Typicality Graphs and Their Properties

Ali Nazari, Dinesh Krithivasan, S. Sandeep Pradhan, Achilleas Anastasopoulos
Dept. of Electrical Engineering and computer science
University of Michigan, Ann Arbor
E-mail: {anazari,dineshk,pradhanv,anastas}@umich.edu

Ramji Venkataramanan
Dept. of Electrical Engineering
Stanford University
E-mail: vramji@stanford.edu

***Abstract*—Let $\mathcal{X}$ and $\mathcal{Y}$ be finite alphabets and P_{XY} a joint distribution over them, with P_X and P_Y representing the marginals. For any $\epsilon > 0$, the set of n-length sequences x^n and y^n that are jointly typical [1] according to P_{XY} can be represented on a bipartite graph. We present a formal definition of such a graph, known as a *typicality* graph, and study some of its properties. These properties arise in the study of several multiuser communication problems.**

I. Introduction

The concept of typicality and typical sequences is central to information theory. It has been used to develop computable performance limits for several communication problems.

Consider a pair of correlated discrete memoryless information sources[1] X and Y characterized by a generic joint distribution p_{XY} defined on the product of two finite sets $\mathcal{X} \times \mathcal{Y}$. An length n X-sequence x^n is typical if the empirical histogram of x^n is close to p_X. A pair of length n sequences $(x^n, y^n) \in \mathcal{X}^n \times \mathcal{Y}^n$ is said to be jointly typical if the empirical joint histogram of (x^n, y^n) is close to the joint distribution p_{XY}. The set of all jointly typical sequence pairs is called the typical set of p_{XY}.

Given a sequence length n, the typical set can be represented in terms of the following undirected, bipartite graph. The left vertices of the graph are all the typical X-sequences, and the right vertices are all the typical Y-sequences. From well-known properties of typical sets, there are (approximately) $2^{nH(X)}$ left vertices and $2^{nH(Y)}$ right vertices. A left vertex is connected to a right vertex through an edge if the corresponding X and Y-sequences are *jointly* typical. From the properties of joint typicality, we know that the number of edges in this graph is roughly $2^{nH(X,Y)}$. Further, every left vertex (a typical X-sequence) has degree roughly equal to $2^{nH(Y|X)}$, i.e., it is jointly typical with $2^{nH(Y|X)}$ Y-sequences. Similarly, each right vertex has degree roughly equal to $2^{nH(X|Y)}$.

In this paper we formally characterize the typicality graph and look at some subgraph containment problems. In particular, we answer three questions concerning the typicality graph:

- When can we find subgraphs such that the left and right vertices of the subgraph have specified degrees, say R'_X and R'_Y, respectively ?
- What is the maximum size of subgraphs that are complete, i.e., every left vertex is connected to every right vertex? One of the main contributions of this paper is a sharp answer to this question.
- If we create a subgraph by randomly picking a specified number of left and right vertices, what is the probability that this subgraph has far fewer edges than expected?

These questions arise in a variety of multiuser communication problems. Transmitting correlated information over a multiple-access channel (MAC) [2], and communicating over a MAC with feedback [3] are two problems where the first question plays an important role. The techniques used to answer the second question have been used to develop tighter bounds on the error exponents of discrete memoryless multiple-access channels [4], [5], [6]. The third question arises in the context of transmitting correlated information over a broadcast channel [7]. Moreover, the evaluation of performance limits of a multiuser communication problem can be thought of as characterizing certain properties of typicality graphs of random variables associated with the problem.

The paper is organized as follows. Some definitions are introduced in Section II. In section III, typicality graphs are formally defined and some of their properties (regarding the number of vertices, edges, and degrees) are stated. The main results of the paper- four propositions concerning subgraphs of typicality graphs- are presented in section IV. The proof of one of these results is given in the Appendix. Due to constraints of space, the other proofs will be presented in an extended version of this paper.

II. Preliminaries

We first review the definitions of δ-typical sets and their properties from [1].

Definition 1. *For any distribution P on $\mathcal{X}$, A sequence $x^n \in \mathcal{X}^n$ is called X-typical with constant δ if*

1) $|\frac{1}{n}N(a|x^n) - P_X(a)| \le \delta, \quad \forall a \in \mathcal{X}$
2) *No $a \in \mathcal{X}$ with $P_X(a) = 0$ occurs in x^n.*

The set of such sequences is denoted by $T^n_\delta(P_X)$ or $T^n_\delta(X)$, when the distribution being used is unambiguous.

Definition 2. *Given a conditional distribution $P_{Y|X} : \mathcal{X} \to \mathcal{Y}$, a sequence $y^n \in \mathcal{Y}^n$ is conditionally $P_{Y|X}$-typical with $x^n \in \mathcal{X}^n$ with constant δ if*

1) $|\frac{1}{n}N(a, b|x^n, y^n) - \frac{1}{n}N(a|x^n)P_{Y|X}(b|a)| \le \delta, \quad \forall a \in \mathcal{X}, b \in \mathcal{Y}$.

[1] We use the following notation throughout this work. Script capitals $\mathcal{U}$, $\mathcal{X}$, $\mathcal{Y}$, $\mathcal{Z}$,... denote finite, nonempty sets. To show the cardinality of a set $\mathcal{X}$, we use $|\mathcal{X}|$. We also use the letters P, Q,... for probability distributions on finite sets, and U, X, Y,... for random variables.

2) $N(a, b|x^n, y^n) = 0$ *whenever* $P_{Y|X}(b|a) = 0$.

The set of such sequences is denoted by $T_\delta^n(P_{Y|X}|x^n)$ *or* $T_\delta^n(Y|x^n)$*, when the distribution being used is unambiguous.*

We will repeatedly use the following results, which we state below as facts:

Fact 1. *[1, Lemma 2.10]: (a) If* $x^n \in T_\delta^n(X)$ *and* $y^n \in T_{\delta'}^n(Y|x^n)$*, then* $(x^n, y^n) \in T_{\delta+\delta'}^n(X, Y)$ *and* $y^n \in T_{(\delta+\delta')|\mathcal{X}|}^n(Y)$. [2]

(b) If $x^n \in T_\delta^n(X)$ *and* $(x^n, y^n) \in T_\epsilon^n(X, Y)$*, then* $y^n \in T_{\delta+\epsilon}^n(Y|x^n)$.

Fact 2. *[1, Lemma 2.13]* [3]*: There exists a sequence* $\epsilon_n \to 0$ *depending only on* $|\mathcal{X}|$ *and* $|\mathcal{Y}|$ *such that for every joint distribution* $P_X \cdot P_{Y|X}$ *on* $\mathcal{X} \times \mathcal{Y}$,

$$\left|\frac{1}{n}\log|T^n(X)| - H(X)\right| \leq \epsilon_n \tag{1}$$

$$\left|\frac{1}{n}\log|T^n(Y|x^n)| - H(Y|X)\right| \leq \epsilon_n, \quad \forall x^n \in T^n(X). \tag{2}$$

The next fact deals with the continuity of entropy with respect to probability distributions.

Fact 3. *[1, Lemma 2.7] If* P *and* Q *are two distributions on* X *such that*

$$\sum_{x\in\mathcal{X}} |P(x) - Q(x)| \leq \epsilon \leq \frac{1}{2}$$

then

$$|H(P) - H(Q)| \leq -\epsilon \log \frac{\epsilon}{|\mathcal{X}|}$$

III. Typicality graphs

Consider any joint distribution $P_X \cdot P_{Y|X}$ on $\mathcal{X} \times \mathcal{Y}$. Consider sequences $\epsilon_{1n}, \epsilon_{2n}, \lambda_n$ satisfying the 'delta convention' property [1, Convention 2.11], i.e.,

$$\epsilon_{1n} \to 0, \quad \sqrt{n}\cdot\epsilon_{1n} \to \infty \text{ as } n \to \infty, \tag{3}$$

$$\epsilon_{2n} \to 0, \quad \sqrt{n}\cdot\epsilon_{2n} \to \infty \text{ as } n \to \infty, \tag{4}$$

$$\lambda_n \to 0, \quad \sqrt{n}\cdot\lambda_n \to \infty \text{ as } n \to \infty. \tag{5}$$

The delta convention ensures that the typical sets have 'large probability'.

Definition 3. *For every* n, $G_n(\epsilon_{1n}, \epsilon_{2n}, \lambda_n)$ *is defined as a bipartite graph, with its left vertices consisting of all* $x^n \in T_{\epsilon_{1n}}^n(X)$ *and the right vertices consisting of all* $y^n \in T_{\epsilon_{2n}}^n(Y)$*. A vertex on the left (say* $\tilde{x}^n$*) is connected to a vertex on the right (say* $\tilde{y}^n$*) if and only if* $(\tilde{x}^n, \tilde{y}^n) \in T_{\lambda_n}^n(X, Y)$.

We will use the notation $V_X(.), V_Y(.)$ to denote the vertex sets of any bipartite graph. From the facts mentioned in Section II, the following properties hold:

1) From Fact 2, we know that for any sequence of typicality graphs $\{G_n(\epsilon_{1n}, \epsilon_{2n}, \lambda_n)\}$, the cardinality of the vertex sets satisfies

$$\left|\frac{1}{n}\log|V_X(G_n)| - H(X)\right| \leq \epsilon_n, \tag{6}$$

$$\left|\frac{1}{n}\log|V_Y(G_n)| - H(Y)\right| \leq \epsilon_n, \tag{7}$$

for some sequence $\epsilon_n \to 0$.

2) The degree of each vertex $x^n \in V_X(G_n)$ and $y^n \in V_Y(G_n)$ satisfies

$$\text{degree}(x^n) \leq 2^{n(H(Y|X)+\epsilon_n)}, \tag{8}$$

$$\text{degree}(y^n) \leq 2^{n(H(X|Y)+\epsilon_n)}, \tag{9}$$

for some $\epsilon_n \to 0$.

Property 2 gives upper bounds on the degree of each vertex in the typicality graph. Since we have not imposed any relationships between the typicality constants $\epsilon_{1n}, \epsilon_{2n}$ and λ_n, in general it cannot be said that the degree of *every* X-vertex (resp. Y-vertex) is close to $2^{nH(Y|X)}$ (resp. $2^{nH(X|Y)}$). However, such an assertion holds for *almost* every vertex in G_n. Specifically, we can show that the above degree conditions hold for a subgraph with exponentially the same size as G_n.

Proposition 1. *Every sequence of typicality graphs* $G_n(\epsilon_{1n}, \epsilon_{2n}, \lambda_n)$ *has a sequence of subgraphs* $A_n(\epsilon_{1n}, \epsilon_{2n}, \lambda_n)$ *satisfying the following properties for some* $\delta_n \to 0$.

1) *The vertex set sizes* $|V_X(A_n)|$ *and* $|V_Y(A_n)|$*, denoted* θ_X^n *and* θ_Y^n*, respectively, satisfy*

$$\left|\frac{1}{n}\log\theta_X^n - H(X)\right| \leq \delta_n, \quad \forall n$$

$$\left|\frac{1}{n}\log\theta_Y^n - H(Y)\right| \leq \delta_n \quad \forall n. \tag{10}$$

2) *The degree of each* X*-vertex* x^n*, denoted* ${\theta'}^n(x^n)$ *satisfies*

$$\left|\frac{1}{n}\log{\theta'}^n(x^n) - H(Y|X)\right| \leq \delta_n \quad \forall x^n \in V_X(A_n). \tag{11}$$

3) *The degree of each* Y*-vertex* y^n*, denoted* ${\theta'}^n(y^n)$*, satisfies*

$$\left|\frac{1}{n}\log{\theta'}^n(y^n) - H(X|Y)\right| \leq \delta_n \quad \forall y^n \in V_Y(A_n). \tag{12}$$

Proof: The proof is provided in a more complete version [8]. ■

IV. Sub-graphs contained in typicality graphs

In this section, we study the subgraphs contained in a sequence of typicality graphs.

[2] The typical sets are with respect to distributions $P_X, P_{Y|X}$ and P_{XY}, respectively.

[3] The constants of the typical sets for each n, when suppressed, are understood to be some δ_n with $\delta_n \to 0$ and $\sqrt{n}\cdot\delta_n \to \infty$ (delta convention).

A. Subgraphs of general degree

Definition 4. *A sequence of typicality graphs $G_n(\epsilon_{1n}, \epsilon_{2n}, \lambda_n)$ is said to contain a sequence of subgraphs Γ_n of rates (R_X, R_Y, R'_X, R'_Y) if for each n, there exists a sequence $\delta_n \to 0$ such that*

1) *The vertex sets of the subgraphs have sizes (denoted Δ_X^n and Δ_Y^n) that satisfy*

$$\left|\frac{1}{n}\log\Delta_X^n - R_X\right| \le \delta_n, \quad \left|\frac{1}{n}\log\Delta_Y^n - R_Y\right| \le \delta_n, \forall n.$$

2) *The degree of each vertex x^n in $V_X(\Gamma_n)$, denoted ${\Delta'}^n(x^n)$ satisfies*

$$\left|\frac{1}{n}\log{\Delta'}^n(x^n) - R'_Y\right| \le \delta_n, \quad \forall x^n \in V_X(\Gamma_n),\ \forall n.$$

3) *The degree of each vertex y^n in the $V_Y(\Gamma_n)$, denoted ${\Delta'}^n(y^n)$ satisfies*

$$\left|\frac{1}{n}\log{\Delta'}^n(y^n) - R'_X\right| \le \delta_n, \quad \forall y^n \in V_Y(\Gamma_n),\ \forall n.$$

The following proposition gives a characterization of the rate-tuple of a sequence of subgraphs in the sequence of typicality graphs of P_{XY}.

Proposition 2. *Let $G_n(\epsilon_{1n}, \epsilon_{2n}, \lambda_n)$ be a sequence of typicality graphs of P_{XY}. Let us define $\mathcal{R}$ as all tuples, (R_X, R_Y, R'_X, R'_Y), such that $G_n(\epsilon_{1n}, \epsilon_{2n}, \lambda_n)$ contains subgraphs of rates (R_X, R_Y, R'_X, R'_Y). Then,*

$$\mathcal{R}' \subseteq \mathcal{R} \tag{13}$$

where $\mathcal{R}'$ is defined as follows:

$$\mathcal{R}' \triangleq \big\{(R_X, R_Y, R'_X, R'_Y) : R_X \le H(X|U),\ R_Y \le H(Y|U),\ R'_X \le H(X|YU),\ R'_Y \le H(Y|XU) \text{ for some } P_{U|XY}\big\}.$$

Proof: The proof is provided in a more complete version [8]. ■

B. Nearly complete subgraphs

A complete bipartite graph is one in which each vertex of the first set is connected to every vertex on the other set. We next consider nearly complete subgraphs of the sequence of typicality graphs. For this class of subgraphs, we have a converse result that fully characterizes the set of nearly complete subgraphs present in any typicality graph.

Definition 5. *A sequence of typicality graphs $G_n(\epsilon_{1n}, \epsilon_{2n}, \lambda_n)$ is said to contain a sequence of nearly complete subgraphs Γ_n of rates (R_X, R_Y) if for each n, there exists a sequence $\delta_n \to 0$ such that*

1) *The sizes of the vertex sets of the subgraphs, denoted Δ_X^n and Δ_Y^n, satisfy*

$$\left|\frac{1}{n}\log\Delta_X^n - R_X\right| \le \delta_n, \quad \left|\frac{1}{n}\log\Delta_Y^n - R_Y\right| \le \delta_n, \forall n.$$

2) *The degree of each vertex x^n in the X-set of the subgraph, denoted ${\Delta'}^n(x^n)$ satisfies*

$$\frac{1}{n}\log{\Delta'}^n(x^n) \ge R_Y - \delta_n, \quad \forall x^n \in V_X(\Gamma_n),\ \forall n.$$

3) *The degree of each vertex y^n in the Y-set of the subgraph, denoted ${\Delta'}^n(y^n)$ satisfies*

$$\frac{1}{n}\log{\Delta'}^n(y^n) \ge R_X - \delta_n, \quad \forall y^n \in V_Y(\Gamma_n),\ \forall n.$$

Proposition 3. *Let $G_n(\epsilon_{1n}, \epsilon_{2n}, \lambda_n)$ be a sequence of typicality graphs for P_{XY}. Again, let us define $\mathcal{R}$ as all tuples, (R_X, R_Y), such that $G_n(\epsilon_{1n}, \epsilon_{2n}, \lambda_n)$ contains subgraphs of rates (R_X, R_Y). Then,*

$$\mathcal{R}'' \subseteq \mathcal{R} \tag{14}$$

where $\mathcal{R}''$ is defined as follows:

$$\mathcal{R}'' \triangleq \big\{(R_X, R_Y) : R_X \le H(X|U),\ R_Y \le H(Y|U),\ \text{for some } P_{U|XY} \text{ such that } X - U - Y\big\}$$

Moreover, For all sequences of nearly complete subgraphs of G_n such that the sequence δ_n (in Definition 5) converges to 0 faster than $1/\log n$ (i.e.,$\lim_{n\to\infty}\delta_n \log n = 0$), the rates of the subgraph (R_X, R_Y) satisfy

$$R_X \le H(X|U),\ R_Y \le H(Y|U)$$

for some $P_{U|XY}$ such that $X - U - Y$. [4] *In this case,*

$$\mathcal{R}'' = \mathcal{R} \tag{15}$$

Proof: The proof is provided in the Appendix. ■

C. Nearly Empty Subgraphs

So far, we have discussed properties of subgraphs of the typicality graph $G_n(\epsilon_{1n}, \epsilon_{2n}, \lambda_n)$ such as the containment of nearly complete subgraphs and subgraphs of general degree. We characterized these subgraphs based on the degrees of their vertices. We now turn our attention to the presence of nearly empty subgraphs in the typicality graph. We analyze the probability that a randomly chosen subgraph of the typicality graph has far fewer edges than expected. In particular, we focus attention on the case when the random subgraph has no edges.

Consider a pair (X, Y) of discrete memoryless stationary correlated sources with finite alphabets $\mathcal{X}$ and $\mathcal{Y}$ respectively. Suppose we sample 2^{nR_1} sequences from the typical set $T^n_{\epsilon_{1n}}(X)$ of X independently with replacement and similarly sample 2^{nR_2} sequences from the typical set $T^n_{\epsilon_{2n}}(Y)$ of Y. The underlying typicality graph $G_n(\epsilon_{1n}, \epsilon_{2n}, \lambda_n)$ induces a bipartite graph on these $2^{nR_1} + 2^{nR_2}$ sequences. We provide a characterization of the probability that this graph is sparser than expected.

With these preliminaries, we are ready to state the main result of this section.

[4] We use the notation $X - U - Y$ to indicate X, U, Y form a Markov chain in that order.

Proposition 4. *Suppose X and Y are correlated finite alphabet memoryless random variables with joint distribution $p(x,y)$. Let $\epsilon_{1n}, \epsilon_{2n}, \lambda_n$ satisfy the 'delta convention' and R_1, R_2 be any positive real numbers such that $R_1 + R_2 > I(X;Y)$. Let $\mathcal{C}_X$ be a collection of 2^{nR_1} sequences picked independently and with replacement from $T^n_{\epsilon_{1n}}(X)$ and let $\mathcal{C}_Y$ be defined similarly. Let U be the cardinality of the set*

$$\mathcal{U} \triangleq \{(x^n, y^n) \in \mathcal{C}_X \times \mathcal{C}_Y : (x^n, y^n) \in T^n_{\lambda_n}(X,Y)\} \quad (16)$$

Assume, without loss of generality that $R_1 \geq R_2$. Then, for any $\gamma \geq 0$, we have

$$\begin{aligned}\lim_{n\to\infty} \frac{1}{n} \log\log \left[\mathbb{P}\left(\frac{\mathbb{E}(U) - U}{\mathbb{E}(U)} \geq e^{-n\gamma}\right)\right]^{-1} \\ \geq \begin{cases} R_1 + R_2 - I(X;Y) - \gamma & \text{if } R_1 < I(X;Y) \\ R_2 - \gamma & \text{if } R_1 \geq I(X;Y) \end{cases}\end{aligned} \quad (17)$$

Setting $\gamma = 0$ in the above equation gives us

$$\lim_{n\to\infty} \frac{1}{n} \log\log \frac{1}{\mathbb{P}(U=0)} \geq \min\left(R_2, R_1 + R_2 - I(X;Y)\right) \quad (18)$$

This inequality holds with equality when $R_2 \leq R_1 \leq I(X;Y)$.

Proof: The proof can be obtained using a version of Suen's inequalities [9] and the Lovasz local lemma [10]. It is provided in a more complete version [8] ■

V. Appendix

The first part of the proposition follows directly from Proposition 2 by choosing $P_{U|XY}$ such that $X - U - Y$ form a Markov chain. We now prove the converse under the stated assumption that the sequence δ_n satisfies $\lim_{n\to\infty} \delta_n \log n = 0$.

Suppose that a sequence of typicality graphs $G_n(\epsilon_{1n}, \epsilon_{2n}, \lambda_n)$ contains nearly complete subgraphs Γ_n of rates R_X, R_Y. The total number of edges in Γ_n can be lower bounded as

$$\begin{aligned} |\text{Edges}(\Gamma_n)| &\geq \Delta^n_X \cdot \text{ min degree of a vertex in } V_X(\Gamma_n) \\ &\geq \Delta^n_X \cdot 2^{n(R_Y - \delta_n)} \\ &\geq \Delta^n_X \cdot 2^{n(R_Y - \delta_n)} \Delta^n_Y \cdot 2^{-n(R_Y + \delta_n)} \\ &= \Delta^n_X \cdot \Delta^n_Y \cdot 2^{-2n\delta_n}. \end{aligned} \quad (19)$$

Each of these edges represent a pair (x^n, y^n) that is jointly λ_n-typical with respect to the distribution P_{XY}. In other words, each of these pairs (x^n, y^n) belongs to a joint type[1] that is 'close' to P_{XY}. Since the number of joint types of a pair of sequences of length n is at most $(n+1)^{|\mathcal{X}||\mathcal{Y}|}$, the number of edges belonging to the dominant joint type, say $\bar{P}_{XY}$ satisfies

$$|\text{Edges}(\Gamma_n) \text{ having joint type } \bar{P}_{XY}| \geq \frac{\Delta^n_X \cdot \Delta^n_Y 2^{-2n\delta_n}}{(n+1)^{|\mathcal{X}||\mathcal{Y}|}}. \quad (20)$$

Define a subgraph $\mathcal{A}_n$ of Γ_n consisting only of the edges having joint type $\bar{P}_{XY}$. A word about the notation used in the sequel: We will use i, j to index the vertices in $V_X(\Gamma_n), V_Y(\Gamma_n)$, respectively. Thus $i \in \{1, \ldots, \Delta^n_X\}$ and $j \in \{1, \ldots, \Delta^n_Y\}$. The actual sequences corresponding to these vertices will be denoted $x^n(i), y^n(j)$ etc. Using this notation,

$$\begin{aligned}\mathcal{A}_n \triangleq \{(i,j) : i \in V_X(\Gamma_n), j \in V_Y(\Gamma_n) \\ \text{s.t. } (x^n(i), y^n(j)) \text{ has joint type } \bar{P}_{XY}\}\end{aligned} \quad (21)$$

From (20),

$$|\mathcal{A}_n| \geq \frac{\Delta^n_X \cdot \Delta^n_Y 2^{-2n\delta_n}}{(n+1)^{|\mathcal{X}||\mathcal{Y}|}} \quad (22)$$

We will prove the converse result using a series of lemmas concerning $\mathcal{A}_n$. Some of the lemmas are similar to those required to prove in [4, Theorem 1]. We only sketch the proofs of such lemmas, referring the reader to [4] for details.

Define random variables X'^n, Y'^n with pmf

$$\Pr((X'^n, Y'^n) = (x^n(i), y^n(j)) = \frac{1}{|\mathcal{A}_n|}, \ \text{ if } (i,j) \in \mathcal{A}_n. \quad (23)$$

Lemma 1. $I(X'^n; Y'^n) \leq 2n\delta_n + |\mathcal{X}||\mathcal{Y}| \log(n+1)$.

Proof: Follow steps similar to the proof of [4, Lemma 1], using (22) to lower bound the size of $\mathcal{A}_n$.

The next lemma is Ahlswede's version of the 'wringing' technique. Roughly speaking, if it is known that the mutual information between two random sequences is small, then the lemma gives an upper bound on the per-letter mutual information terms (conditioned on some values).

Lemma 2. *[11] Let A^n, B^n be RV's with values in $\mathcal{A}^n$, $\mathcal{B}^n$ resp. and assume that*

$$I(A^n; B^n) \leq \sigma$$

Then, for any $0 < \delta < \sigma$ there exist $t_1, t_2, ..., t_k \in \{1, ..., n\}$ where $0 \leq k < \frac{2\sigma}{\delta}$ such that for some $\bar{a}_{t_1}, \bar{b}_{t_1}, \bar{a}_{t_2}, \bar{b}_{t_2}, ..., \bar{a}_{t_k}, \bar{b}_{t_k}$

$$\begin{aligned} I(A_t; B_t | A_{t_1} = \bar{a}_{t_1}, B_{t_1} = \bar{b}_{t_1}, ..., A_{t_k} = \bar{a}_{t_k}, B_{t_k} = \bar{b}_{t_k}) \leq \delta \\ \textit{for } t = 1, 2, ..., n, \end{aligned} \quad (24)$$

and

$$\begin{aligned} Pr(A_{t_1} = \bar{a}_{t_1}, B_{t_1} = \bar{b}_{t_1}, ..., A_{t_k} = \bar{a}_{t_k}, B_{t_k} = \bar{b}_{t_k}) \\ \geq \left(\frac{\delta}{|\mathcal{A}||\mathcal{B}|(2\sigma - \delta)}\right)^k. \end{aligned} \quad (25)$$

In our case, we will apply Lemma 2 to random variables X'^n and Y'^n. Lemma 1 indicates $\sigma = 2n\delta_n + |\mathcal{X}||\mathcal{Y}| \log(n+1)$, and δ shall be specified later. Hence we have that for some

$$k \leq \frac{2\sigma}{\delta} = \frac{2(n\delta_n + |\mathcal{X}||\mathcal{Y}| \log(n+1))}{\delta},$$

there exist $\bar{x}_{t_1}, \bar{y}_{t_1}, \bar{x}_{t_2}, \bar{y}_{t_2}, ..., \bar{x}_{t_k}, \bar{y}_{t_k}$ such that for all $t = 1, 2, ..., n$

$$I(X'_t; Y'_t | X'_{t_1} = \bar{x}_{t_1}, Y'_{t_1} = \bar{y}_{t_1}, ..., X'_{t_k} = \bar{x}_{t_k}, Y'_{t_k} = \bar{y}_{t_k}) \leq \delta. \quad (26)$$

We now define a subgraph of $\mathcal{A}_n$ consisting of all edges (X'^n, Y'^n) that have

$$X'_{t_1} = \bar{x}_{t_1}, Y'_{t_1} = \bar{y}_{t_1}, ..., X'_{t_k} = \bar{x}_{t_k}, Y'_{t_k} = \bar{y}_{t_k}$$

The subgraph denoted as $\bar{\mathcal{A}}_n$ is given by: [5]

$$\bar{\mathcal{A}}_n \triangleq \{(i,j) \in \mathcal{A}_n : \quad (27)$$
$$X'_{t_1}(i) = \bar{x}_{t_1}, Y'_{t_1}(j) = \bar{y}_{t_1}, ..., X'_{t_k}(i) = \bar{x}_{t_k}, Y'_{t_k}(j) = \bar{y}_{t_k}\}.$$

On the same lines as [4, Lemma 3], we have

$$|\bar{\mathcal{A}}_n| \geq \left(\frac{\delta}{|\mathcal{X}||\mathcal{Y}|(2\sigma - \delta)}\right)^k |\mathcal{A}_n|. \quad (28)$$

Define random variables $\bar{X}^n$, $\bar{Y}^n$ on $\mathcal{X}^n$ resp. $\mathcal{Y}^n$ by

$$Pr((\bar{X}^n, \bar{Y}^n) = (x^n(i), y^n(j)) = \frac{1}{|\bar{\mathcal{A}}_n|} \text{if } (i,j) \in \bar{\mathcal{A}}_n. \quad (29)$$

If we denote $\bar{X}^n = (\bar{X}_1, ..., \bar{X}_n)$, $\bar{Y}^n = (\bar{Y}_1, ..., \bar{Y}_n)$, the Fano-distribution of the graph $\bar{\mathcal{A}}_n$ induces a distribution $P_{\bar{X}_t, \bar{Y}_t}$ on the random variables $\bar{X}_t, \bar{Y}_t, t = 1, \ldots, n$. One can show that [4] for all $t = 1, 2, ..., n$

$$\begin{aligned} &P(\bar{X}_t = x, \bar{Y}_t = y) = \\ &\quad P\Big(X'_t = x, \bar{Y}'_t = y | X'_{t_1}(i) = \bar{x}_{t_1}, ..., X'_{t_k}(i) = \bar{x}_{t_k} \\ &\qquad\qquad Y'_{t_1}(j) = \bar{y}_{t_1}, ..., Y'_{t_k}(j) = \bar{y}_{t_k}\Big). \quad (30) \end{aligned}$$

Using (30) in Lemma 2, we get the bound $I(\bar{X}_t; \bar{Y}_t) < \delta$. By applying Pinsker's inequality for I-divergences [12], we conclude that for all $t = 1, 2, ..., n$,

$$\sum_{x,y} |Pr(\bar{X}_t = x, \bar{Y}_t = y) - Pr(\bar{X}_t = x) Pr(\bar{Y}_t = y)| \leq 2\delta^{1/2}. \quad (31)$$

We are now ready to present the final lemma required to complete the proof of the converse.

Lemma 3.

$$R_X \leq \frac{1}{n} \sum_{t=1}^{n} H(\bar{X}_t | \bar{Y}_t) + \delta_{1n}$$
$$R_Y \leq \frac{1}{n} \sum_{t=1}^{n} H(\bar{Y}_t | \bar{X}_t) + \delta_{2n}$$
$$R_X + R_Y \leq \frac{1}{n} \sum_{t=1}^{n} H(\bar{X}_t \bar{Y}_t) + +\delta_{3n}$$

for some $\delta_{1n}, \delta_{2n}, \delta_{3n} \to 0$ and the distributions of the RV's are determined by the Fano-distribution on the codewords $\{(x^n(i), y^n(j)) : (i,j) \in \bar{\mathcal{A}}_n\}$.

Proof: The proof is provided in a more complete version [8]. ■

We can rewrite Lemma 3 using new variables $\bar{X}, \bar{Y}, Q$, where $Q = t \in \{1, 2, ..., n\}$ with probability $\frac{1}{n}$ and $P_{\bar{X}\bar{Y}|Q=t} = P_{\bar{X}_t \bar{Y}_t}$. So we now have (for all sufficiently large n),

$$R_X \leq H(\bar{X} | \bar{Y} Q) + \delta_{1n} \quad (32)$$
$$R_Y \leq H(\bar{Y} | \bar{X} Q) + \delta_{2n} \quad (33)$$
$$R_X + R_Y \leq H(\bar{X}\bar{Y} | Q) + \delta_{3n}, \quad (34)$$

[5]The heirarchy of subgraphs is $G_n \supset \Gamma_n \supset \mathcal{A}_n \supset \bar{\mathcal{A}}_n$

for some $\delta_{1n}, \delta_{2n}, \delta_{3n} \to 0$.

Finally, by using (31), we conclude that for all t, $\bar{X}_t$ and $\bar{Y}_t$ are almost independent for large n. Consequently, using the continuity of mutual information with respect to the joint distribution, Lemma 3 holds with for any joint distribution $P_Q P_{\bar{X}|Q} P_{\bar{Y}|Q}$ such that the marginal on $(\bar{X}, \bar{Y})$ is $P_{\bar{X}\bar{Y}}$. Recall that $P_{\bar{X}\bar{Y}}$ is the dominant joint type that is λ_n-close to $P_{X,Y}$. Using suitable continuity arguments [4], we can now argue that Lemma 3 holds with for any joint distribution $P_Q P_{X|Q} P_{Y|Q}$ such that the marginal on (X, Y) is P_{XY}, completing the proof of the converse.

References

[1] I. Csiszar and J. Korner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. New York: Academic Press,, 1981.

[2] S. S. Pradhan, S. Choi, and K. Ramchandran, "A graph-based framework for transmission of correlated sources over multiple-access channels," *IEEE Trans Inf Theory*, vol. 53, no. 12, pp. 4583–4604, 2007.

[3] R. Venkataramanan and S. S. Pradhan, "A new achievable rate region for the discrete memoryless multiple-access channel with feedback," in *Proc. IEEE Int. Symp. on Inf. Theory*, June 2009.

[4] A. Nazari, S. S. Pradhan, and A. Anastasopoulous, "A new sphere-packing bound for maximal error exponent for multiple-access channels," in *Proc. IEEE Int. Symp. Inf. Theory*, July 2008. Online: http://arxiv.org/abs/0803.3645.

[5] A. Nazari, S. S. Pradhan, and A. Anastasopoulous, "New bounds on the maximal error exponent for multiple-access channels," in *Proc. IEEE Int. Symp. Inf. Theory*, July 2009.

[6] A. Nazari, S. S. Pradhan, and A. Anastasopoulous, "A new upper bound for the average error exponent for discrete memoryless multiple-access channels," in *Online: http://arxiv.org*, Dec 2009.

[7] S. Choi and S. S. Pradhan, "A graph-based framework for transmission of correlated sources over broadcast channels," *IEEE Transactions on Information Theory*, vol. 54, no. 7, pp. 2841–2856, 2008.

[8] A. Nazari, R. Venkataramanan, D. Krithivasan, S. S. Pradhan, and A. Anastasopoulos, "Typicality graphs and their properties," in *Online: http://arxiv.org/*, 2010.

[9] S. Janson, "New versions of Suen's correlation inequality," *Random Structures Algorithms*, vol. 13, pp. 467–483, 1998.

[10] N. Alon and J. Spencer, *The Probabilistic Method*. John Wiley and Sons, Inc., 1992.

[11] R. Ahlswede, "An elementary proof of the strong converse theorem for the multiple-access channel," 1982.

[12] A. A. Fedotov, F. Topsoe, and P. Harremoes, "Refinements of pinsker's inequality," *IEEE Trans. Inf. Theory*, vol. 49, pp. 1491–1498, June 2003.