

Zero-error Relaying for Primitive Relay Channels

Yanying Chen, Natasha Devroye
 University of Illinois at Chicago, Chicago IL 60607, USA
 Email: ychen90, devroye @ uic.edu

Abstract—In a primitive relay channel, a new one-shot relaying scheme termed “color-and-forward” is proposed that guarantees a probability of error equal to zero. This relaying scheme constructs a relaying compression graph of relay outputs based on the joint conditional distribution of the relay and destination outputs and forwards a minimum coloring of this graph. The n -letter extension of the proposed color-and-forward scheme is shown to be optimal in the sense that it results in the smallest needed out-of-band relay to destination link rate for the overall message rate to equal the single input multiple output outer bound for any fixed number of channel uses. This is used to obtain an upper bound on the asymptotic minimal relay to destination link rate needed to achieve the single input multiple output outer bound.

I. INTRODUCTION

A primitive relay channel (PRC) $(\mathcal{X}, p(y, y_R|x), \mathcal{Y} \times \mathcal{Y}_R)$, illustrated in Figure 1(b), is a special type of relay channel $(\mathcal{X} \times \mathcal{X}_R, p(y, y_R|x, x_R), \mathcal{Y} \times \mathcal{Y}_R)$, shown in Figure 1(a). Both consist of a source terminal S that communicates a message to a destination terminal D aided by a relay terminal R. In a primitive relay channel, the broadcasting links $(\mathcal{X}, p(y, y_R|x), \mathcal{Y} \times \mathcal{Y}_R)$ from the source to the relay and destination terminals are out-of-band, or orthogonal to the error-free but rate-limited relay to destination (R-D) link. Quantities of interest may then be the maximal number of codewords (the *message rate*) that can be reliably communicated for a given R-D link rate r (the *relay rate*), or the minimal relay rate needed to transmit at a desired message rate (provided the desired message rate is feasible at all).

This paper characterizes the relaying scheme that, for a fixed number of channel uses, minimizes the relay rate needed to achieve the maximal possible message rate. When the R-D link rate is large enough, the relay can forward its entire observation to the destination terminal. Thus, the primitive relay channel effectively turns into a point-to-point channel with a single input and two outputs, say $(\mathcal{X}, p(y, y_R|x), \mathcal{Y} \times \mathcal{Y}_R)$, shown in Figure 1(c), for which we have finite n (or n -shot, where n is the number of channel uses) and asymptotic expressions for the zero-error capacity. This capacity is an upper bound to the message rate achievable for any finite relay rate. The question addressed here is how large the relay rate should be to ensure that overall message rate achieved meets the capacity of the single-input multiple output (SIMO) channel $(\mathcal{X}, p(y, y_R|x), \mathcal{Y} \times \mathcal{Y}_R)$.

The work of the authors was partially funded by NSF under award 1216825; the contents of this article are solely the responsibility of the authors and do not necessarily represent the official views of the NSF. The work was presented in part at Allerton 2014 and ISIT 2015. The authors would like to sincerely thank the anonymous reviewer for their help.

The *small-error*¹ version of this question was considered in [1], [2], [3] and remains open. Recent work by Wu and Özgür [4] has shown that in general, the cut-set bound is loose for such channels. We propose and study the zero-error version of this problem. We propose a new one-shot (or $n = 1$) “color-and-forward” relaying scheme whose n -shot extension yields the minimum relay rate needed to achieve the n -letter SIMO bound for any given number of channel uses n . This is used to obtain an upper bound on the asymptotic relay rate needed to achieve the asymptotic SIMO upper bound.

A. A motivating example

Consider a PRC with $p(y, y_R|x) = p(y|x)p(y_R|x)$ as in Figure 2, noting that in zero-error communication the values of the conditional probabilities is immaterial, only whether they are non-zero or not matters. The destination, upon receiving Y can tell whether $\{1, 2\}$ or $\{3, 4\}$ were sent, but not which input within those sets. The relay can resolve this ambiguity by forwarding E or O , i.e. whether the X was even or odd. This amounts to considerable savings for the R-D link capacity with respect to sending Y_R directly, and allows the destination to fully resolve which X was sent as long as the R-D link capacity is at least 1 bit. It may be checked that this channel does not fall into any class of PRCs for which capacity is known, i.e. it is not a degraded, semideterministic, orthogonal-component, or semideterministic PRC [2]. The “color-and-forward” relaying scheme proposed here generalizes this example to arbitrary PRCs and yields achievable (message rate, relay rate) pairs for any number of channel uses n .

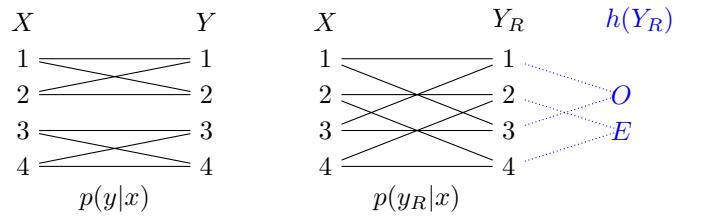


Fig. 2: Toy Problem: $p(y, y_R|x) = p(y|x)p(y_R|x)$. A solid link indicates the probability value $p(+|x)$ is positive, where $+$ indicates y or y_R .

B. Background on zero-error communication over a PRC

Zero-error communication over a primitive relay channel at first glance seems to be a combination of two notoriously

¹Communication allowing a vanishing probability of error is called *small-error* or ϵ -error communication, while communication without error is called *zero-error* or 0-error communication.

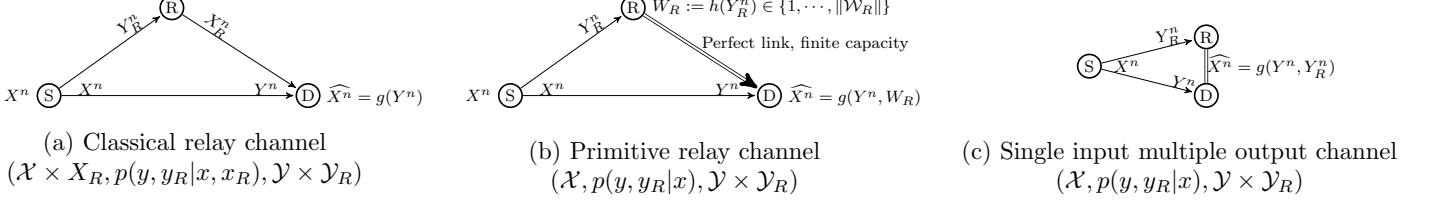


Fig. 1: The relay, primitive relay, and single-input multiple output channels. The question posed here is, in a zero error setting, how large the perfect link in (b) needs to be to render (b)'s capacity equal to that of (c).

difficult and open communication problems in information theory: computing the zero-error capacity over a point-to-point channel, and deriving the small-error capacity of a relay channel. As outlined in the excellent survey [2], the small-error capacity of primitive relay channels is in general unknown, except for degraded [5], semideterministic [6], orthogonal-component [7], and semideterministic primitive [8] relay channels.

The zero-error capacity of the PRC has not been studied besides the authors' initial results [9], [10]. This problem does however generalize a point-to-point zero-error source coding problem with correlated side information available only at the receiver end, a problem studied by Witsenhausen [11]. This connection will be explained in Section VII.

The zero-error capacity of a point-to-point channel $(\mathcal{X}, p(y|x), \mathcal{Y})$ with discrete finite channel input and output alphabets is characterized as the limit as the number of channel uses $n \rightarrow \infty$ of the normalized independence number $\alpha([G_{X|Y}]^{\boxtimes n})$ of the n -fold AND (or strong) product of the confusability graph $G_{X|Y}$ associated with $p(y|x)$. This generally uncomputable limiting expression may be unsatisfying. Even for small alphabet sizes, this is a challenging problem: Shannon's conjecture that the capacity of the famous "pentagon graph" channel is $\frac{1}{2} \log 5$ was only formally proven by Lovasz [12] 23 years later by proposing the θ function, which serves as an upper bound for the independence number of a graph and can be computed in polynomial time [13]. Thus, a computable expression for the zero-error capacity for even the simplest, point-to-point channel remains open, except for a small class of channels with *perfect* graphs²[14].

C. Contributions and organization

In Section II, notation and zero-error preliminaries are introduced. The main contributions follow in:

- Section III, where zero-error communication over a primitive relay channel is formally defined.
- Section IV: where a new one-shot $n = 1$ color-and-forward relaying scheme is proposed that achieves zero error. Achievable (message rate, relay rate) pairs are characterized. The extension to n channel uses is immediate. The color-and-forward scheme is shown to be optimal in the sense that for any fixed number of channel uses n , this relaying scheme results in the smallest relay rate such that

²A *perfect graph* is a graph where the chromatic number of every induced subgraph is that subgraph's largest clique size.

the overall message rate equals the n -letter SIMO upper bound.

- Section V: where an upper bound on the minimum required asymptotic (in blocklength) relay rate to achieve the asymptotic SIMO upper bound is developed.
- Section VI: where one example is worked out, and a class of channels for which the bound is tight and easily characterized is presented.
- Section VII: where connections with Witsenhausen's problem on zero-error source-coding with receiver side information are made, and conclusions are made.

II. NOTATION AND ZERO-ERROR PRELIMINARIES

Throughout the paper we will use subscripts z to emphasize the zero-error context. We use upper and lower cases to differentiate the message rate R_z and the R-D link rate, or relay rate r_z . Graphs use the letter G – subscripts are used to differentiate relaying compression graphs (e.g. G_R) from confusability graphs (e.g. $G_{X|Y, Y_R}$), where additional subscripts and superscripts will indicate the number of channel uses or restrictions on alphabets, and will be defined later. Sequences of length n are denoted using superscript n as x^n . When a conditional joint pmf $p(y, y_R|x)$ with support \mathcal{X} and output $\mathcal{Y} \times \mathcal{Y}_R$ is restricted to input \mathcal{K} , we denote its *induced conditional pmf, support, and output* by $p_{\mathcal{K}}(y, y_R|x)$, \mathcal{K} and $\mathcal{Y}|_{\mathcal{K}} \times \mathcal{Y}_R|_{\mathcal{K}}$ respectively. All logarithms are base 2.

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. Note that one graph can have multiple maximum independent sets. A *coloring* of graph G is any function c over the vertex set such that c^{-1} induces a partition of the vertex set into independent sets of G . The *chromatic number* $\chi(G)$ of the graph G is the least number of colors in any coloring. A *minimum coloring* of graph G uses $\chi(G)$ colors.

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 function $p(y|x)$ with support \mathcal{X} and output \mathcal{Y} , is a graph whose vertex set is \mathcal{X} and an edge is placed when two different nodes $x, x' \in \mathcal{X}$ may be “confused” at the output, that is, if $\exists y \in \mathcal{Y} : p(y|x) \cdot p(y|x') > 0$.

B. Zero-error preliminaries

The zero-error capacity of a point-to-point discrete memoryless channel was studied by Shannon in [15] in 1956; see [12], [14] for further insight. We outline the results we will build upon next.

Consider zero-error communication over a point-to-point channel $(\mathcal{X}, p(y|x), \mathcal{Y})$. An n -shot protocol $(n, \underline{\mathcal{X}}, g)$ for communicating over a point-to-point channel without error is composed of a codebook $\underline{\mathcal{X}} \subseteq \mathcal{X}^n$ and a decoding function $g : \mathcal{Y}^n \rightarrow \underline{\mathcal{X}}$ that estimates the transmitted codeword as $\hat{X} = g(Y^n)$. Message rate $R_z^{(n)} := \frac{1}{n} \log \|\underline{\mathcal{X}}\|$, where $\|\mathcal{X}\|$ is the cardinality of set \mathcal{X} is called *achievable* if there exists an n -shot protocol $(n, \underline{\mathcal{X}}, g)$ for which $g(Y^n) = \underline{X}$ for all codewords $\underline{X} \in \underline{\mathcal{X}}$. The capacity is the supremum over all n of all achievable rates.

First, note that only whether $p(y|x)$ is zero or not matters for communication without error. Next, consider communicating using a single channel use: the maximal number of channel inputs the destination can distinguish without error is $\alpha(G_{X|Y})$, the maximum number of vertices that are non-adjacent, or pairwise distinguishable. When multiple channel uses are allowed, $\alpha([G_{X|Y}]^{\boxtimes n})$ is the number of distinguishable channel inputs X^n , where $[G_{X|Y}]^{\boxtimes n}$ is the strong product of n copies of graph $G_{X|Y}$.³ Thus, the zero-error capacity of a point-to-point channel $(\mathcal{X}, p(y|x), \mathcal{Y})$ is defined as the supremum of all achievable message rates, i.e., $\sup_n \frac{1}{n} \log \alpha([G_{X|Y}]^{\boxtimes n})$.

The zero-error capacity may then be shown to be characterized as [12]

$$\begin{aligned} \sup_n \frac{1}{n} \log \alpha([G_{X|Y}]^{\boxtimes n}) &= \lim_{n \rightarrow \infty} \frac{1}{n} \log \alpha([G_{X|Y}]^{\boxtimes n}) \\ &= \lim_{n \rightarrow \infty} \log \sqrt[n]{\alpha([G_{X|Y}]^{\boxtimes n})}, \end{aligned}$$

which may be upper and lower bounded as [15], [12]:

$$\log \alpha(G_{X|Y}) \leq \lim_{n \rightarrow \infty} \log \sqrt[n]{\alpha([G_{X|Y}]^{\boxtimes n})} \leq \log \|\mathcal{X}\|.$$

The behavior of the sequence of independence numbers for strong product graphs, say $\{\alpha(G^{\boxtimes n})\}_{n=1}^{\infty}$, is a long standing, notoriously difficult open question and attracts attention in graph theory and combinatorics [16], [17], [18].

III. ZERO-ERROR COMMUNICATION OVER A PRIMITIVE RELAY CHANNEL AND THE MINIMUM R-D LINK RATES r_z^* AND $r_z^{*(n)}$

We first define zero-error communication over a PRC, then introduce finite n and asymptotic upper bounds on the message

³Note that the n -fold strong product graph $[G_{X|Y}]^{\boxtimes n}$ is equivalent to the confusability graph $G_{X^n|Y^n}$ constructed from the *extended channel* $(\mathcal{X}^n, p(y^n|x^n), \mathcal{Y}^n)$ with $p(y^n|x^n) = \prod_{i=1}^n p(y_i|x_i)$.

rate $SIMO(n)$ and $SIMO$, as well as the quantities of interest here – the minimum needed relay rate $r_z^{*(n)}$ and r_z^* to achieve these finite n and asymptotic upper bounds.

A. Zero-error communication over a primitive relay channel

An n -shot protocol $(n, \underline{\mathcal{X}}, h, g)$ for $n \geq 1$ channel uses, for zero-error communication over a PRC $(\mathcal{X}, p(y, y_R|x), \mathcal{Y} \times \mathcal{Y}_R)$ shown in Figure 3 is composed of a codebook $\underline{\mathcal{X}} \subseteq \mathcal{X}^n$, a relaying function $h : \mathcal{Y}_R^n \rightarrow \mathcal{W}_R$, and a decoding function $g : \mathcal{Y}^n \times \mathcal{W}_R \rightarrow \underline{\mathcal{X}}$.

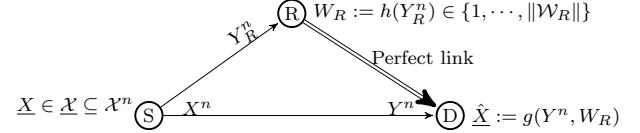


Fig. 3: An n -shot protocol $(n, \underline{\mathcal{X}}, h, g)$ for zero-error communication over a PRC $(\mathcal{X}, p(y, y_R|x), \mathcal{Y} \times \mathcal{Y}_R)$, with a codebook $\underline{\mathcal{X}}$, a relaying function h and a decoding function g .

A rate pair $(R_z^{(n)}, r_z^{(n)})$, where $R_z^{(n)} := \frac{1}{n} \log \|\underline{\mathcal{X}}\|$ is the *message rate*, and $r_z^{(n)}$ is the *relay rate* is said to be *achievable* if there exists an n -shot protocol $(n, \underline{\mathcal{X}}, h, g)$ over a PRC $(\mathcal{X}, p(y, y_R|x), \mathcal{Y} \times \mathcal{Y}_R)$ such that the relaying function uses on average less than $r_z^{(n)}$ bits per channel use, i.e. $\|\mathcal{W}_R\| \leq 2^{n \cdot r_z^{(n)}}$, and it achieves zero error, i.e. $\Pr[g(Y^n, W_R) \neq \underline{X} | \underline{X} \text{ sent}] = 0$ for all $\underline{X} \in \underline{\mathcal{X}}$.

For n channel uses, define $C_z^{(n)}(r_z^{(n)})$ to be the maximum $R_z^{(n)}$ such that there exists an achievable $(R_z^{(n)}, r_z^{(n)})$ pair for which $r^{(n)} \leq r_z^{(n)}$. We define $C_z(r_z)$, the zero-error capacity of the relay channel at R-D link rate r_z , to be the supremum over n of all $R_z^{(n)}$ such that there exists an achievable $(R_z^{(n)}, r_z^{(n)})$ pair for which $r_z^{(n)} \leq r_z$.

It is easy to see the following properties of $C_z(r_z)$:

- $C_z(r_z) \leq \log \|\mathcal{X}\|$;
- $C_z(r_z)$ is a non-decreasing function of r_z ;
- $C_z(r) = C_z(\infty)$ whenever $r \geq \log \|\mathcal{Y}_R\|$;

We next present upper bounds to the asymptotic $C_z(r_z)$ and define the problem of interest – finding the minimum R-D link rate to achieve this bound, either asymptotically or for fixed n .

B. SIMO bounds and the minimum R-D link rates r_z^* and $r_z^{*(n)}$

We first present upper bounds on $C_z(r_z)$ and $C_z^{(n)}(r_z^{(n)})$, analogous to the cut-set bound, and termed the single input (in this case \mathcal{X}) multiple output (in this case $\mathcal{Y}, \mathcal{Y}_R$) outer bound, or SIMO bound.

Proposition 1 (Zero-error capacity SIMO upper bound).

$$C_z(r_z) \leq C_z(\infty) := SIMO := \log \lim_{n \rightarrow \infty} \sqrt[n]{\alpha([G_{X|Y, Y_R}]^{\boxtimes n})}. \quad (1)$$

Proof. By giving the destination terminal perfect access to the full sequence of received Y_R values, we cannot decrease the value of $C_z(r_z)$, as we are making the destination more

capable. With this knowledge of Y_R , we obtain a single-input, multiple-output point-to-point channel with input X and vector output (Y, Y_R) , whose zero-error capacity is given by the SIMO value defined above. \square

We are interested in determining the smallest r_z such that $C_z(r_z) = \text{SIMO}$, defined formally next as r_z^* .

Definition 1 (The minimum R-D link rate r_z^* to achieve the SIMO bound).

$$r_z^* := \inf\{r_z : C_z(r_z) = \text{SIMO}\}. \quad (2)$$

Clearly, $r_z^* \leq \log |\mathcal{Y}_R|$; this work revolves around finding a tighter upper bound.

For finite n , we have analogous bounds and definitions, below. We initially focus on characterizing these quantities for fixed n , and then derive an upper bound on r_z^* based on $r_z^{*(n)}$.

Proposition 2 (The n -shot SIMO bound).

$$C_z^{(n)}(r_z) \leq \text{SIMO}(n) := \log \sqrt[n]{\alpha([G_{X|Y, Y_R}]^{\boxtimes n})}. \quad (3)$$

We remind the reader that $\text{SIMO} := \lim_{n \rightarrow \infty} \text{SIMO}(n) = \sup_n \text{SIMO}(n)$.

Definition 2 (The n -shot minimum R-D link rate $r_z^{*(n)}$ to achieve the n -shot SIMO bound).

$$r_z^{*(n)} := \min\{r_z : C_z^{(n)}(r_z) = \text{SIMO}(n)\}. \quad (4)$$

In the following we characterize $r_z^{*(n)}$ exactly as an n -letter extension of the one-shot ‘‘color-and-forward’’ scheme presented next.

IV. A ONE-SHOT ‘‘COLOR-AND-FORWARD’’ SCHEME AND ITS n -SHOT EXTENSION

In the following, we simplify the problem and propose a one-shot relaying scheme – that we term ‘‘color-and-forward’’ – on a restricted input alphabet. Extensions to arbitrary input alphabets and n channels uses are straightforward and will be outlined thereafter.

The main problem of the paper can be stated and solved as follows. For the PRC $(\mathcal{X}, p(y, y_R|x), \mathcal{Y} \times \mathcal{Y}_R)$, let \mathcal{K} be a maximal independent set of the confusability graph $G_{X|Y, Y_R}$. Note that \mathcal{K} need not be unique. Then, by definition, for each (y, y_R) there exists at most one $x \in \mathcal{K}$ for which $p(y, y_R|x) > 0$. Assume that every symbol $x \in \mathcal{K}$ has a positive and equal probability. This condition guarantees that the maximum amount of information is transmitted from $X \in \mathcal{K}$ to $(Y, Y_R) \in \mathcal{Y}_K \times \mathcal{Y}_R|_{\mathcal{K}}$.

Question. Find the minimum cardinality relaying function $h : \mathcal{Y}_R|_{\mathcal{K}} \rightarrow \mathcal{W}_R$ (and its cardinality) such that $(Y, h(Y_R))$ can distinguish X without error, namely, for each $(y, h(y_R))$ with positive probability, there exists at most one $x \in \mathcal{K}$ for which $p(y, h(y_R)|x) > 0$. This is a reformulation of the problem of finding $r_z^{*(1)}$ for the PRC $(\mathcal{K}, p_{\mathcal{K}}(y, y_R|x), \mathcal{Y}_K \times \mathcal{Y}_R|_{\mathcal{K}})$.

To answer this question, we construct a new graph, for which a coloring provides the optimal compression function, i.e. the minimal R-D link rate needed to achieve a message rate equal to $\text{SIMO}(n)$.

Construction of the relaying compression graph $G_R^{(1)}|_{\mathcal{K}}$.

The relaying compression graph $G_R^{(1)}|_{\mathcal{K}}$ is defined as follows:

- Vertices: $Y_R|_{\mathcal{K}}$
- Edges: vertices $y_R \neq y'_R$ both in $Y_R|_{\mathcal{K}}$ share an edge when $\exists x \in \mathcal{K}$ and $\exists x' \in \mathcal{K}$, $x \neq x'$ such that $p(y, y_R|x) > 0$ and $p(y, y'_R|x') > 0$ for some $y \in \mathcal{Y}|_{\mathcal{K}}$.

Answer: Coloring of the relaying compression graph $G_R^{(1)}|_{\mathcal{K}}$. Color the graph at its chromatic number and let $h(y_R)$ be one of the corresponding minimal colorings. We term this processing and forwarding at the relay as ‘‘color-and-forward’’ relaying. We claim that this relaying function, which requires $\log \chi(G_R^{(1)}|_{\mathcal{K}})$ bits per channel use for the R-D link to transmit, recovers all $x \in \mathcal{K}$ with zero error and does so with minimal cardinality. We use superscript (1) to indicate that is graph corresponds to $n = 1$.

Lemma 3. For the PRC $(\mathcal{K}, p_{\mathcal{K}}(y, y_R|x), \mathcal{Y}|_{\mathcal{K}} \times \mathcal{Y}_R|_{\mathcal{K}})$, for \mathcal{K} an independent set of $G_{X|Y, Y_R}$, the relaying function corresponding to a minimal coloring of $G_R^{(1)}|_{\mathcal{K}}$ described above, $r_z^{*(1)} = \log \chi(G_R^{(1)}|_{\mathcal{K}})$.

Proof. To show that $(Y, h(Y_R))$ can recover X with zero error, consider every y_R with $h(y_R) = h$ for some color h . We know that at least one of them, say, y_R^* , must satisfy $p(y, y_R^*|x^*) > 0$ for some x^* , since $p(y, y_R|x) > 0$ for the actual sent x and actual observation (y, y_R) . This x^* must be unique even when y_R^* may not be so. To see this, by the definition of $h(y_R)$ above, any other y'_R with $h(y'_R) = h$ must satisfy $p(y, y'_R|x') = 0$ for $x' \neq x^*$. Otherwise, y_R^* and y'_R are connected and $h(y_R^*) \neq h(y'_R)$. Furthermore, since the possible transmitted x ’s form an independent set of $G_{X|Y, Y_R}$, $p(y, y_R^*|x') = 0$ for every $x' \neq x^*$. Hence, we have established the uniqueness of x^* . This is valid regardless of the transmitted x and hence the message rate achieved is $\log \alpha(G_{X|Y, Y_R})$.

Conversely, suppose that the cardinality of $h(\cdot)$ is less than the chromatic number of the graph. Then, there are $y_R \neq y'_R$ with $h(y_R) = h(y'_R) = h$ such that $p(y, y_R|x) > 0$ and $p(y, y'_R|x') > 0$ for some $x \neq x'$. Since $p(x) = p(x')$ by the assumption and $p(y) > 0$ as y is the actual observation, the conditional probabilities $p(x, y_R|y)$ and $p(x', y'_R|y)$ are both positive, which implies that $p(x, h|y)$ and $p(x', h|y)$ are both positive, which, in turn, implies that $p(x|h, y)$ and $p(x'|h, y)$ are both positive. Hence, there is a positive probability of error. \square

From this construction, we can express $r_z^{*(1)}$ as follows.

Lemma 4. For the PRC $(\mathcal{X}, p(y, y_R|x), \mathcal{Y} \times \mathcal{Y}_R)$

$$r_z^{*(1)} = \min_{\mathcal{K} : \mathcal{K} \text{ is a maximum independent set of graph } G_{X|Y, Y_R}} \log \chi(G_R^{(1)}|_{\mathcal{K}}), \quad (5)$$

where $\chi(G_R^{(1)}|_{\mathcal{K}})$ is the chromatic number of graph $G_R^{(1)}|_{\mathcal{K}}$, constructed via the algorithm described in the Answer above with restricted input / codebook \mathcal{K} .

Proof. Note that there are may be multiple maximal independent sets of confusability graph $G_{X|Y, Y_R}$. Restricting the input

to one of these, denoted by \mathcal{K} , yields a zero-error relaying scheme achieving the rate pair

$$(R_z^{(1)}, r_z^{(1)}) = (\log \alpha(G_{X|Y, Y_R}), \log \chi(G_R^{(1)}|_{\mathcal{K}})).$$

Taking the minimum over all maximal independent sets yields the theorem. \square

It is easy to see how this scheme may be extended to n channel uses. In the following, note that $\mathcal{K}^{(n)}$ corresponds to a maximum independent set of graph $G_{X|Y, Y_R}^{\boxtimes n}$, and should not be confused with the n -fold Cartesian product of the set \mathcal{K} .

Corollary 5. *For the PRC $(\mathcal{X}, p(y, y_R|x), \mathcal{Y} \times \mathcal{Y}_R)$*

$$r_z^{*(n)} = \min_{\mathcal{K}^{(n)}: \mathcal{K}^{(n)} \text{ is a max. ind. set of } G_{X|Y, Y_R}^{\boxtimes n}} \frac{1}{n} \log \chi(G_R^{(n)}|_{\mathcal{K}^{(n)}}), \quad (6)$$

where $\chi(G_R^{(n)}|_{\mathcal{K}^{(n)}})$ is the chromatic number of graph $G_R^{(n)}|_{\mathcal{K}^{(n)}}$, constructed via the algorithm described in the Answer above with restricted input / codebook $\mathcal{K}^{(n)}$.

In working towards the asymptotic r_z^* , it may be natural to ask whether there exists a strong product relationship between $G_R^{(n)}|_{\mathcal{K}^n}$ and $G_R^{(1)}|_{\mathcal{K}}$ for \mathcal{K} an independent set of $G_{X|Y, Y_R}$ and \mathcal{K}^n the n -fold strong product of \mathcal{K} . Note that even if \mathcal{K}^n is a maximal independent set of $G_{X|Y, Y_R}^{\boxtimes n}$ (it need not be), that in general, $G_R^{(n)}|_{\mathcal{K}^n}$ is not the n -fold strong product of $G_R^{(1)}|_{\mathcal{K}}$. The following can however be shown.

Lemma 6. *Let \mathcal{K} be an independent set of graph $G_{X|Y, Y_R}$, and let \mathcal{K}^n denote its n -fold Cartesian product. Then:*

- 1) \mathcal{K}^n is an independent set of $G_{X^n|Y^n, Y_R^n}$, though it need not be a maximal independent set;
- 2) the n -fold strong product of $G_R^{(1)}|_{\mathcal{K}}$ is a sub-graph of the n -shot graph $G_R^{(n)}|_{\mathcal{K}^n}$, i.e.

$$E\left([G_R^{(1)}|_{\mathcal{K}}]^{\boxtimes n}\right) \subseteq E\left(G_R^{(n)}|_{\mathcal{K}^n}\right);$$

- 3) if $G_R^{(1)}|_{\mathcal{K}}$ is fully connected,

$$[G_R^{(1)}|_{\mathcal{K}}]^{\boxtimes n} = G_R^{(n)}|_{\mathcal{K}^n}.$$

Remark 3. Note that when $G_R^{(1)}|_{\mathcal{K}}$ is fully connected, its n -fold strong product $[G_R^{(1)}|_{\mathcal{K}}]^{\boxtimes n}$ is also fully connected, meaning the corresponding n -shot color-and-forward graph $G_R^{(n)}|_{\mathcal{K}^n}$ is fully connected.

Proof of Lemma 6. The statement in 1) is obvious. To show 2) consider any two different nodes $[y_{R1}, \dots, y_{Rn}], [y'_{R1}, \dots, y'_{Rn}] \in [\mathcal{Y}_R|_{\mathcal{K}}]^n$, which are connected by an edge in graph $[G_R^{(1)}|_{\mathcal{K}}]^{\boxtimes n}$. By definition of the strong product, this implies that y_{Ri} and y'_{Ri} are either equal or connected by an edge in graph $G_R^{(1)}|_{\mathcal{K}}$, for any $i = 1, 2, \dots, n$. This also implies there exist $[y_1, \dots, y_n], [x_1, \dots, x_n]$ and $[x'_1, \dots, x'_n]$ satisfying:

- For y_{Ri} and y'_{Ri} pairs that are connected in graph $G_R^{(1)}|_{\mathcal{K}}$, we know there exists some $y_i \in \mathcal{Y}|_{\mathcal{K}}$ and $x_i \neq x'_i \in \mathcal{X}|_{\mathcal{K}}$ such that $p(y_i, y_{Ri}|x_i) \cdot p(y_i, y'_{Ri}|x'_i) > 0$.

- For y_{Ri} and y'_{Ri} pairs where $y_{Ri} = y'_{Ri}$, there exists at least one $y_i \in \mathcal{Y}|_{\mathcal{K}}$ and one $x_i \in \mathcal{X}|_{\mathcal{K}}$ such that $p(y_i, y_{Ri}|x_i) > 0$. In this scenario, we let $x'_i = x_i$.

It can then be checked that

$$\begin{aligned} & p([y_1, \dots, y_n], [y_{R1}, \dots, y_{Rn}] | [x_1, \dots, x_n]) \\ & \quad \cdot p([y_1, \dots, y_n], [y'_{R1}, \dots, y'_{Rn}] | [x'_1, \dots, x'_n]) \\ & = \prod_{i=1}^n p(y_i, y_{Ri}|x_i) \cdot p(y_i, y'_{Ri}|x'_i) \end{aligned}$$

is positive, and $[x_1, \dots, x_n] \neq [x'_1, \dots, x'_n]$ because there is at least one i such that y_{Ri} and y'_{Ri} are connected in graph $G_R^{(1)}|_{\mathcal{K}}$. Hence $[y_{R1}, \dots, y_{Rn}], [y'_{R1}, \dots, y'_{Rn}] \in [\mathcal{Y}_R|_{\mathcal{K}}]^n$ must be connected in $G_R^{(n)}|_{\mathcal{K}^n}$, and 2) follows.

To show 3), notice that when $G_R^{(1)}|_{\mathcal{X}}$ is fully connected, $E\left(G_R^{(n)}|_{\mathcal{K}^n}\right) \subseteq E\left([G_R^{(1)}|_{\mathcal{K}}]^{\boxtimes n}\right)$, and hence 3) follows. \square

Since $E\left([G_R^{(1)}|_{\mathcal{K}}]^{\boxtimes n}\right) \subseteq E\left(G_R^{(n)}|_{\mathcal{K}^n}\right)$, Lemma 6 guarantees that a minimal coloring of $[G_R^{(1)}|_{\mathcal{K}}]^{\boxtimes n}$ cannot be larger than a minimal coloring of $G_R^{(n)}|_{\mathcal{K}^n}$. However, Corollary 5 minimizes $G_R^{(n)}|_{\mathcal{K}^n}$ over all *maximal* independent sets \mathcal{K}^n of $G_{X|Y, Y_R}^{\boxtimes n}$, which may not even contain \mathcal{K}^n (consider for example $G_{X|Y, Y_R}^{\boxtimes n}$ a pentagon).

V. UPPER BOUND ON THE ASYMPTOTIC RATE VIA THE PRESENTED ONE-SHOT SCHEME

We now use Corollary 5 to obtain our main result, an upper bound on the asymptotic r_z^* of Definition 1.

Lemma 7 (Upper bound on r_z^*). $r_z^* \leq U_1 := \sup_n r_z^{*(n)}$.

Proof of Lemma 7. Let $r_z = U_1$. First, $U_1 \geq r_z^{*(n)}$ holds for any n , because U_1 is the supremum of all $r_z^{*(n)}$'s. It is also true that $C_z^{(n)}(r_z) \geq C_z^{(n)}(r_z^{*(n)})$, because $C_z^{(n)}(r_z)$ is non-decreasing with respect to r_z . Note that $C_z^{(n)}(r_z^{*(n)}) = \text{SIMO}(n)$ by definition. So $C_z^{(n)}(r_z) \geq \text{SIMO}(n)$ holds for any n . Thus we have $\sup_n C_z^{(n)}(r_z) \geq \sup_n \text{SIMO}(n)$. Equivalently, $C_z(r_z) \geq \text{SIMO}$. This implies $C_z(r_z) = \text{SIMO}$, because $C_z(r_z) \leq \text{SIMO}$ always holds. Thus, $r_z^* \leq U_1$. \square

VI. EXAMPLES

In this section we first show one detailed example of how to compute $r_z^{*(1)}$, and how this minimization over different maximal independent sets of $G_{X|Y, Y_R}$ cannot be ignored. We then show an example of a class of channels for which one can characterize the asymptotic r_z^* exactly.

A. A detailed computation of $r_z^{*(1)}$

Table I enumerates a conditional joint probability mass function $p(y, y_R|x)$, where $\|\mathcal{X}\| = \|\mathcal{Y}\| = \|\mathcal{Y}_R\| = 5$. An entry at position (x, y, y_R) is denoted by “+” (the actual value does not matter), when its probability $p(y, y_R|x)$ is positive and by “0” when $p(y, y_R|x) = 0$. To obtain $r_z^{*(1)}$:

$s(p(y, y_R x))$	Y_R																								
	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
Y	1	0	0	+	+	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	+	0	0	0	0
	2	+	+	0	0	0	0	+	0	0	0	0	0	0	0	0	0	+	0	0	0	0	0	0	0
	3	0	0	0	0	0	0	0	+	0	0	+	0	0	0	0	0	0	0	0	0	0	0	0	0
	4	0	0	0	0	0	0	0	0	0	0	0	0	+	0	0	0	0	+	0	0	0	0	0	0
	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	+	0	0	0	0	+
	$X = 1$					$X = 2$					$X = 3$					$X = 4$					$X = 5$				

TABLE I: Conditional joint probability mass function: $p(y, y_R|x)$, where $\|\mathcal{X}\| = \|\mathcal{Y}\| = \|\mathcal{Y}_R\| = 5$. Note that $s(p(y, y_R|x))$ equals to + when $p(y, y_R|x) > 0$ (actual value is unimportant) and 0, otherwise.

1) *Compute all possible codebooks*: first enumerate all maximal independent sets of $G_{X|Y, Y_R}$. The confusability graph $G_{X|Y, Y_R}$ is shown on the left of Fig. 4, and this has two maximal independent sets (which are used as codebooks for zero-error communication), $\mathcal{K}_1 = \{1, 3, 4, 5\}$ and $\mathcal{K}_2 = \{2, 3, 4, 5\}$.

2) *Compute color-and-forward graph $G_R|_{\mathcal{K}_1}$* : When codebook $\mathcal{K}_1 = \{1, 3, 4, 5\}$ is chosen, $G_R|_{\mathcal{K}_1}$ and one possible minimum coloring are shown in the middle of Fig. 4. It is easy to verify that this relaying function recovers all $x \in \mathcal{K}_1$ perfectly, and does so with only 3 colors.

3) *Compute color-and-forward graph $G_R|_{\mathcal{K}_2}$* : When codebook $\mathcal{K}_2 = \{2, 3, 4, 5\}$ is chosen, $G_R|_{\mathcal{K}_2}$ and one possible minimum coloring are shown on the right of Fig. 4. It is easy to verify that this relaying function recovers all $x \in \mathcal{K}_2$ perfectly, and does so with only 2 colors.

4) *Compute $r_z^{*(1)}$* : We thus conclude that $r_z^{*(1)} = \log \min\{\chi(G_R|_{\mathcal{K}_1}), \chi(G_R|_{\mathcal{K}_2})\} = \log \min\{3, 2\} = 1$. We note that in order to have the smallest number of colors (smallest relay rate) we must use \mathcal{K}_2 and not \mathcal{K}_1 .

B. A class of PRC channels where no information lossless compression is possible at the relay

In this subsection, we study a class of PRC channels where, to achieve the SIMO bound, the relay has to forward everything that it has observed regardless of the blocklength n .

Proposition 8. *For the class of zero-error PRCs which satisfy:*

- (a) $G_{X|Y, Y_R}$ is edge-free
- (b) $G_R^{(1)}|_{\mathcal{X}}$ is fully connected

then $r_z^ = r_z^{*(n)} = \log \|\mathcal{Y}_R\|$, for any n .*

This statement is strong in the sense that it shows that a one-shot scheme is optimal and that multiple channel uses cannot increase the capacity or decrease the required R-D link rate: 1) the SIMO upper bound rate satisfies $SIMO = SIMO(n) = \log \|\mathcal{X}\|$, for any n ; 2) the minimum required R-D link rate to achieve the SIMO upper bound message rate satisfies $r_z^* = r_z^{*(n)} = \log \|\mathcal{Y}_R\|$, for any n . Figure 5 shows an example of such a channel.

Proof of Proposition 8. Condition (a) implies that \mathcal{X} itself is the (unique) maximum independent set for graph $G_{X|Y, Y_R}$, and \mathcal{X}^n is the unique maximum independent set for $[G_{X|Y, Y_R}]^{\boxtimes n}$. Applying Lemma 6, and by Condition (b), we

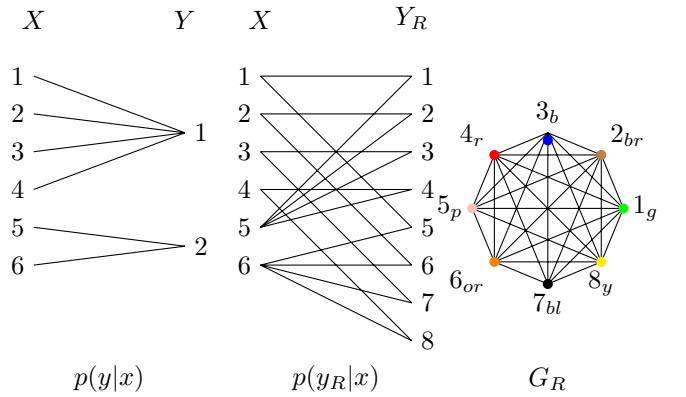


Fig. 5: An example where compression at the relay is impossible. Here, $p(y, y_R|x) = p(y|x)p(y_R|x)$.

$s(p(y, y_R x))$	Y_R					Y_R				
	1	2	3	4	5	1	2	3	4	5
Y	1	+	0	0	0	0	+	0	0	+
	2	0	+	0	+	0	0	+	0	0
	3	0	0	0	+	0	0	0	0	+
	$X = 1$					$X = 2$				

TABLE II: Conditional joint probability mass function: $p(y, y_R|x)$ for Remark 4.

know that graph $G_R^{(n)}|_{\mathcal{X}^n}$ is the same as graph $[G_R^{(1)}|_{\mathcal{X}}]^{\boxtimes n}$, which is fully-connected, for any n . Thus, $r_z^* = r_z^{*(n)} = \log \|\mathcal{Y}_R\|$, for any n . \square

Remark 4. *We note that even if $SIMO(1) = SIMO(2)$, that $r_z^{*(2)}$ can be strictly smaller than $r_z^{*(1)}$, that is, compressing over blocks can reduce the minimal R-D link rate needed. To see this, consider a relay channel for which $G_{X|Y, Y_R}$ is edge free (hence the entire \mathcal{X} is a maximal independent set and $SIMO(1) = SIMO(2) = SIMO(n) = \log(\|\mathcal{X}\|)$ for any n), and for which the relay compression graph $G_R^{(1)}$ is a pentagon. An example of such primitive relay channel distribution $p(y, y_R|x)$ is provided in Table II. In this case, 1 channel use will require a R-D link rate of $\log(3)$ bits, while two will only require a link-rate of $\frac{1}{2} \log(7) < \log(3)$ bits.*

VII. CONNECTIONS AND CONCLUSIONS

In this section, we compare our color-and-forward, or relaying compression graph with Witsenhausen's graph on source

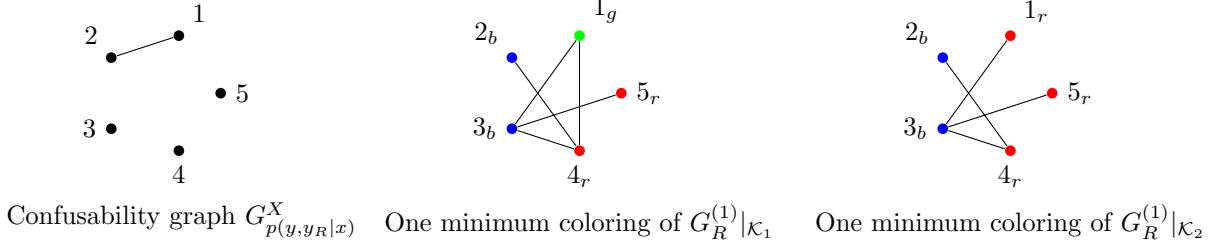


Fig. 4: Construction of confusability graph $G_{X|Y,Y_R}$ and two relaying compression graphs $G_R^{(1)}|_{\mathcal{K}_1}$ and $G_R^{(1)}|_{\mathcal{K}_2}$ for its two maximal independent sets \mathcal{K}_1 and \mathcal{K}_2 .

coding problem [11] and conclude this paper by presenting open questions.

A. Connection with Witsenhausen's source coding problem [11]

Readers might be reminded of Witsenhausen's work in [11], as shown in Fig. 6, where a point-to-point zero-error source coding problem with correlated side information available only at the receiver's end is studied. Using our notation (but with a subscript W to indicate the distributions in Witsenhausen's problem), the problem solved in [11, Section III, Proposition 2] may be posed as follows. Given are random variables (X, Y) , distributed i.i.d. according to $p_W(x, y)$. The problem is to transmit $X \in \mathcal{X}$, to a receiver which has knowledge of $Y \in \mathcal{Y}$ (the side information) by means of a discrete signal W_R (notation used to parallel our problem) taking as few values as possible. The solution is to transmit a minimal coloring of graph G_X with vertices \mathcal{X} and edges between $x \neq x'$ if $\exists y : p_W(y|x) \cdot p_W(y|x') > 0$. Their main result is:

Proposition 9. [11, Proposition 2] *When Y is not known at the transmitter, the minimum signal alphabet size, for encoding a sequence of n independent pairs with $n \geq 1$, is $\chi(G_X^{\boxtimes n})$.*

The $r_z^{*(n)}$ obtained in Corollary 5 for the case $Y_R = X$ coincides with the answer in Proposition 9. To see this, consider a PRC for which $p(y, y_R|x)$ is such that $p(y_R|x) = 1$ whenever $y_R = x$ (i.e. random variable $Y_R = X$), and for which the marginal $\sum_{y_R} p(y, y_R|x)$ is equal to the $p_W(y|x)$ of Witsenhausen⁴. Since $Y_R = X$, we know that graph $G_{X|Y,Y_R}$ is edge-free, and hence the maximal independent set for any n is \mathcal{X}^n (and hence the minimization in Corollary 5 disappears). In this case, our solution in Corollary 5 obtains the smallest relay rate, for any n , for which the network message rate attains the maximal $\frac{1}{n} \log \|\mathcal{X}\|$. We now show that our construction $G_R^{(n)}|_{\mathcal{X}^n} = G_X^{\boxtimes n}$, where $G_X^{\boxtimes n}$ denotes the n fold strong product of the graph G_X constructed by Witsenhausen. Note that both $G_R^{(n)}|_{\mathcal{X}^n}$ and $G_X^{\boxtimes n}$ have vertices \mathcal{X}^n . Next, note that $x^n := (x_1, x_2, \dots, x_n) \neq (x'_1, x'_2, \dots, x'_n) =: x'^n \subset \mathcal{X}^n$ share an edge if there exists a y^n such that $p(y^n|x^n) \cdot p(y^n|x'^n) > 0$. Since the channels are memoryless i.e. $p(y^n|x^n) = \prod_{i=1}^n p(y_i|x_i)$, this means $G_R^{(n)}|_{\mathcal{X}^n} = G_X^{\boxtimes n}$,

⁴Note that only $p_W(y|x)$ matters in Witsenhausen's Proposition 9, and hence the choice of $p(x)$ in our problem is immaterial.

and hence the $r_z^{*(n)}$ obtained in Corollary 5 for the case $Y_R = X$ coincides with the answer in Proposition 9.

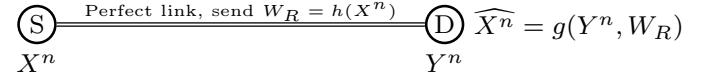


Fig. 6: Witsenhausen's problem: How to transmit X^n to a receiver which has knowledge of Y^n (the side information) by means of a discrete signal W_R taking as few values as possible. (X^n, Y^n) are generated i.i.d. according to $p_W(x, y)$.

We note that in general, finding the minimum R-D link rate $r_z^{*(n)}$ is different from Witsenhausen's source coding problem with receiver side information. This is because not all PRCs have $Y_R = X$ with probability 1. In general, the SIMO bound is not the absolute maximum $\log \|\mathcal{X}\|$, and hence there may be more than one choice of maximal independent sets of the confusability graph $G_{X|Y,Y_R}^{\boxtimes n}$. Since for each n , $G_R^{(n)}|_{\mathcal{K}^{(n)}}$ is constructed from a maximal independent set $\mathcal{K}^{(n)}$ of $G_{X|Y,Y_R}^{\boxtimes n}$, and these maximal independent sets in general do not form a "nice" structure [18], it follows that in general $G_R^{(n)}|_{\mathcal{K}^{(n)}} \neq [G_R^{(1)}|_{\mathcal{K}}]^{\boxtimes n}$, and cannot be constructed via any standard graph product operations surveyed in [19]. Lemma 6 and the surrounding remarks comment on this.

B. Conclusions and future work

In this paper, the problem of communicating over a primitive relay channel without error is for the first time proposed. A relaying scheme termed "color-and-forward" is proposed. This scheme is shown to be the most efficient way to compress signals at the relay terminal, for any fixed number of channel uses, to achieve the single-input multi-output (SIMO) upper bound. We also provide bounds on the asymptotic r_z^* – the minimum required R-D link rate such that the given PRC channel can achieve the SIMO upper bound message rate. The general relationship between $r_z^{*(n)}$ and r_z^* depends on the behavior of the sequence of SIMO bounds $\{SIMO(n)\}$. This is a sequence of normalized independence numbers of the strong products of a graph specified by the broadcasting component of the PRC, whose behavior is a long-standing open question [16], [17]. We believe understanding the behavior of the color-and-forward graphs $G_R^{(n)}|_{\mathcal{K}^{(n)}}$ plays a central role in determining how the relay terminal can contribute to the overall communication.

REFERENCES

- [1] T. Cover, *Open Problems in Communication and computation*. Springer-Verlag, 1987, ch. The capacity of the relay channel, pp. 72–73.
- [2] Y. Kim, “Coding techniques for primitive relay channels,” in *Proc. Forty-Fifth Annual Allerton Conf. Commun., Contr. Comput.*, 2007.
- [3] X. Wu, L. P. Barnes, and A. Özgür, “Cover’s open problem: ‘The capacity of the relay channel’,” <https://arxiv.org/abs/1701.02043>, 2017.
- [4] X. Wu and A. Özgür, “Cut-set bound is loose for gaussian relay networks,” [arXiv:1608.08540v2](https://arxiv.org/abs/1608.08540v2), 2016.
- [5] T. M. Cover and A. El Gamal, “Capacity theorems for the relay channel,” *IEEE Trans. Inf. Theory*, vol. 25, no. 5, pp. 572–584, Sep. 1979.
- [6] A. Gamal and M. Aref, “The capacity of the semideterministic relay channel (corresp.),” *IEEE Trans. Inf. Theory*, vol. 28, no. 3, pp. 536–536, May 1982.
- [7] A. El Gamal and S. Zahedi, “Capacity of a class of relay channels with orthogonal components,” *IEEE Trans. Inf. Theory*, vol. 51, no. 5, pp. 1815–1817, May 2005.
- [8] Y. Kim, “Capacity of a class of deterministic relay channels,” *IEEE Trans. Inf. Theory*, vol. 54, no. 3, pp. 1328–1329, 2008.
- [9] 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.
- [10] 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.
- [11] H. Witsenhausen, “The zero-error side information problem and chromatic numbers (corresp.),” *IEEE Trans. Inf. Theory*, vol. 22, no. 5, pp. 592–593, 1976.
- [12] L. Lovasz, “On the shannon capacity of a graph,” *IEEE Trans. Inf. Theory*, vol. 25, no. 1, pp. 1–7, 1979.
- [13] D. E. Knuth, “The sandwich theorem,” *Electron. J. Combin.*, vol. 1, 1994.
- [14] J. Korner and A. Orlitsky, “Zero-error information theory,” *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2207–2229, 1998.
- [15] C. Shannon, “The zero error capacity of a noisy channel,” *Information Theory, IRE Transactions on*, vol. 2, no. 3, pp. 8–19, September 1956.
- [16] E. Sonnemann and O. Krafft, “Independence numbers of product graphs,” *Journal of Combinatorial Theory, Series B*, vol. 17, no. 2, pp. 133 – 142, 1974.
- [17] P. Jha and G. Slutzki, “Independence numbers of product graphs,” *Applied Mathematics Letters*, vol. 7, no. 4, pp. 91 – 94, 1994.
- [18] 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.
- [19] S. Klavzar, “Coloring graph products: A survey,” *Discrete Mathematics*, vol. 155, no. 1 - 3, pp. 135 – 145, 1996.