SCIENCE CHINA
Information Sciences

• Supplementary File •

Global Optimization of Double Protograph LDPC Codes for JSCC Scheme

Qiwang Chen[1,2], Chen Chen[1], Yu-Cheng He[1] & Zhiping Xu[2,3*]

[1]*Xiamen Key Laboratory of Mobile Multimedia Communications, College of Information Science and Engineering, Huaqiao University, Xiamen* 361021*, P. R. China;*
[2]*Key Laboratory of Underwater Acoustic Communication and Marine Information Technology (Xiamen University), Ministry Education, Xiamen 361005, P. R. China;*
[3]*School of Ocean Information Engineering, Jimei University, Xiamen 361021, P. R. China*

Appendix A Background

Separate source and channel design [1] can reach the optimal when the code length is very long. By comparison, joint source-channel coding (JSCC) [2]- [5] scheme with finite length can make full use of source redundancy information and channel state information so as to achieve coding gains. As a typical JSCC scheme, double low-density parity-check (D-LDPC) coding system was proposed in [2], where two LDPC coding matrices perform source and channel coding respectively. An important idea in D-LDPC coding system is the introducing of a linking matrix between the check nodes (CNs) of source coding matrix and the variable nodes (VNs) of channel LDPC coding matrix because it sets up the fundamental connection and realizes the exchange of information between source decoding and channel decoding. Due to the simple and structured modality, protograph LDPC (P-LDPC) codes are introduced into the D-LDPC coding system, named as DP-LDPC coding system [6].

A mountain of research work have been proceed for the optimization of D-LDPC coding system, which can be summarized in two aspects, i.e., partial optimization and global optimization. The partial optimization focused on the single component element, such as the re-design of channel protograph [7] and the optimization of source protograph [8]. In order to lower error-floor further, an extending matrix was introduced [9] to connect the VNs of source coding matrix and the CNs of channel coding matrix and the extending matrix was optimized in [10]. In addition, the linking matrix was optimized [11] assuming that the source and channel coding matrices were fixed. As a rule, the partial optimal cannot reach the global optimal. Thus, several global optimizations are studied including (i) the optimization of source and channel code pair under a curve-fitting algorithm [12], (ii) joint component design based on multi-objective optimization differential evolution [13] and mesh model-based merging method [14], (iii) the quasi-cyclic D-LDPC codes based on an algebraic construction [15], (iv) the allocation of degree-2 VNs for joint protograph [16] and (v) a joint-decoding-complexity-oriented collaborative design [17].

However, these optimizations are all based on a condition that the linking matrix is an identity matrix. Recently, a non-identity linking matrix was introduced into the D-LDPC coding system to improve the performance in water-fall region [18]. But these work only focused on the optimization of the non-identity linking matrix when the extending matrix is a zero matrix, which only belonged to a partial optimization. In addition, the optimization only considered that the extending matrix was a zero matrix. Therefore, a class of global optimization algorithm for DP-LDPC codes with non-zero extending matrix and non-identity linking matrix will be proposed in this paper. The main contributions can be summarized: firstly, the general encoding and decoding method for D-LDPC codes are proposed; secondly, the characteristics of the structure for the DP-LDPC codes are analyzed and new design strategy and guideline are put forward; and lastly, an iterative code searching method is proposed, by which several DP-LDPC codes are optimized.

Appendix B DP-LDPC Code Structure

A protograph can be represented by a base matrix $\boldsymbol{B}=[b^{ij}]$, where $b^{ij}\in\mathbb{N}$ represents the edges connecting i-th CN and j-th VN. The corresponding LDPC coding matrix can be obtained by a two-step "copy and permute" operation, such as progressive edge growth (PEG) algorithm [19]. First, the base matrix is lifted by a factor of z_1 to remove all parallel edges. Then, the resultant matrix is lifted by a factor of z_2 to form a quasi-cyclic matrix.

For DP-LDPC coding system, a joint protograph $\boldsymbol{B}_J$ with $(m_S+m_C)\times(n_S+n_C)$ can be defined by

$$\boldsymbol{B}_J=\left[\begin{array}{c:c}\boldsymbol{B}_S & \boldsymbol{B}_L\\ \hdashline \boldsymbol{B}_E & \boldsymbol{B}_C\end{array}\right], \tag{B1}$$

where $\boldsymbol{B}_S$ with $m_S\times n_S$ is the source component, $\boldsymbol{B}_C$ with $m_C\times n_C$ is the channel component, $\boldsymbol{B}_L=[\boldsymbol{0}\ \boldsymbol{B}_L^*]$ with $m_S\times n_C$ is the linking component and $\boldsymbol{B}_E$ with $m_C\times n_S$ is the extending[1)] component. $\boldsymbol{B}_L^*$ with $m_S\times m_S$ is the actual linking part. After

* Corresponding author (email: xzpxmu@gmail.com)

1) It was called as second kind of linking matrix [7] or source-check-channel-variable (SCCV) linking matrix [10] [18] in previous work. Here considering that the part is different from the other three parts, because $\boldsymbol{B}_E$ can be zero but the other parts in joint matrix must be non-zero. Thus it is called as the extending component appropriately.

the two-step "copy and permute" operations, the joint parity-check matrix $\boldsymbol{H}_J$ with $(M_S + M_C) \times (N_S + N_C)$ can be given by

$$\boldsymbol{H}_J = \left[\begin{array}{c:c} \boldsymbol{H}_S & \boldsymbol{H}_L \\ \hdashline \boldsymbol{H}_E & \boldsymbol{H}_C \end{array}\right] = \left[\begin{array}{c:c} \boldsymbol{H}_S & \boldsymbol{0}\ \boldsymbol{H}_L^* \\ \hdashline \boldsymbol{H}_E & \boldsymbol{H}_C \end{array}\right], \tag{B2}$$

where $\boldsymbol{H}_S$ with $M_S \times N_S$, $\boldsymbol{H}_C$ with $M_C \times N_C$, $\boldsymbol{H}_L$ with $M_S \times N_C$ ($\boldsymbol{H}_L^*$ with $M_S \times M_S$) and $\boldsymbol{H}_E$ with $M_C \times N_S$ are the corresponding parity-check matrices, respectively. Here, we assume that $\boldsymbol{H}_E = [\boldsymbol{0}\ \boldsymbol{H}_E^*](\boldsymbol{B}_E = [\boldsymbol{0}\ \boldsymbol{B}_E^*])$ although the extending component can be a zero matrix or a full column rank matrix. $\boldsymbol{H}_E^*$ and $\boldsymbol{B}_E^*$ are with size $M_E \times N_E$ and $m_E \times n_E$, respectively. The punctured $\boldsymbol{H}_c$ is considered here and the number of the punctured bits is N_P, so the channel rate is $R_{cc} = \dfrac{N_C - M_C}{N_C - N_P}$.

Appendix C General DP-LDPC Encoding and Decoding Systems

Here we generalize the encoding and decoding method regardless of (i) whether the extending matrix is zero or not and (ii) whether the linking matrix is identity or non-identity. The detailed encoding and decoding procedures are depicted as follows.

Let a source sequence $\boldsymbol{s}$ with $1 \times N_s$ take value from a binary i.i.d Bernoulli ($\xi_1 < 0.5$) with entropy

$$\mathbb{H} = -\xi_1 \log_2(\xi_1) - (1 - \xi_1)\log_2(1 - \xi_1). \tag{C1}$$

For the encoding, a generated matrix in systematic form $\boldsymbol{G}_S = [\boldsymbol{I}_S\ (\boldsymbol{P}_S)^T]$ can firstly be obtained from $[\boldsymbol{H}_S\ \boldsymbol{H}_L^*]$ by Gaussian elimination, where $\boldsymbol{I}_S$ with $N_s \times N_s$ is an identity matrix, $\boldsymbol{P}_S$ has a size of $N_s \times M_s$ and $(\cdot)^T$ is the matrix transpose operation. Thus, the compressed bits $\boldsymbol{b}$ can be obtained by

$$[\boldsymbol{s}\ \boldsymbol{b}] = \boldsymbol{s} \cdot \boldsymbol{G}_S = \boldsymbol{s} \cdot [\boldsymbol{I}_S\ (\boldsymbol{P}_S)^T]. \tag{C2}$$

When $\boldsymbol{H}_L^*$ is an identity matrix, $\boldsymbol{P}_S = \boldsymbol{H}_S$. Then the source sequence connected by $\boldsymbol{H}_E^*$, denoted as $\boldsymbol{s}_p$, and the compressed bits $\boldsymbol{b}$ are combined into new sequence $[\boldsymbol{s}_p\ \boldsymbol{b}]$ and a generated matrix $\boldsymbol{G}_C = [\boldsymbol{I}_C\ (\boldsymbol{P}_C)^T]$ is obtained from $[\boldsymbol{H}_E^*\ \boldsymbol{H}_C]$, where $\boldsymbol{I}_C$ is an identity matrix with $(M_C + N_E) \times (M_C + N_E)$. The $[\boldsymbol{s}_p\ \boldsymbol{b}]$ is encoded by $\boldsymbol{G}_C$, i.e.,

$$\boldsymbol{c} = [\boldsymbol{s}_p\ \boldsymbol{b}] \cdot \boldsymbol{G}_C = [\boldsymbol{s}_p\ \boldsymbol{b}] \cdot [\boldsymbol{I}_C\ (\boldsymbol{P}_C)^T], \tag{C3}$$

where $\boldsymbol{c} = [\boldsymbol{s}_p\ \boldsymbol{b}\ \boldsymbol{p}]$ and $\boldsymbol{p}$ with $1 \times M_C$ is the parity bits sequence. Then the sequence $[\boldsymbol{b}\ \boldsymbol{p}]$ is sent into an additive white Gaussian noise (AWGN) channel after BPSK modulation. The noise variance is given by $\sigma^2 = \dfrac{1}{2 \cdot R_{cc} \cdot E_b/N_0}$, where E_b is the average transmitted energy per source information bit and N_0 is the noise power spectral density.

Considering that the whole codeword $\boldsymbol{u} = [\boldsymbol{s}\ \boldsymbol{b}\ \boldsymbol{p}]$ satisfies

$$\begin{aligned}
\boldsymbol{u}(\boldsymbol{H}_J)^T &= [\boldsymbol{s}\ \boldsymbol{b}\ \boldsymbol{p}] \cdot (\boldsymbol{H}_J)^T \\
&= [\boldsymbol{s}_{1-p}\ \boldsymbol{s}_p\ \boldsymbol{b}\ \boldsymbol{p}] \cdot \left[\begin{array}{c:c} \boldsymbol{H}_S & \boldsymbol{H}_L^*\ \boldsymbol{0} \\ \hdashline \boldsymbol{0}\ \boldsymbol{H}_E^* & \boldsymbol{H}_C \end{array}\right]^T \\
&= [\boldsymbol{s}_{1-p}\ \boldsymbol{s}_p\ \boldsymbol{b}\ \boldsymbol{p}] \cdot \left[\begin{array}{c:c} (\boldsymbol{H}_S)^T & \boldsymbol{0} \\ \hdashline (\boldsymbol{H}_L^*)^T & (\boldsymbol{H}_E^*)^T \\ \boldsymbol{0} & (\boldsymbol{H}_C)^T \end{array}\right] \\
&= \left[[\boldsymbol{s}\ \boldsymbol{b}] \cdot \left[\begin{array}{c} \boldsymbol{H}_S \\ \boldsymbol{H}_L^* \end{array}\right]^T\ [\boldsymbol{s}_p\ \boldsymbol{b}\ \boldsymbol{p}] \cdot \left[\begin{array}{c} \boldsymbol{H}_E^* \\ \boldsymbol{H}_C \end{array}\right]^T\right] \\
&= [\boldsymbol{0}\ \boldsymbol{0}] = [\boldsymbol{0}],
\end{aligned} \tag{C4}$$

where $\boldsymbol{s}_{1-p}$ is the remaining part of $\boldsymbol{s}$, i.e., $\boldsymbol{s} = [\boldsymbol{s}_{1-p}\ \boldsymbol{s}_p]$. Thus, the corrupted sequence $\boldsymbol{y}$ at receiver can be decoded by joint belief propagation (JBP) algorithm when considering the joint Tanner graph as a whole, which is shown in Fig. 1.

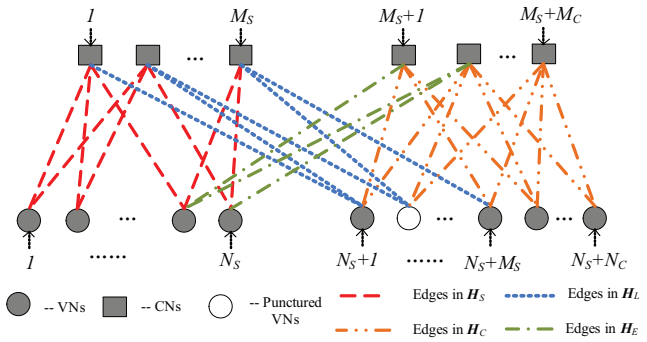

Figure C1 Joint tanner graph of DP-LDPC codes, where different types of lines represent the edges in different matrices.

Appendix D Global Optimization of DP-LDPC Codes

Appendix D.1 Design Strategy and Guideline

In order to analyze the effects of each parts in joint protograph, some separate components, i.e., $\boldsymbol{B}_S$, $\boldsymbol{B}_C$, $\boldsymbol{B}_E$ and $\boldsymbol{B}_L$, are as follows.

Source component:

$$\boldsymbol{B}_S^{R4JA} = \begin{bmatrix} 1\,1\,3\,1\,3\,1\,3\,1 \\ 1\,2\,1\,3\,1\,3\,1\,2 \end{bmatrix} \text{ and } \boldsymbol{B}_S^{Reg} = \begin{bmatrix} 2\,1\,2\,1\,2\,1\,2\,1 \\ 1\,2\,1\,2\,1\,2\,1\,2 \end{bmatrix}$$

Channel component:

$$\boldsymbol{B}_C^{IARA} = \begin{bmatrix} 1\,0\,0\,3\,0 \\ 0\,1\,1\,2\,1 \\ 0\,1\,1\,2\,1 \end{bmatrix} \text{ and } \boldsymbol{B}_C^{AR3A} = \begin{bmatrix} 1\,0\,0\,2\,1 \\ 0\,1\,1\,2\,1 \\ 0\,1\,1\,1\,2 \end{bmatrix}.$$

Extending component:

$$\boldsymbol{B}_E^{v1} = \begin{bmatrix} 0\,0\,0\,0\,0\,0\,0\,0 \\ 3\,1\,0\,0\,0\,0\,0\,0 \\ 1\,3\,0\,0\,0\,0\,0\,0 \end{bmatrix} \text{ and } \boldsymbol{B}_E^{v2} = \begin{bmatrix} 0\,0\,0\,0\,0\,0\,0\,0 \\ 0\,1\,0\,0\,0\,0\,0\,0 \\ 1\,0\,0\,0\,0\,0\,0\,0 \end{bmatrix}.$$

Linking component:

$$\boldsymbol{B}_L^{E} = \begin{bmatrix} 0\,0\,0\,1\,0 \\ 0\,0\,0\,0\,1 \end{bmatrix}, \boldsymbol{B}_L^{v1} = \begin{bmatrix} 0\,0\,0\,1\,0 \\ 0\,0\,0\,2\,1 \end{bmatrix} \text{ and } \boldsymbol{B}_L^{v2} = \begin{bmatrix} 0\,0\,0\,2\,0 \\ 0\,0\,0\,0\,2 \end{bmatrix}.$$

Several joint protographs by combining these components are shown in Fig. 2, where the $\boldsymbol{B}_{J1}^{0.02/0.03}$ is proposed in [7]. Further more, the $\boldsymbol{B}_L^{v1}$ in $\boldsymbol{B}_{J2}^{0.02/0.03}$ is the optimized result according to the algorithm proposed in [18] at $\xi_1 = 0.02$ when the other components are fixed. The $\boldsymbol{B}_C^{AR3A}$, $\boldsymbol{B}_E^{v2}$ and $\boldsymbol{B}_L^{v2}$ are the compared counterparts.

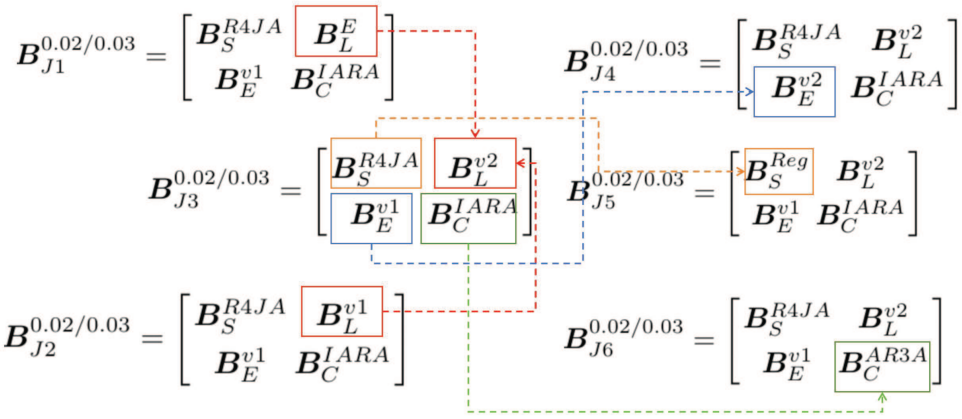

Figure D1 The differences among the joint base matrices $\boldsymbol{B}_{J1}^{0.02/0.03}$, $\boldsymbol{B}_{J2}^{0.02/0.03}$, $\boldsymbol{B}_{J3}^{0.02/0.03}$,$\boldsymbol{B}_{J4}^{0.02/0.03}$,$\boldsymbol{B}_{J5}^{0.02/0.03}$ and $\boldsymbol{B}_{J6}^{0.02/0.03}$.

Their channel decoding threshold $(E_b/N_0)_{th}$ and source decoding threshold ξ_{th} are calculated by JPEXIT algorithm [7] shown in Table. I. Firstly, the $\boldsymbol{B}_{J2}^{0.02/0.03}$ has a coding gain in channel decoding threshold, which has non-identity $\boldsymbol{B}_L^*$ compared with $\boldsymbol{B}_{J1}^{0.02/0.03}$. However, the $(E_b/N_0)_{th}$ of $\boldsymbol{B}_{J3}^{0.02/0.03}$ becomes infinite, although it also has non-identity $\boldsymbol{B}_L^*$. This extreme case implies that $\boldsymbol{B}_{J3}^{0.02/0.03}$ is a "bad" code[2)]. If the source component or the extending component is adjusted, the $\boldsymbol{B}_{J3}^{0.02/0.03}$ changes to be $\boldsymbol{B}_{J4}^{0.01/0.02}$ or $\boldsymbol{B}_{J5}^{0.02/0.03}$ respectively, which have the same ξ_{th} but also belong to "bad" codes. If the channel component is adjusted, the $\boldsymbol{B}_{J3}^{0.02/0.03}$ changes to be $\boldsymbol{B}_{J6}^{0.02/0.03}$, which has a reasonable $(E_b/N_0)_{th}$ and keeps the same ξ_{th}. It can be found that the adjustment of $\boldsymbol{B}_L^*$ and $\boldsymbol{B}_C$ has little effects on the ξ_{th} and the adjustment of $\boldsymbol{B}_S$ or $\boldsymbol{B}_E$ cannot improve the $(E_b/N_0)_{th}$ of $\boldsymbol{B}_{J3}^{0.01/0.02}$, $\boldsymbol{B}_{J4}^{0.02/0.03}$ and $\boldsymbol{B}_{J5}^{0.02/0.03}$.

2) A code has catastrophic performance in the communication system.

Table D1 The channel and source decoding thresholds of several representative $\boldsymbol{B}_J^{0.02/0.03}$.

$\boldsymbol{B}_J$	$(\xi_1)_{th}$	$(E_b/N_0)_{th}$	
		0.02	0.03
$\boldsymbol{B}_{J1}^{0.02/0.03}$	0.041	−0.85 dB	0.70 dB
$\boldsymbol{B}_{J2}^{0.02/0.03}$	0.041	−1.03 dB	0.75 dB
$\boldsymbol{B}_{J3}^{0.02/0.03}$	0.041	∞	∞
$\boldsymbol{B}_{J4}^{0.02/0.03}$	0.041	∞	∞
$\boldsymbol{B}_{J5}^{0.02/0.03}$	0.043	∞	∞
$\boldsymbol{B}_{J6}^{0.02/0.03}$	0.041	−0.26 dB	1.06 dB

After the above analysis, the design strategies are given as follows.

- The base matrices $\boldsymbol{B}_S$ and $\boldsymbol{B}_E$ are considered as a whole, denoted as source-extending coding (SEC) base matrix $\boldsymbol{B}_{S\&E}$, since it determines the source decoding threshold $(\xi_1)_{th}$;
- The $\boldsymbol{B}_C$ and $\boldsymbol{B}_L$ are considered as a whole, i.e., linking channel coding (LCC) base matrix $\boldsymbol{B}_{L\&C}$, since the mismatch between them will result in a "bad" code;
- Now the joint base matrix becomes

$$\boldsymbol{B}_J = \left[\begin{array}{c:c} \boldsymbol{B}_S & \boldsymbol{B}_L \\ \boldsymbol{B}_E & \boldsymbol{B}_C \end{array}\right] = \left[\begin{array}{c:c} \boldsymbol{B}_{S\&E} & \boldsymbol{B}_{L\&C} \end{array}\right]. \tag{D1}$$

- The design of $\boldsymbol{B}_J$ is converted into the design of $\boldsymbol{B}_{S\&E}$ and $\boldsymbol{B}_{L\&C}$, as well as the optimal match between them.

According to the design strategies, some properties of a good joint protograph is analyzed as follows. Firstly, a structure of degree-2 VN is very important in its quantity and allocation. Although more degree-2 VNs can lower the channel decoding threshold, it matters whether or not the linear minimum distance growth property is satisfied, which relates to the error-floor level. The property requires the maximum quantity of degree-2 VNs in the protograph to be limited by total number of checks minus 1. For a fixed quantity of degree-2 VNs, they should be assigned to the channel component as much as possible [16], i.e., $\boldsymbol{B}_{L\&C}$. Secondly, the channel decoding threshold can be further improved by a high degree VN and several degree-1 VNs (usually called as a pre-coder), which exist in $\boldsymbol{B}_{L\&C}$. With the pre-coder structure, the high degree VN are optionally punctured, denoted as n_{punc}, which is equal to the quantity of degree-1 VNs. Thirdly, all the degrees of VNs expect the pre-coder structure and degree-2 VNs should be large than 2. Lastly, the number of non-zero column in $\boldsymbol{B}_E$ is positively related to ξ_{th} but negatively related to $(E_b/N_0)_{th}$, so the structure of $\boldsymbol{B}_{S\&E}$ should be set appropriately.

The detailed design principles are summarized as follows.

Design Guidelines

(a) The number of degree-2 VNs in $\boldsymbol{B}_J$ is at most $m_S + m_C - n_{punc} - 1$;
(b) The VN with the high degree exists in $\boldsymbol{B}_{L\&C}$;
(c) The degree-1 VNs exist in $\boldsymbol{B}_{L\&C}$;
(d) The left VNs except the pre-coder structure and degree-2 VNs in base matrix $\boldsymbol{B}_J$ have the degree larger than 2;
(e) The non-zero columns of $\boldsymbol{B}_E$ in $\boldsymbol{B}_{S\&E}$ should be set appropriately.

Appendix D.2 Optimization Method

For a good joint protograph, the error floor level satisfying the essential requirements of the communication system (e.g., BER of 10^{-6} in wireless communications) is indispensable and the lowering of the channel decoding threshold is more important. Therefore, we will focus on the improvement of the channel decoding threshold on the premise of ensuring the essential requirements of error-floor level.

Protograph design involves the search for a base matrix which has the best possible threshold while preserving some preset structure and conditions. The size of the search space is at most $(1+b_{max})^{mn}-1$, where $m = m_S + m_C$ and $n = n_S + n_C$. Due to the exponential growth of the search space as a function of the size of the base matrix, it is difficult to find an optimal base matrix of large size in a brute force manner. Differential evolution (DE) is a parameter optimization algorithm that iteratively tries to improve candidate solutions with regard to a given cost function, which has a lower complexity. Thus, an iterative code searching method based on differential evolution (DE) algorithm is described in **Algorithm D1**.

Compared with the conventional coding searching methods [13] that search the whole joint protograph based on the DE algorithm, the proposed iterative code searching algorithm takes full advantage of the characteristics of $\boldsymbol{B}_{S\&E}$ and $\boldsymbol{B}_{L\&C}$, and the complexity of the proposed method will decrease significantly.

Example-1: Consider the optimized results cases in [13] and [16], where source statistic is $\xi_1^{obj} = 0.04$, the parameters are $m_S = 4$, $n_S = 2$ for source base matrix and $m_C = 3$, $n_C = 5$ for channel protograph. In addition, there is one degree-1 VN and and one punctured VN with the high degree existed in channel component $\boldsymbol{B}_{L\&C}$. The maximum number of degree-2 VNs can be calculated as 3. According to the allocation principles, these VNs should be assigned to the channel component $\boldsymbol{B}_{L\&C}$, so it can be initialized as

$$\boldsymbol{B}_{L\&C}^{0.04-ini} = \begin{bmatrix} 0\,0\,0\,1\,b_{lc}^{15} \\ 0\,0\,0\,0\,b_{lc}^{25} \\ 1\,0\,0\,0\,b_{lc}^{35} \\ 0\,1\,1\,0\,b_{lc}^{45} \\ 0\,1\,1\,1\,b_{lc}^{55} \end{bmatrix} \tag{D2}$$

Algorithm D1 Iterative code searching method based on DE algorithm for designing the joint protograph $\boldsymbol{B}_J$.

GIVEN DE ALGORITHM: For given parameters including the number of candidates in one generation, the maximum number of generations, crossover probability and mutation probability, the algorithm has *INITIALIZATION*, *MUTATION*, *CROSSOVER* and *SELECTION* processes. The detailed DE algorithm is described in [12] and [13].

Step 1: For an objective source statistics ξ_1^{obj}, m_S, n_S, m_C, n_C, n_{punc} and b_{max}, generate an initial $\boldsymbol{B}_{S\&E}^{ini}$ and an initial $\boldsymbol{B}_{L\&C}^{ini}$ with values randomly chosen from $\{0, 1, ..., b_{max}\}$. Both of them need to satisfy the design principles (a)-(e).

Step 2: At the k-th iteration process of optimizing $\boldsymbol{B}_{S\&E}$:

- Fix $\boldsymbol{B}_{L\&C}^{opt(k-1)}$ and optimize $\boldsymbol{B}_{S\&E}^{(k)}$ based on DE algorithm.
- The optimized $\boldsymbol{B}_{S\&E}^{opt(k)}$ conforms the design guideline (d) and (e) and has a source decoding threshold that is larger than $T((\xi_1)_{th})$. The parameter $T((\xi_1)_{th})$ indicates the minimum source decoding threshold satisfying the requirements of error-floor level [12].
- The corresponding $\boldsymbol{B}_J$ conforms the design guidelines (a) and (d), and possesses the lowest $(E_b/N_0)_{th}$ among all candidates $\boldsymbol{B}_{S\&E}^{candi(k)}$.

Step 3: At the k-th iteration process of optimizing $\boldsymbol{B}_{L\&C}$:

- Fix $\boldsymbol{B}_{S\&E}^{opt(k)}$ and optimize $\boldsymbol{B}_{L\&C}^{(k)}$ based on DE algorithm.
- The optimized $\boldsymbol{B}_{L\&C}^{opt(k)}$ conforms the design guidelines (b) and (c).
- The corresponding $\boldsymbol{B}_J$ conforms the design guidelines (a) and (d), and possesses the lowest $(E_b/N_0)_{th}$ among all candidates $\boldsymbol{B}_{L\&C}^{candi(k)}$;

Step 4: If the obtained $\boldsymbol{B}_J$ satisfies the *END CONDITION*, we will stop the iterative searching process; otherwise, set k=k+1 and go to **Step 2**.

END CONDITION: The value of $(E_b/N_0)_{th}^{min(k)}$ of the optimized $\boldsymbol{B}_J$ tends to be stable, i.e., the coding gain $|(E_b/N_0)_{th}^{min(k-1)} - (E_b/N_0)_{th}^{min(k)}| < 0.01$ dB.

where $b_{lc}^{15} + b_{lc}^{25} + b_{lc}^{35} + b_{lc}^{45} + b_{lc}^{55} > 2$. For the source component, there is no degree-2 VNs and one non-zero , so the $\boldsymbol{B}_{S\&E}$ can be initialized as

$$\boldsymbol{B}_{S\&E}^{0.04-ini} = \begin{bmatrix} b_{se}^{11} & b_{se}^{12} & b_{se}^{13} & b_{se}^{14} \\ b_{se}^{21} & b_{se}^{22} & b_{se}^{23} & b_{se}^{24} \\ b_{se}^{31} & 0 & 0 & 0 \\ b_{se}^{41} & 0 & 0 & 0 \\ b_{se}^{51} & 0 & 0 & 0 \end{bmatrix}, \tag{D3}$$

where

$$\begin{cases} b_{se}^{11} + b_{se}^{21} + b_{se}^{31} + b_{se}^{41} + b_{se}^{51} > 2 \\ b_{se}^{1j} + b_{se}^{2j} > 2 \quad (j = 2, 3, 4) \end{cases} \tag{D4}$$

In order to control the complexity, the maximum value of the elements is $b_{max} = 3$. Then the Algorithm 1 is applied and the optimized result can be obtained as

$$\boldsymbol{B}_{opt}^{0.04} = \left[\begin{array}{cccc:cccc} 2 & 1 & 1 & 2 & 0 & 0 & 0 & 1\ 2 \\ 1 & 2 & 2 & 1 & 0 & 0 & 0 & 0\ 2 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0\ 3 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0\ 2 \\ 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1\ 1 \end{array}\right], \tag{D5}$$

which has a source decoding threshold $(\xi_1)_{th} = 0.135$ and a channel decoding $(E_b/N_0)_{th} = -2.91$ dB for $\xi_1 = 0.04$.

Example-2: The case for $\xi_1 = 0.02$ in [7] is considered, i.e., $\boldsymbol{B}_{J1}^{0.02/0.03}$, which is rewritten as $\boldsymbol{B}_{conv1}^{0.02}$. The detailed parameters are$m_S = 2$, $n_S = 8$, $m_C = 3$, $n_C = 5$ and $b_{max} = 3$, where the initialized structure of $\boldsymbol{B}_{L\&C}$ is the same as the case of $\xi_1 = 0.04$. Here, we set the non-zero column in $\boldsymbol{B}_E$ as 2 for the sake of fairness. So the initialized $\boldsymbol{B}_{S\&E}$ can be given by

$$\boldsymbol{B}_{S\&E}^{0.02-ini} = \begin{bmatrix} b_{se}^{11} & b_{se}^{12} & b_{se}^{13} & b_{se}^{14} & b_{se}^{15} & b_{se}^{16} & b_{se}^{17} & b_{se}^{18} \\ b_{se}^{21} & b_{se}^{22} & b_{se}^{23} & b_{se}^{24} & b_{se}^{25} & b_{se}^{26} & b_{se}^{27} & b_{se}^{28} \\ b_{se}^{31} & b_{se}^{32} & 0 & 0 & 0 & 0 & 0 & 0 \\ b_{se}^{41} & b_{se}^{42} & 0 & 0 & 0 & 0 & 0 & 0 \\ b_{se}^{51} & b_{se}^{52} & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \tag{D6}$$

where

$$\begin{cases} b_{se}^{1j}+b_{se}^{2j}+b_{se}^{3j}+b_{se}^{4j}+b_{se}^{5j}>2 & (j=1,2) \\ b_{se}^{1j}+b_{se}^{2j}>2 & (j=3,4,\cdots,8) \end{cases}. \tag{D7}$$

The optimized result can be obtained by the Algorithm 1 and given by

$$\boldsymbol{B}_{opt}^{0.02}=\left[\begin{array}{cccccccc:cccc} 1&0&3&1&3&1&2&1&0&0&0&2 \\ 1&2&1&3&1&3&1&2&0&0&0&1&1 \\ \hdashline 0&0&0&0&0&0&0&0&1&0&0&3 \\ 1&2&0&0&0&0&0&0&0&1&1&2 \\ 2&1&0&0&0&0&0&0&0&1&1&3 \end{array}\right] \tag{D8}$$

which has a source decoding threshold $(\xi_1)_{th}=0.043$ and a channel decoding $(E_b/N_0)_{th}=-1.77$ dB for $\xi_1=0.02$.

Example-3: The proposed algorithm is also suitable for the case $\boldsymbol{B}_E=0$. The case in [16] is considered, where the source statistic $\xi_1^{obj}=0.01$, $m_S=2$, $n_S=8$, $m_C=3$ and $n_C=5$. It has the same initialized structure of $\boldsymbol{B}_{L\&C}$ as the case of $\xi_1=0.02$ and 0.04. So the initialized $\boldsymbol{B}_{S\&E}$ can be given by

$$\boldsymbol{B}_{S\&E}^{0.01-ini}=\begin{bmatrix} b_{se}^{11}&b_{se}^{12}&b_{se}^{13}&b_{se}^{14}&b_{se}^{15}&b_{se}^{16}&b_{se}^{17}&b_{se}^{18} \\ b_{se}^{21}&b_{se}^{22}&b_{se}^{23}&b_{se}^{24}&b_{se}^{25}&b_{se}^{26}&b_{se}^{27}&b_{se}^{28} \\ 0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \end{bmatrix}, \tag{D9}$$

where

$$b_{se}^{1j}+b_{se}^{2j}>2(j=1,2,\cdots,8) \tag{D10}$$

Thus the optimized result can be obtained, i.e.,

$$\boldsymbol{B}_{opt}^{0.01}=\left[\begin{array}{cccccccc:cccc} 2&1&2&1&2&1&3&1&0&0&0&2 \\ 1&2&1&2&1&2&2&2&0&0&0&1&1 \\ \hdashline 0&0&0&0&0&0&0&0&1&0&0&3 \\ 0&0&0&0&0&0&0&0&0&1&1&2 \\ 0&0&0&0&0&0&0&0&0&1&1&2 \end{array}\right], \tag{D11}$$

which has a source decoding threshold $(\xi_1)_{th}=0.028$ and a channel decoding $(E_b/N_0)_{th}=-4.52$ dB for $\xi_1=0.01$.

Appendix E Simulation and Comparison Results

In order to show the superiority of the proposed joint protographs, several state-of-the-art ones as the compared counterparts are given as follows,

$$\boldsymbol{B}_{conv1}^{0.04}=\left[\begin{array}{ccccc:ccccc} 3&2&2&1&0&0&0&1&0 \\ 3&1&1&1&0&0&0&0&1 \\ 0&0&0&0&1&0&0&1&2 \\ 3&0&0&0&0&1&1&0&2 \\ 3&0&0&0&0&1&1&1&1 \end{array}\right], \boldsymbol{B}_{conv2}^{0.04}=\left[\begin{array}{ccccc:ccccc} 3&2&2&1&0&0&0&1&0 \\ 3&1&1&1&0&0&0&0&1 \\ 0&0&0&0&1&0&0&0&3 \\ 3&0&0&0&0&1&1&0&2 \\ 3&0&0&0&0&1&1&1&1 \end{array}\right],$$

$$\boldsymbol{B}_{conv1}^{0.01}=\left[\begin{array}{ccccccccc:ccccc} 2&1&2&1&2&2&2&2&0&0&0&1&0 \\ 1&2&1&2&1&3&1&3&0&0&0&0&1 \\ \hdashline 0&0&0&0&0&0&0&0&1&0&0&0&3 \\ 0&0&0&0&0&0&0&0&0&1&1&1&2 \\ 0&0&0&0&0&0&0&0&0&1&1&0&3 \end{array}\right], \boldsymbol{B}_{conv2}^{0.01}=\left[\begin{array}{ccccccccc:ccccc} 1&1&2&1&3&1&3&1&0&0&0&1&0 \\ 1&2&1&2&1&2&1&2&0&0&0&3&1 \\ \hdashline 0&0&0&0&0&0&0&0&1&0&0&0&3 \\ 0&0&0&0&0&0&0&0&0&1&1&2&1 \\ 0&0&0&0&0&0&0&0&0&1&1&1&2 \end{array}\right],$$

where both of $\boldsymbol{B}_{conv2}^{0.04}$ and $\boldsymbol{B}_{conv1}^{0.01}$ proposed in [16], $\boldsymbol{B}_{conv1}^{0.04}$ in [13] and $\boldsymbol{B}_{conv2}^{0.01}$ in [18].

For the Monte Carlo simulation, the codelength of these DP-LDPC codes is 6400, the simulated environment is performed over AWGN channels and the maximum iteration number of joint BP decoding algorithm is set to 100.

Appendix E.1 Performance Comparison

The channel decoding threshold of different joint protographs at the respective source statistic is analyzed and compared in Table II, where the optimized ones all have certain coding gain. For source statistic $\xi_1=0.04$, the BER performance of the $\boldsymbol{B}_{conv1}^{0.04}$, $\boldsymbol{B}_{conv2}^{0.04}$ and $\boldsymbol{B}_{opt}^{0.04}$ is plotted in Fig. E1. At BER=1×10^{-6}, the $\boldsymbol{B}_{opt}^{0.04}$ outperforms $\boldsymbol{B}_{conv1}^{0.04}$ and $\boldsymbol{B}_{conv2}^{0.04}$ by 0.25 dB and 0.15 dB,

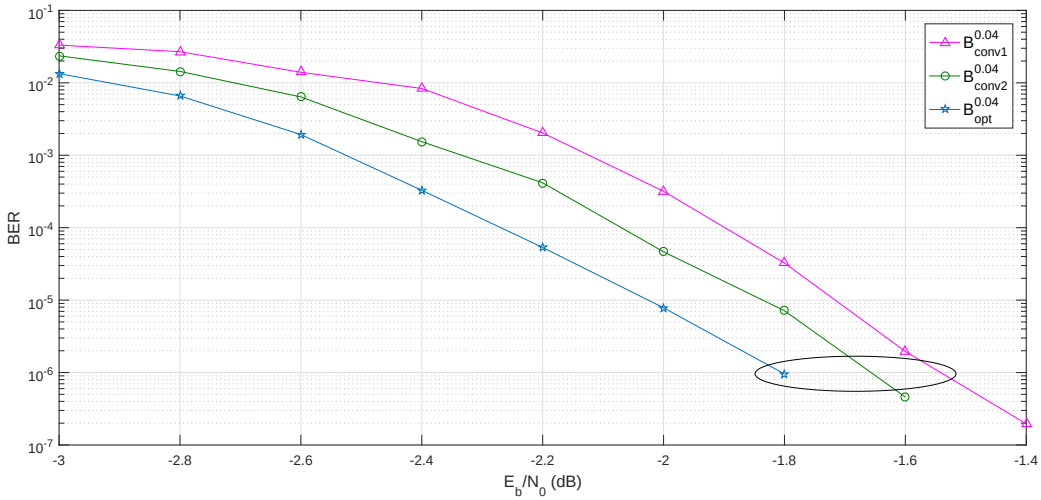

Figure E1 BER performance of $\boldsymbol{B}^{0.04}_{conv1}$, $\boldsymbol{B}^{0.04}_{conv2}$ and $\boldsymbol{B}^{0.04}_{opt}$ at $\xi_1 = 0.04$.

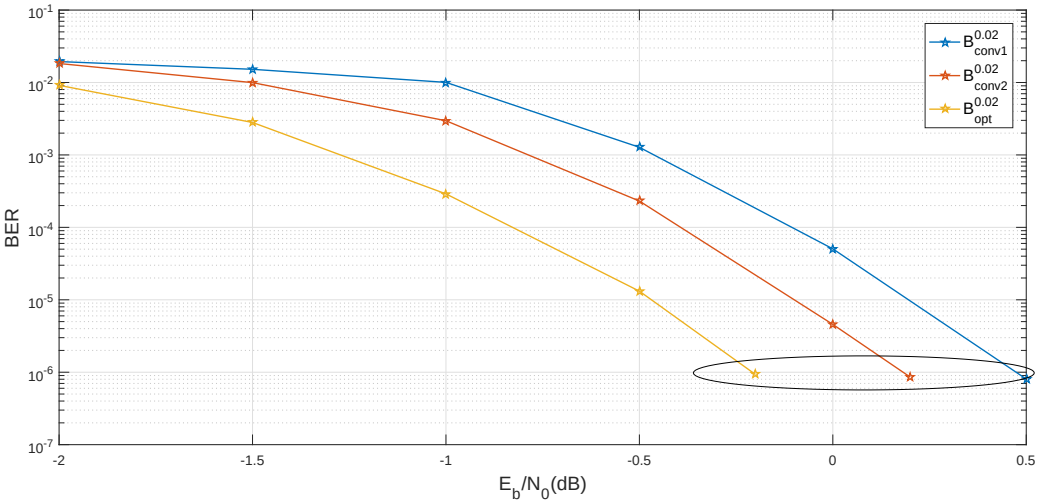

Figure E2 BER performance of $\boldsymbol{B}^{0.02}_{conv1}$, $\boldsymbol{B}^{0.02}_{conv2}$ and $\boldsymbol{B}^{0.02}_{opt}$ at $\xi_1 = 0.02$.

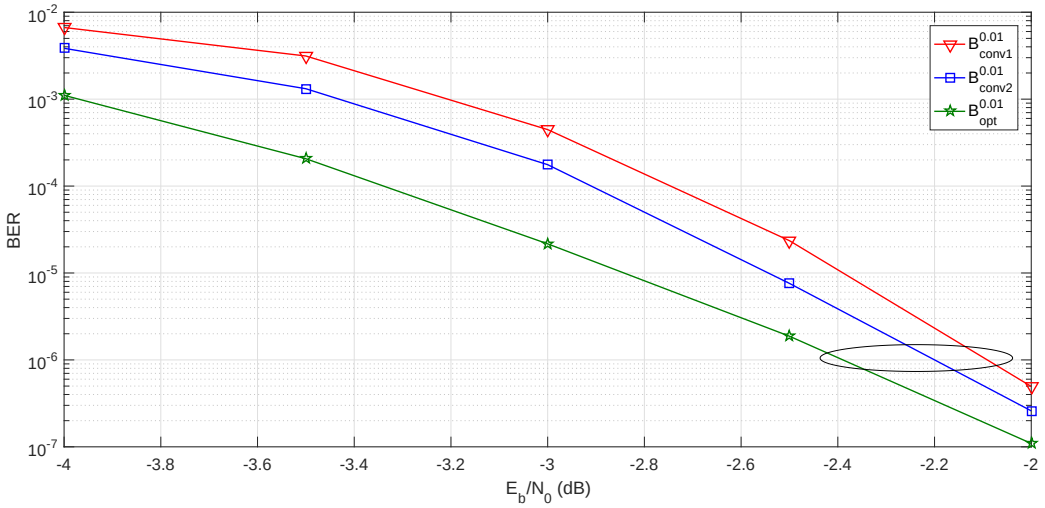

Figure E3 BER performance of $\boldsymbol{B}_{conv1}^{0.01}$, $\boldsymbol{B}_{conv2}^{0.01}$ and $\boldsymbol{B}_{opt}^{0.01}$ at $\xi_1 = 0.01$.

Table E1 The Channel Decoding Threshold and Average Degree of Check Nodes $\overline{d_c}$ of Different Joint Protographs.

	$\boldsymbol{B}_{conv1}^{0.01}$	$\boldsymbol{B}_{conv2}^{0.01}$	$\boldsymbol{B}_{opt}^{0.01}$
$(E_b/N_0)_{th}$	-3.72 dB	-3.74 dB	-4.51 dB
$\overline{d_c}$	8.8	8.8	8.6
	$\boldsymbol{B}_{conv1}^{0.02}$	$\boldsymbol{B}_{conv2}^{0.02}$	$\boldsymbol{B}_{opt}^{0.02}$
$(E_b/N_0)_{th}$	-0.85 dB	-1.03 dB	-1.73 dB
$\overline{d_c}$	10.4	10.8	9.4
	$\boldsymbol{B}_{conv1}^{0.04}$	$\boldsymbol{B}_{conv2}^{0.04}$	$\boldsymbol{B}_{opt}^{0.04}$
$(E_b/N_0)_{th}$	-2.57 dB	-2.75 dB	-2.91 dB
$\overline{d_c}$	6.8	6.8	6.0

respectively, which is in line with the channel decoding threshold analysis. Then the performance of $\boldsymbol{B}_{conv1}^{0.02}$, $\boldsymbol{B}_{conv2}^{0.02}$ and $\boldsymbol{B}_{opt}^{0.02}$ is illustrated in Fig. E2 at source statistics $\xi_1 = 0.02$. At BER=1×10^{-6}, the $\boldsymbol{B}_{opt}^{0.02}$ has a coding gain of 0.7 dB and 0.5 dB compared with $\boldsymbol{B}_{conv1}^{0.02}$ and $\boldsymbol{B}_{conv2}^{0.02}$, respectively, which is also in line with the channel decoding threshold analysis. The coding gains are also reflected in the superiority of the $\boldsymbol{B}_{opt}^{0.01}$ compared with the $\boldsymbol{B}_{conv1}^{0.01}$ and $\boldsymbol{B}_{conv2}^{0.01}$, which is shown in Fig. E3.

Appendix E.2 Complexity Comparisons

For the JBP decoding algorithm, the decoding complexity is dominated by the update calculation of the information from the CNs to VNs, which can be measured by the average degree $\overline{d_c}$ of a protograph. As shown in Table II, the decoding complexity of $\boldsymbol{B}_{opt}^{0.04}$ decreases by 11.7% compared with the $\boldsymbol{B}_{conv1}^{0.04}$ and $\boldsymbol{B}_{conv2}^{0.04}$,as well as that of $\boldsymbol{B}_{opt}^{0.02}$ decreases by 9.6% and 13.0% compared with the $\boldsymbol{B}_{conv1}^{0.02}$ and $\boldsymbol{B}_{conv2}^{0.02}$, respectively. The $\boldsymbol{B}_{opt}^{0.01}$ has a certain decrease in complexity compared with $\boldsymbol{B}_{conv1}^{0.01}$ and $\boldsymbol{B}_{conv2}^{0.01}$.

Therefore, it can be concluded that the optimized DP-LDPC codes have better performance and lower decoding complexity, compared with the conventional ones, which verifies the superiority of the proposed global optimization algorithm.

Appendix F Conclusion

Considering that the introducing of non-identity linking matrix results in the change of the code structure, general encoding and decoding methods are proposed in this paper. By analyzing several different joint protographs, it is found that the mismatch between linking base matrix and channel base matrix will result in a catastrophic channel decoding threshold. Meanwhile, the source and extending base matrix determine the source decoding threshold. Thus, a class of iterative code search algorithm with new structure i.e., source-extending coding matrix and linking-channel coding matrix, are is proposed under several new design principles. By comparison, the optimized DP-LDPC codes show their advantages reflected in better BER performance and lower decoding complexity.

References

1 P. Chen, L. Shi, Y. Fang, F. C. M. Lau, J. Cheng, "Rate-diverse multiple access over Gaussian channels," *IEEE Trans. Wire. Commun.*, vol. 22, no. 8, Aug. 2023, pp. 5399-5413

2 M. Fresia, F. Perez-cruz, H. V. Poor and S. Verdu, "Joint source and channel coding," *IEEE Signal Processing Mag.*, vol. 27, no. 6, pp. 104-113, Nov. 2010

3 L. Deng, X. Yu, Y. Wang, Md. Noor-A-Rahim, Y. Guan, Z. Shi and Z. Zhang, "Joint source channel anytime coding based on spatially coupled repeat-accumulate codes," *IEEE Tran. Commun.*, early access, 2023, doi:10.1109/TCOMM.2023.3303955

4 Y. Dong, K. Niu, J. Dai, S. Wang, Y. Yuan, "Joint source and channel coding using double polar codes," *IEEE Commun. Letters*, vol. 25, no. 9, pp. 2810-2814, Sept. 2021

5 A. Golmohammadi and D. G. M. Mitchell, "Concatenated spatially coupled LDPC codes with sliding window decoding for joint source-channel coding," *IEEE Trans. Commun.*, vol. 70, no. 2, pp. 851-864, Feb. 2022.

6 J. He, L. Wang and P. Chen, "A joint source and channel coding scheme based on simple protograph structured codes," *in Proc. Symp. Commun. and Info. Technologies (ISCIT)*, pp. 65-69, Sydney, Oct. 2012

7 Q. Chen, L. Wang, S. Hong, and Z. Xiong, "Performance improvement of JSCC scheme through redesigning channel codes," *IEEE Commun. Letters*, vol. 20, no. 6, pp. 1088-1091, June 2016

8 C. Chen, L. Wang and S. Liu, "The design of protograph LDPC codes as source codes in a JSCC system," *IEEE Commun. Letters*, vol. 22, no. 4, pp. 672- 675, Apr. 2018.

9 H. Neto and W. Henkel, "Multi-edge optimization of low-density parity-check codes of joint source-channel coding," *IEEE Systems, Commun. and Coding (SCC)*, January 21-24, 2013 in Munich, Germany, pp. 1-6

10 Q. Chen, F. C. M. Lau, H. Wu, and C. Chen,"Analysis and improvement of error-floor performance for JSCC scheme based on double protograph LDPC codes," *IEEE Trans. Vehi. Tech.*, vol. 69, no. 12, pp. 14316-14329, 2020

11 S. Liu, C. Chen, L. Wang, and S. Hong, "Edge connection optimization for JSCC system based on DP-LDPC codes," *IEEE Wireless Commun. Letters*, vol. 8, no. 4, pp. 996-999, Aug. 2019.

12 C. Chen, L. Wang, F. C. M. Lau, "Joint optimization of protograph LDPC code pair for joint source and channel coding," *IEEE Trans. on Commum.*, vol. 66, no. 8, pp. 3255-3267, Aug. 2018.

13 S. Liu, L. Wang, J. Chen, S. Hong, "Joint component design for the JSCC system based on DP-LDPC codes," *IEEE Trans. on Commum.*, vol. 68, no. 9, pp. 5808-5818, Sept. 2020.

14 D. Song, L. Wang and P. Chen, "Mesh model-based merging method for DP-LDPC code pair," *IEEE Trans. Commun.*, vol. 71, no. 3, pp. 1296-1308, March 2023

15 H. Lin, S. Lin, K. Abdel-Ghaffar, "Integrated code design for a joint source and channel LDPC coding scheme, " *In Proceedings of the IEEE Information Theory and Applications Workshop (ITA)*, San Diego, CA, USA, 12-17 February 2017; pp. 1-9.

16 S. Hong, J. Ke, L. Wang, "Global design of double protograph LDPC Codes for joint source-channel coding," *IEEE Commun. Letters*, vol. 27, no. 2, pp. 424-427, Feb. 2023.

17 Z. Xu, D. Song, J. Zheng, and L. Wang," Joint-decoding-complexity-oriented collaborative design for joint source-channel coding system based on double protograph-LDPC codes," *Science China Information Sciences*, 2023, 66(8)

18 J. Zhan, and F. Lau, "Joint design of source-channel codes with linear source encoding complexity and good channel thresholds based on double-protograph LDPC codes," *IEEE Commun. Letters*, vol. 27, no. 11, pp. 2909-2913, Nov. 2023.

19 X.-Y. Hu, E. Eleftheriou, and D. M. Arnold, "Regular and irregular progressive edge-growth Tanner graphs," *IEEE Trans. Inform. Theory*, vol. 51, pp. 386–398, Jan. 2005.