Staircase and other structured linear-time encodable LDPC codes: analysis and design

Federica Garin, Giacomo Como and Fabio Fagnani
Dipartimento di Matematica, Politecnico di Torino
C.so Duca degli Abruzzi 24, 10129 Torino, Italy
Email: fabio.fagnani@polito.it

Abstract

We consider a family of codes which can be seen both as a special kind of serial turbo codes and as LDPC codes having a parity check matrix which is partly random and partly structured. These codes are linear-time encodable, thanks to the turbo structure, and can be decoded as LDPC codes. We provide an ensemble analysis for the waterfall region, on the line of classical results for serial turbo codes, and we find some design parameters.

1 Introduction

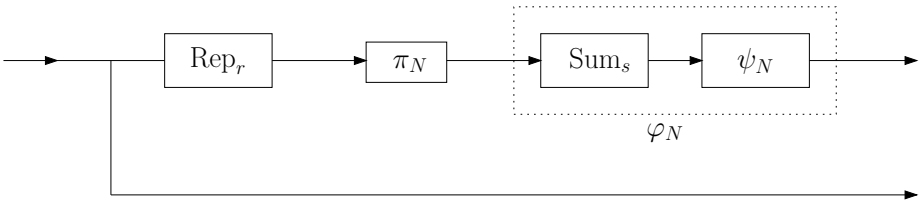
One of the main problems related to LDPC codes is their encoding complexity, which is in general quadratic in the block length, as the generating matrix is not low density. This issue has been addressed in two different ways. On one side there are the results in [8], which allow to construct, for given generic LDPC matrix, equivalent generating matrices with lower encoding complexity. On the other side, there are the constructions of parity check matrices structured in a such a way that allows easy encoding. A successful construction is the one using matrices with a staircase part (i.e. a sub-matrix with ones on the diagonal and on the lower diagonal, and zeros everywhere else), so that the encoder can be seen as a serial concatenation of a repetition code, an interleaver and an accumulator: this gives Repeat-Accumulate codes and their generalization, the Irregular Repeat-Accumulate (IRA) codes, introduced in [5].

In this paper, we follow this second approach, studying LDPC codes which can be encoded with a serial turbo structure. There is a wide literature on analysis and design of IRA (see particularly [9]), but previous work focuses on the design of the degree distribution of the variable nodes (the time-varying number of repetitions) and of the check nodes (the so-called grouping factor). On the contrary, here we investigate the possibility to vary the structured part of the matrix, which is equivalent to choosing a different inner encoder instead of the accumulator. To do so, we focus on the simpler case when the degrees are

constant and we analyze the performance following the classical results for serial turbo codes in [1]. We analyze the performance of schemes with different inner encoders in the waterfall region, showing at first that there is an interleaver gain, i.e. for large enough SNR the average error probability goes to zero when the interleaver length grows to infinity. Then we look at the behavior of the main term when the SNR goes to infinity, as was done in [1] to underline the role of the effective free distance of the inner encoder. The results in [1] generalize to our setting in a non-trivial way, as the relevance of the inner encoder can be shown only expurgating some codes from the ensemble. Our results are theoretical and are coding theorems in the same sense as in [6]: they hold true under ML decoding, on a memoryless binary-input symmetric-output channel (e.g. BSC or BIAWGN).

2 Encoding schemes and parity check matrices

Consider the family of serially concatenated turbo encoders which have the following structure:



By $\mathrm{Rep}_r : \mathbb{Z}_2^N \to \mathbb{Z}_2^{rN}$ we denote the repetition code with rate $1/r$; $\mathrm{Sum}_s : \mathbb{Z}_2^{rN} \to \mathbb{Z}_2^{rN/s}$ is defined by

$$\mathrm{Sum}_s(\boldsymbol{x}) = (x_1 + \ldots + x_s,\, x_{s+1} + \ldots + x_{2s},\, \ldots)$$

i.e. it gives the modulo-2 sum of every block of s bits (s is the grouping factor). Finally, let $\psi(D) : \mathbb{Z}_2^k((D)) \to \mathbb{Z}_2^k((D))$ be a rate-1 non-catastrophic and recursive convolutional encoder, and let $\psi_N : \mathbb{Z}_2^{rN/s} \to \mathbb{Z}_2^{rN/s}$ be the truncated encoder obtained by using the trellis of $\psi(D)$ for $rN/(sk)$ time steps. We will always assume that rN is a multiple of sk, so that the above construction can be properly made (this will be implicitly assumed also when taking limits for $N \to \infty$). As a reminder of properties of convolutional encoders, notice that $\psi(D)$ can be seen as a $k \times k$ matrix whose entries are fractions of polynomials, and that $\psi(D)$ is non-catastrophic if and only if this matrix has an inverse whose entries are Laurent polynomials. Recursiveness of $\psi(D)$ is equivalent to the recursiveness of at least one entry in each column of the matrix. In particular, if $k = 1$, our assumptions imply that $\psi(D) = 1/p(D)$ for some polynomial $p(D)$.

The encoding scheme we are considering is a particular kind of systematic serial turbo encoder; the outer encoder is Rep_r, the inner encoder is $\varphi_N = \psi_N \circ \mathrm{Sum}_s$. The inner encoder φ_N can be considered as the truncation of a

proper convolutional encoder, which is not injective, but the transmission of the systematic bits ensures injectivity and non-catastrophicity of the overall coding scheme. Also notice that φ_N is recursive, in the sense that inputs of weight one produce outputs with weight growing to infinity when $N \to \infty$; this will be essential to our result about the interleaver gain.

The representation as serial turbo codes allows linear-time encoding, and it is also useful for some performance analysis, as stated in the next sections. The decoding can be performed exploiting the fact that these same codes can also be seen as LDPC codes: a parity check matrix can be constructed in the following way. Notice that a pair $(\boldsymbol{u}, \boldsymbol{c}) \in \mathbb{Z}_2^N \times \mathbb{Z}_2^{rN/s}$ belongs to our code if and only if $\boldsymbol{c} = \psi_N \circ \mathrm{Sum}_s \circ \pi_N \circ \mathrm{Rep}_r(\boldsymbol{u})$, which is equivalent to $\mathrm{Sum}_s \circ \pi_N \circ \mathrm{Rep}_r(\boldsymbol{u}) + \psi_N^{-1}(\boldsymbol{c}) = \boldsymbol{0}$ and can be represented with matrices as $[H_N\, K_N]\left[\begin{smallmatrix}\boldsymbol{u}\\ \boldsymbol{c}\end{smallmatrix}\right] = \boldsymbol{0}$.

Notice that H_N is a low-density matrix depending only on r, s and on the permutation π_N, and has at most s ones per row and r ones per column, while K_N is a matrix depending on the choice of ψ, and is also low density, having a number of ones per row and per column bounded by $k(\deg \psi^{-1}(D) + 1)$, where by 'deg' we denote the difference between the largest and the smallest exponent of a Laurent polynomial.

We give here some examples of encoders $\psi(D)$ satisfying our assumptions, and of the corresponding matrices K_N. The properties peculiar to these encoders will be commented later.

(E1) If $k = 1$ and $\psi(D)$ is the accumulator $\psi(D) = 1/(1 + D)$, we have the so-called 'staircase' LDPC codes: K_N has ones on the diagonal and on the lower diagonal, and zeros everywhere else.

(E2) With $k = 1$, $\psi(D) = \frac{1}{1+D+D^3}$ gives

$$K_N = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & \dots \\ 1 & 1 & 0 & 0 & 0 & 0 & \dots \\ 0 & 1 & 1 & 0 & 0 & 0 & \dots \\ 1 & 0 & 1 & 1 & 0 & 0 & \dots \\ 0 & 1 & 0 & 1 & 1 & 0 & \dots \\ 0 & 0 & 1 & 0 & 1 & 1 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \end{bmatrix}$$

(E3) Let $k = 3$ and $\psi(D) = \frac{1}{1+D^3}\begin{bmatrix} 1 & D & D^2 \\ D^2 & 1 & D \\ D & D^2 & 1 \end{bmatrix}$.

Its inverse is $\psi^{-1}(D) = \begin{bmatrix} 1 & D & 0 \\ 0 & 1 & D \\ D & 0 & 1 \end{bmatrix}$, which gives:

$$K_N = \begin{bmatrix} 1 & 0 & 0 & & & & & & \\ 0 & 1 & 0 & & & & & & \\ 0 & 0 & 1 & & & & & & \\ 0 & 1 & 0 & 1 & 0 & 0 & & & \\ 0 & 0 & 1 & 0 & 1 & 0 & & & \\ 1 & 0 & 0 & 0 & 0 & 1 & & & \\ & & & 0 & 1 & 0 & 1 & 0 & 0 \\ & & & 0 & 0 & 1 & 0 & 1 & 0 \\ & & & 1 & 0 & 0 & 0 & 0 & 1 \\ & & & & & & & & \dots \end{bmatrix}$$

(E4) Let $k = 3$ and $\psi(D) = \begin{bmatrix} \frac{1+D+D^2}{1+D^2} & \frac{1}{1+D} & \frac{1}{1+D^2} \\ \frac{1}{1+D} & 1 & \frac{1}{1+D} \\ \frac{D}{1+D} & 0 & \frac{D}{1+D} \end{bmatrix}$.

Its inverse is $\psi^{-1}(D) = \begin{bmatrix} 1+D^{-1} & D^{-1} & 0 \\ 0 & 1 & D^{-1} \\ 1+D^{-1} & D^{-1} & 1+D^{-1} \end{bmatrix}$, which gives:

$$K_N = \begin{bmatrix} \boxed{\begin{matrix} 1&0&0 \\ 0&1&0 \\ 1&0&1 \end{matrix}} & \boxed{\begin{matrix} 1&1&0 \\ 0&0&1 \\ 1&1&1 \end{matrix}} & & \\ & \boxed{\begin{matrix} 1&0&0 \\ 0&1&0 \\ 1&0&1 \end{matrix}} & \boxed{\begin{matrix} 1&1&0 \\ 0&0&1 \\ 1&1&1 \end{matrix}} & \\ & & \boxed{\begin{matrix} 1&0&0 \\ 0&1&0 \\ 1&0&1 \end{matrix}} & \cdots \end{bmatrix}$$

3 Interleaver gain: average error probability

To analyze the performance of the coding schemes we have introduced, we will follow the analysis of serial turbo codes in [1, 6]: we focus on the behavior in the waterfall region and under the assumption that the decoding is Maximum Likelihood. We build an ensemble by fixing ψ and letting the interleaver Π be a random variable uniformly distributed over S_{rN} (the set of all permutations of rN elements), and then we study the average error probability and particularly its behavior when $N \to \infty$.

The coding ensemble presented here is included in the wide class of generalized serial turbo codes studied in [4]; here we state the results as applied to this particular case and we give a rough idea of the proofs, referring the interested reader to [4] for detailed proofs.

Our main result is that, for sufficiently large SNR, there is an interleaver gain: the average bit and word error probabilities go to zero when $N \to \infty$, provided that $r \geq 2$ and $r \geq 3$ respectively. The decay is polynomial in $1/N$, with an exponent increasing with r. More formally, denoting by $\overline{P_b(e)}$ and $\overline{P_w(e)}$ the average bit and word error probabilities:

Theorem 1 *Take $s \geq 2$ and $r \geq 2$. Define $\mu = \lfloor \frac{r+1}{2} \rfloor$ and*

$$d^* = \begin{cases} 1 & \text{if } r \text{ is even} \\ 2 & \text{if } r = 3 \\ 1 + d^{\psi}_{1,tr} & \text{otherwise} \end{cases}$$

where $d^{\psi}_{1,tr}$ is the smallest weight of a truncated error event of ψ_N having an input weight 1 (if $k = 1$, then $d^{\psi}_{1,tr} = 1$).

There exist positive constants γ_0, c_1 and c_2 (depending only on the ensemble, i.e. on $r, s, \psi(D)$) such that, for all $\gamma \leq \gamma_0$:

- $c_1 p^{d^*} N^{-\mu} \leq \overline{P_b(e)} \leq c_2 \gamma^{d^*} N^{-\mu} + O(N^{-\mu-1})$
- $c_1 p^{d^*} N^{-\mu+1} \leq \overline{P_w(e)} \leq c_2 \gamma^{d^*} N^{-\mu+1} + O(N^{-\mu})$

where p is the equivocation probability and γ is the Bhattacharyya noise parameter of the channel. □

We prove the upper bound for $\overline{P_b(e)}$ and the lower bound for $\overline{P_w(e)}$; then Thm. 1 follows as $\overline{P_b(e)} \geq \frac{1}{N}\overline{P_w(e)}$.
The upper bound is based on the Union-Bhattacharyya bound:

$$\overline{P_b(e)} \leq \sum_{w=1}^{N} \sum_{d=w}^{(r+s)N/s} \frac{w}{N} \overline{A_{w,d}}^N \gamma^d$$

where $\overline{A_{w,d}}^N$ is the average number of codewords of a serial ensemble with input weight w and output weight d. For $d \geq w$,

$$\overline{A_{w,d}}^N = \frac{1}{\binom{rN}{rw}} \binom{N}{w} |V_{rw,d-w}^{\varphi_N}| \tag{1}$$

where $V_{h,k}^{\varphi_N}$ is the set of words $\boldsymbol{v} \in \mathbb{Z}_2^{rN}$ such that $\mathrm{w_H}(\boldsymbol{w}) = h$ and $\mathrm{w_H}(\varphi_N(\boldsymbol{v})) = k$.

Then we use some properties of the convolutional encoders to estimate $|V_{rw,d-w}^{\varphi_N}|$. First of all, we need to define $V_{h,k,n}^{\varphi_N}$ to be the set of words $\boldsymbol{v} \in V_{rw,d-w}^{\varphi_N}$ producing exactly n error events of φ_N, plus possibly a final truncated error event not counted by n. Clearly

$$|V_{rw,d-w}^{\varphi_N}| = \sum_{n=0}^{n_{\max}} |V_{rw,d-w,n}^{\varphi_N}|$$

Then, for some constants $a, b > 0$, we have the estimation:

$$|V_{rw,d-w,n}^{\varphi_N}| \leq \binom{N+n}{n} a^{rw} b^{d-w}$$

The recursiveness of φ_N comes into the picture for bounding $n_{\max}$: it ensures $n_{\max} \leq \lfloor rw/2 \rfloor$, as every error event must have input weight at least two. When $r \geq 4$, we use the bound $n_{\max} \leq \lfloor rw/2 \rfloor$ for all terms, except some terms with $w = 1$. We do not give here the proof for $r = 2$ and $r = 3$, which is different because more values of w contribute to the main term of the estimations.

For $r \geq 4$, notice that we have defined d^* in such a way that $d^* - 1$ is exactly the smallest output weight of φ_N that can be obtained when the input weight is r and there are $\lfloor r/2 \rfloor$ error events (plus possibly a truncated one). This comes from the fact that every pair of ones in the output of Rep_r can be permuted by some interleaver in such a way that they are summed up by Sum_s, producing a zero output. The consequence is that if $w = 1$ and $d < d^*$ we know that $n_{\max} \leq \lfloor rw/2 \rfloor - 1$.

Substituting these estimations in Eq. (1) and separating the term with $w = 1$,

we get:

$$\begin{aligned}\overline{P_b(e)} \le & \sum_{d=d^*}^{(r+s)N/s} \frac{1}{\binom{rN}{r}} \sum_{n=0}^{\lfloor r/2 \rfloor} \binom{N+n}{n} a^r b^{d-1} \gamma^d \\ & + \sum_{d=1}^{d^*-1} \frac{1}{\binom{rN}{r}} \sum_{n=0}^{\lfloor r/2 \rfloor - 1} \binom{N+n}{n} a^r b^{d-1} \gamma^d \\ & + \sum_{d=1}^{(r+s)N/s} \sum_{w=2}^{d} \frac{w}{N} \frac{\binom{N}{w}}{\binom{rN}{rw}} \sum_{n=0}^{\lfloor rw/2 \rfloor} \binom{N+n}{n} a^{rw} b^{d-w} \gamma^d\end{aligned}$$

Now we refer to [4] for a formal proof that, for sufficiently small γ, all these series are convergent, and the first is bounded by $c\gamma^{d^*} N^{-\mu}$ while the second and third are bounded by $c(\gamma) N^{-\mu-1}$.

Now we give a sketch of the proof of the lower bound. The key idea is that, for all fixed d,

$$\overline{P_w(e)} \ge p^d \mathbb{P}(\mathrm{d}_N^{\min} \le d)$$

where $\mathrm{d}_N^{\min}$ is the minimum distance of the overall coding scheme (which is a r.v. as the encoder is a r.v.).

We choose $d = d^*$ and we find a lower bound for $\mathbb{P}(\mathrm{d}_N^{\min} \ge d^*)$; for simplicity of notation we consider here only even r. We fix some codewords of the repetition code: $\boldsymbol{c}_a^* = \mathrm{Rep}_r(D^a)$ for all $a \in \mathcal{A} = \{0, \dots, N-1\}$. We also fix an error event of φ_N with input weight 2 and output weight 0, for example with input $1 + D$ and then we construct the following inputs for φ_N: let $\mathcal{B} = \{0, \dots, 2N/s - 1\}^{r/2}$ and for any $\boldsymbol{b} = (b_0, \dots, b_{r/2-1}) \in \mathcal{B}$ define

$$\boldsymbol{u}_{\boldsymbol{b}}^* = \sum_{j=0}^{r/2-1} D^{sb_j + j2N}(1 + D)\,.$$

Clearly $\boldsymbol{c}_a^*, \boldsymbol{u}_{\boldsymbol{b}}^* \in \mathbb{Z}_2^{rN}$ and both have weight r. Also notice that $\mathrm{w_H}\left(\varphi_N(\boldsymbol{u}_{\boldsymbol{b}}^*)\right) = 0$, so that if $\Pi(\boldsymbol{c}_a^*) = \boldsymbol{u}_{\boldsymbol{b}}^*$ then $\mathrm{d}_N^{\min} \le d^*$. Define the events $E_{a,\boldsymbol{b}} = \{\Pi(\boldsymbol{c}_a^*) = \boldsymbol{u}_{\boldsymbol{b}}^*\}$ and $E_a = \bigcup_{\boldsymbol{b} \in \mathcal{B}} E_{a,\boldsymbol{b}}$, so that

$$\begin{aligned}\mathbb{P}(\mathrm{d}_N^{\min} \le d^*) &\ge \bigcup_{a \in \mathcal{A}} \mathbb{P}(E_a) \\ &\ge \sum_a \mathbb{P}(E_a) - \sum_a \sum_{a' \ne a} \mathbb{P}(E_a \cap E_{a'})\end{aligned}$$

Then $\sum_a \mathbb{P}(E_a) = |\mathcal{A}|\,|\mathcal{B}|\,\frac{1}{\binom{rN}{r}} \ge cN^{-\mu+1}$.

For the term with intersections notice that $E_{a,\boldsymbol{b}} \cap E_{a',\boldsymbol{b}} = \emptyset$ if $a \ne a'$ but $b_j = b'_j$ for some j, while

$$\mathbb{P}(E_{a,\boldsymbol{b}} \cap E_{a',\boldsymbol{b}'}) \le \mathbb{P}\big(\Pi(\boldsymbol{c}_a^* + \boldsymbol{c}_{a'}^*) = \boldsymbol{u}_{\boldsymbol{b}}^* + \boldsymbol{u}_{\boldsymbol{b}'}^*\big) = \frac{1}{\binom{rN}{2r}}$$

if $a \neq a'$ and $b_j \neq b'_j$ for all j. So:

$$\sum_a \sum_{a' \neq a} \mathbb{P}(E_a \cap E_{a'}) < |\mathcal{A}|^2 \, |\mathcal{B}|^2 \, \frac{1}{\binom{rN}{2r}} \leq C N^{-2\mu+2}$$

If $r \geq 3$, then $\mu \geq 2$ and this concludes the proof.

For $r = 2$, the proof is slightly different: $\mathcal{A}$ and $\mathcal{B}$ must be chosen smaller by some constant factor, ensuring that $|\mathcal{A}| \, |\mathcal{B}| \, \frac{1}{\binom{2N}{2}} - |\mathcal{A}|^2 \, |\mathcal{B}|^2 \, \frac{1}{\binom{2N}{4}} > 0$.

4 A better smaller ensemble and a design parameter

In the result given in Thm. 1, notice that there is essentially no dependency of the exponents μ and d^* on the choice of the encoder ψ. Looking at traditional serial turbo codes [1], we see that it is natural that μ depends only on the free distance of the outer encoder, but we expect a dependency of the effective free distance d^* on the inner encoder too. What happens with our schemes is that pairs of bits which are repetition of a same information bit can be permuted by some interleaver in such a way that they are summed up by Sum_s, producing a zero output. The value of d^* is given by this 'worse case' scenario.

This remark suggests to consider a smaller family of interleavers, enforcing that ones coming from the same error event of Rep_r cannot end up in positions where they would be summed up by Sum_s. More precisely, we define the set

$$R^N_{r,s} := \left\{ \pi \in S_{rN} : \lfloor i/r \rfloor = \lfloor j/r \rfloor \Rightarrow \lfloor \pi(i)/s \rfloor \neq \lfloor \pi(j)/s \rfloor \right\}$$

What we want to consider is an ensemble of encoders constructed as in Section 2, except that now the permutation is uniformly distributed on $R^N_{r,s}$ instead of all S_{rN}. Additionally to the motivation of finding a more interesting effective free distance, this ensemble turns out to be a natural choice in analogy with classical results for regular LDPC codes: restricting the permutation to $R^N_{r,s}$ is the same as enforcing that the Tanner graph corresponding to the regular part of the matrix, H_N, does not have cycles of length two. This new ensemble is also equivalent to pick H_N uniformly at random in the set of $N \times N$ binary matrices with exactly s ones per row and r ones per column.

As $R^N_{r,s}$ is not a group, we cannot directly apply results from [4] (i.e. use the same proof techniques sketched in the previous section). However, we can slightly modify our techniques for estimating $\mathbb{E}(P_b(e)|R^N_{r,s})$, where $\mathbb{E}$ is taken in the ensemble with Π uniformly distributed in S_{rN}; notice that $\mathbb{E}(P_b(e)|R^N_{r,s})$ is equal to the average $P_b(e)$ when Π is uniformly distributed in $R^N_{r,s}$ which is what we would like to estimate; we will also denote it $\overline{P_b(e)}_{\mathrm{exp}}$.

The key remark is that the probability that a permutation uniformly extracted from S_{rN} belongs to $R^N_{r,s}$ is non-vanishing: $\mathbb{P}(R^N_{r,s}) \to e^{-(r-1)(s-1)/2}$ when $N \to \infty$ (see e.g. [2] Exercise 2.12 p. 59).

Notice that $\mathbb{P}(R^N_{r,s})$ tends to a constant which is strictly smaller than one, so even though the techniques we use are the same usually known as expurgation, the result we will get is not the typical behavior of the ensemble introduced in Sect. 3: we will find the average behavior of a subensemble which is neither vanishing nor typical, but is well characterized.

Our main result is the following:

Theorem 2 *Take $s \geq 2$ and $r \geq 2$. Define $\mu = \lfloor \frac{r+1}{2} \rfloor$ and*

$$d^*_{\text{exp}} = \begin{cases} 2 & r = 2,\, 3 \\ 1 + \frac{r}{2} d^{\psi}_{f,2} & \text{even } r \geq 4 \\ 1 + \frac{r-3}{2} d^{\psi}_{f,2} + \min\left\{ d^{\psi}_{f,2} + d^{\psi}_{1,tr},\, d^{\psi}_{f,3} \right\} & \text{odd } r \geq 5 \end{cases}$$

where $d^{\psi}_{1,tr}$ is defined as in Thm. 1, while $d^{\psi}_{f,2}$ and $d^{\psi}_{f,3}$ are the smallest weight of an error event of $\psi(D)$ having input weight two and three respectively.

There exist positive constants γ_0, c_1 and c_2 (depending only on the ensemble, i.e. on $r, s, \psi(D)$) such that, for all $\gamma \leq \gamma_0$:

- $c_1 p^{d^*_{\text{exp}}} N^{-\mu} \leq \overline{P_b(e)}_{\text{exp}} \leq c_2 \gamma^{d^*_{\text{exp}}} N^{-\mu} + O(N^{-\mu-1})$
- $c_1 p^{d^*_{\text{exp}}} N^{-\mu+1} \leq \overline{P_w(e)}_{\text{exp}} \leq c_2 \gamma^{d^*_{\text{exp}}} N^{-\mu+1} + O(N^{-\mu})$

□

Now we show how the proof sketched in Section 3 for Theorem 1 can be adapted to prove Theorem 2.

For the upper bound, by the union-Bhattacharyya bound:

$$\overline{P_b(e)} \leq \sum_w \sum_d \frac{w}{N} \mathbb{E}(A^N_{w,d}(\Pi) | R^N_{r,s}) \gamma^d$$

where $A^N_{w,d}(\pi)$ is the number of codewords of the concatenated scheme with input Hamming weight w and output Hamming weight d for a given permutation $\pi \in S_{rN}$.

For most of the terms, we will use the estimation

$$\mathbb{E}(A^N_{w,d}(\Pi) | R^N_{r,s}) \leq \frac{\mathbb{E}\big(A^N_{w,d}(\Pi)\big)}{\mathbb{P}(R^N_{r,s})} = \frac{\overline{A_{w,d}}^N}{\mathbb{P}(R^N_{r,s})} \tag{2}$$

and the fact that $\mathbb{P}(R^N_{r,s})$ is bounded away from zero, so that we can exploit all what we know about $\sum_w \sum_d \overline{A_{w,d}}^N \gamma^d$.

We consider separately the term with $w = 1$ (as in the previous section, we are writing the proof for $r \geq 4$). First notice that

$$\mathbb{E}\big(A^N_{1,d}(\Pi) | R^N_{r,s}\big) = N \sum_{\boldsymbol{v} \in V^{\varphi N}_{r,d-1}} \mathbb{P}\big(\Pi(\text{Rep}_r(1)) = \boldsymbol{v} | R^N_{r,s}\big)$$

Then let $S_s^N = \{\boldsymbol{v} \text{ s.t. } \lfloor i/s \rfloor \neq \lfloor j/s \rfloor \, \forall i \neq j : v_i = v_j = 1\}$ and notice that $\boldsymbol{v} \notin S_s^N$ gives $\mathbb{P}\big(\Pi(\mathrm{Rep}_r(1)) = \boldsymbol{v} \cap R_{r,s}^N\big) = 0$, so that

$$\begin{aligned}\mathbb{E}\big(A_{1,d}^N(\Pi)|R_{r,s}^N\big) &= N \sum_{\boldsymbol{v} \in V_{r,d-1}^{\varphi_N} \cap S_s^N} \mathbb{P}\big(\Pi(\mathrm{Rep}_r(1)) = \boldsymbol{v}|R_{r,s}^N\big) \\ &\leq N \, |V_{r,d-1}^{\varphi_N} \cap S_s^N| \, \frac{1}{\binom{rN}{r}\mathbb{P}(R_{r,s}^N)}\end{aligned}$$

Then, $|V_{r,d-1}^{\varphi_N} \cap S_s^N| = \sum_{n=0}^{n_{\max}} |V_{r,d-1,n}^{\varphi_N} \cap S_s^N|$.

The recursiveness of φ_N ensures $n_{\max} \leq \lfloor r/2 \rfloor$, but also notice that if $\mathrm{w_H}(\boldsymbol{v}) = r$ and $\boldsymbol{v} \in V_{r,d-1,n}^{\varphi_N} \cap S_s^N$, then $\mathrm{w_H}(\varphi_N(\boldsymbol{v})) \geq d_{\mathrm{exp}}^*$, so that for $d < d_{\mathrm{exp}}^*$ we have the tighter bound $n_{\max} \leq \lfloor r/2 \rfloor - 1$. Finally, we estimate $|V_{r,d-1,n}^{\varphi_N} \cap S_s^N| \leq |V_{r,d-1,n}^{\varphi_N}|$ and we end the proof as in the previous section.

Let's see how to adapt the proof of the lower bound (again for simplicity let r be even). We take the same $\mathcal{A}$ and $\boldsymbol{c}_a^*$ as in the previous proof. On the contrary, we have to choose different $\boldsymbol{u}_{\boldsymbol{b}}^*$, to produce output weight d_{exp}^*. Let $\boldsymbol{v} \in \mathbb{Z}_2^{rN/s}$ be an input for ψ_N with weight 2 producing an error event of output weight $d_{f,2}^{\psi}$; let L be the length of the error event (the number of trellis steps where it diverges from the all-zero state); also assume that the error event starts at time zero: $\boldsymbol{v} = 1 + D^t$ for some $1 < t \leq kL$. Now define $\mathcal{B} = \{0, \ldots, \lfloor \frac{N}{2skL} \rfloor - 1\}^{r/2}$ and

$$\boldsymbol{u}_{\boldsymbol{b}}^* = \sum_{j=0}^{r/2-1} D^{skLb_j + j2skLN}(1 + D^{st})$$

so that $\mathrm{w_H}(\varphi_N(\boldsymbol{u}_{\boldsymbol{b}}^*)) = d_{\mathrm{exp}}^* - 1$. Re-defining E_a with these new $\boldsymbol{u}_{\boldsymbol{b}}^*$, we have

$$\begin{aligned}\mathbb{E}\left(P_w(e)\middle|R_{r,s}^N\right) &\geq p^{d_{\mathrm{exp}}^*} \mathbb{P}(\mathrm{d}_N^{\min} \leq d_{\mathrm{exp}}^*|R_{r,s}^N) \\ &\geq p^{d_{\mathrm{exp}}^*} \mathbb{P}\Big(\bigcup_{a \in \mathcal{A}} E_a \Big| R_{r,s}^N\Big)\end{aligned}$$

Then we use again the union-intersection bound.

First of all, notice that $\mathbb{P}(E_{a,\boldsymbol{b}} \cap R_{r,s}^N)$ does not depend on a and $\boldsymbol{b}$ and is non-zero. Then, to find a lower bound for $\sum_a \mathbb{P}(E_a|R_{r,s}^N)$, define the events

$$F_{a,\boldsymbol{l},\boldsymbol{i}}^N = \Big\{\Pi(\boldsymbol{c}_a^*) = \sum_{j=0}^{r-1} D^{l_j + i_j}\Big\}$$

and notice that, for any $a \in \mathcal{A}$, $\boldsymbol{b} \in \mathcal{B}$, we have

$$\begin{aligned}\mathbb{P}(R_{r,s}^N) &= \sum_{\substack{l_0 < \cdots < l_{r-1} \\ 0 \leq l_j \leq \frac{rN}{s} - 1}} \sum_{\substack{i_0, \ldots, i_{r-1} \\ 0 \leq i_j \leq s-1}} \mathbb{P}\big(R_{r,s}^N \cap F_{a,\boldsymbol{l},\boldsymbol{i}}^N\big) \\ &= \binom{rN/s}{r} s^r \, \mathbb{P}(R_{r,s}^N \cap E_{a,\boldsymbol{b}})\end{aligned}$$

so that $\mathbb{P}(E_a|R^N_{r,s}) = \dfrac{\mathbb{P}(E_a \cap R^N_{r,s})}{\mathbb{P}(R^N_{r,s})} = \dfrac{|\mathcal{B}|}{\binom{rN/s}{r} s^r}$ and finally

$$\sum_{a\in\mathcal{A}} \mathbb{P}(E_a|R^N_{r,s}) = |\mathcal{A}|\,|\mathcal{B}|\,\frac{1}{\binom{rN}{r} s^r} \geq cN^{-\mu+1}\,.$$

For the term with intersections, we use the simple bound

$$\mathbb{P}(E_{\boldsymbol{a}} \cap E_{\boldsymbol{a}'}|R^N_{r,s}) = \frac{\mathbb{P}(E_{\boldsymbol{a}} \cap E_{\boldsymbol{a}'} \cap R^N_{r,s})}{\mathbb{P}(R^N_{r,s})} \leq \frac{\mathbb{P}(E_{\boldsymbol{a}} \cap E_{\boldsymbol{a}'})}{\mathbb{P}(R^N_{r,s})}$$

Then we exploit the fact that $\mathbb{P}(R^N_{r,s})$ is bounded away from zero and we estimate $\mathbb{P}(E_{\boldsymbol{a}} \cap E_{\boldsymbol{a}'})$ as in the previous section, ending the proof with

$$\sum_{\substack{\boldsymbol{a},\boldsymbol{a}'\in\mathcal{A}\\ \boldsymbol{a}\neq\boldsymbol{a}'}} \mathbb{P}(E_{\boldsymbol{a}} \cap E_{\boldsymbol{a}'}|R^N_{r,s}) \leq \frac{1}{\mathbb{P}(R^N_{r,s})} \sum_{\substack{\boldsymbol{a},\boldsymbol{a}'\in\mathcal{A}\\ \boldsymbol{a}\neq\boldsymbol{a}'}} \mathbb{P}(E_{\boldsymbol{a}} \cap E_{\boldsymbol{a}'}) \leq \tilde{c}N^{-2\mu+2}$$

5 Conclusion, conjectures and open problems

In this paper we have presented an analysis of the average error probability of the ensemble of LDPC codes obtained by a serial interconnection of a regular repetition code with a generic recursive inner code. We have also studied the sub-ensemble obtained by preventing the appearance of 2-cycles in the Tanner graph. We have proved that both ensembles have the same interleaver gain: they have average error probability polynomially going to zero when $1/N \to 0$, with the same exponent. We have found that their behavior when the SNR goes to infinity is not the same, and in the second ensemble it is influenced by a parameter depending on the choice of the inner encoder ψ, providing a design parameter for such schemes.

Our results leave space for further interesting investigations. For classical serial turbo codes, the ensemble analysis has been done not only studying the average error probability, but also finding the typical behavior [3], which turned out to have a sub-exponential decay (much better than the polynomial decay of the average code, but worse than the exponential decay of typical error probability of LDPC regular ensemble). A careful adaptation of the proofs in [7, 3], which is beyond the scope of this paper and will be discussed elsewhere, allows to extend those results to the ensemble described in Section 3, in the following way. Consider the ensemble described in Sect. 3 and define the random variables $X_N = \frac{\log(\mathrm{d}^{\min}_N)}{\log(N(1+r/s))}$ and $Y_N = \frac{\log(-\log(P_w(e)))}{\log(N)}$. When $N \to \infty$, the result is that X_N and Y_N converge in probability to the constant β (the latter only for sufficiently high SNR), where $\beta = 1 - 2/r$. Even for classical serial turbo codes the parameter β depends only on the free distance of the outer encoder. However, we are working on a more detailed analysis which can underline the role played by $d^i_{f,2}$, the smallest output weight of the inner encoder restricted to

inputs of weight two. We conjecture that for the ensemble of codes considered in this paper the key parameter would not be $d^{\varphi}_{f,2}$, which is always zero, but $d^{\psi}_{f,2}$, without the need to restrict the ensemble as in Section 4.

Another important further study concerns the decoding. The simplest idea is to run the Sum-Product iterative decoding on the Tanner graph exactly as it is done for LDPC codes. We are currently investigating the real significance of our distance parameter in real simulations with such decoding, and the first results do not show the clear hierarchy we would expect. We conjecture that this is related to the fact that some encoders have many cycles of small length in the structured part of their Tanner graph, and this can make their performance significantly worse. For example, with $k = 1$, the accumulator has $d^{\psi}_{f,2} = 1$, while the encoder in Example (E2) has $d^{\psi}_{f,2} = 4$, but the first one has no cycles in the structured part of the graph, while the latter has $O(N)$ cycles of length six which can explain why it does not outperform the accumulator. We are currently exploring the possibility to overcome this problem, either by constructing encoders with cycles of reasonably large length, or by focusing on encoders with $k > 1$. This second approach allows both to get more encoders without cycles in the structured part (an example is (E3), which however has only $d^{\psi}_{f,2} = 1$) and to construct encoders which do have cycles of small length on the bitwise level, but if considered blockwise (with symbols of k bits) have a staircase structure: see example (E4), which has $d^{\psi}_{f,2} = 3$. We think that this last kind of codes can provide a better performance when a proper decoder acting on symbols is applied to them.

Acknowledgment

The authors would like to thank Roberto Garello for motivation and helpful discussions.

References

[1] S. Benedetto, D. Divsalar, G. Montorsi and F. Pollara, "Serial concatenation of interleaved codes: Performance analysis, design and iterative decoding", *IEEE Trans. on Inf. Th.*, vol. 44, pp. 909–926, May 1998.

[2] B. Bollobás, *Random Graphs*, Cambridge University Press, 2001.

[3] G. Como, F. Fagnani and F. Garin, "ML Performances of Serial Turbo Codes do not Concentrate", *Proc. of the 4th International Symposium on Turbo Codes and Related Topics*, Munich, Germany, April 2006.

[4] F. Fagnani and F. Garin, "Analysis of serial turbo codes over Abelian groups for Geometrically Uniform constellations", *submitted to SIAM Journal on Discrete Mathematics* (2007).

[5] H. Jin, A. Khandekar and R. J. McEliece, "Irregular Repeat-Accumulate Codes", *Proc. of the 2nd International Symposium on Turbo Codes and Related Topics*, Brest, France, Sept. 2000.

[6] H. Jin and R. J. McEliece, "Coding theorems for turbo code ensembles", *IEEE Trans. on Inf. Th.*, vol. 48, no. 6, pp. 1451–1461, 2002.

[7] N. Kahale and R. Urbanke, "On the minimum distance of parallel and serially concatenated codes", *submitted to IEEE Trans. on Information Theory* (1997), available online: `http://lthcwww.epfl.ch/papers/KaU.ps`

[8] T. Richardson and R. Urbanke, "Efficient Encoding of Low-Density Parity-Check Codes", *IEEE Transactions on Information Theory*, vol. 47, no. 2, pp. 638–656, Feb. 2001.

[9] A. Roumy, S. Guemghar, G. Caire and S. Verdú, "Design Methods for Irregular Repeat-Accumulate Codes", *IEEE Transactions on Information Theory*, vol. 50, no. 8, pp. 1711–1727, August 2004.