

Irregular Polar Coding for Complexity-Constrained Lightwave Systems

Toshiaki Koike-Akino, *Senior Member, IEEE, Member, OSA*, Congzhe Cao, Ye Wang, Stark C. Draper, *Senior Member, IEEE, Member, OSA*, David S. Millar, *Member, IEEE, Member, OSA*, Keisuke Kojima, *Senior Member, IEEE, Fellow, OSA*, Kieran Parsons, *Senior Member, IEEE, Member, OSA*, Lidia Galdino, *Member, IEEE*, Daniel J. Elson, Domaniç Lavery, *Member, IEEE*, Polina Bayvel, *Fellow, IEEE*

(Invited Paper)

Abstract—Next-generation fiber-optic communications call for ultra-reliable forward error correction (FEC) codes that are capable of low-power and low-latency decoding. In this paper, we propose a new class of polar codes, whose polarization units are irregularly pruned to reduce computational complexity and decoding latency without sacrificing error correction performance. We then experimentally demonstrate that the proposed irregular polar codes can outperform state-of-the-art LDPC codes, while decoding complexity and latency can be reduced by at least 30% and 70%, respectively, versus regular polar codes, while also obtaining a marginal performance improvement.

Index Terms—Polar coding, FEC, coherent fiber-optic communications, irregular degree, complexity/latency reduction

I. INTRODUCTION

CAPACITY-approaching forward error correction (FEC) based on low-density parity-check (LDPC) codes [1]–[8] have made a great contribution to increasing data rates of coherent optical transceivers achieving beyond Tb/s [9]–[13]. However, the pursuit of high FEC performance has led to a significant increase in power consumption and circuit size. Hence, attaining a good trade-off between performance and computational complexity is of great importance. In addition, recent high-performance FEC codes usually require very large codeword lengths, typically longer than 30,000 bits, whereas shorter FEC codes are preferred for latency-constrained systems including short-reach optical intercon-

nects and front/back-hauls. This paper introduces reduced-complexity polar codes as potential alternatives to LDPC codes, with particular emphasis on short block lengths for latency-constrained lightwave systems.

Polar codes [15]–[22] have drawn much attention in the coding theory community due to their ability to achieve capacity via low-complexity successive cancellation (SC) decoding for any arbitrary discrete-input memoryless channels (DMCs). However, in spite of the theoretical strength, polar codes have not been adopted until recently in practical systems due to their poor performance at short block lengths in comparison to LDPC codes. The recent breakthrough in algorithm development of successive cancellation list (SCL) decoding [18] has made polar codes competitive with state-of-the-art LDPC codes, in particular for low-complexity and latency-constrained systems [35]. In consequence, polar codes have just been adopted in the fifth-generation wireless standards.

To date there have been relatively few studies [3] comparing LDPC codes with polar codes in the optical research community. We first re-confirm that polar codes employing SCL decoding can outperform recent LDPC codes [7] for latency-constrained lightwave systems when the block length is below 10,000 bits. For comparison, we also consider turbo product codes (TPC) [24]–[26], which employ iterative message passing for Bose–Chaudhuri–Hocquenghem (BCH) constituent codes to achieve near maximum-likelihood performance.

We then propose a new family of polar codes, whose polarization units are irregularly pruned to reduce the computational complexity at both encoding and decoding. In the context of LDPC codes, it is well-known that irregular codes with specific degree distributions outperform regular ones. For polar codes, there exist few investigations incorporating analogous irregularity. Mixed-kernel polar codes [32] first showed the potential of irregular substitution with different kernels, leading to high flexibility. However, optimization of kernel mixtures has been challenging due to the huge design space. Relaxed polar codes [33] introduced a limited amount of irregularity, where some polarization units are deactivated when the bit-channels are sufficiently good or bad. However, relaxed polar codes [33] do not consider all possible combinations of inactivation; specifically once a polarization unit is inactivated, further polarizations after the inactivated unit are not considered. This limits the flexibility in the selection of polarization units that

Manuscript received October 23, 2017; revised January 14, 2018; revised January 30, 2018; accepted January 30, 2018.

T. Koike-Akino, Y. Wang, D. S. Millar, K. Kojima, and K. Parsons are with Mitsubishi Electric Research Laboratories (MERL), 201 Broadway, Cambridge, MA 02139, USA (e-mail: koike@merl.com; yewang@merl.com; millar@merl.com; kojima@merl.com; parsons@merl.com).

C. Cao is with Department of Electrical & Computer Engineering, University of Alberta, Edmonton, AB T6G 2V4, Canada (e-mail: congze@ualberta.ca). He conducted this study at MERL as a research intern.

S. C. Draper is with Electrical & Computer Engineering, University of Toronto, Toronto, ON M5S 3G4, Canada (e-mail: stark.draper@utoronto.ca).

L. Galdino, D. J. Elson, D. Lavery, and P. Bayvel are with Optical Networks Group, University College London (UCL), Torrington Place, London, WC1E 7JE, UK (e-mail: l.galdino@ucl.ac.uk; daniel.elson.13@ucl.ac.uk; d.lavery@ee.ucl.ac.uk; p.bayvel@ee.ucl.ac.uk).

This paper contains in part our previous work [11], [12], [35]–[41].

This work was in part funded by the UK EPSRC Programme Grant UNLOC EP/J017582/1.

Color versions of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier xx.xxx/JLT.2018.xxx

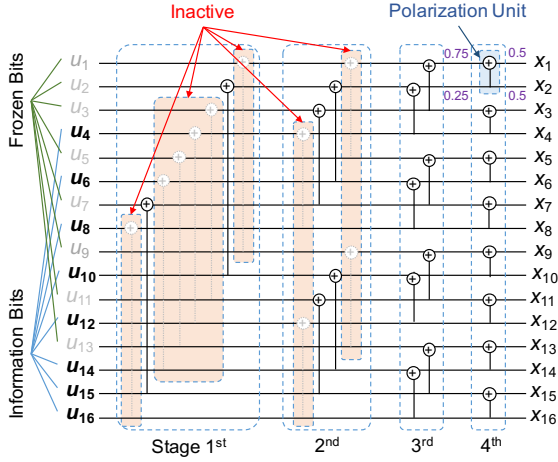


Fig. 1. Four-stage polarization with inactivated units: irregular polar codes (16, 8) having 10 inactivations over 32 polarization units (10/32 = 31.25% complexity reduction).

can be inactivated and does not exploit the possibility of further complexity reduction.

In this paper, we propose fully irregular polar codes by taking the non-uniform reliability of bit-channels into consideration when performing inactivation of polarization units. We show that performance can be slightly improved when compared to conventional (regular) polar codes, while a significant reduction in decoding complexity and latency can be achieved. Experimental validation of the proposed irregular polar codes is demonstrated over 400 km transmission of Corning® Vascade® EX2000 fiber, for two 60 GBd subchannels of dual-polarization 64-ary quadrature-amplitude modulation (DP-64QAM) under full C-band loading [12].

The contributions of this paper are summarized as follows:

- **Irregular inactivation of polarization units:** We generalize the idea in [33] to consider all possible selection of the inactivated polarization units for significant reduction of encoding/decoding complexity.
- **Efficient code construction algorithm:** We propose a greedy algorithm to construct irregular polar codes that minimize bit error rate (BER) through the use of extrinsic information transfer (EXIT) analysis [28] to jointly optimize frozen bit and inactivation unit locations. The proposed algorithm facilitates the complicated optimization of mixed-kernel polar codes [32].
- **Complexity and latency analysis:** We analyze the complexity and latency reductions of the proposed irregular polar codes to verify the advantage in power-constrained optical communications.
- **Experimental validation:** We experimentally demonstrate that our irregular polar coding can offer better performance than regular counterparts as well as state-of-the-art LDPC codes in fiber-optic transmission measurement.
- **Error pattern analysis:** We also provide a brief analysis of the Hamming weight spectrum and error patterns when the polar codes have irregular inactivations.

We note that this paper is distinguished from our previous reports [35]–[40] as follows. We previously studied interleaver

design for polar-coded high-order QAM in [35], achieving a 0.5 dB gain. This work was extended with constellation shaping for wireless fading channels in [36], [37], achieving a 2.5 dB improvement. The concept of irregular polar coding was first proposed for optical communications in [38], later extended to wireless massive antenna systems in [39], and applied to TPC for parallel decoding in [40] and [41] for optical and wireless communications, respectively. In this paper, we extend the preliminary report in [38] with more detailed analyses of the decoding error union bound (UB), decoding complexity, and latency, in addition to a description of the code construction algorithm.

II. POLAR CODES WITH SCL DECODING

A. Polar Encoding

An (N, k) polar code with k information bits and N encoded bits ($N = 2^n$) uses an $N \times N$ generator matrix $F^{\otimes n}$ for encoding, where

$$F = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \quad (1)$$

is a binary kernel matrix and $[\cdot]^{\otimes n}$ denotes the n -fold Kronecker power. Let $\mathbf{u} = [u_1, u_2, \dots, u_N]^T$ and $\mathbf{x} = [x_1, x_2, \dots, x_N]^T$ respectively denote the vectors of input bits and encoded bits. The codeword (for non-systematic polar codes) is given by $\mathbf{x} = F^{\otimes n} \mathbf{B} \mathbf{u}$, where the matrix multiplications are carried out over the binary field (i.e., modulo-2 arithmetic), and \mathbf{B} denotes an $N \times N$ bit-reversal permutation matrix [15]. Due to the nature of Kronecker product, polar encoding and decoding can be performed at a complexity on the order of $\mathcal{O}(N \log_2 N)$ over the n -stage polarization shown in Fig. 1. The multi-stage operation of the Kronecker products gives rise to the so-called polarization phenomenon to approach capacity in arbitrary DMCs [15].

The polar coding maps the information bits to the k most reliable locations in \mathbf{u} . The remaining $N - k$ input bits are frozen bits, fixed to values known to both encoder and decoder. We use \mathbb{K} and $\bar{\mathbb{K}}$ to denote the subsets of $\{1, 2, \dots, N\}$ that correspond to the information bit and frozen bit locations, respectively. By means of various design methods, such as Bhattacharyya parameter [15], density evolution (DE) [20], [30], and Gaussian approximation [31], the locations in \mathbf{u} with the lowest reliability can be selected to be in $\bar{\mathbb{K}}$ for frozen bits.

For systematic polar codes, we employ an encoding procedure [17] which writes the k data bits into a vector \mathbf{u} at the bit-reversal permutation of the locations \mathbb{K} , with the other locations set to zero, and then applies the polar encoding procedure twice, while setting the frozen bit locations in $\bar{\mathbb{K}}$ to zero on the intermediate result between the encodings. This procedure for systematic coding can be expressed as follows:

$$\mathbf{x} = F^{\otimes n} \mathbf{B} \phi_{\bar{\mathbb{K}}}(F^{\otimes n} \mathbf{B} \mathbf{u}), \quad (2)$$

where $\phi_{\bar{\mathbb{K}}}(\cdot)$ denotes the setting of the frozen bit locations in $\bar{\mathbb{K}}$ to zero. We also use a concatenation of cyclic redundancy check (CRC) to increase the minimum Hamming distance [23].

B. Polar Decoding

Let $\mathbf{y} = [y_1, y_2, \dots, y_N]^T$ denote the vector of decoder inputs. Assuming memoryless channels, the transition probability $W_N(\mathbf{y}|\mathbf{x})$ between \mathbf{x} and \mathbf{y} is written as $W_N(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^N W(y_i|x_i)$. The original SC decoding proposed in [15] was recently improved by the SCL decoder in [18], which incorporates list decoding and an embedded CRC code to improve the performance for short-block polar codes. The SC decoder proceeds sequentially over the bits, from index 1 to N . For each index $i \in \{1, 2, \dots, N\}$, an estimate \hat{u}_i for bit u_i is made as follows. If $i \notin \mathbb{K}$, then \hat{u}_i is set to the known, fixed value of u_i , otherwise, when $i \in \mathbb{K}$, \hat{u}_i is set to the most likely value for u_i given the channel outputs, assuming that the previous estimates $[\hat{u}_1, \hat{u}_2, \dots, \hat{u}_{i-1}]$ are correct.

The SCL+CRC decoder proceeds similarly to the SC decoder, except that for each data bit index $i \in \mathbb{K}$, the decoder retains both possible estimates, $\hat{u}_i = 0$ and $\hat{u}_i = 1$, in subsequent decoding paths. To avoid handling an exponentially increasing number of paths, the list-decoding approach limits the number of paths to a fixed-size list of the most likely partial paths. The SCL decoder also employs a CRC code embedded in the data bits, which allows it to select the final decoding as the most likely path with a valid CRC. The combination of list-decoding with embedded CRC code to reject invalid paths yields significantly improved performance [18].

C. Polar Codes vs. LDPC Codes

As reported in [17]–[19], polar codes can outperform turbo codes and LDPC codes that are used in wireless standards when SCL+CRC decoding is adopted. Moreover, it was shown in [19] that polar codes with SCL+CRC decoding achieve near the Polyanskiy bound for finite-length codes. In [35], the authors verified the advantage of short block-length polar codes in additive white Gaussian noise (AWGN) channels even when compared to recently proposed Pareto-optimal LDPC codes [7], which show the best trade-off between threshold and decoding complexity by optimizing the degree distribution for finite-iteration decoding via EXIT trajectory.

For self-consistency, we here re-evaluate the bit-error rate (BER) performance of short-block polar codes to compare with state-of-the-art LDPC codes [7], [35] in AWGN channels for 4QAM transmission, before moving on to investigating irregular polar codes. Fig. 2(a) compares the BER performance of Pareto-optimal LDPC codes and systematic polar codes. Here, we use $I = 32$ iterations for LDPC layered decoding algorithm (LDA) [6], and a list size of $L = 32$ for polar SCL decoding (with CRC-8). The Pareto-optimal LDPC codes [7] were further improved for LDA in [35]. We consider five block lengths, $N \in \{256, 1024, 2048, 4096, 16384\}$. The code rates are identical to be $R = 0.8$ (including CRC overhead). As shown in Fig. 2(a), the BER performance can significantly degrade when the block length is limited. It is also verified that polar codes with SCL+CRC decoding can outperform LDPC codes at block lengths shorter than 3,000 bits.

The BER performance depends on decoding complexity, more specifically, constraints on the available number of iterations I for LDPC codes and the list size L for polar codes.

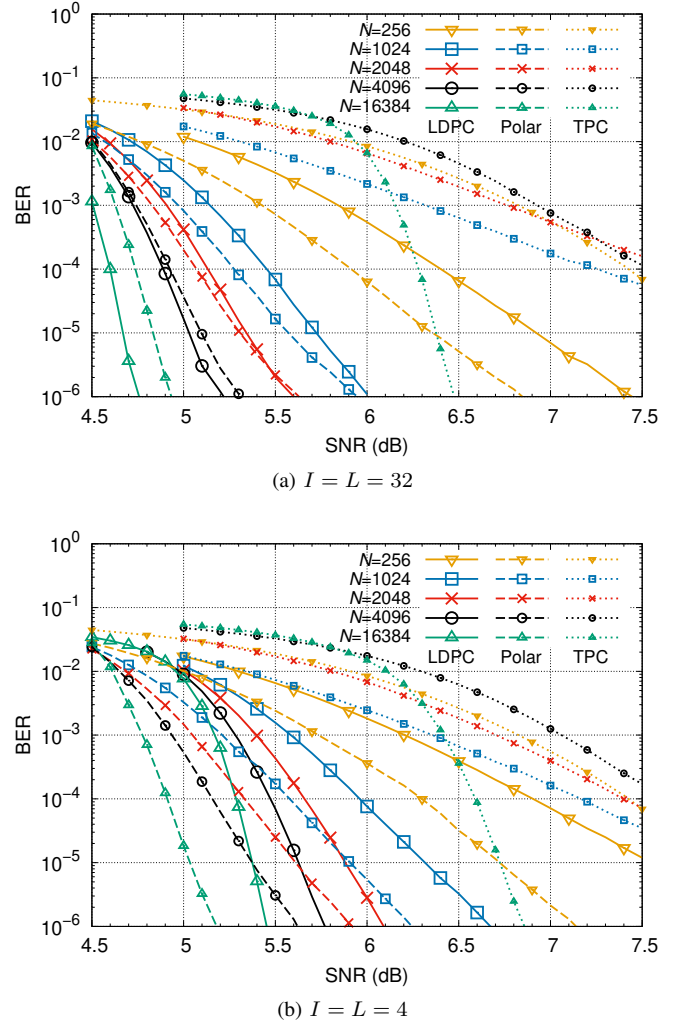


Fig. 2. BER comparison of LDPC codes with LDA BP decoding and polar codes with SCL+CRC decoding for various short block lengths.

In Fig. 2(b), we compare BER curves for the lower complexity cases of $I = L = 4$. Compared to the case of $I = L = 32$ in Fig. 2(a), most curves shift by approximately 0.5 dB due to the reduced decoder complexity. Nevertheless, the performance loss of polar codes is relatively small when compared to that suffered by LDPC codes. Consequently, LDPC codes perform worse than polar codes for all block lengths we considered. These results suggest that polar codes can be a viable FEC candidate for latency- and power-constrained lightwave systems. We note that the considered LDPC codes have potential to be improved further, e.g., by employing scattered EXIT [42], to account for finite-length characteristics. Nonetheless, our primary goal here is not to guarantee the absolute superiority of polar codes against all possible LDPC codes but rather to demonstrate that polar codes are able to compete with recent LDPC codes by evaluating general trends over different block lengths and decoding complexities.

For reference, Fig. 2 also presents the BER performance of TPC [24]–[26], which has been studied for short-reach optical interface, with $I = 32$ iterations for turbo decoding. TPC is constructed with BCH codes as listed in Table I,

TABLE I
TPC CONSTRUCTION PARAMETERS (TARGET RATE: 0.8)

Length	BCH-TPC	Actual rate
256	BCH(16, 11) \times BCH(16, 11)	0.473
1024	BCH(32, 26) \times BCH(32, 26)	0.660
2048	BCH(64, 57) \times BCH(32, 26)	0.724
4096	BCH(64, 57) \times BCH(64, 57)	0.732
16384	BCH(128, 113) \times BCH(128, 113)	0.779

where code rates are chosen to be no higher than 0.8 from narrow-sense BCH codes. We observe that the TPC suffers more than a 1.5 dB loss relative to LDPC codes and polar codes. This is because the short-length TPC constitutes of even shorter and higher-rate BCH component codes, whose minimum Hamming distance is small. Note that we considered only narrow-sense BCH codes without any puncturing or shortening and, hence, TPC has the potential to be further improved, in particular for longer-length and lower-rate cases. Further comparisons between polar and BCH codes to construct moderate-size TPC codes will be discussed in [40], [41].

D. Computational Complexity

It is already known that short LDPC codes do not perform well, e.g., in [43], where various other FEC schemes such as tail-biting convolutional codes (TB-CC) are compared at short block lengths. It was demonstrated that nonbinary (NB) LDPC codes as well as TB-CC have excellent performance in such cases. However, the computational complexity of NB-LDPC codes is generally higher than binary counterparts, in particular for large Galois field sizes. The complexity of TB-CC also increases exponentially with the constraint length. Therefore, we must consider computational complexity as well as BER performance. We now evaluate the computational complexity of SCL polar decoding to show that it is competitive to LDPC decoding even in the sense of complexity.

The nonlinear complexity of polar SCL decoding (specifically, $\mathcal{O}[LN \log_2(N)/2]$) is a major drawback in comparison to the linear complexity of LDPC BP decoding (specifically, $\mathcal{O}[2Id_v N]$). Here, \bar{d}_v denotes the average degree of variable nodes (VNs). Note that the factor of $2Id_v$ comes from the bidirectional message passing, whereas the SCL decoding uses unidirectional message passing over $N \log_2(N)/2$ VNs. Due to the nonlinear complexity, polar codes can eventually be less effective than LDPC codes as we increase the block lengths. However, this nonlinearity can become an advantage when we aim to reduce the block sizes in order to decrease latency. This is illustrated in Fig. 3, where complexity per coded bit (i.e., divided by N) is plotted as a function of block length N for polar and LDPC codes. Because per-bit complexity is constant depending on average degree \bar{d}_v for LDPC codes, there is no motivation to decrease the block length. In contrast, polar decoding becomes simpler when we reduce block sizes. Moreover, polar decoding is more efficient than minimum-complexity LDPC codes, that have the smallest possible degree of $\bar{d}_v = 2$, at short block lengths of $N < 256$. Since such ultra-sparse LDPC codes perform poorly in general, most typical irregular LDPC codes have an average degree between 3 and

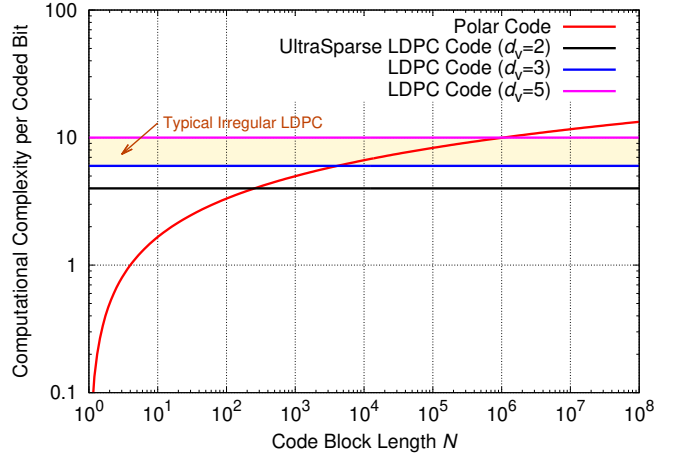


Fig. 3. Computational complexity per coded bit as a function of block length for polar SCL decoding (per list) and LDPC BP decoding (per iteration).

5. The LDPC codes used for 4 and 32 iterations in Fig. 2 have $\bar{d}_v = 4.00$ and 4.97 , respectively. Therefore, the decoding complexity of LDPC codes can be higher than polar decoding at short lengths up to $N < 65,000$ and 10^6 , respectively, for the LDPC codes having $\bar{d}_v = 4.00$ and 4.97 . This promotes polar codes as a good candidate for latency-critical systems.

Since we considered only VN processing without including any extra operations (e.g., sorting), reduced complexity techniques (e.g., [19], [45]), hardware architecture *etc.*, the actual computational complexity can vary greatly depending on hardware implementation. Nevertheless, the general trend described above holds at short block lengths, and most prototyping studies such as [17] have already revealed that the polar codes can compete favorably with LDPC codes in terms of complexity. Note that the complexity of LDPC decoding can be further increased when we include check node (CN) processing because the average CN degree is large for high-rate FEC, i.e., $\bar{d}_c = \bar{d}_v/(1-R) \gg 2$, while polar codes have only two incoming edges at CNs as shown in Fig. 4. On the other hand, polar SCL decoding is not very amenable to parallel implementation. Our irregular polar codes enable partially parallel computation as in [17], [33]. The feasibility of highly parallel polar decoding is discussed in [40], [41].

III. IRREGULAR POLAR CODES

A. Basic Concept of Irregular Polarization

In this section, we explain the basic idea of irregular polar codes, which is to inactivate some polarization units to obtain potential error-rate performance improvement in addition to complexity reduction. Consider a polarization stage of Fig. 4(a), which can be represented by a factor graph as shown in Fig. 4(b). The upper-branch CN of each polarization unit has three connected edges (thus degree-3). Each lower-branch CN is degree-2, and no computation is required to pass messages from right to left (or left to right). Our proposed polar codes have irregular pruning as shown in Fig. 4(c), where the top CN becomes degree-2 by disconnecting one of edges. By pruning the edge, we can reduce the encoding and decoding complexity

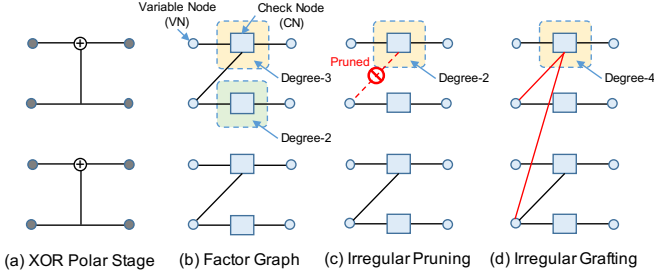


Fig. 4. Factor graph representation of polarization stage: (a) regular polar units, (b) corresponding factor graph, (c) irregular pruning with degree-2 check node, and (d) irregular polar with degree-4 check node.

without destroying the radix-2 Cooley–Tukey like factor graph (and thus maintain the basic SCL decoding structure that has complexity order $\mathcal{O}[N \log_2 N]$). Because the degree of upper-branch CNs are no longer all 3, we refer to the pruned polar codes as *irregular* polar codes, analogous to irregular LDPC codes. Many other types of irregular polar codes are possible, such as the example shown in Fig. 4(d) in which the top CN adds edges to be degree-4. Such irregularity (involving size-4 mixed-kernels [32]) requires major modification of SCL decoding and its computational complexity can be increased. Therefore, in this paper, we focus on irregular polar coding based on pruning (or inactivating) for complexity-constrained fiber-optic communications. We first consider the conventional Bhattacharyya parameter analysis [15] to discuss irregular polar coding, and will later propose a more sophisticated code construction method using EXIT [28] in the next section.

For an (N, k) polar code, there are $n = \log_2(N)$ polarization stages, each of which contains $N/2$ polarization units (thus $N_U \triangleq N \log_2(N)/2$ polarization units in total), as shown in Fig. 1 for $N = 16$. Polar codes exploit the so-called polarization phenomenon, where each polarization unit provides degraded and improved reliability. For example, when the coded bits x_1 and x_2 have uniform reliability having erasure rate (Bhattacharyya parameter) of 0.5, the upper branch of the polarization unit becomes unreliable with erasure rate of 0.75, whereas the lower branch improves the reliability to 0.25. Conventional polar codes have no flexibility in coding architecture except in the location of frozen bits. To increase the design degrees of freedom while reducing decoding complexity, our proposed irregular polar codes inactivate some of the polarization units. Fig. 1 illustrates an example with 10 inactivations among 32 polarization units. For convenience of analysis, we next consider an example in Fig. 5 for $N = 4$. We denote $\mathcal{U}(r, l)$ as the r th polarization unit from top to bottom in the l th polarization stage from left to right. For example, irregular polar codes in Figs. 5(b), (c), and (d) have an inactivated polarization unit at $\mathcal{U}(1, 1)$, $\mathcal{U}(2, 1)$, and $\mathcal{U}(1, 2)$, respectively. At an inactivated polarization unit, the exclusive-or (XOR) operation is removed.

The key benefits of such inactivations are three-fold: i) complexity reduction in encoding/decoding computations, ii) decoding latency reduction, and iii) potential performance improvement by adjusting the weight distribution. For the code in Fig. 1, we can achieve a $10/32 \simeq 31\%$ complexity

reduction since no computation is required for inactivated units in both encoding and decoding. Note that careful choice of polarization units to be inactivated can cause no performance penalty, as discussed below.

B. BEC Analysis

In this section, we specifically focus on the binary erasure channel (BEC) to exemplify potential benefits of irregular polar coding, and later discuss a design method for generic channels in the next section. The word error rate (WER) of the irregular polar codes for the BEC can be expressed via Bhattacharyya parameter evolution [15]. Starting from the channel side (n th polarization stage), the Bhattacharyya parameter $Z_i^{[n]}$ is given as

$$Z_i^{[n]} \triangleq \sum_{y_i \in \mathcal{Y}} \sqrt{W_i(y_i|0)W_i(y_i|1)} = \epsilon, \quad (3)$$

where ϵ is the erasure probability of the BEC. Let $Z_{r_U}^{[l]}$ and $Z_{r_L}^{[l]}$ be the incoming Bhattacharyya parameters, respectively, for upper and lower branches at polarization unit of $\mathcal{U}(r, l)$. If a polarization unit $\mathcal{U}(r, l)$ is inactivated, the bit reliabilities do not change:

$$Z_{r_U}^{[l-1]} = Z_{r_U}^{[l]}, \quad Z_{r_L}^{[l-1]} = Z_{r_L}^{[l]}. \quad (4)$$

On the other hand, for active polarization units, the evolution of the Bhattacharyya parameters is given as follows:

$$Z_{r_U}^{[l-1]} \leq Z_{r_U}^{[l]} + Z_{r_L}^{[l]} - Z_{r_U}^{[l]} Z_{r_L}^{[l]}, \quad Z_{r_L}^{[l-1]} = Z_{r_U}^{[l]} Z_{r_L}^{[l]}, \quad (5)$$

where the equality holds in the BEC. Note that the above evolution is identical to that of the original polar codes [15] if the incoming Bhattacharyya parameters at upper and lower branches of polarization units are equal, i.e., $Z_{r_U}^{[l]} = Z_{r_L}^{[l]}$. In order to account for possible non-uniformity of bit reliabilities, we shall use the modified evolution in (5).

In fact, non-uniform bit reliabilities occur in various situations, e.g., when polar coding is used for high-order modulation schemes [21], [35], and/or in frequency-selective channels [36], [37]. Moreover, our proposed polar codes with irregular inactivation can inherently produce non-uniform reliabilities even when channels are identical. This is illustrated in Fig. 5(d), where an inactivated polarization unit at $\mathcal{U}(1, 2)$ involves non-equal Bhattacharyya parameters for the first polarization stage for $\mathcal{U}(1, 1)$ and $\mathcal{U}(2, 1)$.

Given the left-most Bhattacharyya parameters $Z_i^{[0]}$, the WER P_ϵ of SC decoding is upper-bounded as follows [29]:

$$P_\epsilon \leq 1 - \prod_{i \in \mathbb{K}} \left(1 - \frac{1}{2} Z_i^{[0]}\right) \leq \frac{1}{2} \sum_{i \in \mathbb{K}} Z_i^{[0]}. \quad (6)$$

The last inequality is known as the UB, which can be loose at high erasure rates. In Fig. 5, we present the (tight) upper bound of WER and the Bhattacharyya parameters at every polarization unit for a channel erasure rate of $\epsilon = 0.5$. Note that the exact WER and BER can be easily derived for such short polar codes. The exact WER and BER of the conventional regular polar code in Fig. 5(a) are $P_\epsilon = \epsilon^2(1 - \epsilon^2/4) = 0.234375$ and $P_b = \epsilon^2(3 - \epsilon^2)/4 = 0.171875$, respectively. Whereas the irregular polar codes in Figs. 5(b), (c), and (d) perform no worse

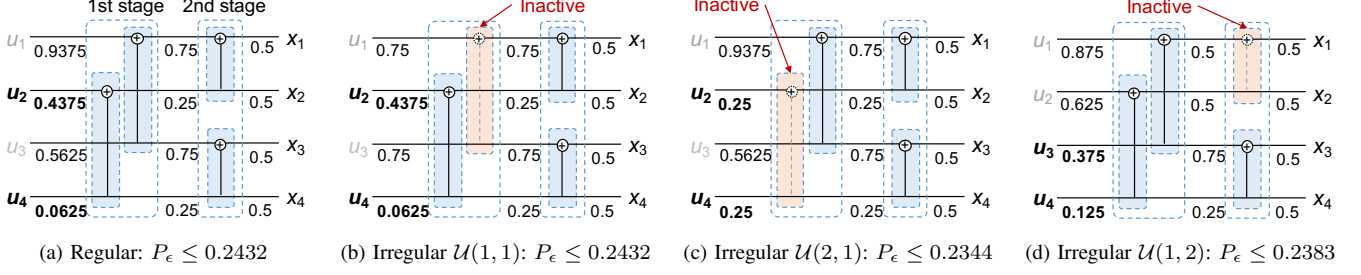


Fig. 5. Examples of inactivated polarization units for $(4, 2)$ polar codes with a corresponding WER of P_e at BEC channels with an erasure rate of $\epsilon = 0.5$.

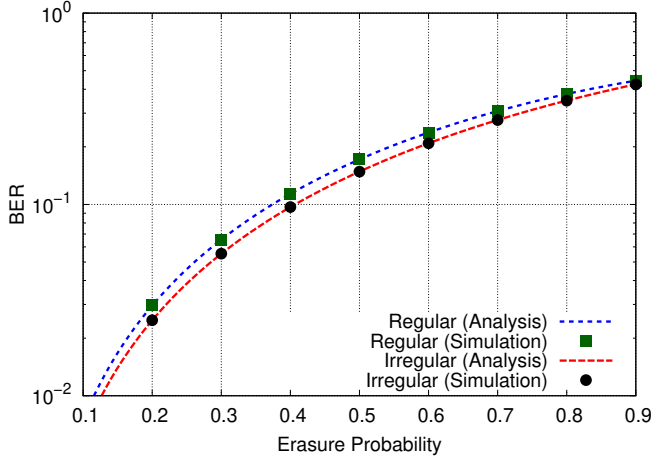


Fig. 6. BER analysis of regular and irregular polar codes at BEC channels.

than the regular polar code. For example, the code in Fig. 5(d) has the exact BER of $P_b = \epsilon^2(5 - \epsilon^2)/8 = 0.1484375$, which is better than the regular one. It is more clearly shown in Fig. 6, where simulated and analytical BER curves with SC decoding are plotted for those two polar codes, i.e., Fig. 5(a) vs. (d), at BEC channels. This indicates that with appropriate inactivation of polarization units, irregular polar codes can outperform regular polar codes. More importantly, since no computation is required for inactivated polarization units during encoding and decoding, the computational complexity of irregular codes can be significantly reduced without any performance penalty. Note that the systematic encoding procedure in [17] and SCL decoding in [18] can be applied without major modifications even in the presence of irregular polarization pruning.

C. Motivations

The above-mentioned benefits were partly discussed in analyses and theories derived for relaxed polar coding [33], which inactivates a series of polarization units when the incoming Bhattacharyya parameters are sufficiently good or bad. For example, the polarization unit at $\mathcal{U}(1, 1)$ in Fig. 5(b) already has poor incoming reliability ($Z_{1U}^{[1]} = Z_{1L}^{[1]} = 0.75$), and thus no performance degradation is incurred from inactivation as u_1 and u_3 will be chosen to be frozen bits. For another example in Fig. 5(c), the polarization unit $\mathcal{U}(2, 1)$ is inactivated since the second polarization stage already created good messages

($Z_{2U}^{[1]} = Z_{2L}^{[1]} = 0.25$). This irregular code achieves not only reduced complexity but also reduced latency without sacrificing the WER performance because partially parallel decoding is possible for u_2 and u_4 .

However, relaxed polar coding [33] has limited flexibility in inactivation patterns, preventing non-uniform reliabilities, in order to accommodate conventional code construction methods [15]. For example, relaxed polar coding excludes the case in Fig. 5(d), where the first polarization units $\mathcal{U}(1, 1)$ and $\mathcal{U}(2, 1)$ are still active even after inactivating the polarization unit at $\mathcal{U}(1, 2)$. This irregular code achieves improved performance while the complexity is reduced at the same time. Motivated by this benefit, in this paper, we generalize the concept in [33] by considering arbitrary irregular inactivation with the help of a modified code construction method, which takes non-uniform reliability into consideration, in order to further reduce computational complexity. Our design method facilitates efficient optimization of complicated kernel mixtures [32] in practice.

As discussed in [33], irregular polar codes with pruned polarization units are partly related to a simplified SC (SSC) decoding [17], [45], in which partially parallel decoding is enabled by pruning/merging tree branches, involving either frozen or information bits, to avoid unnecessary computation. However, our work directly modifies the degree distribution of the generator matrix to significantly reduce both encoding and decoding complexity without limitation of the tree structure, regardless of frozen bit locations. Therefore, our approach has the flexibility to incorporate any other methods including SSC to reduce the decoding complexity further.

IV. DESIGN METHOD FOR IRREGULAR POLAR CODING

A. EXIT Evolution with Non-Uniform Reliability

We now describe a design method to construct fully irregular polar codes, where there are a maximum of $N_U = N \log_2(N)/2$ polarization units that can be inactivated. We address how to accommodate the non-uniform bit reliability to design irregular polar codes by appropriately inactivating polarization units, that leads to the best performance and minimum complexity. As discussed in the previous section, the conventional construction methods [15], [20], [30], [31] assuming uniform bit reliability cannot be directly applied to optimize polar codes in the presence of irregular inactivation (and/or high-order modulations). Motivated by the fact that polar codes can be represented by a factor graph as shown in

TABLE II
JOINT INACTIVATED POLARIZATION AND FROZEN BIT LOCATION DESIGN

```

1: activate all polarization units
2: while  $N_{\text{inact}} \in \{1, 2, \dots, N_U\}$  do
3:   for all active polarization units do
4:     inactivate the target polarization unit
5:      $\mathcal{I}' = \text{UpdateMI}(\mathcal{I})$  according to (7)
6:     select frozen bits  $\mathbb{K}$  having the  $N - k$  smallest  $\mathcal{I}'$ 
7:     calculate the upper bound  $P_e$  according to (9)
8:     reactivate the target polarization unit
9:   end for
10:  inactivate the polarization unit having smallest  $P_e$ 
11: end while
12: return best frozen bit and inactivated polarization unit locations achieving the smallest  $P_e$ 

```

Fig. 4, we propose to introduce EXIT analysis [28] for tracking the non-uniform mutual information at every polarization unit.

Letting $\mathcal{I}_{r_U}^{[l]}$ and $\mathcal{I}_{r_L}^{[l]}$ denote mutual information of the upper and lower branches at the r th active polarization unit of l th stage, we use the following evolution:

$$\begin{aligned}\mathcal{I}_{r_U}^{[l-1]} &= 1 - J\left(\sqrt{[J^{-1}(1 - \mathcal{I}_{r_U}^{[l]})]^2 + [J^{-1}(1 - \mathcal{I}_{r_L}^{[l]})]^2}\right), \\ \mathcal{I}_{r_L}^{[l-1]} &= J\left(\sqrt{[J^{-1}(\mathcal{I}_{r_U}^{[l]})]^2 + [J^{-1}(\mathcal{I}_{r_L}^{[l]})]^2}\right),\end{aligned}\quad (7)$$

where $J(\cdot)$ is ten Brink's J-function [28], defined as

$$J(x) = 1 - \int_{-\infty}^{\infty} \frac{e^{-(t-x^2/2)^2/2x^2}}{\sqrt{2\pi x^2}} \log_2(1 + e^{-t}) dt, \quad (8)$$

and $J^{-1}(\cdot)$ is its inverse function. For inactive polarization units, mutual information is propagated without modification. Note that in [28] the closed-form approximations of the forward and inverse J-functions are derived through the Marquardt-Levenberg algorithm.

We perform the EXIT evolution in (7) iteratively to trace the mutual information of each stage until we obtain the output mutual information $\mathcal{I}_i^{[0]}$ for all $i \in \{1, 2, \dots, N\}$. Once the output mutual information $\mathcal{I}_i^{[0]}$ are obtained, the UB of error rate P_e is calculated as below:

$$P_e = \frac{1}{|\mathbb{K}|} \sum_{i \in \mathbb{K}} Q\left(\frac{1}{2} J^{-1}(\mathcal{I}_i^{[0]})\right), \quad (9)$$

where $Q(\cdot)$ is the Q-function defined as

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} \exp\left(-\frac{t^2}{2}\right) dt. \quad (10)$$

Note that the EXIT evolution in (7) assumed SC decoding, where extrinsic information at the $(l-1)$ th polarization stage are not propagated back to the l th polarization stage. Nonetheless, the EXIT evolution can be readily modified for polar BP decoding as well.

B. Joint Inactivation and Frozen Bit Design

We describe our proposed design method of irregular polar codes in Table II. Given an input mutual information array $\mathcal{I} = [I_1^{[n]}, I_2^{[n]}, \dots, I_N^{[n]}]$ for every bit of modulated symbols s_i in fiber-optic channels having an equivalent signal-to-noise ratio (SNR) ρ , the algorithm performs joint optimization of

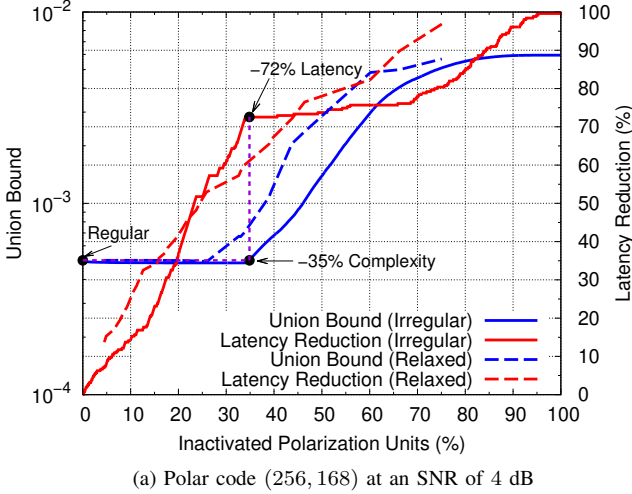
inactivation and frozen bit locations to minimize the UB for P_e . Since a brute-force search for all possible inactivations is unrealistic (i.e., the maximum search space for such a two-kernel mixture is scaled to $2^{N_U} = N^{N/2}$, which becomes more than 10^{308} even for a short block length of $N = 256$), we use a greedy method, which looks for the best polarization unit to be inactivated in a successive manner until we have inactivated N_{inact} polarization units.

For each candidate for inactivation, the mutual informations \mathcal{I} are updated to obtain the output mutual informations \mathcal{I}' according to evolution in (7). The frozen bit locations \mathbb{K} are decided by choosing the $N - k$ indices having the smallest values in \mathcal{I}' . For each iteration up to N_{inact} , the next polarization unit to be inactivated is successively decided by analyzing the UB for P_e . Through the iterations over successive inactivations, we finally select the best irregular polar code with the set of inactivated units as well as frozen bit locations, that result in the minimum P_e as the output of algorithm in Table II. Although the global optimum is not guaranteed, excellent performance is empirically observed using this greedy method. In [39], the authors also extended the greedy algorithm to jointly optimize the frozen bit locations and inactivations as well as the interleaver for wireless multi-antenna channels.

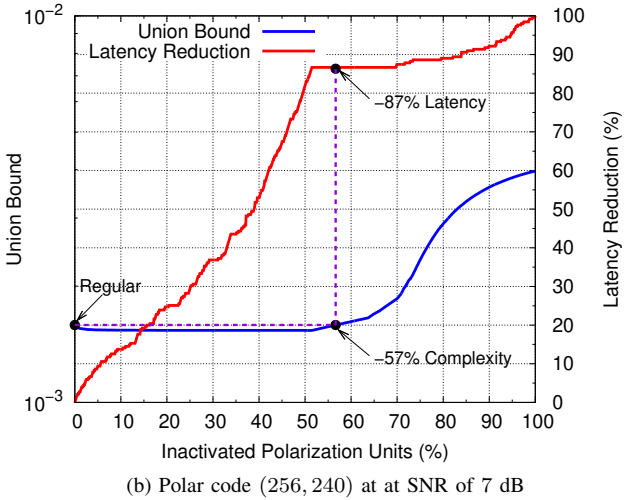
C. Union Bound and Latency Analysis

We now evaluate our greedy design method in Fig. 7, where the UB and decoding latency reduction versus the number of inactivated polarization units are shown for low-rate and high-rate polar codes (256, 168) and (256, 240). Note that the number of inactive units $N_{\text{inact}} = 0\%$ corresponds to the conventional (regular) polar codes, whereas $N_{\text{inact}} = 100\%$ corresponds to the uncoded case. Decoding complexity can be linearly decreased by pruning polarization units, whereas decoding latency is not always be reduced by pruning. More specifically, pruning the polarization units in the latter stages usually does not reduce latency unless the former stages in the decoding tree are all pruned as in Fig. 1. For example, the irregular pruning in Fig. 5(c) can reduce latency since decoding u_4 does not need to wait for decoding u_2 , whereas the pruning in Fig. 5(d) does not reduce latency.

It is confirmed in Fig. 7(a) that our proposed irregular polar codes can reduce computational complexity by at least 35% with a small improvement in the UB. The corresponding latency reduction can be 72%. More than 35% pruning of polarization units rapidly degrade the UB towards the uncoded BER. In this figure, we also present the UB and latency analysis of relaxed polar codes [33] for comparison. We can see that the relaxed polar codes degrade the UB if pruning exceeds 25%. This confirms that our irregular polar codes are more advantageous in reducing complexity and latency by exploiting higher degrees of design freedom in the selection of inactivations. It should be noted that the complexity and latency reduction can be more significant for higher code rates. Specifically, 57% complexity reduction and 87% latency reduction are achieved by (256, 240) irregular polar codes as shown in Fig. 7(b). This is because there are lower interactions between the frozen bits and information bits, resulting in less impact on the UB by the selection of polarization pruning.



(a) Polar code (256, 168) at an SNR of 4 dB



(b) Polar code (256, 240) at an SNR of 7 dB

Fig. 7. BER upper bound and latency reduction for irregular polar codes: (a) improved UB by up to 35% pruning for an overhead of 52.4% and (b) improved UB by up to 57% pruning for an overhead of 6.7%.

V. EXPERIMENTAL PERFORMANCE VALIDATION

A. Experimental Setup

For performance validation of our irregular polar codes, we use an experimental setup with dual-carrier 60 GBd DP-64QAM transmission [12], whose schematic is shown in Fig. 8. A digital root-raised cosine (RRC) filter with 0.1% roll-off was used for 64-QAM signals, and pre-emphasis was applied to compensate for the frequency response of the transmitter components (achieving approximately a 16.5 dB SNR ceiling). The signals were transmitted by four synchronized 92 GS/s digital-to-analog converters (DACs). The radio-frequency (RF) signals were then amplified by 30 GHz bandwidth modulator drivers. Two external cavity lasers (ECLs) with 100 kHz nominal linewidth and 62 GHz spacing were used as carriers for two independent I/Q modulators, with bandwidth of approximately 25 GHz. The combined attenuation of the DAC, RF drivers and modulators was approximately 7.3 dB at 30 GHz. The wavelength channels were combined, and polarization multiplexing emulated with an interferometer with 14.1 ns delay between arms. Spectrally shaped amplified

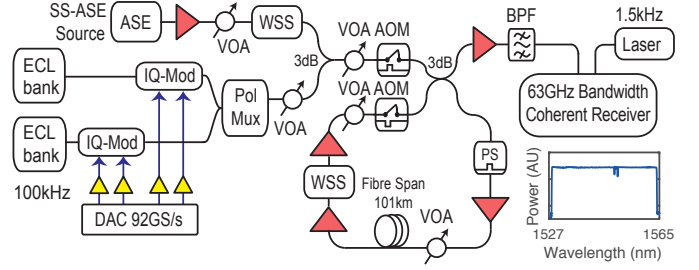


Fig. 8. Experimental setup [12] for dual-carrier 60GBd DP-64QAM. Full C-band loading is emulated by ASE [34].

spontaneous emission (SS-ASE) noise was used as an ultra-broadband source to emulate fully-loaded C-band transmission [34]. The SS-ASE source had a bandwidth of 4.5 THz and a notch centered at 1550 nm with a bandwidth of 140 GHz was carved out to accommodate the channel of interest (see inset Fig. 8). A 9 GHz guard-band on each side of the 122 GHz dual-carrier signal was used to minimize linear crosstalk [34].

For transmission, a recirculating loop was used, with a loop-synchronous polarization scrambler (PS), and a single span of 101.39 km of Vascade® EX2000 fiber with a total loss of 16.2 dB. The span was followed by an Erbium-doped fiber amplifier (EDFA) with 18 dBm output power and 5 dB noise figure to overcome fiber attenuation. A third ECL with 16 dBm output power and 100 kHz nominal linewidth was used as local oscillator in the optical receiver. Detection was carried out by balanced photodetectors with 70 GHz electrical bandwidth and without trans-impedance amplifiers. Finally, the received signals were captured by a real-time digital oscilloscope with an analog electrical bandwidth of 63 GHz at 160 GSa/s. The receiver digital signal processing (DSP) was of the same structure as that described in our previous work [12]. We used a training sequence followed by pilot-aided equalization [11] with a pilot overhead of 1.43%.

B. Back-to-Back and Transmission Performance

In our previous experiments [12], a 1 Tb/s transmission was demonstrated using a high-power LDPC code with very long block size of 52,800 bits at a code rate of 0.71. Because the scope of this paper is low-power and low-latency systems, we focus on relatively short FEC codes with $N \in \{256, 1024, 4096\}$. We compare irregular polar codes (having 33%, 37%, and 31% inactivations of polarization units, respectively, for $N = 256, 1024$, and 4096) with state-of-the-art Pareto-optimal LDPC codes [7] (whose analytical threshold at infinite length is within 0.5 dB of the Shannon limit) at a code rate of 0.625. Although the code rate of 0.625 is relatively low compared to typical lightwave systems, such a mid-rate FEC plays an important role in maximizing spectral efficiency for various modulation formats including DP-64QAM, as discussed in [44]. Random interleaving was carried out across all bit positions, polarizations and wavelength subchannels although a well-designed interleaver suited for polar codes can offer additional gain as investigated in [21], [35]–[37]. The LDPC decoder used the 32-layer LDA with $I = 8$ iterations, whereas the polar decoder used SCL decoding with CRC-8

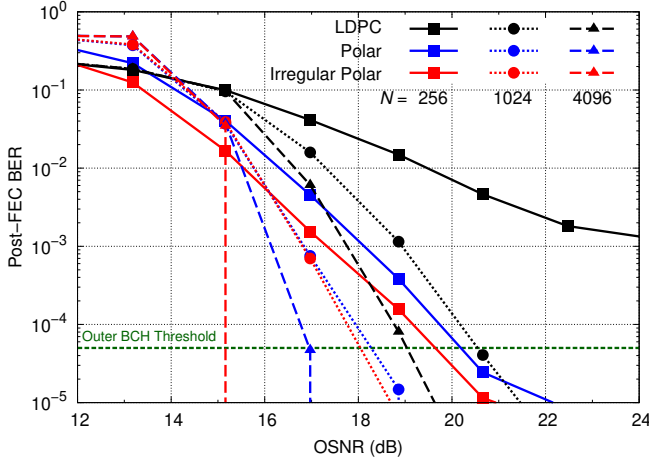


Fig. 9. Experimental BER in back-to-back configuration.

and a list size of $L = 32$. An outer hard-decision high-rate BCH code [12] (whose threshold is 5×10^{-5}) is assumed to realize a BER below 10^{-15} .

The back-to-back results are shown in Fig. 9. Note that when measurements with no bit errors are observed after decoding, those are plotted straight down towards the BER limit of 0. It was verified that polar codes outperform LDPC codes, since LDPC codes do not perform well for short block lengths and a limited number of iterations [35]. Our irregular polar codes can further improve the BER performance against regular polar codes while reducing decoding complexity and latency significantly. Specifically, measurements at an optical signal-to-noise ratio (OSNR) above 17 dB resulted in no errors with irregular polar codes of $N = 4096$ over the approximately 3×10^6 bits in a single capture, whereas errors were observed for regular polar codes.

The results after transmission over 300 km and 400 km are shown in Figs. 10 and 11, respectively. The experimental results demonstrate that the system margin to realize the BER underneath the BCH threshold can be significantly improved by polar codes versus LDPC codes. In particular for 400 km, the Pareto-optimal LDPC code cannot reach the target BER below the BCH threshold even for the block length of $N = 4096$ bits, whereas the polar codes achieved the target at a shorter block length of $N = 1024$. Although the system margin was not significantly improved by irregular polar coding compared to the regular counterparts, the remarkable reduction in decoding complexity and latency is of great advantage for high-throughput fiber-optic communications.

C. Burst Error Analysis

We assumed that an outer BCH code at a very high rate and with sufficiently long interleaving is employed. However, shorter interleaver depth is preferred in practice. Moreover, the error behavior of polar codes is less well-understood to date, compared with LDPC codes. Due to the nature of SCL decoding, potential error propagation may cause long burst errors. This section analyzes the burst error behavior.

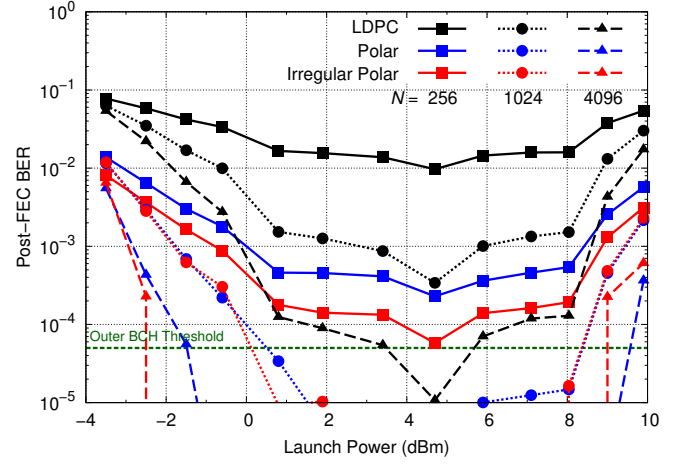


Fig. 10. Experimental BER in 300 km transmission.

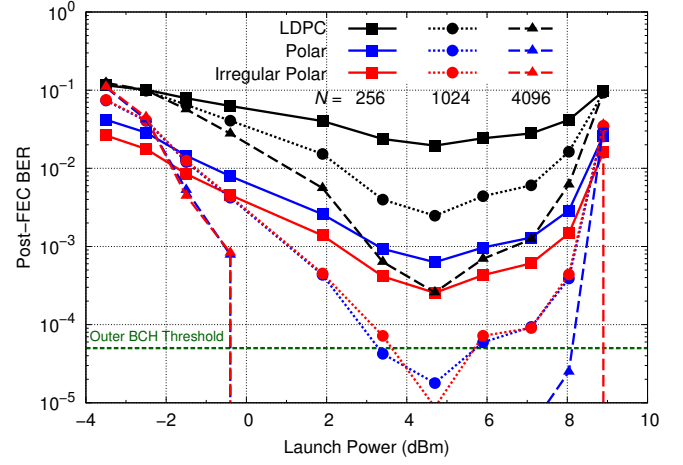


Fig. 11. Experimental BER in 400 km transmission.

In [23], [46], the Hamming weight spectrum of polar codes was analyzed by using SCL decoding with a very large list size. Fig. 12 illustrates the weight spectrum of regular and irregular polar+CRC codes (256, 160) via the same method. It was found that the regular codes have relatively larger distance compared to the irregular codes. As shown in experiments, this does not immediately mean that irregular polar codes perform worse because the Hamming distance larger than necessary does not always contribute to improving BER performance.

Fig. 13 shows empirical density of the number of bits in error per erroneous block for the regular and irregular polar+CRC codes. We consider SCL decoding with $L = 1$ and 32 for systematic and non-systematic codes. The SNR is adjusted for different list sizes so that the BER is roughly comparable to the BCH threshold of 5×10^{-5} . From Fig. 13, we observe the following points:

- Systematic codes have a concentrated error distribution, whereas non-systematic codes have a broad error distribution. Similar results were well-explained in [46].
- For systematic codes, irregularity does not cause degradation of the burst error distribution.
- For non-systematic codes, irregular codes have slightly

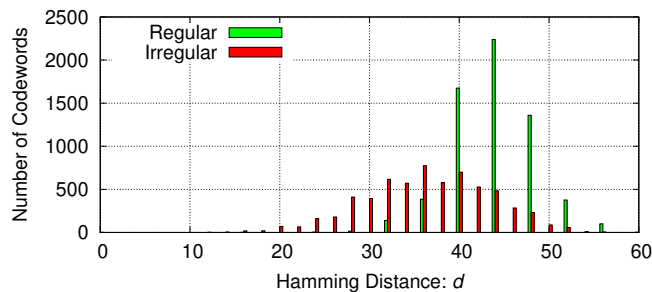


Fig. 12. Hamming weight spectrum for regular and irregular polar+CRC codes (256, 160), analyzed by the method in [23], [46] ($L = 1.6 \times 10^6$).

less-disperse distribution.

- Larger list sizes result in a more dispersed error distribution. This may be because a larger list can improve BER sufficiently under noisier channel conditions.

These results suggest that we should use systematic codes to constrain burst errors, not only to reduce the BER by half, that is known in [16]. The limited lengths of bit errors can relax the requirement of interleaver depth for outer BCH codes. We leave detailed analysis of practical interleaver and outer BCH design to remove burst errors as future work.

VI. CONCLUSIONS

We proposed a novel class of polar codes called irregular polar codes, which inactivate some of the polarization units. With careful inactivation of polarization units, the computational complexity for encoding and decoding can be significantly reduced, while the BER performance is slightly improved. We introduced an EXIT-based greedy code construction method addressing non-uniform bit reliability to jointly optimize inactivation and frozen bit locations. With the proposed design method, greater than 30% complexity reduction and 70% latency reduction were achieved without degrading BER performance. We experimentally demonstrated that irregular polar codes can outperform conventional regular polar codes and state-of-the-art LDPC codes, when codeword lengths and decoding complexity are constrained for high-speed optical communications. Note that the heuristic design method proposed in this paper is applicable for optimizing general mixed-kernel polar codes.

ACKNOWLEDGMENT

We would like to thank all of the researchers involved in the experiments, and Corning® for supplying the fiber.

REFERENCES

- [1] I. B. Djordjevic, "Advanced coded-modulation for ultra-high-speed optical transmission," in *Proc. Opt. Fiber Commun.*, San Francisco, CA, Mar. 2014, Paper W3J-4.
- [2] K. Sugihara, Y. Miyata, T. Sugihara, K. Kubo, H. Yoshida, W. Matsumoto, and T. Mizuochoi, "A spatially-coupled type LDPC code with an NCG of 12 dB for optical transmission beyond 100 Gb/s," in *Proc. Opt. Fiber Commun.*, Anaheim, CA, Mar. 2013, Paper OM2B.4.
- [3] L. Schmalen, V. Aref, J. Cho, D. Suikat, D. Rösener, and A. Leven, "Spatially coupled soft-decision error correction for future lightwave systems," *J. Lightw. Technol.*, vol. 33, no. 5, pp. 1109–1116, Mar. 2015.
- [4] F. Buchali, L. Schmalen, A. Klekamp, K. Schuh, and A. Leven, "5 × 50 Gb/s WDM transmission of 32 Gbaud DP-3-PSK over 36,000 km fiber with spatially coupled LDPC coding," in *Proc. Opt. Fiber Commun.*, San Francisco, CA, Mar. 2014, Paper W1A-1.
- [5] 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.
- [6] D. Chang, F. Yu, Z. Xiao, N. Stojanovic, F. N. Hauske, Y. Cai, C. Xie, L. Li, X. Xu, and Q. Xiong, "LDPC convolutional codes using layered decoding algorithm for high speed coherent optical transmission," in *Proc. Opt. Fiber Commun.*, Los Angeles, CA, Mar. 2012, Paper OW1H-4.
- [7] T. Koike-Akino, D. S. Millar, K. Kojima, K. Parsons, Y. Miyata, K. Sugihara, and W. Matsumoto, "Iteration-aware LDPC code design for low-power optical communications," *J. Lightw. Technol.*, vol. 34, no. 2, pp. 573–581, Jan. 2016.
- [8] B. Smith, M. Ardakani, W. Yu, and F. R. Kschischang, "Design of irregular LDPC codes with optimized performance-complexity tradeoff," *IEEE Trans. Commun.*, vol. 58, no. 2, pp. 489–499, Feb. 2010.
- [9] R. Rios-Müller, J. Renaudier, P. Brindel, H. Mardoyan, P. Jennevé, L. Schmalen, and G. Charlet, "1-terabit/s net data-rate transceiver based on single-carrier Nyquist-shaped 124 Gbaud PDM-32QAM," in *Proc. Opt. Fiber Commun.*, Los Angeles, CA, Mar. 2015, Paper Th5B1.
- [10] T. Richter, C. Schmidt-Langhorst, R. Elschner, L. Molle, S. Alreesh, T. Kato, T. Tanimura, S. Watanabe, J. K. Fischer, and C. Schubert, "Distributed 1-Tb/s all-optical aggregation capacity in 125-GHz optical bandwidth by frequency conversion in fiber," in *Proc. Eur. Conf. Opt. Commun.*, London, UK, Sept. 2013, Paper PDP.2.5.
- [11] D. S. Millar, R. Maher, D. Lavery, T. Koike-Akino, M. Pajovic, A. Alvarado, W. Paskov, K. Kojima, K. Parsons, B. C. Thomsen, S. J. Savory, and P. Bayvel, "Design of a 1 Tb/s superchannel coherent receiver," *J. Lightw. Technol.*, vol. 34, no. 6, pp. 1453–1463, Mar. 2016.
- [12] D. S. Millar, L. Galdino, R. Maher, M. Pajovic, T. Koike-Akino, G. Saavedra, D. J. Elson, D. Lavery, K. Shi, M. S. Erkilinc, and E. Sillekens, "A simplified dual-carrier DP-64QAM 1 Tb/s transceiver," in *Proc. Opt. Fiber Commun.*, Los Angeles, CA, Mar. 2017, Paper M3D-2.
- [13] R. Maher, A. Alvarado, D. Lavery, and P. Bayvel, "Increasing the information rates of optical communications via coded modulation: A study of transceiver performance," *Nat. Sci. Rep.*, vol. 6, 21278, Feb. 2016.
- [14] G. Raybon, J. Cho, A. Adamiecki, P. Winzer, A. Konczykowska, F. Jorge, J. Y. Dupuy, M. Riet, B. Duval, K. Kim, and S. Randel, "Single carrier high symbol rate transmitter for data rates up to 1.0 Tb/s," in *Proc. Opt. Fiber Commun.*, Anaheim, CA, Mar. 2016, Paper Th3A2.
- [15] E. Arkan, "Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels," *IEEE Trans. Inform. Theory*, vol. 55, no. 7, pp. 3051–3073, July 2009.
- [16] E. Arkan, "Systematic polar coding," *IEEE Commun. Lett.*, vol. 15, no. 8, pp. 860–862, Aug. 2011.
- [17] G. Sarkis, I. Tal, P. Giard, A. Vardy, C. Thibault, and W. J. Gross, "Flexible and low-complexity encoding and decoding of systematic polar codes," *IEEE Trans. Commun.*, vol. 64, no. 7, pp. 2732–2745, July 2016.
- [18] I. Tal and A. Vardy, "List decoding of polar codes," *IEEE Trans. Inform. Theory*, vol. 61, no. 5, pp. 2213–2226, May 2015.
- [19] B. Li, H. Shen, and D. Tse, "An adaptive successive cancellation list decoder for polar codes with cyclic redundancy check," *IEEE Commun. Lett.*, vol. 16, no. 12, pp. 2044–2047, Dec. 2012.
- [20] I. Tal and A. Vardy, "How to construct polar codes," *IEEE Trans. Inform. Theory*, vol. 59, no. 10, pp. 6562–6582, Oct. 2013.
- [21] D. M. Shin, S. C. Lim, and K. Yang, "Mapping selection and code construction for 2^m -ary polar-coded modulation," *IEEE Commun. Lett.*, vol. 16, no. 6, pp. 905–908, June 2012.
- [22] M. Seidl, A. Schenk, C. Stierstorfer, and J. B. Huber, "Polar-coded modulation," *IEEE Trans. Commun.*, vol. 61, no. 10, pp. 1302–1306, Oct. 2013.
- [23] Q. Zhang, A. Liu, X. Pan, and K. Pan, "CRC code design for list decoding of polar codes," *IEEE Commun. Lett.*, vol. 21, no. 6, pp. 1229–1232, June 2017.
- [24] R. M. Pyndiah, "Near-optimum decoding of product codes: Block turbo codes," *IEEE Trans. Commun.*, vol. 46, no. 8, pp. 1003–1010, Aug. 1998.
- [25] S. Dave, L. Escher, F. Mo, W. Thesling, J. Keszenheimer, and R. Fuerst, "Soft-decision forward error correction in a 40-nm ASIC for 100-Gbps OTN applications," in *Proc. Opt. Fiber Commun.*, Los Angeles, Mar. 2011, Paper JWA014.
- [26] M. Ferrari and S. Bellini, "Importance sampling simulation of turbo product codes," in *Proc. IEEE Int. Conf. Commun.*, Helsinki, Finland, June 2001.

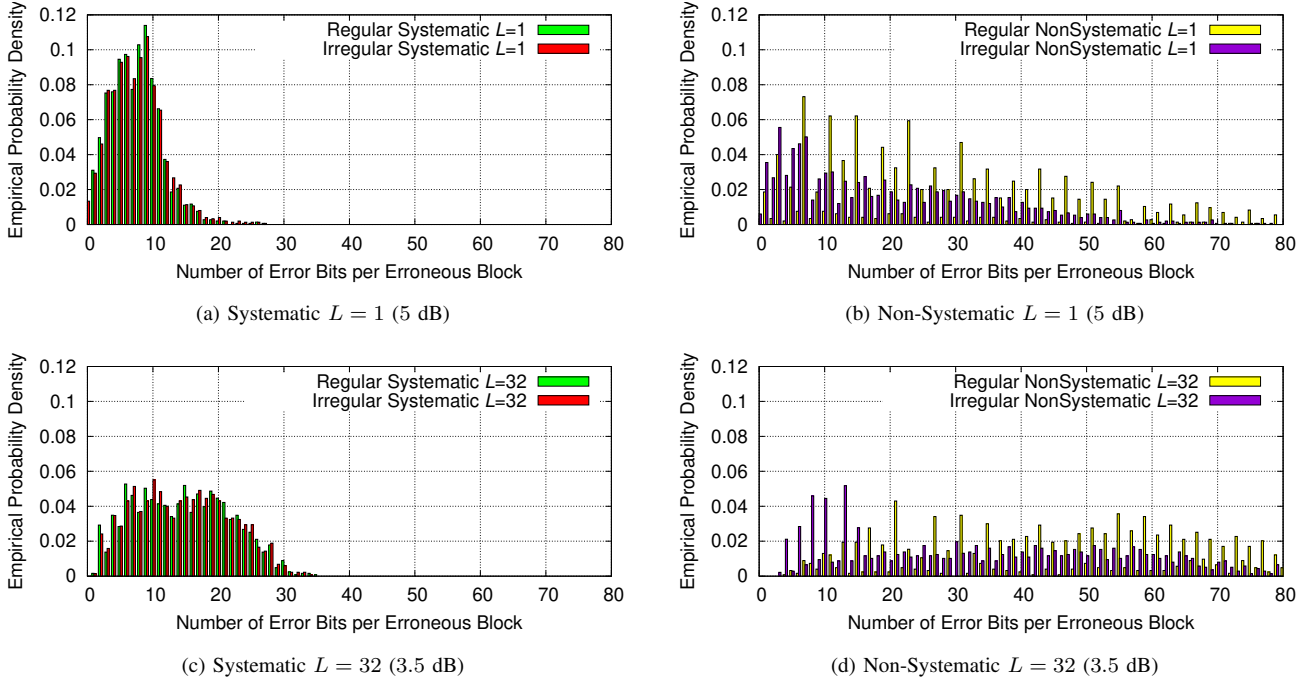


Fig. 13. Empirical probability density of the number of error bits inside erroneous blocks for polar+CRC codes (256, 160).

- [27] O. Y. Takeshita, "On maximum contention-free interleavers and permutation polynomials over integer rings," *IEEE Trans. Inform. Theory*, vol. 52, no. 3, pp. 1249–1253, Mar. 2006.
- [28] S. ten Brink, G. Kramer, and A. Ashikhmin, "Design of low-density parity-check codes for modulation and detection," *IEEE Trans. Commun.*, vol. 52, no. 4, pp. 670–678, Apr. 2004.
- [29] R. Mori and T. Tanaka, "Performance and construction of polar codes on symmetric binary-input memoryless channels," in *Proc. IEEE Int. Sym. Inform. Theory*, pp. 1496–1500, Seoul, South Korea, June 2009.
- [30] R. Mori and T. Tanaka, "Performance of polar codes with the construction using density evolution," *IEEE Comm. Lett.*, vol. 13, no. 7, pp. 519–521, July 2009.
- [31] P. Trifonov, "Efficient design and decoding of polar codes," *IEEE Trans. Commun.*, vol. 60, no. 11, pp. 3221–3227, Nov. 2012.
- [32] N. Presman, O. Shapira, and S. Litsyn, "Polar codes with mixed kernels," in *Proc. IEEE Int. Sym. Inform. Theory*, pp. 6–10, St. Petersburg, Russia, July 2011.
- [33] M. El-Khamy, H. Mahadavifar, G. Feygin, J. Lee, and I. Kang, "Relaxed polar codes," *IEEE Trans. Inform. Theory*, vol. 63, no. 4, pp. 1986–2000, Apr. 2017.
- [34] D. J. Elson, L. Galdino, R. Maher, R. I. Killey, B. C. Thomsen, and P. Bayvel, "High spectral density transmission emulation using amplified spontaneous emission noise," *Opt. Lett.*, vol. 41, no. 1, pp. 68–71, Jan. 2016.
- [35] T. Koike-Akino, Y. Wang, S. C. Draper, K. Sugihara, W. Matsumoto, D. S. Millar, K. Parsons, and K. Kojima, "Bit-interleaved polar-coded modulation for low-latency short-block transmission," in *Proc. Opt. Fiber Commun.*, Los Angeles, CA, Mar. 2017, Paper W1J.6.
- [36] T. Koike-Akino, Y. Wang, S. C. Draper, K. Sugihara, and W. Matsumoto, "Bit-interleaved polar-coded OFDM for low-latency M2M wireless communications," in *Proc. IEEE Int. Conf. Commun.*, Paris, France, May 2017.
- [37] T. Koike-Akino, W. Wang, and S. C. Draper, "Bit-interleaved polar-coded OFDM transmission for machine-to-machine wireless networks," *invited submission to Elsevier Ad Hoc Networks*, Oct. 2017.
- [38] T. Koike-Akino, C. Cao, Y. Wang, S. C. Draper, D. S. Millar, K. Parsons, K. Kojima, M. Pajovic, L. Galdino, D. J. Elson, D. Lavery, and P. Bayvel, "Irregular polar coding for multi-level modulation in complexity-constrained lightwave systems," in *Proc. Eur. Conf. Opt. Commun.*, Gothenburg, Sweden, Sept. 2017, Paper M.1.D.3.
- [39] C. Cao, T. Koike-Akino, Y. Wang, and S. C. Draper, "Irregular polar coding for massive MIMO," in *Proc. IEEE Global Commun. Conf.*, Singapore, Dec. 2017.
- [40] 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," to appear in *Proc. Opt. Fiber Commun.*, San Diego, CA, Mar. 2018.
- [41] T. Koike-Akino, C. Cao, and Y. Wang, "Turbo product codes with irregular polar coding for high-throughput parallel decoding in wireless OFDM transmission," to appear in *Proc. IEEE Int. Conf. Commun.*, Kansas City, MO, May 2018.
- [42] M. Ebada, A. Elkelesh, S. Cammerer, and S. ten Brink, "Scattered EXIT charts for finite length LDPC code design," arXiv preprint arXiv:1706.09239, June 2017.
- [43] G. Liva, L. Gaudio, T. Ninacs, and T. Jerkovits, "Code design for short blocks: A survey," arXiv preprint arXiv:1610.00873, 2016.
- [44] T. Koike-Akino, K. Kojima, D. S. Millar, K. Parsons, T. Yoshida, and T. Sugihara, "Pareto optimization of adaptive modulation and coding set in nonlinear fiber-optic systems," *J. Lightw. Technol.*, vol. 35, no. 4, pp. 1041–1049, Apr. 2017.
- [45] A. Alamdar-Yazdi and F. R. Kschischang, "A simplified successive-cancellation decoder for polar codes," *IEEE Commun. Lett.*, vol. 15, no. 12, pp. 1378–1380, Oct. 2011.
- [46] Z. Liu, K. Chen, K. Niu, and Z. He, "Distance spectrum analysis of polar codes," in *Proc. Wireless Commun. Netw. Conf.*, pp. 490–495, Istanbul, Turkey, Apr. 2014.