Colour-and-Forward: relaying “what the destination needs” in the zero-error primitive relay channel

Yanying Chen, Sara Shahi, Natasha Devroye
University of Illinois at Chicago, Chicago IL 60607, USA

***Abstract*— Zero-error communication over a primitive relay channel is for the first time proposed and studied. This model is used to highlight how one may exploit the channel structure to design a relaying strategy that explicitly provides “what destination needs”. We propose the *Colour-and-Forward* relaying scheme which constructs a graph G_R of relay outputs based on the joint conditional distribution of the relay and destination outputs given the channel input. The colours of this graph G_R are sent over the out-of-band link in the primitive relay channel and are shown to be information lossless in the zero-error sense; they result in the same confusability graph as if the destination had the relay’s received signal. This allows us to obtain an upper bound[1] on the minimum required conference rate required for the relay and destination terminals to be effectively fully cooperative for any number of channel use n. It also leads to an achievable zero-error communication rate for the primitive relay channel, which may be shown to be capacity for a class of channels.**

I. Background and motivation

Motivation. The core function of a relay is to help the destination in disambiguating the inputs, i.e. to provide “what the destination needs”. A relay’s goal is not to decode the message - this is why Decode-and-Forward fails in general; it is not to provide “what the destination does not want”, i.e. the noise, - this is why Amplify-and-Forward fails in general; nor is it desirable to waste its communication to send “what destination already possesses”. One might argue that Partial Decode-and-Forward and Compress-and-Forward embody the idea of providing “what the destination needs” to some extent. However, we are not aware of any *explicit* attempt to characterize and quantify this intuition, which could potentially lead to a new relaying strategy with improved rates.

In this paper, we attempt to quantify intuition about relaying “what the destination needs” in the context of communicating over a primitive relay channel (PRC) without error, because 1) PRC is the simplest [2] network that contains a relay and 2) the imposition of zero-error constraint turns the problem into a combinatorial one. We believe this makes it easier to formalize and hope that insights may be borrowed to inspire new relaying strategies for a vanishing probability of error.

Related work. Zero-error communication over a primitive relay channel at first glance seems to be a combination of two notoriously difficult and open communication problems in information theory: computing the zero-error capacity over a point-to-point channel[2] , and the small-error capacity of a relay channel, whose capacity is unknown in general.

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. The small-error capacity and the zero-error capacity of a point-to-point discrete memoryless channel were both initially studied by Claude E. Shannon, in [4] in 1948 and in [5] in 1956. 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 $n \to \infty$ of the normalized independence number $\alpha(G^n_{X|Y})$ of the n-fold AND product of the confusability graph $G_{X|Y}$ associated with $p(y|x)$. This generally uncomputable limiting expression is rather 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 [6] 23 years later by proposing a computable-in-polynomial-time upper bound for the independence number of a graph. 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* graphs[3][3].

The *primitive relay channel* (PRC) proposed in [2] is a three-node relay channel introduced to decouple the multiple access and broadcast components of the standard relay channel by having the link from the relay to the destination be out of band and of fixed capacity C_0, as shown in Figure 1. In paper [2], an intensive case study of the small-error communication over a PRC is provided and it is shown that the classical relaying strategies Amplify-and-Forward, Decode-and-Forward and Compress-and-Forward are optimal in some classes of channels, but sub-optimal in general.

We first define the new problem of zero-error communication over a primitive relay channel in Section II and state our two main questions in Section III. In Section IV, we present our main results: the construction of a new Colour-and-Forward relaying scheme based on a novel compression graph G_R. This scheme is “information-lossless”: together with the observation at the destination, the colour sent by the relay yields the same confusability graph as the original

[1]The optimality of this upper bound is proved in paper [1].

[2]Multi-letter capacity expressions are available, but these are not generally computable except for a small class of channels with perfect graphs [3].

[3]A *perfect graph* is a graph where the chromatic number of every induced subgraph is that subgraph’s largest clique size.

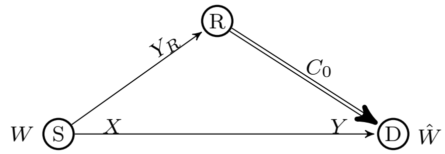

Fig. 1. A primitive relay channel $((\mathcal{X}, p(y, y_R|x), \mathcal{Y} \times \mathcal{Y}_R),\, C_0)$, where the broadcasting links $(\mathcal{X}, p(y, y_R|x), \mathcal{Y} \times \mathcal{Y}_R)$ from the source to the relay and destination terminals are orthogonal to the conferencing link with maximum rate C_0 bits/channel use from the relay to the destination terminal.

received signal at the relay. In Section V, we present a class of primitive relay channels for which this scheme achieves capacity. In Section VI, we provide interesting case studies.

II. Problem statement

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. Finally, the problem of zero-error communication over a primitive relay channel is formally defined. 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 or lines, 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 required in any colouring. A *minimum colouring* of graph G uses $\chi(G)$ colours.

The *strong product* or *AND* product $G \cdot H$ of two graphs G and H is defined as the graph with vertex set $V(G \cdot H) = V(G) \times V(H)$, 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. We denote G^n 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

Consider the 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 $\alpha(G^n_{X|Y})$ is the number of distinguishable channel inputs X^n, where $G^n_{X|Y}$ is the strong product or AND product of n copies of graph $G_{X|Y}$.[4] The zero-capacity is then characterized as [6]

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

which may be upper and lower bounded as [5], [6]:

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

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 due to the super-multiplicativity of the independence number [7].

C. Zero-error communication over a primitive relay channel

As shown in Fig. 2, a *primitive relay channel (PRC)* $((\mathcal{X}, p(y, y_R|x), \mathcal{Y} \times \mathcal{Y}_R),\, C_0)$ consists of: a source terminal S that wants to communicate a message W to a destination terminal D aided by a relay terminal R. This network is defined by a discrete memoryless broadcast channel $(\mathcal{X}, p(y, y_R|x), \mathcal{Y} \times \mathcal{Y}_R)$ from terminal S to terminals (R, D) and an out-of-band conference link with finite capacity C_0 from terminal R to terminal D allowing the relay to communicate at most C_0 error-free bits to D per channel use.

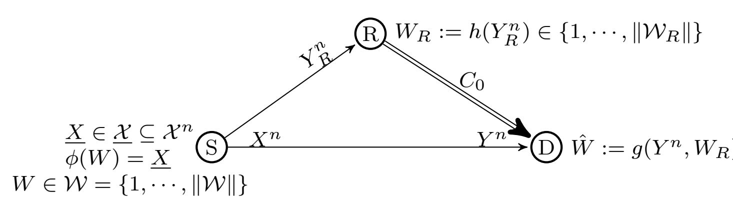

Fig. 2. A primitive relay channel $((\mathcal{X}, p(y, y_R|x), \mathcal{Y} \times \mathcal{Y}_R),\, C_0)$, with an encoder ϕ, a codebook $\underline{\mathcal{X}}$, a relaying function h and a decoding function g.

Definition 1: An encoder $\phi(W)$ at the source terminal S consists of a message set $\mathcal{W} = \{1, \cdots, \|\mathcal{W}\|\}$, a codebook $\underline{\mathcal{X}} \subseteq \mathcal{X}^n$ and the mapping:

$$\phi : \mathcal{W} \to \underline{\mathcal{X}}\,.$$

Remark 2: Codebook $\underline{\mathcal{X}}$ is a subset of the extended channel input alphabet $\mathcal{X}^n$ and consists of distinct sequences of length n. The mapping ϕ is bijective and decoding message

[4]Note that the n-fold strong product graph $G^n_{X|Y}$ 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 $p(y^n|x^n) = \prod_{i=1}^n p(y_i|x_i)$.

$W \in \mathcal{W}$ is equivalent to decoding codeword $\underline{X} \in \underline{\mathcal{X}}$. We will not distinguish these two concepts and abuse notation $\hat{w} \in \underline{\mathcal{X}}$ to denote the decoding result at the destination.

Definition 3: A *conference* $h(Y_R^n)$ *at rate* R_R from the relay to the destination is defined by a message set $\mathcal{W}_R$ and a mapping function h which processes the relay's observation Y_R^n into an index $w_R \in \mathcal{W}_R$:

$$h : \mathcal{Y}_R^n \to \mathcal{W}_R := \{1, 2, \cdots, \|\mathcal{W}_R\|\},$$

where $R_R := \frac{1}{n} \log \|\mathcal{W}_R\|$.

Definition 4: A *C_0-admissible conference* is a conference for which $R_R \leq C_0$.

Definition 5: A decoding function $g(Y^n, W_R) = \underline{\hat{X}}$ at the destination terminal takes in its own observation Y^n and the index W_R received from the conference link and produces an estimate for the transmitted codeword:

$$g : \mathcal{Y}^n \times \mathcal{W}_R \to \underline{\mathcal{X}}.$$

Note that the estimate of the transmitted message is $\hat{W} = \phi^{-1}(\underline{\hat{X}}) = \phi^{-1}(g(Y^n, W_R))$.

Protocols for 0-error communication over PRCs, are the counterparts of channel codes in ϵ-error communications.

Definition 6: An *n-shot protocol* (or scheme) $(n, \underline{\mathcal{X}}, h, g)$ for zero-error communication over a PRC $((\mathcal{X}, p(y, y_R|x), \mathcal{Y} \times \mathcal{Y}_R), C_0)$ is composed of a codebook $\underline{\mathcal{X}} \subseteq \mathcal{X}^n$, a C_0-admissible conference $h : \mathcal{Y}_R^n \to \mathcal{W}_R$ and a decoding function $g : \mathcal{Y}^n \times \mathcal{W}_R \to \underline{\mathcal{X}}$.

Definition 7: 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 $((\mathcal{X}, p(y, y_R|x), \mathcal{Y} \times \mathcal{Y}_R), C_0)$ achieving zero error, i.e. $\Pr[g(y, w_R) \neq w] = 0$ for all values $w \in \underline{\mathcal{X}}$.

Definition 8: The capacity C_z of zero-error communication over a PRC $((\mathcal{X}, p(y, y_R|x), \mathcal{Y} \times \mathcal{Y}_R), C_0)$ is the supremum of all possible achievable rates R_z for any n. Clearly, $C_z \leq \|\mathcal{X}\|$.

III. Two main questions

We are interested in and make contributions towards:

(1) *The zero-error capacity.* What is the zero-error capacity C_z for a given PRC $((\mathcal{X}, p(y, y_R|x), \mathcal{Y} \times \mathcal{Y}_R), C_0)$? We first show the cut-set bounds for C_z in Subsection III-A and then show the zero-error capacity for a special class of primitive relay channels in Section V.

(2) *Minimum required conference rate.* How small may the conference link rate C_0 be to achieve a given zero-error rate $R_z = R_z^*$, when $(\mathcal{X}, p(y, y_R|x), \mathcal{Y} \times \mathcal{Y}_R)$ is fixed? Slightly abusing notation, we use $C_z(C_0)$ to denote the zero error capacity of a PRC channel $((\mathcal{X}, p(y, y_R|x), \mathcal{Y} \times \mathcal{Y}_R), C_0)$ for a given C_0. We are particularly interested in the smallest C_0, denoted as $C_{0,z}^*$, that can achieve the zero-error capacity $C_z(C_0 = \infty)$, as discussed in detail in Subsection III-B. A novel upper bound on $C_{0,z}^*$ is derived in Section IV.

A. "Cut-set" bound for C_z

By allowing full cooperation between the relay and destination terminals (e.g. a genie argument), we have the following "cut-set"-like bound for the PRC zero-error capacity:

Proposition 1 (Zero-error capacity cut-set bound): The capacity $C_z(C_0)$ of the 0-error PRC is upper bounded by

$$\begin{aligned} C_z(C_0) \leq \min\{ &\log \lim_{n\to\infty} \sqrt[n]{\alpha(G^n_{X|Y})} + C_0, \\ &\log \lim_{n\to\infty} \sqrt[n]{\alpha(G^n_{X|Y,Y_R})}\}. \end{aligned} \tag{1}$$

Proof: Note that $\log \lim_{n\to\infty} \sqrt[n]{\alpha(G^n_{X|Y})}$ is the zero-error capacity of the direct link from the source to the destination terminal; if this is orthogonal to what is received from the relay, we obtain the first bound. The second bound is obtained by recognizing $\log \lim_{n\to\infty} \sqrt[n]{\alpha(G^n_{X|Y,Y_R})}$ as the zero-error capacity of a point to point channel $p(\tilde{y}|x)$ with $\tilde{y} = (y, y_R)$, obtained by giving (genie) y_R to the destination. ∎

B. An effectively fully-cooperative scenario

Let $C_z(\infty)$ be the zero-error capacity of the PRC $((\mathcal{X}, p(y, y_R|x), \mathcal{Y} \times \mathcal{Y}_R), C_0)$ when $C_0 = \infty$, i.e., the destination terminal knows exactly what the relay terminal observes (and hence the system is said to be "fully cooperative"). We have $C_z(\infty) = \log \lim_{n\to\infty} \sqrt[n]{\alpha(G^n_{X|Y,Y_R})}$. However, to achieve capacity $C_z(\infty)$, it is not necessary for C_0 to be infinity (a straightforward upper bound is $C_0 \leq \log \|\mathcal{Y}_R\|$), nor does the destination need to know *exactly* what the relay observed. To formally present this idea, we propose the quantity $C_{0,z}^*$ and the concept of an *effectively fully-cooperative* PRC.

Definition 9 ($C_{0,z}^*$ in 0-error setting): Define

$$C_{0,z}^* := \inf\{C_0 \geq 0 : C_z(C_0) = C_z(\infty) = \log \lim_{n\to\infty} \sqrt[n]{\alpha(G^n_{X|Y,Y_R})}\} \tag{2}$$

We call a PRC $((\mathcal{X}, p(y, y_R|x), \mathcal{Y} \times \mathcal{Y}_R), C_0)$ *effectively fully-cooperative* when its zero-error capacity is $C_z(\infty)$, i.e.

$$C_0 \geq C_{0,z}^*.$$

Note that $C_{0,z}^*$ is the smallest conferencing rate needed so that the upper bound $C_z(\infty)$ may be achieved. The ϵ-error communication analogy, $C_{0,\epsilon}^*$ is defined in [2] as

$$C_{0,\epsilon}^* := \inf\{C_0 : C_\epsilon(C_0) = C_\epsilon(\infty) = \max_{p(x)} I(X; Y, Y_R)\}.$$

Remark 10: Note that a straight forward upper bound for $C_{0,z}^*$ is $\log \|\mathcal{Y}_R\|$. Together with the the cut-set bound in Proposition 1, we have the following bounds on $C_{0,z}^*$:

$$\log \frac{\lim_{n\to\infty} \sqrt[n]{\alpha(G^n_{X|Y,Y_R})}}{\lim_{n\to\infty} \sqrt[n]{\alpha(G^n_{X|Y})}} \leq C_{0,z}^* \leq \log \|\mathcal{Y}_R\|.$$

Next we try to solve $C_{0,z}^*$ by exploring upper bounds on $C_{0,z}^{*(n)}$ [5], for some fixed number of channel uses, which is defined as:

$$C_{0,z}^{*(n)} = \inf\{C_0 \geq 0 : C_z^{(n)}(C_0) = C_z^{(n)}(\infty)\},$$

[5] We use (n) in the superscript to indicate n-shot channel usage.

where $C_z^{(n)}(\infty) := \log \sqrt[n]{\alpha(G^n_{X|Y,Y_R})}$.

Our plan is to derive an upper bound on $C_{0,z}^{*(n)}$ for any given number of channel uses n and then to derive an upper bound on $C_{0,z}^*$ based on these bounds.

In order to establish an effectively fully-cooperative scenario, it is sufficient to require the relaying function to be "information lossless" (and let C_0 be $\log \|\hat{\mathcal{Y}}_R^{(n)}\|$) in the sense that:

Definition 11 (Information lossless relaying): A relay mapping $\hat{Y}_R^{(n)} = h(Y_R^n) \in \{1, 2, \cdots, \|\hat{\mathcal{Y}}_R^{(n)}\|\}$ is called information lossless if the confusability graph on $\mathcal{X}^n$ from $p(y^n, y_R^n|x^n)$ is the same as the one from $p(y^n, \hat{y}_R^{(n)}|x^n)$, i.e.

$$G_{X^n|Y^n,Y_R^n} = G_{X^n|Y^n,\hat{Y}_R^{(n)}}.$$

Note that any valid information lossless relaying function $\hat{Y}_R^{(n)}$ provides an upper bound ($\log \|\hat{\mathcal{Y}}_R^{(n)}\|$) to $C_{0,z}^{*(n)}$.

IV. A new upper bound on $C_{0,z}^*$ via Colour-and-Forward

We now propose a general upper bound on $C_{0,z}^*$ (no looser than $\log \|\mathcal{Y}_R\|$) based on a novel compression graph G_R which depends on the channel structure. A colouring of G_R is transmitted by the relay in our new relaying scheme which we term the "Colour-and-Forward" scheme. This scheme captures the intuition of providing the terminal D with "what it needs" to resolve (with zero error) the transmitted symbol. We first focus on the one-shot Colour-and-Forward scheme for simplicity and to emphasize the intuition behind our strategy. We then state the n-shot version briefly in Subsection IV-C and state the novel Colour-and-Forward upper bound in Theorem 7 in Subsection IV-D.

A. One-shot Colour-and-Forward relaying:

Our relaying strategy is based on the intuition of providing "what the destination needs", i.e. remaining information lossless, while trying to minimize the number of bits needed to do so. We construct a new upper bound on $C_{0,z}^{*(1)}$ in Theorem 4, which would generalize to upper bounds on $C_{0,z}^{*(n)}$ for any n, and finally to provide an upper bound on $C_{0,z}^*$.

Sitting at the destination terminal, for a given a conditional joint pmf $p(y, y_R|x)$ with support $\mathcal{X}$ and output $\mathcal{Y} \times \mathcal{Y}_R$, we consider an arbitrary observation $Y = y$. Given this observation $Y = y$, the destination knows that the channel input symbol lies in the corresponding conditional support $S_{X|Y}(y)$. What the destination needs is to resolve the ambiguity among which x out of $S_{X|Y}(y)$ was sent. Furthermore, according to the joint pmf $p(y, y_R|x)$, the destination knows what the relay could have observed when the channel input symbol is $X = x$ given observation $Y = y$, i.e.

$$B_{Y_R}(x, y) := \{y_R : p(y, y_R|x) > 0 \text{ for given } x \text{ and } y\}.$$

In order to help D distinguish which channel input symbol x was actually transmitted, the relay terminal needs to differentiate different collections of y_R, i.e., $B_{Y_R}(x, y)$ in terms of the first index x for a given second index y. We propose to do so through the **Construction of the graph** $G_R(V, E)$:

1)	Vertices: $V = \mathcal{Y}_R := \{y_{R1}, y_{R2}, \cdots y_{R\|\mathcal{Y}_R\|}\}$;
2)	Edges: for every $y \in \mathcal{Y}$, construct a sequence of subsets of $\mathcal{Y}_R$, $B_{Y_R}(x, y)$, indexed by x, where $x \in S_{X|Y}(y)$. Edges are placed by fully connecting any two subsets $B_{Y_R}(x, y)$ and $B_{Y_R}(x', y)$, where $x \neq x'$ (i.e. put an edge between every pair (y_R, y_R') where $y_R \in B_{Y_R}(x, y)$ and $y_R' \in B_{Y_R}(x', y)$.) Note that for a given $Y = y$, the y_R vertices that are inside one $B_{Y_R}(x, y)$ need not be connected.

TABLE I
CONSTRUCTION OF THE GRAPH $G_R(V, E)$

This graph G_R can also be formally defined as:

Definition 12 (Colour-and-Forward graph G_R): Given a conditional joint pmf $p(y, y_R|x)$ with support $\mathcal{X}$ and output $\mathcal{Y} \times \mathcal{Y}_R$, graph G_R is an undirected graph with vertex set $\mathcal{Y}_R$ and an edge $y_{R1} - y_{R2}$ is imposed when for some y, $x_1 \neq x_2$, $\Pr(Y = y, Y_R = y_{R1}|X = x_1) \cdot \Pr(Y = y, Y_R = y_{R2}|X = x_2) > 0$.

Note that these two ways of specifying compression graph G_R are equivalent: defining graph G_R directly by imposing constraints on the joint pmf $p(y, y_R|x)$ is beneficial for the proof of Theorem 2 in Subsection IV-B, while constructing graph G_R via differentiating $B_{Y_R}(x, y)$ in terms of different x for a given y emphasizes the intuition of providing "what definition needs".

One example is provided in Subsection VI-A to illustrate the construction procedure in detail. We now propose a novel relaying index W_R^* based on Y_R, obtained from our *Colour-and-Forward* strategy derived from G_R.

Definition 13 (Colour-and-Forward relaying W_R^*): Given a conditional joint pmf $p(y, y_R|x)$ with support $\mathcal{X}$ and output $\mathcal{Y} \times \mathcal{Y}_R$, we define the Colour-and-Forward relaying W_R^* as a function of Y_R by a minimum colouring c with $\chi(G_R)$ colours on graph G_R:

$$W_R^* := c(Y_R)$$

where graph G_R is defined in Definition 12 and can be equivalently constructed by the iterative algorithm in Table I. (Note that c is not unique.)

We now propose Theorem 2 which states that the Colour-and-Forward relaying W_R^* is information lossless in terms of discriminating, together with Y, channel input X and establishes an effectively fully cooperative scenario for the PRC.

Theorem 2: $G_{X|Y,Y_R} = G_{X|Y,W_R^*}$, i.e. the confusability graph on $\mathcal{X}$ from $p(y, y_R|x)$ is the same as the one from $p(y, w_R^*|x)$, where $W_R^* = h(Y_R) = c(Y_R)$ is defined in Definition 13.

That is, provided C_0 is large enough to carry W_R^*, we may achieve $\log \alpha(G_{X|Y,Y_R})$.

Corollary 3: For a PRC $((\mathcal{X}, p(y, y_R|x), \mathcal{Y} \times \mathcal{Y}_R), C_0)$, when $C_0 \geq \log \chi(G_R)$, $C_z^{(1)}(C_0) = \log \alpha(G_{X|Y,Y_R})$.

Before proceeding to the proof for Theorem 2, we note that a direct application of Theorem 2 leads to a new upper bound on $C_{0,z}^{*(1)}$. 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.

Theorem 4: A new upper bound for $C_{0,z}^{*(1)}$ is:

$$C_{0,z}^{*(1)} \leq T_u^{(1)} ,$$

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

where $\chi(G_R|_{\mathcal{K}})$ is the chromatic number of graph $G_R|_{\mathcal{K}}$, constructed via the algorithm described in Table I from the induced conditional joint pmf $p_{\mathcal{K}}(y, y_R|x)$.

We provide some intuition behind the above theorem. Let the codebook $\underline{\mathcal{X}}$ be some maximum independent set $\mathcal{K}$ of graph $G_{X|Y,Y_R}$. For each given valid codebook $\underline{\mathcal{X}} = \mathcal{K}$, by Theorem 2, compressing Y_R into W_R^* according to the minimum colouring function c on graph $G_R|_{\mathcal{K}}$, is information lossless in terms of discriminating, together with Y, channel input $X \in \mathcal{K}$. To forward the W_R^* to the destination, a conference link with rate $\log \chi(G_R|_{\mathcal{K}})$ suffices. An example of how to compute $T_u^{(1)}$ is provided in Subsection VI-B.

Note that the vertex set of graph $G_R|_{\mathcal{K}}$ is $\mathcal{Y}_R|_{\mathcal{K}}$, which is a subset of $\mathcal{Y}_R$, i.e. $\mathcal{Y}_R|_{\mathcal{K}} \subseteq \mathcal{Y}_R$. Thus,

$$\chi(G_R|_{\mathcal{K}}) \overset{(a)}{\leq} \|\mathcal{Y}_R|_{\mathcal{K}}\| \overset{(b)}{\leq} \|\mathcal{Y}_R\| .$$

By Brooks' Theorem [8], the chromatic number of a graph is at most the maximum degree Δ (the largest vertex degree), unless the graph is complete or an odd cycle. So inequality (a) can be strict and as low as 1. The equality in (b) is obtained only when the restriction of support from $\mathcal{X}$ to $\mathcal{K}$ does not prohibit any $Y_R = y_R$ from showing up. One extreme case is when graph $G_{X|Y,Y_R}$ is edge free, then $\mathcal{K}$ equals to the whole vertex set and $\mathcal{Y}_R|_{\mathcal{K}} = \mathcal{Y}_R$. Please refer to the examples in the case study in section VI.

B. Proof of Theorem 2

Proof: Note $W_R^* = c(Y_R)$ is a deterministic function of Y_R by Definition 13, thus given the conditional pmf $p(y, y_R|x)$, $p(y, w_R^*|x)$ is computable. Since a confusability graph by definition is characterized by the collection of conditional joint supports $T := \{S_{X|Y,Y_R}(y, y_R), (y, y_R) \in \mathcal{Y} \times \mathcal{Y}_R\}$ and does not depend on the actual probability values, it suffices to show that the conditional supports $S_{X|Y,W_R^*}(y, w_R)$[6] (to be formally defined later) form the same collection T. We show this condition is true by pointing out

$$S_{X|Y,W_R^*}(y, w_R) = \bigcup_{y_R \in c^{-1}(w_R)} S_{X|Y,Y_R}(y, y_R)$$

and show that $\{S_{X|Y,W_R^*}(y, w_R), (y, w_R) \in \mathcal{Y} \times \mathcal{W}_R^*\} = \{S_{X|Y,Y_R}(y, y_R), (y, y_R) \in \mathcal{Y} \times \mathcal{Y}_R\} = T$.

It suffices to show every non-empty $S_{X|Y,Y_R}(y_0, y_{R0})$ is equal to $S_{X|Y,W_R^*}(y_0, w_{R0})$, where $w_{R0} = c(y_{R0})$. For every (y_0, y_{R0}) such that $S_{X|Y,Y_R}(y_0, y_{R0}) \neq \emptyset$, we denote $c(y_{R0}) = w_{R0}$ and let $c^{-1}(w_{R0}) = \{y_{R0}, y_{R1}, \cdots, y_{R(K-1)}\}$, where $K \geq 1$ is the number of y_R's that are mapped to the same colour index w_{R0}. When $K = 1$, $S_{X|Y,W_R^*}(y_0, w_{R0}) = S_{X|Y,Y_R}(y_0, y_{R0})$. When $K \geq 2$, $S_{X|Y,W_R^*}(y_0, w_{R0}) = S_{X|Y,Y_R}(y_0, y_{R0}) \cup S_{X|Y,Y_R}(y_0, y_{R1}) \cup \cdots \cup S_{X|Y,Y_R}(y_0, y_{R(K-1)})$. Note that $S_{X|Y,Y_R}(y_0, y_{R0})$ is non-empty:

- when $S_{X|Y,Y_R}(y_0, y_{R0})$ has only one element, say x_0, we know $\Pr(Y = y_0, Y_R = y_{R0}|X = x_0) > 0$. By the construction of W_R^* in Definition 13, we know that $\Pr(Y = y_0, Y_R = y_{Rt}|X = x_q) = 0$ for all $t = 1, \cdots, K-1$ and $x_q \neq x_0$. Otherwise, the presumption that y_{R0} and y_{Rt} share the same colour index w_{R0} leads to a contradiction. As a result, for all $t = 1, \cdots, K-1$, $S_{X|Y,Y_R}(y_0, y_{Rt}) = \{x_0\}$ when $\Pr(Y = y_0, Y_R = y_{Rt}|X = x_0) > 0$ and $S_{X|Y,Y_R}(y_0, y_{Rt}) = \emptyset$ otherwise. Thus, we have $S_{X|Y,Y_R}(y_0, y_{R0}) \cup S_{X|Y,Y_R}(y_0, y_{R1}) \cup \cdots \cup S_{X|Y,Y_R}(y_0, y_{R(K-1)}) = S_{X|Y,Y_R}(y_0, y_{R0})$, and $S_{X|Y,W_R^*}(y_0, w_{R0}) = S_{X|Y,Y_R}(y_0, y_{R0})$.
- when $S_{X|Y,Y_R}(y_0, y_{R0})$ has more than one element, i.e., $x_0, x_0' \in S_{X|Y,Y_R}(y_0, y_{R0})$ and $x_0 \neq x_0'$. Applying the argument above twice, we have $\Pr(Y = y_0, Y_R = y_{Rt}|X = x_q) = 0$ for all $t = 1, \cdots, K-1$ when $x_q \neq x_0$ and $x_q \neq x_0'$. Thus, $\Pr(Y = y_0, Y_R = y_{Rt}|X = x) = 0$ for all $t = 1, \cdots, K-1$ and all $x \in \mathcal{X}$, i.e., $S_{X|Y,Y_R}(y_0, y_{Rt}) = \emptyset$ for all $t = 1, \cdots, K-1$. So, $S_{X|Y,W_R^*}(y_0, w_{R0}) = S_{X|Y,Y_R}(y_0, y_{R0})$.

Thus, every non-empty $S_{X|Y,Y_R}(y_0, y_{R0})$ is equal to $S_{X|Y,W_R^*}(y_0, w_{R0})$, where $w_{R0} = c(y_{R0})$ and hence $\{S_{X|Y,Y_R}(y, y_R), (y, y_R) \in \mathcal{Y} \times \mathcal{Y}_R\}$ and $\{S_{X|Y,W_R^*}(y, w_R), (y, w_R) \in \mathcal{Y} \times \mathcal{W}_R^*\}$ are equal. ∎

C. n-shot Colour-and-Forward relaying:

When block coding or multiple uses of the channel is allowed, all results from the previous section may be extended in a straightforward manner. That is, we focus on the joint conditional pmf $p(y^n, y_R^n|x^n)$ with support $\mathcal{X}^n$ and output $\mathcal{Y}^n \times \mathcal{Y}_R^n$. The compression graph $G_R^{(n)}$ now has y_R^n or its subset as vertices, and is analogously defined. For simplicity, we state the key theorems without proofs.

Theorem 5: $G_{X^n|Y^n,Y_R^n} = G_{X^n|Y^n,W_R^*}$, i.e. the confusability graph on $\mathcal{X}^n$ from $p(y^n, y_R^n|x^n)$ equals that from $p(y^n, w_R^*|x^n)$. W_R^* is generated by Definition 13 from $p(y^n, y_R^n|x^n)$ with support $\mathcal{X}^n$ and output $\mathcal{Y}^n \times \mathcal{Y}_R^n$.

Theorem 6: A new upper bound for $C_{0,z}^{*(n)}$ is:

$$C_{0,z}^{*(n)} \leq T_u^{(n)} ,$$

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

where $\chi(G_R^{(n)}|_{\mathcal{K}})$ is the chromatic number of graph $G_R^{(n)}|_{\mathcal{K}}$, constructed via the algorithm described in Table I from the joint pmf $p(y^n, y_R^n|x^n)$ with restricted input $\mathcal{K}$.

[6]Throughout this proof, we drop the superscript of w_R^* for simplicity and we mean $w_R \in \mathcal{W}_R^* = \{1, \cdots, \chi(G_R)\}$.

D. A new upper bound on $C_{0,z}^*$ via Colour-and-Forward

Theorem 7 (A Colour-and-Forward upper bound on $C_{0,z}^*$): $C_{0,z}^* \le T_u$, where

$$T_u := \min\left\{T_u^{(n^*)} : n^* = \arg\max_n \{C_z^{(n)}(\infty), n = 1, 2, \cdots\}\right\},$$

when $\max_n C_z^{(n)}(\infty)$ exists.

This follows directly from Theorem 6, the definition $C_z(\infty) = \sup_n \log \sqrt[n]{\alpha(G^n_{X|Y,Y_R})}$ and the assumption that the supremum and maximum of $\log \sqrt[n]{\alpha(G^n_{X|Y,Y_R})}$ is equal.

Note that to infer the behavior of $C_{0,z}^*$ from the upper bounds on $C_{0,z}^{*(n)}$ requires knowing how $C_z(\infty)$ depends on $C_z^{(n)}(\infty)$. By the super-multiplicity of the independence number sequence of the strong product graphs and Fekete's lemma, we know that sequence $C_z^{(n)}(\infty)$ converges to its supremum. In general, the maximum need not exist.

V. Zero-error capacity of a special class of primitive relay channels

Applying Theorem 2 using our Colour-and-Forward relay strategy W_R^* defined in Definition 13, we may obtain the zero-error capacity of a special class of primitive relay channels.

Recall that in Definition 9, we call a PRC channel $((\mathcal{X}, p(y, y_R|x), \mathcal{Y} \times \mathcal{Y}_R), C_0)$ *effectively fully-cooperative*, when its zero-error capacity equals $C_z(\infty)$, i.e., when $C_0 \ge C_{0,z}^*$, which is guaranteed when $C_0 \ge T_u$, where T_u is the upper bound in Theorem 7.

We now look at a particular class of primitive relay channels for which we will be able to show capacity. We term these *perfect primitive relay channels* as (1) like in the point-to-point channel, we can characterize the zero-error capacity exactly, and not because any of the associated graphs are *perfect graphs* necessarily; (2) the zero-error capacity of such PRCs is the maximal possible rate – the logarithm of the channel input alphabet size $\|\mathcal{X}\|$.

Definition 14: A PRC channel $((\mathcal{X}, p(y, y_R|x), \mathcal{Y} \times \mathcal{Y}_R), C_0)$ is perfect if 1) $\max_n C_z^{(n)}(\infty)$ exists; 2) $C_0 \ge T_u$; and 3) $G_{X|Y,Y_R}$ is edge free.

Theorem 8: The zero-error capacity of the perfect primitive relay channel satisfying conditions in Definition 14, is

$$C_{\text{z,perfect}} = \log ||\mathcal{X}||.$$

Proof: The converse is trivial: the zero-error capacity is always upper bounded by $\log \|\mathcal{X}\|$. The achievability follows by default:

(1) graph $G_{X|Y,Y_R}$ being edge-free implies that block coding brings no gain than a single-shot coding scheme. Thus, $C_z(\infty) = \log \alpha(G_{X|Y,Y_R}) = \log \|\mathcal{X}\|$.

(2) $C_z(\infty) = \log \|\mathcal{X}\|$ is achievable because $C_0 \ge C_{0,z}^*$ is guaranteed by $C_0 \ge T_u$.

To be explicit, the 1-shot protocol $(n = 1, \underline{\mathcal{X}} = \mathcal{X}, h, g)$ achieves zero error when $C_0 \ge T_u$, with the codebook $\underline{\mathcal{X}}$ being the whole channel input alphabet as desired, the relaying $h(y_R) := c(y_R)$ as in Definition 13 and the decoding function $g(y, w_R)$ as in the proof of Theorem 2:

$$g(y, w_R) := S_{X|Y,W_R^*}(y, w_R) = \bigcup_{y_R \in c^{-1}(w_R)} S_{X|Y,Y_R}(y, y_R).$$

■

Note that we do not claim the $C_0 \ge T_u$ is necessary, but merely sufficient to achieve $\log \|\mathcal{X}\|$.

	$S_{X\|Y}(y)$	$B_{Y_R}(x, y)$	edges
$Y = 1$	$X = 1$ $X = 5$	$\{3, 4\}$ $\{1\}$	$1-3, 1-4$
$Y = 2$	$X = 1$ $X = 2$ $X = 4$	$\{1, 2\}$ $\{2\}$ $\{4\}$	$1-2, 1-4, 2-4$
$Y = 3$	$X = 2$ $X = 3$	$\{3\}$ $\{1\}$	$1-3$
$Y = 4$	$X = 3$ $X = 4$	$\{3\}$ $\{4\}$	$3-4$
$Y = 5$	$X = 4$ $X = 5$	$\{3\}$ $\{5\}$	$3-5$

TABLE III

CONSTRUCTING COMPRESSION GRAPH G_R FROM $p(y, y_R|x)$ IN TABLE II.

VI. Case studies

We first give an example to show how to construct the compression graph G_R, illustrated in Table I and how to use this to find the Colour-and-Forward relay mapping $W_R^* = c(Y_R)$, defined in Definition 13. Then, in Subsection VI-B, we show how to obtain the upper bound $T_u^{(1)}$ on $C_{0,z}^{*(1)}$ in Theorem 4 and how to construct protocols for the whole network. Last, we provide three examples to further illustrate the intuition and potential benefit of relaying to provide "what the destination needs". For the sake of simplifying the description for a conditional joint pmf, we let $p(y, y_R|x) = p(y_R|x)p(y|x)$ in the these three examples. An edge in a bipartite graph between X and Y (or X and Y_R) indicates $p(y|x) > 0$ (or $p(y_R|x) > 0$).

A. Construction of the compression graph G_R

Table II 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 $p(y, y_R|x) = 0$.

Table III illustrates the iterative algorithm: for each $Y \in [1:5]$, construct a sequence of $B_{Y_R}(x, y) \subseteq \mathcal{Y}_R = [1:5]$, where $x \in S_{X|Y}(y) = \{x : p(y|x) > 0\}$ and put an edge between every pair (y_R, y'_R) where $y_R \in B_{Y_R}(x, y)$ and $y'_R \in B_{Y_R}(x', y)$. Superimposing these edges, we have the compression graph G_R as shown in Figure 3. In graph G_R, different colours are used to denote one choice of minimum colouring function c. These colours specify the relay's mapping $W_R^* = c(Y_R)$. As shown in Figure 4, $G_{X|Y,Y_R} = G_{X|Y,W_R^*}$, i.e. compressing Y_R into W_R^* is information lossless in the sense that together with Y, W_R^* provides the same ability to distinguish different $X = x$'s as Y_R.

$s(p(y,y_R\|x))$		Y_R					Y_R					Y_R					Y_R					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
	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 II

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.

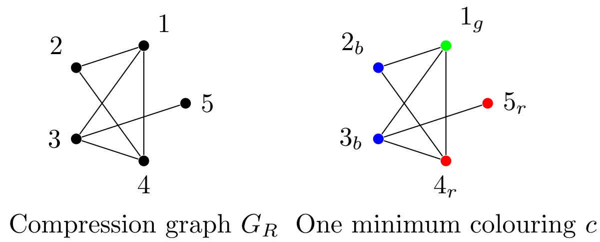

Fig. 3. The compression graph G_R and one choice of minimum colouring function c, for the joint conditional pmf $p(y,y_R|x)$ in Table II. Note that the least number of colours required is: $\chi(G_R)=3$.

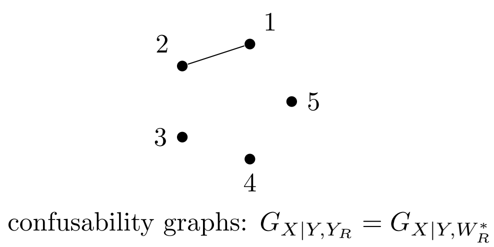

Fig. 4. Compressing Y_R into W_R^* according to the minimum colouring function c on graph G_R in Figure 3, is information lossless in the sense that together with Y, W_R^* provides as much information of about X as Y_R. Note that the independence number is: $\alpha(G_{X|Y,Y_R})=4$.

	$S_{X\|Y}(y)$	$B_{Y_R}(x,y)$	edges
$Y=1$	$X=1$ $X=5$	$\{3,4\}$ $\{1\}$	$1-3, 1-4$
$Y=2$	$X=1$ $\cancel{X=2}$ $X=4$	$\{1,2\}$ $\cancel{\{2\}}$ $\{4\}$	$\cancel{1-2, 1-4, 2-4}$ $1-4, 2-4$
$Y=3$	$\cancel{X=2}$ $X=3$	$\cancel{\{3\}}$ $\{1\}$	$\cancel{1-3}$ $\emptyset$
$Y=4$	$X=3$ $X=4$	$\{3\}$ $\{4\}$	$3-4$
$Y=5$	$X=4$ $X=5$	$\{3\}$ $\{5\}$	$3-5$

TABLE IV

CONSTRUCTING COMPRESSION GRAPH $G_R|_{\mathcal{K}_1}$ FOR INDUCED CONDITIONAL JOINT PMF $(\mathcal{K}_1, p_{\mathcal{K}_1}(y,y_R|x), \mathcal{Y}|_{\mathcal{K}_1}\times\mathcal{Y}_R|_{\mathcal{K}_1})$.

B. Construction of protocols for the PRC

Let the broadcasting component of this primitive relay channel, i.e. $(\mathcal{X}, p(y,y_R|x), \mathcal{Y}\times\mathcal{Y}_R)$, be the conditional joint pmf specified in Table II. We now illustrate how to compute our upper bound $T_u^{(1)}$ on $C_{0,z}^*$ in Theorem 4 and when $C_0 \geq T_u^{(1)}$, how to construct protocols to achieve the message rate of Corollary 3, $R_z = \log\alpha(G_{X|Y,Y_R}) = \log 4$ (Figure 4).

As shown in Figure 4, graph $G_{X|Y,Y_R}$ has two possible maximal independent sets: $\mathcal{K}_1 = \{1,3,4,5\}$ and $\mathcal{K}_2 = \{2,3,4,5\}$.

1) When codebook $\mathcal{K}_1 = \{1,3,4,5\}$ is chosen, the *induced* broadcasting component is $(\mathcal{K}_1, p_{\mathcal{K}_1}(y,y_R|x), \mathcal{Y}|_{\mathcal{K}_1}\times\mathcal{Y}_R|_{\mathcal{K}_1})$.
   - The iterative algorithm: in Table IV.[7]
   - The compression graph $G_R|_{\mathcal{K}_1}$: in Figure 5.[8]
   - One choice of minimum colouring on compression graph $G_R|_{\mathcal{K}_1}$: in Figure 5 with chromatic number $\chi(G_R|_{\mathcal{K}_1})=3$.
2) When codebook $\mathcal{K}_2 = \{2,3,4,5\}$ is chosen and the *induced* broadcasting component is $(\mathcal{K}_2, p_{\mathcal{K}_2}(y,y_R|x), \mathcal{Y}|_{\mathcal{K}_2}\times\mathcal{Y}_R|_{\mathcal{K}_2})$.
   - The iterative algorithm: in Table V.[9]
   - The compression graph $G_R|_{\mathcal{K}_2}$: in Figure 6.[10]
   - One choice of minimum colouring on compression graph $G_R|_{\mathcal{K}_2}$: in Figure 6 with chromatic number $\chi(G_R|_{\mathcal{K}_2})=2$.

So $T_u^{(1)} = \log\min\{\chi(G_R|_{\mathcal{K}_1}), \chi(G_R|_{\mathcal{K}_2})\} = \log\min\{3,2\} = 1$. Thus, by Theorem 4, we know $C_{0,z}^*$ is upper bounded by 1. Also, when $C_0 \geq T_u^{(1)}$, by choosing codebook $\mathcal{K}_2 = \{2,3,4,5\}$, relaying function h to be the minimum colouring for graph $G_R|_{\mathcal{K}_2}$ in Figure 6 and the decoding function $g(y,w_R)$ to be the same one used in the proof of Theorem 8 or Theorem 2, we can achieve the message rate $R_z = \log\alpha(G_{X|Y,Y_R}) = \log 4$.

C. The pentagon problem

We now consider a channel where the direct link between the source and destination consists of Shannon's "pentagon problem", which was notoriously difficult to solve. If the relay link is such that the corresponding channel forms a perfect PRC (this relay link described by $p(y_R|x)$ is not unique) an example of which is shown in Figure 7, we have $\alpha(G_{X|Y,Y_R})=5$ and rate $\log 5$ can be achieved, when $C_0 \geq \log 3$ (in a 1-shot scheme). Note that smaller values

[7]We retain the cross-out items in Tables IV, V to serve a comparison with the construction algorithm in Table III.

[8]We retain the dotted edges in Figure 5, 6 to serve as a comparison with the compression graph in Figure 3.

[9]See footnote 7.

[10]See footnote 8.

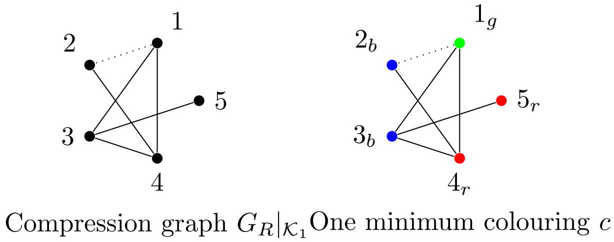

Fig. 5. The compression graph $G_R|_{\mathcal{K}_1}$ and one choice of minimum colouring function c, for induced conditional joint pmf $(\mathcal{K}_1, p_{\mathcal{K}_1}(y, y_R|x), \mathcal{Y}|_{\mathcal{K}_1} \times \mathcal{Y}_R|_{\mathcal{K}_1})$. Note that the least number of colours required is: $\chi(G_R|_{\mathcal{K}_1}) = 3$.

	$S_{X\|Y}(y)$	$B_{Y_R}(x, y)$	edges
$Y = 1$	$\cancel{X = 1}$ $X = 5$	$\cancel{\{3, 4\}}$ $\{1\}$	$\cancel{1 - 3, 1 - 4}$ $\emptyset$
$Y = 2$	$\cancel{X = 1}$ $X = 2$ $X = 4$	$\cancel{\{1, 2\}}$ $\{2\}$ $\{4\}$	$\cancel{1 - 2, 1 - 4, 2 - 4}$ $2 - 4$
$Y = 3$	$X = 2$ $X = 3$	$\{3\}$ $\{1\}$	$1 - 3$
$Y = 4$	$X = 3$ $X = 4$	$\{3\}$ $\{4\}$	$3 - 4$
$Y = 5$	$X = 4$ $X = 5$	$\{3\}$ $\{5\}$	$3 - 5$

TABLE V

CONSTRUCTING COMPRESSION GRAPH $G_R|_{\mathcal{K}_2}$ FOR INDUCED CONDITIONAL JOINT PMF $(\mathcal{K}_2, p_{\mathcal{K}_2}(y, y_R|x), \mathcal{Y}|_{\mathcal{K}_2} \times \mathcal{Y}_R|_{\mathcal{K}_2})$.

of C_0 might still be able to guarantee the maximal rate $\log \|\mathcal{X}\| = \log 5$ when multiple channel uses are allowed, but this is left open.

We compare the rate achieved by our strategy with that achieved by a "Decode-and-Forward" (DF) relaying strategy. In a DF strategy, the relay would like to decode every codeword $w \in \underline{\mathcal{X}}$, in which case the message rate is constrained by $R_z \le \log \alpha(G_{X|Y_R})$. In this example, $\alpha(G_{X|Y_R}) = 3$. Thus, $R_z \le \log 3$ is a hard constraint on the message rates that can be achieved by Decode-and-Forward, which is clearly inferior to that achieved by our scheme. Our scheme might be seen as a "channel-aware" (depends on the conditional $p(y, y_R|x)$) compression of Y_R, and thus might be seen as a smart way of implementing Compress-and-Forward.

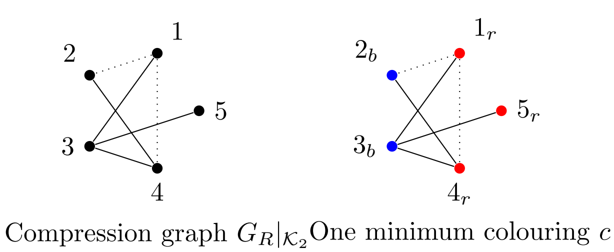

Fig. 6. The compression graph $G_R|_{\mathcal{K}_2}$ and one choice of minimum colouring function c, for induced conditional joint pmf $(\mathcal{K}_2, p_{\mathcal{K}_2}(y, y_R|x), \mathcal{Y}|_{\mathcal{K}_2} \times \mathcal{Y}_R|_{\mathcal{K}_2})$. Note that the least number of colours required is: $\chi(G_R|_{\mathcal{K}_2}) = 2$.

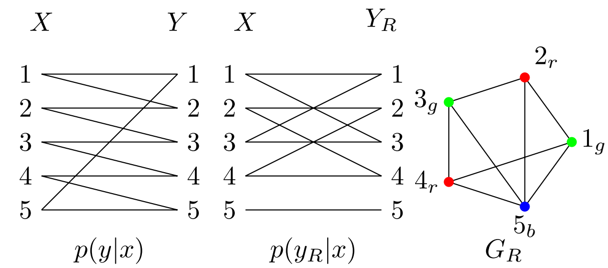

Fig. 7. Pentagon problem: marginals and G_R graph used for relaying. The capacity is $\log 5$ and may be achieved if $C_0 \ge \log 3$, in 1-shot.

D. An example where no compression is possible

We now provide an example in Figure 8 to show that there exist channels for which no information lossless compression is possible at the relay and the relay has to forward everything that it has observed, i.e, $C^*_{0,z} = \log \|\mathcal{Y}_R\|$. Our relaying scheme W^*_R captures this phenomenon by requiring 8 different colours for 8 y_R's, as shown in Figure 8.

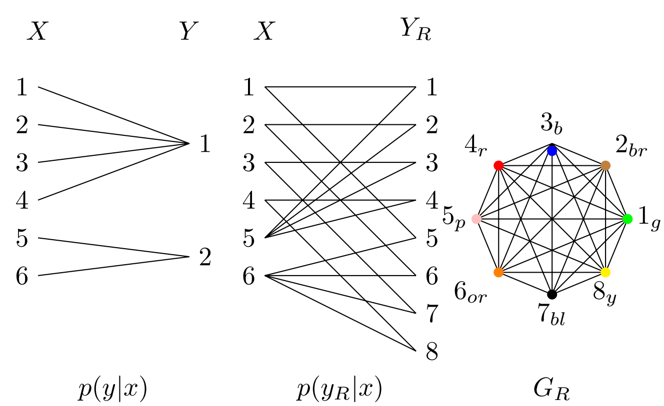

Fig. 8. An example where information lossless compression at the relay is impossible.

E. An example where much compression is possible

Finally, in Figure 9 we show an example of a channel where Y_R may be highly compressed without losing the *needed* information about X – i.e. we do *not* need to reconstruct Y_R at the destination, but only need to use the conferencing link to resolve any remaining ambiguity from the direct link. Here, one may verify that by sending only one of the two colours over the conferencing link, that a capacity of $\log 8$ may be achieved when $C_0 \ge \log 2$.

VII. CONCLUSION AND FUTURE WORK

This paper introduces and formally defines the problem of zero-error communication over a primitive relay channel, which serves as an example of the essential role of a relay in a relay channel: providing "what the destination terminal needs" to disambiguate the transmitted symbols. We develop a compression graph to capture this notion of "what the destination needs" and propose a novel information lossless

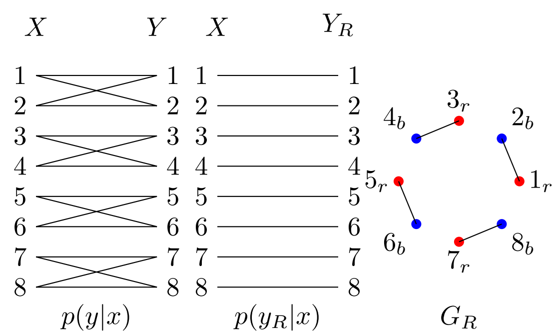

Fig. 9. An example where much compression is possible.

relaying scheme based on some minimum colouring on this compression graph, termed "Colour-and-Forward" relaying. We hope that insights from this zero-error communication relaying strategy may be borrowed to better understand how to exploit the channel structure to design new relaying schemes in the small-error setting.

Due to the space limitations, we leave the following discussions to future work: (1) Whether the presented Colour-and-Forward relaying mapping is optimal, i.e. yields the minimum needed conference capacity to support an overall rate of $\lim_{n\to\infty} \log \sqrt[n]{\alpha(G_{X^n|Y^n,Y_R^n})}$; (2) The connection with Witsenhausen's source coding graph [9]. We note that, for the special channel in which $Y_R = X$ with probability 1, our problem is different from the source coding problem studied by Witsenhausen in [9].

References

[1] Y. Chen and N. Devroye, "The optimality of colour-and-forward relaying for a class of zero-error primitive relay channels," in submitted to ISIT, 2015.

[2] Y. Kim, "Coding techniques for primitive relay channels," in Proc. Forty-Fifth Annual Allerton Conf. Commun., Contr. Comput, 2007.

[3] J. Korner and A. Orlitsky, "Zero-error information theory," IEEE Trans. Inf. Theory, vol. 44, no. 6, pp. 2207–2229, 1998.

[4] C. Shannon, "A mathematical theory of communication," Bell Syst. Tech. J., vol. 27, no. 379-423, 623-656, Jul., Oct. 1948.

[5] ——, "The zero error capacity of a noisy channel," Information Theory, IRE Transactions on, vol. 2, no. 3, pp. 8–19, September 1956.

[6] L. Lovasz, "On the shannon capacity of a graph," IEEE Trans. Inf. Theory, vol. 25, no. 1, pp. 1–7, 1979.

[7] J. Nayak, E. Tuncel, and K. Rose, "Zero-error source-channel coding with side information," IEEE Trans. Inf. Theory, vol. 52, no. 10, pp. 4626–4629, Oct 2006.

[8] R. L. Brooks, "On colouring the nodes of a network," Mathematical Proceedings of the Cambridge Philosophical Society, vol. 37, pp. 194–197, 4 1941. [Online]. Available: http://journals.cambridge.org/article_S0305004100002168X

[9] H. Witsenhausen, "The zero-error side information problem and chromatic numbers (corresp.)," IEEE Trans. Inf. Theory, vol. 22, no. 5, pp. 592–593, 1976.