Bounds for the capacity error function for unidirectional channels with noiseless feedback

Christian Deppe

Institute for Communications Engineering Technical University of Munich D-80333 Munich, Germany Email: christian.deppe@tum.de

Vladimir Lebedev

Kharkevich Institute for Information Transmission Problems Russian Academy of Sciences 127051 Moscow, Russia Email: lebedev37@mail.ru

Georg Maringer

Institute for Communications Engineering Technical University of Munich D-80333 Munich, Germany Email: georg.maringer@tum.de

Abstract

In digital systems such as fiber optical communications, the ratio between probability of errors of type $1 \rightarrow 0$ and $0 \rightarrow 1$ can be large. Practically, one can assume that only one type of error can occur. These errors are called asymmetric. Unidirectional errors differ from asymmetric type of errors; here both $1 \rightarrow 0$ and $0 \rightarrow 1$ type of errors are possible, but in any submitted codeword all the errors are of the same type. This can be generalized for the q-ary case.

We consider q-ary unidirectional channels with feedback and give bounds for the capacity error function. It turns out that the bounds depend on the parity of the alphabet q. Furthermore, we show that for feedback, the capacity error function for the binary asymmetric channel is different from the symmetric channel. This is in contrast to the behavior of the function without feedback. *Keywords:* error-correcting codes, uni-directional errors, feedback

1. Introduction

Shannon analyzed the capacity of a channel based on probabilistic notions. The arguably easiest example is a binary symmetric channel (BSC) with crossover probability ε . In a probabilistic model the number of errors in a block of symbols can be (having different probability) any number between 0 and the blocklength n. In this paper we analyze the situation when the fraction of erroneous symbols is upper bounded by $\tau = \frac{t}{n}$, where t denotes the maximal number of errors per block. Our focus in this work lies in the analysis of the capacity error function of unidirectional channels.

The analysis in this paper is purely combinatorial. That means that we are not distinguishing between likely and unlikely errors. We only consider whether a certain output may occur given that a certain symbol was chosen at the input of the channel.

The difference between symmetrical, asymmetrical and unidirectional errors is described in [12]. A unidirectional channel is composed of two special asymmetric channels. During the transmission of each codeword, one of the two channels is randomly selected. Within the transmission of those codeword symbols, the channel behaves like the selected channel and remains unchanged. After each transmitted codeword, the channel is reselected. Precise definitions are given in Section 2.

Asymmetrical channels model photon communication systems better than symmetrical channels. Due to losses within the transmission medium, photons may not reach the receiver. However, the channel cannot generate photons on its own. Therefore the channel is of asymmetric nature [12, 1.3]. For a transmission without feedback, the binary Z-channel (Figure 3) has the same capacity error function as the symmetric channel (Figure 2). This has already been shown by Bassalygo in [3]. Asymmetric and unidirectional channels without feedback are analyzed in numerous publications. The major results and references can be found in [10] and [12]. In this paper we give lower and upper bounds on the capacity error function of asymmetric and unidirectional channels in the presence of noiseless feedback. For a wide range of τ , we show that the capacity error function for the unidirectional channel composed out of a Z-channel and an inverse Z-channel is larger compared to the one of the symmetric channel.

2. Q-ary codes with feedback

In our transmission scheme, a sender wants to transmit a finite number of messages over a channel with noiseless feedback (see Figure 1). The discrete channel has a q-ary alphabet $\mathcal{X} = \mathcal{Q} = \{0, \ldots, q-1\}$ at its input and at its output, $\mathcal{Y} = \mathcal{Q}$. Each message m is encoded into a block of length n. For channels without feedback, the codeword c to be transmitted over the channel only depends on the message m. In a feedback channel the situation is different.



Figure 1: Channel with feedback

Definition 1. 1. Let the set of possible messages be denoted as $\mathcal{M} = \{1, \ldots, M\}$. Then an encoding algorithm for a feedback channel with blocklength n is a set of functions

$$c_i: \mathcal{M} \times \mathcal{Q}^{i-1} \to \mathcal{Q}, \quad i \in \{1, \dots, n\}$$
 (1)

The respective coding algorithm is then constructed as

$$c(m, y^{n-1}) = ((c_1(m), c_2(m, y_1), \dots, c_n(m, y^{n-1}))) , \qquad (2)$$

where $y^k = (y_1, ..., y_k)$.

 An encoding algorithm is (n, M, t)_f-successful if the corresponding decoder decodes correctly for any transmitted message m and any error pattern with less than or equal t errors caused by the channel.

Definition 1 shows that the encoder may adjust its algorithm for sending the kth symbol c_k according to the previously received symbols y^{k-1} . This extra flexibility can be used to increase the achievable rate of the system.

Definition 2. Let $2 \leq q \in \mathbb{N}$. The capacity error function $C_q^f(\Gamma, \tau)$ of a channel Γ with noiseless feedback denotes the supremum on the rates for which a successful algorithm exists depending on $\tau = \frac{t}{n}$ as the blocklength n goes to infinity. t denotes the maximum number of errors inflicted by the channel noise and q specifies the alphabet size of the channel.

The capacity error function for $\tau = 1$ is denoted as $C_{q,0}^f(\Gamma)$ and was introduced by Shannon in [11] under the term zero error capacity of the channel.

If the sender and the receiver share a channel with noiseless passive feedback (Figure 1), the capacity error function of the symmetric channel is only completely known for the binary case.

The most frequently analyzed communication channels are of symmetric nature. As an example, we give the previously mentioned symmetric channel Γ^2 (Figure 2). In this case it is possible to receive a 1 when a 0 was transmitted and a 0 when a 1 was transmitted. Both events denote an error, because the correct reception of a symbol is considered to be the event that the sent symbol matches the received symbol. In the binary case, the capacity error function of the symmetric channel was found by Berlekamp [4] and Zigangirov [13]

$$C_2^f(\Gamma^2, \tau) = \begin{cases} 1 - h(\tau) & \text{if } 0 \le \tau \le (3 + \sqrt{5})^{-1} \\ (1 - 3\tau) \log\left(\frac{1 + \sqrt{5}}{2}\right) & \text{if } (3 + \sqrt{5})^{-1} \le \tau \le \frac{1}{3} \\ 0 & \text{if } \tau > \frac{1}{3}. \end{cases}$$
(3)

In general, discrete channels can be specified by bipartite graphs.

Definition 3 (Discrete channel/bipartite graph). A discrete channel corresponds to a bipartite graph in the following way. Let \mathcal{V}_{in} denote the set of possible input symbols and let \mathcal{V}_{out} denote the set of possible output symbols. Then if it is possible that the channel maps input symbol $i \in \mathcal{V}_{in}$ to output symbol $j \in \mathcal{V}_{out}$, the pair (i, j) is part of the set of edges $\mathcal{E} \subset \mathcal{V}_{in} \times \mathcal{V}_{out}$. Conversely, a bipartite graph between input symbols \mathcal{V}_{in} and \mathcal{V}_{out} defines a discrete channel. We denote by (x_1, x_2, \ldots, x_n) the sequence of input symbols and by (y_1, y_2, \ldots, y_n) the sequence of output symbols, where x_i depends on the message and (y_1, \ldots, y^{i-1}) (see Definition 1).

We define (e_1, e_2, \ldots, e_n) as the corresponding error vector, where $e_i = y_i - x_i$.

Discrete channels with finite input and output alphabets correspond to their respective bipartite graphs by a one to one mapping. Therefore, in this paper we frequently speak about the graphs when we mean the respective channels and vice versa.

Definition 4 (Asymmetric channel). An asymmetric channel is a discrete channel whose specifying bipartite graph is not the full graph in the sense that the set of edges $\mathcal{E} \neq \mathcal{V}_{in} \times \mathcal{V}_{out}$.

As an example for an asymmetric channel we propose the Z-channel which is specified by the bipartite graph on the left hand side of Figure 3.

Definition 5 (Unidirectional channel). A unidirectional channel is specified by two asymmetric channels having the same set of input symbols \mathcal{V}_{in} and output symbols \mathcal{V}_{out} . One of the channels allows only positive error vectors $(e_i \geq 0 \ \forall i)$, whereas the other one only allows negative error vectors $(e_i \leq 0 \ \forall i)$. Within each transmission of a codeword, the channel is specified by one of the asymmetric channel. Sender and receiver know both channels, but they do not know to which asymmetric channel the unidirectional channel corresponds. This may change for each codeword.



Figure 3: binary Z-channel and inverse Z-channel

In Figure 3 the binary unidirectional channel is shown. It is denoted as Γ_U^2 and is composed of the binary Z-channel and its counterpart, the inverse Z-channel.

Definition 6 (Generalized Z/inverse Z-channel). The generalized Z-channel Γ_Z^q is specified by the bipartite graph with input nodes $\mathcal{V}_{in} = \mathcal{V}_{out} = \{0, \dots, q-1\}$ and the set of edges $\mathcal{E}_Z^q = \{(i, i-1) : i \in \{1, \dots, q-1\}\} \cup \{(i, i) : i \in \{0, \dots, q-1\}\}.$

The inverse Z-channel Γ_{Σ}^{q} is specified by the bipartite graph with input nodes $\mathcal{V}_{in} = \mathcal{V}_{out} = \{0, \dots, q-1\}$ and the set of edges $\mathcal{E}_{\Sigma}^{q} = \{(i, i+1) : i \in \{0, \dots, q-2\} \cup \{(i, i) : i \in \{0, \dots, q-1\}\}.$

Remark 1. The capacity error function of the generalized Z-channel Γ_Z^q is equal to the capacity error function of the inverse generalized Z-channel $\Gamma_{\overline{\chi}}^q$.

Proof: The inverse Z-channel $\Gamma_{\mathbf{X}}^q$ can be obtained from the Z-channel $\Gamma_{\mathbf{Z}}^q$ by a bijective mapping. The labelling of the nodes at the input and the output is simply in reversed order.

Both channels are depicted in Figure 4. Combined they form a unidirectional channel which we denote as Γ_{U}^{q} .

The symbols 0 and q-1 are of special interest for the generalized Z-channel and the inverse generalized Z-channel. For the generalized Z-channel, the symbol



Figure 4: Generalized Z-channel and generalized inverse Z-channel

0 has the property that the sender knows that this symbol cannot be changed by the channel. If the symbol q - 1 is received, the receiver knows that the transmitter indeed sent this symbol. The properties of 0 and q - 1 are swapped for the inverse generalized Z-channel compared to the generalized Z-channel. All other symbols do not have these special properties.

3. Zero-error capacity of the generalized Z-channel

The method we use to obtain the zero error capacity of the generalized Zchannel was already introduced by Shannon in [11, Theorem 7].

The idea is to split the input symbols into sets according to their connected outputs.

Theorem 1 (Shannon [11]). Let the set S_j denote the set of input symbols which are connected to the channel output $j \in \{0, ..., q-1\}$ and let P_i denote the probability of symbol *i* at the input of the channel. Let P_o be defined as

$$P_o := \min_{P_i} \max_j \sum_{i \in \mathcal{S}_j} P_i \tag{4}$$

over all possible input distributions P_i .

The zero error capacity $C_{0,q}^f$ is then obtained by

$$C_{0,q}^f = \log_q \left(\frac{1}{P_o}\right) \quad . \tag{5}$$

Theorem 1 suggests that in order to find the zero error capacity, it is possible to find the input distribution which minimizes equation (4) and to take the logarithm of the reciprocal of the result.

Property 1. The zero error capacity of the generalized Z-channel and the inverse generalized Z-channel is

$$C_{0,q}^f(\Gamma_Z^q) = C_{0,q}^f(\Gamma_{\underline{\lambda}}^q) = \log_q\left(\left\lceil \frac{q}{2} \right\rceil\right) \,. \tag{6}$$

Proof: We apply Theorem 1 to the inverse generalized Z-channel $\Gamma_{\mathbf{X}}^{q}$ shown in Figure 4.

$$S_{0} = \{0\}$$

$$S_{1} = \{0, 1\}$$

$$S_{2} = \{1, 2\}$$

$$\vdots$$

$$S_{q-1} = \{q - 2, q - 1\}$$

Consider now an arbitrary distribution on the set of input letters $\{0, \ldots, q-1\}$. It is obvious that

$$\sum_{i \in \mathcal{S}_0} P_i \le \sum_{i \in \mathcal{S}_1} P_i \tag{7}$$

with equality if and only if $P_1 = 0$. Assuming that $P_1 > 0$, we get an input probability which is at least as good as before according to equation (4) if we change the input probability to

$$P'_0 = P_0 + P_1$$

 $P'_1 = 0$
 $P'_2 = P_2$
 \vdots
 $P'_{q-1} = P_{q-1}$.

This is because $\sum_{i \in S_2} P_i = P_1 + P_2$ can only get smaller, while for $j \in \{3, \ldots, q-1\}, \sum_{i \in S_j} P_i$ remains unchanged. Now we consider the distribution P'. As $P'_1 = 0$, the set S_2 effectively only contains the element 2. As $S_3 = \{2, 3\}$ we can use the same method as before to show that the zero error capacity cannot get smaller if we change the distribution such that $P''_2 = P'_2 + P'_3$ and $P'_3 = 0$ while all other input probabilities remain the same.

Using this procedure inductively gives us the result that it is optimal to only have non-zero probability for every second input symbol. Looking at equation (4) it is easy to see that the optimal input distribution is the uniform distribution over all input symbols k fulfilling $k \equiv 0 \mod 2$. Therefore, the zero error capacity of the inverse generalized Z-channel is

$$C_{0,q}^{f}(\Gamma_{\mathbf{\Sigma}}^{q}) = \log_{q}\left(\left\lceil \frac{q}{2} \right\rceil\right) \,. \tag{8}$$

Remark 1 implies that the zero error capacities of the generalized Z-channel and the inverse generalized Z-channel are equal.

Furthermore, Property 1 can be used to show the following.

Corollary 1. It holds $C_{0,q}^f(\Gamma_U^q) = \log_q(\lceil \frac{q}{2} \rceil)$.

Proof: Because of Property 1, it holds that $C_{0,q}^f(\Gamma_Z^q) = C_{0,q}^f(\Gamma_{\overline{\Sigma}}^q) = \log_q(\lceil \frac{q}{2} \rceil)$. Therefore, $\mathcal{C}_{0,q}^f(\Gamma_U^q) \leq \log_q(\lceil \frac{q}{2} \rceil)$. It remains to be shown that $C_{0,q}^f(\Gamma_U^q) \geq \log_q(\lceil \frac{q}{2} \rceil)$. In order to prove this, we propose the following encoding and strategies.

For this we define the set $\mathcal{S} := \{k \in \{0, \dots, q-1\} : k \equiv 0 \pmod{2}\}.$

The encoding function

$$c: \mathcal{M} \to \mathcal{S}^{n-1} \tag{9}$$

can be any bijective function between \mathcal{M} and \mathcal{S}^{n-1} . If the sender would like to send message $m \in \mathcal{M}$, he computes $x^{n-1} = c(m)$ and sends it over the channel. In order to send the last symbol, the sender distinguishes two cases:

1. The channel caused an error within at least one of the previously transmitted symbols. In this case the sender knows the channel and sends $x_n = 0$ if the active channel is Γ_Z^q and q - 1 otherwise. 2. No error has occurred within the transmission. In this case the sender sends $x_n = 0$.

We denote the vector of symbols at the channel output as y^n . In order to determine which channel was active during the transmission, the decoder checks the value of y_n . There are three cases to be distinguished:

- 1. $y_n = 0$: Either no transmission error occurred at all or the active channel was Γ_Z^q .
- 2. $y_n = q 1$: The active channel was $\Gamma_{\mathbf{x}}^q$.
- 3. $y_n = 1$: No error occurred within the first n 1 symbols and the active channel was Γ^q_{∇} .

Let \mathcal{K}_i denote the set of possible outputs for the channel if the symbol *i* is transmitted. The sets \mathcal{K}_i for different inputs are disjoint for Γ_Z^q and Γ_{Σ}^q , respectively. From the last symbol y_n , the decoder knows which channel has been active. Therefore our coding strategy is successful for a message set \mathcal{M} of size

$$|\mathcal{M}| = \left\lceil \frac{q}{2} \right\rceil^{n-1}$$

achieving asymptotically the rate

$$R := \log_q\left(\left\lceil \frac{q}{2} \right\rceil\right).$$

4. An upper bound on capacity error functions

The analysis in the previous section deals with the zero error capacity, assuming that the number of errors in a block of length n can be from 0 to n. This section deals with the problem of allowing the channel to only change a fraction of the symbols during the transmission. Let this fraction be denoted as $\tau = \frac{t}{n}$, where t denotes the number of possible errors.

An upper bound on the capacity error function of the Z-channel can be obtained by using an upper bound on the maximal set of messages, which can successfully be transmitted over the channel Γ^{q*} depicted in Figure 5. **Definition 7.** The channel Γ^{q*} is specified by the graph with nodes $\mathcal{V}_{in} = \mathcal{V}_{out} = \{*\} \cup \{0, \dots, q-1\}$ and the set of edges $\mathcal{E} = \{(*, *), (i, i) : i \in \{0, \dots, q-1\}\} \cup \{(i, i+1) : i \in \{0, \dots, q-2\}\} \cup \{(*, q-1), (0, *)\}\}.$

Let $\mathcal{M}_q^Z(n,t)$ denote a maximal set of messages which can successfully be transmitted over Γ_Z^q , let $\mathcal{M}_q^U(n,t)$ denote the maximal set of messages which can successfully be transmitted over Γ_U^q and let $\mathcal{M}_{q*}(n,t)$ denote the maximal set of messages which can successfully be transmitted over Γ^{q*} .

Theorem 2.

$$C_q^f(\Gamma_U^q,\tau) \le 1 + h_q\left(\min\left(\tau,\frac{1}{q+1}\right)\right) - \min\left(\tau,\frac{1}{q+1}\right) - h_q\left(\min\left(\tau,\frac{1}{2}\right)\right),$$

where h_q denotes the binary entropy function with logarithms to the base q.

Proof: We will show that the following inequalities hold:

$$\mathcal{M}_{q}^{U}(n,t) \stackrel{(a)}{\leq} \mathcal{M}_{q}^{Z}(n,t) \stackrel{(b)}{\leq} \mathcal{M}_{q}^{upper}(n,t) \quad , \tag{10}$$

where

$$\mathcal{M}_q^{upper}(n,t) := \frac{\sum_{i=0}^t \binom{n}{i} q^{n-i}}{\sum_{i=0}^t \binom{n}{i}} . \tag{11}$$

(a) follows because every successful algorithm for Γ_U^q is also a successful algorithm for Γ_Z^q . In order to show (b), we first note that every successful strategy on Γ_Z^q is also successful on Γ^{q*} by not using the symbol * at the input and interpreting each * at the output as 0. At the output of Γ^{q*} , there are at most t *-symbols using such an encoding strategy. In order to compute the cardinality of the set of output sequences S, they are partitioned into the sets S_0, \ldots, S_t , where S_i denotes the set of output sequences containing * exactly *i* times. The cardinality of these sets is

$$|\mathcal{S}_i| = \binom{n}{i} q^{n-i} \ . \tag{12}$$

(b) results from the fact that the S_i are disjoint and a sphere packing argument. We define

$$C_q^{upper}(\tau) := 1 + h_q \left(\min\left(\tau, \frac{1}{q+1}\right) \right) - \min\left(\tau, \frac{1}{q+1}\right) - h_q \left(\min\left(\tau, \frac{1}{2}\right) \right)$$

Next we show that $\lim_{n\to\infty} \frac{M_q^{upper}(n,t)}{n} = C_q^{upper}(\tau).$

This limit is determined by the largest terms in both sums of equation (11).

The largest term of the sum in the denominator is equal to $\binom{n}{\min(t,n/2)}$. In order to deal with the sum in the numerator, we use Stirling's approximation to estimate the binomial coefficients of

$$\lim_{n \to \infty} \frac{\log_q \left(\sum_{i=0}^t \binom{n}{i} q^{n-i} \right)}{n}$$

neglecting terms which are asymptotically irrelevant. The sum in the numerator omitting all asymptotically irrelevant terms and using Stirling's approximation is given by

$$q^{n} \sum_{i=0}^{\iota} q^{n(h_{q}(\xi_{i})-\xi_{i})} \leq q^{n} \tau n \max_{\xi_{i} \in [0,\tau]} q^{n(h_{q}(\xi_{i})-\xi_{i})} , \qquad (13)$$

where we defined $i := \xi_i n$ and $\tau := \frac{t}{n}$.

Using elementary differential calculus, one obtains that the exponent is maximized for $\xi_i = \min\left(\tau, \frac{1}{q+1}\right)$.

Therefore,

$$\lim_{n \to \infty} \frac{\log_q(\sum_{i=0}^t {n \choose i} q^{n-i})}{n} = 1 + h_q\left(\min\left(\tau, \frac{1}{q+1}\right)\right) - \min\left(\tau, \frac{1}{q+1}\right) \quad . \tag{14}$$

As n goes to infinity, the error due to Stirling's approximation vanishes and using similar arguments for the sum in the denominator, we get

$$\lim_{n \to \infty} \frac{\log_q(\mathcal{M}_q^{upper}(n,t))}{n} = 1 + h_q\left(\min\left(\tau,\frac{1}{q+1}\right)\right) - \min\left(\tau,\frac{1}{q+1}\right) - h_q\left(\min\left(\tau,\frac{1}{2}\right)\right) ,$$

completing the proof.

Remark 2. The proof of Theorem 2 even shows that

$$C_q^f(\Gamma_Z^q,\tau) \le 1 + h_q\left(\min\left(\tau,\frac{1}{q+1}\right)\right) - \min\left(\tau,\frac{1}{q+1}\right) - h_q\left(\min\left(\tau,\frac{1}{2}\right)\right) \quad .$$

Remark 3. In the theorem, we follow the idea of Hamming to get an upper bound. The new idea we use is to extend the channel without using all symbols at the input. Therefore we reduce the number of output sequences, obtaining a



Figure 5: Bipartite graph Γ^{q*}

better upper bound on the number of messages using the sphere packing argument. To the best of our knowledge, this method is new. Perhaps it can be used for other problems to get new upper bounds.

5. Lower bound on the capacity error function

In this section a method introduced in [8, Lemma 2] is recapitulated. It is then used to construct a lower bound on the capacity error function of Γ_Z^q and Γ_{Σ}^q . In the following, the set of graphs with vertices $\mathcal{V}_{in} = \mathcal{V}_{out} = \{0, \ldots, q-1\}$ all having at most degree 2 is denoted as $\tilde{\Gamma}_2(q)$.

Lemma 1 (DL A[8]). Let R > 0 be the rate of a successful algorithm for a channel $\Delta \in \tilde{\Gamma}_2(q)$ and $\tau = 1$. Then there exists a successful algorithm with rate $1 - \log_q(2)$ for all τ .

Lemma 2 (DL B[8]). If there exists a successful algorithm for a channel $\Delta \in \tilde{\Gamma}_2(q)$ with rate R > 0 for all τ , then there exists a successful algorithm with rate $1 - h(\tau) \log_q(2)$ for Δ and $0 \leq \tau \leq \frac{1}{2}$, where $h(\cdot)$ denotes the binary entropy function.

We omit the proof of Lemma 1 and give only the idea of the construction achieving the rate proposed in Lemma 2. A full proof is given in [8]. **Proof sketch:** The strategy of the proof of [8] is to give an algorithm which can transmit $\frac{q^n}{\binom{n}{t}}$ messages. This algorithm then achieves the rate claimed in the Lemma. The algorithm uses the first $k_1 = n - \log_q \binom{n}{t}$ symbols to transmit the information symbols and to use the rest of the block to correct potential errors caused by the channel. It is sufficient for the receiver to know the error positions because the nodes at the receiving side of the graph only have a degree of at most 2. Let t_1 denote the number of errors within the information symbols and $n_1 := n - k_1 = \log_q \binom{n}{t}$.

If $\binom{k_1}{t_1} \leq q^{Rn_1}$, then the successful algorithm for $\tau = 1$ can be used to transmit the error positions, and since this works for any number of errors, the receiver can decode the message correctly. Otherwise, let $k_2 = \log_q \binom{k_1}{t_1}$ and use the interval $[k_1 + 1, k_1 + k_2]$ to send the error positions of the t_1 errors within the first interval. This process continues by defining $k_{i+1} = \log_q \binom{k_i}{t_i}$ and $n_{i+1} = n_i - k_{i+1}$ and is stopped when $\binom{k_i}{t_i} \leq q^{Rn_i}$. Then the successful algorithm for $\tau = 1$ can be used to transmit the error positions with certainty. In the last block, the k_i s have to be transmitted as well. The amount of symbols that need to be reserved in order to do this is of order $\mathcal{O}(\log(n)\log(t))$. This does not change the asymptotically achievable rate of the algorithm.

The capacity error function for the Z-channel Γ_Z^q with an additional edge (0, q - 1) was analyzed by Deppe and Lebedev in [8] for q > 3. We denote this channel as Γ_{DL}^q and its error capacity function as $C_q^f(\Gamma_{DL}^q, \tau)$. Since the additional edge can only make the channel worse,

$$C_q^f(\Gamma_Z^q, \tau) \ge C_q^f(\Gamma_{DL}^q, \tau).$$
(15)

We denote the asymptotic rate of the algorithm achieving $C_q^f(\Gamma_{DL}^q, \tau)$ as

$$R_{DL}^{q} := \begin{cases} 1 - h(\tau) \log_{q} 2 & \text{if } 0 \le \tau \le \frac{1}{2} \\ 1 - \log_{q} 2 & \text{if } \tau > \frac{1}{2}. \end{cases}$$
(16)

The algorithm achieving R_{DL}^q uses the methods of the proofs of Lemma 1 and Lemma 2. The algorithm in the proof of Lemma 2 requires the possibility of error-free transmission at a positive rate for $\tau = 1$. For the channel Γ_{DL}^3 , this is not possible. However for Γ_U^3 the proposed method works, as it is possible to transmit data with a positive rate for $\tau = 1$ by Corollary 1.

Corollary 2. Let $q \ge 3$. Then the capacity error functions of Γ_Z^q , $\Gamma_{\underline{X}}^q$ and Γ_U^q are lower bounded by R_{DL}^q .

$$C_q^f(\Gamma_Z^q, \tau) \ge C_q^f(\Gamma_U^q, \tau) \ge R_{DL}^q \tag{17}$$

We note that Corollary 2 does not give information on the capacity error functions for the binary channels.

Corollary 3. For odd q the following equations hold for $\tau \geq \frac{1}{2}$:

$$C_{q}^{f}(\Gamma_{Z}^{q},\tau) = C_{q}^{f}(\Gamma_{U}^{q},\tau) = C_{0,q}^{f}(\Gamma_{Z}^{q}) = C_{0,q}^{f}(\Gamma_{U}^{q})$$
(18)

Proof: In accordance with the definition in the proof of Theorem 2, it holds that $C_q^{upper}(\Gamma_Z^q, \tau) = \log_q\left(\frac{q+1}{2}\right)$ for $\tau \geq \frac{1}{2}$, which is equal to $\log_q(\lceil \frac{q}{2} \rceil) = C_{0,q}^f(\Gamma_U^q)$ for odd q by Property 1.

Figure 6 illustrates the results obtained in Theorem 2, Property 1 and Corollaries 2 and 3. Property 1 shows that the methodology proposed in [8] (Lemma 1 in this work) only achieves the zero error capacity of Γ_Z^q if q is even. However, the strategy proposed in the proof of Corollary 1 is able to achieve the zero error capacity for Γ_U^q and Γ_Z^q for any q. The results in Figure 6 are in accordance with Corollary 3 for q = 5. For even q we only know the zero error capacity which is achieved by the lower bound. For $\tau \neq 1$ our results do not show whether upper and lower bounds on $C_q^f(\Gamma_U^q, \tau)$ are tight. q = 6 was chosen as an example for even q.

6. Constructing lower bounds using rubber methods

The rubber method was developed in [2]. It is used to determine a lower bound on the capacity error function of a channel. Basically it reserves a sequence of symbols to notify the receiver that his previously received symbol



Figure 6: Lower and upper bound on the capacity error function of the unidirectional channels Γ_U^5 and Γ_U^6

has been an erroneous one. Therefore, after receiving the rubber sequence, the receiver deletes the rubber plus the previous symbol. In [2], a rubber sequence consists of r equal rubber symbols $b \in Q$. The sender then retransmits the previously erroneous symbol again. Because the number of errors t is fixed and smaller than the blocklength n, he will succeed after at most t trials. In [2] it is shown that the rubber method is a successful algorithm. This means the decoder is able to decode each message correctly. In the error analysis in [2], the authors consider two kinds of errors: a standard error (which means a symbol is changed to another symbol and the sender sends the rubber sequence) and a towards rubber error (which means a symbol is changed to a rubber symbol such that the receiver obtains a rubber sequence). If a towards rubber error occurs, a correctly received symbol is deleted and has to be sent again. For the generalized Z-channel Γ_Z^q , a towards rubber error is not possible if we use r times q-1 as the rubber sequence. Also, for the inverse generalized Z-channel $\Gamma^q_{\mathbf{x}}$, a towards rubber error is not possible if we use r times 0 as the rubber sequence. Therefore, the sender does not have to retransmit the

previously erroneous symbol again, because for those channels there is only the possibility of a **standard error**. We denote this modified algorithm (without retransmissions) by A(r, b) if we use $b^r = (b, \ldots, b)$ as the rubber and we get:

Lemma 3. The modified rubber strategy A(r, q - 1) [A(r, 0)] is a successful algorithm for the generalized [inverse] Z-channel Γ_Z^q $[\Gamma_X^q]$.

In the following, we denote this rubber method without retransmission as the modified rubber method.

To calculate the rate of this algorithm we need, as in [2], the following definitions. Let $z_r^{r+1} = qz_r^r - q + 1$. It is well known that for $n \to \infty$ the number of sequences of length n not containing a block $b^r = (b, \ldots, b)$ is asymptotically equal to z_r^n (see [2, 5], how to choose the initial value for the iteration).

Theorem 3. Let z_r be defined as above. Then for the generalized [inverse] Z-channel Γ_Z^q $[\Gamma_{\nabla}^q]$ for $q \ge 2$ we get

$$C_q^f(\Gamma_Z^q, \tau) \ge R_{mr} := \begin{cases} \max_{2 \le r \in \mathbb{N}} (1 - r\tau) \log_q z_r & \text{if } 0 \le \tau \le \frac{1}{2} \\ 0 & \text{if } \tau > \frac{1}{2}. \end{cases}$$

Proof: The modified *r*-rubber method can be used to transmit information vectors not containing the rubber sequence. The blocklength is *n* and the number of erroneous symbols is at most *t*. For each error we require *r* symbols for correction. Therefore the length of the information vectors can be at most n - rt. The number of these sequences is asymptotically equal to z_r^{n-rt} . Therefore the asymptotic rate achieved by the modified *r*-rubber method is

$$(1-r\tau)\log_a z_r$$

This gives the following lower bound on the capacity error function

$$C_q^f(\Gamma_Z^q, \tau) \ge (1 - r\tau) \log_q z_r.$$

We use the modified rubber method with the rubber length r that achieves the highest asymptotic for the respective value of τ and obtain the lower bound R_{mr} .

Remark 4. For the generalized Z-channel Γ_Z^q , the modified rubber method with a single 0 as the rubber symbol achieves the rate

$$R = (1 - \tau) \log_q(q - 1).$$
(19)

Proof: Using Lemma 3, it is easy to see that for every erroneous symbol one extra symbol has to be transmitted in order to achieve error free transmission. Therefore it is possible to achieve a message set \mathcal{M} of cardinality

$$M = (q-1)^{n-t},$$
 (20)

leading to a rate

$$R = (1 - \tau) \log_q(q - 1).$$
(21)

By an adjustment of the modified rubber method for the channels Γ^q_Z and $\Gamma^q_{\overline{X}}$ we obtain the following result:

Theorem 4. Let Γ_U^q be a unidirectional channel consisting of the generalized Z-channel Γ_Z^q and the inverse generalized Z-channel Γ_{∇}^q $(q \ge 2)$. Then we have

$$C_q^f(\Gamma_U^q, \tau) \ge R_{mr}$$

Proof: We adapt the modified rubber method to make it usable for Γ_U^q . Furthermore, we show that this method is successful and achieves the same asymptotic rate as the modified rubber method. The sender starts by using the encoding strategy for the generalized Z-channel Γ_Z^q and then adapts if he discovers that the active channel is Γ_X^q . After all the data symbols are transmitted, an extra symbol (either 0 or q-1 according to the channel) is added to tell the receiver which channel was present. The receiver can thus adjust his decoding algorithm accordingly. The previously described steps are now elaborated in more detail.

The only additional difficulty compared to the situation in Theorem 3 is that it is not known to the sender and the receiver whether Γ_Z^q or Γ_{Σ}^q is used for the transmission of each codeword but once the first error occurs the sender knows which channel is in action.

At the start of the transmission, the sender assumes that Γ_Z^q is the active channel of Γ_U^q . He therefore uses the modified rubber method for Γ_Z^q of length r as his encoding strategy.

When the channel inflicts the first error to the transmission, the sender knows which channel is active. We denote the transmitted symbol corresponding to the first error with x_{fe} . According to the actual channel, the sender adjusts his encoding strategy henceforth. We distinguish two cases:

- 1. The active channel is Γ_Z^q . In this case the channel keeps its encoding strategy as it is. After the sender knows that the all information symbols can be correctly retrieved by the decoder, he fills the remaining block with 0 symbols. Notice that this last step involves knowledge about the decoding algorithm at the sender.
- 2. The active channel is Γ_{Σ}^{q} . In this case there are two cases to be distinguished
 - (a) A standard error occurred: In this case the transmitter sends r times the symbol q 1. If there are not enough symbols left in the block, the sender sends q 1 for the remaining symbols.
 - (b) a **towards rubber error** occurred: In this case the transmitter sends r - 1 times the symbol q - 1. If there are not enough symbols left in the block, the sender sends q - 1 for the remaining symbols.

The modified rubber method for Γ_{Σ}^{q} is used from now on to retransmit x_{fe} and to send the remaining information symbols. But before sending them, they are adapted to the channel Γ_{Σ}^{q} by using the mapping

$$f: \mathcal{Q} \to \mathcal{Q}$$
$$k \mapsto k+1 \mod q.$$

Potential remaining symbols within the block are filled by sending q-1 symbols.

Let the received sequence be denoted as $y^n = (y_1, \ldots, y_n)$. The decoding strategy depends on the symbol y_n :

- 1. $y_n = 0$: The decoder uses the decoding procedure for the modified rubber method for Γ_Z^q until he retrieves the message m. Potential remaining symbols within the block are ignored by the decoder.
- y_n = q 1: The decoder checks for the first occurrence of the sequence r times q-1. All previously sent symbols, with exception of the symbol that was sent right before this sequence, have been received correctly. If those information symbols are sufficient to obtain the message with certainty, the decoder outputs the message m. Otherwise, the receiver continues its decoding procedure by using the decoding algorithm of the modified rubber method for Γ^q_Σ on the remaining symbols. The function

$$f^{-1}: \mathcal{Q} \to \mathcal{Q}$$
$$k \mapsto k - 1 \mod q$$

is then to be applied on all information symbols which were sent after the first occurrence of the sequence r times q - 1 to obtain the information vector. This vector uniquely determines the message m.

If the first error occurs within the last r symbols, this error cannot be corrected anymore, but the receiver knows that all previous symbols have been received correctly. Those contain the information about the message, and thus this case is not problematic for successful decoding.

The following result follows directly from Remark 4 and Theorem 4.

Corollary 4. The asymptotic rate achieved by using the modified rubber method on the unidirectional channel Γ_U^q using a single symbol rubber is

$$R = (1 - \tau) \log_q(q - 1).$$
(22)

Remark 5. The values for z_r in Theorem 3 and Theorem 4 can be computed. An important case is q = 2. It was proved that

$$C_2^f(\Gamma^2, \tau) = (1 - 3\tau) \log_2\left(\frac{1 + \sqrt{5}}{2}\right)$$

for $(3 + \sqrt{5})^{-1} \leq \tau \leq 1/3$ (see equation (3)). In [2] it was shown that this capacity can be achieved with the rubber method.

If we apply the modified rubber method to the binary case we get the following result. The capacity error function of the unidirectional channel is lower bounded by

$$C_2^f(\Gamma_U^2, \tau) \ge \begin{cases} \max_{2 \le r \in \mathbb{N}} (1 - r\tau) \log_2 z_r & \text{if } 0 \le \tau \le \frac{1}{2} \\ 0 & \text{if } \tau > \frac{1}{2} . \end{cases}$$

The result for r = 2 is

$$(1-2\tau)\log_2\left(\frac{1+\sqrt{5}}{2}\right).$$

The major change compared to the result given in equation (3) is the change of the factor $(1-3\tau)$ to $(1-2\tau)$. It occurs because retransmissions after erroneous symbols are unnecessary. Figure 7 shows the lower bound on $C_2^f(\Gamma_U^2, \tau)$ obtained by using the modified rubber method and $C_2^f(\Gamma^2, \tau)$ for comparison. This result is different without feedback, since the capacity error functions of symmetrical and unidirectional channels are the same in that case.

7. Conclusion

In this work we have analyzed channels with feedback with a fixed maximal fraction of erroneous symbols. A major focus of this work has been asymmetric and unidirectional channels. Notably, we have shown that the capacity error function with feedback for the unidirectional channel composed of Z-channel and inverse Z-channel is larger than the capacity error function of the symmetric channel for a significant proportion of the possible values of τ . This is not the case for the same channels without feedback. The channels analyzed in this work have at most two input symbols connected to each output. As the knowledge of an erroneous position at the receiver implies knowing the value of the respective symbol, retransmissions are not necessary if the encoding strategy is chosen in a way such that the receiver is able to obtain the error positions. This can be



Figure 7: Comparison of the lower bound on $C_2^f(\Gamma_U^2, \tau)$ obtained by using the modified rubber method and the capacity error function of the symmetric channel $C_2^f(\Gamma^2, \tau)$

used to create encoding strategies achieving a higher rate for many channels using the modified rubber method. Furthermore it was shown how to change the modified rubber method for unidirectional channels. The method proposed shows that it is possible to achieve the same rate for the unidirectional channel consisting of the generalized Z-channel and the inverse generalized Z-channel as for its components. Obtaining tighter bounds on the capacity error functions for several of the channels analyzed in this paper is an interesting topic for further research.

Acknowledgement

The authors would like to thank Frans Willems for making them aware that the methods proposed in [8] might be adapted for use on unidirectional channels.

Christian Deppe was supported by the Bundesministerium für Bildung und Forschung (BMBF) through Grant 16KIS1005. Vladimir Lebedev was supported in part by the Russian Foundation for Basic Research, project no. 19-01-00364. Georg Maringer's work was supported by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) under Grant No.WA3907/4-1.

References

- R. Ahlswede, H. Aydinian, and L.H. Khachatrian, Unidirectional Error Control Codes and Related Combinatorial Problems, Proceedings of Eight International Workshop on Algebraic and Combinatorial Coding Theory, Russia, 2002.
- [2] R. Ahlswede, D. Deppe, and V. Lebedev, Non-binary error correcting codes with noiseless feedback, localized errors, or both, Annals of the European Academy of Sciences 1, 285-309, 2005.
- [3] L.A. Bassalygo, New upper bounds for error correcting codes, Problems Information Transmission, 1:4, 32-35, 1965.
- [4] E.R. Berlekamp, Block coding for the binary symmetric channel with noiseless delayless feedback, Proc. Symposium on Error Correcting Codes, Univ. Wisconsin, 1968.
- [5] R.A. Brualdi, Introductory Combinatorics, Elsevier North-Holland, 1977.
- [6] F. Cicalese, Fault-Tolerant Search Algorithms, Springer, EATCS Series, 2013.
- [7] C. Deppe, Searching with lies and coding with feedback, Entropy, Search, Complexity, Bolyai Society Mathematical Studies, Katona (Ed.), Vol. 16, 27-70, 2007.
- [8] C. Deppe, and V. Lebedev, Algorithms for Q-ary Error-Correcting Codes with Partial Feedback and Limited Magnitude, ISIT 2019.
- [9] V. Lebedev, Coding with noiseless feedback, Problems of Information Transmission, Vol. 52, No. 2, 103-113, 2016.
- [10] F.J. MacWilliams, and N.J.A. Sloane, The theory of error-correcting codes, Vol. 16, Elsevier 1977.

- [11] C.E. Shannon, The zero error capacity of a noisy channel, IRE Trans. Inform. Theory, IT-2, 3, 8–19, 1956.
- [12] Weber J. H., Bounds and Constructions for Binary Block Codes Correcting Asymmetric or Unidirectional Errors, TU Delft, Dissertation 1989.
- [13] K.Sh. Zigangirov, On the number of correctable errors for transmission over a binary symmetrical channel with feedback, Problems Information Transmission 12, No. 2, 85-97, 1976.