

Two-User Erasure Interference Channels with Local Delayed CSIT

Alireza Vahid, Robert Calderbank

Abstract—We study the capacity region of two-user erasure interference channels with *local* delayed channel state information at the transmitters. In our model, transmitters have local mismatched outdated knowledge of the channel gains. We propose a transmission strategy that only relies on the delayed knowledge of the *outgoing* links at each transmitter and achieves the outer-bound for the scenario in which transmitters learn the *entire* channel state with delay. Our result reveals the subset of the channel state information that affects the capacity region the most.

We also identify cases in which local delayed knowledge of the channel state does not provide any gain over the zero knowledge assumption. To do so, we revisit a long-known intuition about interference channels that as long as the marginal distributions at the receivers are conserved, the capacity remains the same. We take this intuition and impose a certain spatial correlation among channel gains such that the marginal distributions remain unchanged. Then we provide an outer-bound on the capacity region of the channel with correlation that matches the capacity region when transmitters do not have access to channel state information.

Index Terms—Interference channel, local delayed CSIT, capacity, no CSIT.

I. INTRODUCTION

The canonical two-user interference channel (IC) introduced in [2] is a fundamental building block in wireless communications and information theory. The behavior and the capacity of multi-terminal wireless networks could not be understood without a good grasp of the two-user interference channel. Subsequently there developed a significant body of work aimed at understanding the capacity region of this problem (e.g., [3]–[5]). Again, when it comes to fading interference channels, understanding the capacity of the canonical two-user IC is of great importance. Recent results [6]–[9] address the capacity region of the canonical two-user fading IC under a specific channel distribution: the two-user erasure Interference Channel depicted in Fig. 1 where the channel gains at each time are drawn from the binary field according to some Bernoulli distributions. The input-output relation of this channel at time t is given by

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

Alireza Vahid is with the information initiative at Duke (iID), Duke University, Durham, NC, USA. Email: alireza.vahid@duke.edu.

Robert Calderbank is with departments of Electrical and Computer Engineering, Mathematics, and Computer Science, Duke University, Durham, NC, USA. Email: robert.calderbank@duke.edu.

The work of A. Vahid and R. Calderbank was supported in part by AFOSR under award No FA 9550-13-1-0076.

Preliminary parts of this work were presented at the 2015 International Symposium on Information Theory (ISIT) [1].

where $\bar{i} = 3 - i$, $G_{ii}[t], G_{\bar{i}i}[t] \in \{0, 1\}$, $X_i[t] \in \{0, 1\}$ is the transmit signal of transmitter i at time t , and $Y_i[t] \in \{0, 1\}$ is the observation of receiver i at time t . All algebraic operations are in \mathbb{F}_2 . This model is a very good abstraction of wireless packet networks as discussed in [10].

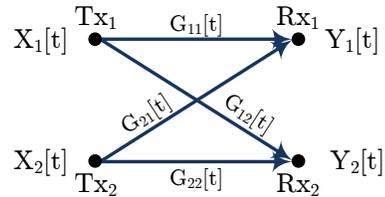


Fig. 1. Two-user Erasure Interference Channel.

In [8], the capacity region of the two-user erasure IC was characterized under the assumption of *global* delayed channel state information at the transmitters (CSIT) where each transmitter at time t knows all the channel realizations up to time $(t - 1)$. The result with global delayed CSIT includes novel transmission strategies and provides a new technique for deriving outer-bounds. However from the achievability perspective, the results rely strongly on the fact that transmitters learn the *entire* channel state information (CSI) with delay. While this assumption might be justified for the small canonical two-user IC, for large-scale networks such assumption might not be feasible at all. Thus, we aim to understand whether it is possible to achieve the same performance of global delayed CSIT with strictly smaller *local* delayed CSIT.

We consider several possible choices for the available delayed CSIT at each transmitter as shown in Fig. 2 (such locality of knowledge was considered in [11] for a different setup). We demonstrate that it is sufficient for each transmitter to only have the knowledge of its outgoing channel gains with delay in order to achieve the same performance of global delayed CSIT. From this result we learn that each transmitter has to resolve the interference it creates at the unintended receiver. In other words, “everyone should clean up their own mess!” We identify those cases when local delayed CSIT provides no gain in capacity over the baseline where transmitters have no knowledge of CSI. Basically, we identify the most important subset or “the most significant bits” (MSBs) of the delayed channel state information.

Our contributions are thus multi-fold. We propose a new transmission strategy that solely relies on local delayed knowledge of the outgoing links at each transmitter. We show

that this transmission strategy achieves the capacity region of the problem under the global delayed assumption. While this transmission strategy achieves the same region as the one proposed in [8], it has a much simpler structure and incorporates new ingredients. To be precise, the strategy of [8] cannot be applied when transmitters have access to local delayed CSI. Our proposed scheme smartly takes advantage of the statistics of the channel to compensate for the lack of global knowledge. Moreover, the proposed scheme reduces the complexity of the transmission protocol by reducing the number of virtual queues it generates at each transmitter when compared to that of [8]. Our result provides a better intuition and a deeper understanding of the coding opportunities that arise from delayed CSIT.

In order to identify the cases where local delayed CSIT does not provide any gain over no CSIT assumption, we borrow the intuition provided by Sato [3]: “the capacity region of all interference channels that have the same marginal distributions is the same.” We take this intuition and create a certain spatial correlation among channel gains such that the marginal distributions remain unchanged. Then, we provide an outer-bound on the capacity region of the channel with correlation which can be achieved with no CSIT.

There is some prior work assessing the value of delayed CSIT. It was used in [12] to create transmitted signals that are simultaneously useful for multiple users in a broadcast channel. These ideas were then extended to different wireless networks with delayed CSIT. Some examples are the study of erasure broadcast channels [13], the DoF region of broadcast channels [14], the approximate capacity region of Gaussian broadcast channels [15], [16], and the DoF region of multi-antenna multi-user Gaussian ICs and X channels [17]–[20].

The rest of the paper is organized as follows. In Section II, we formulate our problem. In Section III, we present our main results. Sections IV and V are dedicated to the proof of the main results. Next, in Section VI, we discuss the implications of our results for more general settings. Section VII concludes the paper.

II. PROBLEM SETTING

To study the capacity region of the two-user fading interference channels with local delayed CSIT, we consider an erasure model. In the erasure model, the channel gain from transmitter Tx_i to receiver Rx_j at time t is the binary field element denoted by $G_{ij}[t] \in \{0, 1\}$, $i, j \in \{1, 2\}$. Channel gains are distributed as independent Bernoulli random variables (independent across time and space). We refer to this model as binary fading. We define the channel state information at time instant t to be the set

$$G[t] \triangleq \{G_{11}[t], G_{12}[t], G_{21}[t], G_{22}[t]\}. \quad (2)$$

In this work, we focus on a homogeneous setting for simplicity of notations where

$$G_{ij}[t] \stackrel{d}{\sim} \mathcal{B}(p), \quad i, j = 1, 2, \quad (3)$$

for $0 \leq p \leq 1$. Define $q \triangleq 1 - p$.

In the sequel, we assume that each receiver has instantaneous knowledge of the channel state information. However, transmitter Tx_i will be aware of a subset of the CSI $\mathcal{S}_{\text{Tx}_i}$ with unit delay, $i = 1, 2$. More precisely,

$$\mathcal{S}_{\text{Tx}_i} \subseteq \{(1, 1), (1, 2), (2, 1), (2, 2)\}, \quad (4)$$

meaning that if $(k, \ell) \in \mathcal{S}_{\text{Tx}_i}$, then Tx_i at time t has access to $G_{k\ell}[1], G_{k\ell}[2], \dots, G_{k\ell}[t-1]$ (or simply $G_{k\ell}^{t-1}$). Moreover, we consider the symmetric setting where we have

$$(k, \ell) \in \mathcal{S}_{\text{Tx}_1} \Leftrightarrow (\bar{k}, \bar{\ell}) \in \mathcal{S}_{\text{Tx}_2}, \quad k, \ell = 1, 2, \quad (5)$$

for $\bar{k} \triangleq 3 - k$ and $\bar{\ell} \triangleq 3 - \ell$.

We consider a total of 8 possible choices of local delayed CSIT for each transmitter as follows.

- View V.0: This case would be our base-line that we refer to as no CSIT where

$$\mathcal{S}_{\text{Tx}_i} = \emptyset, \quad i = 1, 2. \quad (6)$$

- View V.1: In this case each transmitter is aware of channel value to its corresponding receiver with unit delay, *i.e.*

$$\mathcal{S}_{\text{Tx}_i} = \{(i, i)\}, \quad i = 1, 2. \quad (7)$$

- View V.2: In this case each transmitter is aware of channel value of the outgoing links with unit delay, *i.e.*

$$\mathcal{S}_{\text{Tx}_i} = \{(i, i), (i, \bar{i})\}, \quad i = 1, 2. \quad (8)$$

- View V.3: In this case each transmitter is aware of channel value of the direct links with unit delay, *i.e.*

$$\mathcal{S}_{\text{Tx}_i} = \{(i, i), (\bar{i}, \bar{i})\}, \quad i = 1, 2. \quad (9)$$

- View V.4: In this case each transmitter is aware of channel value of the links connected to its receiver with unit delay, *i.e.*

$$\mathcal{S}_{\text{Tx}_i} = \{(i, i), (\bar{i}, i)\}, \quad i = 1, 2. \quad (10)$$

- View V.5: In this case we have

$$\mathcal{S}_{\text{Tx}_i} = \{(i, i), (i, \bar{i}), (\bar{i}, \bar{i})\}, \quad i = 1, 2. \quad (11)$$

- View V.6: In this case we have

$$\mathcal{S}_{\text{Tx}_i} = \{(i, \bar{i}), (i, i), (\bar{i}, i)\}, \quad i = 1, 2. \quad (12)$$

- View V.7: In this case we have

$$\mathcal{S}_{\text{Tx}_i} = \{(i, i), (\bar{i}, i), (\bar{i}, \bar{i})\}, \quad i = 1, 2. \quad (13)$$

- View V.8: Finally View 8 corresponds to the scenario where each transmitter is aware of the entire CSI with delay. We shall refer to this view as the global delayed CSIT where

$$\mathcal{S}_{\text{Tx}_i} = \{(1, 1), (1, 2), (2, 1), (2, 2)\}, \quad i = 1, 2. \quad (14)$$

Fig. 2 pictorially depicts these 8 different views. We note that the transmitters’ knowledge in View V.0 is a subset of V.1; in V.1 is a subset of V.2, V.3, and V.4; in V.2 and V.3 is

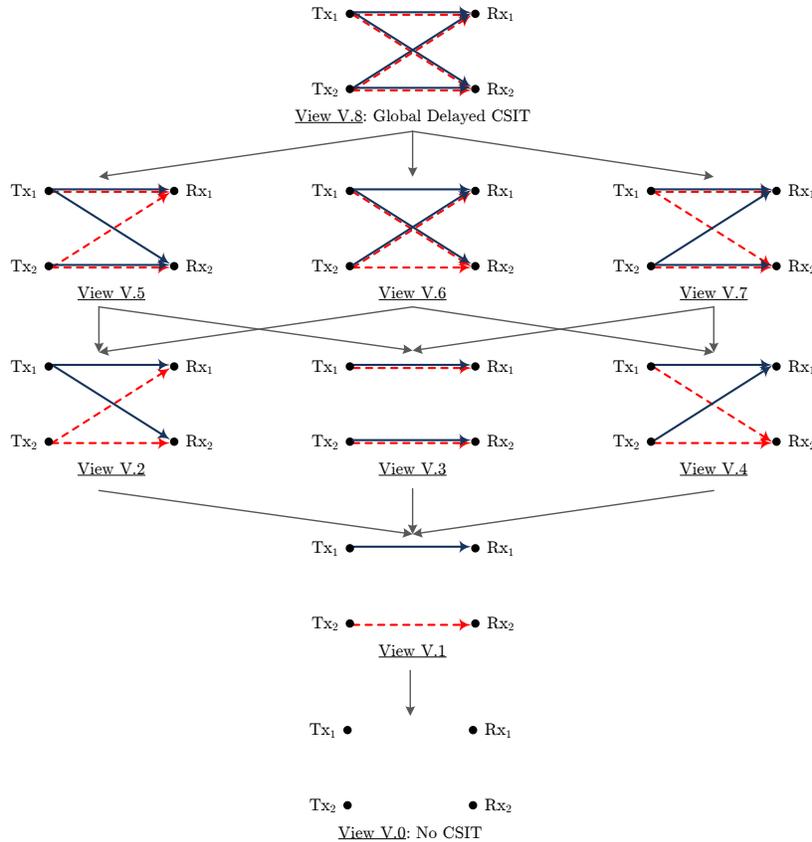


Fig. 2. Different choices of local delayed CSIT. Dark blue solid arrows denote the links for which Tx_1 learns the channel gains with unit delay; while red dashed arrows denote the links for which Tx_2 learns the channel gains with unit delay.

a subset of V.5; and so on. This hierarchical structure is also shown in Fig. 2 using downward arrows¹.

We consider the scenario in which Tx_i wishes to reliably communicate message $W_i \in \{1, 2, \dots, 2^{nR_i}\}$ to Rx_i during n channel uses, $i = 1, 2$. We assume that the messages and the channel gains are *mutually* independent and the messages are chosen uniformly. Let message W_i be encoded as X_i^n at transmitter Tx_i using the encoding function $f_i(W_i, \mathcal{S}_{\text{Tx}_i})$, which depends on the available channel state information at the transmitter. Receiver Rx_i is only interested in decoding W_i , and it will decode the message using the decoding function $\hat{W}_i = g_i(Y_i^n, G^n)$. An error occurs when $\hat{W}_i \neq W_i$. The average probability of decoding error is given by

$$\lambda_{i,n} = \mathbb{E}[P[\hat{W}_i \neq W_i]], \quad i = 1, 2, \quad (15)$$

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 exists 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 for the

given choice of $\mathcal{S}_{\text{Tx}_1}$ and $\mathcal{S}_{\text{Tx}_2}$. The capacity region for View $V.j$, *i.e*

$$\mathcal{C}(V.j), \quad j = 1, 2, \dots, 8, \quad (16)$$

is the closure of all achievable rate-tuples. In the following section, we present our main results.

III. STATEMENT OF THE MAIN RESULTS

Our main objective is to understand the ramification of local delayed CSIT on the capacity region of two-user erasure Interference Channels. We establish the capacity region of two-user erasure ICs with no CSIT and global delayed CSIT as our benchmarks. Then we are interested in finding the answer to the following questions:

- 1) What is the minimum amount of delayed CSIT required to outperform no CSIT?
- 2) Is it possible to achieve the performance of global delayed CSIT with a strictly smaller subset of knowledge at each transmitter?

Our work essentially identifies the ‘‘MSBs’’ of the delayed CSIT in two-user erasure ICs.

A. Benchmarks

Our base-line is the no CSIT scenario. In other words, the only available knowledge at the transmitters is the distribution

¹A similar set of local views for the channel state information was studied in [11] in the context of two-user Gaussian interference channel (not fading) to identify the views in which one could outperform TDMA.

from which the channel gains are drawn. In this case, it is easy to see that for any input distribution, the two received signals are *statistically* the same. Therefore, the capacity region in this case is the same as the intersection of the capacity regions of the multiple-access channels (MACs) formed at the two receivers. Thus, $\mathcal{C}(V.0)$, is the set of all rate-tuples (R_1, R_2) satisfying

$$\begin{cases} 0 \leq R_i \leq p, & i = 1, 2, \\ R_1 + R_2 \leq 1 - q^2. \end{cases} \quad (17)$$

The other extreme point is the global delayed CSIT model. From [8], we know that the capacity region of the two-user Binary Fading IC with global delayed CSIT, $\mathcal{C}(V.8)$, is the set of all rate-tuples (R_1, R_2) satisfying

$$\begin{cases} 0 \leq R_i \leq p, & i = 1, 2, \\ R_i + (1 + q) R_i \leq p(1 + q)^2, & i = 1, 2. \end{cases} \quad (18)$$

These benchmarks are depicted in Fig. 3 for $p = 0.5$. We note that $\mathcal{C}(V.0)$ is a strict subset of $\mathcal{C}(V.8)$, and we are interested in understanding the impact of local delayed CSIT on the capacity region.

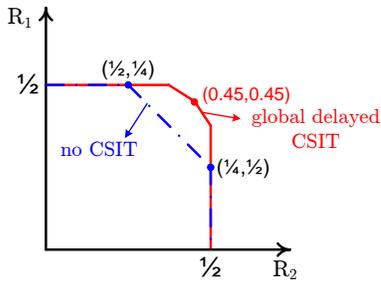


Fig. 3. Capacity Region of the Two-user Binary Fading Interference Channel with no and global delayed CSIT and $p = 0.5$.

B. Main Results

Theorem 1 highlights the cases where no performance gain over the no CSIT assumption is feasible. Then in Theorem 2, we show that the performance gain of global delayed CSIT (View V.8) can be obtained with Views V.2, V.5, and V.6. Note that the achievable region for View V.2 is a subset of Views V.5 and V.6 due to the hierarchical structure shown in Fig. 2.

Theorem 1: For the two-user binary fading interference channel with local delayed CSIT of Views V.1, V.3, and V.4, the capacity region coincides with the capacity region of no CSIT (View V.0), *i.e.*

$$\mathcal{C}(V.j) \subseteq \mathcal{C}(V.0) \quad \text{for } j = 1, 3, 4. \quad (19)$$

The key in proving Theorem 1 is to derive an outer-bound that matches that of the no CSIT assumption. In doing that, we use the intuition given by Sato [3]: the capacity region of all interference channels that have the same marginal distributions is the same. We take this intuition and impose a certain

spatial correlation among channel gains such that the marginal distributions remain unchanged. Then we provide an outer-bound on the capacity region of the channel with correlation and we show that it can be achieved with no CSIT.

Remark 1: In this paper, we only focus on the impact of local *delayed* CSIT on the capacity region of the two-user erasure ICs. Interestingly, the proof of Theorem 1 holds even when channels are learned instantaneously.

Theorem 2: For the two-user binary fading interference channel with local delayed CSIT of Views V.2, V.5, and V.6, the capacity region coincides with the capacity region of global delayed CSIT (V.8), *i.e.*

$$\mathcal{C}(V.8) \subseteq \mathcal{C}(V.j) \quad \text{for } j = 2, 5, 6. \quad (20)$$

For View V.2, we need to provide an achievability strategy that achieves the same performance as V.8. Then the result for View V.5 and V.6 follows. The capacity region $\mathcal{C}(V.8)$ was established in [8] and a novel transmission strategy was introduced. The strategy in [8] is carried on over several phases. Each channel realization creates multiple coding opportunities which can be exploited in the following phases. To achieve the capacity region, an efficient arrangement of combination, concatenation, and merging of the opportunities is needed. Our result for View V.2 provides a better intuition and a deeper understanding of such opportunities and reveals redundancies in the prior work.

Remark 2: The capacity region under View V.7 remains open. While it seems the coding opportunities cannot be detected with such channel state information, the lack of an outer-bound does not allow us to solve the problem for this case. In a sense, each transmitter has “too much” knowledge eliminating the choices for correlation that were needed to obtain the result in Theorem 1. We discuss this case in more details in Section VI.

In the remaining of the paper, we provide the proof of Theorem 1 and Theorem 2. In Section VI, we discuss the challenges for View V.7 and provide some connections to the k -user setting ($k > 2$).

IV. PROOF OF THEOREM 1: CHOOSING SPATIAL CORRELATIONS

In this section we provide the proof of Theorem 1. In particular, we demonstrate that with local delayed CSIT according to Views V.1, V.3, and V.4 the capacity region is the same as View V.0 (no CSIT) as given in (17). Since the knowledge in View V.1 is a subset of other cases, we do not need to provide a separate proof for this case. Thus, we only need to provide the proof for Views V.3 and V.4.

Here, we introduce a certain spatial correlation among channel gains such that the marginal distributions remain unchanged. Then, we provide an outer-bound on the capacity region of the channel with correlation that matches the no CSIT region of (17). The following facts will help us throughout the proof. For any choice of local delayed CSIT, we have

$$\Pr(X_1^n, X_2^n | G^n) = \Pr(X_1^n | G^n) \Pr(X_2^n | G^n). \quad (21)$$

Moreover, if $\mathcal{S}_{\mathcal{T}_{x_1}} \cap \mathcal{S}_{\mathcal{T}_{x_2}} = \emptyset$, then we have

$$\Pr(X_1^n, X_2^n) = \Pr(X_1^n) \Pr(X_2^n). \quad (22)$$

Derivation of (21) and (22) is a straightforward exercise and is omitted here.

In [21], it was shown that under local CSI at the receivers (CSIR):

$$\mathcal{C}(\text{V.1}) \subseteq \mathcal{C}(\text{V.0}). \quad (23)$$

In fact, the proof relied heavily on the assumption of local CSIR and cannot be extended to our setting where the receivers know the *entire* CSI instantaneously.

A. Proof for View V.4

We know that the error probabilities are solely a function of marginal distributions at the receivers. Thus, as long as the marginal distributions remain the same, the capacity region remains the same. Consider the two-user erasure IC with local delayed CSIT according to View V.4. We have

$$\mathcal{S}_{\mathcal{T}_{x_1}} = \{(1, 1), (2, 1)\} \quad \text{and} \quad \mathcal{S}_{\mathcal{T}_{x_2}} = \{(1, 2), (2, 2)\}. \quad (24)$$

Thus writing the marginal distribution at receiver Rx_1 , we get

$$\begin{aligned} & \Pr(Y_1^n, G^n | X_1^n, X_2^n) \\ & \stackrel{(a)}{=} \Pr(G^n | X_1^n, X_2^n) \Pr(Y_1^n | X_1^n, X_2^n, G^n) \\ & = \frac{\Pr(G^n, X_1^n, X_2^n)}{\Pr(X_1^n, X_2^n)} \Pr(Y_1^n | X_1^n, X_2^n, G^n) \\ & \stackrel{(b)}{=} \frac{\Pr(G^n) \Pr(X_1^n | G^n) \Pr(X_2^n | G^n)}{\Pr(X_1^n) \Pr(X_2^n)} \Pr(Y_1^n | X_1^n, X_2^n, G^n) \\ & \stackrel{(c)}{=} \frac{\Pr(G^n) \Pr(X_1^n | G_{11}^n, G_{21}^n) \Pr(X_2^n | G_{12}^n, G_{22}^n)}{\Pr(X_1^n) \Pr(X_2^n)} \\ & \quad \times \Pr(Y_1^n | X_1^n, X_2^n, G^n) \\ & = \left[\frac{\Pr(G_{11}^n, G_{21}^n) \Pr(X_1^n | G_{11}^n, G_{21}^n)}{\Pr(X_1^n)} \right] \\ & \quad \times \left[\frac{\Pr(G_{12}^n, G_{22}^n) \Pr(X_2^n | G_{12}^n, G_{22}^n)}{\Pr(X_2^n)} \right] \\ & \quad \times \mathbf{1}_{\{Y_1^n = G_{11}^n X_1^n \oplus G_{21}^n X_2^n\}}, \end{aligned} \quad (25)$$

where (a) follows from the chain rule; (b) follows from the chain rule and the fact that X_1^n and X_2^n are independent in View V.4 as given in (22); and (c) holds since X_i^n is independent of G_{ii}^n and G_{ii}^n , $i = 1, 2$. Similarly, we can write the marginal distribution at receiver Rx_2 .

$$\begin{aligned} & \Pr(Y_2^n, G^n | X_1^n, X_2^n) = \left[\frac{\Pr(G_{11}^n, G_{21}^n) \Pr(X_1^n | G_{11}^n, G_{21}^n)}{\Pr(X_1^n)} \right] \\ & \quad \times \left[\frac{\Pr(G_{12}^n, G_{22}^n) \Pr(X_2^n | G_{12}^n, G_{22}^n)}{\Pr(X_2^n)} \right] \mathbf{1}_{\{Y_2^n = G_{12}^n X_1^n \oplus G_{22}^n X_2^n\}}. \end{aligned} \quad (26)$$

We conclude that as long as the joint distributions

$$\Pr(G_{11}^n, G_{21}^n) \quad \text{and} \quad \Pr(G_{12}^n, G_{22}^n) \quad (27)$$

remain the same, the marginal distributions remain unchanged. This immediately implies that the capacity region $\mathcal{C}(\text{V.4})$ coincides with the capacity region of any channel that has

same marginal distributions $\Pr(G_{11}^n, G_{21}^n)$ and $\Pr(G_{12}^n, G_{22}^n)$. We use this result to create a channel with a specific correlation that helps us deriving the desired result.

Consider a binary fading interference channel similar to the channel described in Section II, but where channel gains have certain spatial correlation. We distinguish the RVs in this channel using (\cdot) notation (e.g., $\tilde{X}_1[t]$). The input-output relation of this channel at time instant t is given by

$$\tilde{Y}_i[t] = \tilde{G}_{ii}[t] \tilde{X}_i[t] \oplus \tilde{G}_{\bar{i}i}[t] \tilde{X}_{\bar{i}}[t], \quad i = 1, 2. \quad (28)$$

We assume that the channel gains are distributed independently over time. However, we have

$$\tilde{G}_{ii}[t] = \tilde{G}_{\bar{i}\bar{i}}[t] \quad i = 1, 2. \quad (29)$$

In other words, the channel gains corresponding to incoming links at each receiver are still independent but the outgoing links at each transmitter are correlated. We know that the capacity region of this channel coincides with $\mathcal{C}(\text{V.4})$. We have included a more detailed discussion in Appendix A.

Suppose rate-tuple $(\tilde{R}_1, \tilde{R}_2)$ is achievable. The derivation of individual bounds for this channel

$$\tilde{R}_i \leq p, \quad i = 1, 2, \quad (30)$$

is a straight forward exercise and omitted here. For the sum-rate bound, we have

$$\begin{aligned} & n(\tilde{R}_1 + \tilde{R}_2 - \epsilon_n) \stackrel{(a)}{\leq} I(\tilde{W}_1; \tilde{Y}_1^n | \tilde{W}_2, \tilde{G}^n) + I(\tilde{W}_2; \tilde{Y}_2^n | \tilde{G}^n) \\ & = H(\tilde{Y}_1^n | \tilde{W}_2, \tilde{G}^n) - \underbrace{H(\tilde{Y}_1^n | \tilde{W}_1, \tilde{W}_2, \tilde{G}^n)}_{=0} \\ & \quad + H(\tilde{Y}_2^n | \tilde{G}^n) - H(\tilde{Y}_2^n | \tilde{W}_2, \tilde{G}^n) \\ & = H(\tilde{Y}_2^n | \tilde{G}^n) + H(\tilde{G}_{11}^n \tilde{X}_1^n | \tilde{X}_2^n, \tilde{W}_2, \tilde{G}^n) \\ & \quad - H(\tilde{G}_{12}^n \tilde{X}_1^n | \tilde{X}_2^n, \tilde{W}_2, \tilde{G}^n) \\ & = H(\tilde{Y}_2^n | \tilde{G}^n) + H(\tilde{G}_{11}^n \tilde{X}_1^n | \tilde{G}^n) - H(\tilde{G}_{12}^n \tilde{X}_1^n | \tilde{G}^n) \\ & \stackrel{(b)}{=} H(\tilde{Y}_2^n | \tilde{G}^n) \leq (1 - q^2) n, \end{aligned} \quad (31)$$

where $\epsilon_n \rightarrow 0$ as $n \rightarrow \infty$; (a) follows from Fano's inequality and the fact that messages and channel gains are mutually independent; (b) holds since according to (28), we have

$$H(\tilde{G}_{11}^n \tilde{X}_1^n | \tilde{G}^n) = H(\tilde{G}_{12}^n \tilde{X}_1^n | \tilde{G}^n). \quad (32)$$

Dividing both sides by n and letting $n \rightarrow \infty$, we obtain

$$\tilde{R}_1 + \tilde{R}_2 \leq 1 - q^2. \quad (33)$$

Note that the region described by (30) and (33) matches the no CSIT region of (17). This completes the converse proof for View V.4 since $\mathcal{C}(\text{V.4})$ is included in the region described by (30) and (33).

Remark 3: The proof does not rely on fact that the CSI is obtained with delay. In fact, if we assume instantaneous local CSIT, the proof still holds. However in this paper, we only focus on local *delayed* CSIT for which we present achievability strategy as well.

B. Proof for View V.3

The proof is very similar to that of View V.4. Under local delayed CSIT of View V.3, we have

$$\mathcal{S}_{\text{Tx}_1} = \{(1, 1), (2, 2)\} \quad \text{and} \quad \mathcal{S}_{\text{Tx}_2} = \{(1, 1), (2, 2)\}. \quad (34)$$

Thus writing the marginal distribution for receiver Rx_1 , we get

$$\begin{aligned} & \Pr(Y_1^n, G^n | X_1^n, X_2^n) \\ &= \Pr(X_1^n, X_1^n)^{-1} [\Pr(G_{12}^n, G_{21}^n) \Pr(X_1^n | G_{11}^n, G_{22}^n)] \\ & \times [\Pr(G_{11}^n, G_{22}^n) \Pr(X_2^n | G_{11}^n, G_{22}^n)] \mathbf{1}_{\{Y_1^n = G_{11}^n X_1^n \oplus G_{21}^n X_2^n\}}, \end{aligned} \quad (35)$$

Similarly, we can write the marginal distribution for receiver Rx_2 .

$$\begin{aligned} & \Pr(Y_2^n, G^n | X_1^n, X_2^n) \\ &= \Pr(X_1^n, X_1^n)^{-1} [\Pr(G_{12}^n, G_{21}^n) \Pr(X_1^n | G_{11}^n, G_{22}^n)] \\ & \times [\Pr(G_{11}^n, G_{22}^n) \Pr(X_2^n | G_{11}^n, G_{22}^n)] \mathbf{1}_{\{Y_2^n = G_{12}^n X_1^n \oplus G_{22}^n X_2^n\}}. \end{aligned} \quad (36)$$

We conclude that as long as the joint distributions

$$\Pr(G_{11}^n, G_{21}^n), \quad \Pr(G_{12}^n, G_{22}^n), \quad \text{and} \quad \Pr(G_{11}^n, G_{22}^n), \quad (37)$$

remain the same, the capacity region remains the same. We note that the same correlation introduced in (28) can be applied here. Thus, the rest of the proof is identical to the previous subsection.

V. PROOF OF THEOREM 2: OPPORTUNISTIC RETRANSMISSIONS BASED ON LOCAL DELAYED CSIT

In this section, we provide an achievability strategy that solely relies on what the transmitters know in View V.2, *i.e.* the outdated knowledge of the channel gains associated with the outgoing links at each transmitter. We show that the capacity region with global delayed CSIT can be achieved with this local delayed knowledge. This result is surprising since the transmission strategy in prior work [8] for the case of global delayed CSIT relies heavily on the delayed knowledge of the entire channel state information at each transmitter. In fact, the transmission strategy of [8] cannot be applied to the case where transmitter learn the CSI locally.

We note that the channel knowledge of View V.2 is a subset of View V.5 or View V.6; thus there is no need to prove the result separately for those cases. Prior to providing the achievability proof, we discuss some techniques and coding opportunities that we utilize later in this section.

A. Techniques and Coding Opportunities

In this subsection, we describe the coding opportunities that arise from the delayed knowledge of the channel state information. We first assume that nodes have global delayed CSIT to clearly describe the opportunities. Then we discuss the challenges that stem from the locality of the channel state information at the transmitters. Next, we demonstrate how to overcome these challenges. Finally in the following

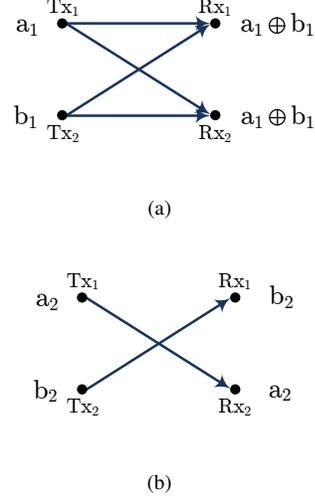


Fig. 4. Providing $a_1 \oplus a_2$ available at Tx_1 and $b_1 \oplus b_2$ available at Tx_2 to both receivers is sufficient to decode the bits.

subsection, we describe the transmission strategy in detail. We categorize the opportunities into three groups as follows.

[Pairing Across Realizations Type-I] Suppose at a time instant, each one of the transmitters simultaneously sends one data bit. The bits of Tx_1 and Tx_2 are denoted by a_1 and b_1 respectively. Assume that the channel realization was according to Fig. 4(a). In another time instant, each one of the transmitters sends one data bit, say a_2 and b_2 respectively. Assume that the channel realization was according to Fig. 4(b). Now, we observe that providing $a_1 \oplus a_2$ and $b_1 \oplus b_2$ to both receivers is sufficient to decode the bits. For instance if Rx_1 is provided with $a_1 \oplus a_2$ and $b_1 \oplus b_2$, then it will use b_2 to decode b_1 , from which it can obtain a_1 , and finally using a_1 and $a_1 \oplus a_2$, it can decode a_2 . The linear combinations $a_1 \oplus a_2$ and $b_1 \oplus b_2$ that are available at transmitters one and two respectively, can be thought of as bits of “common interest.”

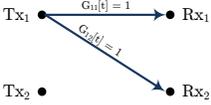
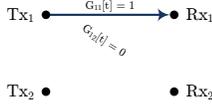
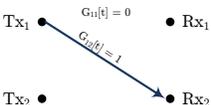
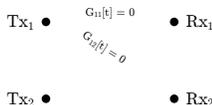
[Pairing Across Realizations Type-II] Suppose the scenario depicted in Fig. 5 is realized. Now, we observe that providing $a_3 \oplus a_4$ available at transmitter Tx_1 to both receivers is useful. For instance if Rx_2 is provided with $a_3 \oplus a_4$, it will use a_4 to decode a_3 , from which it can obtain b_3 . It is easy to visualize the similar opportunity for transmitter Tx_2 .

[Pairing Across Realizations Type-III] Suppose the scenario depicted in Fig. 6(a) has occurred. It is easy to see that bit a_5 is a useful bit for *both* receivers after this point. A similar situation is depicted in Fig. 6(b) where bit b_6 at transmitter Tx_2 becomes a bit of common interest.

Challenges with local delayed CSIT: Given local delayed CSIT of View V.2, each transmitter can identify a total of 4 possible configurations as summarized in Table I for Tx_1 (the configurations known to Tx_2 are simply derived by interchanging user IDs). With the limited knowledge available at each transmitter, the aforementioned opportunities may not be detected properly. For instance consider the scenario depicted in Fig. 7. From transmitter Tx_1 's point of view, this scenario is identical to the one depicted in Fig. 4 and Fig. 5. However,

TABLE I

THE FOUR POSSIBLE CONFIGURATIONS THAT CAN BE IDENTIFIED BY TRANSMITTER Tx_1 . THE BIT TRANSMITTED BY Tx_1 IS DENOTED “ a .” DEPENDING ON THE IDENTIFIED CONFIGURATION, THE STATUS OF THE TRANSMITTED BIT IS UPDATED TO A QUEUE DEFINED IN SECTION V-B.

case ID	channel realization at time instant n	state transition	case ID	channel realization at time instant n	state transition
1		$a \rightarrow Q_{1,1}$	3		$a \rightarrow Q_{1 \rightarrow F}$
2		$a \rightarrow Q_{1,2}$	4		$a \rightarrow Q_{1 \rightarrow 1}$

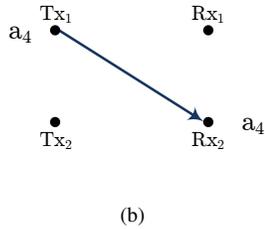
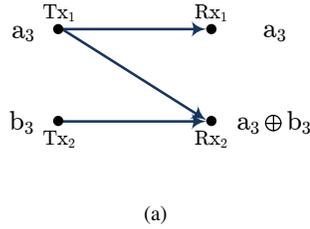


Fig. 5. The combination $a_3 \oplus a_4$ available at Tx_1 is of interest of both receivers.

providing $a_5 \oplus a_6$ to Rx_2 is not useful anymore. The proposed scheme of [8] cannot overcome this challenge and fails with local knowledge as a result. Therefore, one must be careful on how to identify the opportunities and how to exploit them for future communications.

Remark 4: It is important to keep in mind that while transmitters have only local knowledge of the CSI, receivers have global knowledge. This enables receivers to figure out future actions taken by the transmitters based on their past observations of the channel realizations.

Overcoming the challenges with local delayed CSIT: Consider the example demonstrated in Fig. 7. Transmitter Tx_1 can overcome the challenge described above by relying on the fact that statistically a fraction q of the bits that at the time of transmission faced $G_{12}[t] = 1$, are already known to Rx_2 (since with probability q we have $G_{22}[t] = 0$). Keeping this fact in mind, transmitter Tx_1 will create enough linearly independent combinations of the bits such that receiver Rx_2

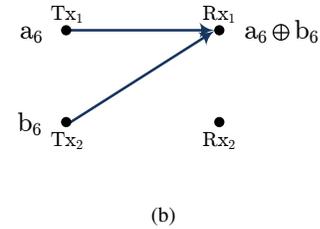
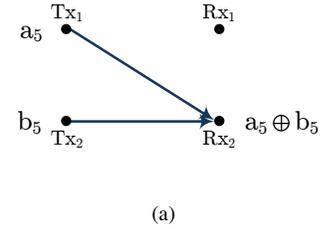


Fig. 6. In each of the channel realizations, one bit a_5 becomes of interest of both receivers. In (a) bit a_5 is a useful bit for both receivers, while in (b) bit b_6 is a useful bit for both receivers.

can recover the required bits. The global knowledge at the receivers is essential for them to know which bits are going to be retransmitted and in what order. This technique is described in more detail in the following subsection.

We also note that out of the four configurations in Table I, in configurations 3 and 4 the future task of Tx_1 is easy. Suppose at time t , transmitter Tx_1 sends one data bit. Later, using local delayed CSIT, Tx_1 figures out that $G_{11}[t] = 1$ and $G_{12}[t] = 0$. In this case, we say that the bit is delivered and if interference was created at Rx_1 then it would be the responsibility of Tx_2 to resolve it in future. If $G_{11}[t]$ and $G_{12}[t]$ were both equal to 0, then the bit must be retransmitted.

We show that with only local delayed CSIT of View V.2, we can achieve \mathcal{C} (V.8).

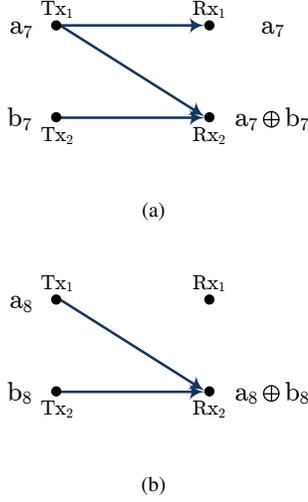


Fig. 7. From transmitter Tx_1 's point of view, this scenario is identical to the one depicted in Fig. 4 and Fig. 5. However, $a_7 \oplus a_8$ is no longer useful for Rx_2 .

B. Transmission Strategy

In this Section, we focus on the achievability strategy for $p = 0.5$. This would simplify the transmission strategy and allows us to focus on the key issue of local delayed CSIT.

We have again depicted the capacity region with global delayed CSIT $\mathcal{C}(\text{V.8})$ in Fig. 8 for $p = 0.5$. We show that with only local delayed CSIT of View V.2, we can achieve $\mathcal{C}(\text{V.8})$. To do so, it suffices to prove achievability for the corner points $(0.45, 0.45)$ and $(0.375, 0.5)$.

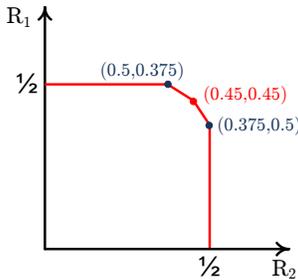


Fig. 8. To show that with only local delayed CSIT of View V.2, we can achieve $\mathcal{C}(\text{V.8})$, it suffices to prove the achievability for corner points $(0.45, 0.45)$ and $(0.375, 0.5)$.

We focus on the corner point $(R_1, R_2) = (0.45, 0.45)$. The achievability strategy for the other corner points is presented in Appendix B. Then in Appendix C, we describe the changes needed in the transmission strategy when considering $0 \leq p \leq 1$.

Suppose each transmitter wishes to communicate m bits to its intended receiver. We show that this task can be accomplished (with vanishing error probability as $m \rightarrow \infty$) in

$$\frac{20}{9}m + \frac{35}{3}m^{\frac{2}{3}} \quad (38)$$

time instants. This immediately implies the achievability for the corner point $(R_1, R_2) = (0.45, 0.45)$. Our transmission strategy comprises two phases as described below.

Phase 1: At the beginning of the communication block, we assume that the m bits at Tx_i are in queue $Q_{i \rightarrow i}$ (the initial state of the bits), $i = 1, 2$. At each time instant t , Tx_i sends out a bit from $Q_{i \rightarrow i}$, and this bit will either stay in the initial queue or transition to one of the queues listed in Table I. If at time instant t , $Q_{i \rightarrow i}$ is empty, then Tx_i , $i = 1, 2$, remains silent until the end of Phase 1.

- (A) $Q_{i \rightarrow F}$: The bits for which no retransmission is required and thus we consider delivered;
- (B) $Q_{i,1}$: The bits for which at the time of communication, all channel gains known to Tx_i with unit delay were equal to 1;
- (C) $Q_{i,2}$: The bits for which at the time of communication, we have $G_{ii}[t] = 0$ and $G_{i\bar{i}}[t] = 1$.

Each transmitter can identify a total of 4 possible configurations as summarized in Table I for Tx_1 . Phase 1 continues for

$$\frac{4}{3}m + m^{\frac{2}{3}} \quad (39)$$

time instants, and if at the end of this phase, either of the queues $Q_{i \rightarrow i}$ is not empty, we declare error type-I and halt the transmission (we assume m is chosen such that $m^{\frac{2}{3}} \in \mathbb{Z}$). We assume that the queues are column vectors and bits are placed according to the order they join the queue.

Assuming that the transmission is not halted, let $N_{i,1}$ and $N_{i,2}$ denote the number of bits in queues $Q_{i,1}$ and $Q_{i,2}$ respectively at the end of the transitions, $i = 1, 2$. The transmission strategy will be halted and error type-II occurs, if any of the following events happens.

$$\begin{aligned} N_{i,1} &> \mathbb{E}[N_{i,1}] + 2m^{\frac{2}{3}} \triangleq n_{i,1}, \quad i = 1, 2; \\ N_{i,2} &> \mathbb{E}[N_{i,2}] + 2m^{\frac{2}{3}} \triangleq n_{i,2}, \quad i = 1, 2. \end{aligned} \quad (40)$$

From basic probability, we have

$$\mathbb{E}[N_{i,1}] = \mathbb{E}[N_{i,2}] = \frac{m}{3}, \quad (41)$$

thus we get

$$n_{i,1} = n_{i,2} = \frac{m}{3} + 2m^{\frac{2}{3}}. \quad (42)$$

At the end of Phase 1, we add 0's (if necessary) in order to make queues $Q_{i,1}$ and $Q_{i,2}$ of size equal to $n_{i,1}$ and $n_{i,2}$ respectively as given above, $i = 1, 2$.

Moreover since channel gains are distributed independently, statistically half of the bits in $Q_{i,1}$ and half of the bits in $Q_{i,2}$ are known to $\text{Rx}_{\bar{i}}$, $i = 1, 2$. Denote the number of bits in $Q_{i,j}$ known to $\text{Rx}_{\bar{i}}$ by

$$N_{i,j|\text{Rx}_{\bar{i}}}, \quad i, j \in \{1, 2\}. \quad (43)$$

At the end of communication, if we have

$$N_{i,j|\text{Rx}_{\bar{i}}} < \frac{1}{2}n_{i,j} - m^{\frac{2}{3}}, \quad i, j \in \{1, 2\}, \quad (44)$$

we declare error type-III. Note that transmitters cannot detect error type-III, but receivers have sufficient information to do so.

Furthermore using the Bernstein inequality, we can show that the probability of errors of types I, II, and III decreases exponentially with m . For the rest of this subsection, we assume that Phase 1 is completed and no error has occurred.

Transmitter Tx_i creates two matrices $\mathbf{C}_{i,1}$ and $\mathbf{C}_{i,2}$, $i = 1, 2$, of size $\left(\frac{m}{3} + 4m^{\frac{2}{3}}\right) \times \left(\frac{m}{3} + 2m^{\frac{2}{3}}\right)$ each, where entries to each matrix are drawn from i.i.d. $\mathcal{B}(0.5)$ distribution. We assume that these matrices are generated prior to communication and are shared with receivers. Transmitter Tx_i does not need to know $\mathbf{C}_{i,1}$ or $\mathbf{C}_{i,2}$, $i = 1, 2$. Note that as $m \rightarrow \infty$, these matrices have full column-rank with probability 1. We refer the reader for a detailed discussion on the rank of randomly generated matrices in a finite field to [22].

Phase 2 [transmitting random linear combinations]: In this phase, transmitter Tx_i combines the bits in $Q_{i,1}$ and $Q_{i,2}$ to create \tilde{Q}_i using the following equation.

$$\tilde{Q}_i \triangleq \mathbf{C}_{i,1}Q_{i,1} \oplus \mathbf{C}_{i,2}Q_{i,2}, \quad i = 1, 2. \quad (45)$$

Then the goal is to provide the bits in \tilde{Q}_1 and \tilde{Q}_2 to *both* receivers. The problem resembles a network with two transmitters and two receivers where each transmitter Tx_i wishes to communicate an independent message W_i to *both* receivers as depicted in Fig. 9, $i = 1, 2$. The channel gain model is the same as described in Section II. We refer to this problem as the two-multicast problem. It is a straightforward exercise to show that for this problem, a rate-tuple of $(R_1, R_2) = \left(\frac{3}{8}, \frac{3}{8}\right)$ is achievable. In other words, for fixed $\epsilon, \delta > 0$, rate-tuple $(R_1, R_2) = \left(\frac{3}{8} - \frac{\delta}{2}, \frac{3}{8} - \frac{\delta}{2}\right)$ is achievable with error less than or equal to ϵ .

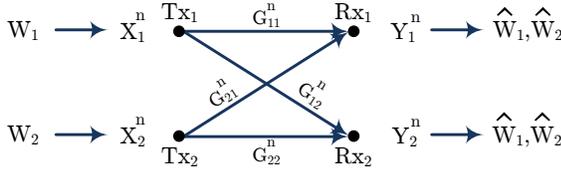


Fig. 9. Two-multicast network. Transmitter Tx_i wishes to reliably communicate message W_i to both receivers, $i = 1, 2$. The capacity region with no or delayed CSIT is the same.

Fix $\epsilon, \delta > 0$. Then, transmitters encode and communicate the bits in \tilde{Q}_1 and \tilde{Q}_2 using the achievability strategy of the two-multicast problem during Phase 2. This phase lasts for

$$\frac{\frac{2m}{3} + 8m^{\frac{2}{3}}}{\frac{3}{4} - \delta} \quad (46)$$

time instants. We assume \tilde{Q}_1 and \tilde{Q}_2 are decoded successfully at both receivers and no error has occurred.

Decoding: At the end of Phase 2, receiver Rx_i removes the known bits from $Q_{i,1}$ and $Q_{i,2}$ (from (44), we know that Rx_i has knowledge of at least $\frac{m}{3}$ bits).

Thus after removing the known bits, receiver Rx_i has access to $\frac{m}{3} + 4m^{\frac{2}{3}}$ random linear combinations of (at most) $\frac{m}{3} + 4m^{\frac{2}{3}}$ unknown bits. Consequently, Rx_i can reconstruct all the

bits in $Q_{i,1}$ and $Q_{i,2}$ with probability 1 as $m \rightarrow \infty$. Then, receiver Rx_i uses the bits in $Q_{i,1}$ and $Q_{i,2}$ to remove the interference. Upon successfully removing interfering bits, the bits intended for Rx_i can be reconstructed from the available linear combinations. The reconstructing of the intended bits can be carried out error free with probability 1 as $m \rightarrow \infty$.

The total communication time is then equal to the length of Phase 1 plus the length of Phase 2. Thus when $\epsilon, \delta \rightarrow 0$, the total communication time is

$$\frac{4}{3}m + m^{\frac{2}{3}} + \frac{4}{3} \left(\frac{2m}{3} + 8m^{\frac{2}{3}} \right) = \frac{20}{9}m + \frac{35}{3}m^{\frac{2}{3}}. \quad (47)$$

Hence, if we let $m \rightarrow \infty$, the decoding error probability at each phase of delivering the bits goes to zero exponentially, and we achieve a symmetric sum-rate of

$$R_1 = R_2 = \lim_{m \rightarrow \infty} \frac{m}{\frac{20}{9}m + \frac{35}{3}m^{\frac{2}{3}}} = 0.45. \quad (48)$$

This completes the achievability proof for the corner point $(R_1, R_2) = (0.45, 0.45)$.

VI. DISCUSSION

In this section, we discuss the problem of two-user erasure IC with local delayed CSIT given by View V.7, and then we try to understand the implications of our results in broader settings.

A. Two-user Erasure IC with local delayed CSIT of View V.7

Consider the two-user erasure IC with local delayed CSIT according to View V.7. We have

$$\begin{aligned} \mathcal{S}_{\text{Tx}_1} &= \{(1, 1), (2, 1), (2, 2)\} \quad \text{and} \\ \mathcal{S}_{\text{Tx}_2} &= \{(1, 1), (1, 2), (2, 2)\}. \end{aligned} \quad (49)$$

Thus writing the marginal distribution at receiver Rx_1 , we get

$$\begin{aligned} &\Pr(Y_1^n, G^n | X_1^n, X_2^n) \\ &= \left[\frac{\Pr(G_{11}^n, G_{12}^n, G_{21}^n, G_{22}^n)}{\Pr(X_1^n, X_2^n)} \right] \\ &\times \Pr(X_1^n | G_{11}^n, G_{21}^n, G_{22}^n) \Pr(X_2^n | G_{11}^n, G_{12}^n, G_{22}^n) \\ &\mathbf{1}_{\{Y_1^n = G_{11}^n X_1^n \oplus G_{21}^n X_2^n\}}. \end{aligned} \quad (50)$$

Here, note that we can no longer use our trick in Section IV. For instance, if we set

$$\tilde{G}_{11}[t] = \tilde{G}_{12}[t], \quad (51)$$

then, we have changed the channel from Tx_2 's point of view and thus, the marginal distributions cannot be preserved.

On the other hand, as discussed in Section V, delayed knowledge of G_{ii}^n has an important role on the future decisions taken by Tx_i , $i = 1, 2$. In fact, we cannot distinguish $Q_{i,1}$ from $Q_{i,2}$ without delayed knowledge of $G_{i\bar{i}}^n$, and thus, our achievability strategy cannot be utilized with local delayed CSIT of View V.7.

In the absence of an achievability that goes beyond the capacity region with no CSIT, or a converse that matches that of no CSIT, the capacity region with local delayed CSIT of View V.7 remains open.

B. k -user Erasure IC with delayed CSIT

Here, we take the results and intuitions obtained for the two-user erasure IC and try to understand the implications in broader settings. We consider the capacity region of the k -user erasure IC (see Fig. 10) and the degrees of freedom (DoF) region of the k -user Gaussian IC with Delayed CSIT. We denote the DoF region of the k -user Gaussian IC with global delayed CSIT by \mathcal{D}_k .

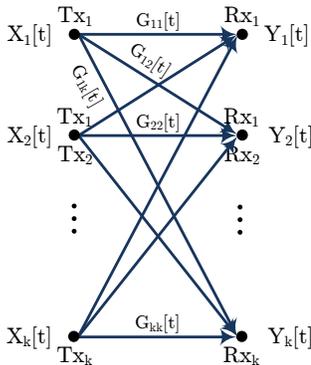


Fig. 10. k -user Erasure Interference Channel. The capacity region with global delayed CSIT is open.

The intuition for the two-user erasure IC was that it is the responsibility of the transmitter who creates interference to resolve it. Characterizing \mathcal{D}_k or the capacity region of the k -user erasure IC with global delayed CSIT are still open. However, there are several results that try to exploit the delayed knowledge of the channel state information for the achievability purposes in the context of k -user Gaussian IC (e.g., see [19], [23] and references therein). In [20], authors have shown that such gains can be also obtained if each transmitter is only aware of the channel gains of the outgoing links from itself with delay. This result matches our intuition for the two-user erasure IC. However, in the lack of a tight outer-bound, a firm conclusion cannot be made.

VII. CONCLUSION AND FUTURE DIRECTIONS

We studied the capacity region of the two-user Binary Fading Interference Channel with *local* delayed channel state information at the transmitters. We showed that in order to achieve the performance of *global* delayed CSIT, it suffices that each transmitter has only access to the delayed knowledge of its *outgoing* links. We also identified the cases in which local delayed CSIT does not provide any gain over the no knowledge assumption. Fig. 11, summarizes our main results.

As discussed in Section VI, an interesting future direction is to extend the result to the k -user Binary Fading Interference Channel and see whether the delayed knowledge of the outgoing links suffices to achieve the capacity with global delayed CSIT. This result, if true, would shed light on finally solving the capacity region (or DoF region) of k -user interference channels with delayed CSIT. Another direction, would be to extend the current results to two-user Rayleigh

fading interference channels (as opposed to the binary fading model). Recently, a new direction was introduced in [24] where spatial correlation between channels were considered. Implications of local delayed CSIT in that setting has great practical importance.

APPENDIX A MORE DISCUSSION ON THEOREM 1

Define

$$\mathcal{E}_{\mathbf{W}_1} \triangleq \left\{ (W_2, G^m) \text{ s.t. } \widehat{\mathbf{W}}_1 \neq \mathbf{W}_1 \right\}, \quad (52)$$

and

$$\begin{aligned} \mathcal{E}_{\mathbf{W}_1}^1 &\triangleq \left\{ (X_1^n, G_{11}^n, G_{21}^n) \text{ s.t. } \widehat{\mathbf{W}}_1 \neq \mathbf{W}_1 \right\}, \\ \mathcal{E}_{\mathbf{W}_1}^2 &\triangleq \left\{ (X_2^n, G_{12}^n, G_{22}^n) \text{ s.t. } \widehat{\mathbf{W}}_1 \neq \mathbf{W}_1 \right\}. \end{aligned} \quad (53)$$

Then, we have

$$\begin{aligned} \lambda_{1,n} &= \sum_{w_1} \Pr(\mathbf{W}_1 = w_1) \Pr(\mathcal{E}_{\mathbf{W}_1 = w_1}) \\ &\stackrel{(25)}{=} \sum_{w_1} \Pr(\mathbf{W}_1 = w_1) \left[\Pr(\mathcal{E}_{\mathbf{W}_1 = w_1}^1) + \Pr(\mathcal{E}_{\mathbf{W}_1 = w_1}^2) \right] \\ &= \sum_{w_1} \Pr(\mathbf{W}_1 = w_1) \left[\Pr(\tilde{\mathcal{E}}_{\mathbf{W}_1 = w_1}^1) + \Pr(\tilde{\mathcal{E}}_{\mathbf{W}_1 = w_1}^2) \right] \\ &\geq \sum_{w_1} \Pr(\mathbf{W}_1 = w_1) \Pr(\tilde{\mathcal{E}}_{\mathbf{W}_1 = w_1}) \\ &= \tilde{\lambda}_{1,n}, \end{aligned} \quad (54)$$

where $\tilde{\mathcal{E}}_{\mathbf{W}_1}^1$, $\tilde{\mathcal{E}}_{\mathbf{W}_1}^2$, and $\tilde{\lambda}_{1,n}$ are respectively the counterparts of $\mathcal{E}_{\mathbf{W}_1}$, $\mathcal{E}_{\mathbf{W}_1}^j$, and $\lambda_{1,n}$ for the new channel, $j = 1, 2$.

Similarly, we can start with $\tilde{\lambda}_{1,n}$ and use the fact that

$$\begin{aligned} \Pr(\tilde{\mathcal{E}}_{\mathbf{W}_1 = w_1}) &\leq \Pr(\tilde{\mathcal{E}}_{\mathbf{W}_1 = w_1}^1) + \Pr(\tilde{\mathcal{E}}_{\mathbf{W}_1 = w_1}^2) \\ &\leq 2 \Pr(\tilde{\mathcal{E}}_{\mathbf{W}_1 = w_1}), \end{aligned} \quad (55)$$

to show that

$$\lambda_{1,n} \rightarrow 0 \Leftrightarrow \tilde{\lambda}_{1,n} \rightarrow 0. \quad (56)$$

Therefore, we conclude that

$$\mathcal{C}(\text{V.4}) \equiv \tilde{\mathcal{C}}(\text{V.4}). \quad (57)$$

APPENDIX B TRANSMISSION STRATEGY FOR THE CORNER POINT (0.375, 0.5)

We now provide the achievability strategy for the corner point

$$(R_1, R_2) = \left(\frac{3}{8}, \frac{1}{2} \right). \quad (58)$$

To achieve this corner point, new challenges arise which are due to the asymmetry of the rates. In this case, Tx_2 (the primary user) communicates at the full rate of 0.5 while Tx_1 (the secondary user) communicates at a lower rate and tries to coexist with the primary user. In fact Tx_1 has to take more responsibility in dealing with interference at *both* receivers.

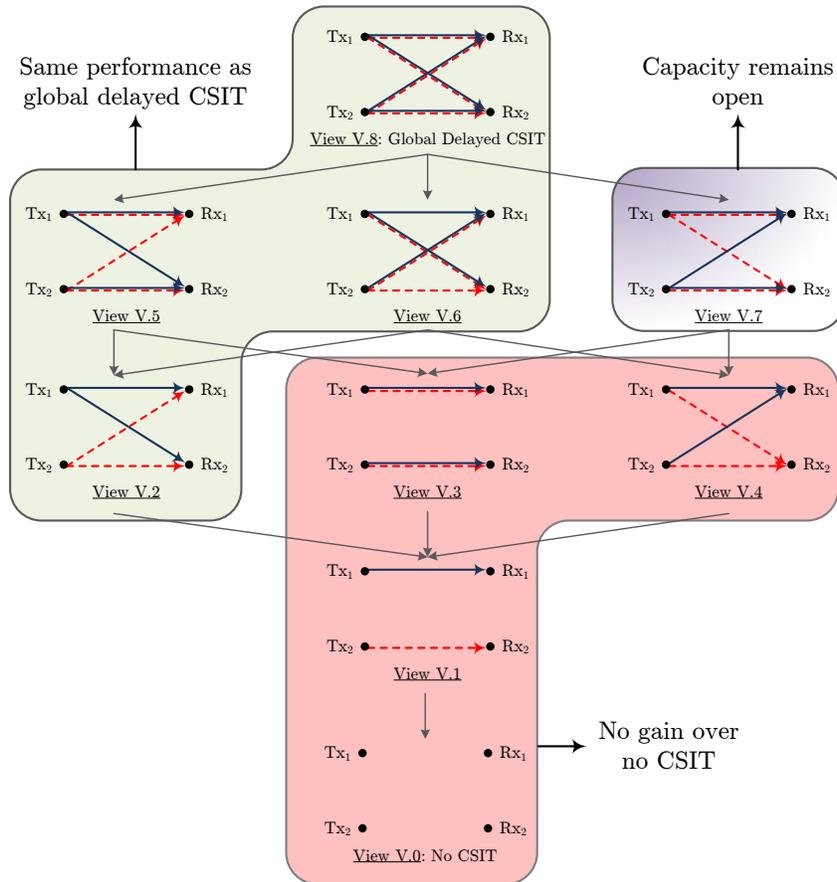


Fig. 11. Summary of the main results. Capacity region with local delayed CSIT of Views V.1, V.3, and V.4 coincides with no CSIT case (V.0); while capacity region with local delayed CSIT of Views V.2, V.5, and V.6 matches that of global delayed CSIT (V.8). The capacity region with local delayed CSIT of View V.7 remains open.

The proposed transmission strategy consists of four phases as described below. We assume that Tx_1 has $\frac{3}{4}m$ bits to communicate while Tx_2 has m bits. We show that all bits can be delivered in $2m + O(m^{2/3})$ time instants with vanishing error probability as $m \rightarrow \infty$. This immediately implies the achievability of the corner point $(0.375, 0.5)$.

Phase 1: This phase is similar to Phase 1 of the achievability of the optimal sum-rate point $(0.45, 0.45)$. The main difference is due to the fact that the transmitters have unequal number of bits at the start. In Phase 1, Tx_1 (the secondary user) transmits all its initial bits while Tx_2 (the primary user) only transmits a fraction of its initial bits. Transmitter two postpones the transmission of its remaining bits to Phase 2.

At the beginning of the communication block, we assume that each transmitter has $\frac{3}{4}m$ bits in queue $Q_{i \rightarrow i}$ (the initial state of the bits), $i = 1, 2$. At each time instant t , Tx_i sends out a bit from $Q_{i \rightarrow i}$, and this bit will either stay in the initial queue or a transition to one of the following possible queues will take place according to the description in Table I. If at time instant t , $Q_{i \rightarrow i}$ is empty, then Tx_i , $i = 1, 2$, remains silent until the end of Phase 1.

Phase 1 continues for

$$m + m^{\frac{2}{3}} \quad (59)$$

time instants, and if at the end of this phase, either of the queues $Q_{i \rightarrow i}$ is not empty, we declare error type-I and halt the transmission.

Assuming that the transmission is not halted, let $N_{i,1}$ and $N_{i,2}$, $i = 1, 2$, denote the number of bits in queues $Q_{i,1}$ and $Q_{i,2}$ respectively at the end of Phase 1. The transmission strategy will be halted and an error type-II will occur, if any of the following events happens.

$$\begin{aligned} N_{i,1} &> \mathbb{E}[N_{i,1}] + 2m^{\frac{2}{3}} \triangleq n_{i,1}, \quad i = 1, 2; \\ N_{i,2} &> \mathbb{E}[N_{i,2}] + 2m^{\frac{2}{3}} \triangleq n_{i,2}, \quad i = 1, 2. \end{aligned} \quad (60)$$

From basic probability, we have

$$\mathbb{E}[N_{i,1}] = \mathbb{E}[N_{i,2}] = \frac{m}{4}, \quad (61)$$

so that

$$n_{i,1} = n_{i,2} = \frac{m}{4} + 2m^{\frac{2}{3}}. \quad (62)$$

At the end of Phase 1, for $i = 1, 2$, we add 0's (if necessary) in order to make queues $Q_{i,1}$ and $Q_{i,2}$ of size equal to $n_{i,1}$ and $n_{i,2}$ respectively.

Since channel gains are distributed independently, statistically half of the bits in $Q_{i,1}$ and half of the bits in $Q_{i,2}$ are

known to $R_{x_{\bar{i}}}$, $i = 1, 2$. Denote the number of bits in $Q_{i,j}$ known to $R_{x_{\bar{i}}}$ by

$$N_{i,j|R_{x_{\bar{i}}}}, \quad i, j \in \{1, 2\}. \quad (63)$$

At the end of communication, if we have

$$N_{i,j|R_{x_{\bar{i}}}} < \frac{1}{2}n_{i,j} - m^{\frac{2}{3}}, \quad i, j \in \{1, 2\}, \quad (64)$$

we declare error type-III.

Moreover, we note that statistically for every two bits in $Q_{\bar{i},1}$, a bit in $Q_{i,1}$ was transmitted simultaneously with one of them. Denote the number of bits in $Q_{i,1}$ that were transmitted simultaneously with a bit in $Q_{\bar{i},1}$ by

$$N_{i \rightarrow \bar{i},1}, \quad i = 1, 2. \quad (65)$$

At the end of communication, if we have

$$N_{i \rightarrow \bar{i},1} < \frac{1}{2}n_{i,1} - m^{\frac{2}{3}}, \quad i = 1, 2, \quad (66)$$

we declare error type-IV. Note that transmitters cannot detect error type-III or error type-IV, but receivers have sufficient information to do so.

Using the Bernstein inequality, we can show that the probability of errors of types I, II, III, and IV decreases exponentially with m . For the rest of this subsection, we assume that Phase 1 is completed and no error has occurred.

Phase 2 [transmission of new bits vs interference management]: In this phase, the primary user T_{x_2} transmits its remaining initial bits while the secondary user T_{x_1} tries to resolve as much interference as it can and deliver some of its bits in $Q_{1,1}$. To do so, the secondary user sends some of its bits in $Q_{1,1}$ at a rate low enough such that both receivers can decode and remove them regardless of what the primary transmitter does. Note that $1/4$ of the time, each receiver obtains an interference-free signal from the secondary transmitter, hence, the secondary transmitter can take advantage of these time instants to deliver its bits during Phase 2.

Transmitter T_{x_1} creates a matrix $C_{1,1}$ of size $(\frac{1}{12}m + m^{2/3}) \times (\frac{m}{4} + 2m^{\frac{2}{3}})$, where the entries of this matrix are drawn from i.i.d. $\mathcal{B}(0.5)$ distribution. We assume that this matrix is generated prior to communication and is shared with receivers. Then T_{x_1} creates $(\frac{1}{12}m + m^{2/3})$ bits by multiplying matrix $C_{1,1}$ and the bits in $Q_{1,1}$. Using point-to-point erasure code of rate $1/4$, transmitter T_{x_1} encodes bits $C_{1,1}Q_{1,1}$ and communicates them during Phase 2. We note that due to the chosen rate (*i.e.* $1/4$) as $m \rightarrow \infty$, each receiver can decode bits $C_{1,1}Q_{1,1}$ with vanishing error probability.

Transmitter T_{x_2} places its remaining $\frac{1}{4}m$ bits in queue $Q_{2 \rightarrow 2}$ (the initial state of the bits). At each time instant t of Phase 2, T_{x_2} sends out a bit from $Q_{2 \rightarrow 2}$, and this bit will either stay in the initial queue or a transition to a new queue will take place according to the description in Table II. Note that here, since the signal of T_{x_1} can be decoded first, we simply consider the bits of T_{x_2} that were transmitted in case 1 (see Table II) to be delivered. At the end of Phase 2, we update the value of $n_{2,2}$ as

$$n_{2,2} = \frac{m}{3} + 3m^{\frac{2}{3}}. \quad (67)$$

Phase 3 [encoding and mixing interfering bits]: Transmitter T_{x_1} creates two matrices: $C_{1,2}$ of size $(\frac{m}{6} + 2m^{\frac{2}{3}}) \times (\frac{m}{4} + 2m^{\frac{2}{3}})$ and $C_{1,3}$ of size $(\frac{m}{4} + 2m^{\frac{2}{3}}) \times (\frac{m}{4} + 2m^{\frac{2}{3}})$, where entries to each matrix are drawn from i.i.d. $\mathcal{B}(0.5)$ distribution. We assume that the matrices are generated prior to communication and are shared with receivers.

Transmitter T_{x_1} creates

$$\begin{aligned} \tilde{Q}_{1,1} &= C_{1,2}Q_{1,1}, \\ \tilde{Q}_{1,2} &= C_{1,3}Q_{1,2}. \end{aligned} \quad (68)$$

Then T_{x_1} encodes bits in $\tilde{Q}_{1,1}$ using a point-to-point erasure code of rate $1/4$ denoted by $\hat{Q}_{1,1}$ and encodes bits in $\tilde{Q}_{1,2}$ using a point-to-point erasure code of rate $1/2$ denoted by $\hat{Q}_{1,2}$. Transmitter T_{x_1} communicates² $\hat{Q}_{1,1} \oplus \hat{Q}_{1,2}$ during Phase 3.

At the same time, transmitter T_{x_2} creates two matrices: $C_{2,1}$ of size $(\frac{m}{8} + 2m^{\frac{2}{3}}) \times (\frac{m}{4} + 2m^{\frac{2}{3}})$ and $C_{2,2}$ of size $(\frac{m}{3} + 2m^{\frac{2}{3}}) \times (\frac{m}{3} + 2m^{\frac{2}{3}})$, where entries to each matrix are drawn from i.i.d. $\mathcal{B}(0.5)$ distribution. We assume that the matrices are generated prior to communication and are shared with receivers.

Transmitter T_{x_2} creates

$$\begin{aligned} \tilde{Q}_{2,1} &= C_{2,1}Q_{2,1}, \\ \tilde{Q}_{2,2} &= C_{2,2}Q_{2,2}. \end{aligned} \quad (69)$$

Then T_{x_2} encodes bits in $\tilde{Q}_{2,1}$ using a point-to-point erasure code of rate $1/4$ denoted by $\hat{Q}_{2,1}$ and encodes bits in $\tilde{Q}_{2,2}$ using a point-to-point erasure code of rate $1/2$ denoted by $\hat{Q}_{2,2}$. Transmitter T_{x_2} communicates³ $\hat{Q}_{2,1} \oplus \hat{Q}_{2,2}$ during Phase 3.

Decoding: Upon completion of the third phase, we show that each receiver has gathered enough linear equations to decode all bits in $Q_{1,1}$, $Q_{1,2}$ and $Q_{2,2}$. Receiver R_{x_i} first removes the known bits from $Q_{\bar{i},1}$ and $Q_{\bar{i},2}$, $i = 1, 2$.

Then, each receiver has

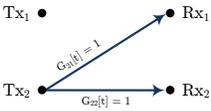
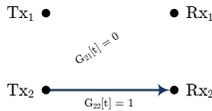
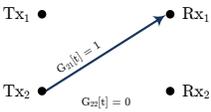
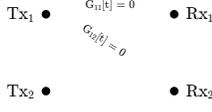
$$\underbrace{\left(\frac{m}{12} + 2m^{\frac{2}{3}}\right)}_{C_{1,1}Q_{1,1}} + \underbrace{\left(\frac{m}{6} + 2m^{\frac{2}{3}}\right)}_{C_{1,2}Q_{1,1}} = \frac{m}{4} + 2m^{\frac{2}{3}}$$

randomly generated equations of $\frac{m}{4} + 2m^{\frac{2}{3}}$ bits in $Q_{1,1}$. Thus, both receivers can recover the bits in $Q_{1,1}$ with vanishing error probability as $m \rightarrow \infty$. Similarly, receivers have sufficient information to recover bits in $Q_{1,2}$ and $Q_{2,2}$.

As opposed to other states, not all bits in $Q_{2,1}$ are provided to the receivers by T_{x_2} . However, transmitter T_{x_2} is not required to provide all bits in $Q_{2,1}$ to both receivers. The reason is that, once $Q_{1,1}$ is known at the receivers, statistically half of the bits in $Q_{2,1}$ can be reconstructed at each receiver. Therefore, transmitter T_{x_2} needs to provide half of the bits in $Q_{2,1}$ to the receivers. That is why in Phase 3, we chose $C_{2,1}$ to have size $(\frac{m}{8} + 2m^{\frac{2}{3}}) \times (\frac{m}{4} + 2m^{\frac{2}{3}})$ in lieu of $(\frac{m}{4} + 2m^{\frac{2}{3}}) \times (\frac{m}{4} + 2m^{\frac{2}{3}})$.

^{2,3} The two sequences are not of equal length, we can simply add deterministic number of zeros to $Q_{1,2}$ to make the two sequences of equal length

TABLE II
TRANSITION OF THE BITS FROM $Q_{2 \rightarrow 2}$ DURING PHASE 2.

case ID	channel realization at time instant n	state transition	case ID	channel realization at time instant n	state transition
1		$a \rightarrow Q_{1,1}$	3		$a \rightarrow Q_{1 \rightarrow F}$
2		$a \rightarrow Q_{1,2}$	4		$a \rightarrow Q_{1 \rightarrow 1}$

We therefore conclude that each receiver can recover its intended bits with vanishing error probability as $m \rightarrow \infty$ in a total of

$$2m + O\left(m^{2/3}\right) \quad (70)$$

time instants.

Hence, if we let $m \rightarrow \infty$, the decoding error probability goes to zero exponentially, and we achieve rate-tuple

$$(R_1, R_2) = \left(\frac{3}{8}, \frac{1}{2}\right). \quad (71)$$

Similarly, we can achieve the corner point

$$(R_1, R_2) = \left(\frac{1}{2}, \frac{3}{8}\right). \quad (72)$$

Together with the results of Section V, we conclude that \mathcal{C} (V.8) is achievable with local delayed CSIT of View V.2.

APPENDIX C

TRANSMISSION STRATEGY FOR $0 \leq p \leq 1$

The reason we considered $p = 0.5$ in Section V was to remain focused on the impact of local delayed CSIT on the capacity region rather than getting involved in the details of the transmission strategy.

Here, we describe the changes needed in the transmission strategy when considering $0 \leq p \leq 1$. The core structure of the achievability strategy remains the same as what we discussed in Section V. However, modifications are needed to ensure optimal performance.

Consider the maximum symmetric sum-rate point as given by

$$R_1 = R_2 = \min \left\{ p, \frac{\beta(1-q^2)}{1+\beta} \right\}, \quad (73)$$

where

$$\beta = 2 - p. \quad (74)$$

The achievability strategy for the other corner points, *i.e.*

$$(R_i, R_{\bar{i}}) = (\min\{p, pq(1+q)\}, p), \quad i = 1, 2, \quad (75)$$

follows similar modifications when compared to the strategy given for $p = 0.5$ in Appendix B. We note that the capacity region is the convex hull of the aforementioned corner points.

Suppose each transmitter wishes to communicate m bits to its intended receiver. We need to show that this task can be accomplished (with vanishing error probability as $m \rightarrow \infty$) in

$$\max \left\{ \frac{1}{p}, \frac{1+\beta}{\beta(1-q^2)} \right\} m + \mathcal{O}\left(m^{\frac{2}{3}}\right) \quad (76)$$

time instants.

In Section V, we considered $p = 0.5$ and that implies

$$\mathbb{E}[N_{i,1}] = \mathbb{E}[N_{i,2}]. \quad (77)$$

However, when $p \neq 0.5$ the above inequality no longer holds. Below, we describe other coding opportunities that were not needed in Section V.

1) Suppose $p > 0.5$, then we have $\mathbb{E}[N_{i,1}] > \mathbb{E}[N_{i,2}]$. After combining the bits in $Q_{i,1}$ and $Q_{i,2}$, one naive solution would be to treat the remaining bits in $Q_{i,1}$ as bits of common interest. However, we can improve upon that scheme as described below.

Suppose each one of the transmitters sends three data bits as depicted in Fig. 12. We observe that providing $a_9 \oplus a_{10}$ and $b_9 \oplus b_{11}$ to both receivers is sufficient to decode the bits. For instance, if R_{x_1} is provided with $a_9 \oplus a_{10}$ and $b_9 \oplus b_{11}$, then it will use a_{10} to decode a_9 , from which it can obtain b_9 ; then using b_9 and $b_9 \oplus b_{11}$, it gains access to b_{11} ; finally using b_{11} , it can decode a_{11} from $a_{11} \oplus b_{11}$. Thus, linear combination $a_9 \oplus a_{10}$ available at T_{x_1} , and linear combination $b_9 \oplus b_{11}$ available at T_{x_2} , are bits of common interest and can be transmitted to both receivers simultaneously in the efficient two-multicast problem. We note that the bits of T_{x_1} in Fig. 12(a) and Fig. 12(b) fall in $Q_{1,1}$ and the bits of T_{x_2} in Fig. 12(a) and Fig. 12(c) fall in $Q_{2,1}$. Thus, we can further combine the bits in $Q_{i,1}$ to improve the achievable rate region.

2) Suppose $p > 0.5$, then we have $\mathbb{E}[N_{i,1}] > \mathbb{E}[N_{i,2}]$. After combining the bits in $Q_{i,1}$ and $Q_{i,2}$, sufficient number

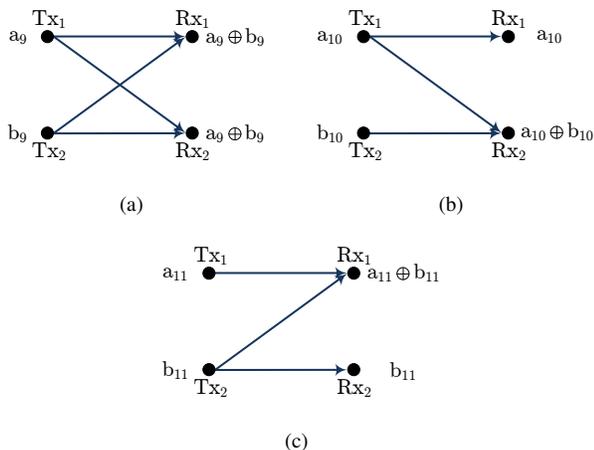


Fig. 12. Providing $a_9 \oplus a_{10}$ and $b_9 \oplus b_{11}$ to both receivers is sufficient to decode the bits.

of linear combinations of the remaining bits in $Q_{i,2}$ should be created and communicated using a point-to-point erasure code of rate p .

It is important to realize that the coding opportunities described above can be identified using the local delayed CSIT. What remains is to pick the right number of linear combinations to communicate in each step of the transmission strategy. It is a straightforward exercise to use the results of [8] to deduce the right number of linear combinations needed in each step of our strategy with local delayed CSIT.

REFERENCES

- [1] A. Vahid and R. Calderbank, "Impact of local delayed CSIT on the capacity region of the two-user interference channel," *submitted to International Symposium on Information Theory (ISIT)*, 2015.
- [2] R. Ahlswede, "The capacity region of a channel with two senders and two receivers," *The Annals of Probability*, pp. 805–814, 1974.
- [3] H. Sato, "Two-user communication channels," *IEEE Transactions on Information Theory*, vol. 23, no. 3, pp. 295–304, 1977.
- [4] T. Han and K. Kobayashi, "A new achievable rate region for the interference channel," *IEEE Transactions on Information Theory*, vol. 27, pp. 49–60, Jan. 1981.
- [5] R. H. Etkin, D. N. Tse, and H. Wang, "Gaussian interference channel capacity to within one bit," *IEEE Transactions on Information Theory*, vol. 54, no. 12, pp. 5534–5562, 2008.
- [6] A. Vahid, M. Maddah-Ali, and A. 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.
- [7] A. Vahid, M. Maddah-Ali, and A. Avestimehr, "Binary fading interference channel with delayed feedback," in *Proceedings of IEEE International Symposium on Information Theory (ISIT)*, pp. 1882–1886, 2012.
- [8] 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.
- [9] S.-W. Jeon and S.-Y. Chung, "Capacity of a class of linear binary field multisource relay networks," *IEEE Transactions on Information Theory*, vol. 59, no. 10, pp. 6405–6420, 2013.
- [10] A. Vahid, M. A. Maddah-Ali, and A. S. Avestimehr, "Communication through collisions: Opportunistic utilization of past receptions," in *IEEE Conference on Computer Communications (INFOCOM)*, pp. 2553–2561, IEEE, 2014.
- [11] D. Kao and A. Sabharwal, "Two-user interference channels with local views: On capacity regions of TDM-dominating policies," *IEEE Transactions on Information Theory*, vol. 59, no. 11, pp. 7014–7040, 2013.
- [12] K. Jolfaei, S. Martin, and J. Matfeldt, "A new efficient selective repeat protocol for point-to-multipoint communication," in *IEEE International Conference on Communications (ICC'93)*, vol. 2, pp. 1113–1117, IEEE, 1993.
- [13] L. Georgiadis and L. Tassiulas, "Broadcast erasure channel with feedback-capacity and algorithms," in *Workshop on Network Coding, Theory, and Applications (NetCod'09)*, pp. 54–61, IEEE, 2009.
- [14] M. A. Maddah-Ali and D. Tse, "Completely stale transmitter channel state information is still very useful," *IEEE Transactions on Information Theory*, vol. 58, no. 7, pp. 4418–4431, 2012.
- [15] A. Vahid, M. A. Maddah-Ali, and A. S. Avestimehr, "Approximate capacity of the two-user MISO broadcast channel with delayed CSIT," in *51st Annual Allerton Conference on Communication, Control, and Computing*, pp. 1136–1143, IEEE, 2013.
- [16] A. Vahid, M. A. Maddah-Ali, and A. S. Avestimehr, "Approximate capacity region of the miso broadcast channels with delayed csit," *IEEE Transactions on Communications*, vol. 64, no. 7, pp. 2913–2924, 2016.
- [17] A. Ghasemi, A. S. Motahari, and A. K. Khandani, "On the degrees of freedom of X channel with delayed CSIT," in *2011 IEEE International Symposium on Information Theory Proceedings*, (Saint-Petersburg, Russia), pp. 909–912, July 2011.
- [18] C. S. Vaze and M. K. Varanasi, "The degrees of freedom region of the two-user MIMO broadcast channel with delayed CSI," Dec. 2010. arxiv.org/abs/1101.0306.
- [19] H. Maleki, S. Jafar, and S. Shamai, "Retrospective interference alignment," *ArXiv preprint arXiv:1009.3593*, 2010.
- [20] A. Ghasemi, A. S. Motahari, and A. K. Khandani, "Interference alignment for the MIMO interference channel with delayed local CSIT," *arXiv preprint arXiv:1102.5673*, 2011.
- [21] A. Vahid, M. A. Maddah-Ali, A. S. Avestimehr, and Y. Zhu, "Binary fading interference channel with no csit," *arXiv preprint arXiv:1405.0203*, 2014.
- [22] J. Bourgain, V. H. Vu, and P. M. Wood, "On the singularity probability of discrete random matrices," *Journal of Functional Analysis*, vol. 258, no. 2, pp. 559–603, 2010.
- [23] M. Abdoli, A. Ghasemi, and A. Khandani, "On the degrees of freedom of k -user SISO interference and X channels with delayed CSIT," *arXiv preprint arXiv:1109.4314*, 2011.
- [24] A. Vahid and R. Calderbank, "When does spatial correlation add value to delayed channel state information?," *arXiv preprint arXiv:1606.07036*, 2016.



Alireza Vahid received the B.Sc. degree in electrical engineering from Sharif University of Technology, Tehran, Iran, in 2009, and the M.Sc. degree and Ph.D. degree in electrical and computer engineering both from Cornell University, Ithaca, NY, in 2012 and 2015 respectively. As of September 2014, he is a postdoctoral scholar at Information Initiative at Duke University, Durham, NC. His research interests include network information theory, wireless communications, statistics and machine learning.

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.



Robert Calderbank (M89 SM97 F98) received the BSc degree in 1975 from Warwick University, England, the MSc degree in 1976 from Oxford University, England, and the PhD degree in 1980 from the California Institute of Technology, all in mathematics.

Dr. Calderbank is Professor of Electrical Engineering at Duke University where he now directs the Information Initiative at Duke (*iID*) after serving as Dean of Natural Sciences (2010-2013). Dr. Calderbank was previously Professor of Electrical

Engineering and Mathematics at Princeton University where he directed the Program in Applied and Computational Mathematics. Prior to joining Princeton in 2004, he was Vice President for Research at AT&T, responsible for directing the first industrial research lab in the world where the primary focus is data at scale. At the start of his career at Bell Labs, innovations by Dr. Calderbank were incorporated in a progression of voiceband modem standards that moved communications practice close to the Shannon limit. Together with Peter Shor and colleagues at AT&T Labs he showed that good quantum error correcting codes exist and developed the group theoretic framework for quantum error correction. He is a co-inventor of space-time codes for wireless communication, where correlation of signals across different transmit antennas is the key to reliable transmission.

Dr. Calderbank served as Editor in Chief of the IEEE TRANSACTIONS ON INFORMATION THEORY from 1995 to 1998, and as Associate Editor for Coding Techniques from 1986 to 1989. He was a member of the Board of Governors of the IEEE Information Theory Society from 1991 to 1996 and from 2006 to 2008. Dr. Calderbank was honored by the IEEE Information Theory Prize Paper Award in 1995 for his work on the Z_4 linearity of Kerdock and Preparata Codes (joint with A.R. Hammons Jr., P.V. Kumar, N.J.A. Sloane, and P. Sole), and again in 1999 for the invention of space-time codes (joint with V. Tarokh and N. Seshadri). He has received the 2006 IEEE Donald G. Fink Prize Paper Award, the IEEE Millennium Medal, the 2013 IEEE Richard W. Hamming Medal, the 2015 Shannon Award, and he was elected to the US National Academy of Engineering in 2005.