MITSUBISHI ELECTRIC RESEARCH LABORATORIES
http://www.merl.com

Iteration-Aware LDPC Code Design for Low-Power Optical Communications

Koike-Akino, T.; Millar, D.S.; Kojima, K.; Parsons, K.; Miyata, Y.; Sugihara, K.; Matsumoto, W.

TR2016-028 January 2016

Abstract

Recent low-density parity-check (LDPC) codes have shown capacity-approaching performance for various communications systems. However, their promising performance cannot always be obtained due to practical constraints such as finite codeword length, finite iteration, finite memory, and finite precision. In this paper, we focus on a practical design method of highperformance LDPC codes under a constraint of finite-iteration decoding for low-power optical communications. We introduce an iteration-aware LDPC code design approach, which is based on decoding trajectory in extrinsic information transfer (EXIT) chart analysis. It is demonstrated that an LDPC code designed by the conventional curve-fitting method exhibits nearly 2 dB of penalty when the maximum number of iterations is limited. The results suggest that the LDPC code should be adaptively changed, e.g., when the number of decoding iterations is decreased to save power consumption.We also extend our design method to a multiobjective optimization concept by taking average degrees into account, so that the threshold and the computational complexity are minimized at the same time. The proposed Pareto-optimal codes can achieve additional 2 dB gain or 50% complexity reduction at maximum, in low-power decoding scenarios.

IEEE/OSA Journal of Lightwave Technology

Iteration-Aware LDPC Code Design for Low-Power Optical Communications

Toshiaki Koike-Akino, *Senior Member, IEEE*, David S. Millar, *Member, IEEE, Member, OSA*, Keisuke Kojima, *Senior Member, IEEE, Fellow, OSA*, Kieran Parsons, *Senior Member, IEEE, Member, OSA*, Yoshikuni Miyata, Kenya Sugihara, and Wataru Matsumoto

(Invited Paper)

***Abstract*—Recent low-density parity-check (LDPC) codes have shown capacity-approaching performance for various communications systems. However, their promising performance cannot always be obtained due to practical constraints such as finite codeword length, finite iteration, finite memory, and finite precision. In this paper, we focus on a practical design method of high-performance LDPC codes under a constraint of finite-iteration decoding for low-power optical communications. We introduce an iteration-aware LDPC code design approach, which is based on decoding trajectory in extrinsic information transfer (EXIT) chart analysis. It is demonstrated that an LDPC code designed by the conventional curve-fitting method exhibits nearly 2 dB of penalty when the maximum number of iterations is limited. The results suggest that the LDPC code should be adaptively changed, e.g., when the number of decoding iterations is decreased to save power consumption. We also extend our design method to a multi-objective optimization concept by taking average degrees into account, so that the threshold and the computational complexity are minimized at the same time. The proposed Pareto-optimal codes can achieve additional 2 dB gain or 50% complexity reduction at maximum, in low-power decoding scenarios.**

***Index Terms*—LDPC codes, coded modulation, EXIT chart, BICM, limited number of BP iterations, Pareto optimum**

I. Introduction

Forward error correction (FEC) codes based on low-density parity-check (LDPC) codes [1]–[6] have realized capacity-achieving performance. For example, an optimized irregular LDPC code reported in [2] already achieved the Shannon limit within 0.04 dB (analytical threshold is within 0.0045 dB). However, such an excellent performance is possible only with high-power processing because it considers a large number of iterations of 2000, a high maximum variable degree of 200, a long codeword length of 10^7, and a high precision of 9-bit quantization. Since transceivers for optical communications call for high-speed operations to accommodate tens/hundreds of Gb/s or even beyond Tb/s, lower-power processing has been of great importance with a limitation in decoding iteration, computational complexity, memory size, latency, and precision.

Manuscript received June 1, 2015; revised Aug. 14, 2015; accepted Sep. 3, 2015.

T. Koike-Akino, 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; millar@merl.com; kojima@merl.com; parsons@merl.com).

Y. Miyata, K. Sugihara, and W. Matsumoto are with Information Technology R&D Center (ITC), Mitsubishi Electric Corporation (MELCO), 5-1-1, Ofuna, Kamakura, Kanagawa 247-8501, Japan (e-mail: miyata.yoshikuni@ak.mitsubishielectric.co.jp; sugihara.kenya@dx.mitsubishielectric.co.jp; matsumoto.wataru@aj.mitsubishielectric.co.jp).

Finite-length code design has been one of the most challenging problems [7]–[9] for memory- and latency-constrained systems. Finite geometry (FG) [10]–[12], nonbinary (NB) [13]–[15], and generalized LDPC (GLDPC) codes [16]–[18] have shown relatively good performance for short codeword lengths. More recently, there has been growing interest in spatially-coupled codes [19]–[21], including LDPC convolutional codes [22]–[28], because low-memory/low-latency window decoding [22] is available. Regarding finite-precision decoding, a quantized version of density evolution (DE) has been used to design LDPC codes, e.g., in [2], [3], [29].

In this paper, we focus on the limited number of iterations for low-power decoding. To optimize LDPC codes for finite-iteration belief-propagation (BP) decoding, we use decoding trajectory in extrinsic information transfer (EXIT) chart [4], [30]. We verify that a significant benefit up to 2 dB can be achieved by re-designing LDPC codes according to the EXIT trajectory for different number of decoding iterations. Although the decoding trajectory across iterations has been already addressed in [3], [4], and a related design has been studied for finite-iteration window decoding in [24], it is the first demonstration of the remarkable advantage provided by the iteration-aware code optimization for finite-iteration BP decoding, to the best of our knowledge. The results suggest that different LDPC codes should be assigned when the number of decoding iterations is decreased to reduce power consumption. Nevertheless, one may often keep using one LDPC code irrespective of the number of iterations, to adjust only power consumption, e.g., in [31]. However, it should be noted that a high penalty up to 2 dB can be imposed without using fully-optimized LDPC codes depending on the limited number of iterations.

The key contributions of this paper are summarized below:

- We introduce a practical design method of LDPC codes, based on the EXIT trajectory under a limited number of iterations for low-power decoding.
- We show a significant gain by at most 2 dB with the iteration-aware LDPC code design method, compared to conventional design approach.
- From our preliminary work in [32], we extend the iteration-aware design method to a novel multi-objective optimization concept, which jointly minimizes the threshold and the computational complexity.

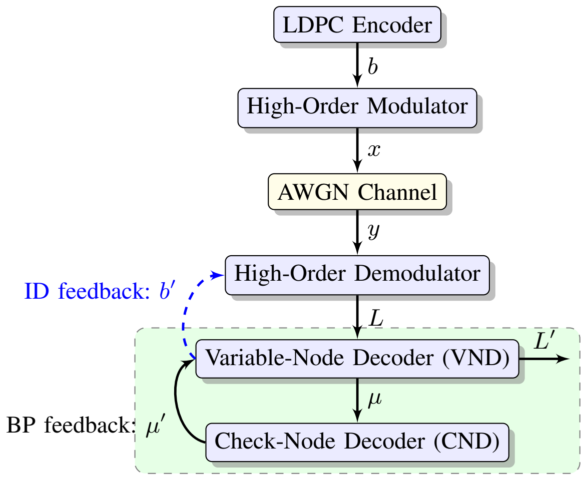

Fig. 1. Schematic of LDPC-coded modulation system for BICM.

- It is demonstrated that our Pareto-optimal LDPC codes can achieve additional 2 dB gain or 50% complexity reduction at maximum.

II. LDPC-Coded Modulation

In this section, we present the LDPC-coded modulation systems and basic concepts underlying the EXIT chart analysis.

A. System Description

Fig. 1 shows a schematic of LDPC-coded modulation systems under consideration. Here, we simplify our coded modulation design procedure under an additive white Gaussian noise (AWGN) channel, which considers the combined effects of amplified spontaneous emission (ASE) noise and nonlinear interference (NLI) over fiber-optic transmission as suggested by a Gaussian noise (GN) model [33]–[36] when dispersion and phase distortion are appropriately equalized.

In the communication system, an LDPC encoder generates a codeword bit sequence $\{b_n\}$. The LDPC codes can be represented by a sparse Tanner graph, according to a parity-check matrix $\boldsymbol{H} \in \mathbb{F}_q^{M\times N}$, where M and N are the numbers of check and variable nodes, respectively ($\mathbb{F}_q$ is a finite field having q distinct elements). The length-N encoded bits $\{b_n\}$ are sequentially mapped to symbol constellations with a Q-ary modulation scheme, e.g., dual-polarization quadrature phase-shift keying (DP-QPSK). The modulated symbols $\{x_n\}$ are transmitted to a receiver over the effective AWGN channel. The received symbols $\{y_n\}$ are demodulated to compute log-likelihood ratio (LLR) messages $\{L_n\}$. The LLR is defined for binary LDPC codes (i.e., $q = 2$) as follows:

$$L_n = \ln \frac{\Pr(y_n|b_n = 1)}{\Pr(y_n|b_n = 0)}, \tag{1}$$

which reduces to $L_n = 2y_n/\sigma_0^2$ for a binary PSK (BPSK) transmission case with a noise variance of σ_0^2.

The LLR messages $\{L_n\}$ computed by the demodulator are fed into an LDPC decoder, which consists of variable-node decoders (VND) and check-node decoders (CND). At VND, belief messages $\{\mu_n\}$ are updated in parallel, given the LLR messages $\{L_n\}$ and BP feedback $\{\mu'_n\}$ from CND. The CND updates the belief messages $\{\mu'_n\}$ from the VND messages $\{\mu_n\}$. The iterative BP decoding can effectively correct potential errors in the original LLR messages $\{L_n\}$. To improve performance, soft-decision messages $\{b'_n\}$ for iterative demodulation (ID) can be fed back from the LDPC decoder to the demodulator to refine the LLR messages.

B. Coded Modulation

Coded modulation is aimed at achieving the channel capacity $C_{\mathsf{ch}} = \sup \mathcal{I}(x; y)$; more specifically, a maximum of mutual information (MI) $\mathcal{I}(x; y)$ between channel input x and output y over all possible signal distribution $\Pr(x)$. For AWGN channels, the capacity is given as $C_{\mathsf{ch}} = \log_2(1 + 1/\sigma_0^2)$ b/s/Hz/pol for a signal-to-noise ratio (SNR) of $1/\sigma_0^2$. However, the achievable rate of coded modulation R_{cm} is bounded by the MI between encoding bit b and channel output y, i.e., $R_{\mathsf{cm}} \leq \mathcal{I}(b; y)$. There have existed various coded modulation schemes as follows:

- **Trellis-coded modulation (TCM)** [37], [40]: joint design of convolutional coding and modulation format via set partitioning *etc*. Concatenation with turbo decoding may be needed to approach capacity. Computational complexity increases exponentially with constraint length.
- **Multi-level coding (MLC)** [41]: proper rate control with a chain rule of MI for multi-level encoders and successive decoding are needed to approach capacity. It theoretically achieves coded modulation bound: $R_{\mathsf{mlc}} \leq \mathcal{I}(b; y)$.
- **Bit-interleaved coded modulation (BICM)** [42]: treating binary input to the modulator and LLR output from the demodulator as an effective channel. No decoding feedback is required, and any modulation format can be uniformly handled. The achievable rate is bounded by the so-called generalized MI (GMI): $R_{\mathsf{bicm}} \leq \mathcal{I}(b; L)$.
- **BICM with iterative demodulation (BICM-ID)** [32], [43]–[45]: better performance than BICM with soft-decision feedback from the decoder to the demodulator. The achievable rate is bounded by the averaged MI conditioned on feedback: $R_{\mathsf{bicmid}} \leq \mathbb{E}_{b'}[\mathcal{I}(b; L|b')]$, where $\mathbb{E}[\cdot]$ denotes the expectation.
- **Nonbinary-input coded modulation (NBICM)** [14], [15], [27], [38], [39]: no feedback is required when the Galois field size q matches the modulation order Q, while more complicated nonbinary decoding is needed. It achieves coded modulation bound.

In order to generate a Gaussian-like distribution for $\Pr(x)$ to approach the AWGN channel capacity, probabilistic shaping [38] and geometric shaping [46] have been investigated.

Since BICM has been widely used in practice, we focus on BICM in this paper. Note that the methodology described for BICM can be adopted for different coded modulation schemes in a straightforward manner.

C. BP Decoding

For BP decoding based on log-domain sum-product algorithm (SPA), the message passing between VND and CND in the Tanner graph is carried out as follows:

- The v-th variable node to c-th check node:
$$\mu_{v\to c} = L_v + \sum_{c' \in \mathcal{M}(v)\backslash c} \mu'_{c'\to v}, \tag{2}$$
where $\mathcal{M}(v)$ denotes the set of check nodes connected to the v-th variable node, and $\backslash c$ denotes exclusion of c from the set. The size of the set $d_v = |\mathcal{M}(v)|$ is called the variable-node degree, which is the number of edges connecting to the v-th variable node in the Tanner graph.
- The c-th check node to v-th variable node:
$$\mu'_{c\to v} = 2\tanh^{-1}\Big(\prod_{v' \in \mathcal{N}(c)\backslash v} \tanh\frac{\mu_{v'\to c}}{2}\Big), \tag{3}$$
where $\mathcal{N}(c)$ denotes the set of variable nodes connected to the c-th check node. The size of the set $d_c = |\mathcal{N}(c)|$ is called the check-node degree, i.e., the number of edges connecting to the c-th check node in the Tanner graph.
- The soft-decision *a posteriori* probability of the v-th bit:
$$L'_v = L_v + \sum_{c' \in \mathcal{M}(v)} \mu'_{c'\to v}. \tag{4}$$

After the BP iteration reaches the predefined maximum number of iterations N_{ite}, a hard decision is taken place based on L'_v. When BICM-ID is employed, the soft-decision extrinsic information (i.e., $b'_v = L'_v - L_v$) are fed back to the demodulator to improve the reliability of LLR messages.

Note that the decoding iteration can be terminated earlier than N_{ite} to avoid unnecessary iterations and to save power consumption once the soft decision provides a valid codeword, which passes a syndrome check. Since the CND requires arithmetic multiplications and nonlinear functions in (3), a number of simplified versions such as min-sum algorithm have been investigated [47]. In order to improve convergence speed of iterative BP decoding, several scheduling methods [48] have also been proposed. Although conventional flooding scheduling (which alternates message updates in all VND at once and all CND at once) does not converge fast compared to other scheduling methods, it has been widely used because of simplicity for implementation. Instead of the regular BP decoding, the LDPC codes can be decoded by, e.g., bit flipping [1], analog decoding [49], divide-and-concur (DC) [50], and linear programming (LP) [51]. In particular, DC and LP decoding showed lower error floor than BP decoding. In order not to over-extend the scope of this paper, we consider the widely used standard BP decoding based on SPA described in (2) through (4) with the flooding scheduling as a benchmark.

D. EXIT Chart

EXIT chart analysis [4] for determining the required SNR, also called as decoding *threshold*, is based on iterative computation of the MI between an edge message and an associated transmitted bit. The EXIT chart analysis examines whether the MI reaches a value of 1 by iterative BP decoding. Note that the value 1 of the MI implies that the transmitted bit is decoded correctly with no error.

As in [4], we let $J(\sigma)$ denote the capacity of binary-input AWGN (BiAWGN) channel, i.e., the MI between a binary random variable $X \in \{-\frac{\sigma^2}{2}, \frac{\sigma^2}{2}\}$ (with equal probabilities) and an output $Y = X + Z$, where Z is a random variable following a zero-mean Gaussian distribution with a variance of σ^2, obtained as follows:
$$J(\sigma) = 1 - \mathbb{E}_Y\big[\log_2(1+\exp(-Y))\big]. \tag{5}$$

In [4], Marquardt–Levenberg algorithm is applied to approximate the $J(\cdot)$-function and inverse $J(\cdot)$-function in closed-form expressions. Let $\lambda(x) = \sum_d \lambda_d\, x^d$ and $\rho(x) = \sum_d \rho_d\, x^d$ be polynomial representations of variable-/check-degree distributions, where λ_d and ρ_d are fractions (edge perspective) of degree-d variable nodes and check nodes, respectively, such that $\sum_d \lambda_d = \sum_d \rho_d = 1$. For irregular LDPC codes having degree distributions $\lambda(x)$ and $\rho(x)$, the averaged MI is updated in the iterative BP decoding as follows:

- The extrinsic MI at VND:
$$\bar{I}_{\mathsf{vnd}} = \sum_{d=1}^{d_{\mathsf{vmax}}} \lambda_d\, I_{\mathsf{Ev}}(\bar{I}_{\mathsf{cnd}}, d, I_{\mathsf{ch}}), \tag{6}$$
where the EXIT function for degree-d_v nodes is given as
$$I_{\mathsf{Ev}}(I_{\mathsf{a}}, d_v, I_{\mathsf{ch}}) = J\big(\sqrt{(d_v-1)[J^{-1}(I_{\mathsf{a}})]^2 + [J^{-1}(I_{\mathsf{ch}})]^2}\big),$$
with I_{ch} being the GMI of the initial LLR.
- The extrinsic MI at CND:
$$\bar{I}_{\mathsf{cnd}} = \sum_{d=1}^{d_{\mathsf{cmax}}} \rho_d\, I_{\mathsf{Ec}}(\bar{I}_{\mathsf{vnd}}, d), \tag{7}$$
where the EXIT function for degree-d_c nodes is approximated by the duality property [4] as follows:
$$I_{\mathsf{Ec}}(I_{\mathsf{a}}, d_c) \simeq 1 - J\big(\sqrt{(d_c-1)[J^{-1}(1-I_{\mathsf{a}})]^2}\big).$$

Here, d_{vmax} and d_{cmax} denote the maximum variable- and check-node degrees, respectively. It is known in [2], [3] that the larger maximum degree offers the better threshold. However, since the computational complexity at variable and check nodes increases linearly with their degree, we should constrain the maximum degree for practical real-time decoding.

If the EXIT curve of VND intersects that of CND, the BP decoding does not converge to an error-free MI of one. The decoding threshold can be determined to find a minimum possible SNR such that the VND curve lies above the CND curve. From the area property [30], the conventional code design method tries to minimize the area between the EXIT curve of VND and that of CND while keeping no intersection, e.g., by means of linear programming for curve fitting. However, this curve-fitting method assumes a large number of decoding iterations to reach an MI of 1. In order to optimize degree distribution under the finite number of iterations, we use the decoding trajectory determined by (6) and (7) in the EXIT chart, instead of the conventional curve fitting.

Note that the EXIT chart can be used for analyzing NB-LDPC codes as studied in [46]. The EXIT chart analysis was also modified to protograph-based LDPC codes in [52], which can provide more accurate threshold by tracking the MI updates for all variable and check nodes in a protograph. This is useful to design LDPC codes for high-order modulation to deal with different reliability of multi-level bits in the modulation.

The protograph-EXIT (P-EXIT) was further extended to NB-LDPC codes in [53]. In [28], this modified EXIT chart analysis was used to evaluate the threshold of NB-LDPC convolutional codes. In this paper, we nonetheless use the decoding trajectory in the standard EXIT chart described in (6) and (7) because it is generally applicable for various communications systems.

III. Iteration-Aware LDPC Code Design

In this section, we analyze degree optimization via the EXIT trajectory for finite-iteration BP decoding, under BICM transmission. We assume DP-QPSK transmission to evaluate the threshold and error-rate performance. Note that the results of DP-QPSK can be readily converted to any other modulation formats for BICM, via the $J(\cdot)$-function with the GMI analysis. Our results suggest that an LDPC code for BICM should be different when the BP decoder changes the maximum number of iterations N_{ite} to control processing throughput and power consumption. Unless we carefully consider the decoding trajectory in designing LDPC codes, we may suffer from up to 2 dB penalty.

A. EXIT Trajectory Analysis

In this paper, we consider check-concentrated triple-degree LDPC codes ensemble, in which the number of distinct variable degrees is at most three and check degrees are consecutive two as follows:

$$\lambda(x) = \lambda_{d_{\mathsf{v1}}} x^{d_{\mathsf{v1}}} + \lambda_{d_{\mathsf{v2}}} x^{d_{\mathsf{v2}}} + \lambda_{d_{\mathsf{v3}}} x^{d_{\mathsf{v3}}}, \tag{8}$$

$$\rho(x) = \rho_{d_{\mathsf{c}}} x^{d_{\mathsf{c}}} + \rho_{d_{\mathsf{c}}+1} x^{d_{\mathsf{c}}+1}. \tag{9}$$

As discussed in [3], [4], such triple-degree LDPC codes perform surprisingly well. For example, rate-0.8 LDPC codes ensemble having degree distributions of $\lambda_\infty(x)$ and $\rho_\infty(x)$ listed in Table I achieves a threshold of 4.17 dB, which is only within 0.06 dB from the Shannon limit of $J^{-1}(0.8)/2 = 4.11$ dB for BICM systems. In fact, these degree distributions, $\lambda_\infty(x)$ and $\rho_\infty(x)$, were obtained by the conventional curve-fitting optimization under the maximum degree constraints of $d_{\mathsf{vmax}} \le 16$ and $d_{\mathsf{cmax}} \le 32$. We chose these maximum degree constraints throughout the paper as it is feasible in some practical codes [21].

Fig. 2 shows the EXIT curves of the optimized LDPC code. Thanks to the curve-fitting optimization method, the VND curve at an SNR of 4.2 dB agrees closely with the CND curve while the VND curve does not go below the CND curve. Although such LDPC codes can provide excellent performance near the Shannon limit, the required number of iterations to reach a top-right corner (i.e., error-free MI) becomes extremely large. For finite-iteration decoding, the required SNR degrades rapidly. For example, the VND curve at an SNR of 6.7 dB in this figure increases the area from the CND curve, and the decoding trajectory after eight iterations can reach the MI of one at the top-right corner of the EXIT chart.

Our code design approach is based on the EXIT trajectory under the finite-iteration decoding. Instead of using the conventional curve-fitting approach, we can optimize the degree distributions $\lambda(x)$ and $\rho(x)$ to minimize the required SNR such that the MI updated after N_{ite}-times iterations, computed via (6) and (7), goes to one. The check-concentrated triple-degree LDPC codes in (8) and (9) have five variables $\{\lambda_d\}$ and $\{\rho_d\}$ given degrees d_{v1}, d_{v2}, d_{v3}, and d_{c}. For this case, we have just two degrees of freedom (DOF) because there are three constraints: i) variable-degree normalization $\sum_d \lambda_d = 1$, ii) check-degree normalization $\sum_d \rho_d = 1$, and iii) target code rate $R = 1 - \bar{d}_{\mathsf{v}}/\bar{d}_{\mathsf{c}}$, where $\bar{d}_{\mathsf{v}}$ and $\bar{d}_{\mathsf{c}}$ are the average variable- and check-node degrees, respectively, written as follows:

$$\bar{d}_{\mathsf{v}} = \frac{1}{\sum_d \lambda_d / d}, \qquad \bar{d}_{\mathsf{c}} = \frac{1}{\sum_d \rho_d / d}. \tag{10}$$

We thus search for the best degree distribution in a two-DOF grid space of $\lambda_{d_{\mathsf{v1}}}$ and $\rho_{d_{\mathsf{c}}}$ for all possible combinations of degrees, $1 < d_{\mathsf{v1}} < d_{\mathsf{v2}} < d_{\mathsf{v3}} \le d_{\mathsf{vmax}}$ and $1 < d_{\mathsf{c}} < d_{\mathsf{cmax}}$. We use scanning grid steps of 0.01 and 0.05 for $\lambda_{d_{\mathsf{v1}}}$ and $\rho_{d_{\mathsf{c}}}$, respectively, from 0 to 1. Although finer grids can potentially provide better solutions, the two-DOF searching can be more time-consuming and the performance improvement is confirmed to be very marginal.

Table I lists examples of the degree distributions optimized by our iteration-aware EXIT trajectory design method for the number of iterations of $N_{\mathsf{ite}} \in \{1, 2, 4, 8, 16, 32, \infty\}$. It can be seen that all the optimized degrees are different depending on the number of iterations N_{ite}. The EXIT curves of the 4-iteration optimal code are also present in Fig. 2 to compare with the one optimized by the conventional curve-fitting method. It is shown in Fig. 2 that the EXIT trajectory for the 4-iteration optimal code achieves the error-free MI after four iterations at an SNR of 6.7 dB, while the conventional code requires eight iterations. Although the conventional curve-fitting method provides capacity-approaching codes when the

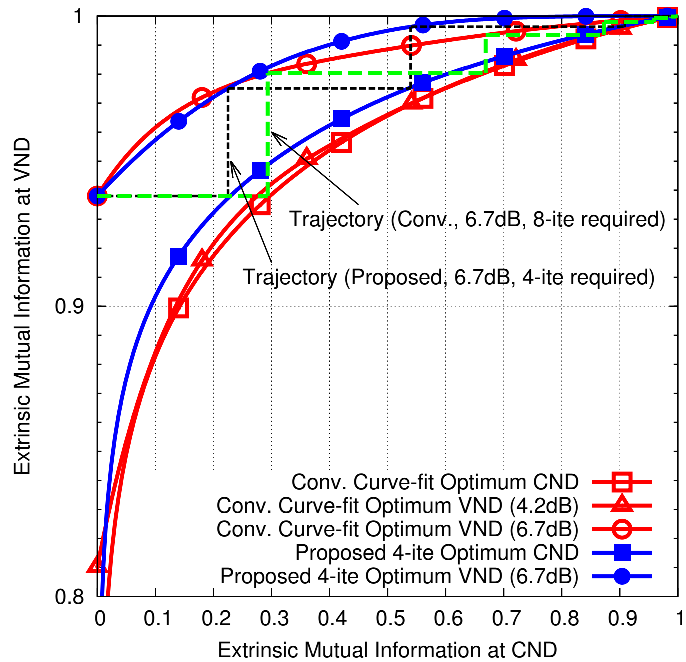

Fig. 2. EXIT curves of two irregular LDPC codes with a rate of 0.8: i) optimized degrees $\lambda_\infty(x) = 0.150x^2 + 0.314x^3 + 0.536x^{12}$ and $\rho_\infty(x) = 0.70x^{22} + 0.30x^{23}$ by conventional curve-fitting method and ii) optimized degrees $\lambda_4(x) = 0.506x^5 + 0.494x^6$ and $\rho_4(x) = 0.75x^{27} + 0.25x^{28}$ by iteration-aware trajectory method.

TABLE I

DEGREE DISTRIBUTIONS OPTIMIZED BY THE ITERATION-AWARE EXIT TRAJECTORY DESIGN METHOD FOR CHECK-CONCENTRATED TRIPLE-DEGREE LDPC CODES WITH A CODE RATE OF $R = 0.8$ (THE CORRESPONDING SHANNON LIMIT IS 4.11 DB).

Maximum number of iterations	Variable degree distribution	Check degree distribution	Average node degrees	Threshold
$N_{\mathsf{ite}} = \infty$ (by curve-fitting)	$\lambda_\infty(x) = 0.150x^2 + 0.314x^3 + 0.536x^{12}$	$\rho_\infty(x) = 0.70x^{22} + 0.30x^{23}$	$\bar{d}_{\mathsf{v}} = 4.45, \bar{d}_{\mathsf{c}} = 22.29$	4.17 dB
$N_{\mathsf{ite}} = 32$	$\lambda_{32}(x) = 0.030x^2 + 0.464x^3 + 0.506x^{16}$	$\rho_{32}(x) = 0.15x^{24} + 0.85x^{25}$	$\bar{d}_{\mathsf{v}} = 4.97, \bar{d}_{\mathsf{c}} = 24.84$	4.45 dB
$N_{\mathsf{ite}} = 16$	$\lambda_{16}(x) = 0.430x^3 + 0.117x^4 + 0.453x^{16}$	$\rho_{16}(x) = 0.10x^{24} + 0.90x^{25}$	$\bar{d}_{\mathsf{v}} = 4.98, \bar{d}_{\mathsf{c}} = 24.90$	4.78 dB
$N_{\mathsf{ite}} = 8$	$\lambda_8(x) = 1.000x^4$	$\rho_8(x) = 1.00x^{20}$	$\bar{d}_{\mathsf{v}} = 4.00, \bar{d}_{\mathsf{c}} = 20.00$	5.39 dB
$N_{\mathsf{ite}} = 4$	$\lambda_4(x) = 0.506x^5 + 0.494x^6$	$\rho_4(x) = 0.75x^{27} + 0.25x^{28}$	$\bar{d}_{\mathsf{v}} = 5.45, \bar{d}_{\mathsf{c}} = 27.24$	6.65 dB
$N_{\mathsf{ite}} = 2$	$\lambda_2(x) = 0.563x^6 + 0.437x^7$	$\rho_2(x) = 1.00x^{32}$	$\bar{d}_{\mathsf{v}} = 6.40, \bar{d}_{\mathsf{c}} = 32.00$	8.89 dB
$N_{\mathsf{ite}} = 1$	$\lambda_1(x) = 0.986x^2 + 0.016x^3$	$\rho_1(x) = 0.95x^{10} + 0.05x^{11}$	$\bar{d}_{\mathsf{v}} = 2.01, \bar{d}_{\mathsf{c}} = 10.05$	12.03 dB

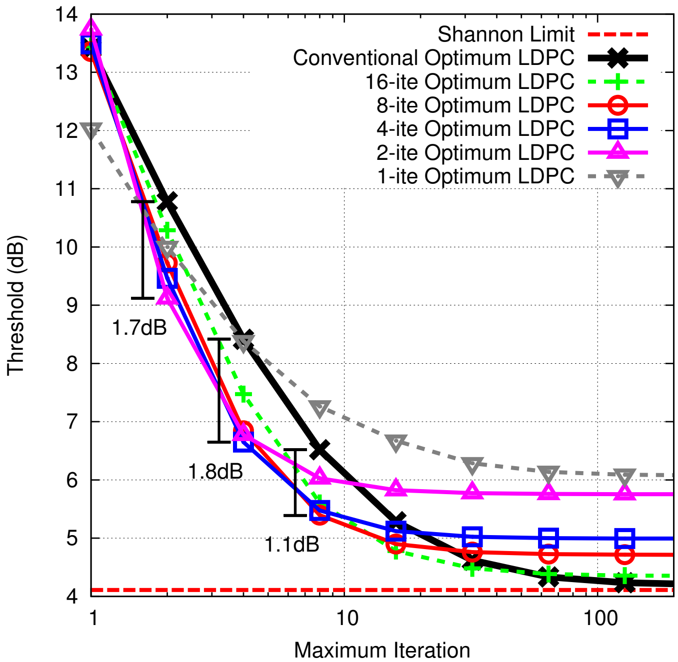

Fig. 3. Threshold vs. maximum iteration of iteration-aware LDPC codes with a code rate of $R = 0.8$ for BICM with finite-iteration BP decoding.

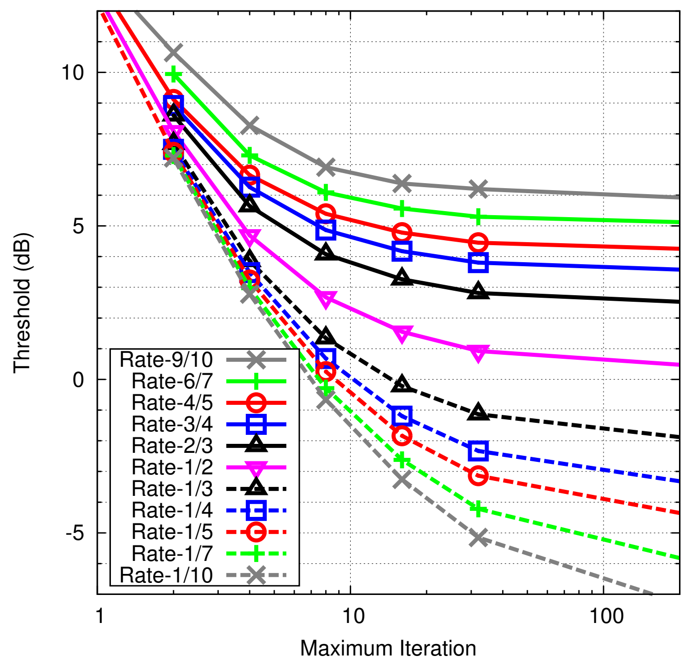

Fig. 4. Threshold of iteration-dependent LDPC codes with various code rates for BICM with finite-iteration BP decoding.

number of iterations is un-limited and codeword length is very large, the obtained codes can be no longer optimal for finite-iteration BP decoding.

B. Threshold Analysis

We now evaluate the iteration-aware LDPC code design in Fig. 3, in which we plot the threshold as a function of the number of BP iterations N_{ite} for the optimized rate-0.8 LDPC codes listed in Table I. Note that the threshold can be computed by a bisection search. The LDPC code designed by the conventional curve-fitting method achieves the best performance near the Shannon limit if the decoder can iterate more than 100 times, while the threshold seriously degrades for the cases of fewer iterations. For such fewer-iteration decoding, we shall use different irregular LDPC codes. Our iteration-aware design method provides the best threshold at the intended number of iterations. For example, the LDPC code optimized for 8-iteration decoder outperforms the conventionally optimized LDPC code by 1.1 dB, and the LDPC code optimized for 4-iteration decoder offers 1.8 dB gain. On the other hand, the 2- and 4-iteration optimized LDPC codes have approximately 1.5 dB and 0.8 dB loss, respectively, from the conventionally optimized LDPC code when the decoder can iterate more than 100 times.

Since an LDPC code optimized for a specific number of iterations can suffer from near 2 dB degradation for the different number of iterations as shown in Fig. 3, we should assign different LDPC codes optimized depending on the number of iterations N_{ite} for BP decoding, which adaptively controls the power consumption. Fig. 4 shows the threshold of such an adaptive LDPC code assignment optimized by our iteration-aware EXIT trajectory design method, according to the limited number of iterations for several code rates of $R \in \{\frac{1}{10}, \frac{1}{7}, \frac{1}{5}, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \frac{4}{5}, \frac{6}{7}, \frac{9}{10}\}$. It is found in this figure that lower-rate codes are more susceptive to the limited number of BP iterations, and require more iterations to converge.

C. BER Performance

The above-described threshold analysis can tell how good the LDPC code ensemble would be, depending on the degree distributions and the number of decoding iterations. However, the EXIT trajectory analysis assumes an infinite-length code-

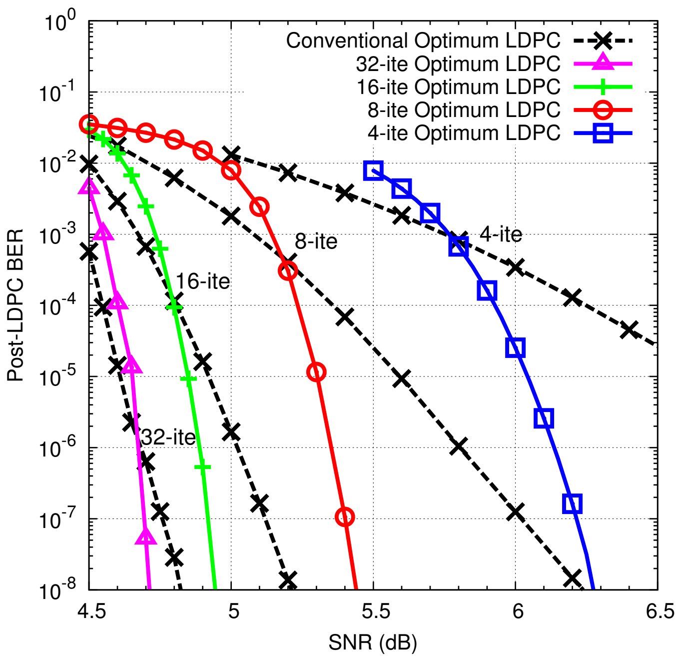

Fig. 5. BER performance of iteration-dependent LDPC codes with a code rate of $R = 0.8$ for BICM with finite-iteration BP decoding.

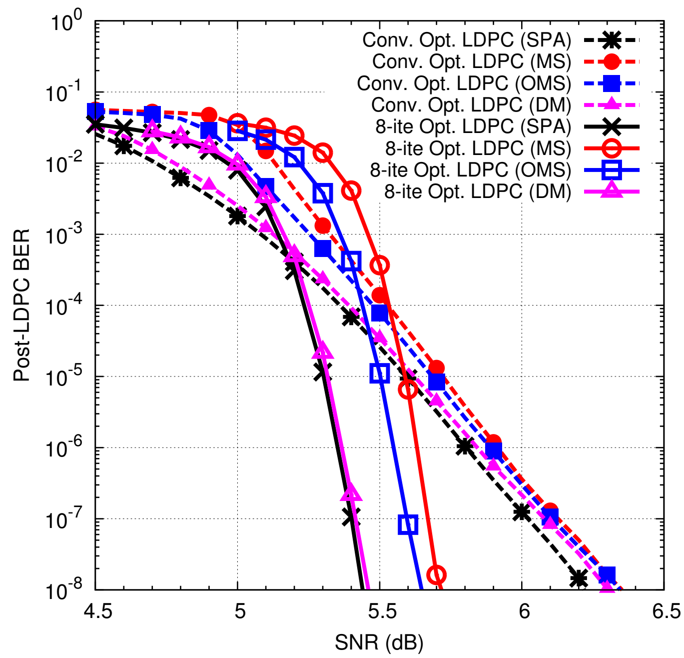

Fig. 6. BER performance of iteration-dependent LDPC code with different decoding algorithms for $N_{\mathrm{ite}} = 8$ (SPA: sum-product algorithm, MS: min-sum, OMS: offset min-sum, DM: delta-min).

word to hold (6) and (7), which rely on the central limit theorem. Hence, after the degree optimization, we need to instantiate a parity-check matrix of finite-length LDPC codes having the corresponding degree distributions. To do so, we use a progressive edge-growth (PEG) [54], which maximizes a minimum length of cycle, referred to as *girth*, in the Tanner graph. The girth maximization is generally important to reduce an error floor, which is an inevitable artifact of loopy BP decoding. We consider a codeword length of 38400 bits as it is the same length used for a state-of-the-art LDPC code achieving a net coding gain (NCG) of 12 dB in [21].

Fig. 5 shows bit-error rate (BER) performance of the iteration-dependent LDPC codes designed by PEG according to the optimized degree distributions listed in Table I. As expected in the threshold analysis, the conventionally optimized LDPC code by the curve-fitting method does not perform well for fewer-iteration BP decoding, in which the BER slope becomes worse. Our iteration-aware LDPC codes perform much better for each cases. For 8-iteration BP decoding, the required SNR at a BER of 10^{-8} of the conventionally optimized code has a loss of 0.8 dB compared to our 8-iteration optimized LDPC code. This gap must be much more significant at a BER of 10^{-15}.

Note that most optical commutations systems require a very low BER around 10^{-15}. In [21], an error floor above a BER of 10^{-8} of an irregular LDPC code was efficiently removed to achieve a BER of 10^{-15}, by using an outer code based on a Bose–Chaudhuri–Hocquenghem (BCH) code with an additional overhead of only 0.78%. Since no error floor is observed in Fig. 5 above a BER of 10^{-8}, it is expected that the optimized LDPC codes in this paper can also achieve a BER of 10^{-15} using an outer code with a small additional overhead in a similar way, to cope with a potential error floor at a BER between 10^{-8} and 10^{-15}.

It should be noticed that the required SNR for a BER of 10^{-8} in Fig. 5 agree well with the analytical threshold derived in Table I. For instance, the required SNR for $N_{\mathrm{ite}} = 8$ in Fig. 5 is 5.43 dB, which is within 0.1 dB from the analytical threshold of 5.39 dB in Table I. This indicates that our EXIT trajectory design method is reasonably applicable for practical finite-length LDPC codes. We also note that the threshold analysis based on the EXIT trajectory can predict more accurate achievability than the GMI analysis, which has been more recently utilized as a better metric than the conventional pre-FEC BER to compare with various modulation formats [55], [56]. We should recall that the GMI assumes un-limited decoding complexity, and that the behavior of the MI updates from the initial GMI in (6) highly depends on which LDPC codes ensemble is available in the communications system (e.g., how large the maximum degrees, average degrees, and distinct degrees are considered). Hence, the GMI metric can often be optimistic in comparison to the threshold metric based on the EXIT trajectory, for practical systems.

We have also confirmed that the optimized codes in Fig. 5 have a lower average number of iterations than the conventional code for the whole SNR regime when early termination with syndrome checking at every iteration is carried out. Although we consider SPA decoding for LDPC code design, it is expected that the designed codes still have a great advantage over the conventional code for various simplified algorithms [47] because those algorithms have relatively small penalty from SPA. In Fig. 6, we compare BER performance of our iteration-aware optimized code and the conventional code optimized by the curve fitting for $N_{\mathrm{ite}} = 8$ using different decoding algorithms; SPA, min-sum (MS), offset min-sum

(OMS), and delta-min (DM) [47]. The LDPC codes are the same ones used in Fig. 5. It is verified from Fig. 6 that our LDPC code designed for SPA still outperforms the conventional code even for different decoding algorithms. However, since the EXIT curves highly depend on decoding algorithms, there is potential improvement by adaptively designing the degree distribution for each specific algorithm. Although our design methodology is applicable to any iterative decoding algorithms by modifying the EXIT curves, we leave detailed analysis of decoder-dependent code design as future work.

D. Pareto-Optimal Code Design

We designed thus far practical LDPC codes under a limited number of iterations N_{ite}, given the maximum degree constraints d_{vmax} and d_{cmax}. Here, we discuss in detail the computational complexity by taking the average node degrees $\bar{d}_{\mathsf{v}}$ and $\bar{d}_{\mathsf{c}}$ into account. In fact, to achieve low-power decoding, we need to consider the average degrees as well as the number of iterations because the computational complexity of the BP decoding in (2) and (3) is of a linear order as a function of the average degree. In this paper, we further introduce a multi-objective optimization concept to design Pareto-optimal LDPC codes so that better threshold and lower complexity are achieved at the same time. For simplicity of analysis, we suppose that the decoding complexity is proportional to the number of iterations N_{ite} multiplied by the number of edges in the Tanner graph, i.e., $N_{\mathsf{ite}} \times \bar{d}_{\mathsf{v}}/R$ per information bit (note that $R = 1 - \bar{d}_{\mathsf{v}}/\bar{d}_{\mathsf{c}}$).

The optimized degree distributions in Table I have relatively larger average degrees of $\bar{d}_{\mathsf{v}} \geq 4.0$, except for the case of $N_{\mathsf{ite}} = 1$, leading to higher complexity in decoding. In the two-DOF search for the iteration-aware degree optimization, there exist a large number of different degree distributions, whose thresholds are comparable to the best codes. In consequence, we may be able to find better codes having good trade-off between the threshold and the complexity. For example, instead of decreasing the number of iterations N_{ite} by half for lower power consumption, we may halve the average degree $\bar{d}_{\mathsf{v}}$ while keeping the number of iterations to achieve better threshold in the end. Our new design criterion considers the following multi-objective optimization for a pair of the threshold and the computational complexity:

$$\min_{\lambda(x),\rho(x),N_{\mathsf{ite}}} \left[\text{threshold},\ N_{\mathsf{ite}} \frac{\bar{d}_{\mathsf{v}}}{1 - \bar{d}_{\mathsf{v}}/\bar{d}_{\mathsf{c}}}\right]. \tag{11}$$

In Fig. 7, we plot the threshold as a function of the computational complexity for some randomly-selected degree distributions in the two-DOF search, varying the number of iterations N_{ite} from 1 to 64. Not only the threshold but also the complexity can scatter in a wide range even for the same number of iterations, according to the degree distributions. The single-objective optimization searches for the best code achieving the minimum possible threshold for each fixed number of iterations N_{ite}. However, if we increase the number of iterations while decreasing the average node degree, we can obtain better codes having lower threshold and lower complexity at the same time.

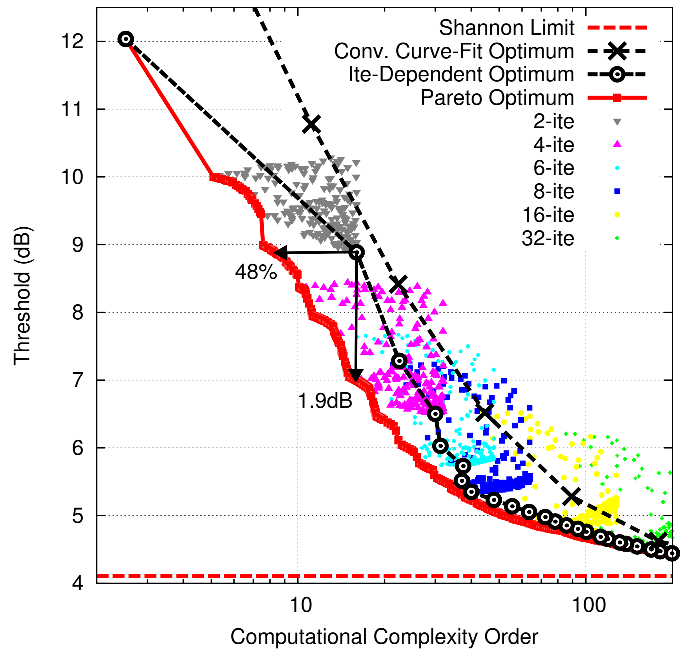

Fig. 7. Threshold vs. computational complexity of Pareto-optimal LDPC codes with a code rate of $R = 0.8$ for BICM with finite-iteration BP decoding.

For example, the 2-iteration optimal code (given in Table I) achieves a threshold of 8.89 dB in a complexity order of $N_{\mathsf{ite}}\bar{d}_{\mathsf{v}}/R = 16.0$, whereas a better threshold of 7.00 dB (thus, 1.89 dB performance improvement) can be achieved in a slightly lower complexity order of 15.89 by a Pareto-optimal code, whose degree distributions are $\lambda(x) = 0.775x^3 + 0.225x^4$ and $\rho(x) = 0.10x^{15} + 0.90x^{16}$ with a lower average node degree of $\bar{d}_{\mathsf{v}} = 3.18$ and the larger number of iterations of $N_{\mathsf{ite}} = 4$. Moreover, another Pareto-optimal code with $\lambda(x) = 0.682x^2 + 0.318x^3$ and $\rho(x) = 0.80x^{11} + 0.20x^{12}$ achieves a slightly better threshold of 8.88 dB while the computational complexity is significantly reduced to 8.39 (thus, 48% complexity reduction) with $N_{\mathsf{ite}} = 3$. For the other examples, we can obtain a better threshold by 0.9 dB and a lower complexity by 33% than the 4-iteration optimal code (threshold: 6.65 dB; complexity order: 27.25), respectively, by using a Pareto-optimal code (threshold: 5.75 dB; complexity order: 27.10) with $N_{\mathsf{ite}} = 7$, $\lambda(x) = 0.875x^3 + 0.125x^4$ and $\rho(x) = 0.50x^{15} + 0.50x^{16}$, and another Pareto-optimal code (threshold: 6.63 dB; complexity order: 18.36) with $N_{\mathsf{ite}} = 5$, $\lambda(x) = 0.043x^2 + 0.957x^3$ and $\rho(x) = 0.30x^{14} + 0.70x^{15}$.

In particular for lower-complexity regimes below 40 ($N_{\mathsf{ite}} < 8$), the Pareto-optimal codes are more advantageous to jointly minimize the threshold and the computational complexity. For higher-complexity regimes above 40, the iteration-dependent single-objective optimization in Table I for $N_{\mathsf{ite}} \geq 8$ can already provide good LDPC codes near the Pareto front.

IV. Conclusions

We have shown a significant benefit of designing iteration-aware LDPC codes, based on EXIT trajectory analysis. We have analyzed thresholds and BER performance of the optimized LDPC codes for BICM under a limited number of decoding iterations. It has been demonstrated that if we use an LDPC code which is optimized at a certain number of iterations, we can suffer from a large penalty close to 2 dB if the number of iterations is changed to control power consumption. The results suggest that we should carefully design LDPC codes depending on the number of iterations to exploit full potential of LDPC codes. We have also introduced a new design concept with multi-objective optimization to jointly minimize the required SNR and the computational complexity by accounting for the average degree. Our Pareto-optimal codes offer an additional gain by 2 dB or a reduced complexity by 50% in the low-complexity regimes, by decreasing the number of edges in the Tanner graph and increasing the number of iterations to keep the total complexity low. Extension to other decoding algorithms and scheduling methods accounting for the average number of iterations remain as future work.

REFERENCES

[1] R. G. Gallager, *Low-Density Parity-Check Codes*, Cambridge, MA: MIT Press, 1963.

[2] S.-Y. Chung, G. D. Forney Jr., 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.

[3] 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.

[4] 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.

[5] I. B. Djordjevic, "Advanced coded-modulation for ultra-high-speed optical transmission," in *Proc. Opt. Fiber Commun.*, San Francisco, CA, USA, Mar. 2014, Paper W3J-4.

[6] L. Schmalen, V. Aref, J. Cho, and K. Mahdaviani, "Next generation error correcting codes for lightwave systems," in *Proc. Eur. Conf. Opt. Commun.*, Cannes, France, Sep. 2014, Paper Th.1.3.3.

[7] Y. Polyanskiy, H. V. Poor, and S. Verdú, "Channel coding rate in the finite blocklength regime," *IEEE Trans. Inf. Theory*, vol. 56, no. 5, pp. 2307–2359, Apr. 2010.

[8] C. Di, D. Proietti, I. E. Telatar, T. J. Richardson, and R. L. Urbanke, "Finite-length analysis of low-density parity-check codes on the binary erasure channel," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1570–1579. Jun. 2002.

[9] I. B. Djordjevic and B. Vasic, "Approaching Shannon's capacity limits of fiber optics communications channels using short LDPC codes," in *Proc. Conf. Lasers Electro-Opt.*, San Francisco, CA, USA, May 2004, Paper CWA7.

[10] Y. Kou, S. Lin, and M. P. C. Fossorier, "Low-density parity-check codes based on finite geometries: A rediscovery and new results," *IEEE Trans. Inf. Theory*, vol. 47, no. 7, pp. 2711–2736, Nov. 2001.

[11] S. Lin, H. Tang, Y. Kou, J. Xu, and K. Abdel-Ghaffar, "Codes on finite geometries," in *Proc. IEEE Inf. Theory Workshop*, Cairns, Queensland, Australia, pp. 2–7. Sep. 2001.

[12] I. B. Djordjevic and B. Vasic, "Projective geometry LDPC codes for ultralong-haul WDM high-speed transmission," *IEEE Photon. Technol. Lett.*, vol. 15, no. 5, pp. 784–786, May 2003.

[13] D. Declercq and M. Fossorier, "Decoding algorithms for nonbinary LDPC codes over GF," *IEEE Trans. Commun.*, vol. 55, no. 4, pp. 633–643, Apr. 2007.

[14] M. Arabaci, I. B. Djordjevic, R. Saunders, and R. M. Marcoccia, "High-rate nonbinary regular quasi-cyclic LDPC codes for optical communications," *J. Lightw. Technol.*, vol. 27, no. 23, pp. 5261–5267, Aug. 2009.

[15] I. B. Djordjevic and B. Vasic, "Nonbinary LDPC codes for optical communication systems," *IEEE Photon. Technol. Lett.*, vol. 17, no. 10, pp. 2224–2226, Oct. 2005.

[16] M. Lentmaier and K. Sh. Zigangirov, "On generalized low-density parity-check codes based on Hamming component codes," *IEEE Commun. Lett.*, vol. 3, no. 8, pp. 248–250, Aug. 1999.

[17] I. B. 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.

[18] I. B. 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.

[19] S. Kudekar, T. Richardson, and R. L. Urbanke, "Spatially coupled ensembles universally achieve capacity under belief propagation," *IEEE Trans. Inf. Theory*, vol. 59, no. 12, pp. 7761–7813, Sep. 2013.

[20] 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, USA, Mar. 2014, Paper W1A-1.

[21] K. Sugihara, Y. Miyata, T. Sugihara, K. Kubo, H. Yoshida, W. Matsumoto, and T. Mizuochi, "A spatially-coupled type LDPC code with an NCG of 12 dB for optical transmission beyond 100 Gb/s," in *Proc. Opt. Fiber Commun.*, San Francisco, CA, USA, Mar. 2013, Paper OM2B.4.

[22] M. Lentmaier, M. M. Prenda, and G. Fettweis, "Efficient message passing scheduling for terminated LDPC convolutional codes," in *Proc. IEEE Int. Sym. Inf. Theory*, St. Petersburg, Russia, Jul.–Aug. 2011.

[23] 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, Apr. 2014.

[24] L. Schmalen, D. Suikat, D. Rosener, and A. Leven, "Evaluation of left-terminated spatially coupled LDPC codes for optical communications," in *Proc. Eur. Conf. Opt. Commun.*, Cannes, France, Sep. 2014, Paper Th.2.3.4.

[25] 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.*, San Francisco, CA, USA, Mar. 2014, Paper OW1H.4.

[26] T. Xia, T. Koike-Akino, D. S. Millar, K. Kojima, K. Parsons, Y. Miyata, K. Sugihara, and W. Matsumoto, "Dynamic window decoding for LDPC convolutional codes in low-latency optical communications," in *Proc. Opt. Fiber Commun.*, San Francisco, CA, USA, Mar. 2015, Paper Th3E.4.

[27] —, "Nonbinary LDPC convolutional codes for high-dimensional modulations," in *Proc. Sig. Proc. Photon. Commun.*, Boston, MA, USA, Jun. 2015, Paper SpS3D.5.

[28] L. Wei, T. Koike-Akino, D. G. M. Mitchell, T. E. Fuja, and D. J. Costello Jr., "Threshold analysis of non-binary spatially-coupled LDPC codes with windowed decoding," in *Proc. IEEE Int. Sym. Inf. Theory*, Honolulu, HI, USA, pp. 881–885, Jun.–Jul. 2014.

[29] D. Declercq and F. Verdier, "Optimization of LDPC finite precision belief propagation decoding with discrete density evolution," in *Proc. Int. Sym. Turbo Codes Related Topics*, Brest, France, Sep. 2003.

[30] A. Ashikhmin, G. Kramer, and S. ten Brink, "Extrinsic information transfer functions: Model and erasure channel properties," *IEEE Trans. Inf. Theory*, vol. 50, no. 11, pp. 2657–2673, Nov. 2004.

[31] C. Dorize, P. Layec, and G. Charlet, "DSP power balancing for multiformat WDM receiver," in *Proc. Eur. Conf. Opt. Commun.*, Cannes, France, Sep. 2014, Paper Mo.3.5.3.

[32] T. Koike-Akino, D. S. Millar, K. Kojima, and K. Parsons, "Coded modulation design for finite-iteration decoding and high-dimensional modulation," in *Proc. Opt. Fiber Commun.*, San Francisco, CA, USA, Mar. 2015, Paper W4K.1.

[33] P. Poggiolini, "The GN model of non-linear propagation in uncompensated coherent optical systems," *J. Lightw. Technol.*, vol. 30, no. 24, pp. 3857–3879, Dec. 2012.

[34] A. Carena, G. Bosco, V. Curri, Y. Jiang, P. Poggiolini, and F. Forghieri, "EGN model of non-linear fiber propagation," *Opt. Exp.*, vol. 22, no. 13, pp. 16335–16362, Jun. 2014.

[35] P. Serena, A. Bononi, and N. Rossi, "The impact of the modulation dependent nonlinear interference missed by the Gaussian noise model," in *Proc. Eur. Conf. Opt. Commun.*, Cannes, France, Sep. 2014, Paper Mo.4.3.1.

[36] E. Agrell, A. Alvarado, G. Durisi, and M. Karlsson, "Capacity of a nonlinear optical channel with finite memory," *J. Lightw. Technol.*, vol. 32, no. 16, pp. 2862–2876, Aug. 2013.

[37] S. Ishimura and K. Kikuchi, "Eight-state trellis-coded optical modulation with signal constellations of four-dimensional M-ary quadrature-amplitude modulation," *Opt. Exp.*, vol. 23, no. 5, pp. 6692–6704, Mar. 2015.

[38] L. Beygi, E. Agrell, J. M. Kahn, and M. Karlsson, "Rate-adaptive coded modulation for fiber-optic communications," *IEEE J. Lightw. Technol.*, vol. 32, no. 2, pp. 333–343, Oct. 2013.

[39] T. Liu and I. B. Djordjevic, "Multidimensional optimal signal constellation sets and symbol mappings for block-interleaved coded-modulation

enabling ultra high-speed optical transport," *IEEE Photon. J.*, vol. 6, no. 4, pp. 1–14, Aug. 2014.

[40] G. Ungerboeck, "Channel coding with multilevel/phase signals," *IEEE Trans. Inf. Theory*, vol. 28, no. 1, pp. 55–67, Jan. 1982.

[41] U. Wachsmann, R. F. H. Fischer, and J. B. Huber, "Multilevel codes: Theoretical concepts and practical design rules," *IEEE Trans. Inf. Theory*, vol. 45, no. 5, pp. 1361–1391, Jul. 1999.

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

[43] T. H. Lotz, R. Urbansky, and W. Sauer-Greff, "Iterative forward error correction decoding for spectral efficient optical OFDM transmission systems," in *Proc. Sig. Proc. Photon. Commun.*, Karlsruhe, Germany, Jun. 2010, Paper SPThA2.

[44] T. Koike-Akino, K. Kojima, D. S. Millar, K. Parsons, Y. Miyata, W. Matsumoto, and T. Mizuochi, "Cycle slip-mitigating turbo demodulation in LDPC-coded coherent optical communications," in *Proc. Opt. Fiber Commun.*, San Francisco, CA, USA, Mar. 2014, Paper M3A-3.

[45] T. Koike-Akino, D. S. Millar, K. Kojima, K. Parsons, T. Yoshida, K. Ishida, Y. Miyata, W. Matsumoto, and T. Mizuochi, "Turbo demodulation for LDPC-coded high-order QAM in presence of transmitter angular skew," in *Proc. Eur. Conf. Opt. Commun.*, Cannes, France, Sep. 2014, Paper Th.1.3.2.

[46] A. Bennatan and D. Burshtein, "Design and analysis of nonbinary LDPC codes for arbitrary discrete-memoryless channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 549–583, Feb. 2006.

[47] W. Ji, M. Hamaminato, H. Nakayama, and S. Goto, "A novel hardware-friendly self-adjustable offset min-sum algorithm for ISDB-S2 LDPC decoder," in *Proc. Eur. Sig. Proc. Conf.* , Aarlborg, Denmark, pp. 1394–1398, Aug. 2010.

[48] A. I. V. Casado, M. Griot, and R. D. Wesel, "LDPC decoders with informed dynamic scheduling," *IEEE Trans. Commun.*, vol. 58, no. 12, pp. 3470–3479. Dec. 2010.

[49] S. Hemati and A. H. Banihashemi, "Dynamics and performance analysis of analog iterative decoding for low-density parity-check (LDPC) codes," *IEEE Trans. Commun.*, vol. 54, no. 1, pp. 61–70, Jan. 2006.

[50] J. S. Yedidia, Y. Wang, and S. C. Draper, "Divide and concur and difference-map BP decoders for LDPC codes," *IEEE Trans. Inf. Theory*, vol. 57, no. 2, pp. 786–802, Feb. 2011.

[51] J. Feldman, M. J. Wainwright, and D. R. Karger, "Using linear programming to decode binary linear codes," *IEEE Trans. Inf. Theorey*, vol. 51, no. 3, pp. 954–972, Mar. 2005.

[52] G. Liva and M. Chiani, "Protograph LDPC codes design based on EXIT analysis," in *Proc. IEEE Global Telecommun. Conf.*, Washington, D.C., USA, pp. 3250–3254, Nov. 2007.

[53] L. Dolecek, D. Divsalar, Y. Sun, and B. Amiri, "Non-binary protograph-based LDPC codes: Enumerators, analysis, and designs," *IEEE Trans. Inf. Theory*, vol. 60, no. 7, pp. 3913–3941, Apr. 2014.

[54] H. Xiao and A. H. Banihashemi, "Improved progressive-edge-growth (PEG) construction of irregular LDPC codes," *IEEE Commun. Lett.*, vol. 8, no. 12, pp. 715–717, Dec. 2004.

[55] A. Alvarado and E. Agrell, "Four-dimensional coded modulation with bit-wise decoders for future optical communications," *J. Lightw. Technol.*, vol. 33, no. 10, pp. 1993–2003, May 2015.

[56] R. Rios-Mueller, J. Renaudier, L. Schmalen, and G. Gharlet, "Joint coding rate and modulation format optimization for 8QAM constellations using BICM mutual information," in *Proc. Opt. Fiber Commun.*, San Francisco, CA, USA, Mar. 2015, Paper W3K.4.

Toshiaki Koike-Akino (M'05–SM'11) received the B.S. degree in electrical and electronics engineering, M.S. and Ph.D. degrees in communications and computer engineering from Kyoto University, Kyoto, Japan, in 2002, 2003, and 2005, respectively. During 2006–2010, he has been a Postdoctoral Researcher at Harvard University, and joined Mitsubishi Electric Research Laboratories, Cambridge, MA, USA, since 2010. His research interest includes digital signal processing for data communications and sensing. He received the YRP Encouragement Award 2005, the 21st TELECOM System Technology Award, the 2008 Ericsson Young Scientist Award, the IEEE GLOBECOM'08 Best Paper Award in Wireless Communications Symposium, the 24th TELECOM System Technology Encouragement Award, and the IEEE GLOBECOM'09 Best Paper Award in Wireless Communications Symposium.

David S. Millar (S'07–M'11) was born in Manchester, U.K., in 1982. He received the M.Eng. degree in electronic and communications engineering from the University of Nottingham, Nottingham, U.K. in 2007, and the Ph.D. degree in optical communications from University College London, London, U.K., in 2011. He is currently working at Mitsubishi Electric Research Laboratories, Cambridge, MA, USA. His research interests include coherent optical transmission systems, digital coherent receiver design, coded modulation for optical communications, and digital nonlinearity mitigation. He has served as a Reviewer for several IEEE publications including IEEE Photonics Technology Letters, IEEE Journal of Selected Topics in Quantum Electronics, IEEE Communications Letters, and the IEEE/OSA Journal of Lightwave Technology. He is also serving on the Technical Program Committee for OFC 2015 and OFC 2016.

Keisuke Kojima (S'82–M'84–SM'13) was born in Hokkaido, Japan, in 1958. He received the B.S., M.S., and Ph.D. degrees in electrical engineering from the University of Tokyo, Tokyo, Japan, in 1981, 1983, and 1990, respectively. He also received the M.S. degree from the University of California, Berkeley, CA, USA, in 1982. He worked for eight years at the Central Research Laboratory, Mitsubishi Electric Corp., from 1983 on the research of narrow linewidth lasers and AlGaAs/GaAs DFB and DBR lasers. He spent nine years at AT&T/Lucent Bell Laboratories on the R&D of uncooled Fabry–Perot and DFB lasers, vertical-cavity surface-emitting lasers, passive optical network systems, and metro optical systems, first as a Member of Technical Staff, and later as a Technical Manager. He also worked at Agere Systems, Denselight Semiconductors, and TriQuint Semiconductors on optical devices and modules, and optical systems testbed. He has been with Mitsubishi Electric Research Laboratories, Cambridge, MA, USA, since 2005, where he is currently working on the R&D of photonic-integrated circuits and coherent optical systems, as a Senior Principal Member Research Staff. He has more than 140 publications in journals and conference proceedings. He is a Fellow of the Optical Society of America.

Kieran Parsons (M'07–SM'09) received the B.Eng. and Ph.D. degrees in electronic and communications engineering from the University of Bristol, Bristol, U.K., in 1992 and 1996, respectively. During 1997–2002, he was with Nortel Networks, Ottawa, Canada, where he worked on wireless and long-haul optical communications system architecture and design. From 2004–2006, he worked on carrier-grade mesh WiFi RF system design at BelAir Networks (now part of Ericsson) and from 2006–2009 on 10-G PHY device development with Applied Micro, both in Kanata, Canada. In 2009, he joined Mitsubishi Electric Research Laboratories, Cambridge, MA, USA, where he is currently Team Leader.

Yoshikuni Miyata, biography not available at the time of publication.

Kenya Sugihara, biography not available at the time of publication.

Wataru Matsumoto, biography not available at the time of publication.