When is the zero-error capacity positive in the relay, multiple-access, broadcast and interference channels?

Natasha Devroye
University of Illinois at Chicago, Chicago IL 60607, USA
Email: devroye @ uic.edu

***Abstract*—Shannon determined that the zero-error capacity of a point-to-point channel whose channel $p(y|x)$ has confusability graph $G_{X|Y}$ is positive if and only if there exist two inputs that are "non-adjacent", or "non-confusable". Equivalently, it is non-zero if and only if the independence number of $G_{X|Y}$ is strictly greater than 1. A multi-letter expression for the zero-error capacity of the channel with confusability graph $G_{X|Y}$ is known, and is given by the normalized limit as the blocklength $n \to \infty$ of the maximum independent set of the n-fold strong product of $G_{X|Y}$. This is not generally computable with known methods. In this paper, we look at the zero-error capacity of four multi-user channels: the relay, the multiple-access (MAC), the broadcast (BC), and the interference (IC) channels. As a first step towards finding a multi-letter expression for the capacity of such channels, we find necessary and sufficient conditions under which the zero-error capacity is strictly positive.**

I. Introduction and notation

Mathematically, one can *define* the zero-error capacity and the ϵ-error capacity of a point-to-point channel with inputs $X \in \mathcal{X}$ and outputs $Y \in \mathcal{Y}$ linked through conditional probability mass functions $p(y|x)$ in a similar fashion. However, the *form of the solutions and analysis* differs: while the $0 < \epsilon < 1$ error capacity leads itself to a probabilistic analysis and the single-letter expression for the capacity C of

$$C = \max_{p(x)} I(X;Y) \text{ bits/channel use,}$$

the zero-error capacity is a combinatorial problem for which only a multi-letter expression is known [1]. While a combinatorial approach may be taken to derive the ϵ-error capacity [2], the reverse is not true; the ϵ and zero-error are different beasts [3], [4] [2, Ch. 11].[1]

A given channel's zero-error capacity, with stricter constraints on the probability of decoding error, is generally smaller than its ϵ-error capacity. In fact, the zero-error capacity of many commonly studied channels is zero – the binary symmetric channel with cross-over probability $p(y = 0|x = 1) = p(y = 1|x = 0) = \delta > 0$ is the simplest channel for which the zero-error capacity is zero, but the ϵ error capacity is strictly positive as long as $\delta \neq \frac{1}{2}$. A natural question to ask is thus, when is the zero-error capacity of a channel positive? For the point-to-point channel, this is well known and will be discussed in Section II. This paper's contribution lies in determining conditions under which the zero-error capacities of the relay, multiple-access, broadcast, and interference channels are positive. This is to the best of our knowledge the first study of such conditions for networks, and acts as a first step towards understanding their zero-error capacities, which are notoriously difficult open problems [4]. The zero-error capacity of the primitive relay channel was considered in [6], [7], [8], the zero-error capacity of the binary adder channel has been a long-standing open problem [9], and the zero-error capacity of a specific interference channel was considered in [4], [10]. However, none of these works consider general channel conditions under which the zero-error capacity of the network is positive.

We next provide some definitions of graphs and of the zero-error capacity of a point-to-point channel. We then proceed to state necessary and sufficient conditions under which the point-to-point (previously known), the relay, the multiple access, the broadcast, and the interference channels have non-zero zero-error capacity. We define the channel models for the multi-user channels directly in their corresponding sections.

A. Graph theoretic notation.

A graph $G(V, E)$ consists of a set V of vertices or nodes together with a set E of edges, which are two-element subsets of V. Two nodes connected by an edge are called *adjacent*. We will usually drop the V, E indices in $G(V, E)$.

An *independent set* of a graph G is a set of vertices, no two of which are adjacent. Let the *independence number* $\alpha(G)$ be the maximum cardinality of all independent sets. A *maximum independent set* is an independent set that has $\alpha(G)$ vertices. One graph can have multiple maximum independent sets.

The *strong product* $G \boxtimes H$ of two graphs G and H is defined as the graph with vertex set $V(G \boxtimes H) = V(G) \times V(H)$, in which two distinct vertices (g, h) and (g', h') are adjacent iff g is adjacent or equal to g' in G and h is adjacent or equal to h' in H. $G^{\boxtimes n}$ denotes the strong product of n copies of G.

A *confusability graph* $G_{X|Y}$ of X given Y, specified by conditional probability mass functions $p(y|x)$ from discrete channel input alphabets $\mathcal{X}$ to discrete channel output alphabets

[1]E.g., a discontinuity often happen between $\epsilon > 0$ to $\epsilon = 0$ capacity [2, Ch.11], and unlike in ϵ-error, Alon [5] disproved Shannon's conjecture [1] that the zero-error capacity of independent, parallel channels is the sum of the zero-error capacities.

$\mathcal{Y}$, is a graph whose vertex set is $\mathcal{X}$ and an edge is placed between vertices $x, x' \in \mathcal{X}$ if they may be "confused", that is, if $\exists y \in \mathcal{Y} : p(y|x) \cdot p(y|x') > 0$.

B. Zero-error capacity definition.

Consider zero-error communication over a point-to-point (P2P) channel $(\mathcal{X}, p(y|x), \mathcal{Y})$, with inputs $X \in \mathcal{X}$ connected to outputs $Y \in \mathcal{Y}$ through the "channel" modeled as a set of conditional probability mass functions $p(y|x)$. This was initially studied by Shannon in 1956 [1]; see [11], [4] for further zero-error capacity details. Communication takes place over n channel uses.

In one channel use, the sender transmits an input $x \in \mathcal{X}$, the receiver receives an output $y \in \mathcal{Y}$ according to $p(y|x)$ and from this must determine which x was transmitted with no error. The largest number of inputs that can be communicated over one channel use is thus $\alpha(G_{X|Y})$, the maximum independent set of the confusability graph $G_{X|Y}$ of the channel described by $p(y|x)$.

If we use the channel n times, with input $x^n := (x_1, x_2, \cdots x_n)$, outputs $y^n := (y_1, y_2, \cdots y_n)$ and where channel output at time i depends only on the channel input at time i according to $p(y_i|x_i)$ and no other inputs and outputs, then $\alpha(G_{X^n|Y^n})$ input vectors will be able to be distinguished with zero error. It is easy to check that the confusability graph of the channel from X^n to Y^n, $G_{X^n|Y^n}$ is given by the strong product of n copies of the confusability graph $G_{X|Y}$, i.e.

$$G_{X^n|Y^n} = G_{X|Y}^{\boxtimes n}.$$

We then call

$$C(G_{X|Y}) := \sup_n \frac{1}{n} \log \left(\alpha(G_{X|Y}^{\boxtimes n}) \right) = \lim_{n\to\infty} \frac{1}{n} \log \left(\alpha(G_{X|Y}^{\boxtimes n}) \right) \tag{1}$$

the *Shannon capacity* C of the channel with confusability graph $G_{X|Y}$. This corresponds to the largest number of bits (if the logarithm is taken to base 2) that may be communicated without error (per channel use). This limits exists by Fekete's lemma and the super-multiplicativity of the strong product of graphs. We note that sometimes, as done by Lovász [11] and Alon [12], one defines the Shannon capacity of the graph G without the logarithm (to represent the number of distinguishable inputs), and that one may easily transfer results from one definition into the other.

Equation (1) is a generally uncomputable limiting expression that may be unsatisfying to some, despite it being quite intuitive. In short, we understand the form of the capacity, but it cannot be computed with current technology, except for special classes such as *perfect* graphs [13], [14]. The difficulty in computing (1) lies not only in the NP-hardness of finding the independence number of a graph, but also in the (possibly) strange behavior of the sequence of independence numbers for strong product graphs, say $\{\alpha(G^{\boxtimes n})\}_{n=1}^{\infty}$. How this sequence behaves is a notoriously difficult open question and attracts attention in graph theory and combinatorics [15], [16], [17].

II. When is the zero-error capacity of a point-to-point channel positive?

Consider zero-error communication over a point-to-point channel $(\mathcal{X}, p(y|x), \mathcal{Y})$. An *$n$-shot protocol* $(n, \mathcal{W}, h, \underline{\mathcal{X}}, g)$ for communicating over a point-to-point channel without error is composed of:

- An input message set $\mathcal{W}$.
- An encoding function $h : \mathcal{W} \to \underline{\mathcal{X}}$ and a codebook $\underline{\mathcal{X}} \subset \mathcal{X}^n$. Message w is mapped to codeword $\underline{x}(w) = h(w)$ and transmitted over the channel $p(y|x)$.
- A decoding function $g : \mathcal{Y}^n \to \mathcal{W}$ that produces an estimate of the transmitted message w.

In the zero-error communication context, transmitting messages over a communication system is equivalent to differentiating different codewords from a codebook, which is a subset of the channel input alphabet, based on the channel output signals. Rate $R = \frac{1}{n} \log ||\mathcal{W}||$ is *achievable* if there exists an n-shot protocol $(n, \mathcal{W}, h, \underline{\mathcal{X}}, g)$ for that achieves zero error, i.e. for which the estimates message equals the true transmitted message for all messages, regardless of the channel realization $p(y|x)$ over the n channel uses. The zero-error capacity is the supremum of all achievable rates (over all blocklengths n).

Shannon showed that the zero-error capacity of the point-to-point channel $(\mathcal{X}, p(y|x), \mathcal{Y})$ is strictly positive if and only if there exist two channel inputs that are *non-adjacent.* Two inputs x, x' are called *non-adjacent* if their *reachable sets*

$$\mathcal{Y}(x) := \{y|p(y|x) > 0\}, \;\; \mathcal{Y}(x') := \{y : p(y|x') > 0\} \tag{2}$$

are disjoint. Equivalently, two inputs are *adjacent* if their reachable sets are not disjoint, in which case Massey [18] prefers calling the inputs *confusable.* We follow Massey's terminology as well. In short then, Shannon showed the following Theorem:

Theorem 1 (Shannon [1], P2P zero-error capacity positive)**.** *The zero-error capacity of the point-to-point channel $(\mathcal{X}, p(y|x), \mathcal{Y})$ is strictly positive if and only if there exist two inputs $x \neq x'$ that are non-confusable, i.e. for which*

$$\mathcal{Y}(x) \cap \mathcal{Y}(x') = \emptyset.$$

Since the zero-error capacity of the P2P channel may be succinctly expressed in terms of its confusability graph $G_{X|Y}$ as in (1), one might hope that Theorem 1 could be rephrased as a condition on $G_{X|Y}$ as well. Indeed, by noting that the graph $G_{X|Y}$ is constructed by placing an edge between any two inputs that are *confusable*, and recalling that an independent set of a graph is a subset of vertices no two of which share an edge, we can re-phrase Shannon's theorem as follows, which we feel is more intuitive. We will present all subsequent theorems in terms of confusability graphs as well for uniformity, though we could just as well have formulated them in terms of the sets $\mathcal{Y}(x)$ and their multi-user analogs.

Corollary 2. *The zero-error capacity of the point-to-point channel $(\mathcal{X}, p(y|x), \mathcal{Y})$ with confusability graph $G_{X|Y}$ is* *strictly positive if and only if*

$$\alpha(G_{X|Y}) > 1. \tag{3}$$

Remark 1. *Notice that this is a "single-letter" condition in some sense – i.e. one need not look at multiple channel uses to determine whether the channel has non-zero capacity. If a channel's confusability graph G has independence number of 1, i.e.* $\alpha(G) = 1$*, then every vertex must be connected to every other vertex. Hence G must be a complete graph, or a clique, which in one channel use is not able to distinguish even 1 bit of information, i.e. has a one-shot zero-error capacity of 0. If G is complete, then every strong product of G is also complete and hence multiple channel uses will not "resolve" any ambiguities, i.e.* $C(G) = 0$ *even as* $n \to \infty$*.*

This single-letter nature of the condition for non-zero capacity is in sharp contrast to the capacity expression in (1), which is a limiting, multi-letter expression. While the condition in (3) is computable in polynomial time (just check for an independent set of size 2), (1) is NP-hard.

However, the two expressions do bear some similarity: the condition for positive capacity asks whether $\alpha(G) > 1$ while the capacity asks for the largest $\alpha(G)$, properly normalized, as $n \to \infty$. It seems both hit on the same concept of confusability of inputs, at the heart of the zero-error capacity. We now look for analogous theorems to Corollary 2 for the relay, MAC, BC and IC channels, with main results highlighted in Fig. 1.

III. When is the zero-error capacity of a relay channel positive?

A relay channel $(\mathcal{X} \times \mathcal{X}_R, p(y, y_R|x, x_R), \mathcal{Y} \times \mathcal{Y}_R)$ consists of a source terminal S that wants to communicate a message W to a destination terminal D aided by a relay terminal R. We first define the zero-error capacity of this channel formally, before finding conditions under which it is non-zero.

An $(n, \mathcal{W}, h, \underline{\mathcal{X}}, \phi_1, \phi_2, \cdots, \phi_n, g)$ protocol for the zero-error relay channel consists of:

- An input message set $\mathcal{W}$.
- An encoding function $h : \mathcal{W} \to \underline{\mathcal{X}}$, mapping messages to codewords in the codebook $\underline{\mathcal{X}} \subset \mathcal{X}^n$.
- Relaying functions ϕ_i, $i \in [1 : n]$ which at time i assigns each past received output $y_R^{i-1} \in \mathcal{Y}_R^{i-1}$ a symbol in $\mathcal{X}_R$ according to $\phi_i \mathcal{Y}_R^{i-1} \to \mathcal{X}_R$.
- A decoding function $g : \mathcal{Y}^n \to \underline{\mathcal{X}}$ which leads to an estimate of the transmitted message $w \in \mathcal{W}$

A message rate $R := \frac{1}{n} \log ||\underline{\mathcal{W}}||$ is said to be achievable if there exists an $(n, \mathcal{W}, h, \underline{\mathcal{X}}, \phi_1, \phi_2, \cdots, \phi_n, g)$ protocol for which for the decoded message exactly equals the true sent message regardless of the channel realizations $p(y, y_R|x, x_R)$ over the n channel uses.

For this channel, the condition for non-zero zero-error capacity boils down to a condition resembling the cut-set bound. Define the following confusability graphs, both defined on vertices $\mathcal{X} \times \mathcal{X}_R$, with edges as follows:

$$\begin{aligned} G_{X,X_R=x_R|Y,Y_R} &: \text{edge between } (x, x_R) \neq (x', x_R) \text{ if} \\ &\exists (y, y_R) : p(y, y_R|x, x_R) \cdot p(y, y_R|x', x_R) > 0 \end{aligned}$$

$$\begin{aligned} G_{X,X_R|Y} &: \text{edge between } (x, x_R) \neq (x', x_R') \text{ if} \\ &\exists y : p(y|x, x_R) \cdot p(y|x', x_R') > 0, \end{aligned}$$

where we recall that $p(y|x, x_R) = \sum_{y_R \in \mathcal{Y}_R} p(y, y_R|x, x_R)$.

Theorem 3. *The zero-error capacity of the relay channel is non-zero if and only if*

$$\exists x_R : \ \alpha(G_{X,X_R=x_R|Y,Y_R}) > 1 \tag{4}$$

$$\textit{AND}$$

$$\alpha(G_{X,X_R|Y}) > 1. \tag{5}$$

Proof: We show that (4) and (5) are both necessary and sufficient conditions for a positive zero-error capacity.

First, to show that they are necessary, we need to show that if the zero-error capacity is positive, then (4) and (5) hold. We show the equivalent contrapositive, i.e. that if either of the conditions is not met, that necessarily the zero-error capacity must be zero.

We notice that the zero-error capacity is upper bounded by the minimum of the capacities of the channels from S to (R,D) and from (S,R) to D (the two usual "cuts" used in the cut-set bound for the relay channel [19]). This follows from arguments similar to the cut-set bound, but for zero-error networks, as follows. Say (4) does not hold. Then for all $x_R \in \mathcal{X}_R$, $\alpha(G_{X,X_R=x_R|Y,Y_R}) = 1$. This means that even if the destination D were to be given Y_R and the input the relay sent $X_R = x_R$ at that channel use (which would increase the capacity of the channel as it could ignore this added knowledge if it does not help in disambiguating inputs x), that still not even one bit of information could be conveyed. Hence the zero-error capacity would be zero. Similar arguments as in Remark 1 may be used to argue that larger blocklengths will not change this statement. Similarly, if (5) does not hold, this means that destination D cannot resolve even two different (x, x_R) pairs. Since resolving a pair of (x, x_R) is less stringent than resolving just x (i.e. it is in general easier to resolve a pair than a singleton), it is hence not possible to resolve just x either, and hence the capacity must be 0.

To show that these conditions are also sufficient, we need to show that if (4) and (5) hold, then the capacity must be positive. Assume the conditions to be true. Then, in one channel use, let the source transmit one of the $x \neq x' \in \mathcal{X}$ that form an independent set of size at least 2, and the relay transmit the dummy x_R that satisfies (4) (exist by presumption). After this channel use, both the source and the relay may thus agree upon 1 bit of information, as this will lead to at least 2 distinct (y, y_R) pairs. If these two pairs have different y's then the destination can distinguish 1 bit by itself (without the relay) and (5) will also automatically hold. If these two pairs have time same y they must have different y_R's

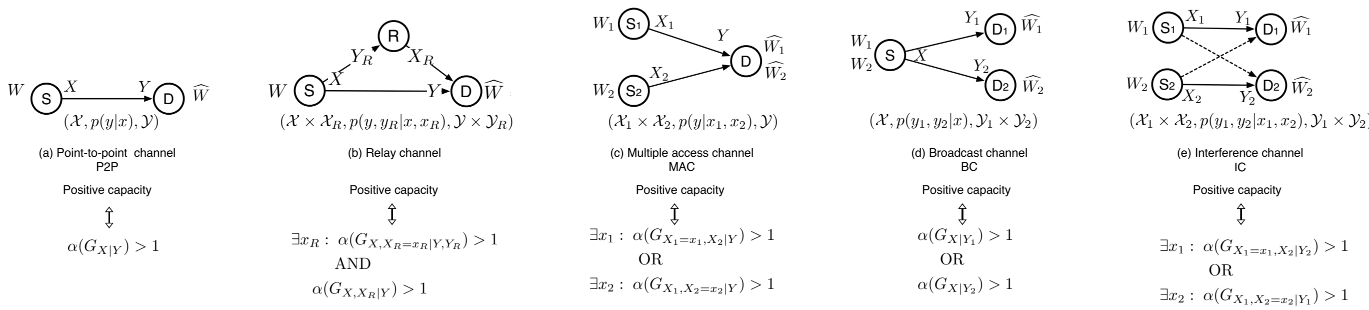

Fig. 1. Channels and main result.

and in this case the relay can distinguish 1 bit. In the second channel use, the source and relay agree to encode the 1 bit they agree upon using the two $(x, x_R) \neq (x', x'_R)$ pairs that form an independent set of size at least 2 that satisfies (5) (which again exist by presumption) and hence are non-confusable at the destination. After this second channel use then, one bit has been unambiguously conveyed to the destination, and hence the capacity is non-zero. ■

Remark 2. *Note that if there exists a dummy x_R^* value for which there exist two inputs $x \neq x'$ that lead to non-confusable outputs at the destination D only (regardless of what is received at the relay's Y_R), i.e. $\{x, x'\}$ forms an independent sets of $G_{X,X_R=x_R^*|Y}$ and hence also of $G_{X,X_R=x_R^*|Y,Y_R}$, then the direct link has non-zero zero-error capacity. In this case, both conditions* (4) *and* (5) *will be satisfied, and the relay need not be involved in information transmission (it may simply transmit the dummy x_R^*) in order to demonstrate a positive zero-error capacity.*

When going after the zero-error capacity of the relay channel in fully generality, one may surmise that the graphs in (4) and (5) will play important roles. This seems to be true for the zero-error capacity of the primitive relay channel, which was studied in [6], [7], [8], and for which a necessary and sufficient condition on the out-of band R-D link capacity was obtained (and a construction of the optimal "Colour-and-Forward" relaying scheme needed) to achieve the single-input multiple-output (cut-set) outer bound. Determining multi-letter expressions for the zero-error capacity of the relay channel is the subject of ongoing investigation.

IV. When is the zero-error capacity of a multiple-access channel positive?

We next consider a multiple access channel $(\mathcal{X}_1, \times\mathcal{X}_2, p(y|x_1,x_2), \mathcal{Y})$ in which two independent transmitters 1 and 2 wish to send independent messages W_1, W_2 to a single destination that wishes to decode both messages. At each channel use the inputs and outputs are related through $p(y|x_1,x_2)$ at that channel use, independent of the inputs and outputs at other times. For this network, the ϵ-error capacity is well known and is equal to the union of all rate pairs satisfying

$$R_1 \leq I(X_1; Y|X_2, Q)$$
$$R_2 \leq I(X_2; Y|X_1, Q)$$
$$R_1 + R_2 \leq I(X_1, X_2; Y|Q)$$

taken over the distributions $p(q)p(x_1|q)p(x_2|q)$. The zero-error capacity of the multiple access channel is in general open, with the binary adder channel being an example of a channel where the zero-error capacity is strictly [20] smaller than the ϵ-error version, and for which the exact zero-error capacity remains unknown. In this section we do not focus on a particular channel such as the binary adder channel, but rather seek conditions under which the zero-error capacity for any MAC channel is positive.

An $(n, \mathcal{W}_1, \mathcal{W}_2, h_1, h_2, \underline{\mathcal{X}_1}, \underline{\mathcal{X}_2}, g)$ protocol for the zero-error MAC consists of:

- Two input message sets $\mathcal{W}_1$ and $\mathcal{W}_2$.
- Two encoding functions $h_1 : \mathcal{W}_1 \to \underline{\mathcal{X}_1}$, $h_2 : \mathcal{W}_2 \to \underline{\mathcal{X}_2}$ and two codebooks $\underline{\mathcal{X}_1} \subset \mathcal{X}_1^n$, $\underline{\mathcal{X}_2} \subset \mathcal{X}_2^n$. Messages w_1, w_2 are mapped to codewords $\underline{x_1}(w_1) = h_1(w_1)$ and $\underline{x_2}(w_2) = h_2(w_2)$ and transmitted over the channel $p(y|x_1, x_2)$.
- A decoding function $g : \mathcal{Y}^n \to \mathcal{W}_1 \times \mathcal{W}_2$ that produces estimates of the transmitted messages w_1 and w_2.

A rate pair $(R_1 := \frac{1}{n}\log||\mathcal{W}_1||, R_2 := \frac{1}{n}\log||\mathcal{W}_2||)$ is *achievable* if there exists an $(n, \mathcal{W}_1, \mathcal{W}_2, h_1, h_2, \underline{\mathcal{X}_1}, \underline{\mathcal{X}_2}, g)$ protocol which achieves zero-error, i.e. for which the estimated messages are equal to the true messages for all messages, regardless of the channel realizations $p(y|x_1, x_2)$ over the n channel uses. The zero-error capacity region is the union of all achievable rate pairs.

Based on the previous section, in which the condition for non-zero zero-error capacity related to being able to find an independent set of the "cut-set" bound for the relay channel of cardinality more than 1, for this channel, the condition $\alpha(G_{X_1,X_2|Y}) > 1$, where $G_{X_1,X_2|Y}$ is a graph with vertices $\mathcal{X}_1 \times \mathcal{X}_2$ and

$$G_{X_1,X_2|Y} : \text{edge between } (x_1,x_2) \neq (x'_1,x'_2) \quad \text{if } \exists y \in \mathcal{Y} : p(y|x_1,x_2)\cdot p(y|x'_1,x'_2) > 0 \tag{6}$$

is *not* immediately relevant in determining whether we have a non-zero zero-error capacity. The main issue is that such a condition would not capture the distributed nature of the encoding process, i.e. that the two transmitters must encode their messages independently. As an example, let $\mathcal{X}_1 = \mathcal{X}_2 = \{0,1\}$ (binary channels). Then the confusability graph $G_{X_1,X_2|Y}$ has 4 vertices $(0,0),(0,1),(1,0),(1,1)$. Consider a possible confusability graph in Fig. 2, in which $\alpha(G_{X_1,X_2|Y}) = 2$. Here, one might suspect that the capacity of this MAC channel is non-zero. However, this is not the case. The two possible independent sets (both of size 2) are $\{(0,0),(1,1)\}$ and $\{(0,1),(1,0)\}$. However, we cannot use these pairs as codewords since the inputs of the two users are correlated rather than independent. Since the users cannot coordinate and jointly decide to transmit only $(0,0)$ and $(1,1)$ for the first independent set (and not the other pairs), we see that this graph is not relevant for the MAC channel. This is not surprising – by looking at $G_{X_1,X_2|Y}$, we inherently assume that pairs of (X_1,X_2) can be sent. However, due to the lack of coordination between the transmitters in a MAC, a Cartesian product of a subset of $\mathcal{X}_1$ and $\mathcal{X}_2$ is more relevant. Indeed, for this example, while $\alpha(G_{X_1,X_2|Y}) = 2$, $\alpha(G_{X_1=x_1,X_2|Y}) = \alpha(G_{X_1,X_2=x_2|Y}) = 1$ for all $x_1, x_2 \in \{0,1\}$, and so the zero-error capacity of this MAC would indeed be zero, as shown next.

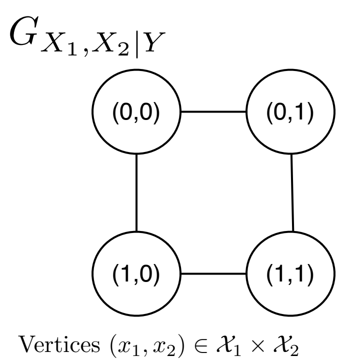

Fig. 2. Example of $G_{X_1,X_2|Y}$ showing why this is not the correct graph to consider for the MAC.

The more relevant conditions relate to the following confusability graphs,

$$G_{X_1=x_1,X_2|Y}: \text{ vertices } (x_1,x_2),\ x_2 \in \mathcal{X}_2 \text{ for fixed } x_1$$
$$\text{edge between } (x_1,x_2) \neq (x_1,x_2') \text{ if } \exists y \in \mathcal{Y}:$$
$$p(y|x_1,x_2)\cdot p(y|x_1,x_2') > 0 \tag{7}$$

$$G_{X_1,X_2=x_2|Y}: \text{ vertices } (x_1,x_2),\ x_1 \in \mathcal{X}_1 \text{ for fixed } x_2$$
$$\text{edge between } (x_1,x_2) \neq (x_1',x_2) \text{ if } \exists y \in \mathcal{Y}:$$
$$p(y|x_1,x_2)\cdot p(y|x_1',x_2) > 0 \tag{8}$$

Then, we may express the condition for non-zero capacity region as :

Theorem 4. *The zero-error capacity region of the MAC is not the point* $(0,0)$ *if and only if either of the following conditions hold:*

$$\exists x_1 \in \mathcal{X}_1 : \alpha(G_{X_1=x_1,X_2|Y}) > 1 \tag{9}$$
$$\exists x_2 \in \mathcal{X}_2 : \alpha(G_{X_1,X_2=x_2|Y}) > 1 \tag{10}$$

Proof: To show that this condition is necessary to achieve a positive zero-error region, we again look at the contrapositive statement. Suppose that neither condition holds. Then we wish to show that the zero-error capacity is necessarily $(0,0)$. Consider (9). Since for each $x_1 \in \mathcal{X}_1$, $G_{X_1=x_1,X_2|Y}$ is complete, no matter what letter (one channel use) or codeword (multiple channel uses) transmitter 1 sends, all of user 2's symbols are confusable. Hence, $R_2 = 0$. Similarly, (10) implies that $R_1 = 0$.

To show that this condition is sufficient to achieve a positive zero-error region, set $X_1 = x_1$ for the x_1 that satisfies (9). Then, let user 2 transmit the two different x_2 values in the independent set of size at least 2 of $G_{X_1=x_1,X_2|Y}$, say x_2^* and x_2^{**}. The receiver will then be able to distinguish the pairs (x_1,x_2^*) and (x_1,x_2^{**}), and hence we can achieve the rate pair $(0,1)$ (and hence a region larger than $(0,0)$). Similarly, if the condition in (10) is satisfied, we can achieve a rate pair of at least $(1,0)$. If both conditions hold, time sharing between having transmitter 1 send the x_1 in (9) and transmitter 2 sending the two symbols in the independent set of $G_{X_1=x_1,X_2|Y}$, and having transmitter 2 send the x_2 in (10) and transmitter 1 sending the two symbols in the independent set of $G_{X_1,X_2=x_2|Y}$ will lead to a whole region in which $R_1 > 0$ and $R_2 > 0$. ■

V. When is the zero-error capacity of a broadcast channel positive?

We next consider a broadcast channel $(\mathcal{X}, p(y_1,y_2|x), \mathcal{Y}_1 \times \mathcal{Y}_2)$ in which one transmitter sends 2 independent messages W_1 and W_2 to two independent receivers, each of which desires only one message. At each channel use the inputs and outputs are related through $p(y_1,y_2|x)$ at that channel use, independent of the inputs and outputs at other times. For this network, the ϵ-error capacity is still unknown in general, but known for degraded, semi-deterministic, and Gaussian channels, among others [19]. In this section we seek conditions under which the zero-error capacity for any BC channel is positive.

An $(n, \mathcal{W}_1, \mathcal{W}_2, h, \underline{\mathcal{X}}, g_1, g_2)$ protocol for the broadcast channel consists of:

- Two input message sets $\mathcal{W}_1$ and $\mathcal{W}_2$.
- An encoding function $h : \mathcal{W}_1 \times \mathcal{W}_2 \to \underline{\mathcal{X}}$ and a codebook $\underline{\mathcal{X}} \subset \mathcal{X}^n$. Messages w_1, w_2 are mapped to codeword $\underline{x} = h(w_1,w_2)$ which is transmitted over the channel $p(y_1,y_2|x)$.
- Two decoding functions $g_1 : \mathcal{Y}_1^n \to \mathcal{W}_1$ and $g_2 : \mathcal{Y}_2^n \to \mathcal{W}_2$ which estimate the transmitted messages w_1 and w_2, respectively.

Rate pair $(R_1 := \frac{1}{n}\log||\mathcal{W}_1||, R_2 := \frac{1}{n}\log||\mathcal{W}_2||)$ is said to be achievable if there exists an $(n, \mathcal{W}_1, \mathcal{W}_2, h, \underline{\mathcal{X}}, g_1, g_2)$

protocol for which the estimated messages are equal to the true messages for all messages, regardless of the channel realizations $p(y_1, y_2|x)$ over the n channel uses. The capacity region is the union of all achievable rate pairs.

Based on the relay channel section, in which the condition for non-zero zero-error capacity related to being able to find an independent set of the "cut-set" bound for the relay channel of cardinality more than 1, for this channel, a similar condition of $\alpha(G_{X|Y_1,Y_2}) > 1$, where $G_{X|Y_1,Y_2}$ is a graph with vertices $\mathcal{X}$ and

$$G_{X|Y_1,Y_2}: \text{ edge between } x \neq x' \quad (11)$$
$$\text{if } \exists (y_1, y_2) : p(y_1, y_2|x) \cdot p(y_1, y_2|x') > 0$$

is *not* sufficient for a non-zero zero-error capacity. The main issue is that such a condition would not capture the distributed nature of the decoding process, i.e. that the receivers cannot cooperate or share their outputs in decoding the messages. As an example of a channel for which $\alpha(G_{X|Y_1,Y_2}) > 1$ but clearly no other rate point than $(0, 0)$ can be achieved by the BC, consider the broadcast channel in Fig. 3 which has binary inputs and outputs. The channel between X and Y_1 is a binary symmetric channel with cross-over probability $p(y = 0|x = 1) = p(y = 1|x = 0) = \frac{1}{2}$. The channel between X and Y_2 is an indicator function which indicates (i.e. outputs $Y_2 = 1$) when either $\{x = 0 \cap y_1 = 1\}$ or $\{x = 1 \cap y_1 = 0\}$ occurs, i.e. indicates whether the channel between X and Y_1 flipped the sent bit. Here, when given both Y_1 and Y_2, one can distinguish whether $X = 0$ or $X = 1$ was sent. However, the channels from X to Y_1 and from X to Y_2, considered on their own, both have zero and ϵ-error capacity equal to 0. Clearly, in the BC, each receiver only has access to its own decoded signal and hence $G_{X|Y_1}$ and $G_{X|Y_2}$ are the relevant graphs to consider, and are defined as:

$$\begin{aligned} G_{X|Y_1}: &\text{ vertices } x \in \mathcal{X} \\ &\text{edge between } x \neq x' \text{ if } \exists y_1 \in \mathcal{Y}_1 : \\ &\quad p(y_1|x) \cdot p(y_1|x') > 0 \end{aligned} \quad (12)$$

$$\begin{aligned} G_{X|Y_2}: &\text{ vertices } x \in \mathcal{X} \\ &\text{edge between } x \neq x' \text{ if } \exists y_2 \in \mathcal{Y}_2 : \\ &\quad p(y_2|x) \cdot p(y_2|x') > 0, \end{aligned} \quad (13)$$

where $p(y_1|x) = \sum_{y_2} p(y_1, y_2|x)$ and $p(y_2|x) = \sum_{y_1} p(y_1, y_2|x)$.

Then, we have the following condition for a positive zero-error capacity region:

Theorem 5. *The zero-error capacity region of the BC is not the point* $(0, 0)$ *if and only if either of the following conditions hold:*

$$\alpha(G_{X|Y_1}) > 1 \quad (14)$$
$$\alpha(G_{X|Y_2}) > 1 \quad (15)$$

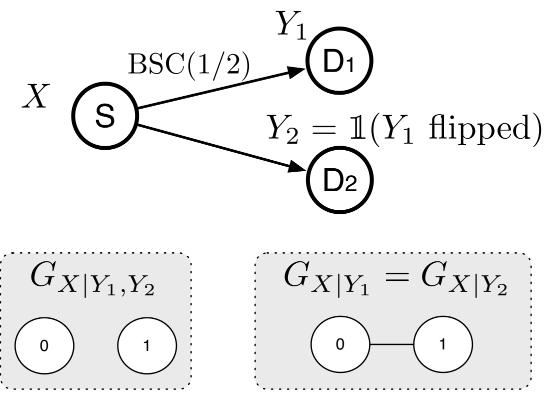

Fig. 3. Example of $G_{X|Y_1,Y_2}$ versus $G_{X|Y_1}$ and $G_{X|Y_2}$ showing why the former is not the correct graph to consider for the BC.

Proof: To show that this condition is necessary to achieve a positive zero-error region, we again look at the contrapositive statement. Suppose that neither condition holds. Then we wish to show that the zero-error capacity is necessarily $(0, 0)$. Consider (14) and assume it does not hold. This implies $\alpha(G_{X|Y_1}) = 1$, in which case given Y_1, not even two X's can be distinguished. Thus, as in Theorem 1 and Remark 1, this implies a rate $R_1 = 0$ no matter what. Similarly for condition (15) and R_2.

To show that this condition is sufficient to achieve a positive zero-error region, clearly if (14) is satisfied then the rate point $(1, 0)$ can be achieved by encoding the message W_1 into the two distinguishable X values in the independent set of $G_{X|Y_1}$ of size at least 2 (ignoring W_2, i.e. making the function h a function of $w_1 \in \mathcal{W}_1$ only). Similarly, if (15) is satisfied, the rate point $(0, 1)$ can be achieved by encoding the message W_2 into the two non-confusable X values (ignoring W_1). Then one can time-share between these two points to achieve a region with positive area. ■

VI. When is the zero-error capacity of a interference channel positive?

In an interference channel $(\mathcal{X}_1 \times \mathcal{X}_2, p(y_1, y_2|x_1, x_2), \mathcal{Y}_1 \times \mathcal{Y}_2)$ two independent transmitters wish to communicate independent messages W_1 and W_2 to to independent receivers. Receiver 1 wishes to decode W_1 and receiver 2 wishes to decode W_2. At each channel use the inputs and outputs are related through $p(y_1, y_2|x_1, x_2)$ at that channel use, independent of the inputs and outputs at other times. The ϵ-error capacity of the interference channel is open in general, but is known for classes of deterministic ICs, strong ICs, and is known to within a constant gap for a class of semi-deterministic ICs which includes the Gaussian IC [19].

First, let us define zero-error communication over the interference channel. An $(n, \mathcal{W}_1, \mathcal{W}_2, h_1, h_2, \underline{\mathcal{X}_1}, \underline{\mathcal{X}_2}, g_1, g_2)$ protocol consists of:

- Two input message sets $\mathcal{W}_1$ and $\mathcal{W}_2$.
- Two encoding functions $h_1 : \mathcal{W}_1 \to \underline{\mathcal{X}_1}$, $h_2 : \mathcal{W}_2 \to \underline{\mathcal{X}_2}$ and two codebooks $\underline{\mathcal{X}_1} \subset \mathcal{X}_1^n$, $\underline{\mathcal{X}_2} \subset \mathcal{X}_2^n$. Messages

w_1, w_2 are mapped to codewords $\underline{x_1}(w_1) = h_1(w_1)$ and $\underline{x_2}(w_2) = h_2(w_2)$ which are transmitted over the channel $p(y_1, y_2|x_1, x_2)$.

- Two decoding functions $g_1 : \mathcal{Y}_1^n \to \mathcal{W}_1$ and $g_2 : \mathcal{Y}_2^n \to \mathcal{W}_2$ that produce estimates of the transmitted messages w_1 and w_2.

A rate pair $(R_1 := \frac{1}{n}\log||\mathcal{W}_1||, R_2 := \frac{1}{n}\log||\mathcal{W}_2||)$ is *achievable* if there exists an $(n, \mathcal{W}_1, \mathcal{W}_2, h_1, h_2, \underline{\mathcal{X}_1}, \underline{\mathcal{X}_2}, g_1, g_2)$ protocol which achieves zero-error, i.e. for which the decoded messages exactly equal the transmitted messages for all possible messages w_1, w_2 in $\mathcal{W}_1 \times \mathcal{W}_2$, regardless of the channel realizations $p(y|x_1, x_2)$ over the n channel uses. The zero-error capacity region is the union of all achievable rate pairs.

Along similar lines as for the MAC and BC channels, we see that the conditions for positive capacity region will depend on the following graphs:

$$\begin{aligned} &G_{X_1=x_1,X_2|Y_2} : \text{ vertices } (x_1, x_2), x_2 \in \mathcal{X}_2 \text{ for fixed } x_1 \\ &\text{edge between } (x_1, x_2) \neq (x_1, x_2'), \text{ if } \exists y_2 \in \mathcal{Y}_2 : \\ &\quad p(y_2|x_1, x_2) \cdot p(y_2|x_1, x_2') > 0 \end{aligned} \tag{16}$$

$$\begin{aligned} &G_{X_1,X_2=x_2|Y_1} : \text{ vertices } (x_1, x_2), x_1 \in \mathcal{X}_1 \text{ for fixed } x_2 \\ &\text{edge between } (x_1, x_2) \neq (x_1', x_2), \text{ if } \exists y_1 \in \mathcal{Y}_1 : \\ &\quad p(y_1|x_1, x_2) \cdot p(y_1|x_1', x_2) > 0. \end{aligned} \tag{17}$$

Then, the conditions for positive capacity region are:

Theorem 6. *The zero-error capacity region of the IC is not the point* $(0, 0)$ *if and only if either of the following conditions hold:*

$$\exists x_1 \in \mathcal{X}_1 : \alpha(G_{X_1=x_1,X_2|Y_2}) > 1 \tag{18}$$

$$\exists x_2 \in \mathcal{X}_2 : \alpha(G_{X_1,X_2=x_2|Y_1}) > 1 \tag{19}$$

Equation (18) will lead to a rate pair of at least $(0, 1)$, while (19) leads to a rate pair of at least $(1, 0)$. The proof follows the same techniques as before.

VII. CONCLUSION

We have presented conditions for the relay, multiple-access, broadcast and interference channels for the zero-error capacity to be larger than the trivial point or region (i.e. for positive zero-error capacity). This is a first step towards finding a useful expression for the capacity region of such channels. For the multi-user channels with rate pairs (MAC, BC, IC), the conditions found boil down to at least one of the users having a positive zero-error capacity. Another interesting question would be to determine necessary and sufficient conditions to ensure that both $R_1 > 0$ and $R_2 > 0$ without time-sharing, i.e. positive rates for both users simultaneously in a single time slot. We hope and expect that the capacity regions may be expressed in terms of some of the multi-user confusability graphs presented here, and this is the subject of ongoing work.

Acknowledgment. The work of Natasha Devroye was partially funded by NSF under award 1216825 and 1645381. The contents of this article are solely the responsibility of the authors and do not necessarily represent the official views of the NSF.

REFERENCES

[1] C. Shannon, "The zero error capacity of a noisy channel," Information Theory, IRE Transactions on, vol. 2, no. 3, pp. 8–19, September 1956.
[2] I. Csiszár and J. Körner, Information Theory: Coding Theorems for Discrete Memoryless Systems, 2nd ed. New York: Cambridge University Press, 2011.
[3] S. Verdú, "Fifty Years of Shannon Theory," IEEE Trans. Inf. Theory, vol. 44, no. 6, pp. 2057–2078, Oct. 1998.
[4] J. Körner and A. Orlitsky, "Zero-error information theory," IEEE Trans. Inf. Theory, vol. 44, no. 6, pp. 2207–2229, 1998.
[5] N. Alon, "The Shannon capacity of a union," Combinatorica, vol. 18, no. 3, pp. 301–310, Mar. 1998.
[6] Y. Chen, S. Shahi, and N. Devroye, "Colour-and-forward: relaying "what the destination needs" in the zero-error primitive relay channel," in Allerton Conference on Communication, Control, and Computing, 2014.
[7] Y. Chen and N. Devroye, "The optimality of colour-and-forward relaying for a class of zero-error primitive relay channels," in Proc. IEEE Int. Symp. Inf. Theory, 2015.
[8] ——, "Relaying in zero-error primitive relay channels that achieves the simo bound," Submitted to IEEE Trans. Inf. Theory, Jan. 2016.
[9] G. H. Khachatrian, ch. A survey of coding methods for the adder channel.
[10] G. Simonyi, "On write-unidirectional memory codes," IEEE Trans. Inf. Theory, vol. 35, no. 3, pp. 663–669, May 1989.
[11] L. Lovasz, "On the shannon capacity of a graph," IEEE Trans. Inf. Theory, vol. 25, no. 1, pp. 1–7, 1979.
[12] N. Alon, "Combinatorial reasoning in information theory," IEEE Information Theory Society Newsletter, vol. 59, pp. 10–13, 2009.
[13] C. Berge, Graphs and Hypergraphs. Amsterdam, The Netherlands: North Holland, 1973.
[14] S. Hougardy, "Classes of perfect graphs," Discrete Mathematics, vol. 306, no. 19-20, pp. 2529–2571, Oct. 2006.
[15] E. Sonnemann and O. Krafft, "Independence numbers of product graphs," Journal of Combinatorial Theory, Series B, vol. 17, no. 2, pp. 133 – 142, 1974.
[16] P. Jha and G. Slutzki, "Independence numbers of product graphs," Applied Mathematics Letters, vol. 7, no. 4, pp. 91 – 94, 1994.
[17] N. Alon and E. Lubetzky, "The Shannon capacity of a graph and the independence numbers of its powers," IEEE Trans. Inf. Theory, vol. 52, no. 5, pp. 2172–2176, May 2006.
[18] J. Massey, "Zero error," in Winter School of Information Theory slides, La Colle sur Loup, Mar. 1997.
[19] A. El Gamal and Y.-H. Kim, Network Information Theory. Cambridge University Press, 2011.
[20] R. Urbanke and Q. Li, "The zero-error capacity region of the 2-user synchronous bac is strictly smaller than its shannon capacity region," in Proc. IEEE Int. Symp. Inf. Theory, 1998.