Binary Fading Interference Channel with No CSIT

Alireza Vahid, Mohammad Ali Maddah-Ali, A. Salman Avestimehr, and Yan Zhu

Abstract—We study the capacity region of the two-user Binary Fading (or Erasure) Interference Channel where the transmitters have no knowledge of the channel state information. We develop new inner-bounds and outer-bounds for this problem. We identify three regimes based on the channel parameters: weak, moderate, and strong interference regimes. Interestingly, this is similar to the generalized degrees of freedom of the two-user Gaussian interference channel where transmitters have perfect channel knowledge. We show that for the weak interference regime, treating interference as erasure is optimal while for the strong interference regime, decoding interference is optimal. For the moderate interference regime, we provide new inner and outer bounds. The inner-bound is based on a modification of the Han-Kobayashi scheme for the erasure channel, enhanced by time-sharing. We study the gap between our inner-bound and our outer-bounds for the moderate interference regime and compare our results to that of the Gaussian interference channel.

Deriving our new outer-bounds has three main steps. We first create a contracted channel that has fewer states compared to the original channel, in order to make the analysis tractable. We then prove the Correlation Lemma that shows an outer-bound on the capacity region of the contracted channel also serves as an outer-bound for the original channel. Finally using the Conditional Entropy Leakage Lemma, we derive our outer-bound on the capacity region of the contracted channel.

Index Terms—Interference channel, binary fading, capacity, channel state information, no CSIT, packet collision.

I. INTRODUCTION

The two-user Interference Channel (IC) introduced in [2] is a canonical example to study the impact of interference in communication networks. There exists an extensive body of work on this problem under various assumptions (*e.g.*, [3]–[11]). In this work, we focus on a specific configuration of this network, named the two-user Binary Fading Interference Channel (BFIC) as depicted in Fig. 1 in which the channel gains at each time instant are in the binary field according to some Bernoulli distribution. The input-output relation of this channel at time instant t is given by

$$Y_i[t] = G_{ii}[t]X_i[t] \oplus G_{\bar{i}i}[t]X_{\bar{i}}[t], \quad i = 1, 2, \qquad (1)$$

where $\bar{i} = 3 - i$, $G_{ii}[t], G_{\bar{i}i}[t] \in \{0, 1\}$, and all algebraic operations are in $\mathbb{F}_2$. This model was first introduced in [12].

A. Vahid is with the School of Electrical and Computer Engineering, Duke University, Durham, NC, USA. Email: alireza.vahid@duke.edu.

Mohammad Ali Maddah-Ali is with the Department of Electrical Engineering, Sharif University of Technology, Tehran, Iran. Email: maddah_ali@sharif.edu.

A. S. Avestimehr is with the School of Electrical and Computer Engineering, University of Southern California, Los Angeles, CA, USA. Email: avestimehr@ee.usc.edu.

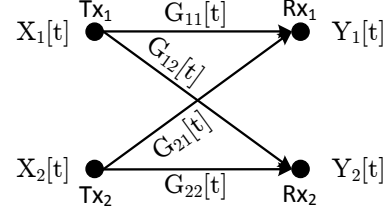Yan Zhu is with Aerohive Networks Inc., Sunnyvale, CA, USA. Email: zhuyan79@gmail.com.

Fig. 1. Two-user Binary Fading Interference Channel (BFIC).

The motivation for studying the Binary Fading (or Erasure) Interference Channel is twofold. First as demonstrated in [13], it provides a simple yet useful physical layer abstraction for wireless packet networks in which whenever a collision occurs, the receiver can store its received analog signal and utilize it for decoding the packets in the future (for example, by successive interference cancellation techniques). In this context, the binary fading model is motivated by a shadow fading environment in which each link is either "on" or "off" (according to the shadow fading distribution), and the multiple access (MAC) is modeled such that if two signals are transmitted simultaneously and the links between the corresponding transmitters and the receiver are not in deep fade, then a linear combination of the signals is available to the receiver. The study of the BFIC in [13] has led to several coding opportunities that can be utilized by the transmitters to exploit the available signal at the receivers for interference management. Moreover, this model allows researchers to focus on other interesting challenges in interference channels such as spatial correlation [14] and locality of channel state knowledge [15], [16].

The second motivation for studying the BFIC is that it can be a first step towards understanding the capacity of the fading interference channels with no knowledge of the channel state information at the transmitters (CSIT). This model was used in [17], [18] to derive the capacity of the one-sided interference channel (also known as Z-Channel). Motivated by the deterministic approach [19], a layered erasure broadcast channel model was introduced in [20] to approximate the capacity of fading broadcast channels. One can view our Binary Fading model as the model introduced in [18], [20] with a single layer.

In this work, we consider the two-user BFIC under the no channel state information at the transmitters assumption. In this no CSIT model, the transmitters are only aware of the distributions of the channel gains but not the actual realizations. We develop new inner-bounds and outer-bounds for this problem and we identify three regimes based on the channel parameters: weak, moderate, and strong interference

regimes. For the weak and the strong interfrence regimes, we show that the entire capacity region is achieved by applying point-to-point erasure codes with appropriate rates at each transmitter, and using either treat-interference-as-erasure or interference-decoding at each receiver. For the moderate interference regime, we provide new inner and outer bounds. The inner-bound is based on a modification of the Han-Kobayashi scheme for the erasure channel, enhanced by time-sharing. In the moderate interference regime, the inner-bounds and the outer-bounds do not match. We provide some further insights and compare this problem to the two-user static (non-fading) Gaussian interference channel where transmitters have perfect channel knowledge.

To derive the outer-bound, we incorporate two key lemmas. The first lemma, the Conditional Entropy Leakage Lemma, establishes how much information is leaked from each transmitter to the unintended receiver. The second lemma, the Correlation Lemma, shows that if the channel gains are correlated under a given set of conditions, the capacity region cannot be smaller than the case of independent channel gains. Using the Correlation Lemma, we create a contracted channel that has fewer states as opposed to the original channel and hence, the problem becomes tractable. Then, using the Conditional Entropy Leakage Lemma, we derive an outer-bound on the capacity region of the contracted channel which in turn, serves as an outer-bound for the original channel.

The rest of the paper is organized as follows. In Section II, we formulate our problem. In Section III, we present our main results. Section IV is dedicated to deriving the outer-bound. We describe our achievability strategy in Section V. Section VI concludes the paper and describes future directions.

II. Problem Setting

We consider the two-user Binary Fading Interference Channel (BFIC) as illustrated in Fig. 1. The channel gain from transmitter Tx_i to receiver Rx_j at time instant t is denoted by $G_{ij}[t]$, $i, j \in \{1, 2\}$. We assume that the channel gains are either 0 or 1 (*i.e.* $G_{ij}[t] \in \{0, 1\}$), and they are distributed as independent Bernoulli random variables (independent from each other and *over time*). Furthermore, we consider the symmetric setting where

$$G_{ii}[t] \overset{d}{\sim} \mathcal{B}(p_d) \quad \text{and} \quad G_{i\bar{i}}[t] \overset{d}{\sim} \mathcal{B}(p_c), \tag{2}$$

for $0 \leq p_d, p_c \leq 1$, $\bar{i} = 3 - i$, and $i = 1, 2$. We define $q_d \overset{\triangle}{=} 1 - p_d$ and $q_c \overset{\triangle}{=} 1 - p_c$.

At each time instant t, the transmit signal at Tx_i is denoted by $X_i[t] \in \{0, 1\}$, $i = 1, 2$, and the received signal at Rx_i is given by

$$Y_i[t] = G_{ii}[t]X_i[t] \oplus G_{i\bar{i}}[t]X_{\bar{i}}[t], \quad i = 1, 2, \tag{3}$$

where all algebraic operations are in $\mathbb{F}_2$. Due to the nature of the channel gains, a total of 16 channel realizations may occur at any given time instant as given in Table I.

The channel state information (CSI) at time instant t is denoted by the quadruple

$$G[t] = (G_{11}[t], G_{12}[t], G_{21}[t], G_{22}[t]). \tag{4}$$

We use the following notations in this paper. We use capital letters to denote random variables (RVs), *e.g.* $G_{ij}[t]$ is a random variable at time instant t, and small letters denote the realizations, *e.g.* $g_{ij}[t]$ is a realization of $G_{ij}[t]$. For a natural number k, we set

$$G^k = [G[1], G[2], \dots, G[k]]^\top. \tag{5}$$

Finally, we set

$$
\begin{aligned}
&G_{ii}^t X_i^t \oplus G_{i\bar{i}}^t X_{\bar{i}}^t \\
&= [G_{ii}[1]X_i[1] \oplus G_{i\bar{i}}[1]X_{\bar{i}}[1], \dots, G_{ii}[t]X_i[t] \oplus G_{i\bar{i}}[t]X_{\bar{i}}[t]]^\top.
\end{aligned} \tag{6}
$$

In this paper, we consider the no CSIT model for the available channel state information at the transmitters. In this model, we assume that transmitters only know the distribution from which the channel gains are drawn, but not the actual realizations of them. Furthermore, we assume that receiver i has instantaneous knowledge of $G_{ii}[t]$ and $G_{\bar{i}i}[t]$ (*i.e.* the incoming links to receiver i), $i = 1, 2$.

Consider the scenario in which Tx_i wishes to reliably communicate message $\mathsf{W}_i \in \{1, 2, \dots, 2^{nR_i}\}$ to Rx_i during n uses of the channel, $i = 1, 2$. We assume that the messages and the channel gains are *mutually* independent and the messages are chosen uniformly. For each transmitter Tx_i, $i = 1, 2$, under no CSIT assumption, let message W_i be encoded as X_i^n using the following equation

$$X_i[t] = f_{i,t}(\mathsf{W}_i), \qquad t = 1, 2, \dots, n, \tag{7}$$

where $f_{i,\cdot}(.)$ is the encoding function at transmitter Tx_i.

Receiver Rx_i is only interested in decoding W_i, and it will decode the message using the decoding function $\widehat{\mathsf{W}}_i = \varphi_i(Y_i^n, G_{ii}^n, G_{\bar{i}i}^n)$. An error occurs when $\widehat{\mathsf{W}}_i \neq \mathsf{W}_i$. The average probability of decoding error is given by

$$\lambda_{i,n} = \mathbb{E}[P[\widehat{\mathsf{W}}_i \neq \mathsf{W}_i]], \quad i = 1, 2, \tag{8}$$

and the expectation is taken with respect to the random choice of the transmitted messages W_1 and W_2. A rate tuple (R_1, R_2) is said to be achievable, if there exist encoding and decoding functions at the transmitters and the receivers respectively, such that the decoding error probabilities $\lambda_{1,n}, \lambda_{2,n}$ go to zero as n goes to infinity. The capacity region $\mathcal{C}(p_d, p_c)$ is the closure of all achievable rate tuples. In the next section, we present the main results of the paper.

III. Main Results

In this section, we present our main contributions. We first need to define the symmetric sum-rate.

Definition 1. *For the two-user BFIC with no CSIT, the symmetric sum-rate is defined as*

$$
\begin{aligned}
R_{\mathrm{sym}}(p_d, p_c) &= R_1 + R_2 \text{ such that} \\
R_1 &= R_2, \quad (R_1, R_2) \in \mathcal{C}(p_d, p_c).
\end{aligned} \tag{9}
$$

The following theorems state our main contributions.

Theorem 1. *For the two-user BFIC with no CSIT the following symmetric sum-rate, $R_{\mathrm{sym}}(p_d, p_c)$, is achievable:*

TABLE I
ALL POSSIBLE CHANNEL REALIZATIONS; SOLID ARROW FROM TRANSMITTER Tx_i TO RECEIVER Rx_j INDICATES THAT $G_{ij}[t] = 1$.

ID	ch. realization	ID	ch. realization	ID	ch. realization	ID	ch. realization
1		2		3		4	
5		6		7		8	
9		10		11		12	
13		14		15		16	

- *For $0 \le p_c \le \frac{p_d}{1+p_d}$:*

$$2p_d q_c;$$

- *For $\frac{p_d}{1+p_d} \le p_c \le p_d$:*

$$p_d + p_c - p_d p_c + \frac{p_d - p_c}{2} C_\delta^*;$$

where

$$C_\delta^* := \frac{p_d p_c - (p_d - p_c)}{p_d p_c - \frac{p_d - p_c}{2}}. \qquad (10)$$

- *For $p_d \le p_c \le 1$:*

$$p_d + p_c - p_d p_c.$$

The following theorem establishes the outer-bound.

Theorem 2. *The capacity region of the two-user BFIC with no CSIT, $\mathcal{C}(p_d, p_c)$, is contained in $\bar{\mathcal{C}}(p_d, p_c)$ given by:*

- *For $0 \le p_c \le \frac{p_d}{1+p_d}$:*

$$\left\{ (R_1, R_2) \middle| \begin{array}{l} 0 \le R_i \le p_d \qquad i = 1, 2 \\ R_i + \beta R_{\bar{i}} \le \beta p_d + p_c - p_d p_c \end{array} \right\} \qquad (11)$$

where

$$\beta = \frac{p_d - p_c}{p_d p_c}. \qquad (12)$$

- *For $\frac{p_d}{1+p_d} \le p_c \le p_d$:*

$$\left\{ (R_1, R_2) \middle| \begin{array}{l} 0 \le R_i \le p_d \qquad i = 1, 2 \\ R_i + R_{\bar{i}} \le 2p_c \\ R_i + \frac{p_d}{p_c} R_{\bar{i}} \le \frac{p_d}{p_c} (p_d + p_c - p_d p_c) \end{array} \right\} \qquad (13)$$

- *For $p_d \le p_c \le 1$:*

$$\left\{ (R_1, R_2) \middle| \begin{array}{l} 0 \le R_i \le p_d \qquad i = 1, 2 \\ R_i + R_{\bar{i}} \le p_d + p_c - p_d p_c \end{array} \right\} \qquad (14)$$
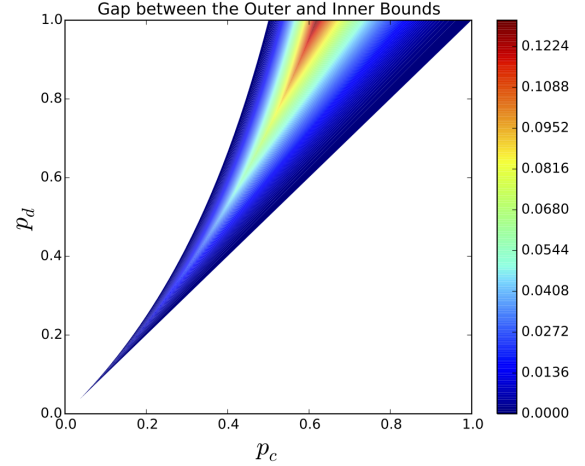

Fig. 2. The gap between the the inner-bounds and the outer-bounds for the two-user BFIC with no CSIT.

As we show later, the inner-bounds and the outer-bounds for the two-user BFIC with no CSIT match when $0 \le p_c \le \frac{p_d}{1+p_d}$ or $p_d \le p_c \le 1$ which correspond to the weak and the strong interference regimes. However, for the moderate interference regime, *i.e.* $\frac{p_d}{1+p_d} \le p_c \le p_d$, the bounds do not meet, thus the capacity region remains open. The gap between the the inner-bounds and the outer-bounds is plotted in Fig. 2 where the bounds match in the white region.

For the weak and the strong interference regimes, the capacity region is obtained by applying point-to-point erasure codes with appropriate rates at each transmitter, and using either treat-interference-as-erasure or interference-decoding at each receiver based on the channel parameters. More precisely for $0 \le p_c < p_d / (1 + p_d)$, the capacity region is obtained by treating interference as erasure, while for $p_d \le p_c \le 1$, the capacity region is obtained by interference-decoding (*i.e.*
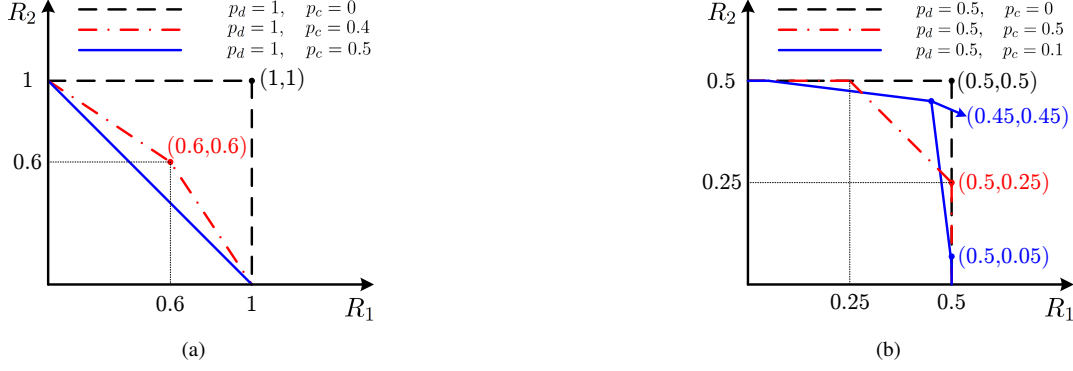
Fig. 3. (a) Capacity region for $p_d = 1$ and $p_c = 0, 0.4, 0.5$; and (b) capacity region for $p_d = 0.5$ and $p_c = 0, 0.1, 0.5$.

the intersection of the capacity regions of the two multiple access channels at receivers). For the moderate interference regime, the inner-bound is based on a modification of the Han-Kobayashi scheme for the erasure channel, enhanced by time-sharing. The detailed proof of the achievability strategy can be found in Section V.

Fig. 3 illustrates the capacity region for several values of p_d and p_c. In Fig. 3(a) the capacity region is depicted for $p_d = 1$ and $p_c = 0, 0.4, 0.5$. From Fig. 3(b), we conclude that unlike the case in Fig. 3(a), decreasing p_c does not necessarily enlarge the capacity region. In fact, for $p_d = 0.5$, we have

$$\mathcal{C}(0.5, 0.5) \nsubseteq \mathcal{C}(0.5, 0.1),$$
$$\mathcal{C}(0.5, 0.1) \nsubseteq \mathcal{C}(0.5, 0.5). \tag{15}$$

Using the results of [21], we have plotted the capacity region of the two-user Binary Fading IC for $p_d = p_c = 0.5$ under three different scenarios in Fig. 4. Under the delayed CSIT, we assume that transmitters become aware of all channel realizations with unit delay and receivers have access to instantaneous CSI, *i.e.*

$$X_i[t] = f_{i,t}\left(\mathrm{W}_i, G^{t-1}\right), \qquad t = 1, 2, \dots, n, \tag{16}$$

and

$$\widehat{\mathrm{W}}_i = \varphi_i\left(Y_i^n, G^n\right); \tag{17}$$

and under instantaneous CSIT model, we assume that at time instant t, transmitters and receivers have access to G^t. As we can see in Fig. 4, when transmitters have access to the delayed knowledge of *all* links in the network, the capacity region is strictly larger than the capacity region under the no CSIT assumption.

To prove Theorem 2, we incorporate two key lemmas as discussed in Section IV-A. The first step in obtaining the outer-bound is to create a "contracted" channel that has fewer states compared to the original channel. Using the Correlation Lemma, we show that an outer-bound on the capacity region of the contracted channel also serves as an outer-bound for the original channel. Finally, using the Conditional Entropy Leakage Lemma, we derive this outer-bound.

The intuition for the Correlation Lemma was first provided by Sato [4]: the capacity region of all interference channels that have the same marginal distributions is the same. We take
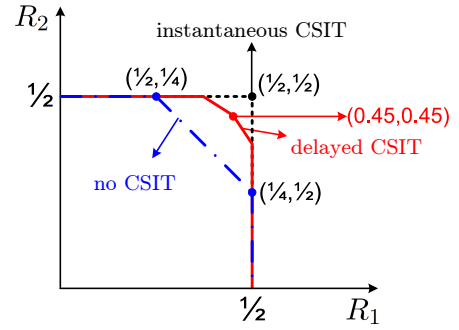

Fig. 4. *Capacity region of the two-user Binary Fading IC for $p_d = p_c = 0.5$, with no CSIT, delayed CSIT, and instantaneous CSIT.*

this intuition and impose a certain spatial correlation among channel gains such that the marginal distributions remain unchanged. For the moderate interference regime, we also incorporate the outer-bounds on the capacity region of the one-sided BFIC with no CSIT [18].

The rest of the paper is dedicated to the proof of our main contributions. We provide the converse proof in Section IV, and we present our achievability strategy in Section V.

IV. CONVERSE

In this section, we provide the converse proof of Theorem 2 for the no CSIT assumption. We incorporate two lemmas in order to derive the outer-bound that we describe in the following subsection.

A. Key Lemmas

1) Entropy Leakage Lemma: Consider a broadcast channel as depicted in Fig. 5 where a transmitter is connected to two receivers through binary fading channels. Suppose $G_i[t]$ is distributed as i.i.d. Bernoulli random variable (*i.e.* $G_i[t] \overset{d}{\sim} \mathcal{B}(p_i)$) where $0 \le p_2 \le p_1 \le 1$, $i = 1, 2$. In this channel the received signals are given as

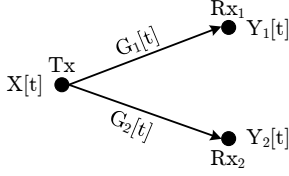$$Y_i[t] = G_i[t]X[t], \qquad i = 1, 2, \tag{18}$$

Fig. 5. A broadcast channel with binary fading links.

where $X[t] \in \{0,1\}$ is the transmit signal at time instant t. Furthermore, suppose $G_3[t] \overset{d}{\sim} \mathcal{B}(p_3)$, $0 \le p_3 \le 1$, such that

$$\Pr[G_i[t] = 1, G_j[t] = 1] = 0, \; i \ne j, \; i,j \in \{1,2,3\}. \quad (19)$$

We further define

$$G_T[t] \overset{\triangle}{=} (G_1[t], G_2[t], G_3[t]). \quad (20)$$

Then, for the channel described above, we have the following lemma.

Lemma 1. *[Conditional Entropy Leakage Lemma] For the channel described above with no CSIT, for $p_1 + p_2 + p_3 \le 1$ and* any *input distribution, we have*

$$H(Y_2^n | G_3^n X^n, G_T^n) \ge \frac{p_2}{p_1} H(Y_1^n | G_3^n X^n, G_T^n). \quad (21)$$

Proof. Let $G_H[t]$ be distributed as $\mathcal{B}(p_2/p_1)$, and be independent of all other parameters in the network. Let

$$Y_H[t] = G_H[t] Y_1[t], \quad t = 1, \dots, n. \quad (22)$$

It is straightforward to see that Y_H^t is statistically the same as Y_2^t under the no CSIT assumption. For time instant t, where $1 \le t \le n$, we have

$H(Y_2[t] | Y_2^{t-1}, G_3^n X^n, G_T^n)$

$\overset{(a)}{=} H(Y_2[t] | Y_2^{t-1}, G_3^n X^n, G_T^n, G_H^{t-1})$

$\overset{(b)}{=} p_2 H(X[t] | Y_2^{t-1}, G_3^n X^n, G_T^n, G_H^{t-1}, G_2[t] = 1, G_3[t] = 0)$

$\overset{(c)}{=} p_2 H(X[t] | Y_2^{t-1}, G_3^n X^n, G_T^n, G_H^{t-1}, G_3[t] = 0)$

$\overset{(d)}{=} p_2 H(X[t] | Y_H^{t-1}, G_3^n X^n, G_T^n, G_H^{t-1}, G_3[t] = 0)$

$\overset{(e)}{\ge} p_2 H(X[t] | Y_1^{t-1}, Y_H^{t-1}, G_3^n X^n, G_T^n, G_H^{t-1}, G_3[t] = 0)$

$\overset{(f)}{=} p_2 H(X[t] | Y_1^{t-1}, G_3^n X^n, G_T^n, G_H^{t-1}, G_3[t] = 0)$

$\overset{(g)}{=} (1 - p_3) \frac{p_2}{p_1} H(Y_1[t] | Y_1^{t-1}, G_3^n X^n, G_T^n, G_H^{t-1}, G_3[t] = 0)$

$\overset{(h)}{=} \frac{p_2}{p_1} H(Y_1[t] | Y_1^{t-1}, G_3^n X^n, G_T^n, G_H^{t-1})$

$\overset{(i)}{=} \frac{p_2}{p_1} H(Y_1[t] | Y_1^{t-1}, G_3^n X^n, G_T^n), \quad (23)$

where (a) holds since G_H^{t-1} is independent of all other parameters in the network; (b) follows from

$$\Pr[G_2[t] = 1 | G_3[t] = 0] = \frac{p_2}{1 - p_3}; \quad (24)$$

(c) holds since the transmit signal is independent of the channel realizations; (d) follows from the fact that

$$\Pr[X[t], Y_2^{t-1}, G_3^n X^n, G_T^n, G_H^{t-1}, G_3[t] = 0]$$
$$= \Pr[X[t], Y_H^{t-1}, G_3^n X^n, G_T^n, G_H^{t-1}, G_3[t] = 0]. \quad (25)$$

This equality holds since starting from one side a simple index exchange of $2 \leftrightarrow H$ gives the other side and $X[t]$ is oblivious of the channel realizations; (e) holds since conditioning reduces entropy; (f) is true since

$$H(Y_H^{t-1} | Y_1^{t-1}, G_3^n X^n, G_T^n, G_H^{t-1}) = 0; \quad (26)$$

(g) follows from the fact that

$$\Pr[G_1[t] = 1 | G_3[t] = 0] = \frac{p_1}{1 - p_3}; \quad (27)$$

(h) is true since $\Pr[G_3[t] = 0] = 1 - p_3$; and (i) holds since G_H^{t-1} is independent of all other parameters in the network. Thus, summing all terms for $t = 1, \dots, n$, we get

$$\sum_{t=1}^{n} H(Y_2[t] | Y_2^{t-1}, G_3^n X^n, G_T^n)$$
$$\ge \frac{p_2}{p_1} \sum_{t=1}^{n} H(Y_1[t] | Y_1^{t-1}, G_3^n X^n, G_T^n), \quad (28)$$

which implies

$$H(Y_2^n | G_3^n X^n, G_T^n) \ge \frac{p_2}{p_1} H(Y_1^n | G_3^n X^n, G_T^n), \quad (29)$$

hence, completing the proof. $\qquad \square$

Remark 1. *In [21], we drived the Entropy Leakage Lemma for the case where the transmitter has the CSI with delay. We observe that, with delayed CSIT, the constant on the RHS of (21) would be smaller, meaning that the transmitter can further favor the stronger receiver.*

2) Correlation Lemma: Consider again a binary fading interference channel similar to the channel described in Section II, but where channel gains have certain correlation. We denote the channel gain from transmitter Tx_i to receiver Rx_j at time instant t by $\tilde{G}_{ij}[t]$, $i,j \in \{1,2\}$. We distinguish the RVs in this channel, using $(\tilde{\cdot})$ notation (*e.g.*, $\tilde{X}_1[t]$). The input-output relation of this channel at time instant t is given by

$$\tilde{Y}_i[t] = \tilde{G}_{ii}[t] \tilde{X}_i[t] \oplus \tilde{G}_{\bar{i}i}[t] \tilde{X}_{\bar{i}}[t], \quad i = 1, 2. \quad (30)$$

We assume that the channel gains are distributed independent over time. However, they can be arbitrary correlated with each other subject to the following constraints.

$$\Pr(\tilde{G}_{ii}[t] = 1) = p_d, \quad \Pr(\tilde{G}_{\bar{i}i}[t] = 1) = p_c,$$
$$\Pr(\tilde{G}_{ii}[t] = 1, \tilde{G}_{\bar{i}i}[t] = 1)$$
$$= \Pr(\tilde{G}_{ii}[t] = 1) \Pr(\tilde{G}_{\bar{i}i}[t] = 1), \quad i = 1, 2. \quad (31)$$

In other words, the channel gains corresponding to incoming links at each receiver are still independent. Similar to the original channel, we assume that the transmitters in this BFIC have no knowledge of the CSI. We have the following result.

Lemma 2. *[Correlation Lemma] For any BFIC that satisfies the constraints in (31), we have*

$$\mathcal{C}(p_d, p_c) \equiv \tilde{\mathcal{C}}(p_d, p_c). \quad (32)$$

Proof. Suppose in the original BFIC messages W_1 and W_2 are encoded as X_1^n and X_2^n respectively, and each receiver

can decode its corresponding message with arbitrary small decoding error probability as $n \to \infty$. Now, we show that if we use the same transmission scheme in the BFIC that satisfies the constraints in (31), *i.e.*

$$\tilde{X}_i[t] = X_i[t], \quad t = 1, 2, \ldots, n, \quad i = 1, 2, \quad (33)$$

then the receivers in this BFIC can still decode W_1 and W_2.

In the original two-user binary fading IC as described in Section II, Rx_i uses the decoding function $\widehat{W}_i = \varphi_i(Y_i^n, G_{ii}^n, G_{i\bar{i}}^n)$. Therefore, the error event $\widehat{W}_i \neq W_i$, only depends on the choice of W_i and marginal distribution of the channel gains G_{ii}^n and $G_{i\bar{i}}^n$.

Define

$$\mathcal{E}_{W_1} = \left\{ (W_2, G_{11}^n, G_{21}^n) \text{ s.t. } \widehat{W}_1 \neq W_1 \right\},$$
$$\tilde{\mathcal{E}}_{W_1} = \left\{ (W_2, \tilde{G}_{11}^n, \tilde{G}_{21}^n) \text{ s.t. } \widehat{W}_1 \neq W_1 \right\}, \quad (34)$$

then, the probability of error is given by

$$p_{\text{error}} = \sum_{w_1} \Pr(W_1 = w_1) \Pr(\mathcal{E}_{w_1})$$
$$= \frac{1}{2^{nR_1}} \sum_{w_1} \sum_{(w_2, g_{11}^n, g_{21}^n) \in \mathcal{E}_{w_1}} \Pr(w_2, g_{11}^n, g_{21}^n) \quad (35)$$
$$\stackrel{(a)}{=} \frac{1}{2^{nR_1}} \sum_{w_1} \sum_{(w_2, \tilde{g}_{11}^n, \tilde{g}_{21}^n) \in \tilde{\mathcal{E}}_{w_1}} \Pr(w_2, \tilde{g}_{11}^n, \tilde{g}_{21}^n) = \tilde{p}_{\text{error}},$$

where p_{error} and $\tilde{p}_{\text{error}}$ are the decoding error probability at Rx_1 in the original and the BFIC satisfying the constraints in (31) respectively; and (a) holds since according to (31), the joint distribution of G_{12}^n and G_{22}^n is the same as $\tilde{G}_{12}^n$ and $\tilde{G}_{22}^n$ and the fact that, as mentioned above, the error probability at receiver one only depends on the marginal distribution of these links. Similar argument holds for Rx_2. $\qquad \square$

B. Deriving the Outer-bounds

The key to derive the outer-bound is the proper application of the two lemmas introduced in Section IV-A. More precisely, we need to find a channel that satisfies the constraints in (31) such that the outer-bound on its capacity region, coincides with the achievable region of the original problem. We provide the proof for three separate regimes.

• **Regime I:** $0 \leq p_c \leq p_d/(1 + p_d)$: The derivation of the individual bounds, *e.g.*, $R_1 \leq p_d$, is straightforward and omitted here. We focus on

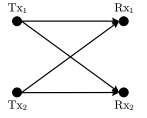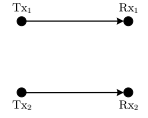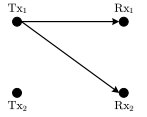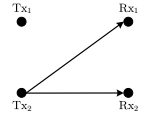$$R_1 + \beta R_2 \leq \beta p_d + p_c - p_d p_c, \quad (36)$$

where

$$\beta = \max\left\{ \frac{p_d - p_c}{p_d p_c}, 1 \right\}. \quad (37)$$

Due to symmetry, the derivation of the other bound is similar.

The first step is to define the appropriate channel that satisfies the constraints in (31). The idea is to construct a channel such that $\tilde{G}_{ii}[t] = 1$ whenever $\tilde{G}_{i\bar{i}}[t] = 1$, $i = 1, 2$. We construct such channel with only five states rather than 16 states and thus, we refer to it as "contracted" channel. The five states are denoted by states $A, B, C, D,$ and E

TABLE II
THE CONTRACTED CHANNEL FOR **Regime I** AND **Regime II**.

ID	channel realization	ID	channel realization
A		B	
	$\Pr[\text{state A}] = p_d p_c$		$\Pr[\text{state B}] = p_d - p_c$
C		D	
	$\Pr[\text{state C}] = q_d p_c$		$\Pr[\text{state D}] = q_d p_c$

with corresponding probabilities $p_d p_c, (p_d - p_c), q_d p_c, q_d p_c$, and $q_d q_c$. These states are depicted in Table II with the exception of state E which corresponds to the case where all channel gains are 0. Here, we have

$$\Pr\left(\tilde{G}_{11}[t] = 1\right) = \sum_{j \in \{A,B,C\}} \Pr(\text{State j}) = p_d,$$
$$\Pr\left(\tilde{G}_{12}[t] = 1\right) = \sum_{j \in \{A,C\}} \Pr(\text{State j}) = p_c,$$
$$\Pr\left(\tilde{G}_{11}[t] = 1, \tilde{G}_{21}[t] = 1\right) = \Pr(\text{State A}) = p_d p_c, \quad (38)$$

thus, this channel satisfies the conditions in (31). From Lemma 2, we have $\mathcal{C}(p_d, p_c) \subseteq \tilde{\mathcal{C}}(p_d, p_c)$, thus, any outer-bound on the capacity region of the contracted channel, provides an outer-bound on the capacity region of the original channel.

We define

$$\tilde{X}_{1A}[t] \triangleq \tilde{X}_1[t] \mathbf{1}_{\{\text{state } A \text{ occurs at time } t\}}, \quad (39)$$

where $\mathbf{1}_{\{\text{state } A \text{ occurs at time } t\}}$ is equal to 1 when at time t state A occurs. Similarly, we define $\tilde{X}_{1B}[t], \tilde{X}_{1C}[t], \tilde{X}_{1D}[t], \tilde{X}_{1E}[t], \tilde{X}_{2A}[t], \tilde{X}_{2B}[t], \tilde{X}_{2C}[t], \tilde{X}_{2D}[t]$ and $\tilde{X}_{2E}[t]$. Therefore, we have

$$\tilde{X}_1[t] = \tilde{X}_{1A}[t] \oplus \tilde{X}_{1B}[t] \oplus \tilde{X}_{1C}[t] \oplus \tilde{X}_{1D}[t] \oplus \tilde{X}_{1E}[t]. \quad (40)$$

Suppose in the contracted channel, there exist encoders and decoders at transmitters and receivers respectively, such that each receiver can decode its corresponding message with arbitrary small decoding error probability as $\epsilon_n \to 0$. The derivation of the outer-bound is given in (42) for $\epsilon_n \to 0$ as $n \to \infty$; and

$$0 \leq p_c \leq p_d/(1 + p_d) \Rightarrow \beta = \frac{p_d - p_c}{p_d p_c} > 1. \quad (41)$$

We have

$$n\left(\tilde{R}_1 + \beta\tilde{R}_2 - \epsilon_n\right)$$

$$\overset{(a)}{\leq} I\left(\tilde{X}_1^n; \tilde{Y}_1^n|\tilde{G}^n\right) + \beta I\left(\tilde{X}_2^n; \tilde{Y}_2^n|\tilde{G}^n\right)$$

$$\overset{(b)}{=} H\left(\tilde{X}_{2D}^n|\tilde{G}^n\right) + H\left(\tilde{X}_{1B}^n, \tilde{X}_{1C}^n|\tilde{X}_{2D}^n, \tilde{G}^n\right)$$

$$+ H\left(\tilde{X}_{1A}^n \oplus \tilde{X}_{2A}^n|\tilde{X}_{1B}^n, \tilde{X}_{1C}^n, \tilde{X}_{2D}^n, \tilde{G}^n\right)$$

$$- H\left(\tilde{X}_{2D}^n|\tilde{X}_1^n, \tilde{G}^n\right) - H\left(\tilde{X}_{2A}^n|\tilde{X}_{2D}^n, \tilde{X}_1^n, \tilde{G}^n\right)$$

$$+ \beta H\left(\tilde{X}_{1C}^n|\tilde{G}^n\right) + \beta H\left(\tilde{X}_{2B}^n, \tilde{X}_{2D}^n|\tilde{X}_{1C}^n, \tilde{G}^n\right)$$

$$+ \beta H\left(\tilde{X}_{1A}^n \oplus \tilde{X}_{2A}^n|\tilde{X}_{2B}^n, \tilde{X}_{2D}^n, \tilde{X}_{1C}^n, \tilde{G}^n\right)$$

$$- \beta H\left(\tilde{X}_{1C}^n|\tilde{X}_2^n, \tilde{G}^n\right) - \beta H\left(\tilde{X}_{1A}^n|\tilde{X}_{1C}^n, \tilde{X}_2^n, \tilde{G}^n\right)$$

$$\overset{(c)}{\leq} H\left(\tilde{X}_{2D}^n|\tilde{G}^n\right) + H\left(\tilde{X}_{1B}^n, \tilde{X}_{1C}^n|\tilde{G}^n\right) + H\left(\tilde{X}_{1A}^n \oplus \tilde{X}_{2A}^n|\tilde{G}^n\right)$$

$$- H\left(\tilde{X}_{2D}^n|\tilde{G}^n\right) - H\left(\tilde{X}_{2A}^n|\tilde{X}_{2D}^n, \tilde{G}^n\right) + \beta H\left(\tilde{X}_{1C}^n|\tilde{G}^n\right)$$

$$+ \beta H\left(\tilde{X}_{2B}^n, \tilde{X}_{2D}^n|\tilde{G}^n\right) + \beta H\left(\tilde{X}_{1A}^n \oplus \tilde{X}_{2A}^n|\tilde{G}^n\right)$$

$$- \beta H\left(\tilde{X}_{1C}^n|\tilde{G}^n\right) - \beta H\left(\tilde{X}_{1A}^n|\tilde{X}_{1C}^n, \tilde{G}^n\right)$$

$$\overset{(d)}{=} H\left(\tilde{X}_{1C}^n|\tilde{G}^n\right) + H\left(\tilde{X}_{1B}^n|\tilde{X}_{1C}^n, \tilde{G}^n\right)$$

$$+ (1+\beta) H\left(\tilde{X}_{1A}^n \oplus \tilde{X}_{2A}^n|\tilde{G}^n\right) - \beta H\left(\tilde{X}_{1A}^n|\tilde{X}_{1C}^n, \tilde{G}^n\right)$$

$$+ \beta H\left(\tilde{X}_{2B}^n, \tilde{X}_{2D}^n|\tilde{G}^n\right) - H\left(\tilde{X}_{2A}^n|\tilde{X}_{2D}^n, \tilde{G}^n\right)$$

$$\overset{(e)}{\leq} H\left(\tilde{X}_{1C}^n|\tilde{G}^n\right) + (1+\beta) H\left(\tilde{X}_{1A}^n \oplus \tilde{X}_{2A}^n|\tilde{G}^n\right)$$

$$+ \beta H\left(\tilde{X}_{2D}^n|\tilde{G}^n\right) + \beta H\left(\tilde{X}_{2B}^n|\tilde{X}_{2D}^n, \tilde{G}^n\right) - H\left(\tilde{X}_{2A}^n|\tilde{X}_{2D}^n, \tilde{G}^n\right)$$

$$\overset{(f)}{\leq} H\left(\tilde{X}_{1C}^n|\tilde{G}^n\right) + (1+\beta) H\left(\tilde{X}_{1A}^n \oplus \tilde{X}_{2A}^n|\tilde{G}^n\right)$$

$$+ \beta H\left(\tilde{X}_{2D}^n|\tilde{G}^n\right) + \left(\beta - \frac{1}{\beta}\right) H\left(\tilde{X}_{2B}^n|\tilde{X}_{2D}^n, \tilde{G}^n\right)$$

$$\overset{(g)}{\leq} nq_dp_c + n(1+\beta) p_dp_c + n\beta q_dp_c + n\left(\beta - \frac{1}{\beta}\right)(p_d - p_c)$$

$$= n(\beta p_d + p_c - p_dp_c), \tag{42}$$

where (a) follows from Fano's inequality and data processing inequality; (b) follows from the definition of the contracted channel and the chain rule; (c) is true since from Claim 1 below, we have

$$I\left(\tilde{X}_1^n; \tilde{X}_2^n|\tilde{G}^n\right) = 0, \tag{43}$$

which results in

$$H\left(\tilde{X}_{2A}^n|\tilde{X}_{2D}^n, \tilde{X}_1^n, \tilde{G}^n\right) = H\left(\tilde{X}_{2A}^n|\tilde{X}_{2D}^n, \tilde{G}^n\right),$$
$$H\left(\tilde{X}_{1A}^n|\tilde{X}_{1C}^n, \tilde{X}_2^n, \tilde{G}^n\right) = H\left(\tilde{X}_{1A}^n|\tilde{X}_{1C}^n, \tilde{G}^n\right); \tag{44}$$

and the fact that conditioning reduces entropy; (d) follows from the chain rule; (e) holds since using Lemma 1, we have (use analogy: $\tilde{X}_{1B}^n \leftrightarrow Y_2^n$, $\tilde{X}_{1C}^n \leftrightarrow G_3^n X^n$, and $\tilde{X}_{1A}^n \leftrightarrow Y_1^n$)

$$H\left(\tilde{X}_{1B}^n|\tilde{X}_{1C}^n, \tilde{G}^n\right) - \beta H\left(\tilde{X}_{1A}^n|\tilde{X}_{1C}^n, \tilde{G}^n\right) \leq 0; \tag{45}$$

and (f) follows since from applying Lemma 1, we have (use analogy: $\tilde{X}_{2A}^n \leftrightarrow Y_1^n$, $\tilde{X}_{2D}^n \leftrightarrow G_3^n X^n$, and $\tilde{X}_{1B}^n \leftrightarrow Y_2^n$)

$$H\left(\tilde{X}_{2A}^n|\tilde{X}_{2D}^n, \tilde{G}^n\right) \geq \frac{1}{\beta} H\left(\tilde{X}_{2B}^n|\tilde{X}_{2D}^n, \tilde{G}^n\right); \tag{46}$$

and (g) follows from the fact that entropy of a binary random variable is maximized by i.i.d. Bernoulli distribution with success probability of half. Dividing both sides by n and let $n \to \infty$, we get the desired result.

Claim 1.

$$I\left(\tilde{X}_1^n; \tilde{X}_2^n|\tilde{G}^n\right) = 0. \tag{47}$$

Proof.

$$0 \leq I\left(\tilde{X}_1^n; \tilde{X}_2^n|\tilde{G}^n\right) \leq I\left(\tilde{W}_1, \tilde{X}_1^n; \tilde{W}_2, \tilde{X}_2^n|\tilde{G}^n\right)$$

$$= I\left(\tilde{W}_1; \tilde{W}_2|\tilde{G}^n\right) + I\left(\tilde{W}_1; \tilde{X}_2^n|\tilde{W}_2, \tilde{G}^n\right)$$

$$+ I\left(\tilde{X}_1^n; \tilde{W}_2, \tilde{X}_2^n|\tilde{W}_1, \tilde{G}^n\right) = 0. \tag{48}$$

where the last equality holds since

$$I\left(\tilde{W}_1; \tilde{W}_2|\tilde{G}^n\right) = 0, \tag{49}$$

due to the fact that the messages and the channel gains are mutually independent;

$$I\left(\tilde{W}_1; \tilde{X}_2^n|\tilde{W}_2, \tilde{G}^n\right) = 0, \tag{50}$$

due to the fact that $\tilde{X}_2^n = f_2(\tilde{W}_2, \tilde{G}^n)$; and

$$I\left(\tilde{X}_1^n; \tilde{W}_2, \tilde{X}_2^n|\tilde{W}_1, \tilde{G}^n\right), \tag{51}$$

due to the fact that $\tilde{X}_1^n = f_1(\tilde{W}_1, \tilde{G}^n)$. $\qquad\square$

• **Regime II:** $p_d/(1 + p_d) \leq p_c \leq p_d$: In this regime for $i = 1, 2$, we borrow the outer-bound

$$R_i + \frac{p_dp_c}{p_d - p_c + p_dp_c}R_{\bar{i}} \leq p_d + \frac{p_dp_c}{p_d - p_c + p_dp_c}p_cq_d, \tag{52}$$

from [18]. Thus, we focus on

$$R_i + R_{\bar{i}} \leq 2p_c, \qquad i = 1, 2. \tag{53}$$

We use the same contracted channel as for the case of Regime I. Let $G_S[t]$ be distributed as i.i.d. Bernoulli RV and

$$G_S[t] \overset{d}{\sim} \mathcal{B}(\frac{p_d - p_c}{p_dp_c}), \tag{54}$$

and for $i = 1, 2$, we define

$$\bar{X}_{iA}[t] \triangleq G_S[t]\tilde{X}_{iA}[t],$$
$$\hat{X}_{iA}[t] \triangleq (1 - G_S[t])\tilde{X}_{iA}[t], \tag{55}$$

We have

$$n\left(\tilde{R}_1 + \tilde{R}_2 - \epsilon_n\right) \overset{(a)}{\leq} I\left(\tilde{X}_1^n; \tilde{Y}_1^n | \tilde{G}^n\right) + I\left(\tilde{X}_2^n; \tilde{Y}_2^n | \tilde{G}^n\right)$$

$$\overset{(b)}{=} H\left(\tilde{X}_{1A}^n \oplus \tilde{X}_{2A}^n, \tilde{X}_{1B}^n, \tilde{X}_{1C}^n, \tilde{X}_{2D}^n | \tilde{G}^n\right)$$

$$- H\left(\tilde{X}_{2A}^n, \tilde{X}_{2D}^n | \tilde{X}_1^n, \tilde{G}^n\right)$$

$$+ H\left(\tilde{X}_{1A}^n \oplus \tilde{X}_{2A}^n, \tilde{X}_{2B}^n, \tilde{X}_{1C}^n, \tilde{X}_{2D}^n | \tilde{G}^n\right)$$

$$- H\left(\tilde{X}_{1A}^n, \tilde{X}_{1C}^n | \tilde{X}_2^n, \tilde{G}^n\right)$$

$$\overset{(c)}{=} H\left(\tilde{X}_{2D}^n | \tilde{G}^n\right) + H\left(\tilde{X}_{1B}^n, \tilde{X}_{1C}^n | \tilde{X}_{2D}^n, \tilde{G}^n\right)$$

$$+ H\left(\tilde{X}_{1A}^n \oplus \tilde{X}_{2A}^n | \tilde{X}_{1B}^n, \tilde{X}_{1C}^n, \tilde{X}_{2D}^n, \tilde{G}^n\right)$$

$$- H\left(\tilde{X}_{2D}^n | \tilde{X}_1^n, \tilde{G}^n\right) - H\left(\tilde{X}_{2A}^n | \tilde{X}_1^n, \tilde{X}_{2D}^n, \tilde{G}^n\right)$$

$$+ H\left(\tilde{X}_{1C}^n | \tilde{G}^n\right) + H\left(\tilde{X}_{2B}^n, \tilde{X}_{2D}^n | \tilde{X}_{1C}^n, \tilde{G}^n\right)$$

$$+ H\left(\tilde{X}_{1A}^n \oplus \tilde{X}_{2A}^n | \tilde{X}_{2B}^n, \tilde{X}_{2D}^n, \tilde{X}_{1C}^n, \tilde{G}^n\right)$$

$$- H\left(\tilde{X}_{1C}^n | \tilde{X}_2^n, \tilde{G}^n\right) - H\left(\tilde{X}_{1A}^n | \tilde{X}_{1C}^n, \tilde{X}_2^n, \tilde{G}^n\right)$$

$$\overset{(d)}{=} H\left(\tilde{X}_{2D}^n | \tilde{G}^n\right) + H\left(\tilde{X}_{1B}^n, \tilde{X}_{1C}^n | \tilde{G}^n\right)$$

$$+ H\left(\tilde{X}_{1A}^n \oplus \tilde{X}_{2A}^n | \tilde{X}_{1B}^n, \tilde{X}_{1C}^n, \tilde{X}_{2D}^n, \tilde{G}^n\right)$$

$$- H\left(\tilde{X}_{2D}^n | \tilde{G}^n\right) - H\left(\tilde{X}_{2A}^n | \tilde{X}_{2D}^n, \tilde{G}^n\right)$$

$$+ H\left(\tilde{X}_{1C}^n | \tilde{G}^n\right) + H\left(\tilde{X}_{2B}^n, \tilde{X}_{2D}^n | \tilde{G}^n\right)$$

$$+ H\left(\tilde{X}_{1A}^n \oplus \tilde{X}_{2A}^n | \tilde{X}_{2B}^n, \tilde{X}_{2D}^n, \tilde{X}_{1C}^n, \tilde{G}^n\right)$$

$$- H\left(\tilde{X}_{1C}^n | \tilde{G}^n\right) - H\left(\tilde{X}_{1A}^n | \tilde{X}_{1C}^n, \tilde{G}^n\right)$$

where $\epsilon_n \to 0$ as $n \to \infty$; (a) follows from Fano's inequality and data processing inequality; (b) holds due to the definition of the contracted channel; (c) follows from the chain rule; (d) is true since from Claim 1, we have

$$I\left(\tilde{X}_1^n; \tilde{X}_2^n | \tilde{G}^n\right) = 0. \tag{56}$$

We use our definition in (55) for the rest of the proof.

$$\overset{(e)}{=} H\left(\tilde{X}_{1C}^n | \tilde{G}^n\right) + H\left(\tilde{X}_{1B}^n | \tilde{X}_{1C}^n, \tilde{G}^n\right)$$

$$+ H\left(\tilde{X}_{1A}^n \oplus \tilde{X}_{2A}^n | \tilde{X}_{1B}^n, \tilde{X}_{1C}^n, \tilde{X}_{2D}^n, \tilde{G}^n\right)$$

$$- H\left(\bar{X}_{1A}^n | \tilde{X}_{1C}^n, \tilde{G}^n\right) - H\left(\hat{X}_{1A}^n | \tilde{X}_{1B}^n, \tilde{X}_{1C}^n, \tilde{G}^n\right)$$

$$+ H\left(\tilde{X}_{2D}^n | \tilde{G}^n\right) + H\left(\tilde{X}_{2B}^n | \tilde{X}_{2D}^n, \tilde{G}^n\right)$$

$$+ H\left(\tilde{X}_{1A}^n \oplus \tilde{X}_{2A}^n | \tilde{X}_{2B}^n, \tilde{X}_{2D}^n, \tilde{X}_{1C}^n, \tilde{G}^n\right)$$

$$- H\left(\bar{X}_{2A}^n | \tilde{X}_{2D}^n, \tilde{G}^n\right) - H\left(\hat{X}_{2A}^n | \tilde{X}_{2B}^n, \tilde{X}_{2D}^n, \tilde{G}^n\right)$$

$$\overset{(f)}{\leq} H\left(\tilde{X}_{1C}^n | \tilde{G}^n\right) + H\left(\tilde{X}_{1A}^n \oplus \tilde{X}_{2A}^n | \tilde{X}_{1B}^n, \tilde{X}_{1C}^n, \tilde{X}_{2D}^n, \tilde{G}^n\right)$$

$$- H\left(\hat{X}_{1A}^n | \tilde{X}_{1B}^n, \tilde{X}_{1C}^n, \tilde{G}^n\right) + H\left(\tilde{X}_{2D}^n | \tilde{G}^n\right)$$

$$+ H\left(\tilde{X}_{1A}^n \oplus \tilde{X}_{2A}^n | \tilde{X}_{2B}^n, \tilde{X}_{2D}^n, \tilde{X}_{1C}^n, \tilde{G}^n\right)$$

$$- H\left(\hat{X}_{2A}^n | \tilde{X}_{2B}^n, \tilde{X}_{2D}^n, \tilde{G}^n\right)$$

$$\tag{57}$$

$$\leq H\left(\tilde{X}_{1C}^n | \tilde{G}^n\right) + H\left(\bar{X}_{1A}^n \oplus \bar{X}_{2A}^n | \tilde{X}_{1B}^n, \tilde{X}_{1C}^n, \tilde{X}_{2D}^n, \tilde{G}^n\right)$$

$$+ H\left(\hat{X}_{1A}^n, \hat{X}_{2A}^n | \bar{X}_{1A}^n \oplus \bar{X}_{2A}^n, \tilde{X}_{1B}^n, \tilde{X}_{1C}^n, \tilde{X}_{2D}^n, \tilde{G}^n\right)$$

$$- H\left(\hat{X}_{1A}^n | \tilde{X}_{1B}^n, \tilde{X}_{1C}^n, \tilde{G}^n\right) + H\left(\tilde{X}_{2D}^n | \tilde{G}^n\right)$$

$$+ H\left(\tilde{X}_{1A}^n \oplus \tilde{X}_{2A}^n | \tilde{X}_{2B}^n, \tilde{X}_{2D}^n, \tilde{X}_{1C}^n, \tilde{G}^n\right)$$

$$- H\left(\hat{X}_{2A}^n | \tilde{X}_{2B}^n, \tilde{X}_{2D}^n, \tilde{G}^n\right)$$

$$\overset{(g)}{\leq} H\left(\tilde{X}_{1C}^n | \tilde{G}^n\right) + H\left(\bar{X}_{1A}^n \oplus \bar{X}_{2A}^n | \tilde{X}_{1B}^n, \tilde{X}_{1C}^n, \tilde{X}_{2D}^n, \tilde{G}^n\right)$$

$$+ H\left(\hat{X}_{2A}^n | \hat{X}_{1A}^n, \bar{X}_{1A}^n \oplus \bar{X}_{2A}^n, \tilde{X}_{1B}^n, \tilde{X}_{1C}^n, \tilde{X}_{2D}^n, \tilde{G}^n\right)$$

$$+ H\left(\tilde{X}_{2D}^n | \tilde{G}^n\right) + H\left(\tilde{X}_{1A}^n \oplus \tilde{X}_{2A}^n | \tilde{X}_{2B}^n, \tilde{X}_{2D}^n, \tilde{X}_{1C}^n, \tilde{G}^n\right)$$

$$\overset{(h)}{\leq} n\left(2p_c + \epsilon_n\right), \tag{58}$$

where (e) follows from (55) and the construction of the signals, we have

$$\Pr\left[\hat{X}_{1A}^n, \bar{X}_{1A}^n, \tilde{X}_{1C}^n, \tilde{G}^n\right] = \Pr\left[\hat{X}_{1A}^n, \tilde{X}_{1B}^n, \tilde{X}_{1C}^n, \tilde{G}^n\right], \tag{59}$$

thus,

$$H\left(\hat{X}_{1A}^n | \bar{X}_{1A}^n, \tilde{X}_{1C}^n, \tilde{G}^n\right) = H\left(\hat{X}_{1A}^n | \tilde{X}_{1B}^n, \tilde{X}_{1C}^n, \tilde{G}^n\right), \tag{60}$$

and similar statement is true for $\bar{X}_{2A}^n$ and $\tilde{X}_{2B}^n$; (f) holds since from Lemma 1, we have

$$H\left(\tilde{X}_{1B}^n | \tilde{X}_{1C}^n, \tilde{G}^n\right) - H\left(\bar{X}_{1A}^n | \tilde{X}_{1C}^n, \tilde{G}^n\right) \leq 0; \tag{61}$$

(g) holds since

$$H\left(\hat{X}_{1A}^n | \bar{X}_{1A}^n \oplus \bar{X}_{2A}^n, \tilde{X}_{1B}^n, \tilde{X}_{1C}^n, \tilde{X}_{2D}^n, \tilde{G}^n\right)$$
$$- H\left(\hat{X}_{1A}^n | \tilde{X}_{1B}^n, \tilde{X}_{1C}^n, \tilde{G}^n\right) \leq 0; \tag{62}$$

(h) holds since

$$H\left(\tilde{X}_{1C}^n | \tilde{G}^n\right) \leq nq_d p_c,$$

$$H\left(\tilde{X}_{2D}^n | \tilde{G}^n\right) \leq nq_d p_c,$$

$$H\left(\tilde{X}_{1A}^n \oplus \tilde{X}_{2A}^n | \tilde{X}_{2B}^n, \tilde{X}_{2D}^n, \tilde{X}_{1C}^n, \tilde{G}^n\right) \leq np_d p_c,$$

$$H\left(\bar{X}_{1A}^n \oplus \bar{X}_{2A}^n | \tilde{X}_{1B}^n, \tilde{X}_{1C}^n, \tilde{X}_{2D}^n, \tilde{G}^n\right) \leq n(p_d - p_c),$$

$$H\left(\hat{X}_{2A}^n | \hat{X}_{1A}^n, \bar{X}_{1A}^n \oplus \bar{X}_{2A}^n, \tilde{X}_{1B}^n, \tilde{X}_{1C}^n, \tilde{X}_{2D}^n, \tilde{G}^n\right)$$
$$\leq n(p_c + p_d p_c - p_d). \tag{63}$$

• **Regime III:** $p_d \leq p_c \leq 1$: In this regime, again we have $\beta = 1$, and the capacity region would be equal to the intersection of capacity regions of the two MACs formed at the receivers. Under no CSIT assumption, we do not need to create a contracted channel.

The derivation of the outer-bound is easier compared to the other regimes. Basically, Rx_i after decoding and removing its corresponding signal, has a stronger channel from $\text{Tx}_{\bar{i}}$ compared to $\text{Rx}_{\bar{i}}$, and thus it must be able to decode both $\tilde{W}_i$ and $\tilde{W}_{\bar{i}}$, $i = 1, 2$. Thus, we have

$$R_1 + R_2 \leq p_d + p_c - p_d p_c. \tag{64}$$

V. ACHIEVABILITY

In this section, we provide the proof of Theorem 1. We show that for the weak and the strong interference regimes, the entire capacity region is achieved by applying point-to-point erasure codes with appropriate rates at each transmitter, using either treat-interference-as-erasure or interference-decoding at each receiver, based on the channel parameters. For the moderate interference regime, *i.e.* $\frac{p_d}{1+p_d} \le p_c \le p_d$, the inner-bound is based on a modification of the Han-Kobayashi scheme for the erasure channel, enhanced by time-sharing.

A. Weak and Strong Interference Regimes

Each transmitter applies a point-to-point erasure random code as described in [22]. On the other hand, at each receiver we have two options: (1) interference-decoding; and (2) treat-interference-as-erasure. When a receiver decodes the interference alongside its intended message, the achievable rate region is the capacity region of the multiple-access channel (MAC) formed at that receiver as depicted in Fig. 6. The MAC capacity at Rx_1, is given by

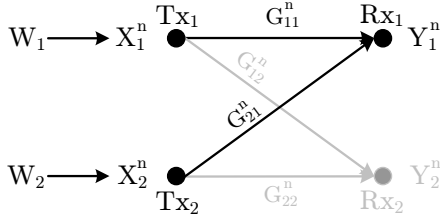$$\begin{cases} R_1 \le p_d, \\ R_2 \le p_c, \\ R_1 + R_2 \le 1 - q_d q_c. \end{cases} \qquad (65)$$

Fig. 6. *The multiple-access channel (MAC) formed at* Rx_1.

As a result, Rx_1 can decode its message by interference decoding, if R_1 and R_2 satisfy the constraints in (65). On the other hand, if Rx_1 treats interference as erasure, it basically ignores the received signal at time instants where $G_{21}[t] = 1$, and $p_d q_c$ fraction of the time, it receives the transmit signal of Tx_1. Thus, Rx_1 can decode its message by treat-interference-as-erasure, if $R_1 \le p_d q_c$.

Similarly, Rx_2 can decode its message by treat-interference-as-erasure, if $R_2 \le p_d q_c$. Also, Rx_2 can decode its message by interference decoding, if R_1 and R_2 are inside the capacity region of the multiple-access channel (MAC) formed at Rx_2, i.e.,

$$\begin{cases} R_1 \le p_c, \\ R_2 \le p_d, \\ R_1 + R_2 \le 1 - q_d q_c. \end{cases} \qquad (66)$$

Therefore, the achievable rate region by either treat-interference-as-erasure or interference-decoding at each receiver is the convex hull of (67) on top of the next page.

In the remaining of this section, we show that the convex hull of $\mathcal{R}$ matches the outer-bound of Theorem 2 for the weak and the strong interference regimes, *i.e.*

- For $0 \le p_c \le \frac{p_d}{1+p_d}$:

$$\bar{\mathcal{C}}(p_d, p_c) = \left\{ (R_1, R_2) \left| \begin{array}{l} 0 \le R_i \le p_d \quad i = 1, 2 \\ R_i + \beta R_{\bar{i}} \le \beta p_d + p_c - p_d p_c \end{array} \right. \right\} \qquad (68)$$

where

$$\beta = \frac{p_d - p_c}{p_d p_c}. \qquad (69)$$

- For $p_d \le p_c \le 1$:

$$\bar{\mathcal{C}}(p_d, p_c) = \left\{ (R_1, R_2) \left| \begin{array}{l} 0 \le R_i \le p_d \quad i = 1, 2 \\ R_i + R_{\bar{i}} \le p_d + p_c - p_d p_c \end{array} \right. \right\} \qquad (70)$$

First, it is easy to verify that the region described in (68), is the convex hull of the following corner points:

$$\begin{aligned} (R_1, R_2) &= (0, 0), \\ (R_1, R_2) &= (p_d, 0), \\ (R_1, R_2) &= (0, p_d), \\ (R_1, R_2) &= (p_d, q_d p_c), \\ (R_1, R_2) &= (q_d p_c, p_d), \\ (R_1, R_2) &= (p_d q_c, p_d q_c). \end{aligned} \qquad (71)$$

Now, when $0 \le p_c \le p_d / (1 + p_d)$, the rate region $\mathcal{R}$ (as defined in (67)) and its convex hull are shown in Fig. 7. As we can note from Fig. 7(b), in this case the convex hull of $\mathcal{R}$ is indeed the convex hull of the corner points in (71).

On the other hand, when $p_d \le p_c \le 1$, the rate region $\mathcal{R}$ (as defined in (67)) is depicted in Fig. 8, which is the convex hull of the first five corner points in (71). In this case, the last point in (71) is strictly inside the region in Fig. 8, hence again, $\mathcal{R}$ coincides with the convex hull of the corner points in (71), and this completes the proof of Theorem 1 under no CSIT assumption for the weak and strong interference regimes.

B. Moderate Interference Regime

In the remaining of this section, a Han-Kobayashi (HK) scheme is proposed and the corresponding sum-rate is investigated. The main focus would be the moderate interference regime, *i.e.* $\frac{p_d}{1+p_d} \le p_c \le p_d$. In this regime, neither treating interference as erasure nor completely decoding the interference would be optimal. The result echoes that the sum-capacity might be a similar "W" curve as that of the non-fading Gaussian interference channel [9].

To explore the power of HK scheme under this channel model, the codebooks of common and private messages are generated based on rate-splitting, which is originally developed for MAC channels [23]. In a MAC channel, each user can split its own message into two sub-messages[1], and the receiver

---

[1] In the view-point of a MAC, each user is split into so-called virtual users, which is equivalent to splitting the corresponding message instead.

$$\mathcal{R} = \left\{ (R_1, R_2) \middle| \underbrace{\left(\left\{ \begin{array}{l} R_1 \leq p_d \\ R_2 \leq p_c \\ R_1 + R_2 \leq 1 - q_d q_c \end{array} \right. \quad \text{or } R_1 \leq p_d q_c \right)}_{\text{decodability constraint at Rx}_1} \text{ and } \underbrace{\left(\left\{ \begin{array}{l} R_1 \leq p_c \\ R_2 \leq p_2 \\ R_1 + R_2 \leq 1 - q_d q_c \end{array} \right. \quad \text{or } R_2 \leq p_d q_c \right)}_{\text{decodability constraint at Rx}_2} \right\}. \quad (67)$$
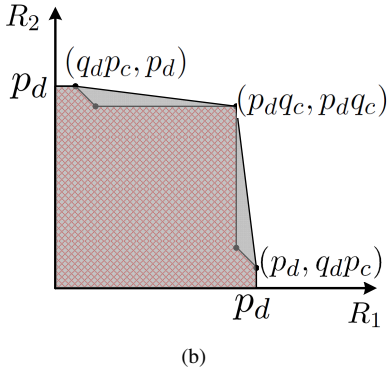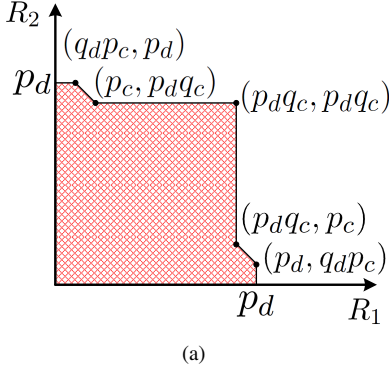

(a)

(b)

Fig. 7. (a) Depiction of the rate region $\mathcal{R}$ for $0 \leq p_c \leq p_d / (1 + p_d)$; and (b) its convex hull.
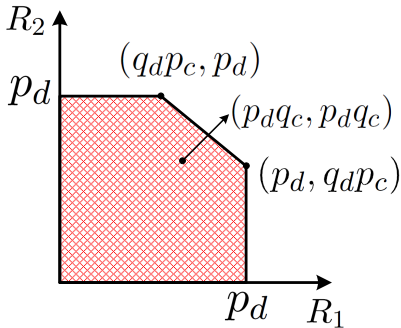

Fig. 8. Depiction of the rate region $\mathcal{R}$ for $p_d \leq p_c \leq 1$. In this case, corner point $(R_1, R_2) = (p_d q_c, p_d q_c)$ is strictly inside the convex hull of the first five points in (71).

decodes all messages and sub-messages through an onion-peeling process. That is the receiver decodes one message by treating others as erasure, and then removes the decoded contribution from the received signal. The receiver decodes the next message from the remaining signal until no message is

left. The key advantage of rate-splitting scheme for the MAC channel is that the entire capacity region can be achieved with single-user codebooks and no time-sharing is required.

We now present the modified HK scheme for the two-user BFIC with no CSIT. Let $\mathcal{B}(p)$ denote Bernoulli distribution with probability p taking value 1. Given $\delta \in [0, 1]$, any random variable X with distribution $\mathcal{B}\left(\frac{1}{2}\right)$ can be split as $X = \max(X_c, X_p)$, where $X_c \sim \mathcal{B}\left(\frac{\delta}{2}\right)$ and $X_p \sim \mathcal{B}\left(\frac{1-\delta}{2-\delta}\right)$. From the view point of the random coding, instead of using X to generate a single codebook, one can generate two codebooks according to the distributions of X_c and X_p respectively, and then the channel input is generated by the $\max$ operator. It is easy to see that $H(X | X_c) = \frac{2-\delta}{2} H\left(\frac{1}{2-\delta}\right) =: C_\delta$. Therefore, we have

$$I(X_c; X) = H(X) - H(X | X_c) = 1 - C_\delta,$$
$$I(X_p; X | X_c) = H(X | X_c) = C_\delta. \quad (72)$$

Intuitively, C_δ can be viewed as the portion of the message carried by codebook generated via X_p. Note that C_δ is continuous and monotonically decreasing with respect to δ. Thus as δ goes from 0 to 1, the coding scheme continuously changes from X_p only scheme ($\delta = 0$) to X_c only scheme ($\delta = 1$).

Although time-sharing does not play a key role in achieving the capacity region of the MAC, it does enlarge achievable rate region of the HK scheme [24]. For this particular channel, we consider a Han-Kobayashi scheme with time-sharing as follows. For any $\delta_1 \in [0, 1]$, we generate one codebook according to distribution $\mathcal{B}\left(\frac{\delta_1}{2}\right)$ for the common message and one random codebook according to distribution $\mathcal{B}\left(\frac{1-\delta_1}{2-\delta_1}\right)$ for the private message. We denote these two codebooks by $\mathcal{C}_c(\delta_1)$ and $\mathcal{C}_p(\delta_1)$ respectively. Similarly for any $\delta_2 \in [0, 1]$, we can generate codebooks $\mathcal{C}_c(\delta_2)$ and $\mathcal{C}_p(\delta_2)$. In addition, let $\{Q[t]\}$ be an i.i.d random sequence with $P(Q[t] = 1) = P(Q[t] = 2) = 1/2$, which generates a particular time-sharing sequence. Before any communication begins, the time-sharing sequence is revealed to the transmitters and the receivers. Then the two transmitters communicate their messages as follows. If $Q[t] = 1$, user 1 encodes its common and private messages according to codebooks $\mathcal{C}_c(\delta_1)$ and $\mathcal{C}_p(\delta_1)$ respectively, and uses the $\max$ operator to generate the transmit signal. Meanwhile user 2 does the same thing except that it uses codebooks $\mathcal{C}_c(\delta_2)$ and $\mathcal{C}_p(\delta_2)$. If $Q[t] = 2$, the two users switch their codebooks.

Equivalently, we can state the coding scheme in another way: Given i.i.d. time-sharing random sequence $\{Q[t]\}$, the two users encode their common and private messages independently according to i.i.d. sequences $\{X_{1c}[t], X_{1p}[t]\}$ and $\{X_{2c}[t], X_{2p}[t]\}$. Given $Q[t]$, the distributions of the other

sequences are defined in Table III. For fixed (δ_1, δ_2), we can

$Q[t]$	$X_{1c}[t]$	$X_{1p}[t]$	$X_{2c}[t]$	$X_{2p}[t]$
1	$\mathcal{B}\left(\frac{\delta_1}{2}\right)$	$\mathcal{B}\left(1-\frac{1}{2-\delta_1}\right)$	$\mathcal{B}\left(\frac{\delta_2}{2}\right)$	$\mathcal{B}\left(\frac{1-\delta_2}{2-\delta_2}\right)$
2	$\mathcal{B}\left(\frac{\delta_2}{2}\right)$	$\mathcal{B}\left(1-\frac{1}{2-\delta_2}\right)$	$\mathcal{B}\left(\frac{\delta_1}{2}\right)$	$\mathcal{B}\left(\frac{1-\delta_1}{2-\delta_1}\right)$

TABLE III
SUMMARY OF THE HK SCHEME WITH TIME-SHARING

compute the corresponding achievable sum-rate, *i.e.*

$$R_{sum}(\delta_1, \delta_2) := R_1(\delta_1, \delta_2) + R_2(\delta_1, \delta_2), \quad (73)$$

and then maximize over all possible values of (δ_1, δ_2). To determine $R_{sum}(\delta_1, \delta_2)$, it is sufficient to make sure that both common messages are decodable at both receivers and each private message is decodable at its corresponding receiver. More specifically, we follow the similar procedure as [9, Section III] except that we explore time-sharing and optimize the rate splitting. That is the rates of the common messages are determined by the two virtual compound-MAC channels at the two receivers, and each private message is decoded only at its corresponding receiver. We have following theorem:

Theorem 3. *The following sum-rate is optimal over δ_1, $\delta_2 \in [0, 1]$:*

$$R_{sum} = \begin{cases} 2p_d(1-p_c) & \text{if } p_c \leq \frac{p_d}{1+p_d} \\ p_d + p_c - p_d p_c + \frac{p_d - p_c}{2} C_\delta^* & \text{if } \frac{p_d}{1+p_d} < p_c \leq p_d \end{cases} \quad (74)$$

where

$$C_\delta^* := \frac{p_d p_c - (p_d - p_c)}{p_d p_c - \frac{p_d - p_c}{2}}. \quad (75)$$

Remark 2. *Treating interference as erasure can achieve a sum-rate as large as $2p_d(1-p_c)$, while the sum-rate of interference-decoding scheme is $p_d + p_c - p_d p_c$. In the weak interference regime where $p_c \leq \frac{p_d}{1+p_d}$, the proposed HK scheme degrades to treating-interference-as-erasure and achieves the sum-capacity. In the moderate interference regime where $\frac{p_d}{1+p_d} < p_c \leq p_d$, partially decoding interference can outperform the other two schemes.*

Proof of Theorem 3:

If $p_c \leq \frac{p_d}{1+p_d}$, the sum-rate in Theorem 3 is the sum-capacity and can be achieved by treating interference as erasure, which corresponds to the HK scheme with $\delta_1 = \delta_2 = 0$. Thus, we assume that $\frac{p_d}{1+p_d} < p_c \leq p_d$.

Let $G[t] = (G_{11}[t], G_{21}[t])$ and let $R_{ic}(\delta_1, \delta_2)$ and $R_{ip}(\delta_1, \delta_2)$ denote achievable rates of the common and the private messages respectively. Since all codebooks are generated according to i.i.d. sequences and the channel is memoryless, we will drop all time indices in what follows.

Since at each receiver we decode private message last, we have

$$R_{1p}(\delta_1, \delta_2) = I(X_{1p}; Y_1 | X_{1c}, X_{2c}, Q, G).$$

For the common messages, the achievable rates fall into the intersection of two virtual MACs at the two receivers. Taking symmetric properties of the channel and the coding

scheme into account, we conclude that any positive rate pair for the common messages satisfying the following conditions is achievable:

$$R_{1c}(\delta_1, \delta_2) + R_{2c}(\delta_1, \delta_2) \leq I(X_{2c}, X_{1c}; Y_1 | Q, G), \quad (76)$$
$$R_{2c}(\delta_1, \delta_2) \leq I(X_{2c}; Y_1 | X_{1c}, Q, G), \quad (77)$$
$$R_{1c}(\delta_1, \delta_2) \leq I(X_{1c}; Y_1 | X_{2c}, Q, G). \quad (78)$$

Since $p_d \geq p_c$, it is easy to see that (77) implies (78) as far as sum-rate is concerned. Therefore, we can achieve the following sum-rate:

$$R_{sum}(\delta_1, \delta_2) \leq 2M_p(\delta_1, \delta_2) + \min(M_{c1}(\delta_1, \delta_2), 2M_{c2}(\delta_1, \delta_2)), \quad (79)$$

where

$$M_p(\delta_1, \delta_2) = I(X_{1p}; Y_1 | X_{1c}, X_{2c}, Q, G),$$
$$M_{c1}(\delta_1, \delta_2) = I(X_{2c}, X_{1c}; Y_1 | Q, G),$$
$$M_{c2}(\delta_1, \delta_2) = I(X_{2c}; Y_1 | X_{c1}, Q, G). \quad (80)$$

The derivation of $M_p(\delta_1, \delta_2)$, $M_{c1}(\delta_1, \delta_2)$, and $M_{c2}(\delta_1, \delta_2)$ relies on the chain rule and the basic equalities (72) as described below.

Starting from $M_{c1}(\delta_1, \delta_2)$, we have

$$M_{c1}(\delta_1, \delta_2) = I(X_{2c}, X_{1c}; Y_1 | Q, G)$$
$$= p_d + p_c - p_c p_d - H(Y_1 | X_{2c}, X_{1c}, Q, G).$$
$$= p_d + p_c - p_c p_d - H(G_{11}X_1 \oplus G_{21}X_2 | X_{1c}, X_{2c}, Q, G)$$
$$= p_d + p_c - p_c p_d$$
$$- \Big(p_d(1-p_c)H(X_1 | X_{1c}, Q) + p_c(1-p_d)H(X_2 | X_{2c}, Q)$$
$$+ p_d p_c H(X_1 \oplus X_2 | X_{2c}, X_{1c}, Q)\Big)$$
$$= p_d + p_c - p_c p_d - (p_d + p_c - 2p_c p_d)\frac{C_{\delta_1} + C_{\delta_2}}{2}$$
$$- p_d p_c \gamma(\delta_1, \delta_2) \quad (81)$$

where

$$\gamma(\delta_1, \delta_2) := H(X_1 \oplus X_2 | X_{2c}, X_{1c}, Q)$$
$$= \frac{\delta_1}{2}C_{\delta_2} + \frac{\delta_2}{2}C_{\delta_1} + \frac{(2-\delta_1)(2-\delta_2)}{4}H(p^*), \quad (82)$$

and $p^* = \frac{2-\delta_2-\delta_1}{(2-\delta_1)(2-\delta_2)}$. In addition, $\gamma(\delta_1, \delta_2)$ is upper bounded by $C_{\delta_1} + C_{\delta_2}$:

$$\gamma(\delta_1, \delta_2) \leq H(X_1 | X_{1c}, Q) + H(X_2 | X_{2c}, Q)$$
$$= C_{\delta_1} + C_{\delta_2}. \quad (83)$$

Next for $M_{c2}(\delta_1, \delta_2)$ using the chain rule and (81), we get

$$M_{c2}(\delta_1, \delta_2) = I(X_{1c}, X_{2c}; Y_1 | Q, G) - I(X_{1c}; Y_1 | Q, G)$$
$$= M_{c1}(\delta_1, \delta_2) - I(X_{1c}; G_{11}X_1 \oplus G_{21}X_2 | Q, G)$$
$$= M_{c1}(\delta_1, \delta_2) - p_d(1-p_c)\left[1 - \frac{C_{\delta_1} + C_{\delta_2}}{2}\right]$$
$$= p_c - (p_c - p_c p_d)\frac{C_{\delta_1} + C_{\delta_2}}{2} + p_d p_c \gamma(\delta_1, \delta_2). \quad (84)$$

Similarly for $M_p(\delta_1, \delta_2)$, we have

$$M_p(\delta_1, \delta_2)$$
$$= I\left(X_{1p}, X_{1c}, X_{2c}; Y_1 \mid Q, G\right) - I\left(X_{1c}, X_{2c}; Y_1 \mid Q, G\right)$$
$$= I\left(X_{1p}, X_{1c}; Y_1 \mid Q, G\right) + I\left(X_{2c}; Y_1 \mid X_{1p}, X_{1c}, Q, G\right)$$
$$- M_{c1}(\delta_1, \delta_2)$$
$$= p_d(1 - p_c) + p_c\left[1 - \frac{C_{\delta_1} + C_{\delta_2}}{2}\right] - M_{c1}(\delta_1, \delta_2)$$
$$= (p_d - 2p_c p_d)\frac{C_{\delta_1} + C_{\delta_2}}{2} + p_d p_c \gamma(\delta_1, \delta_2). \tag{85}$$

Therefore, substituting (81), (84), and (85) into (79), we can get an expression for $R_{sum}(\delta_1, \delta_2)$. If we let $\delta_2 = 1$, then $C_{\delta_2} = 0$ and $\gamma(\delta_1, \delta_2) = C_{\delta_1}$. Furthermore,

$$R_{sum} \geq \max_{\delta_1 \in [0,1]} R_{sum}(\delta_1, 1)$$
$$= \max_{C_{\delta_1} \in [0,1]} p_d C_{\delta_1} + \min\left(p_d + p_c - p_d p_c - \frac{p_d + p_c}{2}C_{\delta_1}\,,\right.$$
$$\left., 2p_c - (p_c + p_c p_d)C_{\delta_1}\right)$$
$$= \max_{C_{\delta_1} \in [0,1]} \min\left(p_d + p_c - p_d p_c + \frac{p_d - p_c}{2}C_{\delta_1}\,,\right.$$
$$\left. 2p_c - (p_c - p_d + p_c p_d)C_{\delta_1}\right), \tag{86}$$

which achieves the maximum value of $p_d + p_c - p_d p_c + \frac{p_d - p_c}{2}C_\delta^*$ at $C_{\delta_1} = C_\delta^*$. So it is sufficient to show that the converse is true.

If $M_{c1}(\delta_1, \delta_2) \geq 2M_{c2}(\delta_1, \delta_2)$, we have

$$R_{sum}(\delta_1, \delta_2) = 2M_p(\delta_1, \delta_2) + 2M_{c2}(\delta_1, \delta_2)$$
$$= 2p_c - (p_c - p_d + p_d p_c)(C_{\delta_1} + C_{\delta_2}). \tag{87}$$

If $M_{c1}(\delta_1, \delta_2) < 2M_{c2}(\delta_1, \delta_2)$, we have

$$R_{sum}(\delta_1, \delta_2) = 2M_p(\delta_1, \delta_2) + M_{c1}(\delta_1, \delta_2)$$
$$= p_d + p_c - p_d p_c + (p_d - p_c - 2p_d p_c)\frac{C_{\delta_1} + C_{\delta_2}}{2}$$
$$+ p_d p_c \gamma(\delta_1, \delta_2)$$
$$\leq p_d + p_c - p_d p_c + \frac{p_d - p_c}{2}(C_{\delta_1} + C_{\delta_2}), \tag{88}$$

where (88) is due to (83).

Now, let $C_{\delta_3} = C_{\delta_1} + C_{\delta_2}$. We have

$$R_{sum} \leq \max_{C_{\delta_3} \in [0,2]} \min\left(p_d + p_c - p_d p_c + \frac{p_d - p_c}{2}C_{\delta_3}\,,\right.$$
$$\left. 2p_c - (p_c - p_d + p_d p_c)C_{\delta_3}\right). \tag{89}$$

Comparing (89) and (86), they are in the same form except that (89) is over a larger domain. However, (89) achieves its maximum when $C_{\delta_3} = C_\delta^*$, and $C_\delta^* \leq 1$ with the assumption of $\frac{p_d}{1+p_d} < p_c$. Therefore, Theorem 3 holds.

HK Scheme with Time-Sharing:

Let Q be a random variable taking values in $\{1, 2\}$ with $\mathsf{P}(Q = i) = \lambda_i$ for $i = 1, 2$. Coding scheme is generated based on

Q	X_{1c}	X_{1p}	X_{2c}	X_{2p}
1	$\mathcal{B}\left(\frac{\delta_1}{2}\right)$	$\mathcal{B}\left(\frac{1-\delta_1}{2-\delta_1}\right)$	$\mathcal{B}\left(\frac{1}{2}\right)$	$\mathcal{B}(0)$
2	$\mathcal{B}\left(\frac{1}{2}\right)$	$\mathcal{B}(0)$	$\mathcal{B}\left(\frac{\delta_2}{2}\right)$	$\mathcal{B}\left(\frac{1-\delta_2}{2-\delta_2}\right)$

The evaluation of achievable rate is similar to the previous case but it breaks the symmetric property in general. Define following short-hand notation:

$$C_{\delta_i} := \frac{2 - \delta_i}{2} H\left(\frac{1}{2 - \delta_i}\right) \qquad i = 1, 2 \tag{90}$$

Any rate pair satisfying the following conditions is achievable for the common messages (see Appendix A for the details).

$$R_{1c} \leq p_c - \lambda_1 C_{\delta_1} p_c - \lambda_2 C_{\delta_2} p_c p_d, \tag{91a}$$
$$R_{2c} \leq p_c - \lambda_2 C_{\delta_2} p_c - \lambda_1 C_{\delta_1} p_d p_c, \tag{91b}$$
$$R_{1c} + R_{2c} \leq p_d + p_c - p_d p_c - \tag{91c}$$
$$\max(\lambda_1 C_{\delta_1} p_d + \lambda_2 C_{\delta_2} p_c, \lambda_1 C_{\delta_1} p_c + \lambda_2 C_{\delta_2} p_d)$$

Rates of the private messages are given by

$$R_{1p} \leq I\left(X_{1p}; Y_1 \mid X_{1c}, X_{2c}, Q, Q, G\right) = \lambda_1 p_d C_{\delta_1}, \tag{92}$$
$$R_{2p} \leq I\left(X_{2p}; Y_2 \mid X_{1c}, X_{2c}, Q, Q, G\right) = \lambda_2 p_d C_{\delta_2}. \tag{93}$$

Finally, we evaluate sum-rate with $\lambda_i = 1/2$ and $C_{\delta_1} = C_{\delta_2} = C_\delta$:

$$R_{sum}(C_\delta) = p_d C_\delta + \min(p_d + p_c - p_d p_c - \frac{p_d + p_c}{2}C_\delta$$
$$, 2p_c - (p_c + p_c p_d)C_\delta) \tag{94}$$

Some numeric results:

For $p_d = 1$ and for $p_c \in [1/2, 1]$, we plotted the HK rate (optimized over δ) in Fig. V-B.
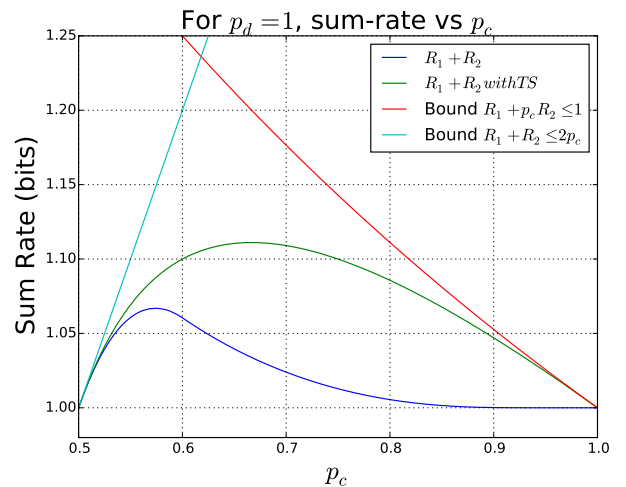

Fig. 9. The optimized Han-Kobayashi scheme for $p_d = 1$ and for $p_c \in [1/2, 1]$.

VI. Conclusion

We studied the capacity region of the two-user Binary Fading Interference Channel with no CSIT. We showed that under the weak and the moderate interference regimes, the entire capacity region is achieved by applying point-to-point erasure codes with appropriate rates at each transmitter, using either treat-interference-as-erasure or interference-decoding at each receiver, based on the channel parameters. For the moderate interference regime, we devised a modified Han-Kobayashi scheme suited for discrete memoryless channels enhanced by time-sharing. Our outer-bounds rely on two key lemmas, namely the Correlation Lemma and the Entropy Leakage Lemma.

Acknowledgement

Appendix A
Proof of (91)

For the virtual MAC at receiver 1:

$$I\left(X_{1c}, X_{2c}; Y_1 \middle| Q, G\right)$$
$$= H\left(Y_1 \middle| Q, G\right) - H\left(Y_1 \middle| X_{1c}, X_{2c}, Q, G\right)$$
$$= p_d + p_c - p_d p_c - \lambda_1 H\left(Y_1 \middle| X_{1c}, X_{2c}, Q=1, G\right)$$
$$+ \lambda_2 H\left(Y_1 \middle| X_{1c}, X_{2c}, Q=2, G\right)$$
$$= p_d + p_c - p_d p_c - \lambda_1 C_{\delta_1} p_d - \lambda_2 C_{\delta_2} p_c \tag{95}$$

$$I\left(X_{2c}; Y_1 \middle| X_{1c}, Q, G\right)$$
$$= I\left(X_{1c}, X_{2c}; Y_1 \middle| Q, G\right) - I\left(X_{1c}; Y_1 \middle| Q, G\right)$$
$$= I\left(X_{1c}, X_{2c}; Y_1 \middle| Q, G\right) - \lambda_1 I\left(X_{1c}; Y_1 \middle| Q=1, G\right)$$
$$+ \lambda_2 I\left(X_{1c}; Y_1 \middle| Q=2, G\right)$$
$$= I\left(X_{1c}, X_{2c}; Y_1 \middle| Q, G\right) - \lambda_1 p_d(1 - p_c)(1 - C_{\delta_1})$$
$$- \lambda_2 p_d(1 - p_c)$$
$$= p_d + p_c - p_d p_c - \lambda_1 C_{\delta_1} p_d - \lambda_2 C_{\delta_2} p_c$$
$$- \lambda_1(1 - C_{\delta_1}) p_d(1 - p_c) - \lambda_2 p_d(1 - p_c)$$
$$= p_c - \lambda_1 C_{\delta_1} p_d - \lambda_2 C_{\delta_2} p_c + \lambda_1 C_{\delta_1} p_d(1 - p_c)$$
$$= p_c - \lambda_2 C_{\delta_2} p_c - \lambda_1 C_{\delta_1} p_d p_c \tag{96}$$

$$I\left(X_{1c}; Y_1 \middle| X_{2c}, Q, G\right)$$
$$= I\left(X_{1c}, X_{2c}; Y_1 \middle| Q, G\right) - I\left(X_{2c}; Y_1 \middle| Q, G\right)$$
$$= I\left(X_{1c}, X_{2c}; Y_1 \middle| Q, G\right) - \lambda_1 I\left(X_{2c}; Y_1 \middle| Q=1, G\right)$$
$$- \lambda_2 I\left(X_{2c}; Y_1 \middle| Q=2, G\right)$$
$$= I\left(X_{1c}, X_{2c}; Y_1 \middle| Q, G\right) - \lambda_1 p_c(1 - p_d)$$
$$- \lambda_2 p_c(1 - p_d)(1 - C_{\delta_2})$$
$$= p_d + p_c - p_d p_c - \lambda_1 C_{\delta_1} p_d - \lambda_2 C_{\delta_2} p_c$$
$$- \lambda_1 p_c(1 - p_d) - \lambda_2 p_c(1 - p_d)(1 - C_{\delta_2})$$
$$= p_d - \lambda_1 C_{\delta_1} p_d - \lambda_2 C_{\delta_2} p_c + \lambda_2 C_{\delta_2} p_c(1 - p_d)$$
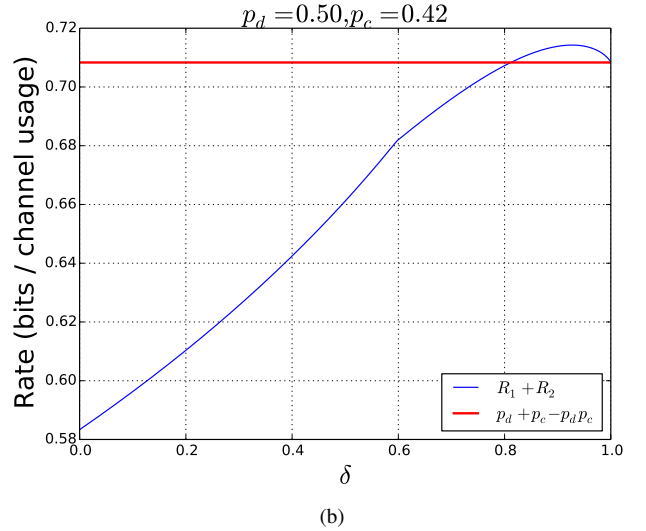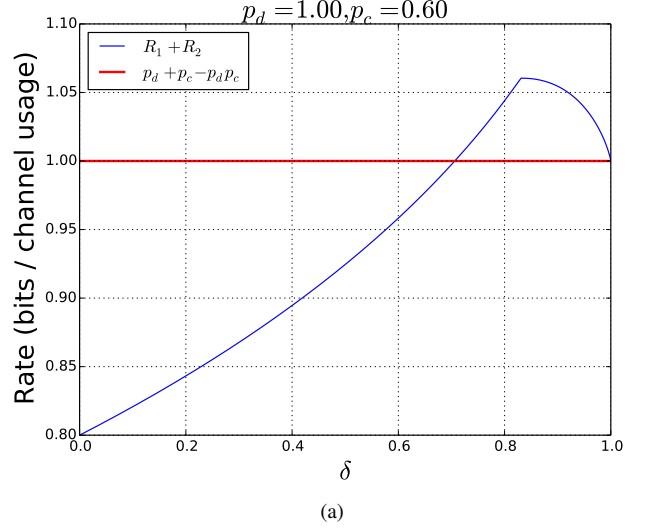$$= p_d - \lambda_1 C_{\delta_1} p_d - \lambda_2 C_{\delta_2} p_c p_d \tag{97}$$

$p_d = 1.00, p_c = 0.60$

(a)

$p_d = 0.50, p_c = 0.42$

(b)

Fig. 10. Numerical Results for the achievable HK Rate

For the virtual MAC at receiver 2:

$$I\left(X_{1c}, X_{2c}; Y_2 \middle| Q, G\right)$$
$$= H\left(Y_2 \middle| Q, G\right) - H\left(Y_2 \middle| X_{1c}, X_{2c}, Q, G\right)$$
$$= p_d + p_c - p_d p_c - \lambda_1 H\left(Y_2 \middle| X_{1c}, X_{2c}, Q=1, G\right)$$
$$- \lambda_2 H\left(Y_2 \middle| X_{1c}, X_{2c}, Q=2, G\right)$$
$$= p_d + p_c - p_d p_c - \lambda_1 C_{\delta_1} p_c - \lambda_2 C_{\delta_2} p_d \tag{98}$$

$$I\left(X_{2c}; Y_2 \middle| X_{1c}, Q, G\right)$$
$$= I\left(X_{1c}, X_{2c}; Y_2 \middle| Q, G\right) - I\left(X_{1c}; Y_2 \middle| Q, G\right)$$
$$= p_d + p_c - p_d p_c - \lambda_1 C_{\delta_1} p_c - \lambda_2 C_{\delta_2} p_d$$
$$- \lambda_1 p_c(1 - p_d)(1 - C_{\delta_1}) - \lambda_2 p_c(1 - p_d)$$
$$= p_d - \lambda_1 C_{\delta_1} p_c - \lambda_2 C_{\delta_2} p_d + \lambda_1 p_c(1 - p_d) C_{\delta_1}$$
$$= p_d - \lambda_2 C_{\delta_2} p_d - \lambda_1 C_{\delta_1} p_c p_d \tag{99}$$

$$I\left(X_{1c};Y_2\middle|X_{2c},Q,G\right)$$
$$= I\left(X_{1c},X_{2c};Y_2\middle|Q,G\right) - I\left(X_{2c};Y_2\middle|Q,G\right)$$
$$= p_d + p_c - p_d p_c - \lambda_1 C_{\delta_1} p_c - \lambda_2 C_{\delta_2} p_d - \lambda_1 p_d(1-p_c)$$
$$\quad - \lambda_2 p_d(1-p_c)(1-C_{\delta_2})$$
$$= p_c - \lambda_1 C_{\delta_1} p_c - \lambda_2 C_{\delta_2} p_d + \lambda_2 p_d(1-p_c)C_{\delta_2}$$
$$= p_c - \lambda_1 C_{\delta_1} p_c - \lambda_2 C_{\delta_2} p_c p_d \tag{100}$$

We reduce the number of constraints by performing the following comparisons.

• Comparing the two dominant faces:

$$I\left(X_{1c},X_{2c};Y_1\middle|Q,G\right) - I\left(X_{1c},X_{2c};Y_2\middle|Q,G\right)$$
$$= p_d + p_c - p_d p_c - \lambda_1 C_{\delta_1} p_d - \lambda_2 C_{\delta_2} p_c$$
$$\quad - \left(p_d + p_c - p_d p_c - \lambda_1 C_{\delta_1} p_c - \lambda_2 C_{\delta_2} p_d\right)$$
$$= (p_c - p_d)(\lambda_1 C_{\delta_1} - \lambda_2 C_{\delta_2}). \tag{101}$$

Thus the sign is determined by $(\lambda_1 C_{\delta_1} - \lambda_2 C_{\delta_2})$.

• Comparing single-user bounds for user 1's common message:

$$I\left(X_{1c};Y_1\middle|X_{2c},Q,G\right) - I\left(X_{1c};Y_2\middle|X_{2c},Q,G\right)$$
$$= p_d - \lambda_1 C_{\delta_1} p_d - \lambda_2 C_{\delta_2} p_c p_d$$
$$\quad - \left(p_c - \lambda_1 C_{\delta_1} p_c - \lambda_2 C_{\delta_2} p_c p_d\right)$$
$$= (p_d - p_c)(1 - \lambda_1 C_{\delta_1}) \geq 0. \tag{102}$$

So only $I\left(X_{1c};Y_2\middle|X_{2c},Q,G\right)$ matters.

• Comparing single-user bounds for user 2's common message:

$$I\left(X_{2c};Y_1\middle|X_{1c},Q,G\right) - I\left(X_{2c};Y_2\middle|X_{1c},Q,G\right)$$
$$= p_c - \lambda_2 C_{\delta_2} p_c - \lambda_1 C_{\delta_1} p_d p_c$$
$$\quad - \left(p_d - \lambda_2 C_{\delta_2} p_d - \lambda_1 C_{\delta_1} p_c p_d\right)$$
$$= (p_c - p_d)(1 - \lambda_2 C_{\delta_2}) \leq 0. \tag{103}$$

So only $I\left(X_{2c};Y_1\middle|X_{1c},Q,G\right)$ matters.

In summary, the following rates of common messages are achievable:

$$R_{1c} \leq p_c - \lambda_1 C_{\delta_1} p_c - \lambda_2 C_{\delta_2} p_c p_d \tag{104a}$$
$$R_{2c} \leq p_c - \lambda_2 C_{\delta_2} p_c - \lambda_1 C_{\delta_1} p_d p_c \tag{104b}$$
$$R_{1c} + R_{2c} \leq p_d + p_c - p_d p_c - \tag{104c}$$
$$\max(\lambda_1 C_{\delta_1} p_d + \lambda_2 C_{\delta_2} p_c, \lambda_1 C_{\delta_1} p_c + \lambda_2 C_{\delta_2} p_d)$$

First, we evaluate cases where $p_d - p_c \leq p_d p_c$. In particular, Fig. 10 shows the achievable rate vs splitting parameter δ for $(p_d, p_c) = (1, 0.6)$ and $(1/2, 5/12)$.

REFERENCES

[1] A. Vahid, M. A. Maddah-Ali, and A. S. Avestimehr, "Binary Fading Interference Channel with No CSIT," in *IEEE International Symposium on Information Theory (ISIT)*, pp. 666–670, 2014.

[2] R. Ahlswede, "The capacity region of a channel with two senders and two receivers," *The Annals of Probability*, pp. 805–814, 1974.

[3] T. Han and K. Kobayashi, "A new achievable rate region for the interference channel," *IEEE Transactions on Information Theory*, vol. 27, no. 1, pp. 49–60, 1981.

[4] H. Sato, "Two-user communication channels," *IEEE Transactions on Information Theory*, vol. 23, no. 3, pp. 295–304, 1977.

[5] A. El-Gamal and M. H. Costa, "The capacity region of a class of deterministic interference channels," *IEEE Transactions on Information Theory*, vol. 28, pp. 343–346, Mar. 1982.

[6] E. C. van der Meulen, "Some reflections on the interference channel," *Communications and Cryptography*, pp. 409–421, 1994.

[7] A. B. Carleial, "A case where interference does not reduce capacity (corresp.)," *IEEE Transactions on Information Theory*, vol. 21, no. 5, pp. 569–570, 1975.

[8] A. B. Carleial, "Outer bounds on the capacity of interference channels (corresp.)," *IEEE Transactions on Information Theory*, vol. 29, no. 4, pp. 602–606, 1983.

[9] R. Etkin, D. Tse, and H. Wang, "Gaussian interference channel capacity to within one bit," *IEEE Transactions on Information Theory*, vol. 54, pp. 5534–5562, Dec. 2008.

[10] C. Suh and D. Tse, "Feedback capacity of the gaussian interference channel to within 2 bits," *IEEE Transactions on Information Theory*, vol. 57, no. 5, pp. 2667–2685, 2011.

[11] A. Vahid, C. Suh, and A. S. Avestimehr, "Interference channels with rate-limited feedback," *IEEE Transactions on Information Theory*, vol. 58, no. 5, pp. 2788–2812, 2012.

[12] A. Vahid, M. A. Maddah-Ali, and A. S. Avestimehr, "Interference channel with binary fading: Effect of delayed network state information," in *49th Annual Allerton Conference on Communication, Control, and Computing*, pp. 894–901, 2011.

[13] A. Vahid, M. A. Maddah-Ali, and A. S. Avestimehr, "Communication through collisions: Opportunistic utilization of past receptions," in *Proceedings IEEE INFOCOM*, pp. 2553–2561, 2014.

[14] A. Vahid and R. Calderbank, "When does spatial correlation add value to delayed channel state information?," in *IEEE International Symposium on Information Theory (ISIT)*, pp. 2624–2628, 2016.

[15] A. Vahid and R. Calderbank, "Impact of local delayed csit on the capacity region of the two-user interference channel," in *IEEE International Symposium on Information Theory (ISIT)*, pp. 2421–2425, 2015.

[16] A. Vahid and R. Calderbank, "Two-user erasure interference channels with local delayed CSIT," *IEEE Transactions on Information Theory*, vol. 62, no. 9, pp. 4910–4923, 2016.

[17] V. Aggarwal, L. Sankar, A. R. Calderbank, and H. V. Poor, "Ergodic layered erasure one-sided interference channels," in *IEEE Information Theory Workshop*, pp. 574–578, IEEE, 2009.

[18] Y. Zhu and D. Guo, "Ergodic fading Z-interference channels without state information at transmitters," *IEEE Transactions on Information Theory*, vol. 57, no. 5, pp. 2627–2647, 2011.

[19] A. S. Avestimehr, S. Diggavi, and D. Tse, "Wireless network information flow: A deterministic approach," *IEEE Transactions on Information Theory*, vol. 57, Apr. 2011.

[20] D. N. Tse and R. D. Yates, "Fading broadcast channels with state information at the receivers," *IEEE Transactions on Information Theory*, vol. 58, no. 6, pp. 3453–3471, 2012.

[21] A. Vahid, M. A. Maddah-Ali, and A. S. Avestimehr, "Capacity results for binary fading interference channels with delayed CSIT," *IEEE Transactions on Information Theory*, vol. 60, no. 10, pp. 6093–6130, 2014.

[22] P. Elias, "The noisy channel coding theorem for erasure channels," *American Mathematical Monthly*, pp. 853–862, 1974.

[23] A. J. Grant, B. Rimoldi, R. L. Urbanke, and P. A. Whiting, "Rate-splitting multiple access for discrete memoryless channels," *IEEE Transactions on Information Theory*, vol. 47, no. 3, pp. 873–890, 2001.

[24] H. Te Sun and K. Kobayashi, "A new achievable rate region for the interference channel," *IEEE Transactions on Information Theory*, vol. 27, no. 1, pp. 49–60, 1981.

Alireza Vahid received his Ph.D. and M.Sc. degrees in Electrical and Computer Engineering both from Cornell University, Ithaca, NY, in 2015 and 2012 respectively. He obtained his B.Sc. degree in Electrical Engineering from Sharif University of Technology, Tehran, Iran, in 2009. He is currently a postdoctoral scholar at Information Initiative at Duke University, Durham, NC. His research interests include network information theory, wireless communications, coding theory, and data storage.

Dr. Vahid received the 2015 Outstanding PhD Thesis Research Award at Cornell University. He also received the Director's Ph.D. Teaching Assistant Award in 2010, Jacobs Scholar Fellowship in 2009, and Qualcomm Innovation Fellowship in 2013.

Yan Zhu received his B.E. and M.S. degrees from Tsinghua University, Beijing, China, in 2002 and 2005, respectively, and his Ph.D. degree from Northwestern University, Evanston, IL, in 2010, all in electrical engineering. From 2010 to 2015, He worked at Broadcom Inc. as a scientist, design staff. He joined Calterah Inc. as a chief system architecturer since 2016. His research interests include wireless communications, information theory, communication network and signal processing. He is a co-recipient of the 2010 IEEE Marconi Prize Paper Award in Wireless Communications (with D. Guo and M. L. Honig).

Mohammad Ali Maddah-Ali (S'03-M'08) received the B.Sc. degree from Isfahan University of Technology, and the M.A.Sc. degree from the University of Tehran, both in electrical engineering. From 2002 to 2007, he was with the Coding and Signal Transmission Laboratory (CST Lab), Department of Electrical and Computer Engineering, University of Waterloo, Canada, working toward the Ph.D. degree. From 2007 to 2008, he worked at the Wireless Technology Laboratories, Nortel Networks, Ottawa, ON, Canada. From 2008 to 2010, he was a post-doctoral fellow in the Department of Electrical Engineering and Computer Sciences at the University of California at Berkeley. Then, he joined Bell Labs, Holmdel, NJ, as a communication research scientist. Recently, he started working at Sharif University of Technology, as a faculty member.

Dr. Maddah-Ali is a recipient of NSERC Postdoctoral Fellowship in 2007, a best paper award from IEEE International Conference on Communications (ICC) in 2014, the IEEE Communications Society and IEEE Information Theory Society Joint Paper Award in 2015, and the IEEE Information Theory Society Joint Paper Award in 2016.

A. Salman Avestimehr (S'03-M'08-SM16) is an Associate Professor at the Electrical Engineering Department of University of Southern California. He received his Ph.D. in 2008 and M.S. degree in 2005 in Electrical Engineering and Computer Science, both from the University of California, Berkeley. Prior to that, he obtained his B.S. in Electrical Engineering from Sharif University of Technology in 2003. He was an Assistant Professor at the ECE school of Cornell University from 2009 to 2013. He was also a postdoctoral scholar at the Center for the Mathematics of Information (CMI) at Caltech in 2008. His research interests include information theory, the theory of communications, and their applications.

Dr. Avestimehr has received a number of awards, including the Communications Society and Information Theory Society Joint Paper Award in 2013, the Presidential Early Career Award for Scientists and Engineers (PECASE) in 2011 for "pushing the frontiers of information theory through its extension to complex wireless information networks", the Young Investigator Program (YIP) award from the U. S. Air Force Office of Scientific Research in 2011, the National Science Foundation CAREER award in 2010, and the David J. Sakrison Memorial Prize in 2008. He is currently an Associate Editor for the IEEE Transactions on Information Theory.