Terminated and Tailbiting Spatially-Coupled Codes with Optimized Bit Mappings for Spectrally Efficient Fiber-Optical Systems

Christian Häger, *Student Member, IEEE*, Alexandre Graell i Amat, *Senior Member, IEEE*, Fredrik Brännström, *Member, IEEE*, Alex Alvarado, *Member, IEEE*, and Erik Agrell, *Senior Member, IEEE*

Abstract—We study the design of spectrally efficient fiber-optical communication systems based on different spatially-coupled (SC) forward error correction (FEC) schemes. In particular, we optimize the allocation of the coded bits from the FEC encoder to the modulation bits of the signal constellation. Two SC code classes are considered. The codes in the first class are protograph-based low-density parity-check (LDPC) codes which are decoded using iterative soft-decision decoding. The codes in the second class are generalized LDPC codes which are decoded using iterative hard-decision decoding. For both code classes, the bit allocation is optimized for the terminated and tailbiting SC cases based on a density evolution analysis. An optimized bit allocation can significantly improve the performance of tailbiting SC codes codes over the baseline sequential allocation, up to the point where they have a comparable gap to capacity as their terminated counterparts, at a lower FEC overhead. For the considered terminated SC codes, the optimization only results in marginal performance improvements, suggesting that in this case a sequential allocation is close to optimal.

Index Terms—Bit mapper optimization, coded modulation, forward error correction, LDPC codes, hard-decision decoding, soft-decision decoding, spatial coupling.

I. INTRODUCTION

Designing spectrally efficient fiber-optical systems that can operate close to the capacity limits [1] has become an important research topic [2]–[4]. Such systems are often implemented according to the pragmatic bit-interleaved coded modulation (BICM) paradigm [5], where a single binary forward error correction (FEC) encoder is used in combination with a nonbinary signal constellation. A random allocation (or interleaving) [5] of the coded bits from the FEC encoder to the modulation bits of the signal constellation is commonly assumed. In this paper, we optimize the allocation to the modulation bits for a coherent long-haul polarization-multiplexed (PM) fiber-optical system. In particular, we consider different

C. Häger, A. Graell i Amat, F. Brännström, and E. Agrell are with the Department of Signals and Systems, Chalmers University of Technology, Gothenburg, Sweden (emails: {christian.haeger, alexandre.graell, fredrik.brannstrom, agrell}@chalmers.se).

A. Alvarado is with the Optical Networks Group, Department of Electronic and Electrical Engineering, University College London, London WC1E7JE, UK (email: alex.alvarado@ieee.org).

spatially-coupled (SC) FEC schemes both with soft-decision decoding (SDD) and hard-decision decoding (HDD).

SC low-density parity-check (SC-LDPC) codes have attracted a great deal of attention in the recent years. They are considered as viable candidates for future spectrally efficient fiber-optical systems [3], [6], [7] due to their capacity-achieving performance for many communication channels [8]. SC-LDPC codes promise excellent belief propagation (BP) performance with a quasi-regular node degree distribution and low node degrees. The BP performance of SC-LDPC codes can further be improved by increasing the node degrees, whereas the decoding performance for regular LDPC codes generally worsens if the node degrees are increased [8]. While irregular LDPC codes can also perform close to capacity [9], the optimal degree distribution depends on the code rate and/or the channel [10]. High node degrees are also often required for good performance which leads to a high decoding complexity.

We consider two different SC code classes taken from the literature. The codes in the first class are protograph-based SC-LDPC codes [11], [12] which are decoded using iterative SDD in the form of BP decoding. BP is a message passing algorithm in which "soft" (i.e., real-valued) messages are exchanged between the variable nodes (VNs) and constraint nodes (CNs) in the Tanner graph representing the code. The codes in the second class are SC generalized LDPC (SC-GLDPC) codes where each coded bit is protected by two t-error correcting Bose–Chaudhuri–Hocquenghem (BCH) component codes [13]. These codes are decoded using iterative HDD with bounded-distance decoding (BDD) of the component BCH codes. Iterative HDD can be seen as a message passing algorithm with "hard" (i.e., binary) messages in the Tanner graph representing the GLDPC code and is significantly less complex than SDD [14].

The adoption of SDD is considered one of the most important factors for increasing the performance of fiber-optical systems [15]. However, SDD poses implementation challenges at very high data rates motivating the use of less complex FEC schemes [14]. The SC-GLDPC codes we consider in this paper were proposed in [13], where it is shown that they can approach the capacity of the binary symmetric channel (BSC) under iterative HDD for high code rates (i.e., low FEC overheads (OHs)). We use these codes because a density evolution (DE) analysis is readily available in [13]. This allows us to apply the optimization techniques for protograph-based SC-LDPC codes we previously presented in [16] to the
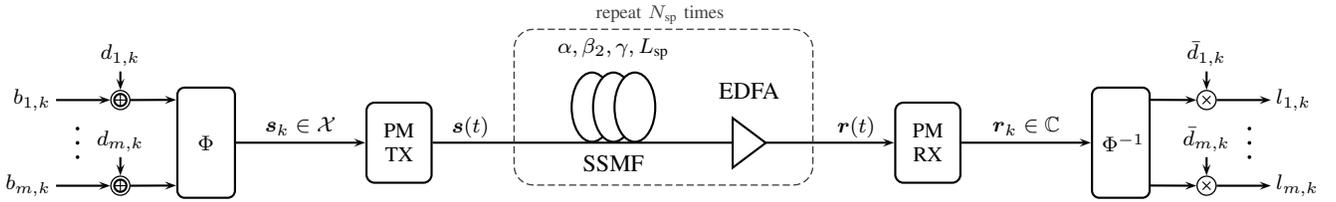
Fig. 1. Block diagram of the considered PM transmission system.

practically relevant case of SC-GLDPC codes with iterative HDD. The SC-GLDPC code ensemble in [13] is closely related to other recently proposed FEC schemes for optical transport networks, such as staircase codes [14] (which are themselves related to block-wise braided block codes [17]), and the modified construction of tightly-braided block codes proposed in [18]. For other related works on GLDPC codes for fiber-optical communications, we refer the interested reader to [19], [20] and references therein.

The outstanding performance of SC codes is due to a termination boundary effect which initiates a wave-like decoding behavior [8]. This behavior of terminated SC codes comes at the price of a rate loss, i.e., a larger FEC OH, compared to the underlying uncoupled codes. So-called tailbiting SC codes provide an interesting solution to this problem, since they do not suffer from an increased OH. However, by default, a tailbiting SC code behaves essentially the same as the underlying uncoupled code due to the absence of a termination boundary. The main aim of this paper is to demonstrate that the unequal error protection offered by the modulation bits of a nonbinary signal constellation can be exploited to create an artificial termination boundary. This significantly improves the performance of tailbiting SC codes, both in the case of SDD and HDD. With an optimized bit allocation, the capacity gap of the considered tailbiting SC codes is comparable to the gap of their terminated counterparts, at a lower FEC OH. For the considered terminated SC codes, the performance gain due to an optimized bit allocation is limited, in particular for the SC-GLDPC codes with HDD. Simulation results for both linear and nonlinear transmission scenarios confirm the DE analysis.

The remainder of the paper is organized as follows. In Section II, the assumed PM transmission system is described. The two SC FEC schemes are covered in Sections III and IV, where we explain the code construction, the decoding algorithms, and the DE analysis with the help of several examples. In Section V, we briefly review the optimization techniques for the bit allocation described in [16], which apply to the considered SC-LDPC codes with SDD. We also discuss how they are easily extended to the SC-GLDPC codes with HDD. Results are presented and discussed in Section VI and the paper is concluded in Section VII.

II. System Model

A block diagram of the considered PM fiber-optical transmission system is shown in Fig. 1. At each discrete time instant k, the modulator Φ takes m bits $b_{i,k}$, $i = 1, \ldots, m$, and maps them to a symbol $\boldsymbol{s}_k = (s_{\mathsf{x},k}, s_{\mathsf{y},k})$ taken with uniform probabilities from a signal constellation $\mathcal{X} \subset \mathbb{C}^2$ ($|\mathcal{X}| = 2^m$)

according to the binary labeling. The modulo-2 addition of the independent and uniformly distributed bits $d_{i,k}$ (and the multiplication by $\bar{d}_{i,k} = (-1)^{d_{i,k}}$ at the receiver) shown in Fig. 1 serves as a symmetrization technique [21].[1]

The baseband signal in polarization x is $s_{\mathsf{x}}(t) = \sum_k s_{\mathsf{x},k} p(t - k/R_s)$ with (real-valued) pulse shape $p(t)$ and symbol rate R_s (and similarly for polarization y). The transmit power $P = \lim_{T \to \infty} \int_{-T}^{T} s_{\mathsf{x}}(t)^2 \mathrm{d}t/(2T)$ is assumed to be equal in both polarizations. The PM signal $\boldsymbol{s}(t) = (s_{\mathsf{x}}(t), s_{\mathsf{y}}(t))$ is launched into the fiber and propagates according to the Manakov equation [23]. The optical link consists of N_{sp} spans of standard single-mode fiber (SSMF) with attenuation coefficient α, group velocity dispersion β_2, nonlinear Kerr parameter γ, length L_{sp}, and a lumped amplification scheme (no optical dispersion compensation is assumed). Each erbium-doped fiber amplifier (EDFA) introduces circularly symmetric complex Gaussian noise with two-sided power spectral density $N_{\mathrm{EDFA}} = (e^{\alpha L_{\mathrm{sp}}} - 1)h\nu_s n_{\mathrm{sp}}$ per polarization [1], where h is Planck's constant, ν_s the carrier frequency, and n_{sp} the spontaneous emission factor. A coherent linear receiver according to $r_{\mathsf{x},k} = r_{\mathsf{x}}(t) * h(t) * p(-t)|_{t=k/R_s}$ is used in each polarization, where $*$ denotes convolution and $h(t)$ is the impulse response of an equalizer which accounts for linear distortions due to chromatic dispersion. The frequency response of the equalizer is given by $H(f) = \exp(j2\beta_2\pi^2 f^2 N_{\mathrm{sp}} L_{\mathrm{sp}})$. The received symbles are denoted by $\boldsymbol{r}_k = (r_{\mathsf{x},k}, r_{\mathsf{y},k})$.

Two different demodulators Φ^{-1} are considered. For SDD, the demodulator computes "soft" reliability information about the bits $b_{i,k}$ in the form of log-likelihood ratios (LLRs) $l_{i,k}$. For HDD, the demodulator performs a minimum distance symbol-by-symbol detection of the received symbols with respect to the signal constellation $\mathcal{X}$ and outputs the binary labeling associated with the detected symbol. Both demodulators are based on the assumption that the discrete-time channel from $\boldsymbol{s}_k$ to $\boldsymbol{r}_k$ is the additive white Gaussian noise (AWGN) channel with signal-to-noise ratio (SNR) denoted by ρ. This assumption is accurate for linear transmission (i.e., $\gamma = 0$) where $\rho = P/(N_{\mathrm{sp}} N_{\mathrm{EDFA}} R_s)$. For the considered setup without optical inline dispersion compensation, it has been shown that this assumption is also justified, provided that dispersive effects are dominant and nonlinear effects are weak [24], [25]. For this case, ρ can be computed using [24, eq. (15)] assuming single channel transmission. Under the Gaussian Noise model assumption, see [25] and references therein, similar expressions for the SNR are also computable

---

[1]The symmetrization makes the bit error probability independent of the transmitted bits. This is an important requirement for the all-zero codeword assumption which is commonly made in DE [22, p. 389].
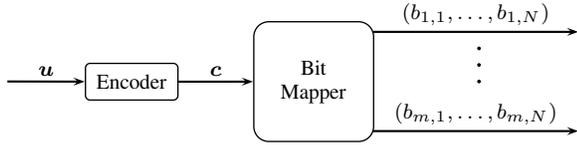
Fig. 2. Illustration of the Bit Mapper.

for wavelength-division multiplexing systems.

We consider a system according to the BICM paradigm, where a binary code $\mathcal{C} \subset \{0,1\}^{n_{\mathcal{C}}}$ of length $n_{\mathcal{C}}$ and dimension $k_{\mathcal{C}}$ is employed and each codeword $\boldsymbol{c} = (c_1, \ldots, c_{n_{\mathcal{C}}})$ is transmitted using $N = n_{\mathcal{C}}/m$ symbols $\boldsymbol{s}_k$. The allocation of the coded bits to the modulation bits is determined by a bit mapper[2] as shown in Fig. 2, where $\boldsymbol{u} = (u_1, \ldots, u_{k_{\mathcal{C}}})$ is the information word. The bit mapper optimization is discussed in Section V. The optimization is based on the AWGN channel model because a direct optimization using DE for the optical channel defined by the nonlinear Schrödinger equation is not feasible. The accuracy of this approach is verified through simulation results for a nonlinear transmission scenario in Section VI-D. To the best of our knowledge, there are no comparable works by other authors on bit mapper optimization for SC codes. Hence, as a baseline for a comparison, we use a sequential mapper according to $b_{i,k} = c_{(k-1)m+i}$ for $1 \le i \le m$, $1 \le k \le N$. For the considered codes, a sequential mapper has the same expected performance as a random mapper.

III. PROTOGRAPH-BASED SC-LDPC CODES

A. Code Construction

An LDPC code of length $n_{\mathcal{C}}$ and dimension $k_{\mathcal{C}}$ is defined via a sparse parity-check matrix $\mathbf{H} = [h_{i,j}] \in \{0,1\}^{r_{\mathcal{C}} \times n_{\mathcal{C}}}$, where $r_{\mathcal{C}} \ge n_{\mathcal{C}} - k_{\mathcal{C}}$ with equality if and only if $\mathbf{H}$ has full rank. One popular method to construct LDPC codes is by using protographs [11]. A protograph is a bipartite graph defined by an adjacency matrix $\mathbf{P} = [p_{i,j}] \in \mathbb{N}_0^{r'_{\mathcal{C}} \times n'_{\mathcal{C}}}$, called the base matrix, where $\mathbb{N}_0$ is the set of nonnegative integers. Given $\mathbf{P}$, the parity-check matrix $\mathbf{H}$ is obtained by replacing each entry $p_{i,j}$ in $\mathbf{P}$ with a random binary M-by-M matrix which contains $p_{i,j}$ ones in each row and column. This procedure is called lifting and $M \ge \max_{i,j} p_{i,j}$ is the so-called lifting factor. The design rate of the code is given by $R = 1 - r_{\mathcal{C}}/n_{\mathcal{C}} = 1 - r'_{\mathcal{C}}/n'_{\mathcal{C}}$, where $r_{\mathcal{C}} = r'_{\mathcal{C}} M$ and $n_{\mathcal{C}} = n'_{\mathcal{C}} M$.

SC-LDPC codes have parity-check matrices with a band-diagonal structure and can be constructed using protographs [12]. The base matrix of a (J, K) regular SC-LDPC code with spatial length T is constructed by specifying matrices $\mathbf{P}_i$, $0 \le i \le m_{\mathrm{s}}$, of dimension J' by K', where m_{s} is referred to as the memory. The matrices are such that $\sum_{i=0}^{m_{\mathrm{s}}} \mathbf{P}_i$ has column weight J and row weight K for all columns and rows. Given the matrices $\mathbf{P}_i$ and the spatial length T, one can construct $\mathbf{P}$ as shown in Fig. 3 (a) for the terminated case and in Fig. 3
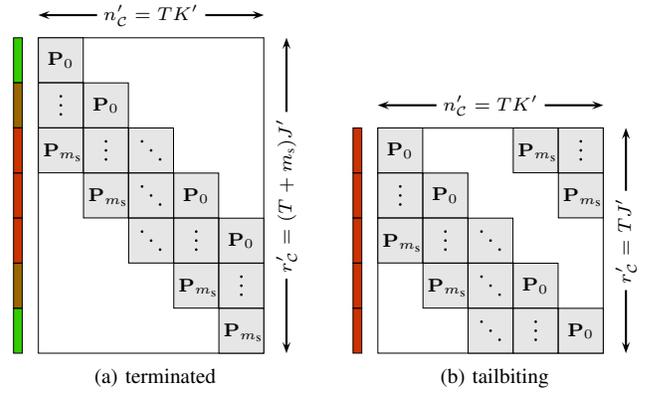

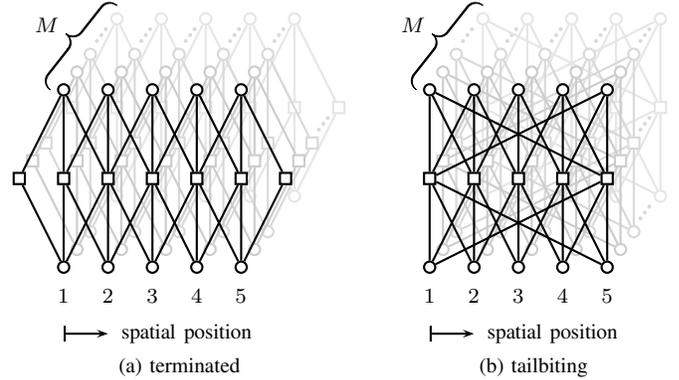Fig. 3. Base matrices $\mathbf{P}$ for protograph-based SC-LDPC codes.

Fig. 4. Protographs for the SC-LDPC code with $T = 5$ in Example 1. The first step of the lifting procedure to obtain the Tanner graph (i.e., copying the protograph M times) is indicated in light gray.

(b) for the tailbiting case.[3] Terminated and tailbiting SC-LDPC codes have design rates $R(T) = 1 - J'/K' - m_{\mathrm{s}}J'/(TK')$ and $R = 1 - J'/K'$, respectively [12]. The rate loss for the terminated code with respect to the tailbiting (or the underlying uncoupled regular) code can be made arbitrary small by letting $T \to \infty$, but this also leads to very long block lengths $n_{\mathcal{C}} = TK'M$ (assuming a fixed lifting factor M).

Example 1: Consider the $(3,6)$ regular SC-LDPC code with $\mathbf{P}_0 = \mathbf{P}_1 = \mathbf{P}_2 = (1,1)$, $T = 5$, $J' = 1$, $K' = 2$, and $m_{\mathrm{s}} = 2$. The two protographs corresponding to the terminated and tailbiting cases are shown in Fig. 4. The design rates are $R(5) = 0.3$ and $R = 0.5$, respectively. △

B. Soft-Decision Decoding and Density Evolution

The protograph-based SC-LDPC codes are decoded using the standard BP decoding [22, Sec. 5.4]. In order to alleviate the long decoding delay and high decoding complexity of SC-LDPC codes under full BP decoding, we employ the windowed decoder (WD) with a window size W developed in [26]. The WD reduces the decoding delay for terminated SC-LDPC codes from TMK' to WMK' coded bits [26]. For tailbiting SC-LDPC codes, additional memory for $(m_{\mathrm{s}} +$

---

[2]The bit mapper should not be confused with the modulator Φ, which is sometimes also referred to as a mapper. In the literature, the term "bit interleaver" is also frequently used instead of "bit mapper".

[3]The terminology originates from the trellis representation of convolutional codes, where the initial and final states are either determined by known bits (terminated) or forced to be identical (tailbiting).

$W-1)MK'$ values is required compared to terminated SC-LDPC codes, in order to take the circular wrap-around of the parity-check matrix into account. In particular, assume that the decoding starts when the channel observations corresponding to spatial positions 1 to $m_s + W$ are received and the first targeted symbols are at position $m_s + 1$. (Due to the circular structure, the last targeted symbols are at position m_s.) Then, the observations corresponding to the first m_s positions as well as the final LLRs for the bits at positions m_s+1 to m_s+W-1 have to be stored. We also point the interested reader to [27], where the decoding of tailbiting SC-LDPC codes based on a pipeline decoder architecture is discussed.

The main tool for the analysis of LDPC codes under BP decoding is DE [28]. DE mimics the decoding process under a cycle-free graph assumption by tracking how the densities of the LLRs evolve with iterations. Tracking the full densities (or quantized densities in practice) is computationally demanding and extrinsic information transfer (EXIT) functions [29] are usually considered to be a good compromise between computational efficiency and accuracy. For the protograph-based codes, we employ a modified protograph EXIT (P-EXIT) analysis [30] which accounts for the different protection levels of a nonbinary signal constellation and the WD, see [16, Algorithm 1].

Example 2: Consider the $(3,6)$ regular SC-LDPC code with $\mathbf{P}_0 = (2,2)$, $\mathbf{P}_1 = (1,1)$, $T = 20$, $J' = 1$, $K' = 2$, and $m_s = 1$, with rates $R(T) = 0.475$ and $R = 0.5$, respectively. This is a slightly different construction compared to the one in Example 1 and the resulting codes are better suited for the use of a WD, see [26, Design Rule 1]. Assume that transmission takes place using PM-QPSK in the linear regime and a WD with $W = 10$ and $l_{\max} = 7$ is used. In Fig. 5 (a), we show the predicted bit error rate (BER) obtained via the P-EXIT analysis (solid lines) together with the actual performance of randomly generated codes (dashed lines) for two different lifting factors $M = 2000$ (crosses) and $M = 10000$ (dots) for both the terminated (blue) and tailbiting (red) cases. Due to graph cycles, there is a mismatch between the actual performance and the DE prediction, in particular for the smaller lifting factor. However, the P-EXIT analysis accurately predicts the SNR region where the finite-length BER curves "bend" into their characteristic waterfall behavior. $\triangle$

In Fig. 5 (a), we also indicate the two points where the P-EXIT performance curves cross a BER of 10^{-5}. We refer to the SNR value of such a point as the decoding threshold ρ^* for a target BER $= 10^{-5}$ and a given finite number of decoding iterations. The thresholds can be numerically computed using a bisection search over a given SNR range. The thresholds are given by $\rho^* \approx 0.82$ dB and $\rho^* \approx 1.19$ dB for the terminated and tailbiting codes, respectively. The better decoding thresholds and finite-length performance of the terminated code can be explained by inspecting the structure of the base matrix Fig. 3 (a). One may verify that the CN degrees corresponding to the first and last couple of rows are lower than the CN degrees corresponding to the rows in between (see also Fig. 4 (a)). The lower degree CNs lead to a locally better decoding capability, which is visualized by the colored scale (green indicates a better correction capability),
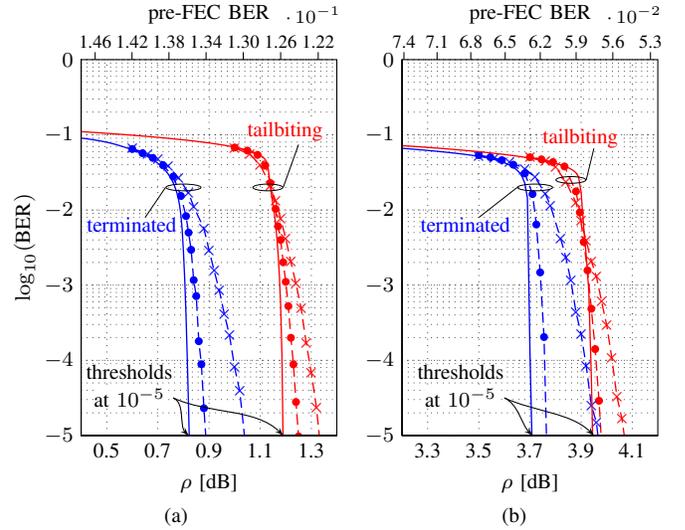

Fig. 5. Predicted (solid lines) and finite-length (dashed lines) performance for the codes in Examples 2 (left) and 6 (right). The codes have lengths 80 000 (a, crosses), 400 000 (a, dots), 168 000 (b, crosses), and 1 260 000 (b, dots).

at the expense of a rate loss. This termination boundary effect initiates the wave-like decoding behavior that is characteristic for terminated SC-LDPC codes [8]. On the other hand, for the tailbiting case, all CNs have the same degree J, hence no rate loss is incurred and all positions are protected equally. However, this also prevents the initiation of a decoding wave.

IV. SC-GLDPC CODES WITH BCH COMPONENT CODES

A. Code Construction

We consider the $(\mathcal{B}, m_c, T, w)$ SC-GLDPC code ensemble proposed in [13], where $\mathcal{B}$ is a binary linear code of length $n_\mathcal{B}$ and dimension $k_\mathcal{B}$ that can correct all error patterns of weight at most t, m_c is the number of CNs per spatial position, T is the spatial length, and w is the coupling size. In the following, we assume that $\mathcal{B}$ is a shortened primitive BCH code with parameters $(n_\mathcal{B}, k_\mathcal{B}) = (2^\nu - 1 - s, 2^\nu - \nu t - 1 - s)$, where ν is the Galois field extension degree and s is the number of shortened information bits. The code $\mathcal{B}$ defines the constraints that have to be satisfied by all CNs in the Tanner graph representing the SC-GLDPC code.

For completeness, we review the construction of the terminated case described in [13, Def. 2] and explain the necessary modifications for the tailbiting case. Assume that m_c CNs with degree $n_\mathcal{B}$ are placed at each of the spatial positions 1 to $T + w - 1$ and $m_c n_\mathcal{B}/2$ VNs of degree 2 are placed at each of the spatial positions 1 to T. Additionally, $m_c n_\mathcal{B}/2$ VNs initialized to a known value are placed at positions $j < 1$ and $j > T$ to terminate the code. The connections between CNs and VNs are as follows. The $m_c n_\mathcal{B}$ VN and CN sockets at each position are partitioned into w groups of equal size $m_c n_\mathcal{B}/w$ via a uniform random permutation. The i-th group at the j-th VN position and the i-th group at the j-th CN position are denoted by $\mathcal{S}_{j,i}^{(v)}$ and $\mathcal{S}_{j,i}^{(c)}$, respectively, where $i \in \{0, 1, \ldots, w - 1\}$. The Tanner graph of one particular code in the ensemble is constructed by using a uniform random permutation to connect $\mathcal{S}_{j,i}^{(v)}$ to $\mathcal{S}_{j+i,w-i-1}^{(c)}$ and mapping the
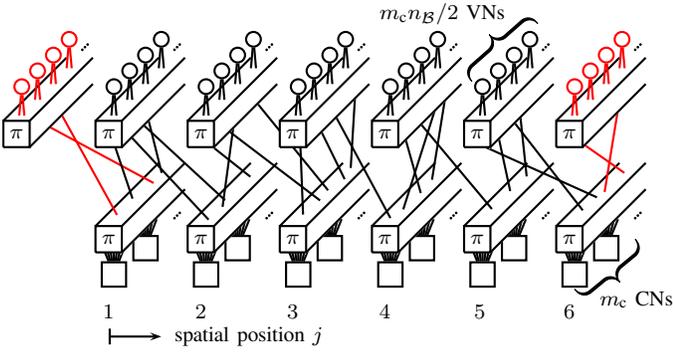
Fig. 6. Tanner graph for a code in the terminated SC-GLDPC ensemble with $T = 5$ and $w = 2$. Known VNs are shown in red.

$m_{\mathrm{c}} n_{\mathcal{B}}/w$ edges between the two groups. For the tailbiting case, the position index $j + i$ is interpreted modulo T and no known VNs are present.

Example 3: Consider the case where $T = 5$ and $w = 2$. The Tanner graph of a code in the terminated ensemble is shown in Fig. 6. The blocks π spread out the edges from the VNs and CNs according to the random permutations in the construction. A code from the tailbiting ensemble is obtained by removing the known VNs and the CNs at position 6, and connecting the lose edges to the CNs at position 1. △

The design rate for the terminated ensemble is lower bounded by [31, eq. (2.2)]

$$R'(T) \geq R' - (1 - R')\frac{w - 1}{T}, \qquad (1)$$

where $R' = 2k_{\mathcal{B}}/n_{\mathcal{B}} - 1$ is the design rate for the tailbiting ensemble. An exact expression for the design rate can be obtained by explicitly considering the possibility that certain CNs are connected exclusively to known VNs, similar to [8, Lemma 3]. However, for the high CN degrees and small coupling factors considered in this paper, one can safely ignore this possibility and we henceforth interpret (1) as an equality.

Example 4: Let $\mathcal{B}$ be a shortened BCH code with $\nu = 7$, $t = 3$, and $s = 43$, i.e., $\mathcal{B}$ has rate 0.75. For the terminated and tailbiting ensembles in Example 3, the design rates are given by $R'(T) = 0.4$ and $R' = 0.5$, respectively. △

Similar to the parity-check matrix of an LDPC code, a GLDPC code can be specified by an incidence matrix [22, p. 220]. The dimensions of the incidence matrix are $m_{\mathrm{c}}(T + w - 1) \times T m_{\mathrm{c}} n_{\mathcal{B}}/2$ and $m_{\mathrm{c}}T \times T m_{\mathrm{c}} n_{\mathcal{B}}/2$ for the terminated and tailbiting ensembles, respectively.

Example 5: Consider the case where $w = 2$ and $T = \infty$. Let $n_{\mathcal{B}}$ be even and $m_{\mathrm{c}} = n_{\mathcal{B}}/2$. If the edge permutations are such that the semi-infinite incidence matrix is the one shown in [32, p. 54], the code corresponds to a staircase code. In other words, staircase codes are contained in the terminated ensemble for a certain choice of parameters w, T, and m_{c}. △

B. Hard-Decision Decoding and Density Evolution

We use the iterative HDD algorithm based on extrinsic message passing of binary messages that is proposed in [13, Sec. II-A] (see also [18, Sec. II-C]). Assume transmission over a BSC with crossover probability p. All outgoing VN

messages are initialized to the channel observation. For each CN, the incoming messages from the VNs are collected in a candidate decoding vector, which is then decoded using BDD. The outgoing CN messages are computed based on the Hamming distance between the candidate vector and the decoded vector, cf. [18, Algorithm 1]. In the next iteration, the outgoing VN message on a particular edge corresponds to the incoming message on the other edge of that VN. Decoding continues for $l_{\max}$ iterations. The final decision for each VN is made based on the channel observation and the two incoming messages. If the two messages agree, the bit is set to the message value. If the messages disagree, the bit is set to the binary complement of the channel observation. As pointed out in [13], extrinsic message passing is different compared to the conventional approach of decoding product-like codes (referred to as intrinsic message passing in [13]) and can be rigorously analyzed via DE even in the event of miscorrection [13], i.e., when undetected errors remain after BDD.

We briefly summarize the DE analysis presented in [13]. Assume that the all-zero codeword is transmitted and let $q_j^{(l)}$ be the average probability that a message emitted by a VN at position j is in error (i.e., the message is "1") after the lth iteration. The DE recursion is given by [13, eq. (5)]

$$q_j^{(l)} = \frac{1}{w} \sum_{k=0}^{w-1} f_{n_{\mathcal{B}}}\left(\frac{1}{w} \sum_{k'=0}^{w-1} q_{j-k'+k}^{(l-1)}; p \right), \qquad (2)$$

with [13, eq. (2)]

$$f_{n_{\mathcal{B}}}(x; p) \triangleq \sum_{i=0}^{n_{\mathcal{B}}-1} \binom{n_{\mathcal{B}} - 1}{i} x^i (1 - x)^{n_{\mathcal{B}}-1-i} \\ \cdot (pP_{n_{\mathcal{B}}}(i) + (1 - p)Q_{n_{\mathcal{B}}}(i)), \qquad (3)$$

where $P_{n_{\mathcal{B}}}(i)$ and $Q_{n_{\mathcal{B}}}(i)$ are defined in [13, eq. (3)] and [13, eq. (4)]. The initial conditions are $q_j^{(0)} = p$ for $j \in \{1, \ldots, T\}$ and $q_j^{(l)} = 0$ for $j \notin \{1, \ldots, T\}$. For tailbiting ensembles, the subscript $j - k' + k$ in (2) is calculated modulo T.

The analysis in [13] is presented for unshortened BCH codes, i.e., $s = 0$. However, CNs connected to known variable nodes are treated as shortened component codes by adjusting the effective error probability of the incoming messages through the boundary condition $q_j^{(l)} = 0$ for $j \notin \{1, \ldots, T\}$. If the component codes are shortened BCH codes, one can therefore apply the same analysis as before, where $n_{\mathcal{B}}$ now denotes the length of the unshortened code and the function $f_{n_{\mathcal{B}}}(x; p)$ is replaced by $f_{n_{\mathcal{B}}}(x(n_{\mathcal{B}} - s)/n_{\mathcal{B}}; p)$.

The BER for the VNs at position j after the lth iteration was not derived in [13], but can be easily found as follows. First, we rewrite (3) in the form

$$f_{n_{\mathcal{B}}}(x; p) = pf_{n_{\mathcal{B}}}^{1\to1}(x) + (1 - p)f_{n_{\mathcal{B}}}^{0\to1}(x), \qquad (4)$$

where $f_{n_{\mathcal{B}}}^{1\to1}(x)$ and $f_{n_{\mathcal{B}}}^{0\to1}(x)$ are implicitly defined via (3). We introduce the two variables

$$a_j^{(l)} \triangleq f_{n_{\mathcal{B}}}^{1\to1}\left(\frac{1}{w} \sum_{k'=0}^{w-1} q_{j-k'}^{(l)} \right), \quad b_j^{(l)} \triangleq f_{n_{\mathcal{B}}}^{0\to1}\left(\frac{1}{w} \sum_{k'=0}^{w-1} q_{j-k'}^{(l)} \right) \qquad (5)$$

and their averages

$$\bar{a}_j^{(l)} \triangleq \frac{1}{w} \sum_{k=0}^{w-1} a_{j+k}^{(l)} \quad \text{and} \quad \bar{b}_j^{(l)} \triangleq \frac{1}{w} \sum_{k=0}^{w-1} b_{j+k}^{(l)}. \quad (6)$$

With these definitions, the recursion (2) becomes

$$q_j^{(l)} = p\bar{a}_j^{(l-1)} + (1-p)\bar{b}_j^{(l-1)} \quad (7)$$

and the decoding error probability after the lth iteration is

$$p_{\mathrm{e},j}^{(l)} = p\left(\bar{a}_j^{(l-1)}\right)^2 + (1-p)\left(1 - \left(1 - \bar{b}_j^{(l-1)}\right)^2\right). \quad (8)$$

The final BER after $l_{\max}$ steps of iterative HDD is computed as $p_e = \frac{1}{T} \sum_{j=1}^{T} p_{\mathrm{e},j}^{(l_{\max})}$.

Since we intend to use the DE analysis in an optimization routine, we approximate the two functions $f_{n_{\mathcal{B}}}^{1\to1}(x)$ and $f_{n_{\mathcal{B}}}^{0\to1}(x)$ with their high-rate scaling limit versions (i.e., for $n_{\mathcal{B}} \to \infty$) which are easier to compute and given by [13]

$$f_{n_{\mathcal{B}}}^{1\to1}(x) \approx \phi(n_{\mathcal{B}}x; t-1) \quad (9)$$

and

$$f_{n_{\mathcal{B}}}^{0\to1}(x) \approx \frac{1}{n_{\mathcal{B}}(t-1)!}\phi(n_{\mathcal{B}}x; t), \quad (10)$$

where $\phi(\lambda; t) = 1 - \sum_{i=0}^{t} \frac{\lambda^i}{i!} e^{-\lambda}$.

It is straightforward to modify the decoding algorithm and the DE analysis if a similar WD as for the protograph-based SC-LDPC codes is used and hence we omit the details.

Example 6: Consider the case where $T = 20$ and $w = 2$. Let $\mathcal{B}$ be the same BCH code as in Example 3. The design rates are $R'(T) = 0.475$ and $R' = 0.5$, respectively. Assume transmission using PM-QPSK in the linear regime and a WD with $W = 5$ and $l_{\max} = 10$. In Fig. 5 (b), we show the predicted BER obtained via DE (solid lines) together with the actual performance of randomly generated codes (dashed lines) for $m_{\mathrm{c}} = 200$ (crosses) and $m_{\mathrm{c}} = 1500$ (dots) for both the terminated (blue) and tailbiting (red) cases. The decoding thresholds at a BER of 10^{-5} are $\rho^* \approx 3.71$ dB and $\rho^* \approx 3.94$ dB, respectively. $\triangle$

V. BIT MAPPER OPTIMIZATION

The different modulation bits of a nonbinary signal constellation have different protection levels, which can be taken advantage of by optimizing the bit mapper. This concept is easiest to understand for HDD, which we describe first.

For the "hard" demodulator, the entire block diagram shown in Fig. 1 can be replaced by m parallel BSCs with different crossover probabilities p_i, $1 \le i \le m$, which depend on the signal constellation, binary labeling, and SNR ρ. Each VN corresponds to a coded bit, and for the SC-GLDPC codes there are $m_{\mathrm{c}}n_{\mathcal{B}}/2$ VNs at each spatial position (see Section IV-A). The baseline bit mapper (see Section II) allocates the same number of coded bits from each spatial position to the different modulation bits (i.e., the m parallel BSCs). In this case, the crossover probability for the VNs at an arbitrary spatial position is simply the average $\bar{p} = \frac{1}{m} \sum_{i=1}^{m} p_i$. More generally, the bit mapper is modeled by specifying the assignment of VNs to the modulation bits via a matrix $\mathbf{A} = [a_{i,j}] \in \mathbb{R}^{m \times T}$, where

$a_{i,j}$, $0 \le a_{i,j} \le 1 \; \forall i, j$, denotes the proportional allocation of the coded bits corresponding to the VNs at spatial position j allocated to the ith modulation bit, and $\sum_{i=1}^{m} a_{i,j} = 1$, for all j. The effective crossover probability for the VNs at spatial position j is therefore a weighted average of the BSC crossover probabilities according to $\varepsilon_j = \sum_{i=1}^{m} a_{i,j}p_i$. To account for different crossover probabilities at the spatial positions in the DE analysis, we can simply replace p in (2) by ε_j.

For the protograph-based SC-LDPC codes with SDD, one can make similar considerations. Each VN in the protograph represents M VNs in the lifted Tanner graph, i.e., M coded bits. If we assume for example that a given protograph is lifted with a lifting factor M which is divisible by m, the baseline bit mapper allocates M/m coded bits for each protograph VN to each modulation bit. The bit mapper is modeled via a matrix $\mathbf{A} = [a_{i,j}] \in \mathbb{R}^{m \times n'_c}$, where $a_{i,j}$ now denotes the proportional allocation of the coded bits corresponding to the jth column in the base matrix allocated to the ith modulation bit. The matrix $\mathbf{A}$ is then used in the modified P-EXIT analysis to predict the iterative performance behavior under SDD [16, Algorithm 1].

We optimize $\mathbf{A}$ based on the decoding threshold with the help of differential evolution [33]. For more details about the optimization procedure, we refer the reader to [16], where we also discuss several techniques to reduce the optimization complexity for SC codes. Once an optimized bit mapping matrix $\mathbf{A}^*$ is found, the finite-length bit mapper is obtained via the rounded matrix $(m_{\mathrm{c}}n_{\mathcal{B}}/2)\mathbf{A}^*$ for the SC-GLDPC codes and via $M\mathbf{A}^*$ for the SC-LDPC codes, from which the index assignment of coded bits to modulation bits is determined.

VI. RESULTS AND DISCUSSION

Since this paper does not deal with code design, we rely on code parameters that have been proposed elsewhere in the literature in order to illustrate the bit mapper optimization technique. For the numerical results, we consider protograph-based SC-LDPC codes with $\mathbf{P}_0 = (1, 2, 1, 2)$ and $\mathbf{P}_1 = (3, 2, 3, 2)$, where $J' = 1$, $K' = 4$, and $m_s = 1$ [34]. The design rate of the tailbiting case is $R = 0.75$ (OH = 33%). For the SC-GLDPC codes, we restrict ourselves to $w = 2$ and use the BCH code parameters in [35, Table I] which are optimized for staircase codes. In particular, we consider $\nu = 9$, $t = 4$, and $s = 223$, which again leads to $R' = 0.75$. The staircase code for these parameters is estimated to perform approximately 1.38 dB away from the BSC capacity (at a BER of 10^{-15}) [35, Table I]. We also consider an example with higher rate and performance closer to capacity. In particular, we consider $\nu = 10$, $t = 4$, $s = 143$ where $R' = 0.91$ (OH = 10%). The gap to capacity of the staircase code for these parameters is estimated to be 0.59 dB [35, Table I].

The bit mapper optimization is performed for the terminated and tailbiting cases of the three code examples for different spatial lengths $T \in \{12, 21, 30, 39, 48, 57, 66, 75, 84, 300\}$. We consider Gray-labeled PM-16-QAM, PM-64-QAM, and PM-256-QAM. In all scenarios, a WD is employed with a window size of $W = 5$ and $l_{\max} = 10$ iterations per window. The target BER for the optimization is set to 10^{-5}. Setting a lower target BER (e.g., 10^{-15}) has virtually no influence on the

(a) SC-LDPC, $R = 0.75$, PM-64-QAM
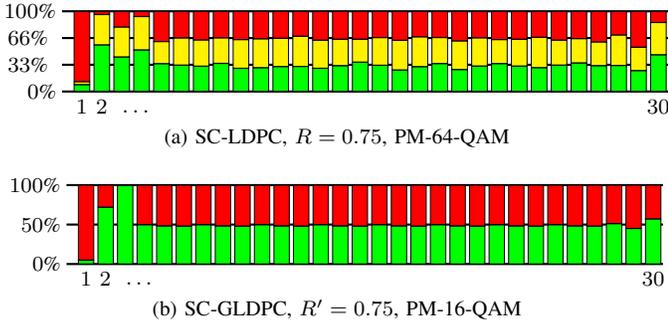

(b) SC-GLDPC, $R' = 0.75$, PM-16-QAM

Fig. 7. Optimized allocation to the modulation bits with the best (green), worst (red), and intermediate (yellow, for PM-64-QAM) protection level for each spatial position of the tailbiting code in two scenarios.

optimization outcome due to the steepness of the predicted DE performance curves (see Fig. 5). This assumes that there are no error floors due to harmful graph structures, which cannot be modeled using DE and are not considered in this paper. An analysis of the error floor for the considered SC-GLDPC codes is an interesting topic for future work and beyond the scope of this paper.

A. Structure of the Optimized Bit Mapper for Tailbiting Codes

For the tailbiting codes, the optimized bit mappers have an interesting structure, which is illustrated in Fig. 7 for two scenarios: (a) the SC-LDPC code with SDD, PM-64-QAM, and $T = 30$; (b) the SC-GLDPC code with HDD, $R' = 0.75$, PM-16-QAM, and $T = 30$. For PM-64-QAM and PM-16-QAM, the modulation bits have three and two different protection levels, respectively. Due to the tailbiting code structure, the bit allocation is invariant to a circular shift, assuming that the scheduling of the WD is modified according to the same shift. For the allocation shown in Fig. 7, it is assumed that the first decoding window begins at the first spatial position. The optimized bit mapper in both scenarios deviates significantly from the baseline mapper in the first few spatial positions. For the SC-LDPC code, the coded bits corresponding to the second, third, and fourth spatial position are proportionally more allocated to the best (green) and intermediate (yellow) protection level of PM-64-QAM. Similarly, for the SC-GLDPC code, the coded bits corresponding to the second and third spatial position are proportionally more allocated to the best modulation bit of PM-16-QAM. In both cases, the optimized allocation leads to a locally improved decoding convergence and initiates a wave-like decoding behavior comparable to that of terminated codes, i.e., the unequal error protection of the signal constellation is exploited to create an artificial termination boundary.

The performance gain due to the optimized bit mapper (which is quantified in the next section) comes at the expense of some increase in system complexity. In particular, one has to account for additional buffering because a symbol cannot be transmitted until all its m bits are encoded. For simplicity, let us assume a model where the FEC encoder outputs coded bits in blocks of MK' or $m_c n_B/2$ bits, i.e., the number of bits per spatial position, and symbols are immediately modulated as soon as all m modulation bits are available. Then, no buffering

is required for the sequential baseline mapper. On the other hand, the "worst-case" bit mapper allocates 100% of the coded bits in the first T/m spatial positions to the first modulation bit, 100% in the next T/m positions to the second bit, and so on (i.e., $a_{i,j} = 1$ for $(i-1)T/m + 1 \leq j \leq iT/m$). Consequently, no bits are allocated to the last modulation bit until spatial position $(m-1)T/m + 1$ and buffering of all coded bits up to position $(m-1)T/m$ is required. In all considered scenarios, however, the required additional buffer size (in terms of the number of spatial positions) due to the optimized bit mappers did not exceed 2.

B. Optimization Gain

In Fig. 8, we show the optimization gain (in dB) as a function of the spatial length T for all considered scenarios. The optimization gain is defined as the difference between the decoding threshold using the baseline bit mapper and the decoding threshold using the optimized bit mapper. The gain quantifies the performance improvement one can expect by employing the optimized bit mappers assuming long codes.

Regardless of the signal constellation or code class, the optimization gain decreases with T for the terminated codes and increases for the tailbiting codes. This behavior can be explained as follows. The optimization gain for the tailbiting codes comes from allocating more coded bits in the beginning of the spatial chain to good modulation bits in order to initiate a decoding wave. This, however, reduces the effective capacity for the bits in the middle part of the spatial chain. As T increases, this reduction becomes negligible and the optimization gain tends to a constant value. For terminated codes, a decoding wave is initiated by default and the optimized bit mapper increases the effective capacity for the bits in the middle part by allocating bits in the beginning and end of the chain proportionally more to modulation bits with lower protection levels. Again, as T increases, this effect becomes negligible and the gain approaches zero. As a result, while the tailbiting codes significantly benefit from the optimization, the gain for the considered terminated codes is limited, i.e., for $T \geq 30$ the gain is < 0.1 dB in all cases.

It can also be observed that the optimization gain generally depends on the signal constellation. The gain increases with the modulation order M due to the increased number of protection levels and stronger unequal error protection. This gain increase can also be observed when optimizing bit mappers for irregular LDPC codes, see, e.g., [36]. It is also important to stress that the optimization relies on the availability of a signal constellation with different protection levels in order to provide a performance gain. In particular, the techniques do not apply to PM-BPSK or (Gray-labeled) PM-QPSK.

C. Gap to Capacity

In order to gain some insight into the performance of the terminated and tailbiting codes relative to each other, the capacity gap (in dB) as a function of the spatial length T is shown in Fig. 9 for PM-64-QAM. For SDD of the SC-LDPC codes, the BICM capacity [5] is taken as a benchmark. For HDD of the GLDPC codes, the capacity of the BSC with
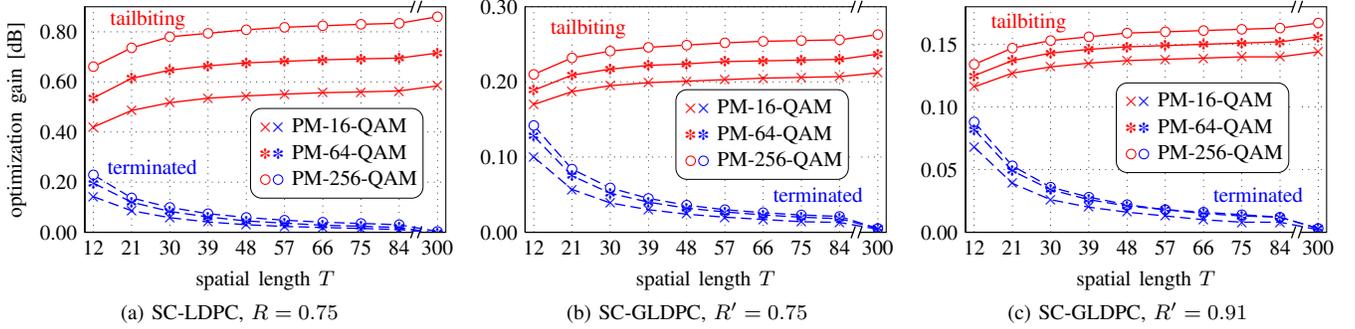
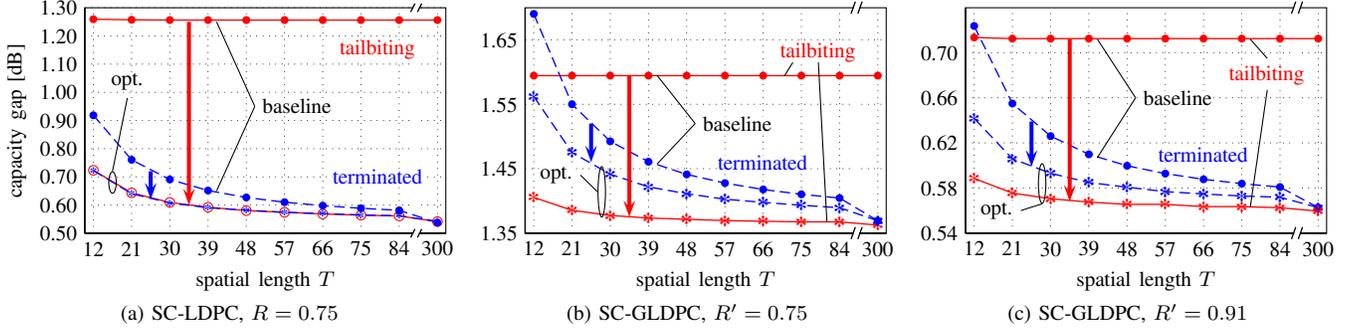Fig. 8. Optimization gain as a function of the spatial length T.

(a) SC-LDPC, $R = 0.75$

(b) SC-GLDPC, $R' = 0.75$

(c) SC-GLDPC, $R' = 0.91$

Fig. 9. Capacity gap as a function of the spatial length T for PM-64-QAM.

(a) SC-LDPC, $R = 0.75$

(b) SC-GLDPC, $R' = 0.75$

(c) SC-GLDPC, $R' = 0.91$

averaged crossover probability is taken as a benchmark, similar to [37]. Alternatively, one may use the capacity of the sum of the m parallel BSCs as a benchmark, which is larger. The gains discussed in the previous subsection are indicated in Fig. 9 with arrows.

The decoding thresholds for the baseline systems are approximately independent of T. Therefore, the capacity gap for the tailbiting codes remains constant in all cases, while the capacity gap for the terminated codes decreases due to the vanishing rate loss. For the baseline systems, the performance difference between terminated and tailbiting codes is most significant for the SC-LDPC codes (up to 0.75 dB), while for the SC-GLDPC codes the difference is lower (up to 0.25 dB for $R' = 0.75$ and up to 0.19 dB for $R' = 0.91$). In all cases, the capacity gap is reduced by employing the optimized bit mappers. If we compare the optimized systems, it can be seen that the gap for the SC-LDPC codes is virtually identical for terminated and tailbiting cases. For the SC-GLDPC codes, the tailbiting codes perform closer to capacity, albeit the difference to the terminated codes for $T \geq 30$ is small. For very long spatial lengths (i.e., $T = 300$), the capacity gap virtually overlaps also for the SC-GLDPC codes.

D. Simulation Results

The results presented in the previous two subsections are based on decoding thresholds, i.e., assume an infinite code length. The deviation of the DE analysis from the finite-length performance is determined by the lifting factor M and the number of CNs per position m_c, see Fig. 5.

As an example, consider the SC-LDPC code with $T = 30$

and $M = 3000$ leading to a code length of $n_{\mathcal{C}} = 360\,000$. The rates are $R(30) \approx 0.742$ and $R = 0.75$, respectively. In Fig. 10, we show simulation results (dashed lines with dots) and the analytical P-EXIT prediction (solid lines) for the AWGN channel, i.e., a linear transmission scenario, assuming PM-64-QAM. As predicted by the optimization gain in Fig. 8 (a), the tailbiting code performs significantly better with an optimized bit mapper and a gain of ≈ 0.55 dB is achieved at a BER of 10^{-5}. The terminated code performs better for the same SNR, but entails a smaller spectral efficiency due to the rate loss. The gap to the BER-constrained BICM capacity [22, p. 17] of the two optimized systems, as indicated by the arrows and predicted from Fig. 9 (a), is approximately the same (as is the gap to the AWGN channel capacity, not shown).

Lastly, we also present simulation results for a nonlinear transmission scenario. We set $\alpha = 0.25$ dB/km, $\beta_2 = -21.668$ ps^2/km, $\gamma = 1.4$ W^{-1} km^{-1}, $\nu_s = 1.934 \times 10^{14}$ Hz, $n_{\mathrm{sp}} = 1.622$, $R_s = 40$ GBaud, and $L_{\mathrm{sp}} = 70$ km. A root-raised cosine pulse $p(t)$ with a roll-off factor of 0.25 is used. We employ the symmetric split-step Fourier method with two samples per symbol and a fixed step size [38, Sec. 2.4.1]. The input power per polarization is set to $P = -2.5$ dBm. In the simulation model, the polarization state is assumed to be known and perfect timing and carrier synchronization is assumed. In Fig. 11, the simulated BER of the PM transmission systems is plotted as a function of the number of fiber spans N_{sp}. For the tailbiting code, the 0.55 dB gain obtained by using the optimized bit mapper translates into an increase of the transmission reach by roughly 3 additional spans or approximately 13%. This gain is obtained at almost
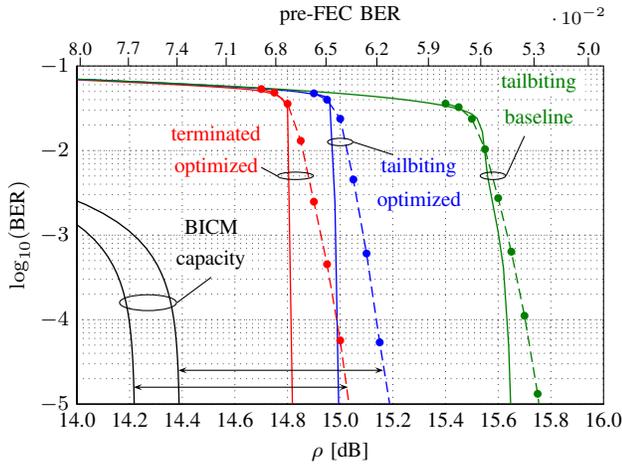
Fig. 10. Simulation results for the SC-LDPC code with PM-64-QAM over the AWGN channel. The codes have length $360\,000$.
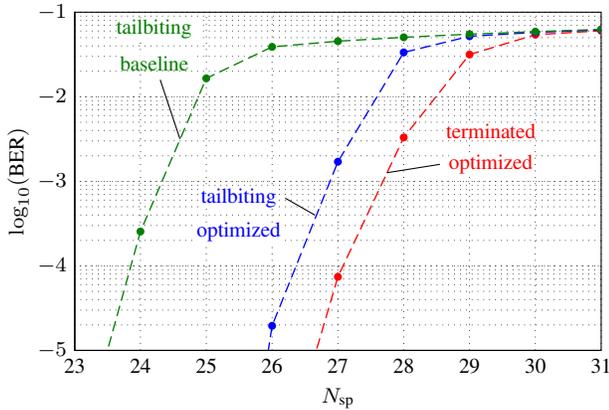

Fig. 11. Simulation results for the SC-LDPC with PM-64-QAM over a dispersion uncompensated transmission link.

no increased system complexity cost, i.e., by simply replacing the baseline bit mapper with an optimized one. This reach extension can also be approximately calculated using the analytical expression for the SNR ρ as a function of the number of spans presented in [24]. The terminated code enables a longer transmission reach of approximately one span, at the expense of a 1.2% decrease in spectral efficiency. The performance of the terminated code with the baseline bit mapper is very close to the performance of the tailbiting code with the optimized bit mapper and is therefore not shown in Figs. 10 and 11 for clarity.

VII. CONCLUSIONS

In this paper, we considered the optimized allocation of the coded bits from the FEC encoder to the modulation bits for terminated and tailbiting SC FEC schemes, assuming both SDD and HDD, as well as different signal constellations. Terminated SC codes generally benefit little from the optimization, particularly for long spatial lengths. However, the performance of tailbiting SC-LDPC codes can be significantly improved. With an optimized bit allocation, the terminated and tailbiting codes are competitive, in the sense that spectral efficiency can

be traded for transmission reach, at approximately the same gap to capacity.

VIII. ACKNOWLEDGEMENTS

REFERENCES

[1] R.-J. Essiambre, G. Kramer, P. J. Winzer, G. J. Foschini, and B. Goebel, "Capacity limits of optical fiber networks," *J. Lightw. Technol.*, vol. 28, no. 4, pp. 662–701, Feb. 2010.

[2] B. P. Smith and F. R. Kschischang, "Future prospects for FEC in fiber-optic communications," *IEEE J. Sel. Topics. Quantum Electron.*, vol. 16, no. 5, pp. 1245–1257, Oct. 2010.

[3] L. Schmalen, A. J. de Lind van Wijngaarden, and S. ten Brink, "Forward error correction in optical core and optical access networks," *Bell Labs Tech. J*, vol. 18, no. 3, pp. 39–66, Mar. 2013.

[4] L. Beygi, E. Agrell, J. M. Kahn, and M. Karlsson, "Coded modulation for fiber-optic networks," *IEEE Signal Processing Mag.*, vol. 31, no. 2, pp. 93–103, Mar. 2014.

[5] G. Caire, G. Taricco, and E. Biglieri, "Bit-interleaved coded modulation," *IEEE Trans. Inf. Theory*, vol. 44, no. 3, pp. 927–946, May 1998.

[6] A. Leven and L. Schmalen, "Status and recent advances on forward error correction technologies for lightwave systems," in *Proc. European Conf. Optical Communication (ECOC)*, London, UK, 2013, p. We.2.C.1.

[7] Y. Miyata, T. Sugihara, K. Kubo, H. Yoshida, and W. Matsumoto, "A spatially-coupled type LDPC code with an NCG of 12 dB for optical transmission beyond 100 Gb/s," in *Proc. Optical Fiber Communication Conf. (OFC)*, Anaheim, USA, Mar. 2013, p. OM2B.4.

[8] S. Kudekar, T. Richardson, and R. Urbanke, "Threshold saturation via spatial coupling: Why convolutional LDPC ensembles perform so well over the BEC," *IEEE Trans. Inf. Theory*, vol. 57, no. 2, pp. 803–834, Feb. 2011.

[9] Sae-Young Chung, G. D. Forney, T. J. Richardson, and R. Urbanke, "On the design of low-density parity-check codes within 0.0045 dB of the Shannon limit," *IEEE Commun. Lett.*, vol. 5, no. 2, pp. 58–60, Feb. 2001.

[10] T. J. Richardson, M. A. Shokrollahi, and R. L. Urbanke, "Design of capacity-approaching irregular low-density parity-check codes," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 619–637, Feb. 2001.

[11] J. Thorpe, "Low-density parity-check (LDPC) codes constructed from protographs," *IPN Progress Report 42-154, JPL*, 2005.

[12] D. G. M. Mitchell, M. Lentmaier, and D. J. Costello, Jr., "AWGN channel analysis of terminated LDPC convolutional codes," in *Proc. Information Theory and Applications Workshop (ITA)*, La Jolla, CA, 2011.

[13] Y.-Y. Jian, H. D. Pfister, and K. R. Narayanan, "Approaching capacity at high rates with iterative hard-decision decoding," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Cambridge, MA, 2012.

[14] B. P. Smith, A. Farhood, A. Hunt, F. R. Kschischang, and J. Lodge, "Staircase codes: FEC for 100 Gb/s OTN," *J. Lightw. Technol.*, vol. 30, no. 1, pp. 110–117, Jan. 2012.

[15] A. Leven and L. Schmalen, "Status and recent advances on forward error correction technologies for lightwave systems," *J. Lightw. Technol.*, vol. 32, no. 16, pp. 2735–2750, Aug. 2014.

[16] C. Häger, A. Graell i Amat, F. Brännström, A. Alvarado, and E. Agrell, "Improving soft FEC performance for higher-order modulations via optimized bit channel mappings," *Opt. Express*, vol. 22, no. 12, pp. 14 544–14 558, Jun. 2014.

[17] M. Lentmaier and G. P. Fettweis, "On the thresholds of generalized LDPC convolutional codes based on protographs," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Austin, TX, 2010, pp. 709–713.

[18] Y.-Y. Jian, H. D. Pfister, K. R. Narayanan, R. Rao, and R. Matahreh, "Iterative hard-decision decoding of braided BCH codes for high-speed optical communication," in *Proc. IEEE Glob. Communication Conf. (GLOBECOM)*, Atlanta, GA, 2014.

[19] I. Djordjevic, O. Milenkovic, and B. Vasic, "Generalized low-density parity-check codes for optical communication systems," *J. Lightw. Technol.*, vol. 23, no. 5, pp. 1939–1946, May 2005.

[20] I. Djordjevic, L. Xu, T. Wang, and M. Cvijetic, "GLDPC codes with Reed-Muller component codes suitable for optical communications," *IEEE Commun. Lett.*, vol. 12, no. 9, pp. 684–686, Sep. 2008.

[21] J. Hou, P. H. Siegel, L. B. Milstein, and H. D. Pfister, "Capacity-approaching bandwidth-efficient coded modulation schemes based on low-density parity-check codes," *IEEE Trans. Inf. Theory*, vol. 49, no. 9, pp. 2141–2155, Sep. 2003.

[22] W. Ryan and S. Lin, *Channel Codes Classical and Modern*. Cambridge University, 2009.

[23] G. P. Agrawal, *Lightwave Technology: Telecommunication Systems*. Wiley-Interscience, 2005.

[24] L. Beygi, E. Agrell, P. Johannisson, M. Karlsson, and H. Wymeersch, "A discrete-time model for uncompensated single-channel fiber-optical links," *IEEE Trans. Commun.*, vol. 60, no. 11, pp. 3440–3450, Nov. 2012.

[25] A. Carena, G. Bosco, V. Curri, Y. Jiang, P. Poggiolini, and F. Forghieri, "EGN model of non-linear fiber propagation." *Opt. Express*, vol. 22, no. 13, pp. 16 335–16 362, Jun. 2014.

[26] A. R. Iyengar, M. Papaleo, P. H. Siegel, J. K. Wolf, A. Vanelli-Coralli, and G. E. Corazza, "Windowed decoding of protograph-based LDPC convolutional codes over erasure channels," *IEEE Trans. Inf. Theory*, vol. 58, no. 4, pp. 2303–2320, Apr. 2012.

[27] M. B. Tavares, K. S. Zigangirov, and G. P. Fettweis, "Tail-biting LDPC convolutional codes," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Nice, Italy, 2007.

[28] T. Richardson and R. Urbanke, "The capacity of low-density parity-check codes under message-passing decoding," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 599–618, Feb. 2001.

[29] S. Ten Brink, "Convergence of iterative decoding," *Electronics Letters*, vol. 35, no. 10, pp. 806–808, May 1999.

[30] G. Liva and M. Chiani, "Protograph LDPC codes design based on EXIT analysis," in *Proc. IEEE Glob. Communication Conf. (GLOBECOM)*, Washington, DC, 2007.

[31] Y.-Y. Jian, "On the analysis of spatially-coupled GLDPC codes and the weighted min-sum algorithm," Ph.D. dissertation, Texas A&M University, 2013.

[32] B. P. Smith, "Error-correcting codes for fibre-optic communication systems," Ph.D. dissertation, University of Toronto, 2011.

[33] R. Storn and K. Price, "Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces," *J. Global Opt.*, vol. 11, pp. 341–359, Nov. 1997.

[34] L. Schmalen and S. ten Brink, "Combining spatially coupled LDPC codes with modulation and detection," in *Proc. Int. Conf. Systems, Communication and Coding (SCC)*, Munich, Germany, Jan. 2013.

[35] L. M. Zhang and F. R. Kschischang, "Staircase codes with 6% to 33% overhead," *J. Lightw. Technol.*, vol. 32, no. 10, pp. 1999–2002, May 2014.

[36] T. Cheng, K. Peng, J. Song, and K. Yan, "EXIT-aided bit mapping design for LDPC coded modulation with APSK constellations," *IEEE Commun. Lett.*, vol. 16, no. 6, pp. 777–780, Jun. 2012.

[37] B. Smith and F. R. Kschischang, "A pragmatic coded modulation scheme for high-spectral-efficiency fiber-optic communications," *J. Lightw. Technol.*, vol. 30, no. 13, pp. 2047–2053, Jul. 2012.

[38] G. P. Agrawal, *Nonlinear Fiber Optics*, 4th ed. Academic Press, 2006.