Oblivious Relaying
for Primitive Interference Relay Channels

O. Simeone
CWCSPR, ECE Dept.,
NJIT, Newark, NJ, USA

E. Erkip
Dept. of ECE, Polytechnic Inst. of NYU
Brooklyn, NY, USA

S. Shamai (Shitz)
Dept. of EE, Technion
Haifa, Israel

Abstract—[1]Consider a relay node that needs to operate without knowledge of the codebooks (i.e., modulation, coding) employed by the assisted source-destination pairs. This paper studies the performance of relaying under this condition, termed oblivious relaying, for the primitive relay channel (PRC) and the primitive interference relay channel (PIRC). "Primitive" refers to the fact that the relay-to-destinations links use orthogonal resources with respect to the other links. Assuming discrete memoryless models, the capacity of a PRC with oblivious relaying is derived, along with the capacity region of the PIRC with oblivious relaying and interference-oblivious decoding (i.e., each decoder is unaware of the codebook used by the interfering transmitter). In all cases, capacity is achieved by Compress-and-Forward (CF) with time-sharing. Performance without time-sharing is discussed as well. Finally, it is shown that for the general (non-oblivious) Gaussian PRC, the achievable rate by CF (with Gaussian inputs and test channels and no time-sharing) is suboptimal by at most half bit with respect to the cut-set bound.

I. INTRODUCTION

A standard, and often implicit, assumption in network-information theoretic analyses is that design of encoding/ decoding functions at all nodes is performed jointly in order to optimize the system performance. This implies, in particular, that all nodes must be aware at all times of the operations carried out by any other node. Moreover, in general, addition of a new node, or even only change of operation at one node, calls for a re-design of the entire network. While this may be reasonable in centrally controlled networks such as conventional[2] cellular system, it becomes impractical in decentralized scenarios. In fact, in the latter cases, nodes operate without extensive signalling capabilities, so that full coordination in the choice of encoding/ decoding functions is typically a prohibitive task.

In this work, we investigate design of basic network building blocks, under the assumption that information about the operations carried out at the source encoders (i.e., of the sources' *codebooks*) is not available throughout the network. We emphasize that this may be due to practical constraints, as discussed above, or simply to the need for simple network protocols that do not require continuous reconfiguration (and
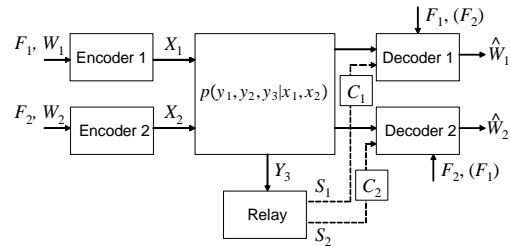
Fig. 1. A Primitive Inteference Relay Channel (PIRC) with oblivious relaying and interference-oblivious or interference-aware decoders.

thus extensive signaling). The analysis is based on the framework of *oblivious processing* first proposed in [1]. We focus on the "primitive" relay channel (PRC, see review in [3]) and on an extension of the PRC to a setting with two source-destination pairs, that we define primitive interference relay channel (PIRC), see Fig. 1. We establish a number of capacity results under the assumption of oblivious processing and the relay and, possibly, at the interfered destinations.

II. SYSTEM MODEL

We study the PRC and PIRC with oblivious processing as depicted in Fig. 2 and 1, respectively. We use the term "primitive" as in [3] to mean that the relay is connected to the destination(s) via finite-capacity orthogonal links. This corresponds to assuming that the relay transmissions occupy a different resource with respect to the other links in the system. As detailed below, oblivious processing, following [1], refers to coding/ decoding operations designed without the knowledge of some of the codebooks in the system.

A discrete memoryless PIRC consists of two source-destination pairs (indexed by subscripts 1 and 2) and is defined by a tuple $(\mathcal{X}_1, \mathcal{X}_2, p(y_1, y_2, y_3|x_1, x_2), \mathcal{Y}_1, \mathcal{Y}_2, \mathcal{Y}_3, C_1, C_2)$ where C_1, C_2 denote the capacities (bits/ channel use) of the links from relay to destination 1 and 2, respectively. Subscript 3 is used for the relay. A special case of the PIRC is the PRC [3], where there is only one source-destination pair, i.e., we set $\mathcal{X}_2 = \mathcal{Y}_2 = \emptyset$. In this case, we drop the subscript 1 for simplicity so that the PRC is defined as $(\mathcal{X}, p(y, y_3|x), \mathcal{Y}, \mathcal{Y}_3, C)$, see Fig. 2. We will also consider a Gaussian model with power constraints, to be introduced below.
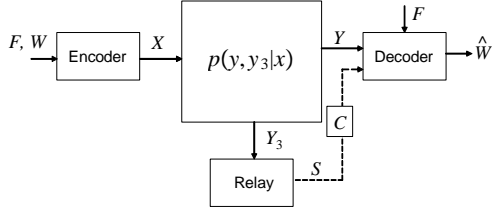
Fig. 2. A Primitive Relay Channel (PRC) with oblivious relaying.

A. Oblivious Processing

In the following, we detail on the considered model of PIRC with oblivious processing. The corresponding model for the PRC with oblivious processing is a special case that follows immediately and will not be detailed explicitly. In the considered model, each source-destination pair agrees on the codebook to be used for communications (i.e., the destination knows the codebook used by the corresponding transmitter), as in regular interference channels. However, we assume that the information about the codebooks may be lacking at the relay (*oblivious relaying*) and possibly at the interfered destination (*oblivious decoding*).

To account for oblivious processing, we follow the model of [1], which we first describe informally in the following. Fix rates R_j [bits/ channel use], $j = 1, 2$, used for transmission between the jth source-destination nodes. According to [1], we assume that the currently employed codebook (say by pair $j = 1, 2$) is identified by an index $F_j \in [1, |\mathcal{X}|^{n2^{nR_j}}]$, which ranges over the set $[1, |\mathcal{X}|^{n2^{nR_j}}]$ of all possible codebooks of rate R_j. Therefore, transmitter j sends a message $W_j \in [1, 2^{nR_j}]$ by transmitting a codeword $x_j^n(F_j, W_j)$ dependent on both message W_j and index F_j. Knowledge of F_j implies awareness of the codebook used by the jth source-destination pair[2]. Moreover, in the absence of knowledge of F_j, it is assumed that the codeword transmitted by the jth source completely lacks any structure, and thus its letters "look" independent identically distributed (i.i.d.) with respect to a given single-letter distribution $p_{X_j}(\cdot)$ over $\mathcal{X}_j$, $j = 1, 2$. Rigorous definitions are given below, highlighting also the role of time-sharing.

B. Formal Setting

Formal definitions are as follows.
Definition 1: A (n, R_1, R_2) code for the PIRC with *oblivious processing* is given by:

a. Message sets $[1, 2^{nR_j}]$ and codebook sets $[1, |\mathcal{X}_j|^{n2^{nR_j}}]$, $j = 1, 2$;

b. Encoding functions: For each user j, the *encoder* is defined by a pair (p_{X_j}, ϕ_j), where p_{X_j} is a single-letter pmf and $\phi^{(j)}$ is a mapping ϕ_j: $[1, |\mathcal{X}_j|^{n2^{nR_j}}] \times [1, 2^{nR_j}] \rightarrow \mathcal{X}_j^n$, that provides the transmitted codeword $x_j^n = \phi_j(F_j, W_j)$ given codebook index F_j and message W_j. The pmf p_{X_j}

[2]This can be seen as a form of randomized encoding.

defines the probability $p_F(f)$ of choosing a certain codebook $F \in [1, |\mathcal{X}_j|^{n2^{nR_j}}]$ as

$$p_{F_j}(f) = \prod_{w \in [1, 2^{nR_j}]} p_{X^n}(\phi_j(f, w)), \qquad (1)$$

where $p_{X^n}(x^n) = \prod_{i=1}^n p_X(x_i)$;

c. Relaying function: The *relay*, unaware of the codebooks F_j with $j = 1, 2$ maps the received sequence $y_3^n \in \mathcal{Y}^n$ into two indices $s_j \in [1, 2^{nC_j}]$ to be sent to destinations $j = 1, 2$ as $\phi_3 : \mathcal{Y}_3^n \rightarrow [1, 2^{nC_1}] \times [1, 2^{nC_2}]$, so that $[s_1, s_2] = \phi_3(y_3^n)$.

d. Decoding functions: For interference-aware decoding, decoding is described by a mapping

$$g_j : [1, |\mathcal{X}_1|^{n2^{nR_1}}] \times [1, |\mathcal{X}_2|^{n2^{nR_2}}] \times \mathcal{Y}_j^n \rightarrow [1, 2^{nR_j}]$$

from the two codebook indices F_1, F_2 and received signal y_j^n to the decoded message $\hat{W}_j = g_j(f_1, f_2, y_j^n)$; Instead, for interference-oblivious decoding we have

$$g_j : [1, |\mathcal{X}_1|^{n2^{nR_j}}] \times \mathcal{Y}_j^n \rightarrow [1, 2^{nR_j}],$$

so that the decoded message $\hat{W}_j = g_j(f_j, y_j^n)$ depends only on the received signal and index of the codebook of the corresponding transmitter alone (not of the interferer).

We say that we have: (i) *Oblivious relaying*: The relay is not aware of both indices F_1 and F_2; (ii) *Interference-oblivious decoding*: Destination j only knows index F_j and not F_i, $i \neq j$.

Definition 2: A rate pair (R_1, R_2) is said to be achievable if there exists a sequence of codes such that $\Pr[(\hat{W}_1, \hat{W}_2) \neq (W_1, W_2)] \rightarrow 0$, where the probability is taken with respect to a uniform distribution of messages W_1 and W_2 and with respect to independent indices F_1 and F_2 whose joint distribution is given by the product of (1) for $j = 1, 2$. The capacity region $\mathcal{C}$ is the closure of the union of all achievable rates.

Remark 1: The definition of oblivious processing obtained from (1), which is the same as in [1], rules out general multiletter input distributions, thus limiting the space of feasible coding schemes, but does not exclude standard "single-letter" coding schemes such as superposition coding and rate-splitting strategies. Moreover, the definition does not allow time-sharing. In fact, in case the transmitters time-share among different codewords, the condition (1) is not satisfied for a given time-sharing schedule. The following alternative definition of oblivious processing instead enables time-sharing.

Definition 3: *Oblivious processing (relaying or decoding) with enabled time-sharing* refers to codes defined as in Definition 1 with the difference that encoders, relay and decoders are all aware of a time-sharing sequence $q^n \in \mathcal{Q}^n$, defined over a finite alphabet $\mathcal{Q}$. Encoding and decoding functions $\{\phi_j, g_j\}$ defined above are modified to depend on q^n. Moreover, codebook generation is constrained so that

$$p_{F_j}(f|q^n) = \prod_{w \in [1, 2^{nR_j}]} p_{X^n|Q^n}(\phi_j(f, w)|q^n), \qquad (2)$$

where $p_{X^n|Q^n}(x^n|q^n) = \prod_{i=1}^n p_{X|Q}(x_i|q_i)$ for a conditional pmf $p_{X|Q}(x_i|q_i)$, instead of (1).

Remark 3: Depending on the application, it may be feasible or not for the relay to acquire the time-sharing sequence q^n decided by sources and destinations. Notice that acquiring the time-sharing sequence is in any case much less demanding that obtaining the full codebook information. If it is possible to acquire q^n, then the definition (2) is appropriate, otherwise the original definition (1) should be adopted.

As a result of the constraints assumed on the coding function, we have the following facts.

Lemma 1 [1]: Given an oblivious processing code for the PIRC, the distribution of a transmitted codeword of source j is given by $p_{X_j^n}(x^n) = \prod_{i=1}^{n} p_{X_j}(x_i)$. In other words, in the absence of information regarding the index F_j and the message W_j, a codeword $x_j^n(F_j, W_j)$ taken from a (n, R_1, R_2) codebook is i.i.d. As a consequence, the received signals at destinations and relay are also i.i.d. vectors.

Lemma 2: Given an oblivious codebook code for the PIRC with enabled time-sharing, the distribution of a transmitted codeword of source j, conditioned on the time-sharing sequence is given by $p_{X^n|Q^n}(x^n|q^n) = \prod_{i=1}^{n} p_{X|Q}(x_i|q_i)$. In other words, in the absence of information regarding the index F_j and the message W_j, a codeword $x_j^n(F_j, W_j)$ taken from a (n, R_1, R_2) codebook has independent, but non-indentically distributed, entries.

Remark 4: While the unconditional pmf $p_{X_j^n}(x^n)$, or $p_{X^n|Q^n}(x^n|q^n)$, factorizes as discussed above, the conditional pmf $p_{X_j^n|F}(x^n|f)$, or $p_{X^n|Q^n,F}(x^n|q^n, f)$, given the key $F_j = f$ does not. In other words, as shown in [2], given a specific "good" code, the empirical distribution with respect to the choice of the message W_j can never be i.i.d. (except for extreme cases such as noiseless channels).

III. PRIMITIVE RELAY CHANNEL WITH OBLIVIOUS RELAYING

We start by analyzing the PRC with oblivious relaying.

Proposition 1: The capacity of a primitive relay channel with oblivious relaying and enabled time-sharing is given by

$$\mathcal{C} = \max I(X; Y\hat{Y}_3|Q) \tag{3a}$$
$$\text{s.t. } C \geq I(Y_3; \hat{Y}_3|YQ) \tag{3b}$$

where maximization is taken with respect to the distribution $p(q)p(x|q)p(\hat{y}_3|y_3, q)$ and the mutual informations are evaluated with respect to

$$p(q)p(x|q)p(\hat{y}_3|y_3, q)p(y, y_3|x). \tag{4}$$

If time-sharing is not allowed, (3) is still an upper bound on the capacity, and the following rate is achievable (i.e., Q =const)

$$\mathcal{C} = \max I(X; Y\hat{Y}_3) \tag{5a}$$
$$\text{s.t. } C \geq I(Y_3; \hat{Y}_3|Y) \tag{5b}$$

Proof: See Appendix A.

Remark 5: Capacity is attained by Compress-and-Forward (CF) with time sharing. This may not be surprising, given that the relay is incapable by design of decoding the codeword

transmitted by the source. However, notice that in the setting of [1] where multiple relays are present but no direct link between source and destination is in place, optimality of (distributed) CF strategies remains elusive. This is in accordance with the current state of the art on the corresponding *source coding* problems, where the source (rather than being an encoded sequence) is a given i.i.d. process to be reconstructed at the destination. In fact, the source coding counterpart of [1] is the (discrete memoryless) CEO problem, which is still generally unsolved [8], while the source coding counterpart of the PRC is the Wyner-Ziv scenario of source coding with side information, whose solution is well-known (see, e.g., [4]). For a discussion on other scenarios where CF was shown to be optimal, we refer to [6].

Remark 6: In (3), variable Q allows time sharing. The fact that the performance of CF can be generally improved by time-sharing was shown in [5, Theorem 2]. In case, time-sharing is not allowed, rate (5) is achievable, which is generally smaller than (3).

A. Gaussian Model

Here we turn to the memoryless Gaussian PRC, that is defined as

$$Y_{3i} = \sqrt{\alpha}X_i + N_{3i}, \ Y_i = X_i + N_i, \tag{6a}$$

where N_{3i}, N_i are independent zero-mean unit-power, and the power constraint is given by $1/n \sum_{i=1}^{n} E[X_i^2] \leq P$. The result of Proposition 1 can be extended using standard arguments to continuous channels and thus to the Gaussian channel (6). However, optimization of the input distribution $p(q)p(x|q)p(\hat{y}_3|y_3, q)$ in (3) remains an open problem. Achievable rates using Gaussian input distribution $p(x|q)$ and quantization test channel $p(\hat{y}_3|y_3, q)$ in (3) can be found in [7] and [5, Theorem 2] without and with time-sharing random variable Q, respectively. However, as discussed in [1], a Gaussian input distribution is generally not optimal and, as seen in [7], non-Gaussian test channels may be advantageous, especially with a non-Gaussian input distribution. Nevertheless, the next proposition shows that the suboptimality of Gaussian channel inputs, Gaussian test channel and no time-sharing, is at most half bit (per (real) channel use), *even if one allows non-oblivious relaying*.

Proposition 2: The rate achievable via CF (and hence oblivious relaying)

$$R_{CF} = \frac{1}{2}\log_2\left(1 + P + \frac{\alpha P}{1 + \frac{1+P+\alpha P}{(2^{2C}-1)(P+1)}}\right) \tag{7}$$

on the Gaussian PRC (6), by employing Gaussian channel inputs, Gaussian test channel and no time-sharing, is at most half bit away from the capacity of the PRC with codebook-aware (and thus also oblivious) relaying.

Proof: The proof is obtained by comparing the achievable rate (7) (that can be found in, e.g., [7]) with the cut-set bound

upper bound (which holds even with non-oblivious relaying)

$$R_{UB} = \min\left\{\frac{1}{2}\log_2(1+P)+C,\ \frac{1}{2}\log_2(1+\alpha P+P)\right\}. \tag{8}$$

See full derivation in Appendix B.

IV. Primitive Interference Relay Channel with Oblivious Relaying

We not turn to the analysis of the PIRC with oblivious relaying. The following proposition shows that in the presence of interference-oblivious decoding, it is optimal for the relay to employ CF and for the destinations to treat the interfering signal as noise.

Proposition 3: The capacity region of the PIRC with oblivious relaying, interference-oblivious decoding and enabled time-sharing is given by the set of all non-negative pairs (R_1, R_2) that satisfy

$$R_j \le I(X_j; Y_j \hat{Y}_3^{(j)} | Q), \text{ for } j=1,2, \tag{9}$$

for some distribution $p(q)\prod_{j=1}^2 p(x_j|q)p(\hat{y}_3^{(j)}|y_3,q)$ $p(y_1,y_2|x_1,x_2)$ that satisfy

$$C_j \ge I(Y_3; \hat{Y}_3^{(j)} | Y_j Q) \text{ for } j=1,2. \tag{10}$$

If time-sharing is not enabled, the above is an outer bound to the capacity region and setting $Q =$const leads to an achievable rate region.

Proof: Follows similarly to the proof of Proposition 1.

V. Appendix

A. Appendix-A: Proof of Proposition 1

Achievability follows by CF with Wyner-Ziv coding and time-sharing determined by variable Q (see, e.g., [3] [7] and [5]). For the converse, consider the first the variable S transmitted by the relay to the destination over the finite-capacity link. Denote as $\tilde{Q}$ the vector of time-sharing variables q^n in Definition 2

$$
\begin{aligned}
nC &\ge& H(S) \ge H(S|\tilde{Q}) \\
&\ge& I(S; X^n Y_3^n | Y^n \tilde{Q}) \ge \sum_{i=1}^n I(S; Y_{3i} | Y^n \tilde{Q} Y_3^{i-1} X^{i-1}) \\
&=& \sum_{i=1}^n H(Y_{3,i}|Y_i\tilde{Q}) - H(Y_{3,i}|\hat{Y}_{3i} Y_i \tilde{Q}) \\
&=& \sum_{i=1}^n I(Y_{3i}; \hat{Y}_{3i} | Y_i \tilde{Q}),
\end{aligned}
$$

where in the third line we used the fact that Y_3^n, Y^n, X^n have conditionally independent entries given $\tilde{Q}$, due to Lemma 2, and we defined $\hat{Y}_{3i} = [SX^{i-1}Y_3^{i-1}Y^{i-1}Y_{i+1}^n]$. Notice that the following Markov chain $(Y_i, X_i) - (Y_{3i}, \tilde{Q}) - \hat{Y}_{3i}$ holds. Now, introducing a variable Q', independent of all other variables and uniformly distributed in $[1,n]$, defining $Y_3 = Y_{3Q'}$ and similarly for the other variables, and $Q = [\tilde{Q}\ Q']$, we get the constraint (3b). Notice that with these definitions we have the Markov chain $(Y, X) - (Y_3, Q) - \hat{Y}_3$. Turning to the destination, using Fano

inequality $H(W|Y^n SF\tilde{Q}) \le n\epsilon_n$ with $\epsilon_n \to 0$ for $n \to \infty$ (for vanishing probability of error), we obtain

$$
\begin{aligned}
nR &\le& I(W; Y^n SF | \tilde{Q}) + n\epsilon_n \\
&=& H(Y^n S|\tilde{Q}) + H(F|Y^n S\tilde{Q}) \\
&& -H(F|W\tilde{Q}) - H(Y^n S|FW\tilde{Q}) + n\epsilon_n \\
&=& I(FW; Y^n S|\tilde{Q}) - I(F; Y^n S|\tilde{Q}) + n\epsilon_n \\
&\le& I(X^n; Y^n S|\tilde{Q}) + n\epsilon_n \\
&=& \sum_{i=1}^n H(X_i|\tilde{Q}) - H(X_i|Y_i\hat{Y}_{3i}\tilde{Q}) \\
&=& \sum_{i=1}^n I(X_i; Y_i\hat{Y}_{3i}|\tilde{Q}) + n\epsilon_n,
\end{aligned}
$$

where in the third equality we have used the fact that F and W are independent and in the last line we have used Lemma 2.

B. Appendix-B: Proof of Proposition 2

We first rewrite (7) as $R_{CF} = \frac{1}{2}\log_2\left(\frac{2^{2C}(1+P)(1+\alpha P+P)}{2^{2C}(1+P)+\alpha P}\right)$, which can be proved by standard algebraic manipulations. Now, assume first that $1 + P(1+\alpha) \le 2^{2C}(1+P)$ so that the upper bound (8) reads $R_{UB} = \frac{1}{2}\log_2(1+\alpha P+P)$. Under this condition, the achievable rate R_{CF} satisfies

$$
\begin{aligned}
R_{CF} &=& R_{UB} - \frac{1}{2}\log_2\left(1 + \frac{\alpha P}{2^{2C}(1+P)}\right) \\
&\ge& R_{UB} - \frac{1}{2}\log_2\left(1 + \frac{(2^{2C}-1)(1+P)}{2^{2C}(1+P)}\right) \\
&\ge& R_{UB} - \frac{1}{2},
\end{aligned}
$$

where the second inequality follows from the assumed condition. The same inequality is proved in a similar way under the complementary condition $1 + P(1+\alpha) \ge 2^{2C}(1+P)$.

References

[1] A. Sanderovich, S. Shamai, Y. Steinberg, and G. Kramer, "Communication via decentralized processing," *IEEE Trans. Inform. Theory,* vol. 54, no. 7, pp. 3008-3023, July 2008.

[2] S. Shamai (Shitz) and S. Verdú, "The empirical distribution of good codes," *IEEE Trans. Inform. Theory*, vol. 43, no. 3, pp. 836-846, May 1997.

[3] Y.-H. Kim, "Coding techniques for primitive relay channels," in *Proc. Forty-Fifth Annual Allerton Conf. Commun., Contr. Comput.*, Monticello, IL, Sept. 2007.

[4] G. Kramer, *Topics in multi-user information theory*, Foundations and Trends in Communications and Information Theory, vol. 4, no. 4-5, pp. 265-444, 2007.

[5] A. El Gamal, M. Mohseni and S. Zahedi, "Bounds on capacity and minimum energy-per-bit for AWGN relay channels," *IEEE Trans. Inform. Theory*, vol. 52, no. 4, pp. 1545-1561, Apr. 2006.

[6] W. Kang and S. Ulukus, "Capacity of a class of diamond channels," submitted [arXiv:0808.0948].

[7] R. Dabora and S.D. Servetto, "On the role of estimate-and-forward with time sharing in cooperative communication," *IEEE Trans. Inform. Theory*, vol. 54, no. 10, pp. 4409-4431, Oct. 2008.

[8] V. Prabhakaran, D. Tse and K. Ramachandran, "Rate region of the quadratic Gaussian CEO problem," in *Proc. IEEE International Symposium on Information Theory* (ISIT 2004), pp. 119, July 2004.

[9] A. Carleial, "Interference channels," *IEEE Trans. Inform. Theory*, vol. 24, no. 1, pp. 60-70, Jan. 1978.