

重量分布式约束的 LDPC 码性能研究

姚春光¹, 张 健², 葛 新², 王建新²

(1. 国防科技大学电子科学与工程学院, 湖南长沙 410073. 2. 中国电子设备系统工程总公司, 北京 100039)

摘 要: 本文对“重量分布式约束的码集合内码性能”这一命题进行了初步研究, 分别得到了码集合性能的上限和下限, 本文给出了性能下限码的 Fill-Shift 构造方法, 而且由 LDPC 码校验矩阵不变性可以对 LDPC 码的校验矩阵作必要的初等变换, 这样可以在保持码性能不变的前提下降低编码复杂度和实现系统编码; 此外, 还可以利用该性质加强对重要信息符号的差错保护。

关键词: LDPC 码; 密度进化; 码集合; 初等变换; 重量分布式

中图分类号: TN911.22 **文献标识码:** A **文章编号:** 0372-2112 (2007) 11-2135-05

Performance Study of LDPC Codes Constrained by Weight Distribution Polynomials

YAO Chun-guang¹, ZHANG Jian², GE Xin², WANG Jian-xin²

(1. Electronic Science and Engineering College, National University of Defense Technology, Changsha, Hunan 410073, China;

2. China Electronic Equipment System Engineering Corporation, Beijing 100039, China)

Abstract: Research on LDPC codes is in the ascendant nowadays. Our paper has done some work on the proposition of “performance of different codes in the same family constrained by weight distribution polynomials”. Different codes of the same weight distribution are listed, and the results indicate the upper performance bound is determined by density evolution theory and the lower performance bound by the Fill-Shift construction method. Furthermore, the codes in the set of elementary matrix transformation share the same error-correction performance but different coding complexity. We can get a conclusion that codes constrained by weight distribution polynomial can be divided into many groups, performance among different groups is not the same, it depends on the loops length and distribution, and performance in the same group is identical. What we can do now is that to find the best group and do suitable transformations if given weight distribution polynomials. Through this paper, we can see that a new idea for code constructing is also presented. If the relations between big girth of LDPC codes and weight distribution polynomial are discovered, new codes can be achieved only by the weight distribution polynomial, the currently used method given in references for code constructing will be less useful. We can think that we have found a completely different method to construct LDPC codes.

Key words: low density parity check (LDPC) codes; density evolution; code family; elementary transformation; weight distribution polynomial

1 引言

LDPC (Low Density Parity Check) 码是一种线性分组码, 由 Gallager 在 1962 年提出^[1]. 1996 年, MacKay 和 Neal 的研究表明: 采用 LDPC 长码可以达到 Turbo 码同样的性能^[2]. 而最近的研究表明, 被优化了的非规则 LDPC 码能得到比 Turbo 码更好的性能^[3].

文献[4]指出, 在同一重量分布式约束的码集合内, 会有较多的具体码型, 这些码的性能各有差异, 并且进一步提出了一个猜测: 在码集合内, 平均环长较长的码

型应该有较好的纠错性能. 本文对这一命题进行了初步研究, 得到了初步定性结论: 密度进化理论可以得到重量分布式约束的 LDPC 码性能的理论上限; 本文给出的填充-移动法 (Fill-Shift) 可以得到该码集合的性能下限; 对校验矩阵做初等矩阵变换得到的 LDPC 码性能保持不变. 下面将对这些结论进行阐述.

2 LDPC 码校验矩阵的重量分布式表示

采用重量分布式来描述 LDPC 码的校验矩阵. 同一类 LDPC 码校验矩阵的列重量分布可以用分布式表示

为:

$$\lambda(x) = \sum_{i=2}^{d_v} \lambda_i x^{i-1} \quad (1)$$

式中 λ_i 表示重量为 i 的列的重量在矩阵中所占的分量比, d_v 为矩阵中列重量的最大值. 同样, 同一类 LDPC 码校验矩阵的行重量分布式可以采用下式描述:

$$\rho(x) = \sum_{j=2}^{d_c} \rho_j x^{j-1} \quad (2)$$

式中 ρ_j 表示重量为 j 的行的重量在矩阵中所占的分量比, d_c 为矩阵中行重量的最大值.

由重量分布式(1)和(2)所约束的 LDPC 码集合中包含数量众多的不同具体码型, 码集合中的不同码型的总量会在数十万甚至百万计, 这些码型的性能是否相同? 下面将对这一问题进行分析.

3 密度进化理论和重量分布式约束的 LDPC 码集合性能上限^[6]

不失一般性, 考虑由式(1)和式(2)约束的 (d_v, d_c) 规则 LDPC 码, d_v 表示和信息节点连接的校验节点数目, d_c 表示和校验节点连接的信息节点数目. 在使用置信传播译码算法情况下, 信息节点和校验节点迭代地传递判决信息, 校验节点从与其连接的 $d_c - 1$ 个信息节点处接收信息, 将其处理后送给余下的另一个信息节点, 信息节点从与其连接的 $d_v - 1$ 个校验节点处接收信息, 将其处理后送给余下的另一个校验节点.

上述过程迭代 n 次后, 信息节点根据这 n 次迭代涉及的校验节点带来的信息对其约束的符号做出判决. 如果校验矩阵的 Girth 足够大, 则这些校验节点互不重复, 因此这些校验节点带给信息节点的信息是互相独立的. 如果编码分组足够长, 根据中心极限定理, 对于几乎所有随机构造的 LDPC 码和几乎所有的输入来说, 译码性能都逼近于理想状态下的译码性能, 即 n 次迭代后涉及的校验节点互不重复条件下的性能.

定义 $v = \log \frac{p(y/x=1)}{p(y/x=-1)}$ 为信息节点的输出消息, $u = \log \frac{p(y'/x'=1)}{p(y'/x'=-1)}$ 为校验节点的输出消息. 根据置信传播算法, 有

$$v = \sum_{i=0}^{d_v-1} u_i \quad (3)$$

$$\tanh \frac{u}{2} = \prod_{j=1}^{d_c-1} \tanh \frac{v_j}{2} \quad (4)$$

u_0 是译码器接输入消息的判决似然比, 服从 $N(2/\sigma^2, 4/\sigma^2)$ 的高斯分布. 使用 m_u 、 m_v 和 m_{u_0} 分别表示 u 、 v 和 u_0 的平均值. 那么, 经过 l 次迭代后对于一个度数为 i 的信息节点有:

$$m_{v,i}^{(l)} = m_{u_0} + (i-1)m_u^{(l-1)} \quad (5)$$

定义函数

$$\phi(x) = \begin{cases} 1 - \frac{1}{\sqrt{4\pi x}} \int_R \tanh \frac{u}{2} e^{-\frac{(u-x)^2}{4x}} du \\ 1 \end{cases} \quad (6)$$

在第 l 次迭代中, 传递到校验节点的信息服从概率密度分布:

$$f_v^{(l)} = \sum_{i=2}^{d_v} \lambda_i N(m_{v,i}^{(l)}, 2m_{v,i}^{(l)}) \quad (7)$$

通过式(7)可以得到:

$$E\left[\tanh \frac{v^{(l)}}{2}\right] = 1 - \sum_{i=2}^{d_v} \lambda_i \phi(m_{v,i}^{(l)}) \quad (8)$$

在第 l 次迭代中, 度数为 j 的校验节点的输出为:

$$m_{u,j}^{(l)} = \phi^{-1}(1 - [1 - \sum_{i=2}^{d_v} \lambda_i \phi(m_{v,i}^{(l)})]^{j-1}) \quad (9)$$

通过线性相加度数为 $2, 3, \dots, d_r$ 的校验节点, 可以得到

$$m_u^{(l)} = \sum_{j=2}^{d_c} \rho_j \phi^{-1}(1 - [1 - \sum_{i=2}^{d_v} \lambda_i \phi(m_{u_0} + (i-1)m_u^{(l-1)})]^{j-1}) \quad (10)$$

至此, 我们可以得到由式(1)和式(2)约束的 LDPC 码的性能上限的定义: $s^* = m_{u_0} = \inf(\mathfrak{R}^+)$, \mathfrak{R}^+ 满足 $m_u^{(l)} \rightarrow \infty$ 当 $l \rightarrow \infty$, 即 LDPC 码能容忍的噪声方差上限为 $2/s^*$. 也就是说, 由 $\lambda(x)$ 和 $\rho(x)$ 约束的全部 LDPC 码中, 性能最好的码可以容忍的噪声方差上限为 $2/s^*$, 当方差低于这个门限时, 只要迭代次数足够多, 误码率可以降低到任意低的数值.

4 Fill-Shift 构造法和码集合性能下限

本文中我们给出了一种构造由重量分布式约束的 LDPC 码的算法, 该算法给出了 LDPC 码纠错性能的下限, 虽然我们无法给出解析的数学证明, 但是可以通过直观的分析来说明.

4.1 Fill-Shift 构造 LDPC 码算法

Fill-Shift 算法思想是首先按照行重量分布式约束条件, 按照行重依次增大的顺序对校验矩阵的每一行从左到右进行填充, 这个过程结束后, 则行重量分布式的要求已经满足, 但是却不符列重量分布式的要求. 对于一个由重量分布如公式(1)、(2)约束的 $M \times N$ 的校验矩阵, 行重量分布式的填充如下:

$$H_{i,j} = 0, 1 \leq i \leq M, 1 \leq j \leq N \quad (11)$$

当 $i = 2, 3, \dots, d_c$ 时,

$$H_{x,y} = 1, M(\rho_2 + \dots + \rho_{i-1}) + 1 \leq x \leq M(\rho_2 + \dots + \rho_{i-1} + \rho_i) \quad (12)$$

不失一般性, 下面我们看第 i 列的移动过程, $\forall 1 \leq i \leq N$, $\exists j$, 满足

$$N(\lambda_2 + \lambda_3 + \cdots + \lambda_{j-1}) < i \leq N(\lambda_2 + \lambda_3 + \cdots + \lambda_{j-1} + \lambda_j) \quad (13)$$

则该列重量为 j ; 算法思想是保持该列的后 j 个‘1’位置不变, 对于该列其他的‘1’则将其移动到它对应行的右首第一个‘0’的位置, 这样第 i 列就会满足列重量分布式的约束条件, 第 i 列的算法流程如图 1 所示。

```

z = M;
do while(  $H_{M,i} + H_{M-1,i} + \cdots + H_{z,i} < j$  ) { z - - ; }
for l = 1 to z - 1 do {
  if  $H_{l,i} = 1$  {
     $H_{l,i} = 0$ ;
    k = i + 1;
    do while  $H_{l,k} \neq 1$  {
       $H_{l,k} = 1$ ;
      k + + ;
    }
    break; } }

```

图 1 Fill-Shift 算法流程

需要指出的是, 算法最后需要根据针对特定的重量分布式做局部微小调整, 如图 4 的右下角不规则部分所示。可以看出, 上述 Fill-Shift 算法构造的 LDPC 码校验矩阵是 (1)、(2) 约束的 LDPC 码集合的子集, 它的纠错性能也反应了重量分布式约束的 LDPC 码的特点。和高斯逼近密度进化理论给出的性能上限不同, Fill-Shift 构造方法给出了 LDPC 码集合的性能下限。虽然构造出纠错性能低的码在实际应用中没有使用价值, 但是对码的构造有指导意义。

4.2 Fill-Shift 构造算法的码纠错性能

这里我们把 Fill-Shift 构造算法的码性能和 PEG 方法构造的码性能进行比较, 可以得到比较明显的结果。从图 2 可以看出, Fill-Shift 算法给出的 3 种码性能都很低, 远低于 PEG 方法的纠错性能, 并且也低于一般随机码的性能, 这里没有给出例子; 此外, 还发生了随着信噪比的上升误码率不仅没有降低, 而且出现了上升的趋势, 并且误码率始终处于一个比较高的水平上, 这种现象是很少见的。但是, 这种 LDPC 码和 PEG 算法构造

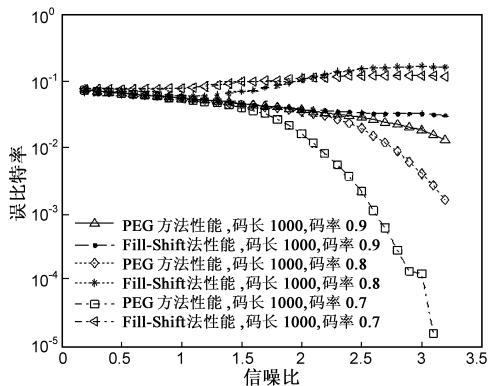


图 2 码长为 1000 的性能比较图

的 LDPC 码却同属于同一对重量分布式所约束的码集合, 这种现象可以借助影响 LDPC 码性能的一个重要参数-环路来分析。

Fill-Shift 方法在构造 LDPC 码的过程中没有考虑环路这一重要参数, 相反, 这一构造方法产生了最大可能的环路数目, 如图 4 所示, 而且这些环路都较短, 长度为 4 的占了很大的比重, 当然, 这些数目很难用有效的数学手段来计算, 但我们可以分别观察两个列填充、行移动的校验矩阵的图, 为了观察的方便我们这里进行的是列填充 (Fill) 行移动 (Shift), 而不是 4.1 所述的行填充列移动, 但二者的原理是完全相同的, 对于一个由

$$\lambda(x) = 0.3x^2 + 0.25x^3 + 0.25x^4 + 0.15x^7 \quad (14)$$

$$\rho(x) = 0.54x^4 + 0.46x^5 \quad (15)$$

约束的 50×60 的校验矩阵的列填充图、行移动分别如图 3 和图 4 所示, 二者中的黑点表示‘1’元素所在的位置。

根据 LDPC 码环路的定义, 该列填充后的校验矩阵有最大可能数目的长度为 4 的环路, 处于任意两行 (列) 相同位置的 4 个‘1’元素就构成一个长度为 4 的环路, 不仅如此, 还有大量的长度为 6、8 等的环路, 还没有找到有效的数学手段来证明它的长度为 4 的环路是最多的, 但是从直观上和 Girth 的定义可以理解这一点。如图 4 所示, 在行移动后其环路会大为减少, 但其环路数仍然保持非常大的

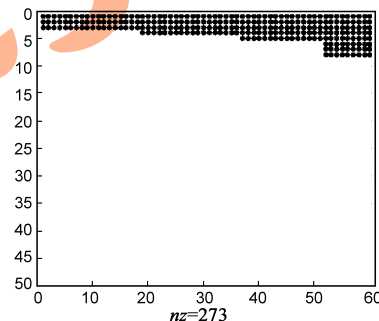


图 3 校验矩阵列填充图

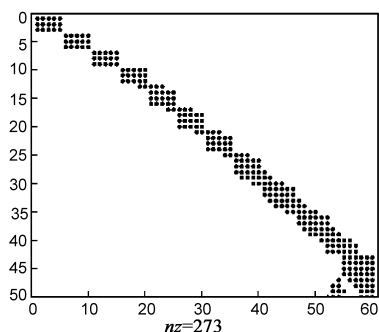


图 4 校验矩阵行移动图

的数量, 不仅如此, 如图示前 3 个校验式参与校验了前 5 个信息位, 而且 3 个校验式是相关的, 而且这 5 个信息位只有这 3 个校验式来约束。而 3 节中所述的高斯逼近密度进化理论中假设不存在环路, 图 2 中所涉及的 PEG 方法中环路实现了局部最大化, 因此 Fill-Shift 算法构造的 LDPC 码的纠错性能最差也就得到了合理的解释。

5 校验矩阵不变性及其应用

对重量分布式和约束的校验矩阵做初等变换得到

的具体 LDPC 码型数量极多,最多会有 $N!$ 个,我们在实际应用中,需要的是一个具体的好码,而不是一个码集合.的确,具体分析如此众多的码型纠错性能的差别,其工作量是非常巨大的,我们这里分析其中的几个有代表性的码型,来推测比较这些码型的性能.下面我们选择 4 种重量分布有代表性的码型,分别是:文献[7]定义的原始码、随机列变换码、列重量递增排列码和下三角最大化码.

5.1 校验矩阵典型初等变换

变换后的校验矩阵在结构上和原始码相比有了很大的变化,但是它们的重量分布式相同:

$$\lambda(x) = 0.175x + 0.45x^2 + 0.375x^5 \quad (16)$$

$$\rho(x) = x^9 \quad (17)$$

因此它们服从于文献[4]提出的命题.

5.2 校验矩阵的性能比较

图 5 给出了 $3a/4$ 码率 LDPC 码 3 种典型初等变换的性能比较,仿真采用 AWGN 信道,假定有理想的信道估计、定时,无码间串扰.可以看出,包括原始码在内的共 4 种码型有几乎完全相同的纠错性能,其他 18 种码率的 LDPC 码也有类似的结果,考虑到噪声的不完全随机性和试验数据量的有限性,我们可以认为,校验矩阵初等变换后得到的不同 LDPC 码型有相同的纠错性能,这与我们在试验之前的设想是一致的.

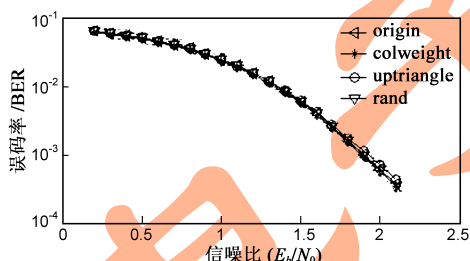


图 5 不同重量分布的 LDPC 码性能

上述结论可以从 LDPC 码的纠错原理得到解释^[10].这可以给我们带来有益的启示.首先,我们可以对校验矩阵作必要的变换,比如把校验矩阵变为下三角结构,这种变换可以降低编码复杂度并能实现系统编码而没有牺牲码的性能^[9].此外,我们可以根据列重不均匀的校验矩阵列对应的信息符号差错概率不同这一点,把校验矩阵进行初等变换,使重要的信息符号对应重量较大的列,以便有更多的校验式对其进行约束,从而使这些重要信息符号保持较低的差错概率^[9],比如卫星传输过程中的一些控制、状态信息等.

6 重量分布式约束的 LDPC 码的性能分类集合

由上面的 3、4 和 5 节可以看出,在同一个重量分布式约束的 LDPC 码集内,不同的码型之间的性能有可能相同,也可能差别很大.性能相同的码型处于一个校验

矩阵初等变换得到的集合中,性能不同的码型则处于不同的校验矩阵集合中.我们可以认为,由重量分布式约束的 LDPC 码依据性能可以划分成不同的很多个集合,这些集合间的性能是不同的,决定性能高低的因素就是环路分布和环路长度,而在同一个集合内部的不同码型之间性能是相同的.因此在给定重量分布式的条件下,寻找纠错性能高的 LDPC 码集合,同时,对已得到的码进行合理的矩阵初等变换,不仅可以实现系统编码和提高对重要信息位的保护,还可以减少耗费资源.

参考文献:

- [1] Gallager R G. Low-density parity-check codes[J]. IRE Transactions on Information Theory, 1962, 8(1): 21–28.
- [2] D J C MacKay, R M Neal. Near Shannon limit performance of low density parity check codes[J]. Electron Lett, 1996, 32(18): 1645–1646.
- [3] S Y Chung, G D Forney Jr, T J Richardson, R Urbanke. On the design of low-density parity-check codes within 0.0045 dB of the Shannon limit[J]. IEEE Commun Lett, 2002, 5(2): 58–60.
- [4] 袁东风,等.宽带移动通信中的先进信道编码技术[M].北京:北京邮电大学出版社,2004.68–70.
- [5] 殷柳国.低密度奇偶校验码编译码技术及应用研究[D].北京:清华大学,2005.
Yin Liu-guo. Coding Technique and Application Studies on Low-Density Parity-Check codes[D]. Beijing: Tsinghua University, 2005. (in Chinese)
- [6] Chung S-Y, Richardson TJ, Urbanke R. L. Analysis of sum-product decoding of low-density parity-check codes using a Gaussian approximation[J]. IEEE Transactions on Information Theory, 2001, 47(2): 657–670.
- [7] IEEE STD 802.16e, Air Interface for Fixed and Mobile Broadband Wireless Access Systems[S].
- [8] 姚春光,等.一种基于均匀环路的 LDPC 码构造方法[J].信号处理,2007,23(2):288–292.
Yao Chun-guang, et al. One method for LDPC codes construction based on average-cycle[J]. Signal Processing, 2007, 23(2): 288–292(in Chinese)
- [9] T Richardson, R L Urbanke. Efficient encoding of low-density parity-check codes[J]. IEEE Trans Information Theory, 2001, 47(2): 638–656.
- [10] 马丕明,等.非规则 LDPC 码的不等错误保护性能研究[J].通信学报,2005,26(11):132–136.
Ma Pi-ming, et al. Research on unequal error protection of irregular LDPC codes[J]. Journal on Communications, 2005, 26(11): 132–136. (in Chinese)
- [11] ETSI EN 302 307-2005, Digital Video Broadcasting (DVB)

[S].

[12] Mao Yong-Yi, Banihashemi A H. A heuristic search for good low-density parity-check codes at short block lengths [A]. IEEE International Conference on Communications [C]. Helsinki, Finland: IEEE, 2001. 41 – 44.

[13] Hu Xiao-yu, Eleftheriou E, Arnold D-M. Progressive edge-growth tanner graphs [A]. IEEE Global Telecommunications Conference [C]. San Antonio, TX, USA: IEEE, 2001. 995 – 1001.

[14] J Campello, D S Modha, S Rajagopalan. Designing LDPC codes using bit-filling [A]. IEEE International Conference on Communications [C]. Helsinki, Finland: IEEE, 2001. 55 – 59.

[15] J Campello, D S Modha. Extended bit-filling and LDPC code design [A]. IEEE Global Telecommunications Conference [C]. San Antonio, TX, USA: IEEE, 2001. 985 – 989.

作者简介:



姚春光 男, 1975 年 9 月生于山东苍山, 国防科技大学博士研究生. 研究方向为卫星通信系统、信道编码技术和信道估计等.
E-mail: ychg75@163.com



张 健 男, 1964 年生于河北衡水, 高级工程师. 主要研究兴趣为卫星通信、光通信和短波通信.

葛 新 女, 1972 年生, 硕士, 工程师. 主要研究方向为卫星通信、无线电频谱管理等.