The Capacity Region of the Fading Interference Channel With a Relay in the Strong Interference Regime

Ron Dabora, *Member, IEEE*

***Abstract*—The interference channel with a relay (ICR) is the fundamental building block of cooperation in wireless networks where there are multiple communicating pairs interfering with each other. This paper considers ICRs in which the links are subject to i.i.d. fading, and each node has channel state information (CSI) only on its incoming links (receive CSI). Two channel models are considered: phase fading and Rayleigh fading. Strong interference conditions are derived for the case where the links from the sources to the relay are good in the sense that the achievable region for decoding both messages at the relay contains the maximal achievable region at the destinations. This leads to the characterization of the capacity region for such scenarios. This is the first time the capacity region of the ICR is characterized for a nondegraded, noncognitive scenario, with a causal relay when all links are active.**

***Index Terms*—Capacity, decode-and-forward, fading, interference channels (ICs), network information theory, relaying.**

I. Introduction

ONE of the main challenges in the design of wireless networks is coping with interference. Node cooperation is one of the key approaches to interference management in future wireless networks. In order to successfully integrate cooperation into the network design, it is essential to analyze the basic "building block" of cooperation in wireless networks which have interference. This building block is obtained by combining two fundamental networks: the interference channel (IC) [1] and the relay channel [2]. The combination of these two networks is a five-node network which consists of two communicating pairs and a relay node. The relay receives a combination of the transmissions from both sources and its signal is received at both destinations. Its role is to assist communication of both pairs. This channel model, referred to as the interference channel with a relay (ICR), was considered in several works (see [3] and references therein).

Relaying in the presence of interference is fundamentally different from relaying in the classic three-node relay channel of [4, Sec. III-A]: in the classic relay channel, there is no interference since all received signals carry information only on the desired message. In the ICR, when the relay forwards desired information to its destination, it also increases the interference at the second destination. Thus, when increasing the rate to one destination, the relay may end up reducing the rate to the second destination. On the other hand, in some situations, increasing the interference is optimal as it facilitates interference cancellation at the assisted receiver [5]. Therefore, in order to apply cooperation in the presence of interference, it is necessary to understand how interference and cooperation affect each other. As studying the classic relay channel does not shed light on the interaction between cooperation and interference, the need to understand this interaction provides a strong motivation for studying the ICR. Indeed, the ICR has attracted significant attention in the last years, and inner bounds [3], [5], [6] as well as outer bounds [7], [8] were derived for this scenario. Different variations such as cognition at the sources and/or the relay, as well as relaying over orthogonal links, have also been considered (see [3], [9] and references therein).

In this paper, we study ICRs when the links are subject to i.i.d. phase fading (PF-ICRs), as well as i.i.d. Rayleigh fading (RF-ICRs). Phase fading models apply to high-speed microwave communications over time-invariant channels where the oscillators' phase noise is a key impairment. Phase fading is also the major impairment in communication systems that employ orthogonal frequency division multiplexing [10], as well as in some applications of naval communications. Phase fading channel models also apply to systems which use dithering to decorrelate signals [11]. Rayleigh fading models are very common in wireless communications and apply to mobile communications in the presence of multiple scatterers without line-of-sight (e.g., dense urban environments) [12]. The key similarity between the two models is the uniformly distributed phase of the fading coefficients. The two fading models differ in the behavior of the fading magnitude, which is fixed for phase fading but varies following the Rayleigh distribution in Rayleigh fading.

The body of work on cooperative communications is vast and we refer to [4] and [13] for a comprehensive reference list. In particular, for cooperative multiuser scenarios, phase-fading models were considered for multiple-access-relay channels (MARCs) [4], [14] as well as for relay channels [4], broadcast-relay channels (BRCs) [4], and ICs [15]. In [15], Rx and Tx cooperation was considered for quasi-static phase fading ICs with two clustered transmitters and two clustered receivers (i.e.,

Manuscript received May 31, 2010; revised January 24, 2012; accepted April 10, 2012. Date of publication May 25, 2012; date of current version July 10, 2012. This work was supported in part by the European Commission's Marie Curie IRG Fellowship PIRG05-GA-2009-246657 under the Seventh Framework Programme. This paper was presented in part at the IEEE Information Theory Workshop, Dublin, Ireland, August 2011.

The author is with the Department of Electrical and Computer Engineering, Ben-Gurion University, 84105, Be'er Sheva, Israel (e-mail: ron@ee.bgu.ac.il).

Communicated by M. Gastpar, Associate Editor for Shannon Theory.

Digital Object Identifier 10.1109/TIT.2012.2201371

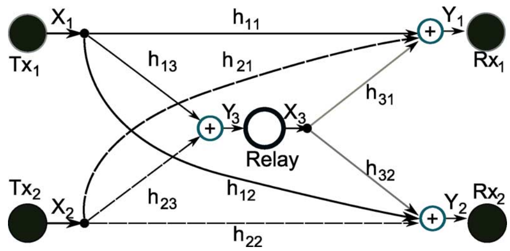

Fig. 1. ICR—a schematic layout. Additive noises are not depicted.

cooperation links are static), and full channel state information (CSI) available at all nodes.

Rayleigh fading were considered for relay channels in [16], [17] and [18]. Parallel relaying for MARCs and BRCs was considered in [19] for block fading channels in asymptotic regimes. In [20], an achievable region was derived for the three-user multiple-access channel (MAC) with cooperating transmitters, with CSI available at the transmitters (Tx-CSI) and at the receiver (Rx-CSI), and in [21], power control for i.i.d. fading cooperative MACs with Tx and Rx CSI was considered. In the context of ICs, the ergodic capacity of ICs subject stationary, ergodic (but not necessarily Gaussian) fading, assuming the transmitters and receivers know the instantaneous channel gains of all users, was obtained in [22].

A. Main Contributions

We characterize the capacity region for fast-fading ICRs in which all links are active (i.e., the channel output at each receiver is a combination of the transmissions from both sources and from the relay, and the channel output at the relay is a combination of the channel inputs from both sources), all transmissions share the same bandwidth, the nodes are full duplex, the relay operation is strictly causal, all signal-to-noise ratios (SNRs) are finite, and the channel is not degraded. Only receive CSI is assumed. The capacity regions are characterized for scenarios in which relay reception is "good" in the sense that the achievable region for decoding both messages at the relay contains the maximal achievable region for decoding the messages at the destinations. The analysis separately treats two interference regimes: "very strong interference" (VSI) and "strong interference" (SI). The results characterize the optimal transmission strategy for each channel configuration and identify the best performance that can be obtained. Specifically,

1) We obtain the capacity regions of PF-ICRs and of RF-ICRs, with Rx-CSI, under VSI. VSI was originally considered for ICs by Carleial [23]. When VSI occurs in ICs, each pair can communicate at a rate equal to its point-to-point (PtP) interference-free capacity. However, in the ICR, the situation is different as the maximal rate for each pair in the absence of interference depends on the relay, and in general, the relay may not be able to simultaneously provide the maximal assistance to each pair. We show how this is resolved in Section IV.
2) We obtain the capacity regions of PF-ICRs and of RF-ICRs, with Rx-CSI, under SI. The concept of SI was initially introduced for ICs in [24]. In interference networks, SI occurs when decoding both messages at each receiver does not constrain the rates of the desired information. In the SI regime, the capacity region of the IC is obtained as the capacity region of the corresponding compound MAC. SI for ICRs was investigated only in the cognitive setup [3], where the relay knows noncausally the sources' signals. In this study, we derive SI conditions for ICRs with *a causal* relay. The derivation of the SI conditions for PF-ICRs and RF-ICRs is focused on how to handle the dual role of the relay signal: sending both information and interference embedded in a single codeword. This is considered in Section V.

Our results demonstrate that a single relay can be simultaneously optimal for several pairs. This implies that when introducing relays into fading wireless networks (subject to the assumptions detailed previously), a few relays can help several pairs optimally, in the sense that each pair observes an interference-free relay channel with a dedicated relay (in SI an additional sum-rate constraint exists). This supports incorporating relays into wireless networks.

The rest of this paper is organized as follows: In Section II, the model and notations are presented, and in Section III relevant preliminary results are reviewed. In Section IV, the capacity regions for VSI are derived for PF and for RF ICRs, and in Section V the capacity regions under SI are derived for the two fading models. Concluding remarks are provided in Section VI.

II. Notations and Model

In the following, we denote random variables with upper case letters, e.g., X, Y, and their realizations with lower case letters, x, y. A random variable (RV) X takes values in the set of complex numbers $\mathfrak{C}$. We use $f_X(x)$ to denote the probability density function (p.d.f.) of a continuous RV X on $\mathfrak{C}$. For brevity, we may omit the subscript X when it is the upper case version of the realization letter x. We denote vectors with boldface letters, e.g., $\mathbf{x}, \mathbf{y}$; the ith element of a vector $\mathbf{x}$ is denoted with x_i and we use x_i^j, where $i \leq j$ to denote the vector $(x_i, x_{i+1}, \ldots, x_{j-1}, x_j)$; x^j is a short-form notation for x_1^j, and $\mathbf{x} \equiv x^n$. We use $I(\cdot;\cdot)$ to denote the mutual information between two random variables, as defined in [25, Ch. 2]. We denote with $A_\epsilon^{(n)}(X)$ the set of jointly typical sequences with respect to $f_X(x)$ as defined in [25, Chs. 3, 8]. We use $\mathcal{CN}(a, \sigma^2)$ to denote a proper, circularly symmetric, complex Normal distribution with mean a and variance σ^2 [26]. We use $X \perp\!\!\!\perp Y$ to denote that X is statistically independent of Y, $\mathfrak{R}$ is used to denote the set of real numbers, and double-stroke letters are used to denote matrices, e.g., $\mathbb{Q}$, with the exception that $\mathbb{E}\{\cdot\}$ is used to denote stochastic expectation. Let $\mathbb{I}_k$ denote the $k \times k$ identity matrix. Finally, $(\cdot)^H$ denotes Hermitian conjugation, and $X \sim f_X(x)$ means "X is a RV distributed according to $f_X(x)$", and $E_1(x)$ is the exponential integral: $E_1(x) = \int_x^\infty \frac{1}{q} e^{-q} dq$.

In fading ICRs, the received signals at time i at Rx_1, Rx_2 and the relay are given by

$$Y_{1,i} = H_{11,i}X_{1,i} + H_{21,i}X_{2,i} + H_{31,i}X_{3,i} + Z_{1,i} \quad \text{(1a)}$$
$$Y_{2,i} = H_{12,i}X_{1,i} + H_{22,i}X_{2,i} + H_{32,i}X_{3,i} + Z_{2,i} \quad \text{(1b)}$$
$$Y_{3,i} = H_{13,i}X_{1,i} + H_{23,i}X_{2,i} + Z_{3,i} \quad \text{(1c)}$$

$i = 1, 2, \ldots, n$, where Z_1, Z_2, and Z_3 are i.i.d., circularly symmetric, complex Normal RVs, $\mathcal{CN}(0,1)$. The channel input signals are subject to per-symbol average power constraints: $\mathbb{E}\{|X_k|^2\} \le P_k$, $k = 1, 2, 3$.

Next, we define the statistical model for each fading type:

- *Phase Fading Channels:* The channel coefficients are given by $H_{lk,i} = a_{lk} e^{j\Theta_{lk,i}}$, $a_{lk} \in \mathfrak{R}_+$ are nonnegative constants representing the attenuation from node l to node k and $\Theta_{lk,i}$ are uniformly distributed over $[0, 2\pi)$, i.i.d., and independent of each other and of the additive noises $\{Z_k, k = 1, 2, 3\}$.
- *Rayleigh Fading Channels:* The channel coefficients are given by $H_{lk,i} = a_{lk} U_{lk,i}$, $a_{lk} \in \mathfrak{R}_+$ are nonnegative constants representing the mean attenuation from node l to node k, and $U_{lk,i}$ are circularly symmetric, complex Normal RVs, $U_{lk,i} \sim \mathcal{CN}(0,1)$, i.i.d., and independent of each other and of the additive noises $\{Z_k, k = 1, 2, 3\}$.

In both models, the values of a_{kl} are fixed and known at all users, therefore we can set $a_{11} = a_{22} = 1$. Note that the magnitude of the phase fading process is constant, $|H_{lk,i}| = a_{lk}$, but for Rayleigh fading the fading magnitude varies between different time instances.

In the following, it is assumed that each destination receiver as well as the relay knows all instantaneous channel coefficients from Tx_1, Tx_2 and the relay to itself. This is referred to as Rx-CSI. Note that each receiver does not have CSI on the links arriving to the other receiver or to the relay, and that the relay does not have CSI on the links arriving to the receivers. It is assumed that the sources and the relay do not know the channel coefficients on their outgoing links (no Tx-CSI). The impact of the lack of Tx-CSI will be highlighted in Sections III-A and IV-A. We represent the CSI at receiver k with $\tilde{H}_k = (H_{1k}, H_{2k}, H_{3k})$, $k = 1, 2$, and at the relay with $\tilde{H}_3 = (H_{13}, H_{23})$. We let $\mathfrak{H}_1 = \mathfrak{H}_2 = \mathfrak{C}^3$ and $\mathfrak{H}_3 = \mathfrak{C}^2$ be the corresponding domains for the channel state vectors, and we define $\tilde{H} \triangleq \Big\{H_{lk} : (l,k) \in \{(1,1),(1,2),(1,3),(2,1),(2,2),(2,3),(3,1),(3,2)\}\Big\}$ and $\tilde{U} \triangleq \Big\{U_{lk} : (l,k) \in \{(1,1),(1,2),(1,3),(2,1),(2,2),(2,3),(3,1),(3,2)\}\Big\}$.

Finally, we provide the definitions of the main terminology:

Definition 1: An (R_1, R_2, n) code for the ICR consists of two message sets $\mathcal{M}_k \triangleq \{1, 2, \ldots, 2^{nR_k}\}$, $k = 1, 2$, two encoders at the sources, e_1, e_2, and two decoders at the destinations, g_1, g_2: $e_k : \mathcal{M}_k \mapsto \mathfrak{C}^n$, $g_k : \mathfrak{H}_k^n \times \mathfrak{C}^n \mapsto \mathcal{M}_k$, $k = 1, 2$. At the relay, there is a causal encoder, $x_{3,i} = e_i^r\left(y_{3,1}^{i-1}, h_{13,1}^{i-1}, h_{23,1}^{i-1}\right) \in \mathfrak{C}$, $i = 1, 2, \ldots, n$.

Observe that the relay operates in full duplex mode and the relay encoding is strictly causal. Also observe that there are no common messages to be decoded at both receivers. The sensitivity of the results to these assumptions will be discussed in Section V-A.

Definition 2: The average probability of error of an (R_1, R_2, n) code is defined as $P_e^{(n)} \triangleq \Pr\Big(\{g_1(\tilde{H}_{1,1}^n, Y_{1,1}^n) \ne M_1\} \cup \{g_2(\tilde{H}_{2,1}^n, Y_{2,1}^n) \ne M_2\}\Big)$, and each source message is selected independently and uniformly from its message set.

Definition 3: A rate pair (R_1, R_2) is called achievable if for any $\epsilon > 0$ and $\delta > 0$ there exists some blocklength $n_0(\epsilon, \delta)$, such that for every $n > n_0(\epsilon, \delta)$ there exists an $(R_1 - \delta, R_2 - \delta, n)$ code with $P_e^{(n)} < \epsilon$.

Definition 4: The capacity region is the convex hull of all achievable rate pairs.

III. Preliminaries

A. Capacity of Phase Fading and of Rayleigh Fading Relay Channels

The fading relay channel with source Tx_1 and destination Rx_1 can be obtained from the ICR by taking only (1a) and (1c) and setting $a_{21} = a_{23} = 0$. For the phase fading relay channel the following capacity theorem was derived in [4]:

Proposition (See [4, Th. 6]): When $a_{21} = a_{23} = 0$ in (1a) and (1c), and the channel coefficients satisfy

$$a_{11}^2 P_1 + a_{31}^2 P_3 \le a_{13}^2 P_1,$$

the capacity of the phase fading relay channel with Rx-CSI is given by

$$C_{\mathrm{PF-relay}} = \log_2(1 + a_{11}^2 P_1 + a_{31}^2 P_3). \tag{2}$$

Capacity is achieved with $X_1 \sim \mathcal{CN}(0, P_1)$, $X_3 \sim \mathcal{CN}(0, P_3)$, $X_1 \perp\!\!\!\perp X_3$, and DF at the relay.

Comment 1: The reason that independent Gaussian inputs achieve capacity (when relay reception is good) follows from a combination of four factors: 1) the fading coefficients are mutually independent and have uniformly distributed phases independent of their magnitudes, 2) the availability of Rx-CSI, 3) the lack of Tx-CSI, and 4) the concavity of the log function. In particular, we note that due to the lack of Tx-CSI, the components of the received signal at the destination, $\{H_{l1,i} X_i\}_{i=1}^n$, $l \in \{1, 3\}$, are uncorrelated, irrespective of the joint distribution $f_{X_1,X_3}(x_1, x_3)$. Furthermore, subject to the four factors mentioned previously, we obtain that each mutual information expression in the cut-set bound [4, eq. (6)] is maximized by a jointly complex Normal input distribution (see [4, Proposition 2]). The fact that the different components of the received signal cannot be correlated implies that there is no point generating correlated channel inputs. Finally, as uncorrelated jointly Gaussian RVs are mutually independent, this implies that independent Gaussian channel inputs maximize all mutual information expressions. A detailed argument can be found in the proof of [4, Th. 8].

For Rayleigh fading relay channels, [4, Th. 8] leads to the characterization of the maximal cut-set bound and the largest DF achievable rate. From this characterization, the following capacity theorem is obtained.

Proposition 1: Let $\tilde{\tilde{U}} \triangleq (U_{11}, U_{13}, U_{31})$. When the condition

$$1 + a_{11}^2 P_1 + a_{31}^2 P_3 \le \frac{a_{13}^2 P_1}{e^{\frac{1}{a_{13}^2 P_1}} E_1\left(\frac{1}{a_{13}^2 P_1}\right)} \tag{3}$$

holds, the capacity of the Rayleigh fading relay channel is

$$C_{\text{Rayleigh,relay}} = \mathbb{E}_{\tilde{U}}\left\{\log_2(1+a_{11}^2|U_{11}|^2P_1 + a_{31}^2|U_{31}|^2P_3)\right\} \tag{4}$$

and it is achieved with $X_1 \sim \mathcal{CN}(0,P_1)$, $X_3 \sim \mathcal{CN}(0,P_3)$ $X_1 \perp\!\!\!\perp X_3$, and DF at the relay.

Proof: The proof follows by applying [4, Th. 8] to the DF achievable rate of [27, Th. 1] and to the cut-set bound [25, Th. 15.10.1]. To guarantee that DF achieves capacity we require that $I(X_1;Y_3|X_3,U_{13}) \geq I(X_1,X_3;Y_1|U_{11},U_{31})$ subject to the maximizing channel inputs. The condition (3) follows by applying Jensen's inequality [25, Th. 2.6.2] to the above inequality, similar to the proof of Corollary 1 (see Appendix B). ■

B. The Multiple-Access Relay Channel

The MARC consists of two sources, a single destination, and a single relay [4, Sec. III-C]. The memoryless MARC is defined by $\{\mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{X}_3,\ p(y_1,y_3|x_1,x_2,x_3), \mathcal{Y}_1 \times \mathcal{Y}_3\}$, where Y_1 is the signal received at the destination, X_3 and Y_3 are the channel input and output at the relay, and X_1 and X_2 are the channel inputs from the sources. The destination uses its received signal vector $\mathbf{Y}_1$ to decode the messages from both sources. The fading Gaussian MARC model is given by (1a) and (1c).

C. Capacity Regions for Phase Fading and for Rayleigh Fading MARCs

The phase fading MARC is defined by (1a) and (1c), with $H_{lk,i} = a_{lk}e^{j\Theta_{lk,i}}$, $i = 1,2,\ldots,n$. The following capacity region of the PF-MARC was characterized in [4, Th. 9]:

Theorem (See [4, Th. 9]): Consider the MARC with phase fading characterized in (1a) and (1c). If the channel coefficients and the source powers satisfy

$$a_{11}^2P_1 + a_{31}^2P_3 \leq a_{13}^2P_1 \tag{5a}$$
$$a_{21}^2P_2 + a_{31}^2P_3 \leq a_{23}^2P_2 \tag{5b}$$
$$a_{11}^2P_1 + a_{21}^2P_2 + a_{31}^2P_3 \leq a_{13}^2P_1 + a_{23}^2P_2, \tag{5c}$$

then the capacity region is characterized by all nonnegative rate pairs $(R_1,R_2) \in \mathfrak{R}_+^2$ s.t.

$$R_1 \leq \log_2(1+a_{11}^2P_1 + a_{31}^2P_3) \tag{6a}$$
$$R_2 \leq \log_2(1+a_{21}^2P_2 + a_{31}^2P_3) \tag{6b}$$
$$R_1 + R_2 \leq \log_2(1+a_{11}^2P_1 + a_{21}^2P_2 + a_{31}^2P_3) \tag{6c}$$

and it is achieved using $X_1 \sim \mathcal{CN}(0,P_1)$, $X_2 \sim \mathcal{CN}(0,P_2)$ and $X_3 \sim \mathcal{CN}(0,P_3)$, mutually independent, and DF at the relay.

Note that achievability of this region also follows from the achievable region in Appendix A.

The Rayleigh fading MARC is defined by (1a) and (1c), with $H_{lk,i} = a_{lk}U_{lk,i}$, $i = 1,2,\ldots,n$. Assuming Rx-CSI, the following achievable rate region is obtained:

Proposition 2: Consider the Rayleigh fading MARC. The achievable rate region obtained by using $X_k \sim \mathcal{CN}(0,P_k)$, $k = 1,2,3$ mutually independent, DF at the relay, and backward decoding at the destination, is given by

$$\mathcal{R}_{\text{Rayleigh,MARC}} = \mathcal{R}_{\text{Rayleigh},1} \cap \mathcal{R}_{\text{Rayleigh},2},$$

$$\mathcal{R}_{\text{Rayleigh},1} = \Big\{(R_1,R_2) \in \mathfrak{R}_+^2 :$$
$$R_1 \leq \mathbb{E}_{\tilde{U}}\left\{\log_2(1+a_{13}^2|U_{13}|^2P_1)\right\} \tag{7a}$$
$$R_2 \leq \mathbb{E}_{\tilde{U}}\left\{\log_2(1+a_{23}^2|U_{23}|^2P_2)\right\} \tag{7b}$$
$$R_1+R_2 \leq \mathbb{E}_{\tilde{U}}\left\{\log_2(1+a_{13}^2|U_{13}|^2P_1 + a_{23}^2|U_{23}|^2P_2)\right\}\Big\}, \tag{7c}$$

$$\mathcal{R}_{\text{Rayleigh},2} = \Big\{(R_1,R_2) \in \mathfrak{R}_+^2 :$$
$$R_1 \leq \mathbb{E}_{\tilde{U}}\left\{\log_2(1+a_{11}^2|U_{11}|^2P_1 + a_{31}^2|U_{31}|^2P_3)\right\} \tag{8a}$$
$$R_2 \leq \mathbb{E}_{\tilde{U}}\left\{\log_2(1+a_{21}^2|U_{21}|^2P_2 + a_{31}^2|U_{31}|^2P_3)\right\} \tag{8b}$$
$$R_1+R_2 \leq \mathbb{E}_{\tilde{U}}\left\{\log_2(1+a_{11}^2|U_{11}|^2P_1 + a_{21}^2|U_{21}|^2P_2 + a_{31}^2|U_{31}|^2P_3)\right\}\Big\}. \tag{8c}$$

Proof: Applying the derivation outlined in Appendix A to the present model, we obtain that the rate constraints for decoding at the relay are given by (7), and the rate constraints for decoding at the destination are given by (8). ■

Now, we can state the following corollary:

Corollary 1: Consider the Rayleigh fading MARC. If the channel coefficients satisfy

$$1 + a_{11}^2P_1 + a_{31}^2P_3 \leq \frac{a_{13}^2P_1}{e^{\frac{1}{a_{13}^2P_1}}E_1\left(\frac{1}{a_{13}^2P_1}\right)} \tag{9a}$$

$$1 + a_{21}^2P_2 + a_{31}^2P_3 \leq \frac{a_{23}^2P_2}{e^{\frac{1}{a_{23}^2P_2}}E_1\left(\frac{1}{a_{23}^2P_2}\right)} \tag{9b}$$

$$1 + a_{11}^2P_1 + a_{21}^2P_2 + a_{31}^2P_3 \leq \frac{a_{23}^2P_2 - a_{13}^2P_1}{\left(e^{\frac{1}{a_{23}^2P_2}}E_1\left(\frac{1}{a_{23}^2P_2}\right) - e^{\frac{1}{a_{13}^2P_1}}E_1\left(\frac{1}{a_{13}^2P_1}\right)\right)} \tag{9c}$$

then the DF achievable region $\mathcal{R}_{\text{Rayleigh},2}$ given in (8) is the capacity region.

Proof: See Appendix B. ■

Comment 2: It is noted that condition (9c) holds also when $a_{23}^2P_2 = a_{13}^2P_1$ by applying L'Hopital's rule [29, Th. 14.1.5]. Note that when (9) holds, reliable decoding of the messages at their destinations implies reliable decoding of both messages at the relay.

IV. The Very Strong Interference Regime

In this section, capacity regions for fading ICRs for the VSI regime are characterized, first for phase fading and then for Rayleigh fading.

A. VSI for Ergodic Phase Fading and Ergodic Rayleigh Fading ICRs

Define $\mathcal{D}_{\mathrm{PF}}$ to be the set of channel coefficients $(a_{31}, a_{13}, a_{21}, a_{23})$ that satisfy

$$\mathcal{D}_{\mathrm{PF}} \triangleq \Big\{ (a_{31}, a_{13}, a_{32}, a_{23}) \in \mathfrak{R}_{+}^{4} :$$

$$P_1 + a_{31}^2 P_3 \le a_{13}^2 P_1 \tag{10a}$$

$$P_2 + a_{32}^2 P_3 \le a_{23}^2 P_2 \tag{10b}$$

$$(1+P_1+a_{31}^2 P_3)(1+P_2+a_{32}^2 P_3) \le 1+a_{13}^2 P_1+a_{23}^2 P_2 \Big\}. \tag{10c}$$

We now state the following capacity theorem for PF-ICRs in VSI (recall $a_{11} = a_{22} = 1$)

Theorem 1: Consider the ICR (1) with all links subject to i.i.d. phase fading. Let the sources and the relay be subject to per-symbol power constraints

$$\mathbb{E}\big\{|X_k|^2\big\} \le P_k, \qquad k = 1, 2, 3 \tag{11}$$

and the channel coefficients satisfy $(a_{31}, a_{13}, a_{32}, a_{23}) \in \mathcal{D}_{PF}$. If the channel coefficients satisfy also

$$a_{12}^2 \ge \left(1 + a_{31}^2 \frac{P_3}{P_1}\right)(1 + P_2 + a_{32}^2 P_3) \tag{12a}$$

$$a_{21}^2 \ge \left(1 + a_{32}^2 \frac{P_3}{P_2}\right)(1 + P_1 + a_{31}^2 P_3) \tag{12b}$$

then the capacity region with Rx-CSI is characterized by all nonnegative rate pairs $(R_1, R_2) \in \mathfrak{R}_{+}^{2}$ s.t.

$$R_1 \le \log_2\left(1 + P_1 + a_{31}^2 P_3\right) \tag{13a}$$

$$R_2 \le \log_2\left(1 + P_2 + a_{32}^2 P_3\right) \tag{13b}$$

and it is achieved by circularly symmetric, complex Normal inputs, $X_k \sim \mathcal{CN}(0, P_k)$, $k = 1, 2, 3$, all mutually independent, and DF at the relay.

Proof: See Appendix C. ■

Comment 3: When $(a_{31}, a_{13}, a_{32}, a_{23}) \in \mathcal{D}_{\mathrm{PF}}$ then the maximal achievable rate region at the destinations is a subset of the achievable region for decoding both messages at the relay. Therefore, decoding at the relay does not constrain the rates, and DF is optimal. When $(a_{31}, a_{13}, a_{32}, a_{23}) \in \mathcal{D}_{\mathrm{PF}}$, we say that relay reception is good. Conditions (12) are the VSI conditions.

For Rayleigh fading ICRs, define $\mathcal{A}_{\mathrm{RF}}$ to be the set of channel coefficients $(a_{31}, a_{13}, a_{21}, a_{23})$ that satisfy

$$\mathcal{A}_{\mathrm{RF}} \triangleq \Bigg\{ (a_{31}, a_{13}, a_{32}, a_{23}) \in \mathfrak{R}_{+}^{4} :$$

$$1 + P_1 + a_{31}^2 P_3 \le \frac{a_{13}^2 P_1}{e^{\frac{1}{a_{13}^2 P_1}} E_1\left(\frac{1}{a_{13}^2 P_1}\right)} \tag{14a}$$

$$1 + P_2 + a_{32}^2 P_3 \le \frac{a_{23}^2 P_2}{e^{\frac{1}{a_{23}^2 P_2}} E_1\left(\frac{1}{a_{23}^2 P_2}\right)} \tag{14b}$$

$$\left(1 + P_1 + a_{31}^2 P_3\right)\left(1 + P_2 + a_{32}^2 P_3\right) \le \frac{a_{23}^2 P_2 - a_{13}^2 P_1}{\left(e^{\frac{1}{a_{23}^2 P_2}} E_1\left(\frac{1}{a_{23}^2 P_2}\right) - e^{\frac{1}{a_{13}^2 P_1}} E_1\left(\frac{1}{a_{13}^2 P_1}\right)\right)} \Bigg\}. \tag{14c}$$

We now have the following characterization of VSI and the associated capacity region.

Theorem 2: Consider the ICR (1) with all links subject to i.i.d. Rayleigh fading. Let the channel coefficients satisfy $(a_{31}, a_{13}, a_{32}, a_{23}) \in \mathcal{A}_{RF}$. If the channel coefficients satisfy also

$$\frac{\frac{a_{12}^2 P_1}{1+P_2+a_{32}^2 P_3}}{e^{\frac{1+P_2+a_{32}^2 P_3}{a_{12}^2 P_1}} E_1\left(\frac{1+P_2+a_{32}^2 P_3}{a_{12}^2 P_1}\right)} \ge (1+P_1+a_{31}^2 P_3) \tag{15a}$$

$$\frac{\frac{a_{21}^2 P_2}{1+P_1+a_{31}^2 P_3}}{e^{\frac{1+P_1+a_{31}^2 P_3}{a_{21}^2 P_2}} E_1\left(\frac{1+P_1+a_{31}^2 P_3}{a_{21}^2 P_2}\right)} \ge (1+P_2+a_{32}^2 P_3) \tag{15b}$$

then the capacity region is given by all nonnegative rate pairs $(R_1, R_2) \in \mathfrak{R}_{+}^{2}$ s.t.

$$R_1 \le \mathbb{E}_{\tilde{U}}\Big\{ \log_2\left(1 + |U_{11}|^2 P_1 + a_{31}^2 |U_{31}|^2 P_3\right)\Big\} \tag{16a}$$

$$R_2 \le \mathbb{E}_{\tilde{U}}\Big\{ \log_2\left(1 + |U_{22}|^2 P_2 + a_{32}^2 |U_{32}|^2 P_3\right)\Big\} \tag{16b}$$

and it is achieved by circularly symmetric, complex Normal inputs, $X_1 \sim \mathcal{CN}(0, P_1)$, $X_2 \sim \mathcal{CN}(0, P_2)$, $X_3 \sim \mathcal{CN}(0, P_3)$, all mutually independent, and DF at the relay.

Proof: See Appendix D. ■

B. Discussion

- Note that without Tx-CSI, it is not possible to achieve nonzero correlation between the components of the received signal generated by the transmissions from Tx_1, Tx_2, and the relay. Therefore, the outer bound is maximized by uncorrelated jointly complex Normal channel inputs, and we conclude that the optimal channel inputs are mutually independent complex Normal RVs. With Tx-CSI, the transmitters and the relay have access to common information through which correlation between the channel inputs can be achieved.

- By comparison of the VSI conditions for phase fading and for Rayleigh fading, we see that the VSI regime for phase fading (12) is larger than for Rayleigh fading (15). This can be attributed to the amplitude fluctuations in the Rayleigh model due to which larger channel coefficients are required to guarantee reliable decoding, and also to the application of Jensen's inequality which is used for simplifying the expressions for Rayleigh fading.
- We note that although there is a single relay, under VSI the ICR behaves like *two parallel relay channels*, with the same relay *optimally assisting each Tx-Rx pair*.
- When the relay is off ($P_3 = 0$), the VSI conditions (12) specialize to $a_{12}^2 \geq 1+P_2, a_{21}^2 \geq 1+P_1$, which are the VSI conditions for the time-invariant Gaussian IC obtained by Carleial [23]. Note that although the models are different (in [23] the channel is constant), the rate expressions are similar, thus the identical conditions.
- The optimal relaying scheme in this case is based on "signal forwarding" [5] rather than "interference forwarding." Namely, we do not take advantage of the fact that the relay signal also carries interference. This is because the interference is strong enough to allow each receiver to decode its interfering message based only on the cross-link component of its received signal (i.e., the signal received from the unintended transmitter). Therefore, there is no benefit in letting the relay increase the interference. Then, the assistance of the relay is through increasing the rate of the desired information at each receiver.

V. The Strong Interference Regime

A. Strong Interference for Ergodic Phase Fading ICRs

Let $\mathcal{D}_k$ be the set of channel coefficients $\mathbf{a}_k = \big(a_{1k}, a_{2k}, a_{3k}, a_{13}, a_{23}\big) \in \mathfrak{R}_+^5$ that satisfy

$$a_{1k}^2 P_1 + a_{3k}^2 P_3 \leq a_{13}^2 P_1 \tag{17a}$$
$$a_{2k}^2 P_2 + a_{3k}^2 P_3 \leq a_{23}^2 P_2 \tag{17b}$$
$$a_{1k}^2 P_1 + a_{2k}^2 P_2 + a_{3k}^2 P_3 \leq a_{13}^2 P_1 + a_{23}^2 P_2 \tag{17c}$$

$k = 1,2$. Let $\mathbf{a} \triangleq \big(a_{11}, a_{12}, a_{13}, a_{21}, a_{22}, a_{23}, a_{31}, a_{32}\big) \in \mathfrak{R}_+^8$ and let $\mathbf{a} \in \mathcal{D}_1 \cap \mathcal{D}_2$ be a short-form notation to denote that $\mathbf{a}_1$ and $\mathbf{a}_2$ satisfy $\big\{\mathbf{a}_1 \cup \mathbf{a}_2 = \mathbf{a}\big\} \bigcap \big\{\mathbf{a}_1 \in \mathcal{D}_1\big\} \bigcap \big\{\mathbf{a}_2 \in \mathcal{D}_2\big\}$. Then, we have the following capacity region:

Theorem 3: Consider the ICR (1) with i.i.d. phase fading whose channel coefficients satisfy $\mathbf{a} \in \mathcal{D}_1 \cap \mathcal{D}_2$. If the channel coefficients satisfy also

$$a_{12}^2 \geq \left(1 + a_{31}^2 \frac{P_3}{P_1}\right)\big(1 + a_{32}^2 P_3\big) \tag{18a}$$
$$a_{21}^2 \geq \left(1 + a_{32}^2 \frac{P_3}{P_2}\right)\big(1 + a_{31}^2 P_3\big) \tag{18b}$$

then the capacity region is characterized by all nonnegative rate pairs $(R_1, R_2) \in \mathfrak{R}_+^2$ s.t.

$$R_1 \leq \log_2(1 + P_1 + a_{31}^2 P_3) \tag{19a}$$
$$R_2 \leq \log_2(1 + P_2 + a_{32}^2 P_3) \tag{19b}$$
$$R_1 + R_2 \leq \log_2(1 + P_1 + a_{21}^2 P_2 + a_{31}^2 P_3) \tag{19c}$$
$$R_1 + R_2 \leq \log_2(1 + a_{12}^2 P_1 + P_2 + a_{32}^2 P_3). \tag{19d}$$

This region is achieved by circularly symmetric, complex Normal inputs, $X_1 \sim \mathcal{CN}(0, P_1)$, $X_2 \sim \mathcal{CN}(0, P_2)$, $X_3 \sim \mathcal{CN}(0, P_3)$, all mutually independent, and DF at the relay.

Proof: See Appendix E. ■

Consider the two MARCs derived from the ICR: the first MARC has Rx_1 as its destination and is given by (1a) and (1c). We refer to it as MARC_1. The second MARC has Rx_2 as its destination and is given by (1b) and (1c). We refer to it as MARC_2. We now make several comments:

Comment 4: Note that the condition $\mathbf{a} \in \mathcal{D}_1 \cap \mathcal{D}_2$ is a combination of the MARC conditions (5) applied to MARC_1 and to MARC_2. Note also that there is no contradiction between the requirement $\mathbf{a} \in \mathcal{D}_1 \cap \mathcal{D}_2$ and conditions (18): making a_{13}, a_{23}, a_{12}, and a_{21} large enough it is possible to find sets of coefficients $\mathbf{a} \in \mathcal{D}_1 \cap \mathcal{D}_2$ that also satisfy (18).

Comment 5: Similarly to Section IV, $\mathbf{a} \in \mathcal{D}_k$, $k = 1,2$, implies that if a rate pair is achievable at destination k, it is also achievable at the relay. The difference between (17c) and (10c) follows as in the VSI regime, each destination decodes its desired message after eliminating the interference, while in the SI regime, joint decoding of the message and interference is applied.

Comment 6: When the relay is off $(P_3 = 0)$, conditions (18) reduce to $a_{21}^2 > 1$, $a_{12}^2 > 1$, which are the standard SI conditions for Gaussian time-invariant ICs [24] with $h_{lk} = a_{lk}$ real constants. This follows as although the models are different, the rate expressions are similar.

Comment 7: When the relay-Rx_1 link is off $(a_{31} = 0)$, the SI condition for Tx_2 (18b) becomes $a_{21}^2 > 1 + \frac{P_3}{P_2} a_{32}^2$. Thus, the interference has to be stronger than the augmented direct path $\mathrm{Tx}_2 - \mathrm{Rx}_2$ with coefficient $\sqrt{1 + \frac{P_3}{P_2} a_{32}^2}$. This path, in turn, can be viewed as if the relay is helping only Tx_2 (best case for Tx_2). On the other hand, if the link relay-Rx_2 is off $(a_{32} = 0)$, then the strong interference condition for Tx_2 becomes $\frac{a_{21}^2}{1 + a_{31}^2 P_3} > 1$. This is the worst case situation for decoding m_2 at Rx_1,[1] because instead of helping, the relay increases the noise at Rx_1. In this case, the interference link has to be strong enough such that the SNR for decoding m_2 at Rx_1 will be higher than the SNR for decoding m_2 at Rx_2 (which equals P_2), overcoming the increase in noise power at Rx_1 caused by the relay.

[1]Recall that we are dealing with a relay whose objective is to help, not with a jammer whose aim is to disturb communications.

When both relay-destination links are active, the SI condition for Tx_2 has two components: one component is the augmented direct path $\mathrm{Tx}_2 - \mathrm{Rx}_2$ with coefficient $\sqrt{1+\frac{P_3}{P_2}a_{32}^2}$. The second component is the increased noise at Rx_1, whose variance is $1+a_{31}^2P_3$. SI is then characterized by the product of these two components, which represents the amplification of the increased noise at Rx_1 by the augmented channel coefficient. Equivalently, we can write the SI condition at Tx_2 as

$$\frac{P_2a_{21}^2}{1+a_{31}^2P_3} > P_2 + P_3a_{32}^2.$$

Observe that the left-hand side is the SNR for decoding the interference at Rx_1 and the right-hand size is the best SNR for decoding the desired message at Rx_2. This is an extension of the original normalization argument used in [24].

Comment 8: In [3, Th. 1], SI was characterized for the "signal-cognitive" time-invariant ICR. The SI condition provided in [3] for, e.g., Tx_2, requires $|a_{21}| \ge 1+(|a_{31}|+|a_{32}|)\sqrt{\frac{P_3}{P_2}}$. This means that the interference has to overcome the augmented direct path $\mathrm{Tx}_2 - \mathrm{Rx}_2$ with coefficient $1+|a_{32}|\sqrt{\frac{P_3}{P_2}}$, further augmented by the $\mathbf{x}_2$ component in the relay signal received at Rx_1, $|a_{31}|\sqrt{\frac{P_3}{P_2}}$. Note that all relay power goes into the effective channel coefficient, and no power from the relay is combined into the noise, contrary to our case. Also note that the relay and the source combine coherently, which is also in contrast to the present case. This is because when the relay is causal, the worst-case interference from the relay manifests itself as an increased noise power (assuming independent codebooks), while when the relay noncausal (i.e., cognitive), the interference it creates plays the same role as the interference the source creates (in the sense that the relay signal consists of two additive components $\mathbf{x}_3 = \alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2$. Then, $\alpha_k\mathbf{x}_k$ coherently combines with the signal transmitted by Tx_k, $k=1,2$). This highlights the difference between cognitive and causal relaying in the context of SI.

Comment 9: Comparing VSI with SI, we see that in VSI, P_1 is included in the effective noise variance when decoding the interference m_2 at Rx_1, while in SI, the SNR for decoding the interference from Tx_2 at Rx_1 is lower bounded by $\frac{a_{21}^2P_2}{1+a_{31}^2P_3}$, which mean that P_1 is not part to the effective noise variance [cf., (12b) and (18b)]. This is the benefit of jointly decoding the interference and the desired message. Symmetric conclusion holds for decoding at Rx_2. Indeed, we see that the SI conditions are satisfied with smaller cross-links coefficients than the VSI conditions. Note that in SI, the relay signal is treated as noise when decoding the interference; therefore, the relaying scheme used is classified as "signal forwarding" and not "interference forwarding," similarly to the situation in VSI.

Comment 10: Note that the results of this study will hold also if some of the scenario assumptions were different. For example, using a long-term power constraint will not affect the results at all, while having a common message for each source would not change the type of argument but will change the VSI and SI conditions and the corresponding rate constraints. Also, if the relay would be capable of instantaneous relaying then the VSI or SI conditions and the rate constraints would not change as we know that in these regimes DF achieves capacity. Note also that the lack of Tx-CSI would not allow correlation between the components of the received signal even with instantaneous relaying. Instantaneous relaying may lead to different results if the channel is not in the SI regime so DF is no longer optimal.

B. Strong Interference for Ergodic Rayleigh Fading ICRs

We next consider SI for Rayleigh fading ICRs. Let $\mathcal{A}_1$ be the set of all vectors $\mathbf{a}_1$ that satisfy (9). Similarly, let $\mathcal{A}_2$ be the set of all vectors $\mathbf{a}_2$ that satisfy

$$1+a_{12}^2P_1+a_{32}^2P_3 \le \frac{a_{13}^2P_1}{e^{\frac{1}{a_{13}^2P_1}}E_1\left(\frac{1}{a_{13}^2P_1}\right)} \quad (20a)$$

$$1+a_{22}^2P_2+a_{32}^2P_3 \le \frac{a_{23}^2P_2}{e^{\frac{1}{a_{23}^2P_2}}E_1\left(\frac{1}{a_{23}^2P_2}\right)} \quad (20b)$$

$$1+a_{12}^2P_1+a_{22}^2P_2+a_{32}^2P_3 \le \frac{a_{23}^2P_2 - a_{13}^2P_1}{\left(e^{\frac{1}{a_{23}^2P_2}}E_1\left(\frac{1}{a_{23}^2P_2}\right) - e^{\frac{1}{a_{13}^2P_1}}E_1\left(\frac{1}{a_{13}^2P_1}\right)\right)}. \quad (20c)$$

Finally, let $\mathbf{a} \in \mathcal{A}_1 \cap \mathcal{A}_2$ be a short-form notation for denoting that $\mathbf{a}_1$ and $\mathbf{a}_2$ satisfy $\{\mathbf{a}_1 \cup \mathbf{a}_2 = \mathbf{a}\} \bigcap \{\mathbf{a}_1 \in \mathcal{A}_1\} \bigcap \{\mathbf{a}_2 \in \mathcal{A}_2\}$. We then have the following capacity region:

Theorem 4: Consider the ICR (1) with i.i.d. Rayleigh fading whose channel coefficients satisfy $\mathbf{a} \in \mathcal{A}_1 \cap \mathcal{A}_2$. If the channel coefficients satisfy also

$$\frac{a_{12}^2P_1}{e^{\frac{1+a_{32}^2P_3}{a_{12}^2P_1}}E_1\left(\frac{1+a_{32}^2P_3}{a_{12}^2P_1}\right)} \ge (1+P_1+a_{31}^2P_3)(1+a_{32}^2P_3) \quad (21a)$$

$$\frac{a_{21}^2P_2}{e^{\frac{1+a_{31}^2P_3}{a_{21}^2P_2}}E_1\left(\frac{1+a_{31}^2P_3}{a_{21}^2P_2}\right)} \ge (1+P_2+a_{32}^2P_3)(1+a_{31}^2P_3) \quad (21b)$$

then the capacity region is characterized by all nonnegative rate pairs $(R_1,R_2) \in \mathfrak{R}_+^2$ s.t.

$$R_1 \le \mathbb{E}_{\tilde{U}}\left\{\log_2\left(1+|U_{11}|^2P_1+a_{31}^2|U_{31}|^2P_3\right)\right\} \quad (22a)$$

$$R_2 \le \mathbb{E}_{\tilde{U}}\left\{\log_2\left(1+|U_{22}|^2P_2+a_{32}^2|U_{32}|^2P_3\right)\right\} \quad (22b)$$

$$R_1+R_2 \le \mathbb{E}_{\tilde{U}}\left\{\log_2\left(1+|U_{11}|^2P_1+a_{21}^2|U_{21}|^2P_2 + a_{31}^2|U_{31}|^2P_3\right)\right\} \quad (22c)$$

$$R_1+R_2 \le \mathbb{E}_{\tilde{U}}\left\{\log_2\left(1+a_{12}^2|U_{12}|^2P_1+|U_{22}|^2P_2 + a_{32}^2|U_{32}|^2P_3\right)\right\}. \quad (22d)$$

This region is achieved by circularly symmetric, complex Normal inputs, $X_1 \sim \mathcal{CN}(0,P_1)$, $X_2 \sim \mathcal{CN}(0,P_2)$, $X_3 \sim \mathcal{CN}(0,P_3)$, all mutually independent, and DF at the relay.

Proof: See Appendix F. ■

Comment 11: It is easy to see that the VSI conditions (15) are stricter than the SI conditions (21). For example, comparing (VSI for Tx_1)

$$\frac{\frac{a_{12}^2 P_1}{1+P_2+a_{32}^2 P_3}}{e^{\frac{1+P_2+a_{32}^2 P_3}{a_{12}^2 P_1}} E_1\left(\frac{1+P_2+a_{32}^2 P_3}{a_{12}^2 P_1}\right)} \geq 1 + P_1 + a_{31}^2 P_3$$

and (SI for Tx_1)

$$\frac{\frac{a_{12}^2 P_1}{1+a_{32}^2 P_3}}{e^{\frac{1+a_{32}^2 P_3}{a_{12}^2 P_1}} E_1\left(\frac{1+a_{32}^2 P_3}{a_{12}^2 P_1}\right)} \geq 1 + P_1 + a_{31}^2 P_3$$

we have

$$\frac{\frac{a_{12}^2 P_1}{1+a_{32}^2 P_3}}{e^{\frac{1+a_{32}^2 P_3}{a_{12}^2 P_1}} E_1\left(\frac{1+a_{32}^2 P_3}{a_{12}^2 P_1}\right)} \geq \frac{\frac{a_{12}^2 P_1}{1+P_2+a_{32}^2 P_3}}{e^{\frac{1+P_2+a_{32}^2 P_3}{a_{12}^2 P_1}} E_1\left(\frac{1+P_2+a_{32}^2 P_3}{a_{12}^2 P_1}\right)}$$

thus VSI guarantees SI.

VI. Conclusion

In this paper, we characterized capacity regions for fading ICRs for different interference regimes. The results were derived for phase fading and for Rayleigh fading channels, under the assumption that each node has CSI only on its incoming links (receive CSI). We first characterized the VSI regime and found the corresponding capacity regions and then characterized the SI regime and obtained the corresponding capacity regions as well. The importance of the results lies in the fact that this is the first time the capacity region is characterized for a cooperative scenario in which a causal relay simultaneously assists two separate communicating pairs, when all links are active and the SNRs are finite. Thus, the results indicate how should the relay optimally operate when there are two communicating pairs that interfere with each other, in the regime in which the relay is able to decode the sources messages without constraining the rates to the destinations. These results have direct implications to practical communication scenarios in two aspects: the results show, for the first time, that it is possible to optimally assist several communicating pairs with a single relay, which is an encouraging result for using relays in wireless networks where there is interference. The results also indicate that independent codebooks are optimal; thus, adding relays does not requires changes in the transmitters. Additionally, our SI and VSI conditions can characterize geographical regions in which DF relaying is optimal. Naturally, implementing relaying in wireless networks which suffer from interference requires considerably more work; in particular, all interference regimes have to be studied as well as the time-invariant case.

Appendix A
Achievable Rate Region for Fading MARCs

In this appendix, an achievable rate region for fading MARCs, based on decode-and-forward at the relay, is presented. This region follows from [4] and [14] with slight modifications (as the backward decoding rule used does not work with independent codebooks).

Fix input distributions $f_{X_1}(x_1)$, $f_{X_2}(x_2)$, $f_{X_3}(x_3)$, and blocklength n, and consider the following coding scheme:

A) Codebooks Construction:

a) For each $m_k \in \mathcal{M}_k$, $k=1,2$, select a codeword $\mathbf{x}_k(m_k)$ with p.d.f. $f_{\mathbf{X}_k}\big(\mathbf{x}_k(m_k)\big) = \prod_{i=1}^n f_{X_k}\big(x_{k,i}(m_k)\big)$.

b) For each $(m_1,m_2) \in \mathcal{M}_1 \times \mathcal{M}_2$ select a codeword $\mathbf{x}_3(m_1,m_2)$ with p.d.f. $f_{\mathbf{X}_3}\big(\mathbf{x}_3(m_1,m_2)\big) = \prod_{i=1}^n f_{X_3}\big(x_{3,i}(m_1,m_2)\big)$. This codebook contains $2^{n(R_1+R_2)}$ codewords and is used by the relay.

As in [4, Appendix A], a block of $B-1$ messages is transmitted using nB channel symbols.

B) Decoding at the Relay at Block b: The relay uses its knowledge of $\mathbf{x}_3(m_{1,b-1},m_{2,b-1})$, $\mathbf{h}_{13}(b)$, and $\mathbf{h}_{23}(b)$ to decode $(m_{1,b},m_{2,b})$ by looking for a unique pair $(m_1,m_2) \in \mathcal{M}_1 \times \mathcal{M}_2$ such that

$$\big(\mathbf{x}_1(m_1),\mathbf{x}_2(m_2),\mathbf{x}_3(m_{1,b-1},m_{2,b-1}),\mathbf{h}_{13}(b), \mathbf{h}_{23}(b),\mathbf{y}_3(b)\big) \in A_\epsilon^{(n)}(X_1,X_2,X_3,H_{13},H_{23},Y_3).$$

Due to Rx-CSI, the achievable rate region for decoding both messages at the relay is given by [25], [28, Ch. 15.3]

$$R_1 \leq I(X_1;Y_3|X_2,X_3,\tilde{H}_3) \quad \text{(A1a)}$$
$$R_2 \leq I(X_2;Y_3|X_1,X_3,\tilde{H}_3) \quad \text{(A1b)}$$
$$R_1+R_2 \leq I(X_1,X_2;Y_3|X_3,\tilde{H}_3). \quad \text{(A1c)}$$

C) Decoding at the Destination: The destination performs backward decoding ([4, Appendix A] with a slight modification). Assume that $m_{1,b+1}$ and $m_{2,b+1}$ were correctly decoded. Define the events

$$\mathcal{E}_{1,b}(m_1,m_2) \triangleq \Big\{\big(\mathbf{x}_1(m_{1,b+1}),\mathbf{x}_2(m_{2,b+1}), \mathbf{x}_3(m_1,m_2),\mathbf{y}_1(b+1),\mathbf{h}_{11}(b+1), \mathbf{h}_{21}(b+1),\mathbf{h}_{31}(b+1)\big) \in A_\epsilon^{(n)}\Big\}$$

$$\mathcal{E}_{2,b}(m_1,m_2) \triangleq \Big\{\big(\mathbf{x}_1(m_1),\mathbf{x}_2(m_2),\mathbf{y}_1(b),\mathbf{h}_{11}(b),\mathbf{h}_{21}(b), \mathbf{h}_{31}(b)\big) \in A_\epsilon^{(n)}\Big\}.$$

The destination decodes $\big(m_{1,b},m_{2,b}\big)$ by choosing a unique message pair $\big(m_1,m_2\big) \in \mathcal{M}_1 \times \mathcal{M}_2$ such that

$$\mathcal{E}_{1,b}(m_1,m_2) \bigcap \mathcal{E}_{2,b}(m_1,m_2) \quad \text{(A2)}$$

holds. It is noted that as the codebooks are generated independently, then $\mathcal{E}_{1,b}(m_1,m_2) \perp\!\!\!\perp \mathcal{E}_{2,b}(m_1,m_2)$.

The probability of error can be analyzed using standard joint-typicality arguments as in [4] and [14]. The analysis results in the following achievable rate constraints:

$$R_1 < I(X_1, X_3; Y_1 | X_2, \tilde{H}_1) \qquad \text{(A3a)}$$
$$R_2 < I(X_2, X_3; Y_1 | X_1, \tilde{H}_1) \qquad \text{(A3b)}$$
$$R_1 + R_2 < I(X_1, X_2, X_3; Y_1 | \tilde{H}_1). \qquad \text{(A3c)}$$

We conclude that all rate pairs (R_1, R_2) satisfying (A1) and (A3) for some $f(x_1, x_2, x_3) = f_{X_1}(x_1) f_{X_2}(x_2) f_{X_3}(x_3)$ are achievable.

Appendix B
Proof of Corollary 1

The capacity region of the Rayleigh fading MARC can be outer bounded using the cut-set theorem [25, Th. 15.10.1], which gives the following three constraints:

$$R_1 \le \min\Big\{ I(X_1; Y_1, Y_3 | X_2, X_3, \tilde{H}), \; I(X_1, X_3; Y_1 | X_2, \tilde{H}) \Big\} \qquad \text{(B1a)}$$
$$R_2 \le \min\Big\{ I(X_2; Y_1, Y_3 | X_1, X_3, \tilde{H}), \; I(X_2, X_3; Y_1 | X_1, \tilde{H}) \Big\} \qquad \text{(B1b)}$$
$$R_1 + R_2 \le \min\Big\{ I(X_1, X_2; Y_1, Y_3 | X_3, \tilde{H}), \; I(X_1, X_2, X_3; Y_1 | \tilde{H}) \Big\} \qquad \text{(B1c)}$$

for the maximizing distribution $f(x_1, x_2, x_3)$. From [4, Th. 8], it follows that all mutual information expressions in (B1) are simultaneously maximized by mutually independent Gaussian inputs.

Since the first expressions in each of the minimums in (B1) satisfy

$$\begin{aligned}
I(X_1; Y_1, Y_3 | X_2, X_3, \tilde{H}) &\ge I(X_1; Y_3 | X_2, X_3, \tilde{H}) \\
&\ge \mathbb{E}_{\tilde{U}}\big\{ \log_2(1 + a_{13}^2 |U_{13}|^2 P_1) \big\} \\
I(X_2; Y_1, Y_3 | X_1, X_3, \tilde{H}) &\ge I(X_2; Y_3 | X_1, X_3, \tilde{H}) \\
&\ge \mathbb{E}_{\tilde{U}}\big\{ \log_2(1 + a_{23}^2 |U_{23}|^2 P_2) \big\} \\
I(X_1, X_2; Y_1, Y_3 | X_3, \tilde{H}) &\ge I(X_1, X_2; Y_3 | X_3, \tilde{H}) \\
&\ge \mathbb{E}_{\tilde{U}}\big\{ \log_2(1 + a_{13}^2 |U_{13}|^2 P_1 + a_{23}^2 |U_{23}|^2 P_2) \big\}
\end{aligned}$$

then, if $\mathcal{R}_{\text{Rayleigh},2} \subseteq \mathcal{R}_{\text{Rayleigh},1}$, it follows that $\mathcal{R}_{\text{Rayleigh},2}$ is the capacity region. Therefore, all that is needed in order to complete the proof is to find conditions on the channel coefficients that guarantee that $\mathcal{R}_{\text{Rayleigh},2} \subseteq \mathcal{R}_{\text{Rayleigh},1}$.

Let R.H.S. denote "right-hand side." Requiring that R.H.S (7a) $\ge$ R.H.S. (8a) and that R.H.S. (7b) $\ge$ R.H.S. (8b), and applying Jensen's inequality leads to (9a) and (9b). We demonstrate the argument by showing that (9c) guarantees R.H.S. (7c) $\ge$ R.H.S. (8c). Comparing (8c) and (7c), we seek to guarantee

$$\begin{aligned}
&\mathbb{E}_{\tilde{U}}\big\{ \log_2(1 + a_{11}^2 |U_{11}|^2 P_1 + a_{21}^2 |U_{21}|^2 P_2 + a_{31}^2 |U_{31}|^2 P_3) \big\} \\
&\qquad \le \mathbb{E}_{\tilde{U}}\big\{ \log_2(1 + a_{13}^2 |U_{13}|^2 P_1 + a_{23}^2 |U_{23}|^2 P_2) \big\} \\
&\Rightarrow \mathbb{E}_{\tilde{U}}\left\{ \log_2\left(\frac{1 + a_{11}^2 |U_{11}|^2 P_1 + a_{21}^2 |U_{21}|^2 P_2 + a_{31}^2 |U_{31}|^2 P_3}{1 + a_{13}^2 |U_{13}|^2 P_1 + a_{23}^2 |U_{23}|^2 P_2} \right) \right\} \\
&\qquad \le 0.
\end{aligned}$$

Applying Jensen's inequality [25, Th. 2.6.2], it follows that this is satisfied if

$$\begin{aligned}
&\log_2\left(\mathbb{E}_{\tilde{U}}\left\{ \frac{1 + a_{11}^2 |U_{11}|^2 P_1 + a_{21}^2 |U_{21}|^2 P_2 + a_{31}^2 |U_{31}|^2 P_3}{1 + a_{13}^2 |U_{13}|^2 P_1 + a_{23}^2 |U_{23}|^2 P_2} \right\} \right) \le 0 \\
&\overset{(a)}{\Rightarrow} (1 + a_{11}^2 P_1 + a_{21}^2 P_2 + a_{31}^2 P_3) \times \\
&\qquad \mathbb{E}_{\tilde{U}}\left\{ \frac{1}{1 + a_{13}^2 |U_{13}|^2 P_1 + a_{23}^2 |U_{23}|^2 P_2} \right\} \le 1
\end{aligned}$$

where (a) follows as all variables are independent.

Let $X \sim \exp(\lambda)$ denote that X is an exponentially distributed RV with parameter λ. Note that if $|U|^2 \sim \exp(1)$ and $Z = c|U|^2$, then, $Z \sim \exp\left(\frac{1}{c}\right)$. Next, we consider the sum of two exponential RVs: let $X_1 \sim \exp(\lambda_1)$, $X_2 \sim \exp(\lambda_2)$ and consider $Z = X_1 + X_2$. From the standard formula for the p.d.f. of the sum of two independent RVs [30, eq. (6–39)] we obtain

$$f_Z(z) = \frac{\lambda_1 \lambda_2}{\lambda_1 - \lambda_2}\left[e^{-\lambda_2 z} - e^{-\lambda_1 z} \right], \qquad z \ge 0, \quad \lambda_1 \ne \lambda_2.$$

Since $|U_{13}|^2, |U_{23}|^2 \sim \exp(1)$, independent of each other, setting $\lambda_1 = \frac{1}{a_{13}^2 P_1}$, $\lambda_2 = \frac{1}{a_{23}^2 P_2}$ we obtain

$$\begin{aligned}
&\mathbb{E}_{\tilde{U}}\left\{ \frac{1}{1 + a_{13}^2 |U_{13}|^2 P_1 + a_{23}^2 |U_{23}|^2 P_2} \right\} \\
&\quad = \frac{1}{a_{23}^2 P_2 - a_{13}^2 P_1}\left(e^{\frac{1}{a_{23}^2 P_2}} E_1\left(\frac{1}{a_{23}^2 P_2}\right) - e^{\frac{1}{a_{13}^2 P_1}} E_1\left(\frac{1}{a_{13}^2 P_1}\right) \right).
\end{aligned} \qquad \text{(B2)}$$

This gives (9c).[2]

Appendix C
Proof of Theorem 1

We begin the proof with the outer bound and then find a matching inner bound. In Appendixes C-1–C-4, we assume only that the fading coefficients have uniformly distributed phases, independent of their magnitudes. In Appendix C-5, the results are specialized to phase fading.

A) Outer Bound: Using only the individual rate constraints of the cut-set bound [25, Th. 15.10.1], we obtain the following outer bound:

$$\mathcal{R}_{\text{VSI}}^{\text{outer}} = \bigcup_{f(x_1, x_2, x_3)} \Big\{ (R_1, R_2) \in \mathfrak{R}_+^2 :$$
$$R_1 \le \min\Big\{ I(X_1; Y_1, Y_3 | X_3, X_2, \tilde{H}), \; I(X_1, X_3; Y_1 | X_2, \tilde{H}) \Big\} \qquad \text{(C1a)}$$
$$R_2 \le \min\Big\{ I(X_2; Y_2, Y_3 | X_3, X_1, \tilde{H}), \; I(X_2, X_3; Y_2 | X_1, \tilde{H}) \Big\} \Big\}. \qquad \text{(C1b)}$$

[2]In case $\lambda_1 = \lambda_2$, $f_Z(z) = \lambda_1^2 z e^{-\lambda_1 z}$, $z \ge 0$, and $\mathbb{E}_{\tilde{U}}\left\{ \frac{1}{1 + a_{13}^2 |U_{13}|^2 P_1 + a_{23}^2 |U_{23}|^2 P_2} \right\} = \lambda_1 - \lambda_1^2 e^{\lambda_1} E_1(\lambda_1)$. This can be obtained by writing $\lambda_2 = \lambda_1 + \epsilon$ and evaluating (B2) noting that $\lim_{\epsilon \to 0} \frac{f(x+\epsilon) - f(x)}{\epsilon} = f'(x)$.

B) Maximizing the Outer Bound: Consider maximizing the mutual information expressions in (C1). Due to the uniform distribution and the independence, both across time and across links, of the phases of the fading processes, as well as the independence of the phase and magnitude of the fading coefficients, it follows from [4, Th. 8] that independent Gaussian inputs with maximum possible powers simultaneously maximize all mutual information expressions in (C1).

C) Inner Bound:

1) Codebooks Construction, Encoding and Relay Decoding: Fix $X_k \sim \mathcal{CN}(0, P_k)$, $k = 1, 2, 3$, mutually independent, and blocklength n. Codebooks construction, encoding at the sources and at the relay and decoding at the relay is done as in Appendixes A-1 and A-2. Decoding at the relay leads to rate constraints (A1). Recall that this construction uses i.i.d. and mutually independent codebooks.

2) Decoding at the Destinations: Decoding at each destination is done in the following sequence:

i) Decoding the interference while treating the desired signal and the relay signal as noise (PtP rule [25, Th. 9.1.1]). This results in constraints on R_1 and R_2.
ii) Decoding the desired information with the help of the relay, while using the knowledge of the interfering message. This results in additional constraints on R_1 and R_2.
iii) Obtaining the conditions on the channel coefficients such that the interference is strong enough to allow this two-step decoding to be carried out without constraining the rate of the desired information.

Decoding the Interference: First note that decoding $m_{2,b}$ at Rx_1 and $m_{1,b}$ at Rx_2, prior to decoding the desired message at each destination receiver, can be done by treating the desired signal and the relay signal as noise. As X_1, X_2, X_3, H_{12}, H_{22}, H_{32}, H_{11}, H_{21}, and H_{31} are mutually independent and the corresponding vectors are i.i.d., it is concluded that $\{H_{32,i}X_{3,i}\}_{i=1}^n$ is an i.i.d. vector, independent of the other components vectors in $\mathbf{Y}_2$, namely, $\{H_{22,i}X_{2,i}\}_{i=1}^n$ and $\{H_{12,i}X_{1,i}\}_{i=1}^n$. This leads to the rate bound

$$R_1 < I(X_1; Y_2|\tilde{H}_2) \tag{C2a}$$

at Rx_2. Similarly, $\{H_{31,i}X_{3,i}\}_{i=1}^n$ is an i.i.d. vector, independent of the other component vectors in $\mathbf{Y}_1$, leading to the constraint

$$R_2 < I(X_2; Y_1|\tilde{H}_1) \tag{C2b}$$

at Rx_1. The constraints (C2) guarantee reliable decoding of the interference.

Decoding the Desired Information: Having decoded the interference, each destination performs backward decoding as described in Appendix A-3. Assuming the interference was decoded successfully, Rx_1 knows $m_{2,b}$ and Rx_2 knows $m_{1,b}$. Consider decoding $m_{1,b}$ at Rx_1: Rx_1 decides on $\hat{m}_{1,b}$ if it is the unique message in $\mathcal{M}_1$ for which

$$\mathcal{E}_{1,b}(\hat{m}_{1,b}, m_{2,b}) \bigcap \mathcal{E}_{2,b}(\hat{m}_{1,b}, m_{2,b})$$

holds.

From standard joint-typicality arguments it follows that Rx_1 can reliably decode $m_{1,b}$ as long as

$$R_1 < I(X_1, X_3; Y_1|X_2, \tilde{H}_1). \tag{C3a}$$

Using parallel arguments it can be shown that Rx_2 can reliably decode $m_{2,b}$ as long as

$$R_2 < I(X_2, X_3; Y_2|X_1, \tilde{H}_2). \tag{C3b}$$

Obtaining the VSI Conditions: Comparing (C2) and (C3), we obtain that decoding the interference first does not constrain the rate of the desired information as long as

$$I(X_1, X_3; Y_1|X_2, \tilde{H}_1) \le I(X_1; Y_2|\tilde{H}_2) \tag{C4a}$$
$$I(X_2, X_3; Y_2|X_1, \tilde{H}_2) \le I(X_2; Y_1|\tilde{H}_1). \tag{C4b}$$

Then, the rate constraints on (R_1, R_2) are given by (C3).

D) Finding Conditions Under Which the Inner and Outer Bounds Coincide: Comparing (A1), (C3), and (C1), we conclude that if

$$I(X_1, X_3; Y_1|X_2, \tilde{H}_1) \le I(X_1; Y_3|X_2, X_3, \tilde{H}_3) \tag{C5a}$$
$$I(X_2, X_3; Y_2|X_1, \tilde{H}_2) \le I(X_2; Y_3|X_1, X_3, \tilde{H}_3) \tag{C5b}$$
$$I(X_1, X_3; Y_1|X_2, \tilde{H}_1) + I(X_2, X_3; Y_2|X_1, \tilde{H}_2) \le I(X_1, X_2; Y_3|X_3, \tilde{H}_3) \tag{C5c}$$

hold together with (C4) for $X_1 \sim \mathcal{CN}(0, P_1)$, $X_2 \sim \mathcal{CN}(0, P_2)$, and $X_3 \sim \mathcal{CN}(0, P_3)$, mutually independent, then (C3) and (C1) are equal and the relay channel capacity is achieved for both users simultaneously, using DF at the relay. Thus, (C3) characterize the capacity region of the fast fading ICR with fading coefficients having uniformly distributed phases, under VSI.

E) Evaluating the Expressions for Phase Fading: The derivation so far assumed only that each fading coefficient is an i.i.d. process, independent across links, and its phase is uniformly distributed over $[0, 2\pi)$ and independent of its magnitude. We now evaluate the resulting expressions, (C3), (C4), and (C5), for the phase fading case:

a) The rate constraints (C3) evaluate to

$$I(X_1, X_3; Y_1|X_2, \tilde{H}_1) = \log_2\left(1 + P_1 + a_{31}^2 P_3\right) = C_{\mathrm{PF-relay},1} \tag{C6a}$$
$$I(X_2, X_3; Y_2|X_1, \tilde{H}_2) = \log_2\left(1 + P_2 + a_{32}^2 P_3\right) = C_{\mathrm{PF-relay},2} \tag{C6b}$$

which are the capacities of the two relay channels Tx_1–Relay–Rx_1 and Tx_2–Relay–Rx_2, given in (13).

b) The VSI conditions (C4) evaluate to

$$\log_2\left(1 + P_1 + a_{31}^2 P_3\right) \le \log_2\left(1 + \frac{a_{12}^2 P_1}{P_2 + a_{32}^2 P_3 + 1}\right)$$
$$\log_2\left(1 + P_2 + a_{32}^2 P_3\right) \le \log_2\left(1 + \frac{a_{21}^2 P_2}{P_1 + a_{31}^2 P_3 + 1}\right)$$

which give conditions (12). Note that for phase fading, the vectors $\{H_{31,i}X_{3,i}\}_{i=1}^n$ and $\{H_{32,i}X_{3,i}\}_{i=1}^n$ are complex Normal, i.i.d. vectors, independent of the other vectors.

c) Conditions (C5) imply that the decoding region for the destinations is contained in the decoding region at the relay, and therefore, the optimality of the DF scheme follows. It is easy to see that (C5) evaluate under phase fading to (10).

APPENDIX D
PROOF OF THEOREM 2

Steps A–E in the proof of Theorem 1 in Appendix C apply also to the RF-ICR. All is left is to evaluate the mutual information expressions in (C3)–(C5) for the Rayleigh fading model.

Using Gaussian inputs, we get from (C3) the rate bounds

$$R_1 \leq \mathbb{E}_{\tilde{U}}\left\{\log_2\left(1+a_{11}^2|U_{11}|^2P_1+a_{31}^2|U_{31}|^2P_3\right)\right\} \quad \text{(D1a)}$$

$$R_2 \leq \mathbb{E}_{\tilde{U}}\left\{\log_2\left(1+a_{22}^2|U_{22}|^2P_2+a_{32}^2|U_{32}|^2P_3\right)\right\} \quad \text{(D1b)}$$

which are (16). Next, we evaluate condition (C4a). Note that for inputs generated according to independent, circularly symmetric, complex Normal RVs, the channel Tx_1 – Rx_2 is not Gaussian. Writing explicitly $I(X_1; Y_2|\tilde{H}_2)$ we obtain:

$$I(X_1;Y_2|\tilde{H}_2) = \mathbb{E}_{\tilde{U}}\left\{\log_2\left(\frac{1+a_{12}^2|U_{12}|^2P_1+a_{22}^2|U_{22}|^2P_2+a_{32}^2|U_{32}|^2P_3}{1+a_{22}^2|U_{22}|^2P_2+a_{32}^2|U_{32}|^2P_3}\right)\right\}.$$

Condition (C4a) can now be written as

$$\mathbb{E}_{\tilde{U}}\left\{\log_2\left(\frac{1+a_{12}^2|U_{12}|^2P_1+a_{22}^2|U_{22}|^2P_2+a_{32}^2|U_{32}|^2P_3}{1+a_{22}^2|U_{22}|^2P_2+a_{32}^2|U_{32}|^2P_3}\right)\right\} \geq \mathbb{E}_{\tilde{U}}\left\{\log_2(1+a_{11}^2|U_{11}|^2P_1+a_{31}^2|U_{31}|^2P_3)\right\}.$$

Using Jensen's inequality [25, Th. 2.6.2], and repeating steps as in the proof of Corollary 1 in Appendix B, we conclude that this is satisfied if (15a) holds. Using a parallel argument, condition (C4b) is satisfied if (15b) holds.

Finally, we need to evaluate (C5). For Gaussian inputs, (C5) can be written as

$$\mathbb{E}_{\tilde{U}}\left\{\log_2\left(1+a_{11}^2|U_{11}|^2P_1+a_{31}^2|U_{31}|^2P_3\right)\right\} \leq \mathbb{E}_{\tilde{U}}\left\{\log_2\left(1+a_{13}^2|U_{13}|^2P_1\right)\right\} \quad \text{(D2a)}$$

$$\mathbb{E}_{\tilde{U}}\left\{\log_2\left(1+a_{22}^2|U_{22}|^2P_2+a_{32}^2|U_{32}|^2P_3\right)\right\} \leq \mathbb{E}_{\tilde{U}}\left\{\log_2\left(1+a_{23}^2|U_{23}|^2P_2\right)\right\} \quad \text{(D2b)}$$

$$\mathbb{E}_{\tilde{U}}\left\{\log_2\left[(1+a_{11}^2|U_{11}|^2P_1+a_{31}^2|U_{31}|^2P_3)\times(1+a_{22}^2|U_{22}|^2P_2+a_{32}^2|U_{32}|^2P_3)\right]\right\} \leq \mathbb{E}_{\tilde{U}}\left\{\log_2\left(1+a_{13}^2|U_{13}|^2P_1+a_{23}^2|U_{23}|^2P_2\right)\right\}. \quad \text{(D2c)}$$

Repeating the same steps as in the proof of Corollary 1, (D2) can be shown to be satisfied when (14) holds.

APPENDIX E
PROOF OF THEOREM 3

A) Achievability Proof: Let $t_k(R_1,R_2)$ denote the coding scheme (codebooks, encoders, and decoders) that achieves the rate pair (R_1,R_2) at MARC_k, $k=1,2$, and let $t_{DF}(R_1,R_2)$ denote the DF scheme described in Appendix A, with mutually independent channel inputs $X_l \sim \mathcal{CN}(0,P_l)$, $l=1,2,3$, and rates (R_1,R_2). Let $\mathcal{C}_k(\mathbf{a}_k)$ denote that capacity region of MARC_k, $k=1,2$. From [4, Th. 9] it follows that if $\mathbf{a}_k \in \mathcal{D}_k$, the capacity region $\mathcal{C}_k(\mathbf{a}_k)$ is achieved with $t_k(R_1,R_2) = t_{DF}(R_1,R_2)$, $k=1,2$. It is only left to verify that when $\mathbf{a} \in \mathcal{D}_1 \cap \mathcal{D}_2$ then the coding strategy of Appendix A achieves any rate pair (R_1,R_2) that belongs to the capacity region of either MARCs. This is shown next.

To show this, recall that as $\mathbf{a} \in \mathcal{D}_1 \cap \mathcal{D}_2$ the coding scheme that achieves the capacity region of either MARCs follows from [4, Th. 9]. This scheme uses the DF coding strategy $t_{DF}(R_1,R_2)$ of Appendix A. For each rate pair (R_1,R_2) in the capacity region, this scheme uses at each source k, $k=1,2$, a codebook whose cardinality is 2^{nR_k} with codewords generated i.i.d. according to $\mathcal{CN}(0,P_k)$ (the same distributions are used for both MARCs), and at the relay it uses a codebook whose cardinality is $2^{n(R_1+R_2)}$ with codewords generated i.i.d. according to $\mathcal{CN}(0,P_3)$. Hence, for all rate pairs $(R_1,R_2) \in \mathcal{C}_1(\mathbf{a}_1) \cap \mathcal{C}_2(\mathbf{a}_2)$ the same relay codebook construction and scheme are used in both MARCs: at block time b the relay transmits $\mathbf{x}_3(m_{1,b-1},m_{2,b-1})$ and decodes $(m_{1,b},m_{2,b})$. This implies that when $\mathbf{a} \in \mathcal{D}_1 \cap \mathcal{D}_2$, then

$$\mathcal{C}_1(\mathbf{a}_1) \cap \mathcal{C}_2(\mathbf{a}_2) \subseteq \mathcal{C}_{\mathrm{PF-ICR}}(\mathbf{a}). \quad \text{(E1)}$$

Next, note that when $\mathbf{a} \in \mathcal{D}_1 \cap \mathcal{D}_2$, the achievable rate pairs in $\mathcal{C}_1(\mathbf{a}_1) \cap \mathcal{C}_2(\mathbf{a}_2)$ are characterized by the rate constraints of [4, Th. 9] applied to each MARC:

$$R_1 \leq \log_2(1+P_1+a_{31}^2P_3) \quad \text{(E2a)}$$

$$R_2 \leq \log_2(1+a_{21}^2P_2+a_{31}^2P_3) \quad \text{(E2b)}$$

$$R_1+R_2 \leq \log_2(1+P_1+a_{21}^2P_2+a_{31}^2P_3) \quad \text{(E2c)}$$

$$R_1 \leq \log_2(1+a_{12}^2P_1+a_{32}^2P_3) \quad \text{(E2d)}$$

$$R_2 \leq \log_2(1+P_2+a_{32}^2P_3) \quad \text{(E2e)}$$

$$R_1+R_2 \leq \log_2(1+a_{12}^2P_1+P_2+a_{32}^2P_3). \quad \text{(E2f)}$$

Note that when (18) hold, the R.H.S. of (E2d) ≥ R.H.S. (E2a); hence, (E2d) can be dropped. Similarly, (E2b) can be dropped and we obtain that the region in (19) is achievable.

B) Converse Proof: We now characterize SI for PF-ICRs with $\mathbf{a} \in \mathcal{D}_1 \cap \mathcal{D}_2$. Fix a blocklength n and let (R_1,R_2) be an achievable rate pair. When SI occurs, decoding m_2 at Rx_1 and m_1 at Rx_2 can be done without constraining the rates R_1 and R_2.

Note that by definition of SI, we have that if $(R_1,R_2) \in \mathcal{C}_{PF-ICR}(\mathbf{a})$ then $(R_1,R_2) \in \mathcal{C}_1(\mathbf{a}_1)$ and $(R_1,R_2) \in \mathcal{C}_2(\mathbf{a}_2)$. This implies $\mathcal{C}_{PF-ICR}(\mathbf{a}) \subseteq \mathcal{C}_1(\mathbf{a}_1) \cap \mathcal{C}_2(\mathbf{a}_2)$. Combining with (E1) it follows that when SI occurs and $\mathbf{a} \in \mathcal{D}_1 \cap \mathcal{D}_2$, then $\mathcal{C}_{PF-ICR}(\mathbf{a}) = \mathcal{C}_1(\mathbf{a}_1) \cap \mathcal{C}_2(\mathbf{a}_2)$. From the achievability proof of Appendix E-A, it follows that capacity of both MARCs is simultaneously achieved by complex Normal codebooks, DF mapping at the relay, and backward decoding at the destination. Thus, we restrict the attention from now on to mutually independent Gaussian codebooks.

Let $\mathbf{a} \cdot \mathbf{b} \triangleq (a_1b_1, a_2b_2, \ldots, a_nb_n)$, and let (R_1,R_2) be an achievable rate pair. Consider decoding m_2 at Rx_2: receiver Rx_2 decodes its information from $\mathbf{y}_2$

$$\mathbf{y}_2 = \mathbf{h}_{12} \cdot \mathbf{x}_1 + \mathbf{h}_{22} \cdot \mathbf{x}_2 + \mathbf{h}_{32} \cdot \mathbf{x}_3 + \mathbf{z}_2.$$

Define $\mathbf{y}_2'$ as $\mathbf{y}_2' = \mathbf{h}_{22} \cdot \mathbf{x}_2 + \mathbf{h}_{32} \cdot \mathbf{x}_3' + \mathbf{z}_2$, where the relay signal $\mathbf{x}_3'$ is *dedicated to assist in decoding m_2 from $\mathbf{y}_2'$,*

$\mathbb{E}\{|X_3'|^2\} \leq P_3$. Clearly, the maximum possible rate for decoding m_2 at Rx_2 is upper bounded by the rate at which m_2 can be decoded from the enhanced signal $\mathbf{y}_2'$. Recalling that $\mathbf{a}_2 \in \mathcal{D}_2$, this rate is given by the capacity of the phase fading relay channel, see eq. (2), and thus

$$R_2 \leq \log_2\left(1 + P_2 + a_{32}^2 P_3\right) \stackrel{\triangle}{=} R_2'.$$

Next, we turn the attention to Rx_1. Here we carry out two steps:

Step 1: Since R_1 is achievable, Rx_1 can decode its message m_1. Rx_1 can, therefore, create the signal

$$\mathbf{y}_1' = \mathbf{y}_1 - \mathbf{h}_{11} \cdot \mathbf{x}_1 = \mathbf{h}_{21} \cdot \mathbf{x}_2 + \mathbf{h}_{31} \cdot \mathbf{x}_3 + \mathbf{z}_1.$$

Clearly, if Rx_1 can decode m_2 from $\mathbf{y}_1'$ it can decode (m_1, m_2) from $\mathbf{y}_1$.

Step 2: Note that $\mathbf{h}_{21} \cdot \mathbf{x}_2$, $\mathbf{h}_{31} \cdot \mathbf{x}_3$, and $\mathbf{z}_1$, are realizations of i.i.d. zero mean, complex Normal vectors. This follows as $X_{2,i}$ and $X_{3,i}$, $i = 1, 2, \ldots, n$ are i.i.d *zero-mean*, circularly symmetric, complex Normal RVs, hence their amplitudes and phases are mutually independent and the phases are uniformly distributed on $[0, 2\pi)$ [30, Ch. 6–3]. Denote this uniform distribution with $U[0, 2\pi)$. Next we recall that for $\theta_1, \theta_2 \sim U[0, 2\pi)$ then $(\theta_1 + \theta_2)_{\mathrm{mod}\, 2\pi} \sim U[0, 2\pi)$. Hence, multiplying the channel inputs by the phase fading channel coefficients does not change the distribution of the phases of the channel inputs, and therefore the elements of the product remain zero-mean, complex Normal. In particular, $H_{31} X_3 \sim \mathcal{CN}(0, a_{31}^2 P_3)$.

Clearly, using a joint typicality decoder that treats the relay signal $H_{31} X_3$ as an additive i.i.d. complex Normal noise, m_2 can be reliably decoded from $\mathbf{y}_1'$ as long as $R_2 \leq I(X_2; Y_1'|H_{21}) = \log_2\left(1 + \frac{a_{21}^2 P_2}{a_{31}^2 P_3 + 1}\right)$. If this rate bound is greater than R_2', then reliable decoding of m_2 at Rx_2 guarantees decoding of m_2 at Rx_1 with a small probability of error. This leads to the requirement $R_2' \leq \log_2\left(1 + \frac{a_{21}^2 P_2}{a_{31}^2 P_3 + 1}\right)$, which results in (18b). Using a parallel argument, the SI conditions for decoding at Rx_2 is shown to be (18a).

APPENDIX F
PROOF OF THEOREM 4

A) Achievability Proof: Similarly to Appendix E-A, we obtain from Corollary 1 that when $\mathbf{a} \in \mathcal{A}_1 \cap \mathcal{A}_2$, $\mathcal{C}_{\mathrm{MARC}_1}(\mathbf{a}_1) \cap \mathcal{C}_{\mathrm{MARC}_2}(\mathbf{a}_2) \subseteq \mathcal{C}_{\mathrm{RF-ICR}}(\mathbf{a})$, where

$$\mathcal{C}_{\mathrm{MARC_k}}(\mathbf{a}_k) \tag{F1a}$$

$$= \Bigg\{ (R_1, R_2) \in \mathfrak{R}_+^2 :$$

$$R_1 \leq \mathbb{E}_{\tilde{U}}\left\{ \log_2(1 + a_{1k}^2 |U_{1k}|^2 P_1 + a_{3k}^2 |U_{3k}|^2 P_3)\right\} \tag{F1b}$$

$$R_2 \leq \mathbb{E}_{\tilde{U}}\left\{ \log_2(1 + a_{2k}^2 |U_{2k}|^2 P_2 + a_{3k}^2 |U_{3k}|^2 P_3)\right\} \tag{F1c}$$

$$R_1 + R_2 \leq \mathbb{E}_{\tilde{U}}\left\{ \log_2(1 + a_{1k}^2 |U_{1k}|^2 P_1 + a_{2k}^2 |U_{2k}|^2 P_2 + a_{3k}^2 |U_{3k}|^2 P_3)\right\} \Bigg\} \tag{F1d}$$

is the capacity region of the Rayleigh fading MARC with destination Rx_k, where $\mathbf{a}_k \in \mathcal{A}_k$, $k = 1, 2$. Each of these regions is obtained with the same coding strategy $t_{\mathrm{DF}}(R_1, R_2)$ described in Appendix A. In particular, the channel inputs $X_k \sim \mathcal{CN}(0, P_k)$, $k = 1, 2, 3$, are mutually independent. In Appendix F-B1, we show that due to (21), (F1c) (with $k = 1$), and (F1b) (with $k = 2$) can be dropped, thus we obtain (22).

B) Converse Proof: Suppose that (R_1, R_2) is an achievable rate pair. Repeating the argument in Appendix E-B, we conclude that as $\mathbf{a} \in \mathcal{A}_1 \cap \mathcal{A}_2$, then $\mathcal{C}_{RF-ICR}(\mathbf{a}) = \mathcal{C}_{\mathrm{MARC}_1}(\mathbf{a}_1) \cap \mathcal{C}_{\mathrm{MARC}_2}(\mathbf{a}_2)$. Similarly to Appendix E-B, as Rx_2 decodes m_2 from $\mathbf{y}$ it can clearly decode m_2 from the signal $\mathbf{y}_2' = \mathbf{h}_{22} \cdot \mathbf{x}_2 + \mathbf{h}_{31} \cdot \mathbf{x}_3' + \mathbf{z}_2$, where $\mathbf{x}_3'$ is a relay signal dedicated to assist in decoding m_2 from $\mathbf{y}_2'$. This follows since the interference from Tx_1 is eliminated. Recalling that $\mathbf{a}_2 \in \mathcal{A}_2$, the maximum rate for decoding m_2 at Rx_2 is the capacity of the Rayleigh fading relay channel given in (4), and thus

$$R_2 \leq \mathbb{E}_{\tilde{U}}\left\{ \log_2(1 + a_{22}^2 |U_{22}|^2 P_2 + a_{32}^2 |U_{32}|^2 P_3)\right\}. \tag{F2}$$

Next, we turn the attention to Rx_1. As R_1 is achievable, Rx_1 can reliably decode m_1, thus it can construct the signal $\mathbf{y}_1'$ as in Appendix E-B. Recall that $\mathbf{x}_2$, $\mathbf{x}_3$, and $\mathbf{z}_1$ are realizations of independent vectors. Therefore, Rx_1 can treat the relay signal as independent noise and proceed to decode m_2. The total noise signal is then $\mathbf{z}_1' = \mathbf{h}_{31} \cdot \mathbf{x}_3 + \mathbf{z}_1$. Then, using joint-typicality decoding Rx_1 can reliably decode m_2 from $\mathbf{y}_1'$ as long as $R_2 \leq I(X_2; Y_1'|U_{21}) = \mathbb{E}_{U_{21}}\{I(X_2; Y_1'|U_{21} = u_{21})\}$. From the independence of H_{31}, X_3, and Z_1 and the fact that their mean is zero, it follows that the variance of $Z_1' = H_{31} X_3 + Z_1$ is $\mathrm{var}(H_{31})\mathrm{var}(X_3) + \mathrm{var}(Z_1) = a_{31}^2 P_3 + 1$. Clearly, the additive noise Z_1' is not Gaussian; however, as given a fixed noise variance, additive Gaussian noise minimizes the achievable rate over all additive noises, subject to average power constraint at the source [31]; then $I(X_2; Y_1'|U_{21} = u_{21}) \geq \log_2\left(1 + \frac{a_{21}^2 |u_{21}|^2 P_2}{1 + a_{31}^2 P_3}\right)$, and hence it is possible to guarantee successful decoding of m_2 at Rx_1 by further restricting R_2:

$$R_2 \leq \mathbb{E}_{U_{21}}\left\{ \log_2\left(1 + \frac{a_{21}^2 |U_{21}|^2 P_2}{1 + a_{31}^2 P_3}\right)\right\}. \tag{F3}$$

Finally, it is required to verify that when reliable decoding of m_2 at Rx_2 is possible, then so is reliable decoding of m_2 at Rx_1. This is verified if the bound on R_2 in (F3) is greater than the bound on R_2 in (F2), i.e.,

$$\mathbb{E}_{U_{21}}\left\{ \log_2\left(1 + \frac{a_{21}^2 |U_{21}|^2 P_2}{1 + a_{31}^2 P_3}\right)\right\} \geq \mathbb{E}_{\tilde{U}}\left\{ \log_2(1 + a_{22}^2 |U_{22}|^2 P_2 + a_{32}^2 |U_{32}|^2 P_3)\right\}. \tag{F4}$$

Repeating steps similar to those in Appendix B, it is possible to show that (F4) is satisfied if

$$\frac{a_{21}^2 P_2}{e^{\frac{1 + a_{31}^2 P_3}{a_{21}^2 P_2}} E_1\left(\frac{1 + a_{31}^2 P_3}{a_{21}^2 P_2}\right)} \geq (1 + a_{22}^2 P_2 + a_{32}^2 P_3)(1 + a_{31}^2 P_3)$$

resulting in (21b). Condition (21a) is obtained by applying similar arguments to decoding at Rx_2.

1) Eliminating Redundant Constraints: Finally, we note that conditions (21) imply that constraints (F1c) (with $k = 1$ and (F1b) (with $k = 2$) are redundant and can be dropped. To see this for (F1c) (with $k = 1$) recall that (21b) guarantees (F4). Note that if

$$\mathbb{E}_{U_{21}}\left\{\log_2\left(1+\frac{a_{21}^2|U_{21}|^2P_2}{1+a_{31}^2P_3}\right)\right\} \leq \mathbb{E}_{\tilde{U}}\left\{\log_2(1+a_{21}^2|U_{21}|^2P_2+a_{31}^2|U_{31}|^2P_3)\right\}, \quad \text{(F5)}$$

(F1c) (with $k = 2$), and (21b) are satisfied, then (F1c) (with $k = 1$) is satisfied as well. It remains to show (F5) is always true. Since the arguments in the expectations are positive and the logarithm function is monotone increasing, (F5) can be shown by comparing the random expressions directly

$$1+\frac{a_{21}^2|U_{21}|^2P_2}{1+a_{31}^2P_3} \leq 1+a_{21}^2|U_{21}|^2P_2+a_{31}^2|U_{31}|^2P_3$$
$$\Rightarrow a_{21}^2|U_{21}|^2P_2 \leq (a_{21}^2|U_{21}|^2P_2+a_{31}^2|U_{31}|^2P_3)(1+a_{31}^2P_3)$$

which always holds. Thus, the capacity region for SI is characterized by the four constraints in (22).

ACKNOWLEDGMENT

The author would like to thank the associate editor and the reviewers for comments which greatly improved this manuscript.

REFERENCES

[1] C. E. Shannon, "Two-way communication channels," in *Proc. 4th Berkeley Symp. Math., Statist. Probabil.*, Jun./Jul. 1960, vol. 1, pp. 611–644.
[2] E. C. van der Meulen, "Transmission of information in a T-terminal discrete memoryless channel," Ph.D. dissertation, Dept. Statist., Univ. California, Berkeley, Jun. 1968.
[3] O. Sahin, E. Erkip, and O. Simeone, "Interference channel with a relay: Models, relaying strategies, bounds," in *Proc. UCSD Inf. Theory Appl. Workshop*, San Diego, CA, Feb. 2009, pp. 90–95.
[4] 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.
[5] I. Maric, R. Dabora, and A. Goldsmith, "On the capacity of the interference channel with a relay," in *Proc. IEEE Int. Symp. Inf. Theory*, Toronto, ON, Canada, Jul. 2008, pp. 554–558.
[6] B. Djeumou, E. V. Belmega, and S. Lasaulce, "Interference relay channels—Part I: Transmission rates," *IEEE Trans. Commun.*, submitted for publication.
[7] I. Maric, R. Dabora, and A. Goldsmith, "An outer bound for the Gaussian interference channel with a relay," in *Proc. IEEE Inf. Theory Workshop*, Taormina, Italy, Oct. 2009, pp. 569–573.
[8] Y. Tian and A. Yener, "The Gaussian interference relay channel with a potent relay," presented at the Proc. IEEE Global Telecommun. Conf., Honolulu, HI, Dec. 2009.
[9] S. Rini, D. Tuninetti, and N. Devroye, "On the capacity of the interference channel with a cognitive relay," *IEEE Trans. Inf. Theory*, Arxiv identifier 1107.4600v1.
[10] S. Wu and Y. Bar-Ness, "OFDM systems in the presence of phase noise: Consequences and solutions," *IEEE Trans. Commun.*, vol. 52, no. 11, pp. 1988–1996, Nov. 2004.
[11] U. Erez, M. D. Trott, and G. W. Wornell, "Rateless coding and perfect rate-compatible codes for Gaussian channels," in *Proc. IEEE Int. Symp. Inf. Theory*, Seattle, WA, Jul. 2006, pp. 528–532.
[12] B. Sklar, "Rayleigh fading channels in mobile digital communication systems part I: Characterization," *IEEE Commun. Mag.*, vol. 35, no. 7, pp. 90–100, Jul. 1997.
[13] Y. Liang and V. V. Veeravalli, "Cooperative relay broadcast channels," *IEEE Trans. Inf. Theory*, vol. 53, no. 3, pp. 900–928, Mar. 2007.
[14] L. Sankaranarayanan, G. Kramer, and N. B. Mandayam, "Capacity theorems for the multiple-access relay channel," in *Proc. Allerton Conf. Commun., Control Comput.*, Monticello, IL, Sep. 2004, pp. 1782–1791.
[15] C. T. K. Ng, N. Jindal, A. J. Goldsmith, and U. Mitra, "Capacity gain from two-transmitter and two-receiver cooperation," *IEEE Trans. Inf. Theory*, vol. 53, no. 10, pp. 3822–3827, Oct. 2007.
[16] G. Farhadi and N. Beaulieu, "On the ergodic capacity of wireless relaying systems over Rayleigh fading channels," *IEEE Trans. Wireless Commun.*, vol. 7, no. 11, pp. 4462–4467, Nov. 2008.
[17] G. Farhadi and N. C. Beaulieu, "On the ergodic capacity of multi-hop wireless relaying systems," *IEEE Trans. Wireless Commun.*, vol. 8, no. 5, pp. 2286–2291, May 2009.
[18] I. Stanojev, O. Simeone, Y. Bar-Ness, and C. You, "Performance of multi-relay collaborative hybrid-ARQ protocols over fading channels," *IEEE Commun. Lett.*, vol. 10, no. 7, pp. 522–524, Jul. 2006.
[19] A. del Coso Sanchez, "Achievable rates for Gaussian channels with multiple relays," Ph.D. dissertation, Dept. Signal Theory Commun., Univ. Politecnica de Catalunya, Barcelona, Spain, 2008.
[20] C. Edemen and O. Kaya, "Achievable rates for the three user cooperative multiple access channel," in *Proc. IEEE Wireless Commun. Network. Conf.*, Las Vegas, NV, Mar. 2008, pp. 1507–1512.
[21] O. Kaya and S. Ulukus, "Power control for fading cooperative multiple access channels," *IEEE Trans. Wireless Commun.*, vol. 6, no. 8, pp. 2915–2923, Aug. 2007.
[22] L. Sankar, X. Shang, E. Erkip, and H. V. Poor, "Ergodic fading interference channels: Sum-capacity and separability," *IEEE Trans. Inf. Theory*, vol. 57, no. 5, pp. 2605–2626, May 2011.
[23] A. B. Carleial, "A case where interference does not reduce capacity," *IEEE Trans. Inf. Theory*, vol. IT-21, no. 5, pp. 569–570, Sep. 1975.
[24] H. Sato, "The capacity of the Gaussian interference channel under strong interference," *IEEE Trans. Inf. Theory*, vol. IT-27, no. 6, pp. 786–788, Nov. 1981.
[25] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.
[26] F. D. Neeser and J. L. Massey, "Proper complex random processes with applications to information theory," *IEEE Trans. Inf. Theory*, vol. 39, no. 4, pp. 1293–1302, Jul. 1993.
[27] T. M. Cover and A. A. El Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inf. Theory*, vol. IT-25, no. 5, pp. 572–584, Sep. 1979.
[28] D. N. C. Tse and S. V. Hanly, "Multiaccess fading channels—Part I: Polymatroid structure, optimal resource allocation and throughput capacities," *IEEE Trans. Inf. Theory*, vol. 44, no. 7, pp. 2796–2815, Nov. 1998.
[29] R. A. Rankin, *An Introduction to Mathematical Analysis*. New York: Dover, 2007.
[30] A. Papoulis, *Probability, Random Variables and Stochastic Processes*, 2nd ed. New York: McGraw-Hill, 1984.
[31] R. L. Dobrushin, "Optimum information transmission through a channel with unknown parameters," *Radiotekh Elektron.*, vol. 4, no. 12, pp. 1951–1956, 1959.

Ron Dabora (M'03) received his B.Sc. and M.Sc. degrees in 1994 and 2000, respectively, from Tel-Aviv University and his Ph.D. degree in 2007 from Cornell University, all in electrical engineering. From 1994 to 2000 he worked as an engineer at the Ministry of Defense of Israel, and from 2000 to 2003, he was with the Algorithms Group at Millimetrix Broadband Networks, Israel. From 2007 to 2009 he was a postdoctoral researcher at the Department of Electrical Engineering at Stanford University. Since 2009 he is an assistant professor at the Department of Electrical and Computer Engineering, Ben-Gurion University, Israel. His research interests include network information theory, wireless communications and power line communications. He currently serves as an associate editor for the IEEE SIGNAL PROCESSING LETTERS.