

Performance of parallel concatenated coding schemes.

Fabio Fagnani*

Abstract

In this paper we study ensembles of parallel concatenated codes and we present precise results on their asymptotic performance. In particular, we prove that in any parallel concatenation scheme where all k encoders are recursive and the rate is sufficiently small, the bit error rate goes to 0 exactly as N^{1-k} . We consider different types of ensembles by changing the subgroup of permutations used to interconnect the various encoders.

Keywords: bit error rate, ensemble of codes, interleaver gain, parallel concatenated codes, recursive convolutional encoders.

1 Introduction

In this paper we study ensembles of parallel concatenated codes and we present precise results on their asymptotic performance. Parallel concatenations have made their appearance in the coding literature in the mid 90's in [4] and they have rapidly gained worldwide attention for their brilliant performances when used in combination of a low complexity iterative decoding scheme. Since then there have been two different lines of analysis of these coding schemes: on one side they have been studied in combination with the suboptimal iterative coding scheme [9]; on the other side they have also been studied in a more classical setting, considering instead optimal ML decoding. This second line plays a fundamental role in trying to separate the analysis of these coding schemes from the use of the suboptimal iterative decoding. Pioneer work in this sense has been the work [2], [3] where for the first time an ensemble style analysis was performed averaging the error probability estimation on all the possible permutations interleaving the two constituent convolutional encoders. The analysis, though not fully rigorous, gave a lot of insight into the question of understanding the performance of such schemes: in particular it put into evidence the fundamental importance of recursivity of the two convolutional encoders, and showed, in the recursive case, the existence of a sort of interleaver gain term $1/N$ (where N was the length of the interleaver) in the union upper bound of the bit error rate. This explained the good performance of such codes for large interleaver length. In [7] for the first time a formal derivation of the upper bound in [3] was derived, in the more general setting of concatenated schemes consisting of k parallel convolutional encoders. The result in [7] actually proves a slightly worse upper bound than in [3] but in a completely rigorous way. Moreover their results also extend to serial concatenations.

This paper follows the theoretical line of the papers [2], [3], and [7] improving and extending the result in [7]. We prove that in any parallel concatenation scheme where all k encoders are recursive the BER has an asymptotic behavior of N^{1-k} : namely we improve on the upper bound presented in [7] showing that exactly the one derived in [3] holds true and moreover we establish an analogous lower bound (result which was considered open in [7]). The techniques we use are basic combinatorics and basic facts on the theory of convolutional codes. For what concerns the upper bounds our analysis follows closely the ideas contained in [3] filling in the formal gaps in their derivation. Lower bounds are instead obtained through an analysis of minimum distances.

The setting considered in this paper is, moreover, more general than the one considered in the previous literature as we allow different permutation ensembles: beyond the classical bit permutations we in fact also consider the case of symbol permutations and also of independent permutations in different channel inputs. This allows a more flexible theory with potential applications in coupling

*Dipartimento di Matematica, Politecnico di Torino, C.so Duca degli Abruzzi 24, 10129 Torino, Italy.

concatenated schemes with non-binary modulations. A deeper analysis in this sense is presented in [6].

Finally, in the classical case when all encoders have scalar inputs and average is done over all bit permutation, we complete the analysis proving that the presence of non-recursive encoders do not play any role in the asymptotic behavior for $N \rightarrow +\infty$ of the BER in the sense that it still goes as N^{1-k} where $k \geq 2$ is the number of recursive encoders. Moreover, we prove that if the number of recursive encoders is 1 or 0, than no interleaver gain shows up, indeed we have that the BER is in this case bounded away from 0: this fact had been noticed in the simulations and conjectured to be true, but, at our knowledge, never explicitly proven before. While for the rest of the paper we assume that the channel used is any binary input symmetric memoryless channel, to obtain this last result we assume we are using a BSC.

We now briefly comment on the structure of the paper. In Section 1 we consider ensembles of concatenated block codes in the context of generic permutation subgroups and we present a simple generalization of the union bound estimation derived in [2]. Section 2 is devoted to the convolutional extension. In this setting it is necessary to make more stringent assumptions on the permutation groups used: we propose various examples showing how a number of interesting case indeed fall in the case considered. We also introduce concepts of non-catastrophicity and recursivity connected to the ensemble of permutation used which will play a fundamental role in the rest of the paper. Section 3 proves the general upper bound estimate while Section 4 contains the lower bound estimation. Finally Section 5 contains further results for the classical case.

2 Ensemble of concatenated block codes

In this paper we will study the performance of ensembles of block codes obtained by concatenating the truncation of convolutional codes through suitable permutations. We start by briefly recalling some classical facts on block codes and we then pass to consider concatenations.

2.1 Preliminaries on block codes for symmetric binary channels

Throughout this paper, except in some part of Section 5, we assume to be working with a binary input memoryless channel with output alphabet Ω which can be either a finite set or the continuous space \mathbb{R}^n . In the first case the channel is completely described by two probabilities $p(\omega|0)$ and $p(\omega|1)$ on the set Ω . In the second case, instead, by two densities $f(\omega|0)$ and $f(\omega|1)$ on \mathbb{R}^n . We will make the assumption that the channel is symmetric in the following sense: if Ω is finite, $p(\omega|0)$ and $p(\omega|1)$ simply differ by a permutation on Ω , if $\Omega = \mathbb{R}^n$, $f(\omega|0)$ and $f(\omega|1)$ differ by an isometric change of variables. Some of the results we will present will actually hold true also without this symmetry assumption, for others instead the lack of symmetry leads to less elegant formulations.

Let $E : \mathbb{Z}_2^N \rightarrow \mathbb{Z}_2^M$ be a linear block encoder and let $\mathcal{C}_E = \text{Im}(E)$ be the associated block code. Let $P_w(e|E)$ and $P_b(e|E)$ be the word and bit error probability, respectively, of the coding scheme associated with E over the above channel, assuming ML decoding and uniform probability on the information words. We will make use of the union bound estimations to obtain upper bounds on the error performances. These estimations are in terms of the so called Battacharyya noise parameter γ of the channel, which is defined, depending if Ω is discrete or continuous, as

$$\gamma = \sum_{\omega \in \Omega} \sqrt{p(\omega|0)p(\omega|1)}, \quad \gamma = \int_{\Omega} \sqrt{f(\omega|0)f(\omega|1)} d\omega.$$

Using the weight enumerating sequences

$$A_{w,d} = |\{u \in \mathbb{Z}_2^M : w_H(u) = w, w_H(Eu) = d\}|, \quad A_d = \sum_w A_{w,d},$$

and the weight enumerating functions

$$A_w(D) = \sum_d A_{w,d} D^d, \quad A(D) = \sum_w A_w(D)$$

we obtain the well known estimations:

$$P_w(e|E) \leq A(\gamma), \quad P_b(e|E) \leq \sum_w \frac{w}{N} A_w(\gamma). \quad (1)$$

On the other hand, lower bounds will be essentially obtained in terms of upper bounds on the minimal distance $d(\mathcal{C}_E)$ of the code \mathcal{C}_E . Define, in the discrete case, the sets

$$\Lambda_0 = \{\omega \in \Omega : p(\omega|0) \geq p(\omega|1)\}, \quad \Lambda_1 = \{\omega \in \Omega : p(\omega|1) \geq p(\omega|0)\}$$

and, the equivocation probability

$$p = \sum_{\omega \in \Lambda_1} p(\omega|0) = \sum_{\omega \in \Lambda_0} p(\omega|1),$$

In the continuous case the sets Λ_0 and Λ_1 are defined by simply replacing the probabilities $p(\cdot)$ with the densities $f(\cdot)$ and the equivocation probability p is then defined as

$$p = \int_{\Lambda_1} f(\omega|0) d\omega = \int_{\Lambda_0} f(\omega|1) d\omega.$$

We have the following simple estimation

Lemma 1

$$d(\mathcal{C}_E) \leq d \Rightarrow P_w(e|E) \geq p^d. \quad (2)$$

Proof Assume that $\bar{y} = (\bar{y}_1, \dots, \bar{y}_M) \in \mathcal{C}_E$ is such that $w_H(\bar{y}) = \bar{d} \leq d$. Since the channel is symmetric the word error probability is independent on the transmitted sequence; we can thus assume that the the word $0 \in \mathcal{C}_E$ has been transmitted. Consider

$$\Gamma = \{\omega \in \Omega^M \mid p(\omega_i|1) \geq p(\omega_i|0) \forall i : \bar{y}_i = 1\}.$$

We can estimate

$$P_w(e|E) \geq P(\Gamma) = p^{\bar{d}} \geq p^d. \quad \blacksquare$$

Remark: Notice that, in the proof of the above lemma, we assume error if the ML estimate is not unique; the result however still holds true (with a possibly different p) if a different decoding option is taken for these matched situations.

2.2 Ensemble of concatenated block codes

We denote by S_N the group of permutations on N elements. S_N acts in the usual way on \mathbb{Z}_2^N by permuting components: if $u \in \mathbb{Z}_2^N$ and $\sigma \in S_N$, the action is simply denoted by σu .

Consider k \mathbb{Z}_2 -linear block encoders

$$E_i : \mathbb{Z}_2^N \rightarrow \mathbb{Z}_2^{M_i} \quad i = 1, \dots, k.$$

and a vector permutation $\sigma = (\sigma_1, \dots, \sigma_k) \in S_N^k$. We can define the concatenated block encoder

$$E_\sigma : \mathbb{Z}_2^N \rightarrow \mathbb{Z}_2^{M_1} \times \dots \times \mathbb{Z}_2^{M_k}$$

$$E_\sigma(u) = (E_1(\sigma_1 u), \dots, E_k(\sigma_k u))$$

We denote by $\mathcal{C}_\sigma = E_\sigma(\mathbb{Z}_2^N)$ the block code associated to E_σ .

We now fix a subgroup of permutations $G \leq S_N$ on which we consider the uniform probability. Denote the average word and bit error probability over the ensemble determined by G as, respectively,

$$P_w(e) = \frac{1}{|G|^k} \sum_{\sigma \in G^k} P_w(e|E_\sigma), \quad P_b(e) = \frac{1}{|G|^k} \sum_{\sigma \in G^k} P_b(e|E_\sigma). \quad (3)$$

Remark: Notice that given any $\sigma = (\sigma_1, \dots, \sigma_k) \in G^k$, if we consider $\tilde{\sigma} = \sigma \cdot \sigma_1^{-1} = (id, \dots, \sigma_k \sigma_1^{-1})$, we have that

$$E_\sigma = E_{\tilde{\sigma}} \circ \sigma_1$$

On the other, it is well known that the right composition of a block encoder with a permutation does not modify the performance of the coding scheme, in particular it does not modify its word and bit error probability. As a consequence we can as well assume that we are in the smaller ensemble described by the vector permutations in G^k having the first component equal to the identity: the averaged error probabilities are exactly the same. It will turn out to be useful in certain circumstances to work with this smaller ensemble.

Notice that G splits \mathbb{Z}_2^N into equivalence classes

$$\langle u \rangle = \{\sigma u \mid \sigma \in G\}$$

The set of these equivalence classes is denoted $\Lambda = \mathbb{Z}_2^N / G$. Given $\lambda = \langle u \rangle \in \Lambda$, \mathcal{M}_λ denotes the cardinality of the class $\langle u \rangle$ and is called the multiplicity of λ . We denote by $G(u)$ the stabilizer of u , namely $G(u) = \{\sigma \in G : \sigma u = u\}$. $G(u)$ is a subgroup (in general not normal) of G and the set of left lateral classes $G/G(u)$ is canonically in bijection with $\langle u \rangle$. In this paper whenever we quotient by subgroups we will assume that we are considering left lateral classes. Given $u \in \mathbb{Z}_2^N$ denotes by $w_H(u)$ the Hamming weight of u , namely the number of non-zero components of u . Clearly, the Hamming weight is invariant by the action of permutations so that all elements in λ will have the same Hamming weight which will be called the Hamming weight of λ and denoted $\|\lambda\|$. Finally, given $\lambda \in \Lambda$ and $d \in \mathbb{N}$, we define the enumerative weights:

$$A_{\lambda,d}^i = |\{u \in \lambda \mid w_H(E_i(u)) = d\}|,$$

and the enumerative weight functions

$$A_\lambda^i(D) = \sum_{d \geq 0} A_{\lambda,d}^i D^d.$$

We have the following result generalizing [2]. The proof is quite similar to the original one and we only present a quick proof.

Theorem 2 *The following estimations of the averaged word and bit error probabilities hold:*

$$P_w(e) \leq \sum_{\lambda \in \Lambda \setminus \{0\}} \frac{A_\lambda^1(\gamma) \cdots A_\lambda^k(\gamma)}{\mathcal{M}_\lambda^{k-1}}, \quad P_b(e) \leq \sum_{\lambda \in \Lambda \setminus \{0\}} \frac{\|\lambda\| A_\lambda^1(\gamma) \cdots A_\lambda^k(\gamma)}{rN \mathcal{M}_\lambda^{k-1}} \quad (4)$$

Proof It follows from the definition of $P_b(e)$ in (3), the remark following it and the union bound estimation (1) that

$$P_b(e) \leq \frac{1}{|G|^{k-1}} \sum_w \frac{w}{N} \sum_\sigma A_w^\sigma(\gamma) \quad (5)$$

where $A_w^\sigma(D) = \sum_d A_{w,d}^\sigma D^d$ denotes the weight enumerating function associated with the block code E_σ and where we are assuming we are only considering vector permutations with the first component equal the identity: $\sigma = (id, \sigma_2, \dots, \sigma_k)$. Notice now that

$$A_{w,d}^\sigma = \sum_{\substack{\lambda \in \Lambda \\ \|\lambda\| = w}} A_{\lambda,d}^\sigma.$$

Moreover, the following combinatorial identity holds true

$$\sum_\sigma A_{\lambda,d}^\sigma = \left(\frac{|G|}{|\mathcal{M}_\lambda|} \right)^{k-1} \sum_{d_1 + \dots + d_k = d} A_{\lambda,d_1}^1 \cdots A_{\lambda,d_k}^k. \quad (6)$$

This can be seen as follows. It is easy to see that the left hand side counts all the possible vector inputs (u_1, \dots, u_k) each belonging to the class λ such that the corresponding outputs (y_1, \dots, y_k) have

total weight $d_1 + \dots + d_k = d$, with a multiplicity due to the stabilizer $G(u_2) \times \dots \times G(u_k)$. Since we know that $G/G(u)$ is in bijection with $\langle u \rangle$, we obtain that $|G(u)| = |G|/|\mathcal{M}_\lambda|$. This yields (6). We can now write

$$\begin{aligned} \sum_{\sigma} A_w^{\sigma}(D) &= \sum_{\substack{\lambda \in \Lambda \\ \|\lambda\| = w}} \sum_d \sum_{\sigma} A_{\lambda,d}^{\sigma} D^d \\ &= \sum_{\substack{\lambda \in \Lambda \\ \|\lambda\| = w}} \left(\frac{|G|}{|\mathcal{M}_\lambda|} \right)^{k-1} \sum_d \left(\sum_{d_1 + \dots + d_k = d} A_{\lambda,d_1}^1 \cdots A_{\lambda,d_k}^k \right) D^d \\ &= \sum_{\substack{\lambda \in \Lambda \\ \|\lambda\| = w}} \left(\frac{|G|}{|\mathcal{M}_\lambda|} \right)^{k-1} A_{\lambda}^1(D) \cdots A_{\lambda}^k(D). \end{aligned}$$

Inserting this identity inside (5) we obtain the wanted estimation for the $P_b(e)$. The estimation for $P_w(e)$ can be obtained along the same line of reasoning. \blacksquare

3 Ensemble of concatenated convolutional codes

3.1 Fundamental facts on convolutional codes

We now briefly recall some basic notion on convolutional codes and we fix some notation which will be used throughout this paper.

Given a \mathbb{Z}_2 -vector space V , we will denote by $V((z))$ the space of Laurent series with coefficients in V . There are several important subspaces of $V((z))$: the subspace of formal power series $V[[z]]$, the subspace of polynomials $V[z]$, the subspace of Laurent polynomials $V[z, z^{-1}]$, the subspace of rational functions $V(z)$. If $v \in V((z))$, $v(t)$ denotes the coefficient in v of z^t , so that we can write $v = \sum_t v(t)z^t$. Given $v \in V((z))$, we define the *support* of v as

$$\text{supp}(v) = \{t \in \mathbb{Z} \mid v(t) \neq 0\}.$$

Given $v_1, v_2 \in V((z))$ and $t_1 \in \mathbb{Z}$ we define the concatenation of v_1 and v_2 at t_1 as the Laurent series $v_1 \vee_{t_1} v_2$ defined by

$$(v_1 \vee_{t_1} v_2)(t) = \begin{cases} v_1(t) & \text{if } t < t_1 \\ v_2(t) & \text{if } t \geq t_1 \end{cases}$$

We will also consider multiple concatenations of Laurent series $v_1 \vee_{t_1} v_2 \vee_{t_2} v_3 \cdots \vee_{t_{m-1}} v_m$ for concatenation times $t_1 < t_2 < \dots < t_{m-1}$. If $v \in V((z))$ and $I \subseteq \mathbb{Z}$, we define the restriction of v to I as the element $v|_I \in V^I$ such that $(v|_I)(t) = v(t)$ for every $t \in I$. If I_1 and I_2 are disjoint consecutive intervals and $v_1 \in V^{I_1}$, $v_2 \in V^{I_2}$, we also write $v_1 \vee v_2$ to denote the element in $V^{I_1 \cup I_2}$ obtained by the concatenation of the two sequences.

In this paper, a convolutional code will be any mapping

$$E : \mathbb{Z}_2^r((z)) \rightarrow \mathbb{Z}_2^s((z))$$

for which there exists $A \in \mathbb{Z}_2^{r \times s}(z) \cap \mathbb{Z}_2^{r \times s}[[z]]$ such that

$$Eu = u \cdot A, \quad u \in \mathbb{Z}_2^r((z)).$$

A is called the symbol of the encoder E . E is called polynomial if its symbol A is of polynomial type, namely, $A = \sum_{i=0}^M A(i)z_i$. Moreover, if $A(M) \neq 0$, M is called the memory length of E . An encoder is said to be recursive if none of the A_{ij} is polynomial. An encoder is said to be non-catastrophic if

$$Eu \in \mathbb{Z}_2^s[z^{-1}, z] \Rightarrow u \in \mathbb{Z}_2^r[z^{-1}, z],$$

equivalently, if there exists $B \in \mathbb{Z}_2^{s \times r}[z, z^{-1}]$ such that $AB = Id$. Also non-catastrophicity can be equivalently expressed as the existence of $\delta > 0$ such that

$$w_H(E(u)) \geq \delta w_H(u), \quad \forall u \in \mathbb{Z}_2^r((z)).$$

The canonical state space of E is defined as the quotient space

$$X = \frac{\mathbb{Z}_2^r[z^{-1}]}{\mathbb{Z}_2^r[z^{-1}] \cap E^{-1}(\mathbb{Z}_2^s[z^{-1}])}.$$

Given $u \in \mathbb{Z}_2^r((z))$, we can define the associated state sequence as

$$x = \sum_k x(k)z^k, \quad x(k) = (z^{k-1}u)^- + (\mathbb{Z}_2^r[z^{-1}] \cap E^{-1}(\mathbb{Z}_2^s[z^{-1}]))^-.$$

Rationality of the A is equivalent to the fact that X is a finite dimensional \mathbb{Z}_2 -space.

A sequence $u \in \mathbb{Z}_2^r((z))$ is called an error event if there exist $t_1 < t_2$ such that $\text{supp}(u) \subseteq [t_1, t_2]$ and the corresponding state sequence x is such that $\text{supp}(x) = [t_1 + 1, t_2]$. Notice that this implies that necessarily $u(t_1) \neq 0$ and $\text{supp}(E(u)) \subseteq [t_1, t_2]$. $t_2 - t_1 + 1$ is called the length of the error event and $[t_1, t_2]$ its window action. If u_1 and u_2 are two error events with windows action, respectively, $[t_1, t_2]$ and $[t_2 + 1, t_3]$ we can consider the concatenation $u_1 \vee_{t_2+1} u_2$ which coincides with $u_1 + u_2$. It is evident that every finite support input sequence u such that $E(u)$ has also finite support, can be obtained by a concatenation of error events and 0 inputs.

The block truncations of a convolutional code E are defined as follows. Fix $N \in \mathbb{N}$ and consider the block code E^N obtained by restricting the inputs of the convolutional encoder E to those inputs supported inside $[0, N - 1]$ and considering the projection of the output on the coordinates also in $[0, N - 1]$. Namely,

$$E^N : \mathbb{Z}_2^{rN} \rightarrow \mathbb{Z}_2^{sN}$$

is defined by

$$E^N(u(0), u(1), \dots, u(N - 1)) = (y(0), y(1), \dots, y(N - 1))$$

if

$$E(u(0) + u(1)z + \dots + u(N - 1)z^{N-1}) = y(0) + y(1)z + \dots + y(N - 1)z^{N-1} + o(z^{N-1}).$$

Whenever we will need it, the input space \mathbb{Z}_2^{rN} will be identified with the subspace of $\mathbb{Z}_2^r[z]$ consisting of the polynomials of degree up to $N - 1$.

Notice that in this way it might be that, at time N , the system is not in the 0 state ($x(N) \neq 0$). In the applications we usually prefer to consider suitable terminations of such codewords in such a way to bring back the system to the zero state. Notice that there always exists an integer $\nu \geq 0$ (not depending on the particular u or on N such that by extending u to a \tilde{u} supported inside $[0, N + \nu - 1]$ we can insure that $x_{N+\nu} = 0$. The output is then also observed in the extended window $[0, N + \nu - 1]$. Notice that the ν extra bits in $[N, N + \nu - 1]$ are functions of u (can be interpreted as parity checks) and, when used in concatenation schemes, they are not passed through the interleaver. It is therefore clear that the use of the terminating sequence and the corresponding extension of the output can only improve the performance of the convolutional encoder and of its concatenations. If we upper bound the error probabilities of the concatenation where we exclusively consider the truncation at level N without terminating sequences we will have also established an upper bound for the case when terminating sequences are instead used. From now on we will consider the truncated block codes E^N without any termination.

3.2 Regular ensembles of concatenated codes

Consider k causal convolutional encoders

$$E_i : \mathbb{Z}_2^r((z)) \rightarrow \mathbb{Z}_2^{s_i}((z)).$$

and consider the families of their truncations $E_i^N : \mathbb{Z}_2^{rN} \rightarrow \mathbb{Z}_2^{s_i N}$.

Suppose we have a family of subgroups $G_N \leq S_{rN}$. We can then consider the family of ensembles of block concatenated codes induced by E_i^N and G_N . The goal of this section is to establish asymptotic estimations on the average word and bit error probability $P_w(e)$ and $P_b(e)$ for $N \rightarrow +\infty$.

In order to achieve general results we will need to make assumptions on the family of subgroups G_N which roughly will insure compatibility when N varies and that the number of invariants for these group actions will not grow up with N . Before giving the exact definition we set some notation: given any subset $\mathcal{I} \subseteq \{0, \dots, N-1\}$ we denote by $S_{rN}^{\mathcal{I}}$ the subgroup of those permutations acting exclusively on the bits of those vectors whose position is in \mathcal{I} . We have a natural isomorphism

$$\Psi_{N,\mathcal{I}} : S_{r|\mathcal{I}|} \rightarrow S_{rN}^{\mathcal{I}}.$$

We can now give the following definition:

Definition 3 *The family of subgroups G_N is said to be regular if*

1. *Given any N and given any subset $\mathcal{I} \subseteq \{0, \dots, N-1\}$ we have that*

$$\Psi_{N,\mathcal{I}}(G_{|\mathcal{I}|}) = G_N \cap S_{rN}^{\mathcal{I}}.$$

2. *There exists a positive integer ν and a map $w : \mathbb{Z}_2^r \rightarrow \mathbb{N}^\nu$ with the following properties:*

- (a) $w(0) = 0$.
- (b) *The elements of the canonical basis e_i are in $w(\mathbb{Z}_2^r)$ for $i = 1, \dots, \nu$.*
- (c) *Consider the vector extension*

$$\begin{aligned} w : \mathbb{Z}_2^{rN} &\rightarrow \mathbb{N}^\nu \\ w(u_0, \dots, u_{N-1}) &= \sum_{j=0}^{N-1} w(u_j). \end{aligned} \tag{7}$$

Then, given any $u, v \in \mathbb{Z}_2^{rN}$, we have that

$$\langle u \rangle^N = \langle v \rangle^N \Leftrightarrow w(u) = w(v).$$

$w(u)$ is called the weight vector of the input word u .

Some comments on the above definition. Property 1. simply says that the permutations in G_N which only act on a set of $L \leq N$ components are essentially given by the permutation in G_L . Property 2. instead says that w is a complete set of invariants for the action of G_N on \mathbb{Z}_2^{rN} . The map w can be easily extended using the sum definition (7) to $\mathbb{Z}_2^r[z]$: such extension will be called the weight function associated with the family G_N . Let $u, v \in \mathbb{Z}_2^r[z]$ be such that $w(u) = w(v)$. Fix N in such a way that $\text{supp}(u), \text{supp}(v) \subseteq [0, N-1]$. Then, property (c) implies that $\langle u \rangle^N = \langle v \rangle^N$. Hence, $w_H(u) = w_H(v)$. As a consequence, there must exist a function $\|\cdot\| : \mathbb{N}^\nu \rightarrow \mathbb{N}$ such that $w_H(u) = \|\cdot\|$ for every $u \in \mathbb{Z}_2^r[z]$. Define also $|w| = w_1 + \dots + w_\nu$.

We present few basic examples of regular families which are the ones mostly considered in the applications.

Example 1 The classical case is when $G_N = S_{rN}$: this is when the interleaver acts on the single bits. In this case the only invariant is given by the Hamming weight: therefore in this case $\nu = 1$ and $w(u) = w_H(u)$. This is the case mostly considered in the literature.

Example 2 Another interesting case is when $G_N = S_N$ the subgroup of permutations acting on the symbols, namely the vectors of \mathbb{Z}_2^r . In this case there are many more invariants: as many as the elements in \mathbb{Z}_2^r . Therefore in this case $\nu = 2^r$ and $w(u)$ is a vector with 2^r components, which can be thought to be indexed by the vectors in \mathbb{Z}_2^r : if $x \in \mathbb{Z}_2^r$, $w(u)_x$ is the number of times, in the block u the vector x appears.

Example 3 Example 2 can be generalized allowing some permutation to take place also on the symbol space \mathbb{Z}_2^r . Let H be a subgroup of S_r . H^N can be thought as the subgroup of S_{rN} consisting of the permutations acting componentwise on the symbol space \mathbb{Z}_2^r . We consider $G_N = S_N \times H^N$. In this case the invariants are as many as the equivalence classes for the action of H in \mathbb{Z}_2^r . In the specific case when $H = S_r$ we have that such invariant are simply the possible Hamming weights of the elements in \mathbb{Z}_2^r : $0, 1, \dots, r$. In this case $\nu = r + 1$ and $w(u) = (w(u)_0, \dots, w(u)_r)$ with $w(u)_j$ the number of symbols of Hamming weight j in the block sequence u .

Example 4 A further possibility is to consider separated permutations: assume that

$$r = r_1 + r_2 + \dots + r_q$$

and decompose

$$\mathbb{Z}_2^r = \mathbb{Z}_2^{r_1} \oplus \dots \oplus \mathbb{Z}_2^{r_q}, \quad \mathbb{Z}_2^{rN} = \mathbb{Z}_2^{r_1N} \oplus \dots \oplus \mathbb{Z}_2^{r_qN}.$$

This yields canonical inclusions $S_{r_jN} \subseteq S_{rN}$. If $G_N^{(j)}$ are subgroups of S_{r_jN} , for $j = 1, \dots, q$ satisfying the properties above, also the subgroup

$$G_N = G_N^{(1)} \times \dots \times G_N^{(q)}$$

of S_{rN} satisfies the above properties. The simplest case is when $G_N^{(j)} = S_{r_jN}$ for all j . In this case $\nu = q$ and

$$w(u_1, \dots, u_q) = (w_H(u_1), \dots, w_H(u_q)).$$

From now we will assume that G_N is a regular family with weight function $w : \mathbb{Z}_2^{rN}[z] \rightarrow \mathbb{N}^\nu$. From now on we will identify an equivalence class $\lambda = \langle u \rangle$ with $w(u)$ and the set of equivalence classes Λ with $\text{Im}(w) \subseteq \mathbb{N}^\nu$. $\mathcal{M}_{w,N}$ will denote the multiplicity of the equivalence class $w(u) = w$ inside \mathbb{Z}_2^{rN} .

We now present some preliminary results on these regular families which will play an important role in the derivation of our results.

Lemma 4 *There exist positive constants $\rho, \tilde{\rho}$ such that*

$$\tilde{\rho}|w| \leq \|w\| \leq \rho|w|, \quad \forall w \in \mathbb{N}^\nu. \quad (8)$$

Proof It follows from property (b) that there exist input elements $\eta_i \in \mathbb{Z}_2^r$ such that $w(\eta_i) = e_i$. Moreover, because of (a), $\eta_i \neq 0$ for every i . Let $w \in \mathbb{N}^\nu$. The input word $u \in \mathbb{Z}_2^{r|w|}$ containing exactly w_i times the element η_i for $i = 1, \dots, \nu$ has weight $w(u) = w$. Then,

$$\|w\| = \|w(u)\| = w_H(u) = \sum_i w_i w_H(\eta_i).$$

The result now easily follows by taking

$$\rho = \max_i w_H(\eta_i), \quad \tilde{\rho} = \min_i w_H(\eta_i). \quad \blacksquare$$

For the multiplicity indices $\mathcal{M}_{w,N}$ we have the following useful estimation.

Lemma 5 *If $w = (w_1, \dots, w_\nu)$, then,*

$$\mathcal{M}_{w,N} \geq \binom{N}{w_1, \dots, w_\nu} \quad (9)$$

Proof If $|w| > N$, it is obvious. Therefore let us assume that $|w| \leq N$. Let $\eta_i \in \mathbb{Z}_2^r$ be chosen as in the proof of previous lemma. The input words $u \in \mathbb{Z}_2^{rN}$ containing exactly w_i times the element η_i for $i = 1, \dots, \nu$ and $N - |w|$ elements equal to 0 has weight w . Hence, the result follows. \blacksquare

Given $u, v \in \mathbb{Z}_2^{rN}$, we denote

$$G_N(u, v) = \{\sigma \in G_N \mid \sigma u = v\}.$$

Since w is a complete invariant for the group action we have that

$$G_N(u, v) \neq \emptyset \Leftrightarrow w(u) = w(v).$$

$G_N(u, u)$ is the stabilizer of u , previously defined and it will be denoted by $G_N(u)$.

The following result gathers some facts which will be used in Section 5.

Lemma 6 (a) *If u and v are such that $w(u) = w(v)$, then*

$$|G_N(u)| = |G_N(u, v)| = |G_N(v)|.$$

In particular all $G_N(u, v)$ for which $w(u) = w(v)$ have all the same cardinality

(b)

$$\left| \frac{G_N}{G_N(u)} \right| = \mathcal{M}_{w(u), N}.$$

(c) *If $u, v \in \mathbb{Z}_2^{rN}$ have disjoint support, then*

$$\left| \frac{G_N}{G_N(u) \cap G_N(v)} \right| \geq \mathcal{M}_{w(u), N-|w(v)|} \mathcal{M}_{w(v), N-|w(u)|}$$

Proof (a): if $\sigma \in G_N(u, v)$,

$$\begin{aligned} \tau \in G_N(u) &\mapsto \sigma\tau \in G_N(u, v) \\ \tau \in G_N(v) &\mapsto \tau\sigma \in G_N(u, v) \end{aligned}$$

are both bijections and this proves the claim.

(b): We can construct a bijection between the quotient set $G_N/G_N(u)$ and the equivalence class $\langle u \rangle^N$ by considering

$$\sigma G_N(u) \mapsto \sigma u$$

It is easy to see that the above mapping is well-defined and surjective. Injectivity follows from the fact that if $\sigma_1 u = \sigma_2 u$, then $\sigma_2^{-1} \sigma_1 \in G_N(u)$.

(c): We first need to set some notation. If $\mathcal{I} \subseteq \{0, \dots, N-1\}$ let $\phi_{\mathcal{I}, N} : \mathbb{Z}_2^{rN} \rightarrow \mathbb{Z}_2^{r|\mathcal{I}|}$ be the mapping obtained by simply deleting the coordinates which are not in \mathcal{I} . Instead, let $j_{\mathcal{I}, N} : \mathbb{Z}_2^{r|\mathcal{I}|} \rightarrow \mathbb{Z}_2^{rN}$ be the mapping which adds zeroes in the coordinate not present in \mathcal{I} . Clearly, $\phi_{\mathcal{I}, N} \circ j_{\mathcal{I}, N} = Id$. Let $\mathcal{I} = \text{supp}(v)^c$. Let $\bar{u} = \phi_{\mathcal{I}, N}(u)$. Choose any $\bar{u}' \in \langle \bar{u} \rangle^{|\mathcal{I}|}$ and let $\bar{\sigma}_1 \in G_{|\mathcal{I}|}$ be such that $\bar{\sigma}_1 \bar{u} = \bar{u}'$. Let $u' = j_{\mathcal{I}, N}(\bar{u}')$ and let $\sigma_1 = \psi_{\mathcal{I}, N}(\bar{\sigma}_1) \in G_N$ by the property (a) of Definition 3. Clearly, by construction, $\sigma_1 u = u'$ and $\sigma_1 v = v$. Moreover, u' and v still have disjoint support. Let now $\mathcal{J} = \text{supp}(u')^c$. Let $\bar{v} = \phi_{\mathcal{J}, N}(v)$. Choose any $\bar{v}' \in \langle \bar{v} \rangle^{|\mathcal{J}|}$ and let $\bar{\sigma}_2 \in G_{|\mathcal{J}|}$ be such that $\bar{\sigma}_2 \bar{v} = \bar{v}'$. Let $v' = j_{\mathcal{J}, N}(\bar{v}')$ and let $\sigma_2 = \psi_{\mathcal{J}, N}(\bar{\sigma}_2) \in G_N$ again by the property (a) of Definition 3. Clearly, by construction, $\sigma_2 v = v'$ and $\sigma_2 u' = u'$. Notice therefore that

$$\sigma_2 \sigma_1 u = u', \quad \sigma_2 \sigma_1 v = v'.$$

From this it immediately follows that the mapping

$$(\bar{u}', \bar{v}') \mapsto \sigma_2 \sigma_1 [G_N(u) \cap G_N(v)]$$

is injective. Notice now that, the choice of \bar{u}' can be done in

$$\mathcal{M}_{w(u), N-|\text{supp}(v)|} \geq \mathcal{M}_{w(u), N-|w(v)|}$$

different ways. Once \bar{u}' has been fixed, \bar{v}' can be chosen in

$$\mathcal{M}_{w(v), N-|\text{supp}(u')|} \geq \mathcal{M}_{w(v), N-|w(u')|} = \mathcal{M}_{w(v), N-|w(u)|}$$

different ways. This yields the result. ■

3.3 Recursivity and non-catastrophicity

The behavior of the concatenated ensemble as N varies depends on structural properties of the constituent convolutional codes and on the concatenation scheme. To this purpose we give two fundamental definitions.

Definition 7 A convolutional encoder E is said to be w -recursive if for every $i = 1, \dots, \nu$, we have that

$$w(u) = e_i \Rightarrow w_H(E(u)) = +\infty.$$

The following remark will play an important role in the sequel

Remark: If a convolutional encoder is w -recursive, there always exist inputs u with $|w(u)| = 2$ and such that $|w_H(E(u))| < +\infty$. Indeed, given any $\eta \in \mathbb{Z}_2^r$ with $w(\eta) = e_1$, consider the input sequence $u = \eta z^0$. Let x be the corresponding state sequence. Since, the state space is finite, for sure there exist $j_1 < j_2$ such that $x_{j_1} = x_{j_2}$. As a consequence, if we consider the input sequence $\tilde{u} = u + z^{j_2-j_1}u$ the corresponding state sequence $\tilde{x} = x + z^{j_2-j_1}x$ yields $\tilde{x}_{j_2} = 0$. This implies that $E\tilde{u}$ is compactly supported inside $[0, j_2]$ and has finite Hamming weight.

Definition 8 The k -uple (E_1, E_2, \dots, E_k) is said to be w -non-catastrophic if there exists $\delta > 0$ such that for every input word vectors $u^1, u^2, \dots, u^k \in \mathbb{Z}_2^r((z))$ such that $w(u^i) = w$ for all i we have that

$$\sum_{i=1}^k w_H(E_i(u_i)) \geq \delta|w| \quad (10)$$

We now discuss the system theoretic meaning of the above definitions showing how they relate to the classical concepts of recursivity and catastrophicity recalled in the Appendix. Let us start with recursivity. In Example 1, w -recursivity corresponds to the classical notion of recursivity. In Example 2 the situation is different: w -recursivity means that the output corresponding to any input supported on just one time instant must have infinite Hamming weight. To capture the difference with respect to the classical notion, if A is the symbol of the convolutional encoder E , we have that E is w -recursive if and only if uA is not polynomial for any $u \in \mathbb{Z}_2^r \setminus \{0\}$. Instead classical recursivity only requires that uA is not polynomial for any $u \in \mathbb{Z}_2^r \setminus \{0\}$ of Hamming weight 1. If

$$A = \begin{bmatrix} (1-z)^{-1} & 1 & 1+z \\ (1-z)^{-1} & 1+z & 1 \end{bmatrix}$$

the corresponding encoder is recursive but not w -recursive. In the Example 4 everything depends on the way the subgroups of permutations acting on each separated set of inputs have been chosen. In the case when these subgroups coincide with the all possible permutations, w -recursivity is again the same than classical recursivity.

Consider now w -non-catastrophicity. Differently from the classical notion, this is a property of the all k -uple of encoders and not of a single one. Notice however, that if a single encoder, say E_1 for the sake of simplicity, happens to be non-catastrophic, it holds (see Appendix), for a suitable $\delta > 0$,

$$w_H(E(u)) \geq \delta w_H(u)$$

for every $u \in \mathbb{Z}_2^r((z))$. Using Lemma 4 we clearly obtain (10). Hence if a single encoder is non-catastrophic, the k -uple is w -non-catastrophic disregarding of the choice of the group of permutations. The following example shows a situation where non-catastrophicity can also emerge in situation when all encoders are catastrophic.

Example 5 Let $k = 2$, $r = 3$, $s_1 = s_2 = 2$. Consider the situation of independent permutations on the three input channels (see Example 4), namely $G_N = S_N^3$. The invariants are the Hamming weights of each single channel: $w(u) = (w_H(u^1), w_H(u^2), w_H(u^3))$ where $u = (u^1, u^2, u^3)$. Consider the encoders E_i corresponding to the two symbols

$$A_1 = \begin{bmatrix} 1 & q_1(z) \\ 0 & q_2(z) \\ 0 & q_3(z) \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0 & q_2(z) \\ 0 & q_3(z) \\ 1 & q_1(z) \end{bmatrix}.$$

The pair (E_1, E_2) is w -non-catastrophic if and only if the encoder represented by the pair $(q_2(z), q_3(z))$ is non-catastrophic. If we consider coprime factorizations

$$q_i(z) = \frac{a_i(z)}{b_i(z)}, a_i(z), b_i(z) \in \mathbb{Z}_2[z],$$

it is well known that non-catastrophicity is equivalent to the fact that the pair $a_2(z), a_3(z)$ does not have common polynomial divisors except eventual trivial shifts z^j . Notice that each single encoder in this case is not injective and catastrophic since there are compact support inputs in the kernel. Notice, finally, that E_1 and E_2 are w -recursive if and only if each $q_i(z)$ is not of polynomial type.

This is the main result we will present:

Theorem 9 *Let $k \geq 2$. Assume that the k -uple (E_1, E_2, \dots, E_k) is w -non-catastrophic and composed of w -recursive convolutional encoders. Then, there exists $C_1, C_2 > 0$ and $\gamma_0 > 0$ such that, for $\gamma < \gamma_0$ and for any N , the following asymptotic estimations hold true:*

$$\frac{C_1}{N^{k-2}} \leq \overline{P_w(e)}^N \leq \frac{C_2}{N^{k-2}}, \quad \frac{C_1}{N^{k-1}} \leq \overline{P_b(e)}^N \leq \frac{C_2}{N^{k-1}}$$

This result will be obtained through an upper bound estimation carried on in Section 4 and a lower bound estimation done in Section 5.

4 Upper bounds on the error probabilities

In this section we will establish upper bound estimations on the error probabilities. From (4) we immediately obtain

$$\overline{P_w(e)}^N \leq \sum_{w \in \mathbb{N}^\nu \setminus \{0\}} \frac{A_w^{1,N}(\gamma) \cdots A_w^{k,N}(\gamma)}{\mathcal{M}_w^{k-1}}, \quad \overline{P_b(e)}^N \leq \sum_{w \in \mathbb{N}^\nu \setminus \{0\}} \frac{\|w\| A_w^{1,N}(\gamma) \cdots A_w^{k,N}(\gamma)}{rN \mathcal{M}_w^{k-1}} \quad (11)$$

where $A_w^{i,N}(D)$ are the enumerative weight functions associated with E_i^N . We now need to exploit the fact that the considered block codes are truncations of fixed convolutional codes to give estimations of $A_w^{i,N}(D)$.

4.1 Combinatorics of convolutional codes

In this paragraph we establish some combinatorial results about convolutional codes which will be instrumental for our upper bounds.

Consider a causal convolutional encoder $E : \mathbb{Z}_2^s((z)) \rightarrow \mathbb{Z}_2^s((z))$. Given $w \in \mathbb{N}^\nu$, $d, n, l \in \mathbb{N}$, we denote by $A'_{w,d,n,l}$ the number of distinct input sequences $u \in \mathbb{Z}_2^s[[z]]$ with $u(0) \neq 0$, with input weight vector w and an output codeword $y = E(u)$ of Hamming weight d , obtained by concatenating n full error events, whose total length is l . We denote by $A_{w,d,n,l}$ the number of distinct input sequences $u \in \mathbb{Z}_2^s[[z]]$ with $u(0) \neq 0$, with input weight vector w and an output codeword $y = E(u)$ of Hamming weight d , obtained by concatenating $n-1$ full error events and an eventual n -th partial part of a last error event, whose total length is l . We also define

$$A'_{w,d,n} = \sum_{l=0}^{+\infty} A'_{w,d,n,l}, \quad A_{w,d,n} = \sum_{l=0}^{+\infty} A_{w,d,n,l}.$$

Notice moreover that the following concatenation identities hold true

$$A_{w,d,n} = \sum_{\substack{w^1, \dots, w^n \\ w^1 + \dots + w^n = w}} \sum_{\substack{d^1, \dots, d^n \\ d^1 + \dots + d^n = d}} \left(\prod_{j=1}^{n-1} A'_{w^j, d^j, 1} \right) A_{w^n, d^n, 1}, \quad (12)$$

$$A_{w,d,n,l} = \sum_{\substack{w^1, \dots, w^n \\ w^1 + \dots + w^n = w}} \sum_{\substack{d^1, \dots, d^n \\ d^1 + \dots + d^n = d}} \sum_{\substack{l^1, \dots, l^n \\ l^1 + \dots + l^n = l}} \left(\prod_{j=1}^{n-1} A'_{w^j, d^j, l^j} \right) A_{w^n, d^n, l^n} . \quad (13)$$

We also introduce the generating functions

$$A_{w,n}(D) = \sum_{d \geq 0} A_{w,d,n} D^d , \quad A_{w,n,l}(D) = \sum_{d \geq 0} A_{w,d,n,l} D^d ,$$

$$A'_{w,n}(D) = \sum_{d \geq 0} A'_{w,d,n} D^d , \quad A'_{w,n,l}(D) = \sum_{d \geq 0} A'_{w,d,n,l} D^d .$$

Lemma 10 *There exists a constant $\beta > 0$ such that*

$$A_{w,d,n,l} = 0 \text{ if } l > \beta(|w| + d) .$$

Proof We start by proving that there exists $\beta > 0$ such that $A'_{w,d,1,l} = 0$ if $l > \beta(|w| + d)$. By contradiction assume that there exists, for E , a sequence of error events u_s (for $s \in \mathbb{N}$) with support inside $[1, l_s]$, input weights w_s , and output weights d_s satisfying

$$\lim_{s \rightarrow +\infty} \frac{l_s}{|w_s| + d_s} = +\infty . \quad (14)$$

As a consequence of (14) there exist sequences $a_s, b_s \in \mathbb{N}$ with $1 \leq a_s \leq b_s \leq l_s$ and

$$\lim_{s \rightarrow +\infty} b_s - a_s = +\infty ,$$

such that

$$\begin{aligned} u^s(t) &= 0 && \text{for } a_s \leq t \leq b_s , \\ \tilde{E}(u^s)(t) &= 0 && \text{for } a_s \leq t \leq b_s , . \end{aligned}$$

It follows from classical considerations that, if s is sufficiently large, the state corresponding to the input/output pair $(u^s, E(u^s))$ is 0 for some $t_s \in [a_s, b_s]$ contradicting the fact that u_s is an error event. Hence, $A'_{w,d,1,l} = 0$ if $l > \beta(|w| + d)$ for some $\beta > 0$. Notice that we also have $A_{w,d,1,l} = 0$ if $l > \beta(|w| + d)$. The result now follows from a straightforward application of the decomposition (13). ■

Lemma 11 *There exist positive constants C, α_1 , and α_2 such that*

$$A_{w,d,n} \leq C \alpha_1^{|w|} \alpha_2^d \quad \forall w \in \mathbb{N}^\nu , d \in \mathbb{N} , n \in \mathbb{N} . \quad (15)$$

Proof We clearly have

$$A'_{w,d,1,l} \leq \alpha^l , \quad A_{w,d,1,l} \leq \alpha^l , \quad \text{where } \alpha = 2^{\sum_{k=1}^m r_k} .$$

Inserting the above estimations in (13), using standard combinatorial arguments and the inequality (consequence of Stirling)

$$\binom{n+m}{m} \leq e^n e^m ,$$

we obtain

$$\begin{aligned} A_{w,d,n,l} &\leq \left[\sum_{\substack{w^1, \dots, w^n \\ w^1 + \dots + w^n = w}} \sum_{\substack{l^1, \dots, l^n \\ l^1 + \dots + l^n = l}} \sum_{\substack{d^1, \dots, d^n \\ d^1 + \dots + d^n = d}} 1 \right] \alpha^l \\ &\leq \left[\prod_{k=1}^m \binom{w_k + n - 1}{n - 1} \right] \binom{l + n - 1}{n - 1} \binom{d + n - 1}{n - 1} \alpha^l \\ &\leq e^{|w|} e^{(m+2)n} e^l e^d \alpha^l . \end{aligned}$$

Using Lemma 10 and the fact that $A_{w,d,n} = 0$ if $n > |w|$, we finally obtain

$$A_{w,d,n} \leq \sum_{l \leq \beta(|w|+d)} A_{w,d,n,l} \leq e^{(m+3)|w|} \left[\sum_{l \leq \beta(|w|+d)} (e\alpha)^l \right] e^d \leq \frac{e\alpha}{e\alpha - 1} [e^{(m+3)}(e\alpha)^\beta]^{|w|} [e(e\alpha)^\beta]^d .$$

This proves the result. ■

The above result in particular implies that the power series $A_{w,n}(D)$ have always positive convergence ratio.

The above estimation are valid for every convolutional encoder. However, if E is w -recursive, we also have that

$$A_{e_i,d,1} = 0, \quad \forall d.$$

This implies that

$$n > |w|/2 \Rightarrow A_{w,d,n} = 0. \quad (16)$$

which, together with estimation (15), will play a crucial role in the sequel.

The conditional weight enumerating functions of the truncated block encoder E^N can be finally estimated as

$$A_w^N(\gamma) = \sum_{l=1}^N \sum_{n=1}^N \binom{N-l+n-1}{n-1} A_{w,n,l}(\gamma) \leq \sum_{n=1}^N \binom{N+n-1}{n-1} A_{w,n}(\gamma). \quad (17)$$

4.2 Upper bounds on the error probabilities

We can now start the estimation of the error probabilities.

From now on we assume that the k -uple (E_1, E_2, \dots, E_k) is w -non-catastrophic.

From (4), substituting (17), we have now an estimation of the bit error probability in the convolutional setting

$$P_b(e) \leq \sum_w \frac{\|w\|}{Nr} \sum_{n_1=1}^N \dots \sum_{n_k=1}^N \frac{\binom{N+n_1-1}{n_1-1} \dots \binom{N+n_k-1}{n_k-1}}{\mathcal{M}_w^{k-1}} A_{w,n_1}^1(\gamma) \dots A_{w,n_k}^k(\gamma) \quad (18)$$

Using the inequality (9) and

$$\frac{N!}{(N-|w|)!} \geq \frac{N^{|w|}}{e^{|w|}}$$

we further obtain

$$P_b(e) \leq \sum_q \frac{\rho q}{Nr} \sum_{n_1=1}^N \dots \sum_{n_k=1}^N \binom{N+n_1-1}{n_1-1} \dots \binom{N+n_k-1}{n_k-1} \frac{e^{(k-1)q} (q!)^{k-1}}{N^{(k-1)q}} \alpha_{q,n_1,\dots,n_k}(\gamma), \quad (19)$$

where

$$\alpha_{q,n_1,\dots,n_k}(D) = \sum_{w \in \mathbb{N}^\nu : |w|=q} \frac{A_{w,n_1}^1(D) \dots A_{w,n_k}^k(D)}{\binom{q}{w_1, w_2, \dots, w_\nu}^{k-1}} .$$

Our aim is to obtain upper bounds for the right side term in (19). We first estimate the term $\alpha_{q,n_1,\dots,n_k}(\gamma)$. We have the following result.

Proposition 12 *There exist constants $\gamma_0 > 0$, $\mu > 0$, and $C > 0$ such that if $0 < \gamma < \gamma_0$, then,*

$$\alpha_{q,n_1,\dots,n_k}(\gamma) \leq C[\mu\gamma]^{\delta q} \quad \forall q \in \mathbb{N}, n_1, \dots, n_k \in \mathbb{N} .$$

Proof In the assumption that $\gamma > 0$, using Lemma 11 we obtain (without loss of generality we assume the constants C , α_1 , and α_2 work for all encoders)

$$\begin{aligned} A_{w,n_1}^1(\gamma) \cdots A_{w,n_k}^k(\gamma) &= \sum_{d \geq \delta|w|} \left[\sum_{d_1 + \cdots + d_k = d} A_{w,d_1,n_1}^1 \cdots A_{w,d_k,n_k}^k \right] \gamma^d \\ &\leq C^k \alpha_1^{k|w|} \sum_{d \geq \delta|w|} \sum_{d_1 + \cdots + d_k = d} [\alpha_2 \gamma]^d \leq C^k \alpha_1^{k|w|} \sum_{d \geq \delta|w|} [2^k \alpha_2 \gamma]^d. \end{aligned} \quad (20)$$

Taking $0 \leq \gamma \leq \gamma_0 = 1/(2^{k+1}\alpha_2)$ we obtain

$$A_{w,n_1}^1(\gamma) \cdots A_{w,n_k}^k(\gamma) \leq 2C^k [2^k \alpha_1^{k/\nu} \alpha_2 \gamma]^{\delta|w|}.$$

Moreover,

$$\alpha_{q,n_1,\dots,n_k}(\gamma) = \sum_{w_1 + \cdots + w_\nu = q} \frac{A_{w,n_1}^1(\gamma) \cdots A_{w,n_k}^k(\gamma)}{\binom{q}{w_1, w_2, \dots, w_\nu}^{k-1}} \leq \left[\sum_{w_1 + \cdots + w_\nu = q} \frac{1}{\binom{q}{w_1, w_2, \dots, w_\nu}^{k-1}} \right] 2C^k [2^k \alpha_1^{k/\nu} \alpha_2 \gamma]^{\delta|w|}.$$

We now conclude showing that the above summation can be bounded above by a constant independent of q . It is clearly sufficient to do it in the case when $k = 2$. This can be done by induction on $\nu \geq 2$. If $\nu = 2$, we have that

$$\sum_{w_1 + w_2 = q} \frac{1}{\binom{q}{w_1, w_2}} \leq 2 + \sum_{w_1=1}^{q-1} \frac{1}{\binom{q}{w_1}} \leq 2 + (q-1)q^{-1} \leq 3.$$

Assume that there exists a constant a_ν not depending on q such that

$$\sum_{w_1 + \cdots + w_\nu = q} \frac{1}{\binom{q}{w_1, w_2, \dots, w_\nu}} \leq a_\nu$$

for every q . We can estimate

$$\begin{aligned} \sum_{w_1 + \cdots + w_{\nu+1} = q} \frac{1}{\binom{q}{w_1, w_2, \dots, w_{\nu+1}}} &= \sum_{w_{\nu+1}=0}^q \frac{1}{\binom{q}{w_{\nu+1}}} \sum_{w_1 + \cdots + w_\nu = q - w_{\nu+1}} \frac{1}{\binom{q - w_{\nu+1}}{w_1, w_2, \dots, w_\nu}} \\ &\leq \sum_{w_{\nu+1}=0}^q \frac{1}{\binom{q}{w_{\nu+1}}} a_\nu \leq a_2 a_\nu \end{aligned}$$

This completes the proof. ■

We now analyze the remaining part of estimation (19):

$$\binom{N + n_1 - 1}{n_1 - 1} \cdots \binom{N + n_k - 1}{n_k - 1} \frac{e^{(k-1)q} (q!)^{k-1}}{N^{(k-1)q}}. \quad (21)$$

It is at this point that the recursivity property comes into the picture. If all constituent encoders are w -recursive we can assume that $n_i - 1 \leq \lfloor q/2 \rfloor$ in the internal summation of (19). Therefore we only need to estimate the above expression for this range of values. In this case the above expression is in particular uniformly bounded with respect to N as shown by the following result

Lemma 13 *There exists $\zeta > 0$ such that for any $2 \leq q \leq \nu N$ the following estimations hold:*

$$\begin{aligned} \left(\frac{N + \lfloor \frac{q}{2} \rfloor}{\lfloor \frac{q}{2} \rfloor} \right)^k \frac{e^{(k-1)q} (q!)^{k-1}}{N^{(k-1)q}} &\leq N^{2-k(\frac{q}{2} - \lfloor \frac{q}{2} \rfloor + 1)} \zeta^q, \\ \left(\frac{N + r_1}{r_1} \right) \cdots \left(\frac{N + r_k}{r_k} \right) \frac{e^{(k-1)q} (q!)^{k-1}}{N^{(k-1)q}} &\leq N^{1-k(\frac{q}{2} - \lfloor \frac{q}{2} \rfloor + 1)} \zeta^q \quad r_k \leq \lfloor \frac{q}{2} \rfloor, \quad r_1 + \cdots + r_k < k \lfloor \frac{q}{2} \rfloor. \end{aligned}$$

Proof Using Stirling approximation we obtain, for some suitable constant $C_1 > 0$,

$$\begin{aligned} \binom{N + \lfloor \frac{q}{2} \rfloor}{\lfloor \frac{q}{2} \rfloor} &\leq C_1 \frac{(N + \lfloor \frac{q}{2} \rfloor)^{N + \lfloor \frac{q}{2} \rfloor}}{\lfloor \frac{q}{2} \rfloor^{\lfloor \frac{q}{2} \rfloor}} \frac{1}{N^N} \\ &\leq C_1 \frac{N^{\lfloor \frac{q}{2} \rfloor}}{\lfloor \frac{q}{2} \rfloor^{\lfloor \frac{q}{2} \rfloor}} \left(\frac{N + \lfloor \frac{q}{2} \rfloor}{N} \right)^{\lfloor \frac{q}{2} \rfloor} \left(\frac{N + \lfloor \frac{q}{2} \rfloor}{N} \right)^N. \end{aligned} \quad (22)$$

Notice now that since $q \leq \nu N$, we have that

$$\left(\frac{N + \lfloor \frac{q}{2} \rfloor}{N} \right)^{\lfloor \frac{q}{2} \rfloor} \leq \left(1 + \frac{\nu}{2} \right)^{\frac{q}{2}}.$$

Using this, together with

$$\left(\frac{N + \lfloor \frac{q}{2} \rfloor}{N} \right)^N \leq e^{\frac{q}{2}}$$

inside (22), we finally obtain

$$\binom{N + \lfloor \frac{q}{2} \rfloor}{\lfloor \frac{q}{2} \rfloor} \leq C_1 \frac{N^{\lfloor \frac{q}{2} \rfloor}}{\lfloor \frac{q}{2} \rfloor^{\lfloor \frac{q}{2} \rfloor}} \left[\left(1 + \frac{\nu}{2} \right) e \right]^{\frac{q}{2}}. \quad (23)$$

On the other hand, Stirling also gives the estimation

$$q! \leq C_2 \frac{q^q}{e^q} q^{1/2}. \quad (24)$$

From (23) and (24) using the fact that, since $q \geq 2$,

$$2 \lfloor \frac{q}{2} \rfloor \leq q \leq 3 \lfloor \frac{q}{2} \rfloor$$

we obtain

$$\begin{aligned} \left(\frac{N + \lfloor \frac{q}{2} \rfloor}{\lfloor \frac{q}{2} \rfloor} \right)^k \frac{e^{(k-1)q} (q!)^{k-1}}{N^{(k-1)q}} &\leq C_1^k C_2^{k-1} N^{(k \lfloor \frac{q}{2} \rfloor - (k-1)q)} \frac{q^{(k-1)q}}{\lfloor \frac{q}{2} \rfloor^{k \lfloor \frac{q}{2} \rfloor}} q^{k-1/2} \left[\left(1 + \frac{\nu}{2} \right)^k e^{1 - \frac{k}{2}} \right]^q \\ &\leq C_1^k C_2^{k-1} N^{(k \lfloor \frac{q}{2} \rfloor - (k-1)q)} q^{((k-1)q - k \lfloor \frac{q}{2} \rfloor)} q^{k-1/2} \left[\left(1 + \frac{\nu}{2} \right)^k e^{1 - \frac{k}{2}} 3^{\frac{k}{2}} \right]^q \end{aligned} \quad (25)$$

Notice now that

$$\begin{aligned} (k-1)q - k \lfloor \frac{q}{2} \rfloor &= (k-1)q - k \frac{q}{2} + k \left(\frac{q}{2} - \lfloor \frac{q}{2} \rfloor \right) \\ &\geq k - 2 + k \left(\frac{q}{2} - \lfloor \frac{q}{2} \rfloor \right) \end{aligned}$$

Hence,

$$\begin{aligned} q^{((k-1)q - k \lfloor \frac{q}{2} \rfloor)} &= q^{((k-1)q - k \lfloor \frac{q}{2} \rfloor) - [k - 2 + k(\frac{q}{2} - \lfloor \frac{q}{2} \rfloor)]} q^{k-2+k(\lfloor \frac{q}{2} \rfloor - \lfloor \frac{q}{2} \rfloor)} \\ &\leq N^{((k-1)q - k \lfloor \frac{q}{2} \rfloor) - [k - 2 + k(\frac{q}{2} - \lfloor \frac{q}{2} \rfloor)]} \nu^{kq} q^{2k} \end{aligned}$$

Inserting this estimation in (25), we obtain

$$\left(\frac{N + \lfloor \frac{q}{2} \rfloor}{\lfloor \frac{q}{2} \rfloor} \right)^k \frac{e^{(k-1)q} (q!)^{k-1}}{N^{(k-1)q}} \leq C_1^k C_2^{k-1} N^{2-k(\frac{q}{2} - \lfloor \frac{q}{2} \rfloor + 1)} q^{3k} \left[\nu \left(1 + \frac{\nu}{2} \right)^k e^{1 - \frac{k}{2}} 3^{\frac{k}{2}} \right]^q \quad (26)$$

This proves the first estimation. To obtain the second one just notice that, if for instance $r_1 < \lfloor q/2 \rfloor$, it holds

$$\binom{N+r_1}{r_1} \leq \frac{\lfloor \frac{q}{2} \rfloor}{N} \binom{N+\lfloor \frac{q}{2} \rfloor}{\lfloor \frac{q}{2} \rfloor}.$$

Using this inequality and the previous estimation, we easily obtain the second estimation as well. ■

We can now give the final estimation on $\overline{P_b(e)}^N$.

Theorem 14 *There exist $\gamma_0 > 0$, $C \geq 0$ such that*

$$P_b(e) \leq C \frac{1}{N^{k-1}}, \quad P_w(e) \leq C \frac{1}{N^{k-2}}, \quad \forall \gamma \in [0, \gamma_0], \quad \forall N \in \mathbb{N}. \quad (27)$$

Proof Using Proposition 12 and Lemma 13 we obtain

$$\begin{aligned} P_b(e) &= \sum_q \frac{\rho q}{Nr} \sum_{n_1=1}^N \cdots \sum_{n_k=1}^N \binom{N+n_1-1}{n_1-1} \cdots \binom{N+n_k-1}{n_k-1} \frac{e^{(k-1)q} (q!)^{k-1}}{N^{(k-1)q}} \alpha_{q, n_1, \dots, n_k}(\gamma) \\ &\leq \sum_q \frac{\rho q}{Nr} N^{2-k(\frac{q}{2}-\lfloor \frac{q}{2} \rfloor+1)} \zeta^q \alpha_{q, \lfloor \frac{q}{2} \rfloor+1, \dots, \lfloor \frac{q}{2} \rfloor+1}(\gamma) \\ &\quad + \sum_q \frac{\rho q}{Nr} N^{1-k(\frac{q}{2}-\lfloor \frac{q}{2} \rfloor+1)} \zeta^q \sum_{n_1+\dots+n_k < k\lfloor \frac{q}{2} \rfloor+k} \alpha_{q, n_1, \dots, n_k}(\gamma) \\ &\leq \frac{1}{N^{k-1}} \sum_s \frac{2\rho s}{r} \zeta^{2s} \alpha_{2s, s+1, \dots, s+1}(\gamma) + \frac{1}{N^k} \left[\sum_s \frac{(2s+1)\rho}{r} \zeta^{2s+1} \alpha_{2s+1, s+1, \dots, s+1}(\gamma) + \right. \\ &\quad \left. + \sum_q \frac{\rho q}{Nr} N^{1-k(\frac{q}{2}-\lfloor \frac{q}{2} \rfloor+1)} \zeta^q \sum_{n_1+\dots+n_k < k\lfloor \frac{q}{2} \rfloor+k} \alpha_{q, n_1, \dots, n_k}(\gamma) \right]. \end{aligned} \quad (28)$$

Notice now that the estimation in Proposition 12 implies that the three function series in γ

$$\begin{aligned} &\sum_s \frac{2\rho s}{r} \zeta^{2s} \alpha_{2s, s+1, \dots, s+1}(\gamma) \quad \sum_s \frac{(2s+1)\rho}{r} \zeta^{2s+1} \alpha_{2s+1, s+1, \dots, s+1}(\gamma) \\ &\quad \sum_q \frac{\rho q}{Nr} N^{1-k(\frac{q}{2}-\lfloor \frac{q}{2} \rfloor+1)} \zeta^q \sum_{n_1+\dots+n_k < k\lfloor \frac{q}{2} \rfloor+k} \alpha_{q, n_1, \dots, n_k}(\gamma) \end{aligned}$$

are uniformly convergent for γ sufficiently small and their limit are power series with positive convergence ratio. This proves the result for $P_b(e)$. The rest follows from the inequality

$$P_w(e) \leq NP_b(e). \quad \blacksquare$$

5 Lower bounds on error probabilities

Lower bounds will be obtained from an analysis of the distribution of minimal distances of the codes of the ensemble. We start with the following simple fact.

Proposition 15 *It holds,*

$$P_w(e) \geq \sup_{d>0} p^d P(d(C_\sigma) \leq d) \quad (29)$$

Proof Fix $d > 0$. We can estimate, using Lemma 1,

$$P_w(e) = \sum_{\sigma \in (G_N)^k} P_w(e|\mathcal{C}_\sigma)P(\mathcal{C}_\sigma) \geq \sum_{\substack{\sigma \in (G_N)^k \\ d(\mathcal{C}_\sigma) \leq d}} P_w(e|\mathcal{C}_\sigma)P(\mathcal{C}_\sigma) \geq p^d P(d(\mathcal{C}_\sigma) \leq d).$$

This yields the result. ■

The final estimation will be obtained by a lower bound estimation of the probability $P(d(\mathcal{C}_\sigma) \leq d)$. The fundamental fact which we are going to use is that every w -recursive encoder admits input sequences of weight 2 whose output is compactly supported (see Remark after Definition 3). More precisely, let $\eta \in \mathbb{Z}_2^r$ be such that $w(\eta) = e_1$. Given any $i = 1, \dots, k$, we can find $t_i \in \mathbb{N}$, such that if we denote $u_i = \eta + \eta z^{t_i}$, we have that $E_i u_i$ is compactly supported inside $[0, t_i]$ and its Hamming weight is, say, ϵ_i . Put $\epsilon = \sum_i \epsilon_i$. The following result holds true.

Proposition 16 *There exists $C > 0$ such that*

$$P(d(\mathcal{C}_\sigma) \leq \epsilon) \geq \frac{C}{N^{k-2}} \quad (30)$$

Proof

Consider

$$E_{s,t}^N = \bigcup_{s'=0}^{N-t-1} G_N(\eta z^s + \eta z^{s+t_1}, \eta z^{s'} + \eta z^{s'+t}).$$

and

$$F_s^N = E_{s,t_2}^N \times \dots \times E_{s,t_k}^N. \quad (31)$$

Notice that

$$\sigma = (\sigma_2, \dots, \sigma_k) \in F_s^N \Rightarrow d(\mathcal{C}_\sigma) \leq \epsilon. \quad (32)$$

We are going to estimate from below the probability of the union, on s , of the subsets F_s^N .

It follows from (a) of Lemma 6 that

$$|E_{s,t}^N| = (N-t)|G_N(\eta z^s + \eta z^{s+t_1})|. \quad (33)$$

Arguing similarly, if $s_2 > s_1$ and $s_2 \neq s_1 + t_1$, we can estimate

$$|E_{s_1,t}^N \cap E_{s_2,t}^N| \leq (N-t)(N-t-1)|G_N(\eta z^{s_1} + \eta z^{s_1+t_1}) \cap G_N(\eta z^{s_2} + \eta z^{s_2+t_1})|. \quad (34)$$

We now want to estimate the number of left lateral classes of these stabilizers inside G_N .

Denote by M_1 the number of elements $x \in \mathbb{Z}_2^r$ such that $w(x) = e_1$ and by M_2 the number of elements $x \in \mathbb{Z}_2^r$ such that $w(x) = 2e_1$. It is then clear that

$$M_{2e_1} = \frac{M_1^2 N(N-1)}{2} + M_2 N$$

Hence, by (b) of Lemma 6 we obtain

$$\frac{|G_N|}{|G_N(\eta z^{s_1} + \eta z^{s_1+t_1})|} = \frac{M_1^2 N(N-1)}{2} + M_2 N. \quad (35)$$

On the other hand, using (c) of Lemma 6 we also obtain,

$$\frac{|G_N|}{|G_N(\eta z^{s_1} + \eta z^{s_1+t_1}) \cap G_N(\eta z^{s_2} + \eta z^{s_2+t_1})|} \geq \left[\frac{M_1^2 (N-2)(N-3)}{2} \right]^2 \quad (36)$$

From (33), (34), (35), and (36) we now obtain the estimation

$$\begin{aligned}
P\left(\bigcup_{s=0}^{\lfloor \frac{N}{t_1+1} \rfloor - 1} F_{(t_1+1)s+1}^N\right) &\geq \sum_{s=0}^{\lfloor \frac{N}{t_1+1} \rfloor - 1} P\left(F_{(t_1+1)s+1}^N\right) - \sum_{s_1 \neq s_2} P\left(F_{(t_1+1)s_1+1}^N \cap F_{(t_1+1)s_2+1}^N\right) \\
&\geq \sum_{s=0}^{\lfloor \frac{N}{t_1+1} \rfloor - 1} \prod_{j=2}^k P\left(E_{(t_1+1)s+1, t_j}^N\right) - \sum_{s_1 \neq s_2} \prod_{j=2}^k P\left(E_{(t_1+1)s_1+1, t_j}^N \cap E_{(t_1+1)s_2+1, t_j}^N\right) \\
&\geq \sum_{s=0}^{\lfloor \frac{N}{t_1+1} \rfloor - 1} \prod_{j=2}^k (N - t_j) \left(\frac{|G_N(\eta z^s + \eta z^{s+t_1})|}{|G_N|}\right)^{k-1} - \sum_{s_1 \neq s_2} \prod_{j=2}^k (N - t_j)(N - t_j - 1) \left(\frac{|G_N(\eta z^{s_1} + \eta z^{s_1+t_1}) \cap G_N(\eta z^{s_2} + \eta z^{s_2+t_1})|}{|G_N|}\right)^{k-1} \\
&\geq \sum_{s=0}^{\lfloor \frac{N}{t_1+1} \rfloor - 1} \prod_{j=2}^k (N - t_j) \left(\frac{2}{M_1^2 N(N-1) + 2M_2 N}\right)^{k-1} - \sum_{s_1 \neq s_2} \prod_{j=2}^k (N - t_j)(N - t_j - 1) \left(\frac{2}{M_1^4 (N-2)^2 (N-3)^2}\right)^{k-1} \\
&\geq \left\lfloor \frac{N}{t_1+1} \right\rfloor \prod_{j=2}^k (N - t_j) \left(\frac{2}{M_1^2 N(N-1) + 2M_2 N}\right)^{k-1} \left[1 - \frac{\lfloor \frac{N}{t_1+1} \rfloor - 1}{2} \prod_{j=2}^k (N - t_j - 1) \left(\frac{2(M_1^2 N(N-1) + 2M_2 N)}{M_1^4 (N-2)^2 (N-3)^2}\right)^{k-1}\right]
\end{aligned} \tag{37}$$

We now distinguish two different cases: $k = 2$ and $k > 2$. In the first case, $k = 2$, the last row in (37) converges, for $N \rightarrow +\infty$, to

$$C = \frac{2}{(t_1 + 1)M_1^2} \left[1 - \frac{1}{(t_1 + 1)M_1^2}\right] > 0.$$

Hence, in this case, for N sufficiently large,

$$P\left(\bigcup_{s=0}^{\lfloor \frac{N}{t_1+1} \rfloor - 1} F_{(t_1+1)s+1}^N\right) \geq \frac{C}{2} \tag{38}$$

In the second case, $k > 2$, it is immediate to obtain the estimation,

$$P\left(\bigcup_{s=0}^{\lfloor \frac{N}{t_1+1} \rfloor - 1} F_{(t_1+1)s+1}^N\right) \geq \frac{C}{N^{k-2}} \tag{39}$$

This clearly yields the result. ■

We can now prove the lower bound result.

Theorem 17 *There exists $C > 0$ such that*

$$P_b(e) \geq C \frac{1}{N^{k-1}}, \quad P_w(e) \geq C \frac{1}{N^{k-2}}, \quad \forall \gamma \in [0, \gamma_0], \quad \forall N \in \mathbb{N}. \tag{40}$$

Proof We first notice that it is sufficient to find lower bound estimations for the word error probability since it holds

$$P_b(e) \geq \frac{1}{N} P_w(e).$$

Using now (29) and (32), together with (38) and (39), we obtain the final estimation

$$P_w(e) \geq p^\epsilon P(d(\mathcal{C}_\sigma \leq \epsilon) \geq p^\epsilon P\left(\bigcup_{s=0}^{\lfloor \frac{N}{t_1+1} \rfloor - 1} F_{(t_1+1)s+1}^N\right) \geq p^\epsilon \frac{C}{N^{k-2}}. \tag{41}$$

■

Remark 1: In the case $k = 2$ and $G_N = S_{rN}$, the above lower bounds can also be obtained as consequence of a result in [8] on the distribution of minimal distances for ensemble of turbo codes. They indeed prove that

$$P(d(\mathcal{C}_\sigma) \geq d) \leq C \frac{N^{k-2}}{d^k} + O(1/N). \quad (42)$$

As a consequence, if $k = 2$ we can find d and \bar{N} such that

$$P(d(\mathcal{C}_\sigma) \leq d) \geq 1/2, \quad \forall N \geq \bar{N}$$

and conclude as before. However, if $k > 2$, (42) is apparently of no help since for fixed d , the right term tends to $+\infty$ for $N \rightarrow +\infty$. Indeed, in the case $k > 2$ the mass of the probability distribution of $d(\mathcal{C}_\sigma)$ is in some sense moving towards $+\infty$, in particular $\mathbb{E}(d(\mathcal{C}_\sigma)) \rightarrow +\infty$. However, the error probability seems to be more influenced by the way the tail of the distribution $d(\mathcal{C}_\sigma)$ goes to 0 for small d and this type of information can not be obtained from (42) if $k > 2$.

Remark 2: For $k = 2$, a different lower bound on error probabilities had also been obtained in [5] where the author proves that for any permutation, the concatenated code has a minimum distance growing at most logarithmically in N and, as a consequence, error probabilities can only decay polynomially in N . The order of decay in N is of course larger in [5] since it treats the worst possible case and actually it explicitly depends on the noise parameter of the channel. We also notice that in [1] it has been shown that minimum distance growing logarithmically can indeed be achieved. It remains an open question if for $k > 2$ we can achieve superpolynomial decays in the error probabilities for some choice of the permutations.

6 Further results in the classical case

In the previous analysis we have not considered the situation when not all encoders are recursive. In particular we have not analyzed the situation when there is none or just one recursive encoder. The analysis in the general case looks quite complicate due to the possibility of encoders which are only partially recursive with respect to part of the input space. However, in the classical case when the input space is scalar $r = 1$ and the group action is given by the all symmetry group (see Example 1) we can complete our analysis showing how the asymptotic behavior of the error probabilities only depends on the number of recursive convolutional encoders and that, in particular, if the number of recursive encoders is strictly smaller than two, than there is no interleaver gain not even for the bit error probability. To derive this result we assume that the channel used is the BSC: extension to other channels which we believe to be possible present however some technical problems.

We put ourselves in the situation of scalar inputs, namely $s = 1$ so that the encoders are of the type $E_i : \mathbb{Z}_2((z)) \rightarrow \mathbb{Z}_2^{s_i}((z))$. Moreover, we assume that $G_N = S_N$ so that $\nu = 1$ and $w(u)$ is simply the Hamming weight of the sequence u , the only invariant by the action of the symmetry group. w -recursivity amounts now simply to recursivity.

Assume that the first $k_{\text{rec}} \leq k$ of the encoders E_i are recursive while the remaining ones are not (and therefore, being scalar, will be of polynomial type).

In the case $k_{\text{rec}} \geq 2$ we can obtain exactly the estimation of Theorem 14 for the bit and word error probability with k replaced by k_{rec} . Indeed the upper bounds can be obtained by considering only the recursive encoders, ignoring the remaining ones. For what concerns the lower bound, we instead proceed as in Section 5 noticing that for each non-recursive encoder E_i , there exists $\epsilon_i \in \mathbb{N}$ such that

$$w_H(E_i u) \leq \omega_i \quad \forall u : w(u) = 2.$$

We can then similarly define $\epsilon = \sum_i \epsilon_i$ and for such an ϵ it is easy to see that Proposition 16 still holds true (with k replaced by k_{rec}). The only difference in the proof consists in replacing the set of permutations (31) with

$$F_s^N = E_{s,t_2}^N \times \dots \times E_{s,t_{k_{\text{rec}}}}^N \times S_N^{k-k_{\text{rec}}}. \quad (43)$$

and We can thus prove Theorem 17 with k replaced by k_{rec} .

In the case when $k_{\text{rec}} \leq 1$, it is clear that the averaged word error probability will be bounded away from 0. We will prove in this case a much stronger result, namely that, for any vector permutation σ , the bit error probability (and thus also the word error probability, is bounded away from 0. Of course, such a lower bound can not be obtained by simply working at the level of the code, inputs must necessarily come up into the picture. We recall that this result will be obtained under the stronger assumption that the channel is the BSC.

Let η be any non-zero input sequence such that $\text{supp}(\eta) \subseteq [0, D]$ and such that $\text{supp}(E_1(\eta)) \subseteq [0, D]$: we know that such an input sequence always exists, of Hamming weight 1 in the non-recursive case, of weight 2 in the recursive case. Let $l \geq 0$ and consider

$$u^m = z^{(m-1)(D+1+2l)} z^l \eta, \quad m = 1, \dots, \left\lfloor \frac{N}{D+1+2l} \right\rfloor.$$

Let

$$I_m = [(m-1)(D+1+2l), m(D+1+2l)-1], \quad \bar{I}_m = [(m-1)(D+1+2l)+l, (m-1)(D+1+2l)+l+D].$$

Clearly,

$$\text{supp}(u_m) \subseteq \bar{I}_m \subseteq I_m.$$

For $i = 2, \dots, k$, E_i is polynomial with symbol $a_i(z) \in \mathbb{Z}_2^{1 \times s_i}[z]$ polynomial of degree l_i . Let

$$S_i^m = \text{supp}(E_i u^m).$$

Clearly, $|S_i^m| \leq (l_i + 1)(D + 1)$.

Define moreover

$$U(\sigma_i, i, m) = \sigma_i^{-1}[\sigma_i(I_m) + [-l_i, l_i]]$$

and

$$U(\sigma, m) = \bigcup_i U(\sigma_i, i, m).$$

Lemma 18 *There exists $L \in \mathbb{N}$ and an increasing sequence $m_h \leq hL$ such that*

$$U(\sigma, m_h) \cap U(\sigma, m_{h'}) = \emptyset \quad \forall h < h'.$$

Proof Notice that

$$\begin{aligned} U(\sigma_i, i, m_h) \cap U(\sigma_{i'}, i', m_{h'}) = \emptyset &\Leftrightarrow \sigma_{i'} U(\sigma_i, i, m_h) \cap [\sigma_{i'}(I_{m_{h'}}) + [-l_{i'}, l_{i'}]] = \emptyset \\ &\Leftrightarrow [\sigma_{i'} U(\sigma_i, i, m_h) + [-l_{i'}, l_{i'}]] \cap \sigma_{i'}(I_{m_{h'}}) = \emptyset \\ &\Leftrightarrow \sigma_{i'}^{-1}[\sigma_{i'} U(\sigma_i, i, m_h) + [-l_{i'}, l_{i'}]] \cap I_{m_{h'}} = \emptyset \end{aligned} \quad (44)$$

Define now

$$\tilde{I}_m = \bigcup_{i, i'} \sigma_{i'}^{-1}[\sigma_{i'} U(\sigma_i, i, m) + [-l_{i'}, l_{i'}]].$$

It is clear, for previous considerations, that

$$U(\sigma, m_h) \cap U(\sigma, m_{h'}) = \emptyset \Leftrightarrow \tilde{I}_{m_h} \cap I_{m_{h'}} = \emptyset.$$

Notice moreover that

$$|\tilde{I}_m| \leq L = \sum_{i, i'} ((D + 2 + 2l + 2l_i)(2l_{i'} + 1))$$

independent of m .

We can now construct the sequence in the following recursive way: Define $m_1 = 1$ and given m_1, \dots, m_h define m_{h+1} as

$$m_{h+1} = \min\{j > m_h \mid \tilde{I}_{m_{h'}} \cap I_j = \emptyset \quad \forall h' \leq h\}.$$

Notice that $m_{h+1} - m_h \leq Lh$. The result is therefore proven. \blacksquare

Put now $\tilde{N} = \lfloor N/L \rfloor$, $y_i^h = E_i(u^{m_h})$. For any $h \leq \tilde{N}$ define

$$\Gamma_h = \{y \in \mathbb{Z}_2^{sN} : y_{1|I_{m_h}} = y_{1|I_{m_h}}^h, y_{i|A_{m_h}^i} = y_{i|A_{m_h}^i}^h, i > 2\}.$$

Clearly,

We now clarify the role of the sets Γ_h . Let y be the received sequence (we recall we are transmitting along a BSC) and let \hat{y} be the corresponding ML estimate, namely any element in \mathcal{C}_σ having minimum distance from y . Let \hat{u} be the corresponding input sequence. We have the following result

Lemma 19 *If l has been chosen sufficiently large, we have that for every $h \leq \tilde{N}$,*

$$y \in \Gamma_h \Rightarrow w_H(\hat{u}_{|U(\sigma, m_h)}) \geq 1.$$

Proof Assume that $y \in \Gamma_h$ and, by contradiction, assume that $\hat{u}_{|U(\sigma, m_h)} = 0$. As a consequence $\hat{y}_{i|S_i^{m_h}} = 0$ for all $i \geq 2$. On the other hand, since E_1 has a finite number of states, say ν , it follows that $\hat{y}_{1|I_{m_h}}$ can be decomposed as

$$\hat{y}_{1|I_{m_h}} = \zeta^1 \vee \bar{\zeta}^{\vee g} \vee \zeta^2$$

where $\zeta^1, \zeta^2, \bar{\zeta}$ are blocks of length at most ν and $g \in \mathbb{N}$.

Suppose first that $\zeta = 0$ and consider the input $\hat{u}' = \hat{u} + u_{m_h}$ and its corresponding outputs $\hat{y}'_i = E_i(\hat{u}')$. Notice that

$$\begin{aligned} \hat{y}'_{1|\bar{I}_{m_h}} &= y_{1|\bar{I}_{m_h}}^h, & \hat{y}'_{1|[\bar{I}_{m_h}]^c} &= \hat{y}_{1|[\bar{I}_{m_h}]^c}^h \\ \hat{y}'_{i|S_i^{m_h}} &= y_{i|S_i^{m_h}}^h, & \hat{y}'_{i|(S_i^{m_h})^c} &= \hat{y}_{i|(S_i^{m_h})^c}^h, \quad i \geq 2. \end{aligned}$$

As a consequence,

$$w_H(\hat{y}' - y) < w_H(\hat{y} - y)$$

and this is absurd by the way \hat{y} had been chosen.

If instead $\zeta \neq 0$ we argue as follows. We consider in this case the new input sequence \hat{u}' defined by

$$\hat{u}'(t) = \begin{cases} \hat{u}(t), & \text{if } t \notin I_{m_h} \\ \alpha(t), & \text{if } t \in I_{m_h} \end{cases}$$

where the block α is chosen in such a way that the corresponding state sequence x_1 of E_1 is 0 for $t = (m_h - 1)(D + 1 + 2l) + \nu$ and $t = m_h(D + 1 + 2l) - 1 - \nu$ and, moreover, $\alpha(t) = 0$ for $t \in [(m_h - 1)(D + 1 + 2l) + \nu, m_h(D + 1 + 2l) - 1 - \nu]$. As a consequence,

$$w_H(\hat{y}'_1 - y_1) = w_H([\hat{y}'_1 - y_1]_{|I_{m_h}}) + w_H([\hat{y}'_1 - y_1]_{|I_{m_h}^c}) \leq 2\nu s_1 + w_H(y_1^h) + w_H([\hat{y}_1 - y_1]_{|I_{m_h}^c}), \quad (45)$$

while

$$w_H(\hat{y}_1 - y_1) = w_H([\hat{y}_1 - y_1]_{|I_{m_h}}) + w_H([\hat{y}_1 - y_1]_{|I_{m_h}^c}) \geq 2(l - \nu)/\nu + w_H([\hat{y}_1 - y_1]_{|I_{m_h}^c}). \quad (46)$$

Also we can estimate

$$w_H(\hat{y}'_i - y_i) \leq w_H(\hat{y}_i - y_i) + \sum_i \nu l_i s_i \quad (47)$$

Putting together (45), (46) and (47) we finally obtain,

$$w_H(\hat{y}' - y) \leq 2\nu s_1 + \sum_i \nu l_i s_i - 2(l - \nu)/\nu + w_H(\hat{y}' - y).$$

If l is chosen sufficiently large, we obtain also in this case

$$w_H(\hat{y}' - y) < w_H(\hat{y} - y).$$

This is again absurd and we have thus proven our result. ■

We now define, for any $\theta \in \mathbb{Z}_2^{\tilde{N}}$,

$$\Gamma_\theta = \left(\bigcap_{h:\theta_h=1} \Gamma_h \right) \cap \left(\bigcap_{h:\theta_h=0} (\Gamma_h)^c \right).$$

Lemmas 18 and 19 immediately implies that

$$y \in \Gamma_\theta \Rightarrow w_H(\hat{u}) \geq w_H(\theta). \quad (48)$$

We can now prove our final result:

Theorem 20 *If $k_{\text{rec}} \leq 1$, there exists $C > 0$ such that*

$$P_b(e|\sigma) \geq C, \quad \forall N, \forall \sigma \in S_N. \quad (49)$$

Proof Fixed N and $\sigma \in S_N$. Since the block encoder E_σ^N is \mathbb{Z}_2 -linear, we can assume that the all 0's input word has been transmitted along the channel. Denote by $p < 1/2$ the transition probability in the channel. We have that

$$P(\Gamma_h|u=0) \geq q = p^{D+1+\sum_i(2l_i+1)(2l+D+1)} > 0 \quad (50)$$

From (48) we can estimate the bit error probability as

$$P_b(e|\sigma) = P_b(e|\sigma, u=0) \geq \sum_{\theta \in \mathbb{Z}_2^{\tilde{N}}} \frac{w_H(\theta)}{N} P(\Gamma_\theta|u=0) = \sum_{\theta \in \mathbb{Z}_2^{\tilde{N}}} \frac{w_H(\theta)}{N} \prod_{h:\theta_h=1} P(\Gamma_h|u=0) \prod_{h:\theta_h=0} (1-P(\Gamma_h|u=0)). \quad (51)$$

Let W_h be the Bernoulli random variable which is 1 if $y \in \Gamma_h$ and 0 otherwise. The W_h 's are clearly independent, since the channel is memoryless, hence,

$$\sum_{\theta \in \mathbb{Z}_2^{\tilde{N}}} w_H(\theta) \prod_{h:\theta_h=1} P(\Gamma_h|u=0) \prod_{h:\theta_h=0} (1-P(\Gamma_h|u=0)) = \mathbb{E}\left(\sum_{h=1}^{\tilde{N}} (W_h)\right) = \sum_{h=1}^{\tilde{N}} \mathbb{E}(W_h) = \sum_{h=1}^{\tilde{N}} P(\Gamma_h|u=0), \quad (52)$$

Using (52) and (50) we can now continue the estimation in (51) as

$$P_b(e|\sigma) \geq \frac{\sum_{h=1}^{\tilde{N}} P(\Gamma_h|u=0)}{N} \geq \frac{q\tilde{N}}{N} \geq \frac{q}{2L},$$

if N is sufficiently large. This clearly yields the result. ■

References

- [1] L. Bazzi, M. Mahdian, D.A. Spielman The minimum distance of turbo-like codes *submitted*.
- [2] S. Benedetto, G. Montorsi Design of parallel concatenated convolutional codes *IEEE Trans. on Communications*, 44 (5):591–600, 1996.
- [3] S. Benedetto, G. Montorsi Unveiling turbo codes: some results on parallel concatenated coding schemes *IEEE Trans. on Information theory*, 42 (2):409–428, 1996.
- [4] C. Berrou, A. Glavieux Near optimum error correcting coding and decoding: turbo-codes *IEEE Trans. on Communications*, 44 (10):1261–1271, 1999
- [5] M. Breiling A logarithmic upper bound on the minimum distance of turbo codes *submitted to IEEE Trans. on Information theory*.

- [6] F. Fagnani, R. Garello, B. Scanavino, S. Zampieri Analysis and design of geometrically uniform parallel concatenated coded modulation schemes *submitted to IEEE Trans. on Information theory*.
- [7] H. Jin and R.J. McEliece Coding theorems for turbo code ensembles *IEEE Trans. on Information theory*, 48 (6):1451–1461, 2002.
- [8] N. Kahale, R. Urbanke On the minimum distance of parallel and serially concatenated codes *submitted to IEEE Trans. on Information theory*.
- [9] T. Richardson The geometry of turbo-decoding dynamics *IEEE Trans. on Information theory*, 46 (1):9–23, 2000.
- [10] I. Sason, E. Telatar, R. Urbanke The asymptotic input-output weight distributions and thresholds of convolutional and turbo-like encoders *IEEE Trans. on Information theory*, 48 (12):3052–3061, 2002.