Cooperative Relaying With State Available Noncausally at the Relay

Abdellatif Zaidi, *Member, IEEE*, Shiva Prasad Kotagiri, J. Nicholas Laneman, *Senior Member, IEEE*, and Luc Vandendorpe, *Fellow, IEEE*

***Abstract*—In this paper, we consider a three-terminal state-dependent relay channel (RC) with the channel state noncausally available at only the relay. Such a model may be useful for designing cooperative wireless networks with some terminals equipped with cognition capabilities, i.e., the relay in our setup. In the discrete memoryless (DM) case, we establish lower and upper bounds on channel capacity. The lower bound is obtained by a coding scheme at the relay that uses a combination of codeword splitting, Gel'fand–Pinsker binning, and decode-and-forward (DF) relaying. The upper bound improves upon that obtained by assuming that the channel state is available at the source, the relay, and the destination. For the Gaussian case, we also derive lower and upper bounds on the capacity. The lower bound is obtained by a coding scheme at the relay that uses a combination of codeword splitting, generalized dirty paper coding (DPC), and DF relaying; the upper bound is also better than that obtained by assuming that the channel state is available at the source, the relay, and the destination. In the case of degraded Gaussian channels, the lower bound meets with the upper bound for some special cases, and, so, the capacity is obtained for these cases. Furthermore, in the Gaussian case, we also extend the results to the case in which the relay operates in a half-duplex mode.**

***Index Terms*—Channel state information, cognitive radio, (generalized) dirty paper coding (DPC), relay channel (RC), user cooperation.**

I. Introduction

We consider a three-terminal state-dependent relay channel (RC) in which, as shown in Fig. 1, the source wants to communicate a message W to the destination through the state-dependent RC in n uses of the channel, with the help of the relay. The channel outputs Y_2 and Y_3 for the relay and the destination, respectively, are controlled by the channel input X_1, the relay input X_2, and the channel state S, through a given memoryless probability law $W_{Y_2,Y_3|X_1,X_2,S}$. The channel state S is generated according to a given memoryless probability law Q_S. It is assumed that the channel state is known, noncausally, to only the relay. The destination estimates the message sent by the source from the received channel output. In this paper, we study the capacity of this communication system. We refer to the model under investigation as state-dependent RC with informed relay.

A. Background

Channels with random parameters or states have received considerable attention due to a wide range of possible applications. Shannon initiated the study of single-user models with state available *causally* at the encoder [1]. For the single-user discrete memoryless (DM) state-dependent models, Gel'fand and Pinsker derive the capacity for the setup in which the channel state is available *noncausally* at the encoder [2]. In this case, a random coding scheme based on binning, known as *Gel'fand–Pinsker coding*, achieves the capacity [2]. Costa considers an additive Gaussian channel with additive Gaussian state known at the encoder and shows that Gel'fand–Pinsker coding with a specific auxiliary random variable, widely known as dirty paper coding (DPC), achieves the trivial upper bound obtained by assuming the channel state available also at the decoder [3]. Interestingly, DPC eliminates the effect of the additive channel state on the capacity, as if there were no channel state present in the model or the channel state were known to the decoder as well. It is worth noting that since DPC achieves the trivial upper bound for this model there is no need to derive tighter upper bounds in this case. In [4], models with channel state available noncausally at the encoder are studied from the perspective of memories with defects. Practical coding realizations using concepts of lattices for the models with noncausal encoder state information are studied, e.g., in [5] and [6]. For a review on the subject of state-dependent channels and related work, the reader may refer to [7].

A growing body of work studies multiuser state-dependent models with noncausal encoder state information [8]–[21]. In the multiuser models, the channel state can be known to all, only some, or none of the users in the communication system. In the case of state-dependent DM models, the multiple access channel (MAC) with partial channel state at all the encoders and full channel state at the decoder is considered in [11], and the broadcast channel (BC) with state available at the encoder but not at the decoders is considered in [12], [22].

In the Gaussian case, the MAC with all encoders being informed, the BC with informed encoder, the physically degraded

Manuscript received December 01, 2008; revised October 02, 2009. Current version published April 21, 2010. The work of A. Zaidi and L. Vandendorpe was supported in part by the EU framework program COOPCOM and the network of excellence NEWCOM++. The work of S. P. Kotagiri and J. N. Laneman was supported in part by the National Science Foundation (NSF) under Grants CNS06-26595 and CCF05-46618.

A. Zaidi and L. Vandendorpe are with the École Polytechnique de Louvain, Université Catholique de Louvain, Louvain-la-Neuve 1348, Belgium (e-mail: abdellatif.zaidi@uclouvain.be; luc.vandendorpe@uclouvain.be).

S. P. Kotagiri was with the Department of Electrical Engineering, University of Notre Dame, Notre Dame, IN 46556 USA. He is now with the SERDES Technology Group, Xilinx Inc., San Jose, CA 95124 USA (e-mail: skotagir@gmail.com).

J. N. Laneman is with the Department of Electrical Engineering, University of Notre Dame, Notre Dame, IN 46556 USA (e-mail: jnl@nd.edu).

Communicated by G. Kramer, Associate Editor for Shannon Theory.

Color versions of Figures 4–7 in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIT.2010.2043782

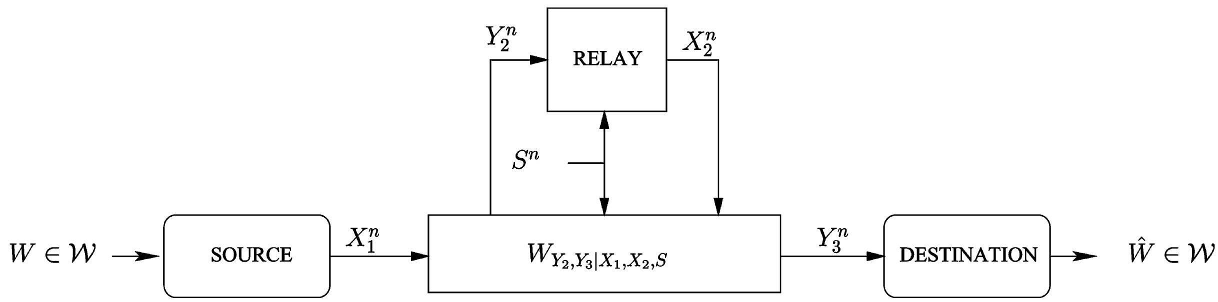


Fig. 1. RC with state information S^n available noncausally at only the relay.

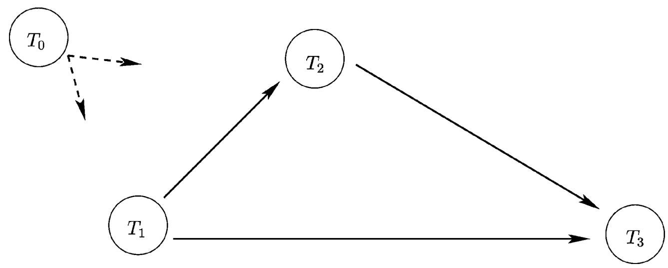


Fig. 2. Example wireless network with cognition capabilities. If the relay T_2 is cognizant of the competing source T_0, it can help the source T_1 cancel the effect of the interference from T_0.

RC with informed source and informed relay, and the physically degraded relay broadcast channel (RBC) with informed source and informed relay are studied in [8], [9], and [18]. In all these cases, it is shown that some variants of DPC achieve the respective capacity or capacity region. Also, since for all these models DPC achieves the trivial upper or outer bound obtained by assuming that the channel state is also available at the decoders, it is not necessary to derive nontrivial upper or outer bounds, i.e., bounds that are tighter than the cut-set bound. For all these models, the key assumption that makes the problem relatively easy is the availability of the channel state at *all* the encoders in the communication model, which allows these encoders to remove the effect of the channel state on their respective communication using variants of DPC. It is interesting to study state-dependent multiuser models in which *some*, but not all, encoders are informed of the channel state, because the uninformed encoders cannot apply DPC.

The state-dependent MAC with some, but not all, encoders informed of the channel state is considered in [10], [13]–[16], and [21] and the state-dependent RC with informed source is considered in [18] and [19]. For the Gaussian cases of these models, the informed encoder applies a slightly generalized DPC (GDPC) in which the channel input and the channel state are negatively correlated. It is interesting to note that in these models the uninformed encoders benefit from the GDPC applied by the informed encoders because the negative correlation can be viewed as partial state cancellation. The capacity region of the DM state-dependent MAC with one informed encoder is characterized in [14] and [16] for the case in which the messages sets are degraded and the informed encoder knows the message of the uninformed encoder. In [16], the authors also study the Gaussian case and they characterize the capacity region by deriving a nontrivial outer bound that is strictly tighter than the cut-set outer bound.

For the study of communication models in which only some of the involved encoders are informed about the channel state, it is important to establish nontrivial upper or outer bounds. These bounds help characterize the rate loss due to not knowing the state at the uninformed encoders, and help assess the effectiveness of the coding schemes that are employed for the achievability results. In this paper, we study a state-dependent RC with the channel state known to only the relay. This model is conceptually different from the model considered previously in [18] and [19] in which the channel state is noncausally known to only the source.

B. Motivation

Channels whose probabilistic input–output relationship depends on random parameters, or channel states, can model a large variety of scenarios. The assumption of noncausal channel state can hold *naturally* or *approximately*. Examples where the assumption of noncausal state holds naturally include information embedding [23]–[28], certain storage applications such as computer memories with defective cells [29], and certain broadcast scenarios such as multiple-input–multiple-output (MIMO) BCs [30]–[32] where DPC is a central ingredient in achieving the capacity region [33]. Examples where the assumption of noncausal state holds approximately include dispersive (ISI) channels [5], block fading in wireless environments [34], network [35], and cooperative networks [36].

Yet, another example application is cooperation in the realm of cognition. Driven by the growing demand for frequency spectrum, cognitive radios, usually defined as smart radio devices that are capable of acquiring some knowledge about the channel state, are introduced into communication systems in order to help noncognitive radios in terms of spectral efficiency [37]. In a wireless interference network in which some terminals compete and some others cooperate, equipping some specific terminals with cognition capabilities that allow them to learn the interference to high accuracy would help other noncognitive terminals. These cognitive radios can exploit the knowledge of the interference or channel state to remove its effect on the transmission of their own messages and also that of the messages of the noncognitive terminals as well. The study of fundamental performance limits of models with only a subset of the encoders being informed is relevant for a better understanding of communication systems that involve cognitive radios. For example, to increase system spectral efficiency, collaboration is investigated in the realm of cognition in [38]–[40]. Also, the problem of collaborative signal transmission in the presence of some cognizant terminals is investigated for a MAC scenario in [13] and [16] and for an interference channel scenario in [41]–[44]. The setup we consider in this paper also models the building block for collaborative wireless networks in which only the relays, but neither sources nor destinations, are cognizant of the channel state. An example of such a scenario is shown in Fig. 2.

C. Main Contributions

For the DM case, we derive lower and upper bounds on the capacity of the general state-dependent RC with informed relay. The lower bound is obtained by a coding scheme at the relay that uses a combination of codeword splitting, Gel'fand–Pinsker coding, and decode-and-forward (DF) relaying. For this model, designing a codebook at the relay is challenging since such a codebook should allow the source to generate codewords that are correlated with the channel input of the relay which exploits the available channel state. In this work, this is accomplished by codeword splitting at the relay. With codeword splitting, the channel input of the relay is generated from two codewords: the first of which is a function of the cooperative information (i.e., the information that is sent cooperatively by the source and the relay using a *joint codebook*) and the channel state, and the second of which is a function of only the cooperative information. Since the source knows the cooperative information, it can generate its channel input in a way such that it is correlated with the latter codeword at the relay, which is a function of only the cooperative information.

Our upper bound on the capacity is tighter than that obtained by assuming that the channel state is also available at the source and the destination. This upper bound is nontrivial and relates to the bounding technique developed in the context of MACs with asymmetric channel state in [16, Th. 2]; however, we note that the present upper bound is proved using techniques that are different from those in [16]. On a related note, we mention that at a high level there is a connection between the multiple access transmission part in the RC with informed relay in this work and the models in [13] and [16]. However, there are also numerous conceptual differences that will be discussed whenever relevant. In particular, in contrast to [13] and [16], here the uninformed encoder (the source) knows the message of the informed encoder (the relay). From this angle, the model in this paper connects more with the state-dependent MAC studied in [21].

Furthermore, we specialize the results to the case in which the channel is degraded. Also, we extend the lower bound for the DF relaying scheme to the case in which the source implements rate splitting and the relay DFs only one part of the source message.

We apply the concepts developed in the DM case to the Gaussian case in which both the noise and the state are additive Gaussian random variables. In our analysis for the Gaussian RC, we first allow the relay to operate in a *full-duplex* mode in which it can transmit and receive simultaneously, and then we constrain it to operate in a *half-duplex* mode in which it can either only transmit or only receive.

In the case of full-duplex transmission, we derive lower and upper bounds on the capacity of the Gaussian RC with informed relay. We obtain two lower bounds by using the concepts of codeword splitting, GDPC [10], [45], and DF relaying. Through codeword splitting, the channel input of the source is *partially* coherent with the channel input of the relay. The first lower bound uses full DF at the relay and the second further enlarges it by allowing rate splitting at the source.

We also point out the loss incurred by the availability of the channel state at only the relay in the upper bound. We show that the lower bound obtained with rate splitting at the source is in general close to the upper bound for general Gaussian channels. In the case of the degraded Gaussian channel, the two lower bounds meet and they meet with the upper bound for some special cases.

In the case of half-duplex transmission, we derive lower and upper bounds for the capacity of the Gaussian RC with informed relay. In this case, we focus on relaying protocols in which the relay either fully or partially decodes the source message, re-encodes and sends it to the destination.

D. Outline and Notation

An outline of the remainder of this paper is as follows. Section II describes in more detail the communication model that we consider in this work. Section III provides lower and upper bounds on the capacity of the general DM RC with informed relay. Section IV provides lower and upper bounds on the capacity of the Gaussian RC with informed relay, and also contains some numerical results and discussions. Finally, Section V concludes the paper.

We use the following notations throughout the paper. Upper case letters are used to denote random variables, e.g., X; lower case letters are used to denote realizations of random variables, e.g., x; and calligraphic letters designate alphabets, i.e., $\mathcal{X}$. The probability distribution of a random variable X is denoted by $P_X(x)$. Sometimes, for convenience, we write it as P_X. We use the notation $\mathbb{E}_X[\cdot]$ to denote the expectation of random variable X. A probability distribution of a random variable Y given X is denoted by $P_{Y|X}$. The set of probability distributions defined on an alphabet $\mathcal{X}$ is denoted by $\mathcal{P}(\mathcal{X})$. The cardinality of a set $\mathcal{X}$ is denoted by $|\mathcal{X}|$. The shorthand notation X_i^j indicates a sequence of random variables $(X_i, X_{i+1}, \ldots, X_j)$ and x_i^j denotes a particular realization of a random sequence X_i^j. For convenience, the length n vector x^n will occasionally be denoted in boldface notation $\mathbf{x}$. Given random variables X, Y, Z, we denote the entropy of X by $H(X)$, the mutual information between X and Y by $I(X;Y)$, and the conditional mutual information between X and Y, conditioned on Z, by $I(X;Y|Z)$ [46]. The Gaussian distribution with mean μ and variance σ^2 is denoted by $\mathcal{N}(\mu, \sigma^2)$. Finally, throughout the paper, logarithms are taken to base 2, and the complement to unity of a scalar $u \in [0,1]$ is denoted by $\bar{u}$, i.e., $\bar{u} = 1 - u$.

II. System Model and Definitions

In this section, we formally present our communication model and the definitions related to it. As shown in Fig. 1, we consider a state-dependent RC denoted by $W_{Y_2,Y_3|X_1,X_2,S}$ whose outputs $Y_2 \in \mathcal{Y}_2$ and $Y_3 \in \mathcal{Y}_3$ for the relay and the destination, respectively, are controlled by the channel inputs $X_1 \in \mathcal{X}_1$ from the source and $X_2 \in \mathcal{X}_2$ from the relay, along with a channel state $S \in \mathcal{S}$. It is assumed that the channel state

S_i at instant i is independently drawn from a given distribution Q_S and the channel state S^n is noncausally known at the relay. Also, each transmitted input block x_1^n from the source and each transmitted input block x_2^n from the relay are subject to additive normalized input constraints

$$\varphi_k^n\left(x_k^n\right) \triangleq \frac{1}{n}\sum_{i=1}^{n}\varphi_k(x_{k,i}) \leq \Gamma_k, \qquad k=1,2 \tag{1}$$

where $\varphi_1 : \mathcal{X}_1 \rightarrow \mathbb{R}^+$ and $\varphi_2 : \mathcal{X}_2 \rightarrow \mathbb{R}^+$ are single-letter input cost functions for the source and the relay, respectively.

The source wants to transmit a message W to the destination with the help of the relay, in n channel uses. The message W is assumed to be uniformly distributed over the set $\mathcal{W} = \{1,\ldots,M\}$. The information rate R is defined as $\log M/n$ bits per transmission.

An $(n, M, \Gamma_1, \Gamma_2)$ code for the state-dependent RC with informed relay consists of an encoding function at the source

$$\phi_1^n : \{1,\ldots,M\} \rightarrow \mathcal{X}_1^n$$

a sequence of encoding functions at the relay

$$\phi_{2,i} : \mathcal{Y}_{2,1}^{i-1} \times \mathcal{S}^n \rightarrow \mathcal{X}_2$$

for $i=1,2,\ldots,n$, and a decoding function at the destination

$$\psi^n : \mathcal{Y}_3^n \rightarrow \{1,\ldots,M\}$$

such that

$$\frac{1}{n}\sum_{i=1}^{n}\varphi_1\left(\phi_1^n(w)_i\right) \leq \Gamma_1$$

and

$$\frac{1}{n}\sum_{i=1}^{n}\varphi_2\left(\phi_{2,i}\left(y_2^{i-1}, s^n\right)\right) \leq \Gamma_2$$

for $w \in \{1,\ldots,M\}$.

From an $(n, M, \Gamma_1, \Gamma_2)$ code, the sequences X_1^n and X_2^n from the source and the relay, respectively, are transmitted across a state-dependent RC modeled as a memoryless conditional probability distribution $W_{Y_2,Y_3|X_1,X_2,S}$. The joint probability mass function on $\mathcal{W}\times\mathcal{S}^n\times\mathcal{X}_1^n\times\mathcal{X}_2^n\times\mathcal{Y}_2^n\times\mathcal{Y}_3^n$ is given by

$$\begin{aligned} P&\left(w, s^n, x_1^n, x_2^n, y_2^n, y_3^n\right) \\ &= P(w)\prod_{i=1}^{n} Q_S(s_i)P(x_{1,i}|w)P\left(x_{2,i}|s^n, y_2^{i-1}\right) \\ &\quad \cdot W_{Y_2,Y_3|X_1,X_2,S}(y_{2,i}, y_{3,i}|x_{1,i}, x_{2,i}, s_i). \end{aligned} \tag{2}$$

The channel is said to be *physically degraded* if the conditional distribution $W_{Y_2,Y_3|X_1,X_2,S}$ factorizes as

$$W_{Y_2,Y_3|X_1,X_2,S} = W_{Y_2|X_1,X_2,S}W_{Y_3|Y_2,X_2,S}. \tag{3}$$

The destination estimates the message sent by the source from the channel output Y_3^n. The average probability of error is defined as $P_e^n = \mathbb{E}_S[\Pr(\psi^n\left(Y_3^n\right) \neq W | S^n = s^n)]$.

An $(\epsilon, n, R, \Gamma_1, \Gamma_2)$ code for the state-dependent RC with informed relay is an $\left(n, \left\lceil 2^{nR}\right\rceil, \Gamma_1, \Gamma_2\right)$ code $(\phi_1^n, \phi_2^n, \psi^n)$ having average probability of error P_e^n not exceeding ϵ.

Given a pair $\mathbf{\Gamma} = (\Gamma_1, \Gamma_2)$, a rate R is said to be $\mathbf{\Gamma}$-achievable if there exists a sequence of $(\epsilon_n, n, R, \Gamma_1, \Gamma_2)$-codes with $\lim_{n\rightarrow\infty}\epsilon_n = 0$. The capacity $C(\mathbf{\Gamma})$ of the state-dependent RC with informed relay is the supremum of the set of $\mathbf{\Gamma}$-achievable rates.

III. The DM RC With Informed Relay

In this section, we assume that all the alphabets in the model, $\mathcal{S}$, $\mathcal{X}_1$, $\mathcal{X}_2$, $\mathcal{Y}_2$ and $\mathcal{Y}_3$, are discrete and finite.

A. Lower Bound on Capacity

The following theorem provides a lower bound on the capacity of the state-dependent DM RC with informed relay.

Theorem 1: Let $\mathbf{\Gamma} = (\Gamma_1, \Gamma_2)$ be given. The capacity $C(\mathbf{\Gamma})$ of the state-dependent DM RC with informed relay satisfies $C(\mathbf{\Gamma}) \geq R^{\text{lo}}(\mathbf{\Gamma})$, where

$$\begin{aligned} R^{\text{lo}}(\mathbf{\Gamma}) = \max\min\Big\{&I(X_1;Y_2|S,U_1,X_2), \\ &I(X_1,U_1,U_2;Y_3) - I(U_2;S|U_1)\Big\} \end{aligned} \tag{4}$$

with the maximization over all probability distributions of the form

$$\begin{aligned} &P_{S,U_1,U_2,X_1,X_2,Y_2,Y_3} \\ &\quad = Q_S P_{U_1} P_{X_1|U_1} P_{U_2|U_1,S} P_{X_2|U_1,U_2,S} W_{Y_2,Y_3|X_1,X_2,S} \end{aligned} \tag{5}$$

and satisfying $\mathbb{E}[\varphi_i(X_i)] \leq \Gamma_i, i=1,2$, and $U_1 \in \mathcal{U}_1, U_2 \in \mathcal{U}_2$ are auxiliary random variables with

$$|\mathcal{U}_1| \leq |\mathcal{S}||\mathcal{X}_1||\mathcal{X}_2| + 1 \tag{6a}$$
$$|\mathcal{U}_2| \leq \left(|\mathcal{S}||\mathcal{X}_1||\mathcal{X}_2| + 1\right)|\mathcal{S}||\mathcal{X}_1||\mathcal{X}_2| \tag{6b}$$

respectively.

Remark 1: The lower bound (4) is based upon a technique at the relay we call *codeword splitting*, combining DF relaying [47, Th. 1] with Gel'fand–Pinsker coding [2]. In conventional DF strategies, the source knows the relay input, allowing the source and relay to utilize a joint codebook to transmit cooperative information. However, in our model there is a tension between the utility of a joint codebook for relaying and the utility of the relay's making use of the channel state, which is unknown to the source. To resolve this tension, we generate two codebooks at the relay. In one codebook, the codewords are generated using a random variable U_1 that is independent of the channel state S. The relay chooses the appropriate codeword from this codebook using only the cooperative information. In the other codebook, the codewords are generated using a random variable U_2 that is correlated with the channel state

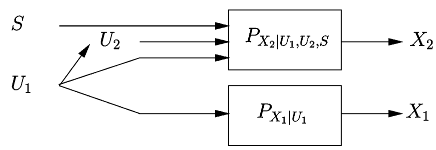


Fig. 3. Dependence diagram of the random variables for the lower bound in Theorem 1.

S and the variable U_1 through $P_{U_2|U_1,S}$. The relay chooses the appropriate codeword from this codebook using both the cooperative information and the channel state, in order to combat the effect of the channel state on the communication. Finally, the relay generates the channel input X_2^n from (U_1^n, U_2^n) using the conditional probability law $P_{X_2|U_1,U_2,S}$. The source knows U_1^n as this is a function of only the cooperative information, and, given U_1^n, it generates the random codeword X_1^n according to the conditional probability law $P_{X_1|U_1}$. Thus, the channel inputs of the source and the relay are correlated through U_1^n. A dependence diagram of the random variables that are involved in the coding scheme is shown in Fig. 3.

Remark 2: The term $[I(X_1, U_1, U_2; Y_3) - I(U_2; S|U_1)]$ in (4) can be interpreted as an achievable sum rate over a state-dependent MAC with one informed encoder and degraded messages, i.e., one common and one individual message. In our model, the informed encoder sends only the common message, i.e., the cooperative information of DF relaying, and the uninformed encoder sends both the common and individual messages. By contrast, [13] and [16] derive the capacity region for the reverse situation in which the informed encoder sends both the common and individual messages, and the uninformed encoder sends only the common message. This swapping of roles makes coding at the relay more involved than in [13] and [16] for the state-dependent MAC and the [18] and [19] for the related state-dependent RC with informed source. As we mentioned earlier, a MAC model that has closer connection with the model in this paper is investigated in [21]. This model is obtained by swapping the roles of the encoders in [13] and [16].

Outline of Proof of Theorem 1: First we generate a random codebook that we use to obtain the lower bound in Theorem 1. Next, we outline the encoding and decoding procedures at the source and the relay. The coding scheme is based on a combination of codeword splitting, regular-encoding backward decoding for DF [48], and a variation of Gel'fand–Pinsker binning. A formal proof with complete error analysis is given in Appendix A. In the formal proof we also show that the input constraints are satisfied.

Codebook generation: Fix a measure $P_{S,U_1,U_2,X_1,X_2,Y_2,Y_3}$ satisfying (5) and $\mathbb{E}[\varphi_i(X_i)] \le \Gamma_i$, $i = 1, 2$. Fix $\epsilon > 0$ and denote

$$J = 2^{n(I(U_2;S|U_1)+2\epsilon)} \tag{7a}$$
$$M = 2^{n(R-4\epsilon)}. \tag{7b}$$

1) We generate M independent and identically distributed (i.i.d.) codewords $\{\mathbf{u}_1(w')\}$ indexed by $w' = 1, \ldots, M$, each with i.i.d. components drawn according to P_{U_1}.
2) For each codeword $\mathbf{u}_1(w')$, we generate M i.i.d. codewords $\{\mathbf{x}_1(w', w)\}$ at the source indexed by $w = 1, \ldots, M$, and J auxiliary codewords $\{\mathbf{u}_2(w', j)\}$ at the relay indexed by $j = 1, \ldots, J$. The codewords $\mathbf{x}_1(w', w)$ and $\mathbf{u}_2(w', j)$ are with i.i.d. components given $\mathbf{u}_1(w')$ drawn according to $P_{X_1|U_1}$ and $P_{U_2|U_1}$, respectively.

Outline of the coding scheme: We outline the coding scheme in the following. The message W to be sent from the source node is divided into B blocks $w_1, w_2, \ldots, w_B$ of nR bits each. For convenience, we let $w_{B+1} = 1$. The transmission is performed in $B+1$ blocks. We denote by $\mathbf{s}[i]$ the channel state in block i, $i = 1, \ldots, B+1$.

Continuing with the strategy, in the first block, the source transmits $\mathbf{x}_1(1, w_1)$. The relay searches for the smallest $j \in \{1, \ldots, J\}$ such that $\mathbf{u}_1(1)$, $\mathbf{u}_2(1, j)$, and $\mathbf{s}[1]$ are jointly typical (the properties of strongly typical sequences guarantee that there exists one such j). Denote this j by $j^\star = j(\mathbf{s}[1], 1)$. Then, the relay transmits a vector $\mathbf{x}_2(1)$ with i.i.d. components given $(\mathbf{u}_1(1), \mathbf{u}_2(1, j^\star), \mathbf{s}[1])$ drawn according to the marginal $P_{X_2|U_1,U_2,S}$ induced by the distribution (5).

The decoder at the relay uses joint typicality. It declares that message $\hat{w}_1$ is sent if there is a unique $\hat{w}_1$ such that $\mathbf{x}_1(1, \hat{w}_1)$ is jointly typical with $(\mathbf{y}_2[1], \mathbf{s}[1])$ given $\mathbf{u}_1(1)$, $\mathbf{u}_2(1, j^\star)$ and $\mathbf{x}_2(1)$, where $\mathbf{y}_2[1]$ denotes the information received at the relay in block 1. One can show that the relay can decode reliably as long as n is large and

$$R < I(X_1; Y_2|S, U_1, X_2). \tag{8}$$

So, suppose the relay correctly obtains w_1. In the second block, the source transmits $\mathbf{x}_1(w_1, w_2)$ and the relay transmits a vector $\mathbf{x}_2(w_1)$ with i.i.d. components given $\mathbf{u}_1(w_1)$, $\mathbf{u}_2(w_1, j(\mathbf{s}[2], w_1))$, $\mathbf{s}[2]$ drawn according to the marginal $P_{X_2|U_1,U_2,S}$; the sequence $\mathbf{u}_2(w_1, j(\mathbf{s}[2], w_1))$ is chosen such that $j(\mathbf{s}[2], w_1)$ is the smallest $j \in \{1, \ldots, J\}$ satisfying $\mathbf{u}_1(w_1)$, $\mathbf{u}_2(w_1, j)$ and $\mathbf{s}[2]$ are jointly typical. Upon observation of $\mathbf{y}_2[2]$, the decoder at the relay declares that $\hat{w}_2$ is sent if there is a unique $\hat{w}_2$ such that $\mathbf{x}_1(w_1, \hat{w}_2)$ is jointly typical with $(\mathbf{y}_2[2], \mathbf{s}[2])$ given $\mathbf{u}_1(w_1)$, $\mathbf{u}_2(w_1, j(\mathbf{s}[2], w_1))$, and $\mathbf{x}_2(w_1)$. Again, it can decode reliably as long as n is large and (8) is true. At the relay, one continues in this way until block $B+1$.

Consider now the destination, and let $\mathbf{y}_3[i]$ be the received information at the destination in block i. Suppose these information are collected until the last block of transmission is completed. The destination can then perform Willem's *backward decoding* [48], by first decoding w_B from $\mathbf{y}_3[B+1]$. Note that $\mathbf{y}_3[B+1]$ depends on $\mathbf{x}_1(w_B, 1)$, $\mathbf{u}_1(w_B)$, and $\mathbf{u}_2(w_B, j(\mathbf{s}[B+1], w_B))$, which in turn depends only on w_B. The decoder at the destination uses joint typicality. It declares that $\hat{w}_B$ is sent if there is a unique $\hat{w}_B$ such that $\mathbf{x}_1(\hat{w}_B, 1), \mathbf{u}_1(\hat{w}_B), \mathbf{u}_2(\hat{w}_B, j_B)$, $\mathbf{y}_3[B+1]$ are jointly typical, for some index $j_B \in \{1, \ldots, J\}$. One can show that the destination can decode reliably as long as n is large and

$$R < I(X_1, U_1, U_2; Y_3) - I(U_2; S|U_1). \tag{9}$$

So, suppose the destination correctly obtains w_B. Next, the destination decodes w_{B-1} from $\mathbf{y}_3[B]$, which depends on

$\mathbf{x}_1(w_{B-1}, w_B)$, $\mathbf{u}_1(w_{B-1})$, and $\mathbf{u}_2(w_{B-1}, j(\mathbf{s}[B], w_{B-1}))$. Since the destination knows w_B, it can again decode reliably as long as n is large and (9) is true. At the destination, one continues in this fashion until all message blocks have been decoded. The average rate over the $B+1$ blocks is $RB/(B+1)$ bits per use, and by making B large one can get the rate as close to R as desired.

Remark 3: In the case of classic RC without state, one can consider three different DF strategies: irregular encoding successive decoding [47], regular encoding sliding-window decoding [49], and regular encoding backward decoding [48]. It is well known that these three strategies achieve the same rate in this case [50]. In the state-dependent case with informed relay, one can show that backward decoding achieves rates higher than those of sliding-window decoding. More precisely, sliding window decoding at the destination at the end of block i is as follows (we use the notation in the proof of Theorem 1). The destination knows w_{i-2} and also the correct index $j(\mathbf{s}[i-1], w_{i-2})$, and decodes w_{i-1} based on the information received in the two adjacent blocks $i-1$ and i. It declares that the message $\hat{w}_{i-1}$ is sent if there is a unique pair $(\hat{w}_{i-1}, \hat{j}_{i-1})$ such that the vectors $\mathbf{u}_1(w_{i-2})$, $\mathbf{u}_2(w_{i-2}, j(\mathbf{s}[i-1], w_{i-2}))$, $\mathbf{x}_1(w_{i-2}, \hat{w}_{i-1})$, $\mathbf{y}_3[i-1]$ are jointly typical, *and* the vectors $\mathbf{u}_1(\hat{w}_{i-1})$, $\mathbf{u}_2(\hat{w}_{i-1}, \hat{j}_{i-1})$, $\mathbf{y}_3[i]$ are jointly typical. Thus, the destination obtains the message w_{i-1} if

$$R < I(X_1, U_1, U_2; Y_3) - I(U_2; S|U_1) \tag{10}$$
$$I(U_2; Y_3|U_1) - I(U_2; S|U_1) > 0. \tag{11}$$

Hence, with window decoding also, the achievable rate is obtained by maximizing the right-hand side of (4). However, unlike the above backward decoding scheme, the maximization is over a set of distributions of the form (5) that satisfy the constraint (11). Because of the additional constraint, this set is smaller than the one used in Theorem 1. Informally speaking, the additional constraint (11) guarantees that, in the decoding of the vectors $\mathbf{u}_1$ and $\mathbf{u}_2$, the destination can actually decode the vector $\mathbf{u}_2$ *fully*, i.e., it can determine not only the bin index (i.e., the message w_{i-1}) but also the correct sequence in the bin (i.e., the index $j(\mathbf{s}[i], w_{i-1})$). □

The achievable rate in (4) requires the relay to *fully* decode the message sent by the source, and this can be rather a severe constraint. We can generalize Theorem 1 by allowing the relay to decode the source message *only partially* [51]. This can be done by implementing rate splitting at the source [52] and introducing a new random variable U that represents the information decoded by the relay. The following corollary gives the resulting rate.

Corollary 1: The capacity $C(\mathbf{\Gamma})$ of the state-dependent DM RC with informed relay satisfies $C(\mathbf{\Gamma}) \geq R'^{\text{lo}}(\mathbf{\Gamma})$, where

$$\begin{aligned} R'^{\text{lo}}(\mathbf{\Gamma}) &= \max\min\Big\{I(U; Y_2|S, U_1, X_2) + I(X_1; Y_3|U, U_1, U_2) \\ &\quad + \min\{0, I(U_2; Y_3|U, U_1) - I(U_2; S|U_1)\}, \\ &\quad I(X_1, U_1, U_2; Y_3) - I(U_2; S|U_1)\Big\} \end{aligned} \tag{12}$$

with the maximization over all probability distributions of the form

$$\begin{aligned} P_{S,U_1,U_2,U,X_1,X_2,Y_2,Y_3} &= Q_S P_{U_1} P_{U|U_1} P_{X_1|U_1,U} P_{U_2|U_1,S} \\ &\quad \times P_{X_2|U_1,U_2,S} W_{Y_2,Y_3|X_1,X_2,S} \end{aligned} \tag{13}$$

and satisfying $\mathbb{E}[\varphi_i(X_i)] \leq \Gamma_i$, $i = 1, 2$, and $U_1 \in \mathcal{U}_1$, $U_2 \in \mathcal{U}_2$, $U \in \mathcal{U}$ are auxiliary random variables with

$$|\mathcal{U}_1| \leq |\mathcal{S}||\mathcal{X}_1||\mathcal{X}_2| + 2 \tag{14a}$$
$$|\mathcal{U}_2| \leq (|\mathcal{S}||\mathcal{X}_1||\mathcal{X}_2| + 2)\,|\mathcal{S}||\mathcal{X}_1||\mathcal{X}_2| + 2 \tag{14b}$$
$$|\mathcal{U}| \leq (|\mathcal{S}||\mathcal{X}_1||\mathcal{X}_2| + 2)\,|\mathcal{S}||\mathcal{X}_1||\mathcal{X}_2| + 2 \tag{14c}$$

respectively.

The proof of Corollary 1 is similar to that of Theorem 1 and, hence, only an outline of it is given in Appendix B. For instance, the particular choice $U = X_1$ in Corollary 1 gives the lower bound in Theorem 1.

An informal interpretation of the rate (12) for the case in which $[I(U_2; Y_3|U, U_1) - I(U_2; S|U_1)] > 0$ is as follows. Since $I(U; Y_3|U_1, U_2, X_1) = 0$ for the distribution considered in (13), the second term of the minimization in (12) can be written as

$$I(U, U_1, U_2; Y_3) - I(U_2; S|U_1) + I(X_1; Y_3|U, U_1, U_2).$$

The rate (12) can then be interpreted as the rate achievable if the message W transmitted by the source is split into two independent parts, one of which is transmitted through the relay, say at rate R_r, and the other is transmitted directly to the destination without the help of the relay, say at rate R_d. The total rate is $R = R_r + R_d$. In (12), the auxiliary variable U stands for the information decoded by the relay and plays the role of X_1 in Theorem 1. Thus, it follows from (4) that the message transmitted through the relay can be decoded correctly at the destination if rate R_r satisfies

$$\begin{aligned} R_r < \min\{&I(U; Y_2|S, U_1, X_2), \\ &I(U, U_1, U_2; Y_3) - I(U_2; S|U_1)\}. \end{aligned} \tag{15}$$

It can also be easily argued (see Appendix B) that the additional information, which is sent on top of the information transmitted through the relay, can be decoded correctly at the destination if rate R_d satisfies

$$R_d < I(X_1; Y_3|U, U_1, U_2). \tag{16}$$

This shows that message W can be sent at the rate (12).

Remark 4: The relay can employ other relaying schemes to assist the source, such as estimate-and-forward [47], amplify-and-forward [53]–[55], or combinations of these schemes. However, none of these schemes achieves capacity even if the channel is state-independent. Hence, though some of these schemes may perform well in terms of achievable rates for some particular channels, we do not focus on these schemes in this paper.

B. Upper Bound on Capacity

The following theorem provides an upper bound on the capacity of the state-dependent DM RC with informed relay.

Theorem 2: Let $\mathbf{\Gamma} = (\Gamma_1, \Gamma_2)$ be given. The capacity $C(\mathbf{\Gamma})$ of the state-dependent DM RC with informed relay satisfies $C(\mathbf{\Gamma}) \leq R^{\mathrm{up}}(\mathbf{\Gamma})$, where

$$R^{\mathrm{up}}(\mathbf{\Gamma}) = \max\min\Big\{I(X_1;Y_2,Y_3|S,X_2), \quad I(X_1,X_2;Y_3|S) - I(X_1;S|Y_3)\Big\} \tag{17}$$

with the maximization over all probability distributions of the form

$$P_{S,X_1,X_2,Y_2,Y_3} = Q_S P_{X_1} P_{X_2|X_1,S} W_{Y_2,Y_3|X_1,X_2,S} \tag{18}$$

and satisfying $\mathbb{E}[\varphi_i(X_i)] \leq \Gamma_i$, $i = 1, 2$.

The proof of Theorem 2 appears in Appendix C.

In the second term of the minimum in (17), $I(X_1;S|Y_3)$ can be interpreted as the rate penalty caused by the source's not knowing the channel state. This rate loss makes the above upper bound tighter than the trivial upper bound obtained by assuming that the channel state is also available at the source and the destination, i.e., the cut set upper bound

$$R^{\mathrm{up}}_{\mathrm{triv}}(\mathbf{\Gamma}) = \max\min\left\{I(X_1;Y_2,Y_3|S,X_2), I(X_1,X_2;Y_3|S)\right\} \tag{19}$$

with the maximization over all distributions of the form

$$P_{S,X_1,X_2,Y_2,Y_3} = Q_S P_{X_1|S} P_{X_2|X_1,S} W_{Y_2,Y_3|X_1,X_2,S} \tag{20}$$

and satisfying $\mathbb{E}[\varphi_i(X_i)] \leq \Gamma_i$, $i = 1, 2$.

If the channel is physically degraded, the upper bound in Theorem 2 reduces to the one in the following corollary.

Corollary 2: The capacity of the state-dependent physically degraded RC with informed relay satisfies $C_{\mathrm{D}}(\mathbf{\Gamma}) \leq R^{\mathrm{up}}_{\mathrm{D}}(\mathbf{\Gamma})$, where

$$R^{\mathrm{up}}_{\mathrm{D}}(\mathbf{\Gamma}) = \max\min\Big\{I(X_1;Y_2|S,X_2), \quad I(X_1,X_2;Y_3|S) - I(X_1;S|Y_3)\Big\} \tag{21}$$

with the maximization over all probability distributions of the form

$$P_{S,X_1,X_2,Y_2,Y_3} = Q_S P_{X_1} P_{X_2|X_1,S} W_{Y_2|X_1,X_2,S} W_{Y_3|Y_2,X_2,S} \tag{22}$$

and satisfying $\mathbb{E}[\varphi_i(X_i)] \leq \Gamma_i$, $i = 1, 2$.

Similar to the general case in Theorem 2, the upper bound in Corollary 2 is tighter than the trivial upper bound in (19) for the degraded case.

IV. The Gaussian RC With Informed Relay

In this section, we consider a state-dependent Gaussian RC in which both the channel state and the noise are additive and Gaussian. We also assume that the additive channel state is noncausally known to only the relay. First, we consider full-duplex transmission at the relay, i.e., the relay transmits and receives at the same time, and we derive lower and upper bounds on channel capacity for this case. Then we extend these results to the half-duplex mode in which the relay is constrained to operate in a time-division (TD) manner.

A. Full-Duplex Channel Model

For the full-duplex state-dependent Gaussian RC, the channel outputs $Y_{2,i}$ and $Y_{3,i}$ at instant i for the relay and the destination, respectively, are related to the channel input $X_{1,i}$ from the source and $X_{2,i}$ from the relay, and the channel state S_i by

$$Y_{2,i} = X_{1,i} + S_i + Z_{2,i} \tag{23a}$$

$$Y_{3,i} = X_{1,i} + X_{2,i} + S_i + Z_{3,i} \tag{23b}$$

where S_i is a zero mean Gaussian random variable with variance Q, $Z_{2,i}$ is zero mean Gaussian with variance N_2, and $Z_{3,i}$ is zero mean Gaussian with variance N_3. The random variables S_i, $Z_{2,i}$ and $Z_{3,i}$ at instant $i \in \{1, 2, \ldots, n\}$ are mutually independent, and are independent of $(S_j, Z_{2,j}, Z_{3,j})$ for $j \neq i$. The random variables $Z_{2,i}$ and $Z_{3,i}$ are also independent of the channel inputs (X_1^n, X_2^n).

For the full-duplex degraded additive Gaussian RC, the channel outputs $Y_{2,i}$ and $Y_{3,i}$ for the relay and the destination, respectively, are related to the channel inputs $X_{1,i}$ and $X_{2,i}$ and the state S_i by

$$Y_{2,i} = X_{1,i} + S_i + Z_{2,i} \tag{24a}$$

$$Y_{3,i} = X_{2,i} + Y_{2,i} + Z'_{3,i} \tag{24b}$$

where $\left(Z'_{3,1}, \ldots, Z'_{3,n}\right)$ is a sequence of i.i.d. zero mean Gaussian random variables with variance $N'_3 = N_3 - N_2$ which is independent of Z_2^n.

The channel inputs from the source and the relay should satisfy the following average power constraints:

$$\sum_{i=1}^{n} X_{1,i}^2 \leq nP_1, \quad \sum_{i=1}^{n} X_{2,i}^2 \leq nP_2. \tag{25}$$

As we indicated previously, we assume that the channel state S^n is noncausally known at only the relay. The definition of a code for this channel is the same as that given in Section II, with the additional constraint that the channel inputs should satisfy the power constraints (25).

B. Lower Bounds on Capacity

In this section, we derive lower bounds on the capacity of the state-dependent full-duplex Gaussian RC with informed relay. The results obtained in Section III for the DM case can be extended to memoryless channels with discrete time and continuous alphabets using standard techniques [56, Ch. 7].

The following theorem provides a lower bound on the capacity of the state-dependent full-duplex Gaussian RC with informed relay.

Theorem 3: The capacity C_{G} of the state-dependent Gaussian RC with informed relay satisfies $C_{\mathrm{G}} \geq R_{\mathrm{G}}^{\mathrm{lo}}$, where

$$R_{\mathrm{G}}^{\mathrm{lo}} = \max_{\rho'_{12}} \min \Bigg\{ \frac{1}{2}\log\left(1+\frac{P_1\left(1-\rho'^2_{12}\right)}{N_2}\right), \max_{\theta,\rho'_{2s}} \frac{1}{2}\log\left(1+\frac{P_1+\bar{\theta}P_2+2\rho'_{12}\sqrt{\bar{\theta}P_1P_2}}{\theta P_2+Q+N_3+2\rho'_{2s}\sqrt{\theta P_2 Q}}\right) + \frac{1}{2}\log\left(1+\frac{\theta P_2\left(1-\rho'^2_{2s}\right)}{N_3}\right)\Bigg\} \tag{26}$$

with the maximization over parameters $\rho'_{12} \in [0,1]$, $\theta \in [0,1]$, and $\rho'_{2s} \in [-1,0]$.

Proof: A formal proof of Theorem 3 is given in Appendix D.

Outline of Proof:

- We compute the lower bound (4) for an appropriate choice of the input distribution that will be specified in the sequel. By extension, Remark 1 also applies for the Gaussian case. More specifically, we should consider two important features in the design of an efficient coding scheme at the relay: obtaining correlation or coherence between the channel inputs from the source and the relay, and exploiting the channel state to remove the effect of the state on the communication. As we already mentioned, it is not obvious to accomplish these features because the channel state is not available at the source. Proceeding like for the code construction in the DM case, we split the relay input X_2^n into two parts, namely U_1^n and $\tilde{X}_2^n$. Furthermore, here we set U_1^n and $\tilde{X}_2^n$ to be *independent*. The first part, U_1^n, is a function of only the cooperative information, and is generated using standard coding. Since the source knows the cooperative information at the relay, it can generate its codeword X_1^n in such a way that it is coherent with U_1^n, by allowing correlation between X_1^n and U_1^n. The second part, $\tilde{X}_2^n$, which is independent of the source input X_1^n, is a function of both the cooperative information and the channel state S^n at the relay, and is generated using a GDPC similar to that in [14], [16], [18].
- More formally, we decompose the relay input random variable X_2 as

$$X_2 = U_1 + \tilde{X}_2 \tag{27}$$

where U_1 is zero mean Gaussian with variance $\bar{\theta}P_2$, is independent of both $\tilde{X}_2$ and S, and is correlated with X_1 with $\mathbb{E}[U_1X_1] = \rho'_{12}\sqrt{\bar{\theta}P_1P_2}$, for some $\theta \in [0,1]$, $\rho'_{12} \in [-1,1]$; and $\tilde{X}_2$ is zero mean Gaussian with variance θP_2, is independent of X_1, and is correlated with the channel state S with $\mathbb{E}[\tilde{X}_2 S] = \rho'_{2s}\sqrt{\theta P_2 Q}$, for some $\rho'_{2s} \in [-1,1]$. Expressed in terms of the covariances $\sigma_{12} = \mathbb{E}[X_1X_2] = \mathbb{E}[X_1U_1]$ and $\sigma_{2s} = \mathbb{E}[X_2S] = \mathbb{E}[\tilde{X}_2S]$, the parameters ρ'_{12}, ρ'_{2s} are given by

$$\rho'_{12} = \frac{\sigma_{12}}{\sqrt{\bar{\theta}P_1P_2}}, \quad \rho'_{2s} = \frac{\sigma_{2s}}{\sqrt{\theta P_2 Q}}. \tag{28}$$

For the GDPC, we choose the auxiliary random variable U_2 as

$$U_2 = \tilde{X}_2 + \alpha_{\mathrm{opt}} S \tag{29}$$

with

$$\alpha_{\mathrm{opt}} = \frac{\theta P_2\left(1-\rho'^2_{2s}\right) - \rho'_{2s}\sqrt{\frac{\theta P_2}{Q}}N_3}{\theta P_2\left(1-\rho'^2_{2s}\right)+N_3}. \tag{30}$$

Similarly to in the DM case, we can generalize Theorem 3 by allowing the relay to decode the source message only partially, through rate splitting at the source. The following corollary gives the resulting rate.

Corollary 3: The capacity C_{G} of the state-dependent Gaussian RC with informed relay satisfies $C_{\mathrm{G}} \geq R'^{\mathrm{lo}}_{\mathrm{G}}$, where

$$R'^{\mathrm{lo}}_{\mathrm{G}} = \max \min\{T_1, T_2, T_3\} \tag{31}$$

with

$$T_1 = \frac{1}{2}\log\left(1+\frac{\bar{\gamma}P_1\left(1-\rho'^2_{12}\right)}{N_2+\gamma P_1}\right) + \frac{1}{2}\log\left(1+\frac{\gamma P_1}{N_3+\Phi\left(\alpha',\theta,\rho'_{2s}\right)}\right) \tag{32}$$

$$T_2 = \frac{1}{2}\log\left(1+\frac{\bar{\gamma}P_1\left(1-\rho'^2_{12}\right)}{N_2+\gamma P_1}\right) + \frac{1}{2}\log\left(\frac{P'_2\left(P'_2+Q'+\gamma P_1+N_3\right)}{P'_2Q'(1-\alpha')^2+N_3\left(P'_2+\alpha'^2Q'\right)}\right) \tag{33}$$

$$T_3 = \frac{1}{2}\log\left(1+\frac{P_1+\bar{\theta}P_2+2\rho'_{12}\sqrt{\bar{\theta}\bar{\gamma}P_1P_2}}{\theta P_2+Q+N_3+2\rho'_{2s}\sqrt{\theta P_2 Q}}\right) + \frac{1}{2}\log\left(\frac{P'_2(P'_2+Q'+N_3)}{P'_2Q'(1-\alpha')^2+N_3\left(P'_2+\alpha'^2Q'\right)}\right) \tag{34}$$

$P'_2 := \theta P_2(1-\rho'^2_{2s})$, $Q' := (\sqrt{Q}+\rho'_{2s}\sqrt{\theta P_2})^2$, $\Phi(\alpha',\theta,\rho'_{2s}) := \frac{P'_2Q'(1-\alpha')^2}{P'_2+\alpha'^2Q'}$; and the maximization is over parameters $\gamma \in [0,1]$, $\theta \in [0,1]$, $\rho'_{12} \in [0,1]$, $\rho'_{2s} \in [-1,0]$, and $\alpha' \in \mathbb{R}$ such that the second logarithm terms in T_2 and T_3 are defined.

Outline of Proof: An informal proof of Corollary 3 is as follows. We decompose the message W to be sent from the source into two independent parts W_r and W_d. The message W_r will be sent through the relay, at rate R_r; and the message W_d will be sent directly to the destination, at rate R_d. The total rate is $R'_{\mathrm{G}} = R_r + R_d$. The input X_1^n from the source is divided accordingly into two independent parts, i.e., $X_1^n = U^n + \tilde{X}_1^n$, where U^n carries message W_r and has power constraint $n\bar{\gamma}P_1$ and $\tilde{X}_1^n$ carries message W_d and has power constraint $n\gamma P_1$, for some $\gamma \in [0,1]$. The relay decodes and forwards only the part U^n, and its input sequence is obtained in a manner which is similar to that in the coding scheme for Theorem 3 (with U^n playing the role of X_1^n therein).

The rest of the proof follows by computing the lower bound in Corollary 1 using an input distribution and techniques that are essentially similar to those in the proof of Theorem 3. An outline of the important steps is given in Appendix E.

C. Upper Bound on Capacity

The following theorem provides an upper bound on the capacity of the state-dependent full-duplex general Gaussian RC with informed relay.

Theorem 4: The capacity C_{G} of the state-dependent general Gaussian RC with informed relay satisfies $C_{\mathrm{G}} \leq R_{\mathrm{G}}^{\mathrm{up}}$, where $R_{\mathrm{G}}^{\mathrm{up}}$ is given by (35), shown at the bottom of the page, and the maximization is over parameters $\rho_{12} \in [0,1]$ and $\rho_{2s} \in [-1,0]$ such that

$$\rho_{12}^2 + \rho_{2s}^2 \leq 1. \tag{36}$$

Proof: The proof of Theorem 4 is given in Appendix F. In the proof, we evaluate [1] the upper bound (17) using an appropriate joint distribution of S, X_1, X_2, Y_2, and Y_3.

Following straightforwardly the proof of Theorem 4 in Appendix F, it can be easily shown that the capacity of the state-dependent degraded Gaussian RC is upper-bounded as in the following corollary.

Corollary 4: The capacity C_{DG} of the state-dependent degraded Gaussian RC with informed relay satisfies $C_{\mathrm{DG}} \leq R_{\mathrm{DG}}^{\mathrm{up}}$, where $R_{\mathrm{DG}}^{\mathrm{up}}$ is given by (37), shown at the bottom of the page, and the maximization is over parameters $\rho_{12} \in [0,1]$ and $\rho_{2s} \in [-1,0]$ such that

$$\rho_{12}^2 + \rho_{2s}^2 \leq 1. \tag{38}$$

[1] In Theorem 4, if the maximizing ρ_{2s} in (35) has absolute value equal to unity then (36) implies that ρ_{12} is zero. In this case, and also in the rest of this paper, we use the convention that $\frac{0}{0} = 0$.

D. Analysis of Some Special Cases

We note that comparing the above lower and upper bounds analytically can be tedious in the general case. In what follows, we identify a few cases in which the lower bound in Theorem 3 and the upper bound in Corollary 4 meet for degraded Gaussian channels, and some extreme cases for which the lower bound in Corollary 3 and the upper bound in Theorem 4 meet for general, i.e., not necessarily degraded, Gaussian channels; and so we obtain the capacity expression for these cases.

In the following corollary, we recast the lower bound (26) into an equivalent form by substituting $\varrho_{12} = \rho_{12}'\sqrt{\theta}$ and $\varrho_{2s} = \rho_{2s}'\sqrt{\theta}$. Also, we recast the upper bound given in Corollary 4 into an equivalent form by substituting $\kappa = \rho_{12}/\sqrt{1-\rho_{2s}^2}$ and $\rho = \rho_{2s}$.

Corollary 5: For the Gaussian RC, the lower bound (26) in Theorem 3 can be written as (39) shown at the bottom of the page, where the maximization is over parameters $\theta \in [0,1]$, $\varrho_{12} \in [0,1]$, $\varrho_{2s} \in [-1,0]$ such that

$$\varrho_{12}^2 + \varrho_{2s}^2 \leq 1. \tag{40}$$

For the physically degraded case, the upper bound in Corollary 4 can be written as (41) shown at the bottom of the page, where the maximization is over parameters $\kappa \in [0,1]$ and $\rho \in [-1,0]$.

By investigating the bounds in Theorem 3 and Corollary 4 and the equivalent expressions of these bounds in Corollary 5, it can be shown that the lower bound for the degraded case is tight for certain values of P_1, P_2, Q, N_2, and N_3. The following observation states some cases for which the lower bound is tight.

Observation 1: For the physically degraded Gaussian RC, we have the following.

$$\begin{aligned} R_{\mathrm{G}}^{\mathrm{up}} = \max\min\Bigg\{ &\frac{1}{2}\log\left(1 + P_1\left(1 - \frac{\rho_{12}^2}{1-\rho_{2s}^2}\right)\left(\frac{1}{N_2} + \frac{1}{N_3}\right)\right), \\ &\frac{1}{2}\log\left(1 + \frac{\left(\sqrt{P_1} + \rho_{12}\sqrt{P_2}\right)^2}{P_2\left(1-\rho_{12}^2-\rho_{2s}^2\right) + \left(\sqrt{Q} + \rho_{2s}\sqrt{P_2}\right)^2 + N_3}\right) + \frac{1}{2}\log\left(1 + \frac{P_2\left(1-\rho_{12}^2-\rho_{2s}^2\right)}{N_3}\right)\Bigg\} \end{aligned} \tag{35}$$

$$\begin{aligned} R_{\mathrm{DG}}^{\mathrm{up}} = \max\min\Bigg\{ &\frac{1}{2}\log\left(1 + \frac{P_1\left(1-\rho_{12}^2-\rho_{2s}^2\right)}{N_2\left(1-\rho_{2s}^2\right)}\right), \\ &\frac{1}{2}\log\left(1 + \frac{\left(\sqrt{P_1} + \rho_{12}\sqrt{P_2}\right)^2}{P_2\left(1-\rho_{12}^2-\rho_{2s}^2\right) + \left(\sqrt{Q} + \rho_{2s}\sqrt{P_2}\right)^2 + N_3}\right) + \frac{1}{2}\log\left(1 + \frac{P_2\left(1-\rho_{12}^2-\rho_{2s}^2\right)}{N_3}\right)\Bigg\} \end{aligned} \tag{37}$$

$$\begin{aligned} R_{\mathrm{G}}^{\mathrm{lo}} = \max\min\Bigg\{ &\frac{1}{2}\log\left(1 + \frac{P_1\left(1-\varrho_{12}^2-\theta\right)}{N_2(1-\theta)}\right), \\ &\frac{1}{2}\log\left(1 + \frac{\left(\sqrt{P_1} + \varrho_{12}\sqrt{P_2}\right)^2 + \left(\bar{\theta}-\varrho_{12}^2\right)P_2}{P_2\left(1-\bar{\theta}-\varrho_{2s}^2\right) + \left(\sqrt{Q} + \varrho_{2s}\sqrt{P_2}\right)^2 + N_3}\right) + \frac{1}{2}\log\left(1 + \frac{P_2\left(1-\bar{\theta}-\varrho_{2s}^2\right)}{N_3}\right)\Bigg\} \end{aligned} \tag{39}$$

1) If P_1, P_2, Q, N_2, N_3 satisfy

$$N_2 \geq \max_{\zeta\in[-1,0]} \frac{P_1N_3(P_2+Q+N_3+2\zeta\sqrt{P_2Q})}{P_1N_3+P_2(1-\zeta^2)(P_1+P_2+Q+N_3+2\zeta\sqrt{P_2Q})} \tag{42}$$

then channel capacity is given by

$$C_{\mathrm{DG}} = \frac{1}{2}\log\left(1+\frac{P_1}{N_2}\right) \tag{43}$$

which is the same as the interference-free capacity, i.e., the capacity if the channel state were not present in the model, or were also known to the source.

2) If the maximizing ρ_{12} and ρ_{2s} in the upper bound in Corollary 4 are such that condition (36) is met with equality, i.e., $\rho_{12}^2+\rho_{2s}^2=1$, then the lower bound (39) is tight and gives the capacity.

Proof: The proof of Observation 1 appears in Appendix G.

Remark 5: The condition in (42) specifies a range of values (P_1,P_2,Q,N_2,N_3) for which the lower bound for the degraded Gaussian case is tight. In this case, the capacity is the same as that of the degraded Gaussian RC with informed relay and informed source or interference-free capacity. Thus, the first statement in Observation 1 also provides a *sufficient* condition for the rate loss incurred by not knowing the interference at the source as well to be zero. At a high level, condition (42) means that there is no rate loss due to the asymmetry when capacity is constrained by the broadcast part in the model, i.e., transmission from the source to the relay and the destination. By investigating the upper bound (41) and comparing it with the interference-free capacity, it can be shown that this condition is also *necessary*. That is, the interference-free capacity is attained *only* if (42) is fulfilled. If the capacity of our model is constrained by the sum rate of the cooperative MAC part, i.e., the cooperative transmission from the source and the relay to the destination, the asymmetry resulting from not knowing the interference at the source as well causes an inevitable rate loss, i.e., the term $I(X_1;S|Y_3)$ in Corollary 2.

Extreme Cases. We now summarize the behavior of the above bounds in some extreme cases.

1) Arbitrarily strong channel state: In the asymptotic case $Q\to\infty$, the lower bound in Theorem 3 and the upper bound in Corollary 5 meet, thus yielding the capacity for degraded Gaussian RC

$$C_{\mathrm{DG}}(Q=\infty) = \min\left\{\frac{1}{2}\log\left(1+\frac{P_1}{N_2}\right), \frac{1}{2}\log\left(1+\frac{P_2}{N_3}\right)\right\}. \tag{44}$$

Equation (44) suggests that traditional multihop transmission achieves the capacity in this case. A two-hop scheme allows to completely cancel the effect of the channel state by subtracting it out upon reception at the relay, and by applying standard DPC for transmission from the relay to the destination.

For arbitrarily strong channel state and general, i.e., not necessarily degraded, Gaussian RC, the lower bound in Corollary 3 and the upper bound in Theorem 4 meet if $P_2N_2\leq P_1N_3$ or $P_2+N_3\leq P_1$, and capacity in these cases is given by

$$C_{\mathrm{G}}(Q=\infty) = \frac{1}{2}\log\left(1+\frac{P_2}{N_3}\right). \tag{45}$$

It is interesting to note that if $P_2+N_3\leq P_1$ the lower bound in Corollary 3 is maximized for $\alpha'=P_2/(P_2+N_3)$ and $\gamma=1$, meaning that direct transmission from the source to the relay is, not only possible, but also optimal in this case. The relay transmits independent information and decoding this information and subtracting it out at the destination, in a sense, clears the channel for the direct transmission.

2) Deaf helper problem: In case the relay is unable to *hear* the source (e.g., due to a very noisy or broken link source-to-relay) and $Q\to\infty$, the lower bound in Corollary 3 and the upper bound in Theorem 4 meet if $N_3\leq|P_1-P_2|$, giving

$$C_{\mathrm{G}}(Q=\infty,N_2=\infty) = \frac{1}{2}\log\left(1+\frac{\min\{P_1,P_2\}}{N_3}\right). \tag{46}$$

If $Q,N_2\longrightarrow\infty$ and $N_3>|P_1-P_2|$ the bounds do not meet. However, the lower bound is "within one bit" from the upper bound if $P_1+N_3>P_2$, and it reaches it asymptotically in the power at the relay if $P_2+N_3>P_1$ and $P_2\gg N_3$, i.e.,

$$\begin{aligned} R'^{\mathrm{lo}}_{\mathrm{G}}&(Q=\infty,N_2=\infty)\\ &= R^{\mathrm{up}}_{\mathrm{G}}(Q=\infty,N_2=\infty)-o(1)\\ &= \frac{1}{2}\log\left(1+\frac{P_1}{N_3}\right)-o(1) \end{aligned} \tag{47}$$

where $o(1)\longrightarrow 0$ as $P_2\longrightarrow\infty$.

3) For $Q=0$, the lower bound in Corollary 3 reduces to the rate achievable using a partial DF scheme in an interference-free RC, i.e.,

$$R_{\mathrm{G}}(Q=0) = \max_{0\leq\rho'_{12},\gamma\leq 1}\min\left\{\frac{1}{2}\log\left(1+\frac{\bar{\gamma}P_1\left(1-\rho'^2_{12}\right)}{N_2+\gamma P_1}\right),\right.$$

$$R^{\mathrm{up}}_{\mathrm{DG}} = \max_{\kappa}\min\left\{\frac{1}{2}\log\left(1+\frac{P_1(1-\kappa^2)}{N_2}\right), \max_{\rho}\frac{1}{2}\log\left(1+\frac{P_2\left(1-\kappa^2(1-\rho^2)-\rho^2\right)}{N_3}\right)\right.$$
$$\left. +\frac{1}{2}\log\left(1+\frac{P_1+\kappa^2(1-\rho^2)P_2+2\kappa\sqrt{1-\rho^2}\sqrt{P_1P_2}}{P_2(1-\kappa^2(1-\rho^2))+Q+2\rho\sqrt{P_2Q}+N_3}\right)\right\} \tag{41}$$

$$\left.\frac{1}{2}\log\left(1+\frac{\bar{\gamma}P_1+P_2+2\rho'_{12}\sqrt{\bar{\gamma}P_1P_2}}{N_3+\gamma P_1}\right)\right\}$$
$$+\frac{1}{2}\log\left(1+\frac{\gamma P_1}{N_3}\right) \quad (48)$$

and the upper bound in Theorem 4 reduces to the cut-set upper bound. Furthermore, if the channel is degraded these bounds meet and give the capacity of standard degraded Gaussian RC [47, Th. 5].

4) If $P_2=0$, capacity for general Gaussian RC is given by

$$C_{\rm G}(P_2=0)=\frac{1}{2}\log\left(1+\frac{P_1}{Q+N_3}\right). \quad (49)$$

E. Numerical Examples and Discussion

In this section, we discuss some numerical examples, for both the degraded Gaussian case and the general Gaussian case. We consider two numerical examples, a) $P_1=P_2=Q=10$ dB, $N_3=20$ dB; and b) $P_1=P_2=Q=N_3=10$ dB.

Fig. 4 illustrates the lower bound (39) and the upper bound (41) as functions of the signal-to-noise-ratio (SNR) at the relay, i.e., SNR $=P_1/N_2$ (in decibels), for a degraded channel.[2] Also shown for comparison are the cut-set upper bound (19) computed for the degraded Gaussian case and the trivial lower bound obtained by considering the channel state as an unknown noise and implementing full-DF at the relay [47, Th. 5].

The curves show that the lower bound and the upper bound do not meet for all SNR regimes. However, as it is visible from the depicted numerical examples, the gap between the two bounds is small for the degraded case. Furthermore, the curves in Fig. 4 also illustrate the results in observation 1, by showing that the lower bound and the upper bound meet for the cases identified in Observation 1. We note that the pentagram marker visible in Fig. 4 indicates capacity when the noise at the relay is equal to the right-hand side of (42); and this illustrates the first case for which the lower bound and the upper bound meet in Proposition 1. Also, Fig. 5 depicts the variation of $\rho_{12}^2+\rho_{2s}^2$, where ρ_{12} and ρ_{2s} are the maximizing for the upper bound, as a function of the SNR for the two numerical examples considered in Fig. 4; and this illustrates the second case for which the lower and upper bounds meet in Proposition 1.

Fig. 6 shows similar curves for the general Gaussian channel. The curves show that the lower bound (31) is close to the upper bound (35) at large SNR, i.e., when capacity of the channel is determined by the sum rate of the MAC formed by transmission from the uninformed source and the informed relay to the destination. At small SNR, the lower bound given in Corollary 3 improves upon that in Theorem 3 due to rate splitting.

Furthermore, Fig. 6 also shows the variation of the maximizing θ, ρ'_{12}, ρ'_{2s} in (26) as function of the SNR at the relay. This shows how the informed relay allocates its power among combating the interference for the source (related to the value of ρ'_{2s}) and sending signals that are coherent with the transmission from the source (related to the values of θ and ρ'_{12}).

[2]Note that for the full-duplex degraded Gaussian RC, the rate in Corollary 3 reduces to that in Theorem 3.

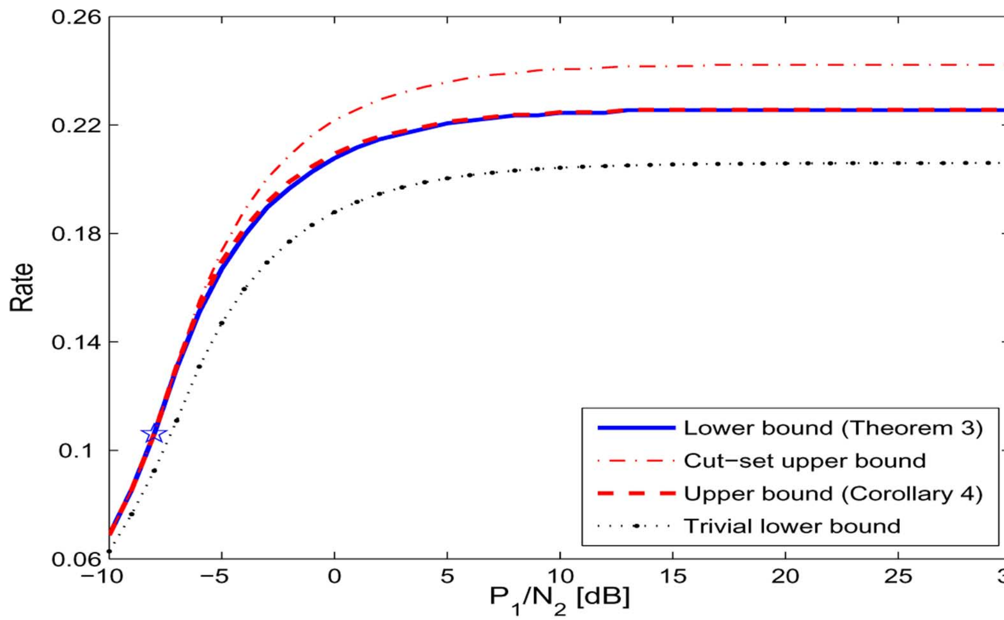
Remark 6: In standard, i.e., state-independent, Gaussian RCs, partial DF simply reduces to direct transmission if the link source-to-relay is too noisy, i.e., at low SNR. For the studied model, however, it is insightful to observe that the relay can still help the source even at very small SNR. This can be seen by observing that the lower bound (31) is better than the trivial lower bound even at this range of SNR (the trivial lower bound in Fig. 6 is obtained by treating the channel state as additional noise and implementing partial DF). This observation has some connection with the aforementioned deaf helper problem (see Case 2, section "Extreme Cases"), and it can be interpreted as follows. The relay does not hear the source and generates its input $X_{2,i}$ using a *dummy* DPC as $X_2=U_2-S$, where $X_2\sim\mathcal{N}(0,P_2)$ is independent of S and U_2 is Costa's auxiliary random variable. Upon reception of $Y_{3,i}=X_{1,i}+X_{2,i}+S_i+Z_{3,i}$ at the destination, the decoder first decodes the codeword $U_{2,i}$ fully, i.e., not only the bin index but also the correct sequence in the bin. This can be done reliably as long as $I(U_2;Y_3)-I(U_2;S)>0$. Then, the decoder at the destination subtracts out $U_{2,i}$ from $Y_{3,i}$ to obtain $\tilde{Y}_{3,i}=X_{1,i}+Z_{3,i}$ from which it decodes the source's message using standard decoding, at full rate $0.5\log(1+P_1/N_3)$. A related scenario for a helper over a state-dependent Gaussian MAC is studied in [17].

Remark 7: The gap between the lower bound and the upper bound which is visible at low SNR is due to that DF relaying (even partial) is not effective at small SNR and also to that our upper bounding technique is efficient on the MAC side but not on the BC side of the RC.

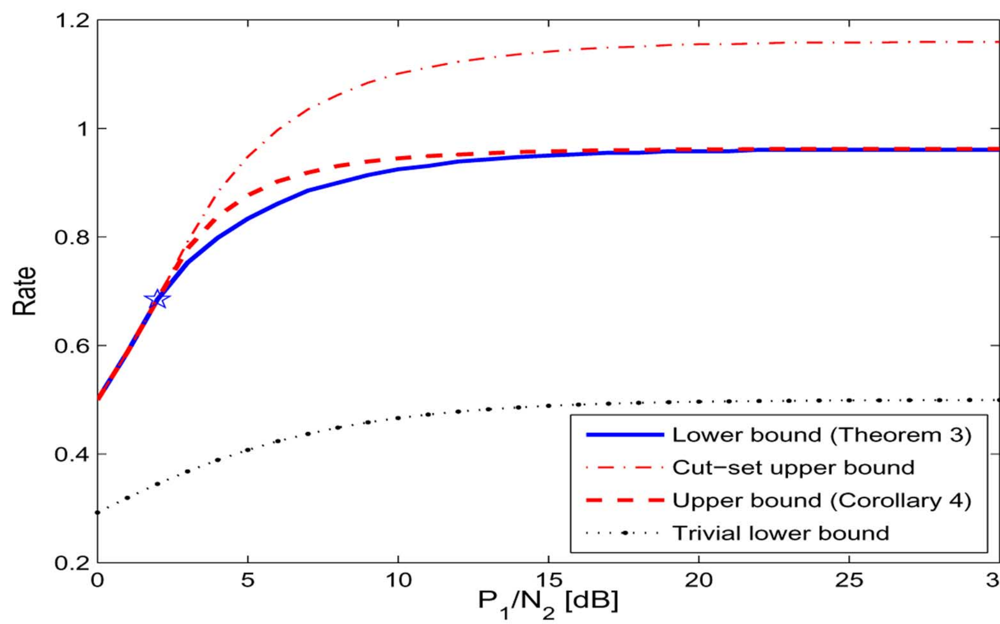
In Fig. 7, the lower and upper bounds are plotted as function of the interference power Q, for fixed value of the power at the relay and several choices of the power at the source. The curves are depicted for two examples of noise configuration: $N_2<N_3$ ($N_2=10$ dB and $N_3=20$ dB), and $N_2>N_3$ ($N_2=20$ dB and $N_3=10$ dB). The curves illustrate the discussion in the above extreme cases analysis. For instance, for both noise configurations, that the rate achievable for very large values of Q is strictly positive illustrates that transmission from the uninformed source to the uninformed destination is possible even in presence of an infinitely strong interference. Furthermore, the lower and upper bounds meet for the cases identified in the "Extreme Cases" section, for both degraded Gaussian and General Gaussian channels.

F. Half-Duplex Channel Model

In this section, we extend the results of Section IV-A to the case of half-duplex relaying, i.e., the relay can either transmit only or receive only. We consider a state-dependent Gaussian RC with informed relay, and we assume that the relay operates in a TD relaying mode. In the TD mode, for a given time window, the relay is in the receive mode for a fraction of the given time and in the transmit mode for the remaining fraction of this time. Since the message from the source is transmitted to



(a)



(b)

Fig. 4. Lower and upper bounds on the capacity of the state-dependent degraded Gaussian RC with informed relay versus the SNR in the link source-to-relay, for two examples of numerical values (a) $P_1 = P_2 = Q = 10$ dB, $N_3 = 20$ dB, and (b) $P_1 = P_2 = Q = N_3 = 10$ dB.

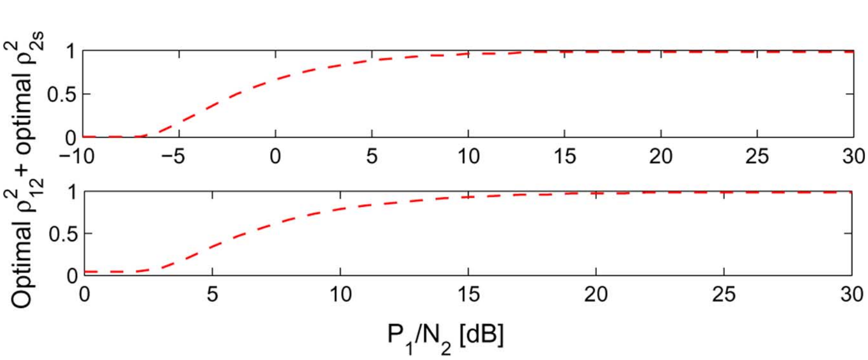


Fig. 5. The sum $\rho_{12}^2 + \rho_{2s}^2$ in the constraint (36). Optimal ρ_{12} and ρ_{2s} are the maximizing for the upper bound for the numerical examples considered in Fig. 4. The upper figure is for the upper bound curve in Fig. 4(a), and the lower figure is for the upper bound curve in Fig. 4(b).

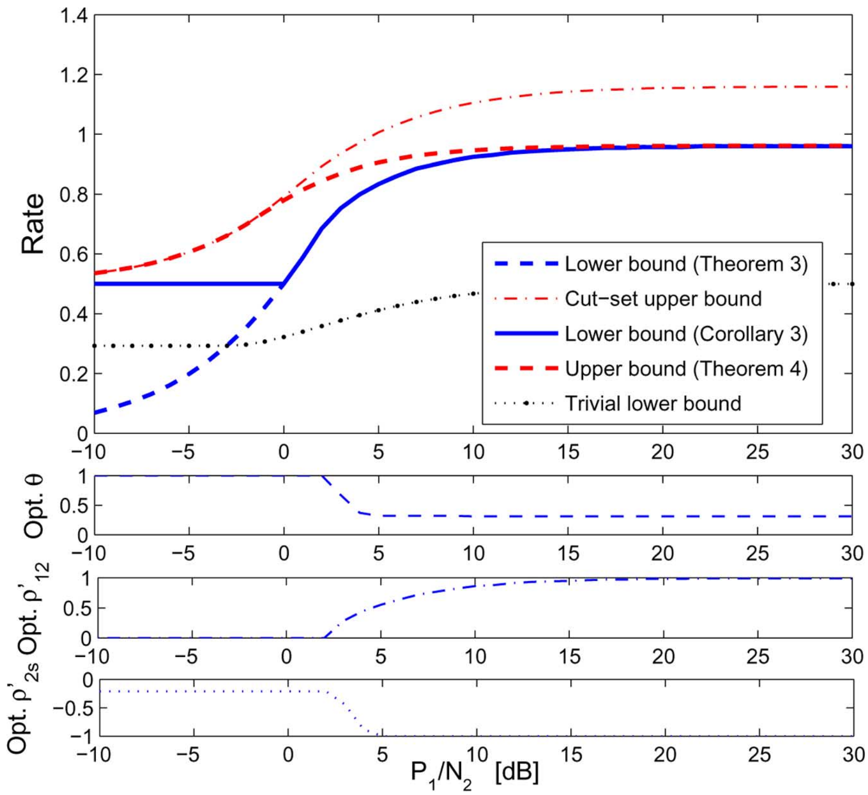


Fig. 6. Lower and upper bounds on the capacity of the state-dependent general Gaussian RC with informed relay and the maximizing $\theta, \rho'_{12}, \rho'_{2s}$ in (26) as functions of the SNR at the relay. Numerical values are $P_1 = P_2 = Q = N_3 = 10$ dB.

the destination in n channel uses, in the remaining of this section, we refer to the time indices from[3] 1 to $\lfloor \nu n \rfloor$ as the *relay-receive period* and the time indices from $\lfloor \nu n \rfloor + 1$ to n as the *relay-transmit period*, for some $\nu \in [0,1]$. Furthermore, to generalize the model, we assume that the channel state $S^{(1)}$ is zero mean Gaussian with variance $Q^{(1)}$ during the relay-receive period, and the channel state $S^{(2)}$ is zero mean Gaussian with variance $Q^{(2)}$ during the relay-transmit period. The channel output $Y_{2,i}$ at instant i at the relay is given by

$$Y_{2,i} = X_{1,i}^{(1)} + S_i^{(1)} + Z_{2,i}$$

during the relay-receive period, and is zero with probability one during the relay-transmit period. The channel output at instant i at the destination is given by

$$Y_{3,i}^{(1)} = X_{1,i}^{(1)} + S_i^{(1)} + Z_{3,i} \quad \text{during the relay-receive period} \tag{50a}$$

$$Y_{3,i}^{(2)} = X_{1,i}^{(2)} + X_{2,i} + S_i^{(2)} + Z_{3,i} \quad \text{during the relay-transmit period.} \tag{50b}$$

Furthermore, the source has average power constraint $P_1^{(1)}$ during the relay-receive period and average power constraint $P_1^{(2)}$ during the relay-transmit period; the relay has average power constraint P_2.

For fixed values of ν, $P_1^{(1)}$, $P_1^{(2)}$, and P_2, we have the following upper and lower bounds on the capacity of the state-dependent half-duplex Gaussian RC with informed relay.

[3]For a scalar x, $\lfloor x \rfloor$ stands for the largest integer small than or equal to x.

Proposition 1: The capacity of the state-dependent TD Gaussian RC with informed relay is upper bounded by

$$R_{\mathrm{G}}^{\mathrm{up}}(\mathrm{TD}) = \max \min \left\{R_1^{\mathrm{up}}, R_2^{\mathrm{up}}\right\} \tag{51}$$

with

$$\begin{aligned} R_1^{\mathrm{up}} =& \frac{\nu}{2}\log\left(1 + P_1^{(1)}\left(\frac{1}{N_2} + \frac{1}{N_3}\right)\right) \\ &+ \frac{\bar{\nu}}{2}\log\left(1 + \frac{P_1^{(2)}\left(1-\rho_{12}^2-\rho_{2s}^2\right)}{N_3\left(1-\rho_{2s}^2\right)}\right) \qquad (52a) \\ R_2^{\mathrm{up}} =& \bar{\nu}\Psi\left(P_1^{(2)}, P_2, Q^{(2)}, \rho_{12}, \rho_{2s}\right) \\ &+ \frac{\nu}{2}\log\left(1 + \frac{P_1^{(1)}}{N_3 + Q^{(1)}}\right) \qquad (52b) \end{aligned}$$

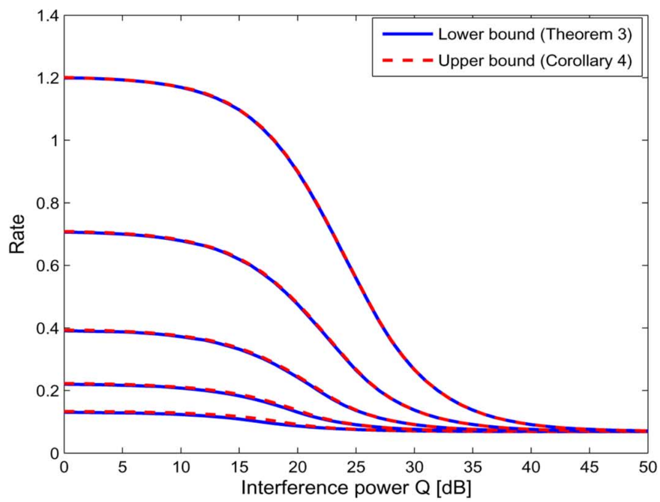
where $\Psi(P_1, P_2, Q, \rho_{12}, \rho_{2s})$ is defined as the second term of the minimization in (35), and the maximization is over parameters $\rho_{12} \in [0,1]$ and $\rho_{2s} \in [-1,0]$ such that $\rho_{12}^2 + \rho_{2s}^2 \le 1$.

Proposition 2: The capacity of the state-dependent TD Gaussian RC with informed relay is lower bounded by

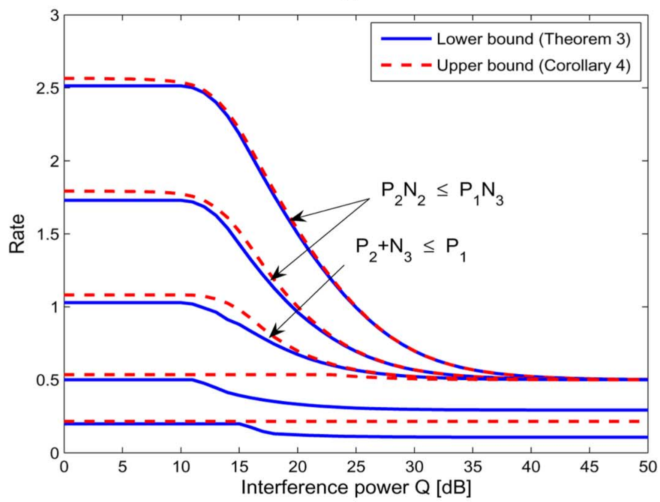
$$R_{\mathrm{G}}^{\mathrm{lo}}(\mathrm{TD}) = \max \min \left\{R_1^{\mathrm{lo}}, R_2^{\mathrm{lo}}, R_3^{\mathrm{lo}}\right\} \tag{53}$$

with

$$R_1^{\mathrm{lo}} = \frac{\nu}{2}\log\left(1 + \frac{P_1^{(1)}}{N_2}\right)$$



(a)



(b)

Fig. 7. Bounds on channel capacity as function of the interference power Q. The curves correspond to different choices of power at the source: from bottom to top P_1 = 5, 10, 15, 20, 25 dB. (a) Degraded Gaussian RC. $P_2 = N_2 = 10$, $N_3 = 20$ dB. (b) General Gaussian RC. $P_2 = N_3 = 10$, $N_2 = 20$ dB.

$$+\frac{\bar{\nu}}{2}\log\left(1+\frac{\left(1-\rho_{12}'^2\right)P_1^{(2)}}{N_3+\Phi\left(\alpha',\theta,\rho_{2s}'\right)}\right) \tag{54a}$$

$$R_2^{\mathrm{lo}}=\frac{\nu}{2}\log\left(1+\frac{P_1^{(1)}}{N_2}\right) + \frac{\bar{\nu}}{2}\log\left(\frac{P_2'\left(P_2'+Q'^{(2)}+\left(1-\rho_{12}'^2\right)P_1^{(2)}+N_3\right)}{P_2'Q'^{(2)}(1-\alpha')^2+N_3\left(P_2'+\alpha'^2Q'^{(2)}\right)}\right) \tag{54b}$$

$$R_3^{\mathrm{lo}}=\frac{\nu}{2}\log\left(1+\frac{P_1^{(1)}}{N_3+Q^{(1)}}\right) + \frac{\bar{\nu}}{2}\log\left(1+\frac{P_1^{(2)}+\bar{\theta}P_2+2\rho_{12}'\sqrt{\bar{\theta}P_1^{(2)}P_2}}{\theta P_2+Q^{(2)}+2\rho_{2s}'\sqrt{\theta P_2Q^{(2)}}+N_3}\right) + \frac{\bar{\nu}}{2}\log\left(\frac{P_2'\left(P_2'+Q'^{(2)}+N_3\right)}{P_2'Q'^{(2)}(1-\alpha')^2+N_3\left(P_2'+\alpha'^2Q'^{(2)}\right)}\right) \tag{54c}$$

where maximization is over parameters $\theta \in [0,1]$, $\rho_{12}' \in [0,1]$, $\rho_{2s}' \in [-1,0]$, and $\alpha' \in \mathbb{R}$ such that the last logarithm terms on the right-hand side of (54b) and (54c) are defined

$$\Phi\left(\alpha',\theta,\rho_{2s}'\right) := \frac{P_2'Q'^{(2)}(1-\alpha')^2}{P_2'+\alpha'^2Q'^{(2)}} \tag{55}$$

and $P_2' := \theta P_2(1-\rho_{2s}'^2)$, $Q'^{(2)} := (\sqrt{Q^{(2)}}+\rho_{2s}'\sqrt{\theta P_2})^2$.

The proofs of Proposition 1 and Proposition 2 appear in Appendix H.

Remark 8: The coding scheme employed for the proof of Proposition 2 preassigns the time slots for the relay's receiving and transmitting modes. All the nodes then know ahead of time when the relay receives and when it transmits. This is relevant for nodes synchronization but suboptimal in general for information rate. Instead, one can let the source and the relay choose the relay's mode and, so, in a sense, transmit additional information to the destination through that choice. This idea is introduced in [57] in the context of wireline and wireless networks without state and is called *mode coding* therein; see also [58, Sec. 4.3]. More specifically, let M denote a random variable that takes on values 1 (receive) and 2 (transmit) with probabilities ν and $\bar{\nu}$, respectively. Also, let us redefine the channel so as to include the relay's operating mode as $W_{Y_2,Y_3|X_1,X_2,S,M}$; set $X_1' = (X_1, M)$, $X_2' = (X_2, M)$, $U' = (U, M)$, $U_1' = (U_1, M)$, $U_2' = (U_2, M)$, and choose $U = X_1^{(1)}$ if $M = 1$ and $U = 0$ if $M = 2$. Then, using $(X_1', X_2', U', U_1', U_2')$ in place of (X_1, X_2, U, U_1, U_2) in (12), it can be shown that this yields a rate which is obtained by maximizing the minimum among R_1^{lo}, R_2^{lo}, and $R_3^{\mathrm{lo}} + I(M;Y_3)$, i.e., larger than (53). However, as mentioned in [58, Sec. 4.3], the improvement is no larger than 1 bit per block and, also, harnessing it in practice requires some challenges in general.

V. CONCLUSION

In this paper, we consider a state-dependent RC with the channel state available noncausally at only the relay, i.e., neither at the source nor at the destination. We refer to this communication model as state-dependent RC with informed relay. This setup may model the basic building block for node cooperation over wireless networks in which some of the terminals may be equipped with cognition capabilities that enable estimating to high accuracy the states of the channel.

We investigate this problem in the DM case and in the Gaussian case, and we derive bounds on the channel capacity. For both cases, the upper bounds are tighter than those obtained by assuming that the channel state is also available at the source and the destination, and they help characterizing the rate loss due to the asymmetry, i.e., having the channel state available at the relay but not the source. Key to the development of the lower bounds is a coding scheme that splits the codeword at the informed relay into two parts: one part depends only on the cooperative information, not on the known channel state, and is used to enable coherent transmission from the source and the relay to the destination; another part is a function of both the cooperative information and the known channel state, and is used to combat the effects of the channel state on the communication

	block 1	block 2	block 3	block 4
Source codewords	$\mathbf{x}_1(1, w_1)$	$\mathbf{x}_1(w_1, w_2)$	$\mathbf{x}_1(w_2, w_3)$	$\mathbf{x}_1(w_3, 1)$
Relay codewords	$\mathbf{u}_1(1)$	$\mathbf{u}_1(w_1)$	$\mathbf{u}_1(w_2)$	$\mathbf{u}_1(w_3)$
	$\mathbf{u}_2(1, j(\mathbf{s}[1], 1))$	$\mathbf{u}_2(w_1, j(\mathbf{s}[2], w_1))$	$\mathbf{u}_2(w_2, j(\mathbf{s}[3], w_2))$	$\mathbf{u}_2(w_3, j(\mathbf{s}[4], w_3))$
	$\mathbf{x}_2(1)$	$\mathbf{x}_2(w_1)$	$\mathbf{x}_2(w_2)$	$\mathbf{x}_2(w_3)$

Fig. 8. Regular encoding for DF for the state-dependent RC with informed relay. At the beginning of block i, the source transmits $\mathbf{x}_1(w_{i-1}, w_i)$ and the relay transmits a codeword $\mathbf{x}_2(w_{i-1})$ with i.i.d. components given $(\mathbf{u}_1(w_{i-1}), \mathbf{u}_2(w_{i-1}, j(\mathbf{s}[i], w_{i-1})), \mathbf{s}[i])$ drawn according to the marginal $P_{X_2|U_1,U_2,S}$.

through a generalized Gel'fand–Pinsker binning scheme. In the Gaussian case, we consider average power constraints at the source and the relay and power allocation at the relay among the two parts of the code, allowing for a tradeoff between the coherence gain obtained through the coherent transmission and the mitigation of the channel state.

Specializing the results to the case in which the channel is physically degraded, we show that the developed lower and upper bounds meet in some cases, thus characterizing the channel capacity. For the general Gaussian case, the bounds are in general close, but they meet only in some extreme cases.

Appendix

Throughout this section, we denote the set of strongly jointly ϵ-typical sequences [46, Ch. 14.2] with respect to the distribution $P_{X,Y}$ as $T_\epsilon^n(P_{X,Y})$.

A. Proof of Theorem 1

Consider the random coding scheme that we outlined in Section III. We now give a formal description of the coding scheme and analyze the average probability of error.

As we outlined after Theorem 1 we transmit in $B+1$ blocks, each of length n. During each of the first B blocks, the source encodes a message $w_i \in [1, 2^{nR}]$ and sends it over the channel, where $i = 1, \dots, B$ denotes the index of the block. For fixed n, the average rate $R\frac{B}{B+1}$ over $B+1$ blocks approaches R as $B \longrightarrow +\infty$.

Encoding: Let w_i be the new message to be sent from the source node at the beginning of block i, and w_{i-1} be the message sent in the previous block $i-1$. At the beginning of block i, the relay has decoded the message w_{i-1} correctly and the source sends $\mathbf{x}_1(w_{i-1}, w_i)$. The relay searches for the smallest $j \in \{1, \dots, J\}$ such that $\mathbf{u}_1(w_{i-1})$, $\mathbf{u}_2(w_{i-1}, j)$ and $\mathbf{s}[i]$ are jointly typical. Denote this j by $j^\star = j(\mathbf{s}[i], w_{i-1})$. If such $j^\star$ is not found, or if the observed state is not typical, an error is declared and $j^\star$ is set to J. Then, the relay transmits a vector $\mathbf{x}_2(w_{i-1})$ with i.i.d. components given $(\mathbf{u}_1(w_{i-1}), \mathbf{u}_2(w_{i-1}, j^\star), \mathbf{s}[i])$ drawn according to the marginal $P_{X_2|U_1,U_2,S}$ induced by the distribution (5).

The encoder at the source declares an error if the chosen codeword exceeds the power constraint, that is, $\varphi_1^n(\mathbf{x}_1(w_{i-1}, w_i)) > \Gamma_1 + \gamma_1(\epsilon)$ for some $\gamma_1(\epsilon) > 0$. Similarly, the encoder at the relay declares an error if $\varphi_2^n(\mathbf{x}_2(w_{i-1})) > \Gamma_2 + \gamma_2(\epsilon)$, for some $\gamma_2(\epsilon) > 0$.

For convenience, we list the codewords at the source and the relay that are used for transmission in the first four blocks in Fig. 8.

Decoding: The decoding procedure at the relay is based on joint typicality. The decoding procedure at the destination is based on a combination of joint typicality and backward-decoding.

1) At the end of block i, the relay knows w_{i-1} and declares that $\hat{w}_i$ is sent if there is a unique $\hat{w}_i$ such that $\mathbf{x}_1(w_{i-1}, \hat{w}_i)$ and $(\mathbf{y}_2[i], \mathbf{s}[i])$ are jointly typical given $\mathbf{u}_1(w_{i-1})$, $\mathbf{u}_2(w_{i-1}, j^\star)$ and $\mathbf{x}_2(w_{i-1})$, where $\mathbf{y}_2[i]$ denotes the output of the channel at the relay in block i and $j^\star = j(\mathbf{s}[i], w_{i-1})$ as mentioned earlier. One can show that the decoding error in this step is small for sufficiently large n if

$$R < I(X_1; Y_2 | S, U_1, X_2). \tag{A-1}$$

2) At the end of the transmission, the destination has collected all the blocks of channel outputs $\mathbf{y}_3[1], \mathbf{y}_3[2], \dots, \mathbf{y}_3[B+1]$, and can then perform Willem's backward-decoding by first decoding w_B from $\mathbf{y}_3[B+1]$.

First, the destination declares that $\hat{w}_B$ is sent if there is a unique $\hat{w}_B$ such that $\mathbf{u}_1(\hat{w}_B), \mathbf{u}_2(\hat{w}_B, j_B), \mathbf{x}_1(\hat{w}_B, 1), \mathbf{y}_3[B+1]$ are jointly typical, for some $j_B \in \{1, \dots, J\}$. One can show that the decoding error in this step is small for sufficiently large n if

$$R < I(X_1, U_1, U_2; Y_3) - I(U_2; S | U_1). \tag{A-2}$$

Next, for b ranging from B to 2, the destination knows w_b and decodes w_{b-1} based on the information received in block b. It declares that $\hat{w}_{b-1}$ is sent if there is a unique $\hat{w}_{b-1}$ such that $\mathbf{u}_1(\hat{w}_{b-1}), \mathbf{u}_2(\hat{w}_{b-1}, j_{b-1}), \mathbf{x}_1(\hat{w}_{b-1}, w_b), \mathbf{y}_3[b]$ are jointly typical, for some $j_{b-1} \in \{1, \dots, J\}$. One can show that the decoding error in this step is small for sufficiently large n if (A-2) is true.

Analysis of Probability of Error: Fix a probability distribution $P_{S,U_1,U_2,X_1,X_2,Y_2,Y_3}$ satisfying (5) and $\mathbb{E}[\varphi_i(X_i)] < \Gamma_i$, $i = 1, 2$. Let $\mathbf{s}[i]$, w_{i-1}, and w_i be the state sequence in block i, the message sent from the source node in block $i-1$, and the message sent in block i, respectively. As we already mentioned above, at the beginning of block i the source transmits $\mathbf{x}_1(w_{i-1}, w_i)$ and the relay transmits a vector $\mathbf{x}_2(w_{i-1})$ with i.i.d. components conditionally given $(\mathbf{u}_1(w_{i-1}), \mathbf{u}_2(w_{i-1}, j^\star), \mathbf{s}[i])$, with $j^\star = j(\mathbf{s}[i], w_{i-1})$, drawn according to the marginal $P_{X_2|U_1,U_2,S}$.

The average probability of error is such that

$$\Pr(\text{Error}) \leq \sum_{(\mathbf{s}, \mathbf{u}_1) \notin T_\epsilon^n(Q_S P_{U_1})} \Pr(\mathbf{s}) \Pr(\mathbf{u}_1)$$

$$+ \sum_{(\mathbf{s},\mathbf{u}_1)\in T_\epsilon^n(Q_S P_{U_1})} \Pr(\mathbf{s})\Pr(\mathbf{u}_1)\Pr(\text{error}|\mathbf{s},\mathbf{u}_1). \quad \text{(A-3)}$$

The first term, $\Pr((\mathbf{s},\mathbf{u}_1)\notin T_\epsilon^n(Q_S P_{U_1}))$, on the right-hand side of (A-3) goes to zero as $n\to\infty$, by the asymptotic equipartition property (AEP) [46, p. 384]. Thus, it is sufficient to upper bound the second term on the right-hand side of (A-3).

We now examine the probabilities of the error events associated with the encoding and decoding procedures. The error event is contained in the union of the error events given below, where the events E_{1i}, E_{2i} and E_{3i} correspond to the encoding step at block i; the events E_{4i} and E_{5i} correspond to decoding at the relay at block i; the events E_{6B} and E_{7B} correspond to decoding at the destination at block $B+1$, and for b ranging from B to 2, the events $E_{8(b-1)}$ and $E_{9(b-1)}$ correspond to decoding at the destination at block b.

- Let E_{1i} be the event that there is no sequence $\mathbf{u}_2(w_{i-1},j)$ jointly typical with $\mathbf{s}[i]$ given $\mathbf{u}_1(w_{i-1})$, i.e.,

$$E_{1i}=\Big\{\nexists\, j\in\{1,\dots,J\}\text{s.t.} \quad (\mathbf{u}_1(w_{i-1}),\mathbf{u}_2(w_{i-1},j),\mathbf{s}[i])\in T_\epsilon^n(P_{U_1,U_2,S})\Big\}.$$

  To bound the probability of the event E_{1i}, we use a standard argument [2]. More specifically, for $\mathbf{u}_2(w_{i-1},j)$ and $\mathbf{s}[i]$ generated independently given $\mathbf{u}_1(w_{i-1})$, with i.i.d. components drawn according to $P_{U_2|U_1}$ and Q_S, respectively, the probability that $\mathbf{u}_2(w_{i-1},j)$ is jointly typical with $\mathbf{s}[i]$ given $\mathbf{u}_1(w_{i-1})$ is greater than $(1-\epsilon)2^{-n(I(U_2;S|U_1)+\epsilon)}$ for sufficiently large n. There is a total of J such $\mathbf{u}_2$'s in each bin. The probability of the event E_{1i}, the probability that there is no such $\mathbf{u}_2$, is therefore bounded as

$$\Pr(E_{1i})\le[1-(1-\epsilon)2^{-n(I(U_2;S|U_1)+\epsilon)}]^J. \quad \text{(A-4)}$$

  Taking the logarithm on both sides of (A-4) and substituting J using (7) we obtain $\ln(\Pr(E_{1i}))\le-(1-\epsilon)2^{n\epsilon}$. Thus, $\Pr(E_{1i})\to 0 \quad \text{as} \quad n\to\infty$.

- Let E_{2i} be the event that the chosen codeword at the source $\mathbf{x}_1(w_{i-1,w_i})$ exceeds the power constraint Γ_1 by $\gamma_1(\epsilon)$

$$E_{2i}=\{\varphi_1^n(\mathbf{x}_1(w_{i-1,w_i}))>\Gamma_1+\gamma_1(\epsilon)\}. \quad \text{(A-5)}$$

  By the weak law of large numbers, we have

$$\Pr(E_{2i})=\Pr\left(\frac{1}{n}\sum_{i=1}^{n}\varphi_1(x_{1,i}(w))>\Gamma_1+\gamma_1(\epsilon)\right) < \epsilon \quad \text{(A-6)}$$

  for n large enough and $\mathbb{E}[\varphi_1(X_1)]<\Gamma_1$.

- Let E_{3i} be the event that the chosen codeword at the relay $\mathbf{x}_2(w_{i-1})$ exceeds the power constraint Γ_2 by $\gamma_2(\epsilon)$

$$E_{3i}=\{\varphi_2^n(\mathbf{x}_2(w_{i-1}))>\Gamma_2+\gamma_2(\epsilon)\}. \quad \text{(A-7)}$$

  Using arguments similar to those for the event E_{2i}, we get $\Pr(E_{3i}|E_{1i}^c)<\epsilon$ for n large enough and $\mathbb{E}[\varphi_2(X_2)]<\Gamma_2$, where E_{1i}^c denotes the event complement of E_{1i}.

- Let E_{4i} be the event that $\mathbf{x}_1(w_{i-1},w_i)$, $\mathbf{y}_2[i]$, $\mathbf{s}[i]$ are not jointly typical given $\mathbf{u}_1(w_{i-1})$, $\mathbf{u}_2(w_{i-1},j^\star)$ and $\mathbf{x}_2(w_{i-1})$, i.e.,

$$E_{4i}=\Big\{\Big(\mathbf{u}_1(w_{i-1}),\mathbf{u}_2(w_{i-1},j^\star), \quad \mathbf{x}_1(w_{i-1},w_i),\mathbf{x}_2(w_{i-1}),\mathbf{y}_2[i],\mathbf{s}[i]\Big) \quad \notin T_\epsilon^n(P_{U_1,U_2,X_1,X_2,Y_2,S})\Big\}.$$

  Conditioned on E_{1i}^c, E_{2i}^c, E_{3i}^c, we have that $(\mathbf{s}[i],\mathbf{u}_1(w_{i-1}))$ is jointly typical with $\mathbf{u}_2(w_{i-1},j^\star)$ and with the source input $\mathbf{x}_1(w_{i-1},w_i)$ and the relay input $\mathbf{x}_2(w_{i-1})$, i.e.,

$$(\mathbf{s}[i],\mathbf{u}_1(w_{i-1}),\mathbf{u}_2(w_{i-1},j^\star),\mathbf{x}_1(w_{i-1},w_i),\mathbf{x}_2(w_{i-1})) \in T_\epsilon^n(Q_S P_{U_1}P_{X_1|U_1}P_{U_2,X_2|S,U_1,X_1}). \quad \text{(A-8)}$$

  For $\mathbf{s}[i]$, $\mathbf{u}_1(w_{i-1})$, $\mathbf{u}_2(w_{i-1},j^\star)$, $\mathbf{x}_1(w_{i-1},w_i)$, and $\mathbf{x}_2(w_{i-1})$ jointly typical, we have $\Pr\left(E_{4i}|\cap_{k=1}^3 E_{ki}^c\right)\longrightarrow 0$ as $n\longrightarrow\infty$, by the Markov Lemma [46, p. 436].

- Let E_{5i} be the event that $\mathbf{x}_1(w_{i-1},w_i')$, $\mathbf{y}_2[i]$, $\mathbf{s}[i]$ are jointly typical given $\mathbf{u}_1(w_{i-1})$, $\mathbf{u}_2(w_{i-1},j^\star)$, $\mathbf{x}_2(w_{i-1})$ for some $w_i'\neq w_i$, i.e.,

$$E_{5i}=\Big\{\exists\, w_i'\in\{1,\dots,M\}\text{s.t.}w_i'\neq w_i, \quad \Big(\mathbf{u}_1(w_{i-1}),\mathbf{u}_2(w_{i-1},j^\star), \quad \mathbf{x}_1(w_{i-1},w_i'),\mathbf{x}_2(w_{i-1}),\mathbf{y}_2(i),\mathbf{s}\Big) \quad \in T_\epsilon^n(P_{U_1,U_2,X_1,X_2,Y_2,S})\Big\}.$$

  Using the union bound and standard arguments on strongly typical sequences, the probability of the event E_{5i} conditioned on E_{1i}^c, E_{2i}^c, E_{3i}^c, E_{4i}^c can be easily bounded as

$$\begin{aligned}\Pr(E_{5i}|E_{1i}^c,E_{2i}^c,E_{3i}^c,E_{4i}^c) &\le M2^{-n(I(X_1;Y_2,S|U_1,U_2,X_2)-\epsilon)} &\text{(A-9a)}\\ &=2^{-n(I(X_1;Y_2|S,U_1,U_2,X_2)-R+3\epsilon)} &\text{(A-9b)}\end{aligned}$$

  where in (A-9b) we used the fact that $I(X_1;S|U_1,U_2,X_2)=0$ under the joint distribution (5). Thus, $\Pr\left(E_{5i}|\cap_{k=1}^4 E_{ki}^c\right)\longrightarrow 0$ as $n\longrightarrow\infty$ if $R<I(X_1;Y_2|S,U_1,U_2,X_2)$. This constraint can be rewritten equivalently as

$$\begin{aligned}R&<I(X_1;Y_2|S,U_1,U_2,X_2)\\ &=H(Y_2|S,U_1,U_2,X_2)-H(Y_2|S,U_1,U_2,X_1,X_2)\\ &=H(Y_2|S,U_1,X_2)-H(Y_2|S,X_1,X_2)\end{aligned}$$

$$= I(X_1; Y_2|S, U_1, X_2) \tag{A-10}$$

where the second equality holds since the measure (5) satisfies $P_{Y_2|S,U_1,U_2,X_2} = P_{Y_2|S,U_1,X_2}$; and Y_2 and (U_1, U_2) are conditionally independent given (S, X_1, X_2).

- For the decoding of message w_B at the destination, let E_{6B} be the event that $\mathbf{u}_1(w_B)$, $\mathbf{u}_2(w_B, j(\mathbf{s}[B+1], w_B))$, $\mathbf{x}_1(w_B, 1)$, $\mathbf{y}_3[B+1]$ are not jointly typical, i.e.,

$$E_{6B} = \Big\{ \Big(\mathbf{u}_1(w_B), \mathbf{u}_2(w_B, j(\mathbf{s}[B+1], w_B)), \mathbf{x}_1(w_B, 1), \mathbf{y}_3[B+1] \Big) \notin T_\epsilon^n(P_{U_1,U_2,X_1,Y_3}) \Big\}.$$

  For $\mathbf{s}[B+1]$, $\mathbf{u}_1(w_B)$, $\mathbf{u}_2(w_B, j(\mathbf{s}[B+1], w_B))$, $\mathbf{x}_1(w_B, 1)$, and $\mathbf{x}_2(w_B)$ jointly typical as shown by (A-8), $\Pr\left(E_{6B}| \cap_{k=1}^5 E_{ki}^c\right) \longrightarrow 0$ as $n \longrightarrow \infty$, by the Markov lemma.

- For the decoding of message w_B at the destination, let E_{7B} be the event that $\mathbf{u}_1(w'_B)$, $\mathbf{u}_2(w'_B, j'_B)$, $\mathbf{x}_1(w'_B, 1)$, $\mathbf{y}_3[B+1]$ are jointly typical for some $w'_B \neq w_B$ and some $j'_B \in \{1, \ldots, J\}$, i.e.,

$$E_{7B} = \Big\{ \exists\, w'_B \in \{1, \ldots, M\}, j'_B \in \{1, \ldots, J\} \text{ s.t. } w'_B \neq w_B, \left(\mathbf{u}_1(w'_B), \mathbf{u}_2(w'_B, j'_B), \mathbf{x}_1(w'_B, 1), \mathbf{y}_3[B+1]\right) \in T_\epsilon^n(P_{U_1,U_2,X_1,Y_3}) \Big\}.$$

  Conditioned on the events E_{1i}^c, E_{2i}^c, E_{3i}^c, E_{4i}^c, E_{5i}^c, E_{6B}^c, the probability of the event E_{7B} can be bounded using the union bound, as

$$\begin{aligned} \Pr\left(E_{7B}| \cap_{k=1}^5 E_{ki}^c, E_{6B}^c\right) &\leq MJ2^{-n(I(X_1,U_1,U_2;Y_3)-\epsilon)} && \text{(A-11a)} \\ &= 2^{-n(I(X_1,U_1,U_2;Y_3)-I(U_2;S|U_1)-R+\epsilon)}. && \text{(A-11b)} \end{aligned}$$

  Thus, $\Pr\left(E_{7B}| \cap_{k=1}^5 E_{ki}^c, E_{6B}^c\right) \longrightarrow 0$ as $n \longrightarrow \infty$ if $R < I(X_1, U_1, U_2; Y_3) - I(U_2; S|U_1)$.

- For the decoding of message w_{b-1} at the destination, $b = B, \ldots, 2$, let $E_{8(b-1)}$ be the event that $\mathbf{u}_1(w_{b-1})$, $\mathbf{u}_2(w_{b-1}, j(\mathbf{s}[b], w_{b-1}))$, $\mathbf{x}_1(w_{b-1}, w_b)$, $\mathbf{y}_3[b]$ are not jointly typical, i.e.,

$$E_{8(b-1)} = \Big\{ \Big(\mathbf{u}_1(w_{b-1}), \mathbf{u}_2(w_{b-1}, j(\mathbf{s}[b], w_{b-1})), \mathbf{x}_1(w_{b-1}, w_b), \mathbf{y}_3[b] \Big) \notin T_\epsilon^n(P_{U_1,U_2,X_1,Y_3}) \Big\}.$$

  For $\mathbf{s}[b]$, $\mathbf{u}_1(w_{b-1})$, $\mathbf{u}_2(w_{b-1}, j(\mathbf{s}[b], w_{b-1}))$, $\mathbf{x}_1(w_{b-1}, w_b)$, and $\mathbf{x}_2(w_{b-1})$ jointly typical as shown by (A-8), $\Pr\left(E_{8(b-1)}| \cap_{k=1}^5 E_{ki}^c, E_{6B}^c, E_{7B}^c\right) \longrightarrow 0$ as $n \longrightarrow \infty$, by the Markov lemma.

- For the decoding of message w_{b-1} at the destination, let $E_{9(b-1)}$ be the event that $\mathbf{u}_1(w'_{b-1})$, $\mathbf{u}_2\left(w'_{b-1}, j'_{b-1}\right)$, $\mathbf{x}_1\left(w'_{b-1}, w_b\right)$, $\mathbf{y}_3[b]$ are jointly typical for some $w'_{b-1} \neq w_{b-1}$ and some $j'_{b-1} \in \{1, \ldots, J\}$, i.e.,

$$E_{9(b-1)} = \Big\{ \exists\, w'_{b-1} \in \{1, \ldots, M\}, j'_{b-1} \in \{1, \ldots, J\}, \text{ s.t. } w'_{b-1} \neq w_{b-1}, \left(\mathbf{u}_1\left(w'_{b-1}\right), \mathbf{u}_2\left(w'_{b-1}, j'_{b-1}\right), \mathbf{x}_1\left(w'_{b-1}, w_b\right), \mathbf{y}_3[b]\right) \in T_\epsilon^n(P_{U_1,U_2,X_1,Y_3}) \Big\}.$$

  Proceeding like for the event E_{7B}, one can easily show that $\Pr(E_{9(b-1)}| \cap_{k=1}^5 E_{ki}^c, E_{6B}^c, E_{7B}^c, E_{8(b-1)}^c)$ can be bounded similarly to in (A-11), and this shows that $\Pr(E_{9(b-1)}| \cap_{k=1}^5 E_{ki}^c, E_{6B}^c, E_{7B}^c, E_{8(b-1)}^c) \longrightarrow 0$ as $n \longrightarrow \infty$ if $R < I(X_1, U_1, U_2; Y_3) - I(U_2; S|U_1)$.

It remains to show that the rate (4) is not altered if one restricts the random variables U_1 and U_2 to have their alphabet sizes limited as indicated in (6). This is done by invoking the support lemma [59, p. 310]. Fix a distribution μ of $(S, U_1, U_2, X_1, X_2, Y_2, Y_3)$ on $\mathcal{P}(\mathcal{S}\times\mathcal{U}_1\times\mathcal{U}_2\times\mathcal{X}_1\times\mathcal{X}_2\times\mathcal{Y}_2\times\mathcal{Y}_3)$ that has the form (5) and satisfies $\mathbb{E}[\varphi_i(X_i)] \leq \Gamma_i$, $i = 1, 2$.

To prove the bound (6a) on $|\mathcal{U}_1|$, note that we have

$$\begin{aligned} I_\mu(X_1; Y_2|S, U_1, X_2) &= I_\mu(X_1; Y_2, S, X_2|U_1) && \text{(A-12a)} \\ &= H_\mu(X_1|U_1) + H_\mu(Y_2, S, X_2|U_1) \\ &\quad - H_\mu(X_1, X_2, Y_2, S|U_1), && \text{(A-12b)} \end{aligned}$$

where (A-12a) follows since $(S, X_2) \leftrightarrow U_1 \leftrightarrow X_1$ under the distribution μ. Also, we have

$$\begin{aligned} &I_\mu(X_1, U_1, U_2; Y_3) - I_\mu(U_2; S|U_1) \\ &\quad = I_\mu(U_1; Y_3) + I_\mu(X_1, U_2; Y_3|U_1) - I_\mu(U_2; S|U_1) && \text{(A-13a)} \\ &\quad = H_\mu(Y_3) - H_\mu(S) - H_\mu(U_2|U_1) + H_\mu(U_2, S|U_1) \\ &\qquad + H_\mu(X_1, U_2|U_1) - H_\mu(X_1, U_2, Y_3|U_1). && \text{(A-13b)} \end{aligned}$$

Hence, it suffices to show that the following functionals of $\mu(S, U_1, U_2, X_1, X_2, Y_2, Y_3)$:

$$\begin{aligned} r_{s,x,x'}(\mu) &= \mu(s, x, x'), \quad \forall\, (s, x, x') \in \mathcal{S}\times\mathcal{X}_1\times\mathcal{X}_2 && \text{(A-14a)} \\ r_1(\mu) &= \int_u d_\mu(u)[H_\mu(X_1|u) + H_\mu(Y_2, S, X_2|u) \\ &\qquad - H_\mu(X_1, X_2, Y_2, S|u)] && \text{(A-14b)} \\ r_2(\mu) &= \int_u d_\mu(u)[H_\mu(X_1, U_2|u) - H_\mu(X_1, U_2, Y_3|u) \\ &\qquad - H_\mu(U_2|u) + H_\mu(U_2, S|u)] && \text{(A-14c)} \end{aligned}$$

can be preserved with another measure μ' that has the form (5). Observing that there is a total of $|\mathcal{S}||\mathcal{X}_1||\mathcal{X}_2| + 1$ functionals in (A-14), this is ensured by a standard application of the support lemma; and this shows that the cardinality of the alphabet of the auxiliary random variable U_1 can be limited as indicated in (6a) without altering the rate (4). We note that the inputs constraints for the source and the relay, which involve $\mu(S,U_1,U_2,X_1,X_2,Y_2,Y_3)$ only through its marginals over (S,U_1,U_2,X_2,Y_2,Y_3) and (S,U_1,U_2,X_1,Y_2,Y_3) respectively, are satisfied.

Once the alphabet of U_1 is fixed, we apply similar arguments to bound the alphabet of U_2, where this time $|\mathcal{S}||\mathcal{X}_1||\mathcal{X}_2|(|\mathcal{S}||\mathcal{X}_1||\mathcal{X}_2|+1)-1$ functionals must be satisfied in order to preserve the joint distribution of S, U_1, X_1, X_2, and one more functional to preserve

$$\begin{aligned}I_\mu(X_1,U_1,U_2;Y_3)-I_\mu(U_2;S|U_1)\\ =H_\mu(Y_3)-H_\mu(S)-H_\mu(U_1|U_2)+H_\mu(U_1,S|U_2)\\ +H_\mu(X_1,U_1|U_2)-H_\mu(X_1,U_1,Y_3|U_2)\end{aligned}\tag{A-15}$$

yielding the bound indicated in (6b).

B. Proof of Corollary 1

The proof combines rate splitting [52] and the techniques used in the proof of Theorem 1. As we already mentioned in the discussion following Corollary 1, we split the message W to be transmitted from the source node into two independent parts W_r and W_d; the relay forwards only the part W_r, at rate R_r, and the part W_d is sent directly to the destination, at rate R_d. The total rate is $R=R_r+R_d$. We transmit in $B+1$ blocks, each of length n. During each of the first B blocks, the source sends a message $w_i=(w_{r,i},w_{d,i})$, with $w_{r,i}\in[1,2^{nR_r}]$ and $w_{d,i}\in[1,2^{nR_d}]$ and $i=1,\dots,B$ denotes the index of the block. For convenience, we let $w_{r,B+1}=w_{d,1}=1$. For fixed n, the average rate $R\frac{B}{B+1}$ over $B+1$ blocks approaches R as $B\longrightarrow+\infty$.

Codebook Generation: Fix a measure $P_{S,U_1,U_2,U,X_1,X_2,Y_2,Y_3}$ satisfying (13) and $\mathbb{E}[\varphi_i(X_i)]\le\Gamma_i$, $i=1,2$. Fix $\epsilon>0$ and let

$$J=2^{n(I(U_2;S|U_1)+2\epsilon)}\tag{B-16a}$$
$$M_r=2^{n(R_r-2\epsilon)}\tag{B-16b}$$
$$M_d=2^{n(R_d-4\epsilon)}.\tag{B-16c}$$

1) We generate M_r i.i.d. codewords $\{\mathbf{u}_1(w_r')\}$ indexed by $w_r'=1,\dots,M_r$, each with i.i.d. components drawn according to P_{U_1}. For each $\mathbf{u}_1(w_r')$, we generate M_r i.i.d. codewords $\{\mathbf{u}(w_r',w_r)\}$ at the source indexed by $w_r=1,\dots,M_r$, and J auxiliary codewords $\{\mathbf{u}_2(w_r',j)\}$ at the relay indexed by $j=1,\dots,J$. The codewords $\mathbf{u}(w_r',w_r)$ and $\mathbf{u}_2(w_r',j)$ are with i.i.d. components given $\mathbf{u}_1(w_r')$ drawn according to $P_{U|U_1}$ and $P_{U_2|U_1}$, respectively.
2) For each $\mathbf{u}_1(w_r')$, for each $\mathbf{u}(w_r',w_r)$, we generate M_d i.i.d. codewords $\{\mathbf{x}_1(w_r',w_r,w_d)\}$ indexed by $w_d=1,\dots,M_d$, each with i.i.d. components given $(\mathbf{u}_1(w_r'),\mathbf{u}(w_r',w_r))$ drawn according to $P_{X_1|U_1,U}$.

Encoding: At the beginning of block i, let $w_i=(w_{r,i},w_{d,i})$ be the new message to be sent from the source and $w_{i-1}=(w_{r,i-1},w_{d,i-1})$ be the message sent in the previous block $i-1$.

At the beginning of block i, the relay has decoded $w_{r,i-1}$ correctly, and the source transmits $\mathbf{x}_1(w_{r,i-1},w_{r,i},w_{d,i})$. The relay searches for the smallest $j\in\{1,\dots,J\}$ such that $\mathbf{u}_2(w_{r,i-1},j)$ and $\mathbf{s}[i]$ are jointly typical given $\mathbf{u}_1(w_{r,i-1})$. Since the vectors $\mathbf{u}_2(w_{r,i-1},j)$ and $\mathbf{s}[i]$ are generated independently given $\mathbf{u}_1(w_{r,i-1})$ according to the memoryless distributions defined by the n-product of $P_{U_2|U_1}$ and the n-product of Q_S, respectively; and there are J sequences in the bin indexed by $w_{r,i-1}$, the probability that there is no such sequence $\mathbf{u}_2$ goes to zero as $n\longrightarrow+\infty$. Denote the found j by $j^\star=j(\mathbf{s}[i],w_{r,i-1})$. The relay then transmits a vector $\mathbf{x}_2(w_{r,i-1})$ with i.i.d. components conditionally given $(\mathbf{u}_1(w_{r,i-1}),\mathbf{u}_2(w_{r,i-1},j^\star),\mathbf{s}[i])$ drawn according to the marginal $P_{X_2|U_1,U_2,S}$ induced by (13). Using arguments similar to those in the proof of Theorem 1, it can be shown that the inputs $\mathbf{x}_1(w_{r,i-1},w_{r,i},w_{d,i})$ and $\mathbf{x}_2(w_{r,i-1})$ satisfy the input constraints.

Decoding: The decoding procedures at the source and the relay are as follows.

1) At the end of block i, the relay knows $w_{r,i-1}$ and declares that $\hat{w}_{r,i}$ is sent if there is a unique $\hat{w}_{r,i}$ such that $\mathbf{u}(w_{r,i-1},\hat{w}_{r,i})$, $\mathbf{y}_2[i]$ and $\mathbf{s}[i]$ are jointly typical given $\mathbf{u}_1(w_{r,i-1})$, $\mathbf{u}_2(w_{r,i-1},j^\star)$ and $\mathbf{x}_2(w_{r,i-1})$. One can show that the decoding error in this step is small for sufficiently large n if
$$R_r<I(U;Y_2|S,U_1,X_2).\tag{B-17}$$
2) At the end of the transmission, the destination has collected all the blocks of channel outputs $\mathbf{y}_3[1],\mathbf{y}_3[2],\dots,\mathbf{y}_3[B+1]$, and can then perform backward-decoding by first decoding $(w_{r,B},w_{d,B+1})$ from $\mathbf{y}_3[B+1]$.
First, it declares that the pair $(\hat{w}_{r,B},\hat{w}_{d,B+1})$ is sent if there is a unique pair $(\hat{w}_{r,B},\hat{w}_{d,B+1})$, with $\hat{w}_{r,B}\in\{1,\dots,M_r\}$ and $\hat{w}_{d,B+1}\in\{1,\dots,M_d\}$, there is $j_B\in\{1,\dots,J\}$, such that $\mathbf{u}_1(\hat{w}_{r,B})$, $\mathbf{u}_2(\hat{w}_{r,B},j_B)$, $\mathbf{u}(\hat{w}_{r,B},1)$, $\mathbf{x}_1(\hat{w}_{r,B},1,\hat{w}_{d,B+1})$, $\mathbf{y}_3[B+1]$ are jointly typical. One can show that the decoding error in this step is small for sufficiently large n if
$$\begin{aligned}R_d&<I(X_1;Y_3|U,U_1,U_2)\\ R_d&<I(X_1,U_2;Y_3|U,U_1)-I(U_2;S|U_1)\\ R_r+R_d&<I(X_1,U,U_1,U_2;Y_3)-I(U_2;S|U_1)\\ &\overset{(a)}{=}I(X_1,U_1,U_2;Y_3)-I(U_2;S|U_1)\end{aligned}\tag{B-18}$$
where in (a) we used the fact that $I(U;Y_3|U_1,U_2,X_1)=0$ under the distribution (13).
Next, for b ranging from B to 2, the destination knows $w_{r,b}$ and decodes $(w_{r,b-1},w_{d,b})$ based on the information received in block b. It declares that the pair $(\hat{w}_{r,b-1},\hat{w}_{d,b})$ is sent if there is a unique pair $(\hat{w}_{r,b-1},\hat{w}_{d,b})$, with $\hat{w}_{r,b-1}\in\{1,\dots,M_r\}$ and $\hat{w}_{d,b}\in\{1,\dots,M_d\}$, there is $j_{b-1}\in\{1,\dots,J\}$, such that $\mathbf{u}_1(\hat{w}_{r,b-1})$, $\mathbf{u}_2(\hat{w}_{r,b-1},j_{b-1})$, $\mathbf{u}(\hat{w}_{r,b-1},w_{r,b})$, $\mathbf{x}_1(\hat{w}_{r,b-1},w_{r,b},\hat{w}_{d,b})$, $\mathbf{y}_3[b]$ are jointly typical. One can show that the decoding error in this step is small for sufficiently large n if (B-18) is true.

It remains to show that the rate (12) is not altered if the sizes of the alphabets of the auxiliary random variables U, U_1 and U_2 are restricted as in (14). This can be easily done by following the steps in the proof of Theorem 1.

C. Proof of Theorem 2

Consider a sequence of $(\epsilon_n, n, R, \mathbf{\Gamma})$-codes with $\epsilon_n \to 0$ as $n \to +\infty$. We show that R must be less than or equal $R^{\mathrm{up}}(\mathbf{\Gamma})$. By Fano's inequality, we have

$$H\left(W|Y_3^n\right) \leq nR\epsilon_n + 1 \triangleq n\delta_n. \tag{C-19}$$

Thus

$$nR = H(W) \leq I\left(W; Y_3^n\right) + n\delta_n. \tag{C-20}$$

We upper bound $I\left(W; Y_3^n\right)$ as in the following lemma, the proof of which follows.

Lemma 1:

$$\text{i) } I\left(W; Y_3^n\right) \leq \sum_{i=1}^{n} I(X_{1,i}, X_{2,i}; Y_{3,i}|S_i) - I(S_i; X_{1,i}|Y_{3,i}) \tag{C-21a}$$

$$\text{ii) } I(W; Y_3^n) \leq \sum_{i=1}^{n} I(X_{1,i}; Y_{2,i}, Y_{3,i}|S_i, X_{2,i}). \tag{C-21b}$$

Proof: To simplify the notation, we use $S^i = (S_1, S_2, \ldots, S_i)$, $Y_k^i = (Y_{k,1}, Y_{k,2}, \ldots, Y_{k,i})$, $k = 2, 3$, and $X_j^i = (X_{j,1}, X_{j,2}, \ldots, X_{j,i})$, $j = 1, 2$.

We obtain the bound on $I\left(W; Y_3^n\right)$ given in i) as follows:

$$\begin{aligned}
&I(W; Y_3^n)\\
&= I\left(W, S^n; Y_3^n\right) - I\left(S^n; Y_3^n|W\right)\\
&= \sum_{i=1}^{n} I\left(W, S^n; Y_{3,i}|Y_3^{i-1}\right) - H(S^n|W) + H\left(S^n|W, Y_3^n\right)\\
&= \sum_{i=1}^{n}\Big[H\left(Y_{3,i}|Y_3^{i-1}\right) - H\left(Y_{3,i}|W, S^n, Y_3^{i-1}\right)\\
&\qquad - H(S_i) + H\left(S_i|W, Y_3^n, S^{i-1}\right)\Big]\\
&\overset{(a)}{\leq} \sum_{i=1}^{n}\Big[H(Y_{3,i}) - H(Y_{3,i}|X_{1,i}, X_{2,i}, S_i)\\
&\qquad - H(S_i) + H\left(S_i|W, Y_3^n, S^{i-1}, X_{1,i}\right)\Big]\\
&\overset{(b)}{\leq} \sum_{i=1}^{n}\Big[I(X_{1,i}, X_{2,i}, S_i; Y_{3,i}) - H(S_i) + H(S_i|X_{1,i}, Y_{3,i})\Big]\\
&= \sum_{i=1}^{n}[I(X_{1,i}, X_{2,i}, S_i; Y_{3,i}) - I(S_i; X_{1,i}, Y_{3,i})]\\
&= \sum_{i=1}^{n}[I(X_{1,i}, X_{2,i}; Y_{3,i}|S_i) - I(S_i; X_{1,i}|Y_{3,i})]
\end{aligned}$$

where (a) follows from $\left(W, S^n, Y_3^{i-1}\right) \leftrightarrow (X_{1,i}, X_{2,i}, S_i) \leftrightarrow Y_{3,i}$ (a Markov chain); and the fact that $X_{1,i}$ is a deterministic function of W; and (b) follows from the fact that conditioning reduces entropy.

We obtain the bound on $I\left(W; Y_3^n\right)$ given in ii) as follows:

$$\begin{aligned}
&I\left(W; Y_3^n\right) \leq I\left(W; Y_2^n, Y_3^n\right)\\
&= H(W) - H\left(W|Y_2^n, Y_3^n\right)\\
&\overset{(c)}{\leq} H(W|S^n) - H\left(W|Y_2^n, Y_3^n, S^n\right)\\
&= \sum_{i=1}^{n} I\left(W; Y_{2,i}, Y_{3,i}|Y_2^{i-1}, Y_3^{i-1}, S^n\right)\\
&\overset{(d)}{=} \sum_{i=1}^{n} I\left(W; Y_{2,i}, Y_{3,i}|Y_2^{i-1}, Y_3^{i-1}, S^n, X_{2,i}\right)\\
&= \sum_{i=1}^{n}\Big[H\left(Y_{2,i}, Y_{3,i}|Y_2^{i-1}, Y_3^{i-1}, S^n, X_{2,i}\right)\\
&\qquad - H\left(Y_{2,i}, Y_{3,i}|Y_2^{i-1}, Y_3^{i-1}, S^n, X_{2,i}, W\right)\Big]\\
&\overset{(e)}{=} \sum_{i=1}^{n}\Big[H\left(Y_{2,i}, Y_{3,i}|Y_2^{i-1}, Y_3^{i-1}, S^n, X_{2,i}\right)\\
&\qquad - H\left(Y_{2,i}, Y_{3,i}|Y_2^{i-1}, Y_3^{i-1}, S^n, X_{2,i}, W, X_{1,i}\right)\Big]\\
&\leq \sum_{i=1}^{n}\Big[H\left(Y_{2,i}, Y_{3,i}|S_i, X_{2,i}\right)\\
&\qquad - H\left(Y_{2,i}, Y_{3,i}|Y_2^{i-1}, Y_3^{i-1}, S^n, X_{2,i}, W, X_{1,i}\right)\Big]\\
&\overset{(f)}{=} \sum_{i=1}^{n}\Big[H(Y_{2,i}, Y_{3,i}|S_i, X_{2,i}) - H(Y_{2,i}, Y_{3,i}|S_i, X_{2,i}, X_{1,i})\Big]\\
&= \sum_{i=1}^{n} I(X_{1,i}; Y_{2,i}, Y_{3,i}|S_i, X_{2,i}),
\end{aligned}$$

where (c) follows from the fact that W and S^n are independent; and $H\left(W|Y_2^n, Y_3^n\right) \geq H\left(W|Y_2^n, Y_3^n, S^n\right)$; (d) follows from the fact that $X_{2,i}$ is a deterministic function of $\left(S^n, Y_2^{i-1}\right)$; (e) follows from the fact that $X_{1,i}$ is a deterministic function of W; and (f) follows from the fact that the channel is DM. □

Consider now the input constraints. By definition the code satisfies

$$\begin{aligned}
\frac{1}{n}\sum_{i=1}^{n} \varphi_1\left(\phi_1^n(w)_i\right) &\leq \Gamma_1\\
\frac{1}{n}\sum_{i=1}^{n} \varphi_2\left(\phi_{2,i}\left(y_2^{i-1}, s^n\right)\right) &\leq \Gamma_2
\end{aligned} \tag{C-22}$$

for $w \in \{1, \ldots, M\}$.

We start with the input constraint of the source. Since each codeword satisfies the input constraint, their average over w_1 also satisfies the input constraint. Thus, we have

$$\begin{aligned}
\Gamma_1 &\geq \sum_{w=1}^{M} P(w)\varphi_1^n\left(x_1^n(w)\right)\\
&= \sum_{w=1}^{M} P(w)\frac{1}{n}\sum_{i=1}^{n}\varphi_1(x_{1,i}(w))\\
&= \sum_{x_1^n}\sum_{w=1}^{M} P(w)P\left(x_1^n|w\right)\frac{1}{n}\sum_{i=1}^{n}\varphi_1(x_{1,i}(w))\\
&= \frac{1}{n}\sum_{i=1}^{n}\sum_{w=1}^{M}\sum_{x_1^n} P(w)P\left(x_1^n|w\right)\varphi_1(x_{1,i}(w))\\
&= \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}_{X_{1,i}}[\varphi_1(X_{1,i})].
\end{aligned} \tag{C-23}$$

Similarly, for the input constraint of the relay, we have

$$\begin{aligned}\Gamma_2 &\geq \sum_{w=1}^{M}\sum_{s^n} P(w)P(s^n)\frac{1}{n}\sum_{i=1}^{n}\varphi_2\left(\phi_{2,i}\left(y_2^{i-1},s^n\right)\right)\\ &= \sum_{w=1}^{M}\sum_{s^n,x_1^n,x_2^n,y_2^n} P(w)P(s^n)P\left(x_1^n|w\right)\\ &\quad\times P\left(x_2^n|s^n,x_1^n\right)P\left(y_2^n|s^n,x_1^n,x_2^n\right)\\ &\quad\times \frac{1}{n}\sum_{i=1}^{n}\varphi_2\left(\phi_{2,i}\left(y_2^{i-1},s^n\right)\right)\\ &= \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}_{X_{2,i}}[\varphi_2(X_{2,i})]. \end{aligned} \tag{C-24}$$

We introduce a random variable T which is uniformly distributed over $\{1,\ldots,n\}$. Set $S=S_T, X_1=X_{1,T}, X_2=X_{2,T}$, $Y_2=Y_{2,T}$, and $Y_3=Y_{3,T}$. We substitute T into the above bounds on the message rate and the input constraints. Considering the bounds given in Lemma 1, we obtain

$$\begin{aligned}&\frac{1}{n}\sum_{i=1}^{n} I(X_{1,i},X_{2,i};Y_{3,i}|S_i) - I(S_i;X_{1,i}|Y_{3,i})\\ &\quad= I(X_1,X_2;Y_3|S,T) - I(S;X_1|Y_3,T)\\ &\quad= I(X_1,X_2,S;Y_3|T) - I(S;X_1,Y_3|T) \end{aligned}\tag{C-25}$$

and

$$\frac{1}{n}\sum_{i=1}^{n} I(X_{1,i};Y_{2,i},Y_{3,i}|S_i,X_{2,i}) = I(X_1;Y_2,Y_3|S,X_2,T) \tag{C-26}$$

where the distribution on (T,S,X_1,X_2,Y_2,Y_3) from a given code is of the form

$$P_{T,S,X_1,X_2,Y_2,Y_3} = P_S P_T P_{X_1|T} P_{X_2|X_1,S,T} W_{Y_2,Y_3|S,X_1,X_2}. \tag{C-27}$$

Similarly, substituting T into the input constraints, we obtain

$$\begin{aligned}\Gamma_k &\geq \frac{1}{n}\sum_{i=1}^{n}\sum_{x_{k,i}} P_{X_k}(x_{k,i})\varphi_k(x_{k,i})\\ &= \sum_{t=1}^{n}\frac{1}{n}\sum_{x_k} P_{X_k|T}(x_k|t)\varphi_k(x_k)\\ &= \mathbb{E}[\varphi_k(X_k)], \qquad k=1,2. \end{aligned}\tag{C-28}$$

We now eliminate the variable T from (C-25) and (C-26) as follows. The right-hand side of (C-25) can be bounded as

$$\begin{aligned}&I(X_1,X_2,S;Y_3|T) - I(S;X_1,Y_3|T)\\ &\overset{(g)}{\leq} H(Y_3) - H(Y_3|X_1,X_2,S)\\ &\qquad - H(S|T) + H(S|X_1,Y_3,T)\\ &= I(X_1,X_2,S;Y_3) - H(S|T) + H(S|X_1,Y_3,T)\\ &\overset{(h)}{\leq} I(X_1,X_2,S;Y_3) - H(S) + H(S|X_1,Y_3)\\ &= I(X_1,X_2,S;Y_3) - I(S;X_1,Y_3)\\ &= I(X_1,X_2;Y_3|S) - I(S;X_1|Y_3) \end{aligned}\tag{C-29}$$

where (g) holds since $H(Y_3|T) \leq H(Y_3)$ and $H(Y_3|X_1,X_2,S,T) = H(Y_3|X_1,X_2,S)$ [by the Markovian relation $T \leftrightarrow (X_1,X_2,S) \leftrightarrow Y_3$]; and (h) holds since S is independent of T and $H(S|X_1,Y_3,T) \leq H(S|X_1,Y_3)$.

Similarly, the right-hand side of (C-26) can be bounded as

$$I(X_1;Y_2,Y_3|S,X_2,T) \leq I(X_1;Y_2,Y_3|S,X_2). \tag{C-30}$$

Finally, combining (C-20), (C-21a), (C-25), (C-29) on the one hand, and (C-20), (C-21b), (C-26), (C-30) on the other hand, we get

$$R \leq I(X_1,X_2;Y_3|S) - I(S;X_1|Y_3) \tag{C-31a}$$

$$R \leq I(X_1;Y_2,Y_3|S,X_2) \tag{C-31b}$$

where the distribution on (S,X_1,X_2,Y_2,Y_3), obtained by marginalizing (C-27) over the variable T, has the form given in (18) and satisfies $\mathbb{E}[\varphi_i(X_i)] \leq \Gamma_i$ for $i=1,2$.

We conclude that, for a given sequence of $(\epsilon_n,n,R,\mathbf{\Gamma})$ codes with ϵ_n going to zero as n goes to infinity, there exists a probability distribution of the form (18) such that the rate R satisfies (C-31) and the input constraints $\mathbb{E}[\varphi_i(X_i)] \leq \Gamma_i$, $i=1,2$, are satisfied. This completes the proof of Theorem 2.

D. Proof of Theorem 3

In this proof, we compute the lower bound in Theorem 1 using an appropriate jointly Gaussian distribution on S, X_1, U_1, U_2, X_2. The techniques used in this section rely strongly on those used in [16, proof of Theorem 6].

We first evaluate the second term of the minimization in (4) because this gives insights about the distribution that we should use to compute the lower bound. The second term of the minimization in (4) can be written as

$$\begin{aligned}&I(X_1,U_1,U_2;Y_3) - I(U_2;S|U_1)\\ &= I(X_1,U_1;Y_3) + I(U_2;Y_3|X_1,U_1) - I(U_2;S|X_1,U_1) \end{aligned}\tag{D-32}$$

which follows from the fact that $I(U_2;S|U_1) = I(U_2;S|U_1,X_1)$ for the considered distribution.

We first focus on the evaluation of the term $[I(U_2;Y_3|X_1,U_1) - I(U_2;S|X_1,U_1)]$. To evaluate it, we assume that X_1 is zero mean Gaussian with variance P_1, U_1 is zero mean Gaussian with variance $\bar{\theta}P_2$, and X_1 and U_1 are jointly Gaussian with $\mathbb{E}[U_1X_1] = \rho'_{12}\sqrt{\bar{\theta}P_1P_2}$, for some $\theta \in [0,1]$, $\rho'_{12} \in [-1,1]$. The random variables X_1 and U_1 are independent of S as shown by the distribution given in Theorem 1. We also consider

$$X_2 = U_1 + \tilde{X}_2 \tag{D-33}$$

where $\tilde{X}_2$ is zero mean Gaussian with variance θP_2, is independent of both X_1 and U_1, and is jointly Gaussian with S with $\mathbb{E}[\tilde{X}_2 S] = \rho'_{2s}\sqrt{\theta P_2 Q}$, for some $\rho'_{2s} \in [-1,1]$. Then, from (23) and (D-33), we can write Y_3 as

$$Y_3 = X_1 + U_1 + \tilde{X}_2 + S + Z_3. \tag{D-34}$$

Let $\hat{\tilde{X}}_2 = \mathbb{E}[\tilde{X}_2|S]$ be the optimal linear estimator of $\tilde{X}_2$ given S under minimum mean square error criterion, and X'_2 be the resulting estimation error. The estimator $\hat{\tilde{X}}_2$ and the estimation error X_2' are given by

$$\hat{\tilde{X}}_2 = \rho'_{2s}\sqrt{\frac{\theta P_2}{Q}}S \tag{D-35}$$

$$X_2' = \tilde{X}_2 - \hat{\tilde{X}}_2. \tag{D-36}$$

We can alternatively write Y_3 in (D-34) as

$$\begin{aligned} Y_3 &= (\tilde{X}_2 - \hat{\tilde{X}}_2) + \hat{\tilde{X}}_2 + X_1 + U_1 + S + Z_3 \\ &= X_2' + X_1 + U_1 + S' + Z_3 \end{aligned} \tag{D-37}$$

where

$$S' = \left(1 + \rho'_{2s}\sqrt{\frac{\theta P_2}{Q}}\right) S.$$

We now consider the following new channel output Y_3' given by

$$Y_3' := Y_3 - \mathbb{E}[Y_3|X_1, U_1] = X_2' + S' + Z_3. \tag{D-38}$$

This new channel output Y_3' is similar to the channel output considered in [3] because X_2' is independent of the state S'. Hence, the capacity of this new channel is achieved if we use an auxiliary random variable

$$U_2 = X_2' + \alpha S' \tag{D-39}$$

where α is Costa's parameter given by

$$\alpha = \frac{\mathbb{E}\left[X_2'^2\right]}{\mathbb{E}\left[X_2'^2\right] + \mathbb{E}\left[Z_3^2\right]} = \frac{\theta P_2\left(1 - \rho'^2_{2s}\right)}{\theta P_2\left(1 - \rho'^2_{2s}\right) + N_3}. \tag{D-40}$$

Then, we can easily show that

$$\begin{aligned} &[I(U_2; Y_3|X_1, U_1) - I(U_2; S|X_1, U_1)] \\ &\qquad = \left[I\left(U_2; Y_3'\right) - I\left(U_2; S'\right)\right]. \end{aligned}$$

The term $\left[I\left(U_2; Y_3'\right) - I\left(U_2; S'\right)\right]$ is maximized if U_2 is chosen as in (D-39). Thus, we obtain

$$\begin{aligned} &I(U_2; Y_3|X_1, U_1) - I(U_2; S|X_1, U_1) \\ &\quad = \frac{1}{2}\log\left(1 + \frac{\mathbb{E}\left[X_2'^2\right]}{N_3}\right) \\ &\quad = \frac{1}{2}\log\left(1 + \frac{\theta P_2\left(1 - \rho'^2_{2s}\right)}{N_3}\right). \end{aligned} \tag{D-41}$$

By substituting X_2' and S' in (D-39), we get

$$\begin{aligned} U_2 &= \tilde{X}_2 - \rho'_{2s}\sqrt{\frac{\theta P_2}{Q}}S + \alpha\left(1 + \rho'_{2s}\sqrt{\frac{\theta P_2}{Q}}\right)S \\ &= \tilde{X}_2 + \alpha_{\text{opt}} S \end{aligned} \tag{D-42}$$

where

$$\begin{aligned} \alpha_{\text{opt}} &= \left(1 + \rho'_{2s}\sqrt{\frac{\theta P_2}{Q}}\right)\alpha - \rho'_{2s}\sqrt{\frac{\theta P_2}{Q}} \\ &= \frac{\theta P_2\left(1 - \rho'^2_{2s}\right) - \rho'_{2s}\sqrt{\frac{\theta P_2}{Q}}N_3}{\theta P_2\left(1 - \rho'^2_{2s}\right) + N_3}. \end{aligned} \tag{D-43}$$

The term $I(X_1, U_1; Y_3)$ on the right-hand side of (D-32) can be computed as

$$\begin{aligned} &I(X_1, U_1; Y_3) \\ &\quad = h(Y_3) - h(Y_3|X_1, U_1) \\ &\quad = h(Y_3) - h(\tilde{X}_2 + S + Z_3|X_1, U_1) \\ &\quad \overset{(b)}{=} h(Y_3) - h(\tilde{X}_2 + S + Z_3) \\ &\quad = \frac{1}{2}\log\left(\frac{\mathbb{E}\left[(X_1 + X_2 + S)^2\right] + \mathbb{E}\left[Z_3^2\right]}{\mathbb{E}\left[(\tilde{X}_2 + S)^2\right] + \mathbb{E}\left[Z_3^2\right]}\right) \\ &\quad = \frac{1}{2}\log\left(1 + \frac{P_1 + \bar{\theta}P_2 + 2\rho'_{12}\sqrt{\bar{\theta}P_1P_2}}{\theta P_2 + Q + N_3 + 2\rho'_{2s}\sqrt{\theta P_2 Q}}\right) \end{aligned} \tag{D-44}$$

where (b) follows from the fact that $\tilde{X}_2$ and S are independent of (X_1, U_1). Then, by adding (D-41) and (D-44), we get the second term of the minimization in (26).

The first term of the minimization in (4) can be written as

$$\begin{aligned} &I(X_1; Y_2|S, U_1, X_2) \\ &\quad = h(Y_2|S, U_1, X_2) - h(Y_2|S, U_1, X_1, X_2) \\ &\quad = h(X_1 + Z_2|S, U_1, X_2) - h(Z_2) \\ &\quad \overset{(a)}{=} h(X_1 + Z_2|U_1) - h(Z_2) \\ &\quad = \frac{1}{2}\log\left(1 + \frac{P_1\left(1 - \rho'^2_{12}\right)}{N_2}\right) \end{aligned} \tag{D-45}$$

where (a) follows from the fact that X_1 and (S, X_2) are independent conditionally on U_1.

Finally, we obtain the rate on the right-hand side of (26) by maximization over all possible values of $\theta \in [0,1]$, $\rho'_{12} \in [-1,1]$, and $\rho'_{2s} \in [-1,1]$. Investigating the two terms of the minimization, we can easily see that it suffices to consider $\rho'_{12} \in [0,1]$ and $\rho'_{2s} \in [-1,0]$.

E. *Proof of Corollary 3*

Recall the outline after Corollary 3. We decompose the source input X_1 and the relay input X_2 as

$$X_1 = U + \tilde{X}_1 \tag{E-1}$$

$$X_2 = U_1 + \tilde{X}_2 \tag{E-2}$$

where U and $\tilde{X}_1$ are independent zero mean Gaussian random variables with variances $\bar{\gamma}P_1$ and γP_1, respectively, for some $\gamma \in [0,1]$; and U_1 and $\tilde{X}_2$ are independent zero mean Gaussian random variables with variances $\bar{\theta}P_2$ and θP_2, respectively, for some $\theta \in [0,1]$. Furthermore, $\tilde{X}_1$ is independent of all other variables; U and U_1 are correlated, with $\mathbb{E}[UU_1] = \rho'_{12}\sqrt{\bar{\theta}\bar{\gamma}P_1P_2}$ for some $\rho'_{12} \in [0,1]$, and are both independent of S; $\tilde{X}_2$ is independent of U, is correlated with S with $\mathbb{E}[\tilde{X}_2 S] = \rho'_{2s}\sqrt{\theta P_2 Q}$ for some $\rho'_{2s} \in [-1,0]$, and is obtained using a GDPC the auxiliary random variable of which is given by

$$U_2 = \tilde{X}_2 + \left[\alpha'\left(1+\rho'_{2s}\sqrt{\frac{\theta P_2}{Q}}\right) - \rho'_{2s}\sqrt{\frac{\theta P_2}{Q}}\right] S \quad \text{(E-3)}$$

for some $\alpha' \in \mathbb{R}$.

Let

$$\begin{aligned} T_0 &:= I(U; Y_2|S, U_1, X_2) && \text{(E-4)}\\ T_1 &:= T_0 + I(X_1; Y_3|U, U_1, U_2) && \text{(E-5)}\\ T_2 &:= T_1 + I(U_2; Y_3|U, U_1) - I(U_2; S|U_1) && \text{(E-6)}\\ T_3 &:= I(X_1, U_1, U_2; Y_3) - I(U_2; S|U_1). && \text{(E-7)} \end{aligned}$$

Also, define the following function and substitutions which we will use throughout the proof:

$$P'_2 := \theta P_2\left(1-\rho'^2_{2s}\right) \quad \text{(E-8)}$$

$$Q' := \left(\sqrt{Q} + \rho'_{2s}\sqrt{\theta P_2}\right)^2 \quad \text{(E-9)}$$

$$\Phi\left(\alpha', \theta, \rho'_{2s}\right) := \frac{P'_2 Q'(1-\alpha')^2}{P'_2 + \alpha'^2 Q'} \quad \text{(E-10)}$$

and let $\tilde{Y}_3 := \tilde{X}_2 + S + Z_3$.

i) The computation of the quantities T_0 and T_3 can be done along the lines of those for the corresponding quantities in the proof of Theorem 3. We obtain

$$T_0 = \frac{1}{2}\log\left(1 + \frac{\bar{\gamma}P_1\left(1-\rho'^2_{12}\right)}{N_2 + \gamma P_1}\right) \quad \text{(E-11)}$$

and T_3 as given by (34).

ii) Now, we compute T_1

$$\begin{aligned} &I(X_1; Y_3|U, U_1, U_2)\\ &= I(\tilde{X}_1; \tilde{X}_1 + \tilde{Y}_3|U, U_1, U_2)\\ &\overset{(a)}{=} I(\tilde{X}_1; \tilde{X}_1 + \tilde{Y}_3|U_2)\\ &\overset{(b)}{=} h(\tilde{X}_1 + \tilde{Y}_3|U_2) - h(\tilde{Y}_3|U_2)\\ &\overset{(c)}{=} \frac{1}{2}\log\left(1 + \frac{\mathbb{E}\left[\tilde{X}_1^2\right]}{\mathbb{E}\left[\tilde{Y}_3^2\right] - \mathbb{E}\left[\tilde{Y}_3\mathbb{E}\left[\tilde{Y}_3|U_2\right]\right]}\right)\\ &\overset{(d)}{=} \frac{1}{2}\log\left(1 + \frac{\gamma P_1}{N_3 + \Phi\left(\alpha', \theta, \rho'_{2s}\right)}\right) \qquad \text{(E-12)} \end{aligned}$$

where (a) holds since U and U_1 are independent of $\tilde{X}_1$, $\tilde{Y}_3$ and U_2; (b) holds since $\tilde{X}_1$ is independent of $\tilde{Y}_3$ and U_2; (c) holds since $\tilde{X}_1$, U_2 and $\tilde{Y}_3$ are jointly Gaussian, and (d) follows by straightforward algebra using the fact that $\mathbb{E}[\tilde{Y}_3|U_2] = \beta U_2$, with

$$\beta = \frac{P'_2 + \alpha' Q'}{P'_2 + \alpha'^2 Q'}. \quad \text{(E-13)}$$

Then, by adding (E-11) and (E-12), we get T_1 as given by (32).

iii) Finally, we compute T_2. It can be shown easily that

$$I(U_2; S|U_1) = \frac{1}{2}\log\left(\frac{P'_2 + \alpha'^2 Q'}{P'_2}\right). \quad \text{(E-14)}$$

Also, we have

$$\begin{aligned} &I(U_2; Y_3|U, U_1)\\ &= I(U_2; \tilde{X}_1 + \tilde{Y}_3|U, U_1)\\ &\overset{(e)}{=} I(U_2; \tilde{X}_1 + \tilde{Y}_3)\\ &= h(\tilde{X}_1 + \tilde{Y}_3) - h(\tilde{X}_1 + \tilde{Y}_3|U_2)\\ &\overset{(f)}{=} \frac{1}{2}\log\left(\frac{\mathbb{E}\left[\tilde{X}_1^2\right] + \mathbb{E}\left[\tilde{Y}_3^2\right]}{\mathbb{E}\left[\tilde{X}_1^2\right] + \mathbb{E}\left[\tilde{Y}_3^2\right] - \mathbb{E}[\tilde{Y}_3\mathbb{E}[\tilde{Y}_3|U_2]]}\right)\\ &\overset{(g)}{=} \frac{1}{2}\log\left(\frac{P'_2 + Q' + \gamma P_1 + N_3}{N_3 + \gamma P_1 + \Phi(\alpha', \theta, \rho'_{2s})}\right) \qquad \text{(E-15)} \end{aligned}$$

where (e) holds since U and U_1 are independent of U_2, $\tilde{X}_1$ and $\tilde{Y}_3$; (f) holds since $\tilde{X}_1$, U_2 and $\tilde{Y}_3$ are jointly Gaussian, and $\tilde{X}_1$ is independent of U_2 and $\tilde{Y}_3$; and (g) follows through straightforward algebra similar to in (E-12).

Adding T_1 [given by (32)] and (E-15) and subtracting (E-14), we get T_2 as given by (33).

F. Proof of Theorem 4

In this section, we use the upper bound for the DM case in Theorem 2 to compute the upper bound on the capacity of the state-dependent full-duplex Gaussian RC with informed relay.

Fix a joint distribution of X_1, X_2, S, Y_2, Y_3 of the form (18) satisfying

$$\begin{aligned} &\mathbb{E}\left[X_1^2\right] = \tilde{P}_1 \le P_1, \quad \mathbb{E}\left[X_2^2\right] = \tilde{P}_2 \le P_2\\ &\mathbb{E}[X_1X_2] = \sigma_{12}, \quad \mathbb{E}[X_2S] = \sigma_{2s}\\ &\mathbb{E}[X_1S] = 0. \end{aligned} \quad \text{(E-16)}$$

We will also use the correlation coefficients ρ_{12} and ρ_{2s} defined as

$$\rho_{12} = \frac{\sigma_{12}}{\sqrt{\tilde{P}_1\tilde{P}_2}}, \quad \rho_{2s} = \frac{\sigma_{2s}}{\sqrt{\tilde{P}_2 Q}}. \quad \text{(E-17)}$$

We first compute the first term in the minimization on the right-hand side of (17). Let $\mathbf{Y} = (X_1 + Z_2, X_1 + Z_3)^T$. We have

$$\begin{aligned} &I(X_1; Y_2, Y_3|S, X_2)\\ &= h(Y_2, Y_3|S, X_2) - h(Y_2, Y_3|S, X_1, X_2)\\ &= h(X_1 + Z_2, X_1 + Z_3|S, X_2) - h(Z_2, Z_3)\\ &\overset{(a)}{\le} \frac{1}{2}\log\left|\mathbb{E}\left[\left(\mathbf{Y} - \mathbb{E}[\mathbf{Y}|S, X_2]\right)\left(\mathbf{Y} - \mathbb{E}[\mathbf{Y}|S, X_2]\right)^T\right]\right|\\ &\quad - \frac{1}{2}\log(N_2N_3)\\ &= \frac{1}{2}\log\frac{\left|\mathbb{E}[\mathbf{Y}\mathbf{Y}^T] - \mathbb{E}\left[\mathbb{E}[\mathbf{Y}|S, X_2]\mathbb{E}[\mathbf{Y}|S, X_2]^T\right]\right|}{N_2N_3}\\ &\overset{(b)}{=} \frac{1}{2}\log\left(1 + \tilde{P}_1\left(1 - \frac{\rho_{12}^2}{1-\rho_{2s}^2}\right)\left(\frac{1}{N_2} + \frac{1}{N_3}\right)\right) \text{(E-18)} \end{aligned}$$

where $|\cdot|$ denotes the determinant operator; (a) follows from the fact that the conditional differential entropy $h(X_1{+}Z_2, X_1{+}Z_3|S, X_2)$ is maximized if (S, X_1, X_2, Z_2, Z_3) are jointly Gaussian; and (b) follows from the fact the vector

(S, X_1, X_2, Z_2, Z_3) is a jointly Gaussian vector and the MMSE estimator of $\mathbf{Y}$ given (S, X_2) is

$$\mathbb{E}[\mathbf{Y}|S, X_2] = \left(-\frac{\sigma_{12}\sigma_{2s}}{\tilde{P}_2 Q - \sigma_{2s}^2} S + \frac{\sigma_{12} Q}{\tilde{P}_2 Q - \sigma_{2s}^2} X_2\right) \cdot (1,1)^T. \quad \text{(E-19)}$$

We now compute the term $[I(X_1, X_2; Y_3|S) - I(X_1; S|Y_3)]$. We have

$$\begin{aligned} &I(X_1, X_2; Y_3|S) - I(X_1; S|Y_3) \\ &\quad = h(Y_3|S) - h(Y_3|X_1, X_2, S) - h(S|Y_3) + h(S|X_1, Y_3) \\ &\quad = h(Y_3) - h(S) + h(S|X_1, Y_3) - h(Z_3). \end{aligned} \quad \text{(E-20)}$$

For fixed second moments (E-16), we have

$$h(Y_3) \leq \frac{1}{2}\log(2\pi e)(\tilde{P}_1 + \tilde{P}_2 + 2\sigma_{12} + 2\sigma_{2s} + Q + N_3) \quad \text{(E-21)}$$

where equality is attained if Y_3 is Gaussian. Similarly, the term $h(S|X_1, Y_3)$ is maximized if (S, X_1, Y_3) are jointly Gaussian. Let $\hat{S}(X_1, Y_3) = \mathbb{E}[S|X_1, Y_3]$ be the minimum mean square error (MMSE) estimator of S given (X_1, Y_3), i.e.,

$$\begin{aligned} \hat{S}(X_1, Y_3) &= \mathbb{E}[S|X_1, X_2 + S + Z_3] \\ &= \gamma_1 X_1 + \gamma_2 (X_2 + S + Z_3) \end{aligned} \quad \text{(E-22)}$$

with

$$\begin{aligned} \gamma_1 &= -\frac{\sigma_{12}(Q + \sigma_{2s})}{\tilde{P}_1(\tilde{P}_2 + 2\sigma_{2s} + Q + N_3) - \sigma_{12}^2} \\ \gamma_2 &= \frac{\tilde{P}_1(Q + \sigma_{2s})}{\tilde{P}_1(\tilde{P}_2 + 2\sigma_{2s} + Q + N_3) - \sigma_{12}^2}. \end{aligned} \quad \text{(E-23)}$$

Then, we have

$$\begin{aligned} &h(S|X_1, Y_3) \\ &\quad = h(S - \hat{S}(X_1, Y_3)|X_1, Y_3) \\ &\quad \leq h(S - \gamma_1 X_1 - \gamma_2(X_2 + S + Z_3)) \\ &\quad = \frac{1}{2}\log(2\pi e)\mathbb{E}\left[(S - \gamma_1 X_1 - \gamma_2(X_2 + S + Z_3))^2\right] \\ &\quad = \frac{1}{2}\log\left((2\pi e)\frac{Q\tilde{P}_1\tilde{P}_2 + \tilde{P}_1 N_3 Q - \sigma_{2s}^2\tilde{P}_1 - \sigma_{12}^2 Q}{\tilde{P}_1(\tilde{P}_2 + 2\sigma_{2s} + Q + N_3) - \sigma_{12}^2}\right) \end{aligned} \quad \text{(E-24)}$$

where the inequality is attained with equality if S, X_1, X_2, Y_3 are jointly Gaussian. From (E-20), (E-21), and (E-24), we obtain

$$\begin{aligned} &I(X_1, X_2; Y_3|S) - I(X_1; S|Y_3) \\ &= \frac{1}{2}\log\left(\frac{\left(\tilde{P}_1 + \tilde{P}_2 + 2\sigma_{12} + 2\sigma_{2s} + Q + N_3\right)}{\left(\tilde{P}_1\tilde{P}_2 + 2\tilde{P}_1\sigma_{2s} + \tilde{P}_1 Q + \tilde{P}_1 N_3 - \sigma_{12}^2\right)} \right. \\ &\qquad \left. \times \frac{\left(Q\tilde{P}_1\tilde{P}_2 + \tilde{P}_1 N_3 Q - \sigma_{2s}^2\tilde{P}_1 - \sigma_{12}^2 Q\right)}{Q N_3}\right) \\ &= \frac{1}{2}\log\left(1 + \frac{\left(\sqrt{\tilde{P}_1} + \rho_{12}\sqrt{\tilde{P}_2}\right)^2}{\tilde{P}_2\left(1 - \rho_{12}^2 - \rho_{2s}^2\right) + \left(\sqrt{Q} + \rho_{2s}\sqrt{\tilde{P}_2}\right)^2 + N_3}\right) \\ &\quad + \frac{1}{2}\log\left(1 + \frac{\tilde{P}_2\left(1 - \rho_{12}^2 - \rho_{2s}^2\right)}{N_3}\right). \end{aligned} \quad \text{(E-25)}$$

For convenience, let us define the function $\Theta_1(\tilde{P}_1, \rho_{12}, \rho_{2s})$ as the right-hand side of (E-18) and the function $\Theta_2(\tilde{P}_1, \tilde{P}_2, \rho_{12}, \rho_{2s})$ as the right-hand side of (E-25). From the above analysis, the capacity of the channel is upper-bounded as

$$C \leq \max\min\{\Theta_1(\tilde{P}_1, \rho_{12}, \rho_{2s}), \Theta_2(\tilde{P}_1, \tilde{P}_2, \rho_{12}, \rho_{2s})\} \quad \text{(E-26)}$$

where the maximization is over all covariance matrices $\Lambda_{X_1, X_2, S, Z_2, Z_3}$ of (X_1, X_2, S, Z_2, Z_3)

$$\Lambda_{X_1, X_2, S, Z_2, Z_3} = \begin{pmatrix} \tilde{P}_1 & \rho_{12}\sqrt{\tilde{P}_1\tilde{P}_2} & 0 & 0 & 0 \\ \rho_{12}\sqrt{\tilde{P}_1\tilde{P}_2} & \tilde{P}_2 & \rho_{2s}\sqrt{\tilde{P}_2 Q} & 0 & 0 \\ 0 & \rho_{2s}\sqrt{\tilde{P}_2 Q} & Q & 0 & 0 \\ 0 & 0 & 0 & N_2 & 0 \\ 0 & 0 & 0 & 0 & N_3 \end{pmatrix} \quad \text{(E-27)}$$

that satisfy

$$\tilde{P}_1 \leq P_1 \quad \tilde{P}_2 \leq P_2 \quad \text{(E-28)}$$

and have nonnegative discriminant

$$Q\tilde{P}_1\tilde{P}_2 N_2 N_3\left(1 - \rho_{12}^2 - \rho_{2s}^2\right) \geq 0 \quad \text{(E-29)}$$

i.e., for $Q > 0$

$$\rho_{12}^2 + \rho_{2s}^2 \leq 1. \quad \text{(E-30)}$$

Furthermore, investigating $\Theta_1(\tilde{P}_1, \rho_{12}, \rho_{2s})$ and $\Theta_2(\tilde{P}_1, \tilde{P}_2, \rho_{12}, \rho_{2s})$, it can be seen that it suffices to consider $\rho_{12} \in [0, 1]$ and $\rho_{2s} \in [-1, 0]$ for the maximization in (E-26).

To complete the proof, we should show that $\Theta_1(\tilde{P}_1, \rho_{12}, \rho_{2s})$ and $\Theta_2(\tilde{P}_1, \tilde{P}_2, \rho_{12}, \rho_{2s})$ are maximized at $\tilde{P}_1 = P_1$ and $\tilde{P}_2 = P_2$. It is easy to show that $\Theta_1(\tilde{P}_1, \rho_{12}, \rho_{2s})$ and $\Theta_2(\tilde{P}_1, \tilde{P}_2, \rho_{12}, \rho_{2s})$ increase monotonically with $\tilde{P}_1$ for fixed ρ_{12}, ρ_{2s}, $\tilde{P}_2$. Then, we can replace $\tilde{P}_1$ with P_1 in both $\Theta_1(\tilde{P}_1, \rho_{12}, \rho_{2s})$ and $\Theta_2(\tilde{P}_1, \tilde{P}_2, \rho_{12}, \rho_{2s})$. To show that $\tilde{P}_2$ can be replaced by P_2, we use the following intuitive argument. Since the term $\Theta_1(P_1, \rho_{12}, \rho_{2s})$ does not depend on $\tilde{P}_2$ for given ρ_{12} and ρ_{2s}, it remains to show that $\tilde{P}_2$ can be replaced with P_2 in only the term $\Theta_2(\tilde{P}_1, \tilde{P}_2, \rho_{12}, \rho_{2s})$. The term $\Theta_2(\tilde{P}_1, \tilde{P}_2, \rho_{12}, \rho_{2s})$ is the sum rate of a two-user MAC with asymmetric CSI in which the informed encoder knows the message of the uninformed encoder [16, Th. 6]. Then, considering this MAC, it can be argued [16] that for the sum-rate to be maximized the informed encoder should use the entire power available, i.e., P_2. This concludes the proof of Theorem 4.

G. Proof of Observation 1

We first prove the first statement in Observation 1. Let us denote $N_2^\star$ as the right-hand side of (42). We have

$$\begin{aligned} R_{\mathrm{G}}^{\mathrm{lo}} \overset{(a)}{\geq} \min\Bigg\{ &\frac{1}{2}\log\left(1+\frac{P_1}{N_2}\right), \\ &\max_{-1\leq\rho'_{2s}\leq 0}\frac{1}{2}\log\left(1+\frac{P_1}{P_2+Q+N_3+2\rho'_{2s}\sqrt{P_2Q}}\right) \\ &+\frac{1}{2}\log\left(1+\frac{P_2\left(1-\rho'^2_{2s}\right)}{N_3}\right)\Bigg\} \\ \overset{(b)}{=} &\frac{1}{2}\log\left(1+\frac{P_1}{N_2}\right) \\ := &R_{\mathrm{DG}} \end{aligned} \tag{F-31}$$

where (a) follows by putting $\rho'_{12}=0$ and $\theta=1$ in (26), and (b) follows if $N_2 \geq N_2^{\star}$.

Then, it is easy to observe that

$$R_{\mathrm{DG}}^{\mathrm{up}} \leq R_{\mathrm{DG}}. \tag{F-32}$$

From (F-31) and (F-32), we get that

$$R_{\mathrm{DG}} \leq R_{\mathrm{G}}^{\mathrm{lo}} \leq C_{\mathrm{DG}} \leq R_{\mathrm{DG}}^{\mathrm{up}} \leq R_{\mathrm{DG}}. \tag{F-33}$$

Then, we can conclude that the lower bound and upper bound meet if $N_2 \geq N_2^{\star}$.

Let us now prove the second statement in Observation 1. If the pair (ρ_{12}, ρ_{2s}) that maximizes the upper bound in Corollary 4 satisfies the condition in (36) with equality, i.e., $\rho_{12}^2+\rho_{2s}^2=1$, then we choose $\varrho_{2s}=\rho_{2s}$, $\varrho_{12}=\rho_{12}$, and $\theta=\varrho_{2s}^2$ (i.e., $\bar{\theta}=\varrho_{12}^2$) in the lower bound (39) to achieve the upper bound, and thus obtain channel capacity in this case.

H. Proofs for Time Division Relaying

Proof of Proposition 1: Let $(X_{1,1}^{(1)}, X_{1,2}^{(1)}, \ldots, X_{1,\lfloor\nu n\rfloor}^{(1)})$ and $(X_{1,\lfloor\nu n\rfloor+1}^{(2)}, X_{1,\lfloor\nu n\rfloor+2}^{(2)}, \ldots, X_{1,n}^{(2)})$ be the transmitted sequences from the source during the relay-receive period and the relay-transmit period, respectively. The relay receives $Y_{2,1}, Y_{2,2}, \ldots, Y_{2,\lfloor\nu n\rfloor}$ during the relay-receive period and transmits a sequence $X_{2,\lfloor\nu n\rfloor+1}, X_{2,\lfloor\nu n\rfloor+2}, \ldots, X_{2,n}$ during the relay-transmit period. From Fano's inequality (C-20) and Lemma 1, we have the following:

$$\begin{aligned} nR \leq \min\Bigg\{ &\sum_{i=1}^{n} I(X_{1,i}; Y_{2,i}, Y_{3,i}|S_i, X_{2,i}), \\ &\sum_{i=1}^{n} I(X_{1,i}, X_{2,i}; Y_{3,i}|S_i) \\ &- I(X_{1,i}; S_i|Y_{3,i})\Bigg\} + n\delta_n. \end{aligned} \tag{G-34}$$

We now specialize this bound to the TD mode for which we have $X_{2,i}=0$ for $i \leq \lfloor\nu n\rfloor$ (as the relay does not transmit during the relay-receive period) and $Y_{2,i}=0$ for $i \geq \lfloor\nu n\rfloor+1$ (as the relay does not receive during the relay-transmit period). This gives

$$\begin{aligned} nR \leq \min\Bigg\{ &\sum_{i=1}^{\lfloor\nu n\rfloor} I\left(X_{1,i}^{(1)}; Y_{2,i}, Y_{3,i}^{(1)}|S_i^{(1)}, X_{2,i}=0\right) \\ &+ \sum_{i=\lfloor\nu n\rfloor+1}^{n} I\left(X_{1,i}^{(2)}; Y_{3,i}^{(2)}|S_i^{(2)}, X_{2,i}\right), \\ &\sum_{i=1}^{\lfloor\nu n\rfloor} I\left(X_{1,i}^{(1)}; Y_{3,i}^{(1)}|S_i^{(1)}, X_{2,i}=0\right) \\ &- I\left(X_{1,i}^{(1)}; S_i^{(1)}|Y_{3,i}^{(1)}\right) \\ &+ \sum_{i=\lfloor\nu n\rfloor+1}^{n} I\left(X_{1,i}^{(2)}, X_{2,i}; Y_{3,i}^{(2)}|S_i^{(2)}\right) \\ &- I\left(X_{1,i}^{(2)}; S_i^{(2)}|Y_{3,i}^{(2)}\right)\Bigg\} + n\delta_n. \end{aligned} \tag{G-35}$$

By letting $n \to \infty$ and using standard arguments [46], we get the single letter upper bound on capacity

$$\begin{aligned} C \leq \max\min\Bigg\{ &\nu I\left(X_1^{(1)}; Y_2, Y_3^{(1)}|S^{(1)}, X_2=0\right) \\ &+ \bar{\nu} I\left(X_1^{(2)}; Y_3^{(2)}|S^{(2)}, X_2\right), \\ &\nu I\left(X_1^{(1)}; Y_3^{(1)}|S^{(1)}, X_2=0\right) \\ &- \nu I\left(X_1^{(1)}; S^{(1)}|Y_3^{(1)}\right) \\ &+ \bar{\nu} I\left(X_1^{(2)}, X_2; Y_3^{(2)}|S^{(2)}\right) \\ &- \bar{\nu} I\left(X_1^{(2)}; S^{(2)}|Y_3^{(2)}\right)\Bigg\} \end{aligned} \tag{G-36}$$

where the maximization is over all joint distributions of the form

$$\begin{aligned} &Q_{S^{(1)}} P_{X_1^{(1)}} W_{Y_2, Y_3^{(1)}|X_1^{(1)}, S^{(1)}} \\ &\quad Q_{S^{(2)}} P_{X_1^{(2)}} P_{X_2|X_1^{(2)}, S^{(2)}} W_{Y_3^{(2)}|X_1^{(2)}, X_2, S^{(2)}}. \end{aligned} \tag{G-37}$$

The bound in (G-36) is the counterpart, to the TD mode, of the upper bound (17) for the full-duplex case. By closely following the arguments and the algebra used in the proof of Theorem 4, it can be shown that this bound is maximized by choosing $S^{(1)}$, $S^{(2)}$, $X_1^{(1)}$, $X_1^{(2)}$, X_2, Y_2, $Y_3^{(1)}$, $Y_3^{(2)}$ that are jointly Gaussian, with $X_1^{(1)}$ with power $P_1^{(1)}$ is independent of $S^{(1)}$, and $X_1^{(2)}$ and X_2 with power $P_1^{(2)}$ and P_2, respectively, are such that

$$\begin{aligned} \mathbb{E}\left[X_1^{(2)} X_2\right] &= \rho_{12}\sqrt{P_1^{(2)} P_2} \\ \mathbb{E}\left[X_1^{(2)} S^{(2)}\right] &= 0 \\ \mathbb{E}\left[X_2 S^{(2)}\right] &= \rho_{2s}\sqrt{P_2 Q^{(2)}}. \end{aligned}$$

Using this distribution, the evaluation of the right-hand side of (G-36) gives the right-hand side of (51).

Proof of Proposition 2: The proof follows by combining the technique of rate splitting [52] and the GDPC described in Section IV-A for the full-duplex mode. Rate splitting has the message W to be transmitted from the source node split into two independent parts: W_d transmitted directly to the destination at rate R_d, and W_r transmitted through the relay at rate R_r, with a total rate $R=R_r+R_d$.

The encoding and transmission scheme is as follows. During the relay-receive period, the source sends a Gaussian signal

$X_{1,i}^{(1)}$ which carries W_r only and is independently drawn with a random variable $X_1^{(1)} \sim \mathcal{N}(0, P_1^{(1)})$ which is independent of the channel state $S^{(1)}$. During the relay-transmit period, the source transmits a Gaussian signal $X_{1,i}^{(2)}$ which carries both W_r and W_d and is independently drawn with $X_1^{(2)} \sim \mathcal{N}(0, P_1^{(2)})$. During the relay-transmit period, the relay sends a Gaussian signal $X_{2,i}$ which carries W_r only and is given by

$$X_{2,i} = U_{1,i} + \tilde{X}_{2,i} \tag{G-38}$$

where $U_{1,i}$ is drawn with $U_1 \sim \mathcal{N}(0, \bar{\theta} P_2)$ and $\tilde{X}_{2,i}$ is obtained via a GDPC considering $S^{(2)}$ as noncausal channel state information during this period.

The random variables U_1 and $X_1^{(2)}$ are jointly Gaussian with $\mathbb{E}[X_1^{(2)} X_2] = \mathbb{E}[X_1^{(2)} U_1] = \rho'_{12}\sqrt{\bar{\theta} P_1^{(2)} P_2}$, and are both independent of the state $S^{(2)}$. For the GDPC, we use the following auxiliary random variable to generate the auxiliary codewords $U_{2,i}$:

$$U_2 = \tilde{X}_2 + \left[\alpha'\left(1 + \rho'_{2s}\sqrt{\frac{\theta P_2}{Q^{(2)}}}\right) - \rho'_{2s}\sqrt{\frac{\theta P_2}{Q^{(2)}}}\right] S^{(2)} \tag{G-39}$$

where $\tilde{X}_2 \sim \mathcal{N}(0, \theta P_2)$ is jointly Gaussian with $S^{(2)}$, with $\mathbb{E}[X_2 S^{(2)}] = \mathbb{E}[\tilde{X}_2 S_2^{(2)}] = \rho' \_2s\sqrt{\theta P_2 Q^{(2)}}$; and α' is a scale parameter. Thus, using the GDPC given by (G-39), $\tilde{X}_{2,i}$ is generated as

$$\tilde{X}_{2,i} = U_{2,i} - \left[\alpha'\left(1 + \rho'_{2s}\sqrt{\frac{\theta P_2}{Q^{(2)}}}\right) - \rho'_{2s}\sqrt{\frac{\theta P_2}{Q^{(2)}}}\right] S_i^{(2)} \tag{G-40}$$

where $U_{2,i}$ is independently drawn with U_2.

Furthermore, we let $X_{1,i}^{(2)} = \rho'_{12}\sqrt{P_1^{(2)}/\bar{\theta} P_2} U_{1,i} + \tilde{X}_{1,i}^{(2)}$, where $\tilde{X}_{1,i}^{(2)}$ is independently drawn with $\tilde{X}_1^{(2)} \sim \mathcal{N}(0, (1 - \rho'^2_{12}) P_1^{(2)})$, is independent of U_1, X_2, $S^{(2)}$, and carries W_d only.

For the decoding procedures at the source and the relay, we give simple arguments based on intuition (the rigorous decoding uses joint typicality arguments). Also, since all the random variables are i.i.d., we sometimes omit the time index. The relay subtracts out $S^{(1)}$ from the received Y_2 and then decodes W_r. Message W_r can be decoded correctly at the relay as long as

$$R_r < \frac{\nu}{2}\log\left(1 + \frac{P_1^{(1)}}{N_2}\right). \tag{G-41}$$

The destination jointly decodes W_r and W_d from $(Y_3^{(1)}, Y_3^{(2)})$. One can show that this can be done reliable as long as

$$R_d < \bar{\nu} I\left(X_1^{(2)}; Y_3^{(2)} | U_1, U_2\right) \tag{G-42}$$

$$R_d < \bar{\nu}\left[I\left(X_1^{(2)}, U_2; Y_3^{(2)} | U_1\right) - I\left(U_2; S^{(2)} | U_1\right)\right] \tag{G-43}$$

$$R_r + R_d < \nu I\left(X_1^{(1)}; Y_3^{(1)}\right) + \bar{\nu}\left[I\left(X_1^{(2)}, U_1, U_2; Y_3^{(2)}\right) - I\left(U_2; S^{(2)} | U_1\right)\right]. \tag{G-44}$$

Adding (G-41) and (G-42) on the one hand, and (G-41) and (G-43) on the other hand, and using (G-44), we obtain

$$R < \frac{\nu}{2}\log\left(1 + \frac{P_1^{(1)}}{N_2}\right) + \bar{\nu} I\left(X_1^{(2)}; Y_3^{(2)} | U_1, U_2\right) \tag{G-45}$$

$$R < \frac{\nu}{2}\log\left(1 + \frac{P_1^{(1)}}{N_2}\right) + \bar{\nu}\left[I\left(X_1^{(2)}, U_2; Y_3^{(2)} | U_1\right) - I\left(U_2; S^{(2)} | U_1\right)\right] \tag{G-46}$$

$$R < \nu I\left(X_1^{(1)}; Y_3^{(1)}\right) + \bar{\nu}\left[I\left(X_1^{(2)}, U_1, U_2; Y_3^{(2)}\right) - I\left(U_2; S^{(2)} | U_1\right)\right]. \tag{G-47}$$

The computation of the mutual information terms in (G-45)–(G-47) involves straightforward algebra which is very similar to that in the proofs of Theorem 3 in Appendix D and of Corollary 1 in Appendix E; and, so, we omit the details for brevity. More specifically, define

$$P'_2 := \theta P_2\left(1 - \rho'^2_{2s}\right) \quad Q'^{(2)} := \left(\sqrt{Q^{(2)}} + \rho'_{2s}\sqrt{\theta P_2}\right)^2.$$

also, recall $\Phi(\alpha', \theta, \rho'_{2s})$ as defined in (55). Then, we have the following.

The mutual information on the right-hand side of (G-45) can be computed as in (E-12) to obtain

$$I\left(X_1^{(2)}; Y_3^{(2)} | U_1, U_2\right) = \frac{1}{2}\log\left(1 + \frac{\left(1 - \rho'^2_{12}\right) P_1^{(2)}}{N_3 + \Phi(\alpha', \theta, \rho'_{2s})}\right). \tag{G-48}$$

The conditional mutual information difference on the right-hand side of (G-46) is similar to T_2 in Appendix E and it gives

$$I\left(X_1^{(2)}, U_2; Y_3^{(2)} | U_1\right) - I\left(U_2; S^{(2)} | U_1\right) = \frac{1}{2}\log\left(\frac{P'_2\left(P'_2 + Q'^{(2)} + N_3 + \left(1 - \rho'^2_{12}\right) P_1^{(2)}\right)}{P'_2 Q'^{(2)}(1-\alpha')^2 + N_3\left(P'_2 + \alpha'^2 Q'^{(2)}\right)}\right). \tag{G-49}$$

The evaluation of the term $[I(X_1^{(2)}, U_1, U_2; Y_3^{(2)}) - I(U_2; S^{(2)} | U_1)]$ is similar to that of (D-32) in Appendix D, and we obtain

$$I\left(X_1^{(2)}, U_1, U_2; Y_3^{(2)}\right) - I\left(U_2; S^{(2)} | U_1\right) = \frac{1}{2}\log\left(1 + \frac{P_1^{(2)} + \bar{\theta} P_2 + 2\rho'_{12}\sqrt{\bar{\theta} P_1^{(2)} P_2}}{\theta P_2 + Q^{(2)} + 2\rho'_{2s}\sqrt{\theta P_2 Q^{(2)}} + N_3}\right) + \frac{1}{2}\log\left(\frac{P'_2\left(P'_2 + Q'^{(2)} + N_3\right)}{P'_2 Q'^{(2)}(1-\alpha')^2 + N_3\left(P'_2 + \alpha'^2 Q'^{(2)}\right)}\right). \tag{G-50}$$

Also, it is easy to show that

$$I\left(X_1^{(1)}; Y_3^{(1)}\right) = \frac{1}{2}\log\left(1 + \frac{P_1^{(1)}}{N_3 + Q^{(1)}}\right). \tag{G-51}$$

Finally, we obtain (54a) using (G-45) and (G-48); we obtain (54b) using (G-46) and (G-49); and we obtain (54c) using (G-47), (G-50) and (G-51). This completes the proof.

REFERENCES

[1] C. E. Shannon, "Channels with side information at the transmitter," *IBM J. Res. Develop.*, vol. 2, pp. 289–293, Oct. 1958.

[2] S. I. Gel'fand and M. S. Pinsker, "Coding for channel with random parameters," *Probl. Control Inf. Theory*, vol. 9, pp. 19–31, 1980.

[3] M. H. M. Costa, "Writing on dirty paper," *IEEE Trans. Inf. Theory*, vol. IT-29, no. 3, pp. 439–441, May 1983.

[4] C. D. Heegard and A. E. Gamal, "On the capacity of computer memory with defects," *IEEE Trans. Inf. Theory*, vol. IT-29, no. 5, pp. 731–739, Sep. 1983.

[5] R. Zamir, S. Shamai (Shitz), and U. Erez, "Nested linear/lattice codes for structured multi-terminal binning," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1250–1276, Jun. 2002.

[6] U. Erez, S. Shamai (Shitz), and R. Zamir, "Capacity and lattice strategies for cancelling known interference," *IEEE Trans. Inf. Theory*, vol. 51, no. 11, pp. 3820–3833, Nov. 2005.

[7] G. Keshet, Y. Steinberg, and N. Merhav, "Channel coding in the presence of side information: Subject review," *Found. Trends Commun. Inf. Theory*, 2008.

[8] S. Gel'fand and M. Pinsker, "On Gaussian channel with random parameters," in *Proc. IEEE Int. Symp. Inf. Theory*, Tashkent, Uzbekistan, 1984, pp. 247–250.

[9] Y.-H. Kim, A. Sutivong, and S. Sigurjonsson, "Multiple user writing on dirty paper," in *Proc. IEEE Int. Symp. Inf. Theory*, Chicago, IL, Jun. 2004, p. 534.

[10] S. Kotagiri and J. N. Laneman, "Achievable rates for multiple access channels with state information known at one encoder," in *Proc. Allerton Conf. Commun., Control Comput.*, Monticello, IL, 2004.

[11] Y. Cemal and Y. Steinberg, "The multiple-access channel with partial state information at the encoders," *IEEE Trans. Inf. Theory*, vol. 51, no. 11, pp. 3992–4003, Nov. 2005.

[12] Y. Steinberg, "Coding for the degraded broadcast channel with random parameters, with causal and noncausal side information," *IEEE Trans. Inf. Theory*, vol. 51, no. 8, pp. 2867–2877, Aug. 2005.

[13] S. Kotagiri and J. Laneman, "Multiaccess channels with state known to one encoder: A case of degraded message sets," in *Proc. IEEE Int. Symp. Inf. Theory*, Nice, France, Jun. 2007, pp. 1566–1570.

[14] S. Kotagiri and J. N. Laneman, "Multiaccess channels with state known to some encoders and independent messages," *EURASIP J. Wireless Commun. Netw.*, 2008, article ID 450680.

[15] A. Somekh-Baruch, S. Shamai (Shitz), and S. Verdù, "Cooperative multiple access encoding with states available at one transmitter," in *Proc. IEEE Int. Symp. Inf. Theory*, Nice, France, Jun. 2007, pp. 1556–1560.

[16] A. Somekh-Baruch, S. Shamai (Shitz), and S. Verdù, "Cooperative multiple access encoding with states available at one transmitter," *IEEE Trans. Inf. Theory*, vol. 54, no. 10, pp. 4448–4469, Oct. 2008.

[17] T. Philosoph, A. Khisti, U. Erez, and R. Zamir, "Lattice strategies for the dirty multiple access channel," in *Proc. IEEE Int. Symp. Inf. Theory*, Nice, France, Jun. 2007, pp. 386–390.

[18] A. Zaidi and L. Vandendorpe, "Rate regions for the partially-cooperative relay-broadcast channel with non-causal side information," in *Proc. IEEE Int. Symp. Inf. Theory*, Nice, France, Jun. 2007, pp. 1246–1250.

[19] A. Zaidi and L. Vandendorpe, "Lower bounds on the capacity of the relay channel with states at the source," *EURASIP J. Wireless Commun. Netw.*, 2009, article ID 634296.

[20] A. Zaidi, S. Kotagiri, J. N. Laneman, and L. Vandendorpe, "Cooperative relaying with state at the relay," in *Proc. IEEE Inf. Theory Workshop*, Porto, Portugal, May 2008, pp. 139–143.

[21] A. Zaidi, S. Kotagiri, J. N. Laneman, and L. Vandendorpe, "Multiaccess channels with state known to one encoder: Another case of degraded message sets," in *Proc. IEEE Int. Symp. Inf. Theory*, Seoul, Korea, Jun. –Jul. 2009, pp. 2376–2380.

[22] Y. Steinberg and S. Shamai (Shitz), "Achievable rates for the broadcast channel with states known at the transmitter," in *Proc. IEEE Int. Symp. Inf. Theory*, Adelaide, Australia, Sep. 2005, pp. 2184–2188.

[23] P. Moulin and J. A. O'Sullivan, "Information-theoretic analysis of information hiding," *IEEE Trans. Inf. Theory*, vol. 49, no. 3, pp. 563–593, Mar. 2003.

[24] A. S. Cohen and A. Lapidoth, "The Gaussian watermarking game," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1639–1667, Jun. 2002.

[25] A. Somekh-Baruch and N. Merhav, "On the capacity game of public watermarking systems," *IEEE Trans. Inf. Theory*, vol. 50, no. 6, pp. 511–524, Jun. 2004.

[26] A. Zaidi, P. Piantanida, and P. Duhamel, "Broadcast- and MAC-aware coding strategies for multiple user information embedding," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2974–2992, Jun. 2007.

[27] A. Zaidi and L. Vandendorpe, "Coding schemes for relay-assisted information embedding," *IEEE Trans. Inf. Security Forens.*, vol. 4, no. 1, pp. 70–85, Jan. 2009.

[28] N. Merhav, "On random coding error exponents of watermarking systems," *IEEE Trans. Inf. Theory*, vol. 46, no. 2, pp. 420–430, Mar. 2000.

[29] A. V. Kusnetsov and B. S. Tsybakov, "Coding in a memory with defective cells," *Probl. Predach. Inf.*, vol. 10, pp. 52–60, 1974.

[30] G. Caire and S. Shamai (Shitz), "On the throughput of a multi-antenna Gaussian broadcast channel," *IEEE Trans. Inf. Theory*, vol. 49, no. 7, pp. 1691–1706, Jul. 2003.

[31] S. Viswanath, N. Jindal, and A. Goldsmith, "Duality, achievable rates and sum rate capacity of Gaussian MIMO broadcast channel," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2658–2668, Oct. 2003.

[32] P. Viswanath and D. N. Tse, "Sum capacity of the vector Gaussian MIMO broadcast channel," *IEEE Trans. Inf. Theory*, vol. 49, no. 8, pp. 1912–1921, Aug. 2003.

[33] H. Weingarten, Y. Steinberg, and S. Shamai (Shitz), "The capacity region of the Gaussian multiple-input multiple-output broadcast channel," *IEEE Trans. Inf. Theory*, vol. 52, no. 1, pp. 3936–3964, Jan. 2006.

[34] E. Biglieri, J. Proakis, and S. Shamai (Shitz), "Fading channels: Information-theoretic and communication aspects," *IEEE Trans. Inf. Theory*, vol. 44, no. 10, pp. 2619–2692, Oct. 1998.

[35] A. Host-Madsen, "On the capacity of cooperative diversity in slow fading channels," in *Proc. Annu. Allerton Conf. Commun. Control Comput.*, Monticello, IL, Oct. 2004.

[36] S. A. Jafar, G. J. Foschini, and A. Goldsmith, "Phantomnet: Exploring optimal multicellular multiple antenna systems," *EURASIP J. Wireless Commun. Netw.*, vol. 5, pp. 591–604, 2004.

[37] J. Mitola, "Cognitive radio: An integrated agent architecture for software defined radio," Ph.D. dissertation, KTH Royal Inst. Technol., Stockholm, Sweeden, 2000.

[38] N. Devroye, P. Mitran, and V. Tarokh, "Achievable rates in cognitive radio channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 5, pp. 1813–1827, May 2006.

[39] N. Devroye, P. Mitran, and V. Tarokh, "Limits on communications in a cognitive radio channel," *IEEE Commun. Mag.*, vol. 44, pp. 44–49, Jun. 2006.

[40] A. Jovicic and P. Viswanath, "Cognitive radio: An information-theoretic perspective," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2006, pp. 2413–2417.

[41] I. Marić, A. Goldsmith, G. Kramer, and S. Shamai (Shitz), "On the capacity of interference channels with one cooperating transmitter," *Eur. Trans. Telecommun.*, pp. 405–420, 2008.

[42] S. Sridharan, S. Vishwanath, S. A. Jafar, and S. Shamai (Shitz), "On the capacity of cognitive relay assisted Gaussian interference channel," in *Proc. IEEE Int. Symp. Inf. Theory*, Toronto, ON, Canada, Jul. 2008, pp. 549–553.

[43] O. Sahin and E. Erkip, "On achievable rates for interference relay channel with interference cancellation," in *Proc. Asilomar Conf. Signals Syst. Comput.*, Monterey, CA, Nov. 2007, pp. 805–809.

[44] A. Somekh-Baruch, S. Shamai (Shitz), and S. Verdù, "Cognitive interference channels with state information," in *Proc. IEEE Int. Symp. Inf. Theory*, Toronto, ON, Canada, Jul. 2008, pp. 1353–1357.

[45] N. Merhav and S. Shamai (Shitz), "Information rates subjected to state masking," in *Proc. IEEE Int. Symp. Inf. Theory*, Seattle, WA, 2006, pp. 1184–1188.

[46] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.

[47] T. M. Cover and A. E. Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inf. Theory*, vol. IT-25, no. 5, pp. 572–584, Sep. 1979.

[48] F. M. J. Willems, "Information theoretical results for the discrete memoryless multiple access channel," Doctor in de Wetenschappen Proefschrift dissertation, Katholieke Universiteit Leuven, Leuven, Belgium, 1982.

[49] A. B. Carleial, "Multiple-access channels with different generalized feedback signals," *IEEE Trans. Inf. Theory*, vol. IT-28, no. 6, pp. 841–850, Nov. 1982.

[50] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative strategies and capacity theorems for relay networks," *IEEE Trans. Inf. Theory*, vol. 51, no. 9, pp. 3037–3063, Sep. 2005.

[51] A. E. Gamal and M. Aref, "The capacity of the semideterministic relay channel," *IEEE Trans. Inf. Theory*, vol. IT-28, no. 3, pp. 536–536, May 1982.

[52] A. Høst-Madsen and J. Zhang, "Capacity bounds and power allocation for wireless relay channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 6, pp. 2020–2040, Jun. 2005.

[53] B. Schein and R. G. Gallager, "The Gaussian parallel relay network," in *Proc. IEEE Int. Symp. Inf. Theory*, Sorrento, Italy, Jun. 2000, pp. 22–22.

[54] M. Gastpar and M. Vetterli, "On the capacity of large Gaussian relay networks," *IEEE Trans. Inf. Theory*, vol. 51, no. 3, pp. 765–779, Mar. 2005.

[55] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.

[56] R. G. Gallager*, Information Theory and Reliable Communication*. New York: Wiley, 1968.

[57] G. Kramer, "Models and theory for relay channels with receive constraints," in *Proc. Annu. Allerton Conf. Commun. Control Comput.*, Monticello, IL, Sep.–Oct. 2004, pp. 1312–1321.

[58] G. Kramer, I. Marić, and R. D. Yates, "Cooperative communications," *Found. Trends Commun. Inf. Theory*, 2006, Katholieke Universiteit Leuven.

[59] I. Csiszár and J. Körner*, Information Theory: Coding Theorems for Discrete Memoryless Systems*. London, U.K.: Academic, 1981.

Abdellatif Zaidi (M'09) received the B.S. degree in electrical engineering from École Nationale Supérieure de Techniques Avancés, ENSTA ParisTech, Paris, France, in 2002 and the M.Sc. and Ph.D. degrees in electrical engineering from École Nationale Supérieure des Télécommunications, TELECOM ParisTech, Paris, France, in 2002 and 2005, respectively.

From December 2002 to December 2005, he was with the Communications an Electronics Department, TELECOM ParisTech, and the Signals and Systems Laboratory, CNRS/Supélec, France, pursuing his Ph.D. degree. Currently, he is with École Polytechnique de Louvain, Université Catholique de Louvain, Louvain-la-Neuve, Belgium, working as a Research Assistant. He was a Visiting Researcher at the University of Notre Dame, Notre Dame, IN, during fall 2007 and Spring 2008. His research interests cover a broad range of topics from signal processing for communication and multiuser information theory. Of particular interest are the problems of coding for side-informed channels, secure communication, coding and interference mitigation in multiuser channels, and relaying problems and cooperative communication with application to sensor networking and *ad hoc* wireless networks.

Shiva Prasad Kotagiri received the M.S. degree in electrical engineering from the University of Missouri, Rolla, in 2003 and the Ph.D. degree in electrical engineering from the University of Notre Dame, Notre Dame, IN, in 2008.

He is a Senior Design Engineer in the SERDES Technology group at Xilinx Inc., San Jose, CA. His research interests include information theory, high-speed wired communications, wireless communications, and signal processing.

Dr. Kotagiri was awarded the Center for Applied Mathematics (CAM) Fellowship of the University of Notre Dame in 2006. He received the Kaneb Graduate Student Award for Excellence in Teaching in 2006.

J. Nicholas Laneman (S'94–M'02–SM'07) received B.S. degrees (*summa cum laude*) in electrical engineering and in computer science from Washington University in St. Louis, St. Louis, MO, in 1995 and the S.M. and Ph.D. degrees in electrical engineering from the Massachusetts Institute of Technology (MIT), Cambridge, in 1997 and 2002, respectively.

Since 2002, he has been on the faculty of the Department of Electrical Engineering, University of Notre Dame, Notre Dame, IN, where he is currently an Associate Professor. His current research interests are in communications architecture—a combination of information theory, signal processing for communications, network protocols, and hardware design— specifically for wireless systems. From 1995 to 2002, he was affiliated with the Department of Electrical Engineering and Computer Science and the Research Laboratory of Electronics, MIT, where he held a National Science Foundation Graduate Research Fellowship and served as both a Research and Teaching Assistant. During 1998 and 1999 he was also with Lucent Technologies, Bell Laboratories, Murray Hill, NJ, both as a Member of the Technical Staff and as a Consultant, where he developed robust source and channel coding methods for digital audio broadcasting. His industrial interactions have led to five U.S. patents.

Dr. Laneman received a PECASE and NSF CAREER award in 2006, an ORAU Ralph E. Powe Junior Faculty Enhancement Award in 2003, and the MIT EECS Harold L. Hazen Teaching Award in 2001. He is currently an Associate Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS and has served as a Guest Editor of Special Issues in the IEEE TRANSACTIONS ON INFORMATION THEORY and the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS.

Luc Vandendorpe (M'93–SM'99–F'06) was born in Mouscron, Belgium, in 1962. He received the electrical engineering degree (*summa cum laude*) and the Ph.D. degree from the Université catholique de Louvain (UCL) Louvain-la-Neuve, Belgium, in 1985 and 1991, respectively.

Since 1985, he has been with the Communications and Remote Sensing Laboratory, UCL. In 1992, he was a Visiting Scientist and Research Fellow with the Telecommunications and Traffic Control Systems Group of the Delft Technical University, The Netherlands. Currently, he is a Professor and Head of the Electrical Engineering Department of the UCL. His main interests are in digital communication systems: equalization, joint detection synchronization for OFDM (multicarrier), MIMO (distributed MIMO), and turbo-based communications systems (UMTS, xDSL, WLAN, etc.), and in localization.

Dr. Vandendorpe was a corecipient of the Biennal Alcatel-Bell Award from the Belgian NSF in 1990. In 2000, he was a corecipient (with J. Louveaux and F. Deryck) of the Biennal Siemens Award from the Belgian NSF. He is or has been a TPC member for several IEEE conferences. He was the Co-Technical Chair (with P. Duhamel) for IEEE ICASSP 2006. He was an Associate Editor of the IEEE TRANSACTIONS ON COMMUNICATIONS FOR SYNCHRONIZATION AND EQUALIZATION between 2000 and 2002, the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS between 2003 and 2005, and the IEEE TRANSACTIONS ON SIGNAL PROCESSING from 2006 to 2008. He was Chair of the IEEE Benelux joint chapter on Communications and Vehicular Technology between 1999 and 2003. He was an elected member of the Sensor Array and Multichannel Signal Processing Committee of the Signal Processing Society from 2006 to 2008. Currently, he is an elected member of the Signal Processing for Communications Committee and Editor-in-Chief of the *Eurasip Journal on Wireless Communications and Networks*.