Transmission of Correlated Messages over the Broadcast Channel

Suhan Choi
Dept. of EECS
Univ. of Michigan, Ann Arbor
1301 Beal Avenue
Ann Arbor, MI, 48109-2122
email: suhanc@eecs.umich.edu

S. Sandeep Pradhan
Dept. of EECS
Univ. of Michigan, Ann Arbor
1301 Beal Avenue
Ann Arbor, MI, 48109-2122
email: pradhanv@eecs.umich.edu

Abstract — **New architectures for transmission of correlated messages over the broadcast channel is presented. We propose to use graphs as digital interface between source coding and channel coding since the correlation structure of sources can be maintained. An achievable rate region for broadcast channels with correlated messages is presented. We consider graphs to represent correlated messages, and to translate the correlation in the given messages into channel encoder. It is shown that such correlated messages can be sent with arbitrarily small error probability over the broadcast channel, given by $p(y_1, y_2|x)$, by using a special channel code which exploits the existing correlation in the messages, if the sizes of messages and the correlation structure of the messages satisfy certain conditions. We prove this by using the random coding argument, random binning, and jointly typical sequence property. If the messages are independent, this rate region is exactly same as Marton region, which is the best known inner bound of the capacity region of the broadcast channel. However, the achievable rate region can be larger than Marton region if the messages are correlated.**

I. Introduction

The broadcast channel was first introduced by Cover [5] in 1972. Since then, many information theorists have studied the broadcast channel, and found the capacity region for several special classes such as degraded broadcast channels (e.g., Bergmans [2] [3], Gallager [10], Ahlswede and Körner [1]), broadcast channel with degraded message sets (Körner and Marton [13]), more capable broadcast channels (El Gamal [8]), deterministic broadcast channels (Marton [14], Pinsker [16]), and broadcast channels with one deterministic component (Marton [15], Gelfand and Pinsker [11]). Recently the capacity region of the Gaussian Multiple Input Multiple Output (MIMO) broadcast channel was found in [22], [4], [20], [21], and [23].

Marton [15] established an inner bound to the capacity region for the discrete memoryless broadcast channel, which contains all the known achievable rate regions. This is a generalization of the results of Cover [6] and Van der Meulen [19]. She proved this bound by using a random coding method which is a combination of the coding techniques of Bergmans [2], Cover [6], and Van der Meulen [19], and the random coding technique used to prove source coding theorems in rate distortion theory. Subsequently El Gamal and Van der Meulen [9] gave a simpler proof of the Marton region involving standard random coding technique, random binning technique, and a jointly typicality lemma.

Later, Han and Costa [12] presented a new coding theorem for the broadcast channel with arbitrarily correlated sources. Their work is very similar to what Cover, El Gamal and Salehi [7] did for the case of multiple access channels with arbitrarily correlated sources. Their results includes Marton's coding theorem as a special case. 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 is closely related to the example given by Cover, El Gamal and Salehi [7]. Both examples reveal that separate source and channel coding is not optimal for the transmission of correlated sources over multiuser channels.

Consider a broadcast channel communication system with one sender and two receivers. The sender wants to simultaneously transmit a pair of messages, one message for each receiver. The messages for each receiver can not be chosen independently, in other words, if a message is selected for receiver 1, the message for receiver 2 can not be arbitrary among the set, and vice versa.

In [17],[18] we recently showed that the achievable rate region of multiple access channels with correlated messages can be increased, if the sizes of the messages and the correlation structure of the messages satisfy certain conditions, by using correlated codewords, which exploits the existing correlation in the given messages.

We apply a similar method to the broadcast channel with correlated messages which can be associated with a "*message-graph*". The main objective of this paper is to establish an achievable rate region of general broadcast channels with correlated messages by adopting the special channel encoding scheme which uses graphs to represent the correlation.

Note that our achievable rate region is exactly same as Marton region if the messages are independent, but, it can be increased if the messages are correlated. As the amount of the correlation in the messages become larger, the achievable rate region also become bigger.

We will consider a graph, referred to as a "*bin-index graph*", which retains the correlation information between the messages, and will also be used to generate a special channel input. Since matching the message-graph and the bin-index graph is a crucial part in our scheme, we briefly recall the concept of graph matching by permutation and relabeling introduced in [17].

The outline of the remaining part is as follows. First, in Section II, we recall the broadcast channel briefly and discuss the previous results connected with our work. Then, we formulate the problem in Section III. Thereafter, in Section IV we give

[1]This work was supported by NSF ITR Grant CCR-0219735.

the main result of this paper, i.e., the achievable rate region for the broadcast channel with correlated messages, given as Theorem 1.

II. Preliminaries

In this section, we provide an overview of the previous study results in the literature on the broadcast channel which are closely related to our work.

II.A Broadcast Channel

A broadcast channel is composed of one sender and many receivers. The objective is to broadcast information from a sender to the many receivers. We consider broadcast channels with only two receivers since multiple receivers cases can be similarly treated. The discrete memoryless broadcast channel with one sender and two receivers, shown in Fig. 1, consists of an input alphabet $\mathcal{X}$ and two output alphabets $\mathcal{Y}_1$ and $\mathcal{Y}_2$ and a probability transition function $p(y_1^n, y_2^n|x^n) = \prod_{i=1}^{n} p(y_{1i}, y_{2i}|x_i)$ when used without feedback.

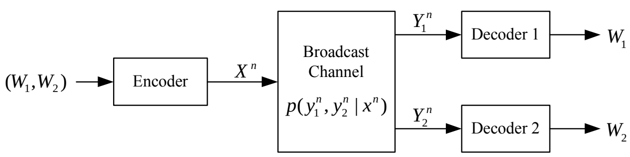

Figure 1: Broadcast channel

Definition 1 *A transmission system with parameters $(n, \Delta_1, \Delta_2, \tau)$ for the given broadcast channel would involve*

- *Two message sets $\mathcal{W}_1 = \{1, 2, \dots, \Delta_1\}$ and $\mathcal{W}_2 = \{1, 2, \dots, \Delta_2\}$*
- *A code $\mathbb{C}$ where $\mathbb{C} \subset \mathcal{X}^n$ and $|\mathbb{C}| = \Delta_1\Delta_2$*
- *An encoder mapping e where $e : \mathcal{W}_1 \times \mathcal{W}_2 \to \mathbb{C}$, i.e., $\forall(i, j) \in \mathcal{W}_1 \times \mathcal{W}_2$, assign $x^n = e(i, j) \in \mathbb{C}$*
- *A set of decoder mappings $\{d^{(1)}, d^{(2)}\}$ where $d^{(1)} : \mathcal{Y}_1^n \to \mathcal{W}_1$, $d^{(2)} : \mathcal{Y}_2^n \to \mathcal{W}_2$*
- *A performance measure given by the average probability of error criterion:*

$$\tau = \sum_{(i,j)\in\mathcal{W}_1\times\mathcal{W}_2} |\mathbb{C}|^{-1} Pr\Big[(d^{(1)}(Y_1^n), d^{(2)}(Y_2^n)) \neq (i, j)|x^n = e(i, j)\Big]$$

Definition 2 *A rate pair (R_1, R_2) is said to be* achievable *for the given broadcast channel if $\forall\epsilon > 0$, and for sufficiently large n, there exists a transmission system as defined above satisfying $\frac{1}{n}\log\Delta_i > R_i - \epsilon$ for $i = 1, 2$ and with the average probability of error $\tau < \epsilon$.*

Definition 3 *The* capacity region *$\mathcal{R}_{BC}$ of the broadcast channel is the closure of the set of all achievable rate pairs (R_1, R_2).*

II.B Marton's Achievable Rate Region for the Broadcast Channel

The capacity region of broadcast channel is still unknown. But Marton [15] found an achievable rate region for the general discrete memoryless broadcast channel, which is the largest known inner bound to the capacity region, given by any rate pairs (R_1, R_2) satisfying

$$R_1 \le I(U; Y_1) \tag{1}$$
$$R_2 \le I(V; Y_2) \tag{2}$$
$$R_1 + R_2 \le I(U; Y_1) + I(V; Y_2) - I(U; V) \tag{3}$$

for some $p(u, v, x)$ on $\mathcal{U} \times \mathcal{V} \times \mathcal{X}$ where U and V are auxiliary random variables with finite alphabets $\mathcal{U}$ and $\mathcal{V}$, respectively such that $(U, V) \to X \to (Y_1, Y_2)$.

III. Problem Formulation

Before we discuss the main problem, let us first define a bipartite graph and related mathematical terms.

Definition 4
- *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 and the second vertex sets of G and $E(G)$ is the edge set of G. Without loss of generality, $V_1 = V_1(G)$ and $V_2 = V_2(G)$ denote the set of vertices on the left and right side of G, respectively.*
- *If $(i, j) \in E(G)$, then i and j are* adjacent, *or* neighboring *vertices of G, and the vertices i and j are* incident *to the edge (i, j).*
- *If each vertex in one set is adjacent to every vertex in the other set, then G is said to be a* complete *bipartite graph. In this case, $E(G) = V_1(G) \times V_2(G)$.*
- *The* degree, *or* valency, *$\deg_G(v)$ of a vertex v in a graph G is the number of edges incident to v.*
- *A* subgraph *of a graph G is a graph whose vertex and edge sets are subsets of those of G. On the contrary, a* supergraph *of a graph G is a graph that contains G as a subgraph.*

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

Definition 5
- *A bipartite graph G is said to have parameters $(\theta_1, \theta_2, \theta_1', \theta_2')$ if it satisfies:*
  - *$|V_i(G)| = \theta_i$ for i=1, 2,*
  - *$\forall u \in V_1(G)$, $\deg_G(u) = \theta_2'$, and $\forall v \in V_2(G)$, $\deg_G(v) = \theta_1'$ such that $\theta_1\theta_2' = \theta_1'\theta_2$.*
- *For two bipartite graphs G_1 and G_2, G_2 is said to "*cover*" G_1 if $E(G_1) \subseteq E(G_2)$.*

Definition 6 *Two message sets $\mathcal{W}_1$ and $\mathcal{W}_2$ can be associated with a bipartite graph G, referred to as a* message-graph, *where $V_1(G) = \mathcal{W}_1$, $V_2(G) = \mathcal{W}_2$, and every edge in $E(G)$ denotes a message pair $(W_1, W_2) \in \mathcal{W}_1 \times \mathcal{W}_2$ which occurs with nonzero equal probability. In general, $E(G) \subseteq \mathcal{W}_1 \times \mathcal{W}_2$.*

Note that $E(G)$ is always a subset of $\mathcal{W}_1 \times \mathcal{W}_2$. If the messages belonging to $\mathcal{W}_1$ and $\mathcal{W}_2$ are "independent", then $\theta_1 = \theta_1'$ and $\theta_2 = \theta_2'$ since all the message pairs can occur jointly. Thus in this case $E(G) = \mathcal{W}_1 \times \mathcal{W}_2$. As the "correlation" between two message sets becomes higher, θ_1' and θ_2' will decrease since there will be less edges in the graph G.

III.A Broadcast Channel with Correlated Messages

Consider a stationary discrete memoryless broadcast channel with one sender and two receivers. The sender wants to deliver a message pair (W_1, W_2) to two receivers which have integer message sets $\mathcal{W}_1 = \{1, 2, \ldots, \Delta_1\}$ and $\mathcal{W}_2 = \{1, 2, \ldots, \Delta_2\}$, respectively, which can be associated with a message-graph $G(\mathcal{W}_1, \mathcal{W}_2, E(G))$. We assume that there is some kind of "*correlation*" between two message sets, i.e., messages for each receiver can not be chosen independently. If the messages of the receivers can be chosen independently, then all possible pairs (W_1, W_2) in the set $\mathcal{W}_1 \times \mathcal{W}_2$ can occur jointly. On the other hand, if they are related, only some pairs (W_1, W_2) in the set $\mathcal{W}_1 \times \mathcal{W}_2$ occur and the other pairs do not. We can also think these messages, without loss of generality, as follows.

- If the messages for receivers are independent, the message pairs (W_1, W_2) are equally likely with probability $\frac{1}{|\mathcal{W}_1 \times \mathcal{W}_2|}$. In this case, each individual message W_1 and W_2 are individually equally likely with probability $\frac{1}{|\mathcal{W}_1|}$ and $\frac{1}{|\mathcal{W}_2|}$, respectively.
- If the messages for receivers are "correlated", the message pairs $(W_1, W_2) \in E(G)$ are equally likely with probability $\frac{1}{|E(G)|}$, and the message pairs $(W_1, W_2) \notin E(G)$ have zero probability where a set $E(G) \subset \mathcal{W}_1 \times \mathcal{W}_2$, and each individual message W_1 and W_2 are individually equally likely with probability $\frac{1}{|\mathcal{W}_1|}$ and $\frac{1}{|\mathcal{W}_2|}$, respectively.

III.B Channel Codes for the Broadcast Channel with Correlated Messages

It was recently shown, in [17] and [18], that the achievable rates for multiple access channel (MAC) with correlated messages can be increased by adopting special channel codes which exploit the existing correlation structure in the messages.

In this work we similarly consider special channel codes and graphs for the broadcast channel with correlated messages. Conventional channel codes in broadcast channels do not consider the existing correlation structure in the given messages for many receivers. In other words, even though messages for receivers are "correlated", they are treated as independent messages. As shown in the example in [12], if we can design special channel codes which translate the existing "correlation" between the messages into the channel input, we might achieve higher transmission rates than is bounded by the conventional codes.

Since there is only one sender in the broadcast channel, we will consider a bipartite graph, which is called a "*bin-index graph*", in order to generate a special channel input which contains the correlation information between the messages. So, the correlation between the messages is translated into a bin-index graph, and then, this bin-index graph is used to generate the special channel code. Here, a bin-index graph should match with the given message-graph to give the best performance, or should at least contain the message-graph as a subgraph for reliable transmission.

III.C Graph Matching by Permutation and Relabeling

We use message-graphs to associate the correlation between the messages, and also consider bin-index graphs to generate special channel codes. Since these two graphs should match for reliable transmission, we discuss the matching of two bipartite graphs which involves permutation and relabeling of vertices of the graphs. Here we will give a simple example and summarize the result. More detail explanation can be found in [17].

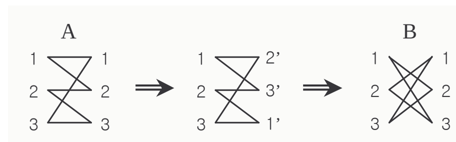

Figure 2: Example of permutation and relabeling of a bipartite graph

Suppose we are given a "correlated" message-graph, characterized by A and a "correlated" bin-index graph characterized by B in Fig. 2. In this case, it is not efficient to use the bin-index graph B to send the message sets associated with A since B does not cover A. However, there exists an interesting property between the two graphs. Note that if we permute right vertices of A, $(1, 2, 3)$ into $(2', 3', 1')$, relabel $(2', 3', 1')$ as $(2, 3, 1)$, and then move right vertices together with their connected edges in natural order $(1, 2, 3)$, then we get graph B. This implies that we can use the bin-index graph B to send the message sets associated with A after permuting and relabeling. This procedure is illustrated in Fig. 2. Clearly, we can also get graph A from graph B similarly.

In general, it may not be easy to construct a specific "correlation" structure in the codebook which can be used to send a given message-graph. But this example motivates us to study the structure of bipartite graphs and its relation to permutation and relabeling.

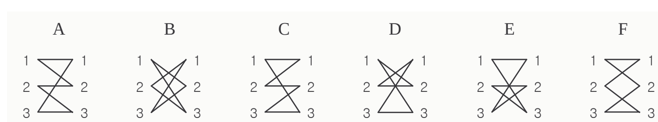

Figure 3: All the possible bipartite graphs in the set $\mathcal{K}_{3,2}$: any graph can be obtained from any other by permutation and relabeling.

Let us consider a set of bipartite graphs $G(n, n, a, a)$, simply denoted by $\mathcal{K}_{n,a}$, $n \in \mathbb{Z}^+$ where $\mathbb{Z}^+$ is the set of positive integers, and $a \in \{1, 2, \ldots, n\}$. For example, Figure 3 illustrates all the elements of $\mathcal{K}_{3,2}$. So there are totally six distinct bipartite graphs in the set $\mathcal{K}_{3,2}$. Note that all the elements in $\mathcal{K}_{3,2}$ can be generated from any one element in the set by permutation and relabeling.

However, in the case of $n = 4$, $a = 2$, we can not get all graphs in $\mathcal{K}_{4,2}$ by just permutation and relabeling of any one graph in the set. The set $\mathcal{K}_{4,2}$ can be partitioned into two *equivalence classes* where the cardinalities of these classes are 72 and 18 respectively, and the *equivalence relation* is characterized by the feasibility of obtaining one element in the class by permutation and relabeling of the vertices of the other. Further, it can be shown that similar partitions exist in $\mathcal{K}_{n,a}$ for general n and a. At this point, we do not have a precise characterization of the number of equivalence classes in $\mathcal{K}_{n,a}$.

To sum up, there is an interesting property in the set of bipartite graphs $\mathcal{K}_{n,a}$ as follows.

- In $\mathcal{K}_{n,a}$, graphs can be partitioned into equivalence classes, where one element in a class can be obtained from the other in the same class by permutation and relabeling. Thus if we have a bin-index graph G_1 that does not cover a given message-graph G_2, but if G_1 covers a graph G_3, where G_2 and G_3 belong to the same equivalence class, then we can use G_1 to transmit the message-graph G_2.
- The elements of $\mathcal{K}_{n,a}$ belonging to different equivalence classes may have different "correlation" structures.

IV. Achievable Rate Region of Broadcast Channels with Correlated Messages

In this section we characterize transmissibility of certain message-graphs over a stationary and memoryless broadcast channel.

IV.A Summary of Results

We consider a stationary discrete memoryless broadcast channel characterized by a conditional probability distribution $p(y_1, y_2|x)$, with input alphabet given by a finite set $\mathcal{X}$, and output alphabets $\mathcal{Y}_1$ and $\mathcal{Y}_2$.

Definition 7 *A transmission system with parameters $(n, \Delta_1, \Delta_2, \Delta_1', \Delta_2', \tau)$ for the given broadcast channel with "*correlated*" messages would involve:*

- $\mathbb{G} = \{G : G$ *has parameters* $(\Delta_1, \Delta_2, \Delta_1', \Delta_2')$ *where* $V_i(G) = \{1, 2, \ldots, \Delta_i\}$ *for* $i = 1, 2\}$,
- *A code* $\mathbb{C}$ *where* $\mathbb{C} \subset \mathcal{X}^n$ *and* $|\mathbb{C}| = \Delta_1\Delta_2'$,
- $\forall G \in \mathbb{G}$, *an encoder mapping* e_G *where* $e_G : E(G) \to \mathbb{C}$, *i.e.,* $\forall (i, j) \in E(G)$, *assign* $x^n = e_G(i, j) \in \mathbb{C}$,
- $\forall G \in \mathbb{G}$, *a set of decoder mappings* $\{d_G^{(1)}, d_G^{(2)}\}$ *such that* $d_G^{(1)} : \mathcal{Y}_1^n \to \mathcal{W}_1$, $d_G^{(2)} : \mathcal{Y}_2^n \to \mathcal{W}_2$ *where* $\mathcal{W}_1 = \{1, 2, \ldots, \Delta_1\}$ *and* $\mathcal{W}_2 = \{1, 2, \ldots, \Delta_2\}$,
- *A performance measure given by the following minimum-average probability of error criterion:*

$$\tau = \min_{G\in\mathbb{G}} \sum_{(i,j)\in E(G)} \frac{1}{|E(G)|} \cdot Pr\Big[(d_G^{(1)}(Y_1^n), d_G^{(2)}(Y_2^n)) \neq (i, j) | x^n = e_G(i, j)\Big]. \quad (4)$$

Definition 8 *A tuple of rates (R_1, R_2, R_1', R_2') is said to be* achievable *for the broadcast channel with correlated messages, if for any $\epsilon > 0$, and for sufficiently large n, there exists a transmission system with parameters $(n, \Delta_1, \Delta_2, \Delta_1', \Delta_2', \tau)$ satisfying: $\frac{1}{n}\log\Delta_1 > R_1 - \epsilon$, $\frac{1}{n}\log\Delta_2 > R_2 - \epsilon$, $\frac{1}{n}\log\Delta_1' > R_1' - \epsilon$, $\frac{1}{n}\log\Delta_2' > R_2' - \epsilon$, and the corresponding minimum-average probability of error $\tau < \epsilon$.*

An achievable rate region of a broadcast channel with correlated messages is given by the following theorem, which is the main result of this paper.

Theorem 1 *For the discrete memoryless broadcast channel $(\mathcal{X}, p(y_1, y_2|x), \mathcal{Y}_1 \times \mathcal{Y}_2)$, an achievable rate region is all tuple of rates (R_1, R_2, R_1', R_2') satisfying*

$$R_1 \le I(U; Y_1), \quad (5)$$
$$R_2 \le I(V; Y_2), \quad (6)$$
$$R_1 + R_2 \le I(U; Y_1) + I(V; Y_2) - I(U; V) + \alpha, \quad (7)$$
$$R_i' \le R_i - \alpha, \quad \textit{for } i = 1, 2 \quad (8)$$

for some distribution $p(u, v, x) = p(u, v)p(x|u, v)$ where α is a nonnegative constant which characterizes the amount of correlation between the messages for two receivers, $0 \le \alpha \le \min\{R_1, R_2\}$ such that $(U, V) \to X \to (Y_1, Y_2)$.

Remark 1 *Note that the achievable rate region given in Theorem 1 is almost the same as Marton's region [15] except for a nonnegative constant α which is determined by the correlation between messages. When the messages are independent, i.e., when all the elements in the set $\mathcal{W}_1 \times \mathcal{W}_2$ can occur, the rate region become exactly same as Marton's since $\alpha = 0$. However, when the messages are correlated, i.e., only some elements in the set $\mathcal{W}_1 \times \mathcal{W}_2$ can occur, the rate region can be larger since $\alpha > 0$ in this case. As the correlation between the messages increases, in other words, as α increases, the achievable rate region also become larger.*

IV.B Proof of Theorem 1

In this section we prove Theorem 1 by using a method similar to that is given in [9] which involves standard random coding, random binning, and the jointly typicality of randomly generated sequences. In addition to the techniques given in [9], we use a concept of a "*super-bin*", which is a group of consecutive bins, to take into account the correlation between the messages.

B.1 Random Sequences and Bin Generation

Let $\epsilon > 0$ and an integer $n \ge 1$ be given. Generate $2^{n(I(U;Y_1)-\epsilon)}$ independent identically distributed (i.i.d.) sequences $u^n \in A_\epsilon^{(n)}(U)$ from the distribution $p(u)$, which is a marginal of the joint distribution $p(u, v)$, each with probability $\frac{1}{|A_\epsilon^{(n)}(U)|}$ where $A_\epsilon^{(n)}(U)$ is the set of ϵ-typical n-sequences. Label these sequences as $u^n(k)$, $k = 1, 2, \ldots, 2^{n(I(U;Y_1)-\epsilon)}$. Similarly, generate $2^{n(I(V;Y_2)-\epsilon)}$ i.i.d. sequences $v^n \in A_\epsilon^{(n)}(V)$ from the distribution $p(v)$, which is also a marginal of $p(u, v)$, each with probability $\frac{1}{|A_\epsilon^{(n)}(V)|}$, and label these as $v^n(l)$, $l = 1, 2, \ldots, 2^{n(I(V;Y_2)-\epsilon)}$. Without loss of generality $2^{n(I(U;Y_1)-\epsilon)}$ and $2^{n(I(V;Y_2)-\epsilon)}$ are assumed to be integers.

Next, for $i \in \{1, 2, \ldots, 2^{nR_1}\}$ and $j \in \{1, 2, \ldots, 2^{nR_2}\}$, define the bins B_i and C_j such that

$$B_i = \{u^n(k) | k \in [(i-1)2^{n(I(U;Y_1)-R_1-\epsilon)} + 1, \ i2^{n(I(U;Y_1)-R_1-\epsilon)}]\} \quad (9)$$

$$C_j = \{v^n(l) | l \in [(j-1)2^{n(I(V;Y_2)-R_2-\epsilon)} + 1, \ j2^{n(I(V;Y_2)-R_2-\epsilon)}]\} \quad (10)$$

where without loss of generality $2^{n(I(U;Y_1)-R_1-\epsilon)}$ and $2^{n(I(V;Y_2)-R_2-\epsilon)}$ are considered to be integers, and $[a,b]$ denotes the set of integers from a to b. Each bin B_i is a set of randomly generated u^n sequences with cardinality $2^{n(I(U;Y_1)-R_1-\epsilon)}$, and similarly each C_j is a set of randomly generated v^n sequences with cardinality $2^{n(I(V;Y_2)-R_2-\epsilon)}$. We also define super-bins $\tilde{B}_p$ and $\tilde{C}_q$ for $p=[1,2^{n(R_1-\alpha)}]$ and $q=[1,2^{n(R_2-\alpha)}]$, each of which is a union of $2^{n\alpha}$ consecutive bins B_i and C_j, respectively, i.e.,

$$\tilde{B}_p = \bigcup_{i=(p-1)2^{n\alpha}+1}^{p\,2^{n\alpha}} B_i, \tag{11}$$

$$\tilde{C}_q = \bigcup_{i=(q-1)2^{n\alpha}+1}^{q2^{n\alpha}} C_i \tag{12}$$

As an example, Figure 4 shows bins and super-bins generated from $p(u,v)$ when a super-bin contains three bins.

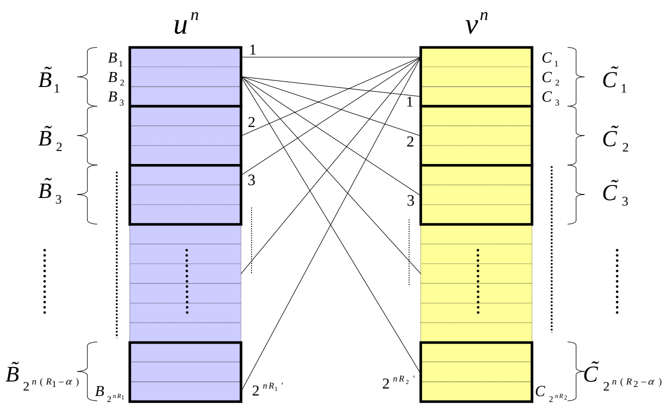

Figure 4: Bin index-graph $G(2^{nR_1},2^{nR_2},2^{nR'_1},2^{nR'_2})$ where $R_i-R'_i=\alpha$, for i=1, 2

For every $(p,j)\in[1,2^{n(R_1-\alpha)}]\times[1,2^{nR_2}]$, define the sets D_{pj},

$$D_{pj}=\{(u^n(k),v^n(l))|u^n(k)\in\tilde{B}_p, v^n(l)\in C_j, (u^n(k),v^n(l))\in A_\epsilon^{(n)}\}, \tag{13}$$

and similarly for every $(i,q)\in[1,2^{nR_1}]\times[1,2^{n(R_2-\alpha)}]$, define the sets E_{iq},

$$E_{iq}=\{(u^n(k),v^n(l))|u^n(k)\in B_i, v^n(l)\in\tilde{C}_q, (u^n(k),v^n(l))\in A_\epsilon^{(n)}\}, \tag{14}$$

where $A_\epsilon^{(n)}=A_\epsilon^{(n)}(U,V)$ is a set of jointly ϵ-typical sequences (u^n,v^n) with respect to the distribution $p(u,v)$.

We now find the condition, given in the following Lemma 1, which guarantees that any set D_{pj} and E_{iq} has at least one jointly ϵ-typical sequence pair (u^n,v^n).

Lemma 1 *For any particular bin B_i and C_j, for any particular super-bin $\tilde{B}_p$ and $\tilde{C}_q$, $\epsilon>0$, and sufficiently large n,*

$$Pr\{|D_{pj}|=0\}\le\frac{\epsilon}{4}, \tag{15}$$

$$Pr\{|E_{iq}|=0\}\le\frac{\epsilon}{4} \tag{16}$$

provided

$$R_1+R_2<I(U;Y_1)+I(V;Y_2)-I(U;V)+\alpha-2\epsilon-\delta(\epsilon) \tag{17}$$

where $\delta(\epsilon)\to 0$ as $\epsilon\to 0$.

B.2 Bin Index-Graph and Codebook Generation

We can generate a bipartite graph from the above randomly generated sequences by taking the indices of the bins as vertices of the bipartite graph and making an edge for an index pair (i,j) when there exists a jointly typical sequence pair (u^n,v^n) in $B_i\times C_j$, i.e., if $\exists(u^n,v^n)\in A_\epsilon^{(n)}\cap(B_i\times C_j)$. The generated bipartite graph is referred to as a "*bin-index graph*", illustrated in the Figure 4, which is a bipartite graph $G(2^{nR_1},2^{nR_2},2^{nR'_1},2^{nR'_2})$ as defined in Definition 5. For i=1 and 2, since 2^{nR_i} bins are grouped into $2^{nR'_i}$ super-bins, each of which is composed of $2^{n\alpha}$ bins, there is a relation between R_i and R'_i, i.e., $R_i-R'_i=\alpha$, for i=1, 2. Note that when $\alpha=0$, i.e., $R_i=R'_i$, a super-bin contains only one bin, so in this case the bin-index graph become a complete bipartite graph where each vertex is connected with all the vertices in the other side. In other words, when $\alpha=0$, the degree of a vertex become maximum. As α increases, the number of edges in the graph decreases and the degree of a vertex also become smaller.

Since our goal is to find minimum probability of error over all message graphs, we assume that the generated bin index-graph and the message-graph belong to the same equivalence class. This means that these two graphs can be matched by using the permutation and relabeling explained in the Section C. For each message pair which is represented as an edge in the message-graph, i.e., $\forall(i,j)\in E(G)$, pick one pair $(u^n(k),v^n(l))\in A_\epsilon^{(n)}\cap(B_i\times C_j)$ and find an $x^n=e_G(i,j)$ which is jointly typical with that pair, i.e., $(u^n(k),v^n(l),x^n)\in A_\epsilon^{(n)}(U,V,X)$.

B.3 Encoding

If a message pair (i,j) is to be transmitted, sender just sends $x^n=e_G(i,j)$ to the channel.

B.4 Decoding

Receiver 1 finds the unique index $\hat{k}$ such that $u^n(\hat{k})$ is jointly ϵ-typical with the received sequence y_1^n, i.e., $(u^n(\hat{k}),y_1^n)\in A_\epsilon^{(n)}(U,Y_1)$. Similarly, Receiver 2 finds the unique index $\hat{l}$ such that $v^n(\hat{l})$ is jointly ϵ-typical with the received sequence y_2^n, i.e., $(v^n(\hat{l}),y_2^n)\in A_\epsilon^{(n)}(V,Y_2)$. Then, each receiver finds decoded messages i and j such that $u^n(\hat{k})\in B_i$ and $v^n(\hat{l})\in C_j$, respectively.

B.5 Probability of Error Analysis

Let us calculate probability of error denoted by $P(E)$. If one or more of the following events occur, it will be declared as an error.

- E_1: The encoding step fails, i.e., for some message pair $(i,j)\in E(G)$, $\nexists(u^n(k),v^n(l))\in(B_i\times C_j)\cap A_\epsilon^{(n)}(U,V)$
- E_2: $(u^n(k),v^n(l),x^n,y_1^n,y_2^n)\notin A_\epsilon^{(n)}(U,V,X,Y_1,Y_2)$
- E_3: Decoding step fails at receiver 1, i.e., $\exists\,\hat{k}\neq k$ such that $(u^n(\hat{k}),y_1^n)\in A_\epsilon^{(n)}(U,Y_1)$
- E_4: Decoding step fails at receiver 2, i.e., $\exists\,\hat{l}\neq l$ such that $(v^n(\hat{l}),y_2^n)\in A_\epsilon^{(n)}(V,Y_2)$

It is easy to see that, for sufficiently large n
$P(E_1)\le\frac{\epsilon}{4}$, if $R_1+R_2<I(U;Y_1)+I(V;Y_2)-I(U;V)+\alpha-2\epsilon-\delta(\epsilon)$ from the result of the Lemma 1;
$P(E_2)\le\frac{\epsilon}{4}$, from the Markov lemma [24];

$P(E_3) \leq \frac{\epsilon}{4}$, if $R_1 < I(U;Y_1) - \epsilon$;
$P(E_4) \leq \frac{\epsilon}{4}$, if $R_2 < I(V;Y_2) - \epsilon$.

Therefore, by applying the union bound

$$P(E) = P(\bigcup_{i=1}^{4} E_i) \leq \sum_{i=1}^{4} P(E_i) \leq \epsilon. \tag{18}$$

Once we have a tuple of achievable rates (R_1, R_2, R_1', R_2') such that $R_i' = R_i - \alpha$ for $i = 1,\ 2$, we can transmit dependent (correlated) message sets with arbitrarily small probability of error. So, a tuple of rates (R_1, R_2, R_1', R_2') such that $R_i' \leq R_i - \alpha$ for $i = 1,\ 2$ is also achievable. Therefore, the Theorem 1 is proved. ■

V. Conclusion

We have discussed the transmission of correlated messages over the broadcast channel. It is shown that if one knows that only certain pairs of messages occurs, the transmission rate can be increased by using the special channel encoding scheme which exploits the known characteristics of the messages. We have assumed that some correlated sources are already mapped into messages sets preserving a predetermined correlation, which also can be associated with a graph. We have established an achievable rate region for the general discrete memoryless broadcast channel, which can be larger than Marton's region when the messages are correlated and satisfy certain conditions.

Acknowledgments

The authors would like to thank Prof. Kannan Ramchandran at University of California, Berkeley for inspiring discussions.

References

[1] R. Ahlswede and J. Körner, "Source coding with side information and a converse for degraded broadcast channels," *IEEE Trans. Inform. Theory*, vol. IT-21, pp. 629–637, Nov. 1975.

[2] P. P. Bergmans, "Random coding theorems for the broadcast channels with degraded components," *IEEE Trans. Inform. Theory*, vol. IT-15, pp. 197–207, Mar. 1973.

[3] ——, "A simple converse for broadcast channels with additive white Gaussian noise," *IEEE Trans. Inform. Theory*, vol. IT-20, pp. 279–280, Mar. 1974.

[4] G. Caire and S. Shamai (Shitz), "On the achievable throughput of a multi-antenna gaussian broadcast channel," *IEEE Trans. Inform. Theory*, vol. 49, no. 7, pp. 1691–1706, July 2003.

[5] T. M. Cover, "Broadcast channels," *IEEE Trans. Inform. Theory*, vol. IT-18, no. 1, pp. 2–14, Jan. 1972.

[6] ——, "An achievable rate region for the broadcast channel," *IEEE Trans. Inform. Theory*, vol. IT-21, no. 4, pp. 399–404, Jul. 1975.

[7] T. M. Cover, A. El Gamal, and M. Salehi, "Multiple-access channel with arbitrarily correlated sources," *IEEE Trans. Inform. Theory*, vol. IT-26, no. 6, pp. 648–657, Nov. 1980.

[8] A. El Gamal, "The capacity of a class of broadcast channels," *IEEE Trans. Inform. Theory*, vol. IT-25, pp. 166–169, Mar. 1979.

[9] A. El Gamal and E. Van der Meulen, "A proof of Marton's coding theorem for the discrete memoryless broadcast channel," *IEEE Trans. Inform. Theory*, vol. IT-27, no. 1, pp. 120–122, Jan. 1981.

[10] R. G. Gallager, "Capacity and coding for degraded broadcast channels," *Probl. Pered. Inform.*, vol. 10, no. 3, pp. 3–14, Jul.-Sep. 1974, translated in *Probl. Inform. Transm.*, pp. 185-193, Jul.-Sep. 1974.

[11] S. I. Gelfand and M. S. Pinsker, "Capacity of a broadcast channel with one deterministic component," *Probl. Pered. Inform.*, vol. 16, no. 1, pp. 24–34, Jan.-Mar. 1980, ; translated in *Probl. Inform. Transm.*, vol. 16, no. 1, pp. 17-25, Jan.-Mar. 1980.

[12] 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.

[13] J. Körner and K. Marton, "General broadcast channels with degraded message sets," *IEEE Trans. Inform. Theory*, vol. IT-23, pp. 60–64, Jan. 1977.

[14] K. Marton, "The capacity region of deterministic broadcast channels," in *Trans. Int. Symp. Inform. Theory*, Paris-Cachan, France, 1977.

[15] ——, "A coding theorem for the discrete memoryless broadcast channel," *IEEE Trans. Inform. Theory*, vol. IT-25, no. 3, pp. 306–311, May 1979.

[16] M. S. Pinsker, "Capacity of noiseless broadcast channels," *Probl. Pered. Inform.*, vol. 14, no. 2, pp. 28–34, Apr.-Jun. 1978, ;translated in *Probl. Inform. Transm.*, pp. 97-102, Apr.-June 1978.

[17] S. S. Pradhan, S. Choi, and K. Ramchandran, "Achievable rates for multiple access channels with correlated messages," Univ. of Michigan, Ann Arbor, Communication and Signal Processing Laboratory (CSPL) Report No. 341, Aug. 2003.

[18] ——, "Achievable rates for multiple-access channels with correlated messages," in *Proc. IEEE Int. Symp. on Inform. Theory (ISIT)*, Jun. 2004.

[19] E. C. Van der Meulen, "Random coding theorems for the general discrete memoryless broadcast channel," *IEEE Trans. Inform. Theory*, vol. IT-21, no. 2, pp. 180–190, Mar. 1975.

[20] P. Viswanath and D. Tse, "Sum capacity of the vector Gaussian channel and uplink-downlink duality," *IEEE Trans. Inform. Theory*, vol. 49, no. 8, pp. 1912–1921, Aug. 2003.

[21] S. Viswanath, N. Jindal, and A. Goldsmith, "Duality, achievable rates and sum-rate capacity of Gaussian MIMO broadcast channels," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2658–2668, Oct. 2003.

[22] H. Weingarten, Y. Steinberg, and S. Sahmai(Shitz), "The capacity region of the Gaussian MIMO broadcast channel," *Proc. CISS 2004*, Mar. 2004.

[23] W. Yu and J. M. Cioffi, "Sum capacity of Gaussian vector broadcast channels," *IEEE Trans. Inform. Theory*, vol. 50, no. 9, pp. 1875–1892, Sep. 2004.

[24] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York:Wiley, 1991.