

# GEOMETRY BASED DESIGNS OF LDPC CODES

Haotian Zhang and José M. F. Moura

Department of Electrical and Computer Engineering  
Carnegie Mellon University, Pittsburgh, PA 15213  
{haotian, moura}@ece.cmu.edu

## ABSTRACT

In this paper we construct three types of low-density parity-check codes with column weight  $j = 3$  based on geometries in graphical models. Low-density parity-check codes with  $j > 2$  are desired because their minimum distance improves linearly with the code block length  $n$ . The codes we present here have girth 8 and girth 10. All codes are regular and well-structured. These codes have flexible block lengths and code rates, and may be used in the area of communications and data storage. Our simulation results show that they have better bit-error-rate decoding performance and lower error floors in additive white Gaussian noise channels than randomly constructed low-density parity-check codes.

## 1. INTRODUCTION

Low-density parity-check (LDPC) codes [1] can perform very close to the Shannon limit when iteratively decoded by the sum-product algorithm [2]. Since it has been well known that the existence of cycles in the Tanner graphs of LDPC codes taxes the computing effort of the sum-product decoding algorithm and prevents it from converging to the optimal result [3], cycles, especially short cycles, affect negatively the performance of LDPC decoders. Therefore, LDPC codes with large girth are desired.

Much work has been devoted to the construction of large-girth LDPC codes. Kou, Lin, and Fossorier, see[4], present a geometric approach to construct four classes of (quasi-)cyclic LDPC codes with girth  $g = 6$ . The construction of these codes are based on the lines and points of Euclidean and projective geometries over finite fields. Recently, structured regular LDPC codes have been constructed from balanced incomplete block designs (BIBD), a thoroughly studied concept of combinatorial designs. Please refer to [5, 6] for details. Song presents in his Ph.D. thesis [7] several classes of regular LDPC codes with column weight  $j = 2$  and girth  $g \leq 12$  and  $j = 3$  and  $g \leq 8$  based on disjoint-difference sets (DDS) and permutation matrices (PM). These codes are quasi-cyclic and well structured.

We present in [8, 9, 10] three classes of structured regular LDPC codes with column weight  $j = 2$  and girth  $g = 12, 16, 20$ , respectively. In this paper, we extend these results and focus on the design of LDPC codes with  $j = 3$  and  $g = 8, 10$ . LDPC codes with column weight  $j > 2$  are desirable since their minimum distance grows linearly with the code block length  $n$ , in contrast with LDPC codes with  $j = 2$ , where the minimum distance grows only logarithmically with  $n$ . In Section 2, we describe the construction of three classes of LDPC codes with column weight  $j = 3$ : two of them have girth 8, and the other one has girth 10. Simulation results and analysis are provided in Section 3. Finally, Section 4 concludes the paper.

## 2. CODE CONSTRUCTIONS

*Preliminaries* Our construction is based on geometries in graphical models. We give here a brief description of the graphical model that is also used in [8, 9, 10] to design LDPC codes with column weight  $j = 2$ . Let  $H$  be the parity check matrix of an LDPC code with  $v$  parity check equations, i.e.,  $H$  is  $v \times n$ . We represent these parity check equations by a set  $X$  of  $v$  points. We call  $X$  the *point set* of the LDPC  $H$  matrix. For LDPC codes with  $j = 3$ , each column of the  $H$  matrix is represented by a triangle composed of three points in  $X$  that correspond to the three nonzero elements in this column. We refer to this as *column triangle* for reasons to be explained shortly. We call the resulting graph the *structure graph* for the LDPC  $H$  matrix.

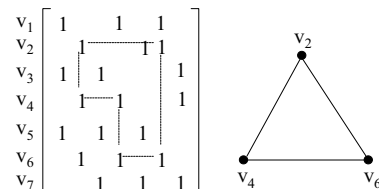


Fig. 1. A 6-cycle in an  $H$  matrix and its structure graph.

The structure graph helps to identify easily cycles in the LDPC  $H$  matrices. Two distinct connection lines between

Research supported by the Data Storage Systems Center (DSSC) at Carnegie Mellon University and by NSF grant # ECS-0225449.

two nodes in a structure graph stand for a 4-cycle. A 6-cycle is just a triangle comprising three points and three connection lines between any two of them. Figure 1 shows the relationship between the  $H$  matrix of an LDPC code with column weight 3 and its structure graph for a 6-cycle. To distinguish this from the column triangle, we refer to it as *cycle triangle*. A triangle is a cycle triangle when its three edges belong to three different columns; otherwise it is a column triangle. Figure 2 shows several examples of triangles in a structure graph for LDPC codes with  $j = 3$ :  $\{a,b,f\}$ ,  $\{b,c,d\}$ , and  $\{d,e,f\}$  are column triangles;  $\{b,d,f\}$  is a cycle triangle, since its edges are all shared with the previous triangles. An 8-cycle is a loop composed of four points and four tail-biting connection lines. It is similar for cycles with length greater than 8.

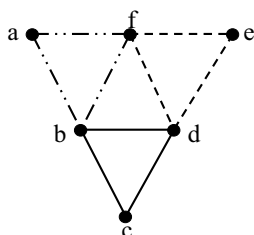


Fig. 2. Triangles in the structure graph.

Before presenting our constructions, we introduce two needed concepts: slope of an edge and admissible slope pair.

Consider a set of points  $X = \{a_1, \dots, a_p, b_1, \dots, b_p\}$  that we divide into two subsets  $X_1 = \{a_1, \dots, a_p\}$  and  $X_2 = \{b_1, \dots, b_p\}$ . Position the points in  $X_1$  and  $X_2$  in a vertical line, as shown in Figure 3 for  $p = 8$ . Label sequentially from bottom to top the points in subsets  $X_1$  and  $X_2$ . Note that Figure 3 repeats twice the set  $X$ .

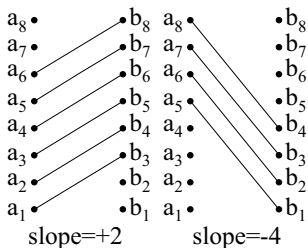


Fig. 3. Edges with specific slopes.

The *slope*  $s$  of an edge between points  $a_i \in X_1$  and  $b_j \in X_2$  is  $s = j - i$ . Slopes take values in the range  $-(p - 1) \leq s \leq (p - 1)$ , where  $p$  is the number of points in each subset. Hence, there are  $(2p - 1)$  possible slopes. Figure 3 illustrates on the left  $s = +2$  and on the right  $s = -4$ . We assign the points in  $X_1$  as reference points when

calculating the slopes. The number of possible edges with slope  $s$  is  $(p - |s|)$ .

A slope pair  $(s, s')$  is an *admissible slope pair (ASP)* iff  $s' = -\text{sgn}(s) \cdot \text{mod}(p - |s|)$ . The slopes  $s$  and  $s'$  are referred to as *mirror slopes*. Each ASP  $(s, s')$  introduces  $(p - |s|) + (p - |s'|) = p$  edges and increases the degree of each point in  $X$  by 1.

*(n, 3, k) LDPC codes with girth 8* To obtain LDPC codes with girth 8, we eliminate 4-cycles and 6-cycles. Cycles of length 4 are avoided as long as each edge belongs to a single column triangle. This can be accomplished easily. Therefore, the major task is to avoid 6-cycles.

To construct 6-cycle-free  $(n, 3, k)$  LDPC codes with high code rates, we should introduce as many columns in  $H$  as possible for the same number  $v$  of parity checks. The selection of columns follows two basic rules: each edge is used at most once to construct column triangles; each new introduced column should not give rise to cycle triangles.

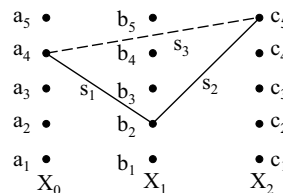


Fig. 4. A triangle in the structure graph when  $p = 5$ .

Assume  $v = 3p$ , with  $p$  an integer. We partition these points into three subsets  $X_0, X_1$ , and  $X_2$  of equal size  $p$ . The points in each subset are aligned in a vertical line, as shown in Figure 4, for  $p = 5$ , and labelled as indicated in the figure. We only consider edges in the structure graph that connect points in different subsets. We introduce the concept of *section*. The section  $S_i$  represents all the edges that are introduced between two neighboring subsets  $X_i$  and  $X_{\text{mod}(i+1,3)}$ ; it corresponds to an ASP set  $\mathcal{A}_i$ . The slope of each edge in section  $S_i$  is calculated using a point in  $X_i$  as the reference point. We can prove that these ASP sets are of the same size. To avoid 6-cycles in  $j = 3$  LDPC codes with high code rates, the task is to construct an admissible slope set  $\mathcal{A}_i$  for each section  $X_i$  that introduces as many column triangles as possible without introducing cycle triangles in the code structure graph.

The following two facts underlie our construction of girth  $g = 8$   $(n, 3, k)$  LDPC codes. The proofs are omitted here.

**Fact 1** *In a structure graph as described above, any column triangle or cycle triangle must be composed of three edges in three different sections with slopes  $s_0, s_1$ , and  $s_2$ ,*

respectively, satisfying  $\text{mod}(s_0 + s_1 + s_2, p) = 0$ , as shown in Figure 4.

**Fact 2** Assume the three ASPs  $(s_0, s'_0)$ ,  $(s_1, s'_1)$ , and  $(s_2, s'_2)$  belong to ASP sets  $\mathcal{A}_0$ ,  $\mathcal{A}_1$ , and  $\mathcal{A}_2$ , respectively, and  $\text{mod}(s_0 + s_1 + s_2, p) = 0$ . Then, after introducing all the edges corresponding to all the three ASPs,  $p$  triangles occur, and any two of these have no common edge.

To construct  $(n, 3, k)$  LDPC codes with girth 8, we need to find three sets  $\mathcal{A}_i$  of admissible slope pairs one for each of the three sections  $S_i$ ,  $i = 0, 1, 2$ . These sets satisfy the following conditions:

1. **Condition 1** The three ASP sets have the same cardinality  $N_s$ , i.e.,  
 $\mathcal{A}_i = \{(s_{i1}, s'_{i1}), \dots, (s_{iN_s}, s'_{iN_s})\}$ .
2. **Condition 2** The slope pairs in each  $\mathcal{A}_i$  must satisfy the following conditions:
  - 2.1 **Condition 2.1** For  $1 \leq j \leq N_s$ ,  $\text{mod}(s_{0j} + s_{1j} + s_{2j}, p) = 0$ .
  - 2.2 **Condition 2.2** If  $v(s_{0i} + s_{1j} + s_{2k}, p) = 0$ , then  $i = j = k$ .

By Fact 2, Condition 2.1 guarantees that the three slope pairs  $(s_{0j}, s'_{0j})$ ,  $(s_{1j}, s'_{1j})$ , and  $(s_{2j}, s'_{2j})$ , one from each ASP set, introduce  $p$  column triangles. Condition 2.2 guarantees that no cycle triangles are introduced.

We present next our construction for LDPC codes with girth  $g = 8$ . Let  $\mathcal{B}$  be the set of all possible admissible slope pairs, and introduce three copies  $\mathcal{B}_i = \mathcal{B}$ ,  $i = 0, 1, 2$ . Call  $\mathcal{B}_i$  the candidate ASP set for  $\mathcal{A}_i$ .

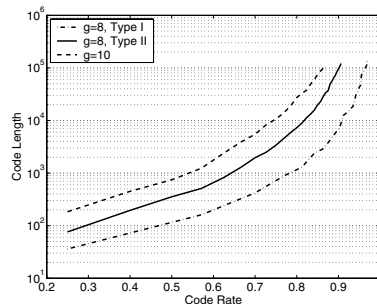
### Construction 1 (Type-I $(n, 3, k)$ LDPC codes, $g = 8$ )

1. **Initialization** Set  $\mathcal{A}_i = \phi$ ,  $i = 0, 1, 2$ , the empty set.
2. Choose two arbitrary ASPs  $(s_{i1}, s'_{i1})$  one from each of the two sets  $\mathcal{B}_i$ ,  $i = 0, 1$ . Let  $s_{21} = \text{mod}(-s_{01} - s_{11}, p)$ . Move  $(s_{i1}, s'_{i1})$  into  $\mathcal{A}_i$ , and delete it from the corresponding  $\mathcal{B}_i$ ,  $i = 0, 1, 2$ . Set  $k = 2$ .
3. If  $\mathcal{B}_0$  is empty, go to step 9; otherwise, choose an arbitrary ASP  $(s_{0k}, s'_{0k})$  from  $\mathcal{B}_0$ , move it to  $\mathcal{A}_0$ , and delete it from  $\mathcal{B}_0$ .
4. If there exist a  $j$ ,  $1 \leq j \leq k - 1$ , such that  $(s_l, s'_l) \in \mathcal{A}_2$ , where  $s_l = \text{mod}(-s_{0k} - s_{1j}, p)$ , then delete  $(s_{0k}, s'_{0k})$  from  $\mathcal{A}_0$ , and go to step 3.
5. If  $\mathcal{B}_1$  is empty, go to step 9; otherwise, let  $\mathcal{B}'_1 = \mathcal{B}_1$ .
6. If  $\mathcal{B}'_1$  is empty, delete  $(s_{0k}, s'_{0k})$  from  $\mathcal{A}_0$ , and go to step 3; otherwise, select an arbitrary ASP  $(s_{1k}, s'_{1k})$  from  $\mathcal{B}'_1$ , move it to  $\mathcal{A}_1$ , and delete it from  $\mathcal{B}'_1$ .
7. If there exist a  $j$ ,  $1 \leq j \leq k - 1$ , such that  $(s_l, s'_l) \in \mathcal{A}_2$ , where  $s_l = \text{mod}(-s_{0j} - s_{1k}, p)$ , then delete  $(s_{1k}, s'_{1k})$  from  $\mathcal{A}_1$  and  $\mathcal{B}_1$ , and go to step 6.

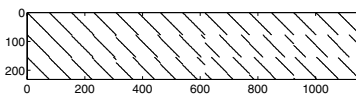
8. Let  $s_{2k} = \text{mod}(-s_{0k} - s_{1k}, p)$ . If  $(s_{2k}, s'_{2k}) \in \mathcal{B}_2$ , delete  $(s_{2k}, s'_{2k})$  from  $\mathcal{B}_2$ , move it to  $\mathcal{A}_2$ , set  $k = k + 1$ , and go to step 3. Otherwise, go to step 6.
9. End.

The cardinality of the sets  $\mathcal{A}_i$  is  $N_s = k - 1$ . The sets  $\mathcal{A}_i$ ,  $i = 0, 1, 2$ , are the desired ASP sets. The corresponding LDPC code has girth  $g = 8$ .

In Construction 1, since each triple of admissible slope pairs from three different admissible slope sets satisfying the above equation will contribute  $p$  column triangles, i.e.,  $p$  columns, to  $\mathbf{H}$ , the block length of the code is  $n = N_s \cdot p$ , and the code rate is  $r = (N_s - 3)/N_s$ . To obtain codes with higher code rates, it is important to maximize  $N_s$  by searching over the admissible slope sets. Attainable code rates and the corresponding code block lengths are shown by the top dashed line in Figure 5, and Figure 6 shows the structure of one parity check matrix  $\mathbf{H}$  when  $v = 231$ ,  $n = 1155$ , and  $r = 0.8$ . The  $\mathbf{H}$  matrix is well-structured.



**Fig. 5.** Code lengths for proposed codes as a function of code rates.



**Fig. 6.**  $\mathbf{H}$  matrix for a Type-I code with  $v = 231$ ,  $n = 1155$ , and  $r = 0.8$ .

Though Type-I  $(n, 3, k)$  LDPC codes with girth 8 have no 6-cycles, they will have in general many 8-cycles. It can be shown that 8-cycles have the two structural forms shown in Figure 7. We refer to the top and bottom 8-cycles in Figure 7 as Form-I and Form-II 8-cycles, respectively. The four edges of a Form-I 8-cycle are all in the same section; these four edges form a 8-cycle when  $|s_{i1} - s_{i2}| = |s_{j1} - s_{j2}|$ . The four edges of a Form-II 8-cycle are in two contiguous sections; with respect to the bottom 8-cycle of Figure 7, the four edges with slopes  $s_{0a}, s_{0b}, s_{1a}$ , and  $s_{1b}$  satisfy  $\text{mod}(s_{0a} + s_{1a}, p) = \text{mod}(s_{0b} + s_{1b}, p)$ , where  $(s_{0a}, s'_{0a}), (s_{0b}, s'_{0b}) \in \mathcal{A}_0$ , and  $(s_{1a}, s'_{1a}), (s_{1b}, s'_{1b}) \in \mathcal{A}_1$ .

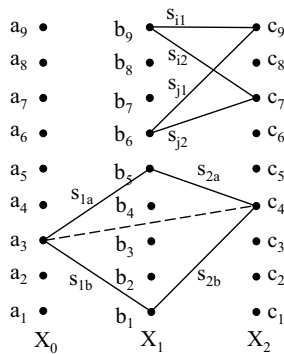
Construction 1 can be modified to eliminate all of these 8-cycles by introducing additional conditions on the ASPs. Of course, the net effect is to reduce the code rate for a given  $v$ . A tradeoff is to eliminate some of these 8-cycles. We refer to the resulting codes as Type-II  $(n, 3, k)$  LDPC codes with girth  $g = 8$ .

The 8-cycles we delete are Form-II cycles with edges in sections  $S_0$  and  $S_1$ . We add this additional condition when searching for admissible slope sets to avoid these 8-cycles.

**Construction 2 (Type-II  $(n, 3, k)$  LDPC codes,  $g = 8$ )** Repeat the steps 1–7 in Construction 1. Then proceed as follows.

8. Let  $s_{2k} = \text{mod}(-s_{0k} - s_{1k}, p)$ . If  $(s_{2k}, s'_{2k}) \in \mathcal{B}_2$ , delete  $(s_{2k}, s'_{2k})$  from  $\mathcal{B}_2$ , and move it to  $\mathcal{A}_2$ , and set  $k = k + 1$ . Otherwise, go to step 6.
9. For  $1 \leq j \leq k - 2$ , let  $s_{1j} = \text{mod}(-s_{0j} - s_{1(k-1)}, p)$ , and  $s_{mj} = \text{mod}(-s_{0(k-1)} - s_{1j}, p)$ . For  $1 \leq j \leq k - 2$ , if  $(s_{1j}, s'_{1j}) \in \mathcal{B}_2$ , delete it from  $\mathcal{B}_2$ ; if  $(s_{mj}, s'_{mj}) \in \mathcal{B}_2$ , delete it from  $\mathcal{B}_2$ . If  $\mathcal{B}_2$  is not empty, go to step 3.
10. End.

The cardinality of the sets  $\mathcal{A}_i$  is  $N_s = k - 1$ . The sets  $\mathcal{A}_i$ ,  $i = 0, 1, 2$ , are the desired ASP sets. The resulting  $\mathbf{H}$  matrix corresponds to a regular  $(n, 3, k)$  LDPC code that still has girth  $g = 8$  but has a lower number of 8-cycles.



**Fig. 7.** Top: Form-I 8-cycle; Bottom: Form-II 8-cycles.

The available code rates and block lengths for Type-II codes are shown by the solid line in Figure 5. The well-structured  $\mathbf{H}$  matrix of a Type-II code is similar to that of a Type-I code.

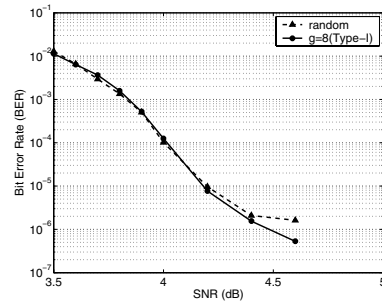
Since Construction 2 deletes some of the Form-II 8-cycles in Type-I codes, we expect Type-II codes to have better decoding performance than Type-I codes. Simulation results confirm this. As mentioned, this performance gain is at the cost of a lower code rate for the same  $v$ , because the more strict conditions in searching for the admissible slope sets

decrease the value of possible  $N_s$ . Therefore, to achieve a given code rate, Type-II codes need larger block lengths than Type-I codes.

*$(n, 3, k)$  LDPC codes with girth 10* By introducing additional conditions on the ASPs to avoid all of 8-cycles, Construction 2 can be modified to obtain LDPC codes with girth  $g = 10$ . In the construction of LDPC codes with girth 10, we need to eliminate all Form-I 8-cycles in sections  $S_0$ ,  $S_1$  and  $S_2$ , and all Form-II 8-cycles with edges in contiguous sections  $S_0$  and  $S_1$ ,  $S_1$  and  $S_2$ , and  $S_2$  and  $S_0$ . The process to eliminate Form-II 8-cycles in contiguous sections  $S_1$  and  $S_2$ , and  $S_2$  and  $S_0$  is similar to the steps 8 and 9 in Construction 2 for avoiding Form-II 8-cycles in  $S_0$  and  $S_1$ . For the process to eliminate Form-I 8-cycles, please refer to the construction of  $(n, 2, k)$  LDPC codes with girth 12 in [9] for details, since they have the same form of 8-cycles. We omit the details of the construction here due to space limitations. Refer to the bottom dashed line in Figure 5 for the available code rates and block lengths for the codes with girth 10. As shown in the figure, the codes with girth 10 have lower code rates than those of Type-I and Type-II codes with girth 8 for given block lengths. The  $\mathbf{H}$  matrix of a code with girth 10 is also well-structured, similar to that of a Type-I code with girth 8.

### 3. SIMULATION RESULTS

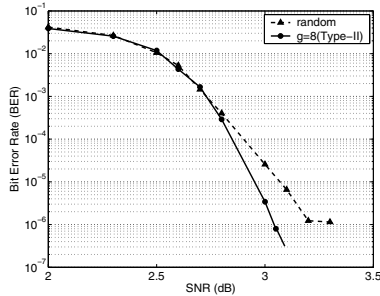
In this section, we compare the bit error rate (BER) of the proposed LDPC codes with the BER of randomly constructed LDPC codes in additive white Gauss noise (AWGN) channels. These codes are decoded by the sum-product algorithm, and we adopt the rate-adjusted signal to noise ratio (SNR) defined in [11]  $\text{SNR}_1 = 10 \log_{10} [E_b / (2r\sigma^2)]$ .



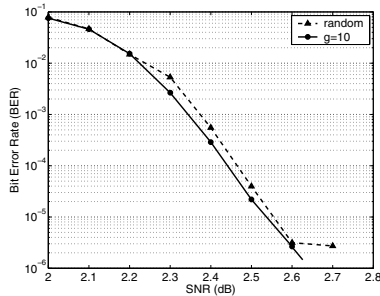
**Fig. 8.** Comparison of BER performance for random code with large-girth LDPC codes (girth 8, Type-I).

Figure 8 compares the BER performance of a Type-I code with girth 8 and a random LDPC code. Both codes have the same block length 6990 and the same code rate 9/10. They show similar BER performance across the SNR region tested, though the Type-I code outperforms the random LDPC code by 0.02dB at  $\text{BER} = 10^{-5}$ .

Figure 9 compares the BER performance of a Type-II



**Fig. 9.** Comparison of BER performance for random code with large-girth LDPC codes (girth 8, Type-II).



**Fig. 10.** Comparison of BER performance for random code with large-girth LDPC codes (girth 10).

code with girth 8 and a random  $(n,3,k)$  LDPC code. Both codes have the same block length 6986 and the same code rate 11/14. They show similar BER performance in the low SNR region; however, in the high SNR region, the Type-II code outperforms the random LDPC code by 0.15dB at  $\text{BER} = 10^{-5}$ . The Type-II code also shows a lower error floor than that of the random code.

Figure 10 compares the BER performance of a code with girth 10 and a random  $(n,3,k)$  LDPC code. Both codes have the same block length 8360 and the same code rate 8/11. The code with girth 10 shows only slightly better BER performance than that of the random code up to  $\text{BER} = 10^{-5}$ ; but, for lower BER, the girth 10 code exhibits lower error floor.

#### 4. CONCLUSION

In this paper, we present three types of regular  $(n, 3, k)$  LDPC codes with girth 8 and 10. These codes are systematically constructed, and their  $H$  matrices are well-structured and completely represented by a set of integers (parity check numbers and slopes in ASP sets). These codes can be used in various areas, e.g., communications and data storage, since they have flexible code rates and block lengths. In the high SNR region, these codes with large girth have better BER decoding performance than the performance of randomly

constructed codes over AWGN channels. These codes also show a lower error floor than that of random codes.

#### 5. REFERENCES

- [1] R. G. Gallager, *Low-Density Parity Check Codes*, MIT Press, Cambridge, MA, 1963.
- [2] D. J. C. Mackay and R. M. Neal, "Good codes based on very sparse matrices," in *Cryptography and Coding, 5th IMA Conference, in Lecture Notes in Computer Science*, C. Boyd, Ed., 1995, vol. 1025, pp. 110–111.
- [3] R. J. McEliece, D. J. C. Mackay, and J. F. Cheng, "Turbo decoding as an instance of Pearl's 'belief propagation' algorithm," *IEEE J. Selected Areas Commun.*, vol. 16, no. 2, pp. 140–152, Feb 1998.
- [4] Yu Kou, Shu Lin, and Marc P. C. Fossorier, "Low-density parity-check codes based on finite geometries: a rediscovery and new results," *IEEE Trans. Inform. Theory*, vol. 47, no. 7, pp. 2711–2736, Nov 2001.
- [5] S. J. Johnson and S. R. Weller, "Regular low-density parity-check codes from combinatorial design," in *Inf. Tech. Workshop (ITW) 2001*, Cairns, Australia, Sep 2001.
- [6] B. Vasic, "Structured iteratively decodable codes based on Steiner systems and their application in magnetic recording," in *IEEE Globecom 2001*, San Antonio, TX, Nov 2001, vol. 5, pp. 2954–2960.
- [7] H. Song, "Iterative soft detection and decoding for data storage channels," in *Ph.D. Thesis*, Dept. of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA, USA, 2003.
- [8] H. Zhang and José M. F. Moura, "Large-girth LDPC codes based on graphical models," in *IEEE Signal Processing and Wireless Communications (SPAWC) Workshop*, Rome, Italy, Jun 2003.
- [9] H. Zhang and José M. F. Moura, "Structured regular LDPC codes with large girth," in *IEEE Globecom 2003*, San Francisco, CA, USA, Dec 2003.
- [10] José M. F. Moura, J. Lu, and H. Zhang, "Structured LDPC codes with large girth," *IEEE Signal Processing Magazine*, vol. 21, no. 1, pp. 42–55, Jan. 2004.
- [11] D. J. C. Mackay, "Good error-correcting codes based on very sparse matrices," *IEEE Trans. Inform. Theory*, vol. 45, no. 2, pp. 399–431, March 1999.