Abstract—The classical interference channel models the communication limits of two independent, interfering streams of one-way data. In this paper we extend the classical interference channel model to a new channel model in which two streams of two-way data interfere with each other. In the absence of interference, this model would result in two parallel two-way channels (a four node channel); in the presence of interference it encompasses two-way, interference, and cooperation tradeoffs. The discrete memoryless “parallel two-way channel with interference” is considered, in which each of the four nodes is the source of one message, the receiver of another, and experiences interference from another in addition to its desired message. The nodes may adapt their transmissions to the past received signals in a fully two-way fashion. We present an outer bound to the four dimensional capacity region which utilizes auxiliary random variables to constrain the input distributions and comment on its relationship to existing outer bounds for related channels.

I. INTRODUCTION

The two-way channel in which two nodes exchange a single two-way data stream (or two messages) over a common channel sheds light on the ability of nodes to adapt their transmissions to past received outputs. While the two-way channel has been extended to two-way relay channels in which a single two-way data stream is exchanged with the help of a relay, little work has considered multiple two-way data streams in a network setting. Two-way networks are particularly interesting not only because they extend one-way networks to allow for the more general class of two-way data, but they also emphasize the relationship between interference, two-way data – which amounts to nodes having additional own-message side-information as well as being able to adapt their transmissions to received data – and cooperation in a network setting.

Towards this goal, we start off by considering the simplest relay-free non-trivial network which contains multiple two-way data streams that interfere. We consider two parallel two-way channels (thus we have four one-way data streams, or two two-way data streams) in which each node’s transmission is heard by its desired receiver as well as one undesired receiver.

A. Past Work

This channel model merges elements of the two-way, interference, and causally cooperative (or forms of generalized feedback) channel models, but has not been explicitly considered before. We consider the general class of discrete memoryless two-way channels first introduced by Shannon [8] who proposed inner and outer bounds of the same form/expression, but taken over independent and fully general input distributions respectively. We will use outer-bounding techniques similar to those used to improve Shannon’s outer bound, as first done by [11] for the two-way channel, and later also used in deriving outer bounds for a two-way multiple-access broadcast-channel model considered in [2]. This last channel is one of the few two-way networks considered in the literature and is most related to the channel considered here. [2] derives outer bounds on the capacity region of a three node network where the two terminal nodes each have a message for the base-station and the base-station has a common message for the terminal nodes. The nodes are half-duplex and may adapt their current transmissions to previously received signals, as in general (non-restricted) two-way channels. The channel model considered also reduces to an interference channel by forbidding node adaptation and turning two nodes into receivers only. Numerous interference channel outer bounds have been derived, including those for Gaussian channels [3], [6], strong interference discrete memoryless channels [7], discrete memoryless channels [4], with feedback [9], and generalized feedback [10].

B. Outline of Paper

We define the discrete memoryless, full-duplex, parallel two-way channel with interference channel model in Section II, derive an outer bound for this channel in Section III, compare the bound to related models in Section IV, and conclude in Section V.
II. CHANNEL MODEL

We consider a multi-user network which we term the parallel two-way channel with interference (PTW-IF), which is shown in Fig.1. This network has four distributed transmitters (encoders) and four distributed receivers (decoders), where four independent messages are communicated in this network. Transmitter 1 and 3 send messages $M_{12}$ and $M_{34}$ to receiver 2 and 4 respectively. Similarly, transmitter 2 and 4 send messages $M_{21}$ and $M_{43}$ to receiver 1 and 3 respectively. Because of the two-way feature with interference, the outputs of receiver 1 and 3 depend (possibly in a noisy fashion) on the inputs at transmitter 1,2,4 and 2,3,4 respectively. Similarly, the outputs of receiver 2 and 4 depend on the inputs at transmitter 1,2,3 and 1,3,4 respectively. We note that a more general channel model would allow for the output at each node to depend on all four inputs, but that we consider this somewhat simplified model at first given its already challenging nature, and expect to be able to carefully generalize these results to the general parallel two-way interference channel (rather than parallel two-way channel with interference).

The channel input and output of user $i \in \{1,2,3,4\}$ at discrete time $k$ are $X_{i,k}$ and $Y_{i,k}$, which lie in the alphabets $X_i$ and $Y_i$ respectively. The channel is discrete and memoryless and is thus fully described by the transition probability matrix $\{p(y_1, y_2, y_3, y_4|x_1, x_2, x_3, x_4)\}$ which we assume has the property that:

$$
Y_1 = T_1(X_1, X_2, X_4)
$$

$$
Y_2 = T_2(X_1, X_2, X_3)
$$

$$
Y_3 = T_3(X_2, X_3, X_4)
$$

$$
Y_4 = T_4(X_1, X_3, X_4)
$$

where $T_l$ ($l \in \{1,2,3,4\}$) are discrete memoryless mappings which may include an element of randomness (or noise). By introducing the notation $A_i^j = (A_{i,1}, A_{i,2}, ..., A_{i,j})$, for any given time $j$, we may describe the encoding functions which yield the channel inputs at time $j$ as follows:

$$
X_{1,j} = f_1(M_{12}, Y_1^{j-1})
$$

$$
X_{2,j} = f_2(M_{21}, Y_2^{j-1})
$$

$$
X_{3,j} = f_3(M_{34}, Y_3^{j-1})
$$

$$
X_{4,j} = f_4(M_{43}, Y_4^{j-1})
$$

where $f_i$ ($i \in \{1,2,3,4\}$) are deterministic functions. For a given blocklength $n$, at time step $0 \leq j \leq n$, the encoder selects the next input $X_{i,j} M_{ik}, Y_i^{j-1}$) out of the $2^{nR_{ik}}$ ($R_{ik} \geq 0$ is the rate of communication between sender $i$ and receiver $k$) codewords $X_i^n(M_{ik}, Y_i^{n-1})$. Receiver $k$ uses a decoding function $g_k : Y_k^n \rightarrow \hat{M}_{ik}$ to obtain an estimate $\hat{M}_{ik}$ of the transmitted message $M_{ik}$. Standard definitions for achievable rate regions and capacity regions for the rates $R_{ik}$ are used [1]. We seek outer bounds to the four-dimensional ($R_{12}$, $R_{21}$, $R_{34}$, $R_{43}$) capacity region.

III. OUTER BOUNDS REGION FOR PTW-IF

We now present an outer bound to the capacity of this channel; remarks and interpretations of this bound are found at the end of this section.

Theorem 1: The capacity region $C$ of the parallel two-way channel with interference is a subset of the region $C^*$:

$$
C^* \equiv \{(R_{12}, R_{21}, R_{34}, R_{43}) : R_{12} \leq \min\{I(X_1, Z_1; Y_2|X_2, Z_2, M_{34}, M_{43}, Q), H(X_1|Z_1, X_2, Q), I(X_2, Y_3|X_3, Z_3, Q) + I(Y_1|Z_1, X_2, Z_2, Q)\}\\
R_{21} \leq \min\{I(X_2, Z_2; Y_1|X_1, Z_1, M_{34}, M_{43}, Q), H(X_2|Z_2, X_1, Q), I(X_1, Y_4|X_4, Z_4, Q) + I(Y_1|X_1, X_4, Z_1, Q, Q)\}\\
R_{34} \leq \min\{I(X_3, Z_3; Y_4|X_4, Z_4, M_{34}, M_{42}, Q), H(X_3|Z_3, X_4, Q), I(X_2, Y_1|X_1, Z_1, Q) + I(Y_4|X_4, X_1, Z_4, Q, Z_4)\}\\
R_{43} \leq \min\{I(X_4, Z_4; Y_3|X_3, Z_3, M_{42}, M_{43}, Q), H(X_4|Z_4, X_3, Q), I(X_1, Y_2|X_2, Z_2, Q) + I(Y_3|X_3, X_2, Z_2, Q)\}\}
$$

Proof: By symmetry, we need only demonstrate bounds (1), (5), (9). For simplicity, we omit the standard time-sharing random variable arguments for $Q$ and refer the reader to standard arguments as in [1].

Proof of bound (1): We obtain three bounds on the single message rates; the first two are in the spirit of [11] for the two-way channel, while the third is reminiscent of the two-way multiple-access / broadcast channel bounds of [2].

$$
nR_{12} = H(M_{12}|M_{21}, M_{34}, M_{43})\\
= H(M_{12}|M_{21}, M_{34}, M_{43}, Y_2) + I(M_{12}; Y_2|M_{34}, M_{43}, M_{21})\\
\leq n \epsilon + I(M_{12}; Y_2|M_{34}, M_{43}, M_{21})\\
= H(Y_2|M_{34}, M_{43}, M_{21}) - H(Y_2|M_{12}, M_{21}, M_{34}, M_{43}) + n \epsilon\\
= \sum_{j=1}^{n} [H(Y_{2,j}|M_{34}, M_{43}, M_{21}, Y_2^{j-1}) - H(Y_{2,j}|M_{12}, M_{21}, M_{34}, M_{43}, Y_2^{j-1})] + n \epsilon
$$

where $X_1, X_2, X_3, X_4, Y_1, Y_2, Y_3, Y_4, Z_1, Z_2, Z_3, Z_4, Q$ are random variables subject to the following input distribution:

$$
p(z_1, z_2, z_3, z_4|q)p(x_1|z_1, q)p(x_2|z_2, q)p(x_3|z_3, q)p(x_4|z_4, q).
$$
where (a) follows from Fano’s inequality, (b) uses $X_2 = f_2(M_{21}, Y_2^{j-1})$ and conditioning reduces entropy when a “genie” provides $Y_1^{j-1}$ in the negative term. We introduce new random variables $Z_{1,j} = (X_1^{j-1}, Y_1^{j-1})$ in (c), and also use $X_1^j = f_1(M_{12}, Y_1^{j-1})$. In (d), the first term follows as conditioning reduces entropy; the second term, follows since given $M_{34}, M_{43}, X_{1,j}, X_{2,j}, Z_{1,j}, Z_{2,j}$, the channel output $Y_{2,j}$ is independent of the other terms. This is illustrated by the Markov chain diagram in Fig. 2, where we see that due to the dependence of $Y_{2,j}$ on $M_{34}$ and $M_{43}$ (since neither $X_{3,j}$ nor $Y_{3,j}$ are given) these terms may not be dropped. By introducing a time sharing random variable $Q$ and using arguments as in [1], we obtain $R_{12} \leq I(X_{1,j}^1, Z_1^1; X_{1,j}^2, Z_2^2, M_{34}, M_{43}, Q)$. One problem with this bound is the presence of the messages. We now derive two other bounds on $R_{12}$ in which this dependency on the messages is removed. To do so, notice

\[
I(M_{12}; Y_1^2 | M_{34}, M_{43}, M_{21}) \\
\leq I(M_{12}; Y_1^2, Y_1^1 | M_{21}, M_{34}, M_{43}) \\
= H(Y_1^2, Y_1^1 | M_{21}, M_{34}, M_{43}) \\
= H(Y_1^2, Y_1^1 | M_{21}, M_{34}, M_{43}) \\
= \sum_{j=1}^{n} [H(Y_{1,j}^2, Y_{1,j}^1, X_{1,j}^2, X_{1,j}^1, Z_{1,j}^1, Z_{2,j}^1) + n\epsilon] \\
= \sum_{j=1}^{n} I(X_{1,j}^1, Z_{1,j}^1; Y_{1,j}^2, X_{2,j}^2, Z_{2,j}^2, M_{34}, M_{43}) + n\epsilon
\]

where (a) uses the chain rule and $X_2^j = f_2(M_{21}, Y_2^{j-1})$, and (b) follows from: 1) the third term in (a) is zero since $X_{1,j}^1 = f_1(M_{12}, Y_1^{j-1})$. 2) The second term and the fourth term in (a) cancel each other by the Markov chain given in Fig. 3, where we see that given $X_{1,j}^1, X_{1,j}^{j-1}, X_{2,j}^j, M_{34}, M_{43}$, the channel outputs $Y_{1,j}^1$ and $Y_{1,j}^2$ are independent of anything else that is given. 3) For the first term in (a), we introduce a new random variable $Z_{1,j} = (X_1^{j-1}, Y_1^{j-1})$, and the inequality comes from conditioning reduces entropy. Again, by introducing a time sharing random variable $Q$ and using arguments as in [1], we obtain $R_{12} \leq H(X_1^1, Z_1^1, X_2^2, Q)$. 

---

**Fig. 2.** First example of the Markov chain used in Theorem 1.

**Fig. 3.** Second example of the Markov chain used in Theorem 1.
In the previous bound, $X_1$ and $Y_1$ were given as genie-aided side-information at node 2. Now we derive another outer bound on $R_{12}$ by giving $Y_3$ as genie-aided side-information at node 2. That is,

$$I(M_{12}; Y_2 | M_{34}, M_{43}, M_{21})$$

$$\leq I(M_{12}; Y_2 | M_{34}, M_{43}, M_{21})$$

$$= I(M_{12}; Y_2 | M_{34}, M_{43}, M_{21})$$

$$+ I(M_{12}; Y_2 | M_{34}, M_{43}, M_{21}, Y_3) \quad (14)$$

Now we upper bound the first term in (14) as follows:

$$I(M_{12}; Y_2 | M_{34}, M_{43}, M_{21})$$

$$= H(Y_2 | M_{34}, M_{43}, M_{21}) - H(Y_2 | M_{12}, M_{21}, M_{34}, M_{43})$$

$$\leq \sum_{j=1}^{n} [H(Y_{3,j} | M_{34}, M_{43}, M_{21}, Y_{j}^{3j-1})$$

$$- H(Y_{3,j} | M_{12}, M_{21}, M_{34}, M_{43}, Y_{j}^{3j-1})]$$

$$\leq \sum_{j=1}^{n} [H(Y_{3,j} | M_{34}, M_{43}, M_{21}, Y_{j}^{3j-1}, X_2^j)$$

$$- H(Y_{3,j} | M_{12}, M_{21}, M_{34}, M_{43}, Y_{j}^{3j-1}, X_2^j, X_4^j, X_3^j)]$$

$$\leq \sum_{j=1}^{n} I(X_{2,j}, X_{4,j}, Y_{j} | X_3^j, Z_{3,j}) \quad (15)$$

where (a) uses $X_2^j = f_2(M_{34}; Y_{j}^{3j-1})$, and the inequality holds as we added $X_2^j, X_3^j$ in the negative term (thereby reducing it). In (b) we introduce a new random variable $Z_{3,j} = (X_2^{j-1}, Y_{j}^{3j-1})$. In (c), for the first term, conditioning reduces entropy. For the second term, we again use the iterated Markov chain properties similar to those in Figs. 2 and 3 to see that given $X_{2,j}, X_{3,j}, X_{4,j}$, the channel output $Y_{3,j}$ is independent of all other terms. Now we proceed to upper bound the second term in (14):

$$I(M_{12}; Y_2 | M_{34}, M_{43}, M_{21}, Y_3)$$

$$= H(Y_2 | M_{34}, M_{43}, M_{21}, Y_3)$$

$$- H(Y_2 | M_{12}, M_{21}, M_{34}, M_{43}, Y_4)$$

$$\leq \sum_{j=1}^{n} [H(Y_{2,j} | M_{34}, M_{43}, M_{21}, Y_{2j}^{3j-1}, Y_3)$$

$$- H(Y_{2,j} | M_{12}, M_{21}, M_{34}, M_{43}, Y_{2j}^{3j-1}, Y_3)]$$

$$\leq \sum_{j=1}^{n} [H(Y_{2,j} | M_{34}, M_{43}, M_{21}, Y_{2j}^{3j-1}, X_2^j, Y_3^{3j-1}, X_3^j)]$$

$$- H(Y_{2,j} | M_{12}, M_{21}, M_{34}, M_{43}, Y_{2j}^{3j-1}, X_2^j, Y_3^{3j-1}, X_3^j, Y_4^j, Z_{2,j}, Z_{4,j})] \quad (16)$$
where (a) follows from Fano’s inequality, (b) uses $X_d^j = f_2(M_{21}, Y_{d}^{j-1})$ and the inequality follows from giving the terms $X_1^j, X_2^j$ as genie-aided information in the negative term. We introduce new random variables $Z_{i,j} = (X_i^{j-1}, Y_{d}^{j-1})$ in (c). In (d), for the first term, conditioning reduces entropy. For the second term, we use Markov chain arguments similar to those in Figs. 2 and 3 to see that given $X_{1,j}, X_{2,j}, X_{3,j}, X_{4,j}$, the channel outputs $Y_{2,j}, Y_{4,j}$ are independent of anything else that is given. By introducing a time sharing random variable $Q$ and using arguments as in [1], we obtain $R_{12} + R_{21} + R_{34} \leq I(X_1, X_3; Y_2, Y_4 | X_2, X_4, Z_2, Z_4, Q)$. That the channel input distribution splits according to (13) is omitted due to space, but follows the lines of [11].

Remark 2: The derivation of the bound closely followed the ideas of [11] for the point-to-point two-way channel as well as those of [2] for the two-way multiple-access and broadcast channel with a common message. The main difference is the structure of our channel, which contains interference but no multiple-access or broadcast elements. In addition, our bounds are derived in full-duplex scenario, which is different from [2]'s half-duplex model.

Remark 3: We note that the double and triple rate bounds (5) – (12) are quite intuitive and follow in a relatively straightforward manner; the key improvement over other cut-set like outer bounds is the somewhat more restrained input distribution due to the auxiliary $Z$ random variables over which this bound is taken. In [11] three outer bounds are presented: a potentially tighter one with auxiliary random variables without bounds on their cardinality – similar in spirit to those presented here – and two potentially looser bounds which are more computationally feasible as their associated auxiliary random variables have cardinality bounds. Altering the above bound in order to bound the cardinality of the auxiliary random variables in terms of the input alphabet cardinality is the subject of ongoing work.

Remark 4: The third term in bound (1) suggests regarding users 3 and 4 as relay nodes for the transmission from user 1 to receiver 2. Indeed, in the PTW-IF, the message $M_{12}$ may be transmitted from 1 to 2 directly, or may intuitively be “routed” through 1 → 2 → 3 → 2, in a cooperative fashion. This ability to cooperate is captured by the third bound in (1).

IV. DISCUSSION

The PTW-IF is a fairly general channel model which combines elements of two-way channels with those of interference channels. As such, we expect an outer bound for this channel to relate to outer bounds for similar channel models and briefly discuss this next.

A. The cut-set outer bound for the PTW-IF

A readily available outer bound to this channel model (may be added to that of Theorem 1) is the standard cut-set outer bound [1], [5] , which, for the PTW-IF, is given by the region:

\begin{align*}
R_{12} & \leq I(X_1; Y_2, Y_4 | X_2, X_4) & \text{(18)} \\
R_{21} & \leq I(X_2; Y_1, Y_3 | X_1, X_3) & \text{(19)} \\
R_{34} & \leq I(X_3; Y_2, Y_4 | X_2, X_4) & \text{(20)} \\
R_{43} & \leq I(X_4; Y_1, Y_3 | X_1, X_3) & \text{(21)} \\
R_{12} + R_{34} & \leq I(X_1, X_3; Y_2, Y_4 | X_2, X_4) & \text{(22)} \\
R_{21} + R_{43} & \leq I(X_2, X_4; Y_1, Y_3 | X_1, X_3) & \text{(23)} \\
R_{12} + R_{34} & \leq I(X_1, X_3; Y_2, Y_4 | X_1, X_4) & \text{(24)} \\
R_{21} + R_{43} & \leq I(X_2, X_4; Y_1, Y_3 | X_2, X_3) & \text{(25)}
\end{align*}
where \( X_1, X_2, X_3, X_4 \) follow the fully general input distribution of \( p(x_1, x_2, x_3, x_4) \). Although it is hard to compare the single-rate bounds (18) – (21) with our bounds (1) – (4), our sum-rate bounds (5) – (12) may intuitively potentially improve upon the cut-set bounds (22) – (25) due to the presence of auxiliary random variables \( Z_i, i \in \{1, 2, 3, 4\} \) and the more constrained input distributions, though this is in general an open problem and the subject of ongoing work.

B. Comparison with Sato’s [4] outer bound for the interference channel

Sato’s outer bound to the capacity region of the interference channel derived in Theorem 2 of [4] (and paraphrased from [6]) is given by the region:

\[
\mathcal{R}^* = \{(R_1, R_2) : \\
R_1 \leq I(X_1; \tilde{Y}_1|X_2), \\
R_2 \leq I(X_2; \tilde{Y}_2|X_1), \\
R_1 + R_2 \leq I(X_1, X_2; \tilde{Y}_1, \tilde{Y}_2)\},
\]

where \( X_1 \) and \( X_2 \) are independent and \( \tilde{Y}_1 \) and \( \tilde{Y}_2 \) are any random variables satisfying \( p(\tilde{y}_1|x_1, x_2) = p(y_1|x_1, x_2) \) and \( p(\tilde{y}_2|x_1, x_2) = p(y_2|x_1, x_2) \) (same marginals as original channel) for all \( x_1, x_2, y_1, y_2 \).

The bounds of Theorem 1 resemble Sato’s bounds for specific choices of random variables. In particular, equating \( R_1 \equiv R_{12}, R_2 \equiv R_{34}, R_{21} = R_{43} = 0 \), setting \( X_2, X_4, Y_1, Y_3, M_{21}, M_{43} = \emptyset \), and noting that in an interference channel the auxiliary random variables used for feedback/two-way \( Z_1, Z_2, Z_3, Z_4 \) do not exist, the bound in (1) reduces to \( R_{12} \leq \min\{I(X_1; Y_2|M_{34}), H(X_1), I(X_1; Y_2|X_3)\} \), where \( I(X_1; Y_2|X_3) \) resembles Sato’s bound \( I(X_1; \tilde{Y}_1|X_2) \). By symmetry, we may obtain \( R_{34} \leq \min\{I(X_3; Y_4|M_{12}), H(X_3), I(X_3; Y_4|X_1)\} \). Our bound (5) resembles Sato’s bound \( R_1 + R_2 \leq I(X_1, X_2; \tilde{Y}_1, \tilde{Y}_2) \). Our bounds are in terms of the original rather than auxiliary “same marginals” output random variables; whether Sato’s “same marginals” idea may be extended to this channel is the subject of ongoing work.

C. Comparison with Zhang and Berger’s [11] outer bound for the two-way channel

Theorem 1 of [11] yields the following outer bound to the capacity region \( \mathcal{R} \) of the discrete memoryless two-way channel:

\[
\mathcal{R}^* = \{(R_1, R_2) : \\
R_1 \leq \min\{H(X_1|Z_1), I(X_1; Y_2|X_2, Z_2)\}, \\
R_2 \leq \min\{H(X_2|Z_2), I(X_2; Y_1|X_1, Z_1)\}\}
\]

where \( Z_1, Z_2, X_1, X_3, Y_1, Y_2 \) are random variables whose joint distribution is of the form \( p(z_1, z_2)p(x_1|z_1)p(x_2|z_2)p(y_1, y_2|x_1, x_2) \).

As the outer bounds of Theorem 1 are based on the techniques of [11] it is natural to expect that they would reduce to [11] for a single two-way channel. Indeed, our outer bound region may be reduced to the above region by setting \( X_3, X_4, Y_3, Y_4, M_{34}, M_{43} = \emptyset \). We note that the first term in (1) and (2) is always larger than the third term (and thus the third term dominates the \( \min\{\cdot, \cdot, \cdot\} \)), and the third term corresponds to the outer bound of [11]. We note that our bounds \( R_{12} \leq H(X_1|Z_1, X_2) \) and \( R_{21} \leq H(X_2|Z_2, X_1) \) are no looser than Zhang and Berger’s bounds \( R_1 \leq H(X_1|Z_1) \) and \( R_2 \leq H(X_2|Z_2) \).

V. Conclusion

In this paper we proposed an outer bound region for the discrete memoryless parallel two-way channel with interference which utilizes auxiliary random variables to constrain the input distribution. A number of questions are the subject of ongoing work, including: 1) whether we can obtain cardinality bounds on the auxiliary random variables, 2) if/when is this bound tighter than the cut-set bound, and 3) whether it may be tightened for specific channel models including deterministic, linear high-SNR deterministic, or Gaussian channels.

References


Acknowledgment

The work of Z. Cheng and N. Devroye was partially supported by NSF under award 1053933. The contents of this article are solely the responsibility of the authors and do not necessarily represent the official views of the NSF.