Possibilities and Limitations of Forward Error Correction in Minimal QSOs

Klaus von der Heide, DJ5HG

This presentation analyses Shannon's famous theoretical result on the maximum achievable information bitrate and applies the results to the case of a minimal QSO. It will be explained why the energy to communicate an optimally encoded callsign in a single transmission is more than twice that given by the Shannon limit and more than eight times in the case of an uncoded transmission.

1. The Channel Capacity

On VHF and above a very good model of the noise is the Additive White Gaussian Noise (AWGN). The power spectral density of the AWGN (the power per Hz of bandwidth) is

$N_0 = k_B T$

$k_B = 1.38 \cdot 10^{-23}$ W / (K*Hz) is the Boltzmann-constant, and T is the equivalent noise temperature in degrees Kelvin. The noise temperature varies with frequency, but it is assumed constant within the received bandwidth (that is the meaning of the term “white“). On 2m, T lies between 200 K and 800 K. It especially depends on which part of our galaxy radiates into the antenna. At zero elevation also the earth surface contributes (and man-made wideband noise).

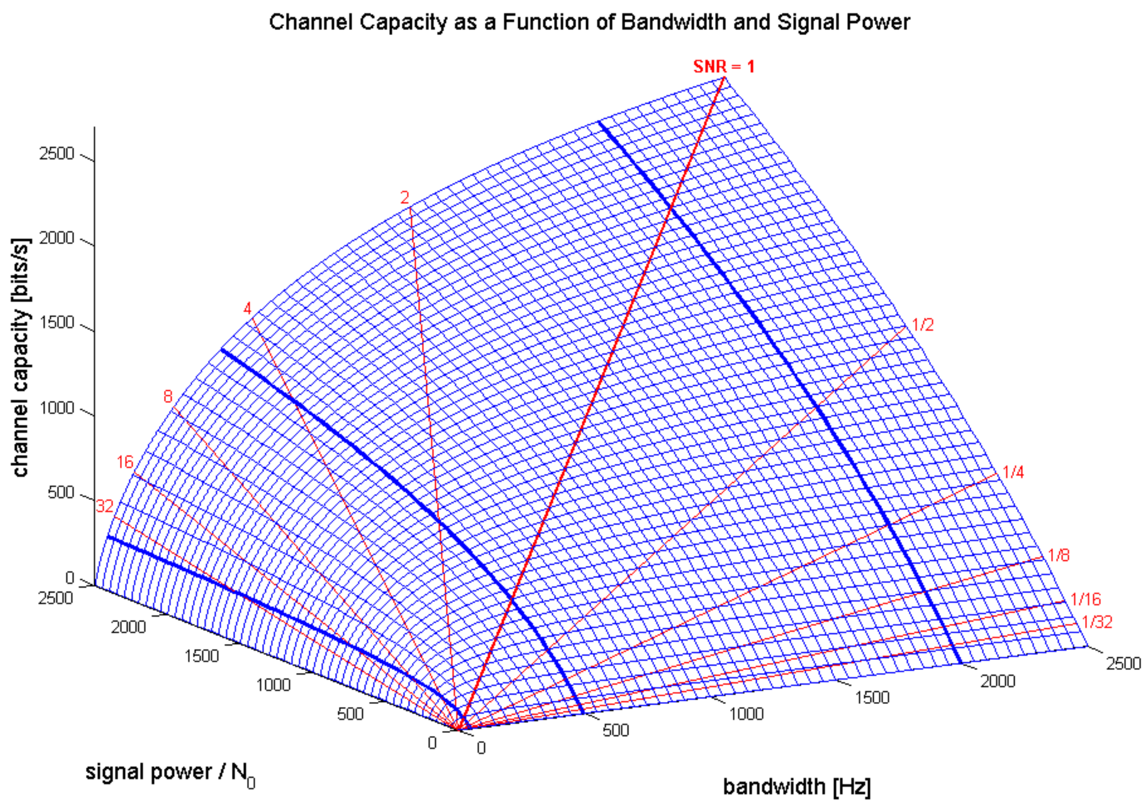

Fig 1 : Theoretical limit of the rate of received information in bits per second (so called channel capacity) as a function of bandwidth and relative signal power or signal to noise ratio (SNR, red lines). The bold blue lines mark usual bandwidths of 50 Hz, 500 Hz, and 2000 Hz.

Let b be the used bandwidth, P the received signal power at the demodulator, and SNR the signal-to-noise ratio at the demodulator. Shannon [1] deduced a theoretical limit of the rate of received information bits which he called the *channel capacity* (ld denotes the base-2 logarithm):

$C = b * ld(1 + P / (b * N_0)) = b * ld(1 + SNR)$

Figure 1 shows this maximum rate C as a function of bandwidth b and signal power P. The red diagonal lines denote constant SNR. On the bold red line we have SNR = 1 and C = b . The three usual bandwidths 50 Hz (CW EME), 500 Hz (Pactor2), and 2000 Hz (FSK441, JT65b etc.) are stressed by bold blue lines.

To ease the interpretation of figure 1 we use the two sectional views of figures 2 and 3.

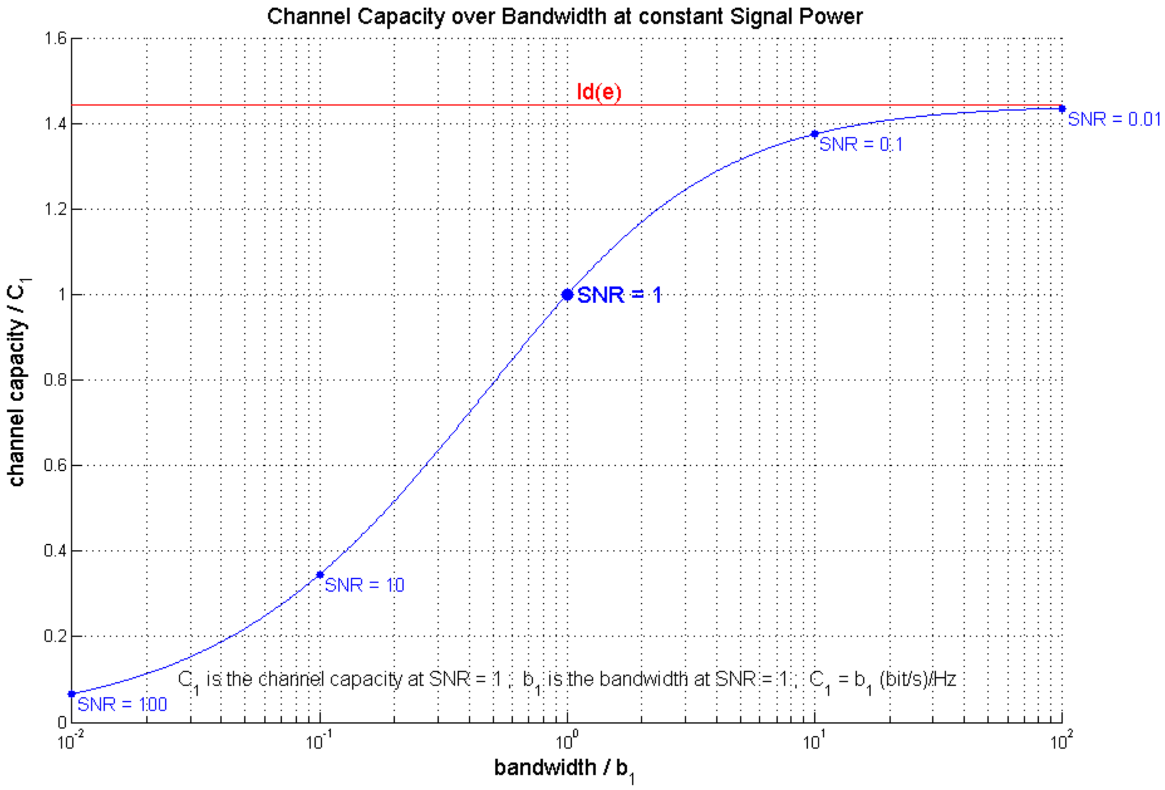

Fig 2 : Sectional view of figure 1: Relative channel capacity as a function of bandwidth at constant signal power. The red line at ld(e) ist the theoretical upper limit found by Claude Elwood Shannon. It is obvious that the upper limit only could be reached at very large bandwidth (resp. very low SNR). In practise, the usable bandwidth often is limited by birdies.

Different from figure 1, the bandwidth is scaled logarithmically in figure 2, and both axes are made relative to their values at SNR = 1 . As a consequence, figure 2 represents the sectional views of figure 1 for all possible constant values of the signal power.

We now discuss the important case of constant signal power, but no restriction in bandwidth. Figure 2 shows that the maximum rate of received information (the channel capacity) is maximized at large bandwidth. The maximum value is

$C_{max} = ld(e^{P/No})$

Figure 2 is normalized such that $P/N_0 = 1$. Thus

$C_{max} = ld(e) = 1.4427$

From the signal power P and the rate of information bits per second C_{max} we get the minimum energy per information bit $E_{b\,min} = P/C_{max}$ and finally

$E_{b\,min} / N_0 = (P/C_{max}) / N_0 = \ln 2 = 0.6931$

Usually this is written as

$E_b / N_0 \geq \ln 2$ or in dB: $E_b / N_0 \geq 10 \log_{10} (\ln 2) = -1.5917$ dB

Using $N_0 = k_B T$ from above with $k_B = 1.38 \cdot 10^{-23}$ W / (K*Hz) and T = 300 K for example we get $E_b \geq \ln(2)\ k_B T = 2.87 \cdot 10^{-21}$ Ws .

This is the minimum necessary energy to receive an information bit confidently. It is important to differentiate between information content measured in bits and the binary digits that represent the information. An information content of k bits can be encoded by at least k binary digits. Forward error correcting codes actually use n>k binary digits to encode k bits of information. The overhead of n – k bits is called *redundancy*, the ratio k/n is called the *code rate*.

Figure 2 makes clear that the channel capacity increases at increasing bandwidth. But the Shannon limit is rapidly reached. Above $b = 10\ b_1$ not more than 0.2 dB could be additionally gained. Let for example be SNRdB = -24 dB at b = 2000 Hz . The bandwidth for SNR = 1 (SNRdB = 0) then is $b_1 = b / (10^{SNRdB/10}) = 7.96$ Hz. Figure 2 is normalized to this value. If we now use b = 80 Hz instead of b = 7.96 Hz then we get an increase by a factor of 1.37 .

Now, we discuss the inverse case of costly bandwidth and cheap signal power. It is profitable then to leave the point of SNR = 1 to the left. Halving the bandwidth only reduces the capacity to $0.8\ C_1$. Further reduction of the bandwidth to $^1/_{50}$ decreases the capacity to $^1/_{10}$ of that at SNR = 1.

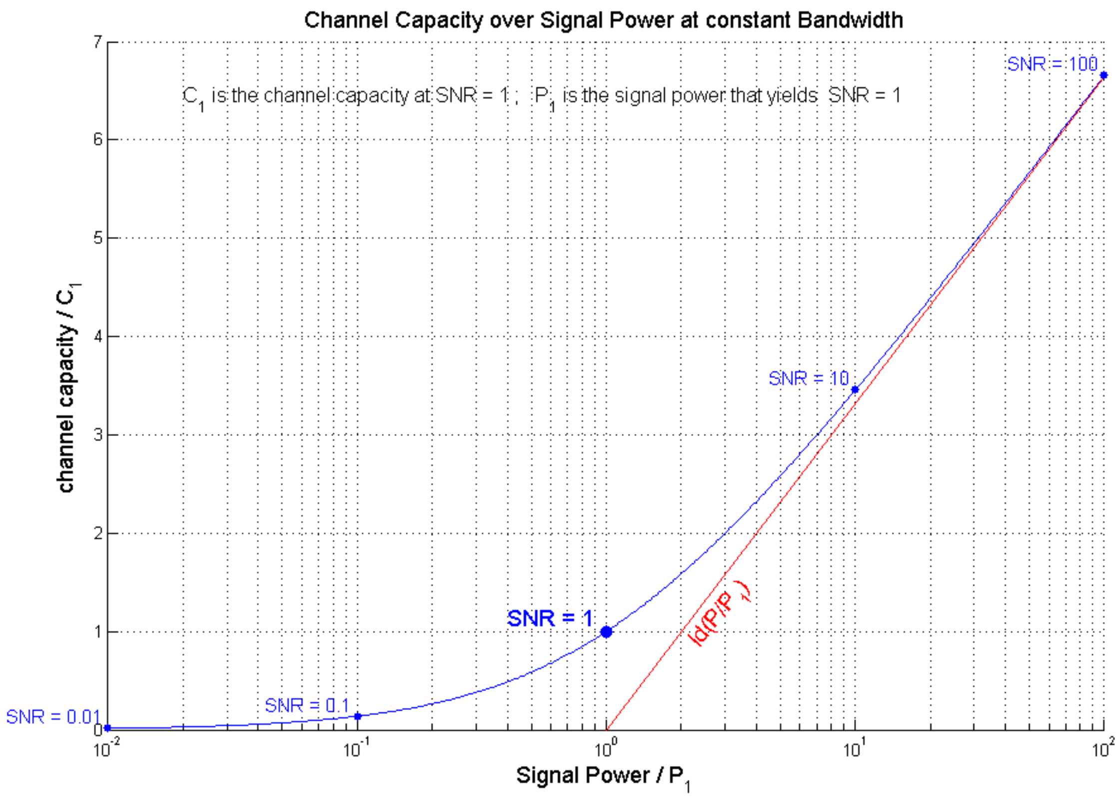

Fig 3 : Sectional view of figure 1: Relative channel capacity as a function of signal power at constant bandwidth. This figure shows that the power must be exponentially increased to get a constant gain of bitrate.

Figure 3 shows the channel capacity as a function of signal power at constant bandwidth. Again starting at the point SNR = 1 , we observe that an increase of the signal power by a decade increases the channel capacity by a factor of 3.5 . Any further increase of the signal power by additional decades only add ld(10) * C_1 = 3.322 C_1 to the capacity.

The theoretical deduction of Shannon is based on the idea of communicating *infinite* amounts of information. The value of the Shannon limit is applicable only to that academic case. If a limited information content of k bits is communicated by a codeword of n binary digits then another limit can be deduced by geometric arguments. It is called the *sphere-packing bound*. Unfortunately, this bound cannot be reached in the case of an AWGN channel. This presentation will use the *Plotkin-bound* instead [3].

Codes usually are rated by values of E_b / N_0 that guarantee a word error rate of 10^{-4} (or 10^{-6}) . This means that on average only one out of 10000 (or 1000000) decoded messages is wrong. Figure 4 shows the sphere-packing bound for code rate r = 0 . At very large information content (k = 10^6) the sphere-packing bound reaches the Shannon limit of -1.59 dB. The figure also contains the values of E_b / N_0 for some codes required for a word error rate of 10^{-4} . The information content of messages within a minimal QSO only lies between 1 and 30 bits. It follows from figure 4 that the required signal power then is at least twice as large as in deep space applications.

http://www331.jpl.nasa.gov/public/AllCodesVsSize.GIF

Fig. 4 Required values of E_b/N_0 to reach a word error rate of P_W=0.0001 for some codes. Codes that encode many information bits (k=10^4 for ex.) need less energy per bit. Beyond 10^4 bits there is no further practical gain that could justify the enormous effort to decode longer words. Deep space missions therefore use codes in this region as is indicated by Voyager, Cassini, and Galileo.

2. The Case of a Minimal QSO

The case of a minimal QSO in Ham Radio is entirely different from commercial applications or deep space:

(a) The main challenge is to succeed in the synchronisation of two stations for at least four periods of subsequent transmissions. This includes at both ends: packet synchronization, symbol synchronization (and carrier synchronization if a coherent demodulation is used).

(b) There is no need to exchange any information except from the control information to identify both stations and to acknowledge correct reception. This accumulates to an extremely low information content of about 4 ... 80 bits transferred in both directions.

(c) The allowed word error rate is much higher than in commercial applications. A word error rate of 0.1 is a good condition. Even at a word error rate of 0.5 and one repetition of each transmission a standard procedure of four steps would work correctly in 30% of all cases.

(d) Because of the low information content of the transmissions very low code rates are practical. If 30 information bits must be transmitted within 60 seconds then the uncoded transmission only needs about 1.0 Hz of bandwidth. If we allow a bandwidth of 500 Hz for the encoded transmission then a code rate of 0.02 is possible, i.e. the 30 information bits may be encoded by a code word of 15000 binary digits.

(e) Switch-over between reception and transmission costs about one second. The decision of the operator which content to send needs even more time. Therefore the time to decode the last received pass also is allowed to take about a second. That is very much more than in commercial applications. This makes codes practical that have a complex decode algorithm.

The literature rarely offers information on codes used under these special conditions. And information on codes optimized to these conditions is even more rare. It is the aim of this presentation to fill that gap.

3. Forward Error Correction

We use the simple case of $k = 4$ information bits to be encoded. Then there exist only 16 different messages. Table 1 lists all messages for six different binary codes.

Table 1 Example of 6 binary codes that all encode 4 information bits. The most left column lists all 16 possible binary patterns of the 4 information bits (uncoded). The separated column within the “uncoded“ section is the parity of the pattern. The four information bits plus the parity bit are the parity code. The (7,4)-Hamming code encodes the 4 information bits into 7 bits of the codeword. If the parity of the codeword is used additionally (the 8th bit to the right in the Hamming-section) the code is called the (8,4)-extended Hamming code. The 1st-order Reed Muller code also is an (8,4)-code, i.e. 4 information bits are encoded by 8 binary digits of the code word. The convolutional code uses the three polynomials `100,110,111` tail-biting resulting in a (12,4)-code. The (15,4)-code to the right is a simplex code which means that all possible pairs of different codewords exactly differ in the same number of bits (8 bits in this case). Repetition codes are not listed here. Repetition codes simply repeat the information bits. A (12,4)-repetion code would transmit each information bit three times.

uncoded	Hamming	Reed-Muller	Convolutional	Simplex
0000 0	0000000 0	00000000	000000000000	000000000000000
0001 1	0001111 0	01010101	000110011101	101010101010101
0010 1	0010110 1	00110011	001000111011	011001100110011
0011 0	0011001 1	01100110	001110100110	110011001100110
0100 1	0100101 1	00001111	010001100111	000111100001111
0101 0	0101010 1	01011010	010111111010	101101001011010
0110 0	0110011 0	00111100	011001011100	011110000111100
0111 1	0111100 0	01101001	011111000001	110100101101001
1000 1	1000011 1	11111111	100011001110	000000011111111
1001 0	1001100 1	10101010	100101010011	101010110101010
1010 0	1010101 0	11001100	101011110101	011001111001100
1011 1	1011010 0	10011001	101101101000	110011010011001
1100 0	1100110 0	11110000	110010101001	000111111110000
1101 1	1101001 0	10100101	110100110100	101101010100101
1110 1	1110000 1	11000011	111010010010	011110011000011
1111 0	1111111 1	10010110	111100001111	110100110010110

Figure 5 presents the word error rates of the codes listed in table 1 as a function of E_b/N_0 . These are results of 2000000 transmissions per point over a simulated AWGN-channel using the encoding and decoding of real transmissions.

To discuss the gain of the codes we first look at the uncoded case for k = 4 information bits. This black line shows at $E_b/N_0 = 4$ dB a word error rate of 0.05 , which means that in 5% of all transmissions of 4 uncoded binary digits at least one of the bits will be received faulty. If the parity code is used then 5 binary digits must be transferred with the same total energy instead of 4. This increases the necessary bandwidth and the noise power by a factor of 5/4 . Nevertheless, the word error rate reduces to 0.02 . The other codes need much more bandwidth, while the word error rate is further decreased until 0.0065. But the conclusion, only the code rate r = k/n must be small to get a low word error rate is wrong. A repetition code for example is not better than the uncoded transmission over an AWGN channel.

The behavior of codes with some more information is shown by the example of a tail biting convolutional code that encodes 28 information bits into 1960 binary digits. While the maximum code gain for 4 information bits at a word error rate of 0.1 only is 2 db it is 5.7 dB for a transmission of 28 information bits (black arrow). With further increasing information content of a codeword these curves approximate the vertical red line indicating the Shannon limit.

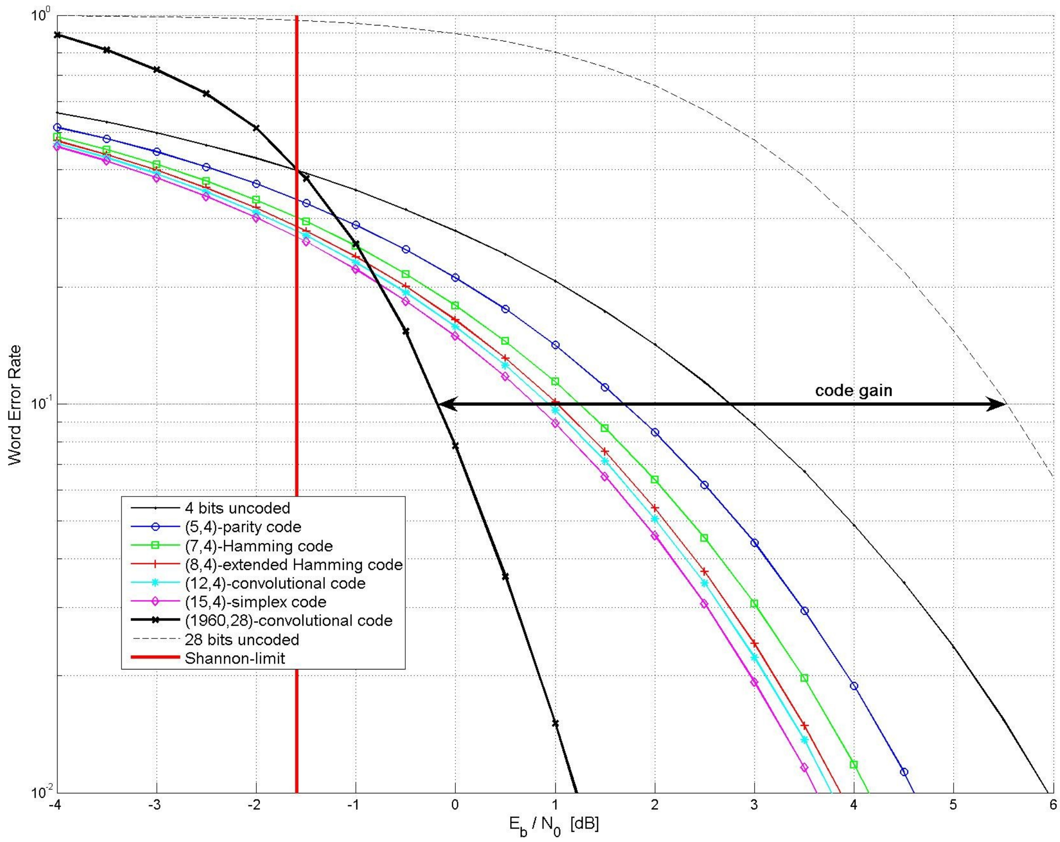

Fig 5 : Word error rates as a function of E_b/N_0 for all codes of table 1. The error rate of the Reed-Muller-code is identical to that of the extended Hamming-code. Additionally, this figure shows the word error rates of the uncoded and the encoded transmission of 28 information bits. While the maximum code gain for 4 information bits at a word error rate of 0.1 only is 2 db it is 5.7 dB for a transmission of 28 information bits. At constant energy per bit of $E_b/N_0 = 0$ dB (i.e. 1.59 dB apart from the Shannon limit) the word error rates for an uncoded and an encoded transmission are 0.9 resp. 0.079 . In the uncoded case a QSO will not run while it will run mostly errorfree in the encoded case.

Left from the Shannon limit the case of 28 information bits seem to be worse than that of 4 bits. But that is not the case. To transmit 28 bits using the codes that encode 4 bits we must send 7 codewords instead of only one. The probability that at least one of them is faulty is larger for all codes of table 1 compared to the word error rate of the (1960,28)-convolutional code shown in figure 5.

To rate a code often a single value of E_b/N_0 that guarantees a fixed word error rate is sufficient. Instead of the usual $P_w = 10^{-4}$ (or even 10^{-6}) we choose the word error rate $P_w = 0.1$. The (15,4)-simplex code in figure 5 crosses the horizontal line of $P_w = 0.1$ at $E_b/N_0 = 0.75$ dB. These values are shown in figure 6 for some codes over the number of information bits to be encoded. This figure compares to figure 4. But it is reduced to the area relevant for minimal QSOs.

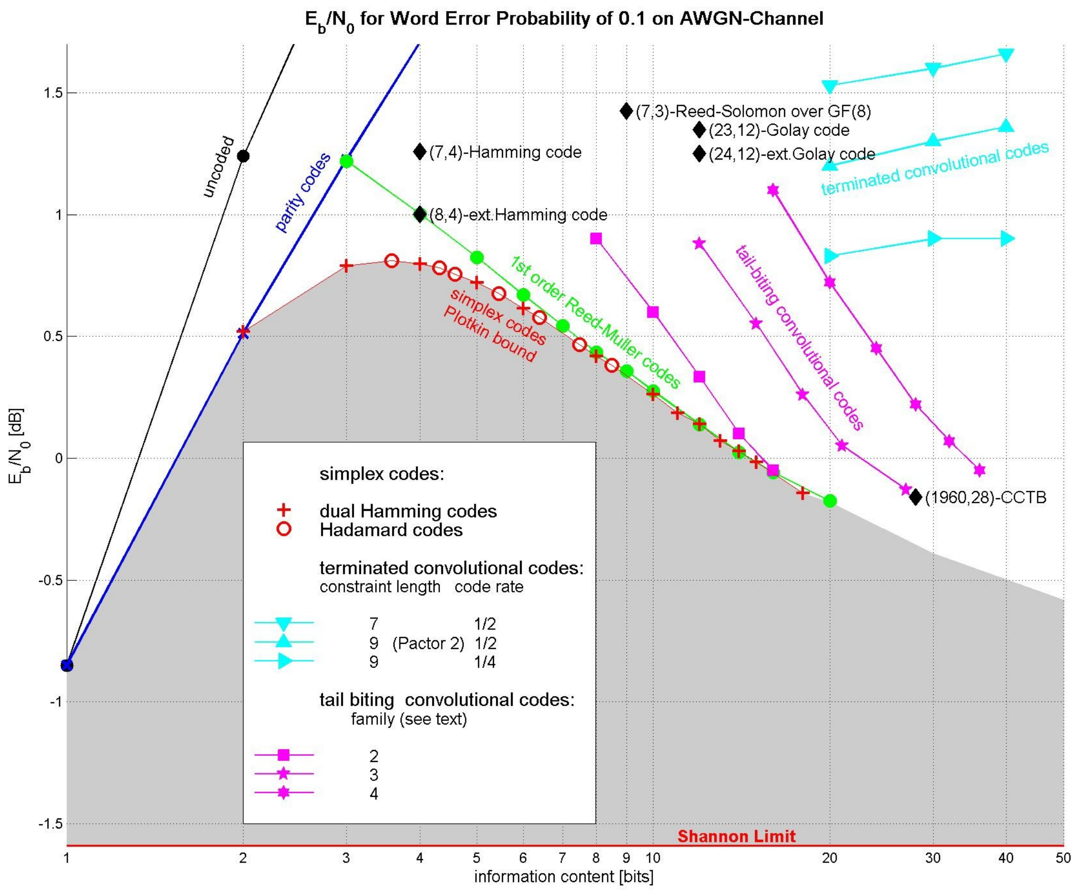

Fig 6 : Necessary values of E_b/N_0 to reach a word error rate of 0.1 for some binary codes on the AWGN-channel. The abscissa is the number of information bits to be encoded.

4. Code Gain

The black arrow in figure 5 indicates the gain of the (1960,28)-code over the uncoded transmission. The corresponding gain of many binary codes at $P_w = 0.1$ is shown in figure 7. The gain rapidly increases with increasing information content of the codewords. In a minimal QSO no more than about 28 bits are transmitted. In that case a code obviously cannot have a gain larger than 6 dB over the uncoded transmission. But, in contrast to commercial applications amateurs can nearly reach that gain, i.e. an encoded minimal QSO only needs ¼ of the power of an uncoded one.

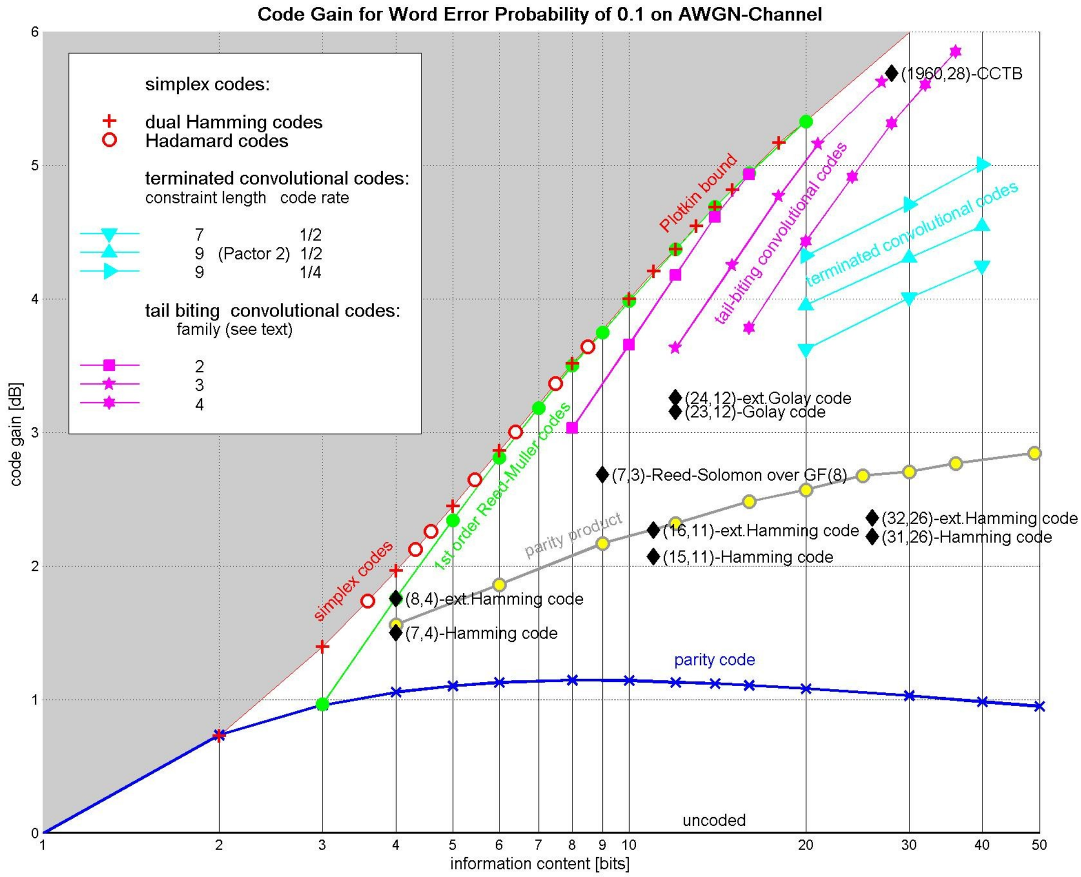

Fig 7 : Gain of some codes over the uncoded transmission (AWGN). There is no possible gain if only one single bit must be transmitted (a final RRR for example). The maximum possible gain of an encoded transmission of 6 (30) information bits is 3 (6) dB resp.. The gain is considerably larger when large information content is encoded (4 k bits for ex.). This figure shows the region relevant for a minimal QSO with transmission of callsigns, reports, and rogers only.

Literature

[1] C.E. Shannon, A Mathematical Theory of Communication,
The Bell System Technical Journal,Vol. 27, pp.397-423,623-656, July, October 1948
http://cm.bell-labs.com/cm/ms/what/shannonday/shannon1948.pdf
or search for shannon1948.pdf (many hits)

[2] http://www331.jpl.nasa.gov/public/AllCodesVsSize.GIF

[3] M. Plotkin, Binary Codes with Specified Minimum Distance, IRE Trans. Inf. Theor. 6 (1960), pp. 445-450.

[4] J. Taylor, How many Bits Are Copied in a JT65 Transmission?, DUBUS 3/2006, pp. 64 – 68