Chapter 1
Bounds on Error-Correction Coding Performance

1.1 Gallager's Coding Theorem

The sphere packing bound by Shannon [18] provides a lower bound to the frame error rate (FER) achievable by an (n, k, d) code but is not directly applicable to binary codes. Gallager [4] presented his coding theorem for the average FER for the ensemble of all random binary (n, k, d) codes. There are 2^n possible binary combinations for each codeword which in terms of the n-dimensional signal space hypercube corresponds to one vertex taken from 2^n possible vertices. There are 2^k codewords, and therefore 2^{nk} different possible random codes. The receiver is considered to be composed of 2^k matched filters, one for each codeword and a decoder error occurs if any of the matched filter receivers has a larger output than the matched filter receiver corresponding to the transmitted codeword. Consider this matched filter receiver and another different matched filter receiver, and assume that the two codewords differ in d bit positions. The Hamming distance between the two codewords is d. The energy per transmitted bit is $E_s = \frac{k}{n}E_b$, where E_b is the energy per information bit. The noise variance per matched filtered received bit, $\sigma^2 = \frac{N_0}{2}$, where N_0 is the single sided noise spectral density. In the absence of noise, the output of the matched filter receiver for the transmitted codeword is $n\sqrt{E_s}$ and the output of the other codeword matched filter receiver is $(n - 2d)\sqrt{E_s}$. The noise voltage at the output of the matched filter receiver for the transmitted codeword is denoted as $n_c - n_1$, and the noise voltage at the output of the other matched filter receiver will be $n_c + n_1$. The common noise voltage n_c arises from correlation of the bits common to both codewords with the received noise and the noise voltages $-n_1$ and n_1 arise, respectively, from correlation of the other d bits with the received noise. A decoder error occurs if

$$(n - 2d)\sqrt{E_s} + n_c + n_1 > n\sqrt{E_s} + n_c - n_1 \tag{1.1}$$

that is, a decoder error occurs when $2n_1 > 2d\sqrt{E_s}$.

M. Tomlinson et al., *Error-Correction Coding and Decoding*,
Signals and Communication Technology,
DOI 10.1007/978-3-319-51103-0_1

The average noise power associated with n_1 is $\mathrm{d}\sigma^2 = \mathrm{d}\frac{N_0}{2}$ and as the noise is Gaussian distributed, the probability of decoder error, p_d, is given by

$$p_d = \frac{1}{\sqrt{\pi d N_0}} \int_{\mathrm{d}\sqrt{E_s}}^{\infty} \mathrm{e}^{\frac{-x^2}{dN_0}} \mathrm{d}x \tag{1.2}$$

This may be expressed in terms of the complementary error function (erfc)

$$\mathrm{erfc}(y) = 2\frac{1}{\sqrt{2\pi}} \int_{y}^{\infty} \mathrm{e}^{\frac{-x^2}{2}} \mathrm{d}x \tag{1.3}$$

and

$$p_d = \frac{1}{2}\mathrm{erfc}\left(\sqrt{d\frac{k}{n}\frac{E_b}{N_0}}\right) \tag{1.4}$$

Each of the other $2^k - 2$ codewords may also cause a decoder error but the weight distribution of the code $\mathscr{C}_i$ is usually unknown. However by averaging over all possible random codes, knowledge of the weight distribution of a particular code is not required. The probability of two codewords of a randomly chosen code $\mathscr{C}_i$, differing in d bit positions, $p(d|\mathscr{C}_i)$ is given by the binomial distribution

$$p(d|\mathscr{C}_i) = \frac{\binom{n}{d}}{2^n}, \tag{1.5}$$

where $\binom{a}{b} = \frac{a!}{(a-b)!b!}$. A given linear code $\mathscr{C}_i$ cannot have codewords of arbitrary weight, because the sum of a subset of codewords is also a codeword. However, for non linear codes, p_d may be averaged over all of the codes without this constraint. Thus, we have

$$\overline{p_C} = \sum_{i=1}^{2^{n2^k}} p(d|\mathscr{C}_i)p(\mathscr{C}_i) < \frac{1}{2^{n2^k}} \sum_{d=0}^{n} \sum_{i=1}^{2^{n2^k}} \frac{\binom{n}{d}}{2^{n+1}} \mathrm{erfc}\left(\sqrt{d\frac{k}{n}\frac{E_b}{N_0}}\right) \tag{1.6}$$

Rearranging the order of summation

$$\overline{p_C} < \frac{1}{2^{n2^k}} \sum_{i=1}^{2^{n2^k}} \sum_{d=0}^{n} \frac{\binom{n}{d}}{2^{n+1}} \mathrm{erfc}\left(\sqrt{d\frac{k}{n}\frac{E_b}{N_0}}\right) \tag{1.7}$$

and

$$\overline{p_C} < \frac{1}{2^{n+1}} \sum_{d=0}^{n} \binom{n}{d} \mathrm{erfc}\left(\sqrt{d\frac{k}{n}\frac{E_b}{N_0}}\right). \tag{1.8}$$

Remembering that any of the $2^k - 1$ matched filters may cause a decoder error, the overall probability of decoder error averaged over all possible binary codes $\overline{p_{\text{overall}}}$, is

$$\overline{p_{\text{overall}}} = 1 - (1 - \overline{p_{\text{C}}})^{2^k - 1} < 2^k \overline{p_{\text{C}}} \tag{1.9}$$

and

$$\overline{p_{\text{overall}}} < \frac{2^k}{2^{n+1}} \sum_{d=0}^{n} \binom{n}{d} \text{erfc}\left(\sqrt{d \frac{k}{n} \frac{E_b}{N_0}}\right). \tag{1.10}$$

An analytic solution may be obtained by observing that $\frac{1}{2}\text{erfc}(y)$ is upper bounded by e^{-y^2} and therefore,

$$\overline{p_{\text{overall}}} < \frac{2^k}{2^n} \sum_{d=0}^{n} \binom{n}{d} \text{e}^{-d \frac{k}{n} \frac{E_b}{N_0}} \tag{1.11}$$

and as observed in [21],

$$\left(1 + \text{e}^{-\frac{k}{n} \frac{E_b}{N_0}}\right)^n = \sum_{d=0}^{n} \binom{n}{d} \text{e}^{-d \frac{k}{n} \frac{E_b}{N_0}} \tag{1.12}$$

and

$$\overline{p_{\text{C}}} < \frac{1}{2^n} \left(1 + \text{e}^{-\frac{k}{n} \frac{E_b}{N_0}}\right)^n \tag{1.13}$$

$$\overline{p_{\text{overall}}} < \frac{2^k}{2^n} \left(1 + \text{e}^{-\frac{k}{n} \frac{E_b}{N_0}}\right)^n \tag{1.14}$$

Traditionally, a cut-off rate R_0 is defined after observing that

$$\frac{2^k}{2^n} \left(1 + \text{e}^{-\frac{k}{n} \frac{E_b}{N_0}}\right)^n = 2^k \left(\frac{1 + \text{e}^{-\frac{k}{n} \frac{E_b}{N_0}}}{2}\right)^n \tag{1.15}$$

with

$$2^{R_0} = \frac{2}{1 + \text{e}^{-\frac{k}{n} \frac{E_b}{N_0}}} \tag{1.16}$$

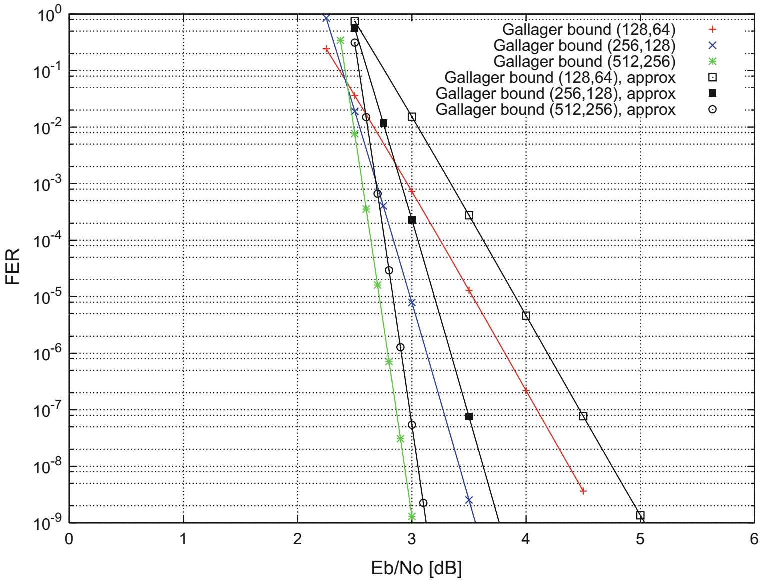

Fig. 1.1 Approximate and exact Gallager bounds for $(128, 2^{64})$, $(256, 2^{128})$ and $(512, 2^{256})$ nonlinear binary codes

then

$$\overline{p_{\text{overall}}} < 2^{k}2^{-nR_0} = 2^{k-nR_0} = 2^{-n(R_0-\frac{k}{n})} \tag{1.17}$$

This result may be interpreted as providing the number of information bits of the code is less than the length of the code times the cut-off rate, then the probability of decoder error will approach zero as the length of the code approaches infinity. Alternatively, provided the rate of the code, $\frac{k}{n}$, is less than the cut-off rate, R_0, then the probability of decoder error will approach zero as the length of the code approaches infinity. The cut-off rate R_0, particularly in the period from the late 1950s to the 1970s was used as a practical measure of the code rate of an achievable error-correction system [11, 20–22]. However, plotting the exact expression for probability of decoder error, Eq. (1.10), in comparison to the cut-off rate approximation Eq. (1.17), shows a significant difference in performance, as shown in Fig. 1.1. The codes shown are the $(128, 2^{64})$, $(256, 2^{128})$ and $(512, 2^{256})$ code ensembles of nonlinear, random binary codes. It is recommended that the exact expression, Eq. (1.10) be evaluated unless the code in question is a long code. As a consequence, in the following sections we shall only use the exact Gallager bound.

Shown in Fig. 1.2 is the sphere packing lower bound, offset by the loss attributable to binary transmission and the Gallager upper bound for the $(128, 2^{64})$, $(256, 2^{128})$ and $(512, 2^{256})$ nonlinear binary codes. For each code, the exact Gallager upper bound given by (1.10), is shown. One reason why Gallager's bound is some way

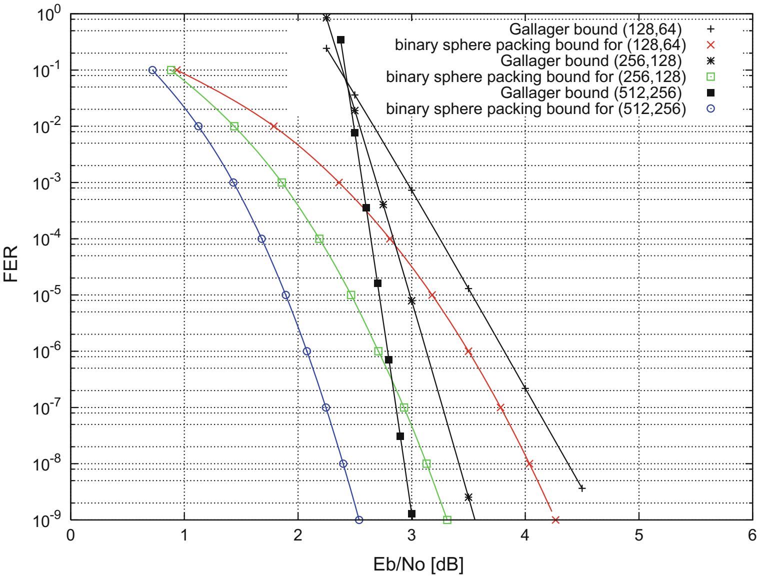

Fig. 1.2 Sphere packing and Gallager bounds for $(128, 2^{64})$, $(256, 2^{128})$ and $(512, 2^{256})$ nonlinear binary codes

from the sphere packing lower bound as shown in Fig. 1.2 is that the bound is based on the union bound and counts all error events as if these are independent. Except for orthogonal codes, this produces increasing inaccuracy as the $\frac{E_b}{N_0}$ is reduced. Equivalently expressed, double counting is taking place since some codewords include the support of other codewords. It is shown in the next section that for linear codes the Gallager bound may be improved by considering the erasure correcting capability of codes, viz. no (n, k) code can correct more than $n - k$ erasures.

1.1.1 Linear Codes with a Binomial Weight Distribution

The weight enumerator polynomial of a code is defined as $A(z)$ which is given by

$$A(z) = \sum_{i=0}^{n} A_i \; z^i \tag{1.18}$$

For many good and exceptional, linear, binary codes including algebraic and quasi-cyclic codes, the weight distributions of the codes closely approximates to a binomial distribution where,

$$A(z) = \frac{1}{2^{n-k}} \sum_{i=0}^{n} \frac{n!}{(n-i)!i!}\, z^i \tag{1.19}$$

with coefficients A_i given by

$$A_i = \frac{1}{2^{n-k}} \frac{n!}{(n-i)!i!} = \frac{1}{2^{n-k}} \binom{n}{i}. \tag{1.20}$$

Tables of the best-known linear codes have been published from time to time [3, 10, 13, 16, 19] and a regularly updated database is maintained by Markus Grassl [5]. Remembering that for a linear code, the difference between any two codewords is also a codeword, and hence the distribution of the Hamming distances between a codeword and all other codewords is the same as the weight distribution of the code. Accordingly, the overall probability of decoder error, for the same system as before using a bank of 2^k matched filters with each filter matched to a codeword is upper bounded by

$$p_{\text{overall}} < \frac{1}{2} \sum_{d=0}^{n} A_d \text{erfc}\left(\sqrt{d \frac{k}{n} \frac{E_b}{N_0}}\right) \tag{1.21}$$

For codes having a binomial weight distribution

$$p_{\text{overall}} < \frac{1}{2} \sum_{d=0}^{n} \frac{1}{2^{n-k}} \binom{n}{d} \text{erfc}\left(\sqrt{d \frac{k}{n} \frac{E_b}{N_0}}\right) \tag{1.22}$$

which becomes

$$p_{\text{overall}} < \frac{2^k}{2^{n+1}} \sum_{d=0}^{n} \binom{n}{d} \text{erfc}\left(\sqrt{d \frac{k}{n} \frac{E_b}{N_0}}\right). \tag{1.23}$$

It will be noticed that this equation is identical to Eq. (1.10). This leads to the somewhat surprising conclusion that the decoder error probability performance of some of the best-known, linear, binary codes is the same as the average performance of the ensemble of all randomly chosen, binary nonlinear codes having the same values for n and k. Moreover, some of the nonlinear codes must have better performance than their average, and hence some nonlinear codes must be better than the best-known linear codes.

A tighter upper bound than the Gallager bound may be obtained by considering the erasure correcting capability of the code. It is shown in Chap. 14 that for the erasure channel, given a probability of erasure, p, the probability of decoder error, $P_{\text{code}}(p)$, is bounded by

$$P_{\text{code}}(p) < \sum_{s=d_{min}}^{n-k} \sum_{j=d_{min}}^{s} A_j \frac{(n-j)!\,(n-s)!}{(s-j)!} p^s(1-p)^{(n-s)} + \sum_{s=n-k+1}^{n} p^s(1-p)^{(n-s)}. \tag{1.24}$$

In Eq. (1.24), the first term depends upon the weight distribution of the code while the second term is independent of the code. The basic principle in the above equation is that an erasure decoder error is caused if an erasure pattern includes the support of a codeword. Since no erasure pattern can be corrected if it contains more than $n-k$ errors, only codewords with weight less than or equal to $n-k$ are involved. Consequently, a much tighter bound is obtained than a bound based on the union bound as there is less likelihood of double counting error events.

Considering the maximum likelihood decoder consisting of a bank of correlators, a decoder error occurs if one correlator has a higher output than the correlator corresponding to the correct codeword where the two codewords differ in s bit positions. To the decoder, it makes no difference if the decoder error event is due to erasures, from the erasure channel, or Gaussian noise from the AWGN channel; the outcome is the same. For the erasure channel, the probability of this error event due to erasures, $P_{\text{erasure}}(p)$ is

$$P_{\text{erasure}}(p) = p^s \tag{1.25}$$

The probability of this error event due to noise, $P_{\text{noise}}\left(\frac{E_b}{N_0}\right)$ is

$$P_{\text{noise}}\left(\frac{E_b}{N_0}\right) = \frac{1}{2}\text{erfc}\left(\sqrt{s\frac{k}{n}\frac{E_b}{N_0}}\right) \tag{1.26}$$

Equating Eqs. (1.25) to (1.26), for these probabilities gives a relationship between the erasure probability, p and $\frac{E_b}{N_0}$ and the Hamming distance, s.

$$p^s = \frac{1}{2}\text{erfc}\left(\sqrt{s\frac{k}{n}\frac{E_b}{N_0}}\right) \tag{1.27}$$

For many codes, the erasure decoding performance is determined by a narrow range of Hamming distances and the variation in $\frac{E_b}{N_0}$ as a function of s is insignificant. This is illustrated in Fig. 1.3 which shows the variation in $\frac{E_s}{N_0}$ as a function of s and p.

It is well known that the distance distribution for many linear, binary codes including BCH codes, Goppa codes, self-dual codes [7, 8, 10, 14] approximates to a binomial distribution. Accordingly,

$$A_j \approx \frac{n!}{(n-j)!\,j!\,2^{n-k}}. \tag{1.28}$$

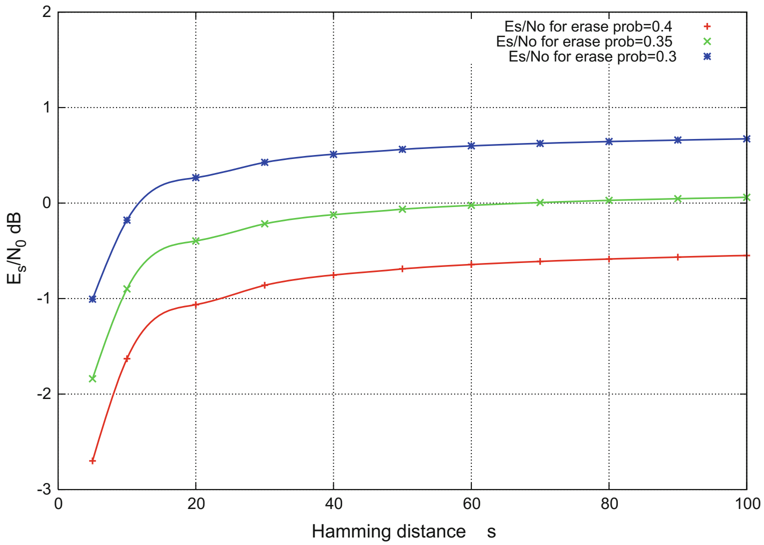

Fig. 1.3 $\frac{E_s}{N_0}$ as a function of Hamming distance s and erasure probability p

Substituting this into Eq. (1.24) produces

$$P_{\text{code}}(p) < \sum_{s=1}^{n-k} \frac{2^s - 1}{2^{n-k}} \binom{n}{s} p^s (1-p)^{(n-s)} + \sum_{s=n-k+1}^{n} p^s (1-p)^{(n-s)} \qquad (1.29)$$

With the assumption of a binomial weight distribution, an upper bound may be determined for the erasure performance of any (n, k) code, and in turn, equating Eq. (1.25) with Eq. (1.26) produces an upper bound for the AWGN channel. For example, Fig. 1.4 shows an upper bound of the erasure decoding performance of a $(128, 64)$ code with a binomial weight distribution.

Using Eq. (1.27), the decoding performance may be expressed in terms of $\frac{E_b}{N_0}$ and Fig. 1.5 shows the upper bound of the decoding performance of the same code against Gaussian noise, as a function of $\frac{E_b}{N_0}$.

The comparison of the sphere packing bound and the Gallager bounds is shown in Fig. 1.6. Also shown in Fig. 1.6 is the performance of the BCH (128, 64, 22) code evaluated using the modified Dorsch decoder. It can be seen from Fig. 1.6 that the erasure-based upper bound is very close to the sphere packing lower bound and tighter than the Gallager bound.

Figure 1.7 gives the bounds for the (512, 256) and (256, 128) codes. It will be noticed that the gap between the sphere packing bound and the erasure-based upper bound increases with code length, but is tighter than the Gallager bound.

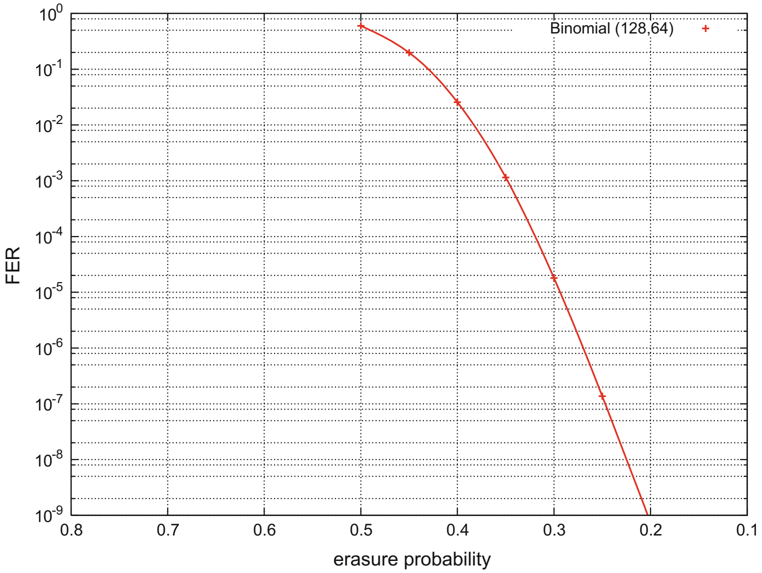

Fig. 1.4 Erasure decoding performance of a (128, 64) code with a binomial weight distribution

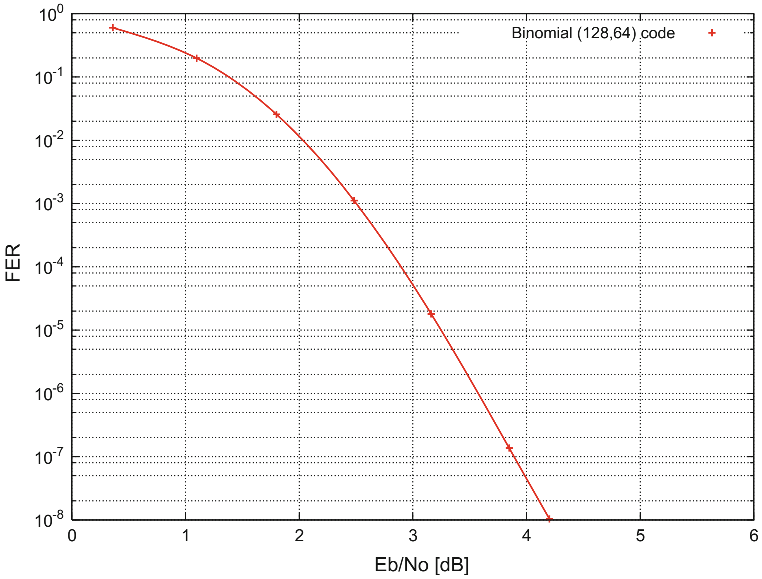

Fig. 1.5 Decoding performance of a (128, 64) code with a binomial weight distribution for Gaussian noise

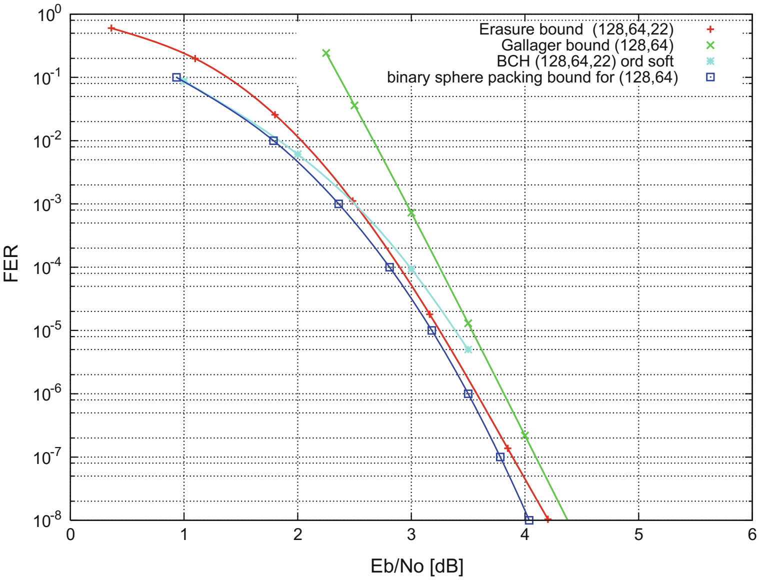

Fig. 1.6 Comparison of sphere packing and Gallager bounds to the upper bound based on erasure performance for the (128, 64) code with a binomial weight distribution

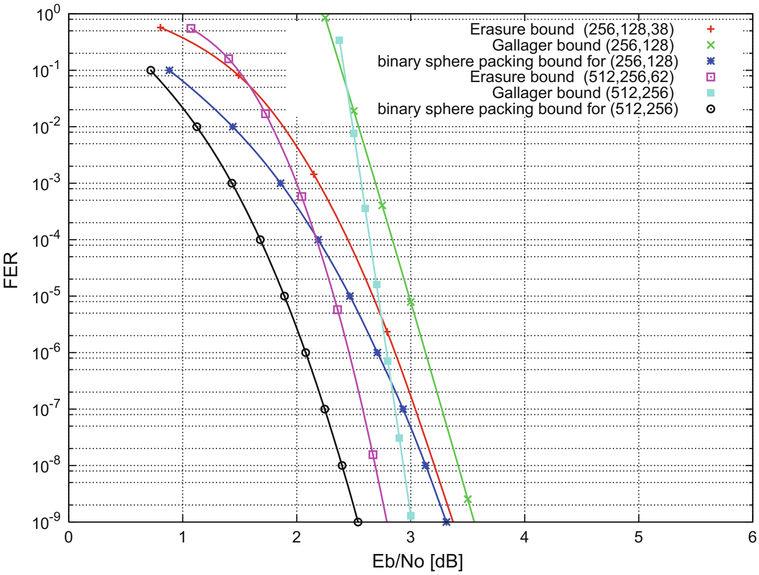

Fig. 1.7 Comparison of sphere packing and Gallager bounds to the upper bound based on erasure performance for (256, 128) and (512, 256) codes with a binomial weight distribution

1.1.2 Covering Radius of Codes

The covering radius of a code, c_r if it is known, together with the weight spectrum of the low-weight codewords may be used to tighten the Union bound upper bound on decoder performance given by Eq. (1.23). The covering radius of a code is defined as the minimum radius which when placed around each codeword includes all possible q^n vectors. Equivalently, the covering radius is the maximum number of hard decision errors that are correctable by the code. For a perfect code, such as the Hamming codes, the covering radius is equal to $\frac{d_{min}-1}{2}$. For the $[2^m-1, 2^m-m-1, 3]$ Hamming codes, the covering radius is equal to 1 and for the (23, 12, 7) Golay code the covering radius is equal to 3. As a corollary, for any received vector in Euclidean space, there is always a codeword within a Euclidean distance of $c_r + 0.5$. It follows that the summation in Eq. (1.23) may be limited to codewords of weight $2c_r + 1$ to produce

$$p_{\text{overall}} < \frac{2^k}{2^{n+1}} \sum_{d=0}^{2c_r+1} \binom{n}{d} \text{erfc}\left(\sqrt{d\frac{k}{n}\frac{E_b}{N_0}}\right). \tag{1.30}$$

1.1.3 Usefulness of Bounds

The usefulness of bounds may be realised from Fig. 1.8 which shows the performance of optimised codes and decoders all (512, 256) codes for a turbo code, LDPC code and a concatenated code.

1.2 Bounds on the Construction of Error-Correcting Codes

A code (linear or nonlinear), $\mathscr{C}$, defined in a finite field of size q can be described with its length n, number of codewords[1] M and minimum distance d. We use $(n, M, d)_q$ to denote these four important parameters of a code. Given any number of codes defined in a field of size q with the same length n and distance d, the code with the maximum number of codewords M is the most desirable. Equivalently, one may choose to fix n, M and q and maximise d or fix M, d and q and maximise n. As a result, it is of interest in coding theory to determine the maximum number of codewords possible of any code defined in a field of size q, with minimum distance d and length n. This number is denoted by $A_q(n, d)$. Bounds on $A_q(n, d)$ are indicators to the maximum performance achievable from any code with parameters $(n, M, d)_q$. As a result, these bounds are especially useful when one constructs good error-correcting codes. The tables in [5] contain the best-known upper and lower bounds on $A_q(n, d)$ for linear codes. The tables in [9] contain bounds on $A_2(n, d)$ for nonlinear binary codes.

[1] Where the code dimension $k = \log_q M$.

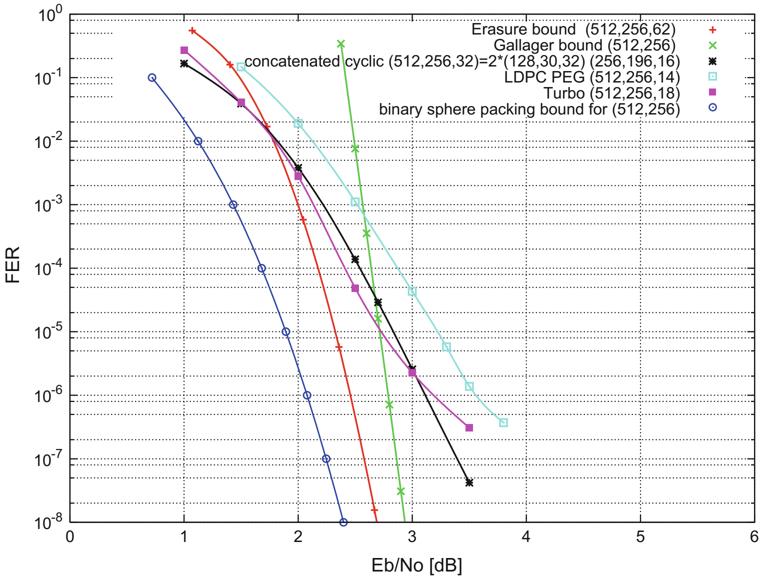

Fig. 1.8 Comparison of sphere packing, Gallager and erasure-based bounds to the performance realised for a (512, 256, 18) turbo code, (512, 256, 14) LDPC code and (512, 256, 32) concatenated code

Lower bounds on $A_q(n, d)$ tend to be code specific; however, there are several generic upper bounds. As an example, consider the best-known upper and lower bounds on $A_2(128, d)$ obtained from the tables in [5]. These are shown in Fig. 1.9 for the range $1 \leq d \leq 128$. Optimal codes of length $n = 128$ are codes whose lower and upper bounds on $A_2(128, d)$ coincide. The two curves coincide when k is small and d is large or vice versa. The gap between the upper and lower bounds that exists for other values of k and d suggests that one can construct good codes with a larger number of codewords and improve the lower bounds. An additional observation is that extended BCH codes count as some of the known codes with the most number of codewords.

It is often useful to see the performance of codes as their code lengths become arbitrarily large. We define the information rate

$$\alpha_q(\delta) = \lim_{n\to\infty} \frac{\log_q(A_q(n, \delta n))}{n}, \tag{1.31}$$

where $\delta = \frac{d}{n}$ is called the relative distance. Since the dimension of the code is defined as $k = \log_q(A_q(n, \delta n))$, then a bound on the information rate $\alpha_q(\delta)$ is a bound on $\frac{k}{n}$, as $n \to \infty$.

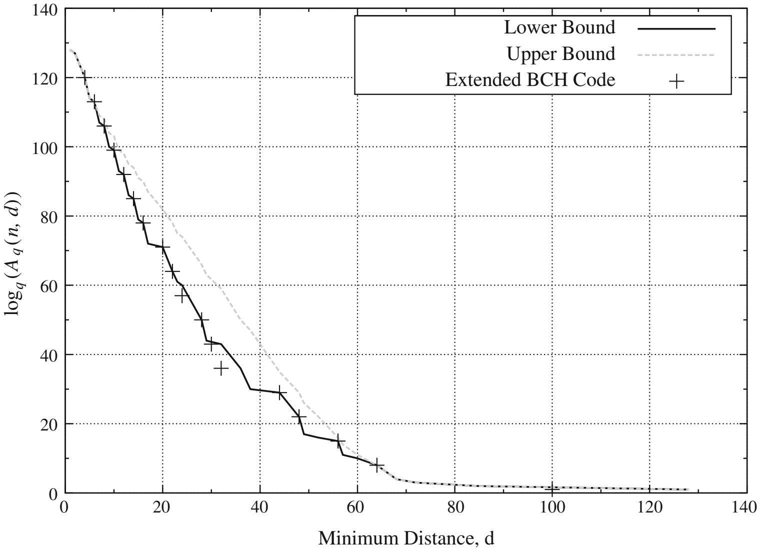

Fig. 1.9 Upper and lower bounds on $A_2(128, d)$

1.2.1 Upper Bounds

1.2.1.1 Sphere Packing (Hamming) Bound

Let $V_q(n, t)$ represent the number of vectors in each sphere then,

$$V_q(n, t) = \sum_{i=0}^{t} \binom{n}{i} (q-1)^i. \tag{1.32}$$

Theorem 1.1 (Sphere Packing Bound) *The maximum number of codewords $A_q(n, d)$ is upper bounded by,*

$$A_q(n, d) \le \frac{q^n}{\sum\limits_{i=0}^{t} \binom{n}{i} (q-1)^i}$$

Proof A code $\mathscr{C}$ is a subset of a vector space $\mathrm{GF}(q)^n$. Each codeword of $\mathscr{C}$ has only those vectors $\mathrm{GF}(q)^n$ but not in $\mathscr{C}$ lying at a hamming distance $t = \left\lfloor \frac{d-1}{2} \right\rfloor$ from it since codewords are spaced at least d places apart. In other words, no codewords lie in a sphere of radius t around any codeword of $\mathscr{C}$. As such, for counting purposes, these spheres can represent individual codewords. The Hamming bound counts the number of such non-overlapping spheres in the vector space $\mathrm{GF}(q)^n$.

Codes that meet this bound are called *perfect* codes. In order to state the asymptotic sphere packing bound, we first define the $q$ary entropy function, $H_q(x)$, for the values $0 \le x \le r$,

$$H_q(x) = \begin{cases} 0 & \text{if} \quad x = 0 \\ x \log_q(q-1) - x \log_q x - (1-x)\log_q(1-x) & \text{if} \quad 0 < x \le r \end{cases} \tag{1.33}$$

Theorem 1.2 (Asymptotic Sphere Packing Bound) *The information rate of a code $\alpha_q(\delta)$ is upper bounded by,*

$$\alpha_q(\delta) \le 1 - H_q\left(\frac{\delta}{2}\right)$$

for the range $0 < \delta \le 1 - q^{-1}$.

1.2.1.2 Plotkin Bound

Theorem 1.3 (Plotkin Bound) *Provided $d > \theta n$, where $\theta = 1 - q^{-1}$, then,*

$$A_q(n, d) \le \left\lfloor \frac{d}{d - \theta n} \right\rfloor$$

Proof Let $S = \sum d(\mathbf{x}, \mathbf{y})$ for all codewords $\mathbf{x}, \mathbf{y} \in \mathscr{C}$, and $\mathbf{x} \ne \mathbf{y}$, and $d(\mathbf{x}, \mathbf{y})$ denotes the hamming distance between codewords $\mathbf{x}$ and $\mathbf{y}$. Assume that all the codewords of $\mathscr{C}$ are arranged in an $M \times n$ matrix D. Since $d(\mathbf{x}, \mathbf{y}) \ge d$,

$$S \ge \frac{M!}{(M-2)!} d = M(M-1)d. \tag{1.34}$$

Let $n_{i,\alpha}$ be the number of times an element α in the defining field of the code GF(q) occurs in the ith column of the matrix D. Then, $\sum\limits_{\alpha \in \mathrm{GF}(q)} n_{i,\alpha} = M$. For each $n_{i,\alpha}$ there are $M - n_{i,\alpha}$ entries of the matrix D in column i that have elements other than α. These entries are a hamming distance 1 from the $n_{i,\alpha}$ entries and there are n possible columns. Thus,

$$\begin{aligned} S &= n \sum_{i=1}^{n} \sum_{\alpha \in \mathrm{GF}(q)} n_{i,\alpha}(M - n_{i,\alpha}) \\ &= nM^2 - \sum_{i=1}^{n} \sum_{\alpha \in \mathrm{GF}(q)} n_{i,\alpha}^2. \end{aligned} \tag{1.35}$$

From the Cauchy–Schwartz inequality,

$$\left(\sum_{\alpha \in \mathrm{GF}(q)} n_{i,\alpha}\right)^2 \leq q \sum_{\alpha \in \mathrm{GF}(q)} n_{i,\alpha}^2. \tag{1.36}$$

Equation (1.35) becomes,

$$S \leq nM^2 - \sum_{i=1}^{n} q^{-1} \left(\sum_{\alpha \in \mathrm{GF}(q)} n_{i,\alpha}\right)^2 \tag{1.37}$$

Let $\theta = 1 - q^{-1}$,

$$\begin{aligned} S &\leq nM^2 - \sum_{i=1}^{n} q^{-1} \left(\sum_{\alpha \in \mathrm{GF}(q)} n_{i,\alpha}\right)^2 \\ &\leq nM^2 - q^{-1} nM^2 \\ &\leq n\theta M^2. \end{aligned} \tag{1.38}$$

Thus from (1.34) and (1.38) we have,

$$M(M-1)d \leq S \leq n\theta M^2 \tag{1.39}$$

$$M \leq \left\lfloor \frac{d}{d - \theta n} \right\rfloor \tag{1.40}$$

and clearly $d > \theta n$.

Corollary 1.1 (Asymptotic Plotkin Bound) *The asymptotic Plotkin bound is given by,*

$$\begin{aligned} \alpha_q(\delta) &= 0 && \text{if} \quad \theta \leq \delta \leq 1 \\ \alpha_q(\delta) &\leq 1 - \frac{\delta}{\theta} && \text{if} \quad 0 \leq \delta \leq \theta. \end{aligned}$$

1.2.1.3 Singleton Bound

Theorem 1.4 (Singleton Bound) *The maximum number of codewords $A_q(n, d)$ is upper bounded by,*

$$A_q(n, d) \leq q^{n-d+1}.$$

Codes that meet this bound with equality, i.e. $d = n - k + 1$, are called maximum distance separable codes (MDS). The asymptotic Singleton bound is given Theorem 1.5.

Theorem 1.5 (Asymptotic Singleton Bound) *The information rate $\alpha_q(\delta)$ is upper bounded by,*

$$\alpha_q(n, \delta) \leq 1 - \delta.$$

The asymptotic Singleton bound does not depend on the field size q and is a straight line with a negative slope in a plot of $\alpha_q(\delta)$ against δ for every field.

1.2.1.4 Elias Bound

Another upper bound is the Elias bound [17]. This bound was discovered by P. Elias but was never published by the author. We only state the bound here as the proof is beyond the scope of this text. For a complete treatment see [6, 10].

Theorem 1.6 (Elias Bound) *A code $\mathcal{C}$ of length n with codewords having weight at most w, $w < \theta n$ with $\theta = 1 - q^{-1}$ has,*

$$d \leq \frac{Mw}{M-1}\left(2 - \frac{w}{\theta n}\right)$$

Theorem 1.7 (Asymptotic Elias Bound) *The information rate $\alpha_q(\delta)$ is upper bounded by,*

$$\alpha_q(\delta) \leq 1 - H_q(\theta - \sqrt{\theta(\theta - \delta)})$$

provided $0 < \delta < \theta$ *where* $\theta = 1 - q^{-1}$.

1.2.1.5 MRRW Bounds

The McEliece–Rodemich–Rumsey–Welch (MRRW) bounds are asymptotic bounds obtained using linear programming.

Theorem 1.8 (Asymptotic MRRW Bound I) *Provided* $0 < r < \theta$, $\theta = 1 - q^{-1}$ *then,*

$$\alpha_q(\delta) \leq H_q\left(\frac{1}{q}(q - 1 - (q-2)\delta - 2\sqrt{\delta(1-\delta)(q-1)})\right)$$

The second MRRW bound applies to the case when $q = 2$.

Theorem 1.9 (MRRW Bound II) *Provided* $0 < \delta < \frac{1}{2}$ *and* $q = 2$ *then,*

$$\alpha_2(\delta) \leq \min_{0 \leq u \leq 1-2\delta}\{1 + g(u^2) - g(u^2 + 2\delta u + 2\delta)\}$$

where

$$g(x) = H_2\left(\frac{1-\sqrt{1-x}}{2}\right).$$

The MRRW bounds are the best-known upper bound on the information rate for the binary case. The MRRW-II bound is better than the MRRW-I bound when δ is small and $q = 2$. An in depth treatment and proofs of the bounds can be found in [12].

1.2.2 *Lower Bounds*

1.2.2.1 Gilbert–Varshamov Bound

Theorem 1.10 (Gilbert–Varshamov Bound) *The maximum number of codewords $A_q(n,d)$ is lower bounded by,*

$$A_q(n,d) \geq \frac{q^n}{V_q(n,d-1)} = \frac{q^n}{\sum\limits_{i=0}^{d-1}\binom{n}{i}(q-1)^i}.$$

Proof We know that $V_q(n,d-1)$ represents the volume of a sphere centred on a codeword of $\mathscr{C}$ of radius $d-1$. Suppose $\mathscr{C}$ has $A_q(n,d)$ codewords. Every vector $\mathbf{v} \in \mathbb{F}_q^n$ lies within a sphere of volume $V_q(n,d-1)$ centred at a codeword of $\mathscr{C}$ as such,

$$\left|\bigcup_{i=1}^{A_q(n,d)} S_i\right| = |\mathbb{F}_q^n|,$$

where S_i is a set containing all vectors in a sphere of radius $d-1$ centred on a codeword of $\mathscr{C}$. The spheres S_i are not mutually disjoint. If we assume S_i are mutually disjoint then,

$$A_q(n,d)V_q(n,d-1) \geq |\mathbb{F}_q^n|.$$

Theorem 1.11 *The information rate of a code is lower bounded by,*

$$\alpha_q(\delta) \geq 1 - H_q(\delta)$$

for $0 \leq \delta \leq \theta$, $\theta = 1 - q^{-1}$.

Figures 1.10 and 1.11 show the asymptotic upper and lower bounds for the cases where $q = 2$ and $q = 32$, respectively. Figure 1.11 shows that the MRRW bounds are the best-known upper bounds when $q = 2$. Observe that the Plotkin bound is the best upper bound for the case when $q = 32$.

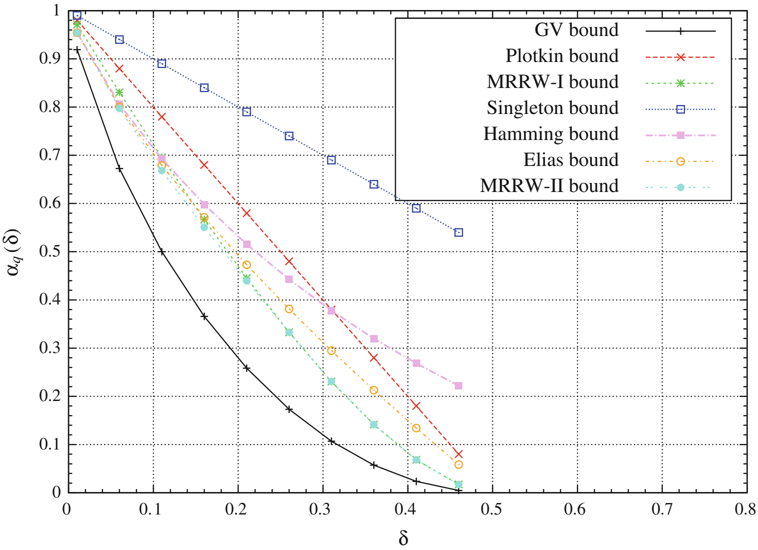

Fig. 1.10 $\alpha_q(\delta)$ against δ for $q = 2$

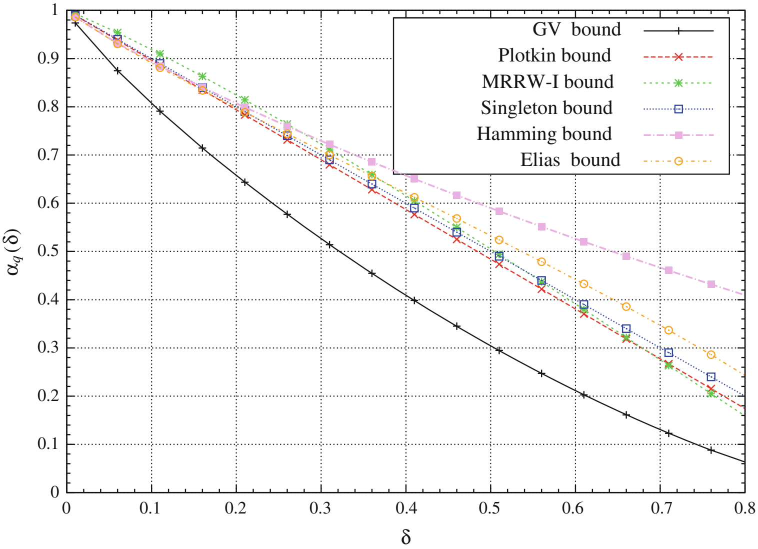

Fig. 1.11 $\alpha_q(\delta)$ against δ for $q = 32$

Table 1.1 Ranges for codes

Finite field	Range
$\mathbb{F}_2$	$1 \le k \le n \le 256$
$\mathbb{F}_3$	$1 \le k \le n \le 243$
$\mathbb{F}_4$	$1 \le k \le n \le 256$
$\mathbb{F}_5$	$1 \le k \le n \le 130$
$\mathbb{F}_7$	$1 \le k \le n \le 100$
$\mathbb{F}_8$	$1 \le k \le n \le 130$
$\mathbb{F}_9$	$1 \le k \le n \le 130$

1.2.3 Lower Bounds from Code Tables

Tables of best-known codes are maintained such that if a code defined in a field q is constructed with an evaluated and verifiable minimum Hamming distance d that exceeds a previously best-known code with the same length n and dimension, the dimension of the new code is a lower bound on $A_q(n, d)$. The first catalogue of best-known codes was presented by Calabi and Myrvaagnes [2] containing binary codes of length n and dimension k in the range $1 \le k \le n \le 24$. Brouwer and Verhoeff [1] subsequently presented a comprehensive update to the tables which included codes with finite fields up to size 9 with the ranges for k and n.

At present, Grassl [5] maintains a significantly updated version of the tables in [1]. The tables now contain codes with k and n in ranges from Table 1.1. Finally, Schimd and Shurer [15] provide an online database for optimal parameters of (t, m, s)-nets, (t, s)-sequences, orthogonal arrays, linear codes and ordered orthogonal arrays. These are relatively new tables and give the best-known codes up to finite fields of size 256. The search for codes whose dimension exceeds the best-known lower bounds on $A_q(n, d)$ is an active area of research with the research community constantly finding improvements.

1.3 Summary

In this chapter we discussed the theoretical performance of binary codes for the additive white Gaussian noise (AWGN) channel. In particular the usefulness of Gallager's coding theorem for binary codes was explored. By assuming a binomial weight distribution for linear codes, it was shown that the decoder error probability performance of some of the best, known linear, binary codes is the same as the average performance of the ensemble of all randomly chosen, binary nonlinear codes having the same length and dimension. Assuming a binomial weight distribution, an upper bound was determined for the erasure performance of any code, and it was shown that this can be translated into an upper bound for code performance in the AWGN channel. Different theoretical bounds on the construction of error-correction codes were discussed. For the purpose of constructing good error-correcting codes, theoretical upper bounds provide fundamental limits beyond which no improvement is possible.

References

1. Brouwer, A., Verhoeff, T.: An updated table of minimum-distance bounds for binary linear codes. IEEE Trans. Inf. Theory **39**(2), 662–677 (1993)
2. Calabi, L., Myrvaagnes, E.: On the minimal weight of binary group codes. IEEE Trans. Inf. Theory **10**(4), 385–387 (1964)
3. Chen, C.L.: Computer results on the minimum distance of some binary cyclic codes. IEEE Trans. Inf. Theory **16**(3), 359–360 (1970)
4. Gallager, R.G.: A simple derivation of the coding theorem and some applications. IEEE Trans. Inf. Theory **11**(1), 459–470 (1960)
5. Grassl, M.: Code Tables: Bounds on the parameters of various types of codes, http://www.codetables.de (2007)
6. Huffman, W.C., Pless, V.S.: Fundamentals of Error-Correcting Codes. Cambridge University Press, Cambridge (2003). ISBN 0 521 78280 5
7. Krasikov, I., Litsyn, S.: On spectra of BCH codes. IEEE Trans. Inf. Theory **41**(3), 786–788 (1995)
8. Krasikov, I., Litsyn, S.: On the accuracy of the binomial approximation to the distance distribution of codes. IEEE Trans. Inf. Theory **41**(5), 1472–1474 (1995)
9. Litsyn, S.: Table of nonlinear binary codes, http://www2.research.att.com/~njas/codes/And/ (1999)
10. MacWilliams, F.J., Sloane, N.J.A.: The Theory of Error-Correcting Codes. North-Holland, Amsterdam (1977)
11. Massey, J.: Coding and modulation in digital communication. In: Proceedings of International Zurich Seminar on Digital Communication, pp. E2(1)–E2(24) (1974)
12. McEliece, R., Rodemich, E., Rumsey, H., Welch, L.: New upper bounds on the rate of a code via the delsarte-macwilliams inequalities. IEEE Trans. Inf. Theory **23**(2), 157–166 (1977)
13. Promhouse, G., Tavares, S.E.: The minimum distance of all binary cyclic codes of odd lengths from 69 to 99. IEEE Trans. Inf. Theory **24**(4), 438–442 (1978)
14. Roychowdhury, V.P., Vatan, F.: Bounds for the weight distribution of weakly self-dual codes. IEEE Trans. Inf. Theory **47**(1), 393–396 (2001)
15. Schimd, W., Shurer, R.: Mint: a database for optimal net parameters, http://mint.sbg.ac.at (2004)
16. Schomaker, D., Wirtz, M.: On binary cyclic codes of odd lengths from 101 to 127. IEEE Trans. Inf. Theory **38**(2), 516–518 (1992)
17. Shannon, C., Gallager, R., Berlekamp, E.: Lower bounds to error probability for coding on discrete memoryless channels, i. Inf. Control **10**(1), 65–103 (1967)
18. Shannon, C.E.: Probability of error for optimal codes in a Gaussian channel. Bell Syst. Tech. J. **38**(3), 611–656 (1959)
19. Tjhai, C., Tomlinson, M.: Results on binary cyclic codes. Electron. Lett. **43**(4), 234–235 (2007)
20. Wozencraft, J.: Sequential decoding for reliable communications. Technical Report No. 325 Research Laboratory of Electronics, MIT (1957)

21. Wozencraft, J., Jacobs, I.: Principles of Communication Engineering. Wiley, New York (1965)
22. Wozencraft, J., Kennedy, R.: Modulation and demodulation for probabilistic coding. IEEE Trans. Inf. Theory IT **12**, 291–297 (1966)