The Value of Local Delayed CSIT
Alireza Vahid, Robert Calderbank

Abstract

We study the capacity region of the two-user interference channel 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 In our model, transmitters have local mismatched outdated knowledge of the channel gains.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 work [6] addresses the capacity region of the canonical two-user fading IC under a specific channel distribution. This paper considers the two-user Binary Fading Interference Channel (BFIC) 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

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

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 \( \mathbb{F}_2 \).

Fig. 1. Two-user Binary Fading Interference Channel (BFIC).
In [5], the capacity region of the two-user BFIC 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. 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. Our result provides a better intuition and a deeper understanding of the coding opportunities that arise from delayed CSIT. Then, 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.

There is some prior work assessing the value of delayed CSIT. It was used in [7] 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 [8], the DoF region of broadcast channels [9], and the DoF region of multi-antenna multi-user Gaussian ICs and X channels [10]–[12].

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 devote to the proof of the main results. We present the key idea for deriving the outer-bound with global delayed CSIT in Section VI. Next, in Section VII we discuss the implications of our results for more general settings. Section VIII concludes the paper.

II. Problem Setting

To study the capacity region of the two-user fading interference channels with local delayed CSIT, we consider a binary fading model. In the binary fading model, the channel gain from transmitter $\text{Tx}_i$ to receiver $\text{Rx}_j$ 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 define the channel state information at time instant $t$ to be the set

$$G[t] \triangleq \{G_{11}[t], G_{12}[t], G_{21}[t], G_{22}[t]\}.$$  \hspace{1cm} (2)

Keeping in mind that our goal is to understand the ramifications of local delayed CSIT on the capacity region; in this work, we focus on a homogeneous setting for simplicity of notations where

$$G_{ij}[t] \overset{d}{\sim} \mathcal{B}(0.5), \quad i, j = 1, 2.$$  \hspace{1cm} (3)

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

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

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

$$(k, \ell) \in \mathcal{S}_{\text{Tx}_i} \iff (\bar{k}, \bar{\ell}) \in \mathcal{S}_{\text{Tx}_2}, \quad k, \ell = 1, 2.$$  \hspace{1cm} (5)

In [6], the capacity region of the two-user BFIC 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. 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. Our result provides a better intuition and a deeper understanding of the coding opportunities that arise from delayed CSIT. Then, 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.

There is some prior work assessing the value of delayed CSIT. It was used in [7] 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 [8], the DoF region of broadcast channels [9], and the DoF region of multi-antenna multi-user Gaussian ICs and X channels [10]–[12].

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 devote to the proof of the main results. We present the key idea for deriving the outer-bound with global delayed CSIT in Section VI. Next, in Section VII we discuss the implications of our results for more general settings. Section VIII concludes the paper.
for $k \triangleq 3 - k$ and $\ell \triangleq 3 - \ell$.

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
  \[ S_{T_x_i} = \emptyset, \quad i = 1, 2. \]  

- **View V.1**: In this case each transmitter is aware of channel value to its corresponding receiver with unit delay, i.e.
  \[ S_{T_x_i} = \{(i, i)\}, \quad i = 1, 2. \]  

- **View V.2**: In this case each transmitter is aware of channel value of the outgoing links with unit delay, i.e.
  \[ S_{T_x_i} = \{(i, i), (i, \bar{i})\}, \quad i = 1, 2. \]  

- **View V.3**: In this case each transmitter is aware of channel value of the direct links with unit delay, i.e.
  \[ S_{T_x_i} = \{(i, i), (\bar{i}, \bar{i})\}, \quad i = 1, 2. \]  

- **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.
  \[ S_{T_x_i} = \{(i, i), (i, \bar{i}), (\bar{i}, \bar{i})\}, \quad i = 1, 2. \]  

- **View V.5**: In this case we have
  \[ S_{T_x_i} = \{(i, i), (i, \bar{i}), (\bar{i}, \bar{i})\}, \quad i = 1, 2. \]
• View V.6: In this case we have
  \[ S_{Tx_i} = \{(i,i), (i,\bar{i}), (\bar{i},i)\}, \quad i = 1, 2. \]  

• View V.7: In this case we have
  \[ S_{Tx_i} = \{(i,\bar{i}), (\bar{i},i), (\bar{i},\bar{i})\}, \quad i = 1, 2. \]  

• 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
  \[ S_{Tx_i} = \{(1,1), (1,2), (2,1), (2,2)\}, \quad i = 1, 2. \]  

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 a subset of V.5; and so on. This hierarchical structure is also shown in Fig. 2 using downward arrows.4

We consider the scenario in which Tx\textsubscript{i} wishes to reliably communicate message W\textsubscript{i} ∈ \{1, 2, \ldots, 2^nR_i\} to Rx\textsubscript{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\textsubscript{i} be encoded as X^n\textsubscript{i} at transmitter Tx\textsubscript{i} using the encoding function f_i(W_i, S_{Tx_i}), which depends on the available channel state information at the transmitter. Receiver Rx\textsubscript{i} is only interested in decoding W\textsubscript{i}, and it will decode the message using the decoding function \hat{W}_i = g_i(Y^n_i, G^n). An error occurs when \hat{W}_i \neq W_i. The average probability of decoding error is given by
  \[ \lambda_{i,n} = E[P(\hat{W}_i \neq W_i)], \quad i = 1, 2, \]  
and the expectation is taken with respect to the random choice of the transmitted messages W\textsubscript{1} and W\textsubscript{2}.

A rate-tuple (R\textsubscript{1}, R\textsubscript{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 S_{Tx_1} and S_{Tx_2}. The capacity region for View V.j, i.e
  \[ C(V.j), \quad j = 1, 2, \ldots, 8, \]  
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 the two-user Binary Fading Interference Channel. We establish the capacity region of the two-user BFIC 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?

By answering these questions, we have essentially identified the “MSBs” of the delayed CSIT in the two-user BFIC.

A. Benchmarks

Our base-line is the no CSIT scenario. In other words, the only available knowledge at the transmitters is the distribution 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, C(V.0), is the set of all rate-tuples (R\textsubscript{1}, R\textsubscript{2}) satisfying
  \[
  \begin{aligned}
  0 &\leq R_i \leq 0.5, \quad i = 1, 2, \\
  R_1 + R_2 &\leq 0.75.
  \end{aligned}
  \]  

A similar set of local views for the channel state information was studied in [13] in the context of two-user Gaussian interference channel (not fading) to identify the views in which one could outperform TDMA.
The other extreme point is the global delayed CSIT model. From [6], we know that the capacity region of the two-user Binary Fading IC with global delayed CSIT, $C(V.8)$, is the set of all rate-tuples $(R_1, R_2)$ satisfying

$$
\begin{align*}
0 &\leq R_i \leq 0.5, \quad i = 1, 2, \\
R_i + \frac{3}{2}R_i &\leq \frac{9}{8}, \quad i = 1, 2.
\end{align*}
$$

(18)

These benchmarks are depicted in Fig. 3. We note that $C(V.0)$ is a strict subset of $C(V.8)$, 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.](image)

**B. Main Results**

Here, we provide answers to the questions posed above. First, 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.

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

(19)

**Remark 1:** 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.

**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.

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

(20)

**Remark 2:** 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 $C(V.8)$ was established in [6] and a novel transmission strategy was introduced. The strategy in [6] 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 3:** 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 VII.

In the remaining of the paper, we provide the proof of Theorem 1 and Theorem 2. In Section VII, 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

\[ \Pr (X_1^n, X_2^n | G^n) = \Pr (X_1^n | G^n) \Pr (X_2^n | G^n). \] (21)

Moreover, if \( \mathcal{S}_{\text{Tx}_1} \cap \mathcal{S}_{\text{Tx}_2} = \emptyset \), then we have

\[ \Pr (X_1^n, X_2^n) = \Pr (X_1^n) \Pr (X_2^n). \] (22)

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

A. Proof for View V.4

Consider the two-user BFIC with local delayed CSIT according to View V.4. We have

\[ \mathcal{S}_{\text{Tx}_1} = \{(1, 1), (2, 1)\} \quad \text{and} \quad \mathcal{S}_{\text{Tx}_2} = \{(1, 2), (2, 2)\}. \] (23)

Thus writing the marginal distribution at receiver \( \text{Rx}_1 \), we get

\[
\Pr (Y_1^n, G^n | X_1^n, X_2^n) = (a) \Pr (G^n | X_1^n, X_2^n) \Pr (Y_1^n | X_1^n, X_2^n, G^n) = \frac{\Pr (G^n, X_1^n, X_2^n)}{\Pr (X_1^n, X_2^n)} \Pr (Y_1^n | X_1^n, X_2^n, G^n) = (b) \frac{\Pr (G^n) \Pr (X_1^n | G^n) \Pr (X_2^n | G^n)}{\Pr (X_1^n, X_2^n) \Pr (X_2^n)} \Pr (Y_1^n | X_1^n, X_2^n, G^n) = (c) \frac{\Pr (G^n) \Pr (X_1^n | G_{11}^n, G_{21}^n) \Pr (X_2^n | G_{12}^n, G_{22}^n)}{\Pr (X_1^n, X_2^n) \Pr (X_2^n)} = \left[ \frac{\Pr (G_{11}^n, G_{21}^n) \Pr (X_1^n | G_{11}^n, G_{21}^n)}{\Pr (X_1^n)} \right] \left[ \frac{\Pr (G_{12}^n, G_{22}^n) \Pr (X_2^n | G_{12}^n, G_{22}^n)}{\Pr (X_2^n)} \right] \mathbf{1}_{\{Y_1^n = G_{11}^n, X_1^n \oplus G_{21}^n, X_2^n\}}, \] (24)

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 (23); and \((c)\) holds since \(X_1^n\) is independent of \(G_{11}^n\) and \(G_{21}^n\), \(i = 1, 2\). Similarly, we can write the marginal distribution at receiver \( \text{Rx}_2 \).

\[
\Pr (Y_2^n, G^n | X_1^n, X_2^n) = \left[ \frac{\Pr (G_{11}^n, G_{21}^n) \Pr (X_1^n | G_{11}^n, G_{21}^n)}{\Pr (X_1^n)} \right] \left[ \frac{\Pr (G_{12}^n, G_{22}^n) \Pr (X_2^n | G_{12}^n, G_{22}^n)}{\Pr (X_2^n)} \right] \mathbf{1}_{\{Y_2^n = G_{12}^n, X_1^n \oplus G_{21}^n, X_2^n\}}. \] (25)

From the intuition provided in [3], and the distributions given by (24) and (25), we conclude that as long as the joint distributions

\[ \Pr (G_{11}^n, G_{21}^n) \quad \text{and} \quad \Pr (G_{12}^n, G_{22}^n) \] (26)

remain the same, and the same input distributions are applied, the capacity region would be the same. This immediately implies that the capacity region \( C(V.4) \) is included in 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 \[\text{VII}\] 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

\[
\bar{Y}_i[t] = \tilde{G}_{ii}[t] \tilde{X}_i[t] + \bar{G}_{ii}[t] \tilde{X}_i[t], \quad i = 1, 2. \tag{27}
\]

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

\[
\tilde{G}_{ii}[t] = G_{ii}[t] \quad i = 1, 2. \tag{28}
\]

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 includes \(\mathcal{C}(\text{V.4})\). Suppose rate-tuple \((\tilde{R}_1, \tilde{R}_2)\) is achievable. The derivation of individual bounds for this channel

\[
\tilde{R}_i \leq 0.5, \quad i = 1, 2, \tag{29}
\]

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

\[
n \left( \tilde{R}_1 + \tilde{R}_2 - \epsilon_n \right) \overset{(a)}{\leq} I \left( \tilde{W}_1; \tilde{Y}_1^n | \tilde{W}_2, G^n \right) + I \left( \tilde{W}_2; \tilde{Y}_2^n | \tilde{G}_n \right) = H \left( \tilde{Y}_1^n | \tilde{W}_2, G^n \right) - H \left( \tilde{Y}_1^n | \tilde{W}_1, \tilde{W}_2, G^n \right) + H \left( \tilde{Y}_2^n | \tilde{G}_n \right) - H \left( \tilde{Y}_2^n | \tilde{W}_2, G^n \right) = H \left( \tilde{Y}_1^n | \tilde{G}_n \right) \overset{(b)}{\leq} \frac{3n}{4}, \tag{30}
\]

where \(\epsilon_n \to 0\) as \(n \to \infty\); (a) follows from Fano’s inequality and the fact that messages and channel gains are mutually independent; (b) holds since according to (27), we have

\[
H \left( \tilde{G}_{11}^n \tilde{X}_1^n | \tilde{G}_n \right) = H \left( \tilde{G}_{12}^n \tilde{X}_1^n | \tilde{G}_n \right). \tag{31}
\]

Dividing both sides by \(n\) and letting \(n \to \infty\), we obtain

\[
\tilde{R}_1 + \tilde{R}_2 \leq \frac{3}{4}. \tag{32}
\]

Note that the region described by (29) and (32) 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 (29) and (32).

**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

\[
\mathcal{S}_{\text{Tx}_1} = \{(1,1), (2,2)\} \quad \text{and} \quad \mathcal{S}_{\text{Tx}_2} = \{(1,1), (2,2)\}. \tag{33}
\]

Thus writing the marginal distribution for receiver Rx_1, we get

\[
\Pr \left( Y_1^n, G^n | X_1^n, X_2^n \right) = \Pr \left( G_{12}^n, G_{21}^n \right) \Pr \left( G_{11}^n, G_{22}^n \right) \Pr \left( X_1^n | G_{11}^n, G_{22}^n \right) \Pr \left( X_2^n | G_{12}^n, G_{22}^n \right) \Pr \left( X_1^n, X_2^n \right) 1_{\{Y_1^n = G_{11}^n X_1^n \oplus G_{22}^n X_2^n\}}. \tag{34}
\]

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

\[
\Pr \left( Y_2^n, G^n | X_1^n, X_2^n \right) = \Pr \left( G_{12}^n, G_{21}^n \right) \Pr \left( G_{11}^n, G_{22}^n \right) \Pr \left( X_1^n | G_{11}^n, G_{22}^n \right) \Pr \left( X_2^n | G_{12}^n, G_{22}^n \right) \Pr \left( X_1^n, X_2^n \right) 1_{\{Y_2^n = G_{12}^n X_1^n \oplus G_{22}^n X_2^n\}}. \tag{35}
\]
From the intuition provided in [3], and the distributions given by [14], and [15], we conclude that as long as the joint distributions
\[
\Pr (G_{11}^n, G_{21}^n), \quad \Pr (G_{12}^n, G_{22}^n), \quad \text{and} \quad \Pr (G_{11}^n, G_{22}^n),
\]
remain the same, and the same joint input distribution is applied, the capacity region would be the same. We note that the same correlation introduced in [27] 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 [6] for the case of global delayed CSIT relies heavily on the delayed knowledge of the entire channel state information at each transmitter.

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

![Fig. 4. Providing \( a_1 \oplus a_2 \) available at \( \text{Tx}_1 \) and \( b_1 \oplus b_2 \) available at \( \text{Tx}_2 \) to both receivers is sufficient to decode the bits.](image)

[Pairing Across Realizations Type-I] Suppose at a time instant, each one of the transmitters simultaneously sends one data bit. The bits of \( \text{Tx}_1 \) and \( \text{Tx}_2 \) 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 \( \text{Rx}_1 \) 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 \( \text{Tx}_1 \) to both receivers is useful. For instance if \( \text{Rx}_2 \) 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 \( \text{Tx}_2 \).

[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 \( \text{Tx}_2 \) becomes a bit of common interest.
Fig. 5. The combination $a_3 \oplus a_4$ available at $T_{x1}$ is of interest of both receivers.

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.

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 $T_{x1}$ (the configurations known to $T_{x2}$ 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 $T_{x1}$’s point of view, this scenario is identical to the one depicted in Fig. 4 and Fig. 5. However, providing $a_5 \oplus a_6$ to $R_{x2}$ is not useful anymore. 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 $T_{x1}$ can overcome the challenge described above by relying on the fact that statistically half of the bits that at the time of transmission faced $G_{12}[t] = 1$, are already known to $R_{x2}$ (since with probability $0.5$ we have $G_{22}[t] = 0$). Keeping this fact in mind, transmitter $T_{x1}$ will create enough linearly independent combinations of the bits such that receiver $R_{x2}$ 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 $T_{x1}$ is
TABLE I
THE FOUR POSSIBLE CONFIGURATIONS THAT CAN BE IDENTIFIED BY TRANSMITTER $T_x_1$. THE BIT TRANSMITTED BY $T_x_1$ 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       | $T_x_1 \rightarrow Q_{1,1}$          |                 | 3       | $T_x_1 \rightarrow Q_{1 \rightarrow F}$|                 |
|         | $G_{11}[t] = 1$                      |                 |         | $G_{12}[t] = 0$                      |                 |
|         | $G_{21}[t] = 0$                      |                 |         | $G_{22}[t] = 0$                      |                 |
| 2       | $T_x_1 \rightarrow Q_{1,2}$          |                 | 4       | $T_x_1 \rightarrow Q_{1 \rightarrow 1}$|                 |
|         | $G_{11}[t] = 0$                      |                 |         | $G_{12}[t] = 1$                      |                 |
|         | $G_{21}[t] = 0$                      |                 |         | $G_{22}[t] = 0$                      |                 |

easy. Suppose at time $t$, transmitter $T_x_1$ sends one data bit. Later, using local delayed CSIT, $T_x_1$ 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 $R_x_1$ then it would be the responsibility of $T_x_2$ to resolve it in future. If $G_{11}[t] = 0$ and $G_{12}[t] = 0$, then the bit must be retransmitted.

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

B. Transmission Strategy

We focus on the corner point $(R_1, R_2) = (0.45, 0.45)$ as shown in Fig. 8. The achievability strategy for the other corner points is presented in Appendix A. 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

$$
\frac{20}{9} m \underbrace{+ \frac{35}{3} m ^\frac{2}{3}}_{(37)}
$$

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 $T_x_i$ are in queue $Q_{i \rightarrow i}$ (the initial state of the bits), $i = 1, 2$. At each time instant $t$, $T_x_i$ 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 $G_{ii}^*[t] = 1$.

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

$$\frac{4}{3}m + m^2 \frac{2}{3}$$

(38)
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 $m^2 \in \mathbb{Z}$). 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.

$$N_{i,1} > \mathbb{E}[N_{i,1}] + 2m^2 \Delta = n_{i,1}, \quad i = 1, 2;$$

$$N_{i,2} > \mathbb{E}[N_{i,2}] + 2m^2 \Delta = n_{i,2}, \quad i = 1, 2.$$  (39)

From basic probability, we have

$$\mathbb{E}[N_{i,1}] = \mathbb{E}[N_{i,2}] = \frac{m}{3},$$  (40)

thus we get

$$n_{i,1} = n_{i,2} = \frac{m}{3} + 2m^2.$$  (41)

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

$$N_{i,j|Rx_i}, \quad i, j \in \{1, 2\}.$$  (42)

At the end of communication, if we have

$$N_{i,j|Rx_i} < \frac{1}{2}n_{i,j} - m^2, \quad i, j \in \{1, 2\},$$  (43)

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 square matrices $C_{i,1}$ and $C_{i,2}$, $i = 1, 2$, of size $\left(\frac{m}{3} + 4m^2\right) \times \left(\frac{m}{3} + 2m^2\right)$ each, where entries to each matrix are drawn from i.i.d. $B(0.5)$ distribution. We assume that these matrices are generated prior to communication and are shared with receivers. Transmitter $Tx_i$ does not need to know $C_{i,1}$ or $C_{i,2}$, $i = 1, 2$. 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 [14].

**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 $Q_i$ using the following equation.

$$Q_i \triangleq C_{i,1}Q_{i,1} \oplus C_{i,2}Q_{i,2}, \quad i = 1, 2.$$  (44)
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 $(R_1, R_2) = (\frac{3}{8}, \frac{3}{8})$ is achievable. In other words, for fixed $\epsilon, \delta > 0$, rate-tuple $(R_1, R_2) = (\frac{3}{8} - \frac{\epsilon}{4}, \frac{3}{8} - \frac{\delta}{4})$ is achievable with error less than or equal to $\epsilon$.

![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.](image)

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

$$\frac{2m}{3} + 8m^2 \frac{2}{3} - \frac{\delta}{3}$$

(45)

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 (43), 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 $\frac{m}{3} + 4m^2$ random linear combinations of (at most) $\frac{m}{3} + 4m^2$ 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

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

(46)

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

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

(47)

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

**VI. Remarks on converse with global delayed CSIT**

In this section, our goal is to provide a better understanding of the converse techniques presented in this paper. To that end, we first describe the converse with global delayed CSIT and then, we highlight the salient techniques that allow us to obtain the outer-bound for local delayed CSIT of Views V.1, V.3, and V.4. The capacity region of the two-user Binary Fading IC with global delayed CSIT was obtained in [6]. The key converse idea there is an extremal entropy inequality for a binary broadcast channel with delayed CSIT.
A. An Extremal Entropy Inequality

Consider a broadcast setting in which a single-antenna transmitter is connected to two single-antenna receivers through binary fading channels as in Fig. 10. For this network, suppose \( G_1[t] \) and \( G_2[t] \) are distributed as i.i.d. Bernoulli RVs \( i.e. \ G_i[t] \sim B(0.5), \ i = 1, 2 \). Then the observed signals are given as

\[
Y_i[t] = G_i[t]X[t], \quad i = 1, 2,
\]

where \( X[t] \) is the transmit signal at time \( t \). We assume that at time \( t \), the transmitter has access to \( (G_1[\ell], G_2[\ell])_{\ell=1}^{t-1} \) (delayed CSIT).

The goal is to find the minimum value of \( \beta \) in this channel that satisfies

\[
H(Y_2^n | G^n) \geq \beta H(Y_1^n | G^n).
\]

In a sense, the goal is to quantify how much the transmitter can favor receiver one over receiver two in terms of the available entropy.

![Fig. 10. A transmitter connected to two receivers through binary fading channels.](image)

**Lemma 1 (Entropy Leakage [6]):** For the channel described above with delayed CSIT and for any input distribution, the minimum value of \( \beta \) satisfying (49) is \( \frac{2}{3} \).

**Remark 5:** The lower-bound \( \beta = \frac{2}{3} \) is in fact achievable using a simple transmission strategy and thus, the bound is tight in the Entropy Leakage Lemma. Furthermore, to achieve this lower-bound, the transmitter only needs to have access to delayed knowledge of one of the channel gains not both.

B. Deriving the Outer-Bound and Remarks

Here, for completeness, we provide the proof of

\[
R_1 + \frac{3}{2} R_2 \leq \frac{9}{8},
\]

as given in [6], for the two-user BFIC with global delayed CSIT. Suppose the rate tuple \((R_1, R_2)\) is achievable. Then we have

\[
n(R_1 + \beta^{-1} R_2) = H(W_1 | W_2, G^n) + \beta H(W_2 | G^n)
\]

(Fano)

\[
\leq I(W_1; Y_1^n | W_2, G^n) + \beta I(W_2; Y_2^n | G^n) + n \epsilon_n
\]

\[
= H(Y_1^n | W_2, G^n) - H(Y_1^n | W_1, W_2, G^n)
\]

\[
+ \beta^{-1} H(Y_2^n | G^n) - \beta^{-1} H(Y_2^n | W_2, G^n) + n \epsilon_n
\]

\[
= \beta^{-1} H(Y_2^n | G^n) + H(Y_1^n | W_2, X_2^n, G^n) - \beta^{-1} H(Y_2^n | W_2, X_2^n, G^n) + n \epsilon_n
\]

\[
= \beta^{-1} H(Y_2^n | G^n) + H(G_{11} X_1^n | W_2, X_2^n, G^n) - \beta^{-1} H(G_{12} X_1^n | W_2, X_2^n, G^n) + n \epsilon_n
\]

\[
\leq n \beta^{-1} H(Y_2^n | G^n) + H(G_{11} X_1^n | W_2, G^n) - \beta^{-1} H(G_{12} X_1^n | G^n) + n \epsilon_n
\]

**Lemma 11**

\[
\leq n \beta^{-1} H(Y_2^n | G^n) + n \epsilon_n \leq \frac{9}{8} n + n \epsilon_n.
\]

(51)
where (a) holds since $X^n_2$ is a deterministic function of $W_2$ and $G^n$; (b) follows from

\[
0 \leq H(G^n_{11}X^n_1|G^n) - H(G^n_{11}X^n_1|W_2,G^n) = I(G^n_{11}X^n_1;W_2|G^n)
\]

\[
\leq I(W_1, G^n_{11}X^n_1; W_2|G^n) = \frac{I(W_1; W_2|G^n)}{2} + \frac{I(G^n_{11}X^n_1; W_2|W_1, G^n)}{2} = 0,
\]

which implies $H(G^n_{11}X^n_1|G^n) = H(G^n_{11}X^n_1|W_2, G^n)$, and similarly $H(G^n_{12}X^n_1|G^n) = H(G^n_{12}X^n_1|W_2, G^n)$. Dividing both sides by $n$ and letting $n \to \infty$, we get

\[
R_1 + \frac{3}{2}R_2 \leq \frac{9}{8}.
\]

**Remark 6:** As stated above, in order to achieve the lower-bound $\beta = \frac{2}{7}$, a transmitter only needs to have access to delayed knowledge of one of the channel gains not both. In local delayed CSIT of Views V.1, V.3, and V.4, transmitter $\text{Tx}_i$ can in fact achieve favor receiver $i$ over receiver $i'$ in terms of the available entropy by the value given in the Entropy Leakage Lemma ($\beta = \frac{2}{3}$). Thus, to obtain a tight outer-bound, we had to create a spatial correlation among channel gains such that: (1) the capacity region is preserved; and (2) in the new channel with correlated links, a transmitter can no longer favor a receiver to the other. This goal was achieved by setting

\[
\tilde{G}_{ii}[t] = \tilde{G}_{ii}[t] \quad i = 1, 2.
\]

Then, we were able to show that under such spatial correlation, the capacity region coincides with that of no CSIT (View V.0).

**VII. DISCUSSION**

In this section, we discuss the problem of two-user BFIC 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 BFIC with local delayed CSIT of View V.7**

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

\[
\mathcal{S}_\text{Tx}_1 = \{(1, 1), (2, 1), (2, 2)\} \quad \text{and} \quad \mathcal{S}_\text{Rx}_2 = \{(1, 1), (1, 2), (2, 2)\}.
\]

Thus writing the marginal distribution at receiver $\text{Rx}_1$, we get

\[
\Pr(Y^n_1, G^n|X^n_1, X^n_2) = \left[ \frac{\Pr(G^n_{11}, G^n_{12}, G^n_{21}, G^n_{22}) \Pr(X^n_1|G^n_{11}, G^n_{21}, G^n_{22}) \Pr(X^n_2|G^n_{12}, G^n_{21}, G^n_{22})}{\Pr(X^n_1, X^n_2)} \right] \mathbf{1}_{\{Y^n_1 = G^n_{11}X^n_1 \oplus G^n_{21}X^n_2\}}.
\]

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

\[
\tilde{G}_{11}[t] = \tilde{G}_{12}[t],
\]

then, we have changed the channel from $\text{Rx}_2$’s point of view and thus, the marginal distributions cannot be preserved.

On the other hand, as discussed in Section [V] delayed knowledge of $G^n_{ii}$ has an important role on the future decisions taken by $\text{Tx}_i$, $i = 1, 2$. In fact, we cannot distinguish $Q_{i, 1}$ from $Q_{i, 2}$ without delayed knowledge of $G^n_{ii}$, 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.
Fig. 11. $k$-user BFIC. The capacity region with global delayed CSIT is open.

B. $k$-user BFIC with delayed CSIT

Here, we take the results and intuitions obtained for the two-user BFIC and try to understand the implications in broader settings. We consider the capacity region of the $k$-user BFIC (see Fig. 11) 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 $D_k$.

The intuition for the two-user BFIC was that it is the responsibility of the transmitter who creates interference to resolve it. Characterizing $D_k$ or the capacity region of the $k$-user BFIC 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 [12], [15] and references therein). In [16], 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 BFIC. However, in the lack of a tight outer-bound, a firm conclusion cannot be made.

VIII. 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. 12 summarizes our main results.

As discussed in Section VII, 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).

APPENDIX A

TRANSMISSION STRATEGY FOR THE CORNER POINT $(0.375, 0.5)$

We now provide the achievability strategy for the corner point

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

(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. The proposed transmission strategy consists of four phases as described below. We assume that
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

$$m + m \frac{2}{3}$$

(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.

\[ N_{i,1} > \mathbb{E}[N_{i,1}] + 2m^{\frac{1}{2}} \triangleq n_{i,1}, \quad i = 1, 2; \]
\[ N_{i,2} > \mathbb{E}[N_{i,2}] + 2m^{\frac{1}{2}} \triangleq n_{i,2}, \quad i = 1, 2. \]  

(60)

From basic probability, we have

\[ \mathbb{E}[N_{i,1}] = \mathbb{E}[N_{i,2}] = \frac{m}{4}, \]

so that

\[ n_{i,1} = n_{i,2} = \frac{m}{4} + 2m^{\frac{1}{2}}. \]  

(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 \( Rx_i \), \( i = 1, 2 \). Denote the number of bits in \( Q_{i,j} \) known to \( Rx_i \) by

\[ N_{i,j | Rx_i}, \quad i, j \in \{1, 2\}. \]  

(63)

At the end of communication, if we have

\[ N_{i,j | Rx_i} < \frac{1}{2} n_{i,j} - m^{\frac{1}{2}}, \quad i, j \in \{1, 2\}, \]

we declare error type-III.

Moreover, we note that statistically for every two bits in \( Q_{i,1} \), 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 \( Q_{i,1} \) by

\[ N_{i \rightarrow i,1}, \quad i = 1, 2. \]  

(65)

At the end of communication, if we have

\[ N_{i \rightarrow i,1} < \frac{1}{2} n_{i,1} - m^{\frac{1}{2}}, \quad i = 1, 2, \]

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 \( Tx_2 \) transmits its remaining initial bits while the secondary user \( Tx_1 \) 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 \( Tx_1 \) creates a matrix \( C_{1,1} \) of size \( \left( \frac{1}{4} m + m^{2/3} \right) \times \left( \frac{m}{4} + 2m^{\frac{1}{2}} \right) \), where the entries of this matrix are drawn from i.i.d. \( B(0.5) \) distribution. We assume that this matrix is generated prior to communication and is shared with receivers. Then \( Tx_1 \) creates \( \left( \frac{1}{4} m + m^{2/3} \right) \) 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 \( Tx_1 \) 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 \( m \to \infty \), each receiver can decode bits \( C_{1,1}Q_{1,1} \) with vanishing error probability.

Transmitter \( Tx_2 \) 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, \( Tx_2 \) 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 \( Tx_1 \) can be
decoded first, we simply consider the bits of Tx2 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

$$n_{2.2} = \frac{m}{3} + 3m^{z.2}. \quad (67)$$

**Phase 3** [encoding and mixing interfering bits]: Transmitter Tx1 creates two matrices: $C_{1,2}$ of size $\left(\frac{m}{6} + 2m^{z.2}\right) \times \left(\frac{m}{4} + 2m^{z.2}\right)$ and $C_{1,3}$ of size $\left(\frac{m}{8} + 2m^{z.2}\right) \times \left(\frac{m}{4} + 2m^{z.2}\right)$, where entries to each matrix are drawn from i.i.d. $B(0.5)$ distribution. We assume that the matrices are generated prior to communication and are shared with receivers.

Transmitter Tx1 creates

$$\hat{Q}_{1,1} = C_{1,2}Q_{1,1},$$
$$\hat{Q}_{1,2} = C_{1,3}Q_{1,2}. \quad (68)$$

Then Tx1 encodes bits in $\hat{Q}_{1,1}$ using a point-to-point erasure code of rate 1/4 denoted by $\hat{Q}_{1,1}$ and encodes bits in $\hat{Q}_{1,2}$ using a point-to-point erasure code of rate 1/2 denoted by $\hat{Q}_{1,2}$. Transmitter Tx1 communicates $\hat{Q}_{1,1} \oplus \hat{Q}_{1,2}$ during Phase 3.

At the same time, transmitter Tx2 creates two matrices: $C_{2,1}$ of size $\left(\frac{m}{8} + 2m^{z.2}\right) \times \left(\frac{m}{4} + 2m^{z.2}\right)$ and $C_{2,2}$ of size $\left(\frac{m}{8} + 2m^{z.2}\right) \times \left(\frac{m}{8} + 2m^{z.2}\right)$, where entries to each matrix are drawn from i.i.d. $B(0.5)$ distribution. We assume that the matrices are generated prior to communication and are shared with receivers.

Transmitter Tx2 creates

$$\hat{Q}_{2,1} = C_{2,1}Q_{2,1},$$
$$\hat{Q}_{2,2} = C_{2,2}Q_{2,2}. \quad (69)$$

Then Tx2 encodes bits in $\hat{Q}_{2,1}$ using a point-to-point erasure code of rate 1/4 denoted by $\hat{Q}_{2,1}$ and encodes bits in $\hat{Q}_{2,2}$ using a point-to-point erasure code of rate 1/2 denoted by $\hat{Q}_{2,2}$. Transmitter Tx2 communicates $\hat{Q}_{2,1} \oplus \hat{Q}_{2,2}$ 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 Rx1 first removes the known bits from $Q_{i,1}$ and $Q_{i,2}$, $i = 1, 2$.

Then, each receiver has

$$\left(\frac{m}{12} + 2m^{z.2}\right) \left(\frac{m}{6} + 2m^{z.2}\right) = \frac{m}{4} + 2m^{z.2}$$

\[C_{1,1}Q_{1,1} + C_{1,2}Q_{1,1} \tag{68} \]

\[C_{1,1}Q_{1,1} \tag{69} \]

$^2^3$ The two sequences are not of equal length, we can simply add deterministic number of zeros to $\hat{Q}_{1,2}$ to make the two sequences of equal length.
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 $m \to \infty$. 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 $Tx_2$. However, transmitter $Tx_2$ 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 $Tx_2$ 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 $(\frac{m}{4} + 2m\frac{2}{3}) \times (\frac{m}{4} + 2m\frac{2}{3})$.

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

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

time instants.

Hence, if we let $m \to \infty$, 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 $C$ (V.8) is achievable with local delayed CSIT of View V.2.

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] H. Te Sun and K. Kobayashi, “A new achievable rate region for the interference channel,” IEEE Transactions on Information Theory, vol. 59, no. 11, pp. 7014–7040, 2013.

[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. 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.

[7] K. Jolfaei, S. Martin, and J. Mattfeldt, “A new efficient selective repeat protocol for point-to-multipoint communication,” in IEEE International Conference on Communications (ICC’09), vol. 2, pp. 1113–1117, IEEE, 1993.

[8] 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.

[9] 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.

[10] 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.

[11] 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.

[12] H. Maleki, S. Jafar, and S. Shamai, “Retrospective interference alignment,” Arxiv preprint arXiv:1009.3593, 2010.

[13] 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.

[14] 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.

[15] 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.

[16] 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.