2.

For solving this related problem, consider a test channel involving three input auxiliary random variables U_1, U_2, and U_3 taking values on finite sets $\mathcal{U}_1, \mathcal{U}_2$ and $\mathcal{U}_3$ respectively. Consider functions $f_1 : \mathcal{U}_1 \times \mathcal{U}_2 \to \mathcal{X}_1$ and $f_2 : \mathcal{U}_2 \times \mathcal{U}_3 \to \mathcal{X}_2$. Now define the test channel $\tilde{W} : \mathcal{U}_1 \times \mathcal{U}_2 \times \mathcal{U}_3 \to \mathcal{Y}$ as

$$\tilde{W}(y|u_1, u_2, u_3) := W(y|f_1(u_1, u_2), f_2(u_2, u_3)). \tag{7}$$

The random variables $(S_1, S_2, S_3, U_1, U_2, U_3)$ are distributed as $p_{S_1,S_2,S_3} p_{U_1|S_1} p_{U_2|S_2} p_{U_3|S_3}$. Further, (7) says that the two inputs to the MAC are $X_1 = f_1(U_1, U_2)$ and $X_2 = f_2(U_2, U_3)$ respectively. To state the next theorem concisely, we introduce the notation $S_{\mathcal{A}}$ to mean the set of random variables indexed by the finite set $\mathcal{A}$. So, for example, $S_{\{1,2\}} = (S_1, S_2)$. Also given a family of subsets $\mathscr{F}$, we use the compact notation $[v_{\mathcal{A}} : \mathcal{A} \in \mathscr{F}]$ to mean the vector indexed by the elements $v_{\mathcal{A}}$ arranged in some consistent order. We have the following analogue of Theorem 1 in [2].

Theorem 1 ((n, ϵ)-Transmissibility for the joint source (S_1, S_2, S_3))**.** *Let* $[3] := \{1, 2, 3\}$ *and* $\mathcal{A}^c := [3] \setminus \mathcal{A}$. *If for sufficiently large* n, *there exists distributions* $p_{U_1|S_1}, p_{U_2|S_2}, p_{U_3|S_3}$ *and functions* f_1 *and* f_2 *such that*

$$\begin{aligned} &[I(U_{\mathcal{A}}; Y|U_{\mathcal{A}^c}, S_{\mathcal{A}^c}) - H(S_{\mathcal{A}}|S_{\mathcal{A}^c}) : \emptyset \neq \mathcal{A} \subset [3]] \\ &\qquad \in \frac{\mathscr{S}_{\mathbf{V}(U_1,U_2,U_3)}(\epsilon)}{\sqrt{n}} + \frac{\log n}{n}\mathbf{1}_7, \end{aligned} \tag{8}$$

where the information dispersion matrix $\mathbf{V}(U_1, U_2, U_3)$ *is the covariance of the random vector*

$$\Bigg[\log \frac{p_{Y|U_{\mathcal{A}},U_{\mathcal{A}^c},S_{\mathcal{A}^c}}(Y|U_{\mathcal{A}}, U_{\mathcal{A}^c}, S_{\mathcal{A}^c})}{p_{Y|U_{\mathcal{A}^c},S_{\mathcal{A}^c}}(Y|U_{\mathcal{A}^c}, S_{\mathcal{A}^c})} - \log \frac{1}{p_{S_{\mathcal{A}}|S_{\mathcal{A}^c}}(S_{\mathcal{A}}|S_{\mathcal{A}^c})} : \emptyset \neq \mathcal{A} \subset [3]\Bigg], \tag{9}$$

then the source (S_1, S_2, S_3) *is* (n, ϵ)*-transmissible over* W.

This result is proved in Section IV. Essentially, we adopt the coding scheme for Theorem 1 in [1] but we modify the decoding rule so as to obtain the second-order term $\mathscr{S}_{\mathbf{V}(U_1,U_2,U_3)}(\epsilon)/\sqrt{n} \subset \mathbb{R}^7$ in (8). Intuitively, if $\epsilon < 1/2$, the second-order term is a subset of vectors in $\mathbb{R}^7$ with positive entries. This means that the "separation" between the mutual information quantities $I(U_{\mathcal{A}}; Y|U_{\mathcal{A}^c}, S_{\mathcal{A}^c})$ and the entropy quantities $H(S_{\mathcal{A}}|S_{\mathcal{A}^c})$ has to be sufficiently large for (n, ϵ)-transmissibility of the source (S_1, S_2, S_3). We observe that the rate at which the difference of the mutual information and entropy vectors can converge to zero is $\Theta(1/\sqrt{n})$, which is in line with the central limit theorem (more precisely the multidimensional Berry-Essèen theorem [20]).

Observe that the exact shape of the region depends on the covariance of the difference between the information density vector and the entropy density vector given in (9). This is different from the recent works on finite blocklength (or dispersion) analysis on (point-to-point) lossy joint source-channel coding [16], [17] where the second-order dispersion term is in fact a sum of two standard deviations—that of the D-tilted information of the source [14], [17] and the information density of the channel. Intuitively, this results from the fact that the source is independent of the channel noise and hence the variance of the sum is equal to the sum of the variances. However, our coding scheme and analysis of error probability used to prove Theorem 1 and Theorem 2 (below) closely parallel the original ones in [1], [2] for the joint source-channel coding problem over a MAC. Hence, we observe that our equivalent dispersion matrix $\mathbf{V}(U_1, U_2, U_3)$ is the covariance of the difference between the information density vector and the entropy density vector.

We remark that Theorem 1 can be strengthened slightly by employing a time-sharing procedure for the MAC part. This would yield a similar result but instead of the mutual information quantities in (8), we have $I(U_{\mathcal{A}}; Y|U_{\mathcal{A}^c}, S_{\mathcal{A}^c}, Q)$ where Q is a finite alphabet time-sharing random variable. The covariance matrix $\mathbf{V}(U_1, U_2, U_3)$ in (9) will then be replaced by a *conditional* covariance matrix involving the auxiliary random variable Q, i.e., the information dispersion is now $\mathbf{V}(U_1, U_2, U_3|Q) := \sum_q p_Q(q)\mathbf{V}(U_1, U_2, U_3|Q = q)$. This is detailed in the achievability proof for the finite blocklength rate region for the discrete memoryless MAC derived recently by Huang and Moulin [21].

B. Specialization of Auxiliary System to Derive a Sufficient Condition for (n, ϵ)-Transmissibility of Sources over a MAC

Now, we specialize Theorem 1 to derive a sufficient condition for (n, ϵ)-transmissibility of the source (S, T) over the MAC W. We use the following identifications:

$$S_1 := S, \quad S_2 := K, \quad S_3 := T, \tag{10}$$

where K is the common part of S and T [6]. This new system is equivalent to the one considered by Cover-El Gamal-Salehi [1] for the problem of transmitting a source (S, T) almost losslessly over a MAC W. The specialization can be stated as follows:

Theorem 2 ((n, ϵ)-Transmissibility for the joint source (S, T) with common part K)**.** *If for sufficiently large* n, *there exists distributions* $p_C(c), p_{X_1|S,C}(x_1|s, c)$ *and* $p_{X_2|T,C}(x_2|t, c)$ *such that*

$$\begin{bmatrix} I(X_1; Y|X_2, T, C) - H(S|T) \\ I(X_2; Y|X_1, S, C) - H(T|S) \\ I(X_1, C; Y|X_2, T) - H(S|T) \\ I(X_2, C; Y|X_1, S) - H(T|S) \\ I(X_1, X_2; Y|K, C) - H(T, S|K) \\ I(X_1, X_2; Y) - H(S, T) \end{bmatrix} \in \frac{\mathscr{S}_{\mathbf{V}(C,X_1,X_2)}(\epsilon)}{\sqrt{n}} + \frac{\log n}{n}\mathbf{1}_6, \tag{11}$$

where the information dispersion matrix $\mathbf{V}(C, X_1, X_2)$ *is the covariance of the random vector*

$$\begin{bmatrix} \log \frac{p_{Y|X_1,X_2,T,C}(Y|X_1,X_2,T,C)}{p_{Y|X_2,T,C}(Y|X_2,T,C)} - \log \frac{1}{p_{S|T}(S|T)} \\ \log \frac{p_{Y|X_1,X_2,S,C}(Y|X_1,X_2,S,C)}{p_{Y|X_1,S,C}(Y|X_1,S,C)} - \log \frac{1}{p_{T|S}(T|S)} \\ \log \frac{p_{Y|X_1,X_2,T,C}(T|X_1,X_2,T,C)}{p_{Y|X_2,T}(Y|X_2,T)} - \log \frac{1}{p_{S|T}(S|T)} \\ \log \frac{p_{Y|X_1,X_2,S,C}(T|X_1,X_2,S,C)}{p_{Y|X_1,S}(Y|X_1,S)} - \log \frac{1}{p_{T|S}(T|S)} \\ \log \frac{p_{Y|X_1,X_2,K,C}(Y|X_1,X_2,K,C)}{p_{Y|K,C}(Y|K,C)} - \log \frac{1}{p_{S,T|K}(S,T|K)} \\ \log \frac{p_{Y|X_1,X_2}(Y|X_1,X_2)}{p_Y(Y)} - \log \frac{1}{p_{S,T}(S,T)} \end{bmatrix}, \quad (12)$$

then the source (S, T) *is* (n, ϵ)*-transmissible over* W.

Note that in (12), we have written the information densities in a somewhat redundant fashion. For example due to the Markov condition $Y - (X_1, X_2) - (C, S, T)$, the numerator in the first conditional information density is in fact $p_{Y|X_1,X_2,T,C}(y|x_1, x_2, t, c) = W(y|x_1, x_2)$ and similarly for the numerators of all the other three conditional information densities. However, for clarity, we choose to write the information densities as in (12) so that the reader observes the similarities between the second-order information dispersion term and the first-order term in (11).

The interpretation of Theorem 2 is similar to that of Theorem 1. It quantifies the difference between the mutual information quantities and the entropies such that (S, T) (with common part K) is (n, ϵ)-transmissible over the MAC W. Note that unlike in Ahlswede and Han's work, we cannot eliminate the third and fourth constraints in (12). This is because the first and second constraints do not imply the third and fourth respectively (unlike the first-order result) since the covariances and the blocklength n characterize the (n, ϵ)-transmissibility of the source.

IV. Proofs

A. Proof of Theorem 1

Assume that the random variables (S_1, S_2, S_3) and the channel W satisfy the sufficient conditions of Theorem 1. We demonstrate that there exists a code for the (n, ϵ)-transmissibility of the correlated sources over the MAC W. We follow the code construction in [2] and apply Gaussian approximations to the general information spectrum analysis presented by Iwata and Oohama in [12].

We will first consider a simpler setup where we encode the three sources $S_j, j \in [3]$ separately using encoders $\tilde{\phi}_n^j : \mathcal{S}_j^n \to \mathcal{U}_j^n, j \in [3]$. The single decoder is $\tilde{\psi}_n : \mathcal{Y}^n \to \mathcal{S}_1^n \times \mathcal{S}_2^n \times \mathcal{S}_3^n$. At the end of the proof, we convert this problem to the one in Theorem 1. Fix the distributions $p_{U_j|S_j}, j \in [3]$. Also fix functions $f_1 : \mathcal{U}_1 \times \mathcal{U}_2 \to \mathcal{X}_1$ and $f_2 : \mathcal{U}_2 \times \mathcal{U}_3 \to \mathcal{X}_2$.

Codebook Generation: For $j \in [3]$ and to every sequence $s_j^n \in \mathcal{S}_j^n$, randomly and independently generate $u_j^n(s_j^n) \in \mathcal{U}_j^n$ from the product distribution $\prod_{i=1}^n p_{U_j|S_j}(u_j|s_j)$.

Encoding: Given s_j^n, encoder j transmits the sequence $u_j^n(s_j^n)$.

Decoding: Define the set

$$\mathscr{T}_{\delta_n}^{(n)}(\mathcal{A}) := \Big\{ (s_1^n, s_2^n, s_3^n, u_1^n, u_2^n, u_3^n, y^n) : \\ \frac{1}{n} \log \frac{p_{Y^n|U_{\mathcal{A}}^n, U_{\mathcal{A}^c}^n, S_{\mathcal{A}^c}^n}(y^n|u_{\mathcal{A}}^n(s_{\mathcal{A}}^n), u_{\mathcal{A}^c}^n(s_{\mathcal{A}^c}^n), s_{\mathcal{A}^c}^n)}{p_{Y^n|U_{\mathcal{A}}^n, U_{\mathcal{A}^c}^n}(y^n|u_{\mathcal{A}}^n(s_{\mathcal{A}}^n), u_{\mathcal{A}^c}^n(s_{\mathcal{A}^c}^n))} \\ - \frac{1}{n} \log \frac{1}{p_{S_{\mathcal{A}}^n|S_{\mathcal{A}^c}^n}(s_{\mathcal{A}}^n|s_{\mathcal{A}^c}^n)} \geq \delta_n \Big\}, \quad (13)$$

as well as the set

$$\mathscr{T}_{\delta_n}^{(n)} := \bigcap_{\emptyset \neq \mathcal{A} \subset [3]} \mathscr{T}_{\delta_n}^{(n)}(\mathcal{A}). \quad (14)$$

Intuitively, the set $\mathscr{T}_{\delta_n}^{(n)}$ takes the role of the typical set [7] but instead of having probability tending to one as n grows, we show that $\mathbb{P}((S_{[3]}^n, U_{[3]}^n(S_{[3]}^n), Y^n) \in \mathscr{T}_{\delta_n}^{(n)}) \approx 1 - \epsilon$. See the steps leading to (27).

Given y^n, the decoder searches for the unique triple of sequences $(\hat{s}_1^n, \hat{s}_2^n, \hat{s}_3^n)$ such that

$$(\hat{s}_1^n, \hat{s}_2^n, \hat{s}_3^n, u_1^n(\hat{s}_1^n), u_2^n(\hat{s}_2^n), u_3^n(\hat{s}_3^n), y^n) \in \mathscr{T}_{\delta_n}^{(n)}. \quad (15)$$

If there is no such unique triple, declare a decoding error. Note that this decoder is different from the jointly typical decoder in [1] and [2]. Choose the sequence

$$\delta_n := \frac{\log n}{2n}. \quad (16)$$

Analysis of Error Probability: We now show that for the random code constructed above and the decoding scheme with $\delta_n = \frac{\log n}{2n}$, the error probability (averaged over the random code ensemble) is no larger than ϵ. Assuming that S_1^n, S_2^n, S_3^n are sent to the encoder, the error events are

$$\mathcal{E}_0 := \{(S_1^n, S_2^n, S_3^n, \\ U_1^n(S_1^n), U_2^n(S_2^n), U_3^n(S_3^n), Y^n) \notin \mathscr{T}_{\delta_n}^{(n)}\} \quad (17)$$

$$\mathcal{E}_{\mathcal{A}} := \{\exists\, \tilde{s}_{\mathcal{A}}^n \neq S_{\mathcal{A}}^n : (\tilde{s}_{\mathcal{A}}^n, S_{\mathcal{A}^c}^n, \\ U_{\mathcal{A}}^n(\tilde{s}_{\mathcal{A}}^n), U_{\mathcal{A}^c}^n(S_{\mathcal{A}^c}^n), Y^n) \in \mathscr{T}_{\delta_n}^{(n)}\}, \quad (18)$$

for all subsets $\emptyset \neq \mathcal{A} \subset [3]$. Clearly,

$$\mathbb{P}\big[(\hat{S}_1^n, \hat{S}_2^n, \hat{S}_3^n) \neq (S_1^n, S_2^n, S_3^n)\big] \leq \mathbb{P}(\mathcal{E}_0) + \sum_{\emptyset \neq \mathcal{A} \subset [3]} \mathbb{P}(\mathcal{E}_{\mathcal{A}}). \quad (19)$$

We bound these error events separately. We will prove that $\mathbb{P}(\mathcal{E}_0) \approx 1 - \epsilon$ and $\mathbb{P}(\mathcal{E}_{\mathcal{A}}) \leq \frac{1}{\sqrt{n}}$. First, consider the probability of the complement of $\mathcal{E}_0$

$$\mathbb{P}(\mathcal{E}_0^c) = \mathbb{P}\big[(S_{[3]}^n, U_{[3]}^n(S_{[3]}^n), Y^n) \in \mathscr{T}_{\delta_n}^{(n)}\big]. \quad (20)$$

Since the source is independent and identically distributed (i.i.d.), the generation of the U_j^n-codewords is conditionally i.i.d. given the sources S_j^n and the channel is memoryless, we see that the probability in (20) can be expressed as

$$\mathbb{P}(\mathcal{E}_0^c) = \mathbb{P}\left[\frac{1}{n} \sum_{k=1}^n (\mathbf{i}_k - \mathbf{h}_k) \geq \delta_n \mathbf{1}_7\right] \quad (21)$$

where $\mathbf{i}_k$ and $\mathbf{h}_k$ are random vectors of information densities and entropy densities of the form in (13) evaluated at the k-th sample $(S_{1k}, S_{2k}, S_{3k}, U_{1k}(S_{1k}), U_{2k}(S_{2k}), U_{3k}(S_{3k}), Y_k)$. The mean of $\mathbf{i}_k$ (resp. $\mathbf{h}_k$) is the vector of mutual information (resp. entropy) quantities in (8). For notational convenience, denote the mutual information (resp. entropy) vector as $\mathbf{I} \in \mathbb{R}^7_+$ (resp. $\mathbf{H} \in \mathbb{R}^7_+$). Subtracting these means in (21), we obtain

$$\mathbb{P}(\mathcal{E}_0^c) = \mathbb{P}\left[\frac{1}{n}\sum_{k=1}^n \mathbf{j}_k \geq \mathbf{H} - \mathbf{I} + \delta_n \mathbf{1}_7\right], \tag{22}$$

where the random vector $\mathbf{j}_k := \mathbf{i}_k - \mathbf{h}_k - (\mathbf{I} - \mathbf{H})$. Multiplying throughout by $-\sqrt{n}$ and using the definition of δ_n in (16), we get

$$\mathbb{P}(\mathcal{E}_0^c) = \mathbb{P}\left[\frac{1}{\sqrt{n}}\sum_{k=1}^n -\mathbf{j}_k \leq \sqrt{n}(\mathbf{I} - \mathbf{H}) - \frac{\log n}{2\sqrt{n}}\mathbf{1}_7\right]. \tag{23}$$

Now, note that the random vectors $-\mathbf{j}_k, k \in \{1, \dots, n\}$, are i.i.d., zero-mean and have covariance matrix $\mathbf{V}(U_1, U_2, U_3) \in \mathbb{S}^7_+$. We can thus apply the multidimensional Berry-Essèen theorem [20] to get the following bound:

$$\mathbb{P}(\mathcal{E}_0^c) \geq \mathbb{P}\left[\mathbf{Z} \leq \sqrt{n}(\mathbf{I} - \mathbf{H}) - \frac{\log n}{2\sqrt{n}}\mathbf{1}_7\right] - O\left(\frac{1}{\sqrt{n}}\right), \tag{24}$$

where $\mathbf{Z} \sim \mathcal{N}(\mathbf{0}, \mathbf{V}(U_1, U_2, U_3))$ Note that the random vector $-\mathbf{j}_1$ is distributed in the same way as that in (9). In (24), we have assumed that the covariance matrix $\mathbf{V}(U_1, U_2, U_3)$ is non-singular and the third moment of the random vector $-\mathbf{j}_1$ is finite. The first assumption here can be dispensed with with an appropriate change of basis to a lower dimensional subspace (see [11] for details). The second assumption holds because all alphabets are finite. Since the sources (S_1, S_2, S_3) and the channel W satisfy the sufficient condition of Theorem 1, i.e., that

$$\mathbb{P}\left[\mathbf{Z} \leq \sqrt{n}(\mathbf{I} - \mathbf{H}) - \frac{\log n}{\sqrt{n}}\mathbf{1}_7\right] \geq 1 - \epsilon \tag{25}$$

by the definition of the set $\mathscr{S}_{\mathbf{V}(U_1,U_2,U_3)}(\epsilon)$, by Taylor expansion of (24), we conclude that as n grows,

$$\mathbb{P}(\mathcal{E}_0^c) \geq 1 - \epsilon + O\left(\frac{\log n}{\sqrt{n}}\right), \tag{26}$$

or equivalently that

$$\mathbb{P}(\mathcal{E}_0) \leq \epsilon - O\left(\frac{\log n}{\sqrt{n}}\right). \tag{27}$$

We now bound $\mathbb{P}(\mathcal{E}_{\mathcal{A}})$. Note from the definition of $\mathscr{T}_{\delta_n}^{(n)}$ in (14) that

$$\mathscr{T}_{\delta_n}^{(n)} \subset \mathscr{T}_{\delta_n}^{(n)}(\mathcal{A}) \tag{28}$$

for every subset $\mathcal{A}$. By monotonicity of measure, it suffices to consider the upper bound

$$\begin{aligned}\mathbb{P}(\mathcal{E}_{\mathcal{A}}) \leq \mathbb{P}\big[\exists\, \tilde{s}_{\mathcal{A}}^n \neq S_{\mathcal{A}}^n : (&\tilde{s}_{\mathcal{A}}^n, S_{\mathcal{A}^c}^n, \\ &U_{\mathcal{A}}^n(\tilde{s}_{\mathcal{A}}^n), U_{\mathcal{A}^c}^n(S_{\mathcal{A}^c}^n), Y^n) \in \mathscr{T}_{\delta_n}^{(n)}(\mathcal{A})\big].\end{aligned} \tag{29}$$

Now, we upper bound the RHS of (29) by conditioning on the events $\{S_{\mathcal{A}}^n = s_{\mathcal{A}}^n\}$ for all $s_{\mathcal{A}}^n \in \mathcal{S}_{\mathcal{A}}^n$ and applying the union-of-events bound:

$$\begin{aligned}\mathbb{P}(\mathcal{E}_{\mathcal{A}}) \leq \sum_{s_{\mathcal{A}}^n} p_{S_{\mathcal{A}}^n}(s_{\mathcal{A}}^n) \sum_{\tilde{s}_{\mathcal{A}}^n \neq s_{\mathcal{A}}^n} &\mathbb{P}\big[(\tilde{s}_{\mathcal{A}}^n, S_{\mathcal{A}^c}^n, \\ U_{\mathcal{A}}^n(\tilde{s}_{\mathcal{A}}^n), U_{\mathcal{A}^c}^n(S_{\mathcal{A}^c}^n), Y^n) &\in \mathscr{T}_{\delta_n}^{(n)} | S_{\mathcal{A}}^n = s_{\mathcal{A}}^n\big]\end{aligned} \tag{30}$$

Now, note that conditioned on $S_{\mathcal{A}}^n = s_{\mathcal{A}}^n$ [7, pp. 341],

$$\begin{aligned}&(S_{\mathcal{A}^c}^n, U_{\mathcal{A}}^n(\tilde{s}_{\mathcal{A}}^n), U_{\mathcal{A}^c}^n(S_{\mathcal{A}^c}^n), Y^n) | \{S_{\mathcal{A}}^n = s_{\mathcal{A}}^n\} \\ &\sim p_{S_{\mathcal{A}^c}^n, U_{\mathcal{A}^c}^n, Y^n | S_{\mathcal{A}}^n}(s_{\mathcal{A}^c}^n, u_{\mathcal{A}^c}^n, y^n | s_{\mathcal{A}}^n) p_{U_{\mathcal{A}}^n | S_{\mathcal{A}}^n}(u_{\mathcal{A}}^n | \tilde{s}_{\mathcal{A}}^n)\end{aligned} \tag{31}$$

Also, note that when $(s_1^n, s_2^n, s_3^n, u_1^n, u_2^n, u_3^n, y^n) \in \mathscr{T}_{\delta_n}^{(n)}(\mathcal{A})$, the sequences satisfy

$$\begin{aligned}&p_{S_{\mathcal{A}^c}^n, U_{\mathcal{A}^c}^n, Y^n}(s_{\mathcal{A}^c}^n, u_{\mathcal{A}^c}^n, y^n) p_{U_{\mathcal{A}}^n | S_{\mathcal{A}}^n}(u_{\mathcal{A}}^n | s_{\mathcal{A}}^n) \\ &\leq p_{Y^n | U_{\mathcal{A}}^n, U_{\mathcal{A}^c}^n, S_{\mathcal{A}^c}^n}(y^n | u_{\mathcal{A}}^n, u_{\mathcal{A}^c}^n, s_{\mathcal{A}^c}^n) p_{U_{\mathcal{A}^c}^n | S_{\mathcal{A}^c}^n}(u_{\mathcal{A}^c}^n | s_{\mathcal{A}^c}^n) \\ &\qquad p_{U_{\mathcal{A}}^n, S_{\mathcal{A}}^n}(u_{\mathcal{A}}^n, s_{\mathcal{A}}^n) p_{S_{\mathcal{A}^c}^n | S_{\mathcal{A}}^n}(s_{\mathcal{A}^c}^n | s_{\mathcal{A}}^n) \exp(-n\delta_n).\end{aligned} \tag{32}$$

This relation can be easily verified by repeatedly applying Bayes' rule to (13). As such, we have the following standard change of measure (Chernoff bound) argument used in information spectrum analysis [18]:

$$\begin{aligned}\mathbb{P}(\mathcal{E}_{\mathcal{A}}) \leq &\sum_{(\tilde{s}_{\mathcal{A}}^n, s_{\mathcal{A}^c}^n, u_{\mathcal{A}}^n, u_{\mathcal{A}^c}^n, y^n) \in \mathscr{T}_{\delta_n}^{(n)}(\mathcal{A})} \\ &p_{S_{\mathcal{A}^c}^n, U_{\mathcal{A}^c}^n, Y^n}(s_{\mathcal{A}^c}^n, u_{\mathcal{A}^c}^n, y^n) p_{U_{\mathcal{A}}^n | S_{\mathcal{A}}^n}(u_{\mathcal{A}}^n | \tilde{s}_{\mathcal{A}}^n) &(33) \\ \leq &\sum_{(\tilde{s}_{\mathcal{A}}^n, s_{\mathcal{A}^c}^n, u_{\mathcal{A}}^n, u_{\mathcal{A}^c}^n, y^n) \in \mathscr{T}_{\delta_n}^{(n)}(\mathcal{A})} \\ &p_{Y^n | U_{\mathcal{A}}^n, U_{\mathcal{A}^c}^n, S_{\mathcal{A}^c}^n}(y^n | u_{\mathcal{A}}^n, u_{\mathcal{A}^c}^n, s_{\mathcal{A}^c}^n) p_{U_{\mathcal{A}^c}^n | S_{\mathcal{A}^c}^n}(u_{\mathcal{A}^c}^n | s_{\mathcal{A}^c}^n) \\ &p_{U_{\mathcal{A}}^n, S_{\mathcal{A}}^n}(u_{\mathcal{A}}^n, s_{\mathcal{A}}^n) p_{S_{\mathcal{A}^c}^n | S_{\mathcal{A}}^n}(s_{\mathcal{A}^c}^n | s_{\mathcal{A}}^n) \exp(-n\delta_n) &(34) \\ \leq &\exp(-n\delta_n), &(35)\end{aligned}$$

where (33) follows by inserting (31) into (30) and marginalizing over $s_{\mathcal{A}}^n$ and (34) follows from (32). From the definition of δ_n in (16), we can assert that

$$\mathbb{P}(\mathcal{E}_{\mathcal{A}}) \leq \exp\left(-n \cdot \frac{\log n}{2n}\right) = \frac{1}{\sqrt{n}}. \tag{36}$$

The above analysis clearly applies to all events $\mathcal{E}_{\mathcal{A}}$. Combining this with (27), we see that the overall probability of error in (19) can be upper bounded as

$$\mathbb{P}(\mathcal{E}) \leq \epsilon - O\left(\frac{\log n}{\sqrt{n}}\right), \tag{37}$$

as n grows. This proves the existence of a length-n code whose probability of error is no larger than ϵ (if n is sufficiently large).

To complete the proof of Theorem 1, we just simply define the functions $\phi_n^1(s_1^n, s_2^n) := f_1(\tilde{\phi}_n^1(s_1^n), \tilde{\phi}_n^2(s_2^n))$ and $\phi_n^2(s_2^n, s_3^n) := f_2(\tilde{\phi}_n^2(s_2^n), \tilde{\phi}_n^3(s_3^n))$. Clearly, encoders ϕ_n^1, ϕ_n^2 and decoder ψ_n yield the same probability as the code given by $(\tilde{\phi}_n^1, \tilde{\phi}_n^2, \tilde{\phi}_n^3, \tilde{\psi}_n)$. $\square$

B. Proof of Theorem 2

Theorem 2 follows by specializing Theorem 1 by the identifications in (10) identifying U_2 to be the auxiliary random variable representing the common part C. Then, we have the condition

$$\mathbf{I} - \mathbf{H} \in \frac{\mathscr{S}_{\mathbf{V}}(\epsilon)}{\sqrt{n}} + \frac{\log n}{n}\mathbf{1}_7 \tag{38}$$

where $\mathbf{I} \in \mathbb{R}_+^7$ and $\mathbf{H} \in \mathbb{R}_+^7$ are the mutual information and entropy vectors as per equations (2.8)-(2.14) in [2]. Since we do not have to recover $K = S_2$ in our setting, the third constraint in (38) is redundant. Also note that because $H(K|S,T) = 0$. Thus, the seven-dimensional region collapses to a six-dimensional region. By using the standard Markov relations, we see, in the same way as in [2], that the information and entropy densities in Theorem 2 are the same as those in Theorem 1. □

V. Conclusions and Further Work

In this work, we used Gaussian approximation-based analysis techniques to derive sufficient conditions (achievability results) for transmitting a correlated discrete memoryless source over a MAC. Theorem 2 provides a sufficient condition and the condition is based on the so-called information dispersion matrix, which is the covariance matrix of the difference between an information density vector and an entropy density vector. Different from the result of Cover-El Gamal-Salehi [1], there are six constraints on the differences between the mutual information and entropy quantities, whereas in [1] there are only four constraints.

There are at least two directions for further research: Firstly, we would like to understand whether a second-order coding result analogous to [16], [17] can be derived using different coding schemes. Secondly, we would like to derive a converse result for the problem perhaps based on the result in [8].

Acknowledgements

I would like to acknowledge Sun Sumei (I2R) for suggesting this problem and providing comments on the manuscript.

References

[1] T. Cover, A. El Gamal, and M. Salehi, "Multiple access channels with arbitrarily correlated sources," *IEEE Trans. on Inf. Th.*, vol. 26, no. 6, pp. 648–57, June 1980.

[2] R. Ahlswede and T. S. Han, "On source coding with side information via a multiple-access channel and related problems in multi-user information theory." *IEEE Trans. on Inf. Th.*, vol. 29, no. 3, pp. 396–412, Mar 1983.

[3] D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. on Inf. Th.*, vol. 19, pp. 471–80, 1973.

[4] H. H. J. Liao, "Multiple access channels," Ph.D. dissertation, University of Hawaii, Honolulu, 1972.

[5] R. Ahlswede, "Multiway communication channels," in *Intl. Symp. on Inf. Th.*, Tsahkadsor, Armenian S.S.R, 1971, pp. 23–52.

[6] P. Gács and J. Körner, "Common information is far less than mutual information," *Prob. Contr. Inform.*, vol. 2, pp. 149–162, 1973.

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

[8] W. Kang and S. Ulukus, "An outer bound for multiple access channels with correlated sources," in *40th Annual Conference on Information Sciences and Systems*, 2006.

[9] M. Hayashi, "Information spectrum approach to second-order coding rate in channel coding," *IEEE Trans. on Inf. Th.*, vol. 55, pp. 4947–66, Nov 2009.

[10] Y. Polyanskiy, H. V. Poor, and S. Verdú, "Channel coding in the finite blocklength regime," *IEEE Trans. on Inf. Th.*, vol. 56, pp. 2307 – 59, May 2010.

[11] V. Y. F. Tan and O. Kosut, "On the dispersions of three network information theory problems," *arXiv:1201.3901*, Feb 2012, [Online].

[12] K. Iwata and Y. Oohama, "Information-Spectrum Characterization of Multiple-Access Channels with Correlated Sources," *IEICE Transactions on Fundamentals of Electronics Communications and Computer Sciences*, vol. E88-A, no. 11, pp. 3196–3202, 2005.

[13] V. Strassen, "Asymptotische Abschätzungen in Shannons Informationstheorie," in *Trans. Third. Prague Conf. Inf. Th.*, 1962, pp. 689–723.

[14] V. Kostina and S. Verdú, "Fixed-length lossy compression in the finite blocklength regime: Discrete memoryless sources," in *Int. Symp. Inf. Th.*, 2011.

[15] A. Ingber and Y. Kochman, "The dispersion of lossy source coding," in *Data Compression Conference (DCC)*, 2011.

[16] D. Wang, A. Ingber, and Y. Kochman, "The dispersion of joint source-channel coding," in *Allerton Conference*, 2011, arXiv:1109.6310.

[17] V. Kostina and S. Verdú, "Lossy joint source-channel coding in the finite blocklength regime," in *Int. Symp. Inf. Th.*, 2012.

[18] T. S. Han, *Information-Spectrum Methods in Information Theory*. Springer Berlin Heidelberg, Feb 2010.

[19] A. T. Campo, G. Vazquez-Vilar, A. Guillén i Fàbregas, and A. Martinez, "Random-coding bounds joint source-channel bounds," in *Int. Symp. Inf. Th.*, 2011.

[20] V. Bentkus, "On the dependence of the Berry-Esseen bound on dimension," *J. Stat. Planning and Inference*, vol. 113, pp. 385 – 402, 2003.

[21] Y.-W. Huang and P. Moulin, "Finite blocklength coding for multiple access channels," in *Int. Symp. Inf. Th.*, 2012.