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

R. Ahlswede, C. Deppe, and V. Lebedev

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 [**?**], 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 τ and all large n, is a central concept:

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

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

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

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 $\{u_i(E) : \ 1 \le i \le M, \ E \in \mathcal{E}_t\}$ 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 [4] for the binary case to equal $H_2(\tau)$. For general q the best known result is

Dagstuhl Seminar Proceedings 06201
Combinatorial and Algorithmic Foundations of Pattern and Association Discover
http://drops.dagstuhl.de/opus/volltexte/2006/784

Theorem Ahlswede/Bassalygo/Pinsker (ABP, [3])

(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 above having as ingredients feedback resp. localized errors give possibilities for code constructions not available in the standard model of error correction and also for probabilistic channel models ([**?**], [**?**]).
For the feedback model we present here a coding scheme, which we call the rubber method, because it is based on erasing letters. **It is the first scheme achieving the capacity curve for** $q \geq 3$**.** It could be discovered only in the q-ary case for $q \geq 3$, because the letter zero is not used as an information symbol, but solely for error correction. However an extension of the method from using single zeros to blocks of zeros also gives Berlekamp's result - by a different scheme.

Lemma 1

(i) $M \sum_{j=0}^{t} \binom{n}{j}(q-1)^j \leq q^n$ *for every* $(M, n, t, q)_f$ *code.*

(ii) $C_q^f(\tau) \leq H_q(\tau)$ *for* $0 \leq \tau \leq 1$.

Lemma 2 $C_q^f(\tau) \leq (1-2\tau)\log_q(q-1)$ *for* $\frac{1}{q} \leq \tau \leq \frac{1}{2}$.

Theorem 1 $C_q^f(\tau) \geq (1-2\tau)\log_q(q-1)$ *for* $\tau = \frac{t}{n}$ *and* $0 < \tau < \frac{1}{2}$.

Theorem 2 $C_q(\tau) = (1-2\tau)\log_q(q-1)$ *for* $\frac{1}{q} \leq \tau \leq \frac{1}{2}$.

Theorem 3 *The rate functions obtained by our strategies are tangents to* $H_q(\tau)$ *going through* $\frac{1}{r+1}$ *for all* $r \geq 1$.

In the 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.

Theorem 4

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

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

Theorem 5 $M_{fl}(n,1) = \lfloor \frac{2^n}{n+1} \rfloor$.

Theorem 6 $M_{fl}(n,1,q) = \lfloor \frac{q^n}{(q-1)n+1} \rfloor$ *if* $n \geq q+1$.

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 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 7 $\sup_{L \in \mathbb{N}} C_q(\tau, L) = H_q(\tau)$ *for* $0 \leq \tau \leq 1$.

Corollary 1 $C_q^{f,\square}(\tau) \geq H_q(\tau)$ *for all* $0 \leq \tau \leq 1$.

Theorem 8 $C_q^{f,\square}(\tau) = H_q(\tau)$ *for all* $0 \leq \tau \leq 1$.

Theorem 9 $C_q^{l,\square}(\tau) = H_q(\tau)$ *for* $0 \leq \tau < 1/2$

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.

However, there is a function $\alpha_q : \left[\frac{1}{2}, \frac{q-1}{q}\right) \to \mathbb{N}$ such that with $n^{\alpha_q(\tau)}$ (that is, polynomially many and thus ratewise zero) additional messages as "protocol" information $H_q(\tau)$ is achievable for every $\tau \in \left[\frac{1}{2}, \frac{q-1}{q}\right)$. For the presently best function

$$\lim_{\tau \to \frac{q-1}{q}} \alpha_q(\tau) = \infty.$$

Theorem 10 *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.*

Also notice that without the marker $\square$ in the range $0 \leq \tau < 1/2$ with feedback the capacity curve is **smaller** than for localized errors. This can be seen already in the case $q = 2$ by comparing the result of [4] and [5]. For general q Theorem 1 and more detailed results of [3] confirm this phenomenon.

Remark. A search model with lies equivalent to Berlekamp's feedback model was first formulated by Rényi [**?**] and rediscovered by Ulam [**?**], who raised great interest in the subject and to whom many papers in search theory refer. We find the priorities most accurately reflected in the name Rényi / Berlekamp / Ulam model and therefore suggest to use it.