# On the Optimality of Colour-and-Forward Relaying for a Class of Zero-error Primitive Relay Channels

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

***Abstract*—Recently a new “Colour-and-Forward” relaying strategy was proposed for the zero-error primitive relay channel, a relay channel in which the relay to destination link is out of band and of fixed, error-free capacity. This “Colour-and-Forward” scheme forwards the colour (from a minimum colouring) of the node corresponding to its received signal. This colouring is of a carefully designed graph based on the joint distribution of the relay and destination outputs given the transmit signal. This scheme was used to provide a non-trivial upper bound on the minimum required conference link capacity to allow the overall network to achieve the single-input multi-output (SIMO) upper bound. In this paper, we strengthen the result and show that this upper bound is tight if one wants to achieve the SIMO bound in the overall network for any fixed number of channel uses.**

## I. Introduction

As shown in Figure 1, a primitive relay channel (PRC) $((\mathcal{X}, p(y, y_R|x), \mathcal{Y} \times \mathcal{Y}_R),\ r)$ consists of a source terminal S that wants to communicate a message $W$ to a destination terminal D aided by a relay terminal R. 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 orthogonal to the error-free conferencing link with maximum rate $r$ bits / channel use from the relay to the destination terminal. This channel model is motivated when a relay terminal cannot simultaneously transmit and receive signals or when the relay has an out-of-band link to the destination.

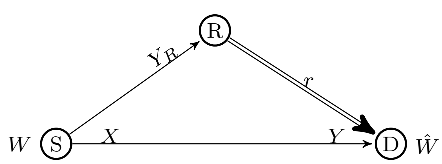



Fig. 1. A primitive relay channel $((\mathcal{X}, p(y, y_R|x), \mathcal{Y} \times \mathcal{Y}_R),\ r)$.

The question we are interested in is how to operate the relay terminal to achieve the maximal possible network message rate while using the least number of bits on the conference link. It has been shown in [1] that relaying strategies like “decode-and-forward”, “compress-and-forward” and “hash-and-forward” are all sub-optimal in general. The core function of the relay is to help the destination in disambiguating the channel inputs, i.e. to provide “what the destination needs”. The relay need not decipher the channel inputs (messages) nor transmit what the destination can infer about the channel inputs from its own received signals.

The work of the authors was partially funded by NSF under awards 1216825 and 1053933. The contents are solely the responsibility of the author and do not necessarily reflect the official views of the NSF.

What the relay should forward depends on both the broadcasting links and the allowable conference rate $r$. When $r$ is infinite or *large enough*, the relay can simply forward everything it has observed to the destination terminal. Thus, the primitive relay channel effectively turns into a point-to-point channel $(\mathcal{X}, p(y, y_R|x), \mathcal{Y} \times \mathcal{Y}_R)$, whose capacity is known. The natural question to ask is how large the conferencing link capacity $r$ should be to ensure that the PRC network can achieve the capacity of the point-to-point channel $(\mathcal{X}, p(y, y_R|x), \mathcal{Y} \times \mathcal{Y}_R)$. We denote this capacity as the *single-input multi-output (SIMO) upper bound* for the given PRC channel. When conference rate $r$ is big enough such that the SIMO upper bound can be achieved, we say that an “effectively full cooperation” between the relay and destination terminals can be established.

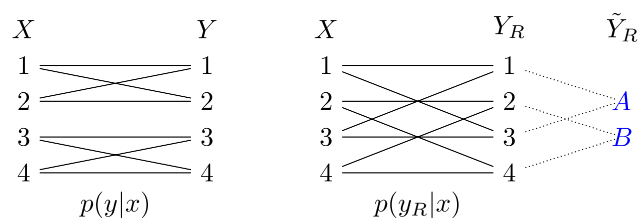



Fig. 2. Toy Problem: $p(y, y_R|x) = p(y|x)p(y_R|x)$. A solid link indicates the probability value $p(*|x)$ is positive, where $*$ indicates $y$ or $y_R$.

Take for example a PRC with $p(y, y_R|x) = p(y|x)p(y_R|x)$ as in Figure 2. The destination, upon receiving $Y$ can tell whether $\{1, 2\}$ or $\{3, 4\}$ were sent, but not which message within those sets. The relay can “provide the destination what it needs” by forwarding $A$ or $B$, i.e. whether the X was even or odd. This amounts to considerable savings for the conferencing link capacity with respect to sending $Y_R$ directly, and allows the destination to fully resolve which $X$ was sent as long as the conferencing link capacity is at least 1 bit.

The Colour-and-Forward scheme first introduced in [2] generalizes this example. Here, we will show that it yields the smallest conferencing link rate *(for any fixed number of channel uses)* that will render the channel effectively fully co-operative, achieving the SIMO bound. It may be checked that this simple channel does not fall into a class of PRCs for which capacity is known, i.e. it is not a degraded, semideterministic, orthogonal-component, or semideterministic PRC.

The threshold for the conferencing link capacity highly depends on the relationship between $Y$ and $Y_R$, or the structure of the conditional joint distribution $p(y, y_R|x)$. The *small-error* [1] version of this question was first proposed in [1] and remains open. The *zero-error* version was studied in [2], which explicitly explored the channel structure using graph theoretic notations and proposed a novel relaying algorithm which depended highly on the structure of $p(y, y_R|x)$ termed "Colour-and-Forward". This scheme thus provided an upper bound on the minimum conferencing link capacity which would render the channel effectively fully cooperative.

**Contribution.** In this paper, we strengthen our prior work and show that the Colour-and-Forward algorithm yields the *smallest* conferencing link capacity required to achieve the SIMO upper bound for any given $n$, i.e. that the upper bound in [2] is tight. This is significant given that relay channel capacity results are rare. The main result is presented in Section IV in Theorem 2. One of the key steps in the converse is the observation and utilization of the zero-error data-processing inequality in Lemma 3. We further discuss connections between the Colour-and-Forward graph and the Witsenhausen graph and state a conjecture about optimality as $n \to \infty$.

## II. Zero-error communication over a primitive relay channel

Note that the Colour-and-Forward relaying algorithm [2] was developed in the context of communication over a PRC without error. Zero-error communication naturally leads to a problem formulation in terms of graphs. We begin this section with some useful graph-theoretic concepts and notation. Next, a preliminary introduction on zero-error point-to-point communication is provided, before the problem of zero-error communication over a PRC is formally defined. Throughout the paper we will use subscripts $_z$ to emphasize this zero-error context. We use upper and lower cases to differentiate the overall network message rate $R_z$ and the conference rate $r_z$. 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 2-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 *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 *colouring* 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 colours in any colouring. A *minimum colouring* of graph $G$ uses $\chi(G)$ colours.

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", that is, if $\exists y \in \mathcal{Y} : p(y|x) \cdot p(y|x') > 0$. For a given conditional probability function $p(y|x)$, we denote $S_{X|Y}(y) := \{x : p(y|x) > 0\}$ as the *conditional support of* $Y = y$. Thus, the confusability graph $G_{X|Y}$ can be equivalently constructed by fully connecting the nodes inside each conditional support $S_{X|Y}(y)$, for all $y \in \mathcal{Y}$.

### B. Zero-error preliminaries

The zero-error capacity of a point-to-point discrete memoryless channel was initially studied by Shannon in [3] in 1956; see [4], [5] for further zero-error capacity details.

Consider zero-error communication over a point-to-point channel $(\mathcal{X}, p(y|x), \mathcal{Y})$. First, note that only whether $p(y|x)$ is zero or not matters for communication without error. Next, consider first communicating over 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, we know that $\alpha(G_{X|Y}^{\boxtimes n})$ is the number of distinguishable channel inputs $X^n$, where $G_{X|Y}^{\boxtimes n}$ is the strong product of $n$ copies of graph $G_{X|Y}$.[2] The zero-error capacity is then characterized as [4]

$$\lim_{n\to\infty} \frac{1}{n} \log \alpha(G_{X|Y}^{\boxtimes n}) = \lim_{n\to\infty} \log \sqrt[n]{\alpha(G_{X|Y}^{\boxtimes n})},$$

which may be upper and lower bounded as [3], [4]:

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

where $\|\mathcal{X}\|$ is the cardinality of the input alphabet, which is the maximal number of possible inputs per channel use. Note the limit exists by Lemma 1.

**Lemma 1.** *Let $G_{X|Y}$ denote the confusability graph specified by $p(y|x)$. Then the sequence $\{\log \sqrt[n]{\alpha(G_{X|Y}^{\boxtimes n})}\}_{n=1}^{\infty}$ converges to $\sup\{\log \sqrt[n]{\alpha(G_{X|Y}^{\boxtimes})}, n = 1, 2, \cdots\}$.*

[1]Communication allowing a vanishing probability of error is called *small-error* or *$\epsilon$−error* communication, while communication without error is called *zero-error* or *0-error* communication.

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

*Proof.* It can be checked that the sequence $\{\alpha(G_{X|Y}^{\boxtimes n})\}_{n=1}^{\infty}$ is super-multiplicative, i.e. $\alpha(G_{X|Y}^{\boxtimes(n_1+n_2)}) \geq \alpha(G_{X|Y}^{\boxtimes n_1})\cdot\alpha(G_{X|Y}^{\boxtimes n_2})$ for any indices $n_1, n_2$. Thus, the sequence $\{\log\alpha(G_{X|Y}^{\boxtimes n})\}_{n=1}^{\infty}$ is super-additive and each item is non-negative. By Fekete's Lemma, the limit $\lim_{n\to\infty} \log \sqrt[n]{\alpha(G_{X|Y}^{\boxtimes n})}$ exists and is equal to $\sup\{\log \sqrt[n]{\alpha(G_{X|Y}^{\boxtimes n})}, n = 1, 2, \cdots\}$. □

### C. Zero-error communication over a primitive relay channel

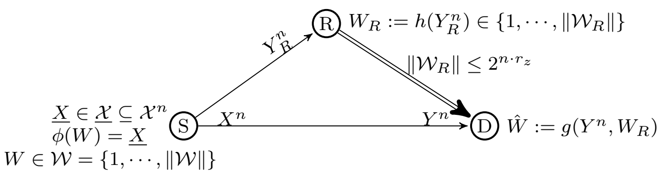



Fig. 3. An $n$-shot protocol $(n, \underline{\mathcal{X}}, h, g)$ for zero-error communication over a PRC $((\mathcal{X}, p(y, y_R|x), \mathcal{Y} \times \mathcal{Y}_R), r_z)$, with an encoder $\phi$, a codebook $\underline{\mathcal{X}}$, a relaying function $h$ and a decoding function $g$.

As shown in Figure 3, an *n-shot protocol* $(n, \underline{\mathcal{X}}, h, g)$ for zero-error communication over a PRC $((\mathcal{X}, p(y, y_R|x), \mathcal{Y} \times \mathcal{Y}_R), r_z)$ is composed of a codebook $\underline{\mathcal{X}} \subseteq \mathcal{X}^n$, a $r_z$-admissible relaying function $h : \mathcal{Y}_R^n \to \mathcal{W}_R$ which satisfies $\|\mathcal{W}_R\| \leq 2^{n\cdot r_z}$ and a decoding function $g : \mathcal{Y}^n \times \mathcal{W}_R \to \underline{\mathcal{X}}$. Let $\hat{\underline{X}}$ and $\hat{W}$ demote the estimate for codeword $\underline{X}$ and the message $W$ respectively. Note that $\hat{W} = \phi^{-1}(\hat{\underline{X}}) = \phi^{-1}(g(Y^n, W_R))$. Because the mapping $\phi(\,)$ is bijective, decoding message $W \in \mathcal{W}$ is equivalent to decoding codeword $\underline{X} \in \underline{\mathcal{X}}$. We will not distinguish these two concepts and abuse notation $\hat{w} \in \underline{\mathcal{X}}$ and $\hat{W} = g(Y^n, W_R)$ for the decoding result at the destination.

A message rate $R_z := \frac{1}{n}\log\|\mathcal{W}\| = \frac{1}{n}\log\|\underline{\mathcal{X}}\|$ is *achievable* if there exists an $n$-shot protocol $(n, \underline{\mathcal{X}}, h, g)$ over a PRC $((\mathcal{X}, p(y, y_R|x), \mathcal{Y} \times \mathcal{Y}_R), r_z)$ achieving zero error, i.e. $\Pr[g(y, w_R) \neq w] = 0$ for all values $w \in \underline{\mathcal{X}}$. The capacity $C_z(r_z)$ of zero-error communication over a PRC $((\mathcal{X}, p(y, y_R|x), \mathcal{Y} \times \mathcal{Y}_R), r_z)$ is the supremum of all possible achievable rates $R_z$ for any $n$. Clearly, $C_z(r_z)$ is at most $\log\|\mathcal{X}\|$.

## III. Minimum conference rate $r_z^*$ and Colour-and-Forward relaying algorithm

In Definition 1, we formally present our problem: finding the minimum required conference rate $r_z^*$ that can enable an effectively full cooperation between the relay and destination terminals. We then state the Colour-and-Forward relaying algorithm [2] in Definitions 2 and 3. In [2], it was further shown that the Colour-and-Forward relaying algorithm is information lossless (Theorem 2 of [2]) and gave a novel upper bound $T_u^{(n)}$ on $r_z^{*(n)}$ (defined below and in Theorem 2). In Section IV, we will strengthen this and show that this upper bound is tight.

### A. The minimum conference rate $r_z^*$

Allowing the relay and destination terminals to fully cooperate, for any fixed number of channel uses $n$, the network message rate can be upper bounded by the maximum achievable rate of the virtual channel $(\mathcal{X}, p(y, y_R|x), \mathcal{Y} \times \mathcal{Y}_R)$:

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

$C_z^{(n)}(r_z)$ and $C_z^{(n)}(\infty)$ denote the ($n$-shot) [3] zero-error PRC channel capacity with conference link of some finite ($r_z$) or infinite rates respectively, while the broadcasting links $(\mathcal{X}, p(y, y_R|x), \mathcal{Y} \times \mathcal{Y}_R)$ remain the same.

Define $C_z(r_z) = \sup_n C_z^{(n)}(r_z)$. Thus, we have $C_z(r_z) \leq \sup_n C_z^{(n)}(\infty)$. By Lemma 1,

$$\sup_n C_z^{(n)}(\infty) = \lim_{n\to\infty} C_z^{(n)}(\infty) =: C_z(\infty).$$

Formally, the minimum conference rate $r_z^*$ that can enable an effectively full cooperation between the relay and destination terminals can be defined as:

**Definition 1** (The minimum conference rate $r_z^*$)**.**

$$r_z^* := \inf\{r_z : C_z(r_z) = C_z(\infty)\}. \tag{1}$$

We approach finding $r_z^*$ by exploring $r_z^{*(n)}$, the minimum conference rate for some fixed number of channel uses:

$$r_z^{*(n)} := \inf\{r_z : C_z^{(n)}(r_z) = C_z^{(n)}(\infty)\}. \tag{2}$$

We conclude the characterization of $r_z^{*(n)}$ in Theorem 2 and discuss our Conjecture 4 on $r_z^*$ at the end of Section IV.

### B. Colour-and-Forward relaying algorithm

The Colour-and-Forward relaying algorithm is defined as a minimum coloring function on the Colour-and-Forward graph, as shown in Definition 3 and Definition 2, for any $n$ channel uses. Note that the bold font is adopted to indicate a sequence of length $n$. 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.

**Definition 2** (Colour-and-Forward graph $G_R^{(n)}$)**.** *Given a conditional joint pmf $p(\mathbf{y}, \mathbf{y}_R|\mathbf{x})$ with support $\mathcal{X}^n$ and output $\mathcal{Y}^n \times \mathcal{Y}_R^n$, graph $G_R^{(n)}$ is an undirected graph with vertex set $\mathcal{Y}_R^n$ and an edge $\mathbf{y}_{R1} - \mathbf{y}_{R2}$ is imposed when for some $\mathbf{y}$, $\mathbf{x}_1 \neq \mathbf{x}_2$, $\Pr(\mathbf{Y} = \mathbf{y}, \mathbf{Y}_R = \mathbf{y}_{R1}|\mathbf{X} = \mathbf{x}_1)\cdot\Pr(\mathbf{Y} = \mathbf{y}, \mathbf{Y}_R = \mathbf{y}_{R2}|\mathbf{X} = \mathbf{x}_2) > 0$.*

**Definition 3** (Colour-and-Forward relaying $W_R^{*(n)}$)**.** *Given a conditional joint pmf $p(\mathbf{y}, \mathbf{y}_R|\mathbf{x})$ with support $\mathcal{X}^n$ and output $\mathcal{Y}^n \times \mathcal{Y}_R^n$, we define the Colour-and-Forward relaying $W_R^{*(n)}$ as a function of $\mathbf{Y}_R$ by a minimum colouring $c$ with $\chi(G_R^{(n)})$ colours on graph $G_R^{(n)}$:*

$$W_R^{*(n)} := c(\mathbf{Y}_R)$$

*where graph $G_R^{(n)}$ is defined in Definition 2. An alternative construction is provided in [2]. (Note that $c$ is not unique.)*

[3] We use the superscript $^{(n)}$ to indicate $n$-shot channel usage.

## IV. Colour-and-Forward is optimal

We now state our main result: that the Colour-and-Forward scheme provides the most efficient compression of $\mathbf{Y}_R$'s, provided one wants to achieve effectively full cooperation between the relay and destination. Achievability was provided in [2] and is not repeated here for sake of space, the converse is the challenging and novel aspect.

**Theorem 2.** *For any fixed $n$, $r_z^{*(n)} = T_u^{(n)}$, where $r_z^{*(n)}$ is specified in (2) and $T_u^{(n)}$ is defined as*

$$T_u^{(n)} := \min_{\mathcal{K} \text{ is a maximum independent set of graph } G_{X^n|Y^n,Y_R^n}} \log \sqrt[n]{\chi(G_R^{(n)}|_{\mathcal{K}})},$$

*where $\chi(G_R^{(n)}|_{\mathcal{K}})$ is the chromatic number of graph $G_R^{(n)}|_{\mathcal{K}}$, constructed via the algorithm described in Definition 2 with restricted input / codebook $\mathcal{K}$.*

**Remark:** To give one a sense of the optimization involved, we provide an example for $n = 1$. We note that the minimization is over the different maximum independent sets of the graph $G_{X^n|Y^n,Y_R^n}$ and that different maximum independent sets may yield different conferencing link rates. To illustrate this, consider the PRC described by the joint distribution $p(y, y_R|x)$ provided in Table I. Its confusability graph, and compression graphs $G_R$ constructed by the Colour-and-Forward algorithm (for inputs in $\mathcal{X}$ or some maximum independent subsets $\mathcal{K}_1$ and $\mathcal{K}_2$ of the confusabilty graph $G_{X|Y,Y_R}$) when $n = 1$, are shown in Table II. We note that in order to have the smallest number of colours for the conferencing link we must use $\mathcal{K}_2$ and not $\mathcal{K}_1$.

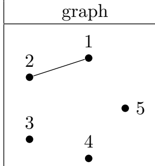
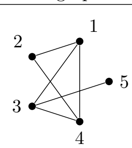
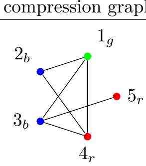
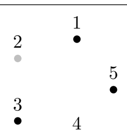
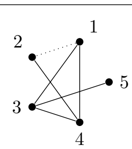
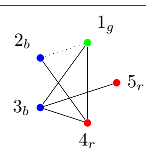
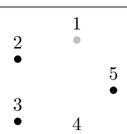
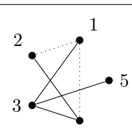
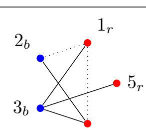
| confusability graph | compression graph | colored compression graph |
|---|---|---|
|  $\mathcal{X} = [1:5]$ |  Compression graph $G_R\|_{\mathcal{X}}$ |  $\chi(G_R\|_{\mathcal{X}}) = 3$ |
|  $\mathcal{K}_1 = [1, 3:5]$ |  Compression graph $G_R\|_{\mathcal{K}_1}$ |  $\chi(G_R\|_{\mathcal{K}_1}) = 3$ |
|  $\mathcal{K}_2 = [2, 3:5]$ |  Compression graph $G_R\|_{\mathcal{K}_2}$ |  $\chi(G_R\|_{\mathcal{K}_2}) = 2$ |

TABLE II

An example to show the impact of the choice of independent sets in Theorem 2. The conditional joint pmf $p(y, y_R|x)$ in discussion is shown in Table I.

To prove optimality of Colour-and-Forward relaying $W_R^{*(n)}$, we require the following zero-error data-processing inequality.

**Lemma 3** (Data-Processing Inequality)**.** *Given a conditional pmf $p(y|x)$, let $Z = f(Y)$ be any deterministic mapping $f : \mathcal{Y} \to \mathcal{Z}$ and denote $p(z|x)$ the induced conditional pmf from $p(y|x)$. Then the confusability graph $G_{X|Y}$ specified by $p(y|x)$ has no more edges than the confusability graph $G_{X|Z}$ specified by $p(z|x)$; i.e., $E(G_{X|Y}) \subseteq E(G_{X|Z})$.*

Recall that the zero-error capacity of a point-to-point channel $(\mathcal{X}, p(y|x), \mathcal{Y})$ is fully characterized by the confusability graph $G_{X|Y}$ and is directly related to the independence number of its $n$-fold strong product. The more densely a graph is connected, the smaller its independence number becomes. Lemma 3 states that the processed observation $Z$ cannot remove any edges from the original confusability graph $G_{X|Y}$ and could potentially add edges, indicating a potential loss of information about the underlying variable $X$. Lemma 3 and its validity follows directly from the definition of confusability graph and the nature of zero-error communication.

Now we present the proof of Theorem 2.

*Proof of Theorem 2.* Achievability follows from our prior work in [2] and is based on showing that when $T_u^{(n)}$ different colours or $W_R^{*(n)}$ can be successfully transmitted to the destination terminal, the destination terminal, together with its own observations $Y^n$, can infer as much information about $X^n$ as if $Y_R^n$ was known. That is, $G_{X^n|Y^n,W_R^{*(n)}} = G_{X^n|Y^n,Y_R^n}$, which implies $\alpha(G_{X^n|Y^n,W_R^{*(n)}}) = \alpha(G_{X^n|Y^n,Y_R^n})$, i.e. $C_z^{(n)}(\infty)$ is achieved.

To establish Theorem 2, it suffices to prove $r_z^{*(n)} \geq T_u^{(n)}$. Let $(n, \underline{\mathcal{X}}, h, g)$ denote any $n$-shot protocol that can achieve the SIMO upper-bound message rate $C_z^{(n)}(\infty)$ without error, say $R_z^{(n)} = \frac{1}{n} \log \|\underline{\mathcal{X}}\| = \log \sqrt[n]{\alpha(G_{X|Y,Y_R}^{\boxtimes n})}$. We will show that $\|\mathcal{W}_R\| \geq \|\mathcal{W}_R^{*(n)}\| = 2^{n \cdot T_u^{(n)}}$ must hold for any such relaying function $h : \mathcal{Y}_R^n \to \mathcal{W}_R$ (for any $n$).

Because rate $\frac{1}{n} \log \|\underline{\mathcal{X}}\|$ can be achieved by the given $n$-shot protocol $(n, \underline{\mathcal{X}}, h, g)$, we know that the induced subgraph $G_{X^n|Y^n,W_R}(\underline{\mathcal{X}})$ [4] must be edge-free. Recall that $W_R$ is a deterministic function of $Y_R^n$, so $(Y^n, W_R)$ is a deterministic function of $(Y^n, Y_R^n)$. According to the data-processing inequality in Lemma 3, we have $E(G_{X^n|Y^n,Y_R^n}(\underline{\mathcal{X}})) \subseteq E(G_{X^n|Y^n,W_R}(\underline{\mathcal{X}})) = \emptyset$. Thus, we know that two induced subgraphs $G_{X^n|Y^n,Y_R^n}(\underline{\mathcal{X}}))$ and $G_{X^n|Y^n,W_R}(\underline{\mathcal{X}})$ must both be free of edges. Consider any triple $(X = x^n, Y = y^n, W_R = w_R) \in \underline{\mathcal{X}} \times \mathcal{Y}^n|_{\underline{\mathcal{X}}} \times \mathcal{W}_R$, we have

$$\begin{aligned} &\Pr[Y = y^n, W_R = w_R | X = x^n] \\ &= \sum_{y_R^n :\, h(y_R^n) = w_R} \Pr[Y = y^n, Y_R^n = y_R^n | X = x^n] \end{aligned}$$

[4]Graph $G(A)$ is the induced subgraph of graph $G$, with vertex set $A \subseteq V(G)$ and edge set $(A \times A) \cap E(G)$.

| $s(p(y,y_R\|x))$ | | $Y_R$ 1 | 2 | 3 | 4 | 5 | $Y_R$ 1 | 2 | 3 | 4 | 5 | $Y_R$ 1 | 2 | 3 | 4 | 5 | $Y_R$ 1 | 2 | 3 | 4 | 5 | $Y_R$ 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 0 | 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 | 0 |
| $Y$ | 3 | 0 | 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 | 0 |
| | 5 | 0 | 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.

Thus,

$$\begin{aligned}
&S_{X^n|Y^n,W_R}(y^n,w_R)=\{x^n\in\underline{\mathcal{X}}:p|_{\underline{\mathcal{X}}}(y^n,w_R|x^n)>0\}\\
=&\{x^n\in\underline{\mathcal{X}}:\sum_{y_R^n:\,h(y_R^n)=w_R}\Pr[Y=y^n,Y_R^n=y_R^n|X=x^n]>0\}\\
=&\bigcup_{y_R^n:\,h(y_R^n)=w_R}\{x^n\in\underline{\mathcal{X}}:\Pr[Y=y^n,Y_R^n=y_R^n|X=x^n]>0\}\\
=&\bigcup_{y_R^n:\,h(y_R^n)=w_R}S_{X^n|Y^n,Y_R^n}(y^n,y_R^n)
\end{aligned} \tag{3}$$

$S_{X^n|Y^n,Y_R^n}(y^n,y_R^n)$ has zero or one element because graph $G_{X^n|Y^n,Y_R^n}(\underline{\mathcal{X}})$ has no edges. Similarly, since graph $G_{X^n|Y^n,W_R}(\underline{\mathcal{X}})$ is edge-free, $S_{X^n|Y^n,W_R}(y^n,w_R)$ shall also at most have one element. So in equation (3), the sets to be unioned can have 0 or 1 element and all non-empty sets shall be the same, i.e., containing one same element. This means that for any fixed $Y^n=y^n$, any two different $y_R^n$'s such that $S_{X^n|Y^n,Y_R^n}(y^n,y_{R1}^n)$ and $S_{X^n|Y^n,Y_R^n}(y^n,y_{R2}^n)$ (which are both either an empty set or a single-element set) have different elements, say $x_1^n$ and $x_2^n$, are prohibited to be mapped into the same color $w_R$. That is, requiring two $y_R^n$'s to be differentiated (via the relaying function $h$) if for some $y^n$, $x_1^n\neq x_2^n$, $\Pr(Y=y^n,Y_R^n=y_{R1}^n|X=x_1^n)\cdot\Pr(Y=y^n,Y_R=y_{R2}^n|X=x_2^n)>0$, is necessary. Equivalently, all edges in the compression graph $G_R^{(n)}$ constructed in the Colour-and-Forward algorithm are necessary; any other valid relay mapping $W_R=h(Y_R^n)$ would result in equally or more strict edge constraints than Colour-and-Forward or graph $G_R^{(n)}$. Note that as more edges are added to a graph, its chromatic number cannot decrease. Therefore, for any valid relay mapping $W_R$, we have $\|\mathcal{W}_R\|\geq\|\mathcal{W}_R^{*(n)}\|$, implying $r_z^{*(n)}\geq T_u^{(n)}$. $\square$

**Connection with other problems.** When $Y_R=X$ with probability 1, we know that graph $G_{X|Y,Y_R}$ is edge-free. In this case, with a large enough conference rate, we can achieve overall network message rate $\log\|\mathcal{X}\|$. This also implies that the channel input codebook has to be the full channel input alphabet $\mathcal{X}$, say$\underline{\mathcal{X}}=\mathcal{K}=\mathcal{X}$. In this specific case, finding $r_z^*$ may be mapped to the source coding problem with receiver side information, which was solved by Witsenhausen [6] and in this case, coincides with the result presented here.

We note that in general, finding the minimum conference rate $r_z^*$ is different from Witsenhausen's source coding problem with receiver side information. This is because: 1) not all PRCs have $Y_R=X$; 2) when the SIMO bound is not the absolute maximum $\log\|\mathcal{X}\|$, only some subset of the channel input alphabet can be transmitted and there may be more than one choice of channel input codebook, as seen in the example in Tables I and II; and 3) in general, $G_R^{(n)}$ is not a $n$-fold strong product of graph $G_R^{(1)}$, i.e. $G_R^{(n)}\neq[G_R^{(1)}]^{\boxtimes n}$, and cannot be constructed via any standard graph product operations surveyed in [7].

The Colour-and-Forward compression graphs $G_R^{(n)}$s' behavior and chromatic numbers as a function of $n$ is not obvious, and is the subject of ongoing work. Understanding this behavior will allow us to go beyond the restriction to $n$ channel uses in Theorem 2 and characterize $r_z^*$. As such, we conservatively propose the following conjecture.

**Conjecture 4.** *$r_z^*=\lim_{n\to\infty}T_u^{(n)}$, where $r_z^*$ is specified in* (1) *and $T_u^{(n)}$ is defined as in Theorem 2.*

## V. CONCLUSION

We have demonstrated that the recently introduced Color-and-Forward algorithm for the zero-error primitive relay channel, is able to optimally – for any fixed number of channel uses, requiring the smallest conferencing link capacity – compress the relay signal if one desires to achieve the SIMO upper bound. The central contribution was the demonstration of the converse. Connections with other problems such as coding for computing [8], understanding the behavior of the compression graph $G_R^{(n)}$, and extensions of the zero-error Color-and-Forward scheme to the $\epsilon$-error setting are the subject of current work.

## REFERENCES


[1] Y. Kim, "Coding techniques for primitive relay channels," in Proc. Allerton Conf. Commun., Contr., Comput., 2007, p. 2007.
[2] Y. Chen, S. Shahi, and N. Devroye, "Colour-and-forward: relaying "what the destination needs" in the zero-error primitive relay channel," in Proc. Allerton Conference on Commun., Contr., and Comput., 2014.
[3] C. Shannon, "The zero error capacity of a noisy channel," Information Theory, IRE Transactions on, vol. 2, no. 3, pp. 8–19, September 1956.
[4] L. Lovasz, "On the shannon capacity of a graph," Information Theory, IEEE Transactions on, vol. 25, no. 1, pp. 1–7, 1979.
[5] J. Korner and A. Orlitsky, "Zero-error information theory," IEEE Trans. Inf. Theory, vol. 44, no. 6, pp. 2207–2229, 1998.
[6] H. Witsenhausen, "The zero-error side information problem and chromatic numbers (corresp.)," IEEE Trans. Inf. Theory, vol. 22, no. 5, pp. 592–593, 1976.
[7] S. Klavzar, "Coloring graph products: A survey," Discrete Mathematics, vol. 155, no. 1?3, pp. 135 – 145, 1996, combinatorics.
[8] A. Orlitsky and J. Roche, "Coding for computing," IEEE Trans. Inf. Theory, vol. 47, no. 3, pp. 903–917, 2001.