Diamond Message Set Groupcasting: From an Inner Bound for the DM Broadcast Channel to the Capacity Region of the Combination Network

Mohamed Salman and Mahesh K. Varanasi, *Fellow, IEEE*

***Abstract*— Multiple groupcasting over the broadcast channel (BC) is studied in a special setting. In particular, an inner bound is obtained for the K-receiver discrete memoryless (DM) BC for the diamond message set which consists of four groupcast messages: one desired by all receivers, one by all but two receivers, and two more desired by all but each one of those two receivers. The inner bound is based on rate-splitting and superposition coding and is given in explicit form herein as a union over coding distributions of four-dimensional polytopes. This inner bound is then shown to be the capacity for a certain class of partially ordered DM BCs with order defined via the less noisy condition. When specialized to the so-called combination network, which is a class of three-layer (two-hop) broadcast networks parameterized by 2^K-1 finite-and-arbitrary-capacity noiseless links from the source node in the first layer to as many nodes of the second layer, our top-down approach from the DM BC to the combination network yields an explicit inner bound as a single polytope via the identification of a single coding distribution. This inner bound consists of inequalities which are then identified to be within the class of generalized cut-set outer bounds recently obtained by Salimi et al for broadcast networks. We hence establish the capacity region of the general K-user combination network for the diamond message set, and do so in explicit and structured form. Such a result implies a certain strength of our inner bound for the DM BC in that it (a) produces a hitherto unknown capacity region when specialized to the combination network and (b) may capture many combinatorial aspects of the capacity region of the K-receiver DM BC itself for the diamond message set. Moreover, we extend that inner bound by adding binning to it. As in the no-binning case, we provide the more general inner bound in explicit form.**

***Index Terms*— Broadcast Channel (BC), capacity, combination network, groupcast.**

I. Introduction

A GENERAL order-theoretic framework for groupcasting over the K-receiver DM BC was proposed in [2, Theorem 1], that allows for a succinct, albeit indirect description of the rate region achieved by *up-set* rate-splitting and superposition coding for simultaneously sending the complete set of 2^K-1 independent messages, each desired by some distinct subset of receivers. In up-set rate-splitting, a message intended for some subset of receivers is split into sub-messages, with each sub-message to be delivered to some distinct subset of receivers for all possible such subsets that include the originally intended set of receivers. The set of sub-messages that are intended for the same group of receivers are then collected to form a new reconstructed message. Among the numerous choices [3], the type of superposition coding used for generating the codebooks for the reconstructed messages is the one with the superposition order taken to the subset inclusion order [2]. Finally, each receiver jointly decodes its desired reconstructed messages which contain the desired messages as well as the partial interference contained in the undesired sub-messages assigned to it via rate-splitting.

The inner bound in [2, Theorem 1] is given in implicit form in that it gives an indirect description of per-distribution polytopes in terms of the split rates (in higher dimension) rather than the original message rates. In principle, the split rates can be projected away with Fourier Motzkin Elimination (FME) [4], but in practice, this is only possible for small settings, i.e., small number of receivers and/or small message sets since projecting away all split rates in this general setting is intractable in general.

In a different line of work on network coding, certain three-layered broadcast (single-source) networks named combination networks were introduced in [5] to demonstrate that network coding for a single multicast session can attain unbounded gain over routing alone. A combination network is defined more generally in [6] to be a class of three-layer (two-hop) broadcast networks (as depicted in Fig. 2 for $K=3$) parameterized by 2^K-1 finite-and-arbitrary-capacity noiseless links from the source node in the first layer to as many nodes of the second layer, with each node in the second layer connected to a distinct subset of K destination nodes via infinite capacity links. For the complete message set, a linear network coding scheme and a matching converse was provided to obtain the capacity region of the three-receiver combination network in [7] (see also [6]). Moreover, the *symmetric* capacity region—in which the rates of the messages desired by the same number of users are equal—of the *symmetric*

Manuscript received 26 May 2020; revised 19 November 2021; accepted 7 January 2022. Date of publication 21 July 2022; date of current version 22 December 2022. This work was supported in part by the 2018 and 2019 Qualcomm Faculty Awards. This paper was presented in part at the 2020 IEEE International Symposium on Information Theory [1]. *(Corresponding author: Mahesh K. Varanasi.)*

Mohamed Salman was with the Electrical, Computer and Energy Engineering Department, University of Colorado Boulder, Boulder, CO 80302 USA. He is now with Qualcomm Inc., Boulder, CO 80301 USA (e-mail: mohamed.salman@colorado.edu).

Mahesh K. Varanasi is with the Electrical, Computer and Energy Engineering Department, University of Colorado Boulder, Boulder, CO 80309-0425 USA (e-mail: varanasi@colorado.edu).

Communicated by M. Costa, Associate Editor for Shannon Theory.

Digital Object Identifier 10.1109/TIT.2022.3193274

combination network[1] was found in indirect and explicit forms in [8] and [6], respectively. The general capacity region of the general combination network for a complete message set (with 2^K-1 messages) for $K>3$ remains an open problem. In this regard, the work of [9] must be mentioned for having produced partial results: in particular, for two nested messages, i.e., a multicast message intended for all receivers and a private message intended for a subset of the receivers, the capacity of the K-user general combination network is established in [9] when the number of receivers that demand only the multicast message is no more than three.

The general combination network can be seen as a special class of deterministic DM BCs as observed in [2]. A top-down approach was initiated therein to study the combination network. In particular, a portion (i.e., an inner bound) of the general inner bound of [2, Theorem 1] based on up-set rate-splitting and superposition coding for the DM BC for a complete message set is proposed as an achievable rate region for the combination network in [2, Theorem 2] through the specification of a single random coding distribution. Its indirect description notwithstanding, it was shown in [2], via its alternative description in [2, Theorem 2], that that achievable rate region recovers the capacity (for $K = 3$) and symmetric capacity (for general K) for symmetric combination networks, results previously obtained via linear network coding in [6]–[8]. Hence, the work in [2] provides a certain validation of the strength of the rate-splitting and superposition coding inner bound of [2, Theorem 1] for the K-receiver DM BC itself for the complete message set.

We continue that study for general (not only symmetric) combination networks for any $K > 3$, but by focusing on the diamond message set. That message set consists of four messages, one desired by all K receivers, one by all but two receivers, and two more desired by all but each one of those two receivers.[2] In particular, we show that adopting the general framework of [2, Theorem 1] even in this incomplete diamond message set case[3] is sufficient to produce the hitherto unknown capacity region of the combination network. The capacity result for the combination network in the special case of three degraded messages, obtained by setting the rate of one of the messages desired by all but one receiver to zero, was also previously unknown. Moreover, special instances of the diamond message set capacity result for various two message set cases, by setting the rates of the other two messages to zero, recover two-message capacity results for the combination network from [9] and [11].

In particular, we present an inner bound for K-receiver DM BC in terms of the actual rates of the four messages instead of the split rates as in [2, Theorem 1] by projecting away the split rates via Fourier-Motzkin elimination [12] as described in Appendix II, in spite of the indeterminate number of inequalities that describe the per-distribution achievable rate region in original- and split-rate space. We then specialize this inner bound first to a class of partially ordered DM BCs (ordered via the less-noisy condition) and show that it is the capacity region for that class of DM BCs.

We then specialize the inner bound for the general DM BC to the combination network by choosing a single distribution for the auxiliary and input random variables to obtain an explicit polyhedral inner bound for the general combination network. We establish the capacity result by identifying the inequalities present in the aforementioned explicit inner bound to be within the class of generalized cut set bounds proposed in [6] as outer bounds for broadcast networks. Notably, our order-theoretic description of the inner bound enables that identification by permitting a description of the inequalities in the inner bound that is similar to the form of the generalized cut-set bounds of [6].

Beyond establishing the capacity of the general K-user combination network for the diamond message set, our top-down approach suggests that the inner bound of [2, Theorem 1] given in a more explicit form here is a good one for the much more general DM BC (with the diamond message set) in that it specializes to the capacity region in the combination network without symmetry assumptions and possibly captures many combinatorial aspects of the capacity region of the DM BC itself. Nevertheless, we extend that inner bound by adding binning to rate-splitting and superposition coding and provide an expression for the resulting inner bound also in explicit form.

The rest of this paper is organized as follows. In Section II, we present the order theory notation and describe the system models for the DM BC and the combination network. In Section III, we present the inner bound for the DM BC and show that it is capacity achieving for (a) a class of partially ordered DM BCs and (b) for the general K-user combination network along with proofs of converses, establishing the capacity regions for the two cases. In Section IV, we extend the inner bound for the DM BC of Section III based on rate-splitting and superposition coding by adding binning to it. While binning is not needed to achieve the capacity of the combination network it is an interesting open question as to whether the more general inner bound of Section IV would be optimal for BCs that are more general than the combination network such as the deterministic broadcast channel. Finally, the paper is concluded in Section V. Proofs of DM BC inner bounds are given in Appendices I, II, and IV, and the proof of the capacity region for the partially ordered DM BCs is given in Appendix III.

II. Notation and System Model

A. Order Theory

We find it convenient to introduce ideas from order theory following the notation in [2] to enable the descriptions of the system model and the results. We consider the ground

[1]A symmetric combination network is one in which the capacity of links to nodes in the second layer that are connected to the same number of receivers are identical.

[2]The name diamond message set is derived from the depiction of the associated ordered set of four message indices, each message index being a subset of the set of all receiver indices at which that message is desired, with the order taken to be the subset inclusion order, in the form of a Hasse diagram, which would hence be diamond-shaped (see Fig. 1).

[3]A more general framework that builds on that of [2] is given in [10] to account for incomplete message sets in general, but that extra generality turns out to be unnecessary for the sake of discovering the capacity of the combination network for the diamond message set, as demonstrated here.

set to be an ordered set of subsets of receiver indices in $\{1, 2, \cdots, K\} \triangleq [1:K]$. We assume the order to be that of set inclusion, i.e., $S \leq S'$ if and only if $S \subseteq S'$ while S and S' are incomparable if neither $S \subseteq S'$ nor $S' \subseteq S$. Let P be such an ordered set of sets and Q be a subset of P. Note that the sans serif letter is used to represent a set of sets like P, Q and E to distinguish it from sets. We say that Q is an *up-set* if $S \in \mathsf{Q}$, $S' \in \mathsf{P}$, and $S' \geq S$ implies $S' \in \mathsf{Q}$ and a *down-set* if $S \in \mathsf{Q}$, $S' \in \mathsf{P}$, and $S' \leq S$ implies $S' \in \mathsf{Q}$.

Moreover, for any subset $\mathsf{Q} \subseteq \mathsf{P}$, we define the smallest *down-set* containing Q as $\downarrow_{\mathsf{P}} \mathsf{Q} = \{S' \in \mathsf{P} : S' \leq S, S \in \mathsf{Q}\}$ and the smallest *up-set* containing Q as $\uparrow_{\mathsf{P}} \mathsf{Q} = \{S' \in \mathsf{P} : S \leq S', S \in \mathsf{Q}\}$. For the sake of simplicity, we abbreviate the set $\{i_1, i_2, .., i_N\} \subseteq [1:K]$ for any positive number $N \leq K$ as $i_1 i_2 \cdots i_N$, adopting the convention that $i_1 < i_2 < \cdots < i_N$. For instance, let P be the set of non-empty subsets of $[1:3]$, i.e., $\mathsf{P} = \{1, 2, 3, 12, 13, 23, 123\}$. Then, for instance, we have $\downarrow_{\mathsf{P}} \{13\} = \{1, 3, 13\}$, $\downarrow_{\mathsf{P}} \{1, 3\} = \{1, 3\}$, $\uparrow_{\mathsf{P}} \{13\} = \{13, 123\}$, and $\uparrow_{\mathsf{P}} \{1, 3\} = \{1, 3, 12, 13, 23, 123\}$. In some cases, especially when the set $S = \{i_1, i_2, .., i_N\}$ has many elements, we find it more convenient to denote it by its complement, i.e., $\overline{S} = [1:K] \backslash S$. For example, the set $\overline{\{i_1\}}$, also denoted $\overline{i_1}$, is the set $[1:K] \backslash \{i_1\}$.

Finally, for any $\mathsf{F} \subseteq \mathsf{P}$ and $i \in [1:K]$, we define $\mathsf{W}_i^{\mathsf{F}}$ as the set of sets in F containing i so that $\mathsf{W}_i^{\mathsf{F}} \triangleq \{S \in \mathsf{F} : i \in S\}$.

From the given notation, we can show that the following relationships are true:

1) For any set $S = i_1 i_2 \cdots i_N \subseteq [1:K]$, we have

$$\cup_{k \in S} \mathsf{W}_k^{\mathsf{P}} = \uparrow_{\mathsf{P}} \{i_1, i_2, \cdots, i_N\} \tag{1}$$

$$\cap_{k \in S} \mathsf{W}_k^{\mathsf{P}} = \uparrow_{\mathsf{P}} \{i_1 i_2 \cdots i_N\} \tag{2}$$

2) For any set $S = i_1 i_2 \cdots i_N \subset [1:K]$ and any $i \in [1:K]$

$$\downarrow_{\mathsf{W}_i^{\mathsf{P}}} \{\overline{i_1}, \overline{i_2}, \cdots, \overline{i_N}\} \cup \uparrow_{\mathsf{W}_i^{\mathsf{P}}} \{S\} = \mathsf{W}_i^{\mathsf{P}} \tag{3}$$

$$\downarrow_{\mathsf{W}_i^{\mathsf{P}}} \{\overline{i_1}, \overline{i_2}, \cdots, \overline{i_N}\} \cap \uparrow_{\mathsf{W}_i^{\mathsf{P}}} \{S\} = \phi \tag{4}$$

$$\downarrow_{\mathsf{W}_i^{\mathsf{P}}} \{\overline{S}\} \cup \uparrow_{\mathsf{W}_i^{\mathsf{P}}} \{i_1, i_2, \cdots, i_N\} = \mathsf{W}_i^{\mathsf{P}} \tag{5}$$

$$\downarrow_{\mathsf{W}_i^{\mathsf{P}}} \{\overline{S}\} \cap \uparrow_{\mathsf{W}_i^{\mathsf{P}}} \{i_1, i_2, \cdots, i_N\} = \phi \tag{6}$$

B. System Model

We consider a DM BC with the transmitter denoted by the transmitted symbol $X \in \mathcal{X}$, K receivers denoted by their respective channel outputs $Y_i \in \mathcal{Y}_i$, and the channel transition probability $W(y_1, y_2, \cdots, y_K | x)$ where the conditional probability of n channel outputs $Y_1^n, \cdots, Y_K^n$ with $Y_k^n \triangleq (Y_{k,1}, \cdots, Y_{k,n})$, conditioned on n channel inputs $X^n \triangleq (X_1, \cdots, X_n)$ is given (in the absence of feedback) by

$$p(y_1^n, \cdots, y_K^n | x^n) = \prod_{j=1}^{n} W(y_{1j}, \cdots, y_{Kj} | x_j)$$

where $X_j, Y_{1,j}, \cdots Y_{K,j}$ are the channel input and outputs in the j^{th} channel use.

The message $M_S \in [1 : 2^{nR_S}]$ of rate R_S is indexed by the subset $S \subseteq [1:K]$ of receivers it is intended for. Hence, $M_{\overline{S}}$ is the message intended for all receivers except the receivers in S. The diamond message set consists of four messages, $M_{\overline{\phi}}$, a message intended for all receivers, $M_{\overline{K-1.K}}$ a message intended for the first $K-2$ receivers[4], and $M_{\overline{K-1}}$ and $M_{\overline{K}}$, two messages intended for all but receivers $K-1$ or K. Define E as the set of all message indices so that $\mathsf{E} = \{\overline{\phi}, \overline{K}, \overline{K-1}, \overline{K-1.K}\}$. Note that receiver Y_K demands $M_{\overline{\phi}}$ and $M_{\overline{K-1}}$, receiver Y_{K-1} demands $M_{\overline{\phi}}$ and $M_{\overline{K}}$, while the rest of receivers $\{Y_i\}_{i=1}^{K-2}$ demand all four messages. See Fig. 1 for an illustration of diamond message sets for $K = 3, 4, 5$ and general K via their diamond-shaped Hasse diagrams.

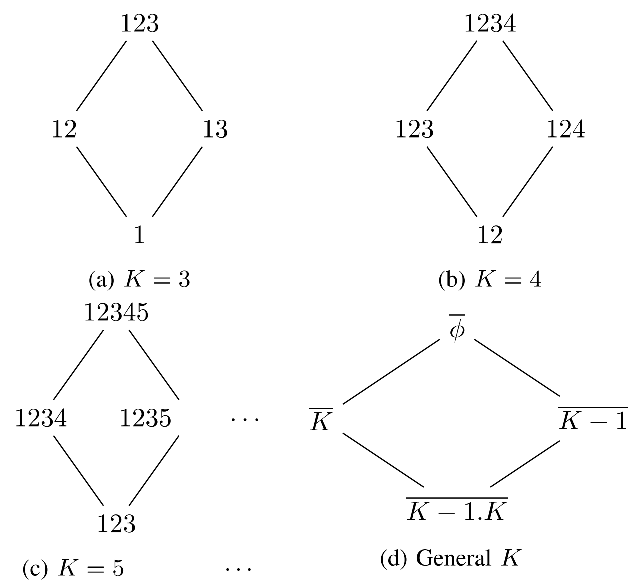

Fig. 1. Hasse diagrams for the four sets of message indices in the diamond message sets for $K = 3, 4$ and 5 and general K. For instance, for $K = 5$, there is a message intended for receivers 1, 2, and 3, a message for receivers 1, 2, 3 and 4, a message for receivers 1, 2, 3 and 5 and one for all five receivers.

Denote the set of the four messages $\{M_S : S \in \mathsf{E}\}$ to be sent over a K-user DM BC as M_{E}. A $(\{2^{nR_S}\}_{S \in \mathsf{E}}, n)$ code consists of (i) an encoder that assigns to each message tuple $m_{\mathsf{E}} \in \prod_{S \in \mathsf{E}} [1 : 2^{nR_S}]$ a codeword $x^n(m_{\mathsf{E}})$ (ii) a decoder at each receiver, with the k^{th} decoder mapping the received sequence Y_k^n for each $k \in [1:K]$ into the set of decoded messages $\{\tilde{m}_S : S \in \mathsf{W}_k^{\mathsf{E}}\} \in \prod_{S \in \mathsf{W}_k^{\mathsf{E}}} [1 : 2^{nR_S}]$, denoted as $\tilde{m}_{\mathsf{W}_k^{\mathsf{E}}}$. It is assumed that all four messages are independent and uniformly distributed over their ranges. The probability of error $P_e^{(n)}$ is the probability that not all receivers decode their intended messages correctly. The rate tuple $(R_S : S \in \mathsf{E})$ is said to be achievable if there exists a sequence of $(\{2^{nR_S}\}_{S \in \mathsf{E}}, n)$ codes with $P_e^{(n)} \to 0$ as $n \to \infty$. The convex closure of the union of achievable rates is the capacity region.

Next, we define a restriction of the DM BC by constraining pairs of receivers via the well-known *less noisy* ordering.

Definition 1: [13, Definition 2] Receiver Y is less noisy than Receiver Z if $I(U;Y) \geq I(U;Z)$ for all $U \multimap X \multimap (Y, Z)$ forming a Markov chain. Henceforth, we denote this condition as $Y \succ Z$.

[4]The dot separating $K-1$ and K in $M_{\overline{K-1.K}}$ is a slight abuse of notation since the subscript in $M_{\overline{K-1K}}$ can be confusing.

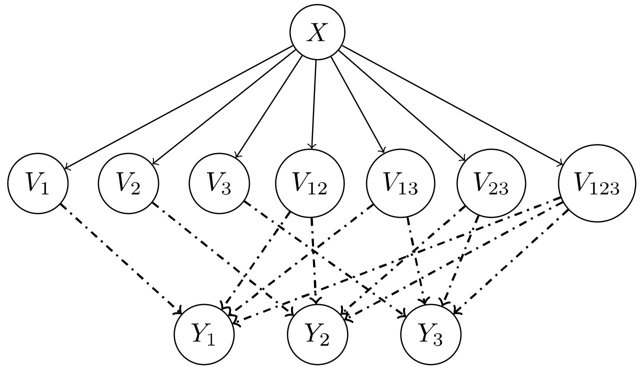

Fig. 2. The three-receiver combination network with 7 intermediate nodes and three receivers. The dark/dashed lines represent finite/infinite capacity links, respectively. The capacity of the dark link connecting the node X to the node V_S is C_S for each $S \in \mathsf{P}$. For brevity, the source/destination nodes are denoted by their transmitted/received symbols and the intermediate nodes by their output symbols. Encoders/decoders are not shown.

The K-receiver DM BC $(Y_1, Y_2, \cdots, Y_K)$ is said to be a less noisy DM BC if it is totally ordered per less noisy ordering, i.e., when $Y_1 \succ Y_2, \cdots \succ Y_{K-1} \succ Y_K$. In this paper, we will be interested in a certain larger class of partially-ordered DM BCs.

The combination network [2], [6] is a special case of the general DM-BC. As shown in Fig. 2 for the three-receiver case, it consists of three layers of nodes. The top layer and bottom layer consist of the single source node X and K receivers $\{Y_i\}_{i=1}^K$, respectively. While, the middle layer consists of $2^K - 1$ intermediate nodes, denoted V_S for all $S \in \mathsf{P}$ where P be the power set of $[1:K]$ excluding the empty set. The source is connected to each of the intermediate nodes V_S through a noiseless link of capacity C_S (per channel use). On the other hand, each receiver Y_i is connected to 2^{K-1} intermediate nodes $V_{\mathsf{W}_i^{\mathsf{P}}} \triangleq \{V_S\}_{S \in \mathsf{W}_i^{\mathsf{P}}}$ via noiseless links of unlimited capacity. An interesting connection between the combination networks and the DM BC is revealed in [2] wherein the authors considered the combination network to be a network of noiseless DM BCs with the channel input vector X connected in different ways to the channel outputs $\{Y_i\}_{i=1}^K$ each through a noiseless BC. In particular, the channel input X contains $2^K - 1$ components V_S, for all $S \in \mathsf{P}$. For each S, the component $V_S \in \mathcal{V}_S$, where $|\mathcal{V}_S| = 2^{C_S}$, is noiselessly received at each receiver Y_i for all $i \in S$ and *not* received at the receivers Y_j with $j \notin S$, i.e., $Y_i = V_{\mathsf{W}_i^{\mathsf{P}}}$. We also define $C_{\mathsf{W}} = \sum_{S \in \mathsf{W}} C_S$ for any $\mathsf{W} \subseteq \mathsf{P}$. For instance, $C_{\mathsf{W}_i^{\mathsf{P}}} = \sum_{S \in \mathsf{W}_i^{\mathsf{P}}} C_S$.

III. RESULTS

In the following theorem, we present the inner bound of [2, Theorem 1] specialized to the diamond message set in the more explicit form of a union of four-dimensional polytopes.

Theorem 1: An inner bound of K-user DM BC for the diamond message set with $\mathsf{E} = \{\overline{\phi}, \overline{K}, \overline{K-1}, \overline{K-1.K}\}$ is the convex closure of the set of non-negative rate tuples $(R_{\overline{\phi}}, R_{\overline{K}}, R_{\overline{K-1}}, R_{\overline{K-1.K}})$ satisfying the following for all $j, j_1, j_2 \in \{1, 2, \cdots, K-2\}$

$$R_{\overline{\phi}} + R_{\overline{K-1}} \leq I(U_{\overline{\phi}}, U_{\overline{K-1}}; Y_K) \tag{7}$$

$$R_{\overline{\phi}} + R_{\overline{K}} \leq I(U_{\overline{\phi}}, U_{\overline{K}}; Y_{K-1}) \tag{8}$$

$$R_{\overline{\phi}} + R_{\overline{K-1}} + R_{\overline{K}} \leq I(U_{\overline{K}}; Y_{K-1} | U_{\overline{\phi}}) + I(U_{\overline{\phi}}, U_{\overline{K-1}}; Y_K) \tag{9}$$

$$R_{\overline{\phi}} + R_{\overline{K-1}} + R_{\overline{K}} \leq I(U_{\overline{K-1}}; Y_K | U_{\overline{\phi}}) + I(U_{\overline{\phi}}, U_{\overline{K}}; Y_{K-1}) \tag{10}$$

$$R_{\overline{\phi}} + R_{\overline{K-1}} + R_{\overline{K}} + R_{\overline{K-1.K}} \leq I(X; Y_j) \tag{11}$$

$$R_{\overline{\phi}} + R_{\overline{K-1}} + R_{\overline{K}} + R_{\overline{K-1.K}} \leq I(X; Y_j | U_{\overline{\phi}}, U_{\overline{K-1}}) + I(U_{\overline{\phi}}, U_{\overline{K-1}}; Y_K) \tag{12}$$

$$R_{\overline{\phi}} + R_{\overline{K-1}} + R_{\overline{K}} + R_{\overline{K-1.K}} \leq I(X; Y_j | U_{\overline{\phi}}, U_{\overline{K}}) + I(U_{\overline{\phi}}, U_{\overline{K}}; Y_{K-1}) \tag{13}$$

$$R_{\overline{\phi}} + R_{\overline{K-1}} + R_{\overline{K}} + R_{\overline{K-1.K}} \leq I(X; Y_j | U_{\overline{\phi}}, U_{\overline{K-1}}, U_{\overline{K}}) + I(U_{\overline{K}}; Y_{K-1} | U_{\overline{\phi}}) + I(U_{\overline{\phi}}, U_{\overline{K-1}}; Y_K) \tag{14}$$

$$R_{\overline{\phi}} + R_{\overline{K-1}} + R_{\overline{K}} + R_{\overline{K-1.K}} \leq I(X; Y_j | U_{\overline{\phi}}, U_{\overline{K-1}}, U_{\overline{K}}) + I(U_{\overline{K-1}}; Y_K | U_{\overline{\phi}}) + I(U_{\overline{\phi}}, U_{\overline{K}}; Y_{K-1}) \tag{15}$$

$$2R_{\overline{\phi}} + R_{\overline{K-1}} + R_{\overline{K}} + R_{\overline{K-1.K}} \leq I(X; Y_j | U_{\overline{\phi}}, U_{\overline{K-1}}, U_{\overline{K}}) + I(U_{\overline{\phi}}, U_{\overline{K}}; Y_{K-1}) + I(U_{\overline{\phi}}, U_{\overline{K-1}}; Y_K) \tag{16}$$

$$2R_{\overline{\phi}} + 2R_{\overline{K-1}} + 2R_{\overline{K}} + R_{\overline{K-1.K}} \leq I(X; Y_j | U_{\overline{\phi}}) + I(U_{\overline{\phi}}, U_{\overline{K}}; Y_{K-1}) + I(U_{\overline{\phi}}, U_{\overline{K-1}}; Y_K) \tag{17}$$

$$2R_{\overline{\phi}} + 2R_{\overline{K-1}} + 2R_{\overline{K}} + 2R_{\overline{K-1.K}} \leq I(X; Y_{j_1} | U_{\overline{\phi}}) + I(X; Y_{j_2} | U_{\overline{\phi}}, U_{\overline{K-1}}, U_{\overline{K}}) + I(U_{\overline{\phi}}, U_{\overline{K}}; Y_{K-1}) + I(U_{\overline{\phi}}, U_{\overline{K-1}}; Y_K) \tag{18}$$

for some joint distribution of the auxiliary and input random variables $(U_{\overline{\phi}}, U_{\overline{K}}, U_{\overline{K-1}}, X)$ and the output random variables that is of the form

$$p(u_{\overline{\phi}}, u_{\overline{K}}, u_{\overline{K-1}}, x, y_1, \cdots, y_K) = p(u_{\overline{\phi}}) \times p(u_{\overline{K}} | u_{\overline{\phi}}) \times p(u_{\overline{K-1}} | u_{\overline{\phi}}) \times p(x | u_{\overline{\phi}}, u_{\overline{K}}, u_{\overline{K-1}}) \times p(y_1, \cdots, y_K | x) \tag{19}$$

Proof: Random coding via rate-splitting and superposition coding is used. A detailed proof is given in Appendices I and II. ∎

A. A Capacity Result for a Class of Partially Ordered DM BCs

In this section, we show that the achievable rate region of Theorem 1 is the capacity region for a certain class of partially ordered DM BCs.

Theorem 2: The capacity region of the class of partially ordered DM BCs with $\mathsf{E} = \{\overline{\phi}, \overline{K}, \overline{K-1}, \overline{K-1.K}\}$ constrained by the $K-2$ less-noisy conditions

$$Y_i \succeq Y_{K-1} \succeq Y_K \qquad i \in [1:K-2]$$

is the convex closure of the set of non-negative rate tuples $(R_{\overline{\phi}}, R_{\overline{K}}, R_{\overline{K-1}}, R_{\overline{K-1.K}})$ satisfying

$$R_{\overline{\phi}} + R_{\overline{K-1}} \leq I(U_{\overline{\phi}}; Y_K) \tag{20}$$
$$R_{\overline{K}} \leq I(U_{\overline{K}}; Y_{K-1}|U_{\overline{\phi}}) \tag{21}$$
$$R_{\overline{K-1.K}} \leq I(X; Y_i|U_{\overline{K}}) \tag{22}$$

for some joint distribution of the auxiliary and input random variables $(U_{\overline{\phi}}, U_{\overline{K}}, X)$ of the form

$$p(u_{\overline{\phi}}, u_{\overline{K}}, x, y_1, \cdots, y_K) = p(u_{\overline{\phi}}) \times p(u_{\overline{K}}|u_{\overline{\phi}}) \times p(x|u_{\overline{K}}) \times p(y_1, \cdots, y_K|x) \tag{23}$$

Proof: Achievability is proved via the inner bound of Theorem 1. A key ingredient of the converse proof is the information inequality of [14, Lemma 1]. The detailed proof of achievability and the converse result are given in Appendix III. ■

B. Capacity of the Combination Network

We next state the inner bound of Theorem 1 when it is applied to the combination network by setting $X = V_{\mathsf{P}}$ and $Y_j = V_{\mathsf{W}^{\mathsf{P}}_j}$. The mutual information terms in the bounds are written as the difference between the entropy and conditional entropy so that we may get insights into optimizing the rate region over the coding distributions.

Corollary 1: An inner bound of K-user combination networks $\mathsf{E} = \{\overline{\phi}, \overline{K}, \overline{K-1}, \overline{K-1.K}\}$ is the convex closure of the set of non-negative rate tuples $(R_{\overline{\phi}}, R_{\overline{K}}, R_{\overline{K-1}}, R_{\overline{K-1.K}})$ satisfying the following for all $j, j_1, j_2 \in \{1, 2, \cdots, K-2\}$

$$R_{\overline{\phi}} + R_{\overline{K-1}} \leq H(V_{\mathsf{W}^{\mathsf{P}}_K}) - H(V_{\mathsf{W}^{\mathsf{P}}_K}|U_{\overline{\phi}}, U_{\overline{K-1}}) \tag{24}$$
$$R_{\overline{\phi}} + R_{\overline{K}} \leq H(V_{\mathsf{W}^{\mathsf{P}}_{K-1}}) - H(V_{\mathsf{W}^{\mathsf{P}}_{K-1}}|U_{\overline{\phi}}, U_{\overline{K}}) \tag{25}$$
$$R_{\overline{\phi}} + R_{\overline{K-1}} + R_{\overline{K}} \leq H(V_{\mathsf{W}^{\mathsf{P}}_{K-1}}|U_{\overline{\phi}}) - H(V_{\mathsf{W}^{\mathsf{P}}_{K-1}}|U_{\overline{\phi}}, U_{\overline{K}}) + H(V_{\mathsf{W}^{\mathsf{P}}_K}) - H(V_{\mathsf{W}^{\mathsf{P}}_K}|U_{\overline{\phi}}, U_{\overline{K-1}}) \tag{26}$$
$$R_{\overline{\phi}} + R_{\overline{K-1}} + R_{\overline{K}} \leq H(V_{\mathsf{W}^{\mathsf{P}}_K}|U_{\overline{\phi}}) - H(V_{\mathsf{W}^{\mathsf{P}}_K}|U_{\overline{\phi}}, U_{\overline{K-1}}) + H(V_{\mathsf{W}^{\mathsf{P}}_{K-1}}) - H(V_{\mathsf{W}^{\mathsf{P}}_{K-1}}|U_{\overline{\phi}}, U_{\overline{K}}) \tag{27}$$
$$R_{\overline{\phi}} + R_{\overline{K-1}} + R_{\overline{K}} + R_{\overline{K-1.K}} \leq H(V_{\mathsf{W}^{\mathsf{P}}_j}) \tag{28}$$
$$R_{\overline{\phi}} + R_{\overline{K-1}} + R_{\overline{K}} + R_{\overline{K-1.K}} \leq H(V_{\mathsf{W}^{\mathsf{P}}_j}|U_{\overline{\phi}}, U_{\overline{K-1}}) + H(V_{\mathsf{W}^{\mathsf{P}}_K}) - H(V_{\mathsf{W}^{\mathsf{P}}_K}|U_{\overline{\phi}}, U_{\overline{K-1}}) \tag{29}$$
$$R_{\overline{\phi}} + R_{\overline{K-1}} + R_{\overline{K}} + R_{\overline{K-1.K}} \leq H(V_{\mathsf{W}^{\mathsf{P}}_j}|U_{\overline{\phi}}, U_{\overline{K}}) + H(V_{\mathsf{W}^{\mathsf{P}}_{K-1}}) - H(V_{\mathsf{W}^{\mathsf{P}}_{K-1}}|U_{\overline{\phi}}, U_{\overline{K}}) \tag{30}$$
$$R_{\overline{\phi}} + R_{\overline{K-1}} + R_{\overline{K}} + R_{\overline{K-1.K}} \leq H(V_{\mathsf{W}^{\mathsf{P}}_j}|U_{\overline{\phi}}, U_{\overline{K-1}}, U_{\overline{K}}) + H(V_{\mathsf{W}^{\mathsf{P}}_{K-1}}|U_{\overline{\phi}}) - H(V_{\mathsf{W}^{\mathsf{P}}_{K-1}}|U_{\overline{\phi}}, U_{\overline{K}}) + H(V_{\mathsf{W}^{\mathsf{P}}_K}) - H(V_{\mathsf{W}^{\mathsf{P}}_K}|U_{\overline{\phi}}, U_{\overline{K-1}}) \tag{31}$$
$$R_{\overline{\phi}} + R_{\overline{K-1}} + R_{\overline{K}} + R_{\overline{K-1.K}} \leq H(V_{\mathsf{W}^{\mathsf{P}}_j}|U_{\overline{\phi}}, U_{\overline{K-1}}, U_{\overline{K}}) + H(V_{\mathsf{W}^{\mathsf{P}}_K}|U_{\overline{\phi}}) - H(V_{\mathsf{W}^{\mathsf{P}}_K}|U_{\overline{\phi}}, U_{\overline{K-1}}) + H(V_{\mathsf{W}^{\mathsf{P}}_{K-1}}) - H(V_{\mathsf{W}^{\mathsf{P}}_{K-1}}|U_{\overline{\phi}}, U_{\overline{K}}) \tag{32}$$
$$2R_{\overline{\phi}} + R_{\overline{K-1}} + R_{\overline{K}} + R_{\overline{K-1.K}} \leq H(V_{\mathsf{W}^{\mathsf{P}}_j}|U_{\overline{\phi}}, U_{\overline{K-1}}, U_{\overline{K}}) + H(V_{\mathsf{W}^{\mathsf{P}}_K}) - H(V_{\mathsf{W}^{\mathsf{P}}_K}|U_{\overline{\phi}}, U_{\overline{K-1}}) + H(V_{\mathsf{W}^{\mathsf{P}}_{K-1}}) - H(V_{\mathsf{W}^{\mathsf{P}}_{K-1}}|U_{\overline{\phi}}, U_{\overline{K}}) \tag{33}$$
$$2R_{\overline{\phi}} + 2R_{\overline{K-1}} + 2R_{\overline{K}} + R_{\overline{K-1.K}} \leq H(V_{\mathsf{W}^{\mathsf{P}}_j}|U_{\overline{\phi}}) + H(V_{\mathsf{W}^{\mathsf{P}}_K}) - H(V_{\mathsf{W}^{\mathsf{P}}_K}|U_{\overline{\phi}}, U_{\overline{K-1}}) + H(V_{\mathsf{W}^{\mathsf{P}}_{K-1}}) - H(V_{\mathsf{W}^{\mathsf{P}}_{K-1}}|U_{\overline{\phi}}, U_{\overline{K}}) \tag{34}$$
$$2R_{\overline{\phi}} + 2R_{\overline{K-1}} + 2R_{\overline{K}} + 2R_{\overline{K-1.K}} \leq H(V_{\mathsf{W}^{\mathsf{P}}_{j_1}}|U_{\overline{\phi}}) + H(V_{\mathsf{W}^{\mathsf{P}}_{j_2}}|U_{\overline{\phi}}, U_{\overline{K-1}}, U_{\overline{K}}) + H(V_{\mathsf{W}^{\mathsf{P}}_K}) - H(V_{\mathsf{W}^{\mathsf{P}}_K}|U_{\overline{\phi}}, U_{\overline{K-1}}) + H(V_{\mathsf{W}^{\mathsf{P}}_{K-1}}) - H(V_{\mathsf{W}^{\mathsf{P}}_{K-1}}|U_{\overline{\phi}}, U_{\overline{K}}) \tag{35}$$

for some joint distribution of the auxiliary and input random variables $(U_{\overline{\phi}}, U_{\overline{K}}, U_{\overline{K-1}}, X = V_{\mathsf{P}})$ that is of the form

$$p(u_{\overline{\phi}}, u_{\overline{K}}, u_{\overline{K-1}}, x) = p(u_{\overline{\phi}}) \times p(u_{\overline{K}}|u_{\overline{\phi}}) \times p(u_{\overline{K-1}}|u_{\overline{\phi}}) \times p(x|u_{\overline{\phi}}, u_{\overline{K}}, u_{\overline{K-1}}) \tag{36}$$

In the following theorem, we show that the inner bound in Theorem 1 achieves the capacity region of the general combination network by identifying the optimal distribution for the auxiliary random variables $U_{\overline{\phi}}, U_{\overline{K-1}}, U_{\overline{K}}$ and the channel input components $\{V_S\}_{S \in \mathsf{P}}$. The optimality of that distribution is proved indirectly by showing that the resulting inequalities are particular members of the generic family of generalized cut-set (outer) bounds found in [6].

Theorem 3: The capacity region of the K-user combination network for the diamond message set with $\mathsf{E} = \{M_{\overline{\phi}}, M_{\overline{K}}, M_{\overline{K-1}}, M_{\overline{K-1.K}}\}$ is the set of non-negative rate tuples $(R_{\overline{\phi}}, R_{\overline{K}}, R_{\overline{K-1}}, R_{\overline{K-1.K}})$ satisfying for all $j \in \{1, 2, \cdots, K-2\}$

$$R_{\overline{\phi}} + R_{\overline{K-1}} \leq C_{\mathsf{W}^{\mathsf{P}}_K} \tag{37}$$
$$R_{\overline{\phi}} + R_{\overline{K}} \leq C_{\mathsf{W}^{\mathsf{P}}_{K-1}} \tag{38}$$
$$R_{\overline{\phi}} + R_{\overline{K-1}} + R_{\overline{K}} \leq C_{\downarrow_{\mathsf{W}^{\mathsf{P}}_{K-1}}\{\overline{K}\}} + C_{\mathsf{W}^{\mathsf{P}}_K} \tag{39}$$
$$R_{\overline{\phi}} + R_{\overline{K-1}} + R_{\overline{K}} + R_{\overline{K-1.K}} \leq C_{\mathsf{W}^{\mathsf{P}}_j} \tag{40}$$
$$2R_{\overline{\phi}} + R_{\overline{K-1}} + R_{\overline{K}} + R_{\overline{K-1.K}} \leq C_{\downarrow_{\mathsf{W}^{\mathsf{P}}_j}\{\overline{K-1.K}\}} + C_{\mathsf{W}^{\mathsf{P}}_{K-1}} + C_{\mathsf{W}^{\mathsf{P}}_K} \tag{41}$$
$$2R_{\overline{\phi}} + 2R_{\overline{K-1}} + 2R_{\overline{K}} + R_{\overline{K-1.K}} \leq C_{\downarrow_{\mathsf{W}^{\mathsf{P}}_j}\{\overline{K}, \overline{K-1}\}} + C_{\mathsf{W}^{\mathsf{P}}_{K-1}} + C_{\mathsf{W}^{\mathsf{P}}_K} \tag{42}$$

where, recall that $C_{\mathsf{W}} = \sum_{S \in \mathsf{W}} C_S$ for any $\mathsf{W} \subseteq \mathsf{P}$.

Remark 1: Note that the right hand side of (39), (41), and (42) remain unchanged when K and $K-1$ are interchanged, as they should.

Remark 2: This capacity result generalizes three previously known results. First, it reproduces the capacity region for two nested (i.e., degraded) messages with at most two receivers demanding only the common messages $M_{\overline{\phi}}$ given in [15, Theorems 4 and 5] and [9, Theorem 3] by setting $R_{\overline{K-1.K}} = R_{\overline{K-1}} = 0$ and $R_{\overline{K}} = R_{\overline{K-1}} = 0$ in (37)-(42). Moreover, Theorem 3 also recovers the capacity region for two messages each intended for $K-1$ receivers given in [15, Theorem 3] by setting $R_{\overline{\phi}} = R_{\overline{K-1.K}} = 0$ in (37)-(42). It must be noted here that the converse proofs in [9] and [15] are established from "first principles", and while they use the sub-modularity of entropy, they do not take a top-down approach as we do here in Section III-D of making the connection to the generalized cut-set bound framework of [6]. Moreover, achievability for up to two common receivers was proved in [9, Proposition 1] using linear network coding tailored to the combination network and two nested messages, whereas we take the top-down approach in Section III-C of [2], [15] of starting from an inner bound for the DM BC and specializing it to the combination network.

The special case of Theorem 3 for three degraded messages (which is more general than [15, Theorem 5] and [9, Theorem 3]) is given next. We simply set $R_{\overline{K-1}} = 0$ in Theorem 3 and note that (39) becomes redundant because of (44) since $C_{\downarrow_{\mathsf{W}^{\mathsf{P}}_{K-1}}\{\overline{K}\}} + C_{\mathsf{W}^{\mathsf{P}}_K} = C_{\mathsf{W}^{\mathsf{P}}_{K-1}\cup\mathsf{W}^{\mathsf{P}}_K} \geq C_{\mathsf{W}^{\mathsf{P}}_{K-1}}$. Moreover, (42) also becomes redundant because the sum of inequalities (38) and (40) is tighter since $C_{\mathsf{W}^{\mathsf{P}}_j} \leq C_{\mathsf{W}^{\mathsf{P}}_j} + C_{\mathsf{W}^{\mathsf{P}}_K} - C_{\mathsf{W}^{\mathsf{P}}_j\cap\mathsf{W}^{\mathsf{P}}_{K-1}\cap\mathsf{W}^{\mathsf{P}}_K = C_{\downarrow_{\mathsf{W}^{\mathsf{P}}_j}\{\overline{K},\overline{K-1}\}} + C_{\mathsf{W}^{\mathsf{P}}_K}}$.

Corollary 2: The capacity region of the K-user combination network for the three degraded messages, i.e., $\mathsf{E} = \{M_{\overline{\phi}}, M_{\overline{K}}, M_{\overline{K-1.K}}\}$ is the set of non-negative rate tuples $(R_{\overline{\phi}}, R_{\overline{K}}, R_{\overline{K-1.K}})$ satisfying for all $j \in \{1, 2, \cdots, K-2\}$

$$R_{\overline{\phi}} \leq C_{\mathsf{W}^{\mathsf{P}}_K} \tag{43}$$

$$R_{\overline{\phi}} + R_{\overline{K}} \leq C_{\mathsf{W}^{\mathsf{P}}_{K-1}} \tag{44}$$

$$R_{\overline{\phi}} + R_{\overline{K}} + R_{\overline{K-1.K}} \leq C_{\mathsf{W}^{\mathsf{P}}_j} \tag{45}$$

$$2R_{\overline{\phi}} + R_{\overline{K}} + R_{\overline{K-1.K}} \leq C_{\downarrow_{\mathsf{W}^{\mathsf{P}}_j}\{\overline{K-1.K}\}} + C_{\mathsf{W}^{\mathsf{P}}_{K-1}} + C_{\mathsf{W}^{\mathsf{P}}_K} \tag{46}$$

Remark 3: The capacity region of the K-user combination network for the two degraded messages, i.e., $\mathsf{E} = \{M_{\overline{\phi}}, M_{\overline{K-1.K}}\}$ is the set of non-negative rate pairs $(R_{\overline{\phi}}, R_{\overline{K-1.K}})$ satisfying for all $j \in \{1, 2, \cdots, K-2\}$ the inequalities (43)-(46) (with $R_{\overline{K}} = 0$). This capacity region coincides with the one found in [15, Theorem 5] and [9, Theorem 3].

C. Proof of Achievability for Theorem 3

One or both of the two negative conditional entropy terms, namely, $H(V_{\mathsf{W}^{\mathsf{P}}_K}|U_{\overline{\phi}}, U_{\overline{K-1}})$ and $H(V_{\mathsf{W}^{\mathsf{P}}_{K-1}}|U_{\overline{\phi}}, U_{\overline{K}})$ appear in all the bounds (except in (28)) of the inequalities of Corollary 1. Hence those bounds can be upper bounded by setting $V_{\mathsf{W}^{\mathsf{P}}_K} = f(U_{\overline{\phi}}, U_{\overline{K-1}})$ and $V_{\mathsf{W}^{\mathsf{P}}_{K-1}} = g(U_{\overline{\phi}}, U_{\overline{K}})$ for some deterministic functions f and g. Next, one or both of the positive entropy terms $H(V_{\mathsf{W}^{\mathsf{P}}_{K-1}})$ and $H(V_{\mathsf{W}^{\mathsf{P}}_K})$ appear in each of the bounds (except in (28)) and all such terms can be maximized by taking the deterministic functions f and g to be identity maps so that $V_{\mathsf{W}^{\mathsf{P}}_K} = (U_{\overline{\phi}}, U_{\overline{K-1}})$ and $V_{\mathsf{W}^{\mathsf{P}}_{K-1}} = (U_{\overline{\phi}}, U_{\overline{K}})$ (thus this choice maximizes the first two bounds (24) and (25) of Corollary 1). Note that, if we make the choice $V_{\mathsf{W}^{\mathsf{P}}_K} = f(U_{\overline{\phi}}, U_{\overline{K-1}})$ and $V_{\mathsf{W}^{\mathsf{P}}_{K-1}} = g(U_{\overline{\phi}}, U_{\overline{K}})$, then the latter choice of letting f and g be identity maps also simultaneously maximizes the positive conditional entropy terms of the form $H(V_{\mathsf{W}^{\mathsf{P}}_j}|U_{\overline{\phi}}, U_{\overline{K-1}})$ and $H(V_{\mathsf{W}^{\mathsf{P}}_j}|U_{\overline{\phi}}, U_{\overline{K}})$ in the bounds of (29) and (30) and $H(V_{\mathsf{W}^{\mathsf{P}}_j}|U_{\overline{\phi}}, U_{\overline{K-1}}, U_{\overline{K}})$ in the bounds of (31)-(34). Next, consider the positive conditional entropy terms of the form $H(V_{\mathsf{W}^{\mathsf{P}}_{K-1}}|U_{\overline{\phi}})$, $H(V_{\mathsf{W}^{\mathsf{P}}_K}|U_{\overline{\phi}})$ in the bounds of (26)-(27) and (31)-(32) and $H(V_{\mathsf{W}^{\mathsf{P}}_j}|U_{\overline{\phi}})$ in (34)-(35). All of these terms are maximized by choosing (given $V_{\mathsf{W}^{\mathsf{P}}_K} = (U_{\overline{\phi}}, U_{\overline{K-1}})$ and $V_{\mathsf{W}^{\mathsf{P}}_{K-1}} = (U_{\overline{\phi}}, U_{\overline{K}})$) $U_{\overline{\phi}}$ to be the just the common part of $V_{\mathsf{W}^{\mathsf{P}}_{K-1}}$ and $V_{\mathsf{W}^{\mathsf{P}}_K}$. After making these choices, given that there are cardinality constraints on the alphabet $\mathcal{V}_S$ of V_S of 2^{C_S}, all the bounds are maximized by choosing the input components $(V_S : S \in \mathsf{F})$ to be uniform and independent with $\mathcal{V}_S$ taken to be $[1 : 2^{C_S}]$.

The above greedy optimization suggests (even though it does not prove the optimality of) the choice of coding distribution resulting by setting

$$U_{\overline{\phi}} = V_{\uparrow_{\mathsf{P}}\{K-1.K\}} \tag{47}$$

$$U_{\overline{K-1}} = V_{\mathsf{W}^{\mathsf{P}}_K} \tag{48}$$

$$U_{\overline{K}} = V_{\mathsf{W}^{\mathsf{P}}_{K-1}} \tag{49}$$

$$X = V_{\mathsf{P}} \tag{50}$$

and choosing the channel input components V_S for all $S \in \mathsf{P}$ to be independent and uniformly distributed over $\mathcal{V}_S$ where $|\mathcal{V}_S| = 2^{C_S}$. The achievability proof now follows from Theorem 1 by choosing the above coding distribution. Note that such a choice of the auxiliary random variables and the channel input components $(U_{\overline{\phi}}, U_{\overline{K}}, U_{\overline{K-1}}, X)$ has a joint distribution that satisfies (19), as it must. For this specific choice, we can compute the mutual information terms in (7)-(18) using $Y_i = V_{\mathsf{W}^{\mathsf{P}}_i}$ and (3)-(6) to obtain

$$I(U_{\overline{\phi}}, U_{\overline{K-1}}; Y_K) = C_{\mathsf{W}^{\mathsf{P}}_K} \tag{51}$$

$$I(U_{\overline{\phi}}, U_{\overline{K}}; Y_{K-1}) = C_{\mathsf{W}^{\mathsf{P}}_{K-1}} \tag{52}$$

$$I(U_{\overline{K}}; Y_{K-1}|U_{\overline{\phi}}) = C_{\downarrow_{\mathsf{W}^{\mathsf{P}}_{K-1}}\{\overline{K}\}} \tag{53}$$

$$I(U_{\overline{K-1}}; Y_K|U_{\overline{\phi}}) = C_{\downarrow_{\mathsf{W}^{\mathsf{P}}_K}\{\overline{K-1}\}} \tag{54}$$

$$I(X; Y_j) = C_{\mathsf{W}^{\mathsf{P}}_j} \tag{55}$$

$$I(X; Y_j|U_{\overline{\phi}}, U_{\overline{K}}) = C_{\downarrow_{\mathsf{W}^{\mathsf{P}}_j}\{\overline{K}\}} \tag{56}$$

$$I(X; Y_j|U_{\overline{\phi}}, U_{\overline{K-1}}) = C_{\downarrow_{\mathsf{W}^{\mathsf{P}}_j}\{\overline{K-1}\}} \tag{57}$$

$$I(X; Y_j|U_{\overline{\phi}}, U_{\overline{K-1}}, U_{\overline{K}}) = C_{\downarrow_{\mathsf{W}^{\mathsf{P}}_j}\{\overline{K-1.K}\}} \tag{58}$$

$$I(X; Y_j|U_{\overline{\phi}}) = C_{\downarrow_{\mathsf{W}^{\mathsf{P}}_j}\{\overline{K},\overline{K-1}\}} \tag{59}$$

where $j \in \{1, 2, \cdots, K-2\}$. Hence, we get (37)-(38) by substituting (51)-(52) in (7)-(8) and (39) by substituting (53)-(54) in (9)-(10) since $C_{\downarrow_{\mathsf{W}^{\mathsf{P}}_{K-1}}\{\overline{K}\}} + C_{\mathsf{W}^{\mathsf{P}}_K} = C_{\downarrow_{\mathsf{W}^{\mathsf{P}}_K}\{\overline{K-1}\}} + C_{\mathsf{W}^{\mathsf{P}}_{K-1}}$. Moreover, we get (40) directly by substituting (55) in (11). Then, we obtain (41) from (16) using (51)-(52) and (58). Finally, we get (42) by substituting (51)-(52) and (59) in (17).

The rest of the inequalities in the inner bound, i.e., (12)-(15) and (18) are redundant, as we show next. By computing the mutual information terms in (12)-(15) and (18) for the specified coding distribution in (47)-(50), it is left to the reader to verify that we obtain the following five categories of bounds (with each j, j_1, j_2 taking all possible values in the set $\{1, 2, \cdots, K-2\}$)

$$R_{\overline{\phi}} + R_{\overline{K-1}} + R_{\overline{K}} + R_{\overline{K-1.K}} \leq C_{\downarrow_{\mathsf{W}^{\mathsf{P}}_j}\{\overline{K}\}} + C_{\mathsf{W}^{\mathsf{P}}_K} \tag{60}$$

$$R_{\overline{\phi}} + R_{\overline{K-1}} + R_{\overline{K}} + R_{\overline{K-1.K}} \leq C_{\downarrow_{\mathsf{W}^{\mathsf{P}}_j}\{\overline{K-1}\}} + C_{\mathsf{W}^{\mathsf{P}}_{K-1}} \tag{61}$$

$$R_{\overline{\phi}} + R_{\overline{K-1}} + R_{\overline{K}} + R_{\overline{K-1.K}} \leq C_{\downarrow_{\mathsf{W}^{\mathsf{P}}_j}\{\overline{K-1.K}\}} + C_{\downarrow_{\mathsf{W}^{\mathsf{P}}_{K-1}}\{\overline{K}\}} + C_{\mathsf{W}^{\mathsf{P}}_K} \tag{62}$$

$$R_{\overline{\phi}} + R_{\overline{K-1}} + R_{\overline{K}} + R_{\overline{K-1.K}} \leq C_{\downarrow_{\mathsf{W}^{\mathsf{P}}_j}\{\overline{K-1.K}\}} + C_{\downarrow_{\mathsf{W}^{\mathsf{P}}_K}\{\overline{K-1}\}} + C_{\mathsf{W}^{\mathsf{P}}_{K-1}} \tag{63}$$

$$2R_{\overline{\phi}} + 2R_{\overline{K-1}} + 2R_{\overline{K}} + 2R_{\overline{K-1.K}} \leq C_{\downarrow_{\mathsf{W}^{\mathsf{P}}_{j_1}}\{\overline{K},\overline{K-1}\}} + C_{\downarrow_{\mathsf{W}^{\mathsf{P}}_{j_1}}\{\overline{K-1.K}\}} + C_{\mathsf{W}^{\mathsf{P}}_{K-1}} + C_{\mathsf{W}^{\mathsf{P}}_K} \tag{64}$$

We next show that these categories of bounds are all redundant. First, (60) and (61) are redundant from (40). Using (79), we can rewrite the right hand side of (60) and lower bound it as follows

$$C_{\downarrow_{\mathsf{W}^{\mathsf{P}}_j}\{\overline{K}\}} + C_{\mathsf{W}^{\mathsf{P}}_K} = C_{\mathsf{W}^{\mathsf{P}}_j} - C_{\mathsf{W}^{\mathsf{P}}_j \cap \mathsf{W}^{\mathsf{P}}_K} + C_{\mathsf{W}^{\mathsf{P}}_K} \geq C_{\mathsf{W}^{\mathsf{P}}_j} \tag{65}$$

where (65) follows from the fact that $C_{\mathsf{W}^{\mathsf{P}}_j \cap \mathsf{W}^{\mathsf{P}}_K} \leq C_{\mathsf{W}^{\mathsf{P}}_K}$. Similarly, we can show that (61) is redundant due to (40).

Moreover, (62) and (63) are also redundant from (40) since we can rewrite the right hand side of (62), using (79), and lower bound it as follows

$$\begin{aligned} &C_{\downarrow_{\mathsf{W}^{\mathsf{P}}_j}\{\overline{K-1.K}\}} + C_{\downarrow_{\mathsf{W}^{\mathsf{P}}_{K-1}}\{\overline{K}\}} + C_{\mathsf{W}^{\mathsf{P}}_K} \\ &= C_{\mathsf{W}^{\mathsf{P}}_j} - C_{\mathsf{W}^{\mathsf{P}}_j \cap (\mathsf{W}^{\mathsf{P}}_K \cup \mathsf{W}^{\mathsf{P}}_{K-1})} + C_{\mathsf{W}^{\mathsf{P}}_{K-1}} - C_{\mathsf{W}^{\mathsf{P}}_{K-1} \cap \mathsf{W}^{\mathsf{P}}_{K-1}} + C_{\mathsf{W}^{\mathsf{P}}_K} \\ &= C_{\mathsf{W}^{\mathsf{P}}_j} - C_{\mathsf{W}^{\mathsf{P}}_j \cap (\mathsf{W}^{\mathsf{P}}_K \cup \mathsf{W}^{\mathsf{P}}_{K-1})} + C_{\mathsf{W}^{\mathsf{P}}_{K-1} \cup \mathsf{W}^{\mathsf{P}}_{K-1}} \qquad (66) \\ &\geq C_{\mathsf{W}^{\mathsf{P}}_j}, \end{aligned}$$

where (66) follows from the modularity of C_{W}.

Finally, (64) is redundant from two times (40) since the right hand side of (62) can be written, using (79)-(80), and lower bounded as follows

$$\begin{aligned} &C_{\downarrow_{\mathsf{W}^{\mathsf{P}}_{j_1}}\{\overline{K},\overline{K-1}\}} + C_{\downarrow_{\mathsf{W}^{\mathsf{P}}_{j_2}}\{\overline{K-1.K}\}} + C_{\mathsf{W}^{\mathsf{P}}_{K-1}} + C_{\mathsf{W}^{\mathsf{P}}_K} \\ &= C_{\mathsf{W}^{\mathsf{P}}_{j_1}} - C_{\mathsf{W}^{\mathsf{P}}_{j_1} \cap \mathsf{W}^{\mathsf{P}}_{K-1} \cap \mathsf{W}^{\mathsf{P}}_K} + C_{\mathsf{W}^{\mathsf{P}}_{j_2}} - C_{\mathsf{W}^{\mathsf{P}}_{j_2} \cap (\mathsf{W}^{\mathsf{P}}_{K-1} \cup \mathsf{W}^{\mathsf{P}}_K)} \\ &\quad + C_{\mathsf{W}^{\mathsf{P}}_{K-1}} + C_{\mathsf{W}^{\mathsf{P}}_K} \\ &= C_{\mathsf{W}^{\mathsf{P}}_{j_1}} - C_{\mathsf{W}^{\mathsf{P}}_{j_1} \cap \mathsf{W}^{\mathsf{P}}_{K-1} \cap \mathsf{W}^{\mathsf{P}}_K} + C_{\mathsf{W}^{\mathsf{P}}_{K-1} \cap \mathsf{W}^{\mathsf{P}}_K} \\ &\quad + C_{\mathsf{W}^{\mathsf{P}}_{j_2}} - C_{\mathsf{W}^{\mathsf{P}}_{j_2} \cap (\mathsf{W}^{\mathsf{P}}_{K-1} \cup \mathsf{W}^{\mathsf{P}}_K)} + C_{\mathsf{W}^{\mathsf{P}}_{K-1} \cup \mathsf{W}^{\mathsf{P}}_K} \qquad (67) \\ &\geq C_{\mathsf{W}^{\mathsf{P}}_{j_1}} + C_{\mathsf{W}^{\mathsf{P}}_{j_2}}, \end{aligned}$$

where (67) follows from the modularity of C_{W}, and hence, $C_{\mathsf{W}^{\mathsf{P}}_{K-1}} + C_{\mathsf{W}^{\mathsf{P}}_K} = C_{\mathsf{W}^{\mathsf{P}}_{K-1} \cap \mathsf{W}^{\mathsf{P}}_K} + C_{\mathsf{W}^{\mathsf{P}}_{K-1} \cup \mathsf{W}^{\mathsf{P}}_K}$.

This concludes the achievability of the polytope described by the inequalities (37)-(42) by the single coding distribution given in (47)-(50).

Remark 4: In Section III-D, we show that the polytope (in the positive orthant) described by (37)-(42) is indeed the capacity region of the combination network. This indirectly shows that when the inner bound of Theorem 1 is specialized to the combination network as a union of polytopes with the union taken over all admissible coding distributions of the form (19) then the distribution given in (47)-(50) is extremal in that the rate region it produces subsumes the rate regions produced by any other distribution that satisfies (19).

D. The Converse Proof for Theorem 3

The converse proof of (37)-(38) and (40) follows directly from the cut-set bounds. In what follows, we show that the inequalities (39) and (41)-(42) are outer bounds by identifying them as belonging to the infinite class of generalized cut-set bounds obtained in [6]. Those bounds provide a generic framework that can be used to obtain upper bounds on the achievable rates sent over any broadcast network that can be described by a directed acyclic graph with a non-negative capacity for each link or arc in the graph (and hence, over the combination network). As shown in [6], these upper bounds take the generic form

$$\sum_{i \in S'} \alpha_i R_{\Phi_i(\mathsf{W}^{\mathsf{E}}_1, \mathsf{W}^{\mathsf{E}}_2, \cdots, \mathsf{W}^{\mathsf{E}}_K)} \leq \sum_{i \in S'} \alpha_i C_{\Phi_i(\mathsf{W}^{\mathsf{P}}_1, \mathsf{W}^{\mathsf{P}}_2, \cdots, \mathsf{W}^{\mathsf{P}}_K)} \tag{68}$$

where S' is a finite non-empty set, α_i's are non-negative real numbers and $\{\Phi_i\}_{i \in S'}$ is a collection of set operators. The set operator is defined as a finite sequence of intersections and unions acting on any K subsets of P to produce a subset of P. Furthermore, these set operators $\{\Phi_i\}_{i \in S'}$ in (68) must satisfy the following extremal inequality

$$\begin{aligned} \sum_{i \in S_1} \alpha_i f(\Phi_i(\mathsf{W}_1, \cdots, \mathsf{W}_K)) &\leq \sum_{j \in S_2} \beta_j f(\Pi_j(\mathsf{W}_1, \cdots, \mathsf{W}_K)) \\ &+ \sum_{l \in S_3} \gamma_l \big(f(\Gamma^+_l(\mathsf{W}_1, \cdots, \mathsf{W}_K)) - f(\Gamma^-_l(\mathsf{W}_1, \cdots, \mathsf{W}_K))\big) \end{aligned} \tag{69}$$

for any submodular function f, with equality holding when f is a modular function, $\{\mathsf{W}_i\}_{i=1}^K \subseteq \mathsf{P}$, nonempty finite sets S_1, S_2, S_3, collection of non-negative real numbers $(\alpha_i, \beta_j, \gamma_l)$, and a collection of set operators $\Pi_j, \Gamma^+_l, \Gamma^-_l$ such that Π_j, Γ^+_l are finite sequences of unions only, and $\Gamma^-_l(\mathsf{W}_1, \mathsf{W}_2, \cdots, \mathsf{W}_K) \subseteq \Gamma^+_l(\mathsf{W}_1, \mathsf{W}_2, \cdots, \mathsf{W}_K)$ for any $l \in S_3$.

The main problem in using the framework of generalized cut-set bounds of the form (68) is to identify the set operators $\Phi_i, \Pi_j, \Gamma^+_l, \Gamma^-_l$ for all $i \in S_1$, $j \in S_2$ and $l \in S_3$ that satisfy the extremal inequality (69) and provide specific upper bounds (68) that exactly match with (39) and (41)-(42), and thereby establishing the converse proof. With the aid of the order-theoretic relations (1)-(6), we show that the following extremal inequalities proposed before in [6, Proposition 1] provide tight upper bounds that match (39) and (41)-(42). From [6, Proposition 1], we get the following three extremal inequalities for $j \in \{1, 2, \cdots, K-2\}$

$$f(\mathsf{W}_{K-1} \cup \mathsf{W}_K) \leq f(\mathsf{W}_{K-1}) + f(\mathsf{W}_K) - f(\mathsf{W}_{K-1} \cap \mathsf{W}_K) \tag{70}$$

$$f(\mathsf{W}_{K-1} \cup \mathsf{W}_K \cup \mathsf{W}_j) + f(\mathsf{W}_{K-1} \cap \mathsf{W}_K) \leq f(\mathsf{W}_j) - f(\mathsf{W}_j \cap (\mathsf{W}_{K-1} \cup \mathsf{W}_K)) + f(\mathsf{W}_{K-1}) + f(\mathsf{W}_K) \tag{71}$$

$$f(\mathsf{W}_{K-1} \cup \mathsf{W}_K \cup \mathsf{W}_j) + f((\mathsf{W}_{K-1} \cap \mathsf{W}_K) \cup (\mathsf{W}_{K-1} \cap \mathsf{W}_j) \cup (\mathsf{W}_K \cap \mathsf{W}_j)) \leq f(\mathsf{W}_j) - f(\mathsf{W}_j \cap \mathsf{W}_{K-1} \cap \mathsf{W}_K) + f(\mathsf{W}_{K-1}) + f(\mathsf{W}_K) \quad (72)$$

Using them, we get immediately from (68) the following generalized cut-set bounds

$$R_{\mathsf{W}^{\mathsf{E}}_{K-1} \cup \mathsf{W}^{\mathsf{E}}_K} \leq C_{\mathsf{W}^{\mathsf{P}}_{K-1} \cup \mathsf{W}^{\mathsf{P}}_K} \quad (73)$$

$$R_{\mathsf{W}^{\mathsf{E}}_{K-1} \cup \mathsf{W}^{\mathsf{E}}_K \cup \mathsf{W}^{\mathsf{E}}_j} + R_{\mathsf{W}^{\mathsf{E}}_{K-1} \cap \mathsf{W}^{\mathsf{E}}_K} \leq C_{\mathsf{W}^{\mathsf{P}}_{K-1} \cup \mathsf{W}^{\mathsf{P}}_K \cup \mathsf{W}^{\mathsf{P}}_j} + C_{\mathsf{W}^{\mathsf{P}}_{K-1} \cap \mathsf{W}^{\mathsf{P}}_K} \quad (74)$$

$$R_{\mathsf{W}^{\mathsf{E}}_{K-1} \cup \mathsf{W}^{\mathsf{E}}_K \cup \mathsf{W}^{\mathsf{E}}_j} + R_{(\mathsf{W}^{\mathsf{E}}_{K-1} \cap \mathsf{W}^{\mathsf{E}}_K) \cup (\mathsf{W}^{\mathsf{E}}_{K-1} \cap \mathsf{W}^{\mathsf{E}}_j) \cup (\mathsf{W}^{\mathsf{E}}_K \cap \mathsf{W}^{\mathsf{E}}_j)} \leq C_{\mathsf{W}^{\mathsf{P}}_{K-1} \cup \mathsf{W}^{\mathsf{P}}_K \cup \mathsf{W}^{\mathsf{P}}_j} + C_{(\mathsf{W}^{\mathsf{P}}_{K-1} \cap \mathsf{W}^{\mathsf{P}}_K) \cup (\mathsf{W}^{\mathsf{P}}_{K-1} \cap \mathsf{W}^{\mathsf{P}}_j) \cup (\mathsf{W}^{\mathsf{P}}_K \cap \mathsf{W}^{\mathsf{P}}_j)} \quad (75)$$

By computing the left hand side of the above three inequalities, we get for any $j \in \{1, 2, \cdots, K-2\}$

$$R_{\mathsf{W}^{\mathsf{E}}_{K-1} \cup \mathsf{W}^{\mathsf{E}}_K} = R_{\overline{\phi}} + R_{\overline{K-1}} + R_{\overline{K}} \quad (76)$$

$$R_{\mathsf{W}^{\mathsf{E}}_{K-1} \cup \mathsf{W}^{\mathsf{E}}_K \cup \mathsf{W}^{\mathsf{E}}_j} + R_{\mathsf{W}^{\mathsf{E}}_{K-1} \cap \mathsf{W}^{\mathsf{E}}_K} = 2R_{\overline{\phi}} + R_{\overline{K-1}} + R_{\overline{K}} + R_{\overline{K-1.K}} \quad (77)$$

$$R_{\mathsf{W}^{\mathsf{E}}_{K-1} \cup \mathsf{W}^{\mathsf{E}}_K \cup \mathsf{W}^{\mathsf{E}}_j} + R_{(\mathsf{W}^{\mathsf{E}}_{K-1} \cap \mathsf{W}^{\mathsf{E}}_K) \cup (\mathsf{W}^{\mathsf{E}}_{K-1} \cap \mathsf{W}^{\mathsf{E}}_j) \cup (\mathsf{W}^{\mathsf{E}}_K \cap \mathsf{W}^{\mathsf{E}}_j)} = 2R_{\overline{\phi}} + 2R_{\overline{K-1}} + 2R_{\overline{K}} + R_{\overline{K-1.K}} \quad (78)$$

These match with the left hand side of (39) and (41)-(42).

Showing that the right hand sides of (73)-(75) match with that of (39) and (41)-(42), respectively, is less straightforward. We first use (1)-(6) to rewrite the right hand side of (39) and (41)-(42) in a form that is closer to the generalized cut-set bounds. From (1)-(6), we can show that the following two equalities hold for any set $S = l_1 l_2 \cdots l_N \subset \{1, 2, \cdots, K\}$ and $j \in [1:K]$

$$C_{\downarrow_{\mathsf{W}^{\mathsf{P}}_j}\{\overline{S}\}} = C_{\mathsf{W}^{\mathsf{P}}_j} - C_{\mathsf{W}^{\mathsf{P}}_j \cap (\mathsf{W}^{\mathsf{P}}_{l_1} \cup \mathsf{W}^{\mathsf{P}}_{l_2} \cup \cdots \cup \mathsf{W}^{\mathsf{P}}_{l_N})} \quad (79)$$

$$C_{\downarrow_{\mathsf{W}^{\mathsf{P}}_j}\{\overline{l_1}, \overline{l_2}, \cdots, \overline{l_N}\}} = C_{\mathsf{W}^{\mathsf{P}}_j} - C_{\mathsf{W}^{\mathsf{P}}_j \cap (\mathsf{W}^{\mathsf{P}}_{l_1} \cap \mathsf{W}^{\mathsf{P}}_{l_2} \cap \cdots \cap \mathsf{W}^{\mathsf{P}}_{l_N})} \quad (80)$$

where $C_{\mathsf{W}} = \sum_{S \in \mathsf{W}} C_S$ for any $\mathsf{W} \subseteq \mathsf{P}$. Using (79)-(80), we can rewrite the achievable bounds in (39) and (41)-(42) as follows

$$R_{\overline{\phi}} + R_{\overline{K-1}} + R_{\overline{K}} \leq C_{\mathsf{W}^{\mathsf{P}}_{K-1}} + C_{\mathsf{W}^{\mathsf{P}}_K} - C_{\mathsf{W}^{\mathsf{P}}_{K-1} \cap \mathsf{W}^{\mathsf{P}}_K} \quad (81)$$

$$2R_{\overline{\phi}} + R_{\overline{K-1}} + R_{\overline{K}} + R_{\overline{K-1.K}} \leq C_{\mathsf{W}^{\mathsf{P}}_j} - C_{\mathsf{W}^{\mathsf{P}}_j \cap (\mathsf{W}^{\mathsf{P}}_{K-1} \cup \mathsf{W}^{\mathsf{P}}_K)} + C_{\mathsf{W}^{\mathsf{P}}_{K-1}} + C_{\mathsf{W}^{\mathsf{P}}_K} \quad (82)$$

$$2R_{\overline{\phi}} + 2R_{\overline{K-1}} + 2R_{\overline{K}} + R_{\overline{K-1.K}} \leq C_{\mathsf{W}^{\mathsf{P}}_j} - C_{\mathsf{W}^{\mathsf{P}}_j \cap (\mathsf{W}^{\mathsf{P}}_{K-1} \cap \mathsf{W}^{\mathsf{P}}_K)} + C_{\mathsf{W}^{\mathsf{P}}_{K-1}} + C_{\mathsf{W}^{\mathsf{P}}_K} \quad (83)$$

By comparing (73)-(74) with (81)-(83), the new representation of the achievable bound, we can see that the right hand side of both group of inequalities are identical using the extremal inequalities (70)-(72) for modular function since C_{W} is modular function for any $\mathsf{W} \subseteq \mathsf{P}$. Hence, the generalized cut-set bounds in (73)-(74) are tight. This concludes the converse proof.

Example 1: The capacity region for the complete message set for the four-receiver combination network is not known. We specify the achievable rate region of Theorem 3 for $K = 4$ for the diamond message set, but without the structure in the order-theoretic specification of the bounds, as being the explicit polytope in the positive orthant described by the set of rate pairs $(R_{12}, R_{123}, R_{124}, R_{1234})$ that satisfy the following nine linear inequalities

$$R_{1234} + R_{124} \leq C_4 + C_{14} + C_{24} + C_{34} + C_{124} + C_{134} + C_{234} + C_{1234} \quad (84)$$

$$R_{1234} + R_{123} \leq C_3 + C_{13} + C_{23} + C_{34} + C_{123} + C_{134} + C_{234} + C_{1234} \quad (85)$$

$$R_{1234} + R_{124} + R_{123} \leq C_3 + C_4 + C_{13} + C_{14} + C_{23} + C_{24} + C_{34} + C_{123} + C_{124} + C_{134} + C_{234} + C_{1234} \quad (86)$$

$$R_{1234} + R_{124} + R_{123} + R_{12} \leq C_1 + C_{12} + C_{13} + C_{14} + C_{123} + C_{124} + C_{134} + C_{1234} \quad (87)$$

$$R_{1234} + R_{124} + R_{123} + R_{12} \leq C_2 + C_{12} + C_{23} + C_{24} + C_{123} + C_{124} + C_{234} + C_{1234} \quad (88)$$

$$2R_{1234} + R_{124} + R_{123} + R_{12} \leq C_1 + C_3 + C_4 + C_{12} + C_{13} + C_{14} + C_{23} + C_{24} + 2C_{34} + C_{123} + C_{124} + 2C_{134} + 2C_{234} + 2C_{1234} \quad (89)$$

$$2R_{1234} + R_{124} + R_{123} + R_{12} \leq C_2 + C_3 + C_4 + C_{12} + C_{13} + C_{14} + C_{23} + C_{24} + 2C_{34} + C_{123} + C_{124} + 2C_{134} + 2C_{234} + 2C_{1234} \quad (90)$$

$$2R_{1234} + 2R_{124} + 2R_{123} + R_{12} \leq C_1 + C_3 + C_4 + C_{12} + 2C_{13} + 2C_{14} + C_{23} + C_{24} + 2C_{34} + 2C_{123} + 2C_{124} + 2C_{134} + 2C_{234} + 2C_{1234} \quad (91)$$

$$2R_{1234} + 2R_{124} + 2R_{123} + R_{12} \leq C_2 + C_3 + C_4 + C_{12} + C_{13} + C_{14} + 2C_{23} + 2C_{24} + 2C_{34} + 2C_{123} + 2C_{124} + 2C_{134} + 2C_{234} + 2C_{1234} \quad (92)$$

IV. Inner Bound for the DM BC With Binning

For the sake of completeness and further investigation, we generalize Theorem 1 by enhancing the achievable scheme therein by adding to it the technique of binning.

Theorem 4: An inner bound of K-user DM BC for the diamond message set with $\mathsf{E} = \{\overline{\phi}, \overline{K}, \overline{K-1}, \overline{K-1.K}\}$ is the convex closure of the set of non-negative rate tuples $(R_{\overline{\phi}}, R_{\overline{K}}, R_{\overline{K-1}}, R_{\overline{K-1.K}})$ satisfying the following for all $j, j_1, j_2 \in \{1, 2, \cdots, K-2\}$

$$R_{\overline{\phi}} + R_{\overline{K-1}} \leq I(U_{\overline{\phi}}, U_{\overline{K-1}}; Y_K) \quad (93)$$

$$R_{\overline{\phi}} + R_{\overline{K}} \leq I(U_{\overline{\phi}}, U_{\overline{K}}; Y_{K-1}) \quad (94)$$

$$R_{\overline{\phi}} + R_{\overline{K-1}} + R_{\overline{K}} \leq I(U_{\overline{K}}; Y_{K-1} | U_{\overline{\phi}}) + I(U_{\overline{\phi}}, U_{\overline{K-1}}; Y_K) - I(U_{\overline{K}}; U_{\overline{K-1}} | U_{\overline{\phi}}) \quad (95)$$

$$R_{\overline{\phi}} + R_{\overline{K-1}} + R_{\overline{K}} \leq I(U_{\overline{K-1}}; Y_K | U_{\overline{\phi}}) + I(U_{\overline{\phi}}, U_{\overline{K}}; Y_{K-1}) - I(U_{\overline{K}}; U_{\overline{K-1}} | U_{\overline{\phi}}) \quad (96)$$

$$2R_{\overline{\phi}} + R_{\overline{K-1}} + R_{\overline{K}} \leq I(U_{\overline{\phi}}, U_{\overline{K}}; Y_{K-1}) + I(U_{\overline{\phi}}, U_{\overline{K-1}}; Y_K) - I(U_{\overline{K}}; U_{\overline{K-1}} | U_{\overline{\phi}}) \quad (97)$$

$$R_{\overline{\phi}} + R_{\overline{K-1}} + R_{\overline{K}} + R_{\overline{K-1.K}} \leq I(X; Y_j) \quad (98)$$

$$R_{\overline{\phi}} + R_{\overline{K-1}} + R_{\overline{K}} + R_{\overline{K-1.K}} \leq I(X;Y_j|U_{\overline{\phi}},U_{\overline{K-1}}) + I(U_{\overline{\phi}},U_{\overline{K-1}};Y_K) \quad (99)$$

$$R_{\overline{\phi}} + R_{\overline{K-1}} + R_{\overline{K}} + R_{\overline{K-1.K}} \leq I(X;Y_j|U_{\overline{\phi}},U_{\overline{K}}) + I(U_{\overline{\phi}},U_{\overline{K}};Y_{K-1}) \quad (100)$$

$$R_{\overline{\phi}} + R_{\overline{K-1}} + R_{\overline{K}} + R_{\overline{K-1.K}} \leq I(X;Y_j|U_{\overline{\phi}},U_{\overline{K-1}},U_{\overline{K}}) + I(U_{\overline{K}};Y_{K-1}|U_{\overline{\phi}}) + I(U_{\overline{\phi}},U_{\overline{K-1}};Y_K) - I(U_{\overline{K}};U_{\overline{K-1}}|U_{\overline{\phi}}) \quad (101)$$

$$R_{\overline{\phi}} + R_{\overline{K-1}} + R_{\overline{K}} + R_{\overline{K-1.K}} \leq I(X;Y_j|U_{\overline{\phi}},U_{\overline{K-1}},U_{\overline{K}}) + I(U_{\overline{K-1}};Y_K|U_{\overline{\phi}}) + I(U_{\overline{\phi}},U_{\overline{K}};Y_{K-1}) - I(U_{\overline{K}};U_{\overline{K-1}}|U_{\overline{\phi}}) \quad (102)$$

$$2R_{\overline{\phi}} + R_{\overline{K-1}} + R_{\overline{K}} + R_{\overline{K-1.K}} \leq I(X;Y_j|U_{\overline{\phi}},U_{\overline{K-1}},U_{\overline{K}}) + I(U_{\overline{\phi}},U_{\overline{K}};Y_{K-1}) + I(U_{\overline{\phi}},U_{\overline{K-1}};Y_K) - I(U_{\overline{K}};U_{\overline{K-1}}|U_{\overline{\phi}}) \quad (103)$$

$$2R_{\overline{\phi}} + 2R_{\overline{K-1}} + 2R_{\overline{K}} + R_{\overline{K-1.K}} \leq I(X;Y_j|U_{\overline{\phi}}) + I(U_{\overline{\phi}},U_{\overline{K}};Y_{K-1}) + I(U_{\overline{\phi}},U_{\overline{K-1}};Y_K) - I(U_{\overline{K}};U_{\overline{K-1}}|U_{\overline{\phi}}) \quad (104)$$

$$2R_{\overline{\phi}} + 2R_{\overline{K-1}} + 2R_{\overline{K}} + 2R_{\overline{K-1.K}} \leq I(X;Y_{j_1}|U_{\overline{\phi}}) + I(X;Y_{j_2}|U_{\overline{\phi}},U_{\overline{K-1}},U_{\overline{K}}) + I(U_{\overline{\phi}},U_{\overline{K}};Y_{K-1}) + I(U_{\overline{\phi}},U_{\overline{K-1}};Y_K) - I(U_{\overline{K}};U_{\overline{K-1}}|U_{\overline{\phi}}) \quad (105)$$

$$2R_{\overline{\phi}} + R_{\overline{K-1}} + 2R_{\overline{K}} + R_{\overline{K-1.K}} \leq I(X;Y_j|U_{\overline{\phi}},U_{\overline{K-1}}) + I(U_{\overline{\phi}},U_{\overline{K-1}};Y_K) + I(U_{\overline{\phi}},U_{\overline{K}};Y_{K-1}) - I(U_{\overline{K}};U_{\overline{K-1}}|U_{\overline{\phi}}) \quad (106)$$

$$2R_{\overline{\phi}} + 2R_{\overline{K-1}} + R_{\overline{K}} + R_{\overline{K-1.K}} \leq I(X;Y_j|U_{\overline{\phi}},U_{\overline{K}}) + I(U_{\overline{\phi}},U_{\overline{K-1}};Y_K) + I(U_{\overline{\phi}},U_{\overline{K}};Y_{K-1}) - I(U_{\overline{K}};U_{\overline{K-1}}|U_{\overline{\phi}}) \quad (107)$$

$$2R_{\overline{\phi}} + 2R_{\overline{K-1}} + 2R_{\overline{K}} + 2R_{\overline{K-1.K}} \leq I(X;Y_{j_1}|U_{\overline{\phi}},U_{\overline{K}}) + I(X;Y_{j_2}|U_{\overline{\phi}},U_{\overline{K-1}}) + I(U_{\overline{\phi}},U_{\overline{K-1}};Y_K) + I(U_{\overline{\phi}},U_{\overline{K}};Y_{K-1}) - I(U_{\overline{K}};U_{\overline{K-1}}|U_{\overline{\phi}}) \quad (108)$$

for some joint distribution of the auxiliary and input random variables $(U_{\overline{\phi}},U_{\overline{K}},U_{\overline{K-1}},X)$ that is of the form

$$p(u_{\overline{\phi}},u_{\overline{K}},u_{\overline{K-1}},x,y_1,\cdots,y_K) = p(u_{\overline{\phi}})\times p(u_{\overline{K}},u_{\overline{K-1}}|u_{\overline{\phi}}) \times p(x|u_{\overline{\phi}},u_{\overline{K}},u_{\overline{K-1}}) \times p(y_1,\cdots,y_K|x) \quad (109)$$

Proof: An outline of the proof is given in Appendix IV. ■

Remark 5: By setting $I(U_{\overline{K}};U_{\overline{K-1}}|U_{\overline{\phi}}) = 0$ (i.e., no binning), we get that (97) is redundant from (93) and (94), (106) is redundant from (99) and (94), (107) is redundant from (100) and (93), and (108) is redundant from (99) and (100), thereby recovering Theorem 1. While the greater generality offered by binning is not necessary in establishing the capacity region of the combination network for the diamond message set it is an interesting open question as to whether the above achievable rate region is tight for say the more general deterministic DM BC where each output Y_i is some deterministic function of $g_i(X)$.

Remark 6: In the achievable rate region for the $K = 2$ case of Theorem 4 obtained for the complete message set $\mathsf{E} = \{1,2,12\}$ by setting $R_{\overline{K-1.K}} = 0$ (in this case X is a function of U_1, U_2 and U_{12}) and $Y_j = \text{const.}$ for $j \in [3:K]$ it is easily seen that all but the first five inequalities (93)-(97) are redundant. In particular, this special case of Theorem 4 recovers [16, Theorem 5], where it was first obtained, as it must, and is also given in [17, Theorem 8.4].

Remark 7: The special case of $K = 3$ and three degraded messages with $\mathsf{E} = \{1,12,123\}$ which is a subset of the diamond message set $\mathsf{E} = \{1,12,13,123\}$ can be deduced by setting $K = 3$ and $R_{13} = 0$ in Theorem 4. This special case was also addressed in [18, Theorem 2]. There are some notable differences between the two regions however[5]. First, the distributions for the input and auxiliary random variables is taken to belong to a more restrictive class in [18, Theorem 2] that can be obtained by setting $U_{123} = U$, $(U_{123},U_{12}) = V_2$ and $(U_{123},U_{13}) = V_3$ and restricting to joint distributions wherein we have the Markov chains $U \multimap V_2 \multimap (V_3,X)$ and $U \multimap V_3 \multimap (V_2,X)$ which is a stronger restriction than the one herein which is that $U_{123} \multimap (U_{12},U_{13}) \multimap X$. However, the restriction in the class of coding distributions in [18, Theorem 2] is without loss of generality as argued in [18]. More importantly however, the achievable region of [18, Theorem 2] corresponds to one wherein Receiver 3 decodes the common message M_{123} non-uniquely via V_3, whereas in the specialization of Theorem 4, Receiver 3 decodes the common message uniquely via U_{123}. This results in more inequalities in the region of Theorem 4 (when specialized to $K = 3$ and $\mathsf{E} = \{1,12,123\}$) than the ones in [18, Theorem 2]. However, it is left to the reader to verify along the lines of [18, Proposition 8] that unique and non-unique decoding at Receiver 3 result in the same region per coding distribution. In effect, the additional inequalities due to unique-decoding are redundant. In particular, [18, Proposition 8] shows that the achievable region of [18, Theorem 2] for the three degraded message set in which Receiver 2 decodes the common message uniquely and Receiver 3 decodes the common message non-uniquely, when specialized to the two degraded message set $\mathsf{E} = \{1,123\}$ (by setting $R_{12} = 0$), is equivalent to the achievable rate region of [18, Proposition 5] obtained by having both Receivers 2 and 3 decode the common message non-uniquely.

V. Conclusion

In this paper, we propose an inner bound based on rate-splitting and superposition coding in explicit form for the K-user DM BC as a union of four-dimensional polytopes described in terms of the rates of the four messages. By specializing it to the combination network through the choice of a single coding distribution we obtain a single polytope which is then shown to be the capacity region of that network for the diamond message set. Our converse proof uses the generalized cut-set bound framework of [6] to recognize, by developing

[5]There is a typographical error in the last inequality in the rate region given in [18, Theorem 2]: the term $I(U_{12};Y_2|U_{123})$ in the bound should instead be $I(U_{12},U_{123};Y_2)$.

key connections between our order-theoretic description of the inner bound and the set-theoretic description of the generalized cut-set bounds, the inequalities in the inner bound as indeed being generalized cut-set outer bounds. Most significantly, our interpretation of this capacity result is that it suggests a certain strength of the coding scheme for the *general* DM BC itself, from whence it was obtained. Put differently, our top-down approach allows us to conclude that rate-splitting and superposition coding (with subset inclusion order) results in a strong inner bound for the DM BC with the diamond message set as opposed to presenting the more conventional view that rate-splitting and superposition coding may not be strong enough to attain the capacity of the entire DM BC for the diamond message set. Nevertheless, we extend the rate-splitting and superposition coding inner bound by adding binning to it and obtain a rate region that is again given as a union of explicit four-dimensional polytopes in terms of the four message rates.

Moreover, we expect that the proposed top-town approach, initiated in [2] and continued in [15] and the present work, can be used to obtain inner bounds for the DM BC for message sets other than the diamond message set, and validate them as being strong, provided that, when specialized to the combination network, they establish its capacity region.

Appendix I
Achievability Proof for Theorem 1

First, up-set rate splitting is used, i.e., each message M_S is divided into a collection of sub-messages $M_{S\to S'}$ where $S' \in\uparrow_{\mathsf{E}} S$. Then, the sub-message $M_{S\to S'}$ will be treated as if it was intended for the receivers in the larger set S' instead of in S. Hence, the messages $M_{\overline{K-1.K}}, M_{\overline{K}}$ and $M_{\overline{K-1}}$ are split into $(M_{\overline{K-1.K}\to\overline{\phi}}, M_{\overline{K-1.K}\to\overline{K}}, M_{\overline{K-1.K}\to\overline{K-1}}, M_{\overline{K-1.K}\to\overline{K-1.K}})$, $(M_{\overline{K}\to\overline{\phi}}, M_{\overline{K}\to\overline{K}})$, $(M_{\overline{K-1}\to\overline{\phi}}, M_{\overline{K-1}\to\overline{K-1}})$, with rates $(R_{\overline{K-1.K}\to\overline{\phi}}, R_{\overline{K-1.K}\to\overline{K}}, R_{\overline{K-1.K}\to\overline{K-1}}, R_{\overline{K-1.K}\to\overline{K-1.K}})$, $(R_{\overline{K}\to\overline{\phi}}, R_{\overline{K}\to\overline{K}})$, and $(R_{\overline{K-1}\to\overline{\phi}}, R_{\overline{K-1}\to\overline{K-1}})$, respectively. The common message and a sub-message of each of the other three messages of the form $M_{S\to\bar{\phi}}$, that is, the reconstructed message $(M_{\overline{\phi}}, M_{\overline{K-1}\to\overline{\phi}}, M_{\overline{K}\to\overline{\phi}}, M_{\overline{K-1.K}\to\overline{\phi}})$ is represented by the cloud center $U_{\overline{\phi}}$. Conditionally independently (conditioned on $U_{\overline{\phi}}$), the reconstructed message pair $(M_{\overline{K-1}\to\overline{K-1}}, M_{\overline{K-1.K}\to\overline{K-1}})$ is represented by $U_{\overline{K-1}}$ and $(M_{\overline{K}\to\overline{K}}, M_{\overline{K-1.K}\to\overline{K}})$ is represented by $U_{\overline{K}}$. Finally, $M_{\overline{K-1.K}\to\overline{K-1.K}}$ is represented by $U_{\overline{K-1.K}} = X$. Receivers $\{Y_i\}_{i=1}^{K-2}$ decode the four intended messages by the joint unique decoding of X (and hence $U_{\overline{\phi}}, U_{\overline{K}}, U_{\overline{K-1}}, U_{\overline{K-1.K}}$), and this happens successfully as long as

$$R_{\overline{\phi}} + R_{\overline{K-1}} + R_{\overline{K}} + R_{\overline{K-1.K}} \leq I(X;Y_j) \tag{110}$$

$$R_{\overline{K-1}} + R_{\overline{K}} + R_{\overline{K-1.K}} - R_{\overline{K-1}\to\overline{\phi}} - R_{\overline{K}\to\overline{\phi}} - R_{\overline{K-1.K}\to\overline{\phi}} \leq I(X;Y_j|U_{\overline{\phi}}) \tag{111}$$

$$R_{\overline{K}} + R_{\overline{K-1.K}} - R_{\overline{K}\to\overline{\phi}} - R_{\overline{K-1.K}\to\overline{K-1}} - R_{\overline{K-1.K}\to\overline{\phi}} \leq I(X;Y_j|U_{\overline{\phi}}, U_{\overline{K-1}}) \tag{112}$$

$$R_{\overline{K-1}} + R_{\overline{K-1.K}} - R_{\overline{K-1}\to\overline{\phi}} - R_{\overline{K-1.K}\to\overline{K}} - R_{\overline{K-1.K}\to\overline{\phi}} \leq I(X;Y_j|U_{\overline{\phi}}, U_{\overline{K}}) \tag{113}$$

$$R_{\overline{K-1.K}} - R_{\overline{K-1.K}\to\overline{K-1}} - R_{\overline{K-1.K}\to\overline{K}} - R_{\overline{K-1.K}\to\overline{\phi}} \leq I(X;Y_j|U_{\overline{\phi}}, U_{\overline{K-1}}, U_{\overline{K}}) \tag{114}$$

for all $j \in \{1, 2, \cdots, K-2\}$.

On the other hand, receiver Y_{K-1} finds the two intended messages $M_{\overline{\phi}}$ and $M_{\overline{K}}$ by decoding $(U_{\overline{\phi}}, U_{\overline{K}})$, and receiver Y_K find the two intended messages $M_{\overline{\phi}}$ and $M_{\overline{K-1}}$ by decoding $(U_{\overline{\phi}}, U_{\overline{K-1}})$. Receivers Y_{K-1} and Y_K decode their intended pair of messages successfully provided the following inequalities hold:

$$R_{\overline{\phi}} + R_{\overline{K}} + R_{\overline{K-1}\to\overline{\phi}} + R_{\overline{K-1.K}\to\overline{K}} + R_{\overline{K-1.K}\to\overline{\phi}} \leq I(U_{\overline{\phi}}, U_{\overline{K}}; Y_{K-1}) \tag{115}$$

$$R_{\overline{K}} - R_{\overline{K}\to\overline{\phi}} + R_{\overline{K-1.K}\to\overline{K}} \leq I(U_{\overline{K}}; Y_{K-1}|U_{\overline{\phi}}) \tag{116}$$

$$R_{\overline{\phi}} + R_{\overline{K-1}} + R_{\overline{K}\to\overline{\phi}} + R_{\overline{K-1.K}\to\overline{K-1}} + R_{\overline{K-1.K}\to\overline{\phi}} \leq I(U_{\overline{\phi}}, U_{\overline{K-1}}; Y_K) \tag{117}$$

$$R_{\overline{K-1}} - R_{\overline{K-1}\to\overline{\phi}} + R_{\overline{K-1.K}\to\overline{K-1}} \leq I(U_{\overline{K-1}}; Y_K|U_{\overline{\phi}}) \tag{118}$$

Besides inequalities (110)-(118) there are eight inequalities to account for the non-negativity of the split rates

$$-R_{\overline{K-1.K}} + R_{\overline{K-1.K}\to\overline{K-1}} + R_{\overline{K-1.K}\to\overline{K}} + R_{\overline{K-1.K}\to\overline{\phi}} \leq 0 \tag{119}$$

$$-R_{\overline{K-1.K}\to\overline{K-1}} \leq 0 \tag{120}$$

$$-R_{\overline{K-1.K}\to\overline{K}} \leq 0 \tag{121}$$

$$-R_{\overline{K-1.K}\to\overline{\phi}} \leq 0 \tag{122}$$

$$-R_{\overline{K}} + R_{\overline{K}\to\overline{\phi}} \leq 0 \tag{123}$$

$$-R_{\overline{K}\to\overline{\phi}} \leq 0 \tag{124}$$

$$-R_{\overline{K-1}} + R_{\overline{K-1}\to\overline{\phi}} \leq 0 \tag{125}$$

$$-R_{\overline{K-1}\to\overline{\phi}} \leq 0 \tag{126}$$

Taken together, these inequalities describe a nine-dimensional polytope (per admissible distribution of (19)) in original- and split-rates space. By projecting away the five split rates using a structured form of the FME method described in Appendix II to follow, we get the achievable region in the positive orthant defined by the inequalities (7)-(18).

Appendix II
The FME for Theorem 1

As described in Appendix I, the coding scheme includes rate-splitting and since we consider arbitrary K receivers, there are an indeterminate number of inequalities that describe the polyhedral achievable rate region per random coding distribution in nine-dimensional original- and split-rate space. In particular, there are $5K + 2$ inequalities given by (110)–(118) and (119)-(126) that describe each such polytope. This means that the FME procedure must deal with groups of inequalities (for general K) at a time instead of exhaustively listing them (for this reason, the FME software of [19] which is designed to take in a definite number of inequalities in exhaustive form cannot be directly used). Moreover, the elimination of the five split rates presents $5! = 120$ possibilities for the orders in which to eliminate them, all of which must lead, in principle, to the right four-dimensional polytope described by the inequalities (7)-(18). It turns out however, that at least some of these orders render the FME procedure too tedious to perform by hand beyond even the second or third steps of the five-step process, including the identification (and hence, elimination) of the redundant inequalities that arise at each step. In this appendix, we provide a "good" order of elimination and the intermediate results obtained after projecting out each of the five split rates in that order that not only makes it possible to complete the five-step task with reasonable effort by hand, but also offers the possibility of verifying if two of the successively reduced-dimensional polytopes at the end of steps 2 and 4 satisfy certain symmetry conditions which, intuitively, they must (and which they do). The outline of the FME procedure is given next.

In particular, we propose to do the projection in five steps in the following order:

1) $R_{\overline{K-1}\to\overline{\phi}}$
2) $R_{\overline{K}\to\overline{\phi}}$
3) $R_{\overline{K-1.K}\to\overline{K}}$
4) $R_{\overline{K-1.K}\to\overline{K-1}}$
5) $R_{\overline{K-1.K}\to\overline{\phi}}$

Note that this order maintains a certain symmetric structure in the problem, in that after projecting away the first two split rates in the above list we expect a symmetry in the description of the resulting seven-dimensional polytope around K and $K-1$ (i.e., exchanging K and $K-1$ should not change it). Similarly, such a symmetry should result again after projecting away the first four split rates in the above list, and eventually after projecting away the fifth split rate that yields the four-dimensional polytopes specified in Theorem 1, which is indeed symmetric in the sense described above.

At the end of each projection to eliminate a split rate, we must remove all redundant inequalities so that the FME procedure remains tractable in future steps. We do not address the issue of why inequalities that are not retained at the end of each step are redundant, leaving this task at each step (beyond step 2, when they arise) to the interested reader. Judicious applications of the data processing inequality and the Markov chain relationships between the involved random variables are all that are needed. The task of identifying which inequalities are redundant is, however, solved in that these are precisely the inequalities that we omit from including in the projected reduced-dimensional region given at the end of each step.

In the first two steps, we project away the split rates $R_{\overline{K-1}\to\overline{\phi}}$ and $R_{\overline{K}\to\overline{\phi}}$ and obtain, after rearranging the inequalities, the following inequalities with $j \in \{1, 2, \cdots, K-2\}$

$$R_{\overline{\phi}} + R_{\overline{K-1}} + R_{\overline{K}} + R_{\overline{K-1.K}} \leq I(X;Y_j|U_{\overline{\phi}},U_{\overline{K}}) + I(U_{\overline{\phi}},U_{\overline{K}};Y_{K-1}) \tag{127}$$

$$R_{\overline{\phi}} + R_{\overline{K-1}} + R_{\overline{K}} + R_{\overline{K-1.K}} \leq I(X;Y_j|U_{\overline{\phi}},U_{\overline{K-1}}) + I(U_{\overline{\phi}},U_{\overline{K-1}};Y_K) \tag{128}$$

$$R_{\overline{\phi}} + R_{\overline{K-1}} + R_{\overline{K}} + R_{\overline{K-1.K}} \leq I(X;Y_j) \tag{129}$$

$$R_{\overline{K-1.K}} - R_{\overline{K-1.K}\to\overline{\phi}} \leq I(X;Y_j|U_{\overline{\phi}}) \tag{130}$$

$$R_{\overline{K-1.K}} - R_{\overline{K-1.K}\to\overline{\phi}} - R_{\overline{K-1.K}\to\overline{K-1}} \leq I(X;Y_j|U_{\overline{\phi}},U_{\overline{K-1}}) \tag{131}$$

$$R_{\overline{K-1.K}} - R_{\overline{K-1.K}\to\overline{\phi}} - R_{\overline{K-1.K}\to\overline{K}} \leq I(X;Y_j|U_{\overline{\phi}},U_{\overline{K}}) \tag{132}$$

$$R_{\overline{K-1.K}\to\overline{K-1}} \leq I(U_{\overline{K-1}};Y_K|U_{\overline{\phi}}) \tag{133}$$

$$R_{\overline{K-1.K}\to\overline{K}} \leq I(U_{\overline{K}};Y_{K-1}|U_{\overline{\phi}}) \tag{134}$$

$$R_{\overline{\phi}} + R_{\overline{K-1}} + R_{\overline{K-1.K}\to\overline{\phi}} + R_{\overline{K-1.K}\to\overline{K-1}} \leq I(U_{\overline{\phi}},U_{\overline{K-1}};Y_K) \tag{135}$$

$$R_{\overline{\phi}} + R_{\overline{K}} + R_{\overline{K-1.K}\to\overline{\phi}} + R_{\overline{K-1.K}\to\overline{K}} \leq I(U_{\overline{\phi}},U_{\overline{K}};Y_{K-1}) \tag{136}$$

$$R_{\overline{\phi}} + R_{\overline{K-1}} + R_{\overline{K}} + R_{\overline{K-1.K}} + R_{\overline{K-1.K}\to\overline{K-1}} \leq I(X;Y_j|U_{\overline{\phi}}) + I(U_{\overline{\phi}},U_{\overline{K-1}};Y_K) \tag{137}$$

$$R_{\overline{\phi}} + R_{\overline{K-1}} + R_{\overline{K}} + R_{\overline{K-1.K}} + R_{\overline{K-1.K}\to\overline{K}} \leq I(X;Y_j|U_{\overline{\phi}}) + I(U_{\overline{\phi}},U_{\overline{K}};Y_{K-1}) \tag{138}$$

$$R_{\overline{\phi}} + R_{\overline{K-1}} + R_{\overline{K}} + R_{\overline{K-1.K}\to\overline{\phi}} + R_{\overline{K-1.K}\to\overline{K-1}} + R_{\overline{K-1.K}\to\overline{K}} \leq I(U_{\overline{K}};Y_{K-1}|U_{\overline{\phi}}) + I(U_{\overline{\phi}},U_{\overline{K-1}};Y_K) \tag{139}$$

$$R_{\overline{\phi}} + R_{\overline{K-1}} + R_{\overline{K}} + R_{\overline{K-1.K}\to\overline{\phi}} + R_{\overline{K-1.K}\to\overline{K-1}} + R_{\overline{K-1.K}\to\overline{K}} \leq I(U_{\overline{K-1}};Y_K|U_{\overline{\phi}}) + I(U_{\overline{\phi}},U_{\overline{K}};Y_{K-1}) \tag{140}$$

$$R_{\overline{K-1.K}} - R_{\overline{K-1.K}\to\overline{\phi}} - R_{\overline{K-1.K}\to\overline{K-1}} - R_{\overline{K-1.K}\to\overline{K}} \leq I(X;Y_j|U_{\overline{\phi}},U_{\overline{K-1}},U_{\overline{K}}) \tag{141}$$

$$2R_{\overline{\phi}} + 2R_{\overline{K-1}} + 2R_{\overline{K}} + R_{\overline{K-1.K}} + R_{\overline{K-1.K}\to\overline{\phi}} + R_{\overline{K-1.K}\to\overline{K-1}} + R_{\overline{K-1.K}\to\overline{K}} \leq I(X;Y_j|U_{\overline{\phi}}) + I(U_{\overline{\phi}},U_{\overline{K}};Y_{K-1}) + I(U_{\overline{\phi}},U_{\overline{K-1}};Y_K) \tag{142}$$

$$-R_{\overline{K-1.K}} + R_{\overline{K-1.K}\to\overline{K-1}} + R_{\overline{K-1.K}\to\overline{K}} + R_{\overline{K-1.K}\to\overline{\phi}} \leq 0 \tag{143}$$

$$-R_{\overline{K-1.K}\to\overline{K-1}} \leq 0 \tag{144}$$

$$-R_{\overline{K-1.K}\to\overline{K}} \leq 0 \tag{145}$$

$$-R_{\overline{K-1.K}\to\overline{\phi}} \leq 0 \tag{146}$$

Observe that the above seven-dimensional polytope remains unchanged if $K-1$ and K are exchanged, as one might expect. Indeed, observe this symmetry in every group of inequalities separated by a new line. There were no redundant inequalities in the first two steps.

In the third step, we project away the split rate $R_{\overline{K-1.K}\to\overline{K}}$. Notice now that we have a group of inequalities in (141) which give lower bounds for $R_{\overline{K-1.K}\to\overline{K}}$ and a group of inequalities in (138) and (142) that give upper bounds for $R_{\overline{K-1.K}\to\overline{K}}$. Hence, to project the split rate $R_{\overline{K-1.K}\to\overline{K}}$ from these, we have to introduce new variables $j_1, j_2 \in \{1, 2, \cdots, K-2\}$ since we get a multiplicative increase in the number of inequalities. By projecting away the split rate $R_{\overline{K-1.K}\to\overline{K}}$, we obtain, by rearranging the inequalities and getting rid of the redundant ones [6], the following for $j, j_1, j_2 \in \{1, 2, \cdots, K-2\}$

$$\begin{aligned}&R_{\overline{\phi}} + R_{\overline{K-1}} + R_{\overline{K}} + R_{\overline{K-1.K}}\\&\quad\le I(X;Y_j|U_{\overline{\phi}},U_{\overline{K}}) + I(U_{\overline{\phi}},U_{\overline{K}};Y_{K-1})\end{aligned}\tag{147}$$

$$\begin{aligned}&R_{\overline{\phi}} + R_{\overline{K-1}} + R_{\overline{K}} + R_{\overline{K-1.K}}\\&\quad\le I(X;Y_j|U_{\overline{\phi}},U_{\overline{K-1}}) + I(U_{\overline{\phi}},U_{\overline{K-1}};Y_K)\end{aligned}\tag{148}$$

$$R_{\overline{\phi}} + R_{\overline{K-1}} + R_{\overline{K}} + R_{\overline{K-1.K}} \le I(X;Y_j)\tag{149}$$

$$\begin{aligned}&R_{\overline{\phi}} + R_{\overline{K-1}} + R_{\overline{K}} + R_{\overline{K-1.K}}\\&\quad\le I(X;Y_j|U_{\overline{\phi}},U_{\overline{K-1}},U_{\overline{K}})\\&\quad+ I(U_{\overline{K}};Y_{K-1}|U_{\overline{\phi}}) + I(U_{\overline{\phi}},U_{\overline{K-1}};Y_K)\end{aligned}\tag{150}$$

$$\begin{aligned}&R_{\overline{\phi}} + R_{\overline{K-1}} + R_{\overline{K}} + R_{\overline{K-1.K}}\\&\quad\le I(X;Y_j|U_{\overline{\phi}},U_{\overline{K-1}},U_{\overline{K}})\\&\quad+ I(U_{\overline{K-1}};Y_K|U_{\overline{\phi}}) + I(U_{\overline{\phi}},U_{\overline{K}};Y_{K-1})\end{aligned}\tag{151}$$

$$\begin{aligned}&2R_{\overline{\phi}} + 2R_{\overline{K-1}} + 2R_{\overline{K}} + 2R_{\overline{K-1.K}}\\&\quad\le I(X;Y_{j_1}|U_{\overline{\phi}}) + I(X;Y_{j_2}|U_{\overline{\phi}},U_{\overline{K-1}},U_{\overline{K}})\\&\quad+ I(U_{\overline{\phi}},U_{\overline{K}};Y_{K-1}) + I(U_{\overline{\phi}},U_{\overline{K-1}};Y_K)\end{aligned}\tag{152}$$

$$R_{\overline{K-1.K}} - R_{\overline{K-1.K}\to\overline{\phi}} \le I(X;Y_j|U_{\overline{\phi}})\tag{153}$$

$$\begin{aligned}&R_{\overline{K-1.K}} - R_{\overline{K-1.K}\to\overline{\phi}}\\&\quad\le I(X;Y_j|U_{\overline{\phi}},U_{\overline{K}}) + I(U_{\overline{K}};Y_{K-1}|U_{\overline{\phi}})\end{aligned}\tag{154}$$

$$\begin{aligned}&R_{\overline{K-1.K}} - R_{\overline{K-1.K}\to\overline{\phi}} - R_{\overline{K-1.K}\to\overline{K-1}}\\&\quad\le I(X;Y_j|U_{\overline{\phi}},U_{\overline{K-1}})\end{aligned}\tag{155}$$

$$\begin{aligned}&R_{\overline{K-1.K}} - R_{\overline{K-1.K}\to\overline{\phi}} - R_{\overline{K-1.K}\to\overline{K-1}}\\&\quad\le I(X;Y_j|U_{\overline{\phi}},U_{\overline{K-1}},U_{\overline{K}}) + I(U_{\overline{K}};Y_{K-1}|U_{\overline{\phi}})\end{aligned}\tag{156}$$

$$\begin{aligned}&R_{\overline{\phi}} + R_{\overline{K}} + R_{\overline{K-1.K}} - R_{\overline{K-1.K}\to\overline{K-1}}\\&\quad\le I(X;Y_j|U_{\overline{\phi}},U_{\overline{K-1}},U_{\overline{K}}) + I(U_{\overline{\phi}},U_{\overline{K}};Y_{K-1})\end{aligned}\tag{157}$$

$$\begin{aligned}&R_{\overline{\phi}} + R_{\overline{K-1}} + R_{\overline{K}} + 2R_{\overline{K-1.K}} - R_{\overline{K-1.K}\to\overline{\phi}} -\\&\quad R_{\overline{K-1.K}\to\overline{K-1}} \le I(X;Y_{j_1}|U_{\overline{\phi}}) +\\&\quad I(X;Y_{j_2}|U_{\overline{\phi}},U_{\overline{K-1}},U_{\overline{K}}) + I(U_{\overline{\phi}},U_{\overline{K}};Y_{K-1})\end{aligned}\tag{158}$$

$$R_{\overline{K-1.K}\to\overline{K-1}} \le I(X;Y_j|U_{\overline{\phi}},U_{\overline{K-1}})\tag{159}$$

$$R_{\overline{K-1.K}\to\overline{K-1}} \le I(U_{\overline{K-1}};Y_K|U_{\overline{\phi}})\tag{160}$$

$$R_{\overline{\phi}} + R_{\overline{K}} + R_{\overline{K-1.K}\to\overline{\phi}} \le I(U_{\overline{\phi}},U_{\overline{K}};Y_{K-1})\tag{161}$$

$$\begin{aligned}&R_{\overline{\phi}} + R_{\overline{K-1}} + R_{\overline{K}} + R_{\overline{K-1.K}} + R_{\overline{K-1.K}\to\overline{K-1}}\\&\quad\le I(X;Y_j|U_{\overline{\phi}}) + I(U_{\overline{\phi}},U_{\overline{K-1}};Y_K)\end{aligned}\tag{162}$$

$$\begin{aligned}&R_{\overline{\phi}} + R_{\overline{K-1}} + R_{\overline{K}} + R_{\overline{K-1.K}} + R_{\overline{K-1.K}\to\overline{K-1}}\\&\quad\le I(X;Y_j|U_{\overline{\phi}},U_{\overline{K}})\\&\quad+ I(U_{\overline{K}};Y_{K-1}|U_{\overline{\phi}}) + I(U_{\overline{\phi}},U_{\overline{K-1}};Y_K)\end{aligned}\tag{163}$$

$$\begin{aligned}&R_{\overline{\phi}} + R_{\overline{K-1}} + R_{\overline{K-1.K}\to\overline{\phi}} + R_{\overline{K-1.K}\to\overline{K-1}}\\&\quad\le I(U_{\overline{\phi}},U_{\overline{K-1}};Y_K)\end{aligned}\tag{164}$$

$$\begin{aligned}&R_{\overline{\phi}} + R_{\overline{K-1}} + R_{\overline{K}} + R_{\overline{K-1.K}\to\overline{\phi}} + R_{\overline{K-1.K}\to\overline{K-1}}\\&\quad\le I(U_{\overline{K}};Y_{K-1}|U_{\overline{\phi}}) + I(U_{\overline{\phi}},U_{\overline{K-1}};Y_K)\end{aligned}\tag{165}$$

$$\begin{aligned}&R_{\overline{\phi}} + R_{\overline{K-1}} + R_{\overline{K}} + R_{\overline{K-1.K}\to\overline{\phi}} + R_{\overline{K-1.K}\to\overline{K-1}}\\&\quad\le I(U_{\overline{K-1}};Y_K|U_{\overline{\phi}}) + I(U_{\overline{\phi}},U_{\overline{K}};Y_{K-1})\end{aligned}\tag{166}$$

$$\begin{aligned}&2R_{\overline{\phi}} + 2R_{\overline{K-1}} + 2R_{\overline{K}} + R_{\overline{K-1.K}} + R_{\overline{K-1.K}\to\overline{\phi}}\\&\quad+ R_{\overline{K-1.K}\to\overline{K-1}} \le I(X;Y_j|U_{\overline{\phi}})\\&\quad+ I(U_{\overline{\phi}},U_{\overline{K}};Y_{K-1}) + I(U_{\overline{\phi}},U_{\overline{K-1}};Y_K)\end{aligned}\tag{167}$$

$$-R_{\overline{K-1.K}} + R_{\overline{K-1.K}\to\overline{K-1}} + R_{\overline{K-1.K}\to\overline{\phi}} \le 0\tag{168}$$

$$-R_{\overline{K-1.K}\to\overline{K-1}} \le 0\tag{169}$$

$$-R_{\overline{K-1.K}\to\overline{\phi}} \le 0\tag{170}$$

In the fourth step, we project away the split rate $R_{\overline{K-1.K}\to\overline{K-1}}$ from the polytope described by the inequalities (147)-(170). Note that the inequalities (155)-(158) and (169) give lower bounds and the inequalities (159)-(160), and (162)-(168) give upper bounds on $R_{\overline{K-1.K}\to\overline{K-1}}$. Hence, we have K^2-K-1 lower bounds and $4K-3$ upper bounds with the rest of the inequalities remaining unchanged.

Projecting away the split rate $R_{\overline{K-1.K}\to\overline{K-1}}$ from the aforementioned inequalities, in which there are five lower bounds and nine upper bounds, counting a group of inequalities as one inequality, we obtain 45 groups of inequalities (including groups of size one). However, and we leave it to the reader to show this, the nine inequalities, besides the unchanged inequalities, shown next, are the only ones that are not redundant. Hence, we get the five-dimensional polytope described by the following inequalities for all $j, j_1, j_2 \in \{1, 2, \cdots, K-2\}$

$$\begin{aligned}&R_{\overline{\phi}} + R_{\overline{K-1}} + R_{\overline{K}} + R_{\overline{K-1.K}}\\&\quad\le I(X;Y_j|U_{\overline{\phi}},U_{\overline{K}}) + I(U_{\overline{\phi}},U_{\overline{K}};Y_{K-1})\end{aligned}\tag{171}$$

$$\begin{aligned}&R_{\overline{\phi}} + R_{\overline{K-1}} + R_{\overline{K}} + R_{\overline{K-1.K}}\\&\quad\le I(X;Y_j|U_{\overline{\phi}},U_{\overline{K-1}}) + I(U_{\overline{\phi}},U_{\overline{K-1}};Y_K)\end{aligned}\tag{172}$$

$$R_{\overline{\phi}} + R_{\overline{K-1}} + R_{\overline{K}} + R_{\overline{K-1.K}} \le I(X;Y_j)\tag{173}$$

$$\begin{aligned}&R_{\overline{\phi}} + R_{\overline{K-1}} + R_{\overline{K}} + R_{\overline{K-1.K}}\\&\quad\le I(X;Y_j|U_{\overline{\phi}},U_{\overline{K-1}},U_{\overline{K}})\\&\quad+ I(U_{\overline{K}};Y_{K-1}|U_{\overline{\phi}}) + I(U_{\overline{\phi}},U_{\overline{K-1}};Y_K)\end{aligned}\tag{174}$$

$$\begin{aligned}&R_{\overline{\phi}} + R_{\overline{K-1}} + R_{\overline{K}} + R_{\overline{K-1.K}}\\&\quad\le I(X;Y_j|U_{\overline{\phi}},U_{\overline{K-1}},U_{\overline{K}})\\&\quad+ I(U_{\overline{K-1}};Y_K|U_{\overline{\phi}}) + I(U_{\overline{\phi}},U_{\overline{K}};Y_{K-1})\end{aligned}\tag{175}$$

$$\begin{aligned}&2R_{\overline{\phi}} + 2R_{\overline{K-1}} + 2R_{\overline{K}} + 2R_{\overline{K-1.K}}\\&\quad\le I(X;Y_{j_1}|U_{\overline{\phi}}) + I(X;Y_{j_2}|U_{\overline{\phi}},U_{\overline{K-1}},U_{\overline{K}})\\&\quad+ I(U_{\overline{\phi}},U_{\overline{K}};Y_{K-1}) + I(U_{\overline{\phi}},U_{\overline{K-1}};Y_K)\end{aligned}\tag{176}$$

$$R_{\overline{\phi}} + R_{\overline{K-1}} + R_{\overline{K-1.K}\to\overline{\phi}} \le I(U_{\overline{\phi}},U_{\overline{K-1}};Y_K)\tag{177}$$

$$R_{\overline{\phi}} + R_{\overline{K-1}} + R_{\overline{K-1.K}\to\overline{\phi}} \le I(U_{\overline{\phi}},U_{\overline{K}};Y_{K-1})\tag{178}$$

$$\begin{aligned}&R_{\overline{\phi}} + R_{\overline{K-1}} + R_{\overline{K}} + R_{\overline{K-1.K}\to\overline{\phi}}\\&\quad\le I(U_{\overline{K}};Y_{K-1}|U_{\overline{\phi}}) + I(U_{\overline{\phi}},U_{\overline{K-1}};Y_K)\end{aligned}\tag{179}$$

[6]Note that we get redundant inequalities by eliminating the split rate $R_{\overline{K-1.K}\to\overline{K}}$ from (132) and (136), (138), (140), and (142), and also from (145) and (138).

$$R_{\overline{\phi}} + R_{\overline{K-1}} + R_{\overline{K}} + R_{\overline{K-1.K} \to \overline{\phi}} \leq I(U_{\overline{K-1}}; Y_K | U_{\overline{\phi}}) + I(U_{\overline{\phi}}, U_{\overline{K}}; Y_{K-1}) \quad (180)$$

$$2R_{\overline{\phi}} + R_{\overline{K-1}} + R_{\overline{K}} + R_{\overline{K-1.K}} + R_{\overline{K-1.K} \to \overline{\phi}} \leq I(X; Y_j | U_{\overline{\phi}}, U_{\overline{K-1}}, U_{\overline{K}}) + I(U_{\overline{\phi}}, U_{\overline{K}}; Y_{K-1}) + I(U_{\overline{\phi}}, U_{\overline{K-1}}; Y_K) \quad (181)$$

$$2R_{\overline{\phi}} + 2R_{\overline{K-1}} + 2R_{\overline{K}} + R_{\overline{K-1.K}} + R_{\overline{K-1.K} \to \overline{\phi}} \leq I(X; Y_j | U_{\overline{\phi}}) + I(U_{\overline{\phi}}, U_{\overline{K}}; Y_{K-1}) + I(U_{\overline{\phi}}, U_{\overline{K-1}}; Y_K) \quad (182)$$

$$R_{\overline{K-1.K}} - R_{\overline{K-1.K} \to \overline{\phi}} \leq I(X; Y_j | U_{\overline{\phi}}) \quad (183)$$

$$R_{\overline{K-1.K}} - R_{\overline{K-1.K} \to \overline{\phi}} \leq I(X; Y_{j_1} | U_{\overline{\phi}}, U_{\overline{K-1}}) + I(X; Y_{j_2} | U_{\overline{\phi}}, U_{\overline{K}}) \quad (184)$$

$$R_{\overline{K-1.K}} - R_{\overline{K-1.K} \to \overline{\phi}} \leq I(X; Y_j | U_{\overline{\phi}}, U_{\overline{K-1}}) + I(U_{\overline{K-1}}; Y_K | U_{\overline{\phi}}) \quad (185)$$

$$R_{\overline{K-1.K}} - R_{\overline{K-1.K} \to \overline{\phi}} \leq I(X; Y_j | U_{\overline{\phi}}, U_{\overline{K}}) + I(U_{\overline{K}}; Y_{K-1} | U_{\overline{\phi}}) \quad (186)$$

$$R_{\overline{K-1.K}} - R_{\overline{K-1.K} \to \overline{\phi}} \leq I(X; Y_j | U_{\overline{\phi}}, U_{\overline{K-1}}, U_{\overline{K}}) + I(U_{\overline{K}}; Y_{K-1} | U_{\overline{\phi}}) + I(U_{\overline{K-1}}; Y_K | U_{\overline{\phi}}) \quad (187)$$

$$- R_{\overline{K-1.K}} + R_{\overline{K-1.K} \to \overline{\phi}} \leq 0 \quad (188)$$

$$- R_{\overline{K-1.K} \to \overline{\phi}} \leq 0 \quad (189)$$

Observe again the symmetry in the above five-dimensional polytope, which is that it remains unchanged when $K-1$ and K are exchanged (even within every group of inequalities separated by new line), as expected.

Finally, in the fifth step, we project away the split rate $R_{\overline{K-1.K} \to \overline{\phi}}$ from the five-dimensional polytope described by the inequalities (171)-(189). After doing this projection, and eliminating redundant inequalities, it is left to the reader to verify that we obtain the four-dimensional polytope in the original rates given in Theorem 1. Note again the presence of symmetry of that polytope when $K-1$ and K are exchanged in that four-dimensional polytope. This concludes the outline of the FME procedure.

Appendix III
Proof of Theorem 2

The achievability proof follows by adopting a special case of the up-set rate-splitting scheme used to prove Theorem 1. In particular, we set the split rates as follows

$$R_{\overline{K-1.K} \to \overline{\phi}} = 0 \quad (190)$$

$$R_{\overline{K-1.K} \to \overline{K}} = 0 \quad (191)$$

$$R_{\overline{K-1.K} \to \overline{K-1}} = 0 \quad (192)$$

$$R_{\overline{K-1.K} \to \overline{K-1.K}} = R_{\overline{K-1.K}} \quad (193)$$

$$R_{\overline{K} \to \overline{\phi}} = 0 \quad (194)$$

$$R_{\overline{K} \to \overline{K}} = R_{\overline{K}} \quad (195)$$

$$R_{\overline{K-1} \to \overline{\phi}} = R_{\overline{K-1}} \quad (196)$$

$$R_{\overline{K-1} \to \overline{K-1}} = 0 \quad (197)$$

For these split rates, the messages $(M_{\overline{\phi}}, M_{\overline{K-1}})$ with rate $R_{\overline{K-1}} + R_{\overline{\phi}}$ are represented by $U_{\overline{\phi}}$, the cloud center codebook. Then the message $M_{\overline{K}}$ is represented by the first satellite codebook $U_{\overline{K}}$ (superposed on $U_{\overline{\phi}}$), and following that $M_{\overline{K-1.K}}$ is represented by the second satellite X codebook which is superposed on $U_{\overline{K}}$. Note that by doing so, we set $U_{\overline{K-1}} = $ *const.* By using the less noisy conditions $Y_i \succ Y_{K-1} \succ Y_K$ for each $i \in \{1 : K-2\}$, we get the stated achievable region.

The converse proof uses the information inequality of [14, Lemma 1] which we state it here for easy reference.

Lemma 1: Let $X \multimap (Y, Z)$ be a DM BC without feedback and $Y \succeq Z$. Consider M to be any random variable such that $M \multimap X^n \multimap (Y^n, Z^n)$ forms a Markov chain. Then,

$$I(Y^{i-1}; Z_i | M) \geq I(Z^{i-1}; Z_i | M) \ 1 \leq i \leq n$$
$$I(Y^{i-1}; Y_i | M) \geq I(Z^{i-1}; Y_i | M) \ 1 \leq i \leq n$$

For achievable rates tuple $\{R_S\}_{S \in \mathsf{E}}$, we show that there exist random variables $U_{\overline{\phi}}, U_{\overline{K}}, X$, where $(U_{\overline{\phi}} \multimap U_{\overline{K}} \multimap X \multimap (Y_1 \cdots Y_K)$ forms a Markov chain, such that inequalities in (20)-(22) holds for $Y_i \succeq Y_{K-1} \succeq Y_K$ for each $i \in \{1 : K-2\}$.

The optimal auxiliary random variables are $U_{\overline{\phi},i} = Y_{K-1,1}^{i-1}, M_{\overline{\phi}}, M_{\overline{K-1}}$ and $U_{\overline{K},i} = Y_{K-1,1}^{i-1}, M_{\overline{\phi}}, M_{\overline{K-1}}, M_{\overline{K}}$. The converse proof for (20)-(21) is similar to that in [20, Theorem 1]. Hence, the proof is omitted and left to the reader.

The proof that (22) is an outer bound is shown as follows: we have each $j \in [1 : K-2]$

$$nR_{\overline{K-1.K}} = H(M_{\overline{K-1.K}}) \leq I(M_{\overline{K-1.K}}; Y_{j,1}^n | M_{\overline{\phi}}, M_{\overline{K-1}}, M_{\overline{K}}) + n\epsilon_n \quad (198)$$

$$\leq \sum_{i=1}^n I(X_i; Y_{j,i} | M_{\overline{\phi}}, M_{\overline{K-1}}, M_{\overline{K}}, Y_{j,1}^{i-1}) + n\epsilon_n \quad (199)$$

$$= \sum_{i=1}^n \left\{ I(X_i; Y_{j,i} | M_{\overline{\phi}}, M_{\overline{K-1}}, M_{\overline{K}}) - I(Y_{j,1}^{i-1}; Y_{j,i} | M_{\overline{\phi}}, M_{\overline{K-1}}, M_{\overline{K}}) \right\} + n\epsilon_n \quad (200)$$

$$\leq \sum_{i=1}^n \left\{ I(X_i; Y_{j,i} | M_{\overline{\phi}}, M_{\overline{K-1}}, M_{\overline{K}}) - I(Y_{K-1,1}^{i-1}; Y_{j,i} | M_{\overline{\phi}}, M_{\overline{K-1}}, M_{\overline{K}}) \right\} + n\epsilon_n \quad (201)$$

$$= \sum_{i=1}^n I(X_i; Y_{j,i} | M_{\overline{\phi}}, M_{\overline{K-1}}, M_{\overline{K}}, Y_{K-1,1}^{i-1}) + n\epsilon_n \quad (202)$$

$$= \sum_{i=1}^n I(X_i; Y_{j,i} | U_{\overline{K},i}) + n\epsilon_n$$

where (198) follows from Fano's inequality and the independence among all messages. Moreover, (199), (200) and (202) follow from the Markov chain

$$(M_{\overline{\phi}}, M_{\overline{K-1}}, M_{\overline{K}}, Y_{j,1}^{i-1}, Y_{K-1,1}^{i-1}) \multimap X_i \multimap Y_{j,i} \quad (203)$$

and (201) follows from Lemma 1.

The rest of the proof proceeds along standard lines, i.e., we define a time-sharing uniform random variable Q that takes the values $\{1, 2, 3, \cdots, n\}$ and independent of all other involved random variables. Identify $U_{\overline{\phi}}=(U_{\overline{\phi},Q}, Q)$, $U_{\overline{K}}=(U_{\overline{K},Q}, Q)$, $X = X_Q$ and $Y_l=Y_{lQ}$ for all $l \in \{1, 2, \cdots, K\}$. It is clear that this choice of auxiliary random variables implies that $U_{\overline{\phi}} \multimap U_{\overline{K}} \multimap X$. Finally, take the limit as $n \to \infty$, so that $\epsilon_n \to 0$, to get (20)-(22). This completes the converse proof.

APPENDIX IV
PROOF OF THEOREM 4

The achievable scheme of Theorem 1 that is based on rate-splitting and superposition coding is enhanced by adding to it the technique of binning that involves generating the $U_{\overline{K}}$ and $U_{\overline{K-1}}$ codebooks at excess rate and creating bins of codewords in place of codewords for the associated messages and selecting a pair of codewords from the product bins selected by the messages that are jointly typical per the allowed joint distribution of $U_{\overline{K}}$ and $U_{\overline{K-1}}$ (conditioned on $U_{\overline{\phi}}$). In particular, the messages $M_{\overline{K}}$, $M_{\overline{K-1}}$, and $M_{\overline{K-1.K}}$ are split as in the proof of Theorem 1. The reconstructed message $(M_{\overline{\phi}}, M_{\overline{K-1}\to\overline{\phi}}, M_{\overline{K}\to\overline{\phi}}, M_{\overline{K-1.K}\to\overline{\phi}})$ is represented by the cloud center $U_{\overline{\phi}}$ again as in Theorem 1. For each $u^n_{\overline{\phi}}(m_{\overline{\phi}}, m_{\overline{K-1}\to\overline{\phi}}, m_{\overline{K}\to\overline{\phi}}, m_{\overline{K-1.K}\to\overline{\phi}})$ randomly and independently generate (a) $2^{n\tilde{R}_{\overline{K}}}$ independent sequences $u^n_{\overline{K}}(m_{\overline{\phi}}, m_{\overline{K-1}\to\overline{\phi}}, m_{\overline{K}\to\overline{\phi}}, m_{\overline{K-1.K}\to\overline{\phi}}, \tilde{m}_{\overline{K}})$, for each $\tilde{m}_{\overline{K}} \in [1 : 2^{n\tilde{R}_{\overline{K}}}]$ with

$$\tilde{R}_{\overline{K}} \geq R_{\overline{K-1.K}\to\overline{K}} + R_{\overline{K}\to\overline{K}} \tag{204}$$

and (b) $2^{n\tilde{R}_{\overline{K-1}}}$ independent sequences $u^n_{\overline{K-1}}(m_{\overline{\phi}}, m_{\overline{K-1}\to\overline{\phi}}, m_{\overline{K}\to\overline{\phi}}, m_{\overline{K-1.K}\to\overline{\phi}}, \tilde{m}_{\overline{K-1}})$ for each $\tilde{m}_{\overline{K-1}} \in [1 : 2^{n\tilde{R}_{\overline{K-1}}}]$ and

$$\tilde{R}_{\overline{K-1}} \geq R_{\overline{K-1.K}\to\overline{K-1}} + R_{\overline{K-1}\to\overline{K-1}} \tag{205}$$

Partition the $2^{n\tilde{R}_{\overline{K}}}$ sequences $u^n_{\overline{K}}$ into $2^{n(R_{\overline{K-1.K}\to\overline{K}}+R_{\overline{K}\to\overline{K}})}$ equi-sized subcodebooks (or bins) and the $2^{n\tilde{R}_{\overline{K-1}}}$ sequences $u^n_{\overline{K-1}}$ into $2^{n(R_{\overline{K-1.K}\to\overline{K-1}}+R_{\overline{K-1}\to\overline{K-1}})}$ equi-sized subcodebooks, with the subcodebooks indexed by the sub-message pairs $(m_{\overline{K-1.K}\to\overline{K}}, m_{\overline{K}\to\overline{K}})$ and $(m_{\overline{K-1.K}\to\overline{K-1}}, m_{\overline{K-1}\to\overline{K-1}})$, respectively. The tuple $(m_{\overline{\phi}}$, $m_{\overline{K-1}\to\overline{\phi}}$, $m_{\overline{K}\to\overline{\phi}}$, $m_{\overline{K-1.K}\to\overline{\phi}}$, $m_{\overline{K-1.K}\to\overline{K}}$, $m_{\overline{K}\to\overline{K}}$, $m_{\overline{K-1.K}\to\overline{K-1}}$, $m_{\overline{K-1}\to\overline{K-1}})$ identifies two subcodebooks of $u^n_{\overline{K}}$ and $u^n_{\overline{K-1}}$ sequences respectively from which a jointly typical $(u^n_{\overline{K}}, u^n_{\overline{K-1}})$ pair is selected. For each such pair, generate $2^{nR_{\overline{K-1.K}\to\overline{K-1.K}}}$ independent x^n sequences indexed by the sub-message $m_{\overline{K-1.K}\to\overline{K-1.K}}$. To ensure that each product bin contains a jointly typical pair with arbitrarily high probability, we require by the mutual covering lemma [17] that

$$\begin{aligned} R_{\overline{K-1.K}\to\overline{K}} \quad &+R_{\overline{K}\to\overline{K}} + R_{\overline{K-1.K}\to\overline{K-1}} + R_{\overline{K-1}\to\overline{K-1}} \\ &\leq \tilde{R}_{\overline{K}} + \tilde{R}_{\overline{K-1}} - I(U_{\overline{K}}; U_{\overline{K-1}}|U_{\overline{\phi}}) \end{aligned} \tag{206}$$

Receivers $\{Y_j\}_{j=1}^{K-2}$ decode all four intended messages by the joint unique decoding of X (and hence of $U_{\overline{\phi}}, U_{\overline{K}}, U_{\overline{K-1}}, U_{\overline{K-1.K}}$), and this happens successfully as long as (110)-(114) hold. On the other hand, receiver Y_{K-1} finds its two intended messages $M_{\overline{\phi}}$ and $M_{\overline{K}}$ by jointly uniquely decoding $(U_{\overline{\phi}}, U_{\overline{K}})$, and similarly, receiver Y_K find its two intended messages $M_{\overline{\phi}}$ and $M_{\overline{K-1}}$ by jointly uniquely decoding $(U_{\overline{\phi}}, U_{\overline{K-1}})$. Receivers Y_{K-1} and Y_K decode their intended pair of messages successfully provided the following inequalities hold:

$$\begin{aligned} R_{\overline{\phi}} + R_{\overline{K}\to\overline{\phi}} + R_{\overline{K-1}\to\overline{\phi}} + R_{\overline{K-1.K}\to\overline{\phi}} + \tilde{R}_{\overline{K}} \\ \leq I(U_{\overline{\phi}}, U_{\overline{K}}; Y_{K-1}) \end{aligned} \tag{207}$$

$$\tilde{R}_{\overline{K}} \leq I(U_{\overline{K}}; Y_{K-1}|U_{\overline{\phi}}) \tag{208}$$

$$\begin{aligned} R_{\overline{\phi}} + R_{\overline{K}\to\overline{\phi}} + R_{\overline{K-1}\to\overline{\phi}} + R_{\overline{K-1.K}\to\overline{\phi}} + \tilde{R}_{\overline{K-1}} \\ \leq I(U_{\overline{\phi}}, U_{\overline{K-1}}; Y_K) \end{aligned} \tag{209}$$

$$\tilde{R}_{\overline{K-1}} \leq I(U_{\overline{K-1}}; Y_K|U_{\overline{\phi}}) \tag{210}$$

At this point, inequalities (110)-(114), (204)- (210) and the non-negativity of message rate constraints define a 11-dimensional polytope. Extending the FME method as outlined in Appendix II for the projection required in Theorem 1 — the details of which are left to the reader —we obtain the rate region given in the statement of Theorem 4.

REFERENCES

[1] M. Salman and M. K. Varanasi, "Diamond message set groupcasting: From an achievable rate region for the DM broadcast channel to the capacity of the combination network," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2020, pp. 1498–1503.
[2] H. Romero and M. Varanasi, "Superposition coding in the combination network," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2016, pp. 2149–2153.
[3] H. P. Romero and M. K. Varanasi, "A unifying order-theoretic framework for superposition coding: Polymatroidal structure and optimality in the multiple-access channel with general message sets," *IEEE Trans. Inf. Theory*, vol. 63, no. 1, pp. 21–37, Jan. 2017.
[4] A. Schrijver, *Theory of Linear and Integer Programming*. Hoboken, NJ, USA: Wiley, 1998.
[5] C. K. Ngai and R. W. Yeung, "Network coding gain of combination networks," in *Proc. IEEE Inf. Theory Workshop*, Oct. 2004, pp. 283–287.
[6] A. Salimi, T. Liu, and S. Cui, "Generalized cut-set bounds for broadcast networks," *IEEE Trans. Inf. Theory*, vol. 61, no. 6, pp. 2983–2996, Jun. 2015.
[7] L. Grokop and D. N. C. Tse, "Fundamental constraints on multicast capacity regions," 2008, *arXiv:0809.2835*.
[8] C. Tian, "Latent capacity region: A case study on symmetric broadcast with common messages," *IEEE Trans. Inf. Theory*, vol. 57, no. 6, pp. 3273–3285, Jun. 2011.
[9] S. S. Bidokhti, V. M. Prabhakaran, and S. N. Diggavi, "Capacity results for multicasting nested message sets over combination networks," *IEEE Trans. Inf. Theory*, vol. 62, no. 9, pp. 4968–4992, Sep. 2016.
[10] H. P. Romero and M. K. Varanasi, "Rate splitting and superposition coding for concurrent groupcasting over the broadcast channel: A general framework," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2017, pp. 1888–1892.
[11] M. Salman and M. K. Varanasi, "An achievable rate region for the K-receiver two nested groupcast DM broadcast channel and a capacity result for the combination network," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2018, pp. 1415–1419.
[12] A. Schrijver, *Theory of Linear and Integer Programming*. Hoboken, NJ, USA: Wiley, 1998.
[13] J. Korner and K. Marton, "A source network problem involving the comparison of two channels II," *IEEE Trans. Inf. Theory*, pp. 411–423, 1975.
[14] C. Nair and Z. V. Wang, "The capacity region of the three receiver less noisy broadcast channel," *IEEE Trans. Inf. Theory*, vol. 57, no. 7, pp. 4058–4062, Jul. 2011.

[15] M. Salman and M. K. Varanasi, "The K-user DM broadcast channel with two groupcast messages: Achievable rate regions and the combination network as a case study," 2018, *arXiv:1812.10525*.

[16] Y. Liang and G. Kramer, "Rate regions for relay broadcast channels," *IEEE Trans. Inf. Theory*, vol. 53, no. 10, pp. 3517–3535, Oct. 2007.

[17] A. El Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge, U.K.: Cambridge Univ. Press, 2011.

[18] C. Nair and A. El Gamal, "The capacity region of a class of three-receiver broadcast channels with degraded message sets," *IEEE Trans. Inf. Theory*, vol. 55, no. 10, pp. 4479–4493, Oct. 2009.

[19] I. B. Gattegno, Z. Goldfeld, and H. H. Permuter, "Fourier-motzkin elimination software for information theoretic inequalities," 2016, *arXiv:1610.03990*.

[20] M. Salman and M. K. Varanasi, "Capacity results for classes of partially ordered K-user broadcast channels with two nested multicast messages," *IEEE Trans. Inf. Theory*, vol. 66, no. 1, pp. 65–81, Aug. 2019.

Mohamed Salman received the bachelor's degree (Hons.) in electrical engineering from Alexandria University, Alexandria, Egypt, in 2012, the master's degree in wireless technologies from Nile University, Giza, Egypt, in 2015, and the Ph.D. degree from the Department of Electrical, Computer, and Energy Engineering, University of Colorado Boulder, Boulder, CO, USA, in 2020. Since March 2020, he has been working as a Modem Systems Engineer at Qualcomm Inc., Boulder. His research interests include information theory and wireless communications.

Mahesh K. Varanasi (Fellow, IEEE) received the Ph.D. degree in electrical engineering from Rice University, Houston, TX, USA, in 1989.

Since 1989, he has been a member of the Faculty of the University of Colorado, Boulder, CO, USA, where he is currently a Professor of electrical, computer, and energy engineering. He has been a Professor (by courtesy) with the Department of Mathematics since 2010 and an affiliated Faculty Member of the Department of Applied Mathematics since 2013, both at the University of Colorado Boulder. His research and teaching interests are in the areas of networks information theory, wireless communications and coding, machine learning, and signal processing.

Dr. Varanasi is a Highly Cited Researcher according to the ISI Web of Science. He was the recipient of the 2017–2019 and 2021 Qualcomm Faculty Awards. He was a General Chair of the 2018 IEEE International Symposium on Information Theory (ISIT), Vail, Colorado. He served as an Editor for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS from 2007 to 2009. He was an Associate Editor of the IEEE TRANSACTIONS ON INFORMATION THEORY from 2014 to 2017.