

# On Parameter Optimization for Staircase Codes

Christian Häger<sup>(1)</sup>, Alexandre Graell i Amat<sup>(1)</sup>, Henry D. Pfister<sup>(2)</sup>,  
Alex Alvarado<sup>(3)</sup>, Fredrik Brännström<sup>(1)</sup>, and Erik Agrell<sup>(1)</sup>

<sup>(1)</sup>Department of Signals and Systems, Chalmers University of Technology, 41296 Göteborg, Sweden,

<sup>(2)</sup>Department of Electrical Engineering, Duke University, Durham, NC 27708,

<sup>(3)</sup>Optical Networks Group, University College London (UCL), Torrington Place, London, WC1E 7JE, UK  
christian.haeger@chalmers.se

**Abstract:** We discuss the optimization of staircase code parameters based on density evolution. An extension of the original code construction is proposed, leading to codes with steeper waterfall performance.

**OCIS codes:** (060.0060) Fiber optics and optical communications, (060.2330) Fiber optics communications.

## 1. Introduction

In [1], Smith *et al.* proposed a new class of forward error correction (FEC) codes for optical transport networks, named staircase codes, which offer a 0.42 dB net coding gain (NCG) improvement over the best code from the ITU-T G.975.1 recommendation, while at the same time maintaining a low implementation complexity [1]. Staircase codes are a class of generalized low-density parity-check (GLDPC) codes where all variable nodes in the underlying Tanner graph have degree two and the check nodes correspond to Bose–Chaudhuri–Hocquenghem (BCH) codes. The code construction described in [1, Sec. IV] is specified in terms of the BCH code parameters  $(v, t, s)$ , where  $v$  is the Galois field extension degree,  $t$  the error-correction capability, and  $s$  the shortening parameter. In [2], these parameters were optimized assuming FEC overheads (OHs) ranging from 6.25% to 33.33%. Given a predefined parameter space, the optimization is based on a brute-force search using extensive software simulations to predict the code performance. Furthermore, in order to reduce the optimization complexity, a simplified iterative hard-decision decoding (HDD) is assumed. This iterative HDD does not account for miscorrections in the bounded-distance decoding of the BCH codes.

In this paper, we discuss a different approach to finding good staircase code parameters for a fixed OH. We observe that, for a certain choice of parameters, staircase codes are contained in the spatially-coupled GLDPC (SC-GLDPC) code *ensemble* proposed in [3], whose asymptotic behavior can be analyzed using density evolution (DE). In particular, we use decoding thresholds (which are defined further below) as the optimization criterion. This approach offers a significantly reduced optimization time (or the possibility to explore a larger parameter space) and can rigorously account for miscorrections. The latter is obviously also possible with simulations, i.e., by performing bounded-distance decoding of the BCH codes, albeit at the cost of an increased optimization time.

A staircase code with parameters optimized with the above approach does not guarantee to provide the best performance for a given pre-FEC bit error rate (BER). This is because the performance of staircase codes is not necessarily well predicted by the asymptotic DE in [3], in particular for small staircase block sizes. To overcome this problem, we propose an extension of staircase codes with larger staircase block sizes by allowing for multiple code constraints per row/column in the staircase array. It is shown that the proposed codes have steeper waterfall performance curves compared to those obtained from the original construction and match the DE prediction more closely.

## 2. Staircase Codes

Let  $\mathcal{C}$  be a shortened primitive BCH code with (even) length  $n = 2^v - 1 - s$  and dimension  $k = 2^v - vt - 1 - s$ . A staircase code with rate  $R = 2k/n - 1$  (and  $\text{OH} = 1/R - 1$ ) is defined as the set of all matrix sequences  $\mathbf{B}_i \in \{0, 1\}^{a \times a}$ ,  $i = 0, 1, 2, \dots$ , such that the rows in  $[\mathbf{B}_{i-1}^T \mathbf{B}_i]$  for all  $i \geq 1$  form valid codewords of  $\mathcal{C}$ , where  $a = n/2$  and  $\mathbf{B}_0$  is initialized to the all-zero matrix. Similar to classical product codes, codewords in a staircase code are naturally represented as two-dimensional arrays and  $a$  is the size of one block in the characteristic staircase array shown in [1, Fig. 4].

Motivated by the work on block-wise braided block codes [4], a SC-GLDPC code ensemble was proposed in [3]. This code ensemble is specified in terms of the parameters  $(\mathcal{C}, m, L, w)$ , where  $m$  is the number of constraint nodes per spatial position,  $L$  the total number of spatial positions, and  $w$  the coupling width. Staircase codes are closely related to block-wise braided block codes [5] and both code classes can be seen as particular codes in the ensemble described in [3, Def. 2] with a proper interleaver choice. In particular, it can be shown that the staircase code for a given  $\mathcal{C}$  is contained in the ensemble for  $m = n/2$ ,  $L \rightarrow \infty$ , and  $w = 2$ .

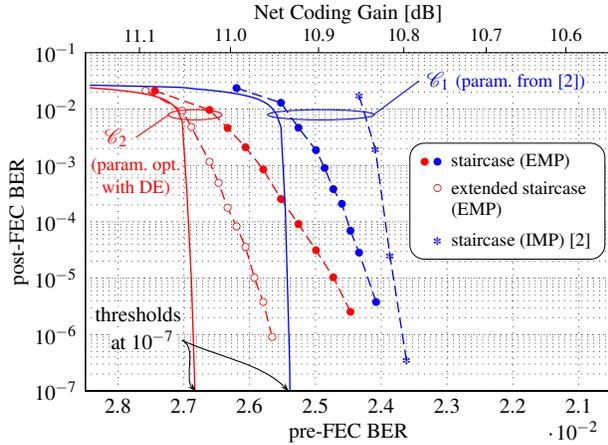


Fig. 1. DE (solid) and simulation (dashed) results

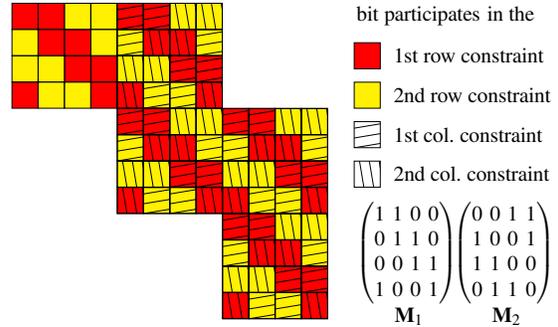


Fig. 2. Modified staircase array

### 2.1. Iterative Hard-Decision Decoding

The decoding of staircase codes described in [1, Sec. IV-A] is performed in a sliding-window fashion by iterating between the BCH decoders for all rows and columns in the staircase array consisting of  $W$  received blocks for a maximum of  $l$  iterations. This decoding scheme can be seen as an iterative message-passing algorithm with “hard” (i.e., binary) messages in the Tanner graph describing the code. However, it is pointed out in [3] that the message-passing rule associated with the decoding algorithm in [1] violates the principle that only extrinsic messages should be exchanged during iterative decoding. Hence, the authors in [3] refer to this decoding as iterative HDD with *intrinsic* message passing (IMP). In this paper, we employ the iterative HDD with *extrinsic* message passing (EMP) proposed in [3], and adapted to operate in a sliding-window fashion for staircase codes. EMP can provide better performance compared to IMP [5], and its performance can be analyzed using DE, even in the event of miscorrection [3]. For more details about this algorithm and the differences with respect to the decoding in [1], we refer the reader to [3] and [5, Algorithm 1].

### 2.2. Density Evolution

DE is a tool to predict the iterative decoding performance of asymptotically long codes in the waterfall region. DE mimics the decoding process under a cycle-free graph assumption by tracking how the probability densities of the exchanged messages evolve with iterations. In [3], a DE analysis is presented for the  $(\mathcal{C}, m, L, w)$  SC-GLDPC code ensemble with iterative HDD and EMP, assuming that  $m \rightarrow \infty$ .

*Example:* Let  $\mathcal{C}_1$  be a BCH code with parameters  $(v, t, s) = (9, 5, 151)$  [2, Table II]. For the  $(\mathcal{C}_1, \infty, 30, 2)$  SC-GLDPC code ensemble, if we adapt the DE analysis in [3, eq. (9)] to a sliding-window decoder with  $W = 7$  and  $l = 8$ , one obtains the solid blue curve in Fig. 1. The simulated performance (assuming a binary symmetric channel (BSC)) of the staircase code with  $\mathcal{C}_1$  (which leads to a block size of  $a = 180$  and  $\text{OH} = 33.33\%$ ) using EMP is shown in Fig. 1 by the dashed blue curve with dots. The performance using IMP is shown by the dashed blue curve with stars, where the data is taken from [2, Fig. 2].  $\triangle$

It can be observed that DE accurately predicts the pre-FEC BER region where the staircase code performance curve “bends” into the characteristic waterfall behavior. This motivates the use of DE to find good staircase code parameters. In particular, we use the decoding threshold (for a finite number of decoding iterations), which is defined as the pre-FEC BER value where the DE curve crosses a certain target post-FEC BER. Decoding thresholds can be computed numerically using a bisection search over a given pre-FEC BER range. The decoding threshold for the code ensemble in Example 1 at a post-FEC BER of  $10^{-7}$  is approximately given by  $2.54 \cdot 10^{-2}$  and also indicated in Fig. 1.

It can also be observed that the actual performance of the staircase code deviates quite significantly from the DE prediction once the pre-FEC BER is smaller than the decoding threshold. In this context, it is important to point out that the DE analysis applies to the SC-GLDPC ensemble for  $m \rightarrow \infty$ , whereas staircase codes belong to the ensemble where  $m$  is fixed to half the length of the BCH code. For a random code taken from the ensemble, the parameter  $m$  determines the steepness of the BER curve. Similar observations can be made also for staircase codes, where the steepness of the BER curves in the waterfall region is determined by the block size  $a$ , see [2, Fig. 2].

## 3. Extended Staircase Codes

By allowing for  $q > 1$  code constraints in each row and column of the staircase array, one can increase the block size of staircase codes to  $\tilde{a} = qa$ . In order to indicate which bit participates in the  $j$ -th code constraint of a particular row or

Table 1. Optimized staircase code parameters based on thresholds (top) and taken from [2] (bottom)

OH [%]	6.25	6.67	7.14	7.69	8.33	9.09	10.00	11.11	12.50	14.29	16.67	20.00	25.00	33.33
$v$	10	10	10	10	10	10	10	10	9	9	9	9	8	8
	11	11	11	11	N/A	10	10	10	10	10	9	9	9	9
$t$	3	3	3	3	3	4	4	5	3	3	4	4	3	3
	4	5	5	5	N/A	4	4	4	4	4	4	4	5	5
$s$	3	63	123	183	243	63	143	23	25	79	7	79	15	63
	551	287	397	507	N/A	63	143	223	303	383	7	79	61	151
$a$	510	480	450	420	390	480	440	500	243	216	252	216	120	96
	748	880	825	770	N/A	480	440	400	360	320	252	216	225	180
NCG [dB]	9.50	9.56	9.62	9.69	9.76	9.89	9.97	10.08	10.21	10.32	10.52	10.64	10.88	11.03
	9.44	9.54	9.60	9.66	N/A	9.82	9.87	9.96	10.05	10.13	10.32	10.41	10.62	10.70
Gap [dB]	0.39	0.41	0.43	0.45	0.48	0.45	0.48	0.51	0.53	0.58	0.58	0.69	0.74	0.95
	0.45	0.43	0.45	0.48	N/A	0.52	0.59	0.63	0.68	0.77	0.78	0.92	1.00	1.28

column, one may use  $q$  masking matrices  $\mathbf{M}_j \in \{0, 1\}^{\tilde{a} \times \tilde{a}}$ ,  $j \in \{1, 2, \dots, q\}$ , with  $a$  ones in each row such that  $\sum_{j=1}^q \mathbf{M}_j$  is the all-one matrix. We informally write  $\mathbf{A} \cap \mathbf{M}_j$  to denote the  $\tilde{a} \times a$  matrix containing the elements from the  $\tilde{a} \times \tilde{a}$  matrix  $\mathbf{A}$  where the corresponding entry in  $\mathbf{M}_j$  is equal to one. The extended staircase code with larger block size is defined as the set of all matrix sequences  $\mathbf{B}_i \in \{0, 1\}^{\tilde{a} \times \tilde{a}}$ ,  $i = 0, 1, 2, \dots$ , such that the rows in  $[\mathbf{B}_{i-1}^T \cap \mathbf{M}_j \ \mathbf{B}_i \cap \mathbf{M}_j]$  for all  $i > 0$  and  $j$  form valid codewords in  $\mathcal{C}$ , where  $\mathbf{B}_0$  is the all-zero matrix. As an example, the staircase array for  $n = 4$  and  $q = 2$  is shown in Fig. 2. We remark that this extension is itself a special case of a more general technique, where the Tanner graph of the staircase code can be interpreted as a protograph which is lifted in order to obtain a larger graph [6]. Our particular choice for the lifting factor (i.e.,  $q^2$ ) and type of lifting preserves many properties of the original staircase codes, e.g., the staircase array structure and time-invariant encoding/decoding operations.

#### 4. Results and Discussion

We consider the same parameter space as in [2], i.e., the product set of  $\text{OH} \in \{1/i : i = 3, 4, \dots, 16\}$ ,  $v \in \{8, 9, 10, 11, 12\}$ , and  $t \in \{2, 3, 4, 5, 6\}$ . In Table 1, we list the parameters that give the best decoding threshold at a BER of  $10^{-15}$  for the SC-GLDPC ensemble with  $L = 30$  and  $w = 2$  according to the DE analysis accounting for miscorrection [3, eq. (9)] and adapted to a sliding-window decoder with  $W = 7$  and  $l = 8$ . We also show the NCG at a BER of  $10^{-15}$  using the obtained decoding thresholds and the corresponding gap to the maximum NCG for the BSC. To allow for a direct comparison, we repeat the values that are reported in [2], where results for  $\text{OH} = 8.33\%$  are unfortunately not available. We would like to stress that in [2], software simulations are extrapolated down to  $10^{-15}$  to estimate the NCG and gap. Hence, those values are not directly comparable to the ones we obtain by using decoding thresholds. Rather, the threshold results should be seen as a way to identify further room for improvement, in particular using the extended staircase code construction. As an example, for  $\text{OH} = 33.33\%$ , we simulated the staircase code based on the BCH code  $\mathcal{C}_2$  with  $(v, t, s) = (8, 3, 63)$  together with the extended staircase code for  $q = 2$ . The results are shown in Fig. 1 by the dashed red lines with dots and circles, respectively. The extended staircase code has a steeper waterfall performance at the expense of a larger block size of  $\tilde{a} = 192$ . The block size is, however, comparable to that of the staircase code with  $\mathcal{C}_1$  and, extrapolating the simulation results to  $10^{-15}$ , we expect a gain of approximately 0.11 dB. If larger block sizes can be tolerated, further improvements (up to around 0.33 dB) can be expected by increasing  $q$ .

#### 5. Conclusions

We have shown that the DE analysis in [3] can be used as an effective tool for finding good staircase code parameters. Compared to the simulation-based method in [2], this approach is less complex (e.g., the parameter space in [2] can be searched within seconds) and can account for miscorrections assuming iterative HDD with EMP. An extended staircase code construction was proposed with steeper waterfall performance at the expense of a larger staircase block size.

#### References

1. B. P. Smith, et. al., "Staircase codes: FEC for 100 Gb/s OTN," J. Lightw. Technol. **30**, 110–117 (2012).
2. L. M. Zhang and F. R. Kschischang, "Staircase codes with 6% to 33% overhead," J. Lightw. Technol. **32**, 1999–2002 (2014).
3. Y.-Y. Jian, et. al., "Approaching capacity at high rates with iterative hard-decision decoding," in "Proc. IEEE Int. Symp. Information Theory (ISIT)," Cambridge, MA, 2012.
4. A. J. Feltström, et. al., "Braided block codes," IEEE Trans. Inf. Theory **55**, 2640–2658 (2009).
5. Y.-Y. Jian, et. al., "Iterative hard-decision decoding of braided BCH codes for high-speed optical communication," in "Proc. IEEE Glob. Communication Conf. (GLOBECOM)," Atlanta, GA, 2014.
6. J. Thorpe, "Low-density parity-check (LDPC) codes constructed from protographs," IPN Progress Report 42-154, JPL (2005).