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 acess 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 theorical 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 theorical 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 developped 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, Valembois [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 citerion 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 theorical 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 theoricaly 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 theorical 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 theoricaly proved.

II. The Algebraic Approach

Let G be the generator matrix of the binary linear error correcting code $\mathcal{C}$, 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 $G^{'} = P \times diag(G)$, where P is the permutation matrix of the interleaver and $diag(G)$ defined by $\begin{pmatrix} G & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & G \end{pmatrix}$.

Let us denote $(y_i)_{i=1...m}$ the codewords generated during the communication, and $(y_i)_{i=1...m}$ the noised codewords

J. Barbier is with the Cryptology Dept, CELAR (France).
Email : johann.barbier@dga.defense.gouv.fr
G. Sicot and S. Houcke are with the Dept Signal and Communication. ENST-Bretagne TAMCIC (CNRS 2658)
Email : {guillaume.sicot, sebastien.houcke}@enst-bretagne.fr.fr

intercepted by the adversary.

$$y_i^j = y_i^j + e_{ij}, \quad \forall i = 1 \dots m, \quad j = 1 \dots n, \qquad (1)$$

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 $\mathcal{H}$ be the dual code of $\mathcal{C}$, so that for all parity checks $h \in \mathcal{H}$, $< h.y_i >= 0$ $\forall i$. Since recognizing the code $\mathcal{C}$ is equivalent to estimate the dual code, we consider now the estimation of $\mathcal{H}$.

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 $[H]_{ij} = y_i^j$. 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 $[H(d)]_{ij} = y_i^{j+d}$ $\forall i = 1 \dots m$, $j = 1 \dots n - d$, $[H(d)]_{ij} = y_{i+1}^{j-d}$, $\forall i = 1 \dots m$, $j = n - d + 1 \dots n$, 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,

$$\begin{aligned} H_1(n_a, d) &= H_1(n_a, d) + E_1, & (2) \\ H_2(n_a, d) &= H_2(n_a, d) + E_2, & (3) \end{aligned}$$

where $E_i = E_i(n_a, d)$. Moreover $\mathcal{H} = Ker(H_1(n, \alpha n))$, $\forall \alpha \in \mathbb{N}$. So, the problem of binary error correcting codes reconstruction is equivalent to determine $Ker(H_1(n, \alpha n))$ observing $H_1(n_a, d)$. The problem of the reconstruction of convolutional codes [1], [5] and of turbocodes [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], Filiol [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 β is an integer.

A. Case n_a= βn

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 $\mathcal{C}$ 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 $n_a \neq \beta n$,

$$\mathcal{P}r(rk(H_1(n_a, d)) < n_a) = 1 - \prod_{i=0}^{n_a - 1} (1 - 2^{i - na}). \qquad (4)$$

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

$$\prod_{i=0}^{l-1} (1 - 2^{i-k}).$$

We denote $\mathcal{P}_{fa}^0 = \mathcal{P}r(rg(H_1(n_a, d)) < n_a)$. 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:

$$2^{-na}(1 - 2^{-n_a}) \leq \mathcal{P}_{fa}^0 \leq (1 - 2^{-n_a}). \qquad (5)$$

Proof : see appendix.
The probability $\mathcal{P}_{det}^0$ to detect a real parity check is trivialy 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.* $h \in Ker(H_1(n_a, d))$ then,

$$h \in Ker(H_1(n_a, d)) \text{ if and only if}$$

$$\sum_{j=1}^{na} [E_1]_{ij} h_j \equiv 0 \mod 2 \ \forall j = 1 \dots n_a.$$

Proof : trivial.

Let $h \in Ker(H_1(n_a, d))$ and $\mathcal{C}_h = \{H_1^i(n_a, d)$ i$^{\text{th}}$ 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 $\mathcal{C}_h$ is even. We can deduce the following theorem.

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

$$\mathcal{P}r(h \in Ker(\tilde{H}_1(n_a, d)) | h \in Ker(H_1(n_a, d)) = \left(\frac{1 + (1 - 2P_e)^{w_h}}{2} \right)^{n_a}. \qquad (6)$$

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

$$\mathcal{P}_{det}^1 = \left(\sum_{i=0}^{\lfloor w_h/2 \rfloor} \binom{w_h}{2i} P_e^{2i} (1 - P_e)^{w_h - 2i} \right)^{n_a}.$$

We have just to write

$$\begin{aligned}
(1-2P_e)^{w_n} &= \sum_{i=0}^{w_h} \binom{w_h}{i} (-1)^i P_e^i (1-P_e)^{w_h-i}, \\
1 = ((1-P_e)+P_e)^{w_h} &= \sum_{i=0}^{w_h} \binom{w_h}{i} P_e^i (1-P_e)^{w_h-i}, \\
\left(\frac{1+(1-2P_e)^{w_h}}{2}\right)^{n_a} &= \left(\sum_{i=0}^{\lfloor w_h/2 \rfloor} \binom{w_h}{2i} P_e^{2i} (1-P_e)^{w_h-2i}\right)^{n_a}
\end{aligned}$$

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 $\mathcal{P}^1_{det}$. Moreover, from corollary 1, the probability $\mathcal{P}^1_{fa}$ to detect $h^{'}$ which is not a parity check is

$$\mathcal{P}^1_{fa} \geq 2^{-n_a}(1-2^{-n_a}) \approx 2^{-n_a}.$$

C. Estimation of the Synchronization

In case of $n_a \neq \beta n$, $Ker(H_1(n_a, d)) = \{0\}$ and so $Ker(H_1(n_a, d)) = \{0\}$ with probability $1 - \mathcal{P}^0_{fa}$ for all d. If $n_a = \beta n$ and $d \neq \alpha n$, then $Ker(H_1(n_a, d)) \subseteq Ker(H_1(n_a, \alpha n))$. Actually, parity checks of $H_1(n_a, \alpha n)$ which implie the last d columns of $H_1(n_a, \alpha n)$ are not in $Ker(H_1(n_a, d))$. So the rank criterion

$$rk(H_1(n_a, \alpha n)) \leq rk(H_1(n_a, d)) \quad \forall d \neq \alpha n, \tag{7}$$

gives us a condition on the right synchronization parameter d_{opt},

$$d_{opt} = Argmin_d \left(rk(H_1(n, d))\right). \tag{8}$$

To estimate d_{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 $H_1^i(n_a, d)$, for $i = 1 \dots l$, by randomly mixing the m rows of $H(n_a, d)$. Then, we compute $\mathcal{A}_l = span(Ker(H_1^1(n_a, d)), \dots, Ker(H_1^l(n_a, d)))$ and define the *l-randomized rank* of $H_1(n_a, d)$, noted l-rrk$(H_1(n_a, d))$ by

$$l\text{-rrk}(H_1(n_a, d)) = n_a - dim(\mathcal{A}_l). \tag{9}$$

l is considered as a parameter of the detection algorithm. Because of theorem 2 and equation 7, we have

$$l\text{-rrk}(H_1(n_a, \alpha n)) \leq l\text{-rrk}(H_1(n_a, d)) \quad \forall d \neq \alpha n, \tag{10}$$

with high probability. So, we have adapted the rank criterion to a *randomized rank criterion*. Then, the estimation of the parameter d_{opt}, d_{opt} is defined by

$$d_{opt} = Argmin_d \left(l\text{-rrk}(H_1(n, d))\right). \tag{11}$$

Moreover, $\frac{\tilde{d}_{opt}}{n}$ is an estimation of $\frac{d_{opt}}{n} = \frac{n-(n-k)}{n} = \frac{k}{n} = \rho$, the rate of $\mathcal{C}$.

D. The Detection Algorithm

Detection Algorithm

Input :(y_i), the intercepted codewords and l, n_{max} parameters.

Output : $\mathcal{H}$, the dual code of $\mathcal{C}$ which has generated the (y_i), its rate ρ and n.

```
$\mathcal{H} \longleftarrow \emptyset$,
$r \longleftarrow 1$,
$n, d_{opt} \longleftarrow 0$,
for $n_a$ from 2 to $n_{max}$ do
    for $d$ from 0 to $n_{max} - 1$ do
        fill $H(n_a, d)$ with $(y_i)$
        if $l$-rrk$(H_1(n_a, d)) < r$ then
            $r \longleftarrow$l-rrk$(H_1(n_a, d))$,
            $n \longleftarrow n_a$,
            $d_{opt} \longleftarrow d$,
            $\mathcal{H} \longleftarrow \mathcal{A}_l$.
        end if
    end for
end for
if $n == 0$ then return "fail"
else return $\mathcal{H}$, $\frac{d_{opt}}{n}$, $n$.
```

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

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 $P_p = \frac{1+(1-2P_e)^{w_h}}{2}$. With the same reasoning we made for the proof of theorem 2, we easily prove the following statement.
If $h \in Ker(H_1(n_a, d))$,

$$\mathcal{P}_r(W_h = x) = \binom{m-n_a}{x} P_p^x (1-P_p)^{(m-x)}, \tag{12}$$

otherwise,

$$\mathcal{P}_r(W_h = x) = 2^{m-n_a} \binom{m-n_a}{x}. \tag{13}$$

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, γ_{opt}, for this distinguisher in order to discriminate the two hypothesis $\mathcal{H}_1$ "$h \in Ker(H_1(n_a, d))$" and $\mathcal{H}_2$ "$h \notin Ker(H_1(n_a, d))$". 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,

$$\begin{aligned} \mathcal{P}^2_{det} &= \mathcal{P}r(W_h \leq (m-n_a)\gamma|\mathcal{H}_1), \\ &= \sum_{i=0}^{\lfloor (m-n_a)\gamma \rfloor} \begin{pmatrix} m-n_a \\ i \end{pmatrix} P_p^i (1-P_p)^{(m-n_a-i)}, \\ \mathcal{P}^2_{fa} &= \mathcal{P}r(W_h \leq (m-n_a)\gamma|\mathcal{H}_2), \\ &= 2^{na-m} \sum_{i=0}^{\lfloor (m-n_a)\gamma \rfloor} \begin{pmatrix} m-n_a \\ i \end{pmatrix}. \end{aligned}$$

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 elimation process onto each $H^i(n_a,d)$ and obtain l new basis, $\mathcal{B}^i$, for $i = 1 \ldots l$. Let us define $\mathcal{S}^i \subset \mathcal{B}^i$ by the set of the vectors of $\mathcal{B}^i$ which are detected by the decision rule $\mathcal{R}_\gamma$ and $\mathcal{S}_l = span(\bigcup_{i=1..l} \mathcal{S}^i)$. 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-rrk$(H(n_a,d))$ is defined by

$$l\text{-rrk}(H(n_a,d)) = n_a - dim(\mathcal{S}_l). \tag{14}$$

V. Theorical 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 $n_a \neq \beta n$ as negligible. When the parity check h is detected by the improved algorithm, the detection probability of the entire process, $\mathcal{P}_{det}$ is better than $\mathcal{P}^1_{det}.\mathcal{P}^2_{det}$ and in the same way $\mathcal{P}_{fa}$ is bounded by $\sum_{n_a=2}^{n} n_a.\mathcal{P}^2_{fa}(n_a)$. Noting $z = 1 - 2P_e$ which only depends on the channel,

$$\mathcal{P}_{det} \geq \sum_{i=0}^{\lfloor (m-n)\gamma \rfloor} \begin{pmatrix} m-n \\ i \end{pmatrix} P(z)^{n+i} \left(1 - P(z)\right)^{(m-n-i)}, \tag{15}$$

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

$$\mathcal{P}_{fa} \leq n.2^{n-m} \sum_{i=0}^{\lfloor (m-n)\gamma \rfloor} \begin{pmatrix} m-n \\ i \end{pmatrix}. \tag{16}$$

Let (h_i) the $(n-k)$ parity checks of Hamming weight of (w_i) with $\forall i,j$ so as $i<j$, $w_i < w_j$ and $w_{n-k} = w_h$. Trivialy, if $(\mathcal{P}^i_{det})$ is the probability to detect h_i then

$$\forall i,j \text{ so as } i<j \text{ then, } \mathcal{P}^j_{det} \leq \mathcal{P}^i_{det}.$$

Parity checks are detected independently, so we have to choose γ which minimizes the probability that h is not detected, which is $\gamma_{opt} =$

$$Argmin_\gamma \left(1 - \sum_{i=0}^{\lfloor (m-n)\beta \rfloor} \begin{pmatrix} m-n \\ i \end{pmatrix} \left(2^{n-m} - P_p^i(1-P_p)^{(m-n-i)}\right)\right). \tag{17}$$

To detect all the parity checks, we need an average of $\frac{1}{\mathcal{P}_{det}}$ distinct matrices $H(n,d)$. This gives us a criterion on the minimal number of bits to intercept, m_{min}. m_{min} must verify

$$\begin{pmatrix} n \\ m_{min} \end{pmatrix} \leq \frac{1}{\mathcal{P}_{det}}. \tag{18}$$

We blindly estimate n and d, so at most $\frac{n^2}{2}$ trials are needed. For each trial we compute $\frac{1}{\mathcal{P}_{det}}$ iterations. For each iteration, we compute a Gauss elimination process onto $H(n_a,d)$, in time $\mathcal{O}(m.n_a^2)$, and n_a decision rules in time $\mathcal{O}(n_a.m)$, that is

$$\mathcal{O}\left(m.\sum_{i=0}^{n} \frac{i}{2}(i^2+i)\frac{1}{\mathcal{P}_{det}}\right) = \mathcal{O}\left(\frac{m.n^4}{4.P_{det}}\right).$$

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 theorical 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 theorical bounds on the complexity and on the detection probabilities of reconstruction algorithms. Our futur 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 probabilty 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 $P(n)$

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

and

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

is true for all $n \geq 1$.

Proof : we will reason by induction. Let us denote

$$\begin{aligned} u_n &= \prod_{i=0}^{n-1}(1-2^{i-M}),\\ v_n &= 1-\sum_{i=0}^{n-1}2^{i-M},\\ w_n &= 1-\sum_{i=0}^{n-1}\left(2^{i-M}-\sum_{j=0}^{n-1}2^{i-M}2^{j-M}\right),\\ &= 1-\sum_{i=0}^{n-1}2^{i-M}\left(1-\sum_{j=i+1}^{n-1}2^{j-M}\right). \end{aligned}$$

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

Assuming (P_n) is true. First, we compute

$$\begin{aligned} w_{n+1}-w_n &= 1-\sum_{i=0}^{n}2^{i-M}\left(1-\sum_{j=i+1}^{n}2^{j-M}\right)-w_n,\\ &= -\sum_{i=0}^{n-1}2^{i-M}\left(1-\sum_{j=i+1}^{n-1}2^{j-M}-2^{n-M}\right)\\ &\quad +1-2^{n-M}-w_n,\\ &= 2^{n-M}\sum_{i=0}^{n-1}2^{i-M}-2^{n-M}+1\\ &\quad -\sum_{i=0}^{n-1}2^{i-M}\left(1-\sum_{j=i+1}^{n-1}2^{j-M}\right)-w_n,\\ &= 2^{n-M}\left(\sum_{i=0}^{n-1}2^{i-M}-1\right)\square \end{aligned}$$

We can now compute an upper bound for u_{n+1},

$$\begin{aligned} &u_{n+1} = (1-2^{n-M})u_n \le (1-2^{n-M})w_n\\ &\le \; w_n-2^{n-M}\left(1-\sum_{i=0}^{n-1}2^{i-M}\left(1-\sum_{j=i+1}^{n-1}2^{j-M}\right)\right),\\ &\le \; w_n-2^{n-M}+2^{n-M}\sum_{i=0}^{n-1}2^{i-M}\left(1-\sum_{j=i+1}^{n-1}2^{j-M}\right),\\ &\le \; w_n+2^{n-M}\left(\sum_{i=0}^{n-1}2^{i-M}-1\right)\\ &\qquad -2^{n-M}\sum_{i=0}^{n-1}2^{i-M}\sum_{j=i+1}^{n-1}2^{j-M},\\ &\le \; w_{n+1}-2^{n-M}\sum_{i=0}^{n-1}2^{i-M}\sum_{j=i+1}^{n-1}2^{j-M},\\ &\le \; w_{n+1} \quad \square \end{aligned}$$

The lower for u_{n+1} is obtained in the same way,

$$\begin{aligned} &u_{n+1} = (1-2^{n-M})u_n \ge (1-2^{n-M})v_n\\ &\ge \; (1-2^{n-M})\left(1-\sum_{i=0}^{n-1}2^{i-M}\right),\\ &\ge \; 1-\left(\sum_{i=0}^{n-1}2^{i-M}+2^{n-M}\right)+2^{n-M}\sum_{i=0}^{n-1}2^{i-M},\\ &\ge \; v_{n+1}+2^{n-M}\sum_{i=0}^{n-1}2^{i-M},\\ &\ge \; v_{n+1} \quad \square \end{aligned}$$

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

Now, we can prove corollary 1.

$$\begin{aligned} 1-u_{n_a} &\le \sum_{i=0}^{n_a-1}2^{i-M} = 2^{-M}(2^{n_a}-1) \text{ and}\\ 1-u_{n_a} &\ge \sum_{i=0}^{n_a-1}\left(2^{i-M}-\sum_{j=i+1}^{n_a-1}2^{i-M}2^{j-M}\right),\\ &\ge 2^{-M}\sum_{i=0}^{n_a-1}2^{i}-2^{-2M}\sum_{i=0}^{n_a-1}2^{i}\sum_{j=i+1}^{n_a-1}2^{j},\\ &\ge 2^{-M}(2^{n_a}-1)-2^{-2M}\sum_{i=0}^{n_a-1}2^{i}\sum_{j=0}^{n_a-1}2^{j},\\ &\ge 2^{-M}(2^{n_a}-1)-\left(2^{-M}(2^{n_a}-1)\right)^2. \end{aligned}$$

For $M = n_a$ we obtain corollary 1 □

References

[1] J. Barbier. Reconstruction of turbo-code encoders. In *Proc. SPIE Security and Defense, Space Communication Technologies Symposium*, volume 5819, pages 463–473, March 28-31 2005.

[2] G. Burel and R. Gautier. Blind estimation of encoder and interleaver characteristics in a non cooperative context. In *Proc. IASTED International Conference on Communications, Internet and Information Technology*, November, 17-19 2003.

[3] A. Canteaut and F. Chabaud. A new algorithm for finding minimum-weight words in a linear code : Application to McEliece's cryptosystem and to narrow-sense BCH codes of length 511. *IEEE Trans. Inform. Theory*, 44(1) :367–378, January 1998.

[4] M. Cluzeau. Block code reconstruction using iterative decoding techniques. In *Proc. 2006 IEEE International Symposium on Information Theory, ISIT06*, Juilly 2006.

[5] E. Filiol. Reconstruction of convolutional encoders over gf(q). In M. Darnell, editor, *Proc. 6th IMA Conference on Cryptography and Coding*, number 1355 in Lecture Notes in Computer Science, pages 100–110. Springer Verlag, 1997.

[6] E. Filiol. Reconstruction of punctured convolutional encoders. In T. Fujiwara, editor, *Proc. 2000 International Symposium on Information Theory and Applications*, pages 4–7. SITA and IEICE Publishing, 2000.

[7] E. Filiol. *Techniques de reconstruction en cryptologie et théorie des codes.* PhD thesis, INRIA Rocquencourt, France, March 2001.

[8] J.S. Leon. A probabilistic algorithm for computing the minimum weight of large error-correcting codes. *IEEE Trans. on Information Theory, IT-34(5)*, pages 1354–1359, September 1988.

[9] R. Lidl and H. Niederreiter. *Finite Fields.* Cambridge University Press, 1983.

[10] G. Planquette. *Identification de trains binaires codés.* PhD thesis, Université de Rennes I, France, December 1996.

[11] B. Rice. Determining the parameters of a rate $\frac{1}{n}$ convolutional encoder over gf(q). In *Proc. Third International Conference on Finite Fields and Applications*, 1995.

[12] G. Sicot and S. Houcke. Blind detection of interleaver parameters. In *Proc. ICASSP 2005*, 2005.

[13] G. Sicot and S. Houcke. Etude statistique du seuil dans la détection d'entrelaceur. In *Proc. GRESTSI 2005*, 2005.

[14] G. Sicot and S. Houcke. Theorical study of the performance of a blind interleaver estimator. In *Proc. ISIVC 2006*, 2006.

[15] J. Stern. A method for finding codewords of small weight. In G. Cohen and J. Wolfmann, editors, *Proc. Coding Theory and Applications*, number 388 in Lecture Notes in Computer Science, pages 106–113. Springer Verlag, 1989.

[16] A. Valembois. *Détection, reconnaissance et décodage de codes linéaires binaires.* PhD thesis, Université de Limoges, France, September 2000.

[17] A. Valembois. Detection and recognition of a binary linear code. *Discrete Applied Mathematics*, (111) :199–218, 2001.