

Article

On the Achievable Rate Region of the K -Receiver Broadcast Channels via Exhaustive Message Splitting

Rui Tang ^{1,2,3}, Songjie Xie ^{1,2,3} and Youlong Wu ^{1,*} 

¹ School of Information Science and Technology, ShanghaiTech University, Shanghai 201210, China; tangrui@shanghaitech.edu.cn (R.T.); xiesj@shanghaitech.edu.cn (S.X.)

² Shanghai Institute of Microsystem and Information Technology, Chinese Academy of Sciences, Shanghai 200083, China

³ University of Chinese Academy of Sciences, Beijing 100049, China

* Correspondence: wuy11@shanghaitech.edu.cn

Abstract: This paper focuses on K -receiver discrete-time memoryless broadcast channels (DM-BCs) with private messages, where the transmitter wishes to convey K private messages to K receivers. A general inner bound on the capacity region is proposed based on an exhaustive message splitting and a K -level modified Marton's coding. The key idea is to split every message into $\sum_{j=1}^K \binom{K}{j-1}$ submessages each corresponding to a set of users who are assigned to recover them, and then send these submessages via codewords chosen from a K -level structure codebooks. To guarantee the joint typicality among all transmitted codewords, a sufficient condition on the subcodebooks' sizes is derived through a newly establishing hierarchical covering lemma, which extends the 2-level multivariate covering lemma to the K -level case with more intricate dependences. As the number of auxiliary random variables and rate conditions both increase exponentially with K , the standard Fourier–Motzkin elimination procedure becomes infeasible when K is large. To tackle this problem, we obtain a *closed form* of achievable rate region with a special observation of disjoint unions of sets that constitute the power set of $\{1, \dots, K\}$. The proposed achievable rate region allows arbitrary input probability mass functions and improves over previously known achievable (closed form) rate regions for K -receiver ($K \geq 3$) BCs.

Keywords: broadcast channel; capacity region; covering lemma



Citation: Tang, R.; Xie, S.; Wu, Y. On the Achievable Rate Region of the K -Receiver Broadcast Channels via Exhaustive Message Splitting. *Entropy* **2021**, *23*, 1408. <https://doi.org/10.3390/e23111408>

Academic Editor: Song-Nam Hong

Received: 10 September 2021

Accepted: 18 October 2021

Published: 26 October 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The 2-receiver discrete-time memoryless broadcast channel (DM-BCs) was first introduced by Cover [1], who proposed the prestigious superposition coding that outperforms the traditional time-division strategy. Superposition coding is optimal for certain classes of BCs such as degraded DM-BCs, less noisy DM-BCs, and more capable DM-BCs [2,3]. The best known inner bound on the capacity region of DM-BCs is achieved by Marton's coding with message splitting [4,5]. The key idea is to split each source message into common and private parts, where the common part is encoded into a cloud-center codeword, and two private parts are encoded into two separate codewords. To enlarge the achievable rate region, the submitted codewords are jointly typical which is guaranteed by a sufficient condition on the sizes of subcodebooks established by the covering lemma [2].

Recently, Gohari [6] improved the outer bounds for two-receiver broadcast channel by introducing an auxiliary receiver to derive the form of the bounds. He also improved Marton's inner bounds on the cardinalities of the auxiliary variables [7]. For a non-asymptotic view, there are also improvements on the maximum coding rate based on a non-asymptotic converse and achievability bounds with feedback [8] on a two-user broadcast channel and also those bounds on the same coding rate of the common-message K -user DM-BC [9] which extends and strengthens the ones in [10].

There have been efforts on general capacity region without common broadcasting on Gaussian MIMO BCs [11] which coincides with dirty-paper coding (DPC) rate region. The capacity regions of physically degraded Gaussian BCs were further studied for the ergodic or composite situations [12–14]. However, the capacity region with a general message set for multi-antenna Gaussian broadcast channels is not known to full extent [15–18].

For the K -receiver ($K \geq 2$) DM-BCs, most previous works are based either on superposition coding (see in [19–21]) which requires certain constraints on Markov chains for auxiliary random variables (RVs), or on the 2-level Marton's coding where each message is split into one common part and one private part [2,22]. The capacity region of broadcast channels of three receivers was studied with degraded message sets [23,24]. The capacity region of less-noisy broadcast channels with three receivers was determined [25], and superposition coding was justified to be sub-optimal for the more capable 3-receiver broadcast channels [26]. For arbitrary K case, the authors of [21] gave a general inner bound based on superposition coding and rate-splitting using notions from order theory and lattices wherein each receiver decodes its intended message (common or private) along with the partial interference designated to it through rate-splitting. The Marton's coding with 2-level superposition coding structure (consisting of one cloud-center codebook and K satellite codebooks) can be easily constructed and analyzed by the multivariate covering lemma and packing lemma [2]. In [27], an evolved Marton's coding scheme based on up-set [28] rate-splitting, superposition coding, and binning was introduced. However, this work failed to derive a *close-form* achievable rate region due to the prohibitive complexity of applying the Fourier-Motzkin elimination procedure.

In this paper, we consider the general K -receiver ($K \geq 2$) DM-BCs with private messages. For the case $K \geq 3$, in order to inherit the characteristics of superposition coding and combine them with Marton's coding, a general inner bound is proposed based on the exhaustive message splitting and a K -level modified Marton's coding. More specifically, every message is split into $\sum_{j=1}^K \binom{K}{j-1}$ submessages, with each corresponding to a set $\mathcal{S}$ belonging to the power set of $\{1, \dots, K\}$. The submessages respecting to $\mathcal{S}$ are encoded into an exclusive codeword and will be decoded by receiver j if $j \in \mathcal{S}$. We enlarge the subcodebooks sizes and use a K -level Marton's coding to send all codewords that are jointly typical with each other. This allows arbitrary dependence among the input RVs, rather than satisfying certain Markov chains as in [19–22]. Based on a newly established hierarchical covering lemma and properties of rate inequalities obtained in the scheme analysis, we obtain a closed-form achievable rate region. To the best of our knowledge, this is the first result of closed-form achievable rate region for general K -receiver DM-BCs allowing arbitrary input probability mass functions (pmfs). The proposed achievable rate region improves over all previously known closed-form achievable rate regions for K -receiver ($K \geq 3$) BCs.

Note that there are mainly two challenges on establishing the closed-form achievable rate region. The first one is how to derive rate conditions such that all transmitted codewords are jointly typical. The solution is related to the covering lemma. Unfortunately, the known multivariate covering lemma in [2] only deals with sequences which are conditionally independent, making it infeasible in our scheme since our inner bound allows arbitrary input pmfs. Additionally, the *recursive* mutual covering lemma in [27] imposes too many intertwined rate constraints, making it prohibitive to apply the Fourier-Motzkin elimination. To overcome these obstacles, we divide all RVs into K levels and process them hierarchically, leading to a new lemma called hierarchical covering lemma. With this lemma, we can keep joint typicality among all selected sequences, while reducing the number of rate conditions compared to the recursive covering lemma in [27]. The second challenge lies on how to apply Fourier-Motzkin elimination procedure to obtain the final form of the inner bound. Even with the newly hierarchical covering lemma, the number of rate conditions incurred by the packing lemma still increases exponentially with K . To tackle this problem, we aggregate the rate of submessages based on special observations

of disjoint unions of sets that constitute the power set of $\{1, \dots, K\}$, and finally establish the final form of achievable rate region.

Notations

Throughout the paper, we use the notation in [2]. In particular, we let

$$\mathcal{K} \triangleq \{1, \dots, K\}$$

for a positive integer K . We use $\mathbb{P}(\mathcal{K})$ to denote the power set of $\mathcal{K}$, e.g., when $K = 3$, $\mathbb{P}(\mathcal{K}) = \{\{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}$.

For any non-empty set $\mathcal{T} = (k_1, k_2, \dots, k_{|\mathcal{T}|})$, where $k_i \in \mathcal{K}$ for $i \in [1 : |\mathcal{T}|]$, denote $\Pi_{\mathcal{T}}$ as the set including all permutations on $\mathcal{T}$, i.e., $\Pi_{\mathcal{T}} \triangleq \{(\pi(1), \pi(2), \dots, \pi(|\mathcal{T}|)) : \pi(j) \in \mathcal{T}, \pi(j) \neq \pi(i), \forall i, j \in [1 : |\mathcal{T}|], i \neq j\}$. We use $\sqcup$ to denote the disjoint union, e.g., $\mathcal{T}_1 \sqcup \mathcal{T}_2$ means $\mathcal{T}_1 \cup \mathcal{T}_2$ with $\mathcal{T}_1 \cap \mathcal{T}_2 = \emptyset$.

Given a set $\mathbb{S} = \{\mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_{|\mathbb{S}|}\}$ with $\mathcal{S}_j \subseteq \mathcal{K}$, $\forall j \in [1 : |\mathbb{S}|]$, the subtuple of RVs with indices from $\mathbb{S}$ is denoted by $X(\mathbb{S}) \triangleq (X_{\mathcal{S}_1}, \dots, X_{\mathcal{S}_{|\mathbb{S}|}})$, and the corresponding realizations by $x(\mathbb{S}) \triangleq (x_{\mathcal{S}_1}, \dots, x_{\mathcal{S}_{|\mathbb{S}|}})$. Let $X^n(\mathbb{S})$ represent $(X_{\mathcal{S}_j}^n : \mathcal{S}_j \in \mathbb{S})$. For positive integers k and j , we define $x_k^j \triangleq (x_{k,1}, \dots, x_{k,j})$, and $X_k^j \triangleq (X_{k,1}, \dots, X_{k,j})$.

We use $\delta(\epsilon) > 0$ to denote a function of ϵ that tends to zero as $\epsilon \rightarrow 0$. When there are multiple functions $\delta_1(\epsilon), \dots, \delta_k(\epsilon)$, we denote them all by a generic function $\delta(\epsilon)$ with the understanding that $\delta(\epsilon) = \max\{\delta_1(\epsilon), \dots, \delta_k(\epsilon)\}$.

2. Channel Model

Consider a K -receiver DM-BC with only private messages depicted in Figure 1. The setup is characterized by an input alphabet $\mathcal{X}$, K output alphabets $(\mathcal{Y}_k : k \in \mathcal{K})$, and a collection of channel transition pmfs $p(y_1, \dots, y_K | x)$. At time $i \in [1 : n]$, the transmitter sends the channel input $x_i \in \mathcal{X}$, receiver $k \in \mathcal{K}$ observes the output $y_{k,i} \in \mathcal{Y}_k$.

The goal of the communication is that the transmitter conveys private messages M_k to receiver k , for $k \in \mathcal{K}$, respectively. Each M_k is independently and uniformly distributed over the set $\mathcal{M}_k \triangleq [1 : 2^{nR_k}]$, where R_k denotes the communication rate of receiver k .

The encoder maps the messages $(M_1, M_2, \dots, M_K)$ to a sequence $x_i \in \mathcal{X}$:

$$X_i = f^{(n)}(M_1, M_2, \dots, M_K), \quad (1)$$

and receiver $k \in \mathcal{K}$ uses channel outputs y_k^n to estimate $\hat{M}_k$ as a guess of messages M_k :

$$\hat{M}_k = g_k^{(n)}(Y_k^n). \quad (2)$$

A rate tuple $(R_k : k \in \mathcal{K})$ is called achievable if there exists an encoding function $f^{(n)}$ and K decoding functions $g_1^{(n)}, \dots, g_K^{(n)}$ such that the error probability

$$P_e^{(n)} = \Pr\{\hat{M}_k \neq M_k, \exists k \in \mathcal{K}\} \quad (3)$$

tends to zero as the block length n tends to infinity. The closure of the set of achievable rate tuple $(R_k : k \in \mathcal{K})$ is called the *capacity region*.

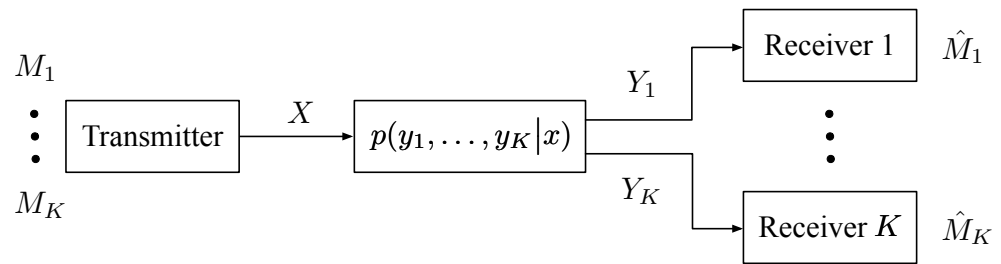


Figure 1. Broadcast channel for K receivers.

3. Preliminary

In this section, we present decomposition of sets which will be used in Sections 4 and 5. Given a set $\mathcal{S}$, an element $k \in \mathcal{S}$ and $l, l' \in \mathcal{K}$ with $l \leq l'$, define

$$\mathbb{A}_{\mathcal{S}}^{l:l'}(\mathcal{T}) \triangleq \{\mathcal{I} \subseteq \mathcal{T} : \mathcal{I} \supseteq \mathcal{S}, |\mathcal{I}| = l, l+1, \dots, l'\}, \quad (4)$$

where $\mathbb{A}_{\mathcal{S}}^{l:l'}(\mathcal{T})$ represents all subsets belonging to $\mathcal{T}$, containing $\mathcal{S}$ and with cardinality between l and l' . For simplicity, we omit the $\mathcal{T}$, $\mathcal{T}$ and l' in $\mathbb{A}_{\mathcal{S}}^{l:l'}(\mathcal{T})$ if $\mathcal{S} = \emptyset$, $\mathcal{T} = \mathcal{K}$ or $l' = l$, i.e.,

$$\mathbb{A}^{l:l'} \triangleq \mathbb{A}_{\emptyset}^{l:l'}(\mathcal{K}), \quad \mathbb{A}_{\mathcal{S}}^{l:l'} \triangleq \mathbb{A}_{\mathcal{S}}^{l:l'}(\mathcal{K}), \quad \mathbb{A}^l \triangleq \mathbb{A}^{l:l}.$$

Given a set $\mathcal{T}$, write its permutation $\pi_{\mathcal{T}} \in \Pi_{\mathcal{T}}$ as $\pi_{\mathcal{T}} = (\pi(1), \pi(2), \dots, \pi(|\mathcal{T}|))$, where $\pi(i)$ denotes the i -th element in $\pi_{\mathcal{T}}$ (i.e., $\pi(i) \in \mathcal{T}$), and $\pi(i) \neq \pi(k)$, if $i \neq k$ for all $i, k \in [|\mathcal{T}|]$. Let $\mathbb{B}_{\pi(i)}$ represent all subsets belonging to $\mathcal{K} \setminus \{\pi(1), \dots, \pi(i-1)\}$ and containing $\pi(i)$ with cardinality between l and l' , i.e.,

$$\mathbb{B}_{\pi(i)}^{l:l'} \triangleq \mathbb{A}_{\{\pi(i)\}}^{l:l'}(\mathcal{K} \setminus \{\pi(1), \dots, \pi(i-1)\}). \quad (5)$$

In particular, when $i = 1$, we have

$$\mathbb{B}_{\pi(1)}^{l:l'} = \mathbb{A}_{\{\pi(1)\}}^{l:l'}. \quad (6)$$

Similarly, we omit the l' if $l' = l$, i.e., $\mathbb{B}_{\pi(i)}^l \triangleq \mathbb{B}_{\pi(i)}^{l:l}$. It is easy to show that $\{\mathbb{B}_{\pi(i)}^{l:l'} : i \in [|\mathcal{T}|]\}$ are disjoint sets for any $l, l' \in [1 : K]$ and any subset $\mathcal{T} \subseteq \mathcal{K}$.

Next, we introduce the following lemma on decomposition among sets.

Lemma 1. *The following decomposition holds:*

$$\bigcup_{k \in \mathcal{T}} \mathbb{A}_{\{k\}}^{l:l'} = \bigsqcup_{i=1}^{|\mathcal{T}|} \mathbb{B}_{\pi(i)}^{l:l'} \quad (7)$$

Proof. Note that

$$\bigcup_{k \in \mathcal{T}} \mathbb{A}_{\{k\}}^{l:l'} = \bigcup_{k \in \mathcal{T}} \{\mathcal{I} \subseteq \mathcal{T} : \mathcal{I} \supseteq \{k\}, |\mathcal{I}| = l, l+1, \dots, l'\} \quad (8)$$

$$\begin{aligned} &= \{\mathcal{I} \subseteq \mathcal{K} : \mathcal{I} \cap \mathcal{T} \neq \emptyset, |\mathcal{I}| = l, \dots, K\} \\ &= \mathbb{A}_{\{\pi(1)\}}^{l:l'} \cup \mathbb{A}_{\{\pi(2)\}}^{l:l'}(\mathcal{K} \setminus \{\pi(1)\}) \cup \dots \cup \mathbb{A}_{\{\pi(|\mathcal{T}|)\}}^{l:l'}(\mathcal{K} \setminus \{\pi(1), \dots, \pi(|\mathcal{T}|-1)\}) \\ &= \bigsqcup_{i=1}^{|\mathcal{T}|} \mathbb{B}_{\pi(i)}^{l:l'}, \end{aligned} \quad (9)$$

where the last equality holds by (5) and as $\{\mathbb{B}_{\pi(i)}^{l:l'} : i \in [|\mathcal{T}|]\}$ are disjoint sets. $\square$

For example, consider the case $K = 3$, $\mathcal{T} = (\pi(1), \pi(2), \pi(3)) = \{1, 2, 3\}$ and $\pi_{\mathcal{T}} = (1, 2, 3)$, we have

$$\begin{aligned}\mathbb{A}_{\{1\}}^{1:2}(\{1, 2, 3\}) &= \{\{1\}, \{1, 2\}, \{1, 3\}\}, \\ \bigcup_{k \in \mathcal{T}} \mathbb{A}_{\{k\}}^{1:K} &= \{\{1\}, \{2, 3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}, \\ \mathbb{B}_{\pi(1)}^{1:3} &= \{\{1\}, \{1, 2\}, \{1, 3\}, \{1, 2, 3\}\}, \mathbb{B}_{\pi(2)}^{1:3} = \{\{2\}, \{2, 3\}\}, \mathbb{B}_{\pi(3)}^{1:3} = \{\{3\}\}.\end{aligned}$$

Figure 2 is given for illustration. To simplify the notation, an arbitrary collection of sets, e.g., $\{1, 12, 123\}$, is recognized as $\{\{1\}, \{1, 2\}, \{1, 2, 3\}\}$

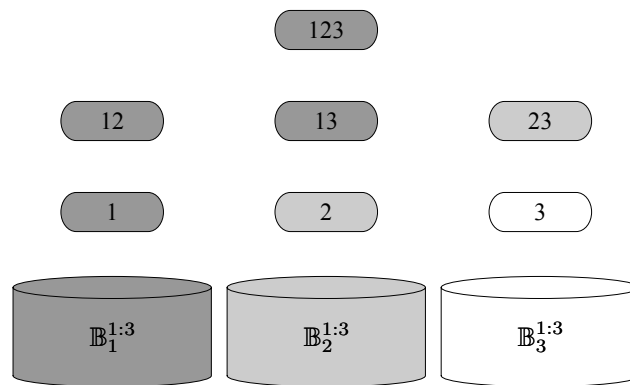


Figure 2. The illustration of the disjoint decomposition of $\bigcup_{k \in \mathcal{T}} \mathbb{A}_{\{k\}}^{1:K}$ in the case for $K = 3$, $\mathcal{T} = \{1, 2, 3\}$ and $\pi = (\pi(1), \pi(2), \pi(3)) = (1, 2, 3)$. The idea is that every time we pick a collection of sets descended from a set in the layer l , the number of unpicked sets in layer $l - 1$ is one, where layer l contains all subsets of $\mathcal{K}$ with cardinality l . Therefore, $\bigcup_{k \in [1:3]} \mathbb{A}_{\{k\}}^{1:3} = \mathbb{B}_1^{1:3} \sqcup \mathbb{B}_2^{1:3} \sqcup \mathbb{B}_3^{1:3}$.

4. Main Results

Theorem 1. Using the notations presented in Section 3, a rate tuple $(R_k : k \in \mathcal{K})$ is achievable for the DM-BC $p(y_1, \dots, y_K | x)$ if for all $\mathcal{T} \subseteq \mathcal{K}$ and permutations $\pi_{\mathcal{T}} = \{\pi(1), \dots, \pi(|\mathcal{T}|)\} \in \Pi_{\mathcal{T}}$,

$$\begin{aligned}\sum_{k \in \mathcal{T}} R_k &\leq \sum_{i=1}^{|\mathcal{T}|} \left(\sum_{\mathcal{S} \in \mathbb{B}_{\pi(i)}^{1:K}} H(U_{\mathcal{S}} | U(\mathbb{A}_{\mathcal{S}}^{1:K} \setminus \mathcal{S})) - H(U(\mathbb{B}_{\pi(i)}^{1:K}) | Y_{\pi(i)} U(\mathbb{A}_{\pi(i)}^{1:K} \setminus \mathbb{B}_{\pi(i)}^{1:K})) \right) \\ &\quad + \sum_{l=1}^{K-1} \left(H\left(U\left(\bigcup_{i=1}^{|\mathcal{T}|} \mathbb{B}_{\pi(i)}^l\right) | U(\mathbb{A}^{(l+1):K})\right) - \sum_{\mathcal{S} \in \bigcup_{i=1}^{|\mathcal{T}|} \mathbb{B}_{\pi(i)}^l} H(U_{\mathcal{S}} | U(\mathbb{A}_{\mathcal{S}}^{1:K} \setminus \mathcal{S})) \right), \quad (10)\end{aligned}$$

for some pmf $p(U(\mathbb{A}^{1:K}))$ and a function $X = f(U(\mathbb{A}^{1:K}))$.

Proof. The achievable scheme is based on an exhaustive message splitting and a K-level Marton's coding. More specifically, each message M_k , for $k \in \mathcal{K}$, is split into $\sum_{j=1}^K \binom{K}{j-1}$ submessages $M_k \triangleq (M_{k,\mathcal{S}} : \mathcal{S} \subseteq \mathcal{K}, |\mathcal{S}| = 1, 2, \dots, K)$ with $M_{k,\mathcal{S}} \in \emptyset$ if $k \notin \mathcal{S}$. The submessages $\mathbf{M}_{\mathcal{S}} = (M_{k,\mathcal{S}} : k \in \mathcal{K})$ are encoded into a codeword $u_{\mathcal{S}}^n \sim p(u_{\mathcal{S}} | u(\mathbb{A}_{\mathcal{S}}^{1:K} \setminus \mathcal{S}))$ and will be decoded by receiver j if $j \in \mathcal{S}$. Then, a K-level Marton's coding is applied to send $(2^K - 1)$ codewords $(u_{\mathcal{S}}^n : \mathcal{S} \subseteq \mathcal{K}, |\mathcal{S}| = 1, \dots, K)$ that are jointly typical with each other. See detailed scheme in Section 5. $\square$

To ensure arbitrary input pmfs $p(U(\mathbb{A}^{1:K}))$ in the inner bound above, we need to establish a sufficient condition on the sizes of subcodebooks, which is related to the covering lemma. The multivariate covering lemma in [2] only consider sequences generated under simple dependence, e.g., $(u_1^n, \dots, u_K^n) \sim p(u_1 | u_0) p(u_1 | u_0) \dots p(u_K | u_0)$ given a sequence

$u_0^n \sim P_{U_0}$. In our scheme there are $(2^K - 1)$ RVs, represented by $(U_S : S \in U(\mathbb{A}^{1:K}))$, and each symbol of codeword U_S^n is generated conditionally independent according to $p(u_S | u(\mathbb{A}_S^{1:K} \setminus S))$. In [27], a recursive mutual covering Lemma was introduced, which allows all transmitted codewords to be jointly typical. However, the recursive mutual covering lemma creates around 2^K intertwined rate equalities, turning the Fourier–Motzkin elimination procedure to a disastrous problem. To obtain the closed-form achievable rate region, we divide subcodebooks into K levels and process them recursively in a hierarchical manner, which greatly reduces the number of rate inequalities and enables to obtain the closed-form rate region.

Lemma 2 (Hierarchical Covering Lemma). *Let $U(\mathbb{A}^{1:K}) \sim p(u(\mathbb{A}^{1:K}))$ and $\epsilon_K < \epsilon_{K-1} \cdots < \epsilon_1$. Let $U_K^n \sim p(u_K^n)$ be a random sequence with $\lim_{n \rightarrow \infty} P\{U_K^n \in \mathcal{T}_{\epsilon_K}^{(n)}\} = 1$. For each $l \in [1 : K - 1]$ and $S \in \mathbb{A}^l$, let $U_S^n(m_S)$, $m_S \in [1 : 2^{nr_S}]$, be pairwise conditionally independent sequences, each distributed according to $\prod_{i=1}^n p_{U_S | U(\mathbb{A}_S^{1:K} \setminus S)}(u_{S,i} | u_{S',i} : S' \in (\mathbb{A}_S^{1:K} \setminus S))$. Assume that $(U_S^n(m_S) : m_S \in [1 : 2^{nr_S}], S \in \mathbb{A}^l)$ are mutually conditionally independent given $(U_S^n(m_S) : S \in (\mathbb{A}_S^{1:K} \setminus S))$. Then, there exists $\delta(\epsilon_l)$ that tends to zero as $\epsilon_l \rightarrow 0$ such that*

$$\lim_{n \rightarrow \infty} P\left\{(U_S^n(m_S) : S \in \mathbb{A}^{1:(K-1)}, U_K^n) \notin \mathcal{T}_{\epsilon_1}^{(n)}\right. \\ \left.\text{for all } (m_S : S \in \mathbb{A}^{1:(K-1)}) \in \prod_{S \in \mathbb{A}^{1:(K-1)}} [1 : 2^{nr_S}]\right\} = 0,$$

if

$$\sum_{S \in \mathcal{J}^l} r_S > \sum_{S \in \mathcal{J}^l} H(U_S | U(\mathbb{A}_S^{1:K} \setminus S)) - H(U(\mathcal{J}^l) | U(\mathbb{A}^{(l+1):K})) + \delta(\epsilon_l), \quad (11)$$

for all $\mathcal{J}^l \subseteq \mathbb{A}^l$ and $l \in [1 : (K - 1)]$.

Proof. See the proof in Appendix A. $\square$

Remark 1. The main difference between recursive mutual covering lemma [27] is that our hierarchical covering lemma reduces the number of variables (this is achieved by assuring the typicality of the codewords top-down from level K to level 1 in the proof of hierarchical covering lemma) in almost each inequalities compared with those in [27], which makes it easier for Fourier–Motzkin elimination to derive a closed-form achievable rate region. However, The accurate performance needs to be justified on specific channels, which also requires an optimal choice of the distribution of the auxiliary variables. Thus, it is hard to evaluate the overall performance, and we leave it in our future work.

To briefly investigate the performance of the two covering lemmas, we give an example of the bound on $r_1 + r_{12} + r_{13} + r_{123}$ of each.

Utilizing our lemma, we get three inequalities to combine it. Namely,

$$\begin{aligned} r_1 &> H(U_1 | U_{12} U_{13} U_{123}) - H(U_1 | U_{12} U_{13} U_{23} U_{123}) \\ r_{12} + r_{13} &> H(U_{12} | U_{123}) + H(U_{13} | U_{123}) - H(U_{12} U_{13} | U_{123}) \\ r_{123} &> 0. \end{aligned}$$

By adding them all together, we get $r_1 + r_{12} + r_{13} + r_{123} > H(U_1 | U_{12} U_{13} U_{123}) + H(U_{12} | U_{123}) + H(U_{13} | U_{123}) - (H(U_1 | U_{12} U_{13} U_{23} U_{123}) + H(U_{12} U_{13} | U_{123}))$.

On the other hand, the bound on it by recursive mutual covering lemma is $H(U_1 | U_{12} U_{13} U_{123}) + H(U_{12} | U_{123}) + H(U_{13} | U_{123}) - H(U_1 U_{12} U_{13} | U_{123})$, namely $H(U_1 | U_{12} U_{13} U_{123}) + H(U_{12} | U_{123}) + H(U_{13} | U_{123}) - (H(U_1 | U_{12} U_{13} U_{123}) + H(U_{12} U_{13} | U_{123}))$, which leads to a lower upper bound (by conditioning reduces entropy) compared with ours, and our scheme may sacrifice some performance to derive a closed-form achievable rate tuple constraints.

Lemma 3. Consider the inequality condition in the hierarchical covering lemma:

$$\sum_{\mathcal{S} \in \mathcal{J}^l} r_{\mathcal{S}} > \sum_{\mathcal{S} \in \mathcal{J}^l} H(U_{\mathcal{S}} | U(\mathbb{A}_{\mathcal{S}}^{1:K} \setminus \mathcal{S})) - H(U(\mathcal{J}^l) | U(\mathbb{A}^{(l+1):K})) + \delta(\epsilon).$$

If the set $\mathcal{J}^l$ is split into $N \in [1 : |\mathcal{J}^l|]$ disjoint pieces $\{\mathcal{J}_1^l, \mathcal{J}_2^l, \dots, \mathcal{J}_N^l\}$ satisfying

$$\mathcal{J}^l = \bigcup_{i=1}^N \mathcal{J}_i^l, \text{ with } \mathcal{J}_i^l \cap \mathcal{J}_j^l = \emptyset, \text{ if } i \neq j, \quad (12)$$

then for all $i \in [N]$, we have

$$\sum_{\mathcal{S} \in \mathcal{J}_i^l} r_{\mathcal{S}} > \sum_{\mathcal{S} \in \mathcal{J}_i^l} H(U_{\mathcal{S}} | U(\mathbb{A}_{\mathcal{S}}^{1:K} \setminus \mathcal{S})) - H(U(\mathcal{J}_i^l) | U(\mathbb{A}^{(l+1):K})) + \delta_{\mathcal{J}_i^l}(\epsilon).$$

With simple justification, the sum of lower bounds for $\sum_{\mathcal{S} \in \mathcal{J}_i^l} r_{\mathcal{S}}, \forall i \in [N]$ is not greater than the lower bound for $\sum_{\mathcal{S} \in \mathcal{J}^l} r_{\mathcal{S}}$, which means combinations of split inequalities will not induce a smaller region bounded by the overall inequalities that simultaneously restrain all variables,

Proof. It is equivalent to prove that

$$\begin{aligned} & \sum_{\mathcal{S} \in \mathcal{J}^l} H(U_{\mathcal{S}} | U(\mathbb{A}_{\mathcal{S}}^{1:K} \setminus \mathcal{S})) - H(U(\mathcal{J}^l) | U(\mathbb{A}^{(l+1):K})) \\ & \stackrel{(a)}{=} \sum_{i=1}^N \sum_{\mathcal{S} \in \mathcal{J}_i^l} \left(H(U_{\mathcal{S}} | U(\mathbb{A}_{\mathcal{S}}^{1:K} \setminus \mathcal{S})) - H(U(\mathcal{J}_i^l) | U(\mathbb{A}^{(l+1):K}), U(\mathcal{J}_1^l), \dots, U(\mathcal{J}_{i-1}^l)) \right) \\ & \geq \sum_{i=1}^N \left(\sum_{\mathcal{S} \in \mathcal{J}_i^l} H(U_{\mathcal{S}} | U(\mathbb{A}_{\mathcal{S}}^{1:K} \setminus \mathcal{S})) - H(U(\mathcal{J}_i^l) | U(\mathbb{A}^{(l+1):K})) \right), \end{aligned}$$

where (a) follows by the chain rule of entropy and (12); (b) holds because conditioning reduces entropy. $\square$

Corollary 1. Define

$$I_{12} := \min_{i \in \{1,2\}} \{I(U_{12}U_{123}; Y_i)\}, I_{13} := \min_{i \in \{1,3\}} \{I(U_{13}U_{123}; Y_i)\}, I_{23} := \min_{i \in \{2,3\}} \{I(U_{23}U_{123}; Y_i)\},$$

and

$$\begin{aligned} \Delta := & I(U_2; U_{13} | U_{12}U_{23}U_{123}) + I(U_3; U_{12} | U_{13}U_{23}U_{123}) + I(U_1; U_3 | U_2U_{12}U_{13}U_{23}U_{123}) \\ & + I(U_1; U_{23} | U_{12}U_{13}U_{123}) + I(U_1; U_2 | U_{12}U_{13}U_{23}U_{123}) + I(U_2; U_3 | U_{12}U_{13}U_{23}U_{123}). \end{aligned}$$

A rate region (R_1, R_2, R_3) is achievable for 3-receiver DM-BC $p(y_1, y_2, y_3|x)$ if

$$R_1 < I(U_1 U_{12} U_{13} U_{123}; Y_1) - I(U_1; U_{23} | U_{12} U_{13} U_{123}), \quad (13a)$$

$$R_2 < I(U_2 U_{12} U_{23} U_{123}; Y_2) - I(U_2; U_{13} | U_{12} U_{23} U_{123}), \quad (13b)$$

$$R_3 < I(U_3 U_{13} U_{23} U_{123}; Y_3) - I(U_3; U_{12} | U_{13} U_{23} U_{123}), \quad (13c)$$

$$\begin{aligned} R_1 + R_2 < & I_{12} + I(U_1 U_{13}; Y_1 | U_{12} U_{123}) + I(U_2 U_{23}; Y_2 | U_{12} U_{123}) \\ & - I(U_2; U_{13} | U_{13} U_{23} U_{123}) - I(U_{13}; U_{23} | U_{12} U_{123}) \\ & - I(U_1; U_{23} | U_{12} U_{13} U_{123}) - I(U_1; U_2 | U_{12} U_{13} U_{23} U_{123}), \end{aligned} \quad (13d)$$

$$\begin{aligned} R_1 + R_3 < & I_{13} + I(U_1 U_{12}; Y_1 | U_{13} U_{123}) + I(U_3 U_{23}; Y_3 | U_{13} U_{123}) \\ & - I(U_{12}; U_{23} | U_{13} U_{123}) - I(U_3; U_{12} | U_{13} U_{23} U_{123}) \\ & - I(U_1; U_{23} | U_{12} U_{13} U_{123}) - I(U_1; U_3 | U_{12} U_{13} U_{23} U_{123}), \end{aligned} \quad (13e)$$

$$\begin{aligned} R_2 + R_3 < & I_{23} + I(U_2 U_{12}; Y_2 | U_{23} U_{123}) + I(U_3 U_{13}; Y_3 | U_{23} U_{123}) \\ & - I(U_{12}; U_{13} | U_{23} U_{123}) - I(U_3; U_{12} | U_{13} U_{23} U_{123}) \\ & - I(U_2; U_{13} | U_{12} U_{13} U_{123}) - I(U_2; U_3 | U_{12} U_{13} U_{23} U_{123}), \end{aligned} \quad (13f)$$

$$\begin{aligned} R_1 + R_2 + R_3 < & I_{12} + I(U_1 U_{13}; Y_1 | U_{12} U_{123}) + I(U_2 U_{23}; Y_2 | U_{12} U_{123}) \\ & + I(U_3; Y_3 | U_{13} U_{23} U_{123}) - I(U_{13} U_{23} | U_{12} U_{123}) - \Delta, \end{aligned} \quad (13g)$$

$$\begin{aligned} R_1 + R_2 + R_3 < & I_{13} + I(U_1 U_{12}; Y_1 | U_{13} U_{123}) + I(U_3 U_{23}; Y_3 | U_{13} U_{123}) \\ & + I(U_2; Y_2 | U_{12} U_{23} U_{123}) - I(U_{12} U_{23} | U_{13} U_{123}) - \Delta, \end{aligned} \quad (13h)$$

$$\begin{aligned} R_1 + R_2 + R_3 < & I_{23} + I(U_2 U_{12}; Y_2 | U_{23} U_{123}) + I(U_3 U_{13}; Y_3 | U_{23} U_{123}) \\ & + I(U_1; Y_1 | U_{12} U_{13} U_{123}) - I(U_{12} U_{13} | U_{23} U_{123}) - \Delta, \end{aligned} \quad (13i)$$

for some pmf $p(u_1 u_2 u_3 u_{12} u_{13} u_{23} u_{123})$ and a function $x = f(u_1 u_2 u_3 u_{12} u_{13} u_{23} u_{123})$.

Proof. The result directly comes from Theorem 1 with $K = 3$. For example, for $\mathcal{T} = \{1, 2\}$ and $\pi_{\mathcal{T}} = (1, 2)$, in (10), then we have $\mathbb{A}^3 = \{123\}$, $\mathbb{A}^{2:3} = \{12, 13, 23, 123\}$, and ,

$$\mathbb{B}_{\pi(1)}^{1:3} = \{123, 12, 13, 1\}, \mathbb{B}_{\pi(2)}^{1:3} = \{23, 2\},$$

which satisfy the decomposition $\cup_{i \in \{1,2\}} \mathbb{A}_{\{i\}}^{1:3} = \{1, 12, 13, 23, 123\} = \mathbb{B}_{\pi(1)}^{1:3} \sqcup \mathbb{B}_{\pi(2)}^{1:3}$. With $\pi_{\mathcal{T}} = (1, 2)$ and $\pi_{\mathcal{T}} = (2, 1)$, we obtain upper bound for $R_1 + R_2$ in (13d). The remaining rate constraints are acquired similarly. $\square$

Remark 2 (Comparison with superposition coding and Marton's coding). Our achievable region generalizes that introduced by standard Marton's coding and superposition coding, i.e., both of them are special cases of our general-form intended coding. Furthermore, the rate region in Theorem 1 contains the regions resulted from the two aforementioned coding schemes.

- Our coding degenerates into Marton's coding by setting $U_S = \text{const}$, $\forall S \in \mathbb{A}^{2:(K-1)}$, which indicates our rate region contains that derived by Marton's coding;
- As superposition coding is optimal for degraded/less noisy/more capable DM-BC, which is barely guaranteed in many cases, i.e., without knowing the concrete relation according to the Markov chain, our rate region contains that derived from superposition coding.

In the future work, we will evaluate our rate region for some specific DM-BCs to show that our coding scheme strictly improves previously known inner bounds.

5. Achievable Coding Scheme for Theorem 1

We first present our scheme for 3-receiver DM-BC as an illustration, and then extend it to general K -receiver DM-BC model for $K \geq 2$.

5.1. Coding Scheme for 3-Receiver DM-BCs

Let an all-one column vector $(1, \dots, 1)$ with a specified dimension be denoted by **1**. Albeit with an abuse of notation, we use a string of elements to imply the singleton notations in Section 3, e.g., $\mathbf{R}_{123} = \mathbf{R}_{\{123\}}$, $\mathbf{M}_{123} = \mathbf{M}_{\{123\}}$, $\mathbb{A}_{123}^{l:l'} = \mathbb{A}_{\{123\}}^{l:l'}$, $l_{123} = l_{\{123\}}$ etc. To simplify notations, we denote $U_S^n(\mathbf{m}_S, l_S | \mathbf{m}_{\mathbb{A}_S^{1:K} \setminus S}, l_{\mathbb{A}_S^{1:K} \setminus S})$ by $U_S^n(\mathbf{m}_S, l_S)$ and its realization $u_S^n(\mathbf{m}_S, l_S | \mathbf{m}_{\mathbb{A}_S^{1:K} \setminus S}, l_{\mathbb{A}_S^{1:K} \setminus S})$ by $u_S^n(\mathbf{m}_S, l_S)$.

5.1.1. Rate Splitting

Divide $M_j \in [1 : 2^{R_j}]$, $j \in \{1, 2, 3\}$, into four independent messages ($M_{j,S} \in [1 : 2^{R_{j,S}}]$: $S \in \mathbb{A}_{\{j\}}^{1:3}$). Therefore, $R_j = \sum_{S \in \mathbb{A}_{\{j\}}^{1:3}} R_{j,S}$. More precisely,

$$\begin{aligned} M_1 &= (M_{1,123}, M_{1,12}, M_{1,13}, M_{1,1}), \\ M_2 &= (M_{2,123}, M_{2,12}, M_{2,23}, M_{2,2}), \\ M_3 &= (M_{3,123}, M_{3,23}, M_{3,13}, M_{3,3}), \\ R_1 &= R_{1,123} + R_{1,12} + R_{1,13} + R_{1,1}, \\ R_2 &= R_{2,123} + R_{2,12} + R_{2,23} + R_{2,2}, \\ R_3 &= R_{3,123} + R_{3,23} + R_{3,13} + R_{3,3}. \end{aligned}$$

For convenience, let $\mathbf{R}_{123} \triangleq R_{1,123} + R_{2,123} + R_{3,123}$, $\mathbf{R}_{12} \triangleq R_{1,12} + R_{2,12}$, $\mathbf{R}_{13} \triangleq R_{1,13} + R_{3,13}$, $\mathbf{R}_{23} \triangleq R_{2,23} + R_{3,23}$, $\mathbf{R}_1 \triangleq R_{1,1}$, $\mathbf{R}_2 \triangleq R_{2,2}$, $\mathbf{R}_3 \triangleq R_{3,3}$ and $\mathbf{M}_{123} \triangleq (M_{1,123}, M_{2,123}, M_{3,123})$, $\mathbf{M}_{12} \triangleq (M_{1,12}, M_{2,12})$, $\mathbf{M}_{23} \triangleq (M_{2,23}, M_{3,23})$, $\mathbf{M}_{13} \triangleq (M_{1,13}, M_{3,13})$, $\mathbf{M}_1 \triangleq M_{1,1}$, $\mathbf{M}_2 \triangleq M_{2,2}$, $\mathbf{M}_3 \triangleq M_{3,3}$.

5.1.2. Codebook Generation

Fix a pmf $p(u_{123}, u_{12}, u_{23}, u_{13}, u_1, u_2, u_3)$. Randomly and independently generate $2^{n\mathbf{R}_{123}}$ sequences $u_{123}^n(\mathbf{m})$ each according to $\prod_{i=1}^n p_{U_{123}}(u_{123,i})$. For each $\mathbf{m}_{12} \in [1 : 2^{n\mathbf{R}_{1,12}}] \times [1 : 2^{n\mathbf{R}_{2,12}}] \times [1 : 2^{n\mathbf{R}_{3,12}}]$ generate a subcodebook $\mathcal{C}_{12}(\mathbf{m}_{12})$ consisting of $2^{n(\bar{\mathbf{R}}_{12} - \mathbf{R}_{12})}$ independent generated sequences $u_{12}^n(\mathbf{m}_{12}, l_{12})$, $l_{12} \in [1 : 2^{n(\bar{\mathbf{R}}_{12} - \mathbf{R}_{12})}]$, each according to $\prod_{i=1}^n p_{U_{12}|U_{123}}(u_{12,i} | u_{123,i})$. In the same way, for each $\mathbf{m}_{23}$ and $\mathbf{m}_{13}$, generate subcodebooks $\mathcal{C}_{23}(\mathbf{m}_{23})$ and $\mathcal{C}_{13}(\mathbf{m}_{13})$, respectively. For each $\mathbf{m}_1 \in [1 : 2^{n\mathbf{R}_{1,1}}]$, generate a subcodebook $\mathcal{C}_1(\mathbf{m}_1)$ consisting of $2^{n(\bar{\mathbf{R}}_1 - \mathbf{R}_1)}$ independent generated sequences $u_1^n(\mathbf{m}_1, l_1)$, $l_1 \in [1 : 2^{n(\bar{\mathbf{R}}_1 - \mathbf{R}_1)}]$, each according to $\prod_{i=1}^n p_{U_1|U_{12}U_{13}U_{123}}(u_{1,i} | u_{12,i}u_{13,i}u_{123,i})$. For $\mathbf{m}_2$ and $\mathbf{m}_3$, generate corresponding subcodebooks similarly.

5.1.3. Encoding

For each $\mathbf{m}_S$, $S \in \{1, 2, 3, 12, 13, 23\}$, find an index tuple $(l_1, l_2, l_3, l_{12}, l_{13}, l_{23})$ such that $u_S^n(\mathbf{m}_S, l_S) \in \mathcal{C}_S(\mathbf{m}_S)$, $S \in \{1, 2, 3, 12, 13, 23\}$, and

$$(u_S^n(\mathbf{m}_S, l_S) : S \in \{1, 2, 3, 12, 13, 23\}) \in \mathcal{T}_{\epsilon'}^{(n)} \quad (14)$$

If more than one such index tuple can be found, choose an arbitrary one among those. If no such tuple exists, just choose $(l_1, l_2, l_3, l_{12}, l_{13}, l_{23}) = (1, 1, 1, 1, 1, 1)$. Then, the transmitter generates $x^n(\mathbf{m}_{123}, \mathbf{m}_{12}, \mathbf{m}_{13}, \mathbf{m}_{23}, m_{1,1}, m_{2,2}, m_{3,3})$ as

$$\begin{aligned} x_i &= x(u_{123,i}(\mathbf{m}_{123}), u_{12,i}(\mathbf{m}_{12}, l_{12}), u_{13,i}(\mathbf{m}_{13}, l_{13}), u_{23,i}(\mathbf{m}_{23}, l_{23}), \\ &\quad u_{1,i}(\mathbf{m}_1, l_1), u_{2,i}(\mathbf{m}_2, l_2), u_{3,i}(\mathbf{m}_3, l_3)), \text{ for } i = 1, \dots, n. \end{aligned}$$

5.1.4. Decoding

Decoder $j = 1$ declares that $(\hat{m}_{1,123}, \hat{m}_{1,12}, \hat{m}_{1,13}, \hat{m}_{1,1})$ is sent if it is the unique message such that

$$(u_{123}^n(\mathbf{m}_{123}), u_S^n(\mathbf{m}_S, l_S) : S \in \{1, 12, 13\}, y_1^n) \in \mathcal{T}_\epsilon^{(n)}.$$

Similarly, receiver $j = 2, 3$ uses joint typicality decoding to find the unique message $(\hat{m}_{2,123}, \hat{m}_{2,12}, \hat{m}_{2,23}, \hat{m}_{2,2})$ and $(\hat{m}_{3,123}, \hat{m}_{3,13}, \hat{m}_{3,23}, \hat{m}_{3,3})$, respectively.

5.1.5. Analysis of the Probability of Error

Assume without loss of generality that

$$(\mathbf{M}_{123}, \mathbf{M}_{12}, \mathbf{M}_{13}, \mathbf{M}_{23}, \mathbf{M}_1, \mathbf{M}_2, \mathbf{M}_3) = (\mathbf{1}, \mathbf{1}, \mathbf{1}, \mathbf{1}, \mathbf{1}, \mathbf{1}, \mathbf{1})$$

is sent and let $L_{12}, L_{13}, L_{23}, L_1, L_2, L_3$ be the index tuple of selected sequences. We note that the subcodebook $\mathcal{C}_S(\mathbf{m}_S)$ consists of $2^{n(\bar{R}_S - R_S)}$ i.i.d. $U_S^n(\mathbf{m}_S, l_S)$ sequences, $\forall S \in \{1, 2, 3, 12, 13, 23\}$. By the hierarchical covering lemma (with $r_S = \bar{R}_S - R_S$), we obtain a set of constraints for $S \in \{12, 13, 23\}$

$$r_{12} + r_{13} > I(U_{12}; U_{13} | U_{123}), \quad (15a)$$

$$r_{12} + r_{23} > I(U_{12}; U_{23} | U_{123}), \quad (15b)$$

$$r_{13} + r_{23} > I(U_{13}; U_{23} | U_{123}), \quad (15c)$$

$$r_{12} + r_{13} + r_{23} > I(U_{12}; U_{23} | U_{123}) + I(U_{13}; U_{23} | U_{123}) + I(U_{12}; U_{13} | U_{123}, U_{23}). \quad (15d)$$

Again, using hierarchical covering lemma for $S \in \{1, 2, 3\}$, we attain the region for r_1, r_2, r_3 :

$$r_1 > I(U_1; U_{23} | U_{123} U_{12} U_{13}), \quad (16a)$$

$$r_2 > I(U_2; U_{13} | U_{123} U_{12} U_{23}), \quad (16b)$$

$$r_3 > I(U_3; U_{12} | U_{123} U_{13} U_{23}), \quad (16c)$$

$$r_1 + r_2 > I(U_1; U_2 | U_{123} U_{12} U_{13} U_{23}) + I(U_1; U_{23} | U_{123} U_{12} U_{13}) + I(U_2; U_{13} | U_{123} U_{23} U_{12}), \quad (16d)$$

$$r_1 + r_3 > I(U_1; U_3 | U_{123} U_{12} U_{13} U_{23}) + I(U_1; U_{23} | U_{123} U_{12} U_{13}) + I(U_3; U_{12} | U_{123} U_{13} U_{23}), \quad (16e)$$

$$r_2 + r_3 > I(U_2; U_3 | U_{123} U_{12} U_{13} U_{23}) + I(U_2; U_{13} | U_{123} U_{12} U_{23}) + I(U_3; U_{12} | U_{123} U_{13} U_{23}), \quad (16f)$$

$$r_1 + r_2 + r_3 > I(U_1; U_2 | U_{123} U_{12} U_{23} U_{13}) + I(U_2; U_3 | U_{123} U_{12} U_{13} U_{23}) + I(U_3; U_{12} | U_{123} U_{13} U_{23}) \\ + I(U_1; U_{23} | U_{123} U_{12} U_{13}) + I(U_2; U_{13} | U_{123} U_{23} U_{12}) + I(U_1; U_3 | U_{123} U_{12}, U_{23} U_{13} U_2). \quad (16g)$$

By the symmetry of decoders, we first consider the average probability of error for decoder 1. To analyze it, the Table 1 list all possible pmfs of message tuple $(U_{123}^n(\mathbf{m}_{123}), U_{12}^n(\mathbf{m}_{12}, l'), U_{13}^n(\mathbf{m}_{13}, l'), U_1^n(\mathbf{m}_1, l'), Y_1^n)$.

$$\mathcal{E}_{1,0} = \{(U_{123}^n(\mathbf{1}), U_{12}^n(\mathbf{1}, l), U_{13}^n(\mathbf{1}, l), U_1^n(\mathbf{1}, l), Y_1^n) \notin \mathcal{T}_\epsilon^{(n)}\},$$

$$\mathcal{E}_{1,1} = \{(U_{123}^n(\mathbf{1}), U_{12}^n(\mathbf{1}, l), U_{13}^n(\mathbf{1}, l), U_1^n(*), Y_1^n) \in \mathcal{T}_\epsilon^{(n)}\},$$

$$\mathcal{E}_{1,13} = \{(U_{123}^n(\mathbf{1}), U_{12}^n(\mathbf{1}, l), U_{13}^n(*), U_1^n(*), Y_1^n) \in \mathcal{T}_\epsilon^{(n)}\},$$

$$\mathcal{E}_{1,12} = \{(U_{123}^n(\mathbf{1}), U_{12}^n(*), U_{13}^n(\mathbf{1}, l), U_1^n(*), Y_1^n) \in \mathcal{T}_\epsilon^{(n)}\},$$

$$\mathcal{E}_{1,12,13} = \{(U_{123}^n(\mathbf{1}), U_{12}^n(*), U_{13}^n(*), U_1^n(*), Y_1^n) \in \mathcal{T}_\epsilon^{(n)}\},$$

$$\mathcal{E}_{1,123} = \{(U_{123}^n(*), U_{12}^n(*), U_{13}^n(*), U_1^n(*), Y_1^n) \in \mathcal{T}_\epsilon^{(n)}\}.$$

Table 1. Table of joint pmfs of all possible error messages. Symbol * denotes the message or index $\mathbf{m}_s \neq \mathbf{1}_s, l'_s \neq l_s, s \in \{123, 12, 13, 1\}$, and p^* denote joint pmf $p(u_{123}^n, u_{12}^n, u_{13}^n, u_1^n)$.

$\mathbf{m}_{123}$	$(\mathbf{m}_{12}, l'_{12})$	$(\mathbf{m}_{13}, l'_{13})$	$m_{1,1}$	Joint pmf
1	$(\mathbf{1}, l_{12})$	$(\mathbf{1}, l_{13})$	1	$p^* p(y_1^n u_{123}^n u_{12}^n u_{13}^n u_1^n)$
1	$(\mathbf{1}, l_{12})$	$(\mathbf{1}, l_{13})$	*	$p^* p(y_1^n u_{123}^n u_{12}^n u_{13}^n)$
1	$(\mathbf{1}, l_{12})$	*	*	$p^* p(y_1^n u_{123}^n u_{12}^n)$
1	*	$(\mathbf{1}, l_{13})$	*	$p^* p(y_1^n u_{123}^n u_{13}^n)$
1	*	*	*	$p^* p(y_1^n u_{123}^n)$
*	*	*	*	$p^* p(y_1^n)$

By the conditional typicality lemma, the first term $P(\mathcal{E}_{1,0})$ tends to zero as $n \rightarrow \infty$ as $(u_{\mathcal{S}}^n(\mathbf{m}_{\mathcal{S}}, l_{\mathcal{S}}) : \mathcal{S} \in \{1, 2, 3, 12, 13, 23\}) \in \mathcal{T}_{\epsilon'}^{(n)}$ and $\epsilon' < \epsilon$.

Considering the second term $P(\mathcal{E}_{1,1})$, note that $U_1^n(l_1) \sim \prod_{i=1}^n P_{U_1}(u_{1,i})$ is independent of Y_1^n for every $U_1(l_1) \notin \mathcal{C}_1(1)$, then the variables $U_1^n(*)$ is independent of Y_1^n on condition of $(U_{123}^n(\mathbf{1}), U_{12}^n(\mathbf{1}, l_{12}), U_{13}^n(\mathbf{1}, l_{13}))$. By the packing lemma, $P(\mathcal{E}_{1,1})$ tends to zero as $n \rightarrow \infty$ if

$$\tilde{R}_1 < I(U_1; Y_1 | U_{123} U_{12} U_{13}).$$

Similarly, $P(\mathcal{E}_{1,12}), P(\mathcal{E}_{1,13}), P(\mathcal{E}_{1,12,13})$ and $P(\mathcal{E}_{1,123})$ tends to zero as $n \rightarrow \infty$ if

$$\begin{aligned} \tilde{R}_{12} + \tilde{R}_1 &< I(U_{12} U_1; Y_1 | U_{123} U_{13}) + I(U_{12}; U_{13} | U_{123}), \\ \tilde{R}_{13} + \tilde{R}_1 &< I(U_{13} U_1; Y_1 | U_{123} U_{12}) + I(U_{12}; U_{13} | U_{123}), \\ \tilde{R}_{12} + \tilde{R}_{13} + \tilde{R}_1 &< I(U_{12} U_{13} U_1; Y_1 | U_{123}) + I(U_{12}; U_{13} | U_{123}), \\ R_{123} + \tilde{R}_{12} + \tilde{R}_{13} + \tilde{R}_1 &< I(U_{123} U_{12} U_{13} U_1; Y_1) + I(U_{12}; U_{13} | U_{123}). \end{aligned}$$

For receiver 2 and 3, the corresponding inequalities can be derived in a similar way.

From the lower bound of linear combinations of $r_{\mathcal{S}}$ and the upper bound of linear combinations of $\tilde{R}_{\mathcal{S}}$ about receivers 1–3, eliminating $\tilde{R}_{123}, \tilde{R}_{12}, \tilde{R}_{13}, \tilde{R}_{23}$ and $\tilde{R}_1, \tilde{R}_2, \tilde{R}_3$ by our elimination procedure yields the characterization (13a).

The procedure of the elimination for the case where $K=3$ is as follows, which is based on Lemma 3 that the sum of lower bounds for $\sum_{\mathcal{S} \in \mathcal{J}_i^l} r_{\mathcal{S}}, \forall i \in [N]$ is not greater than the lower bound for $\sum_{\mathcal{S} \in \mathcal{J}^l} r_{\mathcal{S}}$ and observation (21) and (22) that shows the properties of the bounds derived from packing lemma. The detailed explanation for the procedure is elaborated in Section 5.2.5.

To start with, $\mathcal{T} = \{1, 2, 3\}$ and $\pi_{\{1,2,3\}} = \{1, 2, 3\}$, as $R_1 = R_{1,123} + R_{1,12} + R_{1,13} + R_{1,1}, R_2 = R_{2,123} + R_{2,12} + R_{2,23} + R_{2,2}, R_3 = R_{3,123} + R_{3,23} + R_{3,13} + R_{3,3}$ and $\mathbf{R}_{123} \triangleq R_{1,123} + R_{2,123} + R_{3,123}, \mathbf{R}_{12} \triangleq R_{1,12} + R_{2,12}, \mathbf{R}_{13} \triangleq R_{1,13} + R_{3,13}, \mathbf{R}_{23} \triangleq R_{2,23} + R_{3,23}, \mathbf{R}_1 \triangleq R_{1,1}, \mathbf{R}_2 \triangleq R_{2,2}, \mathbf{R}_3 \triangleq R_{3,3}$. We can express $R_1 + R_2 + R_3$ as

$$\begin{aligned} &R_{1,1} + R_{1,12} + R_{1,13} + R_{1,123} \\ &+ R_{2,2} + R_{2,12} + R_{2,23} + R_{2,123} \\ &+ R_{3,3} + R_{3,13} + R_{3,23} + R_{3,123}, \end{aligned}$$

where with composing order $\pi_{\{1,2,3\}} = \{1, 2, 3\}$, $R_1 + R_2 + R_3$ can be expressed as $(\mathbf{R}_{123} + \mathbf{R}_{12} + \mathbf{R}_{13} + \mathbf{R}_1) + (\mathbf{R}_{23} + \mathbf{R}_2) + (\mathbf{R}_3)$. The additives in the brackets are components from corresponding inequalities resulted from packing lemma, namely,

$$R_1 + R_2 + R_3 = (\mathbf{R}_{123} + \tilde{R}_{12} + \tilde{R}_{13} + \tilde{R}_1) + (\tilde{R}_{23} + \tilde{R}_2) + (\tilde{R}_3),$$

which is remained to combine the lower and upper bounds of the corresponding combination to derive one upper bound as shown in Equations (13g)–(13i).

Then, for $\mathcal{T} = \{1, 2\}$ and $\pi_{\{1,2\}} = \{1, 2\}$, we can write $R_1 + R_2$ as

$$\begin{aligned} & R_{1,1} + R_{1,12} + R_{1,13} + R_{1,123} \\ & + R_{2,2} + R_{2,12} + R_{2,23} + R_{2,123}, \end{aligned}$$

which is equal to $(R_{123} + \tilde{R}_{12} + \tilde{R}_{13} + \tilde{R}_1) + (\tilde{R}_{23} + \tilde{R}_2) - (R_{3,123} + R_{3,13} + R_{3,23})$ following the order $\pi_{\{1,2\}} = \{1, 2\}$, where $(R_{3,123} + R_{3,13} + R_{3,23})$ can be ignored as they are non-negative. Applying the corresponding bounds again, we obtain one from the Equation (13d).

For $\mathcal{T} = \{1\}$ and $\pi_{\{1\}} = \{1\}$, by following the previous process we get Equation (13a). Thus, after consider $\forall \mathcal{T} \subseteq \mathcal{K}$ and all permutations $\pi_{\mathcal{T}} \in \Pi_{\mathcal{T}}$, we get all inequalities in Equations (13a)–(13i).

5.2. Coding Scheme for $K \geq 2$

For each $k \in \mathcal{K}$, split message $M_k \in [1 : 2^{nR_k}]$ into $\sum_{j=1}^K \binom{K}{j-1}$ submessages $M_k \triangleq (M_{k,\mathcal{S}} : \mathcal{S} \subseteq \mathcal{K}, |\mathcal{S}| = 1, 2, \dots, K)$ with $M_{k,\mathcal{S}} \in [1 : 2^{nR_{k,\mathcal{S}}}]$ and $R_{k,\mathcal{S}} = 0$ if $k \notin \mathcal{S}$. Similarly to $K = 3$ case, we use the following notations:

$$\tilde{R}_{\mathcal{S}} \triangleq \sum_{k=1}^K \tilde{R}_{k,\mathcal{S}}, \quad R_{\mathcal{S}} \triangleq \sum_{k=1}^K R_{k,\mathcal{S}}, \quad r_{\mathcal{S}} \triangleq \tilde{R}_{\mathcal{S}} - R_{\mathcal{S}}, \quad (17)$$

and $\mathbf{M}_{\mathcal{S}} \triangleq (M_{1,\mathcal{S}}, \dots, M_{K,\mathcal{S}})$. To simplify notations, we denote $U_{\mathcal{S}}^n(\mathbf{m}_{\mathcal{S}}, l_{\mathcal{S}} | \mathbf{m}_{\mathbb{A}_{\mathcal{S}}^{1:K} \setminus \mathcal{S}}, l_{\mathbb{A}_{\mathcal{S}}^{1:K} \setminus \mathcal{S}})$ by $U_{\mathcal{S}}^n(\mathbf{m}_{\mathcal{S}}, l_{\mathcal{S}})$ and its realization $u_{\mathcal{S}}^n(\mathbf{m}_{\mathcal{S}}, l_{\mathcal{S}} | \mathbf{m}_{\mathbb{A}_{\mathcal{S}}^{1:K} \setminus \mathcal{S}}, l_{\mathbb{A}_{\mathcal{S}}^{1:K} \setminus \mathcal{S}})$ by $u_{\mathcal{S}}^n(\mathbf{m}_{\mathcal{S}}, l_{\mathcal{S}})$.

5.2.1. Codebook Generation

Fix a pmf $p(u(\mathbb{A}^{1:K}))$ and function $x(u(\mathbb{A}^{1:K}))$ and let $\tilde{R}_{\mathcal{S}} \geq R_{\mathcal{S}}, \forall \mathcal{S} \in \mathbb{A}^{1:K}$. Randomly and independently generate $2^{nR_{\mathcal{K}}}$ sequences $u_{\mathcal{K}}^n(\mathbf{m}_{\mathcal{K}}), \mathbf{m}_{\mathcal{K}} \in \prod_{i=1}^K [1 : 2^{nR_{i,\mathcal{K}}}]$ each according to $\prod_{i=1}^n p_{U_{\mathcal{K}}}(u_{\mathcal{K},i})$. For all $\mathcal{S} \in \mathbb{A}^{1:(K-1)}$, construct a subcodebook $\mathcal{C}_{\mathcal{S}}(\mathbf{m}_{\mathcal{S}})$ consisting of $2^{n(\tilde{R}_{\mathcal{S}} - R_{\mathcal{S}})}$ i.i.d. generated sequences $u_{\mathcal{S}}^n(\mathbf{m}_{\mathcal{S}}, j_{\mathcal{S}}), (\mathbf{m}_{\mathcal{S}}, j_{\mathcal{S}}) \in \prod_{i=1}^K [1 : 2^{nR_{i,\mathcal{S}}}] \times [1 : 2^{n(\tilde{R}_{\mathcal{S}} - R_{\mathcal{S}})}]$, each according to $\prod_{i=1}^n p_{U_{\mathcal{S}} | U(\mathbb{A}_{\mathcal{S}}^{1:K} \setminus \mathcal{S})}(u_{\mathcal{S},i} | u_{\mathcal{S}',i} : \mathcal{S}' \in (\mathbb{A}_{\mathcal{S}}^{1:K} \setminus \mathcal{S}))$.

5.2.2. Encoding

We iteratively find all indices $(j_{\mathcal{S}} : \mathcal{S} \in \mathbb{A}^{1:(K-1)})$ such that all selected codewords are jointly typical. In each round $l = K - 1, K - 2, \dots, 1$, given messages $(\mathbf{m}_{\mathcal{S}}, \mathcal{S} \in \mathbb{A}^l)$ and sequences $(u_{\mathcal{S}}^n(\mathbf{m}_{\mathcal{T}}, j_{\mathcal{T}}) : \mathcal{T} \in \mathbb{A}^{l+1:K-1})$, find indices $(j_{\mathcal{S}} : \mathcal{S} \in \mathbb{A}^l)$, such that all

$$\left((u_{\mathcal{S}}^n(\mathbf{m}_{\mathcal{S}}, j_{\mathcal{S}}) : \mathcal{S} \in \mathbb{A}^l), (u_{\mathcal{S}}^n(\mathbf{m}_{\mathcal{T}}, j_{\mathcal{T}}) : \mathcal{T} \in \mathbb{A}^{l+1:K-1}), u_{\mathcal{K}}^n(\mathbf{m}_{\mathcal{K}}) \right) \in \mathcal{T}_{\epsilon_1}^{(n)}.$$

If there is more than one such tuple, pick an arbitrary one among them. If no such tuple exists, pick $(j_{\mathcal{S}} : \mathcal{S} \in \mathbb{A}^{1:K}) = \mathbf{1}$. Then, generate $x^n(\mathbf{m}_{\mathcal{S}} : \mathcal{S} \in \mathbb{A}^{1:K})$ with

$$x_i = x(u_{\mathcal{K},i}(\mathbf{m}_{\mathcal{K}}), u_{\mathcal{S},i}(\mathbf{m}_{\mathcal{S}}, j_{\mathcal{S}}) : \forall \mathcal{S} \in \mathbb{A}^{1:(K-1)}),$$

for $i \in [n]$.

To send the message tuple $(m_1, \dots, m_K) = (\mathbf{m}_{\mathcal{S}} : \mathcal{S} \in \mathbb{A}^{1:K})$, transmit $x^n(\mathbf{m}_{\mathcal{S}} : \mathcal{S} \in \mathbb{A}^{1:K})$.

5.2.3. Decoding

Let $\epsilon > \epsilon'$. Decoder $k, k \in \mathcal{K}$, declares that $\hat{m}_k$ is sent if it is the unique message such that

$$(y_k^n, u_{\mathcal{S}}^n(\mathbf{m}_{\mathcal{S}}, j_{\mathcal{S}}) : \mathcal{S} \in \mathbb{A}^{1:(K-1)}(i), u_{\mathcal{K}}^n(\mathbf{m}_{\mathcal{K}})) \in \mathcal{T}_{\epsilon}^{(n)},$$

for some $(u_S^n(\mathbf{m}_S, j_S) : \mathcal{S} \in \mathbb{A}_{\{k\}}^{1:(K-1)}, u_K^n(\mathbf{m}_K)) \in \prod_{\mathcal{S} \in \mathbb{A}^{1:(K-1)}} \mathcal{C}_S(\mathbf{1}) \times \mathcal{C}_K(\mathbf{1})$.

5.2.4. Analysis of the Probability of Error

Assume without loss of generality that $(M_1, \dots, M_K) = \mathbf{1}$ and let $(J_S : \mathcal{S} \in \mathbb{A}^{1:(K-1)}) = \mathbf{1}$ be the index tuple of the chosen sequences $(U_S : \mathcal{S} \in \mathbb{A}^{1:(K-1)}) \in \prod_{\mathcal{S} \in \mathbb{A}^{1:(K-1)}} \mathcal{C}_S(\mathbf{1})$. Then, decoder i makes an error only if one or more of the following events occur:

$$\begin{aligned}\mathcal{E}_0 &= \{(U_S^n(\mathbf{1}, j_S) : \mathcal{S} \in \mathbb{A}^{1:(K-1)}, U_K^n(\mathbf{1})) \notin \mathcal{T}_e^{(n)}, \forall j_S \in [1 : 2^{nr_S}], \exists l \in [1 : (K-1)]\}, \\ \mathcal{E}_{k,1} &= \{(Y_k^n, U_S^n(\mathbf{1}, 1) : \mathcal{S} \in \mathbb{A}_{\{k\}}^{1:(K-1)}, U_K^n(\mathbf{1})) \notin \mathcal{T}_e^{(n)}\}, \\ \mathcal{E}_{k,2} &= \{(Y_k^n, U_S^n(\mathbf{M}_S, J_S)) \in \mathcal{T}_e^{(n)}, \exists \mathcal{S} \subseteq [1 : K], \mathbf{M}_S \neq \mathbf{1} \text{ or } J_S \neq \mathbf{1}\}.\end{aligned}$$

Therefore, the probability of error for decoder i is upper bounded as

$$P(\mathcal{E}_k) \leq P(\mathcal{E}_0) + P(\mathcal{E}_0^c \cap \mathcal{E}_{k,1}) + P(\mathcal{E}_0^c \cap \mathcal{E}_{k,2}).$$

To bound $P(\mathcal{E}_0)$, we utilize the hierarchical covering lemma in Lemma 2. Therefore, $P(\mathcal{E}_0)$ tends to zero as $n \rightarrow \infty$ if

$$\sum_{\mathcal{S} \in \mathcal{J}^l} r_S > \sum_{\mathcal{S} \in \mathcal{J}^l} H(U_S | U(\mathbb{A}_S^{1:K} \setminus \mathcal{S})) - H(U(\mathcal{J}^l) | U(\mathbb{A}^{(l+1):K})) + \delta(\epsilon_l), \quad (18)$$

holds for all $l \in [1 : (K-1)]$ and $\mathcal{J}^l \in \mathbb{A}^l$.

To bound $P(\mathcal{E}_0^c \cap \mathcal{E}_{k,1})$, by the conditional typicality lemma [2] it tends to zero as $n \rightarrow \infty$.

To bound $P(\mathcal{E}_0^c \cap \mathcal{E}_{k,2})$, note that $U_S^n(\mathbf{1}, 1) \sim \prod_{i=1}^n p_{U_S | U(\mathbb{A}_S^{1:K} \setminus \mathcal{S})}(u_{S,i} | u_{S',i} : \mathcal{S}' \in \mathbb{A}_S^{1:K} \setminus \mathcal{S})$ and Y_k^n are independent of $U_S^n(\mathbf{M}_S, J_S) \notin \mathcal{C}_S(\mathbf{1}), \forall \mathcal{S} \in \mathbb{A}_{\{k\}}^{1:(K-1)}$, as well as $U_K^n(\mathbf{M}_K) \notin \mathcal{C}_K(\mathbf{1})$. Furthermore, if $\mathbf{M}_T \neq \mathbf{1}$ or $J_T \neq \mathbf{1}$ for any subset $\mathcal{T} \in \mathbb{A}_{\{k\}}^{1:K}$, then by the conditional coding distribution we have $U_S^n(\mathbf{M}_S, j_S) \notin \mathcal{C}_S(\mathbf{1})$ for all $\mathcal{S} \in \mathbb{A}_{\{k\}}^{1:K}(\mathcal{T})$. Thus, $\forall k \in \mathcal{K}, \forall \mathcal{J} \subseteq \mathbb{A}_{\{k\}}^{1:K}$ by identifying $\bigcup_{\mathcal{T} \in \mathcal{J}} \mathbb{A}_{\{k\}}^{1:K}(\mathcal{T})$, we obtain inequalities with the help of packing lemma: for all $\forall \mathcal{J} \subseteq \mathbb{A}_{\{k\}}^{1:K}$,

$$\begin{aligned}\sum_{\mathcal{S} \in \bigcup_{\mathcal{T} \in \mathcal{J}} \mathbb{A}_{\{k\}}^{1:K}(\mathcal{T})} \left(\sum_{k=1}^K \tilde{R}_{k,S} + r_S \right) &< \sum_{\mathcal{S} \in \bigcup_{\mathcal{T} \in \mathcal{J}} \mathbb{A}_{\{k\}}^{1:K}(\mathcal{T})} H(U_S | U(\mathbb{A}_S^{1:K} \setminus \mathcal{S})) \\ &- H(U(\bigcup_{\mathcal{S} \in \mathcal{J}} \mathbb{A}_{\{k\}}^{1:K}(\mathcal{S})) | Y_k U(\mathbb{A}_{\{k\}}^{1:K} \setminus \bigcup_{\mathcal{S} \in \mathcal{J}} \mathbb{A}_{\{k\}}^{1:K}(\mathcal{S}))) - \delta(\epsilon).\end{aligned} \quad (19)$$

5.2.5. Eliminating $(\tilde{R}_{k,S}, r_S : k \in \mathcal{K}, \mathcal{S} \in \mathbb{A}^{1:K})$

Due to massive numbers of $\tilde{R}_{k,S}$ and r_S , using standard Fourier–Motzkin elimination to obtain the achievable rate region is disastrous. We first present observations which help the elimination, and then find all valid constraints for $\sum_{k \in \mathcal{T}} R_k$ for all $\mathcal{T} \subseteq \mathcal{K}$.

Observation: Let $\mathbb{S}(k, \mathcal{J}) \triangleq \bigcup_{\mathcal{T} \in \mathcal{J}} \mathbb{A}_{\{k\}}^{1:K}(\mathcal{T})$, denote the right hand term of (19) by $I_{\mathbb{S}(k, \mathcal{J})}$, and from (17), (19) can be rewritten as below.

$$\sum_{\mathcal{S} \in \mathbb{S}(k, \mathcal{J})} \tilde{R}_S < I_{\mathbb{S}(k, \mathcal{J})}. \quad (20)$$

(a) If there exists k', k'' and $\mathcal{J}', \mathcal{J}''$ such that

$$\mathbb{S}(k, \mathcal{J}) = \mathbb{S}(k', \mathcal{J}') \cup \mathbb{S}(k'', \mathcal{J}''), \quad (21a)$$

then

$$I_{\mathbb{S}(k, \mathcal{T})} \leq I_{\mathbb{S}(k', \mathcal{T}')} + I_{\mathbb{S}(k'', \mathcal{T}'')}. \quad (21b)$$

(b) If there exists k' and $\mathcal{T}'$ such that $\mathbb{S}(k, \mathcal{T}) \subseteq \mathbb{S}(k', \mathcal{T}')$, then

$$I_{\mathbb{S}(k, \mathcal{T})} \leq I_{\mathbb{S}(k', \mathcal{T}')}. \quad (22)$$

From the rate-splitting procedure, we have $R_k = \sum_{\mathcal{S} \in \mathbb{A}_{\{k\}}^{1:K}} R_{k, \mathcal{S}}$, for all $k \in \mathcal{K}$. From (17), (20) and the observations (21) and (22), we can obtain that the valid constraints for $R_1, \dots, R_K$ will only have form $\sum_{k \in \mathcal{T}} R_k$, i.e., any rate constraint $\sum_j a_j R_j$, for all $a_j \geq 0$ can be derived from $\sum_{k \in \mathcal{T}} R_k$, for all $\mathcal{T} \subseteq \mathcal{K}$. Note that this will not hold for DM-BCs with common message (there exists an original source message to be sent to all receivers).

Thus, in order to find rate constraints for $\sum_{k \in \mathcal{T}} R_k$, we must find all valid constraints for $\sum_{k \in \mathcal{T}} \sum_{\mathcal{S} \in \mathbb{A}_{\{k\}}^{1:K}} R_{k, \mathcal{S}}$. By definitions in (17), we have

$$\sum_{k \in \mathcal{T}} R_k = \sum_{k \in \mathcal{T}} \sum_{\mathcal{S} \subseteq \mathcal{K}, k \in [\mathcal{S}]} R_{k, \mathcal{S}} = \sum_{\mathcal{S} \in \bigcup_{k \in \mathcal{T}} \mathbb{A}_{\{k\}}^{1:K}} R_{\mathcal{S}} = \sum_{\mathcal{S} \in \bigcup_{k \in \mathcal{T}} \mathbb{A}_{\{k\}}^{1:K}} \tilde{R}_{\mathcal{S}} - \sum_{\mathcal{S} \in \bigcup_{k \in \mathcal{T}} \mathbb{A}_{\{k\}}^{1:K}} r_{\mathcal{S}}. \quad (23)$$

Now we show that how to obtain the closed-form rate region via set decomposition in (7) and the observations in Lemma 3, (21) and (22).

Consider a permutation $\pi_{\mathcal{T}} \in \Pi_{\mathcal{T}}$. For easy reading, from the decomposition (7), we have

$$\bigcup_{k \in \mathcal{T}} \mathbb{A}_{\{k\}}^{1:K} = \bigsqcup_{i=1}^{|\mathcal{T}|} \mathbb{B}_{\pi(i)}^{1:K}. \quad (24)$$

For $R_{\pi(1)}$, it must involve $R_{\pi(1), \mathcal{S}} : \mathcal{S} \in \mathbb{A}_{\pi(1)}^{1:K}$. Since $\mathbb{A}_{\{\pi(1)\}}^{1:K} = \mathbb{B}_{\pi(1)}^{1:K} = \mathbb{S}(\pi(1), \mathcal{K})$ is the largest set with every element containing $\pi(1)$, by observation in (21) and (22), we have the only valid rate constraint

$$\sum_{\mathcal{S} \in \mathbb{A}_{\pi(1)}^{1:K}} \tilde{R}_{\mathcal{S}} \leq I_{\mathbb{A}_{\{\pi(1)\}}^{1:K}} = I_{\mathbb{B}_{\pi(1)}^{1:K}}$$

to support $R_{\pi(1)}$, where the last equality holds by (6).

For $R_{\pi(2)}$, $\sum_{\mathcal{S} \in \mathbb{A}_{\pi(1)}^{1:K}(\pi(1))} \tilde{R}_{\mathcal{S}}$ already contains $\sum_{\mathcal{S} \in \mathbb{A}_{\{\pi(1)\pi(2)\}}^{1:K}} R_{k, \mathcal{S}}$, thus we need to find the rate constraints for $\sum_{\mathcal{S} \in \mathbb{A}_{\pi(2)}^{1:K} \setminus \mathbb{A}_{\{\pi(1)\pi(2)\}}^{1:K}} R_{k, \mathcal{S}}$. Since $\mathbb{A}_{\pi(2)}^{1:K} \setminus \mathbb{A}_{\{\pi(1)\pi(2)\}}^{1:K} = \mathbb{B}_{\pi(2)}^{1:K} = \mathbb{S}(\pi(2), \mathcal{K} \setminus \{\pi(1)\})$, and $\mathbb{B}_{\pi(2)}^{1:K}$ is second largest sets with every element containing $\pi(2)$ and excluding $\pi(1)$, by observation in (21) and (22), we obtain the only effective rate constraints

$$\sum_{\mathcal{S} \in \mathbb{B}_{\pi(2)}^{1:K}} \tilde{R}_{\mathcal{S}} \leq I_{\mathbb{B}_{\pi(2)}^{1:K}}$$

to support $R_{\pi(1)} + R_{\pi(2)}$. Iteratively applying the similar steps above and by (23), we find all valid rate constraints under the permutation π for $\sum_{k \in \mathcal{T}} R_k$:

$$\sum_{k \in \mathcal{T}} R_k = \sum_{\mathcal{S} \in \bigcup_{k \in \mathcal{T}} \mathbb{A}_{\{k\}}^{1:K}} \tilde{R}_{\mathcal{S}} - \sum_{\mathcal{S} \in \bigcup_{k \in \mathcal{T}} \mathbb{A}_{\{k\}}^{1:K}} r_{\mathcal{S}} \quad (25)$$

$$\stackrel{(a)}{\leq} \sum_{i=1}^{|\mathcal{T}|} I_{\mathbb{B}_{\pi(i)}^{1:K}} - \sum_{\mathcal{S} \in \bigcup_{k \in \mathcal{T}} \mathbb{A}_{\{k\}}^{1:K}} r_{\mathcal{S}}, \quad (26)$$

where (a) holds by (20) and $\mathbb{A}_{\pi(i)}^{1:K} \setminus \mathbb{A}_{\{\pi(1), \dots, \pi(i-1)\}}^{1:K} = \mathbb{B}_{\pi(i)}^{1:K} = \mathbb{S}(\pi(i), \mathcal{K} \setminus \{\pi(1), \dots, \pi(i-1)\})$.

Thus, formally we have for all $\mathcal{T} \subseteq \mathcal{K}$ and $\pi \in \Pi_{\mathcal{T}}$:

$$\begin{aligned} \sum_{k \in \mathcal{T}} R_k &\leq \sum_{i=1}^{|\mathcal{T}|} I_{\mathbb{B}_{\pi(i)}^{1:K}} - \sum_{\mathcal{S} \in \bigcup_{k \in \mathcal{T}} \mathbb{A}_k^{1:K}} r_{\mathcal{S}} \\ &\stackrel{(a)}{=} \sum_{i=1}^{|\mathcal{T}|} I_{\mathbb{B}_{\pi(i)}^{1:K}} - \sum_{l=1}^{K-1} \sum_{\mathcal{S} \in \bigsqcup_{i=1}^{|\mathcal{T}|} \mathbb{B}_{\pi(i)}^l} r_{\mathcal{S}} \\ &\stackrel{(b)}{<} \sum_{i=1}^{|\mathcal{T}|} \left(\sum_{\mathcal{S} \in \mathbb{B}_{\pi(i)}^{1:K}} H(U_{\mathcal{S}} | U(\mathbb{A}_{\mathcal{S}}^{1:K} \setminus \mathcal{S})) - H(U(\mathbb{B}_{\pi(i)}^{1:K}) | Y_{\pi(i)} U(\mathbb{A}_{\pi(i)}^{1:K} \setminus \mathbb{B}_{\pi(i)}^{1:K})) \right) \\ &\quad + \sum_{l=1}^{K-1} \left(H\left(U\left(\bigsqcup_{i=1}^{|\mathcal{T}|} \mathbb{B}_{\pi(i)}^l\right) | U(\mathbb{A}^{(l+1):K})\right) - \sum_{\mathcal{S} \in \bigsqcup_{i=1}^{|\mathcal{T}|} \mathbb{B}_{\pi(i)}^l} H(U_{\mathcal{S}} | U(\mathbb{A}_{\mathcal{S}}^{1:K} \setminus \mathcal{S})) \right), \end{aligned}$$

where (a) comes from the disjoint decomposition in (7); (b) holds by letting $\mathcal{J}^l = \bigsqcup_{i=1}^{|\mathcal{T}|} \mathbb{B}_{\pi(i)}^l$ in Lemma 2 and by Lemma 3.

6. Conclusions

In this paper, we propose new scheme for K -receiver DM-BCs with private messages based on exhaustive message splitting and K -level Marton's coding. A hierarchical covering lemma is established which extends the 2-level multivariate covering lemma to K -level case. Our works provides a closed-form achievable rate region for the general K -user DM-BCs. In the future, we will find out some examples to justify that K -level Marton's coding is strictly better than the 2-level one.

Author Contributions: R.T. investigates the related work, carries out the formal analysis, proposes the methodology, and writes the manuscript. Y.W. proposes the main idea, conceptualizes the problem, reviews and edits the manuscripts, and supervises the progress of the work. S.X. carries out the formal analysis, proposes the methodology, and writes some part of the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by the National Nature Science Foundation of China (NSFC) under Grant 61901267.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors have no conflict of interest to disclose.

Appendix A. Proof of Hierarchical Covering Lemma 2

Let

$$\mathcal{M}_{\mathbb{A}^l} \triangleq \prod_{\mathcal{S} \in \mathbb{A}^l} \mathcal{M}_{\mathcal{S}} = \prod_{\mathcal{S} \in \mathbb{A}^l} [1 : 2^{nr_{\mathcal{S}}}] \quad (\text{A1})$$

Divide the messages and codewords into K levels, where at level $l \in [1 : K]$, there are messages sets $\{\mathcal{M}_{\mathcal{S}} : \mathcal{S} \in \mathbb{A}^l\}$, and codewords

$$(U_{\mathcal{S}}^n(m_{\mathcal{S}}) : \mathcal{S} \in \mathbb{A}^l, m_{\mathcal{S}} \in \mathcal{M}_{\mathbb{A}^l}).$$

The proof of hierarchical covering lemma starts at level $l = K - 1$, finding a message tuple $(m_S : S \in \mathbb{A}^{K-1})$ such that $(u_S^n(m_S) : S \in \mathbb{A}^{K-1}, u_K^n) \in \mathcal{T}_{\epsilon_{K-1}}^{(n)}$. Then, with a descending order of $l = K - 2, K - 1, \dots, 1$, we find a message tuple $(m_S : S \in \mathbb{A}^l)$ such that $((u_S^n(m_S) : S \in \mathbb{A}^l), u_K^n) \in \mathcal{T}_{\epsilon_l}^{(n)}$.

For each $l \in [1 : K]$, we use $\mathcal{S}_i^l$ to denote the element of $\mathbb{A}^l$ for $i = 1, \dots, |\mathbb{A}^l|$, i.e.,

$$\mathbb{A}^l = \{\mathcal{S}_1^l, \dots, \mathcal{S}_{|\mathbb{A}^l|}^l\}.$$

We define, for all $l \in [1 : K]$, the random events

$$\begin{aligned} \mathcal{E}_l &= \left\{ \left((U_S^n(m_S) : S \in \mathbb{A}^{l:(K-1)}), U_K^n \right) \notin \mathcal{T}_{\epsilon_l}^{(n)}, \forall (m_S : S \in \mathbb{A}^{l:(K-1)}) \in \prod_{S \in \mathbb{A}^{l:(K-1)}} [1 : 2^{nr_S}] \right\}, \\ \mathcal{A}_l &= \{ (m_{\mathcal{S}_1^l}, \dots, m_{\mathcal{S}_{|\mathbb{A}^l|}^l}) \in \mathcal{M}_{\mathbb{A}^l} : \\ &\quad (U^n(m_{\mathcal{S}_1^l}) \dots U^n(m_{\mathcal{S}_{|\mathbb{A}^l|}^l}), u^n(\mathbb{A}^{l+1:K})) \in \mathcal{T}_{\epsilon_l}^{(n)}, |u^n(\mathbb{A}^{l+1:K}) \in \mathcal{T}_{\epsilon_{l+1}}^{(n)} \}, \\ \mathcal{Q}_l &= (u^n(m_{\mathcal{S}_1^j}) \dots u^n(m_{\mathcal{S}_{|\mathbb{A}^j|}^j}) : j \in [(l+1):K]) \in \mathcal{T}_{\epsilon_l}^{(n)}, \end{aligned}$$

where

$$u^n(\mathbb{A}^{l+1:K}) = (u^n(m_{\mathcal{S}_1^j}) \dots u^n(m_{\mathcal{S}_{|\mathbb{A}^j|}^j}) : j \in [(l+1):K]).$$

Note that $\mathcal{E}_l$ denotes the event that there isn't any tuple of sequences from level l to K such that they are jointly typical. Correspondingly, $\mathcal{E}_l^c$ denotes the event that there exists such tuple of sequences. $\mathcal{A}_l$ denotes the set of messages at level l such that the corresponding codewords are jointly typical with a given tuple of jointly typical codewords from levels $l + 1$ to K . $\mathcal{Q}_l$ denotes a condition that there exists message $(m_{\mathcal{S}_1^l}, \dots, m_{\mathcal{S}_{|\mathbb{A}^l|}^l}) \in \mathcal{M}_{\mathbb{A}^l}$ at level $l + 1, \dots, K$ such that the corresponding codewords are jointly typical.

Our goal is to find rate conditions such that there exists a tuple of messages $(m_S : S \in \mathbb{A}^{1:K})$ satisfying $(U_S^n(m_S) : S \in \mathbb{A}) \in \mathcal{T}_{\epsilon_1}^{(n)}$, which is equivalent to $P\{\mathcal{E}_1\} \rightarrow 0$ as $n \rightarrow \infty$.

The probability of $\mathcal{E}_1$ can be upper bounded as

$$\begin{aligned} P\{\mathcal{E}_1\} &= P\{\mathcal{E}_1 \cap \mathcal{E}_2^c\} + P\{\mathcal{E}_1 \cap \mathcal{E}_2\} \\ &\leq P\{\mathcal{E}_1 | \mathcal{E}_2^c\} + P\{\mathcal{E}_2\}. \end{aligned}$$

With recursively upper bounding $P(\mathcal{E}_l)$, we obtain

$$\begin{aligned} P\{\mathcal{E}_1\} &\leq \sum_{l=1}^{K-1} P\{\mathcal{E}_l | \mathcal{E}_{l+1}^c\} + P\{\mathcal{E}_K\} \\ &= \sum_{l=1}^{K-1} P\{\mathcal{A}_l\} + P\{\mathcal{E}_K\}, \end{aligned} \tag{A2}$$

where the last equality holds by definitions of $\mathcal{A}_l$ and $\mathcal{E}_l$. As $P(\mathcal{E}_K)$ tends to zero as $n \rightarrow \infty$ by the property of typicality, and K is a fixed channel parameter, in order to make $P\{\mathcal{E}_1\}$ tend to 0, it is sufficient to let $P\{\mathcal{A}_l\}$ tend 0 as $n \rightarrow \infty$ for all $l \in [1 : K - 1]$.

To analyze (A2), $P(\mathcal{E}_K)$ tends to zero as $n \rightarrow \infty$ by the condition of the lemma. Each term $P\{\mathcal{E}_l | \mathcal{E}_{l+1}^c\}$ is equivalent to $P\{|\mathcal{A}_l| = 0\}$, which is upper bounded by a variation of Chebyshev inequality:

$$P\{|\mathcal{A}_l| = 0\} \leq P\{(|\mathcal{A}_l| - E|\mathcal{A}_l|)^2 \leq (E|\mathcal{A}_l|)^2\} \leq \frac{\text{Var}(|\mathcal{A}_l|)}{(E|\mathcal{A}_l|)^2}.$$

Using indicator random variables, $|\mathcal{A}_l|$ can be written as

$$|\mathcal{A}_l| = \sum_{j=1}^{|\mathbb{A}^l|} \sum_{m_{S_j^l} \in \mathcal{M}_{S_j^l}} \mathbb{I}(m_{S_1^l}, \dots, m_{S_{|\mathbb{A}^l|}^l}) \quad (\text{A3a})$$

where

$$\mathbb{I}(m_{S_1^l}, \dots, m_{S_{|\mathbb{A}^l|}^l}) = \begin{cases} 1 & \text{if } \mathcal{Q}_l \text{ given } \mathcal{Q}_{l+1}, \\ 0 & \text{otherwise,} \end{cases} \quad (\text{A3b})$$

for each $(m_{S_1^l}, \dots, m_{S_{|\mathbb{A}^l|}^l}) \in \mathcal{M}_{\mathbb{A}^l}$.

For all $\mathcal{J}^l \subseteq \mathbb{A}^l$, define

$$\begin{aligned} p_{\mathcal{J}^l} = P\{ & (U^n(m_{S_1^l}) \dots U^n(m_{S_{|\mathbb{A}^l|}^l}), \mathbf{u}^n(\mathbb{A}^{l+1:K})) \in \mathcal{T}_{\epsilon_l}^{(n)}, \\ & (U^n(m'_{S_1^l}) \dots U^n(m'_{S_{|\mathbb{A}^l|}^l}), \mathbf{u}^n(\mathbb{A}^{l+1:K})) \in \mathcal{T}_{\epsilon_l}^{(n)}, \\ & \text{with } m_{S_j^l} = m'_{S_j^l} \text{ if } S_j^l \in \mathcal{J}^l \text{ and vice versa} | \mathbf{u}^n(\mathbb{A}^{l+1:K}) \in \mathcal{T}_{\epsilon_{l+1}}^{(n)} \}. \end{aligned} \quad (\text{A4})$$

From (A3) and definition of $p_{\mathcal{J}^l}$ in (A4), we have

$$\begin{aligned} \mathbb{E}(|\mathcal{A}_l|) &= 2^{n \sum_{S \in \mathbb{A}^l} r_S} \cdot p_{\mathbb{A}^l}, \\ \mathbb{E}(|\mathcal{A}_l|^2) &= \sum_{m_{S_1^l}, \dots, m_{S_{|\mathbb{A}^l|}^l}} \sum_{\substack{m'_{S_j^l} \neq m_{S_j^l} \\ \forall S \in \mathbb{A}^l \setminus \mathcal{J}^l}} p_{\mathcal{J}^l} \\ &\leq 2^{n \sum_{S \in \mathbb{A}^l} r_S} \cdot \sum_{\mathcal{J}^l \subseteq \mathbb{A}^l} 2^{n \sum_{S \in \mathbb{A}^l \setminus \mathcal{J}^l} r_S} \cdot p_{\mathcal{J}^l}. \end{aligned}$$

From definition of $p_{\mathcal{J}^l}$ and independence of codewords, we have $(p_{\mathcal{A}_l})^2 = p_{\emptyset}$. Therefore,

$$\text{Var}(|\mathcal{A}_l|) \leq 2^{n \sum_{S \in \mathbb{A}^l} r_S} \sum_{\mathcal{J}^l \subseteq \{\mathbb{A}^l \setminus \emptyset\}} 2^{n \sum_{S \in \mathbb{A}^l \setminus \mathcal{J}^l} r_S} \cdot p_{\mathcal{J}^l}.$$

Next, we compute the upper bound of $p_{\mathcal{J}^l}$.

$$\begin{aligned} p_{\mathcal{J}^l} &= P\{ (U^n(m_{S_1^l}) \dots U^n(m_{S_{|\mathbb{A}^l|}^l}), \mathbf{u}^n(\mathbb{A}^{l+1:K})) \in \mathcal{T}_{\epsilon_l}^{(n)}, \\ & (U^n(m'_{S_1^l}) \dots U^n(m'_{S_{|\mathbb{A}^l|}^l}), \mathbf{u}^n(\mathbb{A}^{l+1:K})) \in \mathcal{T}_{\epsilon_l}^{(n)}, \\ & \text{with } m_{S_j^l} = m'_{S_j^l} \text{ if } S_j^l \in \mathcal{J}^l \text{ and vice versa} | \mathcal{Q}_{l+1} \} \\ &= P\{ (U^n(m_{S_1^l}) \dots U^n(m_{S_{|\mathbb{A}^l|}^l}), \mathbf{u}^n(\mathbb{A}^{l+1:K})) \in \mathcal{T}_{\epsilon_l}^{(n)} | \mathcal{Q}_{l+1} \} \\ & \cdot P\{ (U^n(m'_{S_1^l}) \dots U^n(m'_{S_{|\mathbb{A}^l|}^l}), \mathbf{u}^n(\mathbb{A}^{l+1:K})) \in \mathcal{T}_{\epsilon_l}^{(n)} \\ & | (U^n(m_{S_1^l}) \dots U^n(m_{S_{|\mathbb{A}^l|}^l}), \mathbf{u}^n(\mathbb{A}^{l+1:K})) \in \mathcal{T}_{\epsilon_l}^{(n)}, \mathcal{Q}_{l+1} \} \\ &\stackrel{(a)}{\leq} 2^{n(H(U(\mathbb{A}^l) | U(\mathbb{A}^{l+1:K})) + \delta(\epsilon_l))} \cdot 2^{-n(\sum_{S \in \mathbb{A}^l} H(U_S | U(\mathbb{A}_S^{1:K} \setminus S)) - \delta(\epsilon_l))} \\ & \cdot 2^{n(H(U(\mathbb{A}^l \setminus \mathcal{J}^l) | U(\mathcal{J}^l), U(\mathbb{A}^{l+1:K})) + \delta(\epsilon_l))} \cdot 2^{-n(\sum_{S \in \mathbb{A}^l \setminus \mathcal{J}^l} H(U_S | U(\mathbb{A}_S^{1:K} \setminus S)) - \delta(\epsilon_l))} \end{aligned} \quad (\text{A5})$$

where (a) follows by properties of joint typicality and the mutually conditional independence property, i.e., given $\mathbf{u}^n(\mathbb{A}^{(l+1):K})$, $(U_S^n(m'_S) : S \in \mathbb{A}^l \setminus \mathcal{J}^l)$ and $(U_S^n(m_S) : S \in \mathbb{A}^l \setminus \mathcal{J}^l)$ are conditionally independent. The lower bound of $p_{\mathbb{A}^l}$ can be derived as

$$\begin{aligned} p_{\mathbb{A}^l} &= P\{ (U^n(m_{S_1^l}) \dots U^n(m_{S_{|\mathbb{A}^l|}^l}), \mathbf{u}^n(\mathbb{A}^{l+1:K})) \in \mathcal{T}_{\epsilon_l}^{(n)}, \mathbf{u}^n(\mathbb{A}^{l+1:K}) \in \mathcal{T}_{\epsilon_{l+1}}^{(n)} \}, \\ &\geq (1 - \epsilon_l) \cdot 2^{n(H(U(\mathbb{A}^l)|U(\mathbb{A}^{l+1:K})) - \delta(\epsilon_l))} \cdot 2^{-n(\sum_{S \in \mathbb{A}^l} H(U_S|U(\mathbb{A}_S^{1:K} \setminus S)) + \delta(\epsilon_l))}. \end{aligned} \quad (\text{A6})$$

where the last equality holds by the properties of jointly typicality. From (A5) and (A6), we have

$$\begin{aligned} \frac{p_{\mathcal{J}^l}}{(p_{\mathbb{A}^l})^2} &\leq 2^{n(H(U(\mathbb{A}^l)|U(\mathbb{A}^{l+1:K})) + \delta(\epsilon_l))} \cdot 2^{-n(\sum_{S \in \mathbb{A}^l} H(U_S|U(\mathbb{A}_S^{1:K} \setminus S)) - \delta(\epsilon_l))} \\ &\quad \cdot 2^{n(H(U(\mathbb{A}^l \setminus \mathcal{J}^l)|U(\mathcal{J}^l), U(\mathbb{A}^{(l+1):K})) + \delta(\epsilon_l))} \cdot (1 - \epsilon_l)^{-2} \cdot 2^{2n(H(U(\mathbb{A}^l)|U(\mathbb{A}^{l+1:K})) + \delta(\epsilon_l))} \\ &\quad \cdot 2^{-2n(\sum_{S \in \mathbb{A}^l} H(U_S|U(\mathbb{A}_S^{1:K} \setminus S)) + \delta(\epsilon_l))} \\ &\leq (1 - \epsilon_l)^{-2} \cdot 2^{-nH(U(\mathbb{A}^l)|U(\mathbb{A}^{(l+1):K}))} \cdot 2^{n(\sum_{S \in \mathbb{A}^l} H(U_S|U(\mathbb{A}_S^{1:K} \setminus S)))} \\ &\quad \cdot 2^{n(H(U(\mathbb{A}^l \setminus \mathcal{J}^l)|U(\mathcal{J}^l), U(\mathbb{A}^{(l+1):K}))} \cdot 2^{-n\sum_{S \in \mathbb{A}^l \setminus \mathcal{J}^l} H(U_S|U(\mathbb{A}_S^{1:K} \setminus S))} \cdot 2^{n\delta(\epsilon)}. \end{aligned}$$

Therefore, we finally obtain the desired inequality

$$\begin{aligned} \frac{\text{Var}(|\mathcal{A}_l|)}{(\mathbb{E}(|\mathcal{A}_l|))^2} &\leq (1 - \epsilon)^{-2} \cdot 2^{-n\sum_{S \in \mathbb{A}^l} r_S} \cdot \sum_{\mathcal{J}^l \subseteq \mathbb{A}^l \setminus \emptyset} \left(2^{n\sum_{S \in \mathbb{A}^l \setminus \mathcal{J}^l} r_S} \cdot \frac{p_{\mathcal{J}^l}}{(p_{\mathbb{A}^l})^2} \right) \\ &\leq (1 - \epsilon)^{-2} \sum_{\mathcal{J}^l \subseteq \mathbb{A}^l \setminus \emptyset} \left(2^{-n\sum_{S \in \mathcal{J}^l} r_S} \cdot 2^{n\delta(\epsilon)} \right. \\ &\quad \cdot 2^{nH(U(\mathbb{A}^l \setminus \mathcal{J}^l)|U(\mathcal{J}^l), U(\mathbb{A}^{(l+1):K})) - H(U(\mathbb{A}^l)|U(\mathbb{A}^{(l+1):K}))} \\ &\quad \left. \cdot 2^{n(\sum_{S \in \mathbb{A}^l} H(U_S|U(\mathbb{A}_S^{1:K} \setminus S)) - \sum_{S \in \mathbb{A}^l \setminus \mathcal{J}^l} H(U_S|U(\mathbb{A}_S^{1:K} \setminus S)))} \right), \end{aligned}$$

which tends to zero as $n \rightarrow \infty$ if for all $\mathcal{J}^l \subseteq \mathbb{A}^l$

$$\begin{aligned} \sum_{S \in \mathcal{J}^l} r_S &> H(U(\mathbb{A}^l \setminus \mathcal{J}^l) | U(\mathcal{J}^l), U(\mathbb{A}^{(l+1):K})) - H(U(\mathbb{A}^l) | U(\mathbb{A}^{(l+1):K})) \\ &\quad + \sum_{S \in \mathbb{A}^l} H(U_S | U(\mathbb{A}_S^{1:K} \setminus S)) - \sum_{S \in \mathbb{A}^l \setminus \mathcal{J}^l} H(U_S | U(\mathbb{A}_S^{1:K} \setminus S)) + \delta(\epsilon) \\ &= H(U(\mathbb{A}^l \setminus \mathcal{J}^l) | U(\mathcal{J}^l), U(\mathbb{A}^{(l+1):K})) - H(U(\mathbb{A}^l) | U(\mathbb{A}^{(l+1):K})) \\ &\quad + \sum_{S \in \mathcal{J}^l} H(U_S | U(\mathbb{A}_S^{1:K} \setminus S)) + \delta(\epsilon) \\ &= \sum_{S \in \mathcal{J}^l} H(U_S | U(\mathbb{A}_S^{1:K} \setminus S)) - H(U(\mathcal{J}^l) | U(\mathbb{A}^{(l+1):K})) + \delta(\epsilon). \end{aligned} \quad (\text{A7})$$

Therefore, from (A2), $P\{\mathcal{E}_1\}$ tends to zero as $n \rightarrow \infty$ if conditions (A7) hold for all $l \in [1 : (K - 1)]$.

References

1. Cover, T. Broadcast Channels. *IEEE Trans. Inf. Theory* **1972**, *18*, 2–14. [\[CrossRef\]](#)
2. El Gamal, A.; Kim, Y.H. *Network Information Theory*; Cambridge University Press: Cambridge, UK, 2011.
3. Cover, T.M. Comments on Broadcast Channels. *IEEE Trans. Inf. Theory* **1998**, *44*, 2524–2530. [\[CrossRef\]](#)
4. Marton, K. A Coding Theorem for the Discrete Memoryless Broadcast Channel. *IEEE Trans. Inf. Theory* **1979**, *25*, 306–311. [\[CrossRef\]](#)
5. Korner, J.; Marton, K. General Broadcast Channels with Degraded Message Sets. *IEEE Trans. Inf. Theory* **1977**, *23*, 60–64. [\[CrossRef\]](#)
6. Anantharam, V.; Gohari, A.; Nair, C. On the Evaluation of Marton's Inner Bound for Two-Receiver Broadcast Channels. *IEEE Trans. Inf. Theory* **2019**, *65*, 1361–1371. [\[CrossRef\]](#)

7. Gohari, A.; Nair, C. New Outer Bounds for the Two-Receiver Broadcast Channel. In Proceedings of the 2020 IEEE International Symposium on Information Theory (ISIT), Los Angeles, CA, USA, 21–26 June 2020; pp. 1492–1497.
8. Trillingsgaard, K.F.; Yang, W.; Durisi, G.; Popovski, P. Common-Message Broadcast Channels With Feedback in the Nonasymptotic Regime: Full Feedback. *IEEE Trans. Inf. Theory* **2018**, *64*, 7719–7741. [[CrossRef](#)]
9. Trillingsgaard, K.F.; Yang, W.; Durisi, G.; Popovski, P. Common-Message Broadcast Channels With Feedback in the Nonasymptotic Regime: Stop Feedback. *IEEE Trans. Inf. Theory* **2018**, *64*, 7686–7718. [[CrossRef](#)]
10. Trillingsgaard, K.F.; Yang, W.; Durisi, G.; Popovski, P. Broadcasting a Common Message with Variable-Length Stop-Feedback Codes. In Proceedings of the 2015 IEEE International Symposium on Information Theory (ISIT), Hong Kong, China, 14–19 June 2015; pp. 2505–2509.
11. Weingarten, H.; Steinberg, Y.; Shamai, S.S. The Capacity Region of the Gaussian Multiple-Input Multiple-Output Broadcast Channel. *IEEE Trans. Inf. Theory* **2006**, *52*, 3936–3964. [[CrossRef](#)]
12. Tse, D.N. Optimal Power Allocation over Parallel Gaussian Broadcast Channels. In Proceedings of the IEEE International Symposium on Information Theory, Ulm, Germany, 29 June–4 July 1997; p. 27.
13. Li, L.; Goldsmith, A.J. Capacity and Optimal Resource Allocation for Fading Broadcast Channels. I. Ergodic Capacity. *IEEE Trans. Inf. Theory* **2001**, *47*, 1083–1102.
14. Li, L.; Goldsmith, A.J. Capacity and Optimal Resource Allocation for Fading Broadcast Channels. II. Outage Capacity. *IEEE Trans. Inf. Theory* **2001**, *47*, 1103–1127.
15. Viswanath, P.; Tse, D.N.C. Sum Capacity of the Multiple Antenna Gaussian Broadcast Channel. In Proceedings of the IEEE International Symposium on Information Theory, Lausanne, Switzerland, 30 June–5 July 2002; p. 497.
16. Yu, W.; Cioffi, J.M. Sum Capacity of Gaussian Vector Broadcast Channels. *IEEE Trans. Inf. Theory* **2004**, *50*, 1875–1892. [[CrossRef](#)]
17. Vishwanath, S.; Jindal, N.; Goldsmith, A. Duality, Achievable Rates, and Sum-Rate Capacity of Gaussian MIMO Broadcast Channels. *IEEE Trans. Inf. Theory* **2003**, *49*, 2658–2668. [[CrossRef](#)]
18. Caire, G.; Shamai, S. On the Achievable Throughput of a Multiantenna Gaussian Broadcast Channel. *IEEE Trans. Inf. Theory* **2003**, *49*, 1691–1706. [[CrossRef](#)]
19. Salman, M.; Varanasi, M.K. Capacity Results via Message Merging and Superposition Coding in the K-Receiver Broadcast Channel with General Message Sets. In Proceedings of the 2019 IEEE International Symposium on Information Theory (ISIT), Paris, France, 7–12 July 2019; pp. 1642–1646.
20. Romero, H.; Varanasi, M. Superposition Coding in the Combination Network. In Proceedings of the 2016 IEEE International Symposium on Information Theory (ISIT), Barcelona, Spain, 10–15 July 2016; pp. 2149–2153.
21. Romero, H.P.; Varanasi, M.K. Rate Splitting and Superposition Coding for Concurrent Groupcasting over the Broadcast Channel: A General Framework. In Proceedings of the 2017 IEEE International Symposium on Information Theory (ISIT), Aachen, Germany, 25–30 June 2017; pp. 1888–1892.
22. Ganguly, S.; Wang, L. Sliding-Window Gelfand–Pinsker Coding: General K-User Broadcast Channels. In Proceedings of the 2020 IEEE Information Theory Workshop (ITW), Riva del Garda, Italy, 11–15 April 2021; pp. 1–5.
23. Nair, C.; Wang, Z.V. On 3-Receiver Broadcast Channels with 2-Degraded Message Sets. In Proceedings of the 2009 IEEE International Symposium on Information Theory, Seoul, Korea, 28 June–3 July 2009; pp. 1844–1848.
24. Nair, C.; El Gamal, A. The Capacity Region of a Class of 3-Receiver Broadcast Channels with Degraded Message Sets. In Proceedings of the 2008 IEEE International Symposium on Information Theory, St. Petersburg, Russia, 31 July–5 August 2008; pp. 1706–1710.
25. Nair, C.; Wang, Z.V. The Capacity Region of the Three Receiver Less Noisy Broadcast Channel. *IEEE Trans. Inform. Theory* **2011**, *57*, 4058–4062. [[CrossRef](#)]
26. Nair, C.; Xia, L. On Three-Receiver More Capable Channels. In Proceedings of the 2012 IEEE International Symposium on Information Theory Proceedings, Cambridge, MA, USA, 1–6 July 2012; pp. 378–382.
27. Romero, H.P. Fundamental Limits of Network Communication with General Message Sets: A Combinatorial Approach. Ph.D. Thesis, University of Colorado at Boulder, Boulder, CO, USA, 2014.
28. Davey, B.A.; Priestley, H.A. *Introduction to Lattices and Order*; Cambridge University Press: Cambridge, UK, 2002.