Bursty Interference Channel with Feedback

I-Hsiang Wang
EPFL, Lausanne, Switzerland
i-hsiang.wang@epfl.ch

Changho Suh
KAIST, South Korea
chsuh@kaist.ac.kr

Suhas Diggavi
UCLA, Los Angeles, USA
suhasdiggavi@ucla.edu

Pramod Viswanath
UIUC, Urbana, USA
pramodv@illinois.edu

***Abstract*—We explore the benefit of feedback for physical layer interference management in wireless networks without centralized upper layer control mechanisms. Lack of coordination in the upper layer could make the interference experienced in the physical layer bursty. To understand how to harness such burstiness with feedback, we investigate a two-user bursty interference channel (IC), where the presence of interference is governed by a Bernoulli random state. We completely characterize the capacity region of the symmetric two-user linear deterministic bursty IC with feedback. The proposed two-phase scheme exploits feedback either for refining the previous interfered reception or for relaying additional information to the legitimate receiver of the other user. Matching outer bounds are derived by novel techniques that take the effect of delayed state information into account. We also use insights from the deterministic case to characterize the approximate symmetric capacity for the symmetric Gaussian bursty IC with feedback in the weak interference regime.**

I. Introduction

Interference has become the major barrier against efficient utilization of limited spectrum in modern wireless networks. For wireless networks, the simplest information theoretic model for studying interference is the two-user Gaussian *interference channel* (IC), in which it is assumed that the transmitted signal from one transmitter *always* interferes with the receiver of the other user. Under this potentially conservative modeling assumption, the capacity region is characterized to within 1 bit/s/Hz [1], which sheds light on how to manage interference in the high SNR regime.

In many realistic scenarios, however, interference could be *bursty* owing to the distributed medium access control mechanisms and/or the decentralized networking protocols across different users. For example, consider an OFDM-based wireless system, where two neighboring access points (AP's) lack coordination in allocation of subcarriers to the users they serve. Since each cell has full frequency reuse, from time to time a cell-boundary user served by one AP is faced with interference caused by some other cell-boundary user served by the other AP. If these two AP's allocate the same frequency band to the users they serve, they form an interference channel. If not, they form two point-to-point interference-free links.

Hence in the physical layer, it may be too pessimistic to assume that interference is always present. In principle harnessing such burstiness could increase users' data rates.

The work of I.-H. Wang was supported by EU project CONECT FP7-ICT-2009-257616. The work of C. Suh was funded by the MSIP (Ministry of Science, ICT & Future Planning), Korea in the ICT R&D Program 2013. The work of S. Diggavi was supported in part by NSF award 1136174 and MURI award AFOSR FA9550-09-064.

A natural way to harness such burstiness is the degraded-message-set approach proposed in [2], where an opportunistic message can be decoded in addition to the original message when the interference is not present. The degraded-message-set approach does not require any feedback from the receivers. In most wireless systems, however, feedback is an available resource and potentially can be utilized to manage interference and/or exploit the burstiness. For managing interference, it is shown that feedback can provide an unbounded gain in capacity for the non-bursty interference channel [3]. For exploiting the burstiness of interference, one could in principle exploit its temporal correlation to predict the presence of interference beforehand by using feedback, and opportunistically increase the overall throughput. However, when the burstiness is memoryless, it is not clear how feedback can help at the first glance.

In this work, we study the benefit of feedback for managing bursty interference. We focus on the linear deterministic model [4] of a simple two-user bursty Gaussian interference channel, where an i.i.d. Bernoulli process $\{S[t]\}$ controls the presence of the (additive) interference at both receivers. The realization of the *burstiness state* $S[t]$ is known to both receivers at the end of time slot t and can be viewed as part of the channel output. Perfect channel output feedback is available from each receiver to its own transmitter, which includes the received signal and the state. In the literature, our model is closely related to the binary fading interference channel with feedback [5]. In particular, [5] considered a (single-level) binary fading interference channel where the four links are associated with i.i.d. binary fading states, and provided tight characterization of the capacity region under various feedback settings.

Our main contribution is the capacity region characterization of the symmetric linear deterministic bursty IC. We propose a novel two-phase scheme which exploits feedback to harness bursty interference. Depending on the operating regime as well as the target rate pair, the scheme differs in how it utilizes feedback to manage bursty interference. In the weak and the strong interference regime, feedback is used to refine the interfered signals in the previous time slot. In the very strong interference regime, on the other hand, the role of feedback is to relay additional information from the unintended transmitter to the desired receiver. In order to optimally exploit the burstiness of interference using feedback, the transmit strategy in the strong and the very strong interference regime has to be carefully modified from the original scheme proposed in the case of non-bursty IC [3]. For the converse part, we prove a set of outer bounds with novel techniques, which match the inner bounds when the channel is symmetric. Using the insights from the linear deterministic case, we characterize the symmetric capacity of the symmetric Gaussian bursty IC to within 1 bit/s/Hz in the weak interference regime. We compare our result with the non-feedback opportunistic achievable rate [2] and the non-bursty feedback capacity [3] to demonstrate the synergic benefit of feedback in managing bursty interference.

The rest of this paper is organized as follows. In Section II we formulate the problem. In Section III we show the main result. The achievability part is proved in Section IV while the converse part is proved in Section V.

II. Problem Formulation

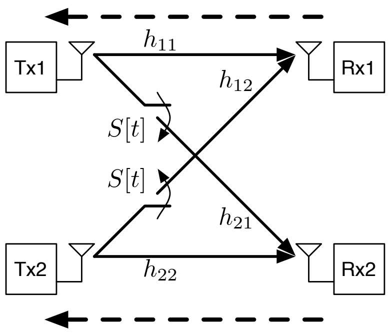


Fig. 1. Two-user Bursty Interference Channel

The two-user bursty Gaussian interference channel is depicted in Fig. 1 and defined as follows. User i has a message W_i to be reliably delivered from transmitter i (Txi) to receiver i (Rxi), $i = 1, 2$. The transmit signal at transmitter i is $X_i \in \mathbb{C}$, for $i = 1, 2$. For $(i, j) = (1, 2), (2, 1)$, the received signal at Rxi at time t is

$$Y_i[t] = h_{ii}X_i[t] + h_{ij}\widetilde{X}_j[t] + Z_i[t],\ \widetilde{X}_j[t] := S[t]X_j[t],$$

where the two independent additive noise terms $Z_1[t], Z_2[t]$ are $\mathcal{CN}(0, 1)$ and i.i.d. over time. The burstiness state $S[t]$ is an i.i.d. Bernoulli process $\mathrm{Ber}(p)$ which governs the presence of interference and is revealed to both receivers at the end of time slot t. The channel inputs are subject to unit power constraint, that is, $\frac{1}{N}\sum_{t=1}^{N} |X_i[t]|^2 \le 1$ for block length N. We denote signal-to-noise ratios $\mathsf{SNR}_i := |h_{ii}|^2$ and interference-to-noise ratios $\mathsf{INR}_i := |h_{ij}|^2$ at receiver $i = 1, 2$ respectively. Perfect feedback from receiver i to transmitter i is available, and the transmit signal from transmitter i at time t, $X_i[t]$, is determined by W_i and $(Y_i[1 : t-1], S[1 : t-1])$.

In this paper, we mostly focus on a linear deterministic model [4] of the bursty Gaussian channel defined above. The transmit signal at transmitter i is $X_i \in \mathbb{F}_2^q$, for $i = 1, 2$. Here $\mathbb{F}_2$ denotes the binary field $\{0, 1\}$. The received signals are (dropping the time index t)

$$Y_1 = \mathbf{H}_{11}X_1 + \mathbf{H}_{12}\widetilde{X}_2, \qquad Y_2 = \mathbf{H}_{22}X_2 + \mathbf{H}_{21}\widetilde{X}_1,$$

where additions are modulo-two component-wise. Channel transfer matrices $\mathbf{H}_{ij} := \mathbf{S}^{q-n_{ij}}$ for $(i, j) \in \{1, 2\}^2$, where $q = \max\{n_{11}, n_{12}, n_{21}, n_{22}\}$, and $\mathbf{S} \in \mathbb{F}_2^{q\times q}$ is the shift matrix $\begin{bmatrix} \mathbf{0}^T & 0 \\ \mathbf{I}_{q-1} & \mathbf{0} \end{bmatrix}$, where $\mathbf{0}$ is the zero vector in $\mathbb{F}_2^{q-1}$ and $\mathbf{I}_{q-1}$ is the identity matrix in $\mathbb{F}_2^{(q-1)\times(q-1)}$.

III. Main Result

In this paper, we focus on the symmetric case where

$$\begin{aligned} &n_{11} = n_{22} = n,\ n_{21} = n_{12} = k, \\ &\mathsf{SNR}_1 = \mathsf{SNR}_2 = \mathsf{SNR},\ \mathsf{INR}_1 = \mathsf{INR}_2 = \mathsf{INR}. \end{aligned}$$

Our main result consists of a general outer bound as well as the characterization of the feedback capacity region of the symmetric linear deterministic bursty IC. We also characterize the symmetric feedback capacity to within $\frac{p(3-p)}{1+p}$ bits/s/Hz of the symmetric Gaussian bursty IC in the weak interference regime. Some of the proof details can be found in [6].

Theorem 3.1 (Outer Bound): If (R_1, R_2) is achievable, it satisfies the bounds in (1) – (3) for $(i, j) = (1, 2), (2, 1)$:

$$R_i \le n_{ii} + p\min\left\{(n_{ij} - n_{ii})^+, (n_{ji} - n_{ii})^+\right\} \tag{1}$$

$$\begin{aligned} R_i + R_j \le\ & (1-p)(n_{11} + n_{22}) \\ & + p(\max(n_{ii}, n_{ij}) + (n_{jj} - n_{ij})^+) \end{aligned} \tag{2}$$

$$\begin{aligned} R_i + pR_j \le\ & (1-p)n_{ii} \\ & + p(\max(n_{ii}, n_{ij}) + (n_{jj} - n_{ij})^+) \end{aligned} \tag{3}$$

Proof: See Section V. ■

Theorem 3.2 (Capacity Region: Symmetric Case): The capacity region of the symmetric bursty linear deterministic interference channel with feedback is characterized by

$$\begin{aligned} R_1, R_2 &\le (1-p)n + p\max(n, k) \\ R_1 + R_2 &\le 2(1-p)n + p\left\{\max(n, k) + (n-k)^+\right\} \\ R_1 + pR_2 &\le (1-p)n + p\left\{\max(n, k) + (n-k)^+\right\} \\ pR_1 + R_2 &\le (1-p)n + p\left\{\max(n, k) + (n-k)^+\right\} \end{aligned}$$

Proof: Theorem 3.1 proves the converse. See Section IV for the achievability part. ■

For the symmetric Gaussian case, we have the following partial result regarding the symmetric capacity[1].

Theorem 3.3 (Symmetric Capacity to within One Bit): In the weak interference regime $\mathsf{SNR} \ge \mathsf{INR}$, we have

$$\overline{C}_{\text{sym}} - \frac{p(3-p)}{1+p} \le C_{\text{sym}}^{\text{FB}} \le \overline{C}_{\text{sym}},$$

where

$$\overline{C}_{\text{sym}} = \left\{ \begin{array}{l} \frac{1-p}{1+p}\log(1+\mathsf{SNR}) + \frac{p}{1+p}\log\left(1 + \frac{\mathsf{SNR}}{1+\mathsf{INR}}\right) \\ +\frac{p}{1+p}\log\left(1 + \left(\sqrt{\mathsf{SNR}} + \sqrt{\mathsf{INR}}\right)^2\right) \end{array} \right\}.$$

Proof: See [6]. ■

Note that the gap $\frac{p(3-p)}{1+p}$ is maximized when $p = 1$ and $\frac{p(3-p)}{1+p} \le 1$. Hence, we recover the one-bit-gap result in [3] for the case $p = 1$.

[1]For the result in the strong interference regime, see [6].

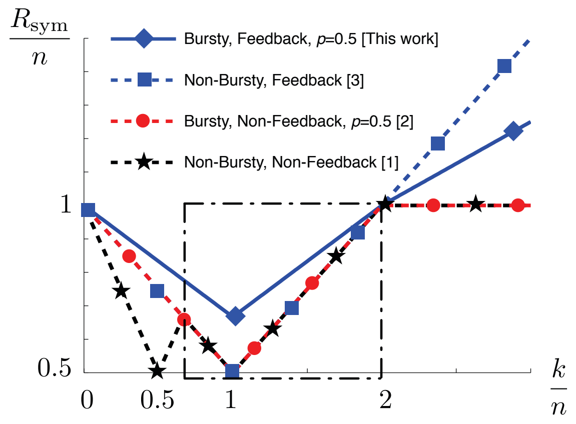


Fig. 2. Comparison of Normalized Symmetric Rates

A natural strategy for the non-feedback case is the degraded message set approach [2]. For comparison, we focus on the symmetric rate point $R_1 = R_2 = R_{\text{sym}}$. The achievable rate by the non-feedback strategy can be derived from the results in [2] (see [6] for more details). On the other hand, from Theorem 3.2 the symmetric feedback capacity is characterized. When $p = 0$, both rates coincide with the interference-free capacity. When $p = 1$, they coincide with the non-feeback IC capacity [1] and the feedback IC capacity [3], respectively. We plot the the above two versus k/n for $p = 0.5$ in Fig. 2 for comparison.

Remark 3.1 (Synergic Benefit): The plot in Fig. 2 shows the synergic benefit of feedback in managing bursty interference. Let us focus on the regime $2/3 \leq k/n \leq 2$, highlighted by the square. Exploiting burstiness in the absence of feedback [2] provides no capacity gain (it provides gain in this regime if $p < 0.5$). For non-bursty IC, feedback provides no gain in this regime either [3]. When the interference is bursty, however, feedback increases capacity. Such gain when translated to the Gaussian case is unbounded.

Remark 3.2 (Optimality in Non-Feedback Case): Note that the degraded message set approach need not be optimal for the non-feedback bursty IC, since it does not harness the possibility to code over time. Nevertheless, we are able to show that indeed it achieves capacity for the case $p \leq 0.5$, which is interesting by itself. See [6] for more details.

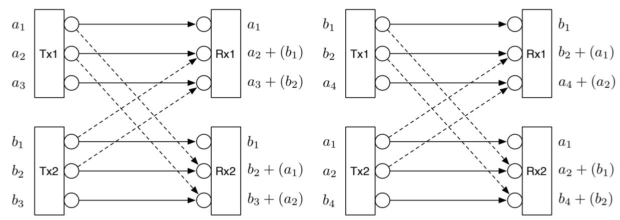
IV. Achieving Symmetric Capacity

In this paper we only present the achievability proof of the symmetric rate point. For the complete proof of achieving the capacity region, please refer to [6].

We shall distinguish into three different regimes according to the ratio k/n: (1) weak interference regime $k/n < 1$, (2) strong interference regime $1 \leq k/n < 2$, and (3) very strong interference regime $k/n \geq 2$. The scheme that we use in different regimes will be different.

A. Weak Interference Regime $k/n < 1$

When $k < n$, we would like to achieve $C_{\text{sym}}^{\text{FB}} = n - \frac{p}{1+p}k$.



(a) Phase F with interference (b) Phase R with interference

Fig. 3. $n = 3, k = 2$. Symbols in $(\cdot)$ only appear when interference is on.

Strategy: We define the following two transmission phases. Both users take the same strategy due to symmetry, and we only describe it for user 1. An illustrative example for the scheme is depicted in Fig. 3.

- Phase F: Send *fresh* information on all n levels. Let $\begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix}^T$ and $\begin{bmatrix} b_1 & \cdots & b_n \end{bmatrix}^T$ denote the binary symbols sent from Tx1 and Tx2 respectively in a particular Phase F at time t. Here a_i, b_i are the bits sent on the i-th MSB level (pictorially).
  If there is no interference at time t, both transmitters remain in Phase F in the next time slot $t+1$. If there is interference, both transit to Phase R at time $t+1$.
- Phase R: Send the past interference on the common k levels to *refine* the previous reception, along with fresh information on the private $(n-k)$ levels. For example, if there is interference in the previous time slot t under Phase F, in the current time slot $t+1$ it should switch to Phase R, and Tx1 shall send $\begin{bmatrix} b_1 & \cdots & b_k & a_{n+1} & \cdots & a_{2n-k} \end{bmatrix}^T$. Regardless of the presence of interference, both transmitters transit to Phase F in the next time slot $t+2$.

Based on the above operation, the random process defined by the transmission phases is a two-state Markov chain depicted in Fig. 4(a).

Analysis: We shall show that in Phase R regardless of the presence of interference, at Rx1 all symbols $\{a_1, a_2, \ldots, a_{2n-k}\} \cup \{b_1, \ldots, b_k\}$ can be decoded. Let us assume that at time t the transmission phase is F while at time $t+1$ the transmission phase is R. If we represent the received linear combinations on all levels as a vector

$$\begin{bmatrix} y_1[t] \\ y_1[t+1] \end{bmatrix} = \mathbf{M}_{\text{w}} \begin{bmatrix} a_1 & \cdots & a_{2n-k} & b_1 & \cdots & b_k \end{bmatrix}^T,$$

it is equivalent to show the following lemma (here the subscript "w" corresponds to the weak interference regime)

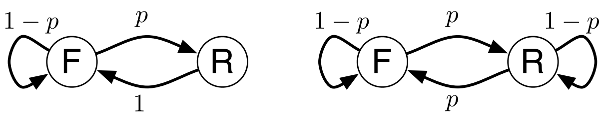
Lemma 4.1: $\mathbf{M}_{\text{w}}$ is full-rank and $\text{rank}(\mathbf{M}_{\text{w}}) = 2n$.

Proof: See Appendix of [6]. ■

Now we analyze the achievable rate $R_1 = R_2 = R_{\text{sym}}$ of the above scheme. Due to symmetry, we focus on user 1. The stationary distribution of the underlying Markov chain is:

$$\pi(\mathsf{F}) = \frac{1}{1+p}, \pi(\mathsf{R}) = \frac{p}{1+p}. \tag{4}$$

During the transition F $\to$ F, n symbols $\{a_1, \ldots, a_n\}$ are successfully decoded at Rx1. During the transition F $\to$ R,



(a) Markov Chain $k/n \leq 2$ (b) Markov Chain $k/n > 2$

Fig. 4. Markov Chains of Transmission Phases

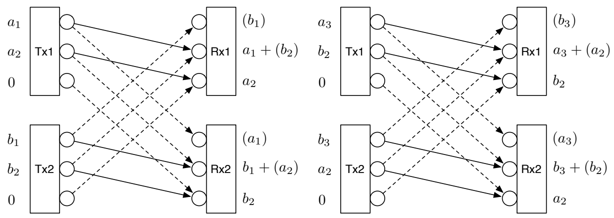
$(n-k)$ symbols $\{a_1, \ldots, a_{n-k}\}$ are decoded. During the transition R $\to$ F, n symbols $\{a_{n-k+1}, \ldots, a_{2n-k}\}$ are decoded.

Hence, the achievable rate for each user is

$$\pi(\mathsf{F}) \cdot (1-p) \cdot n + \pi(\mathsf{F}) \cdot p \cdot (n-k) + \pi(\mathsf{R}) \cdot 1 \cdot n$$
$$= n - \frac{p}{1+p} k.$$

B. *Strong Interference Regime* $1 \leq k/n < 2$

When $n \leq k < 2n$, we would like to achieve $C_{\text{sym}}^{\text{FB}} = \frac{1-p}{1+p} n + \frac{p}{1+p} k$. We shall exclude the case $n = k$ in the following discussion. The proof of achievability for this case will be dealt separately in Appendix of [6].



(a) Phase F (b) Phase R

Fig. 5. $n = 2, k = 3$. Symbols in $(\cdot)$ only appear when interference is on.

Strategy: The strategy in this regime is different from the previous one. Let us begin with redefining the two transmission phases. An illustrative example is depicted in Fig. 5.

- Phase F: Send *fresh* information only on the top n out of the total k levels. Let $\begin{bmatrix} a_1 & \cdots & a_n & 0 & \cdots & 0 \end{bmatrix}^T$ and $\begin{bmatrix} b_1 & \cdots & b_n & 0 & \cdots & 0 \end{bmatrix}^T$ denote the binary symbols sent from Tx1 and Tx2 respectively in a particular Phase F at time t.
  If there is no interference at time t, stay in Phase F in the next time slot $t+1$. If there is interference, transit to Phase R at time $t+1$. Due to the special transmission as well as the channel structure, only b_{k-n+j} interferes with a_j, for $j = 1, \ldots, 2n-k$. (See Fig. 5(b) for an illustration.)
- Phase R: Send $(k-n)$ fresh bits $\{a_{n+1}, \ldots, a_k\}$ on the top $(k-n)$ levels so that they do not interfere with Tx2's transmission. In the next $(2n-k)$ levels, send the $(2n-k)$ interfering bits $\{b_{k-n+1}, \ldots, b_n\}$. Leave the last $(k-n)$ levels empty. In summary, if there is interference in the previous time slot t under Phase F, in the current time slot $t+1$ it should switch to Phase R, and Tx1 shall send $\begin{bmatrix} a_{n+1} & \cdots & a_k & b_{k-n+1} & \cdots & b_n & 0 & \cdots & 0 \end{bmatrix}^T$.
  Regardless of the presence of interference, both transmitters transit to Phase F in the next time slot $t+2$.

The random process defined by the transmission phases is again the two-state Markov chain depicted in Fig. 4(a).

Analysis: Again, the key is to show that, in Phase R regardless of the presence of interference, at Rx1 all symbols $\{a_1, a_2, \ldots, a_k\} \cup \{b_{k-n+1}, \ldots, b_n\}$ can be decoded. Let us assume that at time t transmission phase is F while at time $t+1$ the transmission phase is R. If we represent the received linear combinations on just the *relevant* levels (that is, the bottom n levels discarding the top $(k-n)$ levels) as a vector

$$\mathbf{M}_{\text{s}} \begin{bmatrix} a_1 & \cdots & a_k & b_{k-n+1} & \cdots & b_n \end{bmatrix}^T,$$

it is equivalent to show the following lemma (here the subscript "s" corresponds to the strong interference regime)

Lemma 4.2: $\mathbf{M}_{\text{s}}$ is full-rank and $\operatorname{rank}(\mathbf{M}_{\text{s}}) = 2n$.

Proof: See Appendix of [6]. ∎

Let us analyze the achievable rate $R_1 = R_2 = R_{\text{sym}}$ of the above scheme, focusing on user 1. The stationary distribution of the underlying Markov chain is given in (4).

During the transition F $\to$ F, n symbols $\{a_1, \ldots, a_n\}$ are successfully decoded at Rx1. During the transition F $\to$ R, $(k-n)$ symbols $\{a_{2n-k+1}, \ldots, a_n\}$ are decoded. During the transition R $\to$ F, n symbols $\{a_1, \ldots, a_{2n-k}\} \cup \{a_{n+1}, \ldots, a_k\}$ are decoded.

Hence, the achievable rate for each user is

$$\pi(\mathsf{F}) \cdot (1-p) \cdot n + \pi(\mathsf{F}) \cdot p \cdot (k-n) + \pi(\mathsf{R}) \cdot 1 \cdot n$$
$$= \frac{1-p}{1+p} n + \frac{p}{1+p} k.$$

C. *Very Strong Interference Regime* $k/n \geq 2$

When $k \geq 2n$, our goal is to achieve $C_{\text{sym}}^{\text{FB}} = (1-p)n + \frac{p}{2} k$.

Strategy: Let us define the following two transmission phases.

- Phase F: Send *fresh* information only on the top $(k-n)$ out of total k levels. Let $\begin{bmatrix} a_1 & \cdots & a_{k-n} & 0 & \cdots & 0 \end{bmatrix}^T$ and $\begin{bmatrix} b_1 & \cdots & b_{k-n} & 0 & \cdots & 0 \end{bmatrix}^T$ denote the binary symbols sent from Tx1 and Tx2 respectively in a particular Phase F at time t.
  If there is no interference at time t, stay in Phase F in the next time slot $t+1$. If there is interference, transit to Phase R at time $t+1$. Again, due to the special transmission as well as the channel structure, effectively on all levels there is no interference. Hence for user 1, Phase R (see below) is to *relay* the $(k-2n)$ additional bits $\{a_{n+1}, \ldots, a_{k-n}\}$ from Tx2 to Tx1 rather than to resolve interfered symbols (and vice versa for user 2).
- Phase R: Send n fresh bits $\{a_{k-n+1}, \ldots, a_k\}$ on the top n levels. In the next $(k-2n)$ levels, send the $(k-2n)$ to-be-relayed bits $\{b_{n+1}, \ldots, b_{k-n}\}$. Leave the last n levels empty. In summary, if there is interference in the previous time slot t under Phase F, in the current time slot $t+1$ it should switch to Phase R, and Tx1 shall send $\begin{bmatrix} a_{k-n+1} & \cdots & a_k & b_{n+1} & \cdots & b_{k-n} & 0 & \cdots & 0 \end{bmatrix}^T$.

If there is no interference at the current time slot $t+1$, the to-be-relayed bits remain, and hence both transmitters remain in Phase R in the next time slot $t+2$. If there is interference, on the other hand, the to-be-relayed bits will be successfully delivered, and hence both transmitters transit back to Phase F in the next time slot.

Analysis: Based on the above operation, the random process defined by the transmission phases is the two-state Markov chain depicted in Fig. 4(b) and the stationary distribution is $\pi(\mathsf{F}) = \pi(\mathsf{R}) = 1/2$. Analyzing the achievable rate for both users of the above scheme is straightforward. In Phase F, n symbols $\{a_1, \dots, a_n\}$ are always successfully decoded at Rx1. During the transition $\mathsf{R} \to \mathsf{R}$, n symbols $\{a_{k-n+1}, \dots, a_k\}$ are decoded. During the transition $\mathsf{R} \to \mathsf{F}$, an additional $(k-2n)$ bits $\{a_{n+1}, \dots, a_{k-n}\}$ are decoded. After straightforward computation, we obtain that the achievable rate for each user $= (1-p)n + \frac{p}{2}k$.

Remark 4.1 (Asymmetric Channel): For the general asymmetric setting, the two-phase scheme no longer suffices. Instead, we need to employ a more general scheme with infinite number of phases defined by the number of to-be-resolved interference or to-be-relayed information symbols.

V. Converse Proof

The outer bounds in (1) and (2) are those of the interference channel with feedback where the transmitters know the channel state sequence S^N *non-causally*. Hence the proof follows straightforwardly from [3], and we give the details in Appendix of [6]. The proof of the non-trivial bound (3), on the other hand, needs to make use of the fact the the transmitters only knows the state S with delay. Below we provide the detailed proof of (3) for $(i,j) = (1,2)$.

For any achievable (R_1, R_2), by Fano's inequality we have

$$N(R_1 + pR_2 - \epsilon_N) \le I\left(W_1; Y_1^N, S^N\right) + pI\left(W_2; Y_2^N, S^N\right)$$

where $\epsilon_N \to 0$ as $N \to \infty$.

Let us first manipulate the term $I\left(W_1; Y_1^N, S^N\right)$: (For notational convenience, in the rest of this section, the summation "$\sum$" without specification is the summation over $t \in [1:N]$.)

$$\begin{aligned}
I\left(W_1; Y_1^N, S^N\right) &= \sum I\left(W_1; Y_1[t], S[t] \mid Y_1^{t-1}, S^{t-1}\right) \\
&= \sum I\left(W_1; Y_1[t] \mid Y_1^{t-1}, S^{t-1}, S[t]\right) \\
&\quad + \sum I\left(W_1; S[t] \mid Y_1^{t-1}, S^{t-1}\right) \\
&\overset{(a)}{=} \sum I\left(W_1; Y_1[t] \mid Y_1^{t-1}, S^{t-1}, S[t]\right) \\
&\overset{(b)}{=} \sum (1-p) I\left(W_1; \mathbf{H}_{11} X_1[t] \mid Y_1^{t-1}, S^{t-1}\right) \\
&\quad + \sum pI\left(W_1; \mathbf{H}_{11} X_1[t] + \mathbf{H}_{12} X_2[t] \mid Y_1^{t-1}, S^{t-1}\right) \\
&\le \sum (1-p) H\left(\mathbf{H}_{11} X_1[t]\right) + pH\left(\mathbf{H}_{11} X_1[t] + \mathbf{H}_{12} X_2[t]\right) \\
&\quad - \sum pH\left(\mathbf{H}_{11} X_1[t] + \mathbf{H}_{12} X_2[t] \mid Y_1^{t-1}, S^{t-1}, W_1\right) \\
&\le N\left\{(1-p)n_{11} + p\max(n_{11}, n_{12})\right\} \\
&\quad - p\sum H\left(\mathbf{H}_{11} X_1[t] + \mathbf{H}_{12} X_2[t] \mid Y_1^{t-1}, S^{t-1}, W_1\right) \\
&\overset{(c)}{=} N\left\{(1-p)n_{11} + p\max(n_{11}, n_{12})\right\} \\
&\quad - p\sum H\left(\mathbf{H}_{12} X_2[t] \mid \mathbf{H}_{12} \widetilde{X}_2^{t-1}, S^{t-1}, W_1\right).
\end{aligned}$$

Here (a) is due to the fact that $\left(W_1, Y_1^{t-1}, S^{t-1}\right)$ and $S[t]$ are independent. (b) is due to the evaluation on $S[t]$. (c) is due to the fact that ($A \overset{\mathrm{f}}{=} B$ denotes "A is a function of B")

$$\begin{aligned}
X_1[t] &\overset{\mathrm{f}}{=} \left(W_1, S^{t-1}, Y_1^{t-1}\right) \overset{\mathrm{f}}{=} \left(W_1, S^{t-1}, X_1^{t-1}, \mathbf{H}_{12} \widetilde{X}_2^{t-1}\right) \\
&\overset{\mathrm{f}}{=} \left(W_1, S^{t-1}, \mathbf{H}_{12} \widetilde{X}_2^{t-1}\right).
\end{aligned} \tag{5}$$

Next we manipulate the term $I\left(W_2; Y_2^N, S^N\right)$:

$$\begin{aligned}
I\left(W_2; Y_2^N, S^N\right) &\le I\left(W_2; Y_1^N, Y_2^N, S^N \mid W_1\right) \\
&= \sum I\left(W_2; Y_1[t], Y_2[t], S[t] \mid Y_1^{t-1}, Y_2^{t-1}, S^{t-1}, W_1\right) \\
&\overset{(d)}{=} \sum I\left(W_2; Y_1[t], Y_2[t] \mid Y_1^{t-1}, Y_2^{t-1}, S^{t-1}, W_1, S[t]\right) \\
&\overset{(e)}{=} \sum H\left(\mathbf{H}_{12} \widetilde{X}_2[t], \mathbf{H}_{22} X_2[t] \mid Y_1^{t-1}, Y_2^{t-1}, S^{t-1}, W_1, S[t]\right) \\
&= \sum (1-p) H\left(\mathbf{H}_{22} X_2[t] \mid Y_1^{t-1}, Y_2^{t-1}, S^{t-1}, W_1\right) \\
&\quad + \sum pH\left(\mathbf{H}_{12} X_2[t], \mathbf{H}_{22} X_2[t] \mid Y_1^{t-1}, Y_2^{t-1}, S^{t-1}, W_1\right) \\
&\le \sum H\left(\mathbf{H}_{12} X_2[t], \mathbf{H}_{22} X_2[t] \mid Y_1^{t-1}, S^{t-1}, W_1\right) \\
&\overset{(f)}{=} \sum H\left(\mathbf{H}_{12} X_2[t], \mathbf{H}_{22} X_2[t] \mid \mathbf{H}_{12} \widetilde{X}_2^{t-1}, S^{t-1}, W_1\right).
\end{aligned}$$

Here (d) is due to the fact that $\left(W_1, W_2, Y_1^{t-1}, Y_2^{t-1}, S^{t-1}\right)$ and $S[t]$ are independent. (e) and (f) are due to (5).

Hence, we have

$$\begin{aligned}
N(R_1 + pR_2 - \epsilon_N) &\le I\left(W_1; Y_1^N, S^N\right) + pI\left(W_2; Y_2^N, S^N\right) \\
&\le N\left\{(1-p)n_{11} + p\max(n_{11}, n_{12})\right\} \\
&\quad - p\sum H\left(\mathbf{H}_{12} X_2[t] \mid \mathbf{H}_{12} \widetilde{X}_2^{t-1}, S^{t-1}, W_1\right) \\
&\quad + p\sum H\left(\mathbf{H}_{12} X_2[t], \mathbf{H}_{22} X_2[t] \mid \mathbf{H}_{12} \widetilde{X}_2^{t-1}, S^{t-1}, W_1\right) \\
&= N\left\{(1-p)n_{11} + p\max(n_{11}, n_{12})\right\} \\
&\quad + p\sum H\left(\mathbf{H}_{22} X_2[t] \mid \mathbf{H}_{12} X_2[t], \mathbf{H}_{12} \widetilde{X}_2^{t-1}, S^{t-1}, W_1\right) \\
&\le N\left\{(1-p)n_{11} + p\max(n_{11}, n_{12})\right\} + Np(n_{22} - n_{12})^+,
\end{aligned}$$

and the bound (3) is proved.

References

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

[2] N. Khude, V. Prabhakaran, and P. Viswanath, "Opportunistic interference management," *Proceedings of IEEE International Symposium on Information Theory*, pp. 2076–2080, July 2009.

[3] C. Suh and D. N. C. Tse, "Feedback capacity of the Gaussian interference channel to within 2 bits," *IEEE Transactions on Information Theory*, vol. 57, pp. 2667–2685, May 2011.

[4] A. S. Avestimehr, S. N. Diggavi, and D. N. C. Tse, "Wireless network information flow: A deterministic approach," *IEEE Transactions on Information Theory*, vol. 57, pp. 1872–1905, April 2011.

[5] A. Vahid, M. A. Maddah-Ali, and A. S. Avestimehr, "Capacity results for binary fading interference channels with delayed CSIT," *Submitted to IEEE Transactions on Information Theory*, January 2013, http://arxiv.org/abs/1301.5309.

[6] I.-H. Wang, C. Suh, S. N. Diggavi, and P. Viswanath, "Bursty interference channel with feedback," *Preprint*, January 2013. https://sites.google.com/site/ihsiangw/isit13preprintburstyic.