

Representation of Correlated Sources into Graphs for Transmission over Broadcast Channels

Suhan Choi

Department of Electrical Eng. and Computer Science
University of Michigan,
Ann Arbor, MI 48103, USA
Email: suhanc@eecs.umich.edu

S. Sandeep Pradhan

Department of Electrical Eng. and Computer Science
University of Michigan,
Ann Arbor, MI 48103, USA
Email: pradhanv@eecs.umich.edu

Abstract—In this paper¹ we consider the communication problem that involves transmission of correlated sources over broadcast channels. We consider a graph-based framework for this information transmission problem. The system involves a source coding module and a channel coding module. In the source coding module, the sources are efficiently represented using a nearly semi-regular bipartite graph, and in the channel coding module, the edges of this graph are reliably transmitted over a broadcast channel. We consider nearly semi-regular bipartite graphs as discrete interface between source coding and channel coding in this multiterminal setting. In particular, in this paper, we restrict our attention to the source coding module, building on our earlier work on the channel coding module. We provide an information-theoretic characterization of the rate of growth of the exponent (as a function of the number source samples) of the size of such graphs that can reliably represent a pair of correlated sources.

I. INTRODUCTION

The broadcast channel, where a transmitter wishes to accomplish reliable simultaneous communication with a set of receivers was first introduced by Cover [1] in 1972. Since then, the broadcast channel has been studied extensively, and the capacity region for several special classes of broadcast channels have been found. For the discrete memoryless broadcast channel, Marton [2] established an inner bound to the capacity region which contains all known achievable rate regions. Later, Han and Costa [3] presented a new coding theorem for transmission of correlated sources over broadcast channels. They established a new coding scheme that specifies, instead of achievable rates, a class of source-channel matching conditions between the source and the channel. They also gave an interesting example, which reveals that separate source and channel coding is not optimal for the transmission of correlated sources over broadcast channels.

The essence of Shannon's separation approach in the point-to-point case is to have an efficient architecture for transmission problems by having a discrete interface (a finite set) to represent information sources. The fundamental concept which facilitates this is the notion of typicality, which states that when grouped into large blocks, only those sequences, called typical sequences, coming from a set of size nearly equal to $2^{nH(S)}$ are observed most of the time, and each such sequence

has nearly the same probability, where $H(S)$ denotes Shannon entropy [4].

This notion of typicality can be extended to multiple sources (say S and T) called joint typicality. It says that even though there are roughly $2^{n(H(S)+H(T))}$ individually typical sequence pairs, most of the time only jointly typical sequence pairs with size nearly equal to $2^{nH(S,T)}$ are observed with high probability. Further, using these ideas, it can be shown that for every typical sequence of S (respectively T), there exist roughly $2^{nH(T|S)}$ (respectively $2^{nH(S|T)}$) typical T (respectively S) sequences that are jointly typical, where $H(T|S)$ is the conditional entropy [4] of T given S . This leads us naturally to consider a bipartite undirected graph on the sets of individually typical sequences induced from the property of joint typicality. That is, the vertexes of this graph denote the individually typical sequences, and the jointly typical sequences are connected through an edge. We refer to this graph as the typicality-graph of two correlated sources.

Based on this observation, a graph-based framework for transmission of correlated sources over multiple-access channels was considered in our earlier work [5], [6]. In essence, the transmission system would involve two modules: a source coding module and a channel coding module. In the source coding module, the sources are represented using a bipartite graph, and in the channel coding module, the edges coming from this graph are reliably transmitted over a multiple-access channel.

In this work we consider this approach for the broadcast channels. Our goal is to devise a modular framework based on bipartite graphs for transmission of correlated sources over broadcast channels as shown in Fig. 1. In this transmission system, we have a source coding module and a channel coding module. In the source coding module, the sources are efficiently represented using a bipartite graph, and in the channel coding module, the edges coming from this graph are reliably transmitted over the broadcast channels. The channel coding part was studied in our earlier work in [7], where the rate of growth of the exponent (as a function of the number of channel uses) of the size of the bipartite graphs whose edges can be reliably transmitted over a broadcast channel was characterized in an information-theoretic setting.

In the present work, we provide an information theoretic

¹This work was supported by NSF CAREER Grant CCF-0448115.

characterization of the the rate of growth of the exponent (as a function of the number of source samples) of the size of the bipartite graphs which can reliably represent a pair of correlated sources to be transmitted over a broadcast channel.

In the conventional separate source and channel coding scheme the correlated sources are first represented into one common message W_0 and two private messages W_1 and W_2 in the source coding module, and then (W_0, W_1) and (W_0, W_2) are reliably transmitted to the receiver 1 and the receiver 2, respectively in the channel coding module. Note that here the private messages W_0, W_1 and W_2 are regarded as independent.

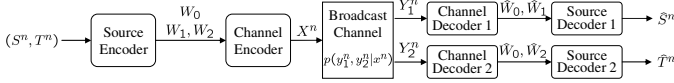


Fig. 1. The broadcast channel communication system

The outline of the remaining part of this paper is as follows. In Section II, we provide definitions of graphs of our interest. Then, we discuss the main result of this paper and its proof in Section III and Section IV, respectively. After that, some interpretations of correlated sources are given in Section V. Finally, Section VI provides concluding remarks.

II. PRELIMINARIES

Before we discuss the main problem, let us first define a bipartite graph and related mathematical terms which will be used in our discussion.

Definition 1: • A *bipartite graph* G is defined as $G = (A_1, A_2, B)$ where A_1 and A_2 are two non-empty sets of vertices, and B is a set of edges where every edge of B joins a vertex in A_1 to a vertex in A_2 . In general, $B \subseteq A_1 \times A_2$.

- If G is a bipartite graph, $V_1(G)$ and $V_2(G)$ are the first (left) and the second (right) vertex sets of G and $E(G)$ is the edge set of G .
- G is said to be a *complete* if $E(G) = V_1(G) \times V_2(G)$.
- The *degree* of a vertex $v \in V_i(G)$ in a graph G , denoted by $\deg_{G,i}(v)$ is the number of edges incident to v for $i = 1, 2$.
- A *subgraph* of a graph G is a graph whose vertex and edge sets are subsets of those of G

Since we consider particular type of bipartite graphs in our discussion, let us define those bipartite graphs.

Definition 2: A bipartite graph G is called *semi-regular* [8] with parameters $(\Delta_1, \Delta_2, \Delta'_1, \Delta'_2)$, denoted by $G(\Delta_1, \Delta_2, \Delta'_1, \Delta'_2)$, if it satisfies:

- $|V_i(G)| = \Delta_i$ for $i=1, 2$,
- $\forall u \in V_1(G), \deg_{G,1}(u) = \Delta'_1$,
- $\forall v \in V_2(G), \deg_{G,2}(v) = \Delta'_2$.

Definition 3: A bipartite graph G is called *nearly semi-regular* with parameters $(\Delta_0, \Delta_1, \Delta_2, \Delta'_1, \Delta'_2, \mu)$, denoted by $G(\Delta_0, \Delta_1, \Delta_2, \Delta'_1, \Delta'_2, \mu)$, if it satisfies:

- G is composed of Δ_0 disjoint subgraphs \tilde{G}_m for $m = 1, 2, \dots, \Delta_0$,

- $\forall m, V_i(\tilde{G}_m) = \{(m-1)\Delta_i+1, (m-1)\Delta_i+2, \dots, m\Delta_i\}$ for $i=1, 2$,
- $V_i(G) = \bigcup_{m=1}^{\Delta_0} V_i(\tilde{G}_m)$ for $i=1, 2$,
- $E(G) = \bigcup_{m=1}^{\Delta_0} E(\tilde{G}_m)$,
- $\forall u \in V_1(G), \Delta'_2\mu^{-1} \leq \deg_{G,1}(u) \leq \Delta'_2\mu$,
- $\forall v \in V_2(G), \Delta'_1\mu^{-1} \leq \deg_{G,2}(v) \leq \Delta'_1\mu$.

Figure 2 illustrates an example of bipartite graph $G(2, 4, 4, 2, 2, 1)$ composed of two disjoint subgraphs.

Definition 4: A triple of messages (random variables) W_0, W_1 and W_2 with message sets $\mathcal{W}_i = \{1, 2, \dots, \Delta_i\}$ for $i = 0, 1, 2$, respectively, is said to be the *message-graph*, denoted by $MG(G)$, associated with a bipartite graph G with parameters $(\Delta_0, \Delta_1, \Delta_2, \Delta'_1, \Delta'_2, \mu)$ if the following condition is satisfied: $\forall w_i \in \mathcal{W}_i$, for $i = 1, 2, 3$, $Pr\{(W_0, W_1, W_2) = (w_0, w_1, w_2)\} = \frac{1}{|E(G)|}$ if $((w_0 - 1)\Delta_1 + w_1, (w_0 - 1)\Delta_2 + w_2) \in E(G)$ and $Pr\{(W_0, W_1, W_2) = (w_0, w_1, w_2)\} = 0$ else. W_0 is referred to as common message and W_1 and W_2 are referred to as private messages. In general, W_0 and (W_1, W_2) are not independent (See Figure 2).

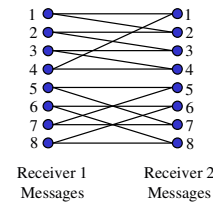


Fig. 2. Message-graph associated with $G(2, 4, 4, 2, 2, 1)$ composed of two disjoint subgraphs \tilde{G}_1 and \tilde{G}_2 where $\mathcal{W}_0 = \{1, 2\}$, $\mathcal{W}_1 = \mathcal{W}_2 = \{1, 2, 3, 4\}$

III. SUMMARY OF RESULTS

In broadcast channels, the source encoder can have access to both sources to be transmitted, but the source decoders can not collaborate with each other. So, this source coding problem can be called distributed reconstruction source coding problem since source decoders reconstruct transmitted sources in a distributed manner.

We consider two correlated sources S and T with a joint probability distribution $p(s, t)$ and the alphabets are given by finite sets \mathcal{S} and \mathcal{T} , respectively. Let $(S_1, T_1), (S_2, T_2), \dots$ be a sequence of jointly distributed random variables i.i.d. $\sim p(s, t)$. For ease of exposition, we assume that the sources do not have a common part [9], [10], [11]. The goal is to represent these sources into message sets which can be associated with nearly semi-regular graphs with parameters $(\Delta_0, \Delta_1, \Delta_2, \Delta'_1, \Delta'_2, \mu)$.

Definition 5: An (n, τ) -transmission system for a bipartite graph $G(\Delta_0, \Delta_1, \Delta_2, \Delta'_1, \Delta'_2, \mu)$ and a pair of correlated sources (S, T) is an ordered tuple (f, g_1, g_2) , consisting of one encoding mapping f and two decoding mappings g_1 and g_2 where

- 1) $f: \mathcal{S}^n \times \mathcal{T}^n \rightarrow E(G)$,
- 2) $g_1: V_1(G) \rightarrow \mathcal{S}^n, g_2: V_2(G) \rightarrow \mathcal{T}^n$,

3) such that a performance measure given by the probability of error satisfies:

$$\tau = \Pr\{(g_1(f(S^n, T^n)), g_2(f(S^n, T^n))) \neq (S^n, T^n)\}. \quad (1)$$

Now we define achievable rates as follows.

Definition 6: A tuple of rates $(R_0, R_1, R_2, R'_1, R'_2)$ is said to be *achievable* for a source coding problem with correlated sources (S, T) (for transmission over broadcast channels), if for any $\epsilon > 0$, and for all sufficiently large n , there exists a bipartite graph G with parameters $(\Delta_0, \Delta_1, \Delta_2, \Delta'_1, \Delta'_2, \mu)$ and an associated (n, τ) -transmission system as defined above satisfying: $\frac{1}{n} \log \Delta_i < R_i + \epsilon$ for $i = 0, 1, 2$, $\frac{1}{n} \log \Delta'_i < R'_i + \epsilon$ for $i = 1, 2$, $\frac{1}{n} \log \mu < \epsilon$ and the corresponding probability of error $\tau < \epsilon$.

The goal is to find the achievable rate region \mathcal{R}_{DR} which is the set of all achievable tuple of rates $(R_0, R_1, R_2, R'_1, R'_2)$. We have found an inner bound to the achievable rate region. It is the main result of this paper, and is given by the following theorem.

Theorem 1: $\mathcal{R}_{DS}^* \subset \mathcal{R}_{DR}$ where

$$\mathcal{R}_{DS}^* = \bigcup_{p(w|s,t)} \{(R_0, R_1, R_2, R'_1, R'_2) : \begin{aligned} R_0 &\geq I(S, T; W), & (2) \\ R_1 &\geq H(S|W), & (3) \\ R_2 &\geq H(T|W), & (4) \\ R'_1 &\geq H(S|T, W), & (5) \\ R'_2 &\geq H(T|S, W), & (6) \\ R_1 + R'_2 &= R'_1 + R_2 \geq H(S, T|W) \}, & (7) \end{aligned}$$

where W is an auxiliary random variable such that $p(w, s, t) = p(w)p(s, t|w)$.

IV. PROOF OF THEOREM 1

In this section, we present the proof of the main result. We use the random binning technique used by Berger [12], and the notion of strongly jointly typical sequences.

Let us consider a fixed joint probability distribution $p(w, s, t)$ on $\mathcal{W} \times \mathcal{S} \times \mathcal{T}$. Also fix $\epsilon > 0$ and real numbers R_0, R_1, R_2, R'_1 , and R'_2 . Without loss of generality, let us assume that $R_i < H(S, T|W)$ for $i = 1, 2$. Let $R'_1 = H(S|T, W)$, and $R'_2 = H(T|S, W)$.

1) Random Sequences and Bin generation: First, draw 2^{nR_0} sequences W_m^n , $m \in \{1, 2, \dots, 2^{nR_0}\}$, of length n , independently from the strongly ϵ -typical set $A_\epsilon^{(n)}(W)$. That is, $P\{W_m^n = w^n\} = \frac{1}{|A_\epsilon^{(n)}(W)|}$ if $w^n \in A_\epsilon^{(n)}(W)$, and $P\{W_m^n = w^n\} = 0$ if $w^n \notin A_\epsilon^{(n)}(W)$. Then, let us define $\alpha = 2^{n(H(S)-R_1+\gamma)}$ and $\beta = 2^{n(H(T)-R_2+\gamma)}$ where γ will be specified shortly. Draw α sequences S^n of length n independently and equally likely with replacement from the strongly ϵ -typical set $A_\epsilon^{(n)}(S)$ with cardinality $2^{n(H(S)-\epsilon_1(\epsilon))} \leq |A_\epsilon^{(n)}(S)| \leq 2^{n(H(S)+\epsilon_1(\epsilon))}$ for large n where $\epsilon_1(\epsilon) \rightarrow 0$ as $\epsilon \rightarrow 0$ [4]. Then, put the all selected α sequences into a bin named B_1 . Repeat the same procedure 2^{nR_1} times independently, resulting in

2^{nR_1} bins denoted by B_i for $i \in \{1, 2, \dots, 2^{nR_1}\}$. Similarly, generate 2^{nR_2} bins denoted by C_j for $j \in \{1, 2, \dots, 2^{nR_2}\}$, each bin contains β sequences T^n , from the strongly ϵ -typical set $A_\epsilon^{(n)}(T)$ with cardinality $2^{n(H(T)-\epsilon_1(\epsilon))} \leq |A_\epsilon^{(n)}(T)| \leq 2^{n(H(T)+\epsilon_1(\epsilon))}$ for large n . Here, we assume $\gamma > \epsilon_1(\epsilon)$.

2) Graph generation: For all $(S^n, T^n) \in A_\epsilon^{(n)}(S, T)$, find an index m , for $m \in \{1, 2, \dots, 2^{nR_0}\}$, such that $(S^n, T^n, W_m^n) \in A_\epsilon^{(n)}(S, T, W)$. If there is no such index m , assign $m = 0$. Also, find a pair of indices (i, j) where i and j are the least indices of bins such that $S^n \in B_i$ and $T^n \in C_j$. If there is no bin containing S^n (or T^n), assign $i = 0$ (or $j = 0$).

Now we can associate a bipartite graph G with the indices m, i , and j where:

- G is composed of 2^{nR_0} disjoint subgraphs \tilde{G}_m for $m \in \{1, 2, \dots, 2^{nR_0}\}$,
- $V_i(G) = \bigcup_{m=1}^{2^{nR_0}} V_i(\tilde{G}_m)$ for $i = 1, 2$,
 $E(G) = \bigcup_{m=1}^{2^{nR_0}} E(\tilde{G}_m)$,
- $\forall m \in \{1, 2, \dots, 2^{nR_0}\}, V_i(\tilde{G}_m) = \{(m-1)2^{nR_i} + 1, (m-1)2^{nR_i} + 2, \dots, m2^{nR_i}\}$ for $i = 1, 2$,
- $\forall m \in \{1, 2, \dots, 2^{nR_0}\}, \forall ((m-1)2^{nR_1} + i, (m-1)2^{nR_2} + j) \in V_1(\tilde{G}_m) \times V_2(\tilde{G}_m), ((m-1)2^{nR_1} + i, (m-1)2^{nR_2} + j) \in E(\tilde{G}_m)$ if $\exists (S^n, T^n) \in (B_i \times C_j) \cap A(S, T|W_m^n)$ where $A(S, T|W_m^n)$ is the set of (S^n, T^n) such that $(S^n, T^n, W_m^n) \in A_\epsilon^{(n)}(S, T, W)$.

For a particular m , a subgraph \tilde{G}_m can be generated as shown in Figure 3.

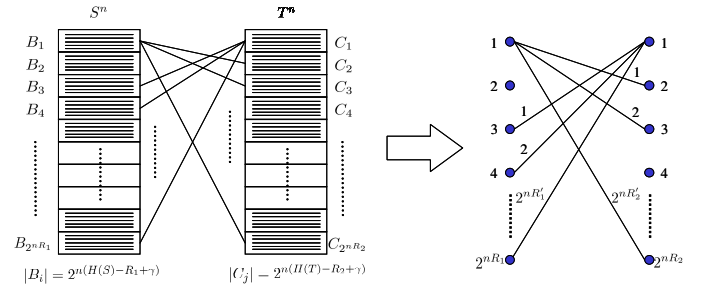


Fig. 3. Generation of a subgraph \tilde{G}_m in a graph G from the bin indices of random sequences S^n and T^n

3) Encoding error events: Before we proceed further, let us make sure that the generated codebooks satisfy certain properties. If the vertices of G do not satisfy certain degree requirements, we may not be able to reliably represent the sources using this graph. So, an encoding error will be declared if either one of the following events occur.

Choose $\gamma > \epsilon_1(\epsilon)$, and ϵ' such that $\epsilon' > 2\gamma + 3\epsilon_1(\epsilon)$. The error events are defined as:

$$\begin{aligned} E_1 &: \exists i \in V_1(G) \text{ such that } \left| \frac{1}{n} \log \deg_{G,1}(i) - R'_2 \right| > \epsilon', \\ E_2 &: \exists j \in V_2(G) \text{ such that } \left| \frac{1}{n} \log \deg_{G,2}(j) - R'_1 \right| > \epsilon'. \end{aligned}$$

We now show that the probability of these events can be made arbitrarily small for sufficiently large n .

Let us define two events $E_{0,1}$ and $E_{0,2}$ as follows:

$$\begin{aligned} E_{0,1} &: \exists i \in V_1(G) \text{ such that } \deg_{G,1}(i) < 2^{n(H(S|T,W)-\epsilon')}, \\ E_{0,2} &: \exists j \in V_2(G) \text{ such that } \deg_{G,2}(j) < 2^{n(H(T|S,W)-\epsilon')}. \end{aligned}$$

Lemma 1: Then for any $\epsilon > 0$, and sufficiently large n :

$$P\{E_{0,i}\} < \frac{\epsilon}{16}, \quad \text{for } i = 1, 2 \quad (8)$$

Proof: The proof of Lemma 1 is omitted.

Let us define two events $E_{0,1}^*$ and $E_{0,2}^*$ as follows.

$$E_{0,1}^* : \exists i \in V_1(G) \text{ such that } \deg_{G,1}(i) > 2^{n(H(S|T,W)+\epsilon')},$$

$$E_{0,2}^* : \exists j \in V_2(G) \text{ such that } \deg_{G,2}(j) > 2^{n(H(T|S,W)+\epsilon')}.$$

Lemma 2: Then for any $\epsilon > 0$, and sufficiently large n :

$$P\{E_{0,i}^*\} < \frac{\epsilon}{16}, \quad \text{for } i = 1, 2 \quad (9)$$

Proof: The proof of Lemma 2 is omitted.

So, by Lemma 1 and Lemma 2, $P(E_1) + P(E_2) < \frac{\epsilon}{4}$. In other words, with high probability we can obtain a bipartite graph G where each vertex in $V_1(G)$ has degree nearly equal to $2^{nH(S|T,W)}$ and each vertex in $V_2(G)$ has degree nearly equal to $2^{nH(T|S,W)}$.

4) *Choosing message-graphs:* If any of E_1 or E_2 occur, then choose any graph with parameters $(2^{nR_0}, 2^{nR_1}, 2^{nR_2}, 2^{nR'_1}, 2^{nR'_2}, 2^{n\epsilon'})$, and assume that the triple of messages W_0, W_1 and W_2 is the message-graph associated with this graph. In this case, no guarantees will be given regarding the probability of error. If none of these events occur, then choose the message-graph $MG(G)$ as the triple of messages.

5) *Encoding:* For a given sources (S^n, T^n) , sender sends (m, i) and (m, j) to receiver 1 and receiver 2, respectively.

6) *Decoding:* Given the received index pair (m, i) , receiver 1 declares $\hat{S}^n = S^n$ if there exists a unique sequence S^n that is jointly typical with the particular sequence W_m^n and belongs to B_i , i.e., $S^n \in B_i \cap A(S|W_m^n)$ where $A(S|W_m^n)$ the set of S^n such that $(S^n, W_m^n) \in A_\epsilon^{(n)}(S, W)$. Otherwise, declare an error. Similarly, given (m, j) , receiver 2 declares $\hat{T}^n = T^n$ if there is a unique sequence T^n such that $T^n \in C_j \cap A(T|W_m^n)$ where $A(T|W_m^n)$ the set of T^n such that $(T^n, W_m^n) \in A_\epsilon^{(n)}(T, W)$. Otherwise, an error is declared.

7) *Probability of error analysis:* Let E denote the event $\{(g_1(m, i), g_2(m, j)) \neq (S^n, T^n)\}$, that the reconstructed sequence pair is not equal to the original source sequence pair. The probability of error $P(E)$ can be given by

$$P(E) = P(E_1 \cup E_2)P(E|E_1 \cup E_2) + P(E \cap E_1^c \cap E_2^c) \quad (10)$$

$$\leq P(E_1 \cup E_2) + P(E \cap E_1^c \cap E_2^c). \quad (11)$$

The second probability in the above equation can be bounded as given in the following lemma.

Lemma 3: For any $\epsilon > 0$, and sufficiently large n ,

$$P(E \cap E_1^c \cap E_2^c) < \frac{3\epsilon}{4} \quad (12)$$

provided

$$R_0 > I(S, T; W) + \epsilon_2(\epsilon), \quad (13)$$

$$R_1 > H(S|W) + \gamma + 3\epsilon_1(\epsilon), \quad (14)$$

$$R_2 > H(T|W) + \gamma + 3\epsilon_1(\epsilon) \quad (15)$$

where $\epsilon_2(\epsilon) \rightarrow 0$ as $\epsilon \rightarrow 0$ and $n \rightarrow \infty$.

Proof: Now let us calculate the probability $P(E \cap E_1^c \cap E_2^c)$. If previous error events E_1 or E_2 do not occur, we define other error events as follows.

$$E_3 : (S^n, T^n) \notin A_\epsilon^{(n)},$$

$$E_4 : \nexists m \in \{1, 2, \dots, 2^{nR_0}\} \text{ such that } (S^n, T^n, W_m^n) \in A_\epsilon^{(n)}(S, T, W), \text{ i.e., } m = 0,$$

$$E_5 : \nexists i \in \{1, 2, \dots, 2^{nR_1}\} \text{ such that } S^n \in B_i, \text{ i.e., } i = 0,$$

$$E_6 : \nexists j \in \{1, 2, \dots, 2^{nR_2}\} \text{ such that } T^n \in C_j, \text{ i.e., } j = 0,$$

$$E_7 : \exists \bar{S}^n \neq S^n \text{ such that } \bar{S}^n \in B_i \cap A(S|W_m^n),$$

$$E_8 : \exists \bar{T}^n \neq T^n \text{ such that } \bar{T}^n \in C_j \cap A(T|W_m^n).$$

Then,

$$P(E \cap E_1^c \cap E_2^c) = P\left(\bigcup_{i=3}^8 E_i \cap E_1^c \cap E_2^c\right) \quad (16)$$

$$\leq \sum_{i=3}^8 P(E_i \cap E_1^c \cap E_2^c) \quad (17)$$

By the property of jointly typical sequences [4], $P(E_3 \cap E_1^c \cap E_2^c) < \frac{\epsilon}{8}$ for sufficiently large n .

$$P(E_4 \cap E_1^c \cap E_2^c) \quad (18)$$

$$= P[\forall m \in \{1, 2, \dots, 2^{nR_0}\}, (S^n, T^n, W_m^n) \notin A_\epsilon^{(n)}(S, T, W)] \quad (19)$$

$$= \left[1 - P\{(S^n, T^n, W_1^n) \in A_\epsilon^{(n)}(S, T, W)\}\right]^{2^{nR_0}} \quad (20)$$

$$\leq \left[1 - 2^{-n(I(S,T;W)+\epsilon_2(\epsilon))}\right]^{2^{nR_0}} \quad (21)$$

$$\leq \exp(-2^{n(R_0 - I(S,T;W) - \epsilon_2)}) \quad (22)$$

$$< \frac{\epsilon}{8} \quad (23)$$

for sufficiently large n , if $R_0 > I(S, T; W) + \epsilon_2(\epsilon)$ where $\epsilon_2(\epsilon) \rightarrow 0$ as $\epsilon \rightarrow 0$ and $n \rightarrow \infty$.

Following [12], for sufficiently large n

$$P(E_5 \cap E_1^c \cap E_2^c) < \frac{\epsilon}{8} \quad (24)$$

$$P(E_6 \cap E_1^c \cap E_2^c) < \frac{\epsilon}{8} \quad (25)$$

if $\gamma > \epsilon_1$.

For sufficiently large n , it can be shown that

$$P(E_7 \cap E_1^c \cap E_2^c) = P\{\exists \bar{S}^n \neq S^n : \bar{S}^n \in B_i \cap A(S|W_m^n)\} \quad (26)$$

$$\stackrel{(a)}{\leq} 2^{n(H(S) - R_1 + \gamma)} 2^{-n(I(S;W) - 3\epsilon_1)} \quad (27)$$

$$= 2^{n(H(S|W) - R_1 + \gamma + 3\epsilon_1)} \quad (28)$$

$$< \frac{\epsilon}{8} \quad (29)$$

provided $R_1 > H(S|W) + \gamma + 3\epsilon_1$, where (a) is from the union bound. Similarly, sufficiently large n , $P(E_8 \cap E_1^c \cap E_2^c) < \frac{\epsilon}{8}$ if $R_2 > H(T|W) + \gamma + 3\epsilon_1$.

Therefore,

$$P(E) \leq P(E_1) + P(E_2) + P(E \cap E_0^c \cap E_1^c \cap E_2^c) < \epsilon. \quad (30)$$

Therefore $P(E) < \epsilon$ for sufficiently large n and under the conditions given by the theorem. In every realization

of random codebooks we have obtained a message-graph $MG(G)$ and an associated graph G with the same parameters, and averaged over this ensemble, we have made sure that the probability of error is within the tolerance level of ϵ . Hence, the proof of the direct coding theorem is completed. ■

V. DIFFERENT MESSAGRAPHS FOR A PAIR OF CORRELATED SOURCES IN THE BROADCAST CHANNEL

We have shown that in the broadcast channel, for sufficiently large block length n , a pair correlated sources (S, T) can be represented using nearly semi-regular graph G with parameters $(2^{nR_0}, 2^{nR_1}, 2^{nR_2}, 2^{nR'_1}, 2^{nR'_2}, 2^{n\epsilon'})$ as shown in Figure 4.

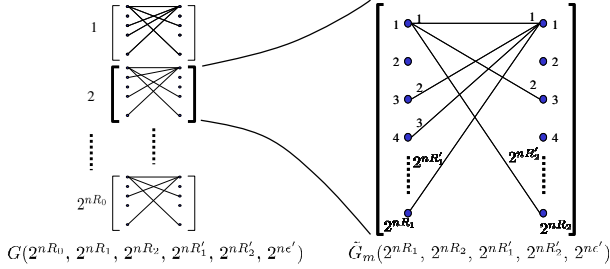


Fig. 4. Graph G with parameters $(2^{nR_0}, 2^{nR_1}, 2^{nR_2}, 2^{nR'_1}, 2^{nR'_2}, 2^{n\epsilon'})$ composed of 2^{nR_0} subgraphs $\tilde{G}_m(2^{nR_1}, 2^{nR_2}, 2^{nR'_1}, 2^{nR'_2}, 2^{n\epsilon'})$

Note that many different message-graphs can represent the same correlated sources without increasing the redundancy. Let us consider a case where the sum rate is minimum, i.e., $R_0 + R_1 + R'_2 = H(S, T)$. In this case, the total number of edges in the graph, $|E(G)| = 2^{n(I(S, T; W) + H(S, T|W))} = 2^{nH(S, T)}$. Since we can vary the mutual information $I(S, T; W)$ such that $0 \leq I(S, T; W) \leq H(S, T)$, we can have many message-graphs without increasing redundancy.

Let us consider some special cases as follows.

- Case A: If $I(S, T; W) = 0$, then $R_0 = 0$, $R_1 = H(S)$, $R_2 = H(T)$, $R'_1 = H(S|T)$, $R'_2 = H(T|S)$. Roughly speaking, this corresponds to the typicality-graph of (S, T) .
- Case B: If $I(S, T; W) = H(S, T)$, then $R_0 = H(S, T)$, $R_1 = 0$, $R_2 = 0$, $R'_1 = 0$, $R'_2 = 0$.
- Case C: If $0 < I(S, T; W) < H(S, T)$, then $R_0 = I(S, T; W)$, $R_1 = H(S|W)$, $R_2 = H(T|W)$, $R'_1 = H(S|T, W)$, $R'_2 = H(T|S, W)$.

In particular, consider the case when $R_0 = C(S; T)$, where $C(S; T)$ denotes the “common information” defined by Wyner [11] as follows. $C(S; T) = \inf I(S, T; W)$ where the infimum is taken over all triples of random variables S, T, W , where auxiliary random variable W take values on a finite set, and such that (1) the marginal distribution for S, T is $p(s, t)$, (2) S and T are conditionally independent given W , i.e., $S \rightarrow W \rightarrow T$ and $p(s, t, w) = p(w)p(s|w)p(t|w)$.

In this case, $R_0 = C(S; T)$, $R_1 = H(S|W)$, $R_2 = H(T|W)$, $R'_1 = H(S|W)$, $R'_2 = H(T|W)$ since S and T are conditionally independent given W . So, each subgraph \tilde{G}_m

become complete. This also means that the private messages for receiver 1 and 2 become independent. Note that a subgraph \tilde{G}_m can not be complete if $R_0 < C(S; T)$ because, by the definition, $C(S; T)$ is the infimum of $I(S, T; W)$ such that $S \rightarrow W \rightarrow T$.

Hence Case A can be thought of as situated at one end of the spectrum, and Case B as situated on the other end of the spectrum. For every value of $I(S, T; W)$, we get an equally efficient representation of the sources into a nearly semi-regular graph. Here, equally efficient representation means that the cardinalities of edge sets of these graphs for the different values of $I(S, T; W)$ are nearly the same.

VI. CONCLUSION

We have considered the task of mapping the outputs of the two correlated sources (for transmission over broadcast channels) into a triple of messages, i.e., a common message and a pair of correlated private messages, which are associated with an undirected nearly semi-regular bipartite graph. We have shown that a pair of correlated sources can be represented into many different triple of messages without increasing the redundancy. This study corresponds to the source coding part of our approach for the transmission of correlated sources over broadcast channels. The goal of this work is to show that a message-graph, which is a triple of messages associated with a nearly semi-regular bipartite graph, can be used as discrete interface in this Shannon-style modular approach to multiterminal communication problems.

REFERENCES

- [1] T. M. Cover, “Broadcast channels,” *IEEE Trans. Inform. Theory*, vol. IT-18, no. 1, pp. 2–14, Jan. 1972.
- [2] K. Marton, “A coding theorem for the discrete memoryless broadcast channel,” *IEEE Trans. Inform. Theory*, vol. IT-25, no. 3, pp. 306–311, May 1979.
- [3] T. S. Han and H. M. Costa, “Broadcast channels with arbitrarily correlated sources,” *IEEE Trans. Inform. Theory*, vol. IT-33, no. 5, pp. 641–650, Sep. 1987.
- [4] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.
- [5] S. S. Pradhan, S. Choi, and K. Ramchandran, “Achievable rates for multiple-access channels with correlated messages,” in *Proc. IEEE Int. Symp. on Inform. Theory (ISIT)*, Jun. 2004.
- [6] —, “A graph-based framework for transmission of correlated sources over multiple access channel,” submitted to *IEEE Trans. Inform. Theory*, Jan. 2006.
- [7] S. Choi and S. S. Pradhan, “Transmission of correlated messages over the broadcast channel,” in *Proc. Conf. on Inform. Sciences and Systems (CISS)*, Mar. 2005.
- [8] J. H. van Lint and R. M. Wilson, *A course in combinatorics*. Cambridge University Press, 2003.
- [9] P. Gács and J. Körner, “Common information is much less than mutual information,” *Problems of Control and Information Theory*, vol. 2, pp. 149–162, 1973.
- [10] H. S. Witsenhausen, “On sequences of pairs of dependent random variables,” *SIAM J. Appl. Math.*, vol. 28, pp. 100–113, Jan. 1975.
- [11] A. D. Wyner, “The common information of two dependent random variables,” *IEEE Trans. Inform. Theory*, vol. IT-21, no. 2, pp. 163–179, Mar. 1975.
- [12] T. Berger, *Multiterminal Source Coding*. In: *The Information Theory Approach to Communications* (ed. G. Longo), CISM Courses and Lecture Notes No. 229. Springer, Wien-New York, 1977.