

General Relayless Networks: Representation of the Capacity Region

Anelia Somekh-Baruch and Sergio Verdú
 Department of Electrical Engineering, Princeton University
 Engineering Quadrangle, Princeton, NJ 08544, USA
 Emails: {anelia,verdu}@princeton.edu

Abstract—Using an Information-Spectrum approach, a limiting expression for the capacity region of a general network without relays combined of M transmitters observing K input messages, and L receivers, is found. This general setup accounts for the broadcast channel with common messages, the general interference channel, and the multiple access channel as special cases. It is demonstrated how the limiting expression can be used to yield a single-letter tight outer bound for the case of a two-user stationary memoryless degraded broadcast channel.

I. INTRODUCTION

Multi-terminal channels, in their various forms, have been the subject of many studies in information theory. In spite of the intensive research efforts, many problems in this area, relating even to the simplest models, remain open. Two such examples are the determination of the capacity region of the memoryless stationary broadcast and interference channels. The problem may lie in the fact that deriving a single-letter characterization for these capacity regions is an extremely complicated task, but it may as well be that no single-letter expressions exist for these channels. The latter possibility, combined with the desire to analyze a generalized network model, having components which are not necessarily stationary and ergodic are what motivated this work. In the absence of single-letter expressions, perhaps other types of expressions can shed some light on general multi-user channels.

A number of limiting expressions characterizing capacity regions of various channels exist in the information theory literature. Shannon, in his 1961 paper [1] on the two-way channel, was the first to present a capacity region in a non-single-letter form. Van der Meulen [2] derived the limiting capacity region for the interference channel, whose special case, the Multiple Access Channel (MAC), appeared also in [3].

In addition to the possibility that even for the general relayless memoryless multiuser channel single-letter expressions simply do not exist, the importance of deriving limiting expressions lies in their potential to be explicitly solved.

The memoryless MAC provides an example where single-letter expressions can be obtained from the limiting expressions [4], [5]. Stationary Gaussian dispersive channels are another example where limiting expressions lead to computable characterizations (e.g. [6]). In addition, limiting expressions can be lower and upper bounded by non-limiting multi-letter

expressions (see [7] in the case of single user channels, and for multiple access channels).

Even in their complex form, there may also be engineering insights that can be inferred from limiting expressions. For example, in the MAC case, the limiting expressions coincide with a convex hull of a union of pentagons. This suggests an optimal encoding-decoding scheme: encoding that is done by time sharing of single user codes (capacity achieving for a channel having no-interference) followed by a successive decoding receiver. A similar intuition can be applied to degraded broadcast channels. From the Van der Meulen [2] limiting expression applied to the Gaussian interference channel and from [4], one learns that there are capacity-achieving encoders (perhaps the only ones) that do not use single-user codes. The decoder structure suggested by the limiting capacity region expressions is single-user decoding (as opposed to successive decoding) in which the decoder does not neglect the structure of the interference, rather it is optimum with respect to the equivalent noise seen at each receiver (sum of Gaussian noise and the interfering signals). With those decoders the empirical distributions of good codes [9] are neither Gaussian nor memoryless (see [4]).

A general single-user channel is defined as a sequence of conditional probability distributions $\{P_{Y^n|X^n}\}_{n=1}^{\infty}$, where $P_{Y^n|X^n}$ has input alphabet \mathcal{X}^n and output alphabet \mathcal{Y}^n . It was shown by Verdú and Han in [10] that for the case of finite alphabet channel, the capacity is given by

$$\sup_{\mathbf{X}} \underline{I}(\mathbf{X}; \mathbf{Y}) \quad (1)$$

where $\underline{I}(\mathbf{X}; \mathbf{Y})$, referred to as the *inf-information rate between \mathbf{X} and \mathbf{Y}* , stands for the liminf in probability of the sequence of normalized information densities $(1/n)i_{X^n Y^n}(X^n; Y^n)$ where $i_{X^n Y^n}(a^n; b^n) = \log \frac{P_{Y^n|X^n}(b^n|a^n)}{P_{Y^n}(b^n)}$. The results of [10] were extended in [5] (see also [8]) where a general formula for the capacity region of a multiple access channel was derived. Recently, a general characterization in terms of information-spectrum for the capacity region of the general broadcast channel was found in [11].

This paper deals with a more general topology of a communication network than the setup investigated in [5], [10]-[11].

The topology we consider consists of M transmitters and L receivers. Each of the transmitters observes a specified subset of K independent messages, and at each of the L receivers, a

specified subset of the messages is to be decoded. The capacity region of such a network is the set of K rate-tuples that can be reliably decoded by the intended receivers. We provide a formula for the capacity region in terms of inf-information rates between the channel outputs and sequences of auxiliary random vectors. The capacity region of the general two-user interference channel is derived as a special case, generalizing [3], which provided the expression for a discrete memoryless interference channel. Furthermore, an alternative expression to the capacity region of the general MAC that was established in [5] is recovered as a special case. Additionally, a limiting expression for the capacity region of the general two-user broadcast channel with a common message is presented as a special case of the general formula.

The rest of the paper is organized as follows. In Section II we present a formal description of our network model. Our main result, a coding theorem for a network, will be stated in Section III. In the two Sections, IV and V, we particularize the result of Theorem 1 to the special cases of the interference channel and the two-user broadcast channel with a common message, respectively. We conclude with Section VI where we demonstrate how a tight single-letter outer bound on the capacity region of the two-user stationary memoryless degraded broadcast channel can be derived from Theorem 1.

II. NETWORK MODEL

Let $W_{(1,n)}, \dots, W_{(K,n)}$ be independent random variables, designating the messages to be transmitted, where $W_{(k,n)}, k = 1, \dots, K$ is uniformly distributed over the set $\{1, \dots, M_n^{(k)}\}$ for some positive integer $M_n^{(k)}$.

A network is composed of $M < \infty$ transmitters and $L < \infty$ receivers. Each of the transmitters observes a subset of the messages. Let $T(m)$ be the subset of $\{1, \dots, K\}$ containing the indices of the messages observed by the m -th transmitter. The output of the m -th transmitter, $X_{(m)}^n$, takes values in some alphabet $\mathcal{X}_{(m)}^n$. Hence, the m -th transmitter is a function

$$\varphi_n^{(m)} : \prod_{k \in T(m)} \{1, \dots, M_n^{(k)}\} \rightarrow \mathcal{X}_{(m)}^n.$$

Denote $\mathbf{X}_{(m)} = \{X_{(m)}^n\}_{n=1}^\infty$. We shall say that two processes $\mathbf{V}_{(1)} = \{V_{(1)}^n\}_{n=1}^\infty$, $\mathbf{V}_{(2)} = \{V_{(2)}^n\}_{n=1}^\infty$ are independent if for all n the random variables $V_{(1)}^n, V_{(2)}^n$ are independent, and we shall say that they are independent given the process $\mathbf{V}_{(3)}$ if for all n the random variables $V_{(1)}^n, V_{(2)}^n$ are independent given $V_{(3)}^n$.

An M -input L -output channel is an arbitrary sequence of n -dimensional conditional distributions from the finite input alphabet¹, $(\mathcal{X}_{(1)}^n, \dots, \mathcal{X}_{(M)}^n)$, to the finite output alphabet $(\mathcal{Y}_{(1)}^n, \dots, \mathcal{Y}_{(L)}^n)$, i.e.,

$$P_{\mathbf{Y}_{(1)}, \dots, \mathbf{Y}_{(L)} | \mathbf{X}_{(1)}, \dots, \mathbf{X}_{(M)}} = \left\{ P_{\mathbf{Y}_{(1)}, \dots, \mathbf{Y}_{(L)} | X_{(1)}^n, \dots, X_{(M)}^n} \right\}_{n=1}^\infty,$$

¹The results are extendable to infinite alphabets.

where $Y_{(\ell)}^n$ is the output of the channel $P_{Y_{(1)}^n, \dots, Y_{(L)}^n | X_{(1)}^n, \dots, X_{(M)}^n}$ observed by the ℓ -th receiver, and $\mathbf{Y}_{(\ell)} = \{Y_{(\ell)}^n\}_{n=1}^\infty$.

Let $D(\ell)$ stand for the set of messages indices that are to be reliably decoded by the ℓ -th receiver. The ℓ -th receiver is thus a function

$$\psi_n^{(\ell)} : \mathcal{Y}_{(\ell)}^n \rightarrow \prod_{k \in D(\ell)} \{1, \dots, M_n^{(k)}\}.$$

Hence, the network, is defined by $P_{\mathbf{Y}_{(1)}, \dots, \mathbf{Y}_{(L)} | \mathbf{X}_{(1)}, \dots, \mathbf{X}_{(M)}}$, $\{T(m)\}_{m=1}^M$, and $\{D(\ell)\}_{\ell=1}^L$.

Given a K -tuple of integers $(M_n^{(1)}, \dots, M_n^{(K)})$, we say that the encoders $(\varphi_n^{(1)}, \dots, \varphi_n^{(M)})$, and the decoders $(\psi_n^{(1)}, \dots, \psi_n^{(L)})$, form an $(n, M_n^{(1)}, \dots, M_n^{(K)}, \epsilon_n)$ -code if

$$\Pr \left\{ \exists \ell : \psi_n^{(\ell)}(Y_{(\ell)}^n) \neq \{W_{(k,n)}\}_{k \in D(\ell)} \right\} \leq \epsilon_n, \quad (2)$$

where the probability law is given by $\prod_{k=1}^K P_{W_{(k,n)}} \prod_{\ell=1}^L P_{Y_{(\ell)}^n | X_{(1)}^n, \dots, X_{(M)}^n}$ with

$$X_{(m)}^n = \varphi_n^{(m)}(\{W_{(k,n)}\}_{k \in T(m)}).$$

Definition 1: A rate K -tuple, (R_1, \dots, R_K) is said to be achievable if there exists an $(n, M_n^{(1)}, \dots, M_n^{(K)}, \epsilon_n)$ -code satisfying $\lim_{n \rightarrow \infty} \epsilon_n = 0$, and $\liminf \frac{1}{n} \log M_n^{(k)} \geq R_k$ for all $k = 1, \dots, K$.

Definition 2: The capacity region is defined as the union of all achievable rate K -tuples, (R_1, \dots, R_K) , such that $R_k \geq 0$ $k = 1, \dots, K$.

Let $\mathcal{R}(P_{\mathbf{U}_{(1)}, \dots, \mathbf{U}_{(K)}, \mathbf{X}_{(1)}, \dots, \mathbf{X}_{(M)}})$ be the union of all non-negative rate K -tuples (R_1, \dots, R_K) satisfying

$$R_k \leq \min_{\ell: k \in D(\ell)} \underline{I}(\mathbf{U}_{(k)}; \mathbf{Y}_{(\ell)}) \quad \forall k \in \{1, \dots, K\} \quad (3)$$

where $\{\mathbf{U}_{(k)}\}_{k=1}^K$ are sequences of auxiliary random variables $\{U_{(1)}^n, \dots, U_{(K)}^n\}_{n=1}^\infty$ with $U_{(k)}^n$ taking values in some alphabet $\mathcal{U}_{(k)}^n$, whose cardinality is upper bounded by

$$|\mathcal{U}_{(k)}^n| \leq \min_{\ell: k \in D(\ell)} \prod_{m: k \in T(m)} |\mathcal{X}_{(m)}^n|, \quad (4)$$

and the joint law defining the quantity $\underline{I}(\mathbf{U}_{(k)}; \mathbf{Y}_{(\ell)})$ in (3) is the marginal distribution of $(\mathbf{U}_{(k)}, \mathbf{Y}_{(\ell)})$ induced by $\left\{ P_{U_{(1)}^n, \dots, U_{(K)}^n, X_{(1)}^n, \dots, X_{(M)}^n} \times P_{Y_{(1)}^n, \dots, Y_{(L)}^n | X_{(1)}^n, \dots, X_{(M)}^n} \right\}_{n=1}^\infty$.

III. CODING THEOREM

Theorem 1: The capacity region of the channel $P_{\mathbf{Y}_{(1)}, \dots, \mathbf{Y}_{(L)} | \mathbf{X}_{(1)}, \dots, \mathbf{X}_{(M)}}$, with $\{T(m)\}_{m=1}^M$ and $\{D(\ell)\}_{\ell=1}^L$ is given by

$$\cup \mathcal{R}(P_{\mathbf{U}_{(1)}, \dots, \mathbf{U}_{(K)}, \mathbf{X}_{(1)}, \dots, \mathbf{X}_{(M)}}), \quad (5)$$

where the union is over distributions $P_{\mathbf{U}_{(1)}, \dots, \mathbf{U}_{(K)}, \mathbf{X}_{(1)}, \dots, \mathbf{X}_{(M)}}$ of the form

$$P_{\mathbf{U}_{(1)}, \dots, \mathbf{U}_{(K)}, \mathbf{X}_{(1)}, \dots, \mathbf{X}_{(M)}} = \prod_{k=1}^K P_{\mathbf{U}_{(k)}} \prod_{m=1}^M P_{\mathbf{X}_{(1)} | \{\mathbf{U}_{(k)}\}_{k \in T(m)}}. \quad (6)$$

The proof of Theorem 1 is a natural extension of [10] and [5], and it is based on the following two lemmas, the first lemma is a generalization of Feinstein's Lemma which serves the basis for the direct part, and the second lemma is a generalization of the Verdú-Han Lemma [10] that establishes the converse part.

Lemma 1: Let $(\mathbf{U}_{(1)}, \dots, \mathbf{U}_{(K)}, \mathbf{X}_{(1)}, \dots, \mathbf{X}_{(M)}) = \{U_{(1)}^n, \dots, U_{(K)}^n, X_{(1)}^n, \dots, X_{(M)}^n\}_{n=1}^{\infty}$ be an arbitrary sequence of random variables such that (6) holds.

For every K -tuple of integers $M_n^{(1)}, \dots, M_n^{(K)}$ there exists an $(n, M_n^{(1)}, \dots, M_n^{(K)}, \epsilon_n)$ -code satisfying

$$\epsilon_n \leq \Pr \left\{ \bigcup_{k=1}^K \left\{ \min_{\ell: k \in D(\ell)} \frac{1}{n} i_{U_{(k)}^n, Y_{(\ell)}^n} (U_{(k)}^n; Y_{(\ell)}^n) \leq \frac{1}{n} \log M_n^{(k)} + \gamma \right\} \right\} + L \cdot K \cdot e^{-n\gamma}, \quad (7)$$

for all $n = 1, 2, \dots$ and $\gamma > 0$.

Lemma 2: Let $P_{\mathbf{Y}_{(1)}, \dots, \mathbf{Y}_{(L)} | \mathbf{X}_{(1)}, \dots, \mathbf{X}_{(M)}}$, $\{T(m)\}_{m=1}^M$ and $\{D(\ell)\}_{\ell=1}^L$ be given. Let $(U_{(1)}^n, \dots, U_{(K)}^n)$ designate the independent messages aimed for the L receivers, let $\{X_{(m)}^n\}_{m=1}^M$ stand for the corresponding channel inputs, and let $\{Y_{(\ell)}^n\}_{\ell=1}^L$ designate the channel outputs. Then for all $\gamma > 0$, every $(n, M_n^{(1)}, \dots, M_n^{(K)}, \epsilon_n)$ -code must satisfy

$$\epsilon_n \geq \Pr \left\{ \bigcup_{k=1}^K \left\{ \min_{\ell: k \in D(\ell)} \frac{1}{n} i_{U_{(k)}^n, Y_{(\ell)}^n} (U_{(k)}^n; Y_{(\ell)}^n) \leq \frac{1}{n} \log M_n^{(k)} - \gamma \right\} \right\} - L \cdot K \cdot e^{-n\gamma}, \quad (8)$$

for all $n = 1, 2, \dots$

Remarks:

- It is possible to derive an alternative equivalent representation of the capacity region (5) in the spirit of [11]. For the sake of brevity, we omit the details.
- Since in Lemma 2, which serves as the core of the converse part of Theorem 1, the random variables $\{U_{(k)}^n\}_{k \in T(m)}$ represent the independent transmitted messages observed by the m -th transmitter, and since the encoder maps $\{U_{(k)}^n\}_{k \in T(m)}$ to $\mathcal{X}_{(m)}^n$ it is clear that the cardinality of $U_{(k)}^n$ can be upper bounded as in (4) by considering a clean channel, i.e., $Y_{(\ell)}^n = (X_{(1)}^n, \dots, X_{(M)}^n)$. Thus, the existence of auxiliary random variables satisfying both (8) and (4) is assured.
- Note that bounding the cardinality of auxiliary random variables is important to guarantee computationally feasible algorithms for single-letter characterizations of capacity regions, and is potential useful in the context of limiting expressions if they turn out to be computable.

IV. A GENERAL REPRESENTATION FOR THE CAPACITY REGION OF THE INTERFERENCE CHANNEL

Let $P_{\mathbf{Y}_{(1)}, \mathbf{Y}_{(2)} | \mathbf{X}_{(1)}, \mathbf{X}_{(2)}} = \{P_{\mathbf{Y}_{(1)}, \mathbf{Y}_{(2)} | \mathbf{X}_{(1)}^n, \mathbf{X}_{(2)}^n}\}_{n=1}^{\infty}$ be an interference channel having two inputs, $\mathbf{X}_{(1)}^n \in \mathcal{X}^n, \mathbf{X}_{(2)}^n \in \tilde{\mathcal{X}}^n$ and two outputs $\mathbf{Y}_{(1)}^n \in \mathcal{Y}^n, \mathbf{Y}_{(2)}^n \in \tilde{\mathcal{Y}}^n$. The first encoder maps a message index (destined to the first decoder) to a channel input, i.e., it is defined by the mapping $\varphi_n^{(1)} : \{1, \dots, e^{nR_1}\} \rightarrow \mathcal{X}^n$. Similarly, the second encoder maps a message (destined to the second decoder) to a channel input using the mapping $\varphi_n^{(2)} : \{1, \dots, e^{nR_2}\} \rightarrow \tilde{\mathcal{X}}^n$. The two decoders are defined by the mappings $\psi_n^{(1)} : \mathcal{Y}^n \rightarrow \{1, \dots, e^{nR_1}\}$ and $\psi_n^{(2)} : \tilde{\mathcal{Y}}^n \rightarrow \{1, \dots, e^{nR_2}\}$.

This is a special case of the general setup with

$$K = M = L = 2, \\ T(1) = 1, T(2) = 2, D(1) = 1, D(2) = 2. \quad (9)$$

Let $\mathcal{R}_{IC}(P_{\mathbf{X}_{(1)}, \mathbf{X}_{(2)}})$ be the union of all non-negative rate pairs (R_1, R_2) satisfying

$$R_1 \leq \underline{I}(\mathbf{X}_{(1)}; \mathbf{Y}_{(1)}) \quad (10)$$

$$R_2 \leq \underline{I}(\mathbf{X}_{(2)}; \mathbf{Y}_{(2)}), \quad (11)$$

where the joint law defining the above quantities is given by $P_{\mathbf{X}_{(1)}, \mathbf{X}_{(2)}} \times P_{\mathbf{Y}_{(1)} | \mathbf{X}_{(1)}, \mathbf{X}_{(2)}} \times P_{\mathbf{Y}_{(2)} | \mathbf{X}_{(1)}, \mathbf{X}_{(2)}}$.

Corollary 1: The capacity region of the interference channel $P_{\mathbf{Y}_{(1)}, \mathbf{Y}_{(2)} | \mathbf{X}_{(1)}, \mathbf{X}_{(2)}}$ is given by

$$\cup \mathcal{R}_{IC}(P_{\mathbf{X}_1, \mathbf{X}_2}), \quad (12)$$

where the union is over the measures $P_{\mathbf{X}_{(1)}, \mathbf{X}_{(2)}}$ such that $\mathbf{X}_{(1)}$ and $\mathbf{X}_{(2)}$ are independent.

Proof: The general formula for the capacity region provided in Theorem 1, in this case, is given by the union of non-negative rate-pairs (R_1, R_2) , satisfying

$$R_1 \leq \underline{I}(\mathbf{U}_{(1)}; \mathbf{Y}_{(1)}) \quad (13)$$

$$R_2 \leq \underline{I}(\mathbf{U}_{(2)}; \mathbf{Y}_{(2)}), \quad (14)$$

for some distribution

$$P_{\mathbf{U}_{(1)}, \mathbf{U}_{(2)}, \mathbf{X}_{(1)}, \mathbf{X}_{(2)}} = P_{\mathbf{U}_{(1)}} P_{\mathbf{U}_{(2)}} P_{\mathbf{X}_{(1)} | \mathbf{U}_{(1)}} P_{\mathbf{X}_{(2)} | \mathbf{U}_{(2)}}.$$

From the data processing Theorem [10, Theorem 9] we have for $i = 1, 2$, $\underline{I}(\mathbf{U}_{(i)}; \mathbf{Y}_{(i)}) \leq \underline{I}(\mathbf{X}_{(i)}; \mathbf{Y}_{(i)})$, with equality whenever $\mathbf{U}_{(i)} = \mathbf{X}_{(i)}$, $i = 1, 2$. Hence, taking $\mathbf{U}_{(i)} = \mathbf{X}_{(i)}$ in (13)-(14) would yield an achievable outer bound on the capacity region, which is given in (12). ■

Discussion

As the MAC can be regarded as a special case of the IC (having two equal outputs, i.e., $\mathbf{Y}_{(1)} = \mathbf{Y}_{(2)}$) we can deduce the capacity region of the general MAC from Corollary 1. Let the MAC be given by $P_{\mathbf{Y} | \mathbf{X}_{(1)}, \mathbf{X}_{(2)}} = \{P_{\mathbf{Y}^n | \mathbf{X}_{(1)}^n, \mathbf{X}_{(2)}^n}\}_{n=1}^{\infty}$. The capacity region of the MAC is given by the union of all non-negative rate pairs (R_1, R_2) satisfying

$$R_1 \leq \underline{I}(\mathbf{X}_{(1)}; \mathbf{Y}) \quad (15)$$

$$R_2 \leq \underline{I}(\mathbf{X}_{(2)}; \mathbf{Y}), \quad (16)$$

for some sequence of distributions

$$P_{\mathbf{X}_{(1)}\mathbf{X}_{(2)}} = P_{\mathbf{X}_{(1)}}P_{\mathbf{X}_{(2)}} = \left\{ P_{X_{(1)}^n} \times P_{X_{(2)}^n} \right\}_{n=1}^{\infty}$$

where $P_{X_{(i)}^n}$, $i = 1, 2$ is some distribution defined on the set $\mathcal{X}_{(i)}^n$. The above expression is in the spirit of [3] and [2] which treated the stationary discrete memoryless MAC.

The expression for the capacity region of the MAC given in [5] (which generalized [12]), is expressed as the union of all non-negative rate pairs (R_1, R_2) satisfying

$$R_1 \leq \underline{I}(\mathbf{X}_{(1)}; \mathbf{Y} | \mathbf{X}_{(2)}) \quad (17)$$

$$R_2 \leq \underline{I}(\mathbf{X}_{(2)}; \mathbf{Y} | \mathbf{X}_{(1)}), \quad (18)$$

$$R_1 + R_2 \leq \underline{I}(\mathbf{X}_{(1)}, \mathbf{X}_{(2)}; \mathbf{Y}) \quad (19)$$

for some sequence of distributions

$$P_{\mathbf{X}_{(1)}\mathbf{X}_{(2)}} = P_{\mathbf{X}_{(1)}}P_{\mathbf{X}_{(2)}} = \left\{ P_{X_{(1)}^n} \times P_{X_{(2)}^n} \right\}_{n=1}^{\infty}$$

where $P_{X_{(i)}^n}$, $i = 1, 2$ is some distribution defined on the set $\mathcal{X}_{(i)}^n$.

Obviously, the two expressions should coincide, so, in fact, these are two alternative expressions for the capacity region of the general MAC: a union of rectangles (15)-(16), and a union of pentagons (17)-(19).

Note that

$$\begin{aligned} \underline{I}(\mathbf{X}_{(1)}; \mathbf{Y}) &\leq \underline{I}(\mathbf{X}_{(1)}; \mathbf{X}_{(2)}, \mathbf{Y}) = \underline{I}(\mathbf{X}_{(1)}; \mathbf{Y} | \mathbf{X}_{(2)}) \\ \underline{I}(\mathbf{X}_{(2)}; \mathbf{Y}) &\leq \underline{I}(\mathbf{X}_{(2)}; \mathbf{X}_{(1)}, \mathbf{Y}) = \underline{I}(\mathbf{X}_{(2)}; \mathbf{Y} | \mathbf{X}_{(1)}) \end{aligned} \quad (20)$$

where the inequalities follow from [10, Theorem 8-f)], and the equalities hold since $\mathbf{X}_{(1)}, \mathbf{X}_{(2)}$ are independent. Moreover,

$$\begin{aligned} &\underline{I}(\mathbf{X}_{(1)}; \mathbf{Y}) + \underline{I}(\mathbf{X}_{(2)}; \mathbf{Y}) \\ &\leq \underline{I}(\mathbf{X}_{(1)}; \mathbf{Y}) + \underline{I}(\mathbf{X}_{(2)}; \mathbf{Y} | \mathbf{X}_{(1)}) = \underline{I}(\mathbf{X}_{(1)}, \mathbf{X}_{(2)}; \mathbf{Y}), \end{aligned} \quad (21)$$

where the inequality follows from (20). Eqs. (20)-(21) imply that the capacity region defined in (15)-(16) is obviously a subset of the capacity region defined in (17)-(19). But the combination of Theorem 1 and Theorem 3 of [5] proves that these expressions coincide.

V. A GENERAL REPRESENTATION FOR THE CAPACITY REGION OF THE BROADCAST CHANNEL WITH A COMMON MESSAGE

Let $P_{\mathbf{Y}, \mathbf{Z} | \mathbf{X}} = \left\{ P_{\mathbf{Y}, \mathbf{Z} | \mathbf{X}}^n \right\}_{n=1}^{\infty}$ be a broadcast channel having one input \mathbf{X} and two outputs, \mathbf{Y} and \mathbf{Z} . The common message, $W_{c,n}$, and the individual messages, $W_{1,n}$ and $W_{2,n}$, are independent random variables, uniformly distributed over $\{1, \dots, M_n^c\}$, $\{1, \dots, M_n^1\}$ and $\{1, \dots, M_n^2\}$, respectively. The common message is to be reliably decoded by both receivers, and the individual messages $W_{1,n}$ and $W_{2,n}$ are to be decoded by the first and second receivers, respectively. Given a triple of integers $M_n^{(1)}$, $M_n^{(2)}$ and $M_n^{(c)}$, we say that the mappings $(\varphi_n, \psi_n^{(1)}, \psi_n^{(2)})$,

$$\begin{aligned} \varphi_n &: \{1, \dots, M_n^{(1)}\} \times \{1, \dots, M_n^{(2)}\} \times \{1, \dots, M_n^{(c)}\} \rightarrow \mathcal{X}^n \\ \psi_n^{(1)} &: \mathcal{Y}^n \rightarrow \{1, \dots, M_n^{(1)}\} \times \{1, \dots, M_n^{(c)}\} \\ \psi_n^{(2)} &: \mathcal{Z}^n \rightarrow \{1, \dots, M_n^{(2)}\} \times \{1, \dots, M_n^{(c)}\} \end{aligned} \quad (22)$$

form an $(n, M_n^{(1)}, M_n^{(2)}, M_n^{(c)}, \epsilon_n)$ -code if

$$\Pr \left\{ \psi_n^{(1)}(Y^n) \neq (W_{1,n}, W_{c,n}) \text{ or } \psi_n^{(2)}(Z^n) \neq (W_{2,n}, W_{c,n}) \right\} \leq \epsilon_n, \quad (23)$$

where the probability law is given by $P_{W_{1,n}, W_{2,n}, W_{c,n}} P_{Y^n Z^n | \varphi_n(W_{1,n}, W_{2,n}, W_{c,n})}$.

Theorem 1 provides us with an expression for the capacity region, which is given by the union of all non-negative rate triples (R_1, R_2, R_c) satisfying

$$R_1 \leq \underline{I}(\mathbf{U}; \mathbf{Y}) \quad (24)$$

$$R_2 \leq \underline{I}(\mathbf{V}; \mathbf{Z}), \quad (25)$$

$$R_c \leq \min \{ \underline{I}(\mathbf{W}; \mathbf{Y}), \underline{I}(\mathbf{W}; \mathbf{Z}) \}, \quad (26)$$

for some distribution $P_{\mathbf{W}, \mathbf{U}, \mathbf{V}, \mathbf{X}, \mathbf{Y}, \mathbf{Z}}$ of the form $P_{\mathbf{W}} P_{\mathbf{U}} P_{\mathbf{V}} P_{\mathbf{X} | \mathbf{W}, \mathbf{U}, \mathbf{V}} P_{\mathbf{Y} \mathbf{Z} | \mathbf{X}}$, where W^n , U^n , and V^n are taking values in some alphabets, whose cardinalities are upper bounded by $|\mathcal{X}|^n$ as follows from (4).

Theorem 3 of [11] establishes an alternative expression for the capacity region of the broadcast channel as the union of all non-negative rate triples (R_1, R_2, R_c) satisfying

$$R_1 \leq \underline{I}(\mathbf{U}; \mathbf{Y} | \mathbf{W}) \quad (27)$$

$$R_2 \leq \underline{I}(\mathbf{V}; \mathbf{Z} | \mathbf{W}), \quad (28)$$

$$R_c \leq \min \{ \underline{I}(\mathbf{W}; \mathbf{Y}), \underline{I}(\mathbf{W}; \mathbf{Z}) \}, \quad (29)$$

for some distribution $P_{\mathbf{W}, \mathbf{U}, \mathbf{V}, \mathbf{X}, \mathbf{Y}, \mathbf{Z}}$ of the form $P_{\mathbf{W}} P_{\mathbf{U}} P_{\mathbf{V}} P_{\mathbf{X} | \mathbf{W}, \mathbf{U}, \mathbf{V}} P_{\mathbf{Y} \mathbf{Z} | \mathbf{X}}$, where W^n , U^n , and V^n are taking values in some alphabets, whose cardinalities are not bounded in [11], but can be bounded similarly to the argument used in Theorem 1.

Note that, similarly to (20)

$$\begin{aligned} \underline{I}(\mathbf{U}; \mathbf{Y}) &\leq \underline{I}(\mathbf{U}; \mathbf{W}, \mathbf{Y}) = \underline{I}(\mathbf{U}; \mathbf{Y} | \mathbf{W}) \\ \underline{I}(\mathbf{V}; \mathbf{Z}) &\leq \underline{I}(\mathbf{V}; \mathbf{W}, \mathbf{Z}) = \underline{I}(\mathbf{V}; \mathbf{Z} | \mathbf{W}), \end{aligned} \quad (30)$$

where the inequalities follow from [10, Theorem 8-f)], and the equalities hold since $\mathbf{W}, \mathbf{U}, \mathbf{V}$ are independent. Hence, the capacity region defined in (24)-(26) is obviously a subset of the capacity region defined in (27)-(29). But the combination of Theorem 1 and Theorem 3 of [11] proves that these expressions coincide.

VI. STATIONARY MEMORYLESS DEGRADED BROADCAST CHANNEL

In this section we demonstrate how a tight single-letter outer bound can be derived from the limiting expression of Theorem 1. This is done for the capacity region of the two-user stationary memoryless degraded broadcast channel.

A stationary memoryless broadcast channel is defined by the input alphabet \mathcal{X} , the output alphabets \mathcal{Y} and \mathcal{Z} , and the probability transition function $\mu_{\mathbf{Y}, \mathbf{Z} | \mathbf{X}}$, that is,

$$\mu^n(Y_1^n = y^n, Z_1^n = z^n | X_1^n = x^n) = \prod_{i=1}^n \mu_{\mathbf{Y}, \mathbf{Z} | \mathbf{X}}(y_i, z_i | x_i).$$

The channel is said to be physically² degraded ([13]) if $\mu_{Y,Z|X} = \mu_{Y|X}\mu_{Z|Y}$.

Let \mathcal{S} be some set³. Consider the pair of random variables, S, X taking values in $\mathcal{S} \times \mathcal{X}$ with \mathcal{X} being the channel input alphabet. For the measure P_{SX} , define the region $\mathcal{R}_\mu(P_{SX})$ as the union of all non-negative rate pairs (R_1, R_2) such that

$$\begin{aligned} R_1 &\leq I(X; Y|S) \\ R_2 &\leq I(S; Z), \end{aligned} \quad (31)$$

where the joint measure which defines $I(X; Y|S)$ and $I(S; Z)$ is given by

$$P_{SXYZ}(s, x, y, z) = P_{SX}(s, x)\mu_{Y,Z|X}(y, z|x).$$

The single-letter expression for the capacity region of the channel μ , $C_{sl,DBC}^\mu$, is given by ([13]) the closure of the convex hull of $\cup_{P_{SX}} \mathcal{R}_\mu(P_{SX})$. Next, it will be shown that this single-letter expression can be derived from Theorem 1, whose application to the capacity region of the two-user broadcast channel (with no common message), denoted $C_{lim,DBC}^\mu$, is given by the union of all non-negative rate-pairs (R_1, R_2) satisfying

$$\begin{aligned} R_1 &\leq \underline{I}(\mathbf{U}; \mathbf{Y}) \\ R_2 &\leq \underline{I}(\mathbf{V}; \mathbf{Y}) \end{aligned} \quad (32)$$

for some distribution $P_{\mathbf{U}, \mathbf{V}, \mathbf{X}, \mathbf{Y}, \mathbf{Z}}$ of the form $P_{\mathbf{U}}P_{\mathbf{V}}P_{\mathbf{X}|\mathbf{U}\mathbf{V}}\mu_{\mathbf{Y}\mathbf{Z}|\mathbf{X}}^n$.

Proposition 1: For the broadcast channel μ ,

$$C_{lim,DBC}^\mu \subseteq C_{sl,DBC}^\mu. \quad (33)$$

Sketch of the proof:

Let P_{U^n, V^n, X^n} be any triple of random variables such that U^n is independent of V^n and X^n is the channel input, let Y^n, Z^n be the channel μ^n outputs corresponding to the input X^n , that is, $(U^n, V^n) \leftrightarrow X^n \leftrightarrow (Y^n, Z^n)$ is a Markov chain. This defines $\mathbf{U}, \mathbf{V}, \mathbf{X}, \mathbf{Y}$ and \mathbf{Z} . Similarly to Theorem 3.5.2. in [8], one can establish the inequalities

$$\begin{aligned} \underline{I}(\mathbf{U}; \mathbf{Y}) &\leq \liminf_{n \rightarrow \infty} \frac{1}{n} I(U^n; Y^n) \\ \underline{I}(\mathbf{V}; \mathbf{Z}) &\leq \liminf_{n \rightarrow \infty} \frac{1}{n} I(V^n; Z^n). \end{aligned} \quad (34)$$

Now, as U^n and V^n are independent we have,

$$\begin{aligned} I(U^n; Y^n) &\leq I(U^n; Y^n, V^n) = I(U^n; Y^n | V^n) \\ &= \sum_{i=1}^n I(U^n; Y_i | V^n, Y^{i-1}) \\ &\leq \sum_{i=1}^n I(X_i; Y_i | V^n, Y^{i-1}), \end{aligned} \quad (35)$$

²A similar derivation applies to the stochastically degraded channel.

³In fact, $|\mathcal{S}| \leq \min\{|\mathcal{X}|, |\mathcal{Y}|, |\mathcal{Z}|\}$

where the last step holds since $(U^n, V^n) \leftrightarrow X^n \leftrightarrow Y^n$ and the channel is memoryless. Furthermore,

$$\begin{aligned} I(V^n; Z^n) &= \sum_{i=1}^n I(V^n; Z_i | Z^{i-1}) \\ &\leq \sum_{i=1}^n I(V^n, Z^{i-1}, Y^{i-1}; Z_i) \stackrel{(*)}{=} \sum_{i=1}^n I(V^n, Y^{i-1}; Z_i) \end{aligned} \quad (36)$$

where $(*)$ is because $(U^n, V^n) \leftrightarrow X^n \leftrightarrow (Y^n, Z^n)$ and the channel is degraded memoryless. Using the standard definition of $T^{(n)}$ as a random variable uniformly distributed over $\{1, \dots, n\}$ and defining $(\bar{S}^{(n)}, X^{(n)}, Y^{(n)}, Z^{(n)})$ by $(Y^{i-1}, V^n, X_i, Y_i, Z_i)$ given $T^{(n)} = i$, we get

$$\begin{aligned} \underline{I}(\mathbf{U}; \mathbf{Y}) &\leq \liminf_{n \rightarrow \infty} I(X^{(n)}; Y^{(n)} | \bar{S}^{(n)}, T^{(n)}) \\ \underline{I}(\mathbf{V}; \mathbf{Z}) &\leq \liminf_{n \rightarrow \infty} I(\bar{S}^{(n)}, T^{(n)}; Z^{(n)}). \end{aligned} \quad (37)$$

Substituting $S^{(n)} = (\bar{S}^{(n)}, T^{(n)})$ in the above expression implies the existence of a probability measure P_{SX} satisfying $\underline{I}(\mathbf{U}; \mathbf{Y}) \leq I(X; Y|S) + \delta$ and $\underline{I}(\mathbf{V}; \mathbf{Z}) \leq I(S; Z) + \delta$ for an arbitrarily small $\delta > 0$, which, combined with standard bounding of $|\mathcal{S}|$, proves that $C_{lim,DBC}^\mu \subseteq C_{sl,DBC}^\mu$.

ACKNOWLEDGMENT

This research was supported by a Marie Curie International Fellowship within the 6th European Community Framework Programme and through collaborative participation in the Collaborative Technology Alliance for Communications and Networks sponsored by the U.S. Army Research Laboratory under Cooperative Agreement DAAD19-01-2-0011.

REFERENCES

- [1] C. E. Shannon, Two-way communication channels, in *Proc. 4th Berkeley Symp. Math. Statist. and Prob.*, 1961, pp. 611-644 (reprinted in Key Papers in the Development of Information Theory).
- [2] E. C. van der Meulen, "The discrete memoryless channel with two senders and one receiver," in *Proc. 2nd Int. Symp. Inform. Theory*, Tsahkadsor, Armenia, U.S.S.R., pp. 95-102, Sept. 1971.
- [3] R. Ahlswede, "Multi-way communication channels," in *Proc. 2nd Int. Symp. Inform. Theory*, Tsahkadsor, Armenia, U.S.S.R., pp. 23-52, Sept. 1971.
- [4] R. S. Cheng and S. Verdú, "On limiting characterizations of memoryless multiuser capacity regions," *IEEE Trans. Inform. Theory*, vol. 39, no. 2, pp. 609-612, Mar. 1993.
- [5] T. S. Han, "An information-spectrum approach to capacity theorems for the general multiple-access channel," *IEEE Trans. Inform. Theory*, vol. 44, no. 7, pp. 2773-2795, Nov. 1998.
- [6] R. G. Gallager, *Information Theory and Reliable Communication*, New York: Wiley, 1968.
- [7] J. Wolfowitz, *Coding Theorem of Information Theory*, 3rd ed. New York: Springer, 1978.
- [8] T. S. Han, *Information-Spectrum Methods in Information Theory*, Springer, 2003.
- [9] S. Shamai (Shitz) and S. Verdú, "The empirical distribution of good codes," *IEEE Trans. Inform. Theory*, vol. 43, no. 3, May, 1997.
- [10] S. Verdú and T. S. Han, "A general formula for channel capacity," *IEEE Trans. Inform. Theory*, vol. 40, no. 4, pp. 1147-1157, July 1994.
- [11] K. Iwata and Y. Oohama, "Information-Spectrum Characterization of Broadcast Channel with General Source," *IEICE-Trans. Fund.* vol. E88-A, no. 10 pp. 2808-2818, Oct. 2005.
- [12] S. Verdú, "Multiple-Access Channels with Memory with and without Frame-Synchronism," *IEEE Trans. Inform. Theory*, vol. IT-35, no. 3, pp. 605-619, May 1989.
- [13] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.