IMPROVED AND UNIFIED APPROXIMATION FOR THE BER OF LINEAR BLOCK CODES

Claude Desset, Fabrice Labeau, Luc Vandendorpe and Benoît Macq*
UCL Communications and Remote Sensing Laboratory
2, place du Levant – B -1348 Louvain-la-Neuve – Belgium
e-mail : desset@tele.ucl.ac.be

ABSTRACT

In this paper, we provide an approximation for the BER achieved by using linear block error-correcting codes. Our work unifies and improves two anterior approximations, in the case of hard-decoding limited to the minimum distance of the code. We compare them to exact values for codes whose weight distribution is known. Our first improvement takes into account undecoded words, enabling the use of our approximation for all linear block codes, while the original ones were only accurate for perfect codes. A second improvement allows us to work in the large channel BER domain, by considering the binomial approximation of weight distributions. Previous approximations were not accurate, as asymptotic expressions based on low-weight error patterns do not hold when dealing with high BER values.

1 INTRODUCTION

Forward error correction has been used for years when dealing with noisy channels. However, this field still contains challenging problems, related not only to finding good codes with respect to Shannon's optimality bounds [1], but also to modeling the ones we know. In the majority of communication systems, the ultimate performance criterion as far as channel coding is concerned is the achieved BER. In this paper, we consider the way to estimate the BER in the case of a classical binary symmetric channel. We use a hard-decoding algorithm, correcting errors in spheres of radius t around codewords.

The exact computation of BER for a given code is a problem of huge complexity, depending mainly on the weight distribution of codewords. For all but the shortest codes, this solution is intractable and we need approximations. Previous work has been achieved in that direction by *Torrieri* [2, 3] and *Labeau* [4]. We describe their work as well as basic ideas in order to compute a BER in section 2. The best approximation is selected in the case of perfect codes, by comparison with the exact value based on the weight distribution. Section 3 extends these approximations for general error-correcting codes (ECCs), considering the code density in a sphere-packing perspective and the binomial weight distribution approximation. Section 4 concludes by presenting some applications and provides ideas for future work.

*This author's work is supported by the Belgian National Fund for Scientific Research (FNRS).

2 APPROXIMATIONS FOR PERFECT CODES

2.1 Basic ideas for BER estimation

Let us suppose that we study an ECC of dimensions (n, k) and of minimum distance $d = 2t+1$, where t is the error-correcting capacity of the code[1]. In this section, we only consider perfect codes in order to benefit from their complete decoding in spheres of radius t. For a linear code, we can always consider that the codeword consisting of all zeros is transmitted and analyse the effect of different error patterns.

The basic principle of BER approximations we are studying is to separate the error patterns into classes of different Hamming weights. Denoting by f_i the average number of erroneous bits in a word after decoding, for error patterns of weight i entering the decoder, we get the following classical post-decoding error probability estimation[2], as all words of weight greater than t are erroneous :

$$p = \sum_{i=t+1}^{n} C_n^i p_e^i (1-p_e)^{n-i} \frac{f_i}{n}, \qquad (1)$$

where p_e is obviously the channel BER. Implicitly, by using $\frac{f_i}{n}$ in this expression, we assume that erroneous bits are equally spread over information and redundancy bits. This point will be discussed in section 3. Expression (1) is completely general, and all what remains to do is to compute the f_i factors. Two different hypotheses have been considered by *Torrieri* and *Labeau*. As depicted on figure 1, they use respectively $f_i = \max(d, i)$[2] and $f_i = \min(i + t, n)$[4], giving the following approximations :

$$p_{Torrieri} = \sum_{i=t+1}^{n} C_n^i p_e^i (1-p_e)^{n-i} \frac{\max(d, i)}{n}, \qquad (2)$$

[1] See e.g. [5] for general information about ECCs.
[2] In this paper, $C_n^i = \frac{n!}{i!(n-i)!}$

$$p_{Labeau} = \sum_{i=t+1}^{n} C_n^i p_e^i (1-p_e)^{n-i} \frac{\min(i+t,n)}{n}. \quad (3)$$

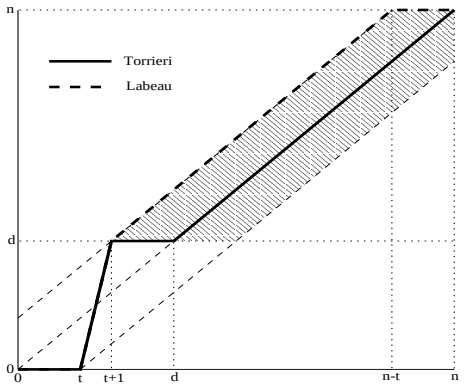

Figure 1: Post-decoding weight f_i with respect to the error pattern weight i for a radius-t sphere decoding. Thick lines represent the two approximations considered. We can see that *Labeau*'s approximation is also an upper bound for the f_i factor, as it has to stay inside the hatched zone limiting to t the number of bits modified by the decoding process.

Both approximations agree on $f_{t+1} = d$, but they begin to differ for $i = t+2$, giving respectively d and $d+1$. These low-weight terms are the most important ones when the BER p_e is small. Intuitively, *Torrieri* assumes that the decoder has no reason to introduce more ones than zeros while decoding, which should be true for words of weight close to $\frac{n}{2}$. On the other hand, *Labeau* considers that most of the time, error patterns will lie at the surface of a codeword sphere, hence leading the decoder to modify t bits, and most of them will be turned from 0 to 1, as the majority of erroneous words have a low weight. An exact analysis will strenghten his intuitions in section 2.3.

2.2 Structure of erroneous words

We have seen that approximating BERs amounts to estimating f_i factors related to the weights of words after decoding. If we know the weight distribution A_i of the code[3], we can use these coefficients in order to get exact values for f_i. Above that, we also have to justify the implicit assumption made in (1) that errors are equally distributed between the n bits of erroneous words. As a matter of fact, the answer depends not only on the weight distribution, but also on the actual implementation — i.e. the generator matrix — of the code. The more "continuous" the encoder is, the smaller values the BER takes [6, 7]. Systematic encoding proves to be optimum from this point of view [8].

In the case of cyclic codes, we can prove that the repartition of bits between information and redundancy ones is equitable [3], in the sense that if a word is decoded into a codeword of weight w, the number of information bits affected will on the average be $\frac{k}{n}w$, leading to a BER for information bits equal to the global BER. For non-cyclic codes, no general result holds, but assuming this repartition is generally close to the truth. In the case of perfect codes under a cyclic implementation, the exact BER can be expressed by the following equation :

$$p = \sum_{i=t+1}^{n} C_n^i p_e^i (1-p_e)^{n-i} \frac{\sum_{j=i-t}^{\min(i+t,n)} j N_{j,i} A_j}{n \sum_{j=i-t}^{\min(i+t,n)} N_{j,i} A_j} \quad (4)$$

$$= \frac{1}{n} \sum_{i=t+1}^{n} p_e^i (1-p_e)^{n-i} \sum_{j=i-t}^{\min(i+t,n)} j N_{j,i} A_j, \quad (5)$$

where $N_{j,i}$ denotes the number of elements of weight i belonging to the radius-t sphere of a codeword of weight j. For a perfect code, $\sum_{j=i-t}^{\min(i+t,n)} N_{j,i} A_j$ is equal to C_n^i, as different ways to count the number of words of weight i. $N_{j,i}$ factors can be easily computed as the following, where maximum values of the summation index j are chosen in order to stay inside the range of coefficients for which combinatorial symbols are defined :

$$N_{i,l} = \sum_{j=0}^{\min(\lfloor \frac{t-\delta}{2} \rfloor, l, n-i)} C_i^{\delta+j} C_{n-i}^{j}, \quad \delta = i-l \geq 0, \quad (6)$$

$$N_{i,l} = \sum_{j=0}^{\min(\lfloor \frac{t-\delta}{2} \rfloor, i, n-l)} C_i^{j} C_{n-i}^{\delta+j}, \quad \delta = l-i \geq 0. \quad (7)$$

2.3 Comparison between *Torrieri* and *Labeau*'s approximations

Considering our analysis based on perfect codes, we can use the exact values we have mentioned to compute the error of *Torrieri* and *Labeau*'s approximations. As both produce exact values for the trivial repetition codes, we compare them for the other perfect codes, i.e. the Hamming family and the Golay code. Figure 2 shows that *Torrieri*'s approximation matches the trivial $(\frac{1}{2}, \frac{1}{2})$ point due to the fact that it follows the line $f_i = i$. However, for lower channel BERs, its asymptotical convergence towards the exact value — i.e. the one only influenced by the first term, of weight $t+1$ — is significantly slower than *Labeau*'s, proving that the latter provides a better approximation of the most important lower-weight terms.

If we take the example of a Hamming code, we can compute that expressing the exact f_{t+2} factor as a linear combination of *Torrieri*'s and *Labeau*'s values, the combination weights are respectively $\frac{1}{n-2}$ and $\frac{n-3}{n-2}$, once again ruling in favour of the second one.

3 EXTENSION TO NON-PERFECT CODES

3.1 Influence of code density

In the case of non-perfect codes, spheres of radius t do not fully cover the space and some words farther than t

[3] A_i, $0 \leq i \leq n$, denotes the number of codewords of weight i, with particular values $A_0 = 1$ and $A_i = 0$, $1 \leq i < d$, for a linear code of minimum distance d

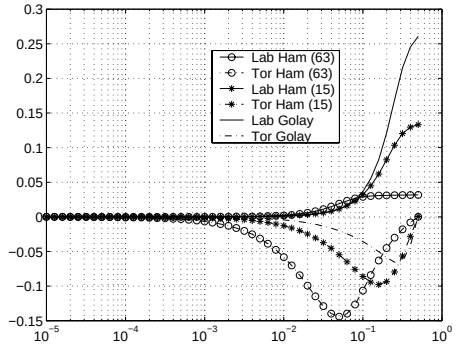

Figure 2: Relative differences between BER approximations of *Labeau* (Lab) and *Torrieri* (Tor) with respect to exact values for Hamming (63, 57), Hamming (15, 11) and Golay (23, 12) codes. Vertical values are computed as $(\frac{p_{approximation}}{p_{exact}} - 1)$, depending on the channel BER.

from any codeword appear. Some advanced decoders are able to decode these words into the nearest codewords, but most of them simply leave these words undecoded. In this case, we simply have $f_i = i$ for such a word. *Torrieri* mentions the corresponding BER approximation, to be used for "loosely packed codes", when "decoding failures — words falling out of radius-t spheres — are the predominant error mechanism" [3].

We have found that one simple but efficient synthesis between the behaviour of decoded and undecoded words is to take into account the "density" of the code, defined as the fraction η of words that belong to spheres of radius t around codewords. Using it as the probability for a word to be decoded, we can approximate the BER by a simple linear combination between *Labeau*'s approximation (3) and the undecoded word case based on $f_i = i$:

$$\eta = \frac{2^{n-k}\sum_{i=0}^{t} C_n^i}{2^n} = 2^{-k}\sum_{i=0}^{t} C_n^i, \quad (8)$$

$$p = \sum_{i=t+1}^{n} C_n^i p_e^i (1-p_e)^{n-i} \cdot \frac{\eta \min(i+t,n) + (1-\eta)i}{n}. \quad (9)$$

Table 1 shows the dramatical improvement introduced by taking into account the code density.

3.2 Binomial weight distribution approximation

By taking into account undecoded words, we have significantly improved the BER approximation for non-perfect codes. Let us now try to find a better approximation than *Torrieri*'s or *Labeau*'s for words being decoded. This is especially important for high channel BERs, when the $t+1$ error term is no more dominant. As the weight distribution is not known, we will use its binomial approximation, which seems more or less accepted for usual codes [10]. This approximation simply states that the weight distribution has a shape $A_i \approx \beta C_n^i$, with β chosen in order to obtain the desired total number of codewords. Actually, we use this approximations for words of weights between d and $n-d$, considering that $A_0 = 1$, $A_i = 0$, $1 \leq i < d$, and assuming a symmetric distribution with respect to $\frac{n}{2}$.

Using this weight distribution, we could simply use it in a straightforward BER computation in which we would include undecoded words. Actually, we would simply use (5) without any modification, but $\sum_{j=i-t}^{\min(i+t,n)} N_{j,i}A_j$ would this time be less than C_n^i. However, even if the binomial approximation approaches the shape of the weight distribution, it is precisely not accurate enough to be used for determining the number of words decoded and undecoded. A far better solution is to use the code density to this aim, while considering the binomial as in (4) to compute the average weight of decoded words :

$$p = \sum_{i=t+1}^{n} C_n^i p_e^i (1-p_e)^{n-i} \cdot \left(\eta \frac{\sum_{j=i-t}^{\min(i+t,n)} jN_{j,i}A_j}{n\sum_{j=i-t}^{\min(i+t,n)} N_{j,i}A_j} + (1-\eta)i \right) \quad (10)$$

This final estimation is accurate for both low-weigh terms — critical at a low channel BER — and middle-weight ones — those of weight close to $\frac{n}{2}$, dominant when the channel BER approaches 0.5. Considering for instance Hamming's $(15, 11)$ code, we can see in figure 2 that *Torrieri* and *Labeau*'s approximations present errors of up to respectively 10 and 15% on the BER. Our new approximation has an error limited to 1%, whatever the channel BER. Figure 4 shows the binomial-based f_i factor, actually very close to the true one, to be compared with figure 1's approximations.

n	k	t	density	no modification	modified
15	7	2	.47	31.88%	4.06%
31	21	2	.49	30.63%	3.74%
31	16	3	.15	54.96%	-1.33%
255	239	2	.50	25.37%	0.20%
255	231	3	.16	54.79%	-0.62%

Table 1: Error percentages — defined as in figure 2 — for *Labeau*'s approximation with and without density modification for various 2 or 3–error correcting BCH codes whose distribution is known [9]. Asymptotic values for small channel error probabilities are given. Code density is defined in (8).

4 CONCLUSIONS, APPLICATIONS AND FURTHER TOPICS

Looking for an accurate BER approximation for linear block codes, we compared two approximations found in

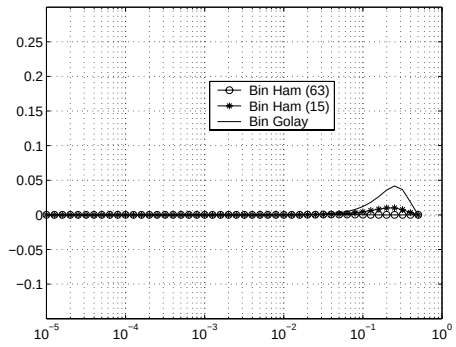

Figure 3: Relative errors of binomial-based BER approximations for Hamming (63, 57), Hamming (15, 11) and Golay (23, 12) codes. Same scale as figure 2.

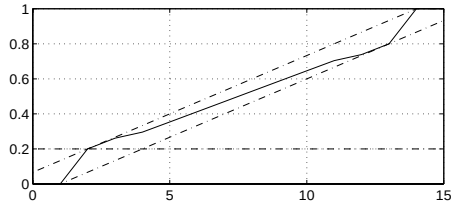

Figure 4: $\frac{f_i}{n}$ factor as a function of i for the Hamming (15, 11) code, to be compared to figure 1.

the literature. *Labeau*'s proved to be better, even if its lack of symmetry appears at high channel BERs. We have first introduced a weighting factor related to the code density, providing a dramatic improvement of the asymptotic accuracy of these approximations by considering undecoded words. Then, considerations based on the binomial weight distribution approximation have enabled a significant improvement in the high channel BER domain, where both existing approximations were poor due to a bad modeling of higher weight terms.

Our work benefits from two kinds of informations found in the weight distribution : the local density providing the probability for an error pattern to be decoded and the weight profile allowing the estimation of post-decoding weights. Separating these two points proved to be critical. Applications include communication systems design, particularly in fields such as joint source-channel coding where ECCs have to be closely matched to source data.

As another possible use, consider that for very noisy channels, correcting less than t errors may produce a better BER [8, 11]. Our expression not only applies straightforwardly to these sub-decoding schemes, but it could also be used in order to find the best decoding schemes by comparing results of different decoding radii.

Further research can be achieved in various directions. Finding bounds on the accuracy of our approximation could be a first challenge. Extending its use is certainly another one, towards other classes of codes (non-binary, convolutional, turbo, ...) or advanced (soft) decoders.

REFERENCES

[1] Claude E. Shannon and W. Weaver. *The Mathematical Theory of Communication*. University of Illinois Press, 1949.

[2] Don J. Torrieri. The information-bit error rate for block codes. *IEEE Transactions on Communications*, 32(4):474–476, April 1984.

[3] Don Torrieri. Information-bit, information-symbol, and decoded-symbol error rates for linear block codes. *IEEE Transactions on Communications*, 36(5):613–617, May 1988.

[4] Fabrice Labeau, Claude Desset, Benoît Macq, and Luc Vandendorpe. Approximating the protection offered by a channel code in terms of bit error rate. In *EUSIPCO'98*, volume 1, pages 33–36, Rhodes, Greece, September 1998.

[5] Richard E. Blahut. *Theory and Practice of Error Control Codes*. Addison-Wesley, 1983.

[6] Ali S. Khayrallah. The expansion factor of error-control codes. *IEEE Transactions on Information Theory*, 41(5):1452–1457, September 1995.

[7] Ali S. Khayrallah. On the minimum bit-error rate of linear codes. *IEEE Transactions on Information Theory*, 41(5):1457–1466, September 1995.

[8] Aaron B. Kiely, John T. Coffey, and Mark R. Bell. Optimal information bit decoding of linear block codes. *IEEE Transactions on Information Theory*, 41(1):130–140, January 1995.

[9] Shu Lin and Daniel J. Costello, Jr. *Error Control Coding: Fundamentals and Applications*. Prentice-Hall, 1983.

[10] Min-Goo Kim and Yae Hong Lee. Undetected error probabilities of binary primitive BCH codes for both error correction and detection. *IEEE Transactions on Communications*, 44(5):575–580, May 1996.

[11] Philippe Delsarte. Partial-optimal piecewise decoding of linear codes. *IEEE Transactions on Information Theory*, 24(1):70–75, January 1978.