

New Lower Bounds on the Error Probability of a Given Block Code

Asaf Cohen

Faculty of Electrical Engineering
Technion - Israel Institute of Technology
Technion City, Haifa 32000, Israel
soofsoof@tx.technion.ac.il

Neri Merhav

Faculty of Electrical Engineering
Technion - Israel Institute of Technology
Technion City, Haifa 32000, Israel
merhav@ee.technion.ac.il

Abstract

A novel technique for deriving lower bounds on the error probability when communicating one of M signals over a communication channel is proposed. At the basis of the technique, stands an improvement on a recent lower bound on the probability of a union of events by de Caen. The new bound includes a function which can be optimized in order to achieve the tightest results. By applying this bound to the problem of lower bounding the error probability, while suggesting an appropriate optimization function, in the spirit of the relevant channel model and type of the code, new bounds on the error probability can be derived. In this talk, we apply the new bound to the problem of lower bounding the error probability of binary linear codes over the Binary Symmetric Channel (BSC). The resulting bound improves on the latest bound appearing in the current literature, by Keren and Litsyn.

1. Introduction

We consider the case of transmitting one of $M = 2^K$ equiprobable binary codewords of length N , $\mathbf{c}_0, \mathbf{c}_1, \dots, \mathbf{c}_{M-1}$, over a BSC channel and maximum likelihood (ML) decoding. Denote by p the channel crossover probability. Let \mathbf{x} be the received word. The optimal decoder, namely, the ML decoder, chooses the closest of the M codewords to \mathbf{x} in the Hamming sense, i.e., $\hat{i} = \arg \min_i d_H(\mathbf{c}_i, \mathbf{x})$, where $d_H(\mathbf{c}_i, \mathbf{x})$ is the Hamming distance between \mathbf{c}_i and \mathbf{x} . Denote by $w(\mathbf{x})$ the Hamming weight of the word \mathbf{x} . Assume \mathbf{c}_0 is the all-zero codeword. The probability of error given that \mathbf{c}_0 was sent is

$$P(\varepsilon|\mathbf{c}_0) = P(\cup_{i \neq 0} \varepsilon_{0i} | \mathbf{c}_0), \quad (1)$$

where

$$\begin{aligned} \varepsilon_{0i} &= \{\mathbf{x} \in GF(2)^N : d_H(\mathbf{x}, \mathbf{c}_i) < d_H(\mathbf{x}, \mathbf{c}_0)\} \\ &= \{\mathbf{x} \in GF(2)^N : w(\mathbf{x} + \mathbf{c}_i) < w(\mathbf{x})\}. \end{aligned} \quad (2)$$

Our goal is to lower bound the error probability given in (1). When the code used is a linear code \mathcal{C} , $P(\varepsilon|\mathbf{c}_0) = P(\varepsilon)$. In this case, we wish to express the bound in terms of the code's weight enumeration. To the author's knowledge, the best known lower bounds, in this context, are Keren and Litsyn's bound [1] and the sphere packing bound (see, for example, [2]). The best known upper bound, [2], is due to Poltyrev.

2. Analysis

Let $\{A_i\}_{i \in \mathcal{I}}$ denote a finite family of events in a probability space (Ω, \mathcal{F}, P) . For each $x \in \Omega$ define $deg(x) = |\{i \in \mathcal{I} : x \in A_i\}|$. The basis of our results is the following

Theorem 1

$$P\left(\bigcup_{i \in \mathcal{I}} A_i\right) \geq \sum_{i \in \mathcal{I}} \frac{(\sum_{x \in A_i} p(x) m_i(x))^2}{\sum_{j \in \mathcal{I}} \sum_{x \in A_i \cap A_j} p(x) m_i^2(x)} \quad (3)$$

where $m_i(x)$ is any real function on Ω such that the sums in the r.h.s of (3) converge. Equality in (3) is achieved when

$$m_i(x) = m(x) = \frac{1}{deg(x)}. \quad (4)$$

The essence of the bound given in Theorem 1 is the ability to choose an appropriate function $m_i(x)$. When seeking such a function, remember that the optimal value of $m_i(x)$ is $1/deg(x)$. While the function $deg(x)$ is complex to evaluate and usually requires more than the available information on the sets $\{A_i\}_{i \in \mathcal{I}}$, its behavior possesses the guidelines for choosing a competent family of approximations. Moreover, since de Caen's bound [3] can be achieved by choosing $m_i(x) \equiv 1$, requiring that any family of approximations will include this choice, and optimizing the bound over this family, results in a bound which is always at least as tight as de Caen's. Although this bound does not depend only on the $P(A_i)$'s and $P(A_i \cap A_j)$'s, the computational complexity depends on our choice of $m_i(x)$. A proper choice of this function may result in the same complexity as de Caen's bound, or yield a more accurate result with only slightly higher complexity, when, for example, chosen to be constant on subsets of the $P(A_i)$'s. Thus, an additional criterion when choosing a family of approximations, is the effort required in evaluating the sums in (3).

We wish to apply the bound in (3) to (1). Define $\mathcal{I} = \{1, \dots, M-1\}$ and $A_i = \varepsilon_{0i}$. $p(\mathbf{x}|\mathbf{c}_0)$ is given by

$$\begin{aligned} p(\mathbf{x}|\mathbf{c}_0) &= p^{d_H(\mathbf{x}, \mathbf{c}_0)}(1-p)^{N-d_H(\mathbf{x}, \mathbf{c}_0)} \\ &= p^{w(\mathbf{x})}(1-p)^{N-w(\mathbf{x})}. \end{aligned} \quad (5)$$

Under these definitions, the computation of (3) requires the summations over $\mathbf{x} \in \varepsilon_{0i}$ and $\mathbf{x} \in \varepsilon_{0i} \cap \varepsilon_{0j}$. The summands are $p(\mathbf{x}|\mathbf{c}_0)m_i(\mathbf{x})$ and $p(\mathbf{x}|\mathbf{c}_0)m_i^2(\mathbf{x})$, respectively. While the dependence of $p(\mathbf{x}|\mathbf{c}_0)$ on \mathbf{x} is only through $w(\mathbf{x})$, $m_i(\mathbf{x})$ is a function to be optimized and hence might, in general, be chosen to be different for each \mathbf{x} . Thus, to avoid a tedious evaluation of these sums, we prefer to reduce the degrees of freedom in choosing $m_i(\mathbf{x})$ by the restriction

$$m_i(\mathbf{x}) = \eta_i(w(\mathbf{x})), \quad \eta_i: \mathbb{Z}^+ \mapsto \mathbb{R}. \quad (6)$$

Clearly, since $\deg(\mathbf{x}|\mathbf{c}_0)$ is not likely to depend only on $w(\mathbf{x})$ when non-trivial codes are discussed, we may assume that the optimal value for $m_i(\mathbf{x})$ cannot be achieved by any function η_i . Nevertheless, we will discover that the function η_i may still be chosen to yield tighter bounds than the one achieved with $m_i(\mathbf{x}) \equiv 1$. Thus, we have

$$P(\varepsilon) \geq \sum_{i=1}^M \frac{(\sum_{\mathbf{x} \in \varepsilon_{0i}} p^{w(\mathbf{x})}(1-p)^{N-w(\mathbf{x})} \eta_i(w(\mathbf{x})))^2}{\sum_{\mathbf{x} \in \varepsilon_{0i}} p^{w(\mathbf{x})}(1-p)^{N-w(\mathbf{x})} \eta_i^2(w(\mathbf{x})) + \sum_{j \neq i} \sum_{\mathbf{x} \in \varepsilon_{0i} \cap \varepsilon_{0j}} p^{w(\mathbf{x})}(1-p)^{N-w(\mathbf{x})} \eta_i^2(w(\mathbf{x}))}, \quad (7)$$

where $\eta_i: \mathbb{Z}^+ \mapsto \mathbb{R}$ is any function to be optimized.

To evaluate the sums in the right hand side (r.h.s.) of (7), suppose that the chosen function η_i depends on the code is only through its weight enumeration. In this case, the same applies to the sums over ε_{0i} . However, the sum over $\varepsilon_{0i} \cap \varepsilon_{0j}$ depends on $w(\mathbf{c}_i \mathbf{c}_j)$. Yet, in [4], we show that the value of this sum is monotonically increasing in $w(\mathbf{c}_i \mathbf{c}_j)$, thus, an upper bound on $w(\mathbf{c}_i \mathbf{c}_j)$, which depends only on $w(\mathbf{c}_i)$ and $w(\mathbf{c}_j)$, yields an upper bound on the sum over $\varepsilon_{0i} \cap \varepsilon_{0j}$ and thus a lower bound on the r.h.s. of (7). The resulting lower bound on the error probability depends on the code only through its weight enumeration.

It is left to choose an appropriate function η_i . As mentioned earlier, the key step is the evaluation of $\deg(\mathbf{x}|\mathbf{c}_0)$. Observe that

$$\begin{aligned} \deg(\mathbf{x}|\mathbf{c}_0) &= |\{\mathbf{c}_i \in \mathcal{C}, i \neq 0 : d_H(\mathbf{x}, \mathbf{c}_i) < d_H(\mathbf{x}, \mathbf{c}_0)\}| \\ &= |\{\mathbf{c}_i \in \mathcal{C}, i \neq 0 : w(\mathbf{x} + \mathbf{c}_i) < w(\mathbf{x})\}|. \end{aligned} \quad (8)$$

Thus, $\deg(\mathbf{x}|\mathbf{c}_0)$ is the number of words in the coset $\mathcal{C} + \mathbf{x}$, with weight smaller than $w(\mathbf{x})$. The computation of the weight enumeration of this coset is complex, usually requiring more than the weight distribution of the code. However, several approximations may be offered. Clearly,

$$\deg(\mathbf{x}|\mathbf{c}_0) = \begin{cases} 0 & w(\mathbf{x}) \leq \lfloor \frac{d-1}{2} \rfloor \\ \sum_{i=0}^{w(\mathbf{x})-1} B_i^{\mathbf{x}} & \lfloor \frac{d-1}{2} \rfloor < w(\mathbf{x}) < N \\ |\mathcal{C}| - 1 & w(\mathbf{x}) = N, \end{cases} \quad (9)$$

where $B_i^{\mathbf{x}}$ is the number of words of weight i in the coset $\mathcal{C} + \mathbf{x}$. Thus, $\deg(\mathbf{x}|\mathbf{c}_0)$ is monotonically increasing in $w(\mathbf{x})$, with known values for $w(\mathbf{x}) = \lfloor \frac{d-1}{2} \rfloor$ and $w(\mathbf{x}) = N$. Consequently, a possible approximation for $\deg(\mathbf{x}|\mathbf{c}_0)$, $\widetilde{\deg}(w(\mathbf{x}))$, which depends only on $w(\mathbf{x})$, the size of the code and its minimum distance d , might be

$$\widetilde{\deg}(w(\mathbf{x})) = (|\mathcal{C}| - 1) \left(\frac{w(\mathbf{x}) - \lfloor \frac{d-1}{2} \rfloor}{N - \lfloor \frac{d-1}{2} \rfloor} \right)^a, \quad w(\mathbf{x}) > \left\lfloor \frac{d-1}{2} \right\rfloor, \quad (10)$$

where $a \geq 0$ is an arbitrary parameter, i.e., when substituting $\eta_i(w(\mathbf{x})) = 1/\widetilde{\deg}(w(\mathbf{x}))$ in (7), the resulting bound can be optimized over a . Note that the bound corresponding to $a = 0$ is the straightforward application of de Caen's bound to this problem. It is important to note that the parametric family given in (10) is only one of several approximations which can be derived. Each approximation, substituted in (7), will result in a new lower bound on the decoding error probability for the BSC.

Although the resulting bound depends only on the weight enumeration of the code, its evaluation is tedious for long codes. We consider two variations of this bound, suggested in [1], which will both reduce the complexity and improve the performance. Denote by t the *covering radius* of the code. Namely, t is the maximum number of errors that can be corrected. More than t errors cannot be corrected. Clearly,

$$P(\varepsilon) = P(\varepsilon, w(\mathbf{x}) \leq t) + P(w(\mathbf{x}) > t). \quad (11)$$

Thus, when evaluating the sums in (7), we may assume no more than t errors were made, then add the probability of *all* the words with weight higher than t . As mentioned in [1], when lower bounds on the error probability are discussed, an upper bound $M \geq t$ can be used if t is not known. Moreover, note that

$$P(\varepsilon, w(\mathbf{x}) \leq M) \geq P_{\mathcal{C}_i^*}(\varepsilon, w(\mathbf{x}) \leq M), \quad (12)$$

where $P_{\mathcal{C}_i^*}(\varepsilon)$ is the error probability when instead of the whole code, only the subcode \mathcal{C}_i^* is used, and $\mathcal{C}_i^* = \{\mathbf{c} \in \mathcal{C} : w(\mathbf{c}) = i\} \cup \{\mathbf{c}_0\}$. Thus, we may compute the bound disregarding all codewords of weight other than i . Although the numerical analysis shows that best results are achieved with $i = d$, we prefer this general form for future reference. Note that when only the subcode \mathcal{C}_i^* is used, the approximation in (10) should be altered by replacing $|\mathcal{C}|$ with $|\mathcal{C}_i^*|$.

As an example, numeric results are given for the code BCH(63,24). Figure 1 includes the new bound, using the approximation given in (10), Keren-Litsyn's lower bound, the sphere packing bound and the upper bound by Poltyrev. It is clear that the new bound improves on Keren-Litsyn's bound, derived using de Caen's inequality, for every value of p . Moreover, in [4], we discuss the asymptotic properties (as $N \rightarrow \infty$) of the new bound. We show that the new bound is exponentially tighter than the de Caen-based bound, and identify the optimal choice of the function η_i . In this talk, present only the main results in this context.

We wish to calculate the bound on the error exponent, resulting from the bound in (7), when only the subcode \mathcal{C}_i^* is used. Clearly, since the denominator of the r.h.s. of (7) is a sum of two expressions, the exponential behavior of the bound in (7) depends on which expression dominates. In [4], we show that this observation translates to a *condition on the code*, which determines the value of the bound in each case. When the difference between the triplets error exponent and the pairwise error exponent is not too small, i.e., the rate of the code is not too large, it can be shown that the first expression dominates, and the resulting bound on the error exponent, valid for any $i \geq d$, is given by

$$-\frac{1}{N} \log P(\varepsilon) \leq -\delta_i \log \sqrt{4p(1-p)} - E_B^{\delta_i}, \quad (13)$$

where $\delta_i = \frac{i}{N}$ and $E_B^{\delta_i} = \frac{1}{N} \log |\mathcal{C}_i^*|$. Namely, one can use the union bound to derive a valid lower bound on the error probability (in this case, we say that the *union bound analysis applies*). Clearly, since (13) applies for any $i \geq d$, by optimizing over i , one can achieve the true error exponent of the code. However, as mentioned earlier, the union bound analysis applies only at low rates. Our main result is that a proper choice of the function η_i can extend the range of rates for which the union bound analysis applies, thus achieving a tighter bound on the error exponent than that achieved by a de Caen-based bound.

Finally, we note that the framework presented here can be used for different channel models. By applying the new bound on the probability of a union, given in (3), with the channel-specific derivations, such as (5) or (8), new lower bounds on the error probability can be derived. For example, [4] includes analogous derivations for the additive white gaussian noise channel. Again, the resulting bound improves on the latest result appearing in the current literature, by Seguin [5].

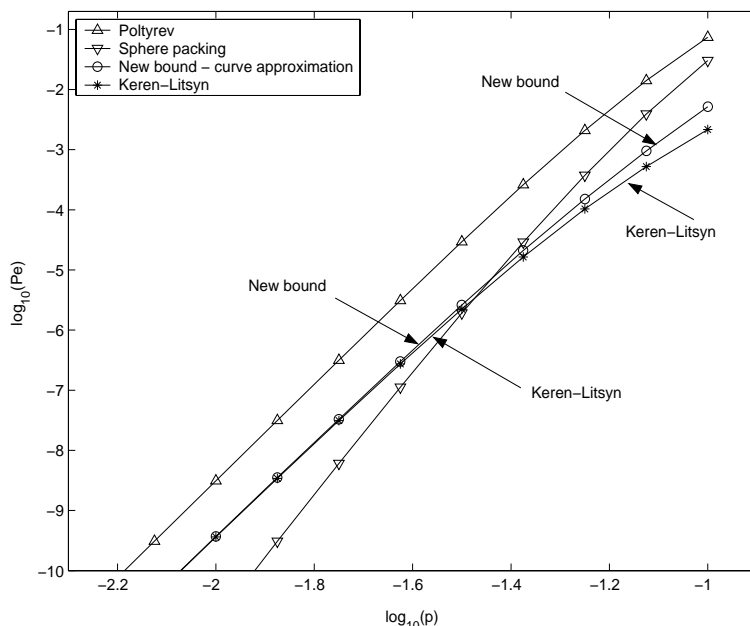


Figure 1. Bounds on the decoding error probability of BCH(63,24) code, BSC.

3. References

- [1] O. Keren and S. Litsyn, "A lower bound on the probability of decoding error over a BSC channel," *The 21st IEEE Electrical and Electronic Engineers in Israel*, pp. 271–273, 2000.
- [2] G. Poltyrev, "Bounds on decoding error probability of binary linear codes via their spectra," *IEEE Trans. Inform. Theory*, vol. 40, no. 4, pp. 1284–1292, July 1994.
- [3] D. de Caen, "A lower bound on the probability of a union," *Discr. Math.*, vol. 169, pp. 217–220, 1997.
- [4] A. Cohen, "Lower bounds on the error probability of a given block code," M.S. thesis, Technion, I.I.T., 2002.
- [5] G. E. Seguin, "A lower bound on the error probability for signals in white Gaussian noise," *IEEE Trans. Inform. Theory*, vol. 44, no. 7, pp. 3168–3175, November 1998.