2018 56th Annual Allerton Conference on Communication, Control, and Computing (Allerton)
Allerton Park and Retreat Center
Monticello, IL, USA, October 2-5, 2018

On the zero-error capacity of channels with rate limited noiseless 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*—While it is known that feedback does not increase the small-error capacity of a discrete memoryless channel, noiseless feedback can increase the zero-error capacity from zero (without feedback) all the way to the small-error capacity. This result depends on the availability of *noiseless output feedback*, which gives the transmitter access to the exact output seen at the destination, as well as the use of variable-length codes. In this work, we consider two more realistic setups: 1) a noiseless feedback link of finite rate (which may not permit transmission of the outputs in their entirety), and 2) a noisy feedback link. We derive rates which may be achieved with zero error. Our results show that the achievable zero-error rate can vary between the zero-undetected-error capacity and the small-error capacity depending on the available feedback link rate and quality.**

I. Introduction

Shannon showed that for fixed block-length coding schemes, noiseless feedback does not increase either the small-error or zero-error capacity. However, for variable-length coding schemes, noiseless feedback can increase the zero-error capacity [1], [2]. As shown by Burnashev [2], it is possible to communicate with zero-error at rates equal to the Shannon small-error capacity over a discrete memoryless channel (DMC) with noiseless feedback if, and only if, there exists at least one channel output (a "disprover") that is reachable from some but not all the channel inputs, *if* one allows for variable-length codes.

The variable-length coding scheme used repeatedly sends a message until the transmitter determines, through perfect output feedback, and subsequently informs the destination that the message has been correctly received at the destination. This variable-length coding scheme relies heavily on the transmitter being able to see exactly what the receiver does. This is for example possible if the zero-error capacity of the noiseless feedback channel is at least $\log_2|\mathcal{Y}|$, where $\mathcal{Y}$ is the set of outputs of the forward DMC. In this case, the outputs can be sent back to the transmitter perfectly. Alternatively, the destination could output its decision, in which case the zero-error feedback link rate would need to be at least the forward link's small-error capacity. In this paper, we present a lower bound on the zero-error capacity when the feedback is noiseless but rate-limited. This is particularly interesting when the feedback rate is below the small-error capacity of the forward channel.

Variable-length noiseless feedback communication schemes have been considered in [2], [3], [4]. Although not targeting zero-error explicitly, these schemes can guarantee zero-error under certain conditions. In our previous work [5], we focused on the zero-error capacity of a DMC when the channel feedback is noisy. We showed that the variable-length zero-error capacity with noisy feedback is lower bounded by the forward channel's zero-undetected-error capacity , and showed that under certain conditions this is tight. We also outlined conditions under which the zero-error capacity without feedback, with perfect feedback, and with noisy feedback, are positive.

In this paper we first carefully define variable-length zero-error communication with perfect, but rate limited feedback. Next, we compare known variable-length noiseless feedback communication schemes in terms of the required feedback rate and their overall achievable rate. We then propose a communication scheme achieving zero-error even for low rate noiseless feedback, as well as noisy feedback links with positive zero-error capacity.

II. Definitions

Let $x_i^j := (x_i, x_{i+1}, \dots, x_j)$ when $i \le j$ and $|x_i^j| = j-i+1$ denote its size. Let $\gamma_n = o(n)$, and $\gamma_n \to \infty$ as $n \to \infty$ (e.g. $\gamma_n = \log(n)$). 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 $W(y|x)$. Let $r_f \in \mathbb{R}^+$ denote the available error-free rate of the link between the destination / decoder and the transmitter / encoder (Figure 1).

Small error fixed-length capacity C without feedback. A $\mathcal{C}(\mathcal{M}, n)$ fixed-length code for DMC $(\mathcal{X}, \mathcal{Y}, W)$ 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

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

The maximum and average, respectively, probabilities of error for a $\mathcal{C}(\mathcal{M}, n)$ are defined as

$$\lambda^{(n)} = \max_{m \in \mathcal{M}} \lambda_m^{(n)}, \qquad P_e^{(n)} = \frac{1}{|\mathcal{M}|} \sum_{m \in \mathcal{M}} \lambda_m^{(n)}.$$

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

Zero-undetected-error fixed-length capacity C_{0u} [6]. A zero-undetected-error code of block-length n, denoted by $\mathcal{C}_{0u}(\mathcal{M}, n)$ 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 $M(y^n)$ denote the set of possible messages corresponding to a received output y^n

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

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

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

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 $\max_{m \in \mathcal{M}} \lambda_m^{(n)}$ 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 zero-error regime for rate limited noiseless feedback.

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

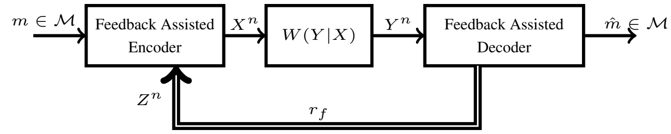

Fig. 1. A DMC $(\mathcal{X}, \mathcal{Y}, W)$ with rate limited perfect feedback of rate r_f.

A variable-length zero-error rate-limited feedback code $\mathcal{C}_0^{VL-PF}(\mathcal{M}, n, r_f)$ 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 $\mathcal{F}_i : \mathcal{M} \times \mathcal{Z}^{i-1} \to \mathcal{X}$, where Z^i is the received sequence through rate limited feedback which generate the inputs

$$X_i = \mathcal{F}_i(M, Z^{i-1}), \quad 1 \leq i \leq l;$$

3) a sequence of encoding function $\mathcal{H}_i : \mathcal{Y}^{i-1} \to \mathcal{Z}$, which generate the feedback

$$Z_i = \mathcal{H}_i(Y^{i-1}), \quad |Z^n| \leq 2^{nr_f} \quad 1 \leq i \leq n;$$

for n to be defined in 5).

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

5) a non-negative integer-valued stopping time N (random variable) defined as the first k that the decoder does not declare an erasure, i.e.

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

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

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

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

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

III. Known variable-length zero-error perfect feedback communication schemes

In this section, we compare three zero-error achievability schemes in terms of their required error-free feedback rate and corresponding (zero-error) achievable rates. These will be combined in the next section in a hybrid communication scheme when the perfect feedback rate is limited to r_f.

A. Burnashev's scheme (observation feedback)

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

$$E_{burn}(\bar{R}) = \frac{C_1}{C}(C - \bar{R}), \quad 0 \leq \bar{R} \leq C \tag{2}$$

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, $C_1 = \infty$, and zero-error communication is possible. In this case, it can be shown that the variable-length zero-error capacity coincides with the small error fixed-length capacity C without feedback, i.e. $\mathcal{C}_0^{VL-PF}(r_f) = C$.

To achieve this, take any capacity achieving code $\mathcal{C}(\mathcal{M}, n)$ for the DMC W whose maximal probability of error $\lambda^{(n)}$ 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 if $r_f \geq \log_2 |\mathcal{Y}|$. The transmitter can thus determine whether the receiver obtained the correct message. The transmitter informs the receiver of whether the decoded message is correct or whether the receiver should expect a repetition of the same message in the next block-length using the disprover (ensuring zero-error). Figure 2 shows an example of this scheme when three repetitions are needed to send the message with zero error. The converse follows as the zero-error capacity is always upper bounded by the small-error capacity.

Remark 1: The rate of the feedback link needed for this (unaltered) scheme is at least $r_f \geq \log_2 |\mathcal{Y}|$, as the entire received sequence is needed at the encoder.

B. Yammamo-Itoh's scheme (tentative decision feedback)

Rather than forwarding the entire received sequence, the receiver sends back its estimated or decoded message. Yamamoto and Itoh demonstrated that the same reliability function as in (2) is achievable with this reduced feedback rate of $r_f \geq R$ [4]. Thus, zero-error is guaranteed as long as the forward channel contains a disprover, and $r_f \geq R$, using the Yamamoto-Itoh scheme.

C. Forney's scheme (erasure feedback)

Forney's scheme reduces the amount of feedback further, to the extreme case where only one bit of perfect feedback is needed per information block (hence feedback of zero rate). This appears to come at the expense of the overall zero-error rate achieved, which drops from the small error capacity to the zero-undetected-error capacity. This is done by sending an "erasure" indicator on the feedback link. That is, this scheme achieves C_{0u} by taking a zero-undetected-error capacity achieving code $C_{0u}(\mathcal{M}, n)$ for channel W whose maximal erasure probability tends to zero and whose rate approaches C_{0u}. To transmit message $m \in \mathcal{M}$, codeword $c_{0u}^{(n)}(m)$ is sent through W. Upon receiving $y^n \in \mathcal{Y}^n$, 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 ($|M(y^n)| > 1$, see (1)). The receiver informs the transmitter of an erasure by sending one bit through the perfect feedback channel. If the transmitter does not see the confirmation of successful decoding, it again transmits $c_{0u}^{(n)}(m)$. Figure 3 shows an example of Forney's scheme for a channel with perfect feedback in which the message is transmitted in 2 iterations. Table I compares all the variable-length described schemes when the perfect feedback is available.

TABLE I
COMPARISON OF ZERO-ERROR VARIABLE-LENGTH SCHEMES FOR CHANNEL WITH PERFECT FEEDBACK

	Burnashev	**Yamamoto-Itoh**	**Forney**
$\mathcal{C}_\mathbf{0}^{\mathbf{VL-PF}}(\mathbf{r_f})$	$C \times 1(C_{0u} \neq 0)$	$C \times 1(C_{0u} \neq 0)$	C_{0u}
$\mathbf{r_f}$	$\geq \log \lvert\mathcal{Y}\rvert$	$\geq C$	only one bit ($r_f = 0$)
Feedback type	passive	active	active
Final decision	encoder	encoder-decoder	decoder

A similar transmission scheme can achieve $C_{0u} > 0$ with zero error *even in the noisy feedback regime* [5] as follows. Since $C_{0u} > 0$ implies the existence of at least one disprover, this disprover can be used to synchronize the receiver and transmitter during the message transmission by assigning 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. A message is re-transmitted until the codeword is not erased and 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).

IV. Hybrid scheme for $r_f < C$

As seen in Table I, differing amounts of feedback may lead to different zero-error achievable rates. In particular, to achieve the small error capacity C, using Yamamoto-Itoh's scheme one needs $r_f \geq C$ to guarantee zero-error communication. Forney's scheme only needs one bit of feedback for any block-length n to guarantee zero-error communication with $\mathcal{C}_0^{VL-PF}(r_f) = C_{0u}$. If $C_{0u} = C$, then Forney's scheme is clearly more efficient than the other two schemes in terms of feedback usage. Note that in general $C_{0u} \leq C$. Here we are interested in the case that $C_{0u} < C$. Our main result is the following:

Theorem 1: Let $\beta := \frac{r_f}{C} < 1$. If $C_{0u} > 0$, then

$$\mathcal{C}_0^{VL-PF}(r_f) \geq \beta C + (1 - \beta) C_{0u}. \tag{3}$$

Fig. 2. Variable-length communication scheme with perfect feedback ($r_f \geq \log|\mathcal{Y}|$), the shaded intervals are used by the transmitter to notify the receiver about the decoding result using disprovers ($\gamma_n = o(n)$). The transmission is completed without error in 3 stages in this example.

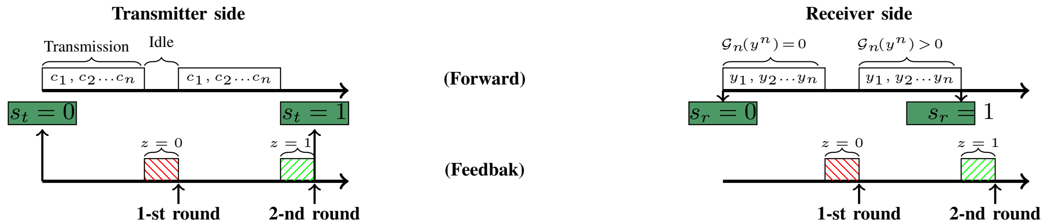

Fig. 3. Forney's communication scheme which provide zero-error with perfect synchronization.

Proof To show the achievability of the rate in (3), we combine Forney's and Yamamoto-Itoh's schemes as follows:

Let $I_n = \{1, 2, \cdots, \lceil \log n \rceil\}$ be the index set (hence $|I_n| = \lceil \log n \rceil$). Let $L_n = \beta \cdot \frac{n}{\lceil \log n \rceil}$ where $\beta \leq 1$ is a constant. Consider a set of capacity achieving codes $\mathcal{C}_1 = \{\mathcal{C}_{1i}(\mathcal{M}_{1i}, L_n), i \in I_n\}$ for the DMC W whose maximal probability of error $\lambda^{(L_n)}$ tends to zero and whose rate approaches the Shannon capacity C as block-length $L_n \to \infty$. Let $n_1 = L_n \times |I_n| = \beta \cdot n$ and $n_2 = n - n_1 = (1-\beta) \cdot n$. Also, consider a zero-undetected-error capacity achieving code $\mathcal{C}_2 = C_{0u}(\mathcal{M}_2, n_2)$ for channel W whose maximal erasure probability tends to zero and whose rate approaches C_{0u}. Then, given the message set $\mathcal{M} = (\prod_{i \in I_n} \mathcal{M}_{1i}) \times \mathcal{M}_2$ let a code $\mathcal{C}'(\mathcal{M}, n)$ be the concatenation of all $\mathcal{C}_{1i}$'s and $\mathcal{C}_2$. Let $c_1^{(L_n)}(m_{1i})$ and $c_2^{(n_2)}(m_2)$ be the corresponding codewords for messages $m_{1i} \in \mathcal{M}_{1i}, i \in I_n$ and $m_2 \in \mathcal{M}_2$. Figure 4 shows our hybrid scheme which consists of the following transmission and confirmation phases. Note that the hybrid scheme repeats the transmission and confirmation phases until all the messages $m_{1i}, i \in I_n$ and m_2 are received without error (e.g. 4 times repetition in Figure. 4).

Transmission Phase, (green and blue in Figure 4): The transmitter sends code $C'^{(n)}(m)$ which is a concatenation of $|I_n| = \log n$ codewords $C_1^{(L_n)}(m_{1i}), i \in I_n$ and one zero-undetected-error codeword $c_2^{(n_2)}(m_2)$. Upon receiving the outputs corresponding to the i-th codeword $c_{1i}^{(L_n)}(m_{1i})$, i.e. $y_{(i-1)\cdot L_n+1}^{i \cdot L_n}$, the receiver decodes $\hat{m}_{1i}$ and sends a lossless, one-to-one function of its decoded message $\mathcal{H}(\hat{m_{1i}})$ through the rate-limited noiseless feedback link in $\frac{L_n \cdot C}{r_f}$ channel uses (the transmitter must be able to retrieve $\hat{m_{1i}}$ from $\mathcal{H}(\hat{m_{1i}})$ without error for any $i \in I_n$). Note that the receiver does not use the first L_n channel uses of feedback link. After receiving $y_{n_1+1}^n$, the receiver uses a list (or erasure) decoder to check whether it is able to decode the second message $\hat{m}_2$ or whether it declares erasure, and sends, using one bit of noiseless feedback, whether the decoding of m_2 resulted in an erasure or not.

Confirmation phase (red in Figure 4): In this case, the transmitter notifies the receiver about the decoding results of all messages $m_{1i}, i \in I_n$ (it has estimates of these from the feedback in the transmission phase). Similar to Burnashev's scheme, the transmitter sends a repetition code of length γ_n for each message $m_{1i}, i \in I_n$. In this phase, the receiver, after receiving each sequence of length γ_n for each of the I_n messages, notifies the transmitter using one bit per message about whether it received a disprover or not. This allows the transmitter and receiver to remain synchronized as follows: if the transmitter receives a bit for for i-th message $m_{1i}, i \in I_n$ that a disprover was seen, then the transmitter and receiver are synchronized to send a new message in the next interval. If the transmitter receives a bit indicating that a disprover was not received, then in the next stage of transmission, the transmitter and receiver are synchronized to repeat the previous message (again send codeword $c_{1i}^{(L_n)}(m_{1i})$). Regarding the message m_2, when the transmitter receives the 1-bit confirmation, showing that the erasure did not happen, the transmitter sends a new message m_2' at the next stage; otherwise, it sends the same code $c_2^{(n_2)}(m_2)$. Thus, using the above hybrid scheme, the sender and receiver can be synchronized using $|I_n| \times \gamma_n$ channel uses in each stage (combination of one transmission and one confirmation phase).

The main reason the hybrid scheme interleaves $|I_n|$ message sets $\mathcal{M}_{1i}, i \in I_n$ is to improve the efficiency of transmission by reducing the amount of time the feedback link is idle. If $|I_n| = 1$, the length of idle feedback link is n_1. Using the hybrid scheme, the idle length become $L_n = o(n)$.

For ease of analysis, we assume that all messages are re-transmitted as long as one of the messages has not been correctly received (i.e. the receiver did not see a disprover for that message). To analyze the achievable rate using the hybrid scheme, note that the number of channel uses to transmit $\mathcal{H}(\hat{m}_{1i})$ for each $i \in I_n$ is $\frac{L_n \cdot C}{r_f}$. Thus, as shown in Figure 4, in order to efficiently transmit the messages the total number of feedback channel uses during the transmission phases should

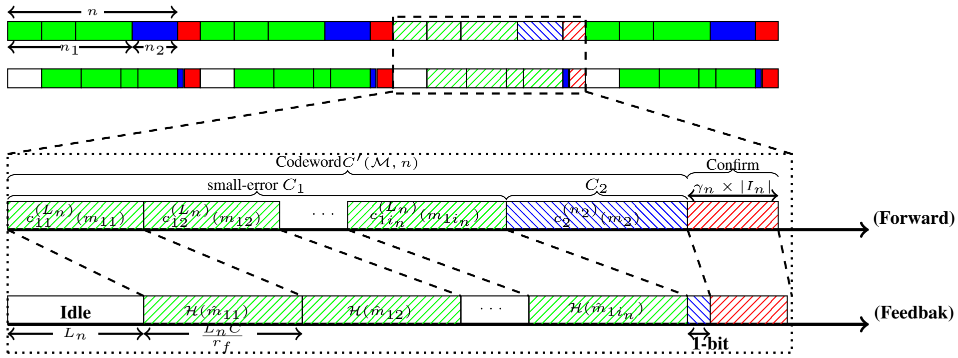

Fig. 4. In the hybrid scheme, the communication is repeated until all the messages are decoded with zero-error. The feedback is mainly used to transmit the tentative message of $\hat{m}_{1i}, i \in I_n$. Message m_2 needs only one-bit of feedback per stage. A repetition code of length γ_n is used for each message $m_{1i}, i \in I_n$.

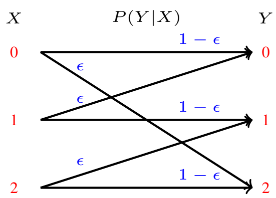

Fig. 5. The DMC whose directed channel graph is the cyclic triangle.

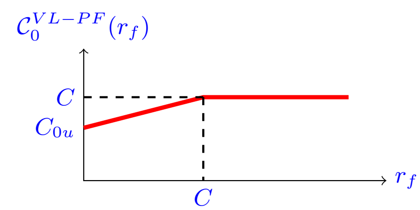

Fig. 6. An achievable rate for the triangle channel in Example 1 as a function of the error-free feedback link rate r_f.

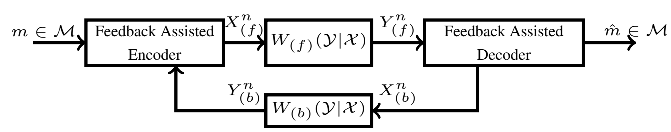

Fig. 7. Communication Scheme for a DMC with active noisy feedback

hold the following relationship:

$$n \geq |I_n| \frac{L_n \cdot C}{r_f} + L_n \approx |I_n| \frac{L_n \cdot C}{r_f},$$

where the L_n on right side of summation is for the idle period. Thus,

$$|\mathcal{M}| = (2^{L_n \cdot C})^{|I_n|} \times 2^{n_2 \cdot C_{0u}}.$$

For sake of simplicity, let $\lambda^{(n)}$ be the maximal probability of error for code $\mathcal{C}'(\mathcal{M}, n)$ ($\lambda^{(n)} \to 0$ as $n \to \infty$). Let N be the stopping time at which the receiver is able to decode all the messages correctly. Similar to the analysis of Burnashev's scheme, the number of repetitions of the two phases follows the geometric distribution with mean $\frac{1}{1-\lambda^{(n)}} \approx 1$. Since the probability of error of all codes can be made sufficiently small as $n \to \infty$ by definition of both capacity achieving codes, we may approximate $E[N] \approx n + |I_n| \times \gamma_n$. Thus,

$$\bar{R} = \lim_{n \to \infty} \frac{\log_2(2^{n_1 \cdot C}) + \log_2(2^{n_2 \cdot C_{0u}})}{E[N]} = \beta C + (1-\beta) C_{0u}.$$

■

Example 1: Figure 5 shows a DMC with a triangular directed channel graph. In [7] it was shown that when ϵ is sufficiently small, $C_{0u} = \log 2$ while $C = \log 3$. Figure 6 shows the achievable zero-error rate using our proposed scheme.

A. Extensions to noisy feedback

In our previous work [5], we presented a lower-bound for the zero-error capacity of a DMC with noisy feedback, shown in Figure 7. We showed that in the presence of noisy feedback with disprovers on both the forward ($W^{(f)}$) and backward ($W^{(b)}$) channels, we can achieve the zero-undetected-error capacity of the forward channel $C_{0u}^{(f)}$ with zero-error, even if the fixed-length zero-error capacity for both the forward channel ($C_0^{(f)}$) and backward channel ($C_0^{(b)}$) are zero. If the zero-error capacity of the backward channel is positive ($C_0^{(b)} > 0$), then using Theorem 3, we can extend our previous result for noisy feedback [5] as follows:

Proposition 1: Let $\beta = \frac{C_0^{(b)}}{C^{(f)}}$. The variable-length zero-error capacity of a DMC $W^{(f)}$ with noisy feedback $W^{(b)}$, denoted by $\mathcal{C}_0^{VL-NF}$ [5], satisfies

$$\mathcal{C}_0^{VL-NF} \geq \begin{cases} C_{0u}^{(f)} & \text{if } C_{0u}^{(b)} > 0 \text{ and } C_0^{(b)} = 0 \\ \beta C + (1-\beta) C_{0u}^{(f)} & \text{if } 0 < C_0^{(b)} < C^{(f)} \\ C & \text{if } C_0^{(b)} \geq C^{(f)} \end{cases}, \tag{4}$$

where $C_{0u}^{(f)}$ and $C_{0u}^{(b)}$ denote the zero-undetected-error capacities of the forward and backward links.

Proof In [5] we proposed a scheme that achieves $\mathcal{C}_0^{VL-NF} \geq C_{0u}^{(f)}$ if $C_{0u}^{(b)} > 0$ and $C_0^{(b)} = 0$. Also, if $C_0^{(b)} \geq C^{(f)}$, then Yamamoto-Itoh's scheme can achieve

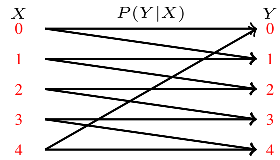

Fig. 8. Pentagon channel with crossover probability $\epsilon > 0$.

$\mathcal{C}_0^{VL-NF} \geq C$. The proof is straightforward when $0 < C_0^{(b)} < C^{(f)}$ using theorem 1.

Example 2: Here we consider a DMC with noisy feedback in which both forward channel $W^{(f)}$ and feedback channel $W^{(b)}$ are pentagon channels as shown in Figure 8. For a pentagon channel with small but positive ϵ, it is known that the small-error capacity $C = \log_2 5$ and the zero-error capacity $C_0 = \frac{1}{2}\log_2 5$ [8]. In the Appendix we show that for the pentagon channel $C_{0u} \geq 2$. Thus, using Proposition 1 and given that $\beta = \frac{1}{2}$ the variable-length zero-error capacity $\mathcal{C}_0^{VL-NF} \geq 2.161$.

V. Conclusion

While it is known that feedback does not increase the small-error capacity, it can increase the zero-error capacity of channels. In the extreme case, availability of perfect noiseless feedback can increase the zero-error capacity of a DMC from zero (without feedback) all the way to the small-error capacity. Here, we provide achievable rates when the amount of noiseless feedback is limited or only noisy feedback is available. Our results show that the achievable zero-error rate can vary between the zero-undetectable capacity and the small-error capacity depending on the rate of available feedback.

Appendix

To the best of our knowledge the zero-undetected-error capacity of the pentagon channel W is not known. Note that Pinsker and Sheverdyaev [9] proved that C_{0u} equals C if the bipartite channel graph is acyclic.[1] However, for the pentagon channel the bipartite graph is cyclic. A lower bound is obtained by using random coding over codes with constant composition [10]

$$C_{0u} \geq \max_{Q} \min_{\substack{V \ll W \\ QV = QW}} I(Q, V), \tag{5}$$

where the maximization is over all distributions Q on 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 ($V \ll W$). Consider the following auxiliary channel V

$$V = \begin{bmatrix} 1-\epsilon & \epsilon & 0 & 0 & 0 \\ 0 & 1-\epsilon & \epsilon & 0 & 0 \\ 0 & 0 & 1-\epsilon & \epsilon & 0 \\ 0 & 0 & 0 & 1-\epsilon & \epsilon \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Note that $V \ll W$ and for any distribution $Q = [q_0, q_1, q_2, q_3, 0]$ it is easy to verify that $QV = QW$. Let $R_{0u}(W)|_Q$ be the zero-undetected error achievable rate for the pentagon channel restricted to the set of all input distributions Q with $(q_4 = 0)$. Note that

$$C_{0u}(W) \geq \max_{Q} R_{0u}(W)|_Q.$$

Given Q, it is easy to verify that $R_{0u}(W)|_Q = R_{0u}(V)|_Q$. Given Q, the bipartite graph for channel V becomes acyclic and thus

$$C(V)|_Q = \max_{Q} R_{0u}(V)|_Q = \log_2 4.$$

[1]The bipartite graph of a channel is acyclic if there does not exist $l > 2$, distinct inputs $x_1, \cdots, x_l$ and distinct outputs $y_1, \cdots, y_l$ such that $W(y_j|x_j) > 0, W(y_j|x_{j+1}) > 0$ for $j = 1, \cdots, l$ and $x_{l+1} = x_1$ [7].

Acknowledgment

The work of the authors was partially supported by NSF under award 1645381. The contents of this article are solely the responsibility of the authors and do not necessarily represent the official views of the NSF.

References

[1] C. E. Shannon, "The zero error capacity of a noisy channel," *IEEE Transactions on Information 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." *Problemy Peredachi Informatsii*, vol. 12, no. 4, pp. 10–30, 1976.

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

[4] H. Yamamoto and K. Itoh, "Asymptotic performance of a modified Schalkwijk-Barron scheme for channels with noiseless feedback," *IEEE Trans. Inform. Theory*, vol. IT-25, pp. 729–733, Nov. 1979.

[5] M. Asadi and N. Devroye, "On the zero-error capacity of channels with noisy feedback," Monticello, IL, Oct. 2017, pp. 642–649.

[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] C. Bunte, A. Lapidoth, and A. Samorodnitsky, "The zero-undetected-error capacity of the low-noise cyclic triangle channel," Istanbul, Turkey, 2013, pp. 91–95.

[8] L. Lovasz, "On the Shannon Capacity of a Graph," *IEEE Transactions on Information Theory*, vol. 25, no. 1, pp. 1–7, 1979.

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

[10] I. Telatar and R. Gallager, "New exponential upper bounds to error and erasure probabilities," *IEEE Symposium of Information Theory*, p. 379, 1994.