Staircase construction with non-systematic polar codes

Carlo Condo, Valerio Bioglio, Ingmar Land
Optical Communications Technology Lab, Huawei Technologies France SASU, Boulogne-Billancourt, France
{*carlo.condo,valerio.bioglio,ingmar.land*}@*huawei.com*

Abstract: We propose staircase codes based on non-systematic polar codes, describing a general framework for encoding and decoding, and presenting simulation results showing the effectiveness of the proposed approach even with short component codes.

1. Introduction

Staircase codes [1] and product codes [2] are powerful concatenated constructions that are particularly attractive for optical communications. They can be viewed as spatially-coupled codes with a specific interleaver shape, and where memory is a single frame, or block, in case of staircase codes and zero blocks for product codes. Their decoding process can be easily parallelized, and thus allows the use of long codes, resulting in good error-correction performance and high throughput. Decoder implementations of staircase and product codes have lower power consumption than inherently more powerful codes, e.g. low-density parity-check codes [3], but the decoding of component codes plays a major role in the total decoding complexity. Polar codes [4] are hardware-friendly, capacity-achieving linear block codes that rely on channel polarization; at infinite code length, polarization leads uncoded bits to be either completely noisy or noise-free. With finite code lengths, polarization is incomplete, and uncoded bits can be sorted in order of reliability. In a code of rate $R = N/K$, the K most reliable uncoded bits carry information, while the remaining $N-K$ are frozen to a known value. Given their low implementation complexity, systematic and non-systematic polar codes alike have been considered as component codes for product codes in [5–7], and a staircase construction with systematic polar codes has been presented in [8,9]. Non-systematic polar codes are characterized by a faster encoding and decoding process. In this work, we propose a method to construct staircase codes with non-systematic polar codes, detailing the encoding and decoding frameworks.

2. Staircase Code With Non-Systematic Polar Codes

Let us define the frozen set of the polar component code of length N with K information bits as $\mathscr{F}$. The N-bit input vector u to the polar encoder is composed of a first half u_{new}, where information and frozen bits are placed, and a second half x' that is composed of previously encoded bits. $N/2$ input vectors u compose an input matrix U. Let us build an input matrix U_0^N, where in every row $0 \leq i < N/2$ each bit position $0 \leq j \leq N/2$ is an information or a frozen bit according to $\mathscr{F}$, while all bit positions $j \geq N/2$ are set to 0. The leftmost half of U_0^N is called U_1, while the rightmost half is X_0'. We encode each row of U_0^N and obtain the encoded matrix X_0^N: we consider the rightmost $N/2 \times N/2$ matrix as X_0, and the leftmost half as X_1'. X_0 is transmitted. We compose U_1^N through a set of column input vectors (U_2) and the columns of X_1'. After encoding, we obtain X_1^N, with X_1 on the lower side and X_2' on the upper side. This is repeated alternating row and column encoding: the encoding process is shown in Fig. 1.

On the receiver side, the decoder relies on a set of logarithmic likelihood ratio (LLR) matrices Λ_k, $0 \leq k \leq W$, representing a decoding window of $W+1$ blocks. While it is implemented as a circular buffer, we can consider the received block being stored in Λ_W and the oldest block being that in Λ_0. Λ_1 is the block output from the decoder, but to decode it we need the information stored in block Λ_0. Let us take Λ_W, whose LLRs are relative to the encoded matrix X_W, that we assume being last encoded by rows. Λ_{W-1} is relative instead to X_{W-1}, encoded by columns. To decode Λ_{W-1} by columns, we have to propagate the LLRs in Λ_W from representing X_W to representing X_W', that was last encoded by columns. This is done by applying a kind of "soft" polar code encoding process to each row of Λ_{W-1}. Polar codes are encoded by multiplying a binary vector with a transformation matrix, an operation commonly represented by a tanner graph constituted of stages of XORs. To propagate LLRs, we use the same tanner graph structure, substituting the $\boxplus$ operation to the XORs, that can be easily expressed as in [10] $a \boxplus b = \text{sgn}(a)\,\text{sgn}(b)\min(a,b)$. The

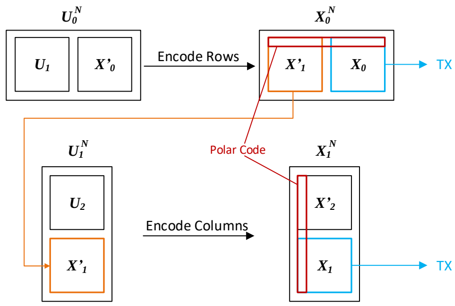

Fig. 1: Encoding process.

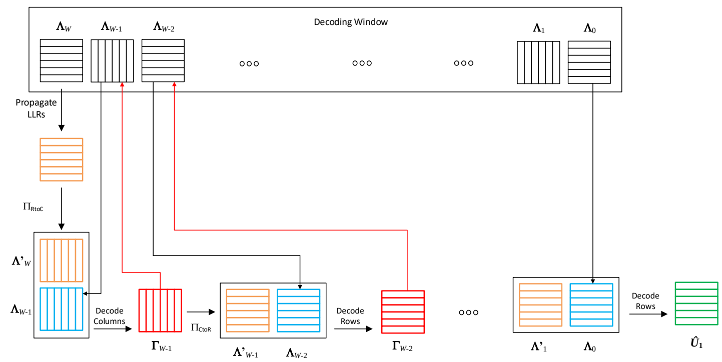

Fig. 2: Decoding process.

LLR propagation is concatenated with an interleaver Π_{RtoC}. After obtaining thus Λ'_W, relative to X'_W, we can prepend it to Λ_{W-1}, and decode each column. The decoding process returns a matrix of extrinsic values Γ_{W-1}: the updated Λ_{W-1} is computed as $\Lambda_{W-1} = \Gamma_{W-1} + Y_{W-1}$, where Y_{W-1} stores the LLRs originally received from the channel after the transmission of X_{W-1}. Λ_W is not updated, as simulation have shown that backpropagating information is damaging, an effect exacerbated by the LLR propagation. In the next decoding step, Λ_{W-1} and Λ_{W-2} have to be considered: the LLRs in Γ_{W-1} already refer to X'_{W-1} save for an interleaving function Π_{CtoR}, and thus enable row decoding of Λ'_{W-1} and Λ_{W-2}. The process is repeated alternating row and column decoding, until Λ'_1 is decoded with Λ_0: the output of the decoding iteration is then the estimated $\hat{U}_1$. The decoding process is shown in Fig. 2.

2.1. *Component Decoding*

Each row (column) component code is a length-N polar code with frozen set $\mathscr{F}$. To be able to obtain soft values out of an inherently soft-in hard-out algorithm, the following is applied. Let us consider the successive-cancellation-list decoding algorithm (SCL) [11], and in particular its LLR-based formulation in [12], where to each of the L candidate paths is assigned a path metric PM, updated after each bit estimation $\hat{u}_i$ as

$$\text{PM}_i = \begin{cases} \text{PM}_{i-1} + |\delta_i|, & \text{if } \hat{u}_i \neq \text{HD}(\alpha_i), \\ \text{PM}_{i-1}, & \text{otherwise,} \end{cases} \tag{1}$$

where δ_i is the LLR associated to $\hat{u}_i$, $\text{PM}_0 = 0$, and $\text{HD}(\alpha_i) = 0$ if $\alpha_i \geq 0$, and 1 otherwise. At the end of the SCL decoding, we take the L estimated input vectors $\hat{u}^0, \ldots, \hat{u}^{L-1}$, having path metrics $M_0, \ldots, M_{L-1}$, and re-encode them obtaining the estimated codewords $\hat{x}^0, \ldots, \hat{x}^{L-1}$. Extrinsic soft information γ_i associated to codeword bit $\hat{x}_i$ is then calculated as (2), where λ_i is the i^{th} LLR input to the decoder, and α_E and α_B are scaling factors ≤ 1. In case the L codewords have the same value for a given bit $\hat{x}_i$, i.e. $\{l = 0, \ldots, L-1 \; s.t. \; \hat{x}_i^l = a\} = \emptyset$, γ_i is computed as (3), where $\hat{\lambda}$ is vector λ sorted in ascending order of magnitude, and k_{min} and k_{max} are two indices devised via simulation. The sign of γ_i is inferred according to $\hat{x}_i$.

$$\gamma_i = \alpha_E \left(\alpha_B \left(\min_{\hat{x}_i^l = 1}(M_l) - \min_{\hat{x}_i^l = 0}(M_l) \right) - \lambda_i \right) \tag{2}$$

$$|\gamma_i| = \alpha_E \left(\alpha_B \left(\sum_{j=k_{min}}^{k_{max}} |\hat{\lambda}|_j \right) - \lambda_i \right) \tag{3}$$

3. Simulation results

Figure 3 and 4 plot the bit error rate (BER) and block error rate (BLER) of the proposed construction, simulated over an additive white Gaussian noise channel. We consider component codes of length $N = \{128, 256\}$, rate $R = \{5/6, 7/8\}$, and window size $W = \{2, 5, 10\}$. The frozen set of the component codes considers a length-N polar code reliability vector and selects the $N/2 - K$ least reliable positions among the first $N/2$. An alternative approach is to directly consider a reliability vector of length-$N/2$, leading to similar error-correction performance with short-to-medium polar codes. Within the decoding process detailed in Section 2, the component codes have been decoded with SCL, $L = 8$. To compute the BLER, every $N/2 \times N/2$ matrix $\hat{U}$ that contains at least a bit in error constitutes an erroneous block. All codes have been decoded with the α_E and α_B that minimize the E_b/N_0 at BER=10^{-7}, obtained via simulation.

The performance of all codes improves of more than 3.5dB when increasing the W from 2 to 5; the larger decoding window allows to correct a higher number of errors, and the waterfall region of the curves starts at lower E_b/N_0. However, within these decoding conditions, increasing W from 5 to 10 does not improve the error correction performance, due to the very local nature of α_E and α_B, that need to be varied across the decoding window as W increases. In case of $W = 5, 10$, for $N = 256$, the lower rate component code yields a 1dB gain with respect to the higher rate at low BER/BLER, as the slope in case of $R = 5/6$ is substantially steeper. On the other hand, the BER and BLER with $N = 128$ for the two rates tend to converge, leading to a smaller gain. Comparing constructions with component codes with the same rate, we can see that increasing the code length when $R = 5/6$ results in 1dB gain at BER=10^{-7}, while 0.5dB are observed for $R = 7/8$. Finally, the brown BER curve in Fig. 4 matches that obtained with the same code parameters in Fig. 4 of [9], but with $L = 8$ instead of $L = 32$, thus requiring substantially lower decoding complexity.

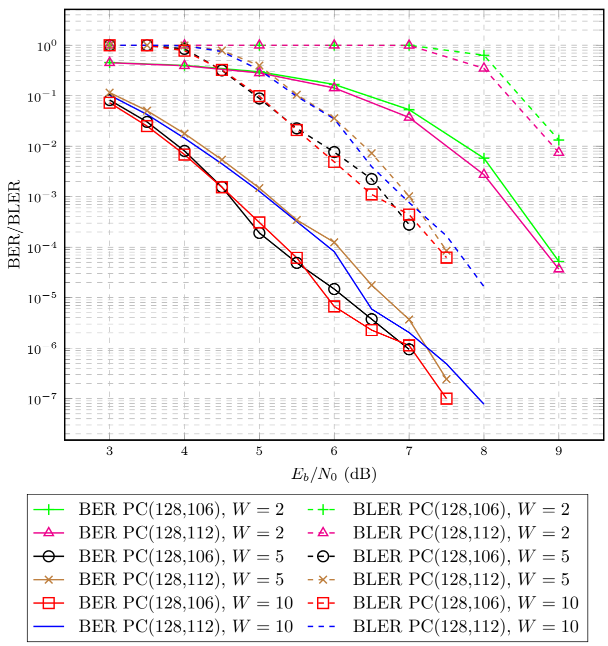

Fig. 3: BER/BLER for $N = 128$, $L = 8$.

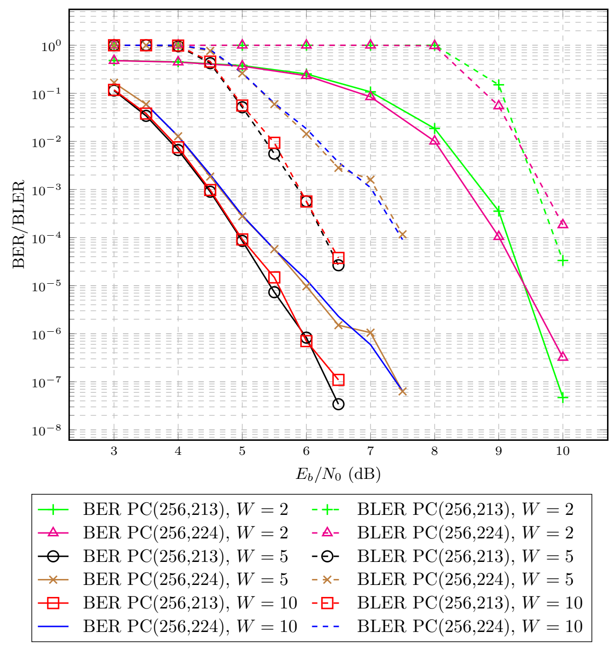

Fig. 4: BER/BLER for $N = 256$, $L = 8$.

4. Conclusions

In this paper, we proposed a staircase construction based on non-systematic polar codes, together with an encoding and decoding framework that match existing solutions with a lower decoding complexity. While we define a specific interleaver as an example, any interleaver with one block of memory can be used straightforwardly.

References

1. B. P. Smith, A. Farhood, A. Hunt, F. R. Kschischang, and J. Lodge, "Staircase codes: FEC for 100 Gb/s OTN," Journal of Lightwave Technology **30**, 110–117 (2012).
2. P. Elias, "Error-free coding," Transactions of the IRE Professional Group on Information Theory **4**, 29–37 (1954).
3. R. Gallager, "Low-density parity-check codes," IRE Transactions on Information Theory **8**, 21–28 (1962).
4. E. Arıkan, "Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels," IEEE Transactions on Information Theory **55**, 3051–3073 (2009).
5. D. Wu, A. Liu, Y. Zhang, and Q. Zhang, "Parallel concatenated systematic polar codes," in "Electronics Letters," , vol. 52 (2015), pp. 43–45.
6. T. Koike-Akino, C. Cao, Y. Wang, K. Kojima, D. S. Millar, and K. Parsons, "Irregular polar turbo product coding for high-throughput optical interface," in "Optical Fiber Communication Conference and Exhibition (OFC)," (San Diego, CA, USA, 2018), p. March.
7. V. Bioglio, C. Condo, and I. Land, "Construction and decoding of product codes with non-systematic polar codes," in "IEEE Wireless Communications and Networking Conference (WCNC)," (Marrakech, Morocco, 2019), arXiv:1901.06892 .
8. B. Feng, J. Jiao, L. Zhou, S. Wu, B. Cao, and Q. Zhang, "A Novel High-Rate Polar-Staircase Coding Scheme," arXiv:1805.10082 (2018).
9. L. Zhou, B. Feng, J. Jiao, K. Liang, S. Wu, and Q. Zhang, "Performance analysis of soft decoding algorithms for polar-staircase coding scheme," in "2018 10th International Conference on Wireless Communications and Signal Processing (WCSP)," (2018), pp. 1–6.
10. C. Leroux, A. Raymond, G. Sarkis, and W. Gross, "A semi-parallel successive-cancellation decoder for polar codes," IEEE Transactions on Signal Processing **61**, 289–299 (2013).
11. I. Tal and A. Vardy, "List decoding of polar codes," IEEE Transactions on Information Theory **61**, 2213–2226 (2015).
12. A. Balatsoukas-Stimming, M. Bastani Parizi, and A. Burg, "LLR-based successive cancellation list decoding of polar codes," IEEE Transactions on Signal Processing **63**, 5165–5179 (2015).