Fiftieth Annual Allerton Conference
Allerton House, UIUC, Illinois, USA
October 1 - 5, 2012

The Zero-Undetected-Error Capacity of Discrete Memoryless Channels with Feedback

Christoph Bunte and Amos Lapidoth
Signal and Information Processing Laboratory
ETH Zurich
E-mail: {bunte,lapidoth}@isi.ee.ethz.ch

Abstract— **We show that the zero-undetected-error capacity (a.k.a. erasures-only capacity and zero-error erasure capacity) of a discrete memoryless channel with feedback is equal to its ordinary capacity whenever its zero-undetected-error capacity without feedback is positive, i.e., whenever it has an output that is reachable from some but not all inputs, and that otherwise its zero-undetected-error capacity is zero. We then show that feedback can increase the zero-undetected-error capacity. Finally, the result is extended to multiple-access and broadcast channels.**

I. Introduction

In typical communication schemes a decoding error is defined as the event that the receiver does not produce the message selected by the transmitter. This may happen in one of two ways: the receiver may produce a message different from the one selected by the transmitter, or it may refuse to decode altogether. In the former case we say that an *undetected error* has occurred, while in the latter case we say that an *erasure* has occurred. In classical channel coding one does not usually distinguish between the two types of errors. This is not always appropriate, in particular in situations where undetected errors may have grave consequences.

In a 1968 paper [1] Forney noticed that if a discrete memoryless channel (DMC) has an output that is reachable from some but not all of its inputs, then positive rates are achievable with arbitrarily small probability of erasure and *zero* probability of undetected errors. The largest such rate is called the *zero-undetected-error capacity* [2], or the *erasures-only capacity* [3], or the *zero-error erasure capacity* [4].

Determining the zero-undetected-error capacity for arbitrary DMCs is still an open problem. However, a number of contributions have been made over the years [1]–[7], the main results of which are summarized below. But first some notation and definitions.

We use W to denote the transition law (or channel matrix) of a generic DMC with finite input alphabet $\mathcal{X}$ and finite output alphabet $\mathcal{Y}$. The corresponding n-fold product channel W^n is defined by

$$W^n(\mathbf{y}|\mathbf{x}) = \prod_{j=1}^{n} W(y_j|x_j), \quad \mathbf{x} \in \mathcal{X}^n, \mathbf{y} \in \mathcal{Y}^n, \tag{1}$$

where x_j and y_j denote the j-th components of $\mathbf{x}$ and $\mathbf{y}$.

A zero-undetected-error code of blocklength n comprises a message set $\mathcal{M}$ and an encoding function

$$f\colon \mathcal{M} \to \mathcal{X}^n.$$

Upon receiving $\mathbf{y} \in \mathcal{Y}^n$ the decoder declares an erasure if the list of messages that cannot be ruled out

$$\mathcal{L}(\mathbf{y}) = \Big\{ m \in \mathcal{M} : W^n\big(\mathbf{y}|f(m)\big) > 0 \Big\}$$

contains more than one message; otherwise it produces the only message on the list. The maximal erasure probability is

$$\max_{m\in\mathcal{M}} \Pr\big[|\mathcal{L}(\mathbf{Y})| > 1 \,\big|\, M = m\big],$$

and the average erasure probability is

$$\frac{1}{|\mathcal{M}|} \sum_{m\in\mathcal{M}} \Pr\big[|\mathcal{L}(\mathbf{Y})| > 1 \,\big|\, M = m\big].$$

The rate of the code is $n^{-1}\log|\mathcal{M}|$. The zero-undetected-error capacity $C_{0\mathrm{u}}$ is defined as the largest number R such that there exists a sequence of zero-undetected-error codes with erasure probability tending to zero and rate approaching R as the blocklength tends to infinity.

For channels with feedback we replace the encoding function f by a finite sequence of functions

$$f_j\colon \mathcal{M} \times \mathcal{Y}^{j-1} \to \mathcal{X}, \quad j = 1, \ldots, n$$

and the list of messages that cannot be ruled out by

$$\mathcal{L}_{\mathrm{f}}(\mathbf{y}) = \left\{ m \in \mathcal{M} : \prod_{j=1}^{n} W\big(y_j|f_j(m, y_1^{j-1})\big) > 0 \right\}, \tag{2}$$

where y_1^{j-1} is shorthand for $y_1, \ldots, y_{j-1}$. The zero-undetected-error capacity for channels with feedback is denoted $C_{0\mathrm{uf}}$.

The definition of $C_{0\mathrm{u}}$ and $C_{0\mathrm{uf}}$ does not depend on whether we use an average or maximal erasure probability criterion. Indeed, by discarding the worst half of the messages of each code in a sequence of codes approaching rate R with average erasure probability tending to zero, we obtain a sequence of codes approaching rate R with maximal erasure probability tending to zero.

Using a random coding argument with letter-by-letter IID codebooks Forney [1] showed that

$$C_{0\mathrm{u}} \geq \max_Q \sum_{y\in\mathcal{Y}} (QW)(y) \log \frac{1}{\sum_{x\in\mathcal{X}: W(y|x)>0} Q(x)}, \tag{3}$$

where the maximization is over all distributions Q on $\mathcal{X}$ and where QW denotes the output distribution induced by Q, i.e.,

$$(QW)(y) = \sum_{x \in \mathcal{X}} W(y|x)Q(x), \quad y \in \mathcal{Y}.$$

Using (3) it is not difficult to show that $C_{0\mathrm{u}} > 0$ if, and only if, there exists a triple $(x, x', y) \in \mathcal{X} \times \mathcal{X} \times \mathcal{Y}$ such that $W(y|x) = 0$ and $W(y|x') > 0$, i.e., there exists an output that is reachable from some but not all inputs.[1] Moreover, Forney's bound and Shannon's expression for the zero-error capacity of DMCs with feedback [8] establish the relation

$$C_0 \leq C_{0\mathrm{f}} \leq C_{0\mathrm{u}} \leq C_{0\mathrm{uf}} \leq C, \tag{4}$$

where $C_{0\mathrm{f}}$ and C_0 denote the zero-error capacity with and without feedback, respectively, and C denotes the (ordinary) capacity (see, e.g., [9] for these concepts). Indeed, Shannon proved that

$$C_{0\mathrm{f}} = \begin{cases} \rho & \text{if } C_0 > 0, \\ 0 & \text{otherwise,} \end{cases} \tag{5}$$

where

$$\rho = \max_Q \min_{y \in \mathcal{Y}} \log \frac{1}{\sum_{x \in \mathcal{X} : W(y|x) > 0} Q(x)}, \tag{6}$$

and the maximization is over all distributions Q on $\mathcal{X}$. Comparing (3) and (6) we see that $C_{0\mathrm{f}} \leq C_{0\mathrm{u}}$. The remaining inequalities in (4) are clear.

Forney's bound is in general not tight. A better lower bound can be obtained by random coding over codes of constant composition [2], [4], [6], [10]:

$$C_{0\mathrm{u}} \geq \max_Q \min_{\substack{V \ll W \\ QV = QW}} I(Q, V), \tag{7}$$

where the maximization is over all distributions Q on $\mathcal{X}$ and where the minimization is over all auxiliary channels V such that $V(y|x) = 0$ whenever $W(y|x) = 0$ and such that V induces the same output distribution under Q as the true channel W. Although (7) is better than (3), it is not always tight [4], [6]. One can obtain tighter non-single-letter bounds by applying (3) or (7) to the n-fold product channel (1) and normalizing the result by n. In fact, it can be shown that both bounds become tight in the limit as n tends to infinity [3], [4]. Although this characterizes $C_{0\mathrm{u}}$ in terms of information theoretic quantities, it does not solve the problem of computing $C_{0\mathrm{u}}$ numerically. It is also worth mentioning that, in general, neither the maximization in (3) nor the maximization in (7) is a concave problem, which diminishes the practical usefulness of these bounds.

Computing $C_{0\mathrm{u}}$ becomes trivial in cases where it is known that $C_{0\mathrm{u}}$ equals C. Pinsker and Sheverdyaev [5] proved that equality holds if the bipartite channel graph is acyclic.[2] This is not a necessary condition. In fact, the class of DMCs for which equality is known to hold was extended by Telatar and Gallager [6] and further by Csiszár and Narayan [3]: It now includes all channels W such that, for some positive functions $A(\cdot)$ and $B(\cdot)$, and some capacity-achieving input distribution $Q^\star$, the relation $W(y|x) = A(x)B(y)$ holds whenever $Q^\star(x)W(y|x) > 0$. It was conjectured in [3] that this condition is also necessary for equality.

While no single-letter expression for $C_{0\mathrm{u}}$ is known, the contribution of this paper is in exhibiting one for $C_{0\mathrm{uf}}$. More specifically, we use a simple two-phase coding scheme, inspired by Burnashev [11], to show that $C_{0\mathrm{uf}}$ coincides with C whenever $C_{0\mathrm{u}}$ is positive, and that otherwise $C_{0\mathrm{uf}}$ is zero; a proof is provided in Section II. In Section III, we use this result to show that feedback can increase the zero-undetected-error capacity. In Sections IV and V we give an extension to multiple-access and broadcast channels. We conclude the paper in Section VI with a brief discussion of the results.

II. Main Result

Theorem 1: For every DMC

$$C_{0\mathrm{uf}} = \begin{cases} C & \text{if } C_{0\mathrm{u}} > 0, \\ 0 & \text{otherwise.} \end{cases} \tag{8}$$

Remark 1: As pointed out in Section I, $C_{0\mathrm{u}}$ is positive if, and only if, there is a triple $(x_\mathrm{e}, x_\mathrm{c}, y_\mathrm{c}) \in \mathcal{X} \times \mathcal{X} \times \mathcal{Y}$ such that $W(y_\mathrm{c}|x_\mathrm{e}) = 0$ and $W(y_\mathrm{c}|x_\mathrm{c}) > 0$.

Proof: Without loss of generality we may assume that every output is reachable from some input, i.e., for every $y \in \mathcal{Y}$ there is some $x \in \mathcal{X}$ such that $W(y|x) > 0$. Consequently, if $C_{0\mathrm{u}}$ is zero, then $W(y|x)$ must be positive for all pairs $(x, y) \in \mathcal{X} \times \mathcal{Y}$ by Remark 1. Thus, no matter which sequence of output letters is received, no message can be ruled out, i.e., $\mathcal{L}_\mathrm{f}(\mathbf{y}) = \mathcal{M}$ for every $\mathbf{y} \in \mathcal{Y}^n$. It follows that the erasure probability is 1 for all zero-undetected-error feedback codes having more than one message, so $C_{0\mathrm{uf}}$ must be zero.

Now assume that $C_{0\mathrm{u}}$ is positive and let the triple $(x_\mathrm{e}, x_\mathrm{c}, y_\mathrm{c})$ be as in Remark 1. Take a capacity-achieving sequence of codes $\{\mathcal{C}_n\}_{n=1}^\infty$ and let $\{\mathcal{M}_n\}_{n=1}^\infty$ be the corresponding sequence of message sets. More precisely, let $\{\mathcal{C}_n\}_{n=1}^\infty$ be a sequence of codes whose maximal probability of error tends to zero and whose rate approaches C as the blocklength n tends to infinity. Choose a sequence $\{\nu_n\}_{n=1}^\infty$ of positive integers such that $\nu_n \to \infty$ and $\nu_n / n \to 0$ as $n \to \infty$ (e.g., $\nu_n = \lceil \log(n+1) \rceil$). We construct for each n a zero-undetected-error feedback code of blocklength $n + \nu_n$ and message set $\mathcal{M}_n$ as follows. Suppose we wish to transmit the message $m \in \mathcal{M}_n$. In the first n channel uses we send the codeword of $\mathcal{C}_n$ corresponding to m. After having received the first n output letters the receiver uses the decoder for the code $\mathcal{C}_n$ to form an estimate $\hat{m}$ of the message. Thanks to the feedback link, the transmitter can do the same. Thus, $\hat{m}$ is available at both the transmitter and the receiver, and the transmitter can compare this estimate to the true message m. If the estimate is correct, i.e., if $\hat{m}$ is equal to m, then the transmitter uses the remaining ν_n channel uses to convey this fact to the

[1] If we assume that every output is reachable from some input, this is equivalent to the channel matrix containing a zero.

[2] This graph is formed by introducing edges between all input-output pairs that have positive transition probability. Thus, an equivalent way of stating that the graph is acyclic is to say that there does not exist $\ell \geq 2$, distinct inputs $x_1, \ldots, x_\ell$ and distinct outputs $y_1, \ldots, y_\ell$ such that $W(y_j|x_j) > 0$, $W(y_j|x_{j+1}) > 0$ for $j = 1, \ldots, \ell$ and $x_{\ell+1} = x_1$.

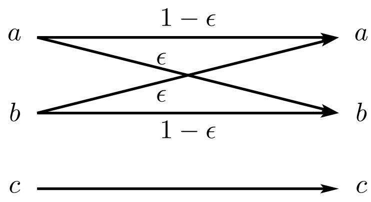

Fig. 1. A DMC with $C_{0\mathrm{u}} < C_{0\mathrm{uf}}$

receiver by sending ν_n times the letter x_{c}. If the estimate is incorrect, i.e., if $\hat{m} \neq m$, then the transmitter sends ν_n times the letter x_{e} to trigger an erasure. Accordingly, if the receiver observes the letter y_{c} in at least one of the last ν_n positions, it knows with certainty that the estimate is correct and produces $\hat{m}$; if the letter y_{c} does not appear in the last ν_n positions, the receiver declares an erasure. Observe that the probability of an undetected error of this coding scheme is zero. Indeed, the transmitter knows when the estimate is wrong, and its mechanism for triggering an erasure by sending the letter x_{e} (ν_n times) is reliable in the sense that the probability of the receiver observing the letter y_{c} in any of the last ν_n positions is zero when $\hat{m} \neq m$. On the other hand, the probability of an erasure conditional on the transmitted message being m is upper bounded by the sum of the probability that the decoder for the code $\mathcal{C}_n$ errs and the probability that none of the last ν_n received output letters equals y_{c} conditional on the last ν_n input letters being equal to x_{c}. The first of these two probabilities tends to zero by the choice of $\{\mathcal{C}_n\}_{n=1}^{\infty}$, while the second probability is

$$\bigl(1 - W(y_{\mathrm{c}}|x_{\mathrm{c}})\bigr)^{\nu_n},$$

which tends to zero as $n \to \infty$ because $W(y_{\mathrm{c}}|x_{\mathrm{c}})$ is positive and $\nu_n \to \infty$. We thus conclude that the maximal erasure probability tends to zero. The proof is completed by noting that the rate of the constructed code is

$$\frac{\log|\mathcal{M}_n|}{n + \nu_n} = \frac{1}{1 + \frac{\nu_n}{n}} \frac{\log|\mathcal{M}_n|}{n} \to C, \quad n \to \infty.$$

■

III. Feedback Can Increase the Zero-Undetected-Error Capacity

Consider the channel in Figure 1 for some $0 < \epsilon < 1/2$. The capacity of this channel is

$$C = \log\Bigl(1 + \exp\bigl(\log 2 - H_b(\epsilon)\bigr)\Bigr),$$

where $H_b(\cdot)$ denotes the binary entropy function

$$H_b(\epsilon) = -\epsilon \log \epsilon - (1 - \epsilon) \log(1 - \epsilon).$$

By Theorem 1, $C_{0\mathrm{uf}}$ is equal to C for this channel and thus

$$C_{0\mathrm{uf}} > \log 2.$$

We claim that $C_{0\mathrm{u}} = \log 2$. Indeed, $C_{0\mathrm{u}} \geq \log 2$ because we can transmit 1 bit per channel use with zero probability of erasure and zero probability of error by using only the inputs a and c. To show that $C_{0\mathrm{u}} \leq \log 2$, note that for codewords of blocklength n there are 2^n distinct patterns of c's so any code with more than 2^n messages will have a pair of distinct messages m and m' such that the corresponding codewords $f(m)$ and $f(m')$ have the same pattern of c's. (That is, the j-th component of $f(m)$ is c if, and only if, the j-th component of $f(m')$ is c.) But this implies that m and m' have the same sequences of output letters that they can induce with positive probability so transmitting m or m' results in an erasure with probability 1. Consequently, no zero-undetected-error code of blocklength n and maximal erasure probability less than 1 can have more than 2^n messages, establishing that $C_{0\mathrm{u}} \leq \log 2$.

IV. Extension to Multiple-Access Channels

In this section we use W to denote the transition law of a generic discrete memoryless multiple-access channel (DM-MAC) with finite input alphabets $\mathcal{X}_1, \mathcal{X}_2$ and finite output alphabet $\mathcal{Y}$. The definition of a zero-undetected-error code for the DM-MAC is a straightforward extension of the definition for DMCs. We use $\mathcal{R}_{\mathrm{f}}^{\mathrm{MAC}}$ and $\mathcal{R}_{0\mathrm{uf}}^{\mathrm{MAC}}$ to denote the capacity region (average error probability criterion) and zero-undetected-error capacity region (average erasure probability criterion) when feedback is available at both transmitters.[3]

Theorem 2: Let $\mathcal{W}_1$ denote the set of all DM-MACs for which there is a quadruple $(x_{1,\mathrm{e}}, x_{1,\mathrm{c}}, x_{2,\mathrm{h}}, y_{1,\mathrm{c}}) \in \mathcal{X}_1 \times \mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{Y}$ such that $W(y_{1,\mathrm{c}}|x_{1,\mathrm{e}}, x_{2,\mathrm{h}}) = 0$ and $W(y_{1,\mathrm{c}}|x_{1,\mathrm{c}}, x_{2,\mathrm{h}}) > 0$. Similarly, let $\mathcal{W}_2$ denote the set of all DM-MACs for which there is a quadruple $(x_{1,\mathrm{h}}, x_{2,\mathrm{e}}, x_{2,\mathrm{c}}, y_{2,\mathrm{c}}) \in \mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{X}_2 \times \mathcal{Y}$ such that $W(y_{2,\mathrm{c}}|x_{1,\mathrm{h}}, x_{2,\mathrm{e}}) = 0$ and $W(y_{2,\mathrm{c}}|x_{1,\mathrm{h}}, x_{2,\mathrm{c}}) > 0$. Then,

$$\mathcal{R}_{0\mathrm{uf}}^{\mathrm{MAC}} = \begin{cases} \mathcal{R}_{\mathrm{f}}^{\mathrm{MAC}} & \text{if } W \in \mathcal{W}_1 \cap \mathcal{W}_2, \\ [0, R_{1,\mathrm{max}}] \times \{0\} & \text{if } W \in \mathcal{W}_1 \cap \mathcal{W}_2^{\mathrm{c}}, \\ \{0\} \times [0, R_{2,\mathrm{max}}] & \text{if } W \in \mathcal{W}_1^{\mathrm{c}} \cap \mathcal{W}_2, \\ \{(0,0)\} & \text{if } W \in \mathcal{W}_1^{\mathrm{c}} \cap \mathcal{W}_2^{\mathrm{c}}, \end{cases}$$

where

$$R_{1,\mathrm{max}} = \max_{Q_{X_1}, x_2 \in \mathcal{X}_2} I(X_1; Y | X_2 = x_2),$$

and

$$R_{2,\mathrm{max}} = \max_{Q_{X_2}, x_1 \in \mathcal{X}_1} I(X_2; Y | X_1 = x_1).$$

Proof: For the converse we show that if $W \in \mathcal{W}_2^{\mathrm{c}}$, then no zero-undetected-error feedback code with average erasure probability less than 1 and more than one message for Transmitter 2 exists. (The case $W \in \mathcal{W}_1^{\mathrm{c}}$ is similar.) Consider the list of message pairs that the receiver cannot

[3] Theorem 2 remains true if both $\mathcal{R}_{\mathrm{f}}^{\mathrm{MAC}}$ and $\mathcal{R}_{0\mathrm{uf}}^{\mathrm{MAC}}$ are defined with a maximal instead of an average error and erasure probability criterion, respectively.

rule out

$$\mathcal{L}_{\mathrm{f}}^{\mathrm{MAC}}(\mathbf{y}) = \Bigg\{ (m_1, m_2) \in \mathcal{M}_1 \times \mathcal{M}_2 : \prod_{j=1}^{n} W\big(y_j | f_{1,j}(m_1, y_1^{j-1}), f_{2,j}(m_2, y_1^{j-1})\big) > 0 \Bigg\}, \tag{9}$$

where

$$f_{1,j}\colon \mathcal{M}_1 \times \mathcal{Y}^{j-1} \to \mathcal{X}_1, \quad j = 1, \ldots, n,$$

and

$$f_{2,j}\colon \mathcal{M}_2 \times \mathcal{Y}^{j-1} \to \mathcal{X}_2, \quad j = 1, \ldots, n,$$

are the encoding functions of Transmitters 1 and 2. Now if $W \in \mathcal{W}_2^{\mathrm{c}}$, then the condition $W(y|x_1, x_2) > 0$ implies that $W(y|x_1, x_2') > 0$ for all $x_2' \in \mathcal{X}_2$. Thus, from (9) we see that if $(m_1, m_2) \in \mathcal{L}_{\mathrm{f}}^{\mathrm{MAC}}(\mathbf{y})$, then $(m_1, m_2') \in \mathcal{L}_{\mathrm{f}}^{\mathrm{MAC}}(\mathbf{y})$ for every $m_2' \in \mathcal{M}_2$, so the average erasure probability is 1 whenever $|\mathcal{M}_2| > 1$.

The direct part is a straightforward extension of the direct part of Theorem 1 and we present only the case $W \in \mathcal{W}_1 \cap \mathcal{W}_2$: Let $\{\eta_n\}_{n=1}^{\infty}$ be a sequence of positive even numbers such that $\eta_n \to \infty$ and $\eta_n/n \to 0$ as $n \to \infty$. We construct a code of blocklength $n + \eta_n$ as follows. In the first n channel uses Transmitter 1 and Transmitter 2 use the n-th code $\mathcal{C}_{\mathrm{f},n}^{\mathrm{MAC}}$ in a sequence of codes achieving the rate pair $(R_1, R_2) \in \mathcal{R}_{\mathrm{f}}^{\mathrm{MAC}}$. The receiver observes the first n channel outputs and, using the decoder for $\mathcal{C}_{\mathrm{f},n}^{\mathrm{MAC}}$, forms the message estimates $\hat{m}_1$ and $\hat{m}_2$. The feedback allows Transmitter 1 to compare the estimate $\hat{m}_1$ to the true message m_1, and likewise for Transmitter 2. In the next $\eta_n/2$ channel uses Transmitter 2 helps Transmitter 1 by sending $\eta_n/2$ times the letter $x_{2,\mathrm{h}}$, while Transmitter 1 sends either $\eta_n/2$ times the letter $x_{1,\mathrm{e}}$ (if $\hat{m}_1 \neq m_1$), or $\eta_n/2$ times the letter $x_{1,\mathrm{c}}$ (if $\hat{m}_1 = m_1$). In the remaining $\eta_n/2$ channel uses the roles of Transmitter 1 and Transmitter 2 are reversed. If the receiver observes the letter $y_{1,\mathrm{c}}$ at least once in the sequence $y_{n+1}, \ldots, y_{n+\eta_n/2}$ and the letter $y_{2,\mathrm{c}}$ at least once in the sequence $y_{n+\eta_n/2+1}, \ldots, y_{n+\eta_n}$, it produces the pair $(\hat{m}_1, \hat{m}_2)$; otherwise it erases both messages. As in the proof of Theorem 1, the probability of an undetected error of this code is zero; the average erasure probability tends to zero; and the sequence of rate pairs tends to (R_1, R_2) as $n \to \infty$. ∎

V. Extension to Broadcast Channels

In this section we use W to denote the transition law of a generic discrete memoryless broadcast channel (DM-BC) with finite input alphabet $\mathcal{X}$ and finite output alphabets $\mathcal{Y}_1$ and $\mathcal{Y}_2$. The transition laws of the marginal channels to Receiver 1 and Receiver 2 are denoted W_1 and W_2, i.e.,

$$W_1(y_1|x) = \sum_{y_2 \in \mathcal{Y}_2} W(y_1, y_2|x), \quad x \in \mathcal{X},\, y_1 \in \mathcal{Y}_1,$$

and

$$W_2(y_2|x) = \sum_{y_1 \in \mathcal{Y}_1} W(y_1, y_2|x), \quad x \in \mathcal{X},\, y_2 \in \mathcal{Y}_2.$$

To simplify the presentation, we focus on the case where the transmitter wants to convey a message $m_1 \in \mathcal{M}_1$ to Receiver 1 and a message $m_2 \in \mathcal{M}_2$ to Receiver 2 (no common message). The definition of a zero-undetected-error code for the DM-BC is a straightforward extension of the definition for DMCs. We use $\mathcal{R}_{\mathrm{f}}^{\mathrm{BC}}$ and $\mathcal{R}_{\mathrm{0uf}}^{\mathrm{BC}}$ to denote the capacity region (average error probability criterion) and zero-undetected-error capacity region (average erasure probability criterion) when feedback from both receivers is available at the transmitter.[4]

Theorem 3: Let $\mathcal{V}_1$ denote the set of DM-BCs for which there is a triple $(x_{1,\mathrm{e}}, x_{1,\mathrm{c}}, y_{1,\mathrm{c}}) \in \mathcal{X} \times \mathcal{X} \times \mathcal{Y}_1$ such that $W_1(y_{1,\mathrm{c}}|x_{1,\mathrm{e}}) = 0$ and $W_1(y_{1,\mathrm{c}}|x_{1,\mathrm{c}}) > 0$. Similarly, let $\mathcal{V}_2$ denote the set of DM-BCs for which there is a triple $(x_{2,\mathrm{e}}, x_{2,\mathrm{c}}, y_{2,\mathrm{c}}) \in \mathcal{X} \times \mathcal{X} \times \mathcal{Y}_2$ such that $W_2(y_{2,\mathrm{c}}|x_{2,\mathrm{e}}) = 0$ and $W_2(y_{2,\mathrm{c}}|x_{2,\mathrm{c}}) > 0$. Then,

$$\mathcal{R}_{\mathrm{0uf}}^{\mathrm{BC}} = \begin{cases} \mathcal{R}_{\mathrm{f}}^{\mathrm{BC}} & \text{if } W \in \mathcal{V}_1 \cap \mathcal{V}_2, \\ [0, C_1] \times \{0\} & \text{if } W \in \mathcal{V}_1 \cap \mathcal{V}_2^{\mathrm{c}}, \\ \{0\} \times [0, C_2] & \text{if } W \in \mathcal{V}_1^{\mathrm{c}} \cap \mathcal{V}_2, \\ \{(0,0)\} & \text{if } W \in \mathcal{V}_1^{\mathrm{c}} \cap \mathcal{V}_2^{\mathrm{c}}, \end{cases}$$

where C_1 and C_2 denote the (ordinary) capacities of the marginal channels W_1 and W_2.

Proof: The direct part is similar to the direct parts of Theorems 1 and 2 and is omitted. For the converse consider the list of messages that cannot be ruled out by Receiver 2

$$\mathcal{L}_{\mathrm{f},2}^{\mathrm{BC}}(\mathbf{y}_2) = \Bigg\{ m_2 \in \mathcal{M}_2 : \exists m_1 \in \mathcal{M}_1,\, \mathbf{y}_1 \in \mathcal{Y}_1^n \text{ s.t. } \prod_{j=1}^{n} W\big(y_{1,j}, y_{2,j} | f_j(m_1, m_2, y_{1,1}^{j-1}, y_{2,1}^{j-1})\big) > 0 \Bigg\}, \tag{10}$$

where

$$f_j\colon \mathcal{M}_1 \times \mathcal{M}_2 \times \mathcal{Y}_1^{j-1} \times \mathcal{Y}_2^{j-1} \to \mathcal{X}, \quad j = 1, \ldots, n,$$

are the encoding functions. If $W \in \mathcal{V}_2^{\mathrm{c}}$, then $W_2(y_2|x) > 0$ implies $W_2(y_2|x') > 0$ for all $x' \in \mathcal{X}$. Consequently, if $m_2 \in \mathcal{L}_{\mathrm{f},2}^{\mathrm{BC}}(\mathbf{y}_2)$, we see from (10) that $W_2(y_{2,j}|x) > 0$ for all $x \in \mathcal{X}$ and all $j \in \{1, \ldots, n\}$, and therefore that for every $x \in \mathcal{X}$ and every $j \in \{1, \ldots, n\}$ there is $y_{1,j} \in \mathcal{Y}_1$ such that $W(y_{1,j}, y_{2,j}|x) > 0$. Thus, for every pair $(m_1, m_2') \in \mathcal{M}_1 \times \mathcal{M}_2$ we can find $y_{1,1} \in \mathcal{Y}_1$ such that $W(y_{1,1}, y_{2,1}|f_1(m_1, m_2')) > 0$, then find $y_{1,2} \in \mathcal{Y}_1$ such that $W(y_{1,2}, y_{2,2}|f_2(m_1, m_2', y_{1,1}, y_{2,1})) > 0$, and so on. Proceeding in this fashion we obtain a sequence $\mathbf{y}_1 \in \mathcal{Y}_1^n$ such that $\prod_{j=1}^{n} W(y_{1,j}, y_{2,j}|f_j(m_1, m_2', y_{1,1}^{j-1}, y_{2,1}^{j-1})) > 0$, which shows that $m_2' \in \mathcal{L}_{\mathrm{f},2}^{\mathrm{BC}}(\mathbf{y}_2)$. Thus, no message for Receiver 2 can be ruled out and we must have $|\mathcal{M}_2| = 1$ for all zero-undetected-error codes with average erasure probability less than 1. The case $W \in \mathcal{V}_1^{\mathrm{c}}$ is similar. ∎

[4]Theorem 3 remains true if both $\mathcal{R}_{\mathrm{f}}^{\mathrm{BC}}$ and $\mathcal{R}_{\mathrm{0uf}}^{\mathrm{BC}}$ are defined with a maximal instead of an average error and erasure probability criterion, respectively.

VI. Concluding Remarks

The observations in Sections II and III show that the zero-undetected-error capacity has a number of features in common with the zero-error capacity: while both are still unknown in general, both admit a single-letter expression in the presence of, and can be increased by, feedback. Moreover, formulas (5) and (8) exhibit a similar zero-nonzero dichotomy.

References

[1] G. Forney Jr, "Exponential error bounds for erasure, list, and decision feedback schemes," *IEEE Trans. Inf. Theory*, vol. 14, no. 2, pp. 206–220, 1968.

[2] I. Telatar and R. Gallager, "New exponential upper bounds to error and erasure probabilities," in *Proc. IEEE Int. Symp. Inf. Theory*, 1994, p. 379.

[3] I. Csiszár and P. Narayan, "Channel capacity for a given decoding metric," *IEEE Trans. Inf. Theory*, vol. 41, no. 1, pp. 35–43, 1995.

[4] R. Ahlswede, N. Cai, and Z. Zhang, "Erasure, list, and detection zero-error capacities for low noise and a relation to identification," *IEEE Trans. Inf. Theory*, vol. 42, no. 1, pp. 55–62, 1996.

[5] M. Pinsker and A. Sheverdyaev, "Transmission capacity with zero error and erasure," *Problemy Peredachi Informatsii*, vol. 6, no. 1, pp. 20–24, 1970.

[6] I. Telatar, "Multi-access communications with decision feedback decoding," Ph.D. dissertation, Massachusetts Institute of Technology, Dept. of Electrical Engineering and Computer Science, May 1992.

[7] ——, "Zero-error list capacities of discrete memoryless channels," *IEEE Trans. Inf. Theory*, vol. 43, no. 6, pp. 1977–1982, 1997.

[8] C. Shannon, "The zero error capacity of a noisy channel," *IRE Trans. Inf. Theory*, vol. 2, no. 3, pp. 8–19, 1956.

[9] I. Csiszár and J. Körner, *Information theory: coding theorems for discrete memoryless systems*. Cambridge University Press, 2011.

[10] J. Hui, "Fundamental issues of multiple accessing," Ph.D. dissertation, Massachusetts Institute of Technology, Dept. of Electrical Engineering and Computer Science, November 1983.

[11] M. Burnashev, "Data transmission over a discrete channel with feedback. random transmission time," *Problemy Peredachi Informatsii*, vol. 12, no. 4, pp. 10–30, 1976.