

一种短小的高码率 LDPC 码设计

巫光福, 王琳

厦门大学通信工程系, 福建厦门 361005

摘要: 提出一种高码率系统——LDPC 码的构造方法. 在列重量一定的情况 (通常列重量 $W_c \geq 3$) 下, 构造满足给定围长条件的子矩阵, 然后将该子矩阵和单位阵合并成 LDPC 码的校验矩阵. 构造了 28×76 和 64×328 的校验矩阵. 仿真表明, 这两个 LDPC 码与校验矩阵为 42×105 、 170×425 、 66×330 的 LDPC 码在 AWGN 信道下的性能相比具有更好的 BER. 同时还基于拟阵理论给出了 Tanner 图的围长的充分条件, 可用于构造给定的短围长 LDPC 码. 短的高码率 LDPC 码可用于未来的手持数字视频广播.

关键词: 高码率; 误码率; 编码增益; LDPC 短码; 围长条件

中图分类号: TN911.22 文章编号: 0255-8297(2013)06-0559-05

Construction of High Rate LDPC Codes with Short Block Length

WU Guang-fu, WANG Lin

Department of Communication Engineering, Xiamen University, Xiamen 361005, Fujian Province, China

Abstract: A method of constructing low-density parity-check (LDPC) codes is presented, where an identity matrix is combined with another sub-matrix constructed under the condition of a given girth and fixed weight of column to generate a parity check matrix. Let W_c denote the fixed weight of column, which is usually not less than 3. Two kinds of check matrices sized 28×76 and 64×328 are constructed. Simulation results over additive white Gaussian noise (AWGN) channels show that performance of the codes is better than that of check matrices sized 42×105 , 170×425 and 66×330 . Meanwhile, sufficient conditions for the girth of Tanner graph based on Matroid theory are present. The girth conditions can be used to construct LDPC codes with given short girths. High-rate LDPC codes with short block length can be applied to the future digital video broadcast-handsets.

Keywords: high rate, bit error rate, coding gain, low-density parity-check codes with short block length, girth condition

低密度奇偶校验码(low-density parity-check, LDPC)早在 1963 年就已提出^[1], 但受当时计算机计算速度的限制没有引起关注, 直到 1996 年才重新得到学者们的重视. 文献 [2-3] 根据图的扩展方法设计 LDPC 码, 使其在编码、解码方面都具有线性时间复杂度且表现出相当好的性能. 随后, 文献 [4] 发现 LDPC 码在码长很长的情况下具有靠近香农限的性能. LDPC 码自从被重新发现后, 因具有靠近香农限的性能、低误码平台以及低译码复杂度而成为信道纠

错编码领域的研究热点.

LDPC 码的构造方法一般分两种: 一种是随机构造方法, 如文献 [5] 列举了一些构造 LDPC 码校验矩阵的随机或半随机方法; 另一种是代数构造方法, 主要构造几何 LDPC 码^[6]和准循环码^[7]. 该类码在码长很长的情况下也具有靠近香农限的性能, 且没有错误地板; 但在码长很短且码率很高的情况下, 上述两种方法构造的 LDPC 码没有明显优势, 如码长小于 500 和码率大于 0.6 的情况. 有关短码的研究主要集

收稿日期: 2012-06-05; 修订日期: 2012-09-01

基金项目: 国家自然科学基金(No.60972053)资助

作者简介: 巫光福, 博士生, 研究方向: 通信系统信道纠错编码, E-mail: wuguangfu@126.com; 王琳, 教授, 博导, 研究方向: 通信物理层算法的设计分析与实现, E-mail: wangling@xum.edu.cn

中在代数码, 如著名的 Golay 码、QR 码, 虽然都具有最大或接近最大值的最小码距, 但因它们校验矩阵中 1 的数目分布不具有稀疏特性, 故不适用于置信传播 (belief propagation, BP) 迭代译码. 因此, 本文基于文献 [5] 构造了一类新的高码率 LDPC 短码, 其性能与当前性能较好的同类码^[8]相比, 具有更大的编码增益.

1 LDPC 码的基本概念

1.1 基本概念

LDPC 码是一类线性分组码, 主要由 $(n-k) \times n$ 的校验矩阵 H 确定. 根据校验矩阵可得到 $k \times n$ 生成矩阵 G , 其中 k 和 n 分别表示信息比特长度和码长. 二进制 LDPC 码的校验矩阵 H 中 1 的数目占总数的比例很少, 故被称为低密度奇偶校验码. 根据每行和每列中 1 的数目可以将 LDPC 码的种类分为规则和非规则 LDPC 码. 如果每行中 1 的数目都为固定整数 W_r , 且每列中 1 的数目都为固定整数 W_c , 则称为规则 (W_c, W_r) LDPC 码, 否则称为非规则 LDPC 码.

除了用校验矩阵表示 LDPC 码外, 还可用 Tanner 图^[9]表示. 例如校验矩阵 H 为

$$H = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 \end{pmatrix}$$

其对应的 Tanner 图如图 1 所示.

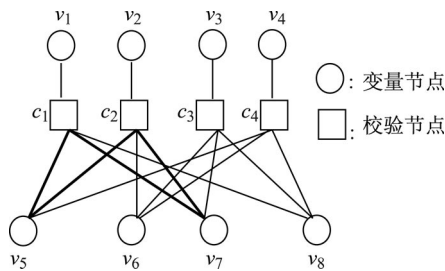


图 1 H 矩阵的 Tanner 图

Figure 1 Tanner graph of H matrix

任何 Tanner 图中的环都是指从其中一个顶点出发经过数个顶点又返回到原点的一条轨迹. 例如, 图 1 中的粗线形成了一个长度为 4 的环, 因为 Tanner 图是二部图, 所以环的长度都是 2 的倍数. 围长是指 Tanner 图中长度最短的环的大小. Tanner 图中存在环, 而在迭代译码过程中消息的传递会受到环的影响, 特别是长度为 4 的环所造成的影响最为显著. 因此, 在码的构造过程中要尽量避免长度为 4 的环, 目

前关于构造较大围长 LDPC 长码的研究不少^[8,10], 但关于 LDPC 短码的研究却较少. 本文通过拟阵理论构造性能较好的给定围长的 LDPC 短码.

1.2 LDPC 码的常用构造方法

构造 $(n-k) \times n$ 校验矩阵 H 的常用方法如下:

方法 1 先产生零矩阵, 然后随机将矩阵的每列 W_c 个位置替换为 1 (每列中 W_c 个 1 的位置没必要完全不同), 一般得到不规则 LDPC 的 H 矩阵.

方法 2 根据校验矩阵的维数随机产生列重量为 W_c 的 H 矩阵.

方法 3 根据校验矩阵的维数随机产生列重量为 W_c 的 H 矩阵, 同时保持行重尽量均匀.

方法 4 根据校验矩阵的维数构造列重量为 W_c , 行重量为 W_r , 且任意两列重叠的位置不大于 1 的 LDPC 码, 这样就可以避免长度为 4 的环.

方法 5 在方法 4 的基础上避免出现短环, 如长度为 6 和 8 的环. 结果可能出现 H 矩阵的秩小于 H 矩阵的行数情况.

方法 6 在方法 5 的基础上把 H 矩阵分割为两部分子矩阵, 即 $H = [H_1|H_2]$, 使得 H_1 是可逆的.

方法 7 根据 Tanner 图的边扩展方法构造校验矩阵 H ^[11].

方法 8 根据组合设计构造校验矩阵 H ^[12].

方法 9 根据准循环结构构造校验矩阵 H ^[13].

以上构造方法都需要已知校验矩阵的行数和列数, 甚至已知 LDPC 码的度分布. 而本文提出的构造方法既不需要知道 LDPC 码的度分布情况, 也不必已知校验矩阵的列数, 只要给定要传输的信息比特位数, 就可以尽量构造列数达到最大值的校验矩阵, 故具有较高的码率.

2 短的高码率 LDPC 码构造

2.1 基本定理

在构造校验矩阵 H 的过程中应尽量避免其对应的 Tanner 图中的短环, 于是需要一些有关围长的定理. 在已有的文献中, 有关围长的定理都是从矩阵或者图论角度加以考虑的, 如 4 环的检验定理比较容易: 如果校验矩阵的两行或两列存在两个或两个以上的位置数值为 1, 则校验矩阵对应的 Tanner 图存在 4 环^[4]. 有关 6 环及以上环的检验定理不容易从校验矩阵建立. 本文则借用拟阵理论把矩阵问题和图论问题转化为集合问题, 并从拟阵的角度给出有关围长的定理. 为更好地阐明定理 1 和 2, 先给出拟阵的定义 1.

定义 1 一个拟阵 M 是一个有序对 (E, I) , E 是有限集, I 是 E 中子集的集合. 它们满足以下 3 个

条件：

- 1) $\emptyset \in I$.
- 2) 若 $I \in I$, 及 $I' \subseteq I$, 则 $I' \in I$.
- 3) 若 $I_1 \in I, I_2 \in I$ 且 $|I_1| < |I_2|$, 则存在 $e \in I_2 - I_1$ 使得 $I_1 \cup e \in I$.

本文根据拟阵理论将校验矩阵的每一列看作一个集合, 把矩阵中列与列之间的关系转化为集合与集合之间的关系, 得到了一系列性质定理.

校验矩阵 H 用集合来表示的方法如下: 设 $r = n - k, H_{r \times n} = [h_1, \dots, h_n], h_i, i = 1, \dots, n$, 是 r 维的列向量, 而向量 h_i 中非零的坐标又可以用集合 L_i 来表示. 如图 1 所对应的校验矩阵可表示为如下集合: $L_1 = \{1\}, L_2 = \{2\}, L_3 = \{3\}, L_4 = \{4\}, L_5 = \{1, 2, 4\}, L_6 = \{2, 3, 4\}, L_7 = \{1, 2, 3\}, L_8 = \{1, 3, 4\}$. 反之, 只要知道 $L_i, i = 1, \dots, 8$, 就可以构造出相应的 H 矩阵. 下面从拟阵的角度给出有关 4 环的定理 1 以及围长大于 6 的定理 2, 并根据定理 1 和 2 推导出了更一般的围长定理.

定理 1 校验矩阵对应的 Tanner 图存在四环当且仅当 $L_i, i = 1, \dots, n$, 中存在两个集合的交集元素的数目大于或等于 2.

证明 从校验矩阵与 Tanner 图的关系可以明显看出, 只要 L_i 中存在两个集合的交集元素的数目大于等于 2, 则这两个集合所对应的两列至少有两个位置同时为 1, 所以存在四环. 反之也成立.

若 L_i 中任意两个集合交集元素的数目小于或等于 1, 则校验矩阵所对应的 Tanner 图的围长至少为 6.

定理 2 如果 $L_i, i = 1, \dots, n$, 满足下列两个条件, 则校验矩阵所对应的 Tanner 图的围长至少是 8.

- 条件 1** $|L_i \cap L_j| < 2, i \neq j, \forall i, j \in \{1, 2, \dots, n\}$.
- 条件 2** $|(L_i \cap L_j) \cup (L_i \cap L_k) \cup (L_j \cap L_k)| < 3, i \neq j \neq k, \forall i, j, k \in \{1, 2, \dots, n\}$.

证明 用反证法证明, 根据定理 1 和条件 1 可推断出校验矩阵所对应的 Tanner 图的围长至少为 6.

如图 2 所示, 若存在围长为 6 的 Tanner 图, 则在集合 $L_i, i = 1, \dots, n$ 中一定存在 3 个集合, 它们两两相交的并的元素数目大于或等于 3. 与条件 2 互相矛盾, 故校验矩阵所对应的 Tanner 图的围长至少是 8.

定理 1 和 2 可以推广到任意的围长的确定.

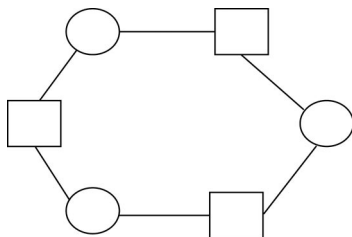


图 2 Tanner 图中长度为 6 的环

Figure 2 A cycle of length six from Tanner graph

定理 3(围长条件) $L(n)$ 表示集合 $\{1, 2, 3, \dots, n\}$, $T(t)$ 表示 $L(n)$ 的任意 t 元子集. 如果 $L_i, i = 1, \dots, n$, 满足式 (1) 中的 t_1 个 ($t_1 > 2$) 条件, 则校验矩阵所对应的 Tanner 图的围长至少是 $2(t_1 + 2)$.

对 $L(n)$ 的任一 t 元子集 $T(t)$ 都有

$$|\cup_{i,j \in T(t) \subseteq L(n)} (L_i \cap L_j)| < t, t = 2, 3, \dots, t_1 + 1 \quad (1)$$

证明 根据定理 2 的证明方法, 由校验矩阵对应的 Tanner 图与集合 $L_i, i = 1, \dots, n$ 的关系可知, 不存在小于 $2(t_1 + 1)$ 的环, 故校验矩阵所对应的 Tanner 图的围长至少是 $2(t_1 + 2)$.

2.2 短的高码率 LDPC 码的构造

根据定理 3 构造给定围长的 LDPC 短码, 但在高码率的情况下构造很大的围长需要很长码长的 LDPC 码. 由于本文只考虑码长在 500 以内的高码率 LDPC 短码, 故只要求避免 4 环的出现即可.

接着构造 $H = [H_1 | H_2]$ 型的 LDPC 码, 其中 H_1 是 $r \times r$ 的单位矩阵. H_2 矩阵是每列的重量为 W_c 通过全面搜索且避免 4 环的条件下得到的. 那么给定初始值 W_c 和 r , 能否确定 H_2 的最大列数? 这个最大列数实际上是存在的, 但难以获得实际的最大列数, 于是根据最大行重的上界得到一个关于列数的上界定理 4.

定理 4 H_2 矩阵的最大列数满足表达式

$$c_{H_2} \leq \left\lfloor \frac{r-1}{w_c-1} \right\rfloor \frac{r}{W_c}$$

证明 由于 H_2 矩阵任意两列中 1 的相同位置不能大于 1, H_2 矩阵的任一行重量小于或等于 $\lfloor \frac{r-1}{W_c-1} \rfloor$.

设 H_2 的列数为 c_{H_2} , 则 H_2 矩阵中所有 1 的数目可表示为 $W_c c_{H_2}$, 故有

$$\left\lfloor \frac{r-1}{W_c-1} \right\rfloor r \geq W_c c_{H_2}.$$

根据下面算法所构造的矩阵的列数很难达到这个上界, 事实上可能不存在最大列数达到上界的矩阵. H_2 矩阵的构造步骤如下:

步骤 1 列举出所有在整数 $1 \sim r$ 之间的任意取 W_c 个不同数的集合.

步骤 2 选取第 1 个集合 L_1 作为矩阵的第 1 列中 1 对应的位置. 从第 2 个集合开始, 根据定理 1 对当前集合与之前所有选取的集合一一进行条件判断, 若都不存在 4 环, 则将该集合作为矩阵新的一列的位置集合.

步骤 3 得到 H_2 矩阵的所有位置集合以及 H_2 矩阵的列数.

构造 H_2 矩阵的伪代码算法

给定 W_c 和 r .

$n_1 = 0$;

for $i_1 = 1$ to $r - W_c + 1$

 for $i_2 = i_1 + 1$ to $r - W_c + 1$

\vdots

 for $i_{W_c} = i_{(W_c-1)} + 1$ to r

$B = \{i_1, i_2, \dots, i_{w_c}\}$

 if ($n_1 = 0$) then

$L_{n_1} = B$

n_1++

 end if

 if ($n_1 > 0$) then

$d = 0$;

 for $j = 1$ to n_1

$n_2 = |L_j \cap B|$

 if ($n_2 > 2$) then

$d = 1$

 end if

 end if

 if ($d = 0$) then

$L_{n_1} = B$

n_1++

 end if

 end for

\vdots

 end for

end for

上述算法不限制 H_2 矩阵的行重量. 如果对 H_2 矩阵的行重量作不同限制, 就可能得到更大或更小的列数的矩阵, 于是可以得到不同码率的 LDPC 码.

3 性能仿真与分析

3.1 仿真条件

本文主要以参数 $W_c = 4$ 进行构造分析, 并与文献 [8-12] 的仿真结果进行对比. 仿真采用 AWGN 信道、BPSK 调制、BP 译码算法, 最大迭代次数为 100. 在低信噪比的情况下, 本文统计 10000 个错误比特为止, 而在高信噪比情况下统计 100 个错误比特为止.

3.2 LDPC 码的性能曲线

首先对比本文设计出的 LDPC 码与文献 [8] 所构造出的 LDPC 码, 得到的仿真结果见图 3 和 4. 在图 3 中, 文献 [8] 中的 LDPC 码的码长为 105, 码率为 0.6, 围长为 12 且最小距离为 6, 它是一种性能比较好的 LDPC 码. 为了进行对比, 本文构造出来的 LDPC 码的码率为 0.6315, 码长为 76, 围长为 6 且最小距离为

5. 码长比文献 [8] 中的 LDPC 码的码长稍短些, 码率略高于文献 [8] 中的 LDPC 码. 但由图 3 可以看出, 本文构造的 LDPC 码的性能优于文献 [8], 在 BER 为 10^{-6} 的数量级下, 约存在 0.2 dB 的编码增益.

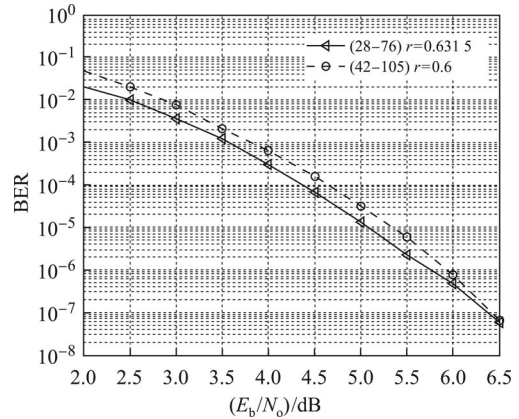


图 3 本文与文献 [8] 所构造的码的误码率对比曲线

Figure 3 Comparison of BER performance between our proposed scheme with that from reference [8]

在图 4 中, 本文构造出码长为 328, 码率为 0.8048 的 LDPC 码, 围长为 6 且最小距离为 5. 该码与文献 [8] 中的码长为 425, 码率为 0.6, 围长为 16 且最小距离为 8 的 LDPC 码性能作比较. 本文构造的码的码长比文献 [8] 中构造的 LDPC 短, 且码率相对文献 [8] 中的码而言略高. 由图 4 可以看出: 当信噪比低于 4.5 dB 时, 文献 [8] 中 LDPC 码的性能稍好, 主要是因为本文构造的 LDPC 码的码率为 0.8048, 而文献 [8] 中的码率为 0.6, 对应的香农门限大约分别为 1.038 和 0.339. 然而, 在高信噪比 4.5 dB 到 5.5 dB 的区间, 本文构造出来的 LDPC 码的性能略优于文献 [8], 在 BER 为 10^{-6} 的数量级下, 大约存在 0.2 dB 的编码增

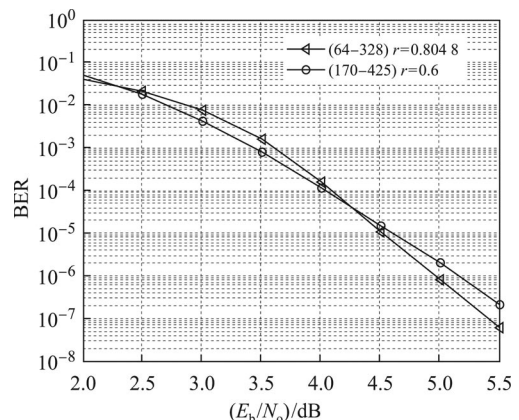


图 4 本文与文献 [8] 所构造的码的误码率对比曲线

Figure 4 Comparison of BER performance between our proposed scheme with that from reference [8]

益,这只能从码的结构上分析原因.文献[8]构造的规则 LDPC 码变量点的度都为 2,校验点的度都为 5.本文构造的 64×328 校验矩阵变量点的度分布如下:64 个变量点的度为 1,264 个变量点的度为 4,校验点的度分布情况见表 1.

表 1 64×328 校验矩阵的校验点的度分布

Table 1 The degree distribution of the check nodes of a 64×328 check matrix

校验点的度	13	15	16	17	18	19	20
校验点数目	7	4	9	10	8	9	17

为了进一步说明本文设计方法的优越性,下面将设计出来的 LDPC 码与文献[12]中的码进行性能比较.文献[12]中 LDPC 码的码长为 330,码率为 0.8,它是根据平衡不完全区组设计构造出来的.本文所构造的是码长为 328,码率为 0.8048,具有围长为 6,最小距离为 5 的 LDPC 码.这两个 LDPC 码具有相近的码长和码率.从性能仿真曲线图 5 可以看出,当信

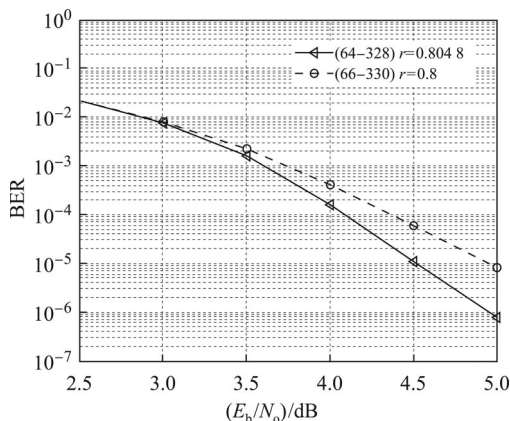


图 5 本文与文献[12]所构造的码的误码率对比曲线

Figure 5 Comparison of BER performance between our proposed scheme with that from reference [12]

噪比低于 3.0 dB 时,本文构造的 LDPC 码与文献[12]中的 LDPC 码的性能相近;但当信噪比高于 3.0 dB 时,本文构造出来的 LDPC 码的性能优于文献[12],且瀑布区比文献[12]的更陡峭.在 BER 为 10^{-5} 的数量级下,大约有 0.45 dB 的编码增益.

4 结 语

本文基于拟阵理论提出了一种新的短的高码率 LDPC 码的设计方法.该方法不同于已有的构造方法,将一个单位矩阵和一个已知列重量为常数且满足一定围长的条件下的子矩阵组合.在 AWGN 信道下

的仿真表明,该类 LDPC 码有较好的性能曲线.但该算法需要对码进行全面搜索,复杂度太高,故只适合构造短的高码率 LDPC 码.未来的研究工作可以考虑引进部分搜索从而降低复杂度,进而能够构造具有很好性能的高码率 LDPC 长码,以期达到靠近香农限的目的.

参考文献:

- [1] GALLAGER R G. Low-density parity check codes [M]. Cambridge, MA: MIT Press, 1963.
- [2] SIPSER M, SPIELMAN D A. Expander codes [J]. IEEE Transactions on Information Theory, 1996, 42(11): 1710-1722.
- [3] SPIELMAN D A. Linear-time encodable and decodable error-correcting codes [J]. IEEE Transactions on Information Theory, 1996, 42(11): 1723-1731.
- [4] MACKAY D J C, NEAL R M. Near Shannon limit performance of low density parity check codes [J]. Electronics Letters, 1997, 33(6): 457-458.
- [5] MACKAY D J C. Good error correcting codes based on very sparse matrices [J]. IEEE Transactions on Information Theory, 1999, 45(2): 399-431.
- [6] KOU Y, LIN S, FOSSORIER M P C. Low-density parity-check codes based on finite geometries: a re-discovery and new results [J]. IEEE Transactions on Information Theory, 2001, 47(7): 2711-2736.
- [7] HUANG Q, DIAO Q, LIN S, ABDEL-GHAFFAR K. Cyclic and quasi-cyclic LDPC codes on constrained parity-check matrices and their trapping sets [J]. IEEE Transactions on Information Theory, 2012, 58(5): 2648-2671.
- [8] ASAMOV T, AYDIN N. LDPC codes of arbitrary girth [C]//Canadian Workshop on Information Theory, 2007: 69-72.
- [9] TANNER R. A recursive approach to low complexity codes [J]. IEEE Transactions on Information Theory, 1981, 27(5): 533-547.
- [10] O'SULLIVAN M E. Algebraic construction of sparse matrices with large girth [J]. IEEE Transactions on Information Theory, 2006, 52(2): 718-727.
- [11] HU X Y, ELEFTHERIOU E, ARNOLD D M. Regular and irregular progressive edge-growth tanner graphs [J]. IEEE Transactions on Information Theory, 2005, 51(1): 386-398.
- [12] FALSAFAIN H, ESMAEILI M. A new construction of structured binary regular LDPC codes based on steiner systems with parameter $t > 2$ [J]. IEEE Transactions on Communications, 2012, 60(1): 74-80.
- [13] HUANG J F, HUANG C M, YANG C C. Construction of one-coincidence sequence quasi-cyclic LDPC codes of large girth [J]. IEEE Transactions on Information Theory, 2012, 58(3): 1825-1836.
- [14] XU Jun, CHEN Lei, ZENG Lan, LAN Lan, LIN Shu. Construction of low-density parity-check codes by superposition [J]. IEEE Transactions on Communications, 2005, 53(2): 243-251.

(编辑:秦 巍)