RESEARCH ARTICLE

Construction of Protograph-Based Partially Doped Generalized LDPC Codes

JAEWHA KIM [ID][1], (Member, IEEE), **JAE-WON KIM** [ID][2], (Member, IEEE),
HEE-YOUL KWAK [ID][3], (Member, IEEE), AND **JONG-SEON NO** [ID][1], (Fellow, IEEE)

[1]Department of Electrical and Computer Engineering, INMC, Seoul National University, Seoul 08826, South Korea
[2]Department of Electronic Engineering, Engineering Research Institute (ERI), Gyeongsang National University, Jinju 52828, South Korea
[3]Department of Electrical Engineering, University of Ulsan, Ulsan 44610, South Korea

Corresponding author: Jae-Won Kim (jaewon07.kim@gnu.ac.kr)

ABSTRACT In this paper, we propose a new code design technique, called partial doping, for protograph-based generalized low-density parity-check (GLDPC) codes. While the conventional construction method of protograph-based GLDPC codes is to replace some single parity-check (SPC) nodes with generalized constraint (GC) nodes applying to multiple variable nodes (VNs) that are connected in the protograph, the proposed technique can select any VNs in the protograph to be protected by GC nodes. In other words, the partial doping technique facilitates finer tuning of doping, which in turn enables a sophisticated code optimization with higher degree of freedom. We construct the proposed partially doped GLDPC (PD-GLDPC) codes using the partial doping technique and optimize the PD-GLDPC codes by the protograph extrinsic information transfer (PEXIT) analysis. In addition, we propose a condition guaranteeing the linear minimum distance growth of the PD-GLDPC codes and use the condition for the optimization. Experimental results show that the optimized PD-GLDPC codes outperform the conventional GLDPC codes and have competitive performance compared to the state-of-the-art protograph-based LDPC codes without the error floor phenomenon over the binary erasure channel (BEC).

INDEX TERMS Generalized low-density parity-check (GLDPC) codes, partial doping, partially doped GLDPC (PD-GLDPC) codes, protograph, protograph extrinsic information transfer (PEXIT), typical minimum distance.

I. INTRODUCTION

Low-density parity-check (LDPC) codes, first introduced in [1], have received much attention due to their low decoding complexity and capacity approaching performance [2]. An LDPC code is defined over a bipartite graph consisting of variable nodes (VNs) and single parity-check (SPC) nodes. As a generalized class of LDPC codes, generalized LDPC (GLDPC) codes were introduced in [3], which are constructed by replacing some SPC nodes with generalized constraint (GC) nodes. GC nodes are defined by code constraints of a linear code with a larger minimum distance [4], which makes GLDPC codes have a larger minimum distance [5].

The associate editor coordinating the review of this manuscript and approving it for publication was Zesong Fei [ID].

In addition, GLDPC codes have several advantages over LDPC codes such as faster decoding convergence [6] and a better asymptotic threshold at the cost of the additional decoding complexity and redundancy introduced by GC nodes [7]. Many types of linear codes for GC nodes, also called as the component codes, are used in the GLDPC codes such as Hamming codes [8], Hadamard codes [9], Bose–Chaudhuri–Hocquenghem (BCH) codes, and Reed-Solomon (RS) codes [10]. The research on GLDPC codes is extended to spatially coupled LDPC codes [11], [12], [13] and doubly GLDPC codes [14], [15], [16]. Moreover, some capacity approaching GLDPC codes were constructed using irregular random GLDPC codes [7], [17].

LDPC codes can be constructed from a small bipartite graph called protograph. Many researches on the

protograph-based LDPC codes were previously carried out under various scenarios [18], [19], [20]. Moreover, the protograph-based GLDPC codes were thoroughly studied in [21], [22], [23], and [24], but they mainly focused on the low-rate codes [21], [22], [23], [24]. Protograph-based GLDPC codes can be constructed from a small protograph [25] using the so called doping technique [26]. Doping a GC node, defined by a (μ, κ) linear code of length μ and dimension κ, means the replacement of an SPC node by the GC node with $\mu - \kappa$ constraints, which causes a rate loss. In the perspective of VNs, μ VNs are selected to be doped by a GC node, assuming that there are no parallel edges in the protograph. Thus, the smallest unit of doping, also called the doping granularity, is μ for the conventional protograph doping technique. In other words, the conventional doping technique has two limitations: 1) the degree of the SPC node to be replaced should be μ, which implies that the doping operation is dependent on the underlying protograph and the parameter μ of component codes and 2) one cannot choose a finer doping granularity less than μ and thus the code design cannot be sophisticated. Due to the limited design flexibility, there has been little works on the well-designed optimization for protograph-based GLDPC codes especially for medium to high code rates.

In this paper, we propose a new doping technique, called partial doping on the VNs, to minimize the doping granularity and enlarge the code design freedom. In detail, the partial doping involves the following three steps: 1) A VN to be doped is selected in the protograph. 2) The Tanner graph is obtained by the lifting operation [25] from the protograph with a lifting factor N. 3) Additional GC nodes are connected to the lifted N VNs in the Tanner graph after lifting the protograph. The main difference from the conventional protograph doping technique is that the partial doping operation is conducted on the Tanner graph instead of the protograph domain. Thus, it is possible to partially dope on a single VN in the protograph and the doping granularity becomes one, which is also independent of μ. In other words, the partial doping enables fine tuning of the code structure regardless of the underlying protograph and the parameter of component codes. Specifically, the selection of VNs to be protected by GC nodes and the rate loss can be adjusted in a more flexible manner.

We denote the proposed protograph-based GLDPC codes constructed using the partial doping as partially doped GLDPC (PD-GLDPC) codes. The structural characteristics of the PD-GLDPC codes have several advantages. First, the PD-GLDPC codes are structurally adequate to adopt the puncturing technique that compensates the rate-loss. Since the partially doped VNs are highly and locally protected by GC nodes, the performance loss occurred by puncturing the doped VNs is relatively small while attaining the code rate gain. Second, the asymptotic performance of the PD-GLDPC codes can be analyzed by the low-complexity extrinsic information transfer (EXIT) analysis. For the conventional protograph doped GLDPC codes [26], the exact

EXIT analysis is provided in [27], where the topology for the a priori and extrinsic mutual information of GC nodes is considered. Since the cases of the topology grow exponentially with the component code length μ, the computational complexity is too high to design a fast optimization algorithm. On the contrary, GC nodes in the PD-GLDPC codes can be analyzed by an average manner EXIT analysis in [28] because GC nodes in the PD-GLDPC codes are incident to VNs lifted from a single VN in the protograph. The a priori and extrinsic mutual information of GC nodes can be evaluated by a single value, which facilitates a fast optimization algorithm. Using this advantage, we propose an efficient optimization algorithm for the PD-GLDPC codes.

In addition, we propose the condition guaranteeing the linear minimum distance growth of the PD-GLDPC codes. We analytically prove that the PD-GLDPC code ensembles satisfying the condition have the typical minimum distance and use this condition for the construction of the PD-GLDPC codes in this paper. Also, we propose novel methods to optimize the asymptotic performance, i.e., the threshold of the code ensemble, by using the protograph EXIT (PEXIT) analysis [29] and differential evolution [30] targeting medium code rate 1/2 and high code rate 2/3. Thus, the optimized PD-GLDPC code ensembles are constructed while satisfying the typical minimum distance condition to have a minimum distance that grows linearly with the block length of the code. Comparison of the PD-GLDPC codes is made with the existing state-of-the-art protograph LDPC codes and conventional GLDPC codes [17]. Threshold analysis and shows that the optimized protograph-based PD-GLDPC codes outperform the well known GLDPC and protograph-based LDPC codes and have a competitive asymptotic performance compared to the optimized protograph-based LDPC codes.

To be specific, the optimized protograph PD-GLDPC codes from a random ensemble with a low doping ratio 0.02439 achieves the coding gain 0.0079 over the binary erasure channel (BEC) compared to the optimized GLDPC codes [17] with a relatively higher doping ratio 0.4. In addition, the optimized protograph PD-GLDPC code by the differential evolution outperforms AR4JA codes [31] with coding gains 0.0477 and 0.032 for code rates 1/2 and 2/3, respectively. Also, the average VN degree of the optimized PD-GLDPC codes, are only 87.2% and 80.5% compared to the state-of-the-art protograph LDPC codes for code rates 1/2 and 2/3, respectively. Similarly, the frame error rate (FER) results show tangible gain in the waterfall performance compared to the existing protograph-based LDPC codes in [31] without the error floor phenomenon up to FER 10^{-4}.

We list the contributions of this paper as follows; 1) We propose a novel doping technique, where the constraints of GC nodes are applied to specific VNs lifted from single protograph node, i.e., partial doping on the VNs after lifting. 2) We propose two design criteria for the optimization of the threshold of the PD-GLDPC codes: the EXIT analysis and the condition for the existence of the typical minimum distance. 3) We propose the optimization method of the asymptotic

performances for the PD-GLDPC codes using differential evolution. 4) We show the finite length performance gain of the optimized PD-GLDPC codes over some well known LDPC and GLDPC codes.

The rest of the paper is organized as follows. In Section II, we introduce some preliminaries on the BEC and protograph-based GLDPC codes. Section III illustrates the proposed PD-GLDPC code structure and derives its PEXIT analysis and the condition for the typical minimum distance. In addition, the comparison of the proposed PD-GLDPC codes and protograph doped GLDPC codes is given. The optimization algorithms of PD-GLDPC codes are given in Section IV. Section V shows the error correcting performance of the proposed codes over the BEC compared with other well known protograph-based LDPC codes. Section VI concludes the paper with some discussion of the results.

II. BACKGROUNDS

In this section, we introduce some notations and concepts of a binary erasure channel, protograph LDPC codes, and the construction method of protograph doped GLDPC codes. The EXIT analysis and the decoding process of protograph doped GLDPC codes are also briefly introduced. The notations mainly used throughout the paper are summarized in Table 1.

A. PROTOGRAPH LDPC CODE AND BEC

Let $\mathbf{x} = \{x_1, \cdots, x_k\}, x_i \in \{0, 1\}$ be a k-bit binary message vector, which is encoded via an (n, k) linear code, forming an n-bit codeword $\mathbf{c} = \{c_1, \cdots, c_n\}, c_i \in \{0, 1\}$. The codeword passes through a memoryless BEC, where each bit is either erased with a probability ϵ or correctly received.

Protograph LDPC codes [25] are defined by a relatively small bipartite graph $G = (V, C, E)$ representing a protograph, where $V = \{v_1, \cdots, v_{n_v}\}$ is a set of VNs and $C = \{c_1, \cdots, c_{n_c}\}$ is a set of check nodes (CNs). Let E be a set of edges e, where $e = (v, c)$ connects a VN $v \in V$ and a CN $c \in C$. The bipartite graph can also be expressed in terms of an $n_c \times n_v$-sized base matrix $\mathbf{B}_{n_c \times n_v} = \{b_{i,j}\}, i \in [n_c], j \in [n_v]$, where $b_{i,j} \in \{0, 1, 2, \cdots\}$ and $[A]$ is a set of positive integers less than or equal to a positive integer A. The rows represent the CNs and the columns represent the VNs in the protograph. Each entry $b_{i,j}$ of the base matrix represents the number of edges connected between a VN and a CN. If there are no edges connected between v_j and c_i, the entry $b_{i,j}$ is zero. The variable (check) node degree $deg(v_j)$ ($deg(c_i)$) is defined as the number of edges incident to itself. A protograph LDPC code is constructed by copy-and-permute operation of G. The bipartite graph G is copied by the lifting factor N and copies of each edge $e = (v, c) \in E$ are permuted among copies of v and c. In general, the large value of N guarantees the sparseness of the code.

B. CONSTRUCTION OF PROTOGRAPH DOPED GLDPC CODES [26]

Conventionally, a protograph doped GLDPC code ensemble is constructed by replacing (doping) a CN of a protograph

Algorithm 1 The PEXIT Analysis of a Protograph Doped GLDPC Code [32]

1: **Step 1) Initialization**
   Initialize $I_{ch}(j) = 1 - \epsilon$ for $j \in [n_v]$.

2: **Step 2) Message update from VN to CN**
   Update $I_{EV}(i, j) = 1 - \epsilon \prod_{t \in N(v_j)} \left(1 - I_{AV}(t, j)\right)^{\delta(t,j)}$ for all $j \in [n_v]$, where $\delta(t, j) = b_{t,j}$ for $t \neq i$ and $\delta(t, j) = b_{t,j} - 1$ for $t = i$. Further, $I_{EV}(i, j) = 0$ if $b_{i,j} = 0$. If c_i is an SPC node, $I_{AV}(i, j) = I_{EC}(i, j)$ and if c_i is a GC node, $I_{AV}(i, j) = I_{EGC}(i, j)$.

3: **Step 3) Message update from CN to VN**
   For all i, if c_i is an SPC node, go to **Step 3-1)** and if c_i is a GC node, go to **Step 3-2)**.
   Step 3-1) $I_{EC}(i, j) = \prod_{t \in N(c_i)} I_{AC}(i, t)^{\delta(i,t)}$, where $\delta(i, t) = b_{i,t}$ for $t \neq j$ and
   $\delta(i, t) = b_{i,t} - 1$ for $t = j$. Further, $I_{AC}(i, t) = I_{EV}(i, t)$.
   Step 3-2) For all $j \in N(c_i)$, compute

$$I_{EGC}(i, j) = \frac{1}{n_i} \sum_{h=1}^{n_i} \left(1 - I_{AGC}(i)\right)^{h-1} \left(I_{AGC}(i)\right)^{n_i - h}$$
$$\times [h\tilde{e}_h - (n_i - h + 1)\tilde{e}_{h-1}], \quad (1)$$

   where $I_{AGC}(i) = \frac{1}{n_i} \sum_{j \in N(c_i)} b_{i,j} \times I_{EV}(i, j)$.

4: **Step 4) APP mutual information computation**
   For all $j \in [n_v]$, $I_{APP}(j) = 1 - \epsilon \prod_{t \in N(v_j)} \left(1 - I_{AV}(t, j)\right)^{b_{t,j}}$. If c_t is an SPC node, $I_{AV}(t, j) = I_{EC}(t, j)$ and if c_t is a GC node, $I_{AV}(t, j) = I_{EGC}(t, j)$.

5: **Step 5) Convergence check of VNs**
   Repeat **Step 2)–4)** until $I_{APP}(j) = 1$, for all $j \in [n_v]$.

with a GC node that has a parity-check constraint from an (n_i, k_i, d_{min}^i) linear code (component code), where n_i (k_i) is the code length (dimension) and d_{min}^i is the minimum distance of the component code for a CN c_i. The condition for replacement is that the CN degree should be exactly equal to the length of the component code, i.e., $deg(c_i) = n_i$. Note that the original CN has the parity-check constraint of an $(n_i, k_i) = (deg(c_i), deg(c_i) - 1)$ SPC code. The code rate R of protograph doped GLDPC codes is $R = 1 - \frac{m_{proto}}{n_v}$, where $m_{proto} = \sum_{i=1}^{n_c} (n_i - k_i)$. While the minimum distance of an SPC node is 2, the VNs connected to the GC node are protected by parity-check constraints of the component code with the minimum distance larger than two. Fig. 1 shows the protograph doped GLDPC code of the code rate 3/7 by replacing an SPC node with the $(7, 4)$ Hamming code constraints.

C. PEXIT ANALYSIS AND DECODING PROCESS OF PROTOGRAPH DOPED GLDPC CODES

The asymptotic performance of the protograph doped GLDPC codes is evaluated by the PEXIT analysis. The PEXIT analysis tracks down the mutual information of extrinsic messages and a priori error probabilities of the VNs,

TABLE 1. Main mathematical notations used in the paper.

| Notation	Explanation
$\mathbf{B}_{n_c \times n_v}$	An $n_c \times n_v$ base matrix defining the protograph
$V = \{v_1, \cdots, v_{n_v}\}$	Set of VNs in a protograph
$C = \{c_1, \cdots, c_{n_c}\}$	Set of CNs in a protograph
E	Set of edges connecting V and C
$deg(v_j) \ (deg(c_i))$	Number of edges incident to v_j (c_i), i.e., variable (check) node degree of v_j (c_i)
$N(c_i) \ (N(v_j))$	Set of variable (check) nodes incident to c_i (v_j), i.e., neighborhood of c_i (v_j)
$I_{ch}(j)$	Channel information of $v_j \in V$
$I_{EV}(i,j) \ (I_{EC}(i,j))$	Extrinsic information sent from v_j (c_i) to c_i (v_j)
$I_{AV}(i,j) \ (I_{AC}(i,j))$	A priori mutual information of v_j (c_i) sent from c_i (v_j), where c_i is an SPC node
$I_{AGC}(i) \ (I_{EGC}(i,j))$	A priori (extrinsic) mutual information of a GC node c_i
$I_{APP}(j)$	A posteriori probability of v_j
(μ, κ) component code	Component code with codelength μ and information size κ
$\mathcal{X}$	Index set of protograph VNs that are partially doped
N	Lifting factor
ν	Doping ratio
$I_{EV}^{(b_j)}(j)$	Extrinsic information from v_j to b_j, $j \in \mathcal{X}$
$I_{AGC}^{(b_j)}(j) \ (I_{EGC}^{(b_j)}(j))$	A priori (extrinsic) mutual information of b_j, $j \in \mathcal{X}$
G_c	Optimized protograph of the LDPC code
G_p	Initial irregular protograph that is used to construct the PD-GLDPC code
$\lambda_{G_c}(x) \ (\rho_{G_c}(x))$	VN (CN) degree distribution to construct G_c
$\mathbf{D^{dv}} = (a_1, \cdots, a_{max})$	$\lvert\mathbf{dv}\rvert$-sized vector defining the numbers of protograph VNs, where a_i is the number of protograph VNs of degree l_i and $\mathbf{dv}$ is a set of VN degrees that exist in the protograph, i.e., $\mathbf{dv} = \{l_1, l_2, \cdots, l_{max}\}$
ρ	Random puncturing ratio of the entire protograph
ρ_d	Random puncturing ratio for partially doped VNs in the protograph
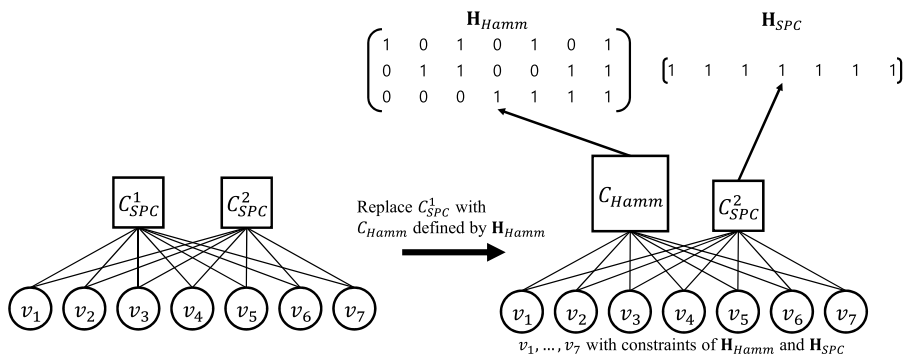
FIGURE 1. An example of protograph doped GLDPC code construction following [26] by replacing an SPC node with a GC node using the $(7, 4)$ Hamming code as the component code.

CNs, and GC nodes of protograph GLDPC codes. For an exact PEXIT analysis, tracking down each mutual information corresponding to edges of the component code is needed, i.e., multi-dimensional EXIT computation [27]. However, in terms of code optimization, where lots of EXIT computation is required, it is beneficial to reduce the complexity of the EXIT computation in GC nodes by averaging the a priori and extrinsic mutual information of the GC nodes. The EXIT and PEXIT analyses of the protograph doped GLDPC codes over

the BEC in terms of average mutual information are given in [28], [29], and [32].

The PEXIT process is given in Alg. 1. Let $I_{ch}(j)$ be the channel information from the erasure channel for the protograph VN v_j. In addition, $I_{EV}(i,j) \ (I_{EC}(i,j))$ is the extrinsic information sent from v_j (c_i) to c_i (v_j) and $I_{AV}(i,j) \ (I_{AC}(i,j))$ is the a priori mutual information of v_j (c_i) sent from c_i (v_j), where c_i is an SPC node. For GC nodes, we use the notations $I_{AGC}(i)$ and $I_{EGC}(i,j)$ for a priori and extrinsic information.

Let $N(c_i)$ $(N(v_j))$ be a set of variable (check) nodes incident to c_i (v_j), i.e., neighborhood of c_i (v_j). Finally, $I_{APP}(j)$ is a posteriori probability of v_j. To explain (1) in Alg. 1, if c_i is a GC node with the (n_i, k_i) Hamming code, the PEXIT of the GC node is computed from a closed form using the property of the simplex code, which is the dual code of a Hamming code. Also, $I_{AGC}(i) = \frac{1}{n_i} \sum_{j \in N(c_i)} b_{i,j} \times I_{EV}(i,j)$ is the average a priori mutual information for a GC node to compute the PEXIT message. In (1), we have

$$\tilde{e}_h = \sum_{t=1}^{h} t \sum_{u=0}^{t-1} (-1)^u 2^{\binom{u}{i}} \begin{bmatrix} k_i \\ t \end{bmatrix} \begin{bmatrix} t \\ u \end{bmatrix} \binom{2^{t-u}}{h}.$$

For two positive integers a and b, we also have $\binom{a}{b} = \prod_{i=0}^{b-1} \frac{a-i}{b-i}$ and $\begin{bmatrix} a \\ b \end{bmatrix} = \prod_{i=0}^{b-1} \frac{2^a - 2^i}{2^b - 2^i}$, where $\binom{a}{0} = 1$ and $\begin{bmatrix} a \\ 0 \end{bmatrix} = 1$. The PEXIT process searches for the minimum ϵ to successfully decode, i.e., $I_{APP}(j) = 1$, for all $j \in [n_v]$, in an asymptotic sense.

Now, we briefly explain the decoding process of GLDPC codes over the BEC [13]. The VNs process the conventional message-passing decoding over the BEC by sending correct extrinsic messages to the CNs if any of the incoming bits from their neighborhood is not erased. The SPC nodes send correct extrinsic messages to the VNs if all of their incoming messages are correctly received, and send erasure messages otherwise. In this paper, the decoding of GC nodes is processed by the maximum likelihood (ML) decoder. For each iteration, a GC node c_i with the (n_i, k_i) component code receives the set of erasure locations $\{e_i\}$ from $N(c_i)$. Let H_{GC} be the parity-check matrix (PCM) of the component code and H_e be the submatrix of H_{GC} indexed with $\{e_i\}$. The decoder computes the Gaussian-elimination operation of H_e, making it into a reduced row echelon form $H_e^{reduced}$. If $\text{rank}(H_e^{reduced}) = |\{e_i\}|$, the GC node solves all the input erasures and otherwise, the decoder corrects the erasures corresponding to the rows with weight 1 from $H_e^{reduced}$. The decoding complexity can be further reduced if the GC node exploits bounded distance decoding; however, the degradation of asymptotic performance is not negligible, as shown in [7].

III. THE PROPOSED PD-GLDPC CODES
In this section, a new construction method of protograph-based GLDPC codes is proposed. While the protograph doped GLDPC codes are constructed by replacing some protograph SPC nodes in the original protograph by GC nodes using the component code, the proposed PD-GLDPC codes are constructed by adding GC nodes for the subset of VNs using component codes after the lifting process of the original protograph, where each GC node is connected to the VNs copied from single protograph VN. A block diagram of the construction process of both codes together with the conventional random GLDPC code is given in Fig. 2.
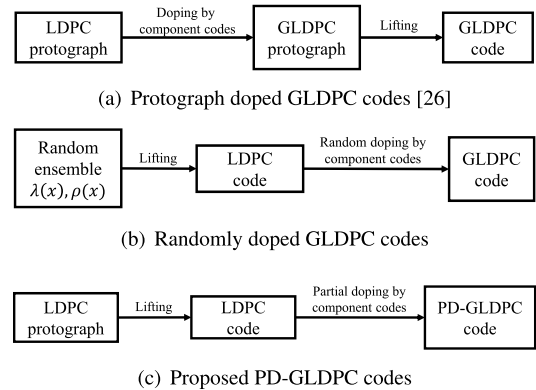

(a) Protograph doped GLDPC codes [26]

(b) Randomly doped GLDPC codes

(c) Proposed PD-GLDPC codes

FIGURE 2. A block diagram of the construction process of protograph doped GLDPC codes [26], randomly doped GLDPC codes, and the proposed PD-GLDPC codes.

A. CONSTRUCTION METHOD OF PD-GLDPC CODES
First, we define the partial doping for VNs using additional GC nodes to a lifted protograph, that is, the addition of rows for the PCM by the component code, where each GC node is incident to VNs copied from single protograph VN. Also, we define a partially doped protograph VN as the aforementioned single protograph VN for the GC node. While the term doping in protograph doped GLDPC codes is used in the perspective of CNs, we use the term partial doping in the perspective of VNs. While it is clear that the connection of GC nodes can be applied to VNs copied from multiple protograph VNs, we have limited the connection to a single protograph VN to compute an accurate EXIT of the PD-GLDPC codes. Let $\mathbf{B}_{n_c \times n_v}$ be an $n_c \times n_v$ base matrix, where some protograph VNs are partially doped with a (μ, κ) component code after the lifting process. Let $\mathcal{X}$ be a set that contains indices of protograph VNs that are partially doped, where each partially doped protograph VN is randomly doped by N/μ component codes after the lifting process. Then, the proposed PD-GLDPC codes are defined with the parameters $(\mathbf{B}_{n_c \times n_v}, \mu, \kappa, \mathcal{X})$. Although any component code can be used, we restrict the component code used in the paper as the (μ, κ) Hamming code and assume that μ divides the lifting factor N such that $N = \mu\beta$, where β is a positive integer. The VNs copied from $|\mathcal{X}|$ VNs in the protograph are partially doped by GC nodes. That is, in the proposed PD-GLDPC code construction, the N/μ GC nodes are randomly connected to N VNs lifted from each partially doped protograph VN. Thus, the proposed construction method can choose any protograph VNs to protect by partial doping. The code rate of the PD-GLDPC code ensemble with $(\mathbf{B}_{n_c \times n_v}, \mu, \kappa, \mathcal{X})$ is $1 - \frac{n_c + |\mathcal{X}| \cdot (\mu - \kappa)/\mu}{n_v}$. An example of the proposed construction is given in Fig. 3, which illustrates the doping process by a $(7, 4)$ component Hamming code over a protograph defined by a 2×3 base matrix.

The basic concept of the proposed construction is to focus on the protection of VNs lifted from single protograph VN. Although only a portion of VNs from a single protograph VN
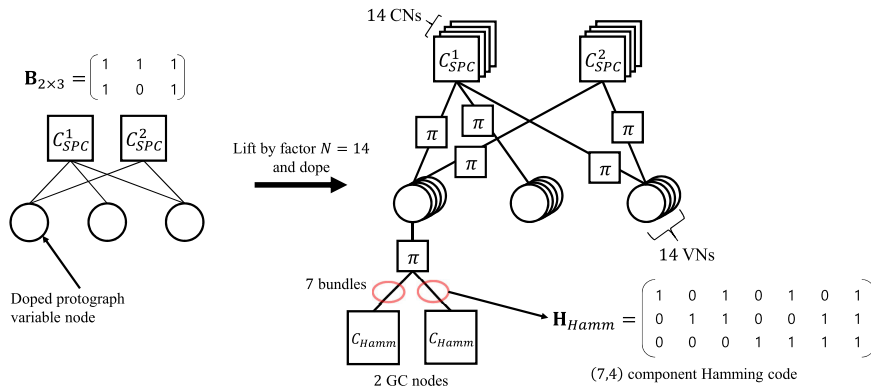
FIGURE 3. An example of a proposed ($\mathbf{B}_{2\times3}$, 7, 4, $\mathcal{X} = \{1\}$) PD-GLDPC code construction, where $\mathbf{B}_{2\times3} = [1\ 1\ 1;\ 1\ 0\ 1]$ and π is the 14 × 14 sized permutation matrix.

can be partially doped, for the exactness of EXIT computation and the typical minimum distance analysis, we have limited the doping process over the entire VNs lifted from a single protograph VN. Since N is the multiple of the component code length, all the VNs lifted from single protograph VN can be protected by using β GC nodes. A simple example for the PCM for β GC nodes, connected to the VNs lifted from single protograph VN, H_H is shown in Fig. 4(a), where H_{Hamm} is the PCM of the Hamming code. Although the PCM of the Hamming code can be applied randomly, a trivial representation of applying generalized constraints sequentially is given. The constructed PD-GLDPC code has a PCM $\mathbf{H}_{PD\text{-}GLDPC}$ as in Fig. 4(b), where the upper part is the PCM of the added βx GC nodes and the lower part $\mathbf{H}_{proto}$ refers to the PCM of the LDPC code lifted from the original protograph. Intuitively, H_H represents the PCM for each partially doped VN in the protograph and thus, the $x = |\mathcal{X}|$ bundles of matrices are diagonally appended to the PCM of the PD-GLDPC code. Since the doping proceeds after the lifting process, PD-GLDPC codes cannot be expressed in terms of a protograph. We define the doping ratio ν as the portion of GC nodes over the entire constraint nodes, i.e., $\nu = \frac{x\beta}{x\beta + n_c N}$. Also, we define the doping granularity as the minimum number of protograph VNs needed for doping. For the protograph doped GLDPC codes with (μ, κ) component code, the doping granularity is μ, whereas the proposed PD-GLDPC code has doping granularity 1. The finer doping granularity of the PD-GLDPC codes allows the construction of protograph-based GLDPC codes with the higher rate.

B. PEXIT ANALYSIS OF PD-GLDPC CODES
The PEXIT of the proposed PD-GLDPC codes is similar to that of the protograph doped GLDPC codes in Alg. 1 except for the EXIT of a GC node. Since the incoming mutual information of each GC node is obtained from only a single protograph VN in the proposed code, the average mutual information sent to each GC node is the same as the extrinsic message of the protograph VN connected to the GC node.

Let $b_j, j \in \mathcal{X}$ be the virtual node representing the set of β GC nodes connected to the protograph VN v_j. An example of the representation of a virtual node over a protograph is given in Fig. 5. Note that although b_j is not a protograph node itself, it is possible to compute the PEXIT of a PD-GLDPC code. Also, let $I_{EV}^{(b_j)}(j)$ be the extrinsic information from v_j to b_j expressed as

$$I_{EV}^{(b_j)}(j) = 1 - \epsilon \prod_{t \in N(v_j)} (1 - I_{AV}(t, j))^{b_{t,j}}, j \in \mathcal{X}.$$

Since b_j is solely connected to v_j, the index term for the extrinsic information from v_j to b_j is expressed by the notation of j only. In order to compute the EXIT of b_j, let $I_{AGC}^{(b_j)}(j)$ and $I_{EGC}^{(b_j)}(j)$ be the a priori and extrinsic mutual informations of b_j, respectively. Note that in an average sense, the EXIT of each GC node is computed from a single a priori mutual information to process the single value of the extrinsic mutual information for the neighboring VNs. Since b_j receives the extrinsic mutual information of v_j only, it is clear that $I_{AGC}^{(b_j)}(j) = I_{EV}^{(b_j)}(j), j \in \mathcal{X}$. We also compute the extrinsic mutual information from b_j to v_j denoted as $I_{EGC}^{(b_j)}(j)$ using (1), given the a priori mutual information $I_{AGC}^{(b_j)}(j)$, which is given as

$$I_{EGC}^{(b_j)}(j) = \frac{1}{\mu} \sum_{h=1}^{\mu} \left(1 - I_{AGC}^{(b_j)}(j)\right)^{h-1} \left(I_{AGC}^{(b_j)}(j)\right)^{\mu-h}$$
$$\times [h\tilde{e}_h - (\mu - h + 1)\tilde{e}_{h-1}]. \quad (2)$$

Note that for the proposed PD-GLDPC codes, the a priori (extrinsic) EXIT of the GC node is computed from the extrinsic (a priori) EXIT of single protograph VN. While the EXIT of VNs and SPC nodes for the proposed PD-GLDPC codes is the same as that of the protograph doped GLDPC codes described in Alg. 1, the EXIT of the GC nodes in the proposed codes is changed to (2) whereas the protograph doped GLDPC codes use (1) from Alg. 1.
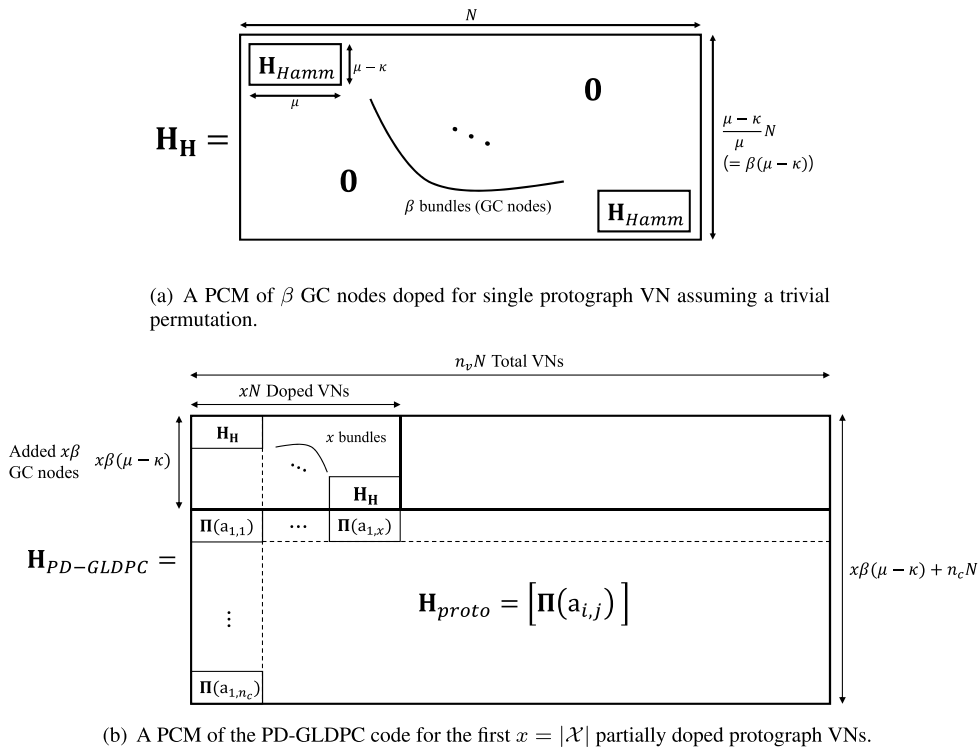
(a) A PCM of β GC nodes doped for single protograph VN assuming a trivial permutation.

(b) A PCM of the PD-GLDPC code for the first $x = |\mathcal{X}|$ partially doped protograph VNs.
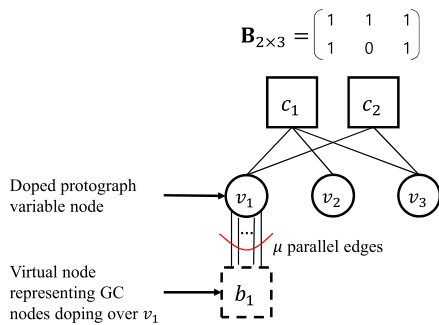
FIGURE 4. An exemplary PCM of a PD-GLDPC code.

FIGURE 5. An example of the Tanner graph representation of the proposed PD-GLDPC codes with a base matrix $\mathbf{B_{2\times3}} = [1\,1\,1; 1\,0\,1]$, where $\mathcal{X} = \{1\}$.

C. CONDITION FOR THE EXISTENCE OF THE TYPICAL MINIMUM DISTANCE OF THE PD-GLDPC CODE ENSEMBLE

The existence of a typical minimum distance in the given LDPC code ensemble defined in [33] guarantees that the minimum distance of its corresponding code grows linearly with the block length in an asymptotic sense [34]. To express it formally, if there exists a small number $\delta^* > 0$ such that the weight enumerators for a given code ensemble with weights less than or equal to $\delta^* n$ vanish as $n \to \infty$, then δ^* is the typical minimum distance of the code ensemble. It was proved in [35] that a protograph LDPC code ensemble has a typical minimum distance if there is no cycle consisting of only degree-2 VNs in the protograph. Furthermore, in [36], the condition for the existence of the typical minimum distance

of the protograph-based GLDPC code ensembles was given.

The proposed PD-GLDPC codes also have a similar approach to that of the protograph doped GLDPC codes in [36]. However, since a GC node of the proposed PD-GLDPC codes is not well defined by a protograph node, the derivation of the weight enumerator of the proposed codeword is quite different from that of the protograph doped GLDPC code. Thus, the condition for the existence of the typical minimum distance of the proposed PD-GLDPC code ensemble is slightly different from that of the protograph doped GLDPC code ensemble. In fact, we can regard the degree-2 VNs to be partially doped as the VNs with higher degrees. The detailed explanation for the existence of the typical minimum distance of the PD-GLDPC code ensemble is given in Appendix A. Then, we have the following theorem for the proposed PD-GLPDC codes.

Theorem 1: For the PD-GLDPC code ensemble of $(\mathbf{B}_{n_c \times n_v}, \mu, \kappa, \mathcal{X})$ without degree-1 VNs in $\mathbf{B}_{n_c \times n_v}$, the property of the typical minimum distance holds if the undoped degree-2 VNs in the protograph have no cycles among themselves.

Proof: The proof is given in the Appendix.

The existence of the typical minimum distance of the proposed PD-GLDPC code ensemble guarantees that the minimum distance of the proposed code grows linearly with the code length, and thus the proposed code is expected to have the low error floor for the large code length. In the next section, we use Theorem 1 as the constraint to optimize

the protograph in order to guarantee the existence of the typical minimum distance of the proposed PD-GLDPC code ensemble.

D. COMPARISON BETWEEN PROPOSED PD-GLDPC CODES AND PROTOGRAPH DOPED GLDPC CODES

The main difference between the proposed PD-GLDPC codes and the protograph doped GLDPC codes is the perspective of doping. While the protograph doped GLDPC codes replace an entire row in the protograph, i.e., a protograph CN by the PCM of the component code, the proposed PD-GLDPC codes append some rows incident to the VNs copied from single protograph VN. The focus of the protograph doped GLDPC codes is to choose a certain protograph CN to be replaced, whereas the PD-GLDPC codes focus on choosing which protograph VNs are further protected by partial doping. The constraint for the protograph doped GLDPC code is that the CNs to be replaced should have the degree equal to the component code length, while the constraint for the proposed PD-GLDPC codes is that the lifting size of a protograph should be the multiple of the component code length.

Furthermore, compared to the proposed PD-GLDPC codes, the protograph doped GLDPC codes have large doping granularity. By generalizing a single CN by a component code with parameters (μ, κ), the μ protograph VNs are doped assuming that the corresponding base entries are all ones for the protograph check node. Whereas, for every partial doping of β GC nodes in the proposed PD-GLDPC code, VNs copied from single protograph VN are partially doped. In other words, the doping granularity is 1, which is smaller than that of the protograph doped GLDPC codes. Since the doping granularity of protograph doped GLDPC codes is large, construction of the small protograph with capacity approaching performance is very difficult. In Section IV.B, we propose the construction method of PD-GLDPC codes with partial doping of $|\mathcal{X}|$ protograph VNs.

IV. OPTIMIZATION OF PD-GLDPC CODES

In this section, we introduce two optimization methods for the PD-GLDPC codes. The first subsection illustrates the construction method of protographs from the degree distribution of a random LDPC code ensemble in order to conduct comparison between LDPC codes and PD-GLDPC codes under the same degree distribution. The second subsection shows the optimization method of the protograph using the differential evolution algorithm in order to conduct comparison between LDPC codes and PD-GLDPC codes without any constraints.

A. DIFFERENTIAL EVOLUTION-BASED CODE CONSTRUCTION FROM THE DEGREE DISTRIBUTION OF RANDOM LDPC CODE ENSEMBLES

In general, as the portion of degree-2 VNs in the LDPC codes increases, the asymptotic performance is enhanced [37], but their minimum distance decreases and then the error floor becomes worse. For the construction of PD-GLDPC codes

in this subsection, we exploit the balance of the portion of degree-2 VNs, where we focus on the partial doping only for degree-2 VNs. The brief construction method is as follows. First, we construct the original base matrix $\mathbf{B}_{n_c \times n_v}$ with the large portion of degree-2 VNs. Then, we partially dope some of the protograph VNs of degree-2 to increase the minimum distance and improve their performance. Thus, irregular protographs with several degree-2 VNs are used for the construction of the proposed PD-GLDPC codes. In terms of irregular LDPC code ensembles, a large portion of degree-2 VNs enables the LDPC code to achieve the capacity approaching performance [38]. On the other hand, by reasonably selecting the number of partially doped VNs of degree-2, the property of the linear minimum distance growth with the length of the LDPC code can be guaranteed. Thus, when we design the proposed PD-GLDPC codes, balancing the partial doping over degree-2 VNs enables both the existence of a typical minimum distance and a good asymptotic performance. Optimization of irregular protograph LDPC code ensembles is made by initially obtaining the degree distribution of the random LDPC code ensemble using differential evolution [30] and constructing the protograph via the progressive edge growth (PEG) [39] algorithm for the construction of the proposed PD-GLPDC codes from irregular protographs. In this subsection, in order to make the CN degrees as even as possible, we try to construct the protograph from the degree distribution of a random LDPC code ensemble. We define G_c as the optimized protograph of the conventional LDPC code and G_p as the initial irregular protograph that is used to construct the PD-GLDPC code. That is, we can regard G_p as the protograph corresponding to H_{proto} in Fig. 4. In order to compare FER performances of the conventional LDPC code and the proposed PD-GLDPC code under the same degree distribution, G_c is constructed to have the same VN degree distribution as the PD-GLDPC code constructed from G_p after lifting by N.

Let $\lambda_{G_c}(x)$ and $\rho_{G_c}(x)$ be the VN and CN degree distributions of an irregular LDPC code ensemble to construct G_c, which is the optimized protograph for the conventional LDPC codes. In this subsection, we assume the degree distributions $\lambda_{G_c}(x) = \lambda_2 x + \lambda_3 x^2 + \lambda_4 x^3 + \lambda_5 x^4 + \lambda_6 x^5 + \lambda_l x^{l-1}$ and $\rho_{G_c}(x) = \rho_{r-1} x^{r-2} + \rho_r x^{r-1}$, where λ_i and ρ_i are the portions of edges of VNs and CNs of degree-i. Using the optimized degree distributions of $\lambda_{G_c}(x)$ and $\rho_{G_c}(x)$, a protograph G_c is constructed by the PEG algorithm. For the description of the protographs that construct the conventional LDPC codes and the proposed PD-GLDPC codes, let $\mathbf{D}^{\mathbf{dv}} = (a_1, \cdots, a_{max})$ be a $|\mathbf{dv}|$-sized vector defining the numbers of protograph VNs, where a_i is the number of protograph VNs of degree l_i and $\mathbf{dv} = \{l_1, l_2, \cdots, l_{max}\}$ is a set of VN degrees that exist in the protograph.

In order to make the same VN degree distributions of the LDPC codes constructed from G_c and the PD-GLDPC codes constructed from G_p after lifting by N, optimization of $\lambda_{G_c}(x)$ and $\rho_{G_c}(x)$ should be constrained by y_{max}, which is the maximum number of bulks of protograph VNs allowed

Algorithm 2 Construction of G_c and the PD-GLDPC Code

Input: $\mu, \kappa, n_v, n_c, R, l, r, y_{max}$

Output: y^{opt}, G_c, G_p

1: **Step 1) Optimize degree distribution of G_c**

Optimize $\lambda_{G_c}(x) = \lambda_2 x + \lambda_3 x^2 + \lambda_4 x^3 + \lambda_6 x^5 + \lambda_l x^{l-1}$ and $\rho_{G_c}(x) = \rho_{r-1} x^{r-2} + \rho_r x^{r-1}$ using differential evolution under constraints (a)~(c):

  (a) rate constraint $R = 1 - \frac{\int_0^1 \rho_{G_c}(x)dx}{\int_0^1 \lambda_{G_c}(x)dx}, 0 \leq \lambda_i \leq 1, 0 \leq \rho_i \leq 1$

  (b) typical minimum distance constraint $\frac{\lambda_2/2}{\Sigma} \times n_v \leq n_c - 1 - y_{max}(\mu - \kappa) \leftrightarrow \lambda_2 \leq \frac{2\Sigma\{n_c - 1 - y_{max}(\mu - \kappa)\}}{n_v}$

  (c) G_p existence constraint $\lambda_3 \geq \frac{12\Sigma y_{max}}{n_v}, \lambda_4 \geq \frac{24\Sigma y_{max}}{n_v}, \lambda_5 \geq \frac{20\Sigma y_{max}}{n_v}, \lambda_6 \geq \frac{6\Sigma y_{max}}{n_v}$

2: **Step 2) Construction of G_c**

From the optimized degree distribution and the random PEG algorithm, construct G_c defined as $\mathbf{D}^{(2,3,4,5,6,l)} = (a, b, c, d, e, f)$ guaranteeing a typical minimum distance.

3: **Step 3) Optimization of G_p**

For each $y = 1, 2, \cdots, y_{max}$, construct G_p defined as $\mathbf{D}^{(2,3,4,5,6,l)} = (a + 15y, b - 4y, c - 6y, d - 4y, e - y, f)$ and choose $y^{opt} \in \{y\}$ with the best threshold.

4: **Step 4) Typical minimum distance check of the PD-GLDPC code**

For the chosen y^{opt} and G_p, if there exists any cycle for the submatrix induced by undoped VNs of degree-2, go to **Step 2)**. Otherwise, output y^{opt} and G_p.

to be partially doped in G_p. Although doping granularity for the proposed PD-GLDPC code is 1, we consider doping for bulks of protograph VNs in order to easily match the code rate and degree distribution because the purpose of this subsection is comparing FER performances between the conventional LDPC code and the proposed PD-GLDPC code under the same degree distribution. A PD-GLDPC code is constructed by partially doping μy protograph VNs in G_p. Construction of a PD-GLDPC code from G_p is optimized by ranging the doping bulk y, $1 \leq y \leq y_{max}$. That is, we search for the optimal value y which maximizes the coding gain between the PD-GLDPC codes constructed from G_p and the conventional protograph LDPC codes constructed from G_c.

Conditions for the degree distributions in order to construct G_c are derived as follows. The conditions need to guarantee two criteria: i) the VN degree distributions of the protograph LDPC code constructed from G_c and the PD-GLDPC code constructed from G_p after lifting by N are the same and ii) a typical minimum distance exists for both code ensembles. In this subsection, we assume that partial doping is conducted for the first μy degree-2 protograph VNs without loss of generality due to randomness of the PEG algorithm. For the y bulks of partially doped protograph VNs using the PCM of the $(15, 11)$ Hamming code, the numbers of

protograph VNs in G_p should be

$$\mathbf{D}^{(2,3,4,5,6,l)} = (a + 15y, b - 4y, c - 6y, d - 4y, e - y, f).$$

Given that G_c is represented as $\mathbf{D}^{(2,3,4,5,6,l)} = (a, b, c, d, e, f)$, for the existence constraint, each element of $\mathbf{D}^{(2,3,4,5,6,l)}$ should be non-negative. The parameters $a \sim f$ are approximated by the PEG construction as

$$a \approx \lfloor n_v \frac{\lambda_2/2}{\Sigma} \rfloor, \quad b \approx \lfloor n_v \frac{\lambda_3/3}{\Sigma} \rfloor, \quad c \approx \lfloor n_v \frac{\lambda_4/4}{\Sigma} \rfloor,$$

$$d \approx \lfloor n_v \frac{\lambda_5/5}{\Sigma} \rfloor, \quad e \approx \lfloor n_v \frac{\lambda_6/6}{\Sigma} \rfloor, \text{ and } f \approx \lfloor n_v \frac{\lambda_l/l}{\Sigma} \rfloor,$$

where $\Sigma = \int_0^1 \lambda_{G_c}(x)dx$. For the realization of the protograph from the degree distribution using the PEG algorithm, if the summation $a + b + c + d + e + f$ is lower than n_v, the values of $a \sim f$ are added by 1 in order starting from the lowest VN degree until the summation is equal to n_v.

If G_c is determined for a given y_{max} as $\mathbf{D}^{(2,3,4,5,6,l)} = (a, b, c, d, e, f)$, where $a + b + c + d + e + f = n_v$, G_p defined by $\mathbf{D}^{(2,3,4,5,6,l)} = (a + 15y, b - 4y, c - 6y, d - 4y, e - y, f)$ can be constructed for $y = 1, \cdots, y_{max}$. By allowing the PEG algorithm of the VN degree distribution over a base matrix with size $\{n_c - (\mu - \kappa)y\} \times n_v$, both the code rate and the VN degree distributions for the LDPC codes constructed from G_c and the proposed PD-GLDPC codes constructed from G_p after lifting by N are matched. We search for the value of y, which has the best PEXIT threshold while having a typical minimum distance. The optimized doping value is denoted as y^{opt}. The construction of G_c and the PD-GLDPC code is described in Alg. 2.

The protograph of the conventional protograph LDPC code, G_c is made for $y_{max} = 5, 10, 15$ for the half-rate protograph LDPC code ensemble. The numerical results are summarized in Table 2, where the coding gain given for the proposed PD-GLDPC code is compared to the conventional protograph LDPC code with the equal degree distribution.

B. OPTIMIZATION OF PD-GLDPC CODES USING PROTOGRAPH DIFFERENTIAL EVOLUTION

In this subsection, we propose the optimization method using the differential evolution algorithm. Similar to the differential evolution algorithm in [37], we use the differential evolution algorithm to find the protograph with the optimized BEC threshold. The parameters for the differential evolution are given as follows. The number of generations of the algorithm g is set to 6000. Each entry of the base matrix can have the integer value varying from 0 to a positive integer t. The number of base matrices examined for each generation instance is defined as N_p. For a given base matrix size $n_c \times n_v$, we fix $N_p = 10 \cdot n_c n_v$. The mutation parameter F is fixed to 0.5 and α is a uniform random variable with the domain $[0, 1]$. Lastly, the crossover probability p_c is fixed to 0.88 in this paper.

We define the optimized PD-LDPC code ensemble as C_1 and the optimization algorithm is given in Alg. 3. It is clear that while the optimization process is the same as that of

TABLE 2. Simulation results for optimized PD-GLDPC codes from irregular protographs using Alg. 2, where $l = 20$, $n_v = 400$, $R = 1/2$.

y_{max}	$\lambda_{G_c}(x), \rho_{G_c}(x)$ (threshold)	G_c protograph $\mathbf{D}^{(2,3,4,5,6,20)}$ / G_c threshold	G_p protograph $\mathbf{D}^{(2,3,4,5,6,20)}, y^{opt}$ / PD-GLDPC threshold	Coding gain
5	$\lambda_{G_c}(x) = 0.2049x + 0.2489x^2 + 0.1150x^3$ $+ 0.074x^4 + 0.0210x^5 + 0.3363x^{19}$ $\rho_{G_c}(x) = 0.9735x^7 + 0.0265x^8$ (0.4815)	$(165, 134, 47, 23, 5, 26)$ / 0.4620	$(240, 114, 17, 3, 0, 26), y^{opt} = 5$ / 0.4699	0.0079
10	$\lambda_{G_c}(x) = 0.1894x + 0.2255x^2 + 0.1431x^3$ $+ 0.1191x^4 + 0.0357x^5 + 0.2872x^{19}$ $\rho_{G_c}(x) = 0.9908x^7 + 0.0012x^8$ (0.4696)	$(152, 121, 57, 38, 9, 23)$ / 0.4523	$(287, 85, 3, 2, 0, 23), y^{opt} = 9$ / 0.4638	0.0115
15	$\lambda_{G_c}(x) = 0.1632x + 0.1758x^2 + 0.2143x^3$ $+ 0.1827x^4 + 0.0543x^5 + 0.2098x^{19}$ $\rho_{G_c}(x) = 0.9940x^7 + 0.0060x^8$ (0.4476)	$(131, 94, 86, 59, 14, 16)$ / 0.4352	$(341, 38, 2, 3, 0, 16), y^{opt} = 14$ / 0.4534	0.0182

Algorithm 3 Differential Evolution Algorithm to Design the Base Matrix of the PD-GLDPC Codes

Input: $\mu, \kappa, n_c, n_v, \mathcal{X}, g, t, N_p, p_c, F, \alpha$

Output: $\mathbf{B}_{n_c \times n_v}$

1: **Initialization:** Set the initial base matrices $(\mathbf{B}_1, \ldots, \mathbf{B}_{N_p})$ each with size $n_c \times n_v$ randomly, where each entry is chosen from $\{0, \ldots, t\}$.

2: **for** $m = 1 : g$ **do**

3: **Mutation:** For each $k \in \{1, \ldots, N_p\}$, the mutation matrices are created through the interpolation as follow:

$$[\mathbf{M}_k]_{i,j} = [\mathbf{B}_{r_1}]_{i,j} + (F + \alpha(1 - F))([\mathbf{B}_{r_2}]_{i,j} - [\mathbf{B}_{r_3}]_{i,j}),$$

where $[\mathbf{A}]_{i,j}$ is the (i, j) element of the matrix $\mathbf{A}$ and indices $r_i \in [N_p]$, $i = 1, 2, 3$ are distinct and randomly selected. Each entry of $\mathbf{M}_k$ is replaced with the nearest integer in $\{0, \ldots, t\}$.

4: **Crossover:** For each $k \in \{1, \ldots, N_p\}$, create the trial matrices $\mathbf{M}'_k$ such that $[\mathbf{M}'_k]_{i,j} = [\mathbf{M}_k]_{i,j}$ with a probability p_c and $[\mathbf{M}'_k]_{i,j} = [\mathbf{B}_k]_{i,j}$ with probability $1 - p_c$. If $\mathbf{M}'_k$ contains any cycles only consisting of undoped degree-2 protograph VNs, $\mathbf{M}'_k$ is regenerated.

5: **Selection:** Each base matrix in the candidates for $(m+1)$th generation is chosen between $\mathbf{B}_k$ and $\mathbf{M}'_k$. If the threshold of $\mathbf{B}_k$ is larger than $\mathbf{M}'_k$, no update is made. Otherwise, update $\mathbf{B}_k$ to $\mathbf{M}'_k$.

6: **end for**

7: From $\mathbf{B}_k, k \in [N_p]$, choose the matrix with the best threshold value and output $\mathbf{B}_{n_c \times n_v}$.

---

the protograph LDPC codes, the indices of the partial doping represented by $\mathcal{X}$ are included, which show the protograph VNs that are doped by GC nodes. Although the indices of $\mathcal{X}$ can be arbitrarily selected for code constructions using the differential evolution algorithms, we fix the number of indices as small as possible. For applications on partially doping over a given protograph, algorithms selecting the indices of $\mathcal{X}$ can be made to optimize the performance of the code ensemble. Also, the criterion for the existence of the typical minimum distance derived in Theorem 1 is used

during the construction of new trial matrices for the proposed PD-GLDPC codes. The component code used in the following optimization is a (15, 11) Hamming code.

We optimize the protographs for the PD-GLDPC codes for base matrices with size 8×16 and 4×12. We set $\mathcal{X} = \{1, 2\}$ and $t = 5$ for $\mathbf{B}_{8 \times 16}$, $\mathcal{X} = \{1\}$ and $t = 3$ for $\mathbf{B}_{4 \times 12}$. Let $\mathbf{B}^{C_1}_{n_c \times n_v}$ be the resulting base matrix of the optimization for both cases. The optimized base matrix result of $\mathbf{B}^{C_1}_{8 \times 16}$ is given in (3), as shown at the bottom of the next page, where the BEC threshold is 0.5227 with the code rate 0.4667. Likewise, the result of $\mathbf{B}^{C_1}_{4 \times 12}$ is given in (4), as shown at the bottom of the next page, where the BEC threshold is 0.3397 with the code rate 0.6444. The bold parts in the matrix represent VNs that are partially doped. The results show that the VN with the highest degree is partially doped. From these optimization results, we can expect that partially doping VNs with high degree and puncturing some portion of them for rate matching can improve the performance of the proposed PD-GLDPC codes.

The approach of partially doping and puncturing is a similar techique to the precoding and puncturing. Precoding and puncturing high degree VNs in a protograph is a well known technique in order to enhance the threshold of protograph LDPC codes [40]. Precoding takes place by placing a CN between a degree-1 VN and a high degree VN. In order to compensate for the rate loss, the high degree VN is punctured. From some intuition of the proposed optimization results and well known concepts of precoding, we apply a similar approach of the precoding technique to the proposed PD-GLDPC codes.

We first define ρ_d as the portion of random puncturing for VNs that are doped. For the BEC, we use the concept in [41] to derive ρ_d. For a target code rate R^*, the random puncturing ratio ρ is $1 - \frac{R}{R^*}$. Thus, ρ_d is derived as $\rho_d = \rho \cdot \frac{n_v}{|\mathcal{X}|}$ and we use it for the computation of the EXIT during the optimization algorithm. The channel values for the partially doped VNs become $I_{ch}(j) = 1 - \{\rho_d + (1 - \rho_d)\epsilon\}$, $j \in \mathcal{X}$. Thus, it is possible to construct the PD-GLDPC codes for the target code rate by using the random puncturing method.

For the construction of PD-GLDPC codes with the target code rate $R^* = 1/2$, the base matrix $\mathbf{B}_{8 \times 16}$ is optimized using

Alg. 3 for $\mathcal{X} = \{1, 2\}$, $\rho_d = 0.5333$, and $t = 5$. Likewise, for the target code rate $R^* = 2/3$, the base matrix $\mathbf{B}_{4 \times 12}$ is optimized for $\mathcal{X} = \{1\}$, $\rho_d = 0.4058$, and $t = 3$. Let $\mathbf{B}_{n_c \times n_v}^{C_2}$ be the resulting base matrix for the optimized results of the protographs constructed by puncturing partially doped VNs. The optimized base matrices for both code rates are given as in (5) and (6), shown at the bottom of the page, where resulting base matrix for $R^* = 1/2$ is given in (5) and the resulting base matrix for $R^* = 2/3$ is given in (6). The resulting thresholds of the optimized base matrices are 0.4857 and 0.319 for target code rates $R^* = 1/2$ and $R^* = 2/3$, respectively. The optimization results show that the constructed PD-GLDPC codes have capacity approaching performances and the average VN density is reduced by huge amount compared to $\mathbf{B}_{n_c \times n_v}^{C_1}$. Since the base matrix $\mathbf{B}_{n_c \times n_v}^{C_2}$ is driven from the random puncturing of partially doped VNs, we define the constructed PD-GLDPC code ensemble with parameters $(\mathbf{B}_{n_c \times n_v}^{C_2}, \mu, \kappa, \mathcal{X}, \rho_d)$.

V. NUMERICAL RESULTS AND ANALYSIS

In this section, we propose the optimized protograph design and show the FER of the proposed PD-GLDPC codes. The performance of the conventional protograph LDPC code is compared with that of the proposed PD-GLDPC code. Two methods of comparison are conducted. The first subsection compares them under the same degree distribution using

Alg. 2. The second subsection compares the performance of the PD-GLDPC codes constructed without the degree distribution constraints using Alg. 3 to the state-of-the-art protograph LDPC codes.

A. SIMULATION RESULT FOR OPTIMIZED PD-GLDPC CODE FROM IRREGULAR RANDOM LDPC CODE ENSEMBLES

As the performance comparison with the existing GLDPC codes, we use the random GLDPC code ensemble with the threshold 0.466 in [17] that is represented as $\lambda(x) = 0.8x^2 + 0.01x^3 + 0.01x^5 + 0.18x^7$ and a doping ratio $\nu = 0.4$ by the Hamming code. Fig. 6 shows performance comparison of four half-rate codes which are AR4JA code [31], the irregular protograph LDPC code constructed from G_c, the random ensemble-based GLDPC code in [17], and the proposed PD-GLDPC code constructed from G_p in Table 2, where $y_{max} = 5$. All four codes in Fig. 6 are $(n, k) = (30000, 15000)$ codes of the half-rate, where G_c is defined as $\mathbf{D}^{(2,3,4,5,6,20)} = (165, 134, 47, 23, 5, 26)$ and has the same VN degree distribution as the PD-GLDPC code after lifting by $N = 75$. $x = \mu y = 75$ protograph VNs are partially doped in the PD-GLDPC code. For the constructed PD-GLDPC code, we have $\nu = \frac{x\beta}{x\beta + n_c N} = \frac{375}{375 + 15000} = 0.02439$. The constructed PD-GLDPC code for $y_{max} = 5$ has a coding gain of 0.0079 and 0.0039 compared to the GLDPC code

$$\mathbf{B}_{8 \times 16}^{C_1} = \begin{bmatrix} 5 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 5 & 5 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 5 & 0 & 1 & 0 & 2 & 1 & 5 & 5 & 0 & 0 & 5 & 5 & 0 & 5 & 0 & 3 \\ 5 & 3 & 0 & 1 & 1 & 1 & 5 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 2 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 4 & 1 & 3 & 0 & 0 & 2 & 1 & 0 \\ 4 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 & 0 & 0 & 1 & 0 & 3 & 0 & 0 & 1 & 1 & 1 & 1 & 0 \end{bmatrix} \tag{3}$$

$$\mathbf{B}_{4 \times 12}^{C_1} = \begin{bmatrix} 3 & 0 & 0 & 3 & 1 & 3 & 0 & 1 & 2 & 2 & 3 & 0 \\ 3 & 0 & 1 & 3 & 1 & 0 & 0 & 0 & 1 & 2 & 0 & 1 \\ 3 & 1 & 0 & 3 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 2 \\ 3 & 3 & 3 & 3 & 0 & 0 & 3 & 1 & 0 & 1 & 0 & 0 \end{bmatrix}. \tag{4}$$

$$\mathbf{B}_{8 \times 16}^{C_2} = \begin{bmatrix} 2 & 0 & 5 & 2 & 1 & 3 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 2 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 2 & 1 & 0 & 2 & 1 & 2 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 2 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 2 & 0 & 4 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 3 & 1 \\ 3 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}, \tag{5}$$

$$\mathbf{B}_{4 \times 12}^{C_2} = \begin{bmatrix} 2 & 0 & 0 & 1 & 0 & 0 & 3 & 2 & 1 & 0 & 2 & 2 \\ 2 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 2 & 2 \\ 3 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ 2 & 1 & 2 & 0 & 3 & 2 & 0 & 0 & 1 & 3 & 0 & 3 \end{bmatrix}, \tag{6}$$
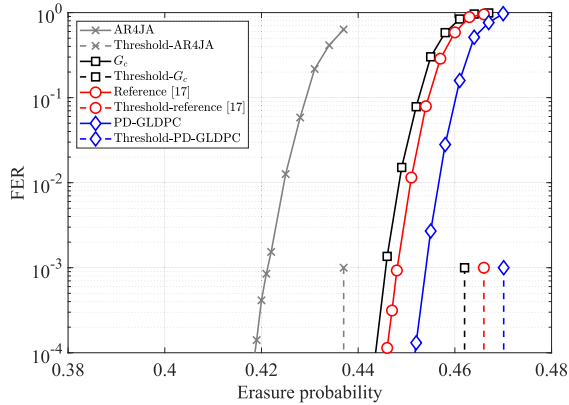
FIGURE 6. Comparison of the BEC threshold and FER for the LDPC codes constructed from AR4JA and G_c, the conventional random GLDPC code from the ensemble in [17], and the PD-GLDPC code from G_p for the code rate 1/2.

in [17] and the irregular protograph LDPC code from G_c, respectively. Fig. 6 shows that the proposed PD-GLDPC code has a good performance both in the waterfall and the low error floor region due to the fact that the code is optimized by increasing the doping as much as possible, and at the same time, the typical minimum distance constraint is satisfied. In terms of the asymptotic analysis, increasing the portion of degree-2 VNs increases the possibility of the code to approach the channel capacity [38]. However, the existence of a typical minimum distance of the protograph is also important, which upper bounds the portion of degree-2 VNs in the LDPC code. Thus, balancing the portion of degree-2 VNs is needed in order to satisfy both a typical minimum distance condition and a good threshold. The proposed PD-GLDPC code guarantees the balance of the degree-2 VNs by carefully choosing the rate of the protograph code and the number of doping on degree-2 VNs.

B. SIMULATION RESULTS FOR PD-GLDPC CODE FROM OPTIMIZED PROTOGRAPH

The proposed PD-GLDPC codes for $R^* = 1/2$ and $R^* = 2/3$ are constructed from the ensembles $(\mathbf{B}^{C_2}_{8\times16}, 15, 11, \{1, 2\}, 0.5333)$ and $(\mathbf{B}^{C_2}_{4\times12}, 15, 11, \{1\}, 0.4058)$, respectively. The protographs are shown in (5) and (6). The AR4JA [31] and block protograph codes in [31] and [37] of the same code rate are used for performance comparison. We first compare the threshold and average VN degree between the proposed PD-GLDPC code ensembles and the aforementioned protograph LDPC code ensembles in Table 3. The average VN degree of the PD-GLDPC codes considers both the base matrix and the edges added from the partial doping. The results show that the asymptotic performance of the proposed PD-GLDPC code ensemble outperforms the AR4JA and block protograph introduced in [31]. The average VN degree of the PD-GLDPC codes is low while having the asymptotic performance comparable to the capacity approaching protographs introduced in [37].
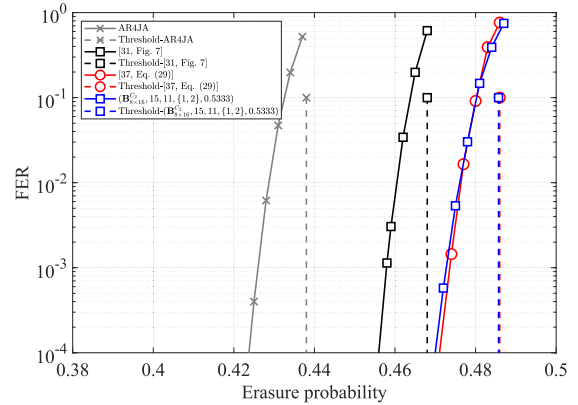
(a) $R^* = \frac{1}{2}$
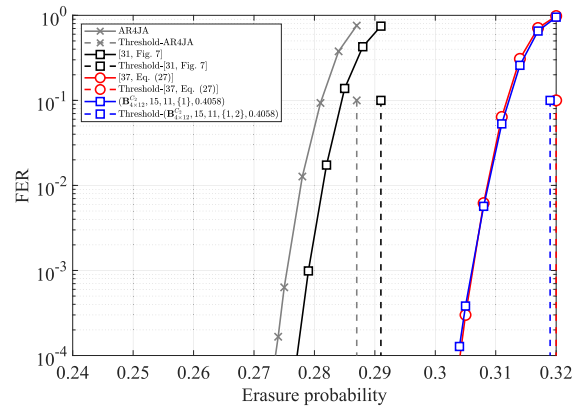
(b) $R^* = \frac{2}{3}$

FIGURE 7. FER comparison for the constructed codes from AR4JA [31], protograph [31, Fig. 7], protograph [37], and PD-GLDPC code ensemble $(\mathbf{B}^{C_2}_{n_C \times n_V}, 15, 11, \mathcal{X}, \rho_d)$.

By using the PEG algorithm, the protographs are lifted to construct (48000, 24000) PD-GLDPC code for $R^* = 1/2$. The protograph AR4JA and protographs in [31] and [37] are lifted to the same code length. The FER results are shown in Fig. 7(a). Likewise, the protograph of the PD-GLDPC, AR4JA, and [37] are lifted to construct (45000, 30000) codes for $R^* = 2/3$. The protograph in [31, Fig. 7] is lifted to blocklength near $n = 45000$. The FER results are shown in Fig. 7(b). The doping ratio ν for the PD-GLDPC codes is 0.016393 for both code rates $R^* = 1/2$ and $R^* = 2/3$. Also, the FER results of the proposed PD-GLDPC codes for both code rates $R^* = 1/2$ and $R^* = 2/3$ show tangible gain compared to the AR4JA code and protograph code in [31]. Also, the performance is comparable to the capacity approaching block LDPC code in [37]. The partial doping and puncturing technique, which is similar to the precoding technique, shows that the capacity approaching PD-GLDPC codes can be constructed with the relatively low average VN degree.

TABLE 3. Comparison for thresholds and average VN degrees of protographs for the BEC.

Code type	Code rate	Protograph size	Threshold	Average VN degree	Gap to capacity
AR4JA [31]	0.5	3×5	0.438	3	0.062
Protograph [31, Fig. 7]	0.5	4×8	0.468	4.25	0.032
Protograph [37]	0.5	8×16	0.486	5.25	0.014
PD-GLDPC	0.5	8×16	0.4857	4.58	0.0143
AR4JA [31]	0.67	3×7	0.287	3.29	0.046
Protograph [31, Fig. 7]	0.67	2×6	0.292	5	0.041
Protograph [37]	0.67	4×12	0.320	5.08	0.013
PD-GLDPC	0.67	4×12	0.319	4.09	0.014

VI. CONCLUSION

We proposed a new class of GLDPC codes called PD-GLDPC codes that has advantages of a finer doping granularity compared to the conventional protograph doped GLDPC codes. Also, we proposed two optimization algorithms for the PD-GLDPC codes: protographs constructed from random LDPC code ensembles and protographs for PD-GLDPC code ensembles constructed from genetic algorithms. Furthermore, we proposed the partially doping and puncturing technique. Using the proposed technique, the constructed PD-GLDPC codes have good FER performances compared to the popular protograph LDPC codes. Since it is possible to partially dope the protograph VNs with a granularity one, the rate loss is reduced from partial doping, and thus, GLDPC codes can have capacity approaching performance in the medium to high code rate regime. For future work, use of other component codes and protographs with degree-1 VNs can be studied. Also, constructions of PD-GLDPC codes by generalizing the partial doping process such as doping over multiple protograph VNs or doping only a portion of a protograph VN can be considered. Furthermore, new constructions of PD-GLDPC codes over additive white Gaussian noise channels can be made.

ABBREVIATIONS
AR4JA	Accumulate-Repeat-4-Jagged-Accumulate
BEC	Binary erasure channel
CN	Check node
EXIT	Extrinsic information transfer
FER	Frame error rate
GC	Generalized constraint
GLDPC	Generalized low-density parity-check
LDPC	Low-density parity-check
ML	Maximum likelihood
PCM	Parity-check matrix
PD-GLDPC	Partially doped GLDPC
PEG	Progressive edge growth
PEXIT	Protograph EXIT
SPC	Single parity-check
VN	Variable node

APPENDIX A
PROOF OF THEOREM 1
A proof for the constraint of the existence of a typical minimum distance for the proposed PD-GLDPC code ensemble

is given in this appendix. Similar to that in [36], a typical minimum distance is derived by the weight enumerator analysis over the lifted protograph. In order to use the notations in [36], we've distinguished the indexing notations during the enumeration for the partially doped VNs using $'$. Also, the c_j and v_i notations are used for the CNs and the VNs, respectively. Suppose that the proposed PD-GLDPC code is constructed from the protograph defined by $G = (V, C, E)$ and x VNs are partially doped, where component codes are identical with the parameters (μ, κ). We assume that the first x protograph VNs are partially doped without loss of generality. Then, we are given a VN set $V = \{v_1, \cdots, v_{n_v}\}$ and a CN set $C_{PD\text{-}GLDPC} = B \cup C = \{b_1, \cdots, b_x\} \cup \{c_1, \cdots, c_{n_c}\}$ for the protograph. It is important to note that the GC node set B is not defined over a protograph. However, the codeword enumeration can be made when the protograph is lifted, where $b_{i'}, i' \in [x]$ is a virtual CN that represents CNs of the component code used for partial doping for $v_{i'}$ in the original protograph. Although $b_{i'}$ is not a protograph CN, we define it for the enumeration of the partially doped protograph VNs. The PD-GLDPC code is constructed by lifting the graph G by N times and permuting the replicated edges. Each v_i (c_j) has degree q_{v_i} (q_{c_j}) in terms of G and each $b_{i'}$ has degree μ in terms of B. For the enumeration of the GC node $b_{i'}$, we can think of it as a protograph node of degree μ that is lifted by a factor of $\frac{N}{\mu}$. The upper bound of the weight enumerator of the proposed PD-GLDPC code ensemble with the weight d, denoted as $A_d^{PD\text{-}GLDPC}$ is derived as follows.

Let $w_{m,u}, u \in [q_{v_m}]$ be the uth edge weight from a VN v_m. For a partially doped VN $v_m, m \in [x]$, there are additional μ weights sent towards the incident GC node, where the uth weight is defined as $w'_{m,u}, u \in [\mu]$. For a given input weight vector $\mathbf{d} = (d_1, \cdots, d_{n_v})$, we need to calculate $A^{PD\text{-}GLDPC}(\mathbf{d})$ and sum it over every instance of $\mathbf{d}$ that satisfies $d = d_1 + \cdots + d_{n_v}$. For input $d_{i'}, i' \in [x]$, it is clear that $\sum_{i=1}^{\mu} w'_{m,i} = d_{i'}$ because the extrinsic weight $w'_{m,i}$ consists of weights solely from v_m. We introduce the following notations:

- $$A_{d_i}^{v_i}(\mathbf{w}_i) = \binom{N}{d_i} \delta_{d_i, w_{i,1}}, \cdots, \delta_{d_i, w_{i,q_{v_i}}}$$
$$= \begin{cases} \binom{N}{d_i}, & \text{if } w_{i,j} = d_i, \forall j \in [q_{v_i}] \\ 0, & \text{otherwise} \end{cases}$$

is the vector weight enumerator for a VN v_i of the protograph [36].

- $A^{c_j}(\mathbf{z}_j)$ is the vector weight enumerator for a CN c_j of the original protograph, for the incoming weight vector $\mathbf{z}_j = [z_{j,1}, \cdots, z_{j,q_{c_j}}]$ [36].

- $B_{d_{i'}}^{v_{i'}}(\mathbf{w}'_{i'}) = \begin{cases} 1, & \text{for } w'_{i',1} + \cdots + w'_{i',\mu} = d_{i'} \\ 0, & \text{otherwise} \end{cases}$ is the vector weight enumerator for partially doped VNs $v_{i'}$, $i' \in [x]$.

- $B^{b_{i'}}(\mathbf{w}'_{i'})$ is the vector weight enumerator for CNs that are created during the lifting process given the weight vector $\mathbf{w}'_{i'}$. $A^{b_{i'}}(d_{i'})$ is the summation of enumerators over all possible $\mathbf{w}'_{i'}$ values given that $w'_{i',1} + \cdots + w'_{i',\mu} = d_{i'}$ satisfying

$$A^{b_{i'}}(d_{i'}) = \sum_{\mathbf{w}'} B^{b_{i'}}(\mathbf{w}'_{i'})$$

$$= \sum_{\mathbf{w}'} \sum_{\{\mathbf{m}\}} C(\frac{N}{\mu}; m_1, \cdots, m_K),$$

where $\mathbf{w}' = (w'_1, \cdots, w'_\mu)$ such that $\sum_{i=1}^{\mu} w'_1 = d'_k$, $w'_i \le \frac{N}{\mu}$.

Then, the weight enumerator is given as

$$A_d^{PD-GLDPC} = \sum_{\{\mathbf{d}\}} A^{PD-GLDPC}(\mathbf{d}),$$

where

$$A^{PD\text{-}GLDPC}(\mathbf{d}) = \frac{\prod_{i=1}^{n_v} A_{d_i}^{v_i}(\mathbf{w}_i) \prod_{j=1}^{n_c} A^{c_j}(\mathbf{z}_j)}{\prod_{s=1}^{n_v} \prod_{r=1}^{q_{v_s}} \binom{N}{w_{s,r}}}$$

$$\times \frac{\prod_{i'=1}^{x} B_{d_{i'}}^{v_{i'}}(\mathbf{w}'_{i'}) B^{b_{i'}}(\mathbf{w}'_{i'})}{\prod_{s'=1}^{x} \prod_{r'=1}^{\mu} \binom{\frac{N}{\mu}}{w'_{s',r'}}}$$

$$= \sum_{\{\mathbf{w}'_{i'} : w'_{i',1} + \cdots + w'_{i',\mu} = d_{i'}\}} \frac{\prod_{j=1}^{n_c} A^{c_j}(\mathbf{d}_j)}{\prod_{i=1}^{n_v} \binom{N}{d_i}^{q_{v_i}-1}}$$

$$\times \frac{\prod_{i'=1}^{x} B^{b_{i'}}(\mathbf{w}'_{i'})}{\prod_{s'=1}^{x} \prod_{r'=1}^{\mu} \binom{\frac{N}{\mu}}{w'_{s',r'}}}.$$

The solution to the equation $\mathbf{w}' = \mathbf{m}\mathbf{M}^C$ is given as $\mathbf{m} = \{m_1, \cdots, m_K\}$. The term $\binom{\frac{N}{\mu}}{w'_{s',r'}}$ is lower bounded by $(\frac{N}{\mu})^{w'_{s',r'}} e^{-w'_{s',r'} \cdot \ln w'_{s',r'}}$. Then, $A^{PD-GLDPC}(\mathbf{d})$ can be upper bounded as in (7), shown at the bottom of the next page, where $P = \sum_{s'=1}^{x} d_{s'}$ is the total weight of the x partially doped VNs. Then $\sum_t(t \cdot \ln t) \le (\sum_t t) \cdot \ln(\sum_t t)$ is used for the second and the third inequalities in (7). It was shown in (18) of [36] that the inequality

$$\frac{\prod_{j=1}^{n_c} A^{c_j}(\mathbf{d}_j)}{\prod_{i=1}^{n_v} \binom{N}{d_i}^{q_{v_i}-1}} \le \prod_{i=1}^{n_v} e^{(q_{v_i}-1-\frac{q_{v_i}}{d_{min}^{(c)}})d_i \ln \frac{d_i}{N} + \frac{q_{v_i}(2+k_{max}^{(c)}\ln 2)}{d_{min}^{(c)}}d_i}$$

holds, where $d_{min}^{(c)}$ is the minimum distance of an SPC component code for the original protograph and $k_{max}^{(c)}$ is the maximum number of codewords of an SPC component code.

Using the similar notations in [36], let $d_{min}^{(b)}$ and $k^{(b)}$ be the minimum distance and the number of codewords of the (μ, κ) component code for the GC nodes. Then, $\prod_{i'=1}^{x} A^{b_{i'}}(d_{i'})$ is upper bounded as in (8), shown at the bottom of the next page. For the inequality in the third line of (8), we use the fact that $\sum_{i=1}^{p} t_i \ln t_i \le s \cdot \ln \frac{s}{p}$ with $s = t_1 + \cdots + t_p$, which is clear by using the derivative on the multivariable function that consists of independent t_i's. The equality is satisfied when all t_i values are the same. Going back to (7), let $f(P) = \frac{1}{\binom{N}{\mu}^P e^{-P \ln P}}$ for convenience. Then we can organize the inequality as in (9), shown at the bottom of the next page. We classify the VNs in the protograph into three groups before doping:

- Protograph VNs of degrees higher than 2
- Protograph VNs of degree-2 to be partially doped
- Protograph VNs of degree-2 not to be partially doped.

We also separate the weights of codewords after lifting into three parts according to the three groups of VNs: u_i, p_z, and l_j, where u_i is the weight of the sub-codeword corresponding to a protograph VN v_i of degree higher than 2 and p_z and l_j are the weights of the sub-codewords of each partially doped and undoped protograph VN v_z and v_j of degree-2 from the protograph, respectively. The sum of sub-codeword weights for each group of VNs is given as $U = \sum_i u_i$, $P = \sum_z p_z$, and $L = \sum_j l_j$. It is clear that for the total codeword weight d, $d = U + P + L$. Then, the upper bound of the first term in (9) is written as

$$\prod_{i=1}^{n_v} e^{(q_{v_i}-1-\frac{q_{v_i}}{d_{min}^{(c)}})d_i \ln \frac{d_i}{N} + \frac{q_{v_i}(2+k_{max}^{(c)}\ln 2)}{d_{min}^{(c)}}d_i}$$

$$\le e^{(2-\frac{3}{d_{min}^{(c)}})(d-P-L)\ln \frac{d-P-L}{N} + \frac{3(2+k_{max}^{(c)}\ln 2)}{d_{min}^{(c)}} \cdot (d-P-L)}$$

$$\times e^{\frac{2(2+k_{max}^{(c)}\ln 2)}{d_{min}^{(c)}} \cdot L} e^{\frac{2(2+k_{max}^{(c)}\ln 2)}{d_{min}^{(c)}} \cdot P}, \quad (10)$$

which is derived by using three weight groups of codewords similar to (20) of [36]. We share the same inequality $u_i < Ne^{-\frac{(2+k_{max}^{(c)}\ln 2)}{d_{min}^{(c)}-1}}$ over the given codeword weight d as in [36]. The upper bound of the second $\prod$ term of (9) can be derived as

$$\prod_{i'=1}^{x} \left(\frac{N}{\mu}\right)^{\frac{1}{d_{min}^{(b)}}d_{i'}} e^{\frac{(2+k^{(b)}\ln 2)}{d_{min}^{(b)}}d_{i'} - \frac{1}{d_{min}^{(b)}}d_{i'}\ln \frac{d_{i'}}{\mu}}$$

$$= \prod_{i'=1}^{x} e^{\frac{1}{d_{min}^{(b)}}d_{i'}\ln \frac{N}{\mu}} e^{\frac{(2+k^{(b)}\ln 2)}{d_{min}^{(b)}}d_{i'} - \frac{1}{d_{min}^{(b)}}d_{i'}\ln \frac{d_{i'}}{\mu}}$$

$$= \prod_{i'=1}^{x} e^{\frac{1}{d_{min}^{(b)}}d_{i'}\ln \frac{N}{d_{i'}}} e^{\frac{(2+k^{(b)}\ln 2)}{d_{min}^{(b)}}d_{i'}}$$

$$= \prod_{z=1}^{x} e^{\frac{1}{d_{min}^{(b)}}p_z\ln \frac{N}{p_z}} e^{\frac{(2+k^{(b)}\ln 2)}{d_{min}^{(b)}}p_z}$$

$$\le e^{\frac{1}{d_{min}^{(b)}}P \cdot \ln \frac{Nx}{P}} e^{\frac{(2+k^{(b)}\ln 2)}{d_{min}^{(b)}}P}. \quad (11)$$

Using (10) and (11), the upper bound of $A^{PD\text{-}GLDPC}(\mathbf{d})$ is derived in terms of $E(d, P, L)$ as follows:

$$A^{PD\text{-}GLDPC}(\mathbf{d})$$
$$\leq e^{\frac{1}{d_{min}^{(b)}} P \cdot \ln \frac{Nx}{P}} e^{\frac{(2+k^{(b)}\ln 2)}{d_{min}^{(b)}} P}$$

$$\times e^{(2-\frac{3}{d_{min}^{(c)}})(d-P-L)\ln \frac{d-P-L}{N} + \frac{3(2+k_{max}^{(c)}\ln 2)}{d_{min}^{(c)}}\cdot(d-P-L)}$$

$$\times e^{\frac{2(2+k_{max}^{(c)}\ln 2)}{d_{min}^{(c)}}\cdot(P+L)} e^{x(\mu-1)} e^{P} f(P). \qquad (12)$$

Let $E(d, P, L)$ be the parameter satisfying $A^{PD\text{-}GLDPC}(\mathbf{d}) \leq e^{x(\mu-1)} \cdot e^{E(d,P,L)}$. Then, from the upper bound in (12),

$$
\begin{aligned}
A^{PD\text{-}GLDPC}(\mathbf{d}) &\leq \sum_{\{\mathbf{w}'_{i'}: w'_{i',1}+\cdots+w'_{i',\mu}=d_{i'}\}} \frac{\prod_{j=1}^{n_c} A^{c_j}(\mathbf{d}_j) \times \prod_{i'=1}^{x} B^{b_{i'}}(\mathbf{w}'_{i'})}{\prod_{i=1}^{n_v} \binom{N}{d_i}^{q_{v_i}-1} \times \prod_{s'=1}^{x}\prod_{r'=1}^{\mu} (\frac{N}{\mu})^{w'_{s',r'}} e^{-w'_{s',r'}\cdot \ln w'_{s',r'}}} \\
&\leq \sum_{\{\mathbf{w}'_{i'}: w'_{i',1}+\cdots+w'_{i',\mu}=d_{i'}\}} \frac{\prod_{j=1}^{n_c} A^{c_j}(\mathbf{d}_j) \times \prod_{i'=1}^{x} B^{b_{i'}}(\mathbf{w}'_{i'})}{\prod_{i=1}^{n_v} \binom{N}{d_i}^{q_{v_i}-1} \times \prod_{s'=1}^{x} (\frac{N}{\mu})^{d_{s'}} e^{-d_{s'}\cdot \ln d_{s'}}} \\
&\leq \sum_{\{\mathbf{w}'_{i'}: w'_{i',1}+\cdots+w'_{i',\mu}=d_{i'}\}} \frac{\prod_{j=1}^{n_c} A^{c_j}(\mathbf{d}_j) \times \prod_{i'=1}^{x} B^{b_{i'}}(\mathbf{w}'_{i'})}{\prod_{i=1}^{n_v} \binom{N}{d_i}^{q_{v_i}-1} \times (\frac{N}{\mu})^{P} e^{-P\cdot \ln P}} \\
&= \frac{\prod_{j=1}^{n_c} A^{c_j}(\mathbf{d}_j) \times \prod_{i'=1}^{x} \sum_{\{\mathbf{w}'_{i'}: w'_{i',1}+\cdots+w'_{i',\mu}=d_{i'}\}} B^{b_{i'}}(\mathbf{w}'_{i'})}{\prod_{i=1}^{n_v} \binom{N}{d_i}^{q_{v_i}-1} \times (\frac{N}{\mu})^{P} e^{-P\cdot \ln P}} \\
&= \frac{\prod_{j=1}^{n_c} A^{c_j}(\mathbf{d}_j) \times \prod_{i'=1}^{x} A^{b_{i'}}(d_{i'})}{\prod_{i=1}^{n_v} \binom{N}{d_i}^{q_{v_i}-1} \times (\frac{N}{\mu})^{P} e^{-P\cdot \ln P}}, \qquad (7)
\end{aligned}
$$

$$
\begin{aligned}
\prod_{i'=1}^{x} A^{b_{i'}}(d_{i'}) &\leq \prod_{i'=1}^{x} \sum_{\{\mathbf{w}'_{i'}: w'_{i',1}+\cdots+w'_{i',\mu}=d_{i'}\}} \prod_{i=1}^{\mu} (\frac{N}{\mu})^{\frac{1}{d_{min}^{(b)}} w'_{i',i}} e^{\frac{(2+k'_{i'}\ln 2)}{d_{min}^{(b)}} w'_{i',i} - \frac{1}{d_{min}^{(b)}} w'_{i',i} \ln w'_{i',i}} \\
&= \prod_{i'=1}^{x} \sum_{\{\mathbf{w}'_{i'}: w'_{i',1}+\cdots+w'_{i',\mu}=d_{i'}\}} (\frac{N}{\mu})^{\frac{1}{d_{min}^{(b)}} d_{i'}} e^{\frac{(2+k'_{i'}\ln 2)}{d_{min}^{(b)}} d_{i'} - \sum_{i=1}^{\mu}\frac{1}{d_{min}^{(b)}} w'_{i',i}\ln w'_{i',i}} \\
&\leq \prod_{i'=1}^{x} \sum_{\{\mathbf{w}'_{i'}: w'_{i',1}+\cdots+w'_{i',\mu}=d_{i'}\}} (\frac{N}{\mu})^{\frac{1}{d_{min}^{(b)}} d_{i'}} e^{\frac{(2+k'_{i'}\ln 2)}{d_{min}^{(b)}} d_{i'} - \frac{1}{d_{min}^{(b)}} d_{i'}\ln \frac{d_{i'}}{\mu}} \\
&\leq \prod_{i'=1}^{x} \binom{d_{i'}+\mu-1}{d_{i'}} (\frac{N}{\mu})^{\frac{1}{d_{min}^{(b)}} d_{i'}} e^{\frac{(2+k'_{i'}\ln 2)}{d_{min}^{(b)}} d_{i'} - \frac{1}{d_{min}^{(b)}} d_{i'}\ln \frac{d_{i'}}{\mu}} \qquad (8)
\end{aligned}
$$

$$
\begin{aligned}
A^{PD\text{-}GLDPC}(\mathbf{d}) &\leq \prod_{i=1}^{n_v} e^{(q_{v_i}-1-\frac{q_{v_i}}{d_{min}^{(c)}})d_i\ln \frac{d_i}{N} + \frac{q_{v_i}(2+k_{max}^{(c)}\ln 2)}{d_{min}^{(c)}}d_i} \\
&\quad \times \prod_{i'=1}^{x} e^{d_{i'}+\mu-1}(\frac{N}{\mu})^{\frac{1}{d_{min}^{(b)}}d_{i'}} e^{\frac{(2+k'_{i'}\ln 2)}{d_{min}^{(b)}}d_{i'} - \frac{1}{d_{min}^{(b)}}d_{i'}\ln \frac{d_{i'}}{\mu}} \times f(P) \\
&\leq \prod_{i=1}^{n_v} e^{(q_{v_i}-1-\frac{q_{v_i}}{d_{min}^{(c)}})d_i\ln \frac{d_i}{N} + \frac{q_{v_i}(2+k_{max}^{(c)}\ln 2)}{d_{min}^{(c)}}d_i} \\
&\quad \times e^{x(\mu-1)} e^{P} \prod_{i'=1}^{x} (\frac{N}{\mu})^{\frac{1}{d_{min}^{(b)}}d_{i'}} e^{\frac{(2+k^{(b)}\ln 2)}{d_{min}^{(b)}}d_{i'} - \frac{1}{d_{min}^{(b)}}d_{i'}\ln \frac{d_{i'}}{\mu}} \times f(P). \qquad (9)
\end{aligned}
$$

$E(d, P, L)$ is given as

$$E(d, P, L) = \frac{1}{d_{min}^{(b)}} P \cdot ln \frac{Nx}{P} + \frac{(2 + k^{(b)} ln 2)}{d_{min}^{(b)}} P$$
$$+ (2 - \frac{3}{d_{min}^{(c)}})(d - P - L) ln \frac{d - P - L}{N}$$
$$+ \frac{3(2 + k_{max}^{(c)} ln 2)}{d_{min}^{(c)}}(d - P - L)$$
$$+ \frac{2(2 + k_{max}^{(c)} ln 2)}{d_{min}^{(c)}}(P + L)$$
$$+ P + P ln P - P ln \frac{N}{\mu}.$$

Assuming that there are no cycles consisting only of undoped VNs of degree-2, we can further ignore the existence of type 1 degree-2 VNs defined in [36] for undoped degree-2 VNs. Thus, we use the result of (22) in [36] such that the inequality $l_{2,k}^{(c_j)} \leq \frac{1}{d_{min}^{(c_j)}}(L_2^{(c_j)} + \sum_i w_i^{(c_j)})$ is satisfied for all $j \in [n_c]$, where $l_{2,k}^{(c_j)}$ is the weight of the degree-2 undoped VN of the original protograph and the total weight of them is denoted as $L_2^{(c_j)}$ for a CN c_j. Similar to the result in [36], we can derive the upper bound $L \leq \gamma(U + P)$, which is the same as $L \leq \frac{\gamma}{1+\gamma} d$. Now, the upper bound of $E(d, P, L)$ needs to be derived for independent values L and P. The first and second partial derivatives of $E(d, P, L)$ by P are given as

$$\frac{dE}{dP} = \frac{1}{d_{min}^{(b)}} ln \frac{Nx}{eP} + \frac{(2 + k^{(b)} ln 2)}{d_{min}^{(b)}}$$
$$- (2 - \frac{3}{d_{min}^{(c)}}) ln \frac{e(d - P - L)}{N} - \frac{(2 + k_{max}^{(c)} ln 2)}{d_{min}^{(c)}}$$
$$+ 1 + ln P + 1 - ln \frac{N}{\mu} < 0,$$
$$\frac{d^2 E}{dP^2} = -\frac{1}{d_{min}^{(b)} P} + (2 - \frac{3}{d_{min}^{(c)}})\frac{1}{d - P - L} + \frac{1}{P} > 0.$$

Since the first derivative over P is negative and the second derivative is positive, $E(d, P, L)$ is upper bounded by

$$\lim_{P \to 0+} E(d, P, L) = (2 - \frac{3}{d_{min}^{(c)}})(d - L) ln \frac{d - L}{N}$$
$$+ \frac{3(2 + k_{max}^{(c)})}{d_{min}^{(c)}}(d - L)$$
$$+ \frac{2(2 + k_{max}^{(c)} ln 2)}{d_{min}^{(c)}} L.$$

Since the resulting upper bound of $E(d, L)$ is the same as (37) in [36], the rest of the proof is the same as that in [36] and thus the proposed constraint guarantees the existence of typical minimum distance of the proposed PD-GLDPC code ensemble.

REFERENCES

[1] R. G. Gallager, "Low-density parity-check codes," *IRE Trans. Inf. Theory*, vol. 8, no. 1, pp. 21–28, Jan. 1962.

[2] D. J. C. MacKay and R. M. Neal, "Near Shannon limit performance of low density parity check codes," *Electron. Lett.*, vol. 32, no. 18, pp. 1645–1646, Aug. 1996.

[3] R. M. Tanner, "A recursive approach to low complexity codes," *IEEE Trans. Inf. Theory*, vol. IT-27, no. 5, pp. 533–547, Sep. 1981.

[4] G. Liva and W. E. Ryan, "Short low-error-floor Tanner codes with Hamming nodes," in *Proc. IEEE Mil. Commun. Conf. (MILCOM)*, Oct. 2005, pp. 208–213.

[5] S. Abu-Surra, D. Divsalar, and W. E. Ryan, "Enumerators for protograph-based ensembles of LDPC and generalized LDPC codes," *IEEE Trans. Inf. Theory*, vol. 57, no. 2, pp. 858–886, Feb. 2011.

[6] I. P. Mulholland, E. Paolini, and M. F. Flanagan, "Design of LDPC code ensembles with fast convergence properties," in *Proc. IEEE Int. Black Sea Conf. Commun. Netw. (BlackSeaCom)*, May 2015, pp. 53–57.

[7] Y. Liu, P. M. Olmos, and T. Koch, "A probabilistic peeling decoder to efficiently analyze generalized LDPC codes over the BEC," *IEEE Trans. Inf. Theory*, vol. 65, no. 8, pp. 4831–4853, Aug. 2019.

[8] M. Lentmaier and K. S. 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.

[9] G. Yue, L. Ping, and X. Wang, "Generalized low-density parity-check codes based on Hadamard constraints," *IEEE Trans. Inf. Theory*, vol. 53, no. 3, pp. 1058–1079, Mar. 2007.

[10] N. Miladinovic and M. Fossorier, "Generalized LDPC codes with Reed–Solomon and BCH codes as component codes for binary channels," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, Nov. 2005, pp. 1–6.

[11] D. G. M. Mitchell, M. Lentmaier, and D. J. Costello, "On the minimum distance of generalized spatially coupled LDPC codes," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2013, pp. 1874–1878.

[12] P. M. Olmos, D. G. M. Mitchell, and D. J. Costello, "Analyzing the finite-length performance of generalized LDPC codes," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2015, pp. 2683–2687.

[13] D. G. M. Mitchell, P. M. Olmos, M. Lentmaier, and D. J. Costello, "Spatially coupled generalized LDPC codes: Asymptotic analysis and finite length scaling," *IEEE Trans. Inf. Theory*, vol. 67, no. 6, pp. 3708–3723, Jun. 2021.

[14] E. Paolini, M. P. C. Fossorier, and M. Chiani, "Generalized and doubly generalized LDPC codes with random component codes for the binary erasure channel," *IEEE Trans. Inf. Theory*, vol. 56, no. 4, pp. 1651–1672, Apr. 2010.

[15] Y. Wang and M. Fossorier, "Doubly generalized LDPC codes," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2006, pp. 669–673.

[16] Y. Wang and M. Fossorier, "Doubly generalized LDPC codes over the AWGN channel," *IEEE Trans. Commun.*, vol. 57, no. 5, pp. 1312–1319, May 2009.

[17] R. Guan and L. Zhang, "Hybrid Hamming GLDPC codes over the binary erasure channel," in *Proc. 11th IEEE Int. Conf. Anti-Counterfeiting, Secur., Identificat. (ASID)*, Oct. 2017, pp. 130–133.

[18] Y. Fang, G. Bi, Y. L. Guan, and F. C. M. Lau, "A survey on protograph LDPC codes and their applications," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 4, pp. 1989–2016, 4th Quart., 2015.

[19] M. Stinner and P. M. Olmos, "On the waterfall performance of finite-length SC-LDPC codes constructed from protographs," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 2, pp. 345–361, Feb. 2016.

[20] Z. Yang, Y. Fang, G. Cai, G. Zhang, and P. Chen, "Design and optimization of tail-biting spatially coupled protograph LDPC codes under shuffled belief-propagation decoding," *IEEE Commun. Lett.*, vol. 24, no. 7, pp. 1378–1382, Jul. 2020.

[21] P. W. Zhang, F. C. M. Lau, and C.-W. Sham, "Spatially coupled PLDPC-Hadamard convolutional codes," *IEEE Trans. Commun.*, early access, Jul. 27, 2022, doi: 10.1109/TCOMM.2022.3194187.

[22] P.-W. Zhang, F. C. M. Lau, and C.-W. Sham, "Protograph-based LDPC Hadamard codes," *IEEE Trans. Commun.*, vol. 69, no. 8, pp. 4998–5013, Aug. 2021.

[23] P. W. Zhang, F. C. M. Lau, and C.-W. Sham, "Layered decoding for protograph-based low-density parity-check Hadamard codes," *IEEE Commun. Lett.*, vol. 25, no. 6, pp. 1776–1780, Jun. 2021.

[24] A. K. Pradhan and A. Thangaraj, "Protograph LDPC codes with block thresholds: Extension to degree-one and generalized nodes," *IEEE Trans. Commun.*, vol. 66, no. 12, pp. 5876–5887, Dec. 2018.

[25] J. Thorpe, "Low-density parity-check (LDPC) codes constructed from protographs," Jet Propulsion Lab., Pasadena, CA, USA, INP Progress Rep. 42–154, 2003.

[26] G. Liva, W. E. Ryan, and M. Chiani, "Quasi-cyclic generalized LDPC codes with low error floors," *IEEE Trans. Commun.*, vol. 56, no. 1, pp. 49–57, Jan. 2008.

[27] M. Lentmaier, M. B. S. Tavares, and G. P. Fettweis, "Exact erasure channel density evolution for protograph-based generalized LDPC codes," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2009, pp. 566–570.

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

[29] G. Liva and M. Chiani, "Protograph LDPC codes design based on EXIT analysis," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, Nov. 2007, pp. 3250–3254.

[30] A. Shokrollahi and R. Storn, "Design of efficient erasure codes with differential evolution," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2000, pp. 413–427.

[31] D. Divsalar, S. Dolinar, and C. Jones, "Construction of protograph LDPC codes with linear minimum distance," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2006, pp. 664–668.

[32] Y. Yu, Y. Han, and L. Zhang, "Hamming-GLDPC codes in BEC and AWGN channel," in *Proc. 6th Int. Conf. Wireless, Mobile Multi-Media (ICWMMN)*, 2015, pp. 103–106.

[33] C. Di, A. Montanari, and R. Urbanke, "Weight distributions of LDPC code ensembles: Combinatorics meets statistical physics," in *Proc. Int. Symp. Inf. Theory, (ISIT)*, 2004, p. 102.

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

[35] S. Abu-Surra, D. Divsalar, and W. E. Ryan, "On the existence of typical minimum distance for protograph-based LDPC codes," in *Proc. Inf. Theory Appl. Workshop (ITA)*, Jan. 2010, pp. 1–7.

[36] S. Abu-Surra, D. Divsalar, and W. E. Ryan, "On the typical minimum distance of protograph-based generalized LDPC codes," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2010, pp. 1–7.

[37] A. K. Pradhan, A. Thangaraj, and A. Subramanian, "Construction of near-capacity protograph LDPC code sequences with block-error thresholds," *IEEE Trans. Commun.*, vol. 64, no. 1, pp. 27–37, Jan. 2016.

[38] C. Di, T. J. Richardson, and R. L. Urbanke, "Weight distribution of low-density parity-check codes," *IEEE Trans. Inf. Theory*, vol. 52, no. 11, pp. 4839–4855, Nov. 2006.

[39] X.-Y. Hu, E. Eleftheriou, and D. M. Arnold, "Regular and irregular progressive edge-growth tanner graphs," *IEEE Trans. Inf. Theory*, vol. 51, no. 1, pp. 386–398, Jan. 2005.

[40] D. Divsalar, C. Jones, S. Dolinar, and J. Thorpe, "Protograph based LDPC codes with minimum distance linearly growing with block size," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, Nov. 2005, pp. 1–5.

[41] D. G. M. Mitchell, M. Lentmaier, A. E. Pusane, and D. J. Costello, "Randomly punctured LDPC codes," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 2, pp. 408–421, Feb. 2016.

JAE-WON KIM (Member, IEEE) received the B.S. and Ph.D. degrees in electrical and computer engineering from Seoul National University, Seoul, South Korea, in 2014 and 2020, respectively. From 2020 to 2021, he was a Staff Engineer with Samsung Electronics, South Korea. In 2021, he joined Gyeongsang National University, Jinju, South Korea, where he is currently an Assistant Professor with the Department of Electronic Engineering. His research interests include error-correcting codes, coding theory, index coding, and DNA storage.

HEE-YOUL KWAK (Member, IEEE) received the B.S. and Ph.D. degrees in electrical and computer engineering from Seoul National University, Seoul, South Korea, in 2013 and 2019, respectively. He was a Staff Engineer with Samsung Electronics, South Korea, from 2019 to 2021, and was a Senior Researcher with the Faculty of the Institute of New Media and Communications, Seoul National University, from 2021 to 2022. He is currently an Associate Professor with the School of Electrical Engineering, University of Ulsan, South Korea. His research interests include error-correcting codes, coding for memory/storage, intelligent communication, and machine learning.

JONG-SEON NO (Fellow, IEEE) received the B.S. and M.S.E.E. degrees in electronics engineering from Seoul National University, Seoul, South Korea, in 1981 and 1984, respectively, and the Ph.D. degree in electrical engineering from the University of Southern California, Los Angeles, CA, USA, in 1988. He was a Senior MTS with Hughes Network Systems, from 1988 to 1990. He was an Associate Professor with the Department of Electronic Engineering, Konkuk University, Seoul, from 1990 to 1999. In 1999, he joined the Faculty of the Department of Electrical and Computer Engineering, Seoul National University, where he is currently a Professor. His current research interests include sequences, LDPC codes, post-quantum cryptography, fully homomorphic encryption, and privacy preserving machine learning. He became a fellow of the IEEE through the IEEE Information Theory Society, in 2012. He became a member of the National Academy of Engineering of Korea (NAEK), in 2015, where he served as the Division Chair for electrical, electronic, and information engineering, from 2019 to 2020. He was a recipient of the IEEE Information Theory Society Chapter of the Year Award in 2007. From 1996 to 2008, he served as the Founding Chair for the Seoul Chapter of the IEEE Information Theory Society. He was the General Chair of Sequence and Their Applications 2004 (SETA 2004), Seoul. He also served as the General Co-Chair for the International Symposium on Information Theory and its Applications 2006 (ISITA 2006) and the International Symposium on Information Theory 2009 (ISIT 2009), Seoul. He served as the Co-Editor-in-Chief for the *Journal of Communications and Networks*, from 2012 to 2013.

JAEWHA KIM (Member, IEEE) received the B.S. degree in electrical engineering from the Pohang University of Science and Technology, in 2014, and the Ph.D. degree in electrical and computer engineering from Seoul National University, in 2022. In 2022, he joined the Faculty of the Institute of New Media and Communications, Seoul National University, as a Researcher. His research interests include error-correcting codes for memory and coding for wireless communication channels.

● ● ●