Capacity of Broadcast Packet Erasure Channels With Single-User Delayed CSI

Shih-Chun Lin, *Senior Member, IEEE*, I-Hsiang Wang, *Member, IEEE*, and Alireza Vahid, *Senior Member, IEEE*

***Abstract*— We characterize the capacity region of the two-user broadcast packet erasure channel (PEC) with single-user delayed channel state information (CSI). More precisely, we assume one receiver does not provide its channel state to the other two nodes (the other receiver and the transmitter), while the other receiver reveals its state globally with unit delay. This is a hybrid CSI at the transmitter (CSIT) setting where the transmitter has the delayed CSI of one user but not the other. Previous results developed opportunistic network coding schemes for this setting, which strictly enlarge the achievable rate region compared to the no-CSIT baseline. Characterization of the capacity region with single-user delayed CSI, however, remained open. In this work, we develop an improved achievability strategy and show that the capacity region, surprisingly, matches that of the broadcast PEC with global delayed CSI of both users. The key to such improvement over previous results is a new precoding strategy for the retransmission phase of the opportunistic network coding scheme. It harnesses the single-user delayed CSI in the retransmission phase, so that interference from the feedback receiver can be aligned at the other receiver. Besides the broadcast PEC with two private messages, an extension to a model with an additional common message is also provided and the corresponding capacity region with single-user CSI also matches that with global delayed CSI. Finally, further extensions to three-user cases are also provided.**

***Index Terms*— Broadcast channel (BC), Automatic Repeat reQuest (ARQ), feedback.**

I. INTRODUCTION

RECENTLY, broadcast channels (BCs) with channel state feedback have been actively studied [1]–[5]. In a packet-based communication network, instead of the classic Gaussian channel, the network-coding-based approaches generally model each communication hop as a packet erasure channel (PEC) [6]. Motivated by the wireless network coding, feedback capacity of memoryless broadcast PEC was studied in [1]–[3]. By sending back the ACKnowledgement (ACK) signal from each receiver to the transmitter, the system can achieve channel state feedback with acceptable overhead. The simple PEC abstraction not only allows tracing information-theoretic capacity with mathematical rigor and but also accelerates the transition from theory to practice. One prominent example is the Automatic Repeat reQuest (ARQ) and its variations, which is conceived by the point-to-point PEC capacity and have been incorporated into 5G mobile networks [7]. Moreover, many capacity-achieving network codes [8] have also been implemented for the wireless multi-hop networks [9], [10]. How to leverage the delayed state feedback to mitigate the inter-session interference and to improve the performance of broadcast PECs then has drawn a lot of attentions.

In contrast to the point-to-point case where feedback cannot increase the capacity region, it was shown that delayed state feedback from the receivers strictly enlarges the capacity region of the broadcast PEC compared to that without feedback [1]–[3], [11]–[13]. The key to such improvement is harnessing the state feedback to create network coding opportunities so that when the transmitter retransmits the erased information, coded packets that are simultaneously beneficial to multiple receivers can be created. Among these works, [1]–[3] assumed that all receivers can provide delayed state feedback, while [11]–[13] discovered that in the context of the two-user broadcast PEC, the erasure state feedback from a single receiver suffices to strictly improve the capacity region. This is in sharp contrast to the fast-fading multiple-input single-output (MISO) BCs with continuous channel gains [5], where it was shown that the sum degrees-of-freedom (DoF) collapses to that of the case without feedback, which is strictly smaller than that of global delayed feedback [4].

For the two-user broadcast PEC, to achieve such capacity improvement with single-user delayed CSI only, [13] proposed a three-phase scheme sketched as follows. In the first two phases, the transmitter sends information bits intended for the two users, respectively, and determines the bits to be recycled in the third phase according to the single-user feedback. In the third phase, bits to be recycled are first re-encoded with linear codes separately, and then, the coded bits are XORed to generate bits to be sent. Consequently,

Manuscript received February 24, 2020; revised March 31, 2021; accepted July 21, 2021. Date of publication August 9, 2021; date of current version September 15, 2021. The work of Shih-Chun Lin was supported in part by the Ministry of Science and Technology (MOST), Taiwan, under Grant 107-2628-E-011-003-MY3. The work of I-Hsiang Wang was supported by the Ministry of Science and Technology (MOST), Taiwan, under Grant MOST 106-2923-E-002-015-MY3. The work of Alireza Vahid was supported in part by the National Science Foundation (NSF), USA, under Grant ECCS-2030285. This article was presented in part at the 2019 IEEE International Symposium on Information Theory (ISIT) and in part at the 2020 Global Communications Conference (GLOBECOM). *(Corresponding author: Shih-Chun Lin.)*

Shih-Chun Lin and I-Hsiang Wang are with the Department of Electrical Engineering, National Taiwan University, Taipei 10617, Taiwan (e-mail: sclin2@ntu.edu.tw; ihwang@ntu.edu.tw).

Alireza Vahid is with the Department of Electrical Engineering, University of Colorado Denver, Denver, CO 80204 USA (e-mail: alireza.vahid@ucdenver.edu).

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

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TIT.2021.3103801.

Digital Object Identifier 10.1109/TIT.2021.3103801

it opportunistically creates the network coding benefits for both receivers as shown in [13]. Prior to [13], it was also found in [12] that this enlarged rate region can be achieved by using the typicality-based nonlinear achievability developed for general discrete memoryless broadcast channels with feedback [14]–[16] specialized to broadcast channels with delayed state information as in [17]. Interestingly, some non-trivial rate pairs on the boundary of the capacity region of the broadcast PEC with global delayed CSI could be achieved. However, it remains an open question whether or not the rate region in [12], [13] is optimal.

Moreover, only the private message for each receiver is considered in [12], [13] and the common message for both receivers is absent. The transmission of common and private messages in BC naturally arises in more complicated problems such as the X-Channel and the two-unicast problem [18], [19], highlighting its importance. Despite all the attention the broadcast PEC has attracted over the decades, the capacity regions with common and private messages have been found in only a few cases when the receivers are allowed to feed back. Note that although the common message was considered in [1], [20], the achievability requires global delayed CSI.

In this paper, we settle the open question and characterize the capacity region of the two-user broadcast PEC with single-user delayed CSI, which, to our surprise, coincides with the one with delayed feedback from both receivers. This result is true not only for the case where there are two private messages but also for the one with an additional common message. Our main contributions are as follows

1) The achievability strategy for the case with private messages only is established by proposing a new precoding strategy for the recycled bits in the third (retransmission) phase of the three-phase opportunistic network coding scheme proposed in [13]. The previous scheme only utilizes the CSI feedback in the retransmission phase. Meanwhile, our new precoding strategy fully utilizes the single-user delayed CSI feedback in the retransmission phase and *dynamically* adapts to it, which is the key to why we can further enlarge the achievable rate region and achieve the capacity. More specifically, in [13], the receiver that never feeds back its channel state needs to decode the recycled bits for *both* users. By efficiently aligning the interference according to the single-user CSI feedback, each receiver in our scheme only decodes its own recycled bits, and the decoding burden is significantly reduced. Moreover, to further improve the three-phase achievability presented in the conference version [21], we propose a two-phase achievability strategy by merging two of the phases. This achievability also simplifies the design for the case with an additional common message.

2) To extend to the case with common and private messages, before transmitting the private messages using the proposed scheme in Contribution 1, one can add a new phase to multicast the common message with rate being the link capacity of the weaker receiver (user with larger erasure probability). Unfortunately, this simple extension cannot achieve the capacity when the erasure probabilities of two links are unequal. The challenge is how to break the limit from the weaker receiver in the first phase and yet ensure decodability. Our solution is to simultaneously create equations of the common message at the weaker receiver and re-transmit private bits after the first phase. Finally, extensions to three-user cases with single-user CSI are also provided.

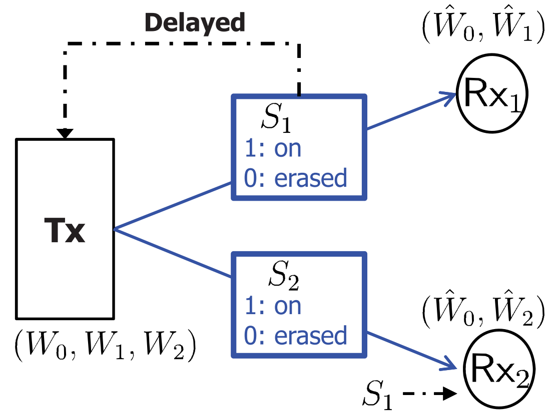

Fig. 1. Broadcast packet erasure channel with single-user delayed CSI, where a common message W_0 is targeted for both receivers and private message W_i is targeted for receiver Rx$i, i = 1, 2$.

The rest of this paper is organized as follows. In Section II, single-user delayed CSI is defined and the problem is formulated. Then in Section III, we present our main results of the two-user cases. Numerical results are also provided. The achievability proof for the case with two private messages only is given in Section IV, by proposing a two-phase opportunistic linear network coding. Section V provides the proof for the two-user case with common and privates messages while Section VI extends the result to the three-user case. Finally, Section VII concludes the paper.

II. Problem Formulation and Preliminaries

A. Problem Formulation

As depicted in Figure 1, we consider the two-user broadcast PEC in which one transmitter wishes to communicate one common message (packet streams) W_0 to two receiving terminals Rx1 and Rx2, and private messages W_1 and W_2 to Rx1 and Rx2, respectively, over n channel uses. Here, we assume messages are independently distributed (from each other and channel parameters), and that each W_i is a nR_i-dimensional vector[1] in a finite field $\mathbb{F}_q$ and uniformly distributed, for $i = 0, 1, 2$. The three messages are mapped to the channel input $X[t] \in \mathbb{F}_q$ and the corresponding received signals at Rx1 and Rx2 are

$$\begin{aligned} Y_1[t] &= S_1[t]X[t] \\ Y_2[t] &= S_2[t]X[t], \end{aligned} \tag{1}$$

respectively, where $\{S_i[t]\}$ denotes the Bernoulli $(1 - \delta_i)$ process that governs the erasure at Rxi, and it is independent and identically distributed (i.i.d.) distributed over time. When $S_i[t] = 1$, Rxi receives $X[t]$ noiselessly; and when $S_i[t] = 0$, it receives an erasure. We also assume

$$\delta_{12} = P\{S_1[t] = 0, S_2[t] = 0\}.$$

[1] As [22], the unit of our rate R_i is packets per time slot and can be converted to the traditional unit bits per time slot by multiplying a factor of $\log_2(q)$.

We assume only Rx1 feeds back its state and thus, the transmitter knows the channel state information (CSI) in scenario "DN", where $S_1[t]$ is known with *unit delay* but $S_2[t]$ is *unknown*. The transmission problem under this scenario is mathematical defined as follows. The constraint imposed at the encoding function $f_t(.)$ at time index t is

$$X[t] = f_t\left(W_0, W_1, W_2, S_1^{[1:\,t-1]}\right), \tag{2}$$

where $S_i^{[1:\,t-1]} = (S_i[1], \ldots, S_i[t-1]), i = 1, 2$. Under scenario DN, each receiver knows its own CSI across the entire transmission block but only receiver 2 knows additional CSI $S_1^{[1:\,n]}$ via the feedback channel. In other words, not only the CSI at the transmitter is hybrid and heterogenous, the CSI at the receivers is also heterogenous. The corresponding error probability constraints at Rx1 and Rx2 respectively are

$$\begin{aligned} &\Pr\left\{(W_1, W_0) \neq g_1(Y_1^{[1:n]}, S_1^{[1:\,n]})\right\} \to 0,\\ &\Pr\left\{(W_2, W_0) \neq g_2(Y_2^{[1:n]}, S_1^{[1:\,n]}, S_2^{[1:\,n]})\right\} \to 0, \end{aligned} \tag{3}$$

as $n \to \infty$, where $g_i(.)$ is the decoding function at receiver i. The channel statistics $\delta_1, \delta_2, \delta_{12}$ and the code blocklength n are known to the entire network.

The capacity region is the closure of the collection of all non-negative rate triples (R_1, R_2, R_0) satisfying the error probability constraints.

B. Recap Results With Only Private Messages: The Opportunistic Three-Phase Network Coding in [13] and Capacity Outer-Bounds in [1], [22]

We briefly recapitulate the baseline three-phase linear network coding scheme in [13] for the DN scenario with only private messages. Assume $q = 2$ (*i.e.* binary inputs), and let us view the two messages W_1 and W_2 as bit vectors $\mathbf{w}_1 \in \mathbb{F}_2^{1\times R_1 n}$ and $\mathbf{w}_2 \in \mathbb{F}_2^{1\times R_2 n}$ respectively, where n is the length of the total three phases. The first two phases, with lengthes n_1 and n_2 respectively, are:

Phase I: The transmitter sends randomly linear coded bits of $\mathbf{w}_1$.

Phase II: Similar to Phase I, but now the transmitter sends coded bits of $\mathbf{w}_2$.

After these two phases, the transmitter identifies two bit sequences to be recycled, called *recycled sequences* hereafter, according to the delayed state feedback from Rx1

- *Recycled sequence* $\bar{\mathbf{v}}_1$: Sequence $\bar{\mathbf{v}}_1$ consists of coded bits of $\mathbf{w}_1$ erased at Rx1 in Phase I
- *Recycled sequence* $\bar{\mathbf{v}}_2$: Sequence $\bar{\mathbf{v}}_2$ consists of coded bits of $\mathbf{w}_2$ target for user 2 but are received at Rx1 in Phase II

Phase III (Retransmission Phase): To recycle $\bar{\mathbf{v}}_1$ and $\bar{\mathbf{v}}_2$, the transmitter precodes them into random linear combinations $(\mathbf{g}_1[t_3])^\intercal \bar{\mathbf{v}}_1$ and $(\mathbf{g}_2[t_3])^\intercal \bar{\mathbf{v}}_2$ respectively, and then sends the XOR

$$(\mathbf{g}_1[t_3])^\intercal \bar{\mathbf{v}}_1 \oplus (\mathbf{g}_2[t_3])^\intercal \bar{\mathbf{v}}_2 \tag{4}$$

at time index t_3. With the side information of $\bar{\mathbf{v}}_1$ and $\bar{\mathbf{v}}_2$ opportunistically obtained in Phase I and Phase II at the receivers, it is proved in [13, Corollary 4] that one can recover the achievable rate region of [12, Theorem 1] as follows, which is the state-of-the-art.

Theorem 2.1 (Inner Bound [12], [13]): For the broadcast packet erasure channel under scenario DN described in Sec. II-A, with private messages only, rate region consists of all nonnegative (R_1, R_2) satisfying

$$\frac{R_1}{1-\delta_1} + \frac{R_2}{1-\delta_{12}} \leq 1, \tag{5}$$

$$\frac{R_1}{\frac{1-\delta_2}{1-\delta_2+\delta_{12}}} + \frac{R_2}{1-\delta_2} \leq 1 \tag{6}$$

is achievable by the three-phase linear network coding aforementioned.

With private messages only, a natural capacity outer bound can be obtained by letting both receivers to feed back their channel states, that is at time t, the transmitter has homogenous global delayed CSI $(S_1^{[1:\,t-1]}, S_2^{[1:\,t-1]})$ and each receiver has homogenous CSI $(S_1^{[1:n]}, S_2^{[1:n]})$ for decoding. Then, from [1], [2], [22], we have

Theorem 2.2 (Outer Bound [1], [22]): For the broadcast packet erasure channel under scenario DN described in Sec. II-A, we have the following outer-bounds for the capacity region with private messages only

$$\frac{R_1}{1-\delta_1} + \frac{R_2}{1-\delta_{12}} \leq 1, \tag{7}$$

$$\frac{R_1}{1-\delta_{12}} + \frac{R_2}{1-\delta_2} \leq 1. \tag{8}$$

Though the boundary of the achievable rate region in Theorem 2.1 coincides partially with that of the outer bound region, namely, (7), the capacity region of scenario DN (even with only private messages) has not been characterized. Furthermore, the region in Theorem 2.1 is strictly larger than the capacity region with no feedback at all in [23],

$$\frac{R_1}{1-\delta_1} + \frac{R_2}{1-\delta_2} \leq 1, \tag{9}$$

only when

$$\frac{1-\delta_2}{1-\delta_2+\delta_{12}} > 1-\delta_1. \tag{10}$$

Thus, it is not known whether single-user delayed CSI helps when the above channel condition is violated.

III. Main Results for the Two-User Cases

We start with our results for the case with private messages only. One drawback of the opportunistic three-phase network coding reviewed in Sec. II-B is that, unlike receiver Rx1 whose interference $(\mathbf{g}_2[t_3])^\intercal \bar{\mathbf{v}}_2$ in (4) is known at decoding, Rx2 needs to jointly decode two recycled sequences $\bar{\mathbf{v}}_1$ and $\bar{\mathbf{v}}_2$. To reduce the decoding burden, we aim to improve the random-linear-combination precoding for the recycling bits such that Rxi only needs to decode $\bar{\mathbf{v}}_i,\ i = 1, 2$. As summarized in the following theorem, the improved scheme is capacity-achieving, and having rate gain over no-feedback case (9) even if channel condition of (10) is violated.

Theorem 3.1: For the broadcast packet erasure channel under scenario DN described in Sec. II-A, the capacity region

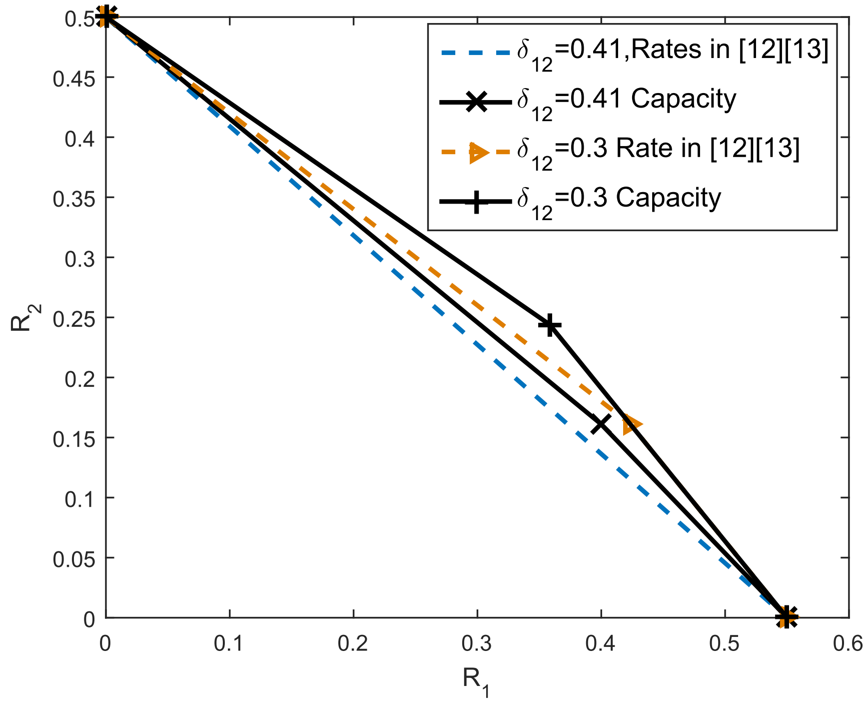

Fig. 2. With private messages only, comparisons of achievable rate regions in [12], [13] and the capacity regions under $\delta_1 = 0.45$ and $\delta_2 = 0.5$.

with private messages W_1 and W_2 only is the collection of all nonnegative (R_1, R_2) satisfying (7) and (8).

Proof: Our first main contribution is to show that outer-bounds (7) (8), obtained under the DD scenario where global delayed state feedback from both users is available, can be achieved by the an opportunistic two-phase network coding improved from [21] which solely relies on the state feedback from Rx1 only. See Section IV for the details. Without loss of generality, throughout the paper we will present the achievability under binary channel input $q = 2$. ■

Now, we extend the results of Theorem 3.1 to the scenario with an additional common message W_0 for both receivers.

Theorem 3.2: For the broadcast packet erasure channel under scenario DN, the capacity region with a common message W_0 and private messages W_1, W_2 is the collection of all nonnegative (R_1, R_2, R_0) satisfying

$$\frac{R_1 + R_0}{1-\delta_1} + \frac{R_2}{1-\delta_{12}} \leq 1 \tag{11}$$

$$\frac{R_1}{1-\delta_{12}} + \frac{R_2 + R_0}{1-\delta_2} \leq 1. \tag{12}$$

Proof: The converse is obtained from scenario DD with common and private messages and already proved in [20, Proposition 1]. The achievability for (11)(12) with delayed CSI from Rx1 only is an extension of that for Theorem 3.1, and is given in Section V. In [24], we provide the achievability proof when global delayed CSI from both receivers is available at the transmitter. Thus, the current capacity results recovers those in [24] as a subset. ■

Our result reveals that for broadcast packet erasure channels with single-user delayed CSI, even with a common message W_0, there is no rate loss compared to the capacity region with global delayed CSI from both receivers.

In Figure 2, we provide numerical examples of the private rate regions in [12], [13], and capacity regions in Theorem 3.1 for $\delta_1 = 0.45$, $\delta_2 = 0.5$, and different values of δ_{12}. Note that when $\delta_{12} = 0.41$, (10) is violated, and thus, the private rate region in [12], [13] is the capacity region with no feedback at all. With common and private messages, in Figure 3, we compare the capacity region of Theorem 3.2 for different values of δ_1 and δ_2 but fixed $\delta_{12} = 0.3$, and for given R_0's. We observe that when the common message rate R_0 is high enough, the shape of the capacity region will no longer be a tetragon but a triangle. In other words, the corner point which both users have non-zero rates will disappear for high enough R_0. Interestingly, even when the capacity region has a triangular shape, only one corner point can be achieved by time sharing between a private message with the common one while the other is achieved by the new scheme leveraging delayed CSI of Rx1. Details are given in Section V.

IV. Private Messages Only: Improved Opportunistic Network Coding for the Achievability in Theorem 3.1

In this section, we prove the achievability for Theorem 3.1. Here we first point out the main drawback of the three-phase scheme of [13]. As recapped in Sec. II-B, in this scheme, Rx1 knows the entire $\bar{\mathbf{v}}_2$ from Phase II. Therefore, in Phase III, the interference caused by the XOR precoding operation (4), $(\mathbf{g}_2[t_3])^\intercal \bar{\mathbf{v}}_2$, can be removed at Rx1 for all t_3's. As a result, equivalently, Rx1 is faced with an interference-free

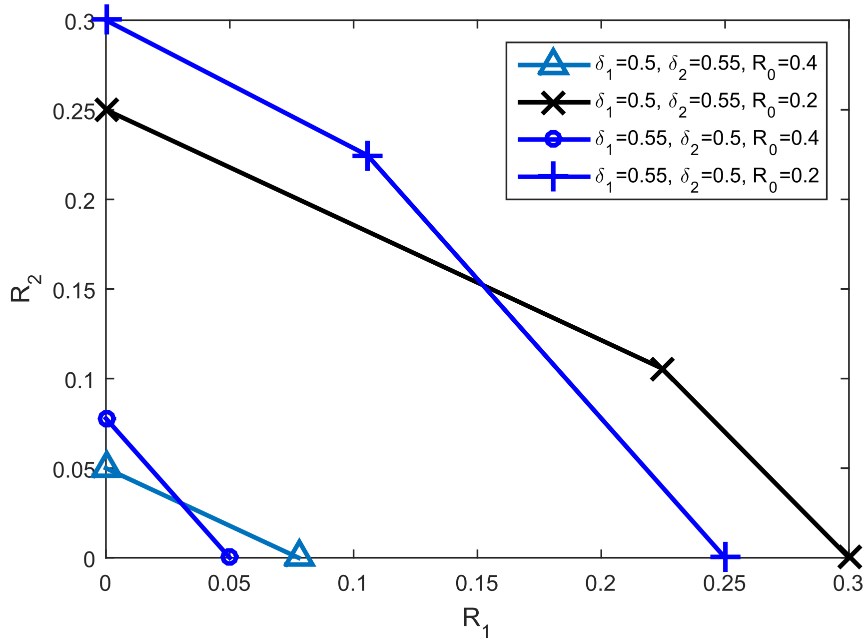

Fig. 3. With common and private messages, capacity regions of broadcast packet erasure channels with single-user delayed CSI under given common message rate R_0s.

point-to-point erasure channel in this phase. For this equivalent point-to-point erasure channel, $\bar{\mathbf{v}}_1$ is encoded into random linear combinations in [13], and hence the state feedback from Rx1 in Phase III is not utilized. However, this causes additional interference at Rx2 in Phase III. This missed opportunity in utilizing the state feedback to further mitigate interference motivates us to modify the precoding before the XOR operation according to the single-user state feedback.

The first idea to improve upon the results of [13] is to replace the random linear combinations of $\bar{\mathbf{v}}_1$ in (4) by the uncoded re-transmission of $\bar{\mathbf{v}}_1$. That is, each bit in $\bar{\mathbf{v}}_1$ is repeated according to the state feedback, and after XOR-ing a random linear combination $(\mathbf{g}_2[t_3])^\intercal\bar{\mathbf{v}}_2$, the resulting superposition is sent. Each bit of $\bar{\mathbf{v}}_1$ is repeated until the corresponding state feedback is $S_1 = 1$. In other words, prior to the superposition via XORing, $\bar{\mathbf{v}}_1$ is repeated as in standard ARQ, while $\bar{\mathbf{v}}_2$ is pre-encoded as a fountain code. The termination of the fountain code for $\bar{\mathbf{v}}_2$ is determined by the state feedback of Rx1, that is whether or not all bits in $\bar{\mathbf{v}}_1$ are successfully delivered to Rx1. Compared with (4), the encoder replaces the randomly-generated basis $\mathbf{g}_1[t_3]$ for $\bar{\mathbf{v}}_1$ with a repetition of a standard basis. The repetition of each standard basis of $\mathbb{F}_2^{1\times|\bar{\mathbf{v}}_1|}$ is controlled by the state feedback of S_1, where $|\bar{\mathbf{v}}_1|$ is the length of $\bar{\mathbf{v}}_1$.

Next, we argue that with the ARQ-type retransmission in the last phase described above, from the viewpoint of user 1, there is no need to separate the recycled $\bar{\mathbf{v}}_1$ from those bits sent in Phase I of [13]. The reason is that if the interference from user 2 can be removed, with only the state feedback of S_1, ARQ-type retransmission is optimal for user 1. Without loss of the optimality, transmission in Phase I of [13] can also be replaced by the ARQ-type one, and then one can merge this new Phase I into the ARQ used in the last phase.

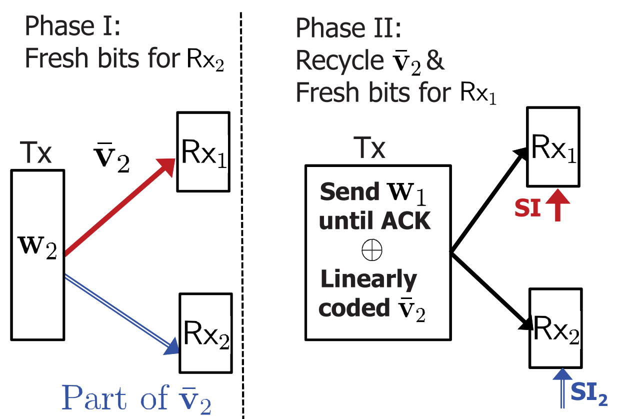

Fig. 4. Capacity achieving opportunistic two-phase network coding for the case with private messages only, under binary channel input.

Finally, we arrive at a two-phase scheme in which the first phase is the same as Phase II of [13] and the transmitter will send coded bits for user 2; while in the second phase, the transmitter will simultaneously send fresh bits of user 1 using ARQ and recycle $\bar{\mathbf{v}}_2$ for user 2. In other words, in the new re-transmission, prior to the superposition via XORing, $\bar{\mathbf{w}}_1$ is repeated as in standard ARQ, while $\bar{\mathbf{v}}_2$ is pre-encoded as a fountain code. As will shown below, these changes significantly improve the achievable rates.

A. Encoder of the New Two-Phase Opportunistic Network Coding

Our new two-phase opportunistic network coding, as depicted in Figure 4, has encoding process comes as follows

Phase I: This phase is the same as Phase II in the three-phase scheme of [13]. Using $\mathbf{g}_2[t_1] \in \mathbb{F}_2^{R_2 n\times 1}$ with each

Algorithm 1 Phase II of the Opportunistic Two-Phase Network Coding in Figure 4

```
1: Initial Set time index t_2 = n_1 + 1 (beginning of Phase II)
2: for i = 1 to R_1 n do
3:   while w_{1,i} is not received by Rx1 do
4:     I: Pre-Encoder (PENC) 1 for w_1
5:       Output the i-th bit w_{1,i} of the input w_1 at t_2
6:     II: Pre-Encoder (PENC) 2 for v̄_2
7:       Output a new linear combination (g_2[t_2])^T v̄_2 of the input v̄_2 at t_2
8:     III: Superposition
9:       Send the XOR of outputs of PENC 1 and 2 at t_2
10:    Increase time index t_2 by 1
11:  end while
12: end for
```

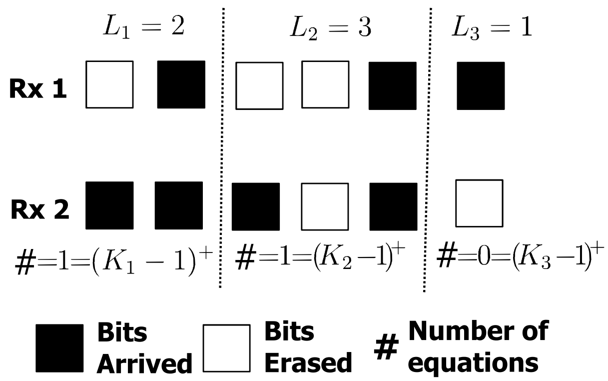

Fig. 5. Example for obtaining pure linear equations of recycled $\bar{\mathbf{v}}_2$ at receiver Rx2 in Phase II of Figure 4, where the number of equations obtained during L_i repetitions of user 1's i-th recycled bit can be calculated with K_i in (13).

entry generated from i.i.d. $\mathrm{Ber}(1/2)$, the transmitter sends coded bits $X[t_1] = (\mathbf{g}_2[t_1])^\intercal \mathbf{w}_2$, $0 \le t_1 \le n_1$, aimed for Rx2. After Phase I, the transmitter knows length $n_1(1-\delta_1)$ recycled sequence $\bar{\mathbf{v}}_2$, which is formed by bits $(\mathbf{g}_2[t_1])^\intercal \mathbf{w}_2$ received at Rx1 in Phase I where $S_1[t_1] = 1$.

Phase II: In this phase, all (fresh) bits of message $\mathbf{w}_1$ for user 1 are repeated according to the state feedback as in the standard ARQ, and after XORing a random linear combination of $\bar{\mathbf{v}}_2$, the resulting superposition is sent. Details are given in Algorithm 1. As described in Algorithm 1, at time index t_2, the i-th bit $\bar{w}_{1,i}$ in $\bar{\mathbf{w}}_1$ is repeated according to the state feedback (lines 3,4,5), and after XORing a random linear combination $(\mathbf{g}_2[t_2])^\intercal \bar{\mathbf{v}}_2$, the resulting superposition is sent (lines 8,9). Each $\bar{w}_{1,i}$ is repeated until the corresponding state feedback is $S_1 = 1$. For easier rate analysis, after the transmitter finishes all bits in $\mathbf{w}_1$, it may send some random linear combinations of $\bar{\mathbf{v}}_2$ without XOR to ensure decodability.

B. Decoding

As for decoding, let us first focus on Rx1. As shown in Figure 4, the side information $\bar{\mathbf{v}}_2$ (marked in red) is known at Rx1 during Phase I, and then the decoding of private $\bar{\mathbf{w}}_1$ at Rx1 is straightforward with large enough n_2 after removing the interference from $\bar{\mathbf{v}}_2$.

For Rx2, it first decodes the recycled sequence $\bar{\mathbf{v}}_2$, and then uses it together with other linear equations arrived during Phase I to decode $\mathbf{w}_2$. However, unlike the side information at Rx1, the transmitter does not know the side information that Rx2 will have for decoding the linearly-precoded $\bar{\mathbf{v}}_2$. It only knows that Rx2 has a random fraction of and $\bar{\mathbf{v}}_2$ respectively during Phase I (marked in blue in Figure 4). If one follow the idea of [13] as recapped in Sec. II-B, Rx2 has to jointly decode the interference $\mathbf{w}_1$ and the desired sequence $\bar{\mathbf{v}}_2$. However, if one can avoid decoding the interference, the length needed for successful decoding can be reduced.

The key to the improvement in our new scheme over [13] is that, in Phase III, since each bit of $\mathbf{w}_1$ is repeated until it successfully arrives at Rx1, such repetition automatically *aligns* the interference at Rx2. Hence, instead of jointly decoding $\mathbf{w}_1$ and $\bar{\mathbf{v}}_2$, Rx2 can opportunistically obtain *pure* linear equations of $\bar{\mathbf{v}}_2$. To be more specific, consider the i-th bit $w_{1,i}$ of recycled sequence $\mathbf{w}_1$. Suppose it is repeated L_i times until its mixture with $\bar{\mathbf{v}}_2$ is successfully arrived at Rx1 in Phase III. Within this span, Rx2 gets

$$K_i \triangleq \sum_{\ell=1}^{L_i} S_{2,i}[\ell] \tag{13}$$

linear equations mixing $w_{1,i}$ and $\bar{\mathbf{v}}_2$, where $S_{2,i}[\ell]$ is the erasure state at Rx2 for the ℓ-th transmission of $w_{1,i}$. Interestingly, even though the interference $w_{1,i}$ is unknown by Rx2, it still has a chance to get pure linear equations of $\bar{\mathbf{v}}_2$ using *interference alignment*. During the retransmissions of $w_{1,i}$, if there are two time slots where $S_{2,i}[\ell]$ are both 1, the XOR of the corresponding received bits at Rx2 yields a pure equation of $\bar{\mathbf{v}}_2$, since the interfering bits cancel themselves. In this case, we get $(K_i - 1)^+$ pure equations, where $(x)^+ \equiv \max\{x, 0\}$.

Let us use an example to illustrate the decoding at Rx2. Suppose there are three bits $\{w_{1,i} \mid i = 1, \ldots, 3\}$ in $\mathbf{w}_1$, and the realization of the channel states are depicted in Figure 5. Bit $w_{1,1}$ and $w_{1,2}$ are repeated $L_1 = 2$ and $L_2 = 3$ times respectively in Phase II. After the processing at Rx2 as mentioned above, one obtains $(K_1 - 1)^+ = 1$ and $(K_2 - 1)^+ = 1$ pure linear equations at Rx2. Meanwhile, bit $w_{1,3}$ is repeated only $L_3 = 1$ time and one obtains $(K_3 - 1)^+ = 0$ pure equations. Hence, we have two pure equations of $\bar{\mathbf{v}}_2$ in total.

C. Rate Analysis

In what follows, we first compute the expected number of linear equations for decoding the messages $\mathbf{w}_1$ and $\mathbf{w}_2$ at Rx1 and Rx2 respectively. Then, we derive the sufficient conditions for the achievable rate pair (R_1, R_2) to prove the achievability part of Theorem 3.1. In the following, we define the normalized lengthes for Phase $i, i = 1, 2$ as

$$\alpha_i := n_i / n. \tag{14}$$

As for Rx1, for the decodability of $\mathbf{w}_1$, as long as the normalized length α_2 of Phase II meets

$$\begin{aligned} \alpha_2 n &> (R_1 n)\mathbb{E}[L_i] \\ &= \frac{R_1 n}{1-\delta_1}, \end{aligned} \tag{15}$$

then, with high probability, $\mathbf{w}_1$ can be reconstructed with long enough $R_1 n$. Here, the last equality is due to the fact that L_i follows a Geometric distribution, and $\mathbb{E}[L_i] = 1/(1-\delta_1)$.

As for Rx2, first consider the decodability of $\bar{\mathbf{v}}_2$. On one hand, some bits of $\bar{\mathbf{v}}_2$ also arrive at Rx2 in Phase I, as the side-information SI_2 for Phase II in Figure 4. It should be pointed out that SI_2 is the side information in terms of decoding $\bar{\mathbf{v}}_2$ from the received bits in Phase II, but it actually carries message W_2 intended for user 2. The expected number of them is

$$\alpha_1 n \Pr\{S_1 = S_2 = 1\} = \alpha_1 n(1-\delta_1-\delta_2+\delta_{12}). \quad (16)$$

Note that when the time index t is not important, we will neglect it and write $S_i[t]$ as $S_i, i = 1, 2$. On the other hand, as mentioned in previous Section IV-B, one can form pure linear equations of $\bar{\mathbf{v}}_2$. The expected number of pure equations of $\bar{\mathbf{v}}_2$ is

$$(R_1 n)\mathbb{E}[(K_i - 1)^+] \quad (17)$$

Note that there is no conditioning in the expectation because the states are memoryless. Next, from Appendix A, we have

$$\mathbb{E}[(K_i-1)^+] = \frac{\delta_1-\delta_2}{1-\delta_1} + \frac{\delta_2-\delta_{12}}{1-\delta_{12}}. \quad (18)$$

Then for Rx2, we obtain the expected number of additional linear equations of $\bar{\mathbf{v}}_2$ produced in Phase II

$$(R_1 n)\left(\frac{\delta_1-\delta_2}{1-\delta_1} + \frac{\delta_2-\delta_{12}}{1-\delta_{12}}\right). \quad (19)$$

Note that due to the memoryless property of the channel erasure state, the number of additional linear equations produced in Phase II will concentrate tightly around the expected value in (19). This is due to the fact that $K_i, i = 1, \ldots R_1 n$ are i.i.d, and from the law of large numbers $\sum (K_i - 1)^+/R_1 n$ will converge to its mean $\mathbb{E}[(K_i-1)^+]$ in (18) almost surely. Such a observation applies to other parts of the proof.

Thus, for decodability of $\bar{\mathbf{v}}_2$, with aids of length $\alpha_1 n(1-\delta_1-\delta_2+\delta_{12})$ side-information SI_2 for $\bar{\mathbf{v}}_2$, it suffices to have

$$\begin{aligned}&\alpha_1 n(1-\delta_1) < \\ &R_1 n\left(\frac{\delta_1-\delta_2}{1-\delta_1} + \frac{\delta_2-\delta_{12}}{1-\delta_{12}}\right) + \alpha_1 n(1-\delta_1-\delta_2+\delta_{12}) \\ &+ (1-\delta_2)\left(\alpha_2 n - \frac{R_1 n}{1-\delta_2}\right)\end{aligned} \quad (20)$$

as n is long enough. The last term $(1-\delta_2)(\alpha_2 n - R_1 n/(1-\delta_2))$ comes from pure random linear equations of $\bar{\mathbf{v}}_2$ sent after the transmitter finishes $\mathbf{w}_1$. Hence, the total expected number of equations for decoding $\mathbf{w}_2$ is $\alpha_1 n(1-\delta_{12})$, and for its decodability, it suffices to have

$$R_2 < \frac{\alpha_1(1-\delta_{12})}{\alpha_1+\alpha_2} \quad (21)$$

Remark 4.1 (Linear Independence): For (20), by collecting linear equations of $\bar{\mathbf{v}}_2$ received in Phase II (with standard basis) and the ones produced in Phase III, one can form a set of linear equations of $\bar{\mathbf{v}}_2$ described by a full (column) rank matrix, when $n \to \infty$. Then $\bar{\mathbf{v}}_2$ can be decoded. With linear equations of message $\mathbf{w}_2$ generated in Phase I, the full rank property to validate (21) is also ensured by letting $n \to \infty$.

Finally, by selecting

$$\alpha_1 = R_1 \frac{\frac{\delta_1-\delta_2}{1-\delta_1} + \frac{\delta_2-\delta_{12}}{1-\delta_{12}}}{\delta_2-\delta_{12}}$$

will validate the decodability (20). Furthermore, one can rewrite α_1 as

$$\alpha_1 = R_1 \frac{(1-\delta_2)}{(\delta_2-\delta_{12})} \frac{(\delta_1-\delta_{12})}{(1-\delta_1)(1-\delta_{12})}, \quad (22)$$

then from $\alpha_1 + \alpha_2 = 1$ and (15),

$$R_1 \frac{(1-\delta_2)}{(\delta_2-\delta_{12})} \frac{(\delta_1-\delta_{12})}{(1-\delta_1)(1-\delta_{12})} + \frac{R_1}{1-\delta_1} < 1. \quad (23)$$

Also from (22) and (15)

$$\frac{\alpha_1}{\alpha_1+\alpha_2} < \frac{\frac{(1-\delta_1)(1-\delta_2)}{\delta_2-\delta_{12}}\left(\frac{\delta_1}{1-\delta_1} - \frac{\delta_{12}}{1-\delta_{12}}\right)}{1 + \frac{(1-\delta_1)(1-\delta_2)}{\delta_2-\delta_{12}}\left(\frac{\delta_1}{1-\delta_1} - \frac{\delta_{12}}{1-\delta_{12}}\right)}. \quad (24)$$

From (23)(24), together with (21), one can show that the following rate pair is achievable:

$$\begin{aligned}R_1 &< \frac{1-\delta_1}{1 + \frac{(1-\delta_1)(1-\delta_2)}{\delta_2-\delta_{12}}\left(\frac{\delta_1}{1-\delta_1} - \frac{\delta_{12}}{1-\delta_{12}}\right)} \\ &= \frac{(1-\delta_2)^{-1} - (1-\delta_{12})^{-1}}{(1-\delta_1)^{-1}(1-\delta_2)^{-1} - (1-\delta_{12})^{-2}},\end{aligned} \quad (25)$$

$$\begin{aligned}R_2 &< (1-\delta_{12}) \frac{\frac{(1-\delta_1)(1-\delta_2)}{\delta_2-\delta_{12}}\left(\frac{\delta_1}{1-\delta_1} - \frac{\delta_{12}}{1-\delta_{12}}\right)}{1 + \frac{(1-\delta_1)(1-\delta_2)}{\delta_2-\delta_{12}}\left(\frac{\delta_1}{1-\delta_1} - \frac{\delta_{12}}{1-\delta_{12}}\right)} \\ &= \frac{(1-\delta_1)^{-1} - (1-\delta_{12})^{-1}}{(1-\delta_1)^{-1}(1-\delta_2)^{-1} - (1-\delta_{12})^{-2}},\end{aligned} \quad (26)$$

The RHSs of (25) and (26) forms the intersection of (7) and (8). Also the rate pair (R_1, R_2) can be arbitrary close to this intersection by choosing $\delta = \alpha_2 - R_1/(1-\delta_1)$ small enough for the normalized length of Phase II (see (15)), when n is large enough.

The other two corner points of $(R_1, R_2), (1-\delta_1, 0)$ and $(0, 1-\delta_2)$, can both be trivially achieved. Then all other achievable rate-points in the region (7)(8) can be obtained by time-sharing. This proves the achievability part of Theorem 3.1.

Remark 4.2: Besides the proposed scheme, another three-phase scheme in our conference version [21] can also achieve the capacity region. The first two phases are identical to those in Sec. II-B while in the third phase recycled $\bar{\mathbf{v}}_1$ is repeated as in standard ARQ and $\bar{\mathbf{v}}_2$ is pre-encoded using a random linear code. These coded bits are then XORed to generate bits to be sent. Indeed, our two-phase scheme comes from merging the Phase I and III of the three-phase one in [21]. The achievability analysis is similar and neglected here.

D. Comparison With Other Works

As discussed in [25, Section VIII-A], capacity of BC with single-user feedback, is unknown for years even without the common message. To the best of authors' knowledge, our result is the first feedback capacity result belonging to this category. Prior to us, only the sum DoF, which approximates the sum rate at high signal-to-noise ratio (SNR), in MISO BC with single-user CSI feedback was identified by Davoodi and Jafar [5]. Moreover, [5] shows that with continuous channel states, single-user CSI feedback does not help at all in terms of sum DoF. In sharp contrast to this recent negative result, our results not only shows that single-user CSI feedback helps in broadcast PEC of which the channel state is discrete, but also the capacity region is the same as that with full two-user feedback. Then the feedback overhead can be significantly reduced. This is the fundamental new technical insight delivered in this paper.

The huge difference between our result and [5] further supports the observation by Shayevitz and Wigger [14], that is, unlike multiple-access channel (MAC), finding a general feedback capacity formula which works for many memoryless BCs is very hard, especially when they are not physically-degraded. The capacities may diverge a lot for different channel settings which makes this problem open for many years. For other channels such as Gaussian BCs, the capacities with single-user feedback are still unknown [25]. Target on Gaussian BC with *global* feedback, the Schalkwijk-Kailath scheme, which is a special case of posterior-matching scheme [26], was generalized by Ozrow and Leung-Yan-Cheong [27]. However, the rate in [27] neither matches the outer-bound nor is always larger than the capacity without feedback. The achievability in [27] is enhanced by [25], and though the rates do not match the Ozrow-type outer bound [27] they are the best known achievability. The linear coding scheme in [25] is developed from the non-feedback duality of the Gaussian MAC and BC. However, this technique is highly specialized for Gaussian BCs and hard to be applied to broadcast PECs.

Finally, the block Markov coding proposed in [16], which includes many other known results such as the Weldon-type scheme for binary symmetric channel [26] as special cases, may potentially also achieve our identified capacity. The issue is that one needs to characterize the complicated auxiliary random variables to maximize its rate, and our achievability may help to identify the optimal ones. However, without the common message, similar direction was already explored in [12] but the reported rate, detailed in Theorem 2.1, is lower than ours. Moreover, even if this difficulty is solved, the typicality-based nonlinear coding in [16] will have much higher complexity and be less explicit than our linear scheme.

V. Common and Private Messages: Achievability Proof of Theorem 3.2

To better illustrate the challenge in developing the achievability when an additional common message is included, we first discuss a simple extension from the scheme with private messages only. We introduce another phase multicasting nR_0 common message bits with length $nR_0/(1-\delta)$, or rate $\min\{1-\delta_1, 1-\delta_2\} = 1-\delta$, before transmitting two private messages with the scheme in Section IV via $n - nR_0/(1-\delta)$ symbols. We can achieve rate region from Theorem 3.1 which matches the converse when $\delta_1 = \delta_2 = \delta$ (can also see from upcoming Lemma 6.1). Note that we do not need any feedback from S_2 for the common-message transmission, and then identify the capacity region for scenario DN with common and private messages under the aforementioned channel parameters. Unfortunately, when $\delta_1 \neq \delta_2$ and two links have different erasure probabilities, we must develop more advanced transmission schemes. For example, when $\delta_1 < \delta_2$, naively adopting the aforementioned scheme by transmitting the common message with rate $\min\{1-\delta_1, 1-\delta_2\} = 1-\delta_2$ in the first phase is no longer capacity-achieving.

Now we show how to modify the aforementioned three-phase scheme to achieve the capacity when $\delta_1 \neq \delta_2$. For the $R_0 n$ common message bits in $\mathbf{w}_0$, the transmitter first generates total

$$\frac{R_0 n}{1-\delta_{12}} + R_0 n \frac{\delta_1 - \delta_{12}}{1-\delta_{12}} \tag{27}$$

linearly coded bits from them, using coding vectors in $\mathbb{F}_2^{R_0 n \times 1}$ with each entry generated from i.i.d. $\mathrm{Ber}(1/2)$. The transmitter sends only the first half in Phase I, and Phase II is the same as Phase I in Figure 4. The transmissions in Phase III are modified from those in Phase II of Figure 4 by replacing $\bar{\mathbf{v}}_2$ and $\mathbf{w_1}$ respectively with $\bar{\mathbf{v}}_2^{(\text{new})}$ and $\mathbf{w}_1^{(\text{new})}$. Both new $\bar{\mathbf{v}}_2^{(\text{new})}$ and $\mathbf{w}_1^{(\text{new})}$ contain equations of common message, and the definitions of them will be revealed later. Recall now the private message vector $\mathbf{w}_i$ for user i has $R_i n$ bits, the detailed encoder is

Phase I: The transmitter sends the first half in (27), $R_0 n/(1-\delta_{12})$ coded bits of $\mathbf{w}_0$, using

$$n_1 = R_0 n/(1-\delta_{12}) \tag{28}$$

time slots. After Phase I, the transmitter knows length $n_1(1-\delta_1)$ recycled sequence $\bar{\mathbf{v}}_2^{[0]}$, which is formed by coded bits received at Rx1 in Phase I where $S_1[t_1] = 1, 1 \leq t_1 \leq n_1$. We also denote the second half of the $R_0 n \frac{\delta_1 - \delta_{12}}{1-\delta_{12}}$ coded bits of $\mathbf{w}_0$, which has not been transmitted in Phase I, by $\bar{\mathbf{v}}_1^{[0]}$.

Phase II: As Phase I in of Figure 4, for $R_2 n$ private message bits for user 2, the transmitter generate

$$n_2 = \frac{R_2 n}{1-\delta_{12}} \tag{29}$$

randomly linearly coded bits of $\mathbf{w}_2$ and transmit them in n_2 time slots. After Phase II, we also denote those bits that have been received by user 1 as $\bar{\mathbf{v}}_2$. Now the transmitter groups it and common $\bar{\mathbf{v}}_2^{[0]}$ as a new super-message for user 2

$$\bar{\mathbf{v}}_2^{(\text{new})} := (\bar{\mathbf{v}}_2^{[0]}, \bar{\mathbf{v}}_2), \tag{30}$$

also it forms a new super-message for user 1

$$\mathbf{w}_1^{(\text{new})} := (\bar{\mathbf{v}}_1^{[0]}, \mathbf{w}_1). \tag{31}$$

Phase III: The transmission is an extension of that in Algorithm 1 by replacing $\mathbf{w}_1$ with $\mathbf{w}_1^{(\text{new})}$ and the input of the Pre-Encoder 2 from $\bar{\mathbf{v}}_2$ to $\bar{\mathbf{v}}_2^{(\text{new})}$. The output of Pre-Encoder 2 at time t_3 becomes the random linear combinations $(\mathbf{g}^{(\text{new})}[t_3])^\intercal \bar{\mathbf{v}}_2^{(\text{new})}$, where each entry of $\mathbf{g}^{(\text{new})}[t_3] \in \mathbb{F}_2^{((n_1+n_2)(1-\delta_1))\times 1}$ is generated from i.i.d. $\mathrm{Ber}(1/2)$. Fresh bits of $\mathbf{w}_1^{(\text{new})}$ are repeated according to the delayed S_1 as described in Pre-Encoder 1, and the XOR of outputs of Pre-Encoder 1 and 2 is sent. Here, $n_1 + n_2 + 1 \leq t_3 \leq n$.

A. Decoding and Rate Analysis

Here we only present the case where both outer bounds (11)(12) are active since the achievability with when one of them is active is trivial. Receiver Rx1 can decode both the common and its own private information within n time slots from (28)(29) if

$$\frac{R_0 n}{(1-\delta_{12})}+\frac{R_2 n}{1-\delta_{12}}+\frac{1}{1-\delta_1}\left(R_0 n\frac{\delta_1-\delta_{12}}{1-\delta_{12}}+R_1 n\right)<n, \tag{32}$$

that is, n must be selected to ensure all bits of $\mathbf{w}_1^{(\text{new})}$ in (31) are delivered to Rx1 via ARQ, with aids of side-information $\bar{\mathbf{v}}_2^{(\text{new})}$. This decodability is ensured from (11).

Now, we turn to the decodability at Rx2. Receiver Rx2 will first decode the super-message $\bar{\mathbf{v}}_2^{(\text{new})}$, from its received bits during the entire three phases. With long enough blocklength for Phase III, during the ARQ transmissions of $\bar{\mathbf{v}}_1^{[0]}$ in (31), Rx2 gets

$$R_0 n\frac{\delta_1-\delta_{12}}{1-\delta_{12}}\frac{1-\delta_2}{1-\delta_1} \tag{33}$$

equations. Since Rx2 is also interested in $\bar{\mathbf{v}}_1^{[0]}$ (generated from the common $\mathbf{w}_0$), every reception of Rx2 is useful during this duration. To see this, we use the following example. Assume the first coded common message bit $\bar{v}_{1,1}^{[0]}$ from $\bar{\mathbf{v}}_1^{[0]}$ is repeated twice until ACK from Rx1. In the meanwhile, Rx2 is always "on" during this retransmission, and has two received bits

$$\bar{v}_{1,1}^{[0]}\oplus(\mathbf{g}^{(\text{new})}[1])^{\intercal}\bar{\mathbf{v}}_2^{(new)} \text{ and } \bar{v}_{1,1}^{[0]}\oplus(\mathbf{g}^{(\text{new})}[2])^{\intercal}\bar{\mathbf{v}}_2^{(new)}, \tag{34}$$

where $\mathbf{g}^{(\text{new})}[1], \mathbf{g}^{(\text{new})}[2]$ are i.i.d. generated random linear coding vectors. We will show that equations from these two received bits are linearly independent. Note that after Phase I, $n_1(1-\delta_2)$ linear equations of $\mathbf{w}_0$ are received at Rx2 and there is $n_1(\delta_2-\delta_{12})$ linear equations of $\mathbf{w}_0$ not received at Rx2 but at Rx1. The former is denoted as $\mathbf{w}'_{0,\mathsf{R2}}$ and the latter as $\mathbf{w}'_{0,\mathsf{R1}\bar{2}}$. From (28), indeed $(\mathbf{w}'_{0,\mathsf{R2}}, \mathbf{w}'_{0,\mathsf{R1}\bar{2}})$ are representations of original common $\mathbf{w}_0$ with a different set of coordinates and can be rewritten as

$$\left((\mathbf{w}'_{0,\mathsf{R2}})^{\intercal},(\mathbf{w}'_{0,\mathsf{R1}\bar{2}})^{\intercal}\right)^{\intercal}=\mathbf{G}'\mathbf{w}_0 \tag{35}$$

where $\mathbf{G}'\in\mathbb{F}_2^{R_0 n\times R_0 n}$ is invertible and known at Rx2 through the feedback. Note that length $n_1(1-\delta_1)$ $\bar{\mathbf{v}}_2^{[0]}$ consists of $\mathbf{w}'_{0,\mathsf{R1}\bar{2}}$ and $n_1(1-\delta_1-\delta_2+\delta_{12})$ bits in $\mathbf{w}'_{0,\mathsf{R2}}$. Now Rx2 removes effects of side-information $\mathbf{w}'_{0,\mathsf{R2}}$ from two bits in (34), and has equations

$$\begin{aligned}&(\bar{\mathbf{g}}_{1,1}^{[0]})^{\intercal}\mathbf{w}'_{0,\mathsf{R1}\bar{2}}\oplus(\mathbf{g}^{(\text{new})}[1])^{\intercal}(\mathbf{w}'_{0,\mathsf{R1}\bar{2}},\bar{\mathbf{v}}_2)\\ \text{and } &(\bar{\mathbf{g}}_{1,1}^{[0]})^{\intercal}\mathbf{w}'_{0,\mathsf{R1}\bar{2}}\oplus(\mathbf{g}^{(\text{new})}[2])^{\intercal}(\mathbf{w}'_{0,\mathsf{R1}\bar{2}},\bar{\mathbf{v}}_2),\end{aligned} \tag{36}$$

where $\bar{\mathbf{g}}_{1,1}^{[0]}$ is the equivalent coding vector after removing the effects of $\mathbf{w}'_{0,\mathsf{R2}}$ in $\bar{v}_{1,1}^{[0]}$ using $(\mathbf{G}')^{-1}$. Since elements of $\mathbf{g}^{(\text{new})}[1]$ and $\mathbf{g}^{(\text{new})}[2]$ are generated from i.i.d. $\text{Ber}(1/2)$, it is easy to see that we have two linearly independent equations of $(\mathbf{w}'_{0,\mathsf{R1}\bar{2}},\bar{\mathbf{v}}_2)$ (also of $\bar{\mathbf{v}}_2^{(\text{new})}$ with side-information from Phase I) with close-to-one probability. Furthermore, by applying proposed interference alignment in Section IV-B on bits received in Phase III during the transmissions of $\mathbf{w_1}$ in (31), Rx2 can get pure linear equations of $\bar{\mathbf{v}}_2^{(\text{new})}$. The expected number for these equations is

$$\begin{aligned}R_1 n\mathbb{E}[(K_i-1)^+]&=R_1 n\left(\frac{\delta_1-\delta_2}{1-\delta_1}+\frac{\delta_2-\delta_{12}}{1-\delta_{12}}\right)\\&=R_1 n\frac{(1-\delta_2)(\delta_1-\delta_{12})}{(1-\delta_1)(1-\delta_{12})},\end{aligned} \tag{37}$$

where the first equality comes from (18).

Finally, by using bits of $\bar{\mathbf{v}}_2^{(\text{new})}$ in (30) received in Phase I and II as additional decoder side-information, together with (37) and (33), we need (38), shown at the bottom of the next page, for successful decoding the length $(n_1+n_2)(1-\delta_1)$ super-message $\bar{\mathbf{v}}_2^{(\text{new})}$, as the codewords are long enough. Note that as in Section IV, after the transmitter finishes all bits in $\mathbf{w}_1^{(\text{new})}$, the transmitter will send random linear combinations of $\bar{\mathbf{v}}_2^{(\text{new})}$ without XOR, which corresponds to the last term in (38). We show in Appendix B that constraint (38) will be met under (11)(12). Then similar to the reasons in Remark 4.1, the ranks to decode $\bar{\mathbf{v}}_2^{(\text{new})}$ will be enough with high probability. Now with $\bar{\mathbf{v}}_2^{[0]}$ from (30) and side-information in Phase I, Rx2 knows $\mathbf{w}'_{0,\mathsf{R1}\bar{2}}$ in (35) can recover common message $\mathbf{w}_0$ using inverse $(\mathbf{G}')^{-1}$. With $\bar{\mathbf{v}}_2$ from (30), the decodability of private message $\mathbf{w}_2$ can be also ensured as in the previous Section.

VI. Extension to More Than One "N" Users

Here we discuss the extension to the channel with more than one "N" users. For simplicity, we consider a three-user case of which only one user has delayed erasure state feedback, that is, scenario DNN. Now besides the two receiver described in (1), we have an additional receiver Rx3 asking a rate R_3 private message W_3 and the common message W_0, with received signal being $Y_3[t]=S_3[t]X[t]$ where $\{S_3[t]\}$ is the i.i.d. Bernoulli $(1-\delta_3)$ process. Since only Rx1 feeds back its state the constraint imposed at the encoding function $f_t(.)$ at time index t is

$$X[t]=f_t\left(W_0,W_1,W_2,W_3,S_1^{[1:\,t-1]}\right). \tag{39}$$

Besides (3), the error probability constraints at Rx3 is similar to that of Rx2 as

$$\Pr\left\{(W_3,W_0)\neq g_3(Y_3^{[1:n]},S_1^{[1:\,n]},S_3^{[1:\,n]})\right\}\to 0, \tag{40}$$

as $n\to\infty$, where $g_3(.)$ is the decoding function at receiver 3.

First, we extend the outer bound for Theorem 3.2 as

Lemma 6.1: For the broadcast packet erasure channel under scenario DNN, we have the following outer-bounds for the capacity region with a common message W_0 and private messages W_1, W_2, W_3

$$\frac{R_0+R_1}{1-\delta_1}+\frac{R_2+R_3}{1-\delta_1\delta_2}\leq 1 \tag{41}$$

$$\frac{R_1}{1-\delta_1\delta_2}+\frac{R_0+R_2+R_3}{1-\delta_2}\leq 1. \tag{42}$$

when the two "N" receiver have same erasure probability $\delta_3=\delta_2$ and all three links are independent.

Proof: The proof is given in Appendix C. ■

Corollary 6.1: For the broadcast packet erasure channel under scenario DNN, the capacity region with a common message W_0 and private messages W_1, W_2, W_3 is the collection of all nonnegative (R_1, R_2, R_3, R_0) satisfying (41)(42) when (i) $R_0 = 0$ (ii) all links have the same erasure probabilities.

Proof: For the first case where $R_0 = 0$, the outer-bound in Lemma 6.1 can be treated as splitting the rate of "N" user in Scenario DN (7)(8) into R_2 and R_3. Then one can split the first phase in Figure 4 into two segments, one for transmitting the coded bits for user 2 while the other for transmitting those for user 3. The ratio between lengths of these two segments is R_2/R_3. The second phase is the same as Figure 4. The rate analysis is similar to those in Section IV and detailed in [28].

For the second case where $R_0 \neq 0, \delta_1 = \delta_2 = \delta$. The achievability is similar to those in the beginning of Section V. Now the outer-bound in Lemma 6.1 becomes

$$\frac{R_1}{1-\delta} + \frac{R_2+R_3}{1-\delta^2} \leq 1 - \frac{R_0}{1-\delta}$$
$$\frac{R_1}{1-\delta^2} + \frac{R_2+R_3}{1-\delta} \leq 1 - \frac{R_0}{1-\delta}$$

A simple three-phase extension of the scheme described in the previous paragraph can achieve this region, that is, we add a first phase that will multicast common message bits with rate $R_0/(1-\delta)$ to the three receivers. It concludes the proof. ■

As a final note, the capacity region in Theorem 6.1 can be simply extended to the case where there is only one "D" user and arbitrary number of same erasure probability "N" users.

VII. Conclusion

We characterized the capacity region of the two-user broadcast PEC with single-user delayed CSI. More precisely, we assumed "DN" hybrid CSIT setting where the transmitter has the delayed CSI of one user but not the other. Compared with previous results, we developed an improved achievability strategy and showed that the capacity region, surprisingly, matches that of the broadcast PEC with global delayed CSI of both users. The key to such improvement over previous results was a new precoding strategy for the retransmission phase of the opportunistic network coding scheme. It leveraged the single-user delayed CSI in the retransmission phase, so that interference from the "D" receiver could be aligned at the "N" receiver. Besides the broadcast PEC with two private messages, extension with an additional common message was also provided and the corresponding capacity region under DN setting also matched that with global delayed CSI. Furthermore, three-user extensions are also provided.

Appendix

A. Computations of $\mathbb{E}[(K_i - 1)^+]$ in (18)

To warmup we calculate $\mathbb{E}[K_i]$ first, and then use it to find $\mathbb{E}[(K_i-1)^+]$. From the definition of K_i in (13), we have

$$\mathbb{E}[K_i] = \mathbb{E}[\mathbb{E}[K_i|L_i]] = \mathbb{E}\left[\left[\sum_{\ell=1}^{L_i} S_{2,i}[\ell] \middle| L_i\right]\right].$$

Note that the following two events are equivalent:

$$\{L_i = l\} \equiv \{S_{1,i}[1] = \ldots = S_{1,i}[l-1] = 0 \text{ and } S_{1,i}[l] = 1\},$$

where erasure state $S_{1,i}[\ell]$ at Rx1 is defined similarly as $S_{2,i}[\ell]$. Furthermore, given this event, $\{S_{2,i}[1], \ldots, S_{2,i}[l]\}$ are independent Bernoulli random variables, and the first $(l-1)$ are Bernoulli with parameter

$$\Pr\{S_2 = 1|S_1 = 0\} = \frac{\delta_1 - \delta_{12}}{\delta_1}$$

while the last one is Bernoulli with parameter

$$\Pr\{S_2 = 1|S_1 = 1\} = \frac{1-\delta_1-\delta_2+\delta_{12}}{1-\delta_1}.$$

Hence,

$$\mathbb{E}[K_i] = \mathbb{E}[L_i - 1]\frac{\delta_1-\delta_{12}}{\delta_1} + \frac{1-\delta_1-\delta_2+\delta_{12}}{1-\delta_1} = \frac{1-\delta_2}{1-\delta_1}, \quad (43)$$

where

$$\mathbb{E}[L_i] = 1/(1-\delta_1)$$

is applied. Similarly,

$$\mathbb{E}[(K_i-1)^+] = \mathbb{E}[\mathbb{E}[(K_i-1)^+|L_i]]$$
$$= \mathbb{E}[K_i - 1] + \mathbb{E}\left[(\frac{\delta_{12}}{\delta_1})^{L_i-1}\frac{\delta_2-\delta_{12}}{1-\delta_1}\right] \quad (44)$$
$$= \frac{\delta_1-\delta_2}{1-\delta_1} + \frac{\delta_2-\delta_{12}}{1-\delta_{12}}. \quad (45)$$

The second term of (44) is because that when $K_i = 0$ we need to add back 1 to $K_i - 1$, which resulting in adding back

$$\Pr\{K_i = 0|L_i\} = (\Pr\{S_2 = 0|S_1 = 0\})^{L_i-1}\Pr\{S_2 = 0|S_1 = 1\}.$$

The equality (45) is due to the following fact

$$\mathbb{E}\left[(\delta_{12}/\delta_1)^{L_i}\right] = \frac{1-\delta_1}{\delta_1}\frac{\delta_{12}}{1-\delta_{12}} \quad (46)$$

which is easy to prove from the Geometric distributed L_i.

$$(n_1+n_2)(1-\delta_1) - (n_1+n_2)(1-\delta_1-\delta_2+\delta_{12}) < \frac{(1-\delta_2)(\delta_1-\delta_{12})}{(1-\delta_1)(1-\delta_{12})}R_1 n + R_0 n\frac{\delta_1-\delta_{12}}{1-\delta_{12}}\frac{1-\delta_2}{1-\delta_1} + (1-\delta_2)\left(n - \left(n_1+n_2+\frac{1}{1-\delta_1}\left(R_0 n\frac{\delta_1-\delta_{12}}{1-\delta_{12}} + R_1 n\right)\right)\right) \quad (38)$$

B. Satisfaction of the Decodability Constraint (38)

Given a common message rate R_0, we aim on the satisfaction of constraint (38) for the following point

$$R_1 = \frac{(1-\delta_1)(\delta_2-\delta_{12})-(\delta_1-\delta_{12})R_0}{(1-\delta_{12})-\frac{(1-\delta_1)(1-\delta_2)}{(1-\delta_{12})}} := R_1^* \tag{47}$$

$$R_2 < \frac{(1-\delta_2)(\delta_1-\delta_{12})-(\delta_2-\delta_{12})R_0}{(1-\delta_{12})-\frac{(1-\delta_1)(1-\delta_2)}{(1-\delta_{12})}} := R_2^* \tag{48}$$

from the intersection of (11) (12), since the satisfaction of other points are trivial. It can be easily checked that rate pair (R_1^*, R_2^*) reduces to the RHS of (25)(26) if $R_0 = 0$. Applying (32) in RHS of (38), together with (28) (29) and the fact that $R_2 < R_2^*$, the decodability constraint (38) is met if

$$(\delta_2-\delta_1)R_0 \le \frac{(1-\delta_2)(\delta_1-\delta_{12})}{1-\delta_{12}}R_1^* - \frac{(1-\delta_1)(\delta_2-\delta_{12})}{1-\delta_{12}}R_2^*. \tag{49}$$

Now we aim to show that constraint is met with equality from R_1^* and R_2^* given in (47) and (48) respectively. First note that from (47)(48),

$$\begin{aligned}&R_1^*\left((1-\delta_2)(\delta_1-\delta_{12})-(\delta_2-\delta_{12})R_0\right)\\ =&R_2^*\left((1-\delta_1)(\delta_2-\delta_{12})-(\delta_1-\delta_{12})R_0\right),\end{aligned}$$

and then (49) turns to

$$(\delta_2-\delta_1)R_0 = \frac{(\delta_2-\delta_{12})}{1-\delta_{12}}R_0R_1^* - \frac{(\delta_1-\delta_{12})}{1-\delta_{12}}R_0R_2^*. \tag{50}$$

To show (50), observe that for any R_0,

$$\begin{aligned}&(\delta_2-\delta_1)R_0 =\\ &\frac{(\delta_2-\delta_{12})}{(1-\delta_{12})-\frac{(1-\delta_1)(1-\delta_2)}{(1-\delta_{12})}}\frac{(1-\delta_1)(\delta_2-\delta_{12})}{1-\delta_{12}}R_0\\ &-\frac{(\delta_1-\delta_{12})}{(1-\delta_{12})-\frac{(1-\delta_1)(1-\delta_2)}{(1-\delta_{12})}}\frac{(1-\delta_2)(\delta_1-\delta_{12})}{1-\delta_{12}}R_0\end{aligned} \tag{51}$$

since

$$\begin{aligned}&(\delta_2-\delta_1)\left((1-\delta_{12})^2-(1-\delta_1)(1-\delta_2)\right)\\ =&(1-\delta_1)(\delta_2-\delta_{12})^2-(1-\delta_2)(\delta_1-\delta_{12})^2.\end{aligned}$$

Furthermore, from (47) and (48),

$$\begin{aligned}&(\delta_2-\delta_{12})R_0R_1^*\\ &=(\delta_2-\delta_{12})R_0\frac{(1-\delta_1)(\delta_2-\delta_{12})-(\delta_1-\delta_{12})R_0}{(1-\delta_{12})-\frac{(1-\delta_1)(1-\delta_2)}{(1-\delta_{12})}},\\ &(\delta_1-\delta_{12})R_0R_2^*\\ &=(\delta_1-\delta_{12})R_0\frac{(1-\delta_2)(\delta_1-\delta_{12})-(\delta_2-\delta_{12})R_0}{(1-\delta_{12})-\frac{(1-\delta_1)(1-\delta_2)}{(1-\delta_{12})}},\end{aligned}$$

and then equality (51) becomes (50), which validates (49).

C. Proof of the Converse in Lemma 6.1

Since $\delta_2 = \delta_3$, the first step involves creating "full correlation" among the two N users. In other words, the erasure links connected to the two receivers that do not provide any feedback, will have the same realizations, *i.e.* $S_2[t] = S_3[t]$. Since the N users do not provide their CSI to the transmitter or the other receivers, the transmit signal (39) is independent from the associated channel state. In other words, the marginal distribution of each link remain unchanged under this induced correlation, and thus, the capacity region remains unchanged under this modification [29]. In essence, we can merge and treat the two N users as a single receiver with a total private rate of

$$R_{23} := R_2 + R_3. \tag{52}$$

To see this, as the two N users are identical, the decoding constraint at Rx3 in (40) becomes

$$\Pr\left\{(W_3, W_0) \neq g_3(Y_2^{[1:n]}, S_1^{[1:n]}, S_2^{[1:n]})\right\} \to 0.$$

By letting $W_{23} \triangleq W_2 \bigcup W_3$, together with (3), W_{23} can be decoded by receiver Rx2 or Rx3.

To derive outer-bound (42), we further let the transmitter in the new broadcast packet erasure channel has global delayed CSI $S^{t-1} = (S_1^{[1:t-1]}, S_2^{[1:t-1]})$ at time t and all receivers have full CSI $S^n = (S_1^{[1:n]}, S_2^{[1:n]})$. We remind the reader that with the induced correlation, we have $S_2[t] = S_3[t]$. Let

$$\beta = \frac{1-\delta_1\delta_2}{1-\delta_2} \tag{53}$$

and then by converting unit of our rate to bits per time slot as footnote in Section II-A, we have (54), shown on the top of the next page, where $\epsilon_n \to 0$ as $n \to \infty$; (a) follows from the independence of messages; (b) follows from Fano's inequality and messages are independent of channel realizations; (c) holds as Y_1^n is a deterministic function of the messages and the global channel state sequence; (d) follows from Claim 1.1 below; (e) holds since

$$H\left(Y_2^n|S^n\right) \le \sum_{t=1}^{n} H\left(Y_2[t]|S^n\right) \le n(1-\delta_2)\log_2(q).$$

Dividing both sides of (54) by $n(1-\delta_1\delta_2)\log_2(q)$ and letting $n \to \infty$, we get (42). Similarly, we can obtain (41). Now, we prove step (d) of (54) in the following claim.

Claim 1.1: Under global delayed CSI, for the broadcast packet erasure channel with two private messages W_1 and W_{23} for receiver 1 and 2 respectively and a common message W_0, by setting the ratio β as (53) the received and state sequences satisfy

$$\begin{aligned}&H\left(Y_1^n|W_0, W_{23}, S^n\right)\\ &-\beta\Big\{H\left(Y_2^n|W_{23}, S^n\right) - I\left(W_0; Y_2^n|W_{23}, S^n\right)\Big\} \le 0.\end{aligned} \tag{55}$$

Proof: We first note that proving (55) is equivalent to proving

$$H\left(Y_1^n|W_0, W_{23}, S^n\right) - \beta H\left(Y_2^n|W_0, W_{23}, S^n\right) \le 0 \tag{56}$$

$$
\begin{aligned}
n\left(R_1+\beta\left\{R_{23}+R_0\right\}\right)\log_2(q) &= H\left(W_1\right)+\beta\left\{H\left(W_{23}\right)+H\left(W_0\right)\right\}\\
&\overset{(a)}{=} H\left(W_1|W_0,W_{23}\right)+\beta\left\{H\left(W_{23}\right)+H\left(W_0|W_{23}\right)\right\}\\
&\overset{(b)}{\leq} I\left(W_1;Y_1^n|W_0,W_{23},S^n\right)+\beta\left\{I\left(W_{23};Y_2^n|S^n\right)+I\left(W_0;Y_2^n|W_{23},S^n\right\}\right)+n\epsilon_n\\
&\overset{(c)}{=} H\left(Y_1^n|W_0,W_{23},S^n\right)-\underbrace{H\left(Y_1^n|W_0,W_1,W_{23},S^n\right)}_{=\,0}+\beta H\left(Y_2^n|S^n\right)\\
&\quad-\beta\left\{H\left(Y_2^n|W_{23},S^n\right)-I\left(W_0;Y_2^n|W_{23},S^n\right)\right\}+n\epsilon_n\\
&\overset{(d)}{\leq} \beta H\left(Y_2^n|S^n\right)+n\epsilon_n \overset{(e)}{\leq} n\beta(1-\delta_2)\log_2(q)+n\epsilon_n=n(1-\delta_1\delta_2)\log_2(q)+n\epsilon_n
\end{aligned}
\tag{54}
$$

since

$$
\begin{aligned}
&H\left(Y_2^n|W_0,W_{23},S^n\right)\\
&=H\left(Y_2^n|W_{23},S^n\right)-I\left(W_0;Y_2^n|W_{23},S^n\right).
\end{aligned}
$$

Then, we have

$$
\begin{aligned}
H\left(Y_2^n|W_0,W_{23},S^n\right) &= \sum_{t=1}^{n} H\left(Y_2[t]|Y_2^{t-1},W_0,W_{23},S^n\right)\\
&\overset{(a)}{=} \sum_{t=1}^{n} H\left(Y_2[t]|Y_2^{t-1},W_0,W_{23},S^t\right)\\
&\overset{(b)}{=} \sum_{t=1}^{n} H\left(X[t]|Y_2^{t-1},W_0,W_{23},S_2[t]=1,S_1[t],S^{t-1}\right)\\
&\qquad\cdot(1-\delta_2)\\
&\overset{(c)}{=} \sum_{t=1}^{n} (1-\delta_2)H\left(X[t]|Y_2^{t-1},W_0,W_{23},S^t\right)
\end{aligned}
\tag{57}
$$

where (a) follows from the fact that signal $Y_2[t]$ at current time t is independent of future channel states; (b) holds since $\Pr\left(S_2[t]=1\right)=(1-\delta_2)$; (c) is true since transmitted signal $X[t]$, as well as $(Y_2^{t-1},W_0,W_{23},S^{t-1})$, are independent of the current channel state at time t. Now, we can lower-bound (57) as

$$
\begin{aligned}
&\sum_{t=1}^{n}(1-\delta_2)H\left(X[t]|Y_2^{t-1},W_0,W_{23},S^t\right)\\
&\overset{(d)}{\geq} \sum_{t=1}^{n}(1-\delta_2)H\left(X[t]|Y_1^{t-1},Y_2^{t-1},W_0,W_{23},S^t\right)\\
&\overset{(e)}{=} \sum_{t=1}^{n}\frac{(1-\delta_2)}{(1-\delta_1\delta_2)}H\left(Y_1[t],Y_2[t]|Y_1^{t-1},Y_2^{t-1},W_0,W_{23},S^t\right)\\
&\overset{(f)}{=} \sum_{t=1}^{n}\frac{(1-\delta_2)}{(1-\delta_1\delta_2)}H\left(Y_1[t],Y_2[t]|Y_1^{t-1},Y_2^{t-1},W_0,W_{23},S^n\right)\\
&\overset{(g)}{=} \frac{(1-\delta_2)}{(1-\delta_1\delta_2)}H\left(Y_1^n,Y_2^n|W_0,W_{23},S^n\right)\\
&\geq \frac{1}{\beta}H\left(Y_1^n|W_0,W_{23},S^n\right),
\end{aligned}
\tag{58}
$$

where (d) holds since conditioning reduces entropy; (e) holds from applying chain rule on $H(Y_1[t],Y_2[t]|Y_1^{t-1},Y_2^{t-1},W_0,W_{23},S^t)$, together with

$$\Pr\left(S_1[t]=1\right)+\Pr\left(S_1[t]=0,S_2[t]=1\right)=1-\delta_1\delta_2$$

for the two erasure links; (f) is true since all signals at current time t are independent of future channel states; (g) follows from the chain rule and the final lower-bound comes from (53) and the non-negativity of the entropy function for discrete random variables. From (57) and (58), (56) is valid and it concludes the proof of the converse.

As a final note, after using the tricks in the first paragraph of this Appendix, the rest of the proof steps are similar to those for the converse of two-user "DD" scenario (full feedback). Thus one can also modify the other scenario "DD" proofs to reach (42)(41). For example, the outer-bounds derived in [20] can be specialized to obtain the two bounds for scenario "DD". Further, if we provide the common message to one user as side-information, then the common message becomes part of the private message for the user with no side-information. Re-writing the outer-bounds of the two-user broadcast packet erasure channel [1] for this setting, we end up with four outer-bounds and ought to determine which two are tight. We believe our proof fits well within the Entropy Leakage lemma framework first introduced in the interference channel [30], which is an alternative method to the Ozarow-type physically degraded arguments [27] and provides an intuitive way to capture the common message in the outer-bounds. Note that in the context of MIMO Gaussian channel, the leakage lemma still holds by replacing the ratio β in the bound of Claim A.1 (without common message) according to the number of antennas [31]. ■

Acknowledgment

The authors would like to thank Prof. Max Costa and the reviewers for their suggestions that greatly improve the presentation of the paper. In particular, the suggestions from Reviewer 2 significantly simplify the presentation of Section V.

References

[1] L. Georgiadis and L. Tassiulas, "Broadcast erasure channel with feedback-capacity and algorithms," in *Proc. Workshop Netw. Coding, Theory, Appl.*, Lausanne, Switzerland, Jun. 2009, pp. 54–61.

[2] C.-C. Wang, "On the capacity of 1-to-k broadcast packet erasure channels with channel output feedback," *IEEE Trans. Inf. Theory*, vol. 58, no. 2, pp. 931–956, Feb. 2012.

[3] M. Gatzianas, L. Georgiadis, and L. Tassiulas, "Multiuser broadcast erasure channel with feedback—Capacity and algorithms," *IEEE Trans. Inf. Theory*, vol. 59, no. 9, pp. 5779–5804, Sep. 2013.

[4] M. A. Maddah-Ali and D. Tse, "Completely stale transmitter channel state information is still very useful," *IEEE Trans. Inf. Theory*, vol. 58, no. 7, pp. 4418–4431, Jul. 2012.

[5] A. G. Davoodi and S. A. Jafar, "Aligned image sets under channel uncertainty: Settling conjectures on the collapse of degrees of freedom under finite precision CSIT," *IEEE Trans. Inf. Theory*, vol. 62, no. 10, pp. 5603–5618, Oct. 2016.

[6] A. F. Dana, R. Gowaikar, R. Palanki, B. Hassibi, and M. Effros, "Capacity of wireless erasure networks," *IEEE Trans. Inf. Theory*, vol. 52, no. 3, pp. 789–804, Mar. 2006.

[7] *Study on New Radio Access Technology Physical Layer Aspects*, document 3GPP TR-38.802, 2017.

[8] T. Ho *et al.*, "A random linear network coding approach to multicast," *IEEE Trans. Inf. Theory*, vol. 52, no. 10, pp. 4413–4430, Oct. 2006.

[9] S. Katti, H. Rahul, W. Hu, D. Katabi, M. Medard, and J. Crowcroft, "XORs in the air: Practical wireless network," in *Proc. ACM Special Interest Group Data Commun. (SIGCOMM)*, 2006, pp. 243–254.

[10] D. Koutsonikolas, C.-C. Wang, and Y. C. Hu, "CCACK: Efficient network coding based opportunistic routing through cumulative coded acknowledgments," in *Proc. IEEE INFOCOM*, San Diego, CA, USA, Mar. 2010, pp. 1–9.

[11] S.-C. Lin and I.-H. Wang, "Single-user CSIT can be quite useful for state-dependent broadcast channels," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2016, pp. 160–164.

[12] C. He and S. Yang, "On the two-user erasure broadcast channel with one-sided feedback," in *Proc. IEEE Inf. Theory Workshop (ITW)*, Nov. 2017, pp. 494–498.

[13] S.-C. Lin and I.-H. Wang, "Gaussian broadcast channels with intermittent connectivity and hybrid state information at the transmitter," *IEEE Trans. Inf. Theory*, vol. 64, no. 9, pp. 6362–6383, Sep. 2018.

[14] O. Shayevitz and M. Wigger, "On the capacity of the discrete memoryless broadcast channel with feedback," *IEEE Trans. Inf. Theory*, vol. 59, no. 3, pp. 1329–1345, Mar. 2013.

[15] R. Venkataramanan and S. S. Pradhan, "An achievable rate region for the broadcast channel with feedback," *IEEE Trans. Inf. Theory*, vol. 59, no. 10, pp. 6175–6191, Oct. 2013.

[16] Y. Wu and M. Wigger, "Coding schemes with rate-limited feedback that improve over the no feedback capacity for a large class of broadcast channels," *IEEE Trans. Inf. Theory*, vol. 62, no. 4, pp. 2009–2033, Apr. 2016.

[17] H. Kim, Y.-K. Chia, and A. El Gamal, "A note on the broadcast channel with stale state information at the transmitter," *IEEE Trans. Inf. Theory*, vol. 61, no. 7, pp. 3622–3631, Jul. 2015.

[18] S. Kamath, V. Anantharam, D. Tse, and C.-C. Wang, "The two-unicast problem," *IEEE Trans. Inf. Theory*, vol. 64, no. 5, pp. 3865–3882, May 2018.

[19] A. Vahid, "On the degrees-of-freedom of two-unicast wireless networks with delayed CSIT," *IEEE Trans. Inf. Theory*, vol. 65, no. 8, pp. 5176–5188, Aug. 2019.

[20] J. Han and C.-C. Wang, "General capacity region for the fully connected three-node packet erasure network," *IEEE Trans. Inf. Theory*, vol. 62, no. 10, pp. 5503–5523, Oct. 2016.

[21] S.-C. Lin, I.-H. Wang, and A. Vahid, "No feedback, no problem: Capacity of erasure broadcast channels with single-user delayed CSI," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2019, pp. 1647–1651.

[22] C.-C. Wang and J. Han, "The capacity region of two-receiver multiple-input broadcast packet erasure channels with channel output feedback," *IEEE Trans. Inf. Theory*, vol. 60, no. 9, pp. 5597–5626, Sep. 2014.

[23] A. F. Dana and B. Hassibi, "The capacity region of multiple input erasure broadcast channels," in *Proc. Int. Symp. Inf. Theory (ISIT)*, 2005, pp. 2315–2319.

[24] A. Vahid, S.-C. Lin, and I.-H. Wang, "Capacity region of erasure broadcast channels with common message and feedback," 2019, *arXiv:1910.05912*. [Online]. Available: http://arxiv.org/abs/1910.05912

[25] A. S. Belhadj, Y. Steinberg, and M. Wigger, "MIMO MAC-BC duality with linear-feedback coding schemes," *IEEE Trans. Inf. Theory*, vol. 61, no. 11, pp. 5976–5998, Nov. 2015.

[26] A. El-Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge, U.K.: Cambridge Univ. Press, 2011.

[27] L. H. Ozarow and S. Leung-Yan-Cheong, "An achievable region and outer bound for the Gaussian broadcast channel with feedback (Corresp.)," *IEEE Trans. Inf. Theory*, vol. IT-30, no. 4, pp. 667–671, Jul. 1984.

[28] Z. Matteo, "On the capacity of erasure broadcast channel with heterogeneous feedback," M.S. thesis, Dept. EE, Nat. Taiwan Univ., Taipei, Taiwan, Aug. 2019.

[29] H. Sato, "An outer bound on the capacity region of broadcast channel," *IEEE Trans. Inf. Theory*, vol. IT-24, pp. 374–377, May 1978.

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

[31] C. S. Vaze and M. K. Varanasi, "The degrees of freedom region and interference alignment for the MIMO interference channel with delayed CSIT," *IEEE Trans. Inf. Theory*, vol. 58, no. 7, pp. 4396–4417, Jul. 2012.

Shih-Chun Lin (Senior Member, IEEE) received the B.S. and Ph.D. degrees in electrical engineering from the National Taiwan University, Taipei, Taiwan, in 2000 and 2007, respectively. He was a Visiting Student with The Ohio State University, Columbus, OH, USA, in 2007. From 2011 to 2012, he was with the National Taipei University of Technology. From 2012 to 2021, he was with the National Taiwan University of Science and Technology, Taipei. In August 2021, he joined the National Taiwan University as an Associate Professor. His research interests include information theory, communications, and cyber-physical security. He was a TPC Member of the IEEE ICC from 2018 to 2022. He received the Best Paper Award for Young Authors from the IEEE IT/COM Society Taipei/Tainan Chapter in 2015 and twice the Project for Excellent Junior Research Investigators from the Ministry of Science and Technology, Taiwan, in 2015 and 2018. He serves as the TPC Co-Chair for IEEE ICC Workshop on B5G-URLLC 2019 and the Publication Chair for the IEEE GLOBECOM 2020.

I-Hsiang Wang (Member, IEEE) received the B.Sc. degree in electrical engineering from the National Taiwan University, Taiwan, in 2006, and the Ph.D. degree in electrical engineering and computer sciences from the University of California at Berkeley, USA, in 2011. From 2011 to 2013, he was a Post-Doctoral Researcher at the Ècole Polytechnique Fèdèrale de Lausanne, Switzerland. Since 2013, he has been with the Department of Electrical Engineering, National Taiwan University, where he is currently an Associate Professor. His research interests include network information theory, networked data analysis, and statistical learning. He was a Finalist of the Best Student Paper Award of IEEE International Symposium on Information Theory in 2011. He received the 2017 IEEE Information Theory Society Taipei Chapter and the IEEE Communications Society Taipei/Tainan Chapters Best Paper Award for Young Scholars.

Alireza Vahid (Senior Member, IEEE) received the B.Sc. degree in electrical engineering from Sharif University of Technology, Tehran, Iran, in 2009, and the M.Sc. and Ph.D. degrees in electrical and computer engineering from Cornell University, Ithaca, NY, USA, in 2012 and 2015, respectively. From 2015 to 2017, he worked as a Post-Doctoral Research Scientist at the Information Initiative, Duke University, Durham, NC, USA. He is currently an Assistant Professor of electrical engineering at the University of Colorado Denver, Denver. His research interests include network information theory, wireless communications, coding theory, and applications of coding theory in high-performance computer memory systems. He received Jacobs Scholar Fellowship in 2009 from Cornell University, the 2010 Director's Ph.D. Teaching Award, and the 2015 Outstanding Ph.D. Thesis Research Award. He has received Qualcomm Innovation Fellowship in 2013 and the Lab Venture Challenge Award in 2019.