# Non–binary error correcting codes with noiseless feedback, localized errors, or both

R. Ahlswede, C. Deppe[*], and V. Lebedev[†]

## 1 Introduction

A famous problem in Coding Theory consists in finding good bounds for the maximal size $M(n,t,q)$ of a $t$-error correcting code over a $q$-ary alphabet $\mathcal{Q}=\{0,1,\ldots,q-1\}$ with block length $n$.
This code concept is suited for communication over a $q$-ary channel with input alphabet $\mathcal{X}=\mathcal{Q}$ and output alphabet $\mathcal{Y}=\mathcal{Q}$, where a word of length $n$ sent by the encoder is changed by the channel in at most $t$ letters. Here neither the encoder nor the decoder knows in advance where the errors, that is changes of letters, occur.
Suppose now that having sent letters $x_1,\ldots,x_{j-1}\in\mathcal{X}$ the encoder knows the letters $y_1,\ldots,y_{j-1}\in\mathcal{Y}$ received before he sends the next letter $x_j$ $(j=1,2,\ldots,n)$. We then have the presence of a noiseless feedback channel.
For $q=2$ this model was considered by Berlekamp [10], who derived striking results for triples of performance $(M,n,t)_f$, that is, the number of messages $M$, block length $n$ and the number of errors $t$. It is convenient to use the notation of relative error $\tau=t/n$ and rate $R=n^{-1}\log M$. We investigate here the $q$-ary case. Again the Hamming bound $H_q(\tau)$ for $C_q^f(\tau)$, the supremum of the rates achievable for $\tau$ and all large $n$, is a central concept:

$$H_q(\tau)=\begin{cases}1-h_q(\tau)-\tau\log_q(q-1) & \text{if } 0\le\tau\le\frac{q-1}{q}\\ 0 & \text{if } \frac{q-1}{q}<\tau\le 1,\end{cases} \tag{1}$$

[*]Supported by the DFG in the project "Allgemeine Theorie des Informationstranfer und Kombinatorik"

[†]Supported in part by the Russian Foundation for Basic Research, project no 06-01-00226.

where $h_q(\tau) = -\tau \log_q(\tau) - (1-\tau)\log_q(1-\tau)$. We also call $C_q^f : [0,1] \to \mathbb{R}_+$ **the capacity error function (or curve)**. One readily verifies that for every $q$

$$C_q^f(\tau) = 0 \text{ for } \tau \geq \frac{1}{2}. \tag{2}$$

We turn now to another model. Suppose that the **encoder**, who wants to encode message $i \in \mathcal{M} = \{1, 2, \ldots, M\}$, knows the $t$-element set $E \subset [n] = \{1, \ldots, n\}$ of positions, in which only errors may occur. He then can make the codeword presenting $i$ dependent on $E \in \mathcal{E}_t = \binom{[n]}{t}$, the family of $t$-element subsets of $[n]$. We call them "a priori error pattern". A family $\{u_i(E) : \ 1 \leq i \leq M, \ E \in \mathcal{E}_t\}$ of $q$-ary vectors with $n$ components is an $(M, n, t, q)_l$ code (for localized errors), if for all $E, E' \in \mathcal{E}_t$ and all $q$-ary vectors $e \in V(E) = \{e = (e_1, \ldots, e_n) : \ e_j = 0 \text{ for } j \notin E\}$ and $e' \in V(E')$

$$u_i(E) \oplus e \neq u_{i'}(E') \oplus e' \text{ for } i \neq i',$$

where $\oplus$ is the addition modulo $q$. We denote now the capacity error function by $C_q^l$. It was determined in [8] for the binary case to equal $H_2(\tau)$. For general $q$ the best known result is

**Theorem** ABP [6]

*(i)* $C_q^l(\tau) \leq H_q(\tau)$, *for* $0 \leq \tau \leq \frac{1}{2}$.

*(ii)* $C_q^l(\tau) = H_q(\tau)$, *for* $0 \leq \tau < \frac{1}{2} - \frac{q-2}{2q(2q-3)}$ .

The two models described have as ingredients feedback resp. localized errors, which give possibilities for code constructions not available in the standard model of error correction (c.f. [16] and also for probabilistic channel models [1], [5]).
For the feedback model we present here a coding scheme based on an idea of deletions. It is easy to analyze and yields also Berlekamp's results for the case $q = 2$. Whereas all this work is for block codes we next investigate variable length codes with all lengths bounded from above by $n$. The end of a word carries the symbol $\square$ and is thus recognizable by the decoder. Very important here is that the lengths carry **sure** data which can be used as a "protocol" information.

For both, the $\square$-model with feedback and the $\square$-model with localized errors, the Hamming bound is the exact capacity curve for $\tau < 1/2$. Whereas with feedback the capacity curve coincides with the Hamming bound also

for $1/2 \leq \tau \leq 1$, somewhat surprisingly in this range for localized errors the capacity curve equals 0.
Also notice that without the marker $\Box$ in the range $0 \leq \tau < 1/2$ with feedback the capacity curve is **smaller** than for localized errors (see Theorem 1 in Section 2 and Theorem ABP above). Also we give constructions in the $\Box$-model with both, feedback and localized errors.
Finally, in the standard model with feedback **and** localized errors the help of feedback is addressed. We give an optimal construction for one-error correcting codes with feedback and localized errors.

# 2 $q$-ary block codes with feedback

We consider a channel with one sender (or encoder) and one receiver (or decoder). Both, the input alphabet $\mathcal{X}$ and the output alphabet $\mathcal{Y}$, equal $\mathcal{Q} = \{0, 1, \ldots, q-1\}$ and they operate in the presence of noiseless feedback. By this is meant that there exists a return channel which sends back from the receiving point to the transmitting point the element of $\mathcal{Y}$ actually received. It is assumed that this information is received at the transmitting point before the next letter is sent, and can therefore be used for choosing the next letter to be sent.
A $t$-error correcting $(M, n, t, q)_f$ code of block length $n$ and $M$ messages for this channel is a system of pairs $\{(f_i^n, \mathcal{D}_i) : i \in \mathcal{M}\}$ which is described as follows:
There is given a finite set of messages $\mathcal{M} = \{1, \ldots, M\}$, one of which will be presented to the sender for transmission. Message $i \in \mathcal{M}$ is encoded by an encoding (vector valued) function $f_i^n = (f_{i1}, \ldots, f_{ij}, \ldots, f_{in})$, where $f_{ij}$ is defined on $\mathcal{Y}^{j-1} = \prod_1^{j-1} \mathcal{Y}$ for $j > 1$ and takes values in $\mathcal{X}_j = \mathcal{X} = \mathcal{Q}$, and $y_1, y_2, \ldots, y_{i-1}$ are received elements of $\mathcal{Y}$ (known to the sender before he sends $f_{ij}(y_1, \ldots, y_{i-1})$; $f_{i1}$ is an element of $\mathcal{X}_1$.
It is assumed that the number of wrongly transmitted letters in $n$ steps does not exceed $t$ and that the receiver has a decoding system $\{\mathcal{D}_i : i \in \mathcal{M}\}$, $\mathcal{D}_i \subset \mathcal{Y}^n$, $\mathcal{D}_i \cap \mathcal{D}_{i'} = \emptyset$ for $i \neq i'$ such that upon receiving $y^n = (y_1, \ldots, y_n)$ he can correctly decide (or decode), which message was sent. Our goal is to determine the capacity error function $C_q^f$ for every $q$. Bassalygo conjectured the following. Let $Ta_q$ be a tangent to $H_q$ with $Ta_q(\frac{q-1}{2q-1}) = 0$ and $a_q$ the

argument with $H_q(a_q) = Ta_q(a_q)$, then

$$C_q^f(\tau) = \begin{cases} H_q(\tau) & \text{if } \ 0 \le \tau \le a_q \\ Ta_q(\tau) & \text{if } \ a_q \le \tau \le \frac{q-1}{2q-1} \\ 0 & \text{otherwise.} \end{cases} \tag{3}$$

One reason for this conjecture was a result of [13], which implies that

$$C_q^f(\frac{q-1}{2q-1}) \ge 0.$$

We will show that Bassalygo's conjecture is not true. We begin with two estimates from above.

## 2.1 Upper bounds on $C_q^f$

First notice that for $t \ge \frac{n}{2}$ (or $\tau \ge \frac{1}{2}$) not even two messages can be transmitted correctly, because for their encoding functions, say

$$f^n = (f_1, \ldots, f_n) \text{ and } g^n = (g_1, \ldots, g_n),$$

at any component $j$ and any output string $y^{j-1} = y_1 \ldots y_{j-1}$ an error for one of the messages can cause

$$f_j(y_1, \ldots, y_{j-1}) = g_j(y_1, \ldots, y_{j-1}).$$

Since $2t \ge n$, there are enough errors to produce this identity for all $j = 1, 2, \ldots, n$ and two messages cannot be decoded correctly:

$$C_q^f(\tau) = 0 \text{ for } \tau \ge \frac{1}{2}. \tag{4}$$

Next we derive the Hamming upper bound.

**Lemma 1**

*(i) For every $(M, n, t, q)_f$ code holds*

$$M \sum_{j=0}^{t} \binom{n}{j} (q-1)^j \le q^n.$$

*(ii) For $0 \le \tau \le 1$*

$$C_q^f(\tau) \le H_q(\tau).$$

**Proof:** We count the number of output sequences. Let $i \in \mathcal{M}$ and $y^n = (y_1, \ldots, y_n)$ be the output sequence with

$$y_1 = f_{i1} \oplus e_1 \text{ and } y_j = f_{ij}(y_1, y_2, \ldots, y_{j-1}) \oplus e_j \text{ for } j = 2, 3, \ldots, n$$

determined by the encoding function $f_i^n$ and the $q$-ary additive noise $e^n = (e_1, \ldots, e_n) \in \mathcal{Q}^n$ occurring in the transmission and so can be regarded as their function $\phi(f_i^n, e^n)$. For the family of encoding functions $F_{\mathcal{M}} = \{f_i^n : i \in \mathcal{M}\}$ and a set $\mathcal{V} \subset Q^n$ of error patterns we write $\Phi(F_{\mathcal{M}}, \mathcal{V}) = \{y^n : \text{there exist } i \in \mathcal{M} \text{ and } e^n \in \mathcal{V} \text{ such that } y^n = \phi(f_i^n, e^n)\}$. If at most $t$ errors occur, we have

$$\mathcal{V} = \bigcup_{E \in \mathcal{E}_t} V(E) = \{e^n = (e_1, \ldots, e_n) : e_j = 0 \text{ for at least } n - t \text{ components}\}.$$

Then we also have

$$\phi(f_i^n, e^n) \neq \phi(f_{i'}^n, e'^n) \text{ for } (i, e^n) \neq (i', e'^n). \tag{5}$$

Indeed this is the case if $i \neq i'$ (because the decoder must be able to distinguish the messages) and if $i = i'$ and $e^n \neq e'^n$ (because the $j$th symbols of $\phi(f_i^n, e^n)$ and $\phi(f_i^n, e'^n)$ are different if $j$ is the first position where $e^n$ and $e'^n$ are different).
Therefore

$$|\Phi(F_{\mathcal{M}}, \mathcal{V})| = M|\mathcal{V}| = M \sum_{j=0}^{t} \binom{n}{j} (q-1)^j \leq q^n \tag{6}$$

and asymptotically we get

$$C_q^f(\tau) \leq H_q(\tau), 0 \leq \tau \leq 1. \tag{7}$$

For the range $\frac{1}{q} \leq \tau \leq \frac{1}{2}$ we derive a second basic upper bound using a result of Aigner.

**Theorem A[7]**
*For every $q \geq 2$ and $t \leq \frac{1}{2}n$, if there exists an $(M, n, t, q)_f$ code, then there exists an $(M, n - 2m, t - m, q)_f$ code for every $0 \leq m \leq t$.*

Substituting in (6) the parameters $M, n, t$ by $M, n - 2m$, and $t - m$ we get

$$M \sum_{j=0}^{t-m} \binom{n-2m}{j} (q-1)^j \leq q^{n-2m}. \tag{8}$$

Consequently $M \cdot \binom{n-2m}{t-m}(q-1)^{t-m} \leq q^{n-2m}$ and, asymptotically, for $0 \leq \mu \leq \tau$

$$C_q^f(\tau) \leq H_q\left(\frac{\tau - \mu}{1 - 2\mu}\right)(1 - 2\mu). \tag{9}$$

Whereas Berlekamp [10] showed in the case $q = 2$ that the tangent at the curve $H_2(\tau)$ running through the point $\left(\frac{1}{k}, 0\right)$ is an upper bound for $C_2^l(\tau)$ if $k = 3$ and a lower bound for $C_2^l(\tau)$ if $k \geq 3$, we show here first for $q > 2$ that the tangent at $H_q(\tau)$ in the point $\left(\frac{1}{q}, H_q\left(\frac{1}{q}\right)\right)$ running through the point $\left(\frac{1}{2}, 0\right)$ gives an upper bound on $C_q^f(\tau)$ for $\frac{1}{q} \leq \tau \leq \frac{1}{2}$.
(In fact this is part of our basic result that in this interval the tangent describes the capacity curve.)
One readily verifies that

$$\frac{d\ H_q(\tau)}{d\tau} = \log_q \frac{\tau}{(1-\tau)(q-1)}. \tag{10}$$

So the tangent at the point with abscissa $a$ has the equation

$$T(\tau) = \tau \log_q \left(\frac{a}{(1-a)(q-1)}\right) + R_0,$$

where $R_0 = T(0)$. Going through $(a, H_q(a))$ implies that

$$\begin{aligned} R_0 &= H_q(a) - a \log_q \frac{a}{(1-a)(q-1)} \\ &= \log_q \left(\frac{q a^a (1-a)^{1-a}}{(q-1)^a}\right) - \log_q \frac{a^a}{(q-1)^a (1-a)^a} = \log_q (q(1-a)) \end{aligned}$$

and therefore

$$T(\tau) = \tau \log_q \left(\frac{a}{(1-a)(q-1)}\right) + \log_q (q(1-a)). \tag{11}$$

Finally, $T\left(\frac{1}{2}\right) = 0$ implies

$$\log_q \left(\frac{a\ q^2(1-a)^2}{(1-a)(q-1)}\right) = 0$$

and therefore

$$a\ q^2(1-a) = q - 1 \text{ and } a = \frac{1}{q}. \tag{12}$$

(The other root is $\frac{q-1}{q}$.)
The form of our tangent is

$$T(\tau) = (1 - 2\tau) \log_q (q-1). \tag{13}$$

We are prepared to state

**Lemma 2** *For* $\frac{1}{q} \leq \tau \leq \frac{1}{2}$

$$C_q^f(\tau) \leq (1-2\tau)\log_q(q-1).$$

**Proof:** By (9) it suffices to show that the equation

$$H_q\left(\frac{\tau-\mu}{1-2\mu}\right)(1-2\mu) = (1-2\tau)\log_q(q-1) \tag{14}$$

has a solution in $\mu$ for $0 \leq \mu \leq \tau$.
This can be written in the form

$$\begin{aligned}(1-2\mu)[1 + \frac{\tau-\mu}{1-2\mu}\log_q\frac{\tau-\mu}{1-2\mu} + \frac{1-\mu-\tau}{1-2\mu}\log_q\frac{1-\mu-\tau}{1-2\mu}\\ -\frac{\tau-\mu}{1-2\mu}\log_q(q-1)]\\ = (1-2\tau)\log_q(q-1)\end{aligned}$$

or in the form

$$(1-2\mu)\log_q\frac{q}{1-2\mu} + (1-\mu-\tau)\log_q\frac{1-\mu-\tau}{q-1} + (\tau-\mu)\log_q(\tau-\mu) = 0.$$

Here the first coefficient equals the sum of the two others.
We equate the arguments in the second and the third log:

$$\begin{aligned} 1-\mu-\tau &= (q-1)(\tau-\mu) && (15)\\ (q-2)\mu &= q\tau - 1 && (16)\\ \mu &= \frac{q\tau-1}{q-2} && (17)\end{aligned}$$

Then the first log has the argument $\left(\frac{1-2\mu}{q}\right)^{-1}$ and since by (15) $\frac{1-2\mu}{q} = \tau-\mu$, the desired equation follows, because $-(1-2\mu)+(1-\mu-\tau)+(\tau-\mu) = 0$.

**Remark 1** *Another way to find* $\mu$ *is to maximize* $bin(m) = q^{2m-n}\binom{n-2m}{t-m}(q-1)^{t-m}$ *by comparing* $bin(m)$ *and* $bin(m+1)$ *like it is done in the binary case in [10].*

## 2.2 Lower bound derived by the new rubber method

In this section we will give a strategy which achieves the upper bound in Lemma 2 for relative errors $\frac{1}{q} \leq \tau \leq \frac{1}{2}$. We show that we can transmit $(q-1)^{n-2t}$ messages in block length $n$. A bijection $b$ of messages $\mathcal{M}$ to the set $\{1, 2, \ldots, q-1\}^{n-2t}$ of used sequences is agreed upon by the sender and the receiver.

Given message $i \in \mathcal{M}$ the sender chooses $b(i) = x^{n-2t} = (x_1, x_2, \ldots, x_{n-2t}) \in \{1, 2, \ldots, q-1\}^{n-2t}$ as a **skeleton for encoding**, which finally will be known to the receiver. The "0" is used for error correction only. For all positions $i \leq n$ not needed dummy $x_i = 1$ are defined to fill the block length $n$.

**Transmission algorithm:** The sender sends $x_1$, continues with $x_2$ and so on until the first error occurs, say in position $p$ with $x_p$ sent. The error can here be of two kinds: a **standard error** (that means symbol $x_p$ is changed to another symbol $y_p \in \{1, 2, \ldots, q-1\}$) and a **towards zero error** (that means $x_p$ is changed to $y_p = 0$).

If a **standard error** occurs, the sender transmits, with smallest $l$ possible, $2l+1$ times 0 (where $l \in \mathbb{N} \cup 0$) until the decoder received $l+1$ zeros (known to the sender via feedback. Such an $l$ exists because the number of errors is bounded by $t$). Then he transmits at the next step $x_p$, again, and continues the algorithm.

If a **towards zero error** occurs, the sender decreases $p$ by one (if it is bigger than 1) and continues (transmits at the next step $x_p$).

**Decoding algorithm:** The decoding is very simple. The receiver just regards the "0" as a kind of deletion symbol - he erases it by a rubber, **who in addition erases the previous symbol.**

This is the reason, why the sender has to repeat sending the symbol according to the skeleton, if a towards zero error occurs.

At the end the first $n-2t$ symbols at the decoder are those of $b(i) = (x_1, x_2, \ldots, x_{n-2t})$.

Indeed, suppose that $t_0$ towards zero errors occur. They are taken care of with loss in block length $2t_0$. So we are left with $t_1 = t - t_0$ possible errors and block length $n - 2t_0$ and only standard errors as well as a third kind of correction errors resulting from a change of a zero to a non-zero. The standard errors $s_1, \ldots, s_r$ cause correction errors $l_1, \ldots, l_r$ resp. and loss in block lengths $2(l_1+1), \ldots, 2(l_r+1)$ and thus a total of $\sum_{i=1}^{r}(1+l_i) \leq t_1$ errors and a total of $\sum_{i=1}^{r} 2(l_i+1) \leq 2t$ block length.

Hence a block length $n - 2(t_0 + t_1) = n - 2t$ to transmit with our strategy $M = (q-1)^{n-2t}$ messages.

Thus $\frac{\log_q M}{n} = (1 - \frac{2t}{n}) \log_q(q-1)$ and we have derived the main result of this section.

**Theorem 1** *For* $\tau = \frac{t}{n}$ *and* $0 < \tau < \frac{1}{2}$

$$C_q^f(\tau) \geq (1 - 2\tau) \log_q(q-1).$$

Now we will show that we can generalize the rubber method in such a way, that we get as the rate function a tangent to the Hamming bound through $(\tau, 0) = (\frac{1}{r+1}, 0)$, where $1 \leq r \in \mathbb{N}$. The generalization also works for $q = 2$ and gives also an alternative optimal strategy to Berlekamp's method for this case. The $r$-rubber method is defined as follows: The communicators map all messages to sequences of the set

$$\mathcal{X}_r^{n-(r+1)t} = \{x^{n-(r+1)t} : \text{the sequence contains} \leq r-1 \text{ consecutive zeros}\}$$

and the sender uses now $r$ consecutive zeros as a deletion symbol. The following result is well known.

**Theorem** B [12]
*Let* $r \geq 2$ *and* $X_r^n = \sum_{j=1}^{r}(q-1)X_r^{n-j}$ *with* $X_r^i = q^i$ *for* $i = 1, \ldots, r-1$ *and* $X_r^r = q^r - 1$. *Then* $|\mathcal{X}_r^{n-(r+1)t}| = X_r^{n-(r+1)t}$.

**Theorem 2** *The rate function of the* $r$*-rubber method is a tangent to* $H_q(\tau)$ *going through* $\frac{1}{r+1}$.

**Proof:**
We know (Theorem 6.2.1 in [12]) that there exists an $z_r \in \mathbb{R}$ for each $r$ and $n \geq n_0$ such that $X_r^n = z_r^n$ for this $z_r$ holds (because of the definition):

$$z_r^{r+1} = (q-1)(\sum_{j=1}^{r} z_r^j) \tag{18}$$

$$z_r^r = (q-1)(\sum_{j=0}^{r-1} z_r^j) \tag{19}$$

From these equations we get:

$$z_r^{r+1} - z_r^r = (q-1)z_r^r - q + 1. \tag{20}$$

With the rubber method we get the following rate function: $R_r(\tau) = (1 - (r+1)\tau) \log_q z_r$.
We know that the tangents have the form: $T_r(\tau) = \tau m + b$, where $m = \log_q \frac{k}{(1-k)(q-1)}$. Let $k$ be the abscissa of the point of contact, then

$$H_q(k) = \log_q \frac{qk^k(1-k)^{1-k}}{(q-1)^k} = k \log_q \frac{k}{(1-k)(q-1)} + b.$$

Thus $b = -\log_q \frac{1}{q(1-k)}$ and $T_r(\tau) = \tau \log_q \frac{k}{(1-k)(q-1)} - \log_q \frac{1}{q(1-k)}$. It holds $T_r(\frac{1}{r+1}) = 0$. Thus we have

$$\frac{1}{r+1} \log_q \frac{k}{(1-k)(q-1)} = \log_q \frac{1}{q(1-k)} \tag{21}$$

$$\frac{(1-k)(q-1)}{k} = (q(1-k))^{r+1}. \tag{22}$$

We set $z_r = q(1-k)$ and get

$$kz_r^{r+1} = (1-k)q - (1-k) \tag{23}$$

$$(1 - \frac{z_r}{q})z_r^{r+1} = z_r - \frac{z_r}{q} \tag{24}$$

$$qz_r^{r+1} - z_r^{r+2} = qz_r - z_r \tag{25}$$

$$z_r^{r+1} = qz_r^r - q + 1. \tag{26}$$

This is equivalent to (19) and thus the two functions are the same.

# 3 Block Codes with feedback and localized errors

We consider in this section block codes with feedback and with localized errors. For each message $i$ of the set $\mathcal{M} = \{1, 2, \ldots, M\}$ depending on $E \in \mathcal{E}_t$, the family of $t$-element subsets of $[n]$ (because we have localized errors), and depending on the error vector $e \in \{0, \ldots, q-1\}^n$ occurring during the transmission (because of the feedback) the receiver gets $y^n = \psi(i, E, e)$ where $e \in V(E) = \{(e_1, \ldots, e_n) : e_j = 0 \; if \; j \notin E\}$. Let $\Psi(i, E) = \bigcup_{e \in V(E)} \psi(i, E, e)$, then for $M_{fl}(n, t, q)$ clearly

$$\Psi(i, E) \cap \Psi(i', E') = \emptyset \; \forall i \neq i'. \tag{27}$$

The following lemma follows from Lemma 1 and Theorem 1 in [4], because the number of output sequences is not reduced by feedback. Thus we follow the same ideas and prove

**Lemma 3** *For any distinct subsets $E(j)$ $(j \in J)$ of $\{1, \ldots, n\}$ and any output function $\Psi$ we have*

$$\left| \bigcup_{j \in J} \Psi(i, E(j)) \right| \geq \sum_{l=0}^{n} \lambda(l)(q-1)^l,$$

*where $\lambda(l) = \{j \in J : 1 \leq j \leq n, |E(j)| = l\}$.*

In the same way as in [4], by using Lemma 3, we get the following

**Theorem 3**

*(i)* $M_{fl}(n,t,q) \leq \frac{q^n}{\sum_{i=0}^{t}\binom{n}{i}(q-1)^i}$

*(ii)* $C_q^{fl}(\tau) \leq H_q(\tau)$.

Pelc showed in [18] that

$$M_f(n,1) = \begin{cases} \lfloor \frac{2^n}{n+1} \rfloor & \text{if } \lfloor \frac{2^n}{n+1} \rfloor \text{ is even} \\ \lfloor \frac{2^n-(n-1)}{n+1} \rfloor & \text{otherwise .} \end{cases} \tag{28}$$

For localized errors Bassalygo stated the conjecture that $M_l(n,1) = \lfloor \frac{2^n}{n+1} \rfloor$ for all $n$. In [14] it was shown that $M_l(n,1) = \lfloor \frac{2^n}{n+1} \rfloor$, if $n = p-1$, such that 2 is a primitive root of $n$ and $p$ is a prime. It is possible to check that $M_l(n,1) = \lfloor \frac{2^n}{n+1} \rfloor$, if $n \leq 8$. But in general the problem is open.
We will construct an optimal code with feedback and localized errors such that $M = \lfloor \frac{2^n}{n+1} \rfloor$. The idea is to construct disjoint sets $A_1(n), \ldots, A_M(n) \subset \{0,1\}^n$ with the following property:
For all $i \in \{1, \ldots, M\}$ and all $1 \leq j \leq n$ there exist $x^n, y^n \in A_i(n)$ such that

$$x_k = y_k \text{ for all } k < j, \tag{29}$$

$$x_j \neq y_j. \tag{30}$$

The sets can be used in the following way for encoding and decoding. The sender and the receiver both know the sets. The sender wants to transmit the message $i$ and gets the a priori error pattern $E = \{j\}$ (because of the localized error). Then he chooses in the set $A_i(n)$ two words $x^n$ and $y^n$ which satisfy (29) and (30). He transmits the first $j$ positions of the alphabetically smaller one and knows whether an error occurred in the $j$-th position, because of the feedback. If no error occurred, he continues sending the word, otherwise he continues with the other one. In each case the receiver gets $x^n$ or $y^n$, a codeword of the set $A_i(n)$ and decodes the message $i$.
Let $A \subset \{0, \ldots, q-1\}^n$ and $h \in \{0, \ldots, q-1\}$, then we set

$$hA = \{(h, x_1, \ldots, x_n) : x^n \in A\} \subset \{0, \ldots, q-1\}^{n+1}.$$

Before we start with the construction we need the following

**Observation 1** *A set $A_i(n)$ can always be represented as a simple graph $(\mathcal{V}, \mathcal{E})$[1]. with $|\mathcal{V}| = n+1$ and $|\mathcal{E}| = n$ without cycles, because for every position $j$ there exists at least 2 words, such that the first $j-1$ positions are pairwise equal and the $j$-th positions are different. We represent each word by a vertex and select two words for each position, which satisfy (29) and (30) and connect them by an edge. This is just a tree. Therefore there exist two subgraphs $(\mathcal{V}_1, \mathcal{E}_1)$ and $(\mathcal{V}_2, \mathcal{E}_2)$ with $|\mathcal{V}_1| = \lceil \frac{n}{2} \rceil$ and $|\mathcal{V}_2| = \lfloor \frac{n}{2} \rfloor + 1$ such that all edges are contained in $\mathcal{E}_1 \cup \mathcal{E}_2$. This means that we can construct from a set $A_i(n)$ a set $A(n+1)$ which contains $\lceil \frac{n}{2} \rceil$ elements of $1A_i(n)$ and $\lfloor \frac{n}{2} \rfloor + 1$ elements of $0A_i(n)$.*

Now we start with the construction. We set $M(n) = \lfloor \frac{2^n}{n+1} \rfloor$. For $n = 3$ and $M(3) = \left\lfloor \frac{2^3}{4} \right\rfloor = 2$ the sets $A_1(3) = \{000, 001, 010, 100\}$ and $A_2(3) = \{111, 110, 101, 011\}$ fulfill all conditions. We will construct the other sets inductively.

**Case 1:** $M(n+1) = \left\lfloor \frac{2^{n+1}}{n+2} \right\rfloor$ is even. We set $A'_{2k}(n+1) = 0A_k(n)$ and $A'_{2k-1}(n+1) = 1A_k(n)$ for $1 \leq k \leq M$. The sets $A'_k(n+1)$ for $1 \leq k \leq M(n+1)$ fulfill all properties for $j \geq 2$. Therefore we set $A_k(n+1) = A'_k(n+1) \cup x^n(k)$ with $x_1(k) = 1$ if $k$ is even and with $x_1(k) = 0$ if $k$ is odd and all $x^n(k) \in A = \bigcup_{l=M(n+1)+1}^{2M(n)} A'_l(n+1)$ are different. This is always possible, because $|A| \geq M(n+1)$ and half of the words start with 1 and the other half start with 0.

**Case 2:** $M(n+1) = \left\lfloor \frac{2^{n+1}}{n+2} \right\rfloor$ is odd. Again we set $A'_{2k}(n+1) = 0A_k(n)$ and $A'_{2k-1}(n+1) = 1A_k(n)$ for $1 \leq k \leq M$. By Observation 1 we can construct $A_M(n+1)$ from the sets $A'_M(n+1)$ and $A'_{M+1}(n+1)$ in such a way that half of the sequences start with 0 and the other half start with 1. Now we can define the sets $A_k(n+1)$ for $k = 1, \ldots, M-1$ as before.

Thus we get the following

**Theorem 4** $M_{fl}(n, 1) = \lfloor \frac{2^n}{n+1} \rfloor$.

The construction can also be generalized to the $q$-ary case. Here the induction works in a very similar way, but we can start it first at $n = q+1$ (use for example the construction of [15]). We have to construct sets $A_1^q(n), \ldots A_M^q(n)$ for $M = \left\lfloor \frac{q^n}{(q-1)n+1} \right\rfloor$ with the following property:

For all $i \in \{1, \ldots, M\}$ and all $1 \leq j \leq n$ there exist $x^0, x^1, \ldots, x^{q-1} \in A_i(n)$ such that

$$x_k^0 = x_k^g \text{ for all } k < j \text{ and all } 0 \leq g \leq q-1, \tag{31}$$

[1]It is customary to use the letters $\mathcal{E}$ and $\mathcal{V}$ for error patterns and for the set of edges and vertices in graphs. Being aware of this no difficulty in understanding should arise

$$\forall x \in \mathcal{X} : \exists l \in \{0, 1, \ldots, q-1\} \text{ such that } x_j^l = x. \tag{32}$$

**Theorem 5**

$$M_{fl}(n, 1, q) = \lfloor \frac{q^n}{(q-1)n+1} \rfloor \quad \textit{if} \quad n \geq q+1.$$

**Proof:** The proof follows the same ideas as in the binary case. It holds $M_{fl}(q+1, 1, q) = \frac{q^n}{(q-1)n+1}$. The induction works in the same way as in the binary case, but the starting point is $n = q+1$. From this point on we can construct the codes inductively. Let $M(n) = \left\lfloor \frac{q^n}{(q-1)n+1} \right\rfloor$.

**Case 1:** $M(n+1) = \left\lfloor \frac{q^{n+1}}{(q-1)(n+1)+1} \right\rfloor$ is divisible by $q$. This is done like in the even case for $q = 2$. We set $A'_{jl}(n+1) = (j-1)A_l(n)$ for all $l = 1, 2, \ldots, M(n)$ and for all $j = 1, 2, \ldots, q$ . For $l = 1, 2, \ldots M(n+1)$ we set $A_l(n+1) = A'_l(n+1) \cup w_l(n+1)$, where $w_l(n+1)$ are $q-1$ words with different first letter. All $w_l(n+1)$ are disjoint.

**Case 2:** $M(n+1) = \left\lfloor \frac{q^{n+1}}{(q-1)(n+1)+1} \right\rfloor = k \mod q$ and $1 \leq k \leq q-1$. This also works similar to the binary case. For $l = 1, 2, \ldots M'(n+1) - k$ as before we set $A_k(n+1) = A'_k(n+1) \cup w_k(n+1)$, where $w_k(n+1)$ are $q-1$ words with different first letter. All $w_k(n+1)$ are disjoint and chosen in such a way that we can construct $k$ more sets out of the remaining sets. This is possible, because all sets $A'_i(n+1)$ for $1 \leq i \leq M'(n+1) + q - k$ have an inherited "feedback structure". This means that in each of these sets there exists one word $x^n$, such that for every localized error position $j > 1$ there exist $q-1$ words in the set with the first $j-1$ symbols pairwise equal to the first $j-1$ symbols of $x^n$ and each element of $\mathcal{Q}$ is contained in one of the $j$-th positions of these words or $x^n$.

**Remark 2** *The case $n < q+1$ does in contrast to the case $n \geq q+1$ not always reach the Hamming bound. The first example is $n = 3$ and $q = 6$. It holds*

$$M_{fl}(3, 6, 1) = 12 < \left\lfloor \frac{6^3}{3 \cdot 5 + 1} \right\rfloor = 13.$$

*We hope that the exact result can be found in the near future.*

**Remark 3** *Generalizing Pelc's result (28) to general q Aigner [7] and Malinowski [17] proved:*

$$M_f(n, 1, q) = \begin{cases} q^{n-2} & \textit{if} \quad n \leq q+1 \\ \lfloor \frac{q^n - r(n-1)(q-1)}{(q-1)n+1} \rfloor & \textit{if} \quad n \geq q+1 \end{cases},$$

$$\text{where } r = \lfloor \frac{q^n}{(q-1)n+1} \rfloor \mod q.$$

*Thus it holds* $M_f(n,q,1) = M_{fl}(n,q,1)$ *if* $\lfloor \frac{q^n}{(q-1)n+1} \rfloor \mod q = 0$.

# 4 The □-model with feedback

## 4.1 List codes for the standard model of error–correcting codes

It was demonstrated in [2] that in **probabilistic** channel coding theory **list codes** are much more adequate than **ordinary codes** in so far as they make it possible to determine capacities for a large class of channels, where they are unknown for ordinary codes.
We show that this is also the case for combinatorial channel coding theory. In fact this is readily verified already for the standard model of $t$-error correcting codes.
For a constant $L$ define $C_q(\tau, L)$ as the supremal rate achievable for all large $n$ with list codes of list size $L$ and block length $n$ correcting $t = \tau n$ errors.

**Theorem 6** *For* $0 \leq \tau < \frac{q-1}{q}$

$$\sup_{L \in \mathbb{N}} C_q(\tau, L) = H_q(\tau).$$

We recall first a result on covering hypergraphs.

**Definition 1** *A covering* $\mathcal{C} = \{E_1, \ldots, E_k\}$ *of a hypergraph* $\mathcal{H} = (\mathcal{V}, \mathcal{E})$ *is called* $c$*–balanced for some constant* $c \in \mathbb{N}$ *if no vertex occurs in more than* $c$ *edges of* $\mathcal{C}$.

**Balanced Covering Lemma** [3] *A hypergraph* $(\mathcal{V}, \mathcal{E})$ *with maximum and minimum degrees* $d_{\max} = \max_{v \in \mathcal{V}} \deg(v)$ *and* $d_{\min} = \min_{v \in \mathcal{V}} \deg(v) > 0$ *has a* $c$*–balanced covering* $\mathcal{C} = \{E_1, \ldots, E_k\}$ *if*
*(a)* $k \geq |\mathcal{E}| \cdot d_{\min}^{-1} \cdot (\log_q |\mathcal{V}| + 1) + 1$,
*(b)* $c \leq k \leq c \cdot |\mathcal{E}| \cdot d_{\max}^{-1}$,
*(c)* $2^{-D(\lambda || \frac{d_{\max}}{|\mathcal{E}|}) \cdot k + \log_q |\mathcal{V}|} < \frac{1}{2}$ *for* $\lambda = \frac{c}{k}$,
*where* $D(P||Q)$ *denotes the Kullback/Leibler distance or relative entropy.*

We now focus our interest on balanced packings. Recall that a packing of a hypergraph $\mathcal{H} = (\mathcal{V}, \mathcal{E})$ is a subset of edges, such that every vertex is

contained in at most one edge. Accordingly, a $c$–balanced packing is a subset of edges in $\mathcal{H}$, such that every vertex is contained in at most $c$ edges.
As every code $\{(u_i, \mathcal{D}_i) : i = 1, \ldots, N, u_i \in Q^n, \mathcal{D}_i \subset \mathcal{Y}^n, \mathcal{D}_i \cap \mathcal{D}_j = \emptyset \text{ for } i \neq j\}$ yields the packing $\{\mathcal{D}_i : i = 1, \ldots, N\}$ of $\mathcal{X}^n$, every list code with $\sum_{i=1}^{N} 1_{\mathcal{D}_i(y^n)} \leq c$ for all $y^n \in Y^n$ corresponds to a $c$–balanced packing. We make use of the following result.

**Packing Lemma** [3] *A hypergraph $\mathcal{H} = (\mathcal{V}, \mathcal{E})$ has a $c$–balanced packing with $k$ edges if (b) and (c) of the Balanced Covering Lemma hold.*

Generally speaking, coverings are easier to handle than packings, because overlap is allowed. On the other hand, $c$–balanced packings are easier to handle than $c$–balanced coverings, since it is not required that $\mathcal{V}$ is covered. This has the effect that condition (a) in the Balanced Covering Lemma can be dropped ((a) just guarantees the existence of a covering), (b) and (c) are proven using the old arguments.
Observe that in typical applications in Information Theory $|\mathcal{V}|$ depends exponentially on the block length $n$ and thus $c$ has to grow with the block length in the Balanced Covering Lemma. Since for $c$–balanced packings condition (a) is no longer required, constant $c$'s are not automatically excluded. That is here the case. The theorem follows from

**Proposition 1** *For $0 \leq \tau < \frac{q-1}{q}$ the rate $R < H_q(\tau)$ is achievable for list size $L = \left\lceil \frac{q}{H_q - R} \right\rceil + 1$.*

**Proof:** Consider the hypergraph $(\mathcal{V}, \mathcal{E}) = (Q^n, (B_{x^n}(n, \tau n))_{x^n \in Q^n})$, where $B_{x^n}(n, \tau n)$ is the ball in $\mathcal{Q}^n$ with radius $\tau n$ and center $x^n$. Write $B(n, \tau n)$ for the ball with center $\underline{0} = (0, 0, \ldots, 0)$.
Here $d_{\min} = d_{\max} = |B(n, \tau n)| = d$, $|\mathcal{C}| = q^n$, $|\mathcal{E}| = q^n$, and

$$\frac{\log_q |B(n, \tau n)|}{n} \to h_q(\tau) - \tau \log_q(q-1) \text{ as } n \to \infty.$$

By the assumptions on $R$ and $L = c$ condition (b) obviously holds for $k = 2^{Rn}$ and $n$ large. To verify (c) we derive an upper bound on the exponent there. We have to show that $-D(\frac{L}{k} || \frac{d}{|\mathcal{E}|}) \cdot k + \log_2 |\mathcal{V}| < -1$.
Evaluation of the relative entropy yields

$$\left[ \frac{L}{k} \left( \log_2 \frac{d}{|\mathcal{E}|} - \log_2 \frac{L}{k} \right) + \left( 1 - \frac{L}{k} \right) \left( \log_2 \left( 1 - \frac{d}{|\mathcal{E}|} \right) - \log_2 \left( 1 - \frac{L}{k} \right) \right) \right] \cdot k$$

$$+\log_2|\mathcal{V}| \le -L\log_2 L + L\log_2 k - k\left(1-\frac{L}{k}\right)\log_2\left(1-\frac{L}{k}\right) + L\log_q\frac{d}{|\mathcal{E}|} + \log_2|\mathcal{V}|,$$

because we have omitted the negative term $k\left(1-\frac{L}{k}\right)\log_2\left(1-\frac{d}{|\mathcal{E}|}\right)$ and use that $\log_2\frac{d}{|\mathcal{E}|} \le \log_q\frac{d}{|\mathcal{E}|}$, as $\frac{d}{|\mathcal{E}|} \le 1$.
Using now $k = 2^{Rn}$ and that $\log_q\frac{d}{|\mathcal{E}|} \to n(H_q(\tau)-1) = -n(1-h_q(\tau)+\tau\log_q(q-1)) = -n(h_q(\tau)-\tau\log_q(q-1)+2\tau\log_q(q-1)) = -nH_q(\tau) - 2t\log_q(q-1) \le -nH_q(\tau)$ as $n\to\infty$, for a $\delta>0$ so small that $L \ge \frac{\log_2 q}{H_q(\tau)-R-\delta}$ and $n > n_0(\delta)$ so that $\log_q\frac{d}{|\mathcal{E}|} \le -n(H_q(\tau)-\delta)$, we continue with

$$\le LnR - 2^{nR}(1-L2^{-nR})\log_2(1-L2^{-nR})$$

$$-nL(H_q(\tau)-\delta) + n\log_2 q - L\log_2 L. \tag{33}$$

Since $-(1-z)\log_2(1-z) \le 2z$ for small $z$ (because $\frac{d}{dz}[-(1-z)\log_2(1-z)] = 1+\log_2(1-z)$ and therefore this function has gradient 1 at $z=0$), we upper bound the expression in (33), using $z = L\cdot 2^{-nR}$, by

$$LnR + 2L - nL(H_q(\tau)-\delta) + n\log_2 q - L\log_2 L$$

$$= 2L - L\log_2 L - nL\left[H_q(\tau)-\delta-R-\frac{\log_2 q}{L}\right].$$

It suffices now to guarantee that the term in square brackets is positive or, equivalently, that

$$L \ge \frac{\log_2 q}{H_q(\tau)-R-\delta},$$

which is the case by the choice of $\delta$. □

Now with $m_0 \ge L$ we can always encode the true message on the list known to the decoder and known to the encoder, because of the feedback, and send it to the decoder. Thus we have the

**Corollary 1** *$C_q^{f,\Box}(\tau) \ge H_q(\tau)$ for all $0 \le \tau < \frac{q-1}{q}$.*

Now we prove the converse by adapting the proof of Lemma 1 to the case of side information. In our □-model the sender gives a message $S \in \{1,\dots,m_\Box\}$ with the number of letters at the end of the transmission of a message $m$. We can right away go to a more general model of side information $S \in \mathcal{S}$

given error free from the sender to the receiver at any time points of the transmission. This way the feedback can be linked with the side information. (Even more generally one can consider a model in which also the receiver can actively send a message $S' \in \mathcal{S}'$ to the sender.)
An encoding is now described by

$$[f_i^n, S_i^n] = \big[(f_{i_1}, S_{i_1}), (f_{i_2}(y^1), S_{i_2}(y^1)), \ldots, (f_{i_i}(y^{i-1}), S_{i_i}(y^{i-1})), \\ \ldots, f_{i_n}(y^{n-1}), S_{i_n}(y^{n-1})\big].$$

For the family of encoding functions with side information $F_{\mathcal{M}} = \{(f_i^n, S_i^n) : i \in \mathcal{M}\}$ and a set $\mathcal{V} \subset Q^n$ of error patterns we write $\Phi(F_{\mathcal{M}}, \mathcal{V}) = \{(y^n, S_i^n)$: there exist $i \in \mathcal{M}$ and $e^n \in \mathcal{V}$ such that $y^n = \phi(f_i^n, e^n)\}$.
In analogy to (5) we have now

$$(\phi(f_i^n, e^n), S_i^n) \neq (\phi(f_{i'}^n, e'^n), S_{i'}^n) \text{ for } (i, e^n, S_m^n) \neq (i', e'^n, S_i'^n).$$

This is the case if $i \neq i'$; $i = i'$ and $e^n \neq e'^n$; and if $i = i'$, $e^n = e'^n$, and $S_i^n$. Therefore

$$|\Phi(F_{\mathcal{M}}, \mathcal{V})| = M|\mathcal{V}| \cdot m_{\Box} = M \sum_{j=0}^{t} \binom{n}{j} (q-1)^j m_{\Box} \leq q^n m_{\Box}$$

and since $m_{\Box}(n)$ has rate 0 asymptotically we get

$$C_q^{f,\Box}(\tau) \leq H_q(\tau), 0 \leq \tau \leq 1.$$

This and Corollary 1 yield the

**Theorem 7** *$C_q^{f,\Box}(\tau) = H_q(\tau)$ for all $0 \leq \tau \leq 1$.*

# 5 The $\Box$-model with localized errors

We follow a key idea of [5] to provide the decoder **some protocol information** about the a priori error pattern $E$. For this we need as auxiliary result the

**Covering Lemma [3]** *For a hypergraph $\mathcal{H} = (\mathcal{V}, \mathcal{E})$ there is a covering $\mathcal{C}$, $\mathcal{C} \subset \mathcal{E}$ of the vertex set $\mathcal{V}$ with*

$$|\mathcal{C}| \leq \lceil |\mathcal{E}| d^{-1} \log_2 |\mathcal{V}| \rceil,$$

*where*

$$d = \min_{v \in \mathcal{V}} |\{E \in \mathcal{E} : v \in E\}|.$$

**Corollary 2** *Let $t < l < n$ be positive integers. For the hypergraph*

$$\mathcal{H} = \left( \binom{[n]}{t}, \binom{[n]}{l} \right)$$

*there is a covering $\mathcal{C}_l \subset \binom{[n]}{l}$ with*

$$|\mathcal{C}_l| \leq \binom{n}{t} \binom{l}{t}^{-1} \cdot n.$$

**Proof:** Since

$$|\mathcal{E}| d^{-1} \log_2 |\mathcal{V}| = \binom{n}{l} \binom{n-t}{l-t}^{-1} \log_2 \binom{n}{t} \leq \binom{n}{t} \binom{l}{t}^{-1} n,$$

the result follows from the Covering Lemma.
The guiding idea in deriving a lower bound on $C_q^{l,\square}(\tau)$ is based on the following calculation for the "useful information". Choose a function $g : \mathcal{E}_t \to \mathcal{C}_l$ with the property $g(E) \supset E$. The encoder, knowing $E$, also knows $g(E)$. Now, if the decoder would also know $g(E)$, then the communicators could transmit $M = q^{n-1}$ messages. However, since $g(E)$ is not known to the decoder, $|\mathcal{C}_l|$ of these messages must be reserved for the "protocol" and there are only

$$M |\mathcal{C}_l|^{-1} \geq q^{n-l} \binom{l}{t} \binom{n}{t}^{-1} n^{-1}$$

"useful messages". An elementary calculation shows that this expression attains its maximum for $l = \frac{q}{q-1} t$. Since

$$q^{\frac{q}{q-1} t} \approx (q-1)^t \binom{qt/(q-1)}{t},$$

its value is

$$q^{n-l} \binom{l}{t} \binom{n}{t}^{-1} n^{-1} = \frac{q^n \binom{qt/(q-1)}{t}}{n q^{qt/(q-1)} \binom{n}{t}} \approx \frac{q^n}{(q-1)^t \binom{n}{t}} \tag{34}$$

and (in rate) corresponds to the Hamming bound.
*How can the information be coded?*
(1) Write the block length $n$ in the form

$$n = m_\square + m = m_\square + m_1 + \cdots + m_r,$$

where

$$m_i = \lfloor \frac{m}{r} \rfloor \quad \text{or} \quad \lceil \frac{m}{r} \rceil, \ i = 1, \ldots, r$$

and $m_\square$ and $r$ are specified later. Furthermore define for $i \geq 1$

$$B_i = \left[ \sum_{j=1}^{i-1} m_j + 1, \sum_{j=1}^{i} m_j \right].$$

Set $F_i = B_i \cap F$. The encoder, knowing $F$, knows also the sets $F_i$ and he orders the intervals $B_i$ $(i = 1, \ldots, r)$ as $B_{i_1}, \ldots, B_{i_r}$ according to increasing cardinalities $t_i = |F_i|$ and, in cases of ties, according to increasing $i$'s.

$$t_{i_1} \leq t_{i_2} \leq \cdots \leq t_{i_r}.$$

(a) The $m_\square$ positions can be used to vary the lengths for being able to send $m_\square$ secure messages at the end! They are used to inform the decoder about the order defined above:

$$r! \leq m_\square \tag{35}$$

(b) For $s \in \mathbb{N}$ determined below we consider the first intervals $B_{i_1}, \ldots, B_{i_s}$. Clearly $B_{i_1} \cup \ldots \cup B_{i_s}$ has an error frequency $\leq \tau$ and a cardinality

$$n_s \sim \frac{m}{r} s. \tag{36}$$

Using for transmission on this block only letters 0 and 1 by the result of [8] – in the notation of [6] the maximal number $M_{AA}$ of messages transmittable satisfies

$$\log_2 M_{AA}(n_s, \tau) \geq n_s \cdot (1 - h_2(\tau)) \frac{1}{2}. \tag{37}$$

We use them to inform the decoder about the values $t_{i_{s+1}}, \ldots, t_{i_r}$. This requires at most $t^r$ messages. Furthermore, this code is used to inform the decoder about

$$E_{i_{s+1}} \in \binom{B_{i_{s+1}}}{t_{i_{s+1}}}.$$

Clearly, a total of

$$M_1 = t^r \cdot 2^{\lceil \frac{m}{r} \rceil}$$

messages suffices for these purposes. Therefore

$$\log_2 M_1 = r \log_2 t + \lceil \frac{m}{r} \rceil \leq r \log_2 n + \frac{n}{r}. \tag{38}$$

Now by (35)–(38) we must have

$$\frac{ms}{2r}(1 - h_2(\tau)) \geq r \log_2 n + \frac{n}{r}$$

and since $r$ is **constant**

$$s \geq \frac{4n}{m} \frac{1}{1-h_2(\tau)}$$

and since $m \geq n/2$, it suffices to have

$$s \geq \frac{8}{1-h_2(\tau)} \tag{39}$$

From block length $n$ we used in (a) and (b) for protocol information

$$r! + \frac{n-m_\square}{r} \frac{8}{1-h_2(\tau)} = r! + \frac{n-r!}{r} \frac{8}{1-h_2(\tau)}$$

positions for $m_\square = r!$, a constant. For any $\delta > 0$ there is an $n_0(\delta)$ and an $r(\delta)$ such that the loss is $\leq \delta n$.
(c) Apply Corollary 2 to each interval

$$B_{i_{s+1}}, B_{i_{s+2}}, \ldots, B_{i_r}.$$

In interval $B_0$ the decoder was informed about $g_1(E_{i_{s+1}}) \subset B_{i_{s+1}}$. In the positions

$$\bigcup_{j=s+1}^{r} [B_{i_j} \setminus g_j(E_{i_j})]$$

the decoder will be informed successively about $g_2(E_{i_{s+2}})$, $g_2(E_{i_{s+3}})$, ... Since the cardinalities

$$l_{i_j} = \frac{q}{q-1} t_{i_j}$$

increase, this is possible. The information about $g_j(E_{i_j})$ is given before we start in $B_{i_j} \setminus g_j(E_{i_j})$. After the total protocol information is conveyed, the decoder will get the useful information in the remaining free positions. The attainable number of useful messages exceeds

$$\prod_{j=s+1}^{r} \frac{q^{m_{i_j}}}{(q-1)^{t_{i_j}} \binom{m_{i_j}}{t_{i_j}}} \geq \frac{q^{n-m_0}}{(q-1)^{t-t_0} \binom{n-m_0}{t-t_0}},$$

because as in (34)

$$q^{m_{i_j}-l_{i_j}} \binom{l_{i_j}}{t_{i_j}} \binom{m_{i_j}}{t_{i_j}}^{-1} m_{i_j}^{-1} \approx \frac{q^{m_{i_j}}}{(q-1)^{t_{i_j}} \binom{m_{i_j}}{t_{i_j}}}$$

and

$$\binom{a+b}{c+d} \geq \binom{a}{c} \cdot \binom{b}{d}.$$

We have proved

**Theorem 8** *In the presence of localized errors for* $0 \leq \tau < 1/2$

$$C_q^{l,\Box}(\tau) = H_q(\tau).$$

# 6 Localized errors with side information – a generalization of the $m_\Box$ model

Suppose that the sender can forward to the receiver one of $M$ messages, where $M$ depends on $n$. We analyze the help of the functions

$$M(n) = n^\alpha,\ \alpha > 0. \tag{40}$$

In particular we show first that for suitable exponent $\alpha = \alpha(\tau)$ the Hamming bound can be achieved for all $0 \leq \tau \leq 1$. Thus polynomial growth suffices – a rate-wise negligible side information.
Following the approach in Section 5, which starts with the partition of $[n]$ into an ordered sequence of intervals $B_{i_1}, \ldots, B_{i_r}$, we have now the problem that we don't have $\tau < 1/2$ and therefore cannot use Theorem BGP. Consequently a constant $m_\Box = r!$ no longer suffices. In fact we have to use very small blocks so that the side information alone gives us enough protocol information to start in the first block $B_{i_1}$. We need here the relatively large

$$r = \beta \frac{n}{\log_q n}, \tag{41}$$

which in turn makes $r!$ very large. We therefore confine ourselves to inform with our side information the receiver first only about $i_1$ and then use the blocks $B_{i_2}, B_{i_3}, \ldots$ iteratively to inform about the next.
Within the blocks we cover the $t/r$ element sets by $l/r$ element sets again, where $l = \frac{q}{q-1}t$. Also the block wise iteration proceeds as before with one additional requirement that in $B_{i_j}$ we have to inform the receiver also about its successor $B_{i_{j+1}}$. This additional protocol information must be provided by the covering and rather precise estimates are necessary. Set

$$n' = \frac{n}{r},\quad t' = \frac{t}{r},\quad l' = \frac{l}{r} = \frac{q}{q-1}\frac{t}{r} \text{ and } k = n' - l'$$

for the number of information positions in a block. Applying the corollary to the hypergraph $\mathcal{H}' = \left(\binom{[n']}{t'}, \binom{[n']}{l'}\right)$ we get a covering $\varphi_{l'}$ with

$$|\varphi_{l'}| \leq \binom{n'}{t'}\binom{l'}{t'}^{-1} \log_2 \binom{n'}{t'}$$

$$\leq \frac{n'(n'-1)\cdots(n'-k+1)}{(n'-t')(n'-t'-1)\cdots(n'-t'-k+1)} \cdot \frac{n}{r} \tag{42}$$

$$\leq \left(\frac{n'-k}{n'-t'-k}\right)^k \frac{n}{r}. \tag{43}$$

Using $r$ possibilities for transmitting $i_1$ we are left with $\frac{n^\alpha}{r}$ side information and it suffices to insure

$$\left(\frac{n'-k}{n'-t'-k}\right)^k \frac{n}{r} \leq \frac{n^\alpha}{r}, \tag{44}$$

and later on

$$\frac{n'(n'-1)\cdots(n'-k+1)}{(n'-t')(n'-t'-1)\cdots(n'-t'-k+1)} \cdot \frac{n}{r} \cdot r \leq q^k. \tag{45}$$

Since $\log_q(1+x) \leq x$, it suffices to have for (44)

$$k \log_q \left(1 + \frac{t'}{n'-t'-k}\right) \leq (\alpha-1)\log_q n \quad \text{or}$$

$$k \frac{t}{n-t-kr} \leq (\alpha-1)\log_q n \quad \text{or}$$

$$k \leq \frac{1-\tau-\left(1-\frac{q}{q-1}\tau\right)}{\tau}(\alpha-1)\log_q n \quad \text{or}$$

$$k \leq \frac{\alpha-1}{q-1}\log_q n. \tag{46}$$

(Sufficient for (45) would be

$$\left(\frac{n'-k}{n'-t'-k}\right)^k n = \left(\frac{l'}{l'-t'}\right)^k n = \left(\frac{l}{l-t}\right)^k n = \left(\frac{qt(q-1)}{(q-1)t}\right)^k n = q^k n \leq q^k,$$

which is impossible!) We have to estimate $|\varphi_{l'}|$ in (42) from above more precisely. We use the following

**Lemma 4** *If for positive integers $A, B, a, b$ we have $A > B$, $a > b$, $A > a$, $B > b$ and*

$$A - a = B - b,$$

*then*

$$\frac{A}{B}\frac{a}{b} > \frac{A-1}{B-1}\frac{a+1}{b+1}.$$

**Proof:** The claim is equivalent with

$$ABab + Aa(B - b - 1) > ABab + Bb(A - a - 1),$$

which holds, because $Aa > Bb$ and $B - b - 1 = A - a - 1$. This lemma implies that

$$\frac{n'(n'-1)\cdots(n'-k+1)}{(n'-t')(n'-t'-1)\cdots(n'-t'-k+1)}n$$

$$\leq \left(\frac{n'(n'-k+1)}{(n'-t')(n'-t'-k+1)}\right)^{k/2} n \leq q^{k/2}n \cdot \left(\frac{1}{1-\tau}\right)^{k/2} = \left(\frac{q}{1-\tau}\right)^{k/2} n$$

Since $H_q(\tau) = 0$ for $\tau \geq \frac{q-1}{q}$, it suffices to consider $\tau = \frac{q-1}{q} - \varepsilon$ and show that

$$\left(\frac{q}{1-\tau}\right)^{k/2} n \leq q^k \quad \text{or} \quad \left(\frac{q}{\frac{1}{q}+\varepsilon}\right)^{k/2} n \leq q^k \quad \text{or}$$

$$\left(\frac{1}{1+\varepsilon q}\right)^{k/2} n \leq 1 \quad \text{or} \quad \log_q n - \frac{k}{2}\log_q(1+\varepsilon q) \leq 0 \quad \text{or}$$

$$k \geq \frac{\log_q n}{\log_q(1+\varepsilon q)}. \tag{47}$$

Now $kr = n - l = n(1 - \frac{q}{q-1}(\frac{q-1}{q} - \varepsilon)) = n(1 + \frac{q\varepsilon}{q-1})$ and $k = \frac{\log_q n}{\beta}(1 + \frac{q\varepsilon}{q-1})$ and (47) holds if we choose

$$\beta = \frac{\log_q(1+q\varepsilon)}{1+\frac{q\varepsilon}{q-1}}, \text{ where } \varepsilon = \frac{q-1}{q} - \tau.$$

Finally, we have obtained from (46) and (47) that

$$\frac{\log_q n}{\log_q(1+\varepsilon q)} \leq \frac{\alpha - 1}{q-1}\log_q n. \tag{48}$$

We summarize our findings.

**Theorem 9** *For $\tau < \frac{q-1}{q}$ and $\alpha(\tau) = 1 + \frac{q-1}{\log_q q(1-\tau)}$ the polynomial side information $n^{\alpha(\tau)}$ gives for $t = \tau n$ localized errors the Hamming bound as capacity curve.*

For $\tau \geq \frac{q-1}{q}$ this is obvious, because $H_q(\tau) = 0$ and $H_q$ is an upper bound for the capacity curve.

# References


[1] R. Ahlswede, A constructive proof of the coding theorem for discrete memoryless channels in case of complete feedback, Sixth Prague Conf. on Inf. Th., Stat. Dec. Fct's and Rand. Proc., Sept. 1971, Publ. House Czechosl. Academy of Sc., 1–22, 1973.

[2] R. Ahlswede, Channel capacities for list codes, J. Appl. Probability, 10, 824–836, 1973.

[3] R. Ahlswede, Coloring hypergraphs: A new approach to multi–user source coding I, J. of Combinatorics, Information and System Sciences, Vol. 4, No. 1, 76–115, 1979;
Coloring hypergraphs: A new approach to multi-user source coding II, J. of Combinatorics, Information and System Sciences, Vol. 5, No. 3, 220-268, 1980.

[4] R. Ahlswede, L.A. Bassalygo, and M.S. Pinsker, Nonbinary codes correcting localized errors, IEEE Trans. Inf. Theory, Vol. 44, No. 4, 1413-1416, 1993.

[5] R. Ahlswede, L.A. Bassalygo, and M.S. Pinsker, Localized random and arbitrary errors in the light of AV channel theory, IEEE Trans. Inf. Theory, Vol. 41, No. 1, 14–25, 1995.

[6] R. Ahlswede, L.A. Bassalygo, and M.S. Pinsker On the Hamming bound for nonbinary localized error–correcting codes, Problems Inform. Transmission 35, No. 2, 117–124, 1999.

[7] M. Aigner, Searching with lies, J. Combinatorial Theory, Ser.A 74, 43–56, 1996.

[8] L.A. Bassalygo, S.I. Gelfand, and M.S. Pinsker, Coding for channels with localized errors, Proc. Fourth Soviet–Swedish Workshop in Information Theory, Gotland, Sweden, pp. 95–99, 1989.

[9] L.A. Bassalygo, S.I. Gelfand, and M.S. Pinsker, Bounds for codes with separately localized errors, Problems Inform. Transmission 28, No. 1, 11–17, 1992.

[10] C.R. Berlekamp, Block Coding with Noiseless Feedback, Doctoral Dissertation, MIT, 1964.

[11] C.R. Berlekamp, Block coding for the binary symmetric channel with noiseless delayless feedback, Proc. Symposium on Error Correcting Codes, Univ. Visconsin, 1968.

[12] R.A. Brualdi, Introductory Combinatorics, Elsevier North-Holland, 1977.

[13] A. Dyachkov, Upper bounds for the probability of error in transmission with feedback for discrete memoryless channels, Problems Inform. Transmission 11, No. 4, 271–283, 1975.

[14] D.N. Gevorkyan, G.A. Kabatyanskii, On the Varshamov-Tenengolts codes and a conjecture of Bassalygo, Problems Inform. Transmission 28, No. 4, 393–395, 1993.

[15] G.A. Kabatyanskii, Construction of nonbinary codes that correct single localized errors, Problems Inform. Transmission 30, No. 2, 175–176, 1994.

[16] F.J. MacWilliams and N.J. Sloane, The Theory of Error-Correcting Codes, North-Holland, Amsterdam, 1977.

[17] A. Malinowski, K-ary searching with a lie, ARS Combin., Vol. 37, 301-308, 1994.

[18] A. Pelc, Solution of Ulam's problem on searching with a lie, J. Combin. Theory Ser. A, Vol. 44, 129-140, 1987.

[19] A. Rényi, On a problem of information theory, MTA Mat. Kut. Int. Kozl. 6B, 505-516, 1961.

[20] S.M. Ulam, Adventures of a Mathematician, Charles Scribner's Sons, New York, 1976.

[21] K.Sh. Zigangirov, Number of correctable errors for transmission over a binary symmetrical channel with feedback, Problems Inform. Transmission 12, No. 2, 85-97, 1976.