Communication with Disturbance Constraints

Bernd Bandemer and Abbas El Gamal
Stanford University, Information Systems Laboratory
350 Serra Mall, Stanford, CA 94305, USA
Email: bandemer@stanford.edu, abbas@ee.stanford.edu

***Abstract*—The problem of communication with disturbance constraints is introduced. The rate–disturbance region is established for the single constraint case. The optimal encoding scheme turns out to be the same as the Han–Kobayashi scheme for the two user-pair interference channel. For communication with two disturbance constraints, a coding scheme and a corresponding inner bound for the deterministic case are presented. The results suggest a natural way to obtain a new inner bound on the capacity region of the interference channel with more than two user pairs.**

I. Introduction

The interference channel can be viewed as a set of coupled multiple access and broadcast channels. However, in each multiple access channel, the receiver is interested in decoding only one of the transmitted messages while treating the transmission from the other senders as interference. On the other hand, in each broadcast channel, the sender wishes to send a message only to one of the receivers while causing the least *disturbance* to the other receivers. This broadcast aspect of the interference channel motivates us to formulate the problem of communication with disturbance constraints in Figure 1, where sender X wishes to reliably communicate a message M at rate R to the intended receiver Y under disturbance constraint $R_{\mathrm{d},j}$ at side receiver Z_j for $j \in [K]$. We measure the disturbance at Z_j by the amount of *undesired* information rate $(1/n)I(X^n; Z_j^n)$ from X. The problem is to determine the optimal trade-off between the data rate R and the disturbance rates $R_{\mathrm{d},j}$.

Several settings involving communication with constraints, such as communication with input cost in [1, 2] and the wiretap channel in [3, 4], have been previously studied. To the best of our knowledge, our setting is new and the resulting coding scheme is quite different from those for the other constrained communication settings.

For a single disturbance constraint, we show that the optimal encoding scheme is rate splitting and superposition coding, which is the same as the Han–Kobayashi scheme for the two user-pair interference channel [5, 6]. This motivates us to study communication with more than one disturbance constraint with the hope of finding good coding schemes for interference channels with more than two user pairs. To this end, we establish an inner bound on the rate–disturbance region for the deterministic channel model with two disturbance constraints.

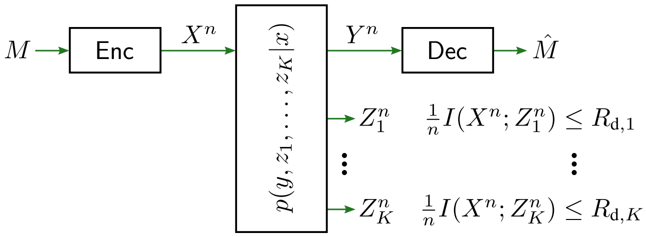

M
Enc
X^n
$p(y, z_1, \dots, z_K|x)$
Y^n
Dec
$\hat{M}$
Z_1^n
$\frac{1}{n}I(X^n; Z_1^n) \le R_{\mathrm{d},1}$
Z_K^n
$\frac{1}{n}I(X^n; Z_K^n) \le R_{\mathrm{d},K}$

Fig. 1. Communication system with disturbance constraints.

Because of space limitation, some of the proofs are deferred to the complete version of this paper posted on Arxiv.

II. Definitions and main results

Consider the communication system in Figure 1. We assume a discrete memoryless channel with K disturbance constraints (DMC-K-DC) that consists of $K+2$ finite sets $\mathcal{X}$, $\mathcal{Y}$, $\mathcal{Z}_j$, $j \in [K]$, and a collection of conditional pmfs $p(y, z_1, \dots, z_K|x)$. A $(2^{nR}, n)$ code for the DMC-K-DC consists of the message set $[2^{nR}]$, an encoding function $x^n : [2^{nR}] \to \mathcal{X}^n$, and a decoding function $\hat{m} : \mathcal{Y}^n \to [2^{nR}]$. We assume that the message M is uniformly distributed over $[2^{nR}]$. A rate–disturbance tuple $(R, R_{\mathrm{d},1}, \dots, R_{\mathrm{d},K}) \in \mathbb{R}_+^{K+1}$ is achievable for the DMC-K-DC if there exists a sequence of $(2^{nR}, n)$ codes such that

$$\lim_{n\to\infty} \mathrm{P}(\hat{M} \neq M) = 0,$$
$$\limsup_{n\to\infty} (1/n)I(X^n; Z_j^n) \le R_{\mathrm{d},j}, \quad j \in [K].$$

The *rate–disturbance region* of the DMC-K-DC is the closure of the set of all achievable tuples $(R, R_{\mathrm{d},1}, \dots, R_{\mathrm{d},K})$.

Remark 1: The measure of disturbance $(1/n)I(X^n; Z_j^n)$ can be expanded as $(1/n)H(Z_j^n) - (1/n)H(Z_j^n \mid X^n)$. The first term is the entropy rate of the received signal Z_j and is caused by both the transmission itself and by noise inherent to the channel. Subtracting the second term separates out the noise part. (For channels with additive white noise, e.g., the Gaussian case, the second term is exactly the differential entropy of each noise sample.)

Remark 2: Our results remain essentially true if disturbance is measured by $(1/n)H(Z_j^n)$ instead. If the channel is deterministic, the two measures coincide.

Remark 3: The disturbance constraint $(1/n)I(X^n; Z_j^n) \le R_{\mathrm{d},j}$ is reminiscent of the information leakage rate constraint for the wiretap channel [3, 4], that is, $(1/n)I(M; Z_j^n) \le R_{\text{leak}}$. Replacing M with X^n, however, dramatically changes the problem and the optimal coding scheme. In the wiretap channel,

This work is partially supported by DARPA ITMANET. Bernd Bandemer is supported by an Eric and Illeana Benhamou Stanford Graduate Fellowship.

the key component of the optimal coding scheme is randomized encoding, which helps control the leakage rate $(1/n)I(M;Z_k^n)$. Such randomization reduces the achievable transmission rate for a given disturbance constraint, hence is not desirable in our setting.

The rate–disturbance region is not known in general. In this paper we establish the following results.

A. Rate–disturbance region for a single disturbance constraint

Consider the case with a single disturbance constraint, i.e., $K = 1$, and relabel Z_1 as Z and $R_{\mathrm{d},1}$ as R_d. We can fully characterize the rate–disturbance region for this case.

Theorem 1: The rate–disturbance region of the DMC-1-DC is the set of rate pairs (R, R_d) such that

$$\begin{aligned} R &\le I(X;Y), \\ R_\mathrm{d} &\ge I(X;Z\,|\,U), \\ R_\mathrm{d} - R &\ge I(X;Z\,|\,U) - I(X;Y\,|\,U), \end{aligned}$$

for some pmf $p(u,x)$ with $|\mathcal{U}| \le |\mathcal{X}| + 1$.

The proof of this theorem is given in Subsection III-A. The region in the theorem is convex. Achievability is established using rate splitting and superposition coding. Receiver Y decodes the satellite codeword while receiver Z distinguishes only the cloud center. Note that this encoding scheme is identical to the Han–Kobayashi scheme for the two user-pair interference channel [5, 6].

We now discuss two interesting special cases.

Deterministic channel. Assume that Y and Z are deterministic functions of X. We show that the rate–disturbance region in Theorem 1 reduces to the following.

Corollary 1: The rate–disturbance region for the deterministic channel with one disturbance constraint is the set of (R, R_d) such that

$$R \le H(Y), \qquad R - R_\mathrm{d} \le H(Y\,|\,Z),$$

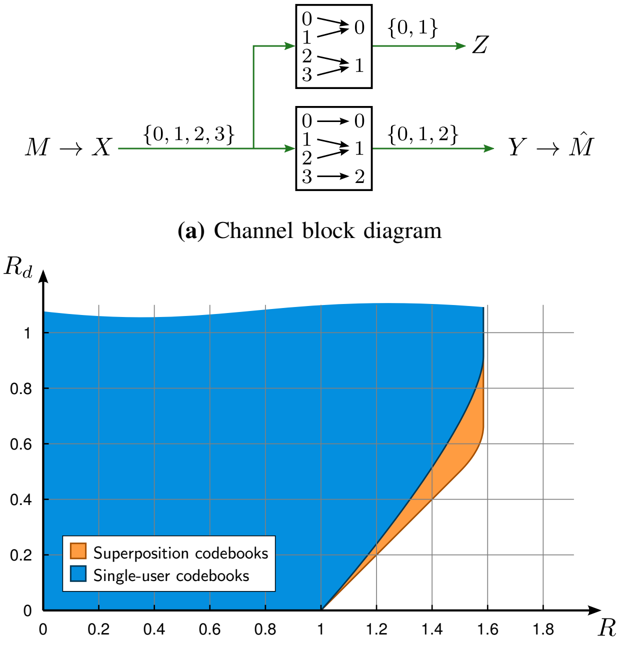
for some pmf $p(x)$.

Clearly, this region is convex. Alternatively, the region can be written as the set of (R, R_d) that satisfy

$$R \le H(Y\,|\,Q), \qquad R_\mathrm{d} \ge I(Z;Y\,|\,Q),$$

for some joint pmf $p(q,x)$ with $|\mathcal{Q}| \le 2$. Corollary 1 and the alternative description of the region are established by substituting $U = Z$ in the region of Theorem 1 and simplifying the resulting region as detailed in Subsection III-B. Observe that this specialization makes the encoding scheme identical to the capacity achieving Han–Kobayashi scheme for the injective deterministic interference channel in [7].

Example 1: Consider the deterministic channel depicted in Figure 2. Note that rates $R \le 1$ can be achieved with zero disturbance rate by restricting the transmission to input symbols $\{0,1\}$ (or $\{2,3\}$), which map into different symbols at Y, but are indistinguishable at Z. On the other hand, for sufficiently large R_d, the disturbance constraint becomes inactive and R is bounded only by the unconstrained capacity $\log_2(3)$. In addition to the optimal region achieved by superposition coding, the figure also shows the strictly suboptimal region achieved by simple non-layered random codes.



(a) Channel block diagram

(b) Rate–disturbance region

Fig. 2. Deterministic example with one disturbance constraint.

Gaussian channel. Consider the problem of communication with disturbance constraints for the AWGN channel

$$Y = X + W_1, \qquad Z = X + W_2,$$

where the noise is $W_1 \sim \mathcal{N}(0,1)$ and $W_2 \sim \mathcal{N}(0,N)$. Assume an average power constraint P on the transmitted signal X.

The case $N \le 1$ is not interesting, since then Y is a degraded version of Z and the disturbance rate is simply given by the data rate R. If $N > 1$, Z is a degraded version of Y, and the rate–disturbance region reduces to the following.

Corollary 2: The rate–disturbance region of the Gaussian channel with parameters $P > 0$ and $N > 1$ is the set of pairs (R, R_d) that satisfy

$$R \le \mathrm{C}(\alpha P), \qquad R_\mathrm{d} \ge \mathrm{C}(\alpha P/N),$$

for some $\alpha \in [0,1]$, where $\mathrm{C}(x) = (1/2)\log(1+x)$ for $x \ge 0$. Achievability is proved using Gaussian codes with power αP. The converse follows by defining $\alpha^\star \in [0,1]$ such that $R = \mathrm{C}(\alpha^\star P)$ and applying the vector entropy power inequality to $Z^n = Y^n + \tilde{W}_2^n$, where $\tilde{W}_2 \sim \mathcal{N}(0, N-1)$ is the excess noise. The details are omitted. Note that this is a degenerate form of the Han–Kobayashi scheme because the constraint from the multiple access side of the interference channel is not taken into consideration.

B. Inner bound for deterministic channel with two disturbance constraints

The correspondence between optimal encoding for the channel with one disturbance constraint and the Han–Kobayashi scheme for the interference channel suggests that the optimal coding scheme for K disturbance constraints may provide an efficient (if not optimal) scheme for the interference channel with more than two user pairs. This is particularly the case for extensions of the two user-pair injective deterministic interference channel for which Han–Kobayashi is optimal [7]. As such, we restrict our attention to the deterministic DMC-2-DC, and establish an inner bound on its rate–disturbance region. The coding scheme used, however, can be readily extended to the general non-deterministic case.

Theorem 2: The rate–disturbance region of the deterministic DMC-2-DC is inner bounded by the set of $(R, R_{\mathrm{d},1}, R_{\mathrm{d},2})$ such that

$$\begin{aligned}
R &\le H(Y),\\
R_{\mathrm{d},1} + R_{\mathrm{d},2} &\ge I(Z_1; Z_2 \,|\, U),\\
R - R_{\mathrm{d},1} &\le H(Y \,|\, Z_1, U),\\
R - R_{\mathrm{d},2} &\le H(Y \,|\, Z_2, U),\\
R - R_{\mathrm{d},1} - R_{\mathrm{d},2} &\le H(Y \,|\, Z_1, Z_2, U) - I(Z_1; Z_2 \,|\, U),\\
2R - R_{\mathrm{d},1} - R_{\mathrm{d},2} &\le H(Y \,|\, Z_1, Z_2, U) + H(Y \,|\, U)\\
&\qquad - I(Z_1; Z_2 \,|\, U),
\end{aligned}$$

for some pmf $p(u, x)$.

The encoding scheme involves rate splitting, superposition coding, and Marton coding. The analysis of the probability of error is complicated by the fact that receiver Y wishes to decode all parts of the message as detailed in Section IV. Receivers Z_1 and Z_2 each observe a satellite codeword from a superposition codebook.

Remark 4: When $U = \emptyset$, receivers Z_1 and Z_2 each observe a codeword from a single-user (non-layered) codebook. In this case, the achievability scheme for Theorem 2 can be combined with the interference decoding scheme in [8] to yield a new inner bound on the capacity region of the three user-pair deterministic interference channel.

III. Proofs for Single Disturbance Constraint

A. Proof of Theorem 1

The proof of the converse for Theorem 1 uses standard techniques. We identify the auxiliary random variable $U = (Q, Y_{Q+1}^n, Z^{Q-1})$, where $Q \sim \mathrm{Unif}[n]$ is a time-sharing random variable, and proceed using Fano's inequality and the Csiszár sum identity. The bound on $|\mathcal{U}|$ is established using the convex cover method in [2].

Achievability is proved as follows.

Codebook generation. Fix a pmf $p(u, x)$.

1) Split the message M into two independent messages M_0 and M_1 with rates R_0 and R_1, respectively. Hence $R = R_0 + R_1$.
2) For each $m_0 \in [2^{nR_0}]$, independently generate a sequence $u^n(m_0)$ according to $\prod_{i=1}^n p(u_i)$.
3) For each $(m_0, m_1) \in [2^{nR_0}] \times [2^{nR_1}]$, independently generate a sequence $x^n(m_0, m_1)$ according to $\prod_{i=1}^n p(x_i \,|\, u_i(m_0))$.

Encoding. For message $m = (m_0, m_1)$, transmit $x^n(m_0, m_1)$.

Decoding. Upon receiving y^n, find the unique $(\hat{m}_0, \hat{m}_1)$ such that $(u^n(\hat{m}_0), x^n(\hat{m}_0, \hat{m}_1), y^n) \in \mathcal{T}_\varepsilon^{(n)}(U, X, Y)$.

Analysis of the probability of error. We are using a superposition code over the channel from X to Y. Using the law of large numbers and the packing lemma in [2], it can be shown that the probability of error tends to zero as $n \to \infty$ if

$$R_1 < I(X; Y \,|\, U), \tag{1}$$

$$R_0 + R_1 < I(X; Y). \tag{2}$$

Analysis of disturbance rate. We analyze the disturbance rate averaged over codebooks $\mathcal{C}$.

$$\begin{aligned}
I(X^n; Z^n \,|\, \mathcal{C}) &\le H(Z^n, M_0 \,|\, \mathcal{C}) - H(Z^n \,|\, X^n, \mathcal{C})\\
&= H(M_0) + H(Z^n \,|\, M_0, \mathcal{C}) - H(Z^n \,|\, X^n)\\
&\overset{\text{(a)}}{\le} nR_0 + H(Z^n \,|\, U^n) - nH(Z \,|\, X)\\
&\le nR_0 + nH(Z \,|\, U) - nH(Z \,|\, X, U)\\
&= nR_0 + nI(X; Z \,|\, U)\\
&\le nR_{\mathrm{d}},
\end{aligned}\tag{3}$$

where (a) follows since U^n is a function of the codebook $\mathcal{C}$ and M_0. Substituting $R = R_0 + R_1$ and using Fourier–Motzkin elimination on inequalities (1), (2), and (3) completes the proof.

B. Proof of Corollary 1

Using the deterministic nature of the channel, the region in Theorem 1 reduces to the set of all (R, R_{d}) such that

$$R \le H(Y),$$

$$R_{\mathrm{d}} \ge H(Z \,|\, U), \tag{4}$$

$$R_{\mathrm{d}} \ge R + H(Z \,|\, U) - H(Y \,|\, U), \tag{5}$$

for some pmf $p(u, x)$. Now fixing a rate R and a pmf $p(x)$ and varying $p(u|x)$ to minimize R_{d}, the right hand sides of (4) and (5) are lower bounded by

$$\begin{aligned}
&H(Z \,|\, U) \ge 0, \quad \text{and}\\
&R + H(Z \,|\, U) - H(Y \,|\, U)\\
&\quad = R + H(Z \,|\, U) - H(Y, Z \,|\, U) + H(Z \,|\, Y, U)\\
&\quad = R - H(Y \,|\, Z, U) + H(Z \,|\, Y, U)\\
&\quad \ge R - H(Y \,|\, Z).
\end{aligned}$$

Note that the particular choice $U = Z$ simultaneously achieves both lower bounds with equality and is therefore sufficient. The rate–disturbance region thus reduces to Corollary 1.

For a fixed pmf $p(x)$, this region has exactly two corner points: $P_1 = (H(Y|Z), 0)$ and $P_2 = (H(Y), I(Y; Z))$. As we vary $p(x)$, there is one corner point P_1 that dominates all other P_1 points. The pmf $p(x)$ for this dominant P_1 can be constructed by maximizing $H(Y|Z)$ as follows. For each $z \in \mathcal{Z}$, define $\mathcal{Y}_z \subseteq \mathcal{Y}$ to be the set of y symbols that are compatible with z. Let $z^\star$ be a symbol that maximizes $|\mathcal{Y}_z|$. For each element of $\mathcal{Y}_{z^\star}$, pick exactly one x that is compatible with it and $z^\star$. Finally, place equal probability mass on each of these x values, and zero mass on all others. This pmf on X yields the dominant corner point P_1, namely $(\log(|\mathcal{Y}_{z^\star}|), 0)$. Moreover, for this distribution, P_2 coincides with P_1. Therefore, the net contribution (modulo convexification) of each pmf $p(x)$ to the rate–disturbance region amounts to its corner point P_2. This implies the alternative description of the region.

IV. Proof of Theorem 2

Codebook generation. Fix a pmf $p(u, x)$. Split the rate as $R = R_0 + R_1 + R_2 + R_3$. Define the auxiliary rates $\tilde{R}_1 \geq R_1$ and $\tilde{R}_2 \geq R_2$, and the set partitions

$$[2^{n\tilde{R}_1}] = \mathcal{L}_1(1) \cup \cdots \cup \mathcal{L}_1(2^{nR_1}),$$
$$[2^{n\tilde{R}_2}] = \mathcal{L}_2(1) \cup \cdots \cup \mathcal{L}_2(2^{nR_2}),$$

where $\mathcal{L}_1(\cdot)$ and $\mathcal{L}_2(\cdot)$ are indexed sets of size $2^{n(\tilde{R}_1 - R_1)}$ and $2^{n(\tilde{R}_2 - R_2)}$, respectively.

1) For each $m_0 \in [2^{nR_0}]$, generate $u^n(m_0)$ according to $\prod_{i=1}^n p(u_i)$.
2) For each $l_1 \in [2^{n\tilde{R}_1}]$, generate $z_1^n(m_0, l_1)$ according to $\prod_{i=1}^n p(z_{1i} \,|\, u_i(m_0))$. Likewise, for each $l_2 \in [2^{n\tilde{R}_2}]$, generate $z_2^n(m_0, l_2)$ according to $\prod_{i=1}^n p(z_{2i} \,|\, u_i(m_0))$.
3) For each (m_0, m_1, m_2), let $\mathcal{S}(m_0, m_1, m_2)$ be the set of all pairs (l_1, l_2) from the product set $\mathcal{L}_1(m_1) \times \mathcal{L}_2(m_2)$ such that $(z_1^n(m_0, l_1), z_2^n(m_0, l_2)) \in \mathcal{T}_{\varepsilon'}^{(n)}(Z_1, Z_2 \,|\, u^n(m_0))$.
4) For each (m_0, l_1, l_2) and $m_3 \in [2^{nR_3}]$, generate $x^n(m_0, l_1, l_2, m_3)$ according to
$$\prod_{i=1}^n p(x_i \,|\, u_i(m_0), z_{1i}(l_1), z_{2i}(l_2))$$
if $(l_1, l_2) \in \mathcal{S}(m_0, m_1, m_2)$. Otherwise, we draw from $\mathrm{Unif}(\mathcal{X}^n)$.
5) Choose $(l_1^{(m_0,m_1,m_2)}, l_2^{(m_0,m_1,m_2)})$ uniformly from $\mathcal{S}(m_0, m_1, m_2)$. If $\mathcal{S}(m_0, m_1, m_2)$ is empty, choose $(1, 1)$.

Encoding. To send message $m = (m_0, m_1, m_2, m_3)$, transmit $x^n(m_0, l_1^{(m_0,m_1,m_2)}, l_2^{(m_0,m_1,m_2)}, m_3)$.

Decoding. Let $\varepsilon > \varepsilon'$. Upon receiving y^n, define the tuple

$$\begin{aligned} &T(m_0, m_1, m_2, m_3) \\ &= \Big(u^n(m_0), z_1^n(m_0, l_1^{(m_0,m_1,m_2)}), z_2^n(m_0, l_2^{(m_0,m_1,m_2)}), \\ &\qquad x^n(m_0, l_1^{(m_0,m_1,m_2)}, l_2^{(m_0,m_1,m_2)}, m_3), y^n\Big) \end{aligned}$$

Declare that $\hat{m} = (\hat{m}_0, \hat{m}_1, \hat{m}_2, \hat{m}_3)$ has been sent if it is the unique message such that $T(\hat{m}_0, \hat{m}_1, \hat{m}_2, \hat{m}_3) \in \mathcal{T}_\varepsilon^{(n)}(U, Z_1, Z_2, X, Y)$.

Message subset	m_0	m_1	m_2	m_3
$\mathcal{M}_0$	1	1	1	1
$\mathcal{M}_1$	1	1	1	$\neq 1$
$\mathcal{M}_2$	1	$\neq 1$	1	any
$\mathcal{M}_3$	1	1	$\neq 1$	any
$\mathcal{M}_4$	1	$\neq 1$	$\neq 1$	any
$\mathcal{M}_5$	$\neq 1$	any	any	any

Table 1. Message subsets for decoding error events.

Analysis of the probability of error. Without loss of generality, assume that $m_0 = m_1 = m_2 = m_3 = 1$ is transmitted. Define the following events.

$$\begin{aligned} \mathcal{E}_{\mathrm{e}1}:&\quad \mathcal{S}(1,1,1) \text{ is empty}, \\ \mathcal{E}_{\mathrm{e}2}:&\quad \mathcal{S}(1,1,1) \text{ contains two distinct pairs with} \\ &\qquad \text{equal first or second component}, \\ \mathcal{E}_i:&\quad \{T(m_0, m_1, m_2, m_3) \in \mathcal{T}_\varepsilon^{(n)}(U, Z_1, Z_2, X, Y) \text{ for} \\ &\quad \text{some } (m_0, m_1, m_2, m_3) \in \mathcal{M}_i\}, \quad i \in \{0, \ldots, 5\}, \end{aligned}$$

where the message subsets $\mathcal{M}_i$ are specified in Table 1. Defining the "encoding error" event $\mathcal{E}_\mathrm{e} = \mathcal{E}_{\mathrm{e}1} \cup \mathcal{E}_{\mathrm{e}2}$ and the "decoding error" event $\mathcal{E}_\mathrm{d} = \mathcal{E}_0^c \cup \mathcal{E}_1 \cup \mathcal{E}_2 \cup \mathcal{E}_3 \cup \mathcal{E}_4$, the probability of error can be upper-bounded as

$$\mathrm{P}(\mathcal{E}) \leq \mathrm{P}(\mathcal{E}_\mathrm{e} \cup \mathcal{E}_\mathrm{d}) \leq \mathrm{P}(\mathcal{E}_\mathrm{e}) + \mathrm{P}(\mathcal{E}_\mathrm{d} \,|\, \mathcal{E}_\mathrm{e}^c).$$

The motivation for introducing $\mathcal{E}_{\mathrm{e}2}$ as an "error" is to simplify the analysis of the second probability term. We bound $\mathrm{P}(\mathcal{E}_\mathrm{e})$ by the following lemma, for which a proof sketch is given in the appendix. Let $r_1 = \tilde{R}_1 - R_1$ and $r_2 = \tilde{R}_2 - R_2$.

Lemma 1: $\mathrm{P}(\mathcal{E}_\mathrm{e}) \to 0$ as $n \to \infty$ if

$$r_1 + r_2 > I(Z_1; Z_1 \,|\, U), \tag{6}$$
$$r_1/2 + r_2 < I(Z_1; Z_2 \,|\, U), \tag{7}$$
$$r_1 + r_2/2 < I(Z_1; Z_2 \,|\, U). \tag{8}$$

We bound the probability $\mathrm{P}(\mathcal{E}_\mathrm{d} \,|\, \mathcal{E}_\mathrm{e}^c)$ by the following lemma, for which we sketch the proof in the appendix.

Lemma 2: $\mathrm{P}(\mathcal{E}_\mathrm{d} \,|\, \mathcal{E}_\mathrm{e}^c) \to 0$ as $n \to \infty$ if

$$R_3 < H(Y \,|\, Z_1, Z_2, U), \tag{9}$$
$$\tilde{R}_1 + R_3 < H(Y \,|\, Z_2, U) + I(Z_1; Z_2 \,|\, U), \tag{10}$$
$$\tilde{R}_2 + R_3 < H(Y \,|\, Z_1, U) + I(Z_1; Z_2 \,|\, U), \tag{11}$$
$$\tilde{R}_1 + \tilde{R}_2 + R_3 < H(Y \,|\, U) + I(Z_1; Z_2 \,|\, U), \tag{12}$$
$$R_0 + \tilde{R}_1 + \tilde{R}_2 + R_3 < H(Y) + I(Z_1; Z_2 \,|\, U). \tag{13}$$

Analysis of disturbance rate. When viewed by receiver Z_1, the codeword for message $m = (m_0, m_1, m_2, m_3)$, namely $x^n(m_0, l_1^{(m_0,m_1,m_2)}, l_2^{(m_0,m_1,m_2)}, m_3)$, appears as $z_1^n(m_0, l_1^{(m_0,m_1,m_2)})$. We can pessimistically assume that all sequences $z_1^n(m_0, l_1)$ as created in step 2 of codebook generation can be seen at the receiver for some message m. Therefore, the number of possible sequences at Z_1, and thus its disturbance rate, is upper-bounded by $H(Z_1^n) \leq n(R_0 + \tilde{R}_1)$. Applying the same argument for Z_2, the proposed scheme achieves

$$R_0 + \tilde{R}_1 \leq R_{\mathrm{d},1}, \qquad R_0 + \tilde{R}_2 \leq R_{\mathrm{d},2}. \tag{14}$$

Conclusion of the proof. Collecting inequalities (6) through (14), recalling $R = R_0 + R_1 + R_2 + R_3$, and using the Fourier-Motzkin procedure to eliminate R_0, R_1, R_2, and R_3 leads to the $(R, R_{\mathrm{d},1}, R_{\mathrm{d},2})$ region claimed in the theorem.

V. Acknowledgments

The authors would like to thank Yeow-Khiang Chia for helpful discussions.

Appendix

A. Proof sketch for Lemma 1

First, consider $\mathcal{E}_{\text{e1}}$. As in the proof of Marton's inner bound for the broadcast channel, the mutual covering lemma [2] implies $\mathrm{P}(\mathcal{E}_{\text{e1}}) \to 0$ as $n \to \infty$ if (6) holds.

Now consider $\mathcal{E}_{\text{e2}}$, for which we need to control the number of typical pairs that can occur in the same "row" or "column" of the product set $\mathcal{L}_1(m_1) \times \mathcal{L}_2(m_2)$, i.e., for the same l_1 or l_2 coordinate. The probability $\mathrm{P}(\mathcal{E}_{\text{e2}})$ tends to zero provided that (7) and (8) hold.

This is akin to the birthday problem [9], where k samples are drawn uniformly and independently from $[N]$, and the interest is in samples that have the same value (collisions). It is well-known that for the probability of collision to be p_c, the number of samples required is roughly $k \approx \sqrt{-2N \ln(1-p_c)}$, which scales with $\sqrt{N}$. In our case, the number of samples is the cardinality of the set $\mathcal{S}(m_0, m_1, m_2)$, which is roughly $k = 2^{n(r_1+r_2-I(Z_1;Z_2\,|\,U))}$. The samples are categorized into $N_1 = 2^{nr_1}$ and $N_2 = 2^{nr_2}$ classes along rows and columns, respectively. To achieve a probability of collision $p_c \to 0$ along both dimensions, we need $k \ll \min\{\sqrt{N_1}, \sqrt{N_2}\}$, which yields exactly the conditions (7) and (8).

B. Proof sketch for Lemma 2

The events of which $\mathcal{E}_{\text{d}}$ is composed are illustrated in Figure 3, which also depicts the structure of the codebook for $m_0 = 1$. The product sets $\mathcal{L}_1(m_1) \times \mathcal{L}_2(m_2)$, for each (m_1, m_2), are represented by shaded squares. In each product set, the sequence pair selected in step 5 of the codebook generation procedure is shown with its superposed x^n codewords, as created in step 4. The correct codeword $x^n(1,1,1,1)$ is shown as a white circle which is connected to the received sequence y^n. The codewords that may be mistakenly detected at the receiver are shown as black circles. The product sets associated with decoding error events $\mathcal{E}_1$, $\mathcal{E}_2$, $\mathcal{E}_3$, and $\mathcal{E}_4$ are labeled 1, 2, 3, and 4, respectively.

We bound the probability of each sub-event of $\mathcal{E}_{\text{d}}$. First, note that by the conditional typicality lemma in [2], $\mathrm{P}(\mathcal{E}_0^c) \to 0$ as $n \to \infty$. The probabilities of the events $\mathcal{E}_1$ through $\mathcal{E}_5$ tend to zero as $n \to \infty$ under conditions (9) through (13), correspondingly.

The events $\mathcal{E}_2$ and $\mathcal{E}_3$ require the most careful analysis, since the true codeword $x^n(1,1,1,1)$ and the codewords with which it may be confused can share the same z_1^n or z_2^n sequence (see dashed line and circles on it in Figure 3). Moreover, even when the chosen pairs in two different product sets do not share one of the two coordinates (see the chosen pairs for $(m_1, m_2) = (1,1)$ and $(2,1)$ in Figure 3), correlation could potentially be caused by the selection procedure in step 5 of codebook generation. We use the following lemma to show that the event $\mathcal{E}_{\text{e}}^c$ prevents this correlation leakage from occurring.

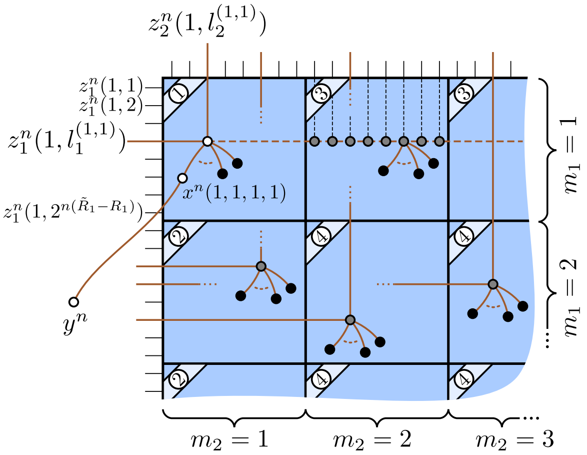


Fig. 3. Illustration of decoding error events, for $m_0 = 1$.

Independence lemma: Consider a finite set $\mathcal{A}$ and a subset $\mathcal{A}' \subset \mathcal{A}$. Let p_A be an arbitrary pmf over $\mathcal{A}$. Let the random vector A^n be distributed proportionally to the product distribution $\prod_{l=1}^n p_A(a_l)$, restricted to the support set $\{(a_1, \ldots, a_n) : a_k \in \mathcal{A}' \text{ for some } k\}$. Let I be drawn uniformly from $\{i : A_i \in \mathcal{A}'\}$. Let $J = ((I + s - 1) \mod n) + 1$ for some integer $s \in [n-1]$. Then, the random variables A_I and A_J are independent.

The application of this lemma is what distinguishes this analysis from the conventional Marton inner bound for broadcast channels [10, 11]. There, analysis of the selection process can be altogether avoided since each receiver decodes only one of the two coordinates.

References

[1] R. McEliece, *The Theory of Information and Coding*. Reading, MA: Addison-Wesley, 1977.

[2] A. El Gamal and Y.-H. Kim, "Lectures on Network Information Theory," 2010, arXiv:1001.3404.

[3] A. D. Wyner, "The wire-tap channel," *Bell System Technical Journal*, vol. 54, no. 8, pp. 1355–1387, Oct. 1975.

[4] I. Csiszár and J. Körner, "Broadcast channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. 24, no. 3, pp. 339–348, May 1978.

[5] T. S. Han and K. Kobayashi, "A new achievable rate region for the interference channel," *IEEE Trans. Inf. Theory*, vol. 27, no. 1, pp. 49–60, Jan. 1981.

[6] H.-F. Chong, M. Motani, H. K. Garg, and H. El Gamal, "On the Han-Kobayashi Region for the Interference Channel," *IEEE Trans. Inf. Theory*, vol. 54, no. 7, pp. 3188–3195, Jul. 2008.

[7] A. A. El Gamal and M. H. M. Costa, "The Capacity Region of a Class of Deterministic Interference Channels," *IEEE Trans. Inf. Theory*, vol. 28, no. 2, pp. 343–346, Mar. 1982.

[8] B. Bandemer and A. El Gamal, "Interference decoding for deterministic channels," *IEEE Trans. Inf. Theory*, vol. 57, no. 5, pp. 2966–2975, May 2011, arXiv:1001.4588.

[9] R. von Mises, "Über Aufteilungs- und Besetzungs-Wahrscheinlichkeiten," *Revue de la Faculté des Sciences de l'Université d'Istanbul*, vol. 4, pp. 145–163, 1939, reprinted in "Selected Papers of Richard von Mises", vol. 2 (Ed. P. Frank, S. Goldstein, M. Kac, W. Prager, G. Szegö, and G. Birkhoff). Providence, RI: American Mathematical Society, pp. 313-334, 1964.

[10] K. Marton, "A coding theorem for the discrete memoryless broadcast channel," *IEEE Trans. Inf. Theory*, vol. 25, no. 3, pp. 306–311, May 1979.

[11] A. El Gamal and E. C. van der Meulen, "A proof of Marton's coding theorem for the discrete memoryless broadcast channel," *IEEE Trans. Inf. Theory*, vol. 27, no. 1, pp. 120–122, Jan. 1981.