Algebraic Approach for the Reconstruction of Linear and Convolutional Error Correcting Codes

Johann Barbier, Guillaume Sicot, and Sébastien Houcke

Abstract—In this paper we present a generic approach for the problem of the blind estimation of the parameters of linear and convolutional error correcting codes. In a non-cooperative context, an adversary has only access to the noised transmission he has intercepted. The intercepter has no knowledge about the parameters used by the legal users. So, before having access to the information he has first to blindly estimate the parameters of the error correcting code of the communication. The presented approach has the main advantage that the problem of reconstruction of such codes can be expressed in a very simple way. This allows us to evaluate theoretical bounds on the complexity of the reconstruction process but also bounds on the estimation rate. We show that some classical reconstruction techniques are optimal and also explain why some of them have theoretical complexities greater than these experimentally observed.

Keywords—Blind estimation parameters, error correcting codes, non-cooperative context, reconstruction algorithm.

I. INTRODUCTION

In all communication where the channel is noisy, the use of an error correcting code is mandatory. This code adds some redundant bits to the information to protect it from noise. In a non-cooperative context, these additional bits make the task of an adversary harder. Indeed, before having an access to the information, the intercepter must first locate the main information and the redundant bits and then, correct the noise introduced by the channel. In such a context, the adversary has only acces to the intercepted noised communication with no knowledge of the parameters of the error correcting code. So, he has to estimate blindly these parameters in order to decode the information. Many papers in cryptography provides techniques to cryptanalyse ciphered texts, but authors always make the hypothesis they have access to noiseless information with no extra bits. In the practical context of communication interception, this hypothesis does not hold. Surprisingly, only few papers deal with the reconstruction of error correcting codes. Rice [11] was the first one to present a technique to determine the parameters of convolutional encoders of rate \( \frac{1}{n} \), then Filiol generalized it [7], [6], [5] for all the rates and also for punctured convolutional encoders. Barbier [1] introduced an algebraic approach to greatly improve Filiol’s technique. He also developed a method to reconstruct turbo-code encoders. The binary linear codes have also been studied. Planquette [10] adapted algorithms for finding codewords of small weight [3], [8], [15] to estimate the parameters of binary linear codes. Then, Valemoibs [17], [16] approaches the problem through the scope of statistical hypothesis tests and pointed out optimal thresholds for these tests. He also introduced a criterion based on the rank of a matrix composed of the intercepted bits. He claimed the criterion is not sufficient and propose to find codewords of small Hamming weight. Its technique has been recently improved by Cluzeau [4]. Then Burel [2] solved the problem for noiseless channels and Sicot and Houcke [14], [12], [13] proposed to blindly estimate the parameters using Gauss pivot for noisy channels.

In this paper, we propose a generic approach for the estimation of binary linear and convolutional codes parameters. This approach is based on linear algebra and turns the problems of error correcting codes reconstruction in a very simple way. This allows us to evaluate theoretical bounds on the complexity of reconstruction algorithms but also bounds on the success rate of these algorithms. We also show that the algorithms proposed by Sicot and Houcke [12], Filiol [5] and Barbier [1] are optimal and explain why the experimental results Filiol and Barbier obtained are better than these they theoretically claimed.

First, we will introduce the notations and translate the reconstruction problem into algebraic equations. Then, we will study the intrinsic probability of detecting the parity checks when the estimated dimension of the code is the right one or not. In the fourth section, we propose a compromise between the detection probability and the false positive rate. In the last section, we measure theoretical detection probability bounds and the complexity of resolving the problem of the reconstruction of linear and convolutional error correcting codes. Finally, we conclude showing the optimality of some techniques presented above and explain why Filiol and Barbier observed better experimental results than these they theoretically proved.

II. THE ALGEBRAIC APPROACH

Let \( G \) be the generator matrix of the binary linear error correcting code \( C \), of rate \( \rho = \frac{k}{n} \), we want to reconstruct. The following technique also fits when the output of the error correcting code is interleaved. In that case, we consider \( G = P \times \text{diag}(G) \), where \( P \) is the permutation matrix of the interleaver and \( \text{diag}(G) \) defined by

\[
\begin{pmatrix}
G & 0 & 0 \\
0 & \ddots & 0 \\
0 & 0 & G
\end{pmatrix}.
\]

Let us denote \( (y_i)_{i=1\ldots m} \) the codewords generated during the communication, and \( (\tilde{y}_i)_{i=1\ldots m} \) the noised codewords
intercepted by the adversary.

\[ y^i_j = y^i_j + e_{ij}, \quad \forall i = 1 \ldots m, \quad j = 1 \ldots n, \quad (1) \]

and \( e_{ij} = 1 \) with probability \( P_e \) and \( e_{ij} = 0 \) with probability \((1 - P_e)\). \( P_e \) is the error rate of the channel. Without loss of generality we will consider \( m \) as constant and \( m \gg n \). Let \( H \) be the dual code of \( C \), so that for all parity checks \( h \in H \), \( < h, y > = 0 \ \forall i \). Since recognizing the code \( C \) is equivalent to estimating the dual code, we consider now the estimation of \( H \).

We define the \( m \times n \) matrices \( H \) with the codewords by \([H]_{ij} = y^i_j \) and \( H \) with the intercepted noised codewords by \([H']_{ij} = y^i_j + e_{ij} \). If \( E = [e_{ij}] \) then \( H = H + E \).

First, we have to consider that the adversary is not synchronized, i.e. the first intercepted bit is not necessary the first of a noised codeword. We denote \( d \), the desynchronization parameter, \( d \leq n \), and define the \( m \times n \) matrices \( H(d) \) and \( H(d) \) by \([H(d)]_{ij} = y^i_j + e_{ij} \) \( \forall i = 1 \ldots m, \quad j = 1 \ldots n - d \), \([H'(d)]_{ij} = y^i_j + e_{ij+1} \), \( \forall i = 1 \ldots m, \quad j = n - d + 1 \ldots n \), and \( H(d) = H(d) + E(d) \).

Since the adversary has to blindly estimate the parameter \( n \), he will have to try all the possible values for \( n \), \( n_a \) from 2 to \( n \). For a tested value \( n_a \), we build the \( m \times n_a \) matrix \( H(n_a, d) \) with the intercepted bits. As previously, \( H(n_a, d) = H(n_a, d) + E(n_a, d) \). \( H(n_a, d) \) is split into two sub-matrices \( H_1(n_a, d) \) and \( H_2(n_a, d) \), where \( H_1(n_a, d) \) is a \( n_a \) square matrix composed of the first \( n_a \) rows of \( H(n_a, d) \). We get two equations,

\[ H_1(n_a, d) = H_1(n_a, d) + E_1, \quad (2) \]
\[ H_2(n_a, d) = H_2(n_a, d) + E_2, \quad (3) \]

where \( E_i = E_i(n_a, d) \). Moreover \( H = \text{Ker}(H(n_a, d)) \), \( \forall n_a \in \mathbb{N} \). So, the problem of binary error correcting codes reconstruction is equivalent to determine \( \text{Ker}(H(n_a, d)) \), the problem of the reconstruction of convolutional codes [1], [5] and of turbo-codes [1] can be expressed in the same way. In these cases, the \( H(n_a, d) \) matrices are slightly different but always built from the intercepted bits.

III. Optimal Detection of the Parity Checks

To detect the parity check we will use the rank criterion introduced by Valemebois [16] and mainly used by Burel [2], Sicot and Houcke [12], Fililo [5] and Barbier [1]. It appears that the rank of \( H_1(n_a, d) \) does not behave in the same way wether \( n_a = \beta n \) or not, where \( \beta \) is an integer.

A. Case \( n_a = \beta n \)

In the case \( n_a \neq \beta n \), no relation may exist between the columns of \( H_1(n_a, d) \), otherwise, the dimension of \( C \) would be smaller than \( n \). \( H_1(n_a, d) \) and then \( H_1(n_a, d) \) can be considered as a random binary matrices.

Theorem 1: In the case of \( n_a \neq \beta n \),

\[ P_r(\text{rk}(H_1(n_a, d)) < n_a) = 1 - \prod_{i=0}^{n_a-1} (1 - 2^{-i-n_a}). \quad (4) \]

Proof: straightforward application of the classical result [9] (p. 455) : the probability that a \( k \times l \) random binary matrix be full rank is

\[ \prod_{i=0}^{n_a-1} (1 - 2^{-i-k}). \]

We denote \( P^0_{fa} = P_r(\text{rg}(H_1(n_a, d)) < n_a) \). This is the probability to detect a relation between the columns of \( H_1(n_a, d) \) which is not a parity check of \( H_1(n_a, d) \). We deduce a corollary which bounds this probability

Corollary 1:

\[ 2^{-n_a}(1 - 2^{-n_a}) \leq P^0_{fa} \leq (1 - 2^{-n_a}). \quad (5) \]

Proof: see appendix.

The probability \( P^0_{det} \) to detect a real parity check is trivally negligible.

B. Case \( n_a = \beta n \)

The following theorem is the central theorem which gives us a necessary and sufficient condition to detect a parity check. Since this condition is necessary and sufficient, computing the kernel of \( H_1(n_a, d) \) is an optimal technique for estimating the parameters of binary linear or convolutional error correcting codes.

Theorem 2: Let \( h \) be a parity check, i.e. \( h \in \text{Ker}(H_1(n_a, d)) \) then,

\[ h \in \text{Ker}(H_1(n_a, d)) \text{ if and only if } \sum_{j=1}^{n_a} [E^*_1]_{ij}h_j \equiv 0 \mod 2 \ \forall j = 1 \ldots n_a . \]

Proof: trivial.

Let \( h \in \text{Ker}(H_1(n_a, d)) \) and \( C_h = \{ H^*_1(n_a, d) \}^{i=1} \) column of \( H_1(n_a, d) \) so as \( h_i = 1 \). The previous theorem claims that \( h \in H_1(n_a, d) \) if and only if for all rows of \( H_1(n_a, d) \), the number of noised bits counted for the columns of \( C_h \) is even. We can deduce the following theorem.

Theorem 3: Let \( h \in \text{Ker}(H_1(n_a, d)) \), of Hamming weight \( w_h \), then the probability that \( h \) is in \( \text{Ker}(H_1(n_a, d)) \) is

\[ P_r(h \in \text{Ker}(H_1(n_a, d))) = \left( \frac{1 + (1 - 2P_e)^w_h}{2} \right)^{n_a} . \quad (6) \]

Proof: let us define \( P^1_{det} = P_r(h \in \text{Ker}(H_1(n_a, d)) | h \in \text{Ker}(H_1(n_a, d))) \). From theorem 2,

\[ P^1_{det} = \left( \sum_{i=0}^{w_h/2} \left( \begin{array}{c} w_h/2 \\ 2i \end{array} \right) P^2_e (1 - P_e)^{w_h/2 - 2i} \right)^{n_a} . \]
We have just to write
\[(1 - 2P_e)^{w_h} = \sum_{i=0}^{w_h} \binom{w_h}{i} (-1)^i P_e^i (1 - P_e)^{w_h-i},\]
\[1 = ((1 - P_e) + P_e)^{w_h} = \sum_{i=0}^{w_h} \binom{w_h}{i} P_e^i (1 - P_e)^{w_h-i},\]
\[\frac{(1 + (1 - 2P_e)^{w_h})}{2} = \sum_{i=0}^{\lfloor w_h/2 \rfloor} \binom{w_h/2}{i} P_e^i (1 - P_e)^{w_h-2i}.\]

Computing the kernel of \(H_1(n_a, d)\) makes possible the detection of a parity check \(h\) of Hamming weight \(w_h\) with probability \(P_{\text{det}}\). Moreover, from corollary 1, the probability \(P_{\text{det}}\) to detect \(h\) which is not a parity check is
\[
P_{\text{det}} \geq 2^{-n_a} (1 - 2^{-n_a}) \approx 2^{-n_a}.\]

C. Estimation of the Synchronization

In case of \(n_a \neq \beta n\), \(Ker(H_1(n_a, d)) = \{0\}\) and so \(Ker(H_1(n_a, d)) \subseteq Ker(H_1(n_a, \alpha))\). Actually, parity checks of \(H_1(n_a, \alpha)\) which imply the last \(d\) columns of \(H_1(n_a, \alpha)\) are not in \(Ker(H_1(n_a, d))\). So the rank criterion
\[rk(H_1(n_a, \alpha)) \leq rk(H_1(n_a, d)) \forall d \neq \alpha n,\]
gives us a condition on the right synchronization parameter \(d_{\text{opt}}\).
\[d_{\text{opt}} = \text{Argmin}_d (rk(H_1(n, d))).\]

To estimate \(d_{\text{opt}}\), we need first to introduce the \(l\)-randomized rank for noised binary matrices. Let \(H(n_a, d)\), \(H_1(n_a, d)\) be two binary noised matrices as defined previously and \(l\) an integer. We generate \(l\) matrices \(H_1(n_a, d)\), for \(i = 1 \ldots l\), by randomly mixing the \(n_a\) rows of \(H(n_a, d)\). Then, we compute \(A_i = \text{span}(Ker(H_1(n_a, d)) \ldots Ker(H_1(n_a, d)))\) and define the \(l\)-randomized rank of \(H_1(n_a, d)\), noted \(l-rkk(H_1(n_a, d))\) by
\[l-rkk(H_1(n_a, d)) = n_a - \text{dim}(A_i).\]

\(l\) is considered as a parameter of the detection algorithm. Because of theorem 2 and equation 7, we have
\[l-rkk(H_1(n_a, \alpha)) \leq l-rkk(H_1(n_a, d)) \forall d \neq \alpha n,\]
with high probability. So, we have adapted the rank criterion to a randomized rank criterion. Then, the estimation of the parameter \(d_{\text{opt}}\), \(d_{\text{opt}}\) is defined by
\[d_{\text{opt}} = \text{Argmin}_d (l-rkk(H_1(n, d))).\]
Moreover, \(\tilde{d}_{\text{opt}}\) is an estimation of \(\frac{d_{\text{opt}}}{n}\) of Hamming weight \(w_h\), then from corollary 1, the probability \(P_{\text{det}}\) to detect \(h\) which is not a parity check is
\[P_{\text{det}} \geq 2^{-n_a} (1 - 2^{-n_a}) \approx 2^{-n_a}.\]

D. The Detection Algorithm

Detection Algorithm

Input: \((y_i)\), the intercepted codewords and \(l, n_{\text{max}}\) parameters.
Output: \(\mathcal{H}\), the dual code of \(C\) which has generated the \((y_i)\), its rate \(\rho\) and \(n\).

1. \(\mathcal{H} \leftarrow \emptyset\).
2. \(r \leftarrow 1\).
3. \(n, d_{\text{opt}} \leftarrow 0\).
4. for \(n_a\) from 2 to \(n_{\text{max}}\) do
5. for \(d\) from 0 to \(n_{\text{max}} - 1\) do
6. fill \(H(n_a, d)\) with \((y_i)\).
7. if \(l-rkk(H(n_a, d)) < r\) then
8. \(r \leftarrow l-rkk(H(n_a, d))\).
9. \(n \leftarrow n_a\).
10. \(d_{\text{opt}} \leftarrow d\).
11. \(\mathcal{H} \leftarrow A_i\).
12. end if
13. end for
14. end for
15. if \(n = 0\) then return "fail".
16. else return \(\mathcal{H}, \frac{d_{\text{opt}}}{n}\), \(n\).

Since the generator matrix of \(C\) is full rank, the parity checks are independent. So, to calibrate our algorithm we have to focus on the hardest parity check to detect, i.e. which has the greatest Hamming weight, \(w_h\). The probability that the detection algorithm succeeds in detecting all the parity checks is \(P_{\text{det}}\). To have a high detection rate, the order of magnitude of \(l\) should be choosen proportional to \(O\left(\frac{1}{P_{\text{det}}}\right)\).

IV. IMPROVING THE DETECTION ALGORITHM

Now, we use the \(H_2(n_a, d)\) matrix in order to increase the probability of parity checks detection. Let be \(h \in Ker(H_1(n_a, d))\) of Hamming weight \(w_h\) and \(\gamma \in [0, 1]\). Let us note \(W_h\) the Hamming weight of \(H_2(n_a, d) \times h\). \(W_h\) can be considered as a random variable in \([0, m - n_a]\). Whether \(h\) is in \(Ker(H_1(n_a, d))\) or not, \(W_h\) does not follow the same probability law. Let us denote \(P_p = \frac{1}{1 - (2P_e)^{n_a}}\). With the same reasoning we made for the proof of theorem 2, we easily prove the following statement.

If \(h \in Ker(H_1(n_a, d))\),
\[P_r(W_h = x) = \frac{m - n_a}{x} P_p (1 - P_p)^{(m-x)},\]
otherwise,
\[P_r(W_h = x) = \frac{m - n_a}{x} P_p (1 - P_p)^{(m-x)},\]
If \(h \in Ker(H_1(n_a, d))\) then \(W_h\) follows a Binomial law of parameter \(P_p\) otherwise it follows a Binomial law of parameter \(1/2\). One can so deduce the optimal threshold, \(\gamma_{\text{opt}}\), for this distinguisher in order to discriminate the two hypothesis \(H_1\) \("h \in Ker(H_1(n_a, d))\)" and \(H_2 \("h \notin Ker(H_1(n_a, d))\)". The
induced decision rule $\mathcal{R}_\gamma$ : one decides $\mathcal{H}_1$ if $W_h \leq (m-n_a)\gamma$ and $\mathcal{H}_2$ otherwise.

We compute the detection and the false alarm probability,

$$P_{det}^2 = P_r(W_h \leq (m-n_a)\gamma|\mathcal{H}_1),$$
$$= \sum_{i=0}^{[(m-n_a)\gamma]} \binom{(m-n_a)}{i} P_i^\gamma (1 - P_i^\gamma)^{(m-n_a-i)},$$
$$P_{fa}^2 = P_r(W_h \leq (m-n_a)\gamma|\mathcal{H}_2),$$
$$= 2^{n_a-m} \sum_{i=0}^{[(m-n_a)\gamma]} \binom{(m-n_a)}{i}.$$

So, to improve our detection algorithm, we first compute a Gauss elimination process onto $H(n_a, d)$ and then, we apply the decision rule $\mathcal{R}_\gamma$ to each vector of the new basis. This strategy is clearly more efficient than applying $\mathcal{R}_\gamma$ only to vectors of $\ker(H_1(n_a, d))$.

As previously, we generate $l$ matrices $H^i(n_a, d)$, for $i = 1 \ldots l$, by randomly mixing the $m$ rows of $H(n_a, d)$. Then, we compute a Gauss elimination process onto each $H^i(n_a, d)$ and obtain $l$ new basis, $\mathcal{B}^i$, for $i = 1 \ldots l$. Let us define $S^i \subset \mathcal{B}^i$ by the set of the vectors of $\mathcal{B}^i$ which are detected by the decision rule $\mathcal{R}_\gamma$ and $S_i = span(\bigcup_{i=1..l} S^i)$. The new detection algorithm is exactly the same as in section III-D, but the refined $l$-randomized rank of $H(n_a, d)$, noted $l$-trk($H(n_a, d)$) is defined by

$$l$-trk($H(n_a, d)$) = \alpha_a - \dim(S_i).$$

V. THEORETICAL COMPLEXITY

Let $h$ be the parity check of greatest Hamming weight, $w_h$. We will now evaluate the complexity of the entire process. We consider that the probability to detect a parity check when $n_a \neq \beta n$ is negligible. When the parity check $h$ is detected by the improved algorithm, the detection probability of the entire process, $P_{det}$ is better than $P_{det}^1$, $P_{det}^2$ and in the same way $P_{fa}$ is bounded by $\sum_{i=n_a}^{2^n} n_a P_i^\gamma(n_a)$. Noting $z = 1 - 2P_e$ which only depends on the channel,

$$P_{det} \geq \sum_{i=0}^{[(m-n)\gamma]} \binom{(m-n)}{i} P(z)^{n+i} (1 - P(z))^{(m-n-i)},$$

where $P(z) = \frac{1 + \gamma^{n_a}}{2}$. Moreover the false alarm probability is bounded by

$$P_{fa} \leq n_a 2^{n-m} \sum_{i=0}^{[(m-n)\gamma]} \binom{(m-n)}{i}.$$ (16)

Let $(h_i)$ the $(n-k)$ parity checks of Hamming weight of $(w_i)$ with $i \leq j$ so as $i < j$, $w_i < w_j$ and $w_{n-k} = w_k$. Trivially, if $(P_{det}^i)$ is the probability to detect $h_i$ then

$$\forall i, j \text{ so as } i < j \text{ then, } P_{det}^i \leq P_{det}^j.$$ (17)

Parity checks are detected independently, so we have to choose $\gamma$ which minimizes the probability that $h$ is not detected, which is $\gamma_{opt} = \underset{\gamma}{\text{argmin}} \left( 1 - \sum_{i=0}^{([(m-n)\gamma])} \binom{(m-n)}{i} (2^{n-m} - P_i^\gamma (1 - P_i^\gamma)^{(m-n-i)}) \right)$.

To detect all the parity checks, we need an average of $\frac{1}{P_{det}}$ trials. For each trial, we compute a Gauss elimination process onto $H(n_a, d)$, in time $O(m \cdot n_a \alpha)$, and $n_a$ decision rules in time $O(n_a \cdot m)$, that is

$$O \left( m \cdot \sum_{i=0}^{n_a} 2(i^2 + i) \frac{1}{P_{det}} \right) = O \left( \frac{m \cdot n_a^4}{4 \cdot P_{det}} \right).$$

VI. CONCLUSION

In this paper we have proposed a generic algebraic approach to unify different techniques for blindly estimating the parameters of binary linear and convolutional encoders. The general problem of reconstructing such error correcting codes appears to be expressed in a very simple way. The theorem 2 gives us an optimality condition to estimate the dual code. Computing the kernel of $H(n_a, d)$ is one optimal method to reconstruct the dual code. Sicot and Houcke [12] have proposed such a method computing the kernels using a Gauss elimination, so as Filiol [5] and Barbier [1] for the convolutional codes. But Filiol and Barbier obtained experimental results better than their theoretical ones since they made the hypothesis that no error must appear in the entire matrix $H_1(n_a, d)$, whereas theorem 2 gives us a lighter condition on the error. Moreover, this approach points out theoretical bounds on the complexity and on the detection probabilities of reconstruction algorithms.

Our further work will be to generalize this approach to soft decision channels in order to design heuristics for selecting good rows for $H_1(n_a, d)$, i.e. for which the probability of being under the application condition of theorem 2 is the highest.

APPENDIX

To prove corollary 1, we first prove the following proposition.

Proposition 1: The property $P(n)$

$$1 - \sum_{i=0}^{n-1} 2^{i-M} \leq \prod_{i=1}^{n-1} (1 - 2^{i-M}),$$

and

$$\prod_{i=0}^{n-1} (1 - 2^{i-M}) \leq 1 - \sum_{i=0}^{n-1} \left( 2^{i-M} - \sum_{j=1}^{n-1} 2^{i-M} 2^{j-M} \right),$$

is true for all $n \geq 1$. 

116
Proof: we will reason by induction. Let us denote
\[ u_n = \prod_{i=0}^{n-1} (1 - 2^{-iM}), \]
\[ v_n = 1 - \sum_{i=0}^{n-1} 2^{-iM}, \]
\[ w_n = 1 - \sum_{i=0}^{n-1} \left( 2^{-iM} - \sum_{j=0}^{n-1} 2^{-iM} 2^{-jM} \right), \]
\[ = 1 - \sum_{i=0}^{n-1} 2^{-iM} \left( 1 - \sum_{j=i+1}^{n-1} 2^{-jM} \right). \]

Since \( v_1 = u_1 = w_1 = 1 - 2^{-M}, (P_1) \) is true.

Assuming \((P_n)\) is true. First, we compute
\[ w_{n+1} - w_n = 1 - \sum_{i=0}^{n} 2^{-iM} \left( 1 - \sum_{j=i+1}^{n} 2^{-jM} \right) - w_n, \]
\[ = - \sum_{i=0}^{n-1} 2^{-iM} \left( 1 - \sum_{j=i+1}^{n-1} 2^{-jM} - 2^{-nM} \right) + 1 - 2^{-nM} - w_n, \]
\[ = 2^{-nM} \sum_{i=0}^{n-1} 2^{-iM} - 2^{-nM} + 1 \]
\[ - \sum_{i=0}^{n-1} 2^{-iM} \left( 1 - \sum_{j=i+1}^{n-1} 2^{-jM} \right) - w_n, \]
\[ = 2^{-nM} \sum_{i=0}^{n-1} 2^{-iM} - 1 \]

We can now compute an upper bound for \( u_{n+1} \),
\[ u_{n+1} = (1 - 2^{-nM}) u_n \leq (1 - 2^{-nM}) w_n \]
\[ \leq w_n - 2^{-nM} \left( 1 - \sum_{i=0}^{n-1} 2^{-iM} \left( 1 - \sum_{j=i+1}^{n-1} 2^{-jM} \right) \right), \]
\[ \leq w_n - 2^{-nM} + 2^{-nM} \sum_{i=0}^{n-1} 2^{-iM} \left( 1 - \sum_{j=i+1}^{n-1} 2^{-jM} \right), \]
\[ \leq w_n + 2^{-nM} \left( \sum_{i=0}^{n-1} 2^{-iM} - 1 \right), \]
\[ - 2^{-nM} \sum_{i=0}^{n-1} 2^{-iM} \sum_{j=i+1}^{n-1} 2^{-jM}, \]
\[ \leq w_{n+1} - 2^{-nM} \sum_{i=0}^{n-1} 2^{iM} \sum_{j=i+1}^{n-1} 2^{jM}, \]
\[ \leq w_{n+1} \]

The lower for \( u_{n+1} \) is obtained in the same way.

\[ u_{n+1} = (1 - 2^{-nM}) u_n \geq (1 - 2^{-nM}) w_n \]
\[ \geq (1 - 2^{-nM}) \left( 1 - \sum_{i=0}^{n-1} 2^{-iM} \right), \]
\[ \geq 1 - \left( \sum_{i=0}^{n-1} 2^{-iM} + 2^{-nM} \right) + 2^{-nM} \sum_{i=0}^{n-1} 2^{-iM}, \]
\[ \geq v_{n+1} + 2^{-nM} \sum_{i=0}^{n-1} 2^{-iM}, \]
\[ \geq v_{n+1} \]

Finally, \( P_{n+1} \) is true and so \((P_n)\) for all \( n \geq 1 \)

Now, we can prove corollary 1
\[ 1 - u_n \leq \sum_{i=0}^{n-1} 2^{-iM} = 2^{-M}(2^{n_a} - 1) \]
\[ 1 - u_n \geq \sum_{i=0}^{n-1} \left( 2^{-iM} - \sum_{i=i+1}^{n-1} 2^{-iM} 2^{-jM} \right), \]
\[ \geq 2^{-M} \sum_{i=0}^{n-1} 2^{iM} - 2^{-2M} \sum_{i=0}^{n-1} 2^{iM} \sum_{j=i+1}^{n-1} 2^{jM}, \]
\[ \geq 2^{-M}(2^{n_a} - 1) - 2^{-2M} \sum_{i=0}^{n-1} 2^{iM} \sum_{j=0}^{n-1} 2^{jM}, \]
\[ \geq 2^{-M}(2^{n_a} - 1) - (2^{-M}(2^{n_a} - 1))^2. \]

For \( M = n_a \) we obtain corollary 1 \( \Box \)

REFERENCES


