On the zero-error capacity of channels with noisy feedback

Meysam Asadi, Natasha Devroye
School of Electrical and Computer Engineering
University of Illinois at Chicago, Chicago IL 60607, USA
Email: masadi, devroye @ uic.edu

***Abstract*—The zero-error capacity of discrete memoryless channels (DMC) with noiseless feedback when variable-length codes are permitted has been shown to be positive whenever there exists at least one channel output "disprover", i.e. a channel output that cannot be reached from at least one of the inputs. Furthermore, whenever there exists a disprover, this variable-length zero-error capacity attains the Shannon (small-error) capacity. Here, we study the zero-error capacity of a DMC when the channel feedback is noisy. We show that the variable-length zero-error capacity with noisy feedback is lower bounded by the forward channel's zero-undetected-error capacity, and show that under certain conditions this is tight. We survey conditions under which the zero-error capacity without feedback, with perfect feedback, and with noisy feedback, are positive.**

I. Introduction

Shannon showed that for fixed block-length coding schemes, noiseless feedback does not increase either the small-error or zero-error capacity. For such channels, Shannon determined that the zero-error capacity C_0 of a point-to-point channel, whose channel $W(y|x)$ has confusability graph $G_{X|Y}$, is positive if and only if there exist two inputs that are "non-confusable" [1], i.e. if there exist two inputs that produce outputs in disjoint sets. Equivalently, it is non-zero if and only if the independence number of $G_{X|Y}$ is strictly greater than 1. This is quite a restrictive condition, and as a consequence, even a simple channel like the the binary symmetric channel which has a positive small-error capacity, has a fixed block-length zero-error capacity of zero with and without noiseless feedback.

However, for variable-length coding schemes, noiseless feedback can increase the zero-error capacity [1], [2]. As shown by Burnashev [2] and illustrated in a set of slides by Massey [3], it is possible to communicate at a non-zero average rate with zero-error over a DMC with noiseless feedback if, and only if, there exists at least one channel output that is reachable from some but not all the channel inputs, *if* one allows for variable-length codes. Such a channel output is called a "disprover". Not only does the existence of a disprover allow for positive rates, but with perfect feedback, the variable-length zero-error capacity of channels attains the small-error Shannon capacity C. Error exponent and finite block-length analyses of variable-length coding with perfect feedback can be found in [2] and [4] respectively.

The binary erasure channel (BEC) and the Z-channel are examples of channels whose zero-error capacity C_0 without feedback is equal zero, but, as both contain a disprover, have zero-error capacity equal to their Shannon capacity (positive in general) in the presence of perfect feedback. In order to achieve such zero-error rates, a variable-length coding scheme is used in which the transmitter repeatedly sends a message until it sees that it has been correctly received. Perfect output feedback allows the transmitter and receiver stay synchronized and in agreement about whether communication of a particular message is completed and another one is ready to start.

The variable-length zero-error capacity in the presence of feedback has strong connections with the zero-undetected-error capacity [5] with noiseless feedback [2], [6]. Two types of communication errors occur: i) erasure errors, when the decoder is unable to uniquely decode any message, and ii) undetected errors, when the decoder uniquely decodes an erroneous message. The zero-undetectable error capacity C_{0u}, first considered by Forney [5], denotes the maximal number of inputs that can be transmitted to ensure that the probability of an undetectable error is exactly zero. Forney derived a lower bound for the zero-undetected-error capacity (C_{0u}) of a channel, which he showed is positive if, and only if, this channel contains a disprover. A tighter lower bound on C_{0u} was later derived by Ahlswede [7], which was shown to be tight for two classes of channels in [8] and [9]. Finally, in [6] it was shown that the zero-undetected-error capacity for a channel with noiseless feedback, denoted by C_{0uf}, is equal to the small-error Shannon capacity C if the channel contains at least one disprover. Note that in general $C_0 \leq C_{0u} \leq C_{0uf} \leq C$.

In this paper, we are interested in the impact of *noisy* feedback on the variable-length zero-error capacity of channels. The central challenge here is that due to the noise in the feedback channel, the transmitter and receiver may not agree about whether the communication (of a particular message or messages) has terminated, causing not only a decoding error of the present message, but also creating problems in subsequent uses of the communication scheme.

Contribution. In this paper we focus on zero-error communication for a general DMC with feedback. We first define zero-error communication with and without perfect and noisy feedback, differentiating between block and variable-length codes. In Theorem 1, we detail the proof of a result attributed to Burnshev [2] and outlined in Massey's slides [3] for the zero error capacity of a channel with noiseless feedback, C_0^{VL-FB}. In Theorem 2, our main result, we consider noisy (rather than noiseless) feedback, and show that the variable-length zero-error capacity of the channel with noisy feedback, C_0^{VL-NFB}, is at least the zero-undetected-error capacity of the forward channel $C_{0u}^{(f)}$. Theorem 2 further outlines a class of channels for which this lower bound is tight.

II. Definitions

Let $x_i^j := (x_i, x_{i+1}, \ldots, x_j)$ when $i \leq j$ and $|x_i^j| = j-i+1$ denote its size. For simplicity we write $x^n = x_1^n$. Let $\mathcal{M}$ be the message set.

Channels. A channel $(\mathcal{X}, \mathcal{Y}, W)$ is used to denote a DMC with finite input alphabet $\mathcal{X}$, finite output alphabet $\mathcal{Y}$, and transition probability $W(y|x)$. We write W^n to denote the channel corresponding to n uses of W:

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

We consider channels with feedback, with a forward channel $(\mathcal{X}_{(f)}, \mathcal{Y}_{(f)}, W_{(f)})$ (subscript (f)) and a backward channel $(\mathcal{X}_{(b)}, \mathcal{Y}_{(b)}, W_{(b)})$ (for feedback, subscript (b)).

Small error fixed-length capacity C without feedback. A $\mathcal{C}(\mathcal{M}, n)$ fixed-length code for DMC W with message set $\mathcal{M}$ without feedback, consists of:

1) a message set $\mathcal{M}$ of size 2^{nR}, for R the rate and block-length n;

2) an encoding function $\mathcal{F}_n : \mathcal{M} \to \mathcal{X}^n$;

3) a decoding function and $\mathcal{G}_n : \mathcal{Y}^n \to \mathcal{M}$.

Let $c^{(n)}(m)$ denote a codeword corresponding to message $m \in \mathcal{M}$, i.e. $c^{(n)}(m) = \mathcal{F}_n(m)$ and let

$$\lambda_m^{(n)} = Pr(\mathcal{G}_n(y^n) \neq m | X^n = c^{(n)}(m)),$$

be the conditional probability of error given that message m was sent. The maximum probability of error for a $\mathcal{C}(\mathcal{M}, n)$ is defined as

$$\lambda^{(n)} = \max_{m \in \mathcal{M}} \lambda_m^{(n)}.$$

The small error capacity C for channel $(\mathcal{X}, \mathcal{Y}, W)$ is defined as the largest number R such that there exists a sequence of $\mathcal{C}(\mathcal{M}, n)$ codes such that $\lambda^{(n)}$ tends to 0 as $n \to \infty$.

Zero-error fixed-length capacity C_0 without feedback. In his 1956 paper, Shannon defined the zero-error capacity

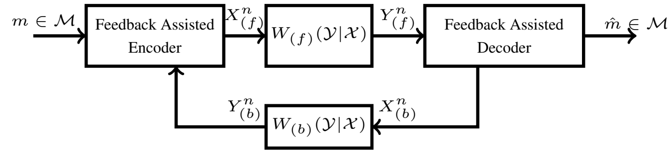

Fig. 1. Communication Scheme for a DMC with active noisy feedback

C_0 as the largest number such that there exists a sequence of $\mathcal{C}(\mathcal{M}, n)$ fixed-length codes such that $\lambda_m^{(n)} = \lambda^{(n)}$ equals 0.

Remark: Note that without feedback, it is known [10, Theorem 16] that the variable-length capacity is equal to the fixed-length capacity for any non-anticipatory channel.

Zero-undetected-error fixed-length capacity C_{0u} [6]. A zero-undetected-error code of block-length n, denoted by $\mathcal{C}_{0u}(\mathcal{M}, n)$, again consists of:

1) a message set $\mathcal{M}$ of size 2^{nR}, for R the rate and block-length n

2) an encoding function $\mathcal{F}_{0u,n} : \mathcal{M} \to \mathcal{X}^n$, that encodes messages m to $c_{ou}^{(n)}(m)$,

3) a decoding function $\mathcal{G}_{0u,n} : \mathcal{Y}^n \to \mathcal{M} \cup \{0\}$ described as follows. Let $M(y_1^n)$ denote the set of probable messages corresponding to a received output y_1^n

$$M(y^n) = \{m \in \mathcal{M} : W^n(y^n | c_{ou}^{(n)}(m)) > 0\}. \tag{1}$$

The decoder declares an erasure, denoted by 0, if there exist more than one possible message that could have yielded output y^n, i.e. $|M(y^n)| > 1$. A zero-undetected-error decoder function is then defined as

$$\mathcal{G}_{0u,n}(y^n) = \begin{cases} M(y^n) & \text{if } |M(y^n)| = 1 \\ 0 & \text{if } |M(y^n)| > 1. \end{cases}$$

4) A zero-error guarantee: a zero-undetected-error code must have no undetected errors, hence the maximal error probability is given only by the probability of erasures as

$$\lambda_m = Pr(\mathcal{G}_{0u,n}(y^n) = 0 | X^n = c_{0u}^{(n)}(m)).$$

The zero-undetected-error capacity C_{0u} for channel W is defined as the largest rate R such that there exist a sequence of $\mathcal{C}_{0u}(\mathcal{M}, n)$ codes that $\max_{m \in \mathcal{M}} \lambda_m$ tends to 0 as $n \to \infty$.

All previous definitions involve fixed-length codes, hence our usage of the "fixed-length" in the definitions. This implies that the codeword length is fixed to n for all messages and channel instances, and decoding is performed after n channel uses. We next define variable-length codes for the small and zero-error regimes for noiseless feedback, followed by the more delicate definition required for noisy feedback.

We use the notation $\mathcal{C}_0^{VL-PF}$ and $\mathcal{C}_0^{VL-NF}$ to distinguish the zero-error variable-length capacities with perfect (noiseless) feedback and with noisy feedback.

Zero-error variable-length capacity with perfect (noiseless) feedback, $\mathcal{C}_0^{VL-PF}$.

Following [4], a variable-length zero-error feedback code $\mathcal{C}_0^{VL-PF}(\mathcal{M}, l)$ for DMC $(\mathcal{X}, \mathcal{Y}, W)$ consists of:

1) a message set $\mathcal{M}$, where messages are equi-probable;

2) a sequence of encoding functions $\mathcal{F}_n : \mathcal{M} \times \mathcal{Y}^{n-1} \to \mathcal{X}$ defining inputs

$$X_n = \mathcal{F}_n(M, Y^{n-1});$$

3) a sequence of decoding functions $\mathcal{G}_n : \mathcal{Y}^n \to \mathcal{M} \cup \{0\}$ yielding the best estimate of the message $m \in \mathcal{M}$ at time n or declaring erasure (denoted by 0);

4) a non-negative integer-valued stopping time N (random variable) defined as the first n that the decoder does not declare an erasure.

$$N = n \text{ if } \forall k < n, \mathcal{G}_k(y^k) = 0 \text{ and } \mathcal{G}_n(y^n) \neq 0$$

which satisfies $E[N] \leq l$;

5) a zero-error guarantee: decoding is performed at time instant N (the stopping time), yielding the message estimate $\widehat{M} = \mathcal{G}_N(Y^N)$ and must satisfy $\lambda_m^{(N)} = \lambda^{(N)} = 0$.

The average-rate $\bar{R}$ is called achievable if there exists a sequence of variable-length zero-error feedback codes $\mathcal{C}_0^{VL-PF}(\mathcal{M}, l)$, where $\mathcal{M}$ may be a function of l, for which

$$\bar{R} \leq \lim_{l \to \infty} \frac{\log_2 |\mathcal{M}|}{E[N]}.$$

The largest average rate $\bar{R}$ achievable by any zero-error variable-length code $\mathcal{C}_0^{VL-PF}$ is called the zero-error variable-length capacity with noiseless feedback, $\mathcal{C}_0^{VL-PF}$.

Zero-error variable-length capacity with noisy feedback, $\mathcal{C}_0^{VL-NF}$.

Recall that in a channel with noisy feedback, the forward channel is denoted by $(\mathcal{X}_{(f)}, \mathcal{Y}_{(f)}, W_{(f)})$ (subscript (f)) and the backward channel by $(\mathcal{X}_{(b)}, \mathcal{Y}_{(b)}, W_{(b)})$.

Variable-length codes achieving zero-error with noisy feedback as in Fig. 1 are more subtle to define, as the transmitter and receiver, due to the noisy channels in both directions, need to take care to stay synchronized – not at the channel use level (channel use level synchronization is always assumed), but at the message level. That is, they need to agree upon when a given message has been correctly received and when a new message has begun. Few formal definitions of such a channel model for the variable-length, noisy feedback regime (fixed block-length, noisy feedback and variable-length, noiseless feedback definitions abound) *with zero-error* exist, though Massey given an informal, intuitive definition in [3], and Draper and Sahai [11] tackle this in the small-error regime.

To capture the subtle effects of synchronization (knowing when communication is terminated and a next message starts) for channels with noisy feedback, we follow the definitions of [11] for the transmission of a sequence of messages. Formulations using this sequence or stream of messages capture the effect of synchronization and scenarios in which the encoder and decoder might not be synchronized even after correctly decoding a particular message.

A zero-error variable-length code with noisy feedback $\mathcal{C}_0^{VL-NF}(\mathcal{M}_1, \mathcal{M}_2, \cdots, \mathcal{M}_s, l)$ consists of:

1) a sequence of s message sets $\mathcal{M}_1, \mathcal{M}_2, \cdots, \mathcal{M}_s$;

2) a sequence of forward channel encoding functions

$$\mathcal{F}_n^{(f)} : \prod_{j=1}^{s} \mathcal{M}_j \times \mathcal{Y}_{(b)}^{n-1} \to \mathcal{X}_{(f)};$$

3) a sequence of feedback channel encoding functions (this allows for active feedback)

$$\mathcal{F}_n^{(b)} : \mathcal{Y}_{(f)}^n \to \mathcal{X}_r;$$

4) a sequence of decoding functions

$$\mathcal{G}_n : \mathcal{Y}_{(f)}^n \to \prod_{j=1}^{s} \mathcal{M}_j \cup \{0\}$$

in which an output 0 means an erasure happened (at least one of the messages cannot be decoded or is decoded incorrectly). Note that the decoder at time n makes an estimate of all the messages sent based on the received symbols $Y_{(f)}^n$;

5) a non-negative integer-valued stopping time N (random variable) defined as the first n that the decoder is able to decode all s messages:

$$N = n \text{ if } \forall k < n,\ \mathcal{G}_k(y^k) = 0 \text{ and } \mathcal{G}_n(y^n) \neq 0,$$

where $E[N] \leq l$;

6) a zero-error guarantee: let $\lambda^{(N,s)}$ denote the maximum error probability among all the s messages when the decoding is performed at stopping time N:

$$\lambda^{(N,s)} \triangleq \max_{1 \leq j \leq s} P(\hat{m}_j \neq m_j).$$

A zero-error variable-length code with noisy feedback must have $\lambda^{(N,s)} = 0$.

The average-rate $\bar{R}$ is called achievable if there exists a sequence of zero-error variable-length codes with noisy feedback such that

$$\bar{R} \leq \lim_{l \to \infty} \lim_{s \to \infty} \frac{\sum_{j=1}^{s} \log_2 |\mathcal{M}_j|}{E[N]}.$$

The largest average rate $\bar{R}$ achievable by any zero-error variable-length code is called the zero-error variable-length capacity with noisy feedback, $\mathcal{C}_0^{VL-NF}$.

III. Fixed-length zero-error communication without feedback

We recap known results on the positivity condition and capacity for zero-error communication without feedback.

When is the zero-error capacity positive? Shannon [1] defined the zero-error capacity of the point- to-point channel $(\mathcal{X}, p(y|x), \mathcal{Y})$ without feedback (denoted by C_0). He obtained a necessary and sufficient condition for C_0 to be positive expressed using "non-adjacency" of inputs. Two inputs x, x' are called non-adjacent if their reachable sets

$$Y(x) := \{y \in \mathcal{Y} | p(y|x) > 0\}, Y(x') := \{y \in \mathcal{Y} | p(y|x') > 0\},$$

are disjoint.

Theorem 1 (Shannon [1]): The zero-error capacity of the point-to-point channel $(\mathcal{X}, p(y|x), \mathcal{Y})$ is strictly positive if and only if there exist two inputs $x \neq x'$ that are non-confusable, i.e. for which

$$Y(x) \cap Y(x') = \emptyset.$$

Zero-error capacity without feedback. Let $\alpha(G_{X|Y})$ be the maximum independent set of the confusability graph $G_{X|Y}$[1] of the channel $W(y|x)$. If we use the channel n times, $\alpha(G_{X^n|Y^n})$ input vectors will be able to be distinguished with zero error. It is easy to check that the confusability graph of the channel from X^n to Y^n, $G_{X^n|Y^n}$ is given by the strong product of n copies of the confusability graph $G_{X|Y}$, i.e.

$$G_{X^n|Y^n} = G_{X|Y}^{\boxtimes n}.$$

We then call

$$C_0 \triangleq \sup_n \frac{1}{n} \log_2 \left(\alpha(G_{X|Y}^{\boxtimes n})\right) = \lim_{n \to \infty} \frac{1}{n} \log_2 \left(\alpha(G_{X|Y}^{\boxtimes n})\right)$$

the Shannon zero-error capacity C_0 of the channel with confusability graph $G_{X|Y}$. Notice that this is not a single-letter quantity, and is challenging to compute in general [12].

IV. Variable-length zero-error communication for channels with feedback

When perfect output feedback is available at the transmitter, it can verify correct reception of the message. One way of ensuring *zero-error* communication in this scenario is to keep repeating a message until it is correctly received. For sake of completeness, we first recall this communication scheme for channels with perfect feedback. Then, we devise a similar communication scheme for channels with noisy feedback.

We denote $x_i^j \doteq x'$ if there exists at least one $k \in [i, j]$ such that $x_k = x'$ (e.g. $1101 \doteq 0$). Let $[x_i]^m$ denote a sequence of m repetitions of letter x_i in some alphabet $\mathcal{X}$, $[x_i]^m = (x_i, x_i, \ldots, x_i), |[x_i]^m| = m$. Let $\gamma_n = o(n)$, and $\gamma_n \to \infty$ as $n \to \infty$ (e.g. $\gamma_n = \log(n)$).

[1] A confusability graph $G_{X|Y}$ of channel W is a graph whose vertex set is $\mathcal{X}$ and an edge is placed between vertices $x, x' \in \mathcal{X}$ if they may be confused, that is, if $\exists y \in \mathcal{Y} : W(y|x)W(y|x') > 0$.

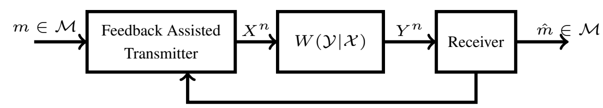

Fig. 2. Communication Scheme for a DMC with noiseless feedback

A. Perfect (noiseless) feedback

When perfect feedback is available (Figure 2), Burnashev [2] showed that the error exponent (maximal exponential decay rate of the probability of error with increasing block-length) when variable-length codes are permitted is

$$E(R) = \frac{C_1}{C}(C - R)$$

for all rates $0 < R < C$, where C_1 is the maximal relative entropy between output distributions,

$$C_1 = \max_{x_i, x_j} \sum_y W(y|x_i) \log \frac{W(y|x_i)}{W(y|x_j)}.$$

Note that when the channel contains at least one disprover, then $C_1 = \infty$ and zero-error communication is possible. In this case, it can be shown that the variable-length zero-error capacity is actually the Shannon capacity of the forward channel, $\mathcal{C}_0^{VL-PF} = C$. Since the backward channel is noiseless, we omit subscript $_{(f)}$ for the forward channel and use $(\mathcal{X}, \mathcal{Y}, W)$.

Theorem 2 (Burnashev and Massey (Elaborated) [2], [3]): The variable-length zero-error capacity $\mathcal{C}_0^{VL-PF}$ for a DMC channel $(\mathcal{X}, \mathcal{Y}, W)$ with noiseless feedback is

$$C_0^{VL-FB} = \begin{cases} C & \text{if } C_{0u} > 0 \\ 0 & \text{otherwise,} \end{cases} \tag{2}$$

where C denotes the Shannon capacity of the channel $(\mathcal{X}, \mathcal{Y}, W)$, and C_{0u} denotes its zero-undetected-error capacity.

Proof If $C_{0u} = 0$ then by [5], channel W does not have a disprover, i.e. for every $x \in \mathcal{X}, y \in \mathcal{Y}, W(y|x) > 0$. Thus, no matter which sequence is sent the receiver is unable to decide anything with zero error and $C_0^{VL-FB} = 0$.

When $C_{0u} > 0$, we may assume that the DMC W contains at least one disprover. Equivalently, there exists at least one triple $(x_c, x_e, y_c) \in (\mathcal{X} \times \mathcal{X} \times \mathcal{Y})$ such that $W(y_c|x_e) = 0$ and $W(y_c|x_c) > 0$.

The converse proof is trivial using $C_0^{VL-FB} \overset{(1)}{\leq} C_f \overset{(2)}{=} C$, where (1) follows as C_f denotes the small-error capacity of the channel with perfect feedback, which is always an outer bound to the more restrictive zero-error setting, and (2) follows from Shannon's result that perfect feedback does not increase the small error capacity of a channel.

For the achievability, let $\mathcal{C}(\mathcal{M}, n)$ be a capacity achieving code for the DMC W whose maximal probability of error

Algorithm 1: Variable-length zero-error communication scheme with complete feedback [2].

Feedback Assisted Encoder;
Input : $M \subseteq \mathcal{M}, \mathcal{C}(\mathcal{M}, n), \gamma_n$, disprover triplet $(x_c, x_e, y_c) \in W$
Output: L
forall the $m \in M$ **do**
  $x^n \leftarrow c^{(n)}(m)$, $I \leftarrow 0$, $L \leftarrow 0$;
  while $I = 0$ **do**
    $L \leftarrow L + 1$;
    Send x^n through channel ;
    $\hat{m} = \mathcal{G}(y^n(c^{(n)}(m)))$;
    ; /* L-th verification */
    if $\hat{m} \neq m$ **then**
      $x^{\gamma_n} \leftarrow [x_e]^{\gamma_n}$;
    else
      $x^{\gamma_n} \leftarrow [x_c]^{\gamma_n}$;
    end
    Send x^{γ_n} through channel ;
    if $y(x^{\gamma_n}) \doteq y_c$ **then**
      $I \leftarrow 1$;
    end
  end
end

$\lambda^{(n)}$ tends to zero and whose rate approaches the Shannon capacity C as block length $n \to \infty$. Note that the output block y^n is available in real time at the transmitter due to the presence of perfect feedback. The transmitter can thus mimic the receiver's decoding rule and determine whether the receiver obtained the correct message. It then tells the receiver this by sending γ_n copies of either x_c (if correct) or x_e (if erroneous) through the noisy W. Since the receiver can only receive a y_c from an x_c (definition of a disprover), once it receives at least one y_c it realizes that its decoded message is correct, and zero-error communication is achieved. We note that variable I in Algorithm 1 is used to synchronize the transmitter and receiver, i.e. indicates when a new message will start.After each iteration ($n + \gamma_n$ channel uses), two cases might occur; case (1), y_c is not received. This happens either when there is a decoding error or when the decoder correctly decodes, but no y_c is received between the n-th and $n + \gamma_n$-th channel uses (synchronization error). In case(2), the decoder correctly decodes and at least one y_c is received between n-th and $n + \gamma_n$-th channel uses. Note that transmission continues until case (2) happens.

To formalize this, which will be useful in the next section with noisy feedback, let $c_n(m) = (c_1, c_2, \cdots, c_n) \in \mathcal{C}(\mathcal{M}, n)$ be the codeword corresponding to $m \in \mathcal{M}$. Let L be the random variable taking on value in $\mathbb{N}$ denoting the number of rounds (a round consists of sending a codeword of length n followed by a "correct" x_c or "erroneous" x_e signaling of length γ_n) needed for the decoder to correctly decode the message, and the transmitter/receiver to synchronize, respectively. Figure 3 shows an example of of when $L = 3$. The distributions of these random variables will be obtained shortly.

Let $N = L \cdot (n + \gamma_n)$ denote the stopping time (time at which the receiver first sees a y_c in the block of length γ_n following a codeword of length n). At time interval $[N - \gamma_n, N]$, the receiver for the first time has received at least one y_c and hence knows it has decoded the message correctly and that the next n channel uses will comprise a new message. Given $N = n^* = l^* * (n + \gamma_n)$, at the i-th channel use the encoding function gives

$$\mathcal{F}_i(m, y^{i-1}) = \begin{cases} c_j & \text{if } 1 \leq j \leq n \\ x_e & \text{if } j \geq n \text{ and } l \geq l^* \\ x_c & \text{if } j \geq n \text{ and } l = l^*, \end{cases}$$

where $j = i \mod (n + \gamma_n)$ and $l = \lfloor i/(n + \gamma_n) \rfloor$. Also,

$$\mathcal{G}_{i \geq n^*}(y^i) \neq 0 \text{ and } \lambda^{(n^*)} = 0.$$

After $n + \gamma_n$ channel uses, the probability of successful transmission of m, denoted by $p_{n+\gamma_n,m}$, is computed as

$$p_{n+\gamma_n,m} = (1 - \lambda_m^{(n)}) \Big[1 - \big(1 - W(y_c|x_c)\big)^{\gamma_n} \Big].$$

Thus, the number of codeword re-transmissions needed to correctly receive message m and wait for the transceivers to synchronize and start a new message is a geometric random variable L with $E[L|M = m] = 1/p_{n+\gamma_n,m}$.

$$E[L] = \frac{1}{|\mathcal{M}|} \sum_{m=1}^{|\mathcal{M}|} (p_{n+\gamma_n,m})^{-1}.$$

Thus, recalling that $N = L \cdot (n + \gamma_n)$

$$E[N] = \frac{(n + \gamma_n)}{|\mathcal{M}|} \sum_{m=1}^{|\mathcal{M}|} (p_{n+\gamma_n,m})^{-1}.$$

Hence, as $n \to \infty$

$$\begin{aligned} \bar{R} &= \lim_{n\to\infty} \frac{\log_2 |\mathcal{M}|}{E[N]} \\ &= \lim_{n\to\infty} \frac{\log_2 |\mathcal{M}|}{n} \frac{|\mathcal{M}|}{(1 + \frac{\gamma_n}{n}) \sum_{m=1}^{|\mathcal{M}|} (p_{n+\gamma_n,m})^{-1}} \stackrel{(2)}{=} C \end{aligned}$$

where (2) follows as we are using a Shannon capacity achieving code $\mathcal{C}(\mathcal{M}, n)$, and by definitions of $\gamma_n = o(n)$ and the fact that $p_{n+\gamma_n,m} \to 1$ as $n \to \infty$.

Fig. 3. Variable-length communication scheme with perfect feedback ($L = 3$).

B. Noisy feedback

When the feedback channel is noisy, the above scheme no longer works as i) the transmitter does not have perfect access to the received signal, and hence cannot mimic the decoding process. It is thus harder to ensure zero error; and ii) synchronizing the transmitter and receiver becomes more challenging as both channels are noisy. How can the receiver know when a codeword is new versus when it is repeated? When feedback is noiseless, the synchronization issue can be completely resolved at the transmitter and communicated to the receiver using a disprover triplet. With noisy feedback, we propose a new synchronization technique that involves the sequential transmission of messages (where we note that our definition of C_0^{VL-NF} allows for multiple messages to be simultaneously transmitted).

Let $s_t, s_r \in \{0,1\}$ be the current states of the transmitter and receiver respectively. In order to synchronize, both transmitter and receiver exchange their state in each round of message transmission. The transmitter and receiver are synchronized based on the following rules; i) Transmission starts when $s_t = s_r$. ii) The receiver only accepts a new message if $s_r = s_t$, otherwise it asks for a retransmission. iii) The receiver flips its state s_t the first time that it decodes correctly (a zero-undetected-error code will be used), and $s_r = s_t$. iv) The transmitter flips its state the first time that it receives y_c through the feedback channel. When the states are synchronized, both transmitter and receiver are working on transmitting a new message; when different, the receiver has decoded the message but the transmitter does not know this yet due to the noisy feedback channel. In [3] a simple version of variable-length zero-error communication scheme for DMC with noisy feedback was proposed. The synchronized feedback assisted transmitter and receiver are described using Algorithms 2 and 3, respectively.

Theorem 3: The variable-length zero-error capacity of a forward DMC $W_{(f)}$ with noisy feedback DMC $W_{(b)}$ shown in Fig. 1, denoted by $\mathcal{C}_0^{VL-NF}$, satisfies

$$\mathcal{C}_0^{VL-NF} \geq C_{0u}^{(f)} \quad \text{if } C_{0u}^{(f)} > 0 \text{ and } C_{0u}^{(b)} > 0, \tag{3}$$

where $C_{0u}^{(f)}$ and $C_{0u}^{(b)}$ denote the zero-undetected-error capacities of the forward and backward links. If furthermore, for some positive functions $A(\cdot)$ and $B(\cdot)$ and some capacity-achieving input distribution Q^*, $W_{(f)}(y|x) = A(x)B(y)$ holds whenever $Q^*(x)W_{(f)}(y|x) > 0$, then $\mathcal{C}_0^{VL-NF} = C^{(f)}$, where

Algorithm 2: Variable-length zero-error communication scheme with Noisy feedback

Synchronized Feedback Assisted Transmitter (Fig. 1);
Input : $m \in \mathcal{M}, \mathcal{C}_{0u}^{(f)}(\mathcal{M}, n), \gamma_n, y_c \in \mathcal{Y}_{(b)}$
Output: $L_n(m)$
$s_t \leftarrow 0$ `/* Transmitter state */`
forall the *m that need to be sent* **do**
  $b_1 \leftarrow s_t$;
  $(b_2, b_3, \cdots, b_k) \leftarrow m$;
  $x^n \leftarrow c_{0u}^{(n)}(b_1^k)$, $I \leftarrow 0$, $L \leftarrow 0$;
  while $I = 0$ **do**
    $L \leftarrow L + 1$;
    ; `/* L-th transmition Stage */`
    Send x^n through channel ;
    $\hat{m} = \mathcal{G}_{0u}(y_1^n(c_{0u}^n(m)))$;
    ; `/* L-th verification stage */`
    Receive $y_1^{\gamma_n}$ through feedback channel ;
    if $y_1^{\gamma_n} \doteq y_c$ **then**
      $I \leftarrow 1$;
    end
  end
  $s_t \leftarrow s_t^c$ `/* Inform receiver about new message */`
end

$C^{(f)}$ denotes the fixed-length capacity without feedback of the forward link.

Proof Note that we require $C_{0u}^{(f)}$ to be positive, else no zero-error communication can take place at all, not even with perfect feedback. When $C_{0u}^{(b)}$ is positive, there exists at least one triple $(x_c', x_e', y_c') \in (\mathcal{X}_{(b)} \times \mathcal{X}_{(b)} \times \mathcal{Y}_{(b)})$ such that $W_{(b)}(y_c'|x_e') = 0$ and $W_{(b)}(y_c'|x_c') > 0$.

To show the achievability of (3), take a zero-undetected-error capacity achieving code $C_{0u}(\mathcal{M}, n)$ for channel $W_{(f)}$ whose maximal erasure probability tends to zero and whose rate approaches $C_{0u}^{(f)}$. In order to synchronize we assign the first message bit b_1 out of the bit stream of length k, b_1^k (that is encoded) to carry the transmitter's state variable s_t. To transmit message $m \in \mathcal{M}$, codeword $c_{0u}^{(n)}(m)$ is sent through $W_{(f)}$. Upon receiving $y^n \in \mathcal{Y}_{(f)}^n$, the zero-undetected-error decoder is used to obtain an estimate of the message. Since the probability of undetected-error is equal to zero, the only type of error that might occur is an erasure ($|M(y^n)| > 1$,

Algorithm 3: Iterative zero-error communication scheme with Noisy feedback

Synchronized Feedback Assisted Receiver (Fig. 1);
Input : $y^n \in \mathcal{Y}^n_{(f)}, \gamma_n, (x_e, x_c) \in \mathcal{X}_{(b)}$ for $W_{(b)}$
Output: $(\hat{b}^k_2)$
$s_r \leftarrow 0$ /* Receiver state */
forall the *Received* y^n **do**
  $M(y^n) = \{m \in \mathcal{M} : W_{(f)}(y^n | c^{(n)}_{0u}(m)) > 0$ or $W_{(f)}(y^n | c^{(n)}_{0u}(m)) > 0\}$;
  if $|M(y^n)| = 1$ **then**
    $\hat{b}^k_1 \leftarrow M(y^n)$;
    if $\hat{b}_1 = s_r$ **then**
      $\hat{m} \leftarrow \hat{b}^k_2$ /* Store message $\hat{m}$ */
      $s_r \leftarrow s^c_r$;
      /* Complement s_r to indicate ready for new message */
    end
    $x^{\gamma_n} \leftarrow [x_c]^\gamma_n$;
  else
    $x^{\gamma_n} \leftarrow [x_e]^\gamma_n$;
  end
  Send x^{γ_n} through feedback channel $W_{(f)}$;
end

see (1)). If there is an erasure, according to Algorithm 3, the receiver informs the transmitter by sending γ_n repetitions of the letter x'_e (i.e. it sends $[x'_e]^{\gamma_n}$). Since $W_{(b)}(y'_c|x'_e) = 0$, it is impossible to receive y'_c at the transmitter through the noisy feedback channel. If the transmitter does not see any y'_c it again transmits $c^{(n)}_{0u}(m)$.

A message is re-transmitted until both of the following conditions are satisfied: i) The codeword is not erased, ii) The receiver and transmitter are synchronized, i.e. $s_t = s_r$ (recalling that the first bit b_1 carries the state s_t and not the message). After the first round that both of these conditions are satisfied, the receiver sets $\hat{m} = M(y^n) = b^k_2$, and flips its own state $s_r = s^c_r$. It knows with probability 1 that this is the correct message (i.e. zero-error in decoding the message by definition of a zero undetected error code). The challenge now is to tell the transmitter, through the noisy channel, that it has received the message and hence that the transmitter can move on to a new message. To do this, the receiver conveys correct decoding by sending γ_n repetitions of the letter x'_c, i.e. transmits $[x'_c]^{\gamma_n}$ through the feedback channel $W_{(b)}$.

Two things can now happen: 1) If the letter y'_c is not received at the transmitter, then the transmitter sends back the same message and the process repeats. At this stage then, $s_t = b_1$ while $s_r = b^c_1$. This process repeats until the decoder uniquely decodes $|M(y^n)| = 1$ AND the state bits match. If the state bits do not match, the receiver does not update the decoded message and still sends back $[x'_c]^{\gamma_n}$. 2) If it does receive y'_c, then it knows the receiver successfully and uniquely decoded the message and hence it sets $s_t = s^c_t$. At this point then the transmitter and receiver states are again equal, $s_t = s_r$. A new message, with the new state again as first bit, is transmitted.

Let L' and L be two random variables taking on values in $\mathbb{N}$ denoting the number of rounds (a round consists of sending a codeword of length n over forward channel W_f followed by a "correct" x'_c or "erroneous" x'_e signaling of length γ_n over the backward channel W_b) needed for the decoder to correctly decode the message, and for the decoder to correctly decode the message and the transmitter/receiver to synchronize, respectively. Let $N' = L' \cdot (n + \gamma_n) - \gamma_n$ denote the first channel use where the decoder correctly decodes the message (though the receiver may not know it, the transmitter does), and let $N = L \cdot (n + \gamma_n)$ denote the stopping time (time at which the transmitter first sees a y'_c over the backward channel in the block of length γ_n). Note that, in general, $L' \leq L$ and $N' < N$. Figure 4 shows an example of the case that $L' = 2, L = 3$. Given $N' = n_1, N = n_2$, at the i-th channel use the forward and backward encoding functions are given as

$$\mathcal{F}^f_i(m, y^{i-1}) = \begin{cases} c_j & \text{if } 1 \leq j \leq n \\ \text{Idle} & \text{if } j > n, \end{cases}$$
$$\mathcal{F}^b_i(m, y^{i-1}) = \begin{cases} \text{Idle} & \text{if } 1 \leq j \leq n \\ x_e & \text{if } j > n \text{ and } 1 \leq i \leq n_1 \\ x_c & \text{if } j > n \text{ and } n_1 \leq i \leq n_2, \end{cases} \quad (4)$$

where $j = i \mod (n+\gamma_n)$. To calculate the average rate, note that the probability that a message is correctly received with zero error is the probability that the message was correctly received at the receiver after seeing the codeword of length n, and then the *transmitter* (now through a noisy channel) seeing at least one y'_c in a block of length γ_n. Hence after $n + \gamma_n$ channel uses, the probability of successfully transmitting and synchronizing message m is

$$p'_{n+\gamma_n,m} = (1 - \lambda^{(n)}_m)\Big[1 - \big(1 - W_{(b)}(y'_c|x'_c)\big)^{\gamma_n}\Big].$$

Viewing this as a probability of success, the number of codeword re- transmissions needed to transmit message m is hence a geometric random variable L with $E[L|M = m] = 1/p'_{n+\gamma_n,m}$. Thus, as before

$$E[N] = \frac{(n + \gamma_n)}{|\mathcal{M}|} \sum_{m=1}^{|\mathcal{M}|} (p'_{n+\gamma_n,m})^{-1}.$$

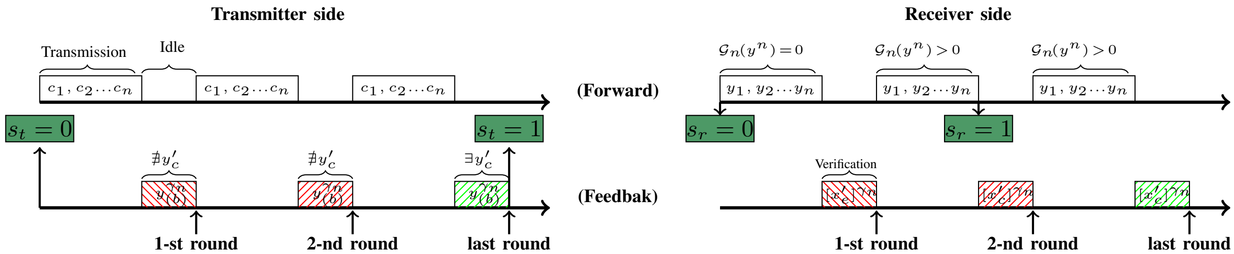

Fig. 4. Variable-length communication scheme with Noisy feedback ($L' = 2$ and $L = 3$). The receiver decodes successfully after 2 rounds and the transmitter and receiver are synchronized after 3 rounds.

Hence, as $n \to \infty$

$$\begin{aligned} \bar{R} &= \lim_{n\to\infty} \frac{\log_2(|\mathcal{M}|)}{E[N]} \\ &= \lim_{n\to\infty} \frac{(k-1)}{n} \frac{2^{k-1}}{(1+\frac{\gamma_n}{n}) \sum_{m=1}^{2^{k-1}} (p'_{n+\gamma_n,m})^{-1}} \overset{(2)}{=} C_{0u}^{(f)}. \end{aligned}$$

The analysis of the achieved average rate is similar to Theorem 1, except that we now use the backward $W_{(b)}(y_c'|x_c')$ in the definition of $p'_{n+\gamma_n,m}$, and the code we use is a zero-undetected-error capacity achieving code, in which case the rate tends to $C_{0u}^{(f)}$ as $n \to \infty$.

To show that our bound is tight for the class of channels stated below (3), note that Csiszár and Narayan showed that if $C_{0u}^{(f)} > 0$, and if the conditions after (3) hold then the zero-undetected-error capacity becomes equal to the small error Shannon capacity ($C_{0u}^{(f)} = C^{(f)}$). For these channels, $\mathcal{C}_0^{VL-NF} \geq C^{(f)}$. This is tight, as we always have $\mathcal{C}_0^{VL-NF} \leq \mathcal{C}_0^{VL-PF} \leq C^{(f)}$.

Remark In order to make use of all channel uses in the forward channel, it is possible to concurrently send more than one message. To be able to synchronize more than one message transmission, a two bit state variable suffices, i.e. $s_t, s_r \in \mathcal{S}, |\mathcal{S}| = 4$. The transmitter can pre-pend the 2 state bits to the message (as before, with 1 state bit) and the above can be appropriately altered but is omitted due to space constraints. This scheme does not improve the overall rate, but may lead to faster decoding for some of the messages.

V. Conclusion

A major difference between our adaptive-zero-error communication schemes with noiseless versus noisy feedback is that the verification sequence (i.e. transmitter and receiver agreeing the receiver has decoded it successfully) is sent by the transmitter in the noiseless case whereas it is sent by the receiver in the noisy case. In the former, the perfect feedback allows us to approach rates up to C as undetected errors can be caught by the transmitter. In the latter, due to the noisy feedback, our scheme must backoff from C to C_{0u} in order to ensure that no undetected errors occur, as they cannot be corrected by the transmitter under our scheme.

Acknowledgment

The work of the authors was partially supported by NSF under awards 1645381 and 1053933. The contents of this article are solely the responsibility of the authors and do not necessarily represent the official views of the NSF. The authors would like to thank the anonymous reviewers of earlier versions of this work.

References

[1] C. E. Shannon, "The zero error capacity of a noisy channel," *IRE Trans. Inform. Theory*, vol. IT-2, no. 3, pp. 8–19, Sep. 1956.
[2] M. Burnashev, "Data transmission over a discrete channel with feedback. random transmission time." *Probl. Inf. Transm.*, vol. 12, no. 4, pp. 10–30, 1976.
[3] J. Massey, "Zero Error," *Winter School of Information Theory, slides*, March 2007.
[4] Y. Polyanskiy, H. Poor, and S. Verdú, "Feedback in the non-asymptotic regime," *IEEE Trans. Inf. Theory*, vol. 57, no. 8, pp. 4903–4925, Aug. 2011.
[5] 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.
[6] C. Bunte and A. Lapidoth, "The Zero-Undetected-Error Capacity of Discrete Memoryless Channels with Feedback," Monticello, IL, Oct. 2012, pp. 1838–1842.
[7] 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.
[8] M. Pinsker and A. Sheverdyaev, "Transmission capacity with zero error and erasure," *Problemy Peredachi Informatsii*, vol. 6, no. 1, pp. 20–24, 1970.
[9] I. Csiszar and P. Narayan, "Channel capacity for a given decoding metric," *IEEE Trans. Inf. Theory*, vol. 41, no. 1, pp. 35–43, 1995.
[10] S. Verdú and S. Shamai, "Variable-rate channel capacity," *IEEE Trans. Inf. Theory*, vol. 56, no. 6, pp. 2651–2667, Jun. 2010.
[11] S. Draper and A. Sahai, "Variable-length channel coding with noisy feedback," *European Trans. on Telecommunications*, vol. 19, no. 4, pp. 355–370, Apr. 2008.
[12] N. Alon and E. Lubetzky, "The Shannon capacity of a graph and the independence numbers of its powers," *IEEE Trans. Inf. Theory*, vol. 52, no. 5, pp. 2172–2176, May 2006.