Interference Forwarding in Multiuser Networks

Ron Dabora, Ivana Marić and Andrea Goldsmith
Dept. of Electrical Engineering
Stanford University

***Abstract*—We study communication in networks with multiple source-destination pairs and relays. In such networks, the channel output at any destination receiver consists of both the desired signal and interference. In this setting the relay can help forward the desired message of a user to the destination receiver, or help forward interference to a receiver to improve its ability to cancel the interference. Focusing on the impact of interference forwarding, we define a new relay-interferer channel (RIC) model, which serves as the basic building block for the study of interference in multiuser networks. Using the RIC we show that correlation between the codebooks of the relay and the interferer (e.g. superposition codebooks) is essential for obtaining performance benefits from interference forwarding. We conclude that in order to achieve rate gains from relaying interference using the decode-and-forward strategy, a superposition codebook is required. Otherwise, this relay strategy has the same rate as interference cancellation at the receiver. We also conclude that compress-and-forward is not useful for forwarding interference and has no better performance than just treating interference as noise at the decoder.**

I. Introduction

Networks with multiple source-destination pairs have multiple possible relaying strategies, since the relay can facilitate communication between any or all of these pairs. However, relaying to assist a given source-destination pair may cause interference to other pairs. In this work we study two aspects of communication in networks with multiple source-destination pairs (MSDP) and a relay: relaying interference and interference cancellation at the receivers. Relaying in MSDP networks is fundamentally different from the classic relay scenario: while in the classic relay scenario the signal received at the destination consists of the information signal and noise, in MSDP networks, the signal received at any destination also contains interference. Here, we define interference as a signal that is selected from a codebook known to all receivers, but carries information intended only for one receiver. Interference differs from noise in that it has a structure that can be utilized by the decoder to help mitigate it. In addition, interference differs from channel states in that the size of the interferer codebook is exponentially smaller than the size of the typical set of the states. We shall expand on this in Section I-B.

In general, a relay in a network is not obliged to help all source-destination pairs – he is more likely to focus on helping only a subset of the communicating pairs due to considerations such as transmit power and processing power. However, when evaluating the rate region for an MSDP network with a relay, it should be noted that the relay transmission affects *all* source-destination pairs, and not only those that he helps. Hence, optimizing the relay strategy for general networks is a challenging problem. In this work we propose a simplified channel model for the study of relaying in the presence of interference, which is described next.

The authors are with the Wireless Systems Lab, Department of Electrical Engineering, Stanford University, Stanford, CA 94305. Email: {andrea,ivanam,ron}@wsl.stanford.edu. This work was supported in part by the DARPA ITMANET program under grant 1105741-1-TFIND and the ARO under MURI award W911NF-05-1-0246.

A. The Relay-Interferer Channel

The analysis of relaying in MSDP networks can be simplified once we identify the basic "building block" for relaying in such networks. Clearly, the classic relay channel [1] is not satisfactory for this purpose as it lacks an interfering signal. We therefore analyze the classic relay channel *with an additional source of interference*. This channel is defined by $\{\mathcal{X} \times \mathcal{X}_1 \times \mathcal{X}_2, p(y, y_1|x, x_1, x_2), \mathcal{Y} \times \mathcal{Y}_1; p(x_2), R_2\}$, see Figure 1. Here, X, X_1 and X_2 are the channel inputs from the source, the relay and the interferer, respectively, Y is the channel output at the destination and Y_1 is the channel output at the relay, R_2 is the transmission rate of the interferer whose codebook is generated in an i.i.d. manner according to $p(x_2)$. Note that this channel does not include a receiver for X_2 which allows us to focus our attention on relaying techniques in the presence of interference. We call this channel the *relay-interferer channel* (RIC). In the RIC X_2 can be a vector of random variables, representing any finite number of interferers.

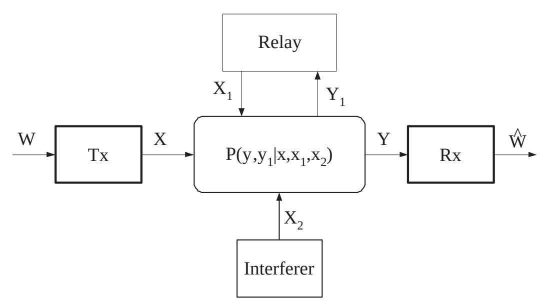

Fig. 1. The relay-interferer channel.

The RIC is the simplest network model that captures the impact of interference. Thus, it provides insight into different aspects of relaying in the presence of interference, which may generalize to more complex network models. A second motivation for studying the RIC is that it represents a practical network relaying scenario. Specifically, the RIC can represent a scenario in which there is a relay station that processes all transmissions and helps several receivers while only the helped receivers adapt their decoding procedure to utilize the information from the relay. The other receivers simply treat the relay transmission as interference. This is indeed more likely to be implemented in practice than the full scenario in which every receiver knows the codebooks of all communicating pairs and processes the transmissions from all sources in order to extract its own information.

Finally, it is important to note that while R_2 is the rate of the interferer in the RIC, the way R_2 and $p(x_2)$ are selected is not considered in the model, and they are treated as fixed scenario parameters. Thus, the RIC is used to facilitate the analysis of relay-assisted communications in the presence of interference with given parameters.

B. Joint-Typicality Compression of a Codebook

In this work we also study interference forwarding and partial interference cancellation using a compressed interference codebook based on joint-typicality compression. We note, however, that there is an inherent loss in this approach. Using joint-typicality based compression implies that (almost) every typical sequence $\mathbf{x}_2$ can be compressed into an estimate sequence $\hat{\mathbf{x}}_2$. This assumption is inherent to the derivation of the minimal rate of the codebook of estimates. To demonstrate this loss assume the interferer has only two messages. Then, it is enough to send only one bit to tell the receiver *the exact interfering sequence*. However, using compression, we need $nI(\hat{X}_2; X_2)$ bits to deliver *an estimate* of the interference, and noiseless compression ($\hat{X}_2 = X_2$) requires $nH(X_2)$ bits.

A related model to the problem considered in this paper is the cognitive interference channel [10]. In this model one of the transmitters (called the cognitive transmitter) has a-priori knowledge of the other transmitter's message. In [10] the Gelfand-Pinsker scheme [2] was applied to encode the message of the cognitive transmitter. In Section IV we show that contrary to information-forwarding using compress-and-forward (CF) and to interference embedding via the Gelfand-Pinsker scheme used in [10], a compressed interference codebook does not exceed the performance obtained by treating interference as noise at the decoder.

C. Related Work and Main Contributions

Capacity of MSDP networks has been investigated in several recent works. Specifically, capacity of the interference channel with a relay has been considered by several authors: in [3] and [4] interference forwarding strategies were first proposed. In [5] and [6] the case where the relay is cognitive was analyzed and in [7], cooperation via DF in Gaussian channels was considered. Joint encoding of the source message and state information was considered in [8] for the arbitrarily varying multiple-access channel with a common message and one informed encoder. Another related scenario, the relay channel with states known only at the relay, was studied in [9]. Encoding interference as channel state information for the cognitive interference channel was considered in [10].

In this work we present new insights on relaying in networks. We first show that in contrast to the classic information relaying of Cover and El-Gamal [1], where correlated codebooks are not essential to achieve rate gains from relaying (and in fact for CF [1, Theorem 6] no such correlation exists), when the relay is forwarding interference, this correlation is essential for achieving a rate increase. In fact, under the assumption that the interferer and the relay do not transmit correlated codewords, it turns out that it is not necessary for the relay to process the interference at all, as the same rates that are obtained by forwarding interference from the relay can be achieved using interference cancellation at the decoder. This holds for both DF and CF.

II. Notations and Definitions

In the following we denote random variables with upper case letters, e.g. X, Y, and their realizations with lower case letters, x, y. A random variable (RV) X takes values in a set $\mathcal{X}$. We use $p_X(x)$ to denote the probability mass function (p.m.f.) of a discrete RV X on $\mathcal{X}$. For brevity we may omit the subscript X when it is obvious from the context. We use $p_{X|Y}(x|y)$ to denote the conditional p.m.f. of X given Y. We denote vectors with boldface letters, e.g. $\mathbf{x}$, $\mathbf{y}$; the i'th element of a vector $\mathbf{x}$ is denoted with x_i and we use x_i^j where $i \leq j$ to denote the vector $(x_i, x_{i+1}, ..., x_{j-1}, x_j)$; x^j is a short form notation for x_1^j, and $\mathbf{x} \equiv x^n$. We use $H(\cdot)$ to denote the entropy and $I(\cdot;\cdot)$ to denote the mutual information as defined in [11, Chapter 2]. $A_\epsilon^{*(n)}(X)$ is the set of strongly-typical i.i.d. sequences of length n w.r.t. $p_X(x)$, see [11, Chapter 13]. The notation $\mathcal{C}(n, p(x), R)$ denotes a codebook that consists of 2^{nR} i.i.d codewords of length n generated according to $p(x)$.

We now introduce the following definitions:

Definition 1: The *discrete, memoryless relay-interferer channel* is defined by $\big\{\mathcal{X} \times \mathcal{X}_1 \times \mathcal{X}_2, p(y, y_1|x, x_1, x_2), \mathcal{Y} \times \mathcal{Y}_1; p(x_2), R_2\big\}$, where $p(y, y_1|x, x_1, x_2)$ is the channel transition function and the sets $\mathcal{X}$, $\mathcal{X}_1$, $\mathcal{X}_2$, $\mathcal{Y}$ and $\mathcal{Y}_1$ are discrete and finite. The interferer transmits at rate R_2 by uniformly selecting codewords from a codebook whose size is 2^{nR_2}. The codewords for the interferer are generated in an i.i.d. manner according to $p(x_2)$. Denoting by m the transmitted message, the distribution of a block of n symbols is given by

$$\begin{aligned}
&p(m, \mathbf{x}, \mathbf{x}_1, \mathbf{x}_2, \mathbf{y}, \mathbf{y}_1)\\
&\overset{(a)}{=} p(m) \prod_{i=1}^{n} p\big(x_i, x_{2,i}|x^{i-1}, x_{1,1}^{i-1}, x_{2,1}^{i-1}, m\big) \times\\
&\qquad p\big(x_{1,i}|x^i, x_{1,1}^{i-1}, x_{2,1}^{i}, y_{1,1}^{i-1}, m\big) p\big(y_i, y_{1,i}|x_i, x_{1,i}, x_{2,i}\big)\\
&= p(m)p(\mathbf{x}|m)p(\mathbf{x}_2)\prod_{i=1}^{n} p\big(x_{1,i}|y_{1,1}^{i-1}\big) p\big(y_i, y_{1,i}|x_i, x_{1,i}, x_{2,i}\big),
\end{aligned}$$

where (a) is because there is no feedback to the transmitter and interferer.

Definition 2: An (R, n) *code for the relay-interferer channel* consists of a message set $\mathcal{M} = \big\{1, 2, ..., 2^{nR}\big\}$ and a collection of mapping functions $(f, \{t_i\}_{i=2}^n, g)$ where $f : \mathcal{M} \mapsto \mathcal{X}^n$, is the mapping at the encoder, $x_{1,i} = t_i\left(y_{1,1}, y_{1,2}, ..., y_{1,i-1}\right)$, $i = 2, 3, ..., n$, are the mapping functions at the relay, and $x_{1,1} = c \in \mathcal{X}_1$ is an arbitrary constant. Finally, $g : \mathcal{Y}^n \mapsto \mathcal{M}$, is the mapping at the decoder.

Definition 3: The *average probability of error* of a code for the RIC is defined as $P_e^{(n)} = \Pr\big(g(Y^n) \neq M\big)$, where M is selected uniformly from $\mathcal{M}$, and the interferer selects its codewords uniformly from the codebook $\mathcal{C}_2(n, p(x_2), R_2)$.

Definition 4: A rate R is *achievable* for the RIC if for every $\epsilon > 0$ and $\delta > 0$ there exists $n(\epsilon, \delta)$ such that for all $n > n(\epsilon, \delta)$ a code $(R - \delta, n)$ with an average probability of error $P_e^{(n)} \leq \epsilon$ can be constructed.

We evaluate the average probability of error $P_e^{(n)}$ averaged over all interferer codebooks $\mathcal{C}_2$. This is because the interferer does not try to deliberately disrupt the communication between the source and the destination. In fact, as long as the interferer can deliver information at rate R_2, its codebook can be selected such that the rate to the destination is maximized (the interferer operates independently of the other nodes). Therefore, we average over all codebooks rather than evaluate $P_e^{(n)}$ for a fixed interferer codebook.

III. Analysis of Decode-and-Forward Applied to Both the Source and the Interference Messages

Consider the case in which the relay decodes both the desired message and the interferer's message. This offers a possibility for the relay to forward both (or either) of the two messages in the next block. As in the classic DF [1, Theorem 1], the relay can use a block transmission strategy in which it decodes messages in one block and forwards information about them in the next block. Therefore, in block b, the encoding scheme at the relay can be described as

$$X_1^n(b) = f_b^n\big(\{W(j), W_2(j)\}_{j=1}^{b-1}\big) \tag{1}$$

where $W(j)$ and $W_2(j)$ denote the respective desired message and the interferer's message sent in block j.

We first considered the application of DF for forwarding both information on the desired message and on the interference. This facilitates both information enhancement and interference cancellation at the destination. The rate achieved with this scheme is given in the following lemma:

Lemma 1: For a given interference rate R_2 and i.i.d distribution $p(x_2)$, any rate R satisfying

$$R \leq I(X; Y_1|X_1, X_2) \tag{2a}$$

$$R \leq I(X, X_2; Y_1|X_1) - R_2 \tag{2b}$$

$$R \leq I(X; Y|X_1, X_2) + I(X_1; Y) \tag{2c}$$

$$R \leq I(X, X_1, X_2; Y) - R_2 \tag{2d}$$

for some input distribution $p(x_2)p(x, x_1)$, is achievable.
We denote the rate given by Lemma 1 with $R_{DF}^{(1)}$. Note that constraints (2a) and (2b) are due to decoding at the relay and (2c) and (2d) are due to decoding at the destination.

We now observe that the constraints (2c) and (2d) can be achieved using interference cancellation at the decoder as in the standard interference channel [10], without forwarding any interference information from the relay. We conclude that for DF, processing interference at the relay does not provide any rate increase over interference cancellation at the destination. This implies that the relay does not need to decode the interferer's message and therefore it saves a rate constraint on R_2 that follows from decoding the interference at the relay.

An alternative to interference cancellation at the destination is to treat interference as noise. If both the relay and the destination treat the interference signal X_2^n as noise and the source message is relayed using DF, the following rate is achievable directly from [1, Theorem 1]:

Lemma 2: For a given interference rate R_2 and i.i.d distribution $p(x_2)$, any rate R satisfying

$$R \leq \sup_{p(x,x_1)} \min\{I(X; Y_1|X_1), I(X, X_1; Y)\} \tag{3}$$

for some input distribution $p(x_2)p(x, x_1)$, is achievable.
We denote the rate given by (8) with $R_{DF}^{(2)}$.

Therefore, depending on whether the relay and destination treat interference as noise or not, the maximum of rates (8) and (2) can be achieved:

$$R_{DF} = \max\{R_{DF}^{(1)}, R_{DF}^{(2)}\}. \tag{4}$$

Observe that the rate (2d) decreases as R_2 increases implying that, as expected, decoding interference is favorable for lower rates of the interferer's codebook.

IV. Analysis of Compress-and-Forward Applied to the Interference Codebook

In Lemma 1 interference cancellation is implemented with the exact interferer codebook $\mathcal{C}_2$. This strategy has a drawback when the rate of the interference is high as the rate is upper bounded by $R \leq I(X, X_1, X_2; Y) - R_2$. We now present an alternative approach to interference cancellation, using a compressed codebook, following the rate-distortion scheme of [14, Chapter 9.5]. The rate of the compressed codebook is adjusted to balance between the gain from interference cancellation and the rate loss due to processing interference at the receiver. We begin with a statement of the lemma:

Lemma 3: For a given interference rate R_2 and i.i.d distribution $p(x_2)$, any rate R satisfying

$$\begin{aligned} R &\leq I(X; Y_1|X_1, X_2) \\ R &\leq I(X, X_2; Y_1|X_1) - R_2 \\ R &\leq I(X; Y|\hat{X}_2, X_1) + I(X_1; Y) - I(\hat{X}_2; X_2|X_1, Y), \end{aligned} \tag{5}$$

for some distribution $p(x, x_1)p(\hat{x}_2, x_2)$, is achievable.
The proof of Lemma 3, outlined in the appendix, shows that processing interference at the relay using joint-typicality compression does not result in a rate increase.

Discussion

Recall that the codebook size is 2^{nR_2} which is exponentially smaller than $2^{nH(X_2)}$. Thus, when performing joint-typicality compression there is an unavoidable loss since such compression inherently considers the entire typical set. The loss is evident from the fact that R_2 is absent from the rate constraint (5). Thus the size of the interference codebook does not matter, only the size of $A_\epsilon^{*(n)}(X_2)$.

We next determine the conditions under which Lemma 3 allows higher rates than Lemma 1, assuming that decoding at the relay does not limit the rate. Comparing the rates (2) and (5) for the same $p(x, x_1)p(x_2)$ we look for:

$$\begin{aligned} &I(X; Y|\hat{X}_2, X_1) + I(X_1; Y) - I(\hat{X}_2; X_2|X_1, Y) \\ &\quad > I(X; Y|X_2, X_1) + I(X_1; Y) + \min\big(0, I(X_2; Y|X_1) - R_2\big) \\ &\Rightarrow \max\big(0, R_2 - I(X_2; Y|X_1)\big) - I(\hat{X}_2; X_2|X_1, Y) \\ &\qquad\qquad > H(X|Y, \hat{X}_2, X_1) - H(X|Y, X_2, X_1). \end{aligned}$$

Note that the right-hand side is non-negative since $H(X|Y, \hat{X}_2, X_1) - H(X|Y, X_2, X_1) = I(X; X_2|Y, \hat{X}_2, X_1)$. If $R_2 \leq I(X_2; Y|X_1)$ then the left-hand side is, in general, negative. *Hence, as long as the receiver can decode* $\mathbf{x}_2$*, the rate of Lemma 1 is higher than the rate of Lemma 3.* Now assume that $R_2 > I(X_2; Y|X_1)$ and proceed with the comparison:

$$\begin{aligned} R_2 - I(X_2; Y|X_1) - I(\hat{X}_2; X_2|X_1, Y) \\ > H(X|Y, \hat{X}_2, X_1) - H(X|Y, X_1, X_2) \\ \Rightarrow R_2 > I(\hat{X}_2; X_2) + I(X_2; Y|X_1, X, \hat{X}_2) \\ = I(X_2; Y, \hat{X}_2|X, X_1). \end{aligned}$$

Since this holds for every $p(x, x_1)p(x_2)$, then it holds also for the input distribution that maximizes $R_{DF}^{(1)}$.

When $\hat{X}_2 = \varnothing$, then for the rate of Lemma 3 to be higher than the rate of Lemma 1 we need

$$R_2 > I(X_2; Y|X_1, X). \tag{6}$$

This means that when (6) holds, it is better for the destination to treat the interference as noise instead of using the coding strategy associated with Lemma 1. Indeed in this case the resulting rate reduces to $R_{DF}^{(2)}$ given by (8). Finally, we check whether the rate defined in Lemma 3 can exceed the DF rate of (8). We look for

$$\begin{aligned} & I(X; Y|\hat{X}_2, X_1) - I(\hat{X}_2; X_2|X_1, Y) > I(X; Y|X_1) \\ \Rightarrow\ & H(X|\hat{X}_2, X_1) - H(X|Y, \hat{X}_2, X_1) \\ & \quad -H(X|X_1) + H(X|Y, X_1) > I(\hat{X}_2; X_2|X_1, Y) \\ \Rightarrow\ & H(\hat{X}_2|X, Y, X_1) < H(\hat{X}_2|X_2, X_1, Y) \\ \Rightarrow\ & I(X_2; \hat{X}_2|X, Y, X_1) < 0, \end{aligned}$$

which contradicts the fact that mutual information is non-negative. The last step is because when X_2 is given, $\hat{X}_2$ is independent of all other variables. This show that using CF on the interference cannot outperform treating the interference as noise at the decoder, which was used to obtain $R_{DF}^{(2)}$.

We now try to give an intuitive explanation for this conclusion: from (5) we observe that there are two opposite effects when using the compressed codebook at the decoder:

- An increase in the point-to-point rate $I(X; Y|\hat{X}_2, X_1) \geq I(X; Y|X_1)$.
- A rate decrease due to uncertainty of the actual $\hat{\mathbf{x}}_2$ to be used. The rate is decreased by $I(\hat{X}_2; X_2|X_1, Y)$.

We see that due to the Markov chain, the effect of the decrease is always greater than the increase.

V. The Importance of Correlated Codebooks in Interference Forwarding

From the previous sections it follows that neither DF nor CF, applied to the interference, provide a rate increase over interference cancellation at the destination receiver. We believe that this conclusion extends beyond the two schemes studied here. We summarize this in the following conjecture:

Conjecture 1: When the relay employs the block transmission strategy (1), there is no gain from relaying the interferer's message in the RIC.

We observe from (1) that information about messages $W(b-1)$ and $W_2(b-1)$ is forwarded at the relay with a block delay, i.e., in block b. However, in block b the interferer sends *no* information about $W_2(b-1)$. Rather, the interference the interferer introduces is *independent* from the one in the previous block and thus the codewords of the relay and the interferer are blockwise independent i.e. $p(x_1^n, x_2^n) = p(x_1^n)p(x_2^n)$. Therefore, the relay *cannot* enhance the interference and facilitate interference cancellation. In contrast, in [3] and [4], we showed that in networks with multiple sources and destinations, interference forwarding results in rate gains and leads to capacity in a certain regime. In these scenarios, the interfering source sends the message $W_2(b-1)$ over two blocks $b-1$ and b. Hence, the relay using a block transmission strategy can facilitate interference cancellation of $W_2(b-1)$ in the second block. In particular, the relay and the interferer use superposition coding and therefore their codewords are *correlated*, thus allowing for interference cancellation and the rate gains. A similar result holds for [10] where the interference is known non-causally at one transmitter.

We note that in the classic DF for the relay channel, superposition coding is not essential to achieve a rate increase from relaying. If the codebooks are independent, i.e.,

$$p(x, x_1) = p(x)p(x_1), \tag{7}$$

the DF rate of [1] is given by

$$R \leq \sup_{p(x)p(x_1)} \min\{I(X; Y_1|X_1), I(X, X_1; Y)\}, \tag{8}$$

which still brings gains over direct transmission. In this case however, the destination is interested in the message that the relay is forwarding. This last example illustrates how the considerations for relaying information to help the destination decode it differ from the considerations for relaying interference to help the destination subtract it.

VI. Conclusions

We have studied two aspects of interference in networks. We first studied the interference-forwarding relay strategy. We showed that correlation between the codewords of the relay and interferer is essential to achieve a rate increase from interference forwarding. This holds for both DF and CF. Without such correlation, the same rates that relaying provides can be achieved with interference cancellation at the destination. As a consequence of our analysis we conclude that CF at the relay. will not be useful for interference forwarding. This situation is completely different from the scenario in which the relay performs its classic role of forwarding desired information to the destination receiver.

Appendix
Proof Outline of Lemma 3

We start by compressing interference at the relay and then conclude that the same maximal rate can be obtained without processing the interference at the relay.

1) Codebook Generation:

- Fix n and the p.m.f.s $p(v)p(x_1|v)p(x|v)p(\hat{x}_2, x_2)$.
- For each $s \in \mathcal{S} \triangleq \{1, 2, ..., 2^{nR_s}\}$ generate a codeword $\mathbf{v}(s)$ according to $\Pr(\mathbf{v}(s)) = \prod_{i=1}^{n} p_V(v_i)$.
- For each $s \in \mathcal{S}$ generate a codebook with 2^{nR} codewords $\mathbf{x}(m,s)$, $m \in \mathcal{M} \triangleq \{1, 2, ..., 2^{nR}\}$, according to $\Pr(\mathbf{x}(m,s)) = \prod_{i=1}^{n} p_{X|V}(x_i|v_i(s))$.
- Uniformly and independently partition the set $\mathcal{M}$ into 2^{nR_s} subsets $\mathcal{B}(s)$, $s \in \mathcal{S}$.
- For each $w \in \mathcal{W} \triangleq \{1, 2, ..., 2^{nR_w}\}$ generate a codeword $\hat{\mathbf{x}}_2(w)$ according to $\Pr(\hat{\mathbf{x}}_2(w)) = \prod_{i=1}^{n} p_{\hat{X}_2}(\hat{x}_{2,i})$.
- Independently and uniformly partition the message set $\mathcal{W}$ into 2^{nR_q} bins denoted $\mathcal{Z}(q)$, $q \in \mathcal{Q} \triangleq \{1, 2, ..., 2^{nR_q}\}$.
- For each $s \in \mathcal{S}$ generate a codebook with 2^{nR_q} codewords $\mathbf{x}_1(q,s)$, $q \in \mathcal{Q}$ according to $\Pr(\mathbf{x}_1(q,s)) = \prod_{i=1}^{n} p_{X_1|V}(x_{1,i}|v_i(s))$.

2) Encoding at the Transmitter at Time i: Let s_i denote the index of the set $\mathcal{B}(s)$ into which m_{i-1} belongs: $m_{i-1} \in \mathcal{B}(s_i)$. For transmitting m_i the transmitter outputs $\mathbf{x}(m_i, s_i)$.

3) Encoding at the Relay at Time i: At time i the relay knows $\mathbf{x}_2(i-1)$. The relay looks for an index $w \in \mathcal{W}$ s.t. $\big(\mathbf{x}_2(i-1), \hat{\mathbf{x}}_2(w)\big) \in A_\epsilon^{*(n)}$. Following [14, Chapter 9.5] this can be done with an arbitrarily small probability of error by taking n large enough as long as

$$R_w \geq I(X_2; \hat{X}_2). \tag{9}$$

Denote the selected w with w_i. Denote the index of the partition $\mathcal{Z}(q)$ into which w_i belongs with q_i: $w_i \in \mathcal{Z}(q_i)$. At time i the relay outputs $\mathbf{x}_1(q_i, s_i)$.

4) Decoding the Source Message and the Interference at the Relay at Time i: At time i the relay decodes $(m_i, \mathbf{x}_2(i))$ by looking for a unique message $m \in \mathcal{M}$ and $\mathbf{x}_2 \in \mathcal{C}_2$ s.t.

$$\big(\mathbf{v}(s_i), \mathbf{x}(m, s_i), \mathbf{x}_1(q_i, s_i), \mathbf{x}_2, \mathbf{y}_1(i)\big) \in A_\epsilon^{*(n)}.$$

As in [11, Chapter 14.3] this can be done with an arbitrarily small probability of error by taking n large enough as long as

$$R < I(X; Y_1|V, X_1, X_2) \tag{10a}$$
$$R_2 \leq I(X_2; Y_1|V, X_1, X) \tag{10b}$$
$$R + R_2 \leq I(X, X_2; Y_1|V, X_1). \tag{10c}$$

5) Decoding at the Destination at Time i:

a) Decoding the Information from the Relay: The destination looks for a unique pair $(s, q) \in \mathcal{S} \times \mathcal{Q}$ such that $\big(\mathbf{v}(s), \mathbf{x}_1(q,s), \mathbf{y}(i)\big) \in A_\epsilon^{*(n)}$. This can be done with an arbitrarily small probability of error for n large enough if

$$R_q + R_s \leq I(V, X_1; Y) \tag{11a}$$
$$R_q \leq I(X_1; Y|V). \tag{11b}$$

b) Decoding the Source Message: At time i the destination decodes m_{i-1}. After decoding the relay information the destination knows s_i and q_i. Thus it knows the set $\mathcal{B}(s_i)$ into which m_{i-1} belongs. The destination generates the set

$$\mathcal{L}(i-1) = \Big\{m \in \mathcal{M} \big| \big(\mathbf{x}(m, s_{i-1}), \mathbf{x}_1(q_{i-1}, s_{i-1}), \mathbf{v}(s_{i-1}), \mathbf{y}(i-1)\big) \in A_\epsilon^{*(n)}\Big\}.$$

The destination now decodes m_{i-1} by looking for a unique $m \in \mathcal{B}(s_i) \bigcap \mathcal{L}(i-1)$ s.t.

$$\big(\mathbf{x}(m, s_{i-1}), \mathbf{v}(s_{i-1}), \hat{\mathbf{x}}_2(w), \mathbf{x}_1(q_{i-1}, s_{i-1}), \mathbf{y}(i-1)\big) \in A_\epsilon^{*(n)},$$

for at least one $w \in \mathcal{Z}(q_i)$. In [12] we show that this can be done with an arbitrarily small probability of error if

$$R \leq I(X; Y|V, X_1, \hat{X}_2) + R_s \tag{12a}$$
$$R \leq I(X; Y|V, X_1, \hat{X}_2) + R_s + R_q - R_w + I(\hat{X}_2; Y|X_1, V) \tag{12b}$$

6) Combining all Bounds: Combining (9) and (11a) with (12) we obtain

$$R \leq I(X; Y|V, X_1, \hat{X}_2) + I(V, X_1; Y) - R_q \tag{13a}$$
$$R \leq I(X; Y|V, X_1, \hat{X}_2) + I(V, X_1; Y) - I(\hat{X}_2; X_2|Y, X_1, V). \tag{13b}$$

Next ,we note that the expression in (13b) can be written as

$$R \leq I(X, X_1, \hat{X}_2; Y) - I(\hat{X}_2; X_2) \tag{14}$$

which is independent of V. Since the rate is the minimum of the two constraints in (13) and since setting $V = X_1$ we achieve (14), we conclude that the maximum rate of this scheme can be achieved with $R_q = 0$, namely not forwarding any interference information from the relay. Finally, as the relay does not send information about the interference, it does not need to select a compressed codeword and also decoding the interference at the relay is not necessary and we can omit constraint (10b). The remaining active constraints are therefore (10a), (10c) and (14).

References

[1] T. M. Cover and A. A. El Gamal. "Capacity Theorems for the Relay Channel". *IEEE Trans. Inform. Theory*, IT-25(5):572–584, 1979.

[2] S. I. Gel'fand and M. S. Pinsker. "Coding for Channel with Random Parameters". *Probl. Inform. & Control*, vol. 9(1):19–31, 1980.

[3] R. Dabora, I. Marić and A. Goldsmith. "Relay Strategies for Interference Forwarding". *Proc. of the Information Theory Workshop (ITW)*, Porto, Portugal, May 2008.

[4] I. Marić, R. Dabora and A. Goldmsmith "On the Capacity of the Interference Channel with a Relay". *Proc. of the IEEE International Symposium on Information Theory (ISIT)*, Toronto, Canada, July 2008.

[5] O. Sahin and E. Erkip. "On Achievable Rates for Interference Relay Channel with Interference Cancellation". *Proc. of 41th Annual Asilomar Conf. on Signals, Systems and Computers*, Pacific Grove, CA, 2007.

[6] S. Sridharan, S. Vishwanath, S. A. Jafar and S. Shamai "On the Capacity of Cognitive Relay Assisted Gaussian Interference Channel". *Proc. of the IEEE International Symposium on Information Theory (ISIT)*, Toronto, Canada, July 2008.

[7] O. Sahin and E. Erkip. "Achievable rates for the Gaussian interference relay channel". *Proceedings of GLOBECOM Communication Theory Symposium*, Washington D.C., November 2007.

[8] A. Somekh-Baruch, S. Shamai and S. Verdu. "Cooperative Encoding with Asymmetric State Information at the Transmitters". *Proc. 44th Allerton Conf. Commun., Contr. and Comp.*, Monticello, IL, Sep. 2006.

[9] A. Zaidi, S. Kotagiri, J. N. Laneman and L. Vandendorpe. "Cooperative Relaying with State Available at the Relay Only". *Proc. of the Information Theory workshop (ITW)*, Porto, Portugal, May 2008.

[10] I. Marić, A. Goldsmith, G. Kramer and S. Shamai. "On the Capacity of Interference Channels with One Cooperating Transmitter". *European Transactions on Telecommunications*, 2008.

[11] T. M. Cover and J. Thomas. *Elements of Information Theory*. John Wiley and Sons Inc., 1991.

[12] R. Dabora, I. Marić and A. Goldsmith. "Relaying in Neworks with Multiple Source-Destination Pairs". In preparation.

[13] I. Marić, N. Liu and A. Goldsmith, "Encoding against an Interferer's Codebook", *Allerton Conference on Communication, Control and Computing*, submitted, 2008.

[14] R. W. Yeung. *A First Course in Information Theory*. Springer, 2002.