

On Dependence Balance Bounds for Two Way Channels

Ravi Tandon Sennur Ulukus
Department of Electrical and Computer Engineering
University of Maryland, College Park, MD 20742
ravit@umd.edu ulukus@umd.edu

Abstract— The idea that no more dependence can be consumed than produced was used to obtain an outer bound to the capacity region of the single output two-way channel in [1]. A parallel channel extension of this bound was used to obtain an outer bound of 0.64628 bits/transmission for the sum-rate point of the binary multiplying channel (BMC) which improved over Zhang-Berger-Schalkwijk bound of 0.64891 bits/transmission [2]. It is not obvious that this selection of parallel channel would yield the best sum-rate outer bound for the BMC using the idea of dependence balance. We show that for all binary adaptive parallel channels, the best sum-rate outer bound for the BMC is 0.64628 bits/transmission and corresponds to the parallel channel selection of [1].

I. INTRODUCTION

The two-way channel (TWC) is the oldest multi-user channel for which finding the information theoretic capacity (\mathcal{C}_{TWC}) still remains an open problem. The TWC was introduced by Shannon [3] in 1961 where he gave inner and outer bounds along with a limiting expression for the capacity region. The limiting expression has the disadvantage of being in the form of an n -letter expression which makes it hard to compute even for simple channels. The binary multiplying channel (BMC) is a binary, deterministic TWC where all alphabets $\mathcal{X}_1 = \mathcal{X}_2 = \mathcal{Y} = \{0, 1\}$ are binary and both the transmitters receive a common output Y where the channel operation is defined by $y = x_1 x_2$. Shannon showed in [3] that his inner and outer bounds for the TWC differ for the BMC. The difference between Shannon's inner and outer bounds for the BMC is the largest for the sum-rate point.

The first attempt to improve upon Shannon's outer bound was made by Zhang, Berger and Schalkwijk [2]. The authors gave a new outer bound [2, Theorem 1] which was in general better than Shannon's outer bound. However, this bound was incomputable due to the inability to bound the cardinality of the involved auxiliary random variables. Hence the authors gave two loose but computable versions [2, Theorems 2 and 3] of their main result. These two computable bounds were shown to be strictly better than Shannon's sum-rate outer bound for the BMC.

Hekstra and Willems [1] obtained a new outer bound on the capacity region of the single output TWC. By relating to the fact that each code must produce the dependence it consumes, the authors named their bound as the dependence balance (DB) bound. This outer bound was different from previous converse techniques in the sense that it limited the correlation between

the channel inputs by giving a non-trivial bound which every allowable input distribution had to satisfy. The authors also developed fixed and adaptive parallel channel extensions of their DB bound. The parallel channel was insinuated to reduce dependence production without introducing excessive information leaks.

It was shown in [1] that the adaptive parallel channel extension of DB resulted in a better outer bound than the computable outer bounds of Zhang, Berger and Schalkwijk [2] for the sum-rate point of the BMC. This improvement was obtained via a particular selection of the adaptive parallel channel which minimized the sum of information leaks.

The introduction of a parallel channel creates the possibility of a tradeoff between the set of allowable input distributions and the excessive rate terms which arise due to the parallel channel output. Hekstra and Willems' selection of the parallel channel for the sum-rate of the BMC corresponds to one extreme of this rate-distribution tradeoff. More specifically, their parallel channel selection corresponds to the smallest set of allowable input distributions. It is not immediately clear whether this is the optimum choice of a parallel channel to obtain the best sum-rate outer bound for the BMC using the idea of dependence balance. Hekstra and Willems [4] posed a question about the existence of other parallel channels which would yield a better outer bound than 0.64628 bits/transmission for the sum-rate point of the BMC. We partially answer this question. We show that among all binary adaptive parallel channels, the adaptive parallel channel selection of [1] yields the best sum-rate outer bound for the BMC. Moreover, for all ternary adaptive parallel channels which operate at the two extremes of the tradeoff, i.e., one where the feasible set of probability distributions is minimized, and the other where the leak terms are minimized, Hekstra and Willems' parallel channel yields the best sum-rate outer bound for the BMC. The unexplored part of this rate-distribution tradeoff for parallel channels with cardinality $|\mathcal{Z}| \geq 3$ remains to be studied.

II. THE SINGLE OUTPUT TWC

A discrete memoryless single output TWC consists of two input alphabets $\mathcal{X}_1, \mathcal{X}_2$, an output alphabet \mathcal{Y} and a transition probability matrix $p(y|x_1, x_2)$ for $x_1 \in \mathcal{X}_1, x_2 \in \mathcal{X}_2$ and $y \in \mathcal{Y}$; see Figure 1.

A $(n, M_1, M_2, P_{e1}, P_{e2})$ code for the single output TWC consists of two sets of encoding functions $f_{1i} : \mathcal{M}_1 \times$

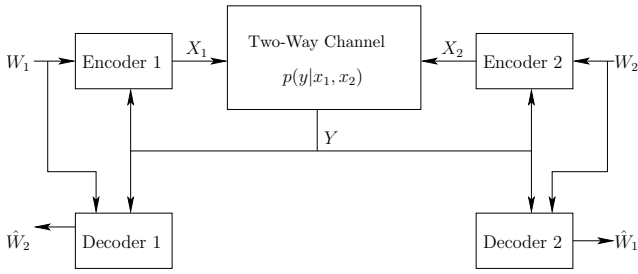


Figure 1: The single-output TWC.

$\mathcal{Y}^{i-1} \rightarrow \mathcal{X}_1$, $f_{2i} : \mathcal{M}_2 \times \mathcal{Y}^{i-1} \rightarrow \mathcal{X}_2$ for $i = 1, \dots, n$ and two decoding functions $g_1 : \mathcal{M}_1 \times \mathcal{Y}^n \rightarrow \mathcal{M}_2$ and $g_2 : \mathcal{M}_2 \times \mathcal{Y}^n \rightarrow \mathcal{M}_1$. The two transmitters produce independent and uniformly distributed messages $W_1 \in \{1, \dots, M_1\}$ and $W_2 \in \{1, \dots, M_2\}$, respectively, and transmit them through n channel uses. The average error probabilities are defined as $P_{ei} = Pr(\hat{W}_i \neq W_i)$, for $i = 1, 2$.

A rate pair (R_1, R_2) is achievable for the single output TWC if and only if for all $\epsilon > 0$, there exists an $(n, M_1, M_2, P_{e1}, P_{e2})$ code with $\frac{\log(M_i)}{n} \geq R_i - \epsilon$ and $P_{ei} \leq \epsilon$ for $i = 1, 2$. The capacity region \mathcal{C}_{TWC} of the single output TWC is the set of all achievable rates (R_1, R_2) .

III. OUTER BOUNDS FOR SINGLE OUTPUT TWC

Shannon's outer bound (S_O) for the single output TWC is given by

$$S_O = \{(R_1, R_2) : R_1 \leq I(X_1; Y|X_2) \\ R_2 \leq I(X_2; Y|X_1)\} \quad (1)$$

where the random variables X_1, X_2 and Y have the joint distribution

$$p(x_1, x_2, y) = p(x_1, x_2)p(y|x_1, x_2). \quad (2)$$

Shannon's outer bound allows for arbitrary input distributions $p(x_1, x_2)$.

Zhang, Berger and Schalkwijk [2] improved upon S_O by imposing conditions on (X_1, X_2) intermediate between independence and complete generality. The two computable outer bounds obtained in [2] are:

$$ZBS_O^{(I)} = \{(R_1, R_2) : R_1 \leq \min(H(X_1|T), I(X_1; Y|X_2)) \\ R_2 \leq I(X_2; Y|X_1, T)\} \quad (3)$$

$$ZBS_O^{(II)} = \{(R_1, R_2) : R_1 \leq I(X_1; Y|X_2, T) \\ R_2 \leq \min(H(X_2|T), I(X_2; Y|X_1))\} \quad (4)$$

where the random variables X_1, X_2, Y and T have the joint distribution

$$p(t, x_1, x_2, y) = p(t)p(x_1|t)p(x_2|t)p(y|x_1, x_2) \quad (5)$$

where T is subject to a cardinality bound $|T| \leq |\mathcal{X}_1||\mathcal{X}_2| + 1$.

In [1], Hekstra and Willems developed the following DB

outer bound on the capacity region of the single output TWC:

$$DB = \{(R_1, R_2) : R_1 \leq I(X_1; Y|X_2, T) \\ R_2 \leq I(X_2; Y|X_1, T)\} \quad (6)$$

where the random variables X_1, X_2, Y and T have the joint distribution

$$p(t, x_1, x_2, y) = p(t, x_1, x_2)p(y|x_1, x_2) \quad (7)$$

and also satisfy the following DB bound

$$I(X_1; X_2|T) \leq I(X_1; X_2|Y, T) \quad (8)$$

where T is subject to a cardinality bound $|T| \leq 3$.

In [1], Hekstra and Willems also developed an adaptive parallel channel extension of the DB bound which is given as: let $\Delta(\mathcal{U})$ denote the set of all distributions of U and $\Delta(\mathcal{U}|\mathcal{V})$ denote the set of all conditional distributions of U given V , then for any mapping $F : \Delta(\mathcal{X}_1 \times \mathcal{X}_2) \rightarrow \Delta(\mathcal{Z}|\mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{Y})$, we have $\mathcal{C}_{TWC} \subset DB_{PC}$ where

$$DB_{PC} = \{(R_1, R_2) : R_1 \leq I(X_1; Y, Z|X_2, T) \\ R_2 \leq I(X_2; Y, Z|X_1, T) \\ R_1 \leq I(X_1; Y|X_2) \\ R_2 \leq I(X_2; Y|X_1)\} \quad (9)$$

where the random variables X_1, X_2, Y and T have the joint distribution

$$p(t, x_1, x_2, y, z) = p(t, x_1, x_2)p(y|x_1, x_2)p^+(z|x_1, x_2, y, t) \quad (10)$$

such that for all t ,

$$p^+(z|x_1, x_2, y, t) = F(p_{X_1, X_2}(x_1, x_2|t)) \quad (11)$$

and such that

$$I(X_1; X_2|T) \leq I(X_1; X_2|Y, Z, T) \quad (12)$$

where T is subject to a cardinality bound $|T| \leq |\mathcal{X}_1||\mathcal{X}_2| + 3$. The idea of introducing a parallel channel can be visualized as in Figure 2.

This parallel channel extension of the DB bound can be thought of in terms of a tradeoff as follows:

- 1) The term $I(X_1; X_2|Y, Z, T)$ on the right hand side of (12) represents the tightness of the restriction of the DB bound. This term is also referred to as the "dependence output" in [4].
- 2) $L_1 = I(X_1; Z|X_2, Y, T)$ and $L_2 = I(X_2; Z|X_1, Y, T)$ represent the rate leak terms at the two transmitters as a consequence of the parallel channel output Z . These information leaks are the additional terms that appear in the rate terms in (9) with respect to the rate terms in (6) due to the introduction of a parallel channel.

To obtain a "good" outer bound for the single output TWC, one would like to choose a parallel channel Z which simultaneously achieves small values of $I(X_1; X_2|Y, Z, T)$ (corresponding to a small set of allowable input distributions $p(t, x_1, x_2)$) as well as small values of L_1 and L_2 .

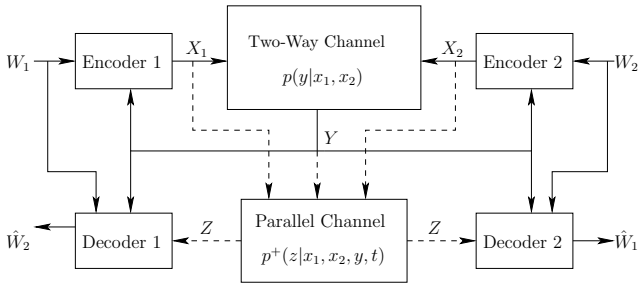


Figure 2: A parallel channel for the single-output TWC.

IV. OUTER BOUNDS FOR SUM-RATE POINT OF THE BMC

The BMC is the only deterministic binary TWC for which Shannon's inner and outer bounds do not coincide as was shown in [5]. Being one of the simplest, deterministic TWCs, the BMC encapsulates all the complexities of the general TWC. It is hence important to fully understand the capacity region of the BMC which in turn might give ideas for proving converse/achievability results for the general single output TWC. Some of the known facts about the sum-rate point of the BMC can be summarized as follows.

- 1) $S_O = 0.69424$ bits/transmission.
- 2) $DB = 0.69424$ bits/transmission.
- 3) $ZBS_O = ZBS_O^{(I)} \cap ZBS_O^{(II)} = 0.64891$ bits/transmission.

It should also be noted that using DB_{PC} , one can obtain the above results by making appropriate choices of the adaptive parallel channel defined by $p^+(z|x_1, x_2, y, t)$ as follows:

- 1) $Z = c$, where c is a constant, yields DB .
- 2) $Z = X_1$ yields $ZBS_O^{(I)}$.
- 3) $Z = X_2$ yields $ZBS_O^{(II)}$.

An improvement over ZBS_O was obtained in [1] by selecting an adaptive parallel channel which is "less informative" than X_1 or X_2 ; see Table I. The parameter p_t when optimized to minimize the sum-leak yields an outer bound of 0.64628 bits/transmission on the sum-rate of the BMC. At this point, it should be noted that the choices of parallel channels in Table I as well as $Z = X_1$ and $Z = X_2$ make the mutual information $I(X_1; X_2|Y, Z, T) = 0$ and hence only allow such input distributions which satisfy $X_1 \rightarrow T \rightarrow X_2$ corresponding to zero "dependence output".

For the choices of $p^+(z|x_1, x_2, y, t)$ where $Z = X_1$, $Z = X_2$ and Z in Table I, the DB bound operates at one extreme of the rate-distribution tradeoff mentioned in the previous section, namely the case where $I(X_1; X_2|Y, Z, T) = 0$. Thus, a natural question arises that can we find such parallel channels which operate at some other point in this tradeoff and yield a better sum-rate outer bound for the BMC. It was posed as a question in [4] whether there are parallel channels which allow non-markov distributions, i.e., corresponding to non-zero "dependence output" and yield a better sum-rate outer bound than 0.64628 bits/transmission. In this paper, we partially answer this question. More specifically, we show that for all binary adaptive parallel channels, and all ternary adaptive parallel channels for which $I(X_1; X_2|Y, Z, T) = 0$, the parallel channel given in Table I yields the best sum-rate outer bound of 0.64628 bits/transmission for the BMC.

TABLE I

ADAPTIVE PARALLEL CHANNEL ASSIGNMENT OF [1] FOR THE BMC

(X_1, X_2, Y)	0	1	2	Z
$(0,0,0)$	0	p_t	$1 - p_t$	
$(0,1,0)$	0	1	0	
$(1,0,0)$	0	0	1	
$(1,1,1)$	1	0	0	$p^+(z x_1, x_2, y, t)$

V. PARALLEL CHANNELS WITH ZERO DEPENDENCE OUTPUT (OR $I(X_1; X_2|Y, Z, T) = 0$)

We will first look at one end of the distribution-rate trade-off: set of all binary adaptive parallel channels which yield $I(X_1; X_2|Y, Z, T) = 0$ or zero "dependence output". First note the fact that

$$\begin{aligned}
 I(X_1; X_2|Y, Z, T) &= \sum_t p(t) I(X_1; X_2|Y, Z, T = t) \\
 &= \sum_t p(t) p(y=0|t) I(X_1; X_2|Y=0, Z, T=t) \quad (13)
 \end{aligned}$$

where (13) holds because when $Y = 1$, we know that $(X_1, X_2) = (1, 1)$, hence $I(X_1; X_2|Y=1, Z, T) = 0$ for any choice of $p^+(z|x_1, x_2, y, t)$ and consequently we only need to consider the case when $Y = 0$. As a result of (13), the condition $I(X_1; X_2|Y, Z, T) = 0$ is equivalent to $I(X_1; X_2|Y=0, Z, T=t) = 0$ for all $t \in \mathcal{T}$. Consider an arbitrary binary adaptive parallel channel $p^+(z|x_1, x_2, y, t)$ which is of the form described in Table II for $Y = 0$ and a fixed $T = t$.

Assume the input distribution $p(x_1, x_2|t)$ as follows: $p((x_1, x_2) = (0, 0)|t) = a_t$, $p((x_1, x_2) = (0, 1)|t) = b_t$, $p((x_1, x_2) = (1, 0)|t) = c_t$ and $p((x_1, x_2) = (1, 1)|t) = d_t$ such that all $0 < a_t, b_t, c_t, d_t < 1$ and $a_t + b_t + c_t + d_t = 1$ for all $t \in \mathcal{T}$. The condition $I(X_1; X_2|Y=0, Z, T=t) = 0$ for all $t \in \mathcal{T}$ is equivalent to

$$\begin{aligned}
 p(x_1, x_2|Y=0, z, T=t) &= \\
 p(x_1|Y=0, z, T=t)p(x_2|Y=0, z, T=t) \quad (14)
 \end{aligned}$$

which gives us the following constraints on (p_t, q_t, r_t) : $b_t c_t q_t (1 - r_t) = b_t c_t (1 - q_t) r_t = 0$. This means that for $b_t, c_t > 0$, the constraints are $q_t (1 - r_t) = (1 - q_t) r_t = 0$ for all $t \in \mathcal{T}$. The only possible pairs (q_t, r_t) satisfying these conditions are $(q_t, r_t) = (0, 0)$ or $(q_t, r_t) = (1, 1)$ which is equivalent to the choice of parallel channel assignment of [1] described in Table I. We have thus shown that among all binary adaptive parallel channels, the choices of $p^+(z|x_1, x_2, y, z, t)$ for which $I(X_1; X_2|Y, Z, T) = 0$ take the form described in Table I. Choosing $p_t = 1$ for all $t \in \mathcal{T}$ yields $Z = X_1$ whereas $p_t = 0$ for all $t \in \mathcal{T}$ yields $Z = X_2$. The optimum p_t which minimizes L is given by

$$p_t^* = \frac{\phi_{1t}(1 - \phi_{2t})}{\phi_{1t}(1 - \phi_{2t}) + \phi_{2t}(1 - \phi_{1t})} \quad (15)$$

and was obtained in [1], where $\phi_{1t} = \Pr(X_1 = 0|T = t)$ and $\phi_{2t} = \Pr(X_2 = 0|T = t)$.

TABLE II
AN ARBITRARY PARALLEL CHANNEL FOR THE BMC WITH $|\mathcal{Z}| = 2$

(X_1, X_2, Y)	1	2	Z
(0,0,0)	p_t	$1 - p_t$	
(0,1,0)	q_t	$1 - q_t$	
(1,0,0)	$1 - r_t$	r_t	$p^+(z x_1, x_2, y = 0, t)$

VI. PARALLEL CHANNELS FOR BMC WITH ZERO SUM-LEAK

We will now look at the other end of the distribution-rate tradeoff. Consider all such $p^+(z|x_1, x_2, y, t)$ for which $L = I(X_1; Z|X_2, Y, T) + I(X_2; Z|X_1, Y, T) = 0$. This is equivalent to the following separate constraints: $I(X_1; Z|X_2, Y, T) = 0$ and $I(X_2; Z|X_1, Y, T) = 0$ by the non-negativity of mutual information.

For the BMC, since $Y = 1$ implies $(X_1, X_2) = (1, 1)$, we will only be concerned for the case when $Y = 0$. By similar arguments used in the previous section, the two constraints can be reduced to $I(X_1; Z|X_2, Y = 0, T = t) = 0$ and $I(X_2; Z|X_1, Y = 0, T = t) = 0$ for all $t \in \mathcal{T}$. It can be easily seen that the only parallel channel $p^+(z|x_1, x_2, y, t)$ satisfying both constraints simultaneously is such that Z is independent of (X_1, X_2) given $(Y = 0, T = t)$ for every $T = t$. For such a parallel channel, the resulting DB bound is $I(X_1; X_2|T) \leq I(X_1; X_2|Y, T)$ which yields (6). Hence for all parallel channels with zero sum-leak, the sum-rate outer bound is the same as Shannon's outer bound of 0.69424 bits/transmission.

VII. PARALLEL CHANNELS WITH NON-ZERO DEPENDENCE OUTPUT

In this section, we consider adaptive parallel channels which produce non-zero "dependence output" and in general produce non-zero rate leak terms. It can be shown that for a fixed input distribution $p(t, x_1, x_2)$, the sum-leak $L(p^+(z|x_1, x_2, y, t))$ is a convex function of $p^+(z|x_1, x_2, y, t)$ although the constraint $I(X_1; X_2|T) - I(X_1; X_2|Y, Z, T) \leq 0$ is not necessarily convex in $p^+(z|x_1, x_2, y, t)$. Thus the overall optimization problem is not necessarily convex and obtaining a complete analytical insight is hard. Hence we will take a "heuristic-analytical" approach as suggested in [4] to show that for all binary adaptive parallel channels with $I(X_1; X_2|Y, Z, T) > 0$, the best sum-rate outer bound for the BMC is 0.64628 bits/transmission. To show that a parallel channel yields a sum-rate outer bound greater than or equal to 0.64628 bits/transmission, it suffices to find a distribution $p(t, x_1, x_2)$ which satisfies the DB bound and yields a sum-rate greater than or equal to 0.64628 bits/transmission. Consider an arbitrary binary adaptive parallel channel parameterized by (p_t, q_t, r_t) as described in Table II.

It was shown in [4] that the distribution which maximizes Shannon's sum-rate outer bound for the BMC is: $p(x_1 = 0, x_2 = 0) = 0$ and $p(x_1 = 0, x_2 = 1) = p(x_1 = 1, x_2 = 0) = 0.276$. We know that this distribution yields a sum-rate of 0.69424 bits/transmission. If this distribution satisfies the

DB bound for a given $p^+(z|x_1, x_2, y, t)$, then it implies that this distribution is not being eliminated by the DB constraint and hence for such parallel channels, the resulting outer bound would be the same as Shannon's outer bound for the sum-rate of BMC. For this distribution we only need to check the dependence balance difference when $|\mathcal{T}| = 1$ which is denoted as:

$$\Delta(q_t, r_t) = I(X_1; X_2|Y, Z) - I(X_1; X_2) \quad (16)$$

This is illustrated in Figure 3 and it shows that for (q_t, r_t) such that $\Delta(q_t, r_t) \geq 0$, the DB bound is satisfied thus the corresponding parallel channels parameterized with such (q_t, r_t) pairs will yield Shannon's sum-rate outer bound. Using this idea, most of the adaptive parallel channels with non-zero dependence output are eliminated. The remaining pairs (q_t, r_t) for which $\Delta^+(q_t, r_t) < 0$ are given by the following equation:

$$\alpha \left((1 - r + q)h \left(\frac{q}{1 - r + q} \right) + (1 - q + r)h \left(\frac{1 - q}{1 - q + r} \right) \right) - h(\alpha) + (1 - \alpha)h \left(\frac{\alpha}{1 - \alpha} \right) > 0 \quad (17)$$

where $\alpha = 0.276$. For such pairs, consider the input distribution $p^*(t, x_1, x_2)$ which maximizes Hekstra and Willems' DB bound. This distribution was obtained in [4] as follows:

$$\begin{aligned} |\mathcal{T}| = 3, \quad p_T(0) = p_T(1) &= 0.164366 \\ p_{X_1|T}(0|0) = p_{X_2|T}(0|1) &= 0.031492 \\ p_{X_1|T}(0|1) = p_{X_2|T}(0|0) &= 0.601822 \\ p_{X_1|T}(0|2) = p_{X_2|T}(0|2) &= 0.289736 \end{aligned}$$

For all pairs (q_t, r_t) given by (17), an exhaustive search over ϵ -vicinity of this distribution, where $\epsilon = 0.01$, yields a sum-rate bound greater than or equal to 0.64628 bits/transmission. An exhaustive search over ϵ -vicinity of the distribution $p^*(t, x_1, x_2)$ is a maximization over those $p(t, x_1, x_2)$ such that

$$|p(t, x_1, x_2) - p^*(t, x_1, x_2)| \leq \epsilon \quad (18)$$

subject to the constraint that $\sum_{(t, x_1, x_2)} p(t, x_1, x_2) = 1$. For every pair (q_t, r_t) , the parameter p_t is chosen as the argument which minimizes the sum-leak, uniqueness of which is justified by the convexity of $L(p^+(z|x_1, x_2, t))$ with respect to $p^+(z|x_1, x_2, t)$. Thus for all binary adaptive parallel channels which yield a non-zero "dependence output", the best sum-rate outer bound is 0.64628 bits/transmission.

VIII. PARALLEL CHANNELS WITH $|\mathcal{Z}| = 3$ AND $I(X_1; X_2|Y, Z, T) = 0$

In previous sections, we considered parallel channels with $|\mathcal{Z}| = 2$. The DB bound can potentially be improved upon by considering parallel channels with larger cardinalities. For the case when $|\mathcal{Z}| = 3$, the parallel channel assignment $p^+(z|x_1, x_2, y, t)$ is of the form described in Table III. Imposing the condition (14), we get the following three constraints $bcq_{1t}r_{1t} = bcq_{2t}r_{2t} = bc(1 - q_{1t} - q_{2t})(1 - r_{1t} - r_{2t}) = 0$. The only non-degenerate case which is not equivalent to $|\mathcal{Z}| = 2$ is described in Table IV. For this parallel channel, the individual

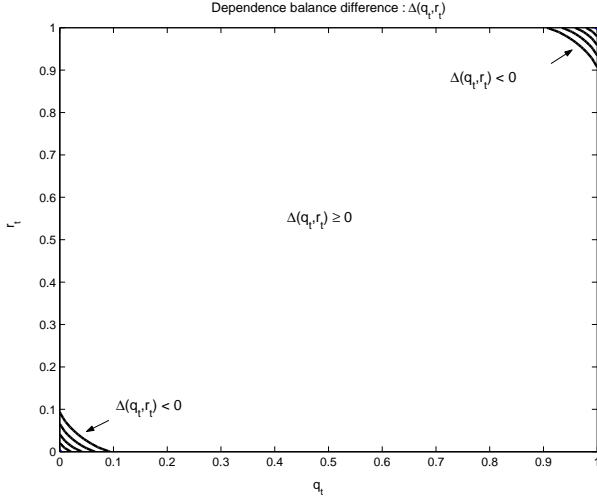


Figure 3: Dependence balance difference for arbitrary binary parallel channels.

leak terms are:

$$L_1 = I(X_1; Z|Y, X_2, T = t) \\ = \phi_{1t} [h^{(3)}(\phi_{1t}p_{1t}, \phi_{1t}p_{2t}, \phi_{1t}(1-p_{1t}-p_{2t})) + 1 - \phi_{1t}] \\ - \phi_{1t}h^{(3)}(p_{1t}, p_{2t}, 1-p_{1t}-p_{2t})]$$

and

$$L_2 = I(X_2; Z|Y, X_1, T = t) \\ = \phi_{1t} [h^{(3)}(\phi_{2t}p_{1t} + (1-\phi_{2t})q_{1t}, \phi_{2t}p_{2t} + \\ (1-\phi_{2t})(1-q_{1t}), \phi_{2t}(1-p_{1t}-p_{2t})) \\ - \phi_{2t}h^{(3)}(p_{1t}, p_{2t}, 1-p_{1t}-p_{2t}) - (1-\phi_{2t})h(q_{1t})]$$

Solving the equation $\nabla L(p_{1t}^*, p_{2t}^*, q_{1t}^*) = 0$ gives

$$q_{1t}^* = \frac{p_{1t}^*}{p_{1t}^* + p_{2t}^*} \quad (19)$$

$$p_{1t}^* + p_{2t}^* = \frac{\phi_{1t}(1-\phi_{2t})}{\phi_{1t}(1-\phi_{2t}) + \phi_{2t}(1-\phi_{1t})} \quad (20)$$

The Hessian evaluated at all critical points $(p_{1t}^*, p_{2t}^*, q_{1t}^*)$ given by (19) and (20) is $H^*(q_{1t}^*, p_{1t}^*, p_{2t}^*) = 0$. Hence the second derivative test is inconclusive. For any fixed allowable input distribution, $p(t, x_1, x_2) = p(t)p(x_1|t)p(x_2|t)$, define $\phi_{1t} = p(x_1 = 0|T = t)$ and $\phi_{2t} = p(x_2 = 0|T = t)$, from (19) and (20), the set of critical points $(p_{1t}^*, p_{2t}^*, q_{1t}^*)$ can be described by just one variable p_{1t}^* , i.e., if we vary p_{1t}^* between $[0, p_t^*]$, then we span all the solutions of (19) and (20). For $p_{1t}^* \in [0, p_t^*]$, substituting $p_{2t}^* = p_t^* - p_{1t}^*$ and $q_{1t}^* = \frac{p_{1t}^*}{p_t^*}$ in the leak terms we obtain

$$L_1(p_{1t}^*) = \phi_{2t}h^{(3)}(\phi_{1t}p_{1t}^*, \phi_{1t}(p_t^* - p_{1t}^*), 1 - \phi_{1t}p_t^*) \\ - \phi_{1t}\phi_{2t}h^{(3)}(p_{1t}^*, p_t^* - p_{1t}^*, 1 - p_t^*) \quad (21)$$

$$L_2(p_{1t}^*) = \phi_{1t}h^{(3)}\left(p_{1t}^*\left(\phi_{2t} + \frac{(1-\phi_{2t})}{p_t^*}\right), (p_t^* - p_{1t}^*)(\phi_{2t} + \frac{(1-\phi_{2t})}{p_t^*}), \phi_{2t}(1-p_t^*)\right) - \phi_{1t}\phi_{2t}.$$

TABLE III
AN ARBITRARY PARALLEL CHANNEL WITH $|Z| = 3$

(X_1, X_2, Y)	1	2	3	Z
(0,0,0)	p_{1t}	p_{2t}	$1 - p_{1t} - p_{2t}$	
(0,1,0)	q_{1t}	q_{2t}	$1 - q_{1t} - q_{2t}$	
(1,0,0)	r_{1t}	r_{2t}	$1 - r_{1t} - r_{2t}$	$p^+(z x_1, x_2, y, t)$

TABLE IV
AN ARBITRARY PARALLEL CHANNEL WITH $I(X_1; X_2|Y, Z, T) = 0$

(X_1, X_2, Y)	1	2	3	Z
(0,0,0)	p_{1t}	p_{2t}	$1 - p_{1t} - p_{2t}$	
(0,1,0)	q_{1t}	$1 - q_{1t}$	0	
(1,0,0)	0	0	1	$p^+(z x_1, x_2, y, t)$

$$h^{(3)}(p_{1t}^*, p_t^* - p_{1t}^*, 1 - p_t^*) - \phi_{1t}(1 - \phi_{2t})h\left(\frac{p_{1t}^*}{p_t^*}\right) \quad (22)$$

Now note the following $L_1(0) = L_1(p_t^*) = \phi_{2t}h(\phi_{1t}p_t^*) - \phi_{1t}\phi_{2t}h(p_t^*)$ and $L_2(0) = L_2(p_t^*) = \phi_{1t}h(\phi_{2t}(1 - p_t^*)) - \phi_{1t}\phi_{2t}h(p_t^*)$. Moreover, it can be easily checked that $\frac{\partial L_1(p_{1t}^*)}{\partial p_{1t}^*} = \frac{\partial L_2(p_{1t}^*)}{\partial p_{1t}^*} = 0$ hence for all $p_{1t}^* \in [0, p_t^*]$, the individual leaks and hence the sum-leak are the same as the ones obtained for Table I. For any two parallel channels with $I(X_1; X_2|Y, Z, T) = 0$, the difference between their sum-rate outer bounds is equal to the difference between their respective sum-leaks, hence we have shown that for $|Z| = 3$ as well, the best possible outer bound is 0.64628 bits/transmission.

IX. CONCLUSION

In this work, we showed that for all binary adaptive parallel channel extensions of the DB bound, the best possible sum-rate outer bound for the BMC is 0.64628 bits/transmission. We also showed that, at one extreme end of the distribution-rate tradeoff, where the set of feasible probability distributions is the smallest, going from $|Z| = 2$ to $|Z| = 3$ does not improve the sum-rate outer bound. By showing these, we partially answered the question posed in [4].

ACKNOWLEDGMENT

This work was supported by NSF Grants CCF 04-47613, CCF 05-14846, CNS 07-16311 and CCF 07-29127.

REFERENCES

- [1] A. P. Hekstra and F. M. J. Willems. Dependence balance bounds for single output two-way channels. *IEEE Trans. on Information Theory*, 35(1):44–53, January 1989.
- [2] Z. Zhang, T. Berger, and J. P. M. Schalkwijk. New outer bounds to capacity regions of two-way channels. *IEEE Trans. on Information Theory*, IT-32:383–386, May 1986.
- [3] C. E. Shannon. Two-way communication channels. In *Proc. 4th Berkeley Symp. Math. Stat. Prob.*, pages 611–644. University of California Press, Berkeley, 1961.
- [4] A. Hekstra. Dependence balance outer bounds for the equal output two-way channel and the multiple access channel with feedback. *Graduate report, Eindhoven University of Technology*, May 1985.
- [5] E. W. Gaal and J. P. M. Schalkwijk. Deterministic binary two-way channels. *Proc. 5th Benelux Symp. on Information Theory, Aalten*, pages 54–63, 1984.