On Codebook Information for Interference Relay Channels With Out-of-Band Relaying

Osvaldo Simeone, *Member, IEEE*, Elza Erkip, *Fellow, IEEE*, and Shlomo Shamai (Shitz), *Fellow, IEEE*

***Abstract*—A standard assumption in network information theory is that all nodes are informed at all times of the operations carried out (e.g., of the codebooks used) by any other terminal in the network. In this paper, information theoretic limits are sought under the assumption that, instead, some nodes are not informed about the codebooks used by other terminals. Specifically, capacity results are derived for a relay channel in which the relay is oblivious to the codebook used by the source (oblivious relaying), and an interference relay channel with oblivious relaying and in which each destination is possibly unaware of the codebook used by the interfering source (interference-oblivious decoding). Extensions are also discussed for a related scenario with standard codebook-aware relaying but interference-oblivious decoding. The class of channels under study is limited to out-of-band (or "primitive") relaying: Relay-to-destinations links use orthogonal resources with respect to the transmission from the source encoders. Conclusions are obtained under a rigorous definition of oblivious processing that is related to the idea of randomized encoding. The framework and results discussed in this paper suggest that imperfect codebook information can be included as a source of uncertainty in network design along with, e.g., imperfect channel and topology information.**

***Index Terms*—Codebook information, femtocells, interference channel, relay channel, robust coding.**

I. Introduction

A STANDARD, and often implicit, assumption in network-information theoretic analyses is that the design of the encoding and decoding functions of all nodes in the network can be performed jointly. As an example, consider the relay channel, in which the operation at the relay terminal is assumed to be jointly designed with the encoding/decoding functions at the source and destination, respectively. Another example is the interference channel, in which each decoder is assumed to be aware not only of the codebook of the intended source, but also of that of the interfering source (see, e.g., [1] and [2]).

Manuscript received March 19, 2010; revised August 18, 2010; accepted December 30, 2010. Date of current version April 20, 2011. O. Simeone was supported by U.S. NSF Grant CCF-0905446. E. Erkip was supported in part by U.S. NSF Grant CNS-0905446 and in part by IIP-1032035. S. Shamai was supported in part by the European Commission in the framework of the FP7 Network of Excellence in Wireless COMmunications, NEWCOM++ and in part by the Israel Science Foundation. The material in this paper was presented in part at the International Zurich Seminar on Communications, Zurich, Switzerland, March 2010.

O. Simeone is with the CWCSPR, ECE Department, New Jersey Institute of Technology, Newark, NJ 07102 USA (e-mail: osvaldo.simeone@njit.edu).

E. Erkip is with the Department of Electrical and CE, Polytechnic Institute of NYU, Brooklyn, NY 11201 USA (e-mail: elza@poly.edu).

S. Shamai (Shitz) is with the Department of Electrical Engineering, Technion, Haifa, 32000, Israel (e-mail: sshlomo@ee.technion.ac.il).

Communicated by A. S. Avestimehr, Associate Editor for the special issue on "Interference Networks".

Digital Object Identifier 10.1109/TIT.2011.2120030

The assumption of joint network-wise design and full codebook knowledge becomes problematic in scenarios when complexity and robustness of network operation and signaling overhead are primary concerns. In fact, this standard approach implies that all nodes must be potentially aware at all times of any change in, say, the codebooks (e.g., modulation) used by any other terminal in the network, or of the addition of new nodes or failure of existing ones. The question thus arises as to whether one can build an information-theoretic understanding of the performance limits of a given network waiving the assumption of full codebook knowledge. In equivalent terms, one may want to impose some form of robustness to the uncertainty at a terminal regarding the codebooks employed by other nodes.

The question posed above was, to the best of the authors' knowledge, first formulated in [3], where the assumption of lack of codebook information was referred to as *oblivious processing*. The scenario studied in [3] consists of a multi-relay channel with a single source and destination without a direct link between them. Transmission from each of the relays is assumed to be *out-of-band*, in the sense that it takes place over an orthogonal, dedicated, link.[1] Under the assumption of oblivious processing (relaying), to be reviewed below, upper and lower bounds on the capacity were derived.

In this work, we further elaborate on the question at hand, by studying the two basic models shown in Fig. 1. The first is a relay channel in which the relay is oblivious to the codebook shared by source and destination (*oblivious relaying*, Fig. 1(a)); The second is an interference channel in which we have oblivious relaying and, furthermore, each destination may be unaware of the codebook of the interfering source [*interference-oblivious relaying*, Fig. 1(b)]. To simplify the analysis, and to model communication scenarios of current interest, such as cellular systems with femtocells (see discussion in Section VI), we assume out-of-band relaying, as in [3]. Following the nomenclature of [4], we refer to these models as Primitive Relay Channel (PRC) and Primitive Interference Relay Channel (PIRC), respectively. Our main contributions are as follows:

- We revisit the definition of oblivious processing of [3] and propose a variation that allows time-sharing (while still enforcing lack of information about the codebooks) in Section II;
- We derive the capacity of the PRC under the constraint of oblivious relaying with enabled time-sharing in Section III. The main contribution here is in the converse part since achievability follows from well-known Compress-and-Forward techniques [11];

[1]Out-of-band relaying is also referred to as primitive in the literature [4]–[6].

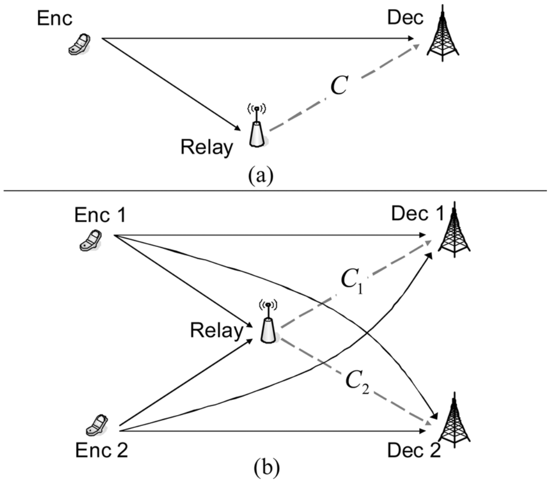

Fig. 1. Two channel models considered in this paper. (a) PRC where the relay is not aware of the codebook shared by source and destination (oblivious relaying). (b) PIRC with oblivious relaying and where each destination may not be aware of the codebook of the interfering source (interference-oblivious decoding).

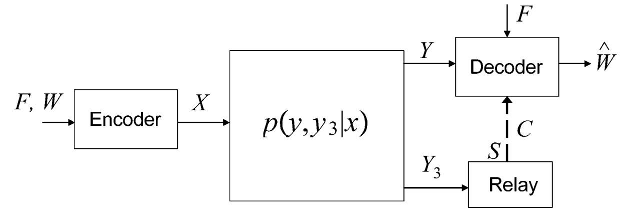

Fig. 2. PRC with oblivious relaying: The relay is not informed about the realization of the codebook, which is indexed by F.

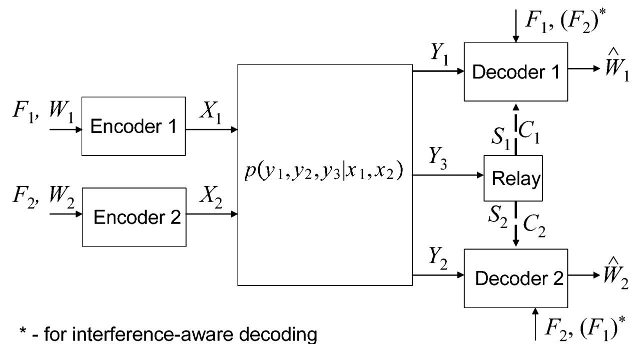

Fig. 3. PIRC with oblivious relaying and interference-oblivious/aware decoding: The relay is not informed about the codebooks F_1, F_2 selected by the two transmitters, and the destinations are unaware/aware of the codebook of the interfering encoder.

- Exploiting the result above for the PRC, we obtain the capacity region of the PIRC under the constraint of oblivious relaying and interference-oblivious decoding with enabled time sharing in Section IV;
- We derive the sum-rate capacity of a symmetric PIRC under the constraint of oblivious relaying with enabled time-sharing in Section IV;
- We discuss a PIRC with interference-oblivious decoding and enabled time-sharing, but where the relay is cognizant of the codebooks (i.e., standard codebook-aware relaying): The capacity region is obtained for a special modulo-additive noise model in Section V.

It should be finally mentioned that general capacity results for the channel models at hand are unavailable (see, e.g., [4]). The conclusive results obtained in this work are made possible by the constraint of oblivious processing, which limits the set of possible transmission strategies.

Notation: Standard definitions $[1,n] = \{1,2,\ldots,n\}$ and $X^n = [X_1, X_2, \ldots, X_n]$ are used; Capital letters generally denote random variables, while a realization thereof is represented by a lowercase letter; Probability mass functions (pmfs) are denoted as $p_X(x) = \Pr[X = x]$ or, for short, by identifying the random variable via the argument as $p(x) = \Pr[X = x]$; $|\mathcal{S}|$ denotes the cardinality of set $\mathcal{S}$.

II. System Model

We study the PRC and PIRC with oblivious processing sketched in Fig. 1 and detailed in Figs. 2 and 3, respectively. In the PRC a single source-destination pair is aided by a relay, whereas in the PIRC two mutually interfering source-destination pairs are aided by a relay. Since the PRC is a special case of the PIRC, we will often detail the system model only for the PIRC. We use the term "primitive" in a similar fashion to [4] to mean that the relay is connected to the destinations via two dedicated finite-capacity links, which are out-of-band (orthogonal) with respect to the transmissions by the sources.

The PIRC consists of: i) A discrete memoryless channel (DMC) $(\mathcal{X}_1, \mathcal{X}_2, p(y_1, y_2, y_3 \,|\, x_1, x_2), \mathcal{Y}_1, \mathcal{Y}_2, \mathcal{Y}_3)$ that relates the transmissions from the two sources, $X_{1,t} \in \mathcal{X}_1$ and $X_{2,t} \in \mathcal{X}_2$, to the symbols received by the two destinations, $Y_{1,t} \in \mathcal{Y}_1$ and $Y_{2,t} \in \mathcal{Y}_2$, and by the relay, $Y_{3,t} \in \mathcal{Y}_3$, $t \in [1,n]$; ii) Two orthogonal links between the relay and each of the two destinations of capacity C_1 and C_2 (bits/channel use),, respectively, see Fig. 3. The operation of the PIRC prescribes standard encoding by the sources over n channel uses of the DMC; The relay receives Y_3^n and maps it into messages S_1 and S_2 of nC_1 and nC_2 bits, respectively, which are sent to the two destinations; Each jth destination, $j = 1, 2$, decodes based on the received signals Y_j^n (on the DMC) and S_j (from the relay).

For the PRC, there is only one source-destination pair, which corresponds to setting $\mathcal{X}_2 = \mathcal{Y}_2 = \emptyset$ in the PIRC: In this case, we drop the subscript 1 for the only source-destination pair for simplicity, so that the PRC is defined by: i) The discrete memoryless channel $(\mathcal{X}, p(y, y_3 \,|\, x), \mathcal{Y}, \mathcal{Y}_3)$; ii) The orthogonal link between relay and destination, of capacity C (bits/channel use), see Fig. 2. The operation of the PRC follows that of the PIRC discussed above. We will also consider a Gaussian model with power constraints, instead of the discrete memoryless model, to be introduced later.

Design of encoding and decoding at sources, destinations and relay must satisfy the oblivious processing constraint, which is discussed next.

A. Oblivious Processing

Oblivious processing was first introduced in [3]. The basic idea is that of using *randomized encoding* (see, e.g., [7]) to model lack of information about the codebooks: The encoders select their codebooks randomly, and "oblivious" nodes are not informed about the currently selected codebook, while nodes that are aware of the codebooks are given such information. Specifically, the codeword $X^n(F, W)$ transmitted by any encoder (we temporarily drop the user subscript to simplify the notation) depends not only on the message W, but also on the index F, which runs over all possible codebooks of the given rate R, i.e., $F \in [1, |\mathcal{X}|^{n2^{nR}}]$, and is selected randomly. Oblivious nodes are not informed about the codebook index F. Moreover, the following condition is imposed on the design of the randomized codes.

Definition 1: A (n, R) *code for oblivious processing* is a pair (p_F, ϕ^n), where $p_F(f)$ is a pmf over the set of codebooks $f \in [1, |\mathcal{X}|^{n2^{nR}}]$ and ϕ^n is a function

$$\phi^n : [1, |\mathcal{X}|^{n2^{nR}}] \times [1, 2^{nR}] \to \mathcal{X}^n \tag{1}$$

which maps a codebook index $f \in [1, |\mathcal{X}|^{n2^{nR}}]$ and a message $w \in [1, 2^{nR}]$ into a codeword $x^n(f, w) = \phi^n(f, w)$. The pair (p_F, ϕ^n) must satisfy

$$\Pr[X^n(F, W) = x^n] = \prod_{i=1}^{n} p_X(x_i) \tag{2}$$

for some pmf $p_X(x), x \in \mathcal{X}$, where $\Pr[\cdot]$ is calculated with respect to the joint distribution of F and W

$$p_{F,W}(f, w) = p_F(f) \cdot 2^{-nR} \tag{3}$$

for $(f, w) \in [1, |\mathcal{X}|^{n2^{nR}}] \times [1, 2^{nR}]$.

Remark 1: Condition (2) states that, when averaged over the probability (3) of selecting a given codebook F and over a uniform distribution on the message set, the transmitted codeword has a product distribution, and, in particular, it has independent identically distributed (i.i.d.) entries with a pmf $p_X(x)$. In essence, this means that an oblivious node, not informed about codebook F and message W, sees a signal $X^n(F, W)$ that lacks of any structure, being i.i.d. This is in contrast to what happens to codebook-aware nodes: In fact, for any fixed (good) code $F = f$ the probability $\Pr[X^n(F = f, W) = x^n | F = f]$ does not have a product form [8]. In other words, a node that is informed about the codebook $F = f$ sees a structured codeword $X^n(F = f, W)$, where the structure is provided by the codebook.

Remark 2: As detailed in Definitions 3 and 4 below, we will define the probability of error on average with respect to the codebook and message distribution $p_{F,W}(f, w)$ (3), similar to the use of randomized encoding in the study of channels with unknown states. However, the discussion above underlines the fact that here randomized encoding is used solely as a means to model oblivious processing, and not as a transmission strategy to improve system performance as for channels with unknown states [2]. In other words, the use of randomized codes here is to be seen merely as a mechanism to model uncertainty about the codebooks at the oblivious nodes: The fact that the transmitters pick their codebooks randomly is not thought of as representing the actual communication scenario at hand, but as a rather natural way to model the obliviousness of some nodes to the actual codebooks.

Remark 3: In [3], a code for oblivious processing was defined by imposing a condition on the distribution $p_F(f)$ of codebook selection, which leads to (2) (see [3, Lemma 1]). While the definition given here is possibly less general, we feel that it is more intuitive, and it does not modify the results. Also, note that function ϕ^n in (1) can be chosen, without loss of generality, to identify all possible $|\mathcal{X}|^{n2^{nR}}$ codebooks of rate R over alphabet $\mathcal{X}$.

Remark 4: The constraint of oblivious processing (2) limits the set of available transmission strategies. For instance, it rules out general "multiletter input distributions" schemes (see, e.g., [9]) and also time-sharing. However, it does not exclude standard "single-letter" coding schemes such as superposition coding and rate-splitting strategies [10].

B. Oblivious Processing Revisited: Enabling Time-Sharing

As discussed in Remark 4, Definition 1 of oblivious processing rules out time-sharing. This is because, conditioned on a given time-sharing sequence q^n, i.e., on the set of time instants where encoders switch among different codebooks [2], (2) is generally not satisfied. Ruling out time-sharing may be justified in that oblivious nodes, that are by definition not informed about the codebooks, may be constrained to be unaware of a time-sharing sequence as well. However, in some scenarios, it may be more appropriate to assume that, while still uniformed about the codebooks, oblivious nodes be at least informed about the time-sharing sequence. Acquiring the latter is generally much less demanding than obtaining the full information about the codebooks.

An example that clarifies the difference between acquiring the codebooks and acquiring the time-sharing sequence is the following. With enabled time-sharing, the two encoders may arrange their transmission according to a Time Division schedule, so that encoder 1 transmits for a certain fraction of time while encoder 2 is silent (this may be marked, say, by setting $q_i = 1$ for the corresponding time instants $i \in [1, n]$), and then encoder 2 transmits for the remaining time instants while encoder 1 is silent (marked by $q_i = 2$). When enabling time-sharing, we assume that the oblivious node may obtain the information about which node is transmitting at a certain time, i.e., about the time-sharing sequence q^n, but not about which specific codebook is being used by the transmitting node.[2]

The following alternative definition of codes for oblivious processing enables time-sharing.

Definition 2: A (n, R) *code for oblivious processing with enabled time-sharing* is a pair $(p_{F|Q^n}, \phi^n)$, where $p_{F|Q^n}(f|q^n)$ is a conditional pmf over the set of codebooks $f \in [1, |\mathcal{X}|^{n2^{nR}}]$ and the set of time-sharing sequences $q^n \in \mathcal{Q}^n$, for some finite

[2]It should be emphasized that the concept of enabled time-sharing is broader than the example above, since encoders may in principle also time-share among different codes, and not merely between transmission and "silence."

alphabet $\mathcal{Q}$, and ϕ^n is a function (1) as in Definition 1. The pair $(p_{F|Q^n}, \phi^n)$ must satisfy

$$\Pr[X^n(F,W) = x^n | Q^n = q^n] = \prod_{i=1}^{n} p_{X|Q}(x_i|q_i) \quad (4)$$

for some conditional pmf $p_{X|Q}(x|q), (x,q) \in \mathcal{X} \times \mathcal{Q}$, where $\Pr[\cdot|Q^n = q^n]$ is calculated with respect to the joint distribution of F and W

$$p_{F,W|Q^n}(f,w|q^n) = p_{F|Q^n}(f|q^n) \cdot 2^{-nR} \quad (5)$$

for $(f,w) \in [1, |\mathcal{X}|^{n2^{nR}}] \times [1, 2^{nR}]$.

Remark 5: Condition (4) states that a node which is not informed about the codebook F and message W, but is aware of the time-sharing sequence Q^n, sees the codeword $X^n(F,W)$ as distributed according to a product pmf. Similar to Definition 1, this models the fact that, in the absence of codebook information, the codeword is unstructured, though not i.i.d. as in Definition 1 (see Remark 1). Notice that codes satisfying Definition 2 also satisfy Definition 1 by choosing an alphabet $\mathcal{Q}$ of cardinality $|\mathcal{Q}| = 1$.

C. Achievable Rates

Here we detail on the definition of achievable rates for oblivious relaying and interference-oblivious/aware decoding.

Definition 3: Rates (R_1, R_2) are said to be achievable for the PIRC with *oblivious relaying and interference-oblivious decoding* if there exist sequences of: i) Pairs of (n, R_1) and (n, R_2) codes for oblivious processing (Definition 1) for users 1 and 2, respectively; ii) Relaying functions

$$\phi_3^n : \mathcal{Y}_3^n \to [1, 2^{nC_1}] \times [1, 2^{nC_2}] \quad (6)$$

which map the received sequence $y_3^n \in \mathcal{Y}_3^n$ into two indices $s_j \in [1, 2^{nC_j}]$ to be sent to destinations $j = 1, 2$ as $[s_1, s_2] = \phi_3(y_3^n)$; and iii) Decoding functions

$$g_j^n : \left[1, |\mathcal{X}_1|^{n2^{nR_j}}\right] \times \mathcal{Y}_j^n \to [1, 2^{nR_j}] \quad (7)$$

which map the codebook index F_j of the intended source and received signal y_j^n to the decoded message $\hat{W}_j = g_j(f_j, y_j^n)$, $j = 1, 2$; such that

$$\Pr[\hat{W}_j \neq W_j] \to 0 \quad (8)$$

as $n \to \infty$ for $j = 1, 2$, where the probability is taken with respect to the joint pmf $p_{F_1,F_2,W_1,W_2}(f_1, f_2, w_1, w_2) = \prod_{j=1}^{2} p_{F_j,W_j}(f_j, w_j)$, with $p_{F_j,W_j}(f_j, w_j)$ given by (3). The capacity region $\mathcal{C}$ is the closure of the union of all achievable rates.

Remark 6: The assumption of oblivious relaying translates in the relaying function (6) not depending on the codebook indices F_1 and F_2. Similarly, the assumption of interference-oblivious decoding consists in imposing that decoding (7) does not depend on the codebook index F_i, $i \neq j$, of the interfering source. With interference-aware decoding, this assumption is waived, as shown in the next definition.

Definition 4: Rates (R_1, R_2) are said to be achievable for the PIRC with *oblivious relaying and interference-aware decoding* if the same conditions as in Definition 3 are satisfied, with the difference that the decoding function (7) is defined as

$$g_j^n : \left[1, |\mathcal{X}_1|^{n2^{nR_1}}\right] \times \left[1, |\mathcal{X}_1|^{n2^{nR_2}}\right] \times \mathcal{Y}_j^n \to [1, 2^{nR_j}] \quad (9)$$

which maps the codebook indices F_1, F_2 of both sources and received signal y_j^n to the decoded message $\hat{W}_j = g_j(f_1, f_2, y_j^n)$, $j = 1, 2$.

The set of achievable rates with oblivious relaying and interference-oblivious/aware decoding but *with enabled time-sharing* is defined similar to Definitions 3 and 4, with the differences that: i) Codes are allowed to employ time-sharing according to Definition 2 with the same time sharing sequence q^n; ii) Relaying (6) and decoding functions (7)–(9) depend also on the given time-sharing sequence q^n. For instance, the relaying function (6) becomes

$$\phi_3^n : \mathcal{Y}_3^n \times \mathcal{Q}^n \to [1, 2^{nC_1}] \times [1, 2^{nC_2}] \quad (10)$$

and similarly for the decoding functions (7)–(9); and iii) The probability of error is calculated for the given time-sharing sequence q^n.

The achievable rate R for PRC with oblivious relaying and with/without enabled time-sharing is defined by specializing the definitions above for the PIRC.

We finally remark that in Section V we also consider the case in which the relay is aware of the codebooks of the sources (i.e., of the indices F_1 and F_2), but the destinations are constrained to perform interference-oblivious decoding.

III. PRC 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

$$\mathcal{C} = \max I(X; Y\hat{Y}_3 \,|\, Q) \quad (11a)$$
$$\text{s.t. } C \geq I(Y_3; \hat{Y}_3 \,|\, YQ) \quad (11b)$$

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

$$p(q)p(x\,|\,q)p(\hat{y}_3\,|\,y_3, q)p(y, y_3\,|\,x). \quad (12)$$

If time-sharing is not allowed, (11) is an upper bound on the capacity, and the following rate is achievable (i.e., Q is set to constant)

$$\mathcal{C} = \max I(X; Y\hat{Y}_3) \quad (13a)$$
$$\text{s.t. } C \geq I(Y_3; \hat{Y}_3 \,|\, Y). \quad (13b)$$

Proof: See Appendix A. □

Remark 7: According to Proposition 1, capacity is attained by Compress-and-Forward (CF) with time sharing. It is recalled that with CF, introduced in [11], the relay treats the received signal as an unstructured random process jointly distributed with the signal received at the destination: This allows the application of Wyner-Ziv compression (see, e.g., [1] and [2]). Optimality of CF is not surprising, given that the relay is incapable by design of decoding the codeword transmitted by the source. In fact, due to the assumption of oblivious relaying, the relay sees an unstructured received signal (recall Remarks 1 and 5). The relevance of the approach here is to formalize this intuition through the notion of oblivious relaying.

Remark 8: The capacity of the PRC without the constraint of oblivious relaying is not known in general [4]. However, in some special cases, the capacity has been found to be attained by CF, namely in [4]–[6] (see also [12]–[14], which treat related models). In these scenarios, clearly, the assumption of oblivious relaying does not cause any loss in performance. It is also noted that the reason for the optimality of CF in the models of [5] and [6] (see also [12] and [13]) is similar to the explanation for the optimality of CF with oblivious relaying given in Remark 7. Specifically, in [5], [6], [12], and [13] the signal received by the relay is unstructured, either because it does not contain information about the transmitted signal [5], [6] or because the transmitted signal is uncoded [12], [13].[3]

Remark 9: In (11), variable Q allows time sharing. The fact that the performance of CF can be generally improved by time-sharing was shown in [15, Theorem 2]. In case time-sharing is not allowed, rate (13) is achievable, which is generally smaller than (11).

Remark 10: The proof of the converse in Proposition 1 follows similar to the Wyner-Ziv theorem (see Appendix A). This is due to the fact that the scenario, as seen by the relay, resembles the source coding setting of Wyner-Ziv compression due to the assumption of out-of-band and oblivious relaying. In particular, the scenario would be the same if the signal sent by the source were i.i.d. and given (instead of a codebook subject to design). On a related note, we remark that, in the model studied in [3] 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 distributed source coding setting between the multiple relays and destinations, i.e., the so called CEO problem (see, e.g., [1] and [2]). It is noted that a solution to the CEO problem would likely directly translate into a solution to the problem of finding the capacity for an oblivious system with multiple relays as well.

Remark 11: A related result is presented in [16], where it is shown that, whenever the source-relay link is used above capacity, if one is interested in lossless compression of the relay's received signal, there is no benefit to be accrued from knowledge of the codebook at the relay. In other words, the rate needed for lossless compression of the relay's signal is the same whether the relay knows the source codebook or not. This is because, in this case, the signal received at the relay is unstructured even when the relay knows the codebook, due to the mix-up of all the relay output sequences corresponding to all possible transmitted codewords. We emphasize that this result does not prove optimality of CF for the relay channel under any specific condition, but only focuses on the rate required for lossless compression at the relay.[4]

A. Gaussian Primitive Relay Channel

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

$$Y_{3i} = \sqrt{\alpha} X_i + N_{3i} \tag{14a}$$

$$\text{and } Y_i = X_i + N_i \tag{14b}$$

where N_{3i}, N_i are independent zero-mean unit-power Gaussian noises, $i \in [1, n]$, 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 (14). However, optimization of the input distribution $p(q)p(x \,|\, q)p(\hat{y}_3 \,|\, y_3, q)$ in (11) 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 (11) can be found in [17] and [15, Theorem 2] without and with time-sharing, respectively. As discussed in [3], a Gaussian input distribution is generally not optimal and, as seen in [17], 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 does not impose oblivious relaying*.[5]

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

$$R_{\mathrm{CF}} = \frac{1}{2}\log_2\left(1 + P + \frac{\alpha P}{1 + \frac{1+P+\alpha P}{(2^{2C}-1)(P+1)}}\right) \tag{15}$$

on the Gaussian PRC (14), 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 (Sketch): The proof is obtained by comparing the achievable rate (15) (that can be found in, e.g., [17]) with the cut-set bound upper bound (which holds even with nonoblivious relaying)

$$R_{\mathrm{UB}} = \min\left\{\frac{1}{2}\log_2(1+P) + C, \frac{1}{2}\log_2(1+\alpha P + P)\right\}. \tag{16}$$

See full derivation in Appendix B. □

[3]We also remark that the characterization of the rate given in Proposition 1, and in previous references, lacks bounds on the cardinality of the auxiliary random variables and is thus, strictly speaking, not computable.

[4]Moroever, the side information available at the receiver is not considered.

[5]It is noted that constant gap results as in Proposition 2 are meaningful only at sufficiently large signal-to-noise ratios.

IV. Primitive Interference Relay Channel With Oblivious Relaying

In this section, we study the PIRC with oblivious relaying.

A. Interference-Oblivious Decoding

The following proposition shows that in the presence of interference-oblivious decoding, it is optimal for the relay to employ CF and for each destination 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 nonnegative pairs (R_1, R_2) that satisfy

$$R_j \leq I\left(X_j; Y_j \hat{Y}_3^{(j)} \;\middle|\; Q\right), \quad \text{for } j = 1, 2 \tag{17}$$

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 satisfies

$$C_j \geq I\left(Y_3; \hat{Y}_3^{(j)} \,|\, Y_j Q\right) \quad \text{for } j = 1, 2. \tag{18}$$

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

Proof: Achievability is obtained by CF and treating interference as noise. The converse follows similar to Proposition 1 (see Appendix A).[6] □

Remark 12: Optimality of CF is explained as in Remark 7, whereas the optimality of treating interference as noise is understood in a similar fashion from the fact that one imposes interference-oblivious decoding. In fact, interference-obliviousness basically makes the interfering signal akin to an unstructured (i.i.d.) channel state due to Definition 2. Again, the framework studied in this paper formalizes the intuitive optimality of the strategies at hand through the concept of oblivious processing.

B. Interference-Aware Decoding

In the presence of interference-aware decoding, single-letter capacity results are rare even for interference channels without a relay. Therefore, here we focus on a *symmetric* PIRC such that: i) The outputs Y_1 and Y_2 are statistically equivalent, in the sense that they have equal marginals $p_{Y_1,Y_3\,|\,X_1X_2}(\cdot, y_3\,|\,x_1, x_2) = p_{Y_2,Y_3\,|\,X_1X_2}(\cdot, y_3\,|\,x_1, x_2)$ [18, Theorem 4]; ii) The relay is constrained to send the same message $S_1 = S_2$ to both destinations in a broadcast fashion, which also implies $C_1 = C_2$ (rather than two distinct messages S_1 and S_2 as in the original model of Section II). In this case, extending the arguments in [18, Theorem 4], it can be seen that any decoding operation carried out at any decoder can be reproduced equivalently (in a statistical sense) by the other decoder. As such, the model is equivalent to the primitive multiple access relay channel (PMARC), shown in Fig. 4, and defined similarly to the PRC and PIRC.

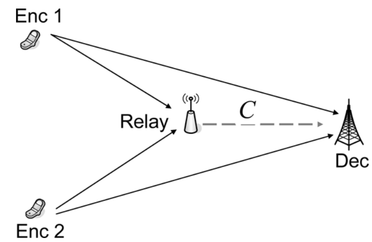

Fig. 4. PMARC considered in Proposition 4.

Proposition 4: The sum-capacity[7] of the symmetric PIRC with oblivious relaying and interference-aware decoding, or equivalently of the PMARC with oblivious relaying, with enabled time sharing, is given by

$$\mathcal{C}_{\text{sum}} = \max I(X_1 X_2; Y_1 \hat{Y}_3 \,|\, Q) \tag{19a}$$

$$\text{s.t. } C \geq I(Y_3; \hat{Y}_3 \,|\, Y_1 Q) \tag{19b}$$

where maximization is taken with respect to the distribution $p(q)p(x_1\,|\,q)p(x_2\,|\,q)p(\hat{y}_3\,|\,y_3, q)$ and the mutual informations are evaluated with respect to

$$p(q)p(x_1\,|\,q)p(x_2\,|\,q)p(\hat{y}_3\,|\,y_3, q)p(y_1, y_3\,|\,x_1, x_2). \tag{20}$$

Proof: Achievability is from CF and joint decoding at each receiver. The converse follows again from the same steps as in proof of Proposition 1 (see Appendix A) with definition $\hat{Y}_{3i} = [SX_1^{i-1}X_2^{i-1}Y_3^{i-1}Y_1^{i-1}Y_{1,i+1}^{n}]$. □

Remark 13: The result of Proposition 2 can be easily extended to symmetric Gaussian PIRC or the equivalent PMARC, noticing that the sum-capacity of a PMARC equals the capacity of a PRC with power constraint at the source given by the sum-power constraint at the sources of the PMARC.

V. Discussion on Codebook-Aware Relaying

In this section, we consider a PIRC in which we have interference-oblivious decoding, but the relay is now aware of the codebooks, i.e., of the codebook indices F_1 and F_2. It is noted that, in this case, the relay can in principle provide partial information about the codebooks employed by the sources (i.e., about F_1 and F_2) to the destinations. While the general problem appears difficult, here we tackle a specific class of channels and demonstrate that optimality of decode-and-forward at the relay coupled with decoding by treating interference as noise at the receivers. In this class of channels, it will be shown that there is no need for the relay to provide codebook information to the receivers.

To elaborate, we focus on a binary PIRC with

$$Y_1 = X_1 \oplus X_2 \tag{21a}$$

$$Y_2 = X_2 \oplus Z_2 \tag{21b}$$

$$\text{and } Y_3 = X_2 \oplus Z_3 \tag{21c}$$

[6] In proving the converse as in Appendix A, one ends up with auxiliary variables $\hat{Y}_3^{(1)}$ and $\hat{Y}_3^{(2)}$ that are not independent given Y_3, Q, unlike what claimed in Proposition 3. However, dependence of $\hat{Y}_3^{(1)}$ and $\hat{Y}_3^{(2)}$ does not affect the mutual informations in (17)–(18) and is therefore immaterial for the result of Proposition 3.

[7] The sum-capacity is given by $\mathcal{C}_{\text{sum}} = \max_{R_1, R_2 \in \mathcal{C}} R_1 + R_2$, where $\mathcal{C}$ is the capacity region.

where $Z_2 \sim Ber(p_2)$, $Z_3 \sim Ber(p_3)$ are independent with $0 \leq p_2, p_3 \leq 1/2$.[8] In the interference channel (21a)–(21b), receiver 1 is affected by interference, while receiver 2 is only impaired by additive noise. The model is thus a special case of a Z-interference channel. Moreover, from (21c), the relay observes only the signal sent by the transmitter of the one link that interferes with the other (i.e., transmitter 2). Denote $x*y = x(1-y)+(1-x)y$.

Proposition 5: If

$$C_2 \leq H(p_2) - H(p_3) \tag{22}$$

the capacity region of the PIRC (21) with interference-oblivious decoding and enabled time-sharing and codebook-aware relaying is given by the closure of the convex hull of the union of all rates that satisfy

$$R_1 \leq \min\{1 - H(p) + C_1, 1\} \tag{23a}$$

$$R_2 \leq \min\{H(p * p_2) - H(p_2) + C_2 \tag{23b}$$

for some $0 \leq p \leq 1/2$.

Proof: The converse follows from the cut-set bound having fixed $X_1 \sim Ber(1/2)$ and $X_2 \sim Ber(p)$ for some $0 \leq p \leq 1/2$ without loss of generality. Notice that the bound (23a) requires the assumption that the sequence X_2^n is i.i.d. at the receiver 1, which is the case here due to the assumption of interference-obliviousness. For the achievability part of the proof, we assume that the relay decodes the message of user 2, which incur in vanishing probability of error provided that

$$R_2 < H(p * p_3) - H(p_3). \tag{24}$$

Then, the relay randomly bins (i.e., hashes) the decoded codeword X_2^n over 2^{nC_1} and 2^{nC_2} bins. The indices of the two bins where the decoded codeword lies are provided to decoder 1 and decoder 2, respectively. Decoding then takes place as in Theorem 1 of [22] for decoder 1 and as in Proposition 2 of [4] for decoder 2. This step is guaranteed to have vanishing error probability if (23) are satisfied [4], [22]. Now, if (22) holds, it can be seen that the right-hand side (RHS) of (24) is always larger than that of (23b),[9] which concludes the proof. □

Remark 14: As discussed in the proof below, the capacity region (23) is achieved by letting the relay decode the message of transmitter 2. The relay then provides partial information about transmitter 2's codeword to both decoder 1 and decoder 2. Decoder 2 uses this information to decode the intended message of user 2 itself. Instead, decoder 1 exploits this information indirectly, not being able to decode the message of user 2 (due to obliviousness) and being only interested in the message of user 1. Notice that this role of the relay towards destination 1 has been referred to as interference forwarding in some related work [21], [20].[10]

[8]The notation $X \sim Ber(p)$ means that X is a binary variable with $\Pr[X = 1] = p$ and $\Pr[X = 0] = 1 - p$.

[9]In fact, the difference between the RHS of (24) and (23b) is increasing with $0 \leq p \leq 1/2$ and therefore the condition should only be checked for $p = 1/2$. This leads to (22).

[10]Interference forwarding in [21], [20] was aimed at aiding decoding of the interference at a destination. However, here such decoding is not possible, and the information from the relay is used at decoder 1 to reduce the uncertainty about the interfering sequence X_2^n [22], [5].

Remark 15: The capacity characterization (23) hinges on the obliviousness of the decoders (in particular, of decoder 1). The capacity region for the channel (21) is in fact generally not known when this assumption is waived. It should also be noted that Proposition 5 is a fairly simple consequence of Proposition 2 of [4] and Theorem 1 of [22] (see also [5]). Moreover, extension to larger-alphabet modulo channels is straightforward.

Remark 16: The model at hand of the PIRC (Fig. 3) is related to the Gaussian interference channel with conferencing decoders studied in [23]. Therein, decoders are connected to one another via out-of-band links. It is proved that a special version of CF in which decoders jointly decompress the quantization index received by the other receiver and the message of interest is almost optimal.

VI. Concluding Remarks

Among the standard assumptions made in network information theoretic analyses, that of full codebook knowledge at all nodes in the network is as ubiquitous as it is questionable, especially when signaling overhead, on the one hand, and robustness and complexity of system design, on the other, are primary concerns. In this paper, we have elaborated on a framework that accounts for modeling of imperfect codebook knowledge, which is referred to as oblivious processing. Based on a rigorous definition of the constraint of oblivious processing on the network encoding/decoding functions, which is extended from [3], we have obtained the capacity region of a class of relay and interference relay channels. The approach in this work suggests to include imperfect codebook knowledge among the practical constraints to be imposed on network design, along with, e.g., imperfect channel and topology state information.

The class of channels studied in this paper assumes out-of-band relaying. Besides simplifying the analysis, out-of-band relaying finds application to the timely scenario of cellular communication in the presence of femtocells: A home base station, serving the femtocell, can indeed be seen as relay connected to the final destination (i.e., the mobile operator controller or the macro base station) via a wired backhaul link, which typically consists of a last-mile connection followed by the Internet [19]. The constraint of oblivious relaying may be of particular interest in these scenarios, especially in the case of open-access femtocells [19], where the home base station serves also outdoor users, whose codebooks may be too expensive to learn or to adapt to.

Finally, the analysis of this paper leaves open a number of directions for future research, such as deriving the capacity region of the PMARC of Fig. 4 or extending the results here to in-band relaying.

Appendix I

A. *Appendix-A: Proof of Proposition 1*

Achievability follows by CF with Wyner-Ziv coding and time-sharing determined by random variable Q (see, e.g., [15] and [17]). Notice that standard randomly generated single-letter codes with a pmf $p_{X\,|\,Q}(x\,|\,q)$, conditioned on a randomly generated time-sharing sequence q^n, can be used at the source encoder, since this form of randomized encoding is a code for oblivious processing in the sense of Definition 2. In fact, on average over the induced codebook generation probability $p_{F\,|\,Q^n}(f\,|\,q^n)$, (4) is satisfied. To be specific, this approach is equivalent to drawing the codebook index F with probability

$$p_{F|Q^n}(f\,|\,q^n) = \prod_{w\in[1,2^{nR}]} p_{X^n\,|\,Q^n}(\phi^n(f,w)\,|\,q^n) \tag{25}$$

where $p_{X^n\,|\,Q^n}(x^n\,|\,q^n) = \prod_{i=1}^n p_{X|Q}(x_i\,|\,q_i)$. That this choice satisfies (4) follows from Lemma 1 in [3].

For the converse, consider 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. We have the series of inequalities

$$\begin{aligned}
nC &\geq H(S) && \text{(26a)}\\
&\geq H(S\,|\,\tilde{Q}) && \text{(26b)}\\
&\geq I(S;X^nY_3^n\,|\,Y^n\tilde{Q}) && \text{(26c)}\\
&= \sum_{i=1}^n I(S;X_iY_{3i}\,|\,Y^nY_3^{i-1}X^{i-1}\tilde{Q}) && \text{(26d)}\\
&\geq \sum_{i=1}^n I(S;Y_{3i}\,|\,Y^nY_3^{i-1}X^{i-1}\tilde{Q}) && \text{(26e)}\\
&= \sum_{i=1}^n H(Y_{3i}\,|\,Y^nY_3^{i-1}X^{i-1}\tilde{Q})\\
&\quad - H(Y_{3i}\,|\,SY^nY_3^{i-1}X^{i-1}\tilde{Q}) && \text{(26f)}\\
&= \sum_{i=1}^n H(Y_{3i}\,|\,Y_i\tilde{Q}) - H(Y_{3,i}\,|\,\hat{Y}_{3i}Y_i\tilde{Q}) && \text{(26g)}\\
&= \sum_{i=1}^n I(Y_{3i};\hat{Y}_{3i}\,|\,Y_i\tilde{Q}) && \text{(26h)}
\end{aligned}$$

where in (26b)–(26c) we have used the fact that conditioning reduces entropy, (26d)–(26e) follow from the chain rule and the positivity of mutual information, in (26g) we used the fact that Y_3^n, Y^n, X^n have conditionally independent entries given $\tilde{Q}$, due to Definition 2 [recall (4)], and we have defined

$$\hat{Y}_{3i} = \left[SX^{i-1}Y_3^{i-1}Y^{i-1}Y_{i+1}^n\right]. \tag{27}$$

It can be proved, as shown in Appendix C, 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 (11b). 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 $H(W\,|\,Y^nSF\tilde{Q}) \leq n\epsilon_n$ with $\epsilon_n \to 0$ for $n \to \infty$ (for vanishing probability of error), we obtain

$$\begin{aligned}
nR &= H(W) = H(W|\tilde{Q}) && \text{(28a)}\\
&\leq I(W;Y^nSF\,|\,\tilde{Q}) + n\epsilon_n && \text{(28b)}\\
&= H(Y^nS\,|\,\tilde{Q}) + H(F\,|\,Y^nS\tilde{Q}) && \text{(28c)}\\
&\quad - H(F\,|\,W\tilde{Q}) - H(Y^nS\,|\,FW\tilde{Q}) + n\epsilon_n && \text{(28d)}\\
&= I(FW;Y^nS\,|\,\tilde{Q}) - I(F;Y^nS\,|\,\tilde{Q}) + n\epsilon_n && \text{(28e)}\\
&\leq I(X^n;Y^nS\,|\,\tilde{Q}) + n\epsilon_n && \text{(28f)}\\
&= \sum_{i=1}^n H(X_i\,|\,X^{i-1}\tilde{Q}) - H(X_i\,|\,SX^{i-1}Y^n\tilde{Q}) && \text{(28g)}\\
&\leq \sum_{i=1}^n H(X_i\,|\,\tilde{Q}) - H(X_i\,|\,Y_i\hat{Y}_{3i}\tilde{Q}) && \text{(28h)}\\
&= \sum_{i=1}^n I(X_i;Y_i\hat{Y}_{3i}\,|\,\tilde{Q}) + n\epsilon_n && \text{(28i)}
\end{aligned}$$

where (28b) follows from the Fano inequality discussed above, (28c) is obtained by first writing $I(W;Y^nSF\,|\,\tilde{Q}) = H(Y^nSF\,|\,\tilde{Q}) - H(Y^nSF\,|\,\tilde{Q}W)$ and then using the chain rule for entropy, in (28f) we have used the fact that X^n is a function of F and W, and in (28h) we have used (4) and the fact that conditioning reduces entropy. Inequality (11a) follows from the definition of Q made above.

B. Appendix-B: Proof of Proposition 2

We first rewrite (15) as $R_{\text{CF}} = \frac{1}{2}\log_2(\frac{2^{2C}(1+P)(1+\alpha P+P)}{2^{2C}(1+P)+\alpha P})$, which can be proved by standard algebraic manipulations. Now, assume first that $1+P(1+\alpha) \leq 2^{2C}(1+P)$ so that the upper bound (16) reads $R_{\text{UB}} = \frac{1}{2}\log_2(1+\alpha P+P)$. Under this condition, the achievable rate R_{CF} satisfies

$$\begin{aligned}
R_{\text{CF}} &= R_{\text{UB}} - \frac{1}{2}\log_2\left(1+\frac{\alpha P}{2^{2C}(1+P)}\right)\\
&\geq R_{\text{UB}} - \frac{1}{2}\log_2\left(1+\frac{(2^{2C}-1)(1+P)}{2^{2C}(1+P)}\right)\\
&\geq R_{\text{UB}} - \frac{1}{2}
\end{aligned} \tag{29}$$

where the second inequality follows from the assumed condition. Finally, assume the complementary condition $1+P(1+\alpha) \geq 2^{2C}(1+P)$, we similarly have

$$\begin{aligned}
R_{\text{CF}} &= R_{\text{UB}} - \frac{1}{2}\log_2\left(\frac{2^{2C}(1+P)+\alpha P}{1+\alpha P+P}\right)\\
&\geq R_{\text{UB}} - \frac{1}{2}\log_2\left(1+\frac{\alpha P}{1+\alpha P+P}\right)\\
&\geq R_{\text{UB}} - \frac{1}{2}
\end{aligned} \tag{30}$$

which concludes the proof.

C. Appendix C: Proof of the Markov Chain for Proposition 1

In order to prove the Markov chain $(Y_i,X_i) - (Y_{3i},\tilde{Q}) - \hat{Y}_{3i}$, we use a graphical method based on d-separation for Bayesian networks, as discussed, e.g., [1, pp. 166–168] or [24]. This is done by first building a Bayesian networks (i.e., a directed graph) that encodes the joint distribution of the variables at hand. This is shown in Fig. 5 and follows easily from the problem setup. Then, one eliminates the dashed arrows in the figures. This step corresponds to conditioning on the variables $(Y_{3i},\tilde{Q})$.

Now, the Markov chain at hand follows from consideration of the undirected graph obtained from the one in the figure by eliminating the dashed arrows and making all the edges non-directed. In particular, the Markov chain at hand is verified since

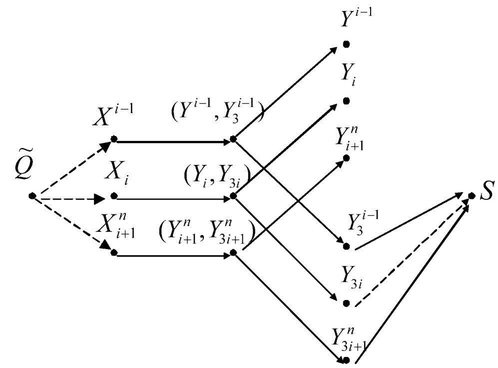

Fig. 5. Illustration of the joint distribution required in Appendix C.

there is no path on such graph between any variable in (Y_i, X_i) and any variable in the auxiliary $\hat{Y}_{3i}$ (27). In other words, one can say that variables $(Y_{3i}, \tilde{Q})$ d-separate (Y_i, X_i) and $\hat{Y}_{3i}$.

REFERENCES

[1] G. Kramer, "Topics in multi-user information theory," *Found. Trends in Commun. Inf. Theory*, vol. 4, no. 4–5, pp. 265–444, 2007.

[2] A. El Gamal and Y.-H. Kim, Lecture Notes on Network Information Theory [Online]. Available: arXiv:1001.3404

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

[4] Y.-H. Kim, "Coding techniques for primitive relay channels," in *Proc. 45th Ann. Allerton Conf. Commun., Contr. Comput.*, Monticello, IL, Sep. 2007.

[5] M. Aleksic, P. Razaghi, and W. Yu, "Capacity of a class of modulo-sum relay channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 3, pp. 921–930, Mar. 2009.

[6] R. Tandon and S. Ulukus, "A new upper bound on the capacity of a class of primitive relay channels," in *Proc. Ann. Allerton Conf. on Commun., Contr., Comput.*, Sep. 2008, pp. 1562–1569.

[7] A. Lapidoth and P. Narayan, "Reliable communication under channel uncertainty," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2148–2177, Oct. 1998.

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

[9] R. S. Cheng and S. Verdú, "On limiting characterizations of memoryless multiuser capacity regions," *IEEE Trans. Inf. Theory*, vol. 39, no. 2, pp. 609–612, Mar. 1993.

[10] T. Han and K. Kobayashi, "A new achievable rate region for the interference channel," *IEEE Trans. Inf. Theory*, vol. 27, no. 1, pp. 49–60, Jan. 1981.

[11] T. M. Cover and A. A. El Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inf. Theory*, vol. IT-25, pp. 572–584, Sep. 1979.

[12] W. Kang and S. Ulukus, Capacity of a class of diamond channels [Online]. Available: [arXiv:0808.0948]

[13] S.-H. Lee and S.-Y. Chung, "When is compress-and-forward optimal?," in *Proc. Inf. Theory Applications (ITA 2010)*, La Jolla, CA, Feb. 2010.

[14] S. H. Lim, Y.-H. Kim, A. El Gamal, and S.-Y. Chung, "Noisy network coding," in *Proc. IEEE Inf. Theory Workshop (IEEE ITW 2010)*, Cairo, Egypt, Jan. 2010, pp. 58–62.

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

[16] X. Wu and L.-L. Xie, Asymptotic equipartition property of output when rate is above capacity [Online]. Available: [arXiv:0908.4445]

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

[18] A. Carleial, "Interference channels," *IEEE Trans. Inf. Theory*, vol. 24, no. 1, pp. 60–70, Jan. 1978.

[19] V. Chandrasekhar, J. G. Andrews, and A. Gatherer, "Femtocell networks: A survey," *IEEE Commun. Mag.*, vol. 46, no. 9, pp. 59–67, Sept. 2008.

[20] I. Maric, R. Dabora, and A. Goldsmith, "On the capacity of the interference channel with a relay," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT 2008)*, Toronto, , Ontario, Canada, Jul. 6–11, 2008, pp. 554–558.

[21] O. Sahin and E. Erkip, "Achievable rates for the Gaussian interference relay channel," in *Proc. Global Telecommun. Conf. (GLOBECOM 2007)*, Wash., DC, Nov. 26–30, 2007, pp. 1627–1631.

[22] Y.-H. Kim, "Capacity of a class of deterministic relay channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 3, pp. 1328–1329, Mar. 2008.

[23] I.-H. Wang and D. N. C. Tse, Interference mitigation through limited receiver cooperation [Online]. Available: [arXiv:0911.2053]

[24] D. Koller and N. Friedman, *Probabilistic Graphical Models: Principle and Techniques*. Cambridge, MA: MIT Press, 2009.

Osvaldo Simeone (M'02) received the M.Sc. degree (with honors) and the Ph.D. degree in information engineering from the Politecnico di Milano, Milan, Italy, in 2001 and 2005, respectively.

He is currently with the Center for Wireless Communications and Signal Processing Research (CWCSPR), New Jersey Institute of Technology (NJIT), Newark, where he is an Assistant Professor. His current research interests concern the cross-layer analysis and design of wireless networks with emphasis on information-theoretic, signal processing, and queuing aspects. Specific topics of interest are: cooperative communications, unlicensed spectrum access, ad hoc, sensor, mesh and hybrid networks, distributed estimation, and synchronization.

Dr. Simeone is the corecipient of the best paper award of the IEEE SPAWC 2007 and the IEEE WRECOM 2007. He was an Editor for IEEE TRANSACTIONS WIRELESS COMMUNICATIONS between 2007–2010 and he currently serves as an Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS.

Elza Erkip (S'93–M'96–SM'05–F'11) received the B.S. degree in electrical and electronics engineering from the Middle East Technical University, Ankara, Turkey, and the M.S. and Ph.D. degrees in electrical engineering from Stanford University, Stanford, CA.

Currently, she is an Associate Professor of electrical and computer engineering at the Polytechnic Institute of New York University, Brooklyn. In the past, she has held positions at Rice University, Houston, TX, and at Princeton University, Princeton, NJ. Her research interests are in information theory, communication theory, and wireless communications.

Dr. Erkip received the National Science Foundation CAREER Award in 2001, the IEEE Communications Society Rice Paper Prize in 2004, and the ICC Communication Theory Symposium Best Paper Award in 2007. She co-authored a paper that received the ISIT Student Paper Award in 2007. She was a Finalist for The New York Academy of Sciences Blavatnik Awards for Young Scientists in 2010. Currently, she is an Associate Editor for the IEEE TRANSACTIONS ON INFORMATION THEORY and a Technical Co-Chair of WiOpt 2011. She was an Associate Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS from 2006–2009, a Publications Editor for the IEEE TRANSACTIONS ON INFORMATION THEORY from 2006–2009, and a Guest Editor of the *IEEE Signal Processing Magazine* in 2007. She was the co-chair of the GLOBECOM Communication Theory Symposium in 2009, the publications chair of ITW Taormina in 2009, the MIMO Communications and Signal Processing Technical Area chair of the Asilomar Conference on Signals, Systems, and Computers in 2007, and the technical program co-chair of the Communication Theory Workshop in 2006.

Shlomo Shamai (Shitz) (F'94) received the B.Sc., M.Sc., and Ph.D. degrees in electrical engineering from the Technion—Israel Institute of Technology, in 1975, 1981, and 1986, respectively.

During 1975–1985, he was with the Communications Research Labs in the capacity of a Senior Research Engineer. Since 1986, he has been with the Department of Electrical Engineering, Technion—Israel Institute of Technology, where he is now the William Fondiller Professor of Telecommunications. His research interests encompasses a wide spectrum of topics in information theory and statistical communications.

Dr. Shamai (Shitz) is the recipient of the 2011 Claude E. Shannon Award. He is the recipient of the 1999 van der Pol Gold Medal of the Union Radio Scientifique Internationale (URSI), and a corecipient of the 2000 IEEE Donald G. Fink Prize Paper Award, the 2003, and the 2004 joint IT/COM societies paper award, the 2007 IEEE Information Theory Society Paper Award, the 2009 European Commission FP7, Network of Excellence in Wireless COMmunications (NEWCOM++) Best Paper Award, and the 2010 Thomson Reuters Award for International Excellence in Scientific Research. He is also the recipient of 1985 Alon Grant for Distinguished Young Scientists and the 2000 Technion Henry Taub Prize for Excellence in Research. He has served as Associate Editor for Shannon Theory of the IEEE TRANSACTIONS ON INFORMATION THEORY, and has also served on the Board of Governors of the Information Theory Society.