

代 号 10701

学 号 0108120476

分类号 TN911.22

密 级 公开

题 (中、英文) 目 LDPC 码的编译码原理及编码设计

Principles and Code-Design of LDPC Codes

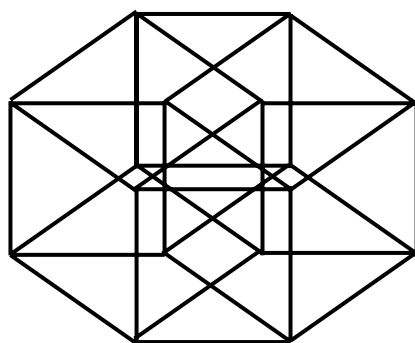
作 者 姓 名 王 鹏 指导教师姓名、职务 王新梅 教授

学 科 门 类 工学 学科、专业 通信与信息系统

提交论文日期 二〇〇四年一月

西安电子科技大学硕士学位论文

LDPC 码的编译码原理及编码设计



作者：王鹏

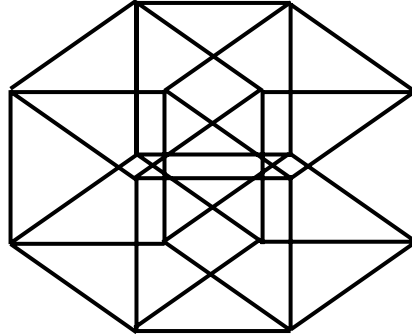
导师：王新梅 教授

学科：通信与信息系统

二〇〇三年十二月

中国 西安

Principles and Code-Design of LDPC Codes



A Dissertation

Presented to XIDIAN University

In candidacy for the Degree of

Master of Engineering

In

Communication and Information System

By

Wang Peng

Xi'an, People's Republic of China

December, 2003

谨 以 此 文

献 给 我 的 外 婆

创新性声明

本人声明所呈交的论文是我个人在导师指导下进行的研究工作及取得的研究成果。尽我所知，除了文中特别加以标注和致谢中所罗列的内容以外，论文中不包含其他人已经发表或撰写过的研究成果；也不包含为获得西安电子科技大学或其它教育机构的学位或证书而使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示了谢意。

本人签名：_____ 日期 _____

关于论文使用授权的说明

本人完全了解西安电子科技大学有关保留和使用学位论文的规定，即：研究生在校攻读学位期间论文工作的知识产权单位属西安电子科技大学。学校有权保留送交论文的复印件，允许查阅和借阅论文；学校可以公布论文的全部或部分内容，可以允许采用影印、缩印或其它复制手段保存论文。

本人签名：_____ 日期 _____

导师签名：_____ 日期 _____

摘 要

低密度校验码以其低复杂度的迭代译码算法和可逼近信道容量限而成为目前最佳的编码技术之一，越来越受到众多编码研究学者的关注。本文在对低密度校验码现有理论的研究基础上，系统地分析了低密度校验码在删除信道下的纠错性能和度序列设计、低密度校验码的围长设计和快速编码设计等编码设计问题，获得了一些研究成果，主要概括为：

1. 系统地阐述了低密度校验码基于图模型的编译码思想，介绍了密度进化理论，对影响低密度校验码纠错性能的两个主要因素——度序列设计和围长设计进行了深入分析；
2. 阐述了应用于删除信道下的纠删码基本原理，介绍了两类标准的 RS 码类纠删码，重点分析了具有线性时间编码和恢复算法的渐近好码—级联型低密度纠删码，分析了正则度分布的阈值，对正则低密度校验码在删除信道下的纠错性能进行了仿真，从理论上证明了基于 $(d, 2d)$ -正则度序列的低密度纠删码都不是渐近最优码 ($d \geq 3$)，同时还分析了非正则低密度校验码的度序列设计，基于右边正则序列提出了一种改进型右边正则序列，证明了此序列为渐近似最优的，对基于几类现有典型度分布序列的级联型低密度纠删码进行了模拟仿真及性能分析；
3. 研究了现有的具有较大围长的低密度校验码设计方法，提出了一种新的构造具有较大围长的正则低密度校验码方法并对其在高斯信道下的纠错性能进行了仿真，提出了渐进边增长算法的改进算法，使采用改进后的算法构造的低密度校验码能够严格满足给定的度序列分布；
4. 对低密度校验码的快速编码问题进行了深入研究，指出了旋风码和重复累积码能够达到线性编码的原因及其与可快速编码的低密度校验码之间的关系，提出了两种可线性编码的低密度校验码的构造方法并对其在高斯信道下的纠错性能进行了仿真。

关键词： 低密度校验码 删除信道 图模型 度分布序列 围长 高斯信道 快速编码

ABSTRACT

Low-density parity-check codes come to be one of the best coding technologies because of their low-complexity iterative decoding algorithm and capacity approaching performance, and they attract more and more researchers' eyes in recent years. In this dissertation, the basic principles of the coding and decoding of LDPC codes are studied systematically, and some code-design problems such as the design of degree distribution sequences, the design of girths and the design of efficient encode-able LDPC codes are analyzed in detail. Based on all these efforts, some positive results are obtained and summarized as follow:

1. The coding and decoding ideas of low-density parity-check codes on graphs are systematically summarized, and the density evolution theory is introduced. The two leading factors on the performance of LDPC codes, i.e. the degree distribution sequences and the girths of these codes, are analyzed in detail;

2. The principles of Erasure codes used under Binary Erasure Channels are summarized and Erasure codes which belong to standard classes of RS codes are introduced with emphasis on cascaded low-density erasure codes with linear time encoding and erasure recover algorithms. Thresholds of regular degree distributions are analyzed. It is shown that low-density erasure codes based on $(d, 2d)$ -regular sequences of degree distribution are not close to optimal ($d \geq 3$). Two pairs of irregular degree distribution sequences are introduced and a pair of improved right regular sequences of low-density erasure codes are presented, It is testified that the new sequences are asymptotically quasi-optimal. In the meantime, simulations of cascaded low-density erasure codes based on a few types of special sequences of degree distribution available are given, together with performance analyses on these codes.

3. The available design methods of LDPC codes with large girth are introduced and a new construction of regular LDPC codes with large girth is brought along with its realization algorithm, and the performances of the LDPC codes generated by this method are analyzed and simulated under AWGN channels. Improved Progressive Edge-Growth algorithm is presented by which the LDPC codes generated can satisfy the given degree distribution strictly.

4. The efficient encoding problem of LDPC codes is discussed in detail, and the reasons that Tornado codes and repeat accumulate codes are linear encode-able and the relationships between them and efficient encode-able LDPC codes are presented. Two constructions of linear encode-able LDPC codes are brought up and their performances under AWGN channels are simulated.

Keywords: low density parity-check codes erasure channel graph model
degree distribution sequences girth AWGN channel
efficient encoding

目 录

第一章 绪论	1
1.1 数字通信系统与信道模型.....	1
1.1.1 数字通信系统.....	1
1.1.2 信道模型.....	2
1.2 纠错编码理论及其发展.....	3
1.3 低密度校验码的提出、发展和现状.....	5
1.4 本文主要研究工作及内容安排.....	6
第二章 LDPC 码的编译码原理	7
2.1 LDPC 码的定义及其 Tanner 图表示.....	7
2.1.1 LDPC 码的定义及其描述.....	7
2.1.2 LDPC 码的 Tanner 图表示及非正则 LDPC 码.....	8
2.2 LDPC 码的译码.....	9
2.2.1 LDPC 码的译码思想.....	9
2.2.2 BIAWGN 信道下的算法描述.....	12
2.2.3 BSC 信道下的算法描述.....	13
2.2.4 BEC 信道下的算法描述.....	14
2.3 LDPC 码的性能分析.....	15
2.3.1 LDPC 码的度序列设计及密度进化理论.....	15
2.3.2 LDPC 码的围长设计.....	17
2.4 本章小结.....	18
第三章 删除信道下的 LDPC 码	19
3.1 纠删码及其发展.....	19
3.1.1 删除信道和纠删码.....	19
3.1.2 RS 码类纠删码.....	21
3.1.3 低密度纠删码.....	21
3.2 删除信道下的密度进化理论.....	23
3.2.1 密度进化理论的直接描述.....	23
3.2.2 密度进化理论的微分方程描述.....	25
3.3 正则 LDPC 码在删除信道下的性能.....	27
3.3.1 正则 LDPC 码阈值的唯一存在性分析.....	25

3.3.2	一类正则LDPC码的性能分析.....	25
3.4	非正则 LDPC 码在删除信道下的性能.....	31
3.4.1	Heavy-Tail/Poisson度序列分布.....	31
3.4.2	右边正则度序列分布.....	32
3.5	基于改进型的右边正则度序列设计.....	34
3.5.1	改进型右边正则度序列设计.....	34
3.5.2	改进型右边正则度序列的性能分析.....	35
3.6	本章小结.....	38
第四章	LDPC 码的围长研究.....	39
4.1	现有的几种围长设计方法.....	39
4.1.1	一种基于RS码的代数构造方法.....	39
4.1.2	一种基于矩阵分裂的代数构造方法.....	40
4.1.3	一种启发式搜索方法—PEG 构造方法.....	41
4.2	一种具有较大围长的正则 LDPC 码构造方法.....	42
4.3	PEG 算法研究.....	47
4.4	本章小结.....	48
第五章	LDPC 码的快速编码研究.....	49
5.1	LDPC 码的快速编码.....	49
5.2	两类可快速编码的码类.....	51
5.2.1	Tornado 码及其编码.....	51
5.2.2	重复累积码及其编码.....	52
5.3	可线性编码的 LDPC 码构造.....	52
5.3.1	直接构造法.....	52
5.3.2	删除构造法.....	53
5.4	关于线性编码的几点设想.....	55
5.5	本章小结.....	56
结束语	57
致谢	59
参考文献	61
硕士期间完成的论文和科研工作	65

第一章 绪 论

本章首先简要介绍了数字通信系统及信道模型，回顾了信道编码理论与技术的发展历程，然后概述了低密度校验码的提出、发展和研究现状，最后总结了作者在攻读硕士学位期间的研究工作，给出了全文的内容安排。

1.1 数字通信系统与信道模型

1.1.1 数字通信系统

通信系统的基本目的在于将信息由信源高效、可靠、有时还需安全地传送到信宿。有扰通信信道中的噪声会不可避免地会对传输信息产生不同程度的干扰，从而可能降低通信可靠性。所以通信系统的核心问题就是在存在随机噪声的信道中如何克服干扰，减小信息传输的差错，同时又不降低信息传输的效率，即如何解决系统的有效性与可靠性之间的矛盾。一般地，通信系统的可靠性用误比特率（BER）来衡量，其有效性则用信息传输速率 R 比特/信道符号来衡量。早期的人们普遍认为^[1]：通信系统的可靠性与有效性之间是一对不可调和的矛盾，一方的改善总是以牺牲另一方为代价，并指出当功率受限时，在有扰通信信道上实现任意小错误概率的信息传输的唯一途径就是把信息传输速率降低至零。Shannon信息和编码理论的奠基性论文“通信的数学理论”于1948年发表之后^[2]，改变了这一观点。他首次阐明了在有扰信道上实现可靠通信的方法，指出实现有效而可靠地信息传输的途径就是通过编码。根据Shannon的信息理论，数字通信系统的基本组成如图1.1所示^{[3][4][5]}。

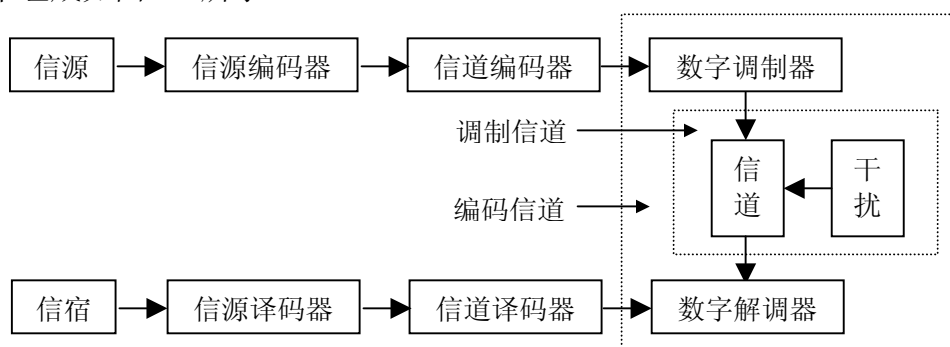


图 1.1 数字通信系统模型

Shannon的信息理论从通信系统的整体最佳化来研究信息的传输和处理。比特是一种通用的信息表示形式，它本身并不依赖于信源或信道特征。这就允许我们分别设计图1.1所示的两个阶段的信息处理，即信源编码和信道编码。Shannon不失最佳性地证明了这种分离性^{[2][3]}。

1.1.2 信道模型

图 1.1 中的信道部分只是信息传输所通过媒介的一种抽象，实际信道是多种多样的，如电缆、光缆、存储设备、甚至我们所处的实际空间及外太空等等。对于通信系统设计者来讲，了解系统中信道的特性是必需的。根据信道的输入输出的取值连续与否可以将其分为离散信道、连续信道和离散输入/连续输出信道；根据信道统计特性是否随时间改变可以将其分为平稳信道和非平稳信道；根据信道的输出之间是否具有相关性可将其分为记忆信道和无记忆信道；根据信道的特性对输入端是否具有对称性可以将其分为对称信道和非对称信道。实际应用中所涉及到的信道大多都是离散输入的平稳无记忆对称信道，下面给出几种常用的编码信道模型^{[4][5][6][7]}。

二进制对称信道 (BSC)：输入为二值变量 0、1，输出也为二值变量 0、1，且传输过程中发生错误（输入为 0 输出为 1 或输入为 1 输出为 0）的概率与输入无关；

二进制删除信道 (BEC)：输入为二值变量 0、1，输出或为输入的二值变量 0、1，或为删除 E ，且通常传输过程中不同输入被删除的概率相同；

二进制输入高斯信道 (BIAWGN)：输入为二值变量，输出为连续变量，且信道中的加性噪声为服从 $N(0, \sigma^2)$ 的高斯随机变量。

图 1.2 至图 1.4 给出了以上三种信道的模型图。

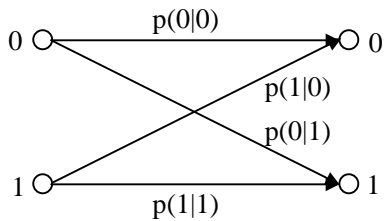


图 1.2 二进制对称信道 (BSC)

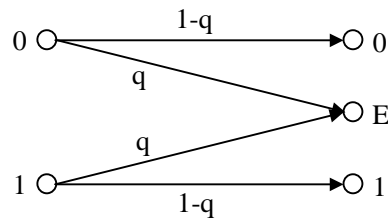


图 1.3 二进制删除信道 (BEC)

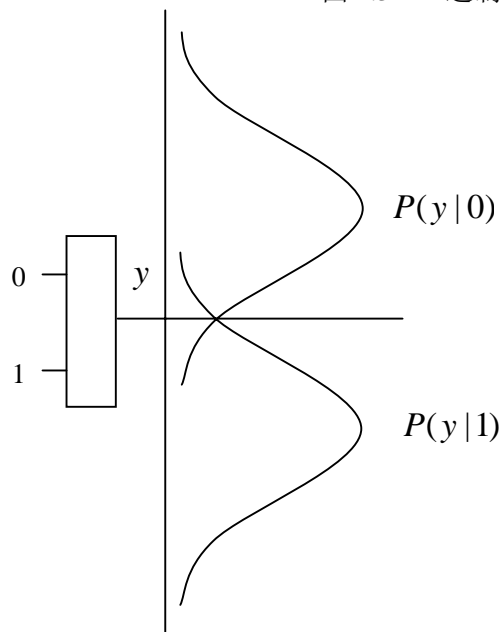


图 1.4 二进制输入高斯信道 (BIAWGN)

1.2 信道编码理论及其发展

图 1.1 中的信道编译码部分是以特定的控制手段, 引入适量冗余比特, 以克服信息在传输过程中受到的噪声和干扰影响。根据 Shannon 提出的信道编码定理, 对任意一个平稳离散无记忆有噪声信源, 都有一个固定的量, 称之为信道容量, 记做 C 。只要信息的传输速率低于信道容量, 就必然存在一种编码方法, 使得信息出现差错的概率随码长的增加趋于任意小; 反之, 当信息传输速率超过信道容量时, 则不存在这样的编码方法。这就是著名的信道编码定理, 它给出了特定信道上信息传输速率的上确界。

定理 1.1 (信道编码定理): 任意离散输入无记忆平稳信道存在信道容量 C , 对预期的任一数据速率 $R < C$ 和任一错误概率 $p > 0$, 有可能设计一对编译码器, 以保证该信道中速率为 R 的数据传输具有小于 p 的译码错误概率。

信道编码定理指出, 在有扰信道中, 只要信息传输速率小于信道容量, 就有可能实现任意可靠的信息传输。这个存在性定理提醒我们可以实现以接近信道容量的传输速率进行通信。但遗憾的是该定理采用的是非构造性的证明方法, 其中并没有给出逼近信道容量的码的具体编译码方法。

Shannon 在信道编码定理的证明中引用了三个基本条件, 即:

- (1)、采用随机的编码方式;
- (2)、码字长度趋近于无穷大;
- (3)、译码采用最佳的极大似然译码。

一般可将信道编译码器所使用的纠错码从性能上分为**坏码**和**好码**。所谓坏码是指只有将码率降至零才能使误码率为任意小的编码方式; 而好码又可以分为当误码率任意小时, 码率逼近信道容量限的**非常好码**和码率可达到的非零最大值小于信道容量限的**一般好码**。虽然 Shannon 指出一个随机选择的码以很高的概率为好码, 但随机码的极大似然译码的复杂度往往与码长呈指数关系, 即在误码率随码长趋于无穷而趋向于零的同时, 译码复杂度以指数增长, 而在实际应用中, 只能够使用以码长多项式的时间和空间复杂度内完成编译码的纠错码, 因而尽管一般的随机码是好码, 但不能看作是**实用码**。

自信道编码定理提出以来, 如何构造一个逼近信道容量限的**实用好码**成了众多研究学者竞相研究的课题, 并逐渐形成信息论的一个重要分支——信道编码理论。五十多年来, 人们构造实用好码的探索基本上是按照 Shannon 所引用的基本条件的后两条为主线展开的。虽然从理论上讲, 除了目前已知的码外, 几乎所有的码都是好码, 但到目前为止, 构造出真正意义上的实用好码还有较长的距离。虽然如此, 通过众多学者, 特别是有关数学和信息论学术界的研究人员五十多年的共同努力, 目前已经取得了很多成果。下面对其进行简要概述。

纠错码^{[5][8]}从构造方法上可分为分组码 (Block Codes) 和卷积码 (Convolutional Codes) 两大部分。在分组码方面, 第一个分组码是 1950 年发现的能纠正单个错误的

Hamming码；在整个 50 年代，基于代数理论又发现了多个短码长的分组码，如 1954 年 Golay 发现的 Golay 码以及 Reed 和 Muller 发现的 RM 码，Prange 在 1957 年发现的循环码等。最有意义的是 Bose 和 Ray-Chaudhuri 在 1960 年及 Hocuenghem 在 1959 年分别独立发现的能纠正多个错误的 BCH 码，以及 Reed 和 Solomon 在 1960 年发现的非二进制 RS 码。实际上，BCH 码可以看作是某个 RS 码的子域子码，而 RS 码又可以看作是 BCH 码的特例。其后发现的分组码主要有 1970 年的 Goppa 码和 1982 年发现的代数几何码。在所有这些分组码中，除了 Goppa 码和代数几何码中存在个别达到 GV 限的渐进好码外，其它均不是好码。分组码的译码主要采用基于代数的硬判决译码。

卷积码最早由 Elias 提出，早期被称为树码（Tree Codes），现在称为格图码（Trellis Codes）或卷积码。卷积码具有动态格图结构，可用有限状态机来描述其状态。由于缺乏有效的理论研究工具，对卷积码的有效研究成果不是很多，目前性能好的卷积码的构造方法主要借助于计算机搜索来获得。卷积码的译码一般采用概率译码，由于译码算法的简单、实用和易于实现，卷积码被广泛应用于实际中。

1966 年，Forney 将分组码和卷积码结合起来，提出了级联码（Concatenated Codes）的概念。级联码一般采用 RS 码作为外码，卷积码作为内码。Forney 的研究表明，级联码在性能得到较大改善的情况下，其译码复杂度并不显著增加。

根据对接收信号处理方式的不同，纠错码的译码可以分为硬判决译码和软判决译码。硬判决译码是基于传统纠错码观点的译码方法，解调器首先对信道输出值进行最佳硬判决，再将判决结果送入译码器，译码器根据解调器的判决结果，利用码字的代数结构来纠正其中的错误。而软判决译码则充分利用了信道输出的波形信息，解调器将匹配滤波器输出的一个实数值送入译码器，由于实数值包含了比硬判决更多的信道信息，译码器能够通过概率译码充分利用这些信息，从而获得比硬判决译码更大的编码增益。

总之，尽管随机码是理论上的好码，但由于其编码实现的困难性和无法承受的译码复杂度而只被用作理论分析的工具，在信道编码定理和后来的许多编码理论成果中，代数编码理论始终占据了主导地位，使得传统的信道编码技术受到临界速率（Critical Rate），也称做截止速率（Cutoff Rate） R_0 的限制^[4]。

1993 年 Turbo 码^{[9][10]}的提出被看作是信道编码理论研究的重要里程碑。Berrou 等人将卷积码和随机交织器相结合，同时采用软输出迭代译码来逼近最大似然译码，取得了超乎寻常的优异性能，并一举超越了截止速率，直接逼近 Shannon 提出的信道容量限。Turbo 码是一种信道编码理论界梦寐以求的可实用非常好码，它的出现标志着信道编码理论研究进入了一个崭新的阶段。Turbo 码成功的根本原因在于其实现方案中长码构造的伪随机性是核心，它通过随机交织器对信息序列的伪随机置换实现了随机编码的思想，从而为 Shannon 随机编码理论的应用研究奠定了基础。

随着 Turbo 码的深入研究，人们重新发现 Gallager 早在 1962 年提出的低密度校验码^{[11][12]}（Low Density Parity-Check Codes，简称 LDPC 码）也是一种具有渐进特性的非常好

码，它的译码性能同样可以逼近Shannon信道容量限^{[13][14][15][16]}。由于LDPC码具有在中长码长时超过Turbo码的性能，并且具有译码复杂度更低，能够并行译码及译码错误可检测等特点，成为目前信道编码理论的研究热点。研究表明，Turbo码只是LDPC码的一个特例^{[17][18][19]}，两者都是基于图构造的低密度码，译码算法具有等价性，从而使两者在基于图模型的编译码研究中得到了统一。

1.3 低密度校验码的提出、发展和现状

1962年，Gallager在他的博士论文中提出了二元正则LDPC码，也被称做Gallager码^[11]。Gallager证明了这类码具有很好的汉明距离特性，是满足GV限的渐进好码，在计算树上进行 $I \propto \log(\log(N))$ （ N 为码长）次后验概率迭代译码可以获得依码字长度指数降低的比特错误概率，但限于当时的计算能力，LDPC码被认为不是实用码，在很长一段时间内没有受到人们的重视。

1981年，Tanner在他的一篇奠基性的文章中正式提出了用图模型来描述码字的概念，从而将LDPC码的校验矩阵对应到被称为Tanner图^[20]的双向二部图上。采用Tanner图构造的LDPC码，通过并行译码可以显著地降低译码复杂度。Tanner还仔细分析了最小和算法（Min-Sum Algorithm）与和积算法（Sum-Product Algorithm）两种信息传递算法，证明了基于有限无环Tanner图的最小和译码算法与和积译码算法的最优性。但Tanner图在实际当中是采用随机图构造的，其中不可避免地存在小环路现象，这些小的环路会造成译码信息的重复传递，使译码过程中的消息之间不满足独立性假设，影响了迭代译码算法的收敛性。

Turbo码的发现重新引发了众多学者对LDPC码的研究兴趣。MacKay和Neal利用随机构造的Tanner图研究了LDPC码的性能，发现采用和积译码算法的正则LDPC码具有和Turbo码相似的译码性能，在长码时甚至超过了Turbo码^{[16][21]}，这一结果引起了信道编码界的极大关注。此后，Davey和MacKay从减少Tanner图上小环路的概念出发提出了基于 $GF(q)$, $q > 2$ 的LDPC码^{[22][23]}，进一步提高了LDPC码的译码性能。

在MacKay和Neal重新发现LDPC码优异性能的同时，Spielman和Sipser提出了基于二部扩展图的扩展码^{[24][25][26][27]}。在对扩展码的研究中，他们证明了一个随机构造的Tanner图以很大的概率为好的扩展图，而由好的扩展图构造的线性纠错码是渐进好码，从而证明了采用随机Tanner图构造的LDPC码以很大概率是渐进好码。Luby等人将采用非正则Tanner图构造的扩展图用于删除信道，称之为Tornado码^{[28][29]}。由于采用了非正则的Tanner图，Tornado码具有更大的扩展性和更好的收敛性，纠删性能更强。此后，采用优化度序列设计的非正则Tanner图被用于构造LDPC码，称为非正则LDPC码，与正则LDPC码相比，非正则LDPC码的性能得到显著的提高^{[30][31][32][33][34]}。

同时, Wiberg结合Turbo码和网格图的研究, 将Tanner图推广到包含隐含状态变量的因子图(Factor Graph) [19][22][35], 对Turbo码和LDPC码的研究在因子图的基础上得到统一。Wiberg对因子图的研究发现, 对任意给定系统, 无环图的状态复杂度是最大的, 有环图的状态复杂度则会大大降低, 从而证明了基于有环Tanner图的LDPC码具有较低的译码复杂度。Wiberg同时还证明了最小和算法和和积算法在本质上的同一性, 在格图译码中, 最小和算法退化为Viterbi译码算法, 和积算法退化为BCJR译码算法。

近两年, Richardson等人应用密度进化理论来测度LDPC码的性能[32][36]。Richardson等人在对LDPC码的研究中发现, 译码信息的迭代传递过程中存在着译码阈值现象, 即当信噪比大于译码阈值时, 迭代译码可使误码率趋于零, 反之无论采用多长的LDPC码, 经过多少次迭代译码, 总存在一定的错误概率。应用中心极限定理, Richardson等人证明了有限随机有环图的译码阈值可以逼近无环图的译码阈值。通过建立在无环图上的密度进化理论, 可以精确地计算无环图上LDPC码的译码阈值, 分析其译码收敛条件, 从而近似估算有环Tanner图上LDPC码的性能。研究表明, 译码阈值的大小与LDPC码的构造参数密切相关, 采用优化度序列设计的非正则LDPC码可以有效地改善阈值, 因此密度进化理论可以用于指导LDPC码的优化设计。

Chung等通过对密度进化理论的研究, 进一步提出了应用高斯逼近原理来简化译码阈值计算和收敛性分析, 从而使测度LDPC码性能模型由多参数动态系统的密度进化理论模型简化为单一参数动态系统的高斯逼近模型[30][33]。

1.4 本文主要研究工作和内容安排

作者结合国家和陕西省自然科学基金、“适合中国的4G及后3G中关键技术研究”(与三星公司合作)及国家自然科学基金委和香港科技局联合资助项目等科研课题, 采用理论分析和仿真相结合的方法, 对LDPC码的理论进行了深入研究, 取得了一些成果。全文共六章, 其余章节安排如下。

第二章系统地阐述了LDPC码基于Tanner图的编译码原理, 分析了LDPC码的各项参数指标对其性能的影响, 给出了合理的解释; 第三章分析了正则LDPC码在删除信道下的理论性能, 证明了 $(d, 2d)$ 正则码都不是渐进最优码, 介绍了几种非正则的度序列分布函数, 提出了一种改进型的右边正则度序列分布函数并对相应LDPC码在删除信道下的性能进行了仿真; 第四章介绍了围长对LDPC码性能的影响, 提出了一种构造具有较大围长的正则LDPC码设计方法; 第五章对LDPC码的快速编码实现进行了分析, 构造了两种可实现线性编码的LDPC码并对它们在高斯信道下的性能进行了仿真; 最后对全文的主要内容和成果进行了总结, 并指出了进一步研究的问题和需要做的工作。

第二章 LDPC 码的编译码原理

本章系统地概述了 LDPC 码的定义及其 Tanner 图表示, 基于图结构阐述了 LDPC 码的译码思想, 并给出了 LDPC 码在几种常见信道下的译码算法, 随后介绍了密度进化理论并对影响 LDPC 码性能的几个因素进行了分析, 指出了优化方向, 最后对本章进行了总结。

2.1 LDPC 码的定义及其 Tanner 图表示

2.1.1 LDPC 码的定义及其描述

一个码长为 n 、信息位个数为 k 的线性分组码可以由一个生成矩阵 $\mathbf{G}_{k \times n}$ 来定义, 信息序列 $\mathbf{s}_{1 \times k}$ 通过 \mathbf{G} 被映射到码字 $\mathbf{x} = \mathbf{s} \cdot \mathbf{G}$ 。线性分组码也可以由一个一致校验矩阵 $\mathbf{H}_{(n-k) \times n}$ 来等效描述, 所有码字均满足 $\mathbf{x} \cdot \mathbf{H}^T = \mathbf{0}$ 。校验矩阵的每一行表示一个校验约束 z_i , 其中所有非零元素对应的码元变量 x_j 构成一个校验集, 用一个校验方程表示; 校验矩阵的每一列表示一个码元变量参与的校验约束, 当列元素不为零时, 表示该码元变量参与了该行的校验约束。码元变量与校验方程之间的关系称为结构。下面主要对二元 LDPC 码进行讨论。

LDPC 码是一种线性分组码, 它的名字来源于其校验矩阵的稀疏性, 即校验矩阵中只有数量很少的元素为“1”, 大部分都是“0”。Gallager 最早给出了正则 LDPC 码的定义, 具体来讲正则 LDPC 码的校验矩阵 \mathbf{H} 满足下面三个条件:

- (1)、 \mathbf{H} 的每行有 ρ 个“1”;
- (2)、 \mathbf{H} 的每列有 λ 个“1”, $\lambda \geq 3$;
- (3)、与码长和 \mathbf{H} 矩阵的行数相比, ρ 和 λ 都很小。

1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1

1	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0	0	0	0
0	1	0	0	0	1	0	0	0	1	0	0	0	0	0	0	1	0	0
0	0	1	0	0	0	1	0	0	0	0	0	0	1	0	0	0	1	0
0	0	0	1	0	0	0	0	0	0	1	0	0	0	1	0	0	0	1
0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	1	0	0

1	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	1	0
0	1	0	0	0	0	1	0	0	0	1	0	0	0	0	1	0	0	0
0	0	1	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	1
0	0	0	1	0	0	0	0	1	0	0	0	0	1	0	0	1	0	0
0	0	0	0	1	0	0	0	0	1	0	0	0	0	1	0	0	0	0
0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	1	0	0	1

图 2.1 (20,3,4)LDPC 码的校验矩阵

矩阵 \mathbf{H} 的每列各自包含 λ 个“1”，表示每个码元变量受到相同数目的校验约束；每行也各自包含 ρ 个“1”，表示每个校验方程对相同数目的码元变量进行校验约束；由于 ρ 和 λ 都很小，校验矩阵 \mathbf{H} 具有很低的“密度”，因此由校验矩阵 \mathbf{H} 所确定的码称为 LDPC 码。Gallager 证明了当 $\lambda \geq 3$ 时，这类码具有很好的汉明距离特性^{[11][22]}。正则 LDPC 码通常用 (n, λ, ρ) 来表示，其中 n 为码长， λ 和 ρ 的含义不变，图 2.1 给出了一个 $(20, 3, 4)$ 正则 LDPC 码的校验矩阵。

当校验矩阵 \mathbf{H} 各列（行）中“1”的个数不全相同时，就得到了非正则 LDPC 码^{[28][29][31][34][37][38][39][40][41]}。非正则 LDPC 码通常用度序列分布函数来表示，我们会在给出 LDPC 码的 Tanner 图描述之后具体介绍非正则 LDPC 码的度序列表示。

2.1.2 LDPC 码的 Tanner 图表示及非正则 LDPC 码

设一个 (n, λ, ρ) 正则 LDPC 码 C 具有校验矩阵 $H = (h_{i,j})_{m \times n}$ ，则其相应的 Tanner 图模型可以表示为一个二部图。其中码字向量 $\mathbf{x} = (x_1, x_2, \dots, x_n)$ 表示为一组变量节点 $\{x_j, j = 1, 2, \dots, n\}$ ，分别对应于校验矩阵的各列，而校验约束则表示为一组校验节点 $\{z_i, i = 1, 2, \dots, m\}$ ，对应于校验矩阵的各行。仅当 $h_{i,j} = 1$ 时，变量节点 x_j 与校验节点 z_i 之间有一条边相连，节点 x_j 与 z_i 之间互称邻接节点，其间的连接边称为两个节点的邻接边。对于 (n, λ, ρ) 正则 LDPC 码，每个变量节点与 λ 个校验节点相连，称该变量节点的度为 λ ；每个校验节点与 ρ 个变量节点相连，称该校验节点的度为 ρ ，度表示与节点相连的边的数目。图 2.2 给出了图 2.1 所示的校验矩阵对应的 Tanner 图结构。

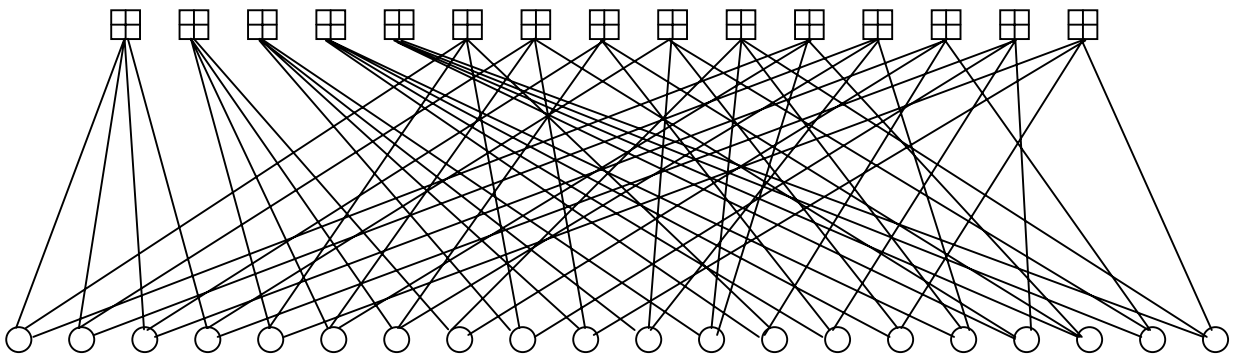


图 2.2 (20,3,4)LDPC 码的 Tanner 图表示

对于非正则 LDPC 码，相应的 Tanner 图中各变量节点或校验节点的度并不相同，通常用序列 $\{\lambda_1, \lambda_2, \dots, \lambda_{d_v}\}$ 和 $\{\rho_1, \rho_2, \dots, \rho_{d_c}\}$ 来表示其中边的度分布，其中 λ_j 表示与度为 j 的变量节点相连的边占总边数的比率， ρ_i 表示与度为 i 的校验节点相连的边占总边数的比率， d_v 和 d_c 分别表示变量节点和校验节点的最大度数。显然应有 $\sum_{j=1}^{d_v} \lambda_j = 1$ 及 $\sum_{i=1}^{d_c} \rho_i = 1$ ，即部分之和等于全部。这里之所以用边的度分布，而不用节点的度分布来

描述 LDPC 码是因为 LDPC 码的译码采用的是基于置信传播 (Belief Propagation) 的软输出迭代译码算法 (Soft Output Iterative Decoding Algorithms), 如和积译码算法 (Sum-Product Algorithm), 在译码过程中, 信息的传递是在边上进行的, 采用边的分布来描述 LDPC 码有助于分析其在给定译码算法下的实际性能和理论性能的上下界。

边的度分布序列可以用多项式来表示, 即:

$$\lambda(x) = \sum_{j=1}^{d_l} \lambda_j x^{j-1} \quad (2-1)$$

$$\rho(x) = \sum_{i=1}^{d_r} \rho_i x^{i-1} \quad (2-2)$$

满足 $\lambda(1) = \sum_{j=1}^{d_l} \lambda_j = 1$ 及 $\rho(1) = \sum_{i=1}^{d_r} \rho_i = 1$ 。

正则 LDPC 码可以看作是非正则 LDPC 码的特例, 例如对于 $(n, 3, 4)$ 正则 LDPC 码, 相应边的度分布多项式分别退化为 $\lambda(x) = x^2$ 和 $\rho(x) = x^3$ 。

设一个 LDPC 码对应的 Tanner 图中边的总数为 E , 根据边的度分布多项式可以得到度为 j 的变量节点个数为 $v_j = E\lambda_j / j$, 度为 i 的校验节点个数为 $c_i = E\rho_i / i$, 则变量节点和校验节点的总数分别为:

$$n = \sum_{j=1}^{d_l} v_j = \sum_{j=1}^{d_l} E\lambda_j / j = E \int_0^1 \lambda(x) dx \quad (2-3)$$

$$m = \sum_{i=1}^{d_r} u_i = \sum_{i=1}^{d_r} E\rho_i / i = E \int_0^1 \rho(x) dx \quad (2-4)$$

当校验矩阵满秩时, 通过度分布多项式 $\lambda(x)$ 和 $\rho(x)$ 构造的非正则 LDPC 码的码率为:

$$R(\lambda, \rho) = \frac{n-m}{n} = 1 - \frac{\int_0^1 \rho(x) dx}{\int_0^1 \lambda(x) dx} \quad (2-5)$$

对于校验矩阵非满秩的情况, 实际的码率要比 $R(\lambda, \rho)$ 略高一些。

2.2 LDPC 码的译码

2.2.1 LDPC 码的译码思想

下面介绍 LDPC 码通用的一类译码算法, 即所谓的消息传递算法 (Message Passing Algorithms) 消息传递算法是一种迭代译码算法 (Iterative Algorithms) [16][17][19][20][22][36][42][43][44], 它的名字来源于其运行机制, 在该算法的每一轮迭代过程中, 关于各个节点的置信消息需要在变量节点和校验节点之间传递。例如由变量节点向校验节点传递的消息是基于变量节点对应的码元变量经过信道后的观察值和由邻接的校验节

点在上一次迭代过程中传递过来的消息联合计算的，其中需要特别注意的是由某个变量节点 v 向校验节点 c 所传递消息的计算中不包含在上一次迭代中由校验节点 c 传递给变量节点 v 的消息，对由校验节点向变量节点传递的消息也有同样情况。

一类比较重要的消息传递算法称做置信传播算法（Belief Propagation Algorithm），该算法在 Gallager 的博士论文中有具体的描述，也经常在人工智能（Artificial Intelligence）等领域内使用。在置信传播算法中，各个节点之间传递的信息是概率或置信信息，比如由变量节点 v 传递给校验节点 c 的信息是 v 取某些特定值的概率信息，该信息的具体取值依赖于 v 的观测值和其它所有与 v 相连的校验节点（除 c 以外）在上一轮迭代中传递给 v 的置信信息，同样由 c 传递给 v 的信息也是 v 取某些特定值的概率信息，该信息的具体取值依赖于其它所有与 c 相连的校验节点（除 v 以外）在上一轮迭代中传递给 c 的置信信息。

置信传播算法的迭代公式是在独立性假设（Independence Assumption）下推导得到的，这里仅考虑码元取自 $GF(2)$ 的情况。设某个二值随机变量 x 取值为 0、1 的概率分别为 $P_r(x=0)$ 和 $P_r(x=1)$ ，为了使所传递的消息中能够同时包含 x 的两个取值概率信息，通常两者的一个函数变量来表示该信息，如概率差 $\Delta_x = P_r(x=0) - P_r(x=1)$ 、概率比 $L(x) = P_r(x=0)/P_r(x=1)$ 或者对数似然比 $LLR(x) = \ln[P_r(x=0)/P_r(x=1)]$ ，这些函数变量称做量度（Metric）。可以证明，不同量度下推导得到的迭代公式是等价的，这里仅给出对数似然比量度下的公式推导^{[45][46][47]}。

设 x 为满足均匀分布的二值随机变量，即 $P_r(x=0) = P_r(x=1) = 1/2$ ， y 为 x 经过信道后的观测值，根据 Bayes 规则有：

$$\begin{aligned} LLR(x|y) &= \ln \frac{P(x=0|y)}{P(x=1|y)} = \ln \frac{P(x=0,y)/P(y)}{P(x=1,y)/P(y)} \\ &= \ln \frac{P(y|x=0)P(x=0)}{P(y|x=1)P(x=1)} = \ln \frac{P(y|x=0)}{P(y|x=1)} \\ &= LLR(y|x) \end{aligned} \quad (2-6)$$

根据独立性假设，若 x_1, x_2, \dots, x_d 为互相独立的二值随机变量， y_1, \dots, y_d 是 x_1, x_2, \dots, x_d 经过无记忆信道后的观测值，则有：

$$LLR(x_1, x_2, \dots, x_d | y_1, \dots, y_d) = \sum_{i=1}^d LLR(x_i | y_i) \quad (2-7)$$

考虑两个二值随机变量 x_1 和 x_2 ， y_1 和 y_2 为 x_1 、 x_2 经过信道后的观测值，则：

$$\begin{aligned} LLR(x_1 \oplus x_2 | y_1, y_2) &= \ln \frac{P(x_1 \oplus x_2 = 0 | y_1, y_2)}{P(x_1 \oplus x_2 = 1 | y_1, y_2)} \\ &= \ln \frac{P(x_1 = 0, x_2 = 0 | y_1, y_2) + P(x_1 = 1, x_2 = 1 | y_1, y_2)}{P(x_1 = 0, x_2 = 1 | y_1, y_2) + P(x_1 = 1, x_2 = 0 | y_1, y_2)} \\ &= \ln \frac{P(x_1 = 0 | y_1)P(x_2 = 0 | y_2) + P(x_1 = 1 | y_1)P(x_2 = 1 | y_2)}{P(x_1 = 0 | y_1)P(x_2 = 1 | y_2) + P(x_1 = 1 | y_1)P(x_2 = 0 | y_2)} \end{aligned}$$

$$\begin{aligned}
&= \ln \frac{1 + e^{LLR(x_1|y_1) + LLR(x_2|y_2)}}{e^{LLR(x_1|y_1)} + e^{LLR(x_2|y_2)}} \\
&= \ln \frac{(e^{LLR(x_1|y_1)} + 1)(e^{LLR(x_2|y_2)} + 1) + (e^{LLR(x_1|y_1)} - 1)(e^{LLR(x_2|y_2)} - 1)}{(e^{LLR(x_1|y_1)} + 1)(e^{LLR(x_2|y_2)} + 1) - (e^{LLR(x_1|y_1)} - 1)(e^{LLR(x_2|y_2)} - 1)} \\
&= \ln \frac{1 + \frac{e^{LLR(x_1|y_1)} - 1}{e^{LLR(x_1|y_1)} + 1} \cdot \frac{e^{LLR(x_2|y_2)} - 1}{e^{LLR(x_2|y_2)} + 1}}{1 - \frac{e^{LLR(x_1|y_1)} - 1}{e^{LLR(x_1|y_1)} + 1} \cdot \frac{e^{LLR(x_2|y_2)} - 1}{e^{LLR(x_2|y_2)} + 1}} \quad (2-8) \\
&= \ln \frac{1 + \tanh(e^{LLR(x_1|y_1)}/2) \cdot \tanh(e^{LLR(x_2|y_2)}/2)}{1 - \tanh(e^{LLR(x_1|y_1)}/2) \cdot \tanh(e^{LLR(x_2|y_2)}/2)}
\end{aligned}$$

现在考虑多个二值随机变量 x_1, x_2, \dots, x_k 的情况，设 y_1, y_2, \dots, y_k 为 x_1, x_2, \dots, x_k 经过信道后得到的观测值，则将(2-8)式推广即可得到：

$$LLR(x_1 \oplus x_2 \oplus \dots \oplus x_k | y_1, y_2, \dots, y_k) = \ln \frac{1 + \prod_{i=1}^k \tanh(LLR(x_i | y_i)/2)}{1 - \prod_{i=1}^k \tanh(LLR(x_i | y_i)/2)} \quad (2-9)$$

利用(2-7)式和(2-9)式便可推导出置信传播算法的迭代公式。设发端为二进制等概信源，即每个发送符号取 0 和 1 的概率均为 1/2，发送码字为 $\mathbf{x} = \{x_1, x_2, \dots, x_n\} \in GF^n(2)$ ，经过信道传输之后在接收端得到的码序列为 $\mathbf{y} = \{y_1, y_2, \dots, y_n\} \in F_y^n$ ，对不同的信道输出，集合 F_y 的含义有所不同。首先，根据信道观测值 \mathbf{y} 计算每个码元变量的后验概率信息，即 $f_j = LLR(x_j | y_j)$ ；在之后的每一轮迭代中，每个校验节点 i 根据(2-9)式计算出相应的对数似然信息 Q_{ij}^l 并传递给相邻的变量节点 j ；每个变量节点 j 也根据(2-7)式计算出相应的信息 R_{ij}^l 并传递给相邻的校验节点 i ，其中 l 为迭代次数。

通常，用 $M(j)$ 表示与变量节点 j 相连的所有校验节点所构成的集合， $M(j) \setminus i$ 表示 $M(j)$ 中除去其中的校验节点 i 后剩下的集合；用 $N(i)$ 表示与校验节点 i 相连的所有变量节点构成的集合， $N(i) \setminus j$ 表示 $N(i)$ 中除去其中的变量节点 j 后剩下的集合，则相应的迭代公式可表示为：

$$Q_{ij}^l = \begin{cases} f_j & l = 0, \\ f_j + \sum_{k \in M(j) \setminus i} R_{kj}^{l-1} & l > 0, \end{cases} \quad (2-10)$$

$$R_{ij}^l = \ln \frac{1 + \prod_{k \in N(i) \setminus j} \tanh(Q_{ik}^l / 2)}{1 - \prod_{k \in N(i) \setminus j} \tanh(Q_{ik}^l / 2)} \quad (2-11)$$

注意到(2-11)式还可以写做：

$$\tanh(R_{ij}^l / 2) = \prod_{k \in N(i) \setminus j} \tanh(Q_{ik}^l / 2) \quad (2-12)$$

可以发现第一个迭代公式主要采用求和运算，而第二个迭代公式则主要采用乘积运算，因此置信传播算法也被称作和积译码算法（Sum-Product Algorithm）。图 2-3 给出了和积译码算法在 Tanner 图上的具体描述。

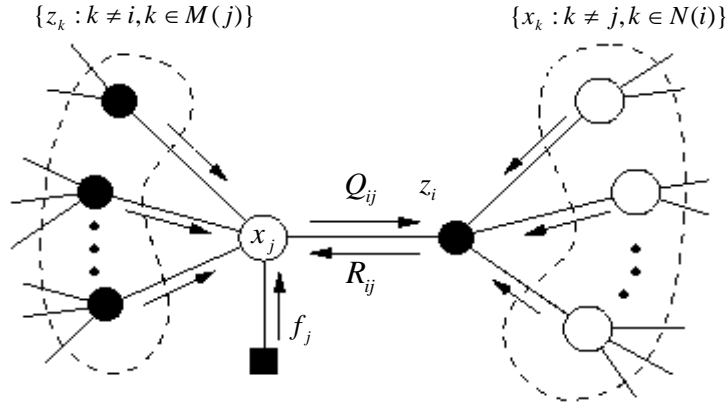


图 2.3 LDPC 码的译码算法示意图

2.2.2 BIAWGN 信道下的算法描述

在 BIAWGN 信道下，由信道编码器产生的码字序列 $\mathbf{x} = \{x_1, x_2, \dots, x_n\} \in GF^n(2)$ 通常需要经过 BPSK 调制，即经过映射 $0 \rightarrow +1, 1 \rightarrow -1$ 得到 $\tilde{\mathbf{x}} = \{\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_n\}$ ，然后才送入信道，并在接收端得到接收序列 $\mathbf{y} = \{y_1, y_2, \dots, y_n\} \in R^n$ 。根据 BIAWGN 信道的特性可以得到：

$$y_j = \tilde{x}_j + n_j \quad j = 1, 2, \dots, n \quad (2-13)$$

其中 n_j 为服从 $N(0, \sigma^2)$ 的一维高斯随机变量，因此可以推知：

$$\begin{aligned} f_j &= \ln \frac{P(x_j = 0 | \mathbf{y})}{P(x_j = 1 | \mathbf{y})} = \ln \frac{P(\tilde{x}_j = +1 | y_j)}{P(\tilde{x}_j = -1 | y_j)} \\ &= \ln \frac{P(y_j | \tilde{x}_j = +1)}{P(y_j | \tilde{x}_j = -1)} = \ln \frac{P(n_j = y_j - 1)}{P(n_j = y_j + 1)} \\ &= \ln \frac{\frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(y_j-1)^2}{2\sigma^2}}}{\frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(y_j+1)^2}{2\sigma^2}}} = 2y_j / \sigma^2 \end{aligned} \quad (2-14)$$

BIAWGN 信道下和积译码算法的整个流程如下：

- i、 初始化：对所有变量节点根据(2-14)式计算 f_j ，对所有 Q_{ij}^l 赋初值 $Q_{ij}^l = f_j$ ；
- ii、 迭代：
 - 水平步骤：根据式(2-11)，对所有校验节点 i ($i = 1, \dots, m$)，计算 R_{ij}^l ；
 - 垂直步骤：根据式(2-10)，对所有变量节点 j ($j = 1, \dots, n$)，计算 Q_{ij}^l ；

- iii、 判决与终止迭代：每次迭代结束后对所有变量节点计算 Q_j^l 并做硬判决，其中 Q_j^l 计算公式为：

$$Q_j^l = f_j + \sum_{k \in M(j)} R_{kj}^l \quad (2-15)$$

判决规则为：

$$\bar{x}_j = \begin{cases} 0 & Q_j^l > 0 \\ 1 & Q_j^l < 0 \end{cases} \quad (2-16)$$

若判决序列 $\bar{\mathbf{x}} = \{\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n\}$ 满足方程 $\bar{\mathbf{x}} \cdot \mathbf{H}^T = \mathbf{0}$ ，则停止迭代并输出判决结果，否则继续迭代至最大迭代次数 l_{\max} 。

2.2.3 BSC 信道下的算法描述

理论上讲，LDPC 码在 BSC 信道下的译码可以直接采用和积译码算法，它与 BIAWNG 信道下和积译码算法的不同之处仅在于初始化公式，设 BSC 信道的转移概率为 p ，某个二值随机变量 x 经过 BSC 信道输出的二值变量为 y ，当 $y=1$ 时有 $LLR(x|y) = \ln p - \ln(1-p)$ ，当 $y=0$ 时有 $LLR(x|y) = \ln(1-p) - \ln p$ ，初始化公式可写做：

$$f_j = \begin{cases} \ln p - \ln(1-p), & y_j = 1 \\ \ln(1-p) - \ln p, & y_j = 0 \end{cases} \quad (2-17)$$

在实际应用中，由于 BSC 信道是硬判决信道，人们总希望有一种合适的、低复杂度的硬判决译码算法。相应的算法有很多，这里仅介绍由 Gallager 提出的一种。

设发端发送的码字序列为 $\mathbf{x} = \{x_1, \dots, x_n\} \in GF^n(2)$ ，在接收端接收到的序列为 $\mathbf{y} = \{y_1, \dots, y_n\} \in GF^n(2)$ ，则 BSC 信道下 LDPC 码的硬判决译码算法流程如下：

- i、 初始化：所有变量节点赋初值 $f_j = y_j$ ，对所有 Q_{ij}^l 赋初值 $Q_{ij}^l = f_j$ ；
 ii、 迭代：
 水平步骤：对所有校验节点 i ($i=1, \dots, m$)，按照下式计算 R_{ij}^l

$$R_{ij}^l = \sum_{k \in N(i) \setminus j} Q_{ik}^{l-1} \quad (2-18)$$

垂直步骤：对所有变量节点 j ($j=1, \dots, n$)，按照下式计算 Q_{ij}^l

$$Q_{ij}^l = \begin{cases} b & \text{若对 } \forall k \in M(j) \setminus i, \text{ 恒有 } R_{kj}^{l-1} = b; \\ f_j & \text{其它情况。} \end{cases} \quad (2-19)$$

- iii、 判决与终止迭代：每次迭代结束后对所有变量节点按下式计算 Q_j^l 并做判决：

$$Q_j^l = (1 - 2 \times f_j) + \sum_{k \in M(j)} (1 - 2 \times R_{kj}^l) \quad (2-20)$$

判决规则为：

$$\bar{x}_j = \begin{cases} 0 & Q_j^l > 0 \\ 1 & Q_j^l < 0 \end{cases} \quad (2-21)$$

若判决序列 $\bar{\mathbf{x}} = \{\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n\}$ 满足方程 $\bar{\mathbf{x}} \cdot \mathbf{H}^T = \mathbf{0}$ ，则停止迭代并输出判决结果，否则继续迭代至最大迭代次数 l_{\max} 。

2.2.4 BEC 信道下的算法描述

类似于 BSC 信道，BEC 信道下 LDPC 码的译码也可以直接采用其在 BIAWGN 信道下的译码算法，下面先从和积译码算法出发，通过针对其在 BEC 信道下的分析推导出相应的硬判决算法。

在 BEC 信道下，和积译码算法的初始化公式应更新为：

$$f_j = \begin{cases} +\infty & y_j = 0 \\ 0 & y_j = E \\ -\infty & y_j = 1 \end{cases} \quad (2-22)$$

现在考虑消息的更新过程。对于消息 R_{ij}^l 的更新公式(2-12)，由于等式的右边为若干双曲正切函数的连乘，且双曲正切函数具有下列性质：

$$\tanh(x) = \begin{cases} +1 & x = +\infty \\ 0 & x = 0 \\ -1 & x = -\infty \end{cases} \quad (2-23)$$

故只要 $\exists k \in N(i) \setminus j$ 使得 $Q_{ik}^{l-1} = 0$ ，等式(2-12)的右边边为 0，相应的必有 $R_{ij}^l = 0$ ；而对于消息 Q_{ij}^l 的更新公式(2-10)，当 $l > 0$ 时，等式右边为连加，只要其中有一项为非 0，相应的 Q_{ij}^l 就为无穷（注意根据 BEC 信道的特性，公式(2-10)的右边不可能同时出现 $+\infty$ 和 $-\infty$ ）。

通过上面分析，可以得到 BEC 信道下 LDPC 码的译码算法流程为：

- i、初始化：所有变量节点根据接收值赋为 0、1 或删除 E ，所有校验节点赋初值 0；
- ii、直接恢复：对所有变量节点，若某一变量节点未被删除，则将该节点的接收值（模二）加到所有与其相连的校验节点上，并从原先的二部图中移去该变量节点和所有与该变量节点相连的边；
- iii、迭代恢复：若剩下的二部图中存在有度为 1 的校验节点，则唯一与其相连的变量节点的值就等于该校验节点的值，这样就恢复出来一个被删除的变量节点，然后再从二部图中去掉恢复出的变量节点及其相连的边，重复替代恢复操作直至所有的变量节点都被恢复出来或剩下的二部图中不存在度为 1 的校验节点。

图 2.4 给出了该算法的译码流程。可以看出，整个译码过程实际上是在二部图上去边的一个过程，而二部图的稀疏性则可以保证该算法具有线性的复杂度。值得一提的是 Luby 于 2001 年首次提出该算法时，并没有从和积译码算法的角度出发。

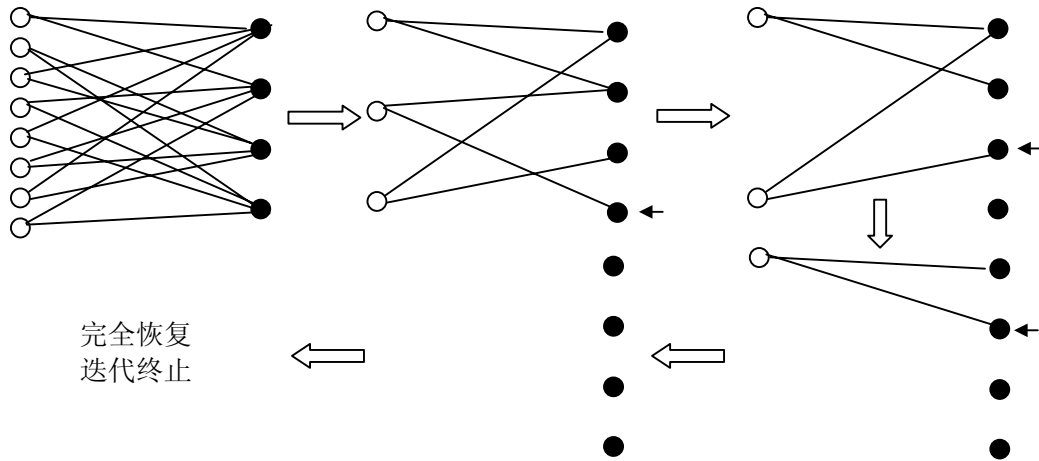


图 2.4 LDPC 码在删除信道下的译码流程

2.3 LDPC 码的性能分析

2.3.1 LDPC 码的度序列设计及密度进化理论

Gallager 虽然证明了正则 LDPC 码是渐进好码，但分析发现正则码在和积译码算法下并不能完全逼近 Shannon 限。1997 年，Luby 针对删除信道给出了一个惊人的结果，即非正则 LDPC 码的性能优于正则 LDPC 码；接着，Richardson 等针对 AWGN 信道从理论上证明了基于和积译码算法的非正则 LDPC 码优于正则 LDPC 码。虽然是否存在以任意接近信道容量的速率进行可靠传输的 LDPC 码仍然是一个非常重要的公开问题，人们已构造出许多非常接近 AWGN 信道容量和任意接近删除信道容量的非正则 LDPC 码。

给定一对度序列分布函数 $\lambda(x)$ 和 $\rho(x)$ ，就定义了一个满足度分布的 LDPC 码集，为了从理论上分析该码集的平均性能，Richardson 等人提出了密度进化理论（Density Evolution Theory）^{[32][46]}，下面进行简要介绍。

假定在每一轮迭代中，每个节点接收到的消息之间统计独立，由于信道的输出值是一些满足一定概率分布、互相独立的随机变量，整个译码过程实际上是求解一个以多个随机变量为自变量的复杂函数，因此如果能够跟踪译码过程中消息的概率密度的变化，则可以从理论上确定一个给定的 LDPC 码在独立性假设下采用和积译码算法所能够达到的纠错性能。

考虑和积译码算法中的(2-12)式，为便于分析，通常将该计算式变换到对数域（Log-Domain）中。由于函数 $\tanh(x)$ 的取值可正可负，而负数不能做真数进行对数运算，因此需要对函数 $\tanh(x)$ 的符号单独跟踪。令 γ 为一个从实数域 $(-\infty, +\infty)$ 到域 $F_2 \times (0, +\infty)$ 上的映射，该映射定义如下：

$$\gamma(x) \equiv (\text{sgn}(x), -\ln \tanh(|x|/2)) \quad (2-24)$$

$$\text{其中 } \text{sgn}(x) = \begin{cases} 1 & x < 0 \\ 0 & x \geq 0 \end{cases}$$

易知 $\gamma(x)$ 是一个双射函数，因此它的逆函数 $\gamma^{-1}(x)$ 存在，且满足

$$\gamma(xy) = \gamma(x) + \gamma(y) \quad (2-25)$$

则(2-12)式可以等效变换为：

$$R_{ij}^l = \gamma^{-1}\left(\sum_{k \in N(i) \setminus j} \gamma(Q_{ik}^l)\right) \quad (2-26)$$

至此可以看出，在和积译码算法中，只有加法操作和 γ 函数运算及其逆运算，而根据概率知识可知若 v_1, \dots, v_d 为某个加群 G 中的 d 个独立随机变量，其概率密度函数均为 f ，令 $\chi = v_1 + \dots + v_d$ ，则随机变量 χ 的概率密度函数则应为 f 的 d 次卷积，记做 $f^{\otimes d}$ 。现在我们来推导译码过程中消息概率密度的演化。

设 f 、 g_l 和 h_l 分别为初始消息 f_j 及消息 R_{ij}^l 和 Q_{ij}^l 的概率密度函数，并设变量节点 j 的度数为 d ，则由(2-10)式可得到传递给变量节点 j 的消息 Q_{ij}^{l+1} 的概率密度函数为：

$$h_{l+1} = f \otimes g_l^{\otimes(d-1)} \quad (2-27)$$

通常假设 LDPC 码对应的 Tanner 图是随机构造的，图中每条边与度为 d 的变量节点相连的概率为 λ_d ，与度为 d 的校验节点相连的概率为 ρ_d ，则消息 Q_{ij}^{l+1} 的期望概率密度函数应为：

$$h_{l+1} = f \otimes \lambda(g_l) = f \otimes \left(\sum_d \lambda_d g_l^{\otimes(d-1)}\right) \quad (2-27)$$

对消息 R_{ij}^l ，需要考虑(2-24)式中定义的函数 $\gamma(x)$ ，设 x 为域 $(-\infty, +\infty)$ 上的随机变量，其概率密度函数为 F ， $\Gamma(F)$ 为随机变量函数 $\gamma(x)$ 的概率密度函数，则消息 R_{ij}^l 的期望概率密度函数应为：

$$g_l = \Gamma^{-1}(\rho(\Gamma(f_l))) = \Gamma^{-1}\left(\sum_d \rho_d (\Gamma(f_l))^{\otimes(d-1)}\right) \quad (2-28)$$

综合(2-27)式和(2-28)式，就得到：

$$h_{l+1} = f \otimes \lambda(\Gamma^{-1}(\rho(\Gamma(f_l)))) \quad (2-29)$$

式(2-29)就是密度进化理论的递归计算式。利用密度进化理论，可以求得具有确定度序列分布函数的 LDPC 码在和积译码算法下能够实现可靠传输的渐进阈值，反过来，也可以利用密度进化理论来优化 LDPC 码的度序列设计，从而获得具有更大渐进阈值的度序列函数。

上面仅仅介绍了密度进化理论的简单推导，并没有给出其具体实现方案，如没有给出函数 Γ 的表达式。事实上，密度进化理论模型是一个多参数的动态系统，在复杂信道上的应用非常困难，Chung 等通过对密度进化理论的研究，进一步提出了应用高斯逼近

原理^{[33][48]}来简化译码阈值计算和收敛性分析，从而将密度进化理论模型简化为单一参数动态系统的高斯逼近模型，这里不再赘述。

2.3.2 LDPC 码的围长设计

在推导和积译码算法和密度进化理论的过程中，我们都假定所传递的消息之间是统计独立的，即满足独立性假设，而对于给定的度序列分布，大部分 LDPC 码在码长为有限时都不能满足该假设。下面分析消息之间的非独立性对 LDPC 码性能的影响以及减弱或消除该影响的方法。

每一个 LDPC 码都有相应的 Tanner 图表示。可以证明，当相应的 Tanner 图为无环图（树图）时，译码过程中所传递的消息之间必然满足独立性假设。但人们通过分析发现，无环 Tanner 图对应的 LDPC 码由于不具有良好的距离谱分布，其纠错性能往往没有有环 Tanner 图对应的 LDPC 码好；而当图中存在有长度为 $2l$ 的环路时，译码过程中传递的消息则只在前 l 轮迭代内满足独立性假设，相应的密度进化理论也只在在前 l 轮迭代内对消息概率密度的分析是准确的。为了研究环的存在对译码性能的影响，人们提出了围长的概念。所谓围长（Girth），是指一个 LDPC 码所对应的 Tanner 图中所有环的最小长度，该值至少为 4 且必为偶数。下图给出了一个长度为 4 的环的例子。

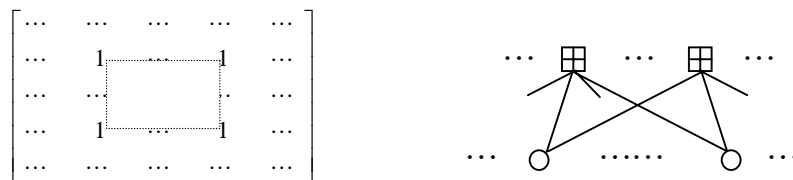


图 2.5 4 环环路在校验矩阵和 Tanner 图中的表示

由于和积译码算法在通过消息传递来降低每个变量节点的模糊程度的过程中，前几轮迭代是最有效的，人们通常希望 LDPC 码的围长尽可能的大，从而使相应的 Tanner 图中不存在长度较小的环，以保证最初几轮迭代中的消息满足独立性假设。例如，在 Gallager 给出的正则 LDPC 码的定义中增加一条：(4)、 H 的任意两列在共同的行上的“1”的个数至多为 1，就可以消除相应 Tanner 图中长度为 4 的环。除此之外，众多研究学者从代数构造和启发式搜索等各种方法出发，提出了许多具有较大围长的正则及非正则 LDPC 码的构造方法^[49~60]，这里不一一介绍，仅给出一个正则 LDPC 码围长上限的理论分析。

若一个 (n, λ, ρ) 正则 LDPC 码的围长为 $2l$ ，则对相应 Tanner 图中的任意一个变量节点而言，图中所有与该节点最短距离不超过 $l-1$ 的节点（包括变量节点和校验节点）以及它们之间相连的边必然构成一个深度（Depth）为 $l-1$ （这里设根节点的深度为 0）的树图，该树图的根节点度数为 λ ，所有深度为奇数的节点（均为校验节点）均有 $\rho-1$ 个子节点，同样所有深度为偶数的节点（均为变量节点）均有 $\lambda-1$ 个子节点，这些节点各不相同。现在考虑该树图中的变量节点个数：深度为 0 的变量节点个数为 1，深度为 2 的变量节点个数为 $\lambda(\rho-1)$ ，深度为 4 的变量节点个数为 $\lambda(\rho-1)^2(\lambda-1)$ ，依次类推，可以得到总共的变量节点个数为：

$$n' = \frac{D^{\lfloor \frac{l+1}{2} \rfloor} - 1}{D-1} + (\rho-1) \frac{D^{\lfloor \frac{l-1}{2} \rfloor} - 1}{D-1} \quad (2-30)$$

其中 $D = (\lambda-1)(\rho-1)$ 。

显然，该树图作为整个 Tanner 图的子图，其中变量节点的个数必然不超过整个 Tanner 图的变量节点数 n ，即有：

$$n' \leq n \quad (2-31)$$

结合(2-30)式和(2-31)式，可得到：

$$2l \leq 4 \left\lceil \log_D \left(\frac{(D-1)n + \rho}{D-1 + \rho} \right) \right\rceil + 2 \quad (2-32)$$

若考虑该树图中的校验节点个数或以任意一个校验节点为根节点的树图，可以推导出类似的围长上限。由(2-32)式可知，围长上限的数量级为 $\log_D n$ 。

2.4 本章小结

本章系统地阐述了 LDPC 码的基本编译码原理，给出了 LDPC 码在常见信道下的译码算法，简要介绍了密度进化理论，并对影响 LDPC 码性能的两个主要因素——度序列函数和围长进行了分析。需要注意的是，基于密度进化理论的度序列设计和围长设计都是针对 LDPC 码采用置信传播算法译码所能达到的性能而言的，相应的理论分析结果只有在码长趋于无穷时才能完全与实际相符。如今，有关 LDPC 码的有限长分析已经成为一个研究的热点。

第三章 删除信道下的 LDPC 码

本章首先阐述了纠删码的基本概念及其发展，介绍了 LDPC 码在删除信道下所能达到性能的理论分析，进而从理论上证明了在删除信道下，所有的 $(d, 2d)$ -正则 LDPC 码都不是码率逼近信道容量限的非常好码，然后介绍了几类码率能够逼近信道容量限的非正则 LDPC 码度序列设计，并基于右边正则 LDPC 码的度序列给出了一种改进型右边正则度序列设计，最后对这些码在删除信道下的纠错性能进行了仿真，给出了本章总结。

3.1 纠删码及其发展

3.1.1 删除信道和纠删码

绪论部分已经给出了删除信道的定义及其模型，它与其它信道有着明显的不同。在一般的通信系统中，前向纠错码纠正的错误所在的位置事先通常是不知道的，而当系统中的信道为删除信道时，错误的的数据被遗弃，丢失的数据在数据流中的位置是知道的，这样的错误纠正起来相对容易些。

删除信道在实际通信系统中是很常见的，如在互联网中进行的多点传输 (Multicast) 和广播传输 (Broadcast) [61][62][63]。当网络公司要通过互联网对众多用户有效地传送大容量的数据 (如大型软件、图像等)，一般都会使用多点传输或广播传输。为保证传输的可靠性，同时有尽可能小的网络开销及支持众多各类用户的随机访问，当用户请求传输的数据丢失时，可以采用 ARQ (Automatic Repeat Request) 技术，即重传丢失的数据，但这种方法很容易造成网络过载和阻塞现象，严重时会使网络瘫痪 [64][65][66][67]。为了解决这一问题，已经提出了随机发送、本地恢复、分层恢复等新技术来有效地提高数据传输的可靠性和避免网络阻塞，但有可能导致大的时延，这一点在大容量数据的实时传输中是不可接受的。为了克服基于重传带来的缺点，许多研究者将基于纠删码的前向纠错 (FEC, Forward Error Correcting) 技术用于可靠的多点传输 [68][69][70][71][81][82][83]。

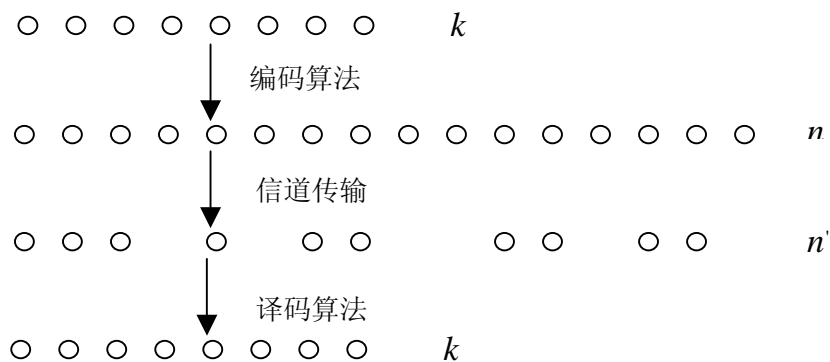


图 3.1 纠删码的基本原理图

纠删码的基本原理如上图所示，要传输的 k 个源数据包经过编码得到 n ($n > k$) 个数据包后发送出去，在接收端，若接收方收到足够数量的数据包，就可以运用适当的译码方法来重构 k 个源数据包。严格来讲，几乎所有的纠错码都可以用做纠删码，但如何设计一种纠删能力强、编译码速度快、同时满足大容量数据实时可靠传输的纠删码则是一个比较困难的问题。根据纠错码理论，对 $[n, k, d]$ 线性分组码，只要其中的删除错误不超过 $d-1$ 个，就可以通过适当的算法将删除错误全部恢复；而如果一个线性分组码从接收到的 n 个编码数据包中的任意 k 个都能够恢复出所有的删除错误，则称此码关于恢复删除错误为最优码。

定理 3.1 给出了最优码的充要条件。

定理 3.1^{[69][71]}: 设 \mathbf{G} 为一 (n, k) 线性纠删码 C 的生成矩阵，则 C 为最优码的充要条件为 \mathbf{G} 的任意 k 列组成的子矩阵 \mathbf{G}' 均可逆。

证明: 设 \mathbf{c}' 为由源数据 \mathbf{x} 生成的码字 \mathbf{c} 的任意 k 个分量组成的向量， \mathbf{G}' 为这 k 个分量对应的 \mathbf{G} 中各列构成的子矩阵，若 C 为最优码，则由 \mathbf{c}' 可以恢复出源数据 \mathbf{x} ，即存在矩阵 \mathbf{M} 满足

$$\mathbf{x} = \mathbf{c}'\mathbf{M} \quad (3-1)$$

而根据码的结构有

$$\mathbf{c}' = \mathbf{x}\mathbf{G}' \quad (3-2)$$

故有

$$\mathbf{x} = \mathbf{c}'\mathbf{M} = \mathbf{x}\mathbf{G}'\mathbf{M} \Rightarrow \mathbf{G}'\mathbf{M} = \mathbf{I} \quad (3-3)$$

即得矩阵 \mathbf{G}' 可逆；

反过来，若矩阵 \mathbf{G}' 可逆，则有

$$\begin{aligned} \mathbf{c}' &= \mathbf{x}\mathbf{G}' \\ \Rightarrow \mathbf{c}'\mathbf{G}'^{-1} &= \mathbf{x}\mathbf{G}'\mathbf{G}'^{-1} \\ \Rightarrow \mathbf{x} &= \mathbf{c}'\mathbf{G}'^{-1} \end{aligned} \quad (3-4)$$

故由 \mathbf{c}' 可以恢复出源数据 \mathbf{x} 。#

对有限域 $GF(p^r)$ (p 为素数， r 为正整数)，有下面引理：

引理 3.1^[72]: 设 C 为 $GF(p^r)$ 上的 (n, k) 线性分组码，则当且仅当其生成矩阵的任意 k 列线性无关时， C 为最大距离可分 (MDS) 码。

由定理 3.1 和引理 3.1，可以得到下面定理：

定理 3.2: $GF(p^r)$ 上的 (n, k) 线性纠删码为 MDS 码，当且仅当利用接收到的任意 k 个数据包可以重构源数据包。

因此，最优线性纠删码为最大距离可分 (MDS) 码。

3.1.2 RS码类纠删码^[72]

RS码的全称是Reed-Solomon码，它是由I.S.Reed和G.Solomon在1960年提出的一种纠错能力很强的多进制BCH码，是一类典型的代数几何码。由于RS码是极大距离可分码，因此可以用做最优纠删码。在RS码类纠删码中，比较常用的是范德蒙码（Vandermonde Code）^[71]和柯西码（Cauchy Code）^[73]。

范德蒙码：若纠删码的生成矩阵 $\mathbf{G}_{k \times n}$ 满足 $\mathbf{G}^T = (g_{ij})$ ，其中 $g_{ij} = x_i^{j-1}$ ， $x_i \in GF(p^r)$ （ p 为素数， r 为正整数），则称该纠删码为范德蒙码。 \mathbf{G} 的任意 k 列组成的子方阵 \mathbf{G}' 的转置矩阵 $(\mathbf{G}')^T$ 为范德蒙矩阵，若 $x_i (i=1, 2, \dots, k)$ 互不相同，则 $|(\mathbf{G}')^T| \neq 0$ ，从而 $|\mathbf{G}'| \neq 0$ ，即 \mathbf{G} 的任意 k 列线性无关，所以这样得到的矩阵满足最优纠删码生成矩阵的特性。

柯西码：设 $X = \{x_1, x_2, \dots, x_m\}$ 和 $Y = \{y_1, y_2, \dots, y_n\}$ 为有限域 F 中的两个元素集，若 (1)、对 $\forall i \in X$ 和 $\forall j \in Y$ ，有 $x_i + y_j \neq 0$ ；(2)、对 $\forall i, j \in \{1, 2, \dots, m\} (i \neq j)$ 有 $x_i \neq x_j$ 和对 $\forall i, j \in \{1, 2, \dots, n\} (i \neq j)$ 有 $y_i \neq y_j$ ，则称矩阵 $\mathbf{C}_{m \times n} = (c_{ij})$ 为有限域 F 上的柯西矩阵，其中 $c_{ij} = 1/(x_i + y_j)$ 。然后设 $I_{m \times m}$ 为有限域 F 上的单位阵，则称以矩阵 $\mathbf{G} = (\mathbf{I} | \mathbf{C})$ 为生成矩阵的纠删码为柯西码，下面引理可以保证柯西码为最优纠删码。

引理 3.2^[72]：设 \mathbf{C} 为某一有限域上的柯西矩阵，则 \mathbf{C} 的任意子方阵均可逆。

需要指出，尽管范德蒙码和柯西码具有强大的恢复删除能力，它们在编译码过程中都要用到矩阵运算，因此具有较高的复杂度。其中范德蒙码的编译码时间复杂度为 $O(n^2)$ ，译码时间复杂度高于 $O(n^2)$ ，而柯西码的编译码时间复杂度均为 $O(n^2)$ 。Elias已经证明删除信道的信道容量为 $1-p$ ，且随机线性码可在删除信道下以任意不超过信道容量的速率传输，其编译码时间复杂度分别为 $O(n^2)$ 和 $O(n^3)$ 。

3.1.3 低密度纠删码（LDEC, Low Density Erasure Code）

当LDPC码在Turbo码提出之后被重新发现和推广时，人们迅速将它应用到了删除信道，称做低密度纠删码。1997年，Luby等人基于LDEC采用级联的方式给出了一种能以任意接近删除信道容量的速率传输的线性时间复损码（Loss-Resilient Codes）——级连型低密度纠删码^{[28][29]}，该类纠删码具有线性编译码时间复杂度，更适合实际应用，进一步的研究说明在大文件的可靠多点传输中复损码比基于RS码的纠删码更有效。

复损码的构造如下：设码 $C(B)$ 有 k 个信息比特位和 βk ($0 < \beta < 1$)个校验比特位，该码所对应的二部图 B 的左边集有 k 个节点（称做信息比特节点），右边集有 βk 个节点（称做校验比特节点），分别对应码 $C(B)$ 的 k 个信息比特位和 βk 个校验比特位，且二部图 B 通常为稀疏图，其结构与LDPC码所对应的Tanner图相似，其中的信息比特节点和校验比特节点分别对应于Tanner图中变量节点和校验节点。 $C(B)$ 型码的编码方式

如图 3.2 所示，每个校验比特位等于二部图 B 中和该校验比特节点相连的所有信息比特的模二和。

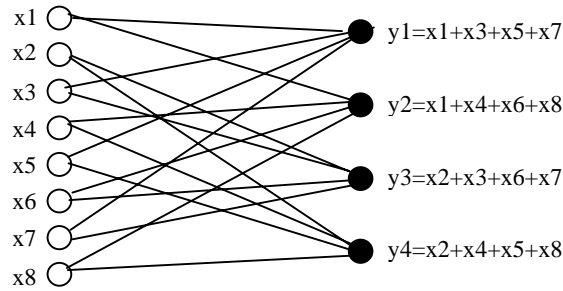


图 3.2 $C(B)$ 型码编码的二部图表示

若设 $C(B)$ 型码的所有校验比特位的值均已知，则相应的译码即可按照 LDPC 码在删除信道下的译码算法进行，不同的是在初始化步骤中所有校验比特节点应赋其接收值，而对于校验比特位有删除的情况，也只需要在除去被删除的校验比特节点后的剩余二部图上进行译码即可。

现在考虑将若干个 $C(B)$ 型码（分别为 $C(B_0)$ ， $C(B_1)$ ， \dots ， $C(B_m)$ ）级联起来。其中码 $C(B_i)$ 有 $\beta^i k$ 个信息比特位和 $\beta^{i+1} k$ 个校验比特位，码 $C(B_i)$ 的校验比特位就是码 $C(B_{i+1})$ 的信息比特位。适当选取 m 使得 $\beta^{m+1} k \approx \sqrt{k}$ ，然后再将一个有 $\beta^{m+1} k$ 个信息比特位、码率为 $1-\beta$ 的传统纠错码（如柯西码） C 与码 $C(B_m)$ 级联，就最终得到了复损码。复损码在编码时由前向后依次对码 $C(B_i)$ 和码 C 进行编码，而译码时则是从后向前依次对码 C 和码 $C(B_i)$ 进行译码。可以计算出该复损码的码率 R 为：

$$R = \frac{k}{n} = \frac{k}{\sum_{i=0}^{m+1} \beta^i k + \beta^{m+2} k / (1-\beta)} = 1-\beta \tag{3-5}$$

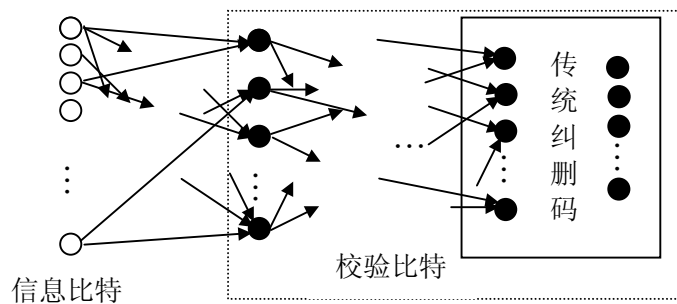


图 3.3 复损码的二部图结构示意图

本质上讲，复损码的编译码算法的核心是求解一个线性方程组，当每一个二部图 B_i 均稀疏时，相应方程组中每个方程平均包含的变量个数很少，因此能够获得很快的编译码速度，而它的局限性是为保证所有的删除错误都被恢复，正确接收到的数据量必须略大于源数据个数 k （即为需付出的代价）。对于复损码，Luby 给出了下面定理。

定理 3.3^[28]：对任意码率 R 和任意小的正数 $\varepsilon > 0$ ，若码长 n 足够大，必存在复损码能以大概率 $O(n \ln(1/\varepsilon))$ 的运行时间从 $(1-R)(1-\varepsilon)$ 的随机损失率中恢复出源数据。

3.2 删除信道下的密度进化理论

3.2.1 密度进化理论的直接描述^[46]

第二章给出了一般信道下 LDPC 码的密度进化理论。对于一般信道，信道输出的变量所满足的概率分布函数往往比较复杂，甚至即使借助计算机也难以计算出密度进化理论的最终迭代结果，而对于删除信道，由于信道的输出只有 0、1 和 E，在对数似然比量度下译码过程中所传递的信息要么是 $\pm\infty$ ，要么是 0。根据删除信道的特性，可以不失一般性地假设每次发送的码字均为全 0 码字，这样消息的取值就具有二值性（ $+\infty$ 和 0），于是消息取值的概率密度函数变得非常简单，只需要用一个参数就可以表示，式 (2-29) 中的卷积运算也相应地可以简化为普通的四则运算。

设 LDPC 码边的度序列分布函数分别为 $\lambda(x) = \sum_i \lambda_i x^{i-1}$ 和 $\rho(x) = \sum_i \rho_i x^{i-1}$ ，信道的删除概率为 p ，在第 l 轮迭代译码操作中，由变量节点向校验节点传递的消息 Q_{ij}^l 为 0 的概率为 $P(Q_{ij}^l = 0) = p_l$ ；由校验节点向变量节点传递的消息 R_{ij}^l 为 0 的概率为 $P(R_{ij}^l = 0) = q_l$ 。若变量节点的度数为 d ，由 (2-10) 式可以推知，第 $l+1$ 轮迭代中只有当 f_j 和所有的 R_{kj}^l 均为 0， Q_{ij}^{l+1} 才能为 0，故 Q_{ij}^{l+1} 为 0 的概率为：

$$P(Q_{ij}^{l+1} = 0) = p_{l+1} = p \cdot q_l^{d-1} \quad (3-6)$$

而一条边与度数为 d 的变量节点相连的概率为 λ_d ，则在整个 Tanner 图上消息 Q_{ij}^{l+1} 为 0 的概率为：

$$P(Q_{ij}^{l+1} = 0) = p_{l+1} = p \cdot \sum_d \lambda_d q_l^{d-1} = p\lambda(q_l) \quad (3-7)$$

同样，若校验节点的度数为 d ，由 (2-11) 式可以推知，第 l 轮迭代中只要有一个 Q_{ik}^l 为 0， R_{ij}^l 就能为 0，故 R_{ij}^l 为 0 的概率为：

$$P(R_{ij}^l = 0) = q_l = 1 - P(R_{ij}^l = 1) = 1 - (1 - p_l)^{d-1} \quad (3-8)$$

而一条边与度数为 d 的校验节点相连的概率为 ρ_d ，则在整体 Tanner 图上消息 R_{ij}^l 为 0 的概率为：

$$P(R_{ij}^l = 0) = q_l = 1 - \sum_d \rho_d (1 - p_l)^{d-1} = 1 - \rho(1 - p_l) \quad (3-9)$$

由 (3-7) 式和 (3-9) 式可知，

$$p_{l+1} = p\lambda(q_l) = p\lambda(1 - \rho(1 - p_l)) \quad (3-10)$$

如果在迭代译码过程中，所传递消息 Q_{ij}^l 为 0 的概率是单调递减的，即对 $\forall l \in N$ 及 $\exists \varepsilon > 0$ ，恒有

$$p_{l+1} < (1 - \varepsilon)p_l \quad (3-11)$$

就可以以很高的概率译码成功。

将 (3-10) 式代入 (3-11) 有：

$$p\lambda(1 - \rho(1 - x)) < x, \quad x \in [0, p) \quad (3-12)$$

同样，由 (3-7) 式和 (3-9) 式可知：

$$q_{l+1} = 1 - \rho(1 - p_{l+1}) = 1 - \rho(1 - p\lambda(q_l)) \quad (3-13)$$

如果在迭代译码过程中，所传递消息 R_{ij}^l 为 0 的概率是单调递减的，即对 $\forall l \in N$ 及 $\exists \varepsilon > 0$ ，恒有

$$q_{l+1} < (1 - \varepsilon)q_l \quad (3-14)$$

就可以以很高的概率译码成功。

将(3-13)式代入(3-14)式有：

$$\rho(1 - p\lambda(x)) > 1 - x, \quad x \in [0, 1) \quad (3-15)$$

下面定理可以保证(3-12)式和(3-15)式是等价的。

定理 3.4^[44]：对给定度序列分布函数 $\lambda(x)$ 和 $\rho(x)$ 和信道删除概率 $p \in [0, 1)$ ，则对 $\forall x \in [0, 1)$ ，恒有 $\rho(1 - p\lambda(x)) > 1 - x$ 当且仅当对 $\forall x \in [0, p)$ ，恒有 $p\lambda(1 - \rho(1 - x)) < x$ 成立。

证明：由度序列分布函数的定义可知， $\lambda(x)$ 为 $x \in [0, 1)$ 上的严格单调递增函数，故它的逆函数 $\lambda^{-1}(x)$ 唯一存在且为严格单调递减函数，于是有：

必要性：若对 $\forall x \in [0, 1)$ ，恒有 $\rho(1 - p\lambda(x)) > 1 - x$ ，令 $y = p\lambda(x)$ ，则 $0 \leq y < p$ 且 $x = \lambda^{-1}(y/p)$ ，代入(3-15)式有：

$$\rho(1 - y) > 1 - \lambda^{-1}(y/p) \quad (3-16)$$

上式变换得到：

$$\begin{aligned} 1 - \rho(1 - y) &< \lambda^{-1}(y/p) \\ p\lambda(1 - \rho(1 - y)) &< y, \quad 0 \leq y < p \end{aligned}$$

即得到(3-12)式；

充分性：若对 $\forall x \in [0, p)$ ，恒有 $p\lambda(1 - \rho(1 - x)) < x$ ，则 $\rho(1 - x) > 1 - \lambda^{-1}(x/p)$ ，令 $u = \lambda^{-1}(x/p)$ ，则有 $0 \leq u < 1$ ，且 $x = p\lambda(u)$ ，代入(3-12)式有

$$p\lambda(1 - \rho(1 - p\lambda(u))) < p\lambda(u) \quad (3-17)$$

注意到 $\lambda(x)$ 为 $x \in [0, 1)$ 上的严格单调递增函数，则有

$$\begin{aligned} 1 - \rho(1 - p\lambda(u)) &< u \\ \rho(1 - p\lambda(u)) &> 1 - u \end{aligned}$$

即得到(3-15)式。#

(3-12)式和(3-15)式就是删除信道下LDPC码的收敛判定公式，它们对在删除信道下给定度序列函数LDPC码的性能分析和度序列函数设计中起着非常重要的作用。对给定度序列分布函数 $\lambda(x)$ 、 $\rho(x)$ 和信道删除概率 p ，若(3-12)式或(3-15)式成立，则以很高的概率译码成功，反之则以很高的概率译码失败。Luby等进一步利用新的概率分析工具对译码过程相关联的“与或（And-Or）树”^[34]进行了分析，得出第 l 次迭代译码后删除错误的比率为：

$$x_l = x_l(p) = p\lambda(1 - \rho(1 - x_{l-1})) \quad (3-18)$$

其中 $x_0 = p \in [0, 1)$ ，同时定义了度分布 (λ, ρ) 的阈值（Threshold） $p^*(\lambda, \rho)$ 为：

$$p^*(\lambda, \rho) = \sup\{p \mid 0 \leq p < 1, \lim_{l \rightarrow \infty} x_l(p) = 0\} \quad (3-19)$$

容易看出若 $\lim_{l \rightarrow \infty} x_l(p) = 0$ ，则对任意 $p' < p$ 必有

$$\lim_{l \rightarrow \infty} x_l(p') = 0 \quad (3-20)$$

所以，如果 $p < p^*(\lambda, \rho)$ ，则删除错误信息的比率以很高的概率收敛于 0，此时纠错译码器以高概率成功译码；反之，若 $p > p^*(\lambda, \rho)$ ，纠错译码器以高概率不成功译码，因此，人们普遍认为对给定码率 R ，阈值 $p^*(\lambda, \rho)$ 越大，具有此度分布 (λ, ρ) 的低密度纠错码的性能越好。通常，对于码率为 R 的级联型低密度纠错码，能使阈值 $p^*(\lambda, \rho)$ 逼近 $1-R$ 的度分布 (λ, ρ) 所构造的码称为渐近最优码。

3.2.2 密度进化理论的微分方程描述^[28]

第二章提到，Luby 最早并没有从和积译码算法的角度出发，自行推导出了删除信道下 LDPC 码的译码算法。同时，Luby 为了分析该算法当码长趋于无穷时的理论性能，也没有基于密度进化理论，而是从微分方程的角度出发，推导出了删除信道下 LDPC 码的收敛判定公式，其推导过程如下。

考虑一个 $C(B)$ 型码，其二部图 B 中共有 k 个信息比特节点和 βk 个校验比特节点，边的度序列分布函数仍为 $\lambda(x) = \sum_i \lambda_i x^{i-1}$ 和 $\rho(x) = \sum_i \rho_i x^{i-1}$ ，设在译码开始之前， B 中边的总数为 E ，则左边信息比特节点的平均度数为

$$a_l = \frac{E}{k} = \frac{E}{E \int_0^1 \lambda(x) dx} = \frac{1}{\sum_i \lambda_i / i} \quad (3-21)$$

相应地，右边校验比特节点的平均度数为

$$a_r = \frac{E}{\beta k} = \frac{E}{E \int_0^1 \rho(x) dx} = \frac{1}{\sum_i \rho_i / i} \quad (3-22)$$

为便于描述，这里将译码时间分段，每一个单位时间段 Δt 表示译码过程中的一步操作，长度正比于 $1/E$ ，即 $\Delta t := 1/E$ 。令 p 表示信道的删除概率，则在译码的直接恢复过程中，每个信息比特节点以概率 $1-p$ 被从原始的二部图 B 中移走（由于相应的信息比特被正确接收），于是迭代恢复操作开始时（这时定义为 0 时刻），剩余二部图中的信息比特节点总数应为 pk ，若整个译码能够成功进行，总共的译码时间应为 $pk\Delta t := pk/E = p/a_l$ 。令 $l_i(t)$ 和 $r_i(t)$ 分别表示在 t 时刻，剩余二部图中与度为 i 的左边信息比特节点（右边校验比特节点）相连的边占总边数 E 的比率， $e(t)$ 表示同一时刻剩余二部图中的边占总边数 E 的比率，则有

$$e(t) = \sum_i l_i(t) = \sum_i r_i(t) \quad (3-23)$$

在迭代恢复的每一步译码操作中，右边某个度为 1 的校验比特节点被选中，随后左边与之相连的信息比特节点连同与该信息比特节点相连的边都被移去，而当剩余二部图

中右边不存在度为 1 的校验比特节点，译码操作就被迫停止。在 t 时刻，被移去的变量节点度数为 i 的概率为 $l_i(t)/e(t)$ ，于是有下面差分方程成立：

$$L_i(t + \Delta t) - L_i(t) = -il_i(t)/e(t) \quad (3-24)$$

其中 $L_i(t)$ 表示 t 时刻剩余二部图中与左边度为 i 的信息比特节点相连的边的期望数量，即有 $l_i(t) = L_i(t)/E = L_i(t)\Delta t$ ，当码长趋于无穷时，有 $E \rightarrow +\infty$ ，所以 $\Delta t := 1/E \rightarrow 0$ ，因此上面的差分方程就可以表示为微分方程：

$$\frac{dl_i(t)}{dt} = -\frac{il_i(t)}{e(t)} \quad (3-25)$$

在移去度为 i 的信息比特节点的同时，不但要移去一条与一个度为 1 的校验比特节点相连的边，还要移去另外 $i-1$ 条和该信息比特节点相连的边。因此，要移去的边数的期望值应为：

$$a(t) = \sum_i il_i(t)/e(t) \quad (3-26)$$

现在考虑被移去的其它 $i-1$ 条边，由于二部图构造的随机性，这 $i-1$ 条边所连的右边校验比特节点的度数是随机的。如果与其中一条边相连的校验比特节点度数为 j ，则经过一步译码操作之后，二部图上就少了 j 条与右边度数为 j 的校验比特节点相连的边，同时多了 $j-1$ 条与右边度数为 $j-1$ 的校验比特节点相连的边，而任意一条边与右边度数为 j 的校验比特节点相连的概率是 $r_j(t)/e(t)$ ，则对 $\forall i > 1$ ，有

$$\begin{aligned} R_i(t + \Delta t) - R_i(t) &= \frac{r_{i+1}(t)}{e(t)}((i+1)-1)(a(t)-1) + \frac{r_i(t)}{e(t)}i(a(t)-1) \\ &= (r_{i+1}(t) - r_i(t))\frac{i(a(t)-1)}{e(t)} \end{aligned} \quad (3-27)$$

其中 $R_i(t)$ 表示 t 时刻剩余二部图中与右边度为 i 的校验比特节点相连的边的期望数量，同样考虑码长趋于无穷的情况，就得到下面的微分方程：

$$\frac{dr_i(t)}{dt} = (r_{i+1}(t) - r_i(t))\frac{i(a(t)-1)}{e(t)} \quad (3-28)$$

易知当 t 趋于无穷时，对 $\forall i \in N$ ，都有 $r_i(t) \rightarrow 0$ 。由于每一次译码操作中，都要先找到一个度为 1 的校验比特节点，才能进行迭代恢复，所以度为 1 的校验比特节点数的变化对译码性能的影响是非常大的。根据(3-28)式可以写出关于 $r_1(t)$ 的微分方程：

$$\frac{dr_1(t)}{dt} = (r_2(t) - r_1(t))\frac{(a(t)-1)}{e(t)} - 1 \quad (3-29)$$

注意等式右边之所以减 1，是因为在每一步译码操作中，都要先去掉一条与右边度为 1 的校验比特节点相连的边。

我们关注的焦点是函数 $r_1(t)$ 随时间的变化，只要 $r_1(t) > 0$ ，右边就存在度为 1 的校验比特节点，译码操作就可以继续进行，当 $r_1(t) = 0$ 时，译码就要被迫停止，因此问题的关键是寻找 $r_1(t)$ 能够在左边所有信息比特节点都被恢复之前始终大于 0 的情况以保证译码成功。

对方程 (3-29) 进行求解的过程较为繁琐，这里不再赘述，仅给出如下的求解结果：

$$r_1(t) = p\lambda(x)[x - 1 + \rho(1 - p\lambda(x))] \quad (3-30)$$

显然，要保证 $r_1(t) > 0$ ，只需要下面不等式成立：

$$\rho(1 - p\lambda(x)) > 1 - x, \quad x \in [0, 1) \quad (3-31)$$

该式与上一部分中得到的(3-15)式是一样的。

3.3 正则 LDPC 码在删除信道下的性能

3.3.1 正则 LDPC 码阈值的唯一存在性分析

现在研究 (ℓ, r) -正则低密度删码的性能，即 $\lambda(x) = x^{\ell-1}$ ， $\rho(x) = x^{r-1}$ ， $r > \ell \geq 3$ 均为整数。由 (3-12) 式和 (3-19) 式可知，对 $\forall p < p^*(\lambda, \rho)$ 及 $\forall x \in [0, 1)$ ，恒有 $p\lambda(1 - \rho(1 - x)) \leq x$ 成立，即

$$p \leq (1 - x) / \lambda(1 - \rho(x)) \quad (3-32)$$

故令 $f(x) = (1 - x) / \lambda(1 - \rho(x)) = (1 - x) / (1 - x^{r-1})^{\ell-1}$ ，则有

$$p^*(\lambda, \rho) = p^*(\ell, r) = \min\{f(x) \mid x \in [0, 1)\} \quad (3-33)$$

考虑函数 $f(x)$ ，易知

$$\begin{aligned} \lim_{x \rightarrow 1^-} f(x) &= \lim_{x \rightarrow 1^-} \frac{(1-x)}{(1-(x^{r-1}))^{\ell-1}} \\ &= \lim_{x \rightarrow 1^-} \frac{1}{(1-x)^{\ell-2} \left(\sum_{i=0}^{r-2} x^i\right)^{\ell-1}} \\ &= \lim_{x \rightarrow 1^-} \frac{1}{(1-x)^{\ell-2} (r-1)^{\ell-1}} \end{aligned} \quad (3-34)$$

由 $r > \ell \geq 3$ 知 $\ell - 2 \geq 1$ 及 $r - 1 > 2$ ，则有 $\lim_{x \rightarrow 1^-} (1-x)^{\ell-2} = 0$ ， $(r-1)^{\ell-1} > 0$ ，故

$$\lim_{x \rightarrow 1^-} f(x) = +\infty \quad (3-35)$$

所以 $\exists \delta > 0$ ，使得 $x \in (\delta, 1)$ 时 $f(x) \geq C$ ，其中 C 为一个较大的正常数，同时考虑到 $\lim_{x \rightarrow 0^+} f(x) = 1$ ，故 $f(x)$ 为闭区间 $[0, \delta]$ 上的连续函数，在 $[0, \delta]$ 上必有最小值，则

$\min\{f(x) \mid x \in [0, 1)\} = \min(\{f(x) \mid x \in [0, \delta]\}, C)$ 必然存在。

事实上，由下面定理可知，该最小值不但是存在的，而且是唯一确定的。

定理 3.5: 对于 (ℓ, r) -正则低密度删码 ($r > \ell \geq 3$ 均为整数), 若 θ 是方程 $((r-1)(\ell-1)-1)x^{r-2} - (x^{r-3} + x^{r-4} + \dots + x + 1) = 0$ 在 $[0, 1)$ 上的惟一正实根, 则 $p^*(\ell, r) = (1-\theta)/((1-\theta^{r-1})^{\ell-1})$ 。

证明: 设 $f(x) = (1-x)/(1-x^{r-1})^{\ell-1}$, 易见其一阶导函数为:

$$\begin{aligned} f'(x) &= \frac{(\ell-1)(r-1)(1-x)x^{r-2} - (1-x^{r-1})}{(1-x^{r-1})^\ell} \\ &= -\frac{((\ell-1)(r-1)-1)x^{r-1} - (\ell-1)(r-1)x^{r-2} + 1}{(1-x^{r-1})^\ell} \end{aligned} \quad (3-36)$$

可以看出, $\forall x \in [0, 1)$, 上式的分母部分始终为正, 因此只需要考虑上式的分子部分。设 $g(x) = ((\ell-1)(r-1)-1)x^{r-1} - (\ell-1)(r-1)x^{r-2} + 1$ 令 $f'(x) = 0$ 得方程 $g(x) = 0$ 。

现在证明方程 $g(x) = 0$ 在 $[0, 1]$ 上至多有两个不同的实根。假如此方程在 $[0, 1]$ 上有三个不同的实根 (多于三个不同的实根的情形可类似证明) x_1, x_2 和 x_3 。不妨假定 $x_1 < x_2 < x_3$, 由罗尔定理必存在 $y_1 \in (x_1, x_2)$ 和 $y_2 \in (x_2, x_3)$ 使得 $g'(y_i) = 0$, 即

$$((\ell-1)(r-1)-1)y_i^{r-1} - (\ell-1)(r-2)y_i^{r-2} = 0 \quad (i=1, 2) \quad (3-37)$$

由 $y_1 \neq 0$ 和 $y_2 \neq 0$ 知

$$y_i = \frac{(\ell-1)(r-2)}{(\ell-1)(r-1)-1} \quad (3-38)$$

即 $y_1 = y_2$, 易见这是不可能的。可验证 $x=1$ 为方程 $g(x) = 0$ 的一实根, 故方程 $((r-1)(\ell-1)-1)x^{r-2} - (x^{r-3} + x^{r-4} + \dots + x + 1) = 0$ 在 $[0, 1)$ 上至多有一个实根, 由已知条件知 θ 就是方程 $g(x) = 0$ 在 $[0, 1)$ 上惟一的一实根, 从而 $f(x)$ 在 $[0, 1)$ 上仅有惟一的极值点, 再由 $\lim_{x \rightarrow 1^-} f(x) = \infty$ 可知该极值点为极小值点, 同时由于 $g(0) \neq 0$, 即 $x=0$ 不是方程 $g(x) = 0$ 的根, 所以 $f(x)$ 在 $[0, 1)$ 上惟一的极小点 θ 不可能为 0, 即 $0 < \theta < 1$ 。#

3.3.2 一类正则 LDPC 码的性能分析

现在研究一类码率为 $R = 1 - 1/n$ 的 (d, nd) -正则低密度删码 ($d \geq 3$ 和 $n \geq 2$ 均为整数), 为了给出 (3, 6) 和 (d, nd) -正则度分布的阈值之间的关系, 首先给出如下引理。

引理 3.3^[74]: 对于码率为 $R = 1 - 1/n$ 的 (d, nd) -正则低密度删码 ($d \geq 3, n \geq 2$) 有 $p^*(3, 3n) \geq p^*(d, nd)$ 。

引理 3.4: 设 $f(x) = (1-x)/(1-x^{r-1})^2$, 整数 $r \geq 6$, 则有在 $[0, 1)$ 上 $f''(x) \geq 0$ 且 $f(x)$ 在 $[0, 1)$ 上有惟一的极小点。

证明: 易见 $f(x)$ 的一阶导数为

$$f'(x) = \frac{2(r-1)(1-x)x^{r-2} - (1-x^{r-1})}{(1-x^{r-1})^3} \quad (3-39)$$

其二阶导数为

$$f''(x) = \frac{2(r-1)x^{r-3}}{(1-x^{r-1})^4} [(2r-1)x^{r-1} - (2r-3)x^r - rx + (r-2)] \quad (3-40)$$

考虑上式右边的后一项因式, 令

$$g(x) = (2r-1)x^{r-1} - (2r-3)x^r - rx + (r-2) \quad (3-41)$$

则有

$$g'(x) = (2r-1)(r-1)x^{r-2} - (2r-3)rx^{r-1} - r \quad (3-42)$$

设 $s(x) = g'(x)$, 则有

$$s'(x) = (r-1)x^{r-3} [(2r-1)(r-2) - r(2r-3)x] \quad (3-43)$$

令 $s'(x) = 0$ 得 $s(x)$ 的稳定点 $x_0 = (2r-1)(r-2)/(r(2r-3))$, 所以

$$x_0 = \frac{2r^2 - 5r + 2}{r(2r-3)} < \frac{2r^2 - 5r + 3}{r(2r-3)} = \frac{(2r-3)(r-1)}{r(2r-3)} = \frac{r-1}{r} \quad (3-44)$$

同时

$$s(x_0) = \left(\frac{(2r-1)(r-2)}{r(2r-3)} \right)^{r-2} (2r-1) - r < \left(\frac{r-1}{r} \right)^{r-2} (2r-1) - r \quad (3-45)$$

而 $((r-1)/r)^{r-2}$ ($r \geq 6$) 为单调递减函数且 $\lim_{r \rightarrow \infty} ((r-1)/r)^{r-2} = 1/e < 1/2$, $r=6$ 时 $(5/6)^4 < 1/2$, 故 $((r-1)/r)^{r-2} < 1/2$ ($r \geq 6$), 所以 $s(x_0) < (2r-1)/2 - r = -1/2 < 0$ 。结合 $s(0) = -r < 0$ 和 $s(1) = -r + 1 < 0$, 即对所有的 $x \in [0, 1]$ 有 $g'(x) = s(x) < 0$, 即 $g(x)$ 为 $[0, 1]$ 上的单调递减函数且 $g(1) = 0$, 从而对所有的 $x \in [0, 1]$ 有 $g(x) \geq 0$ 。因此对所有 $x \in [0, 1]$ 有 $f''(x) \geq 0$ 。于是 $f(x)$ 为 $[0, 1]$ 上的下凸函数, $f(x)$ 在 $[0, 1]$ 上有惟一的极小点。#

由引理 3.3 知对于给定的整数 n ($n \geq 2$), 在码率为 $R = 1 - 1/n$ 的 (d, nd) -正则度低密度纠删码中 ($d \geq 3$ 为整数), $(3, 3n)$ -正则纠删码的纠删性能最好。结合引理 3.4 和阈值 $p^*(\ell, r)$ 的定义可得如下结论。

定理 3.6: 对于码率为 $R = 1 - 1/n$ 的 (d, nd) -正则度低密度纠删码 ($d \geq 3, n \geq 2$), 必有

(a)、 $p^*(3, 6)/(1-R) < 1$, 即 $(3, 6)$ -正则低密度纠删码不是最优码;

(b)、 $(d, 2d)$ -正则度低密度纠删码都不是渐近最优码;

(c)、 $p^*(d, nd) \leq p^*(3, 6)$ 。

证明: (a)、对于 $(3, 6)$ -正则低密度纠删码, 其码率为 $R = 1 - 1/2 = 1/2$, 用计算机搜索的方法对高次不等式 $p(1-x^5)^2 < 1-x, x \in [0, 1]$ 求解得到 $p^*(3, 6) \approx 0.42944$, 故

$$p^*(3, 3n)/(1-R) < 0.43/0.5 = 0.86 < 1 \quad (3-46)$$

所以 $(3, 6)$ -正则低密度纠删码不是最优码;

(b)、由引理 3.3 知 $p^*(3, 6) \geq p^*(d, 2d)$, 因此有

$$\frac{p^*(d, 2d)}{1-R} \leq \frac{p^*(3,6)}{1/2} < 0.86 < 1 \quad (3-47)$$

从而说明这类 $(d, 2d)$ -正则度低密度删码都不是渐近最优码；

(c)、设 $f(x) = (1-x)/(1-x^5)^2$ ，由引理 3.4 知 $f(x)$ 在 $[0, 1)$ 上有唯一的极小点，并设此极小点 $x = \beta \in [0, 1)$ 。易见 $(1-\beta^5)^2 \leq (1-\beta^{3n-1})^2 (n \geq 2)$ ，从而有

$$f(\beta) = \frac{1-\beta}{(1-\beta^5)^2} \geq \frac{1-\beta}{(1-\beta^{3n-1})^2} \quad (3-48)$$

结合(3-33)式得

$$p^*(3,6) \geq p^*(3,3n) \quad (3-49)$$

又由引理 3.3 得

$$p^*(3, 3n) \geq p^*(d, nd) \quad (3-50)$$

所以 $p^*(3,6) \geq p^*(d, nd)$ 。#

由定理 3.6 可知， $(d, 2d)$ -正则低密度删码都不是渐近最优码。仿真结果也证实了结论的正确性。下图为 $(3,6)$ 、 $(4,8)$ 和 $(5,10)$ -正则删码在删码信道下的性能曲线图，其中选择的码长均为 25600。

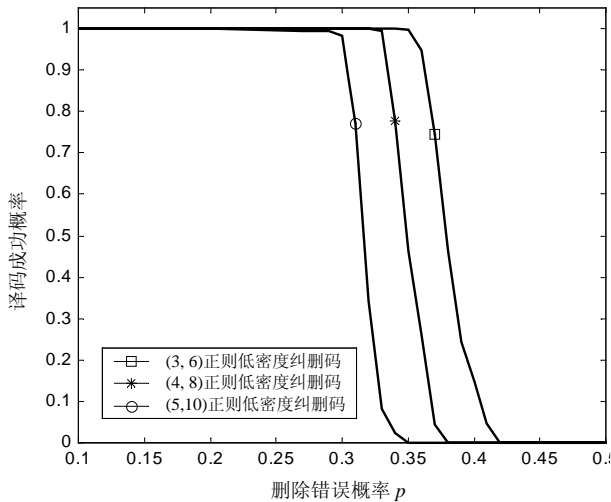


图 3.4 正则低密度删码的性能比较

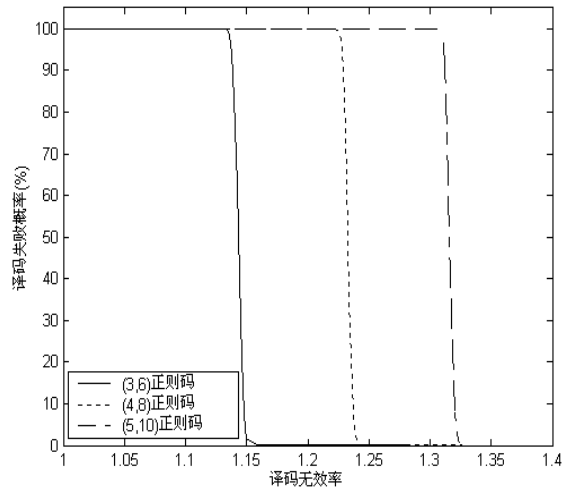


图 3.5 基于正则序列的低密度删码的性能比较

图 3.5 的水平坐标为译码无效率 (Decoding Inefficiency)，即为实现译码成功所需要付出的代价，通常用正确接收到的信息比特数 n' 与该低密度删码的信息位个数 k 的比值来表示。

事实上，正则低密度删码都不是渐近最优码，这一点可从构成低密度删码的二部图得到直观性解释。二部图的结点集包括信息比特结点集和校验比特结点集，对于每一个信息比特结点，其度数越大，它从相关联的校验比特结点得到信息越多，越能准确地判断它的正确值；而对每个校验比特结点情况则相反，校验结点的度数越小，它能反馈给其邻接的信息比特结点的信息越有价值。由于正则二部图不能很好地平衡这两种相反的要求，所以正则低密度删码都不是渐近最优码。

3.4 非正则 LDPC 码在删除信道下的性能

当二部图中的信息比特节点度数或校验比特节点度数不唯一时，相应地就得到了非正则 LDPC 码。相对于正则二部图，非正则二部图能更好、更灵活地平衡上面提到的两种相反的要求，基于非正则二部图的低密度纠错码中具有大度数的信息比特节点能很快地得到它的正确值，从而给某些关联的校验比特节点提供更有价值的信息，而这些校验比特节点又反过来可给小度数的信息比特节点提供更好的信息，如此反复进行，大度数的信息比特节点可促进算法收敛，因此非正则低密度纠错码在性能上要优于正则低密度纠错码。现有的研究已经说明，能够以逼近删除信道容量限的非正则低密度纠错码是存在的，下面介绍两种这样的非正则低密度纠错码的度序列分布函数。

3.4.1 Heavy-Tail/Poisson 度序列分布^[28]

Heavy-Tail/Poisson 度序列分布函数由 Luby 等人提出，采用此度序列分布构造的非正则 LDPC 码在删除信道下能够以任意接近信道容量的速率实现可靠传输，其度序列分布函数设计如下。

设要构造一个非正则二部图 B ，图中共有 k 个信息比特节点， βk 个校验比特节点。令 d 为一个正整数，并构造一个截止于整数 d 的谐函数 $H(d)$ ，即

$$H(d) = \sum_{i=1}^d 1/i \quad (3-51)$$

容易证明当 $d \rightarrow +\infty$ 时有 $H(d) \rightarrow \ln(d)$ 。

根据谐函数 $H(d)$ 即可设计二部图的左边度序列分布函数

$$\lambda(x) = \sum_{i=2}^{d+1} \lambda_i x^{i-1} \quad (3-52)$$

其中
$$\lambda_i = \frac{1}{H(d)(i-1)} \quad (3-53)$$

显然，该度序列分布函数满足 $\lambda(1) = 1$ ，同时左边信息比特节点的平均度数 a_l 为：

$$a_l = \frac{1}{\sum_{i=2}^{d+1} \lambda_i / i} = \frac{H(d)}{\sum_{i=2}^{d+1} 1/(i-1)i} = H(d)(d+1)/d \quad (3-54)$$

相应地右边校验比特节点的平均度数为

$$a_r = a_l / \beta \quad (3-55)$$

对右边度序列函数，将其设为满足 Poisson 分布，即

$$\rho(x) = \sum_i \rho_i x^{i-1} \quad (3-56)$$

其中
$$\rho_i = \frac{e^{-\alpha} \alpha^{i-1}}{(i-1)!}, \quad i \geq 1 \quad (3-57)$$

为保证右边校验比特节点的平均度数满足(3-55)式，(3-57)式中的 α 应满足

$$\frac{\alpha e^\alpha}{e^\alpha - 1} = a_r \quad (3-58)$$

注意到对右边度序列分布有：

$$\sum_{i=1}^{+\infty} \rho_i = \sum_{i=1}^{+\infty} \frac{e^{-\alpha} \alpha^{i-1}}{(i-1)!} = 1 \quad (3-59)$$

即仅当多项式 $\rho(x)$ 有无穷项时才能保证 $\rho(1) = 1$ ，故实际应用中需要进行截断，只取前有限项，再做归一化处理。下面引理说明了当 d 足够大时，所构造码是渐进最优的。

引理 3.5^[28]：若度序列函数满足式 (3-52) 和式 (3-56)，则对 $\forall x \in (0,1)$ 和 $p \leq \beta/(1+1/d)$ 恒有 $\rho(1-p\lambda(x)) > 1-x$ 。

下表给出了度序列函数如上面描述，码率为 $R = 1/2$ 的低密度删码的参数取值，其中 $\bar{p}^*(\lambda, \rho)$ 表示理论的最大容许损失上限。

表 3.1 满足 Heavy-Tail/Poisson 度分布的码参数(码率 $R = 1/2$)

d	a_r	α	$p^*(\lambda, \rho)$	$p^*(\lambda, \rho)/(1-R)$	$\bar{p}^*(\lambda, \rho)$	$p^*(\lambda, \rho)/\bar{p}^*(\lambda, \rho)$
8	5.9266	5.9105	0.45984	0.91968	0.49085	0.93682
16	7.0788	7.0729	0.47796	0.95592	0.49609	0.96345
27	8.0054	8.0027	0.48628	0.97256	0.49799	0.97648
47	9.0256	9.0243	0.49177	0.98354	0.49902	0.98547
79	10.007	10.007	0.49495	0.98990	0.49951	0.99087
132	10.996	10.996	0.49689	0.99378	0.49975	0.99427
221	12.000	12.000	0.49813	0.99626	0.49988	0.99650

根据 Heavy-Tail/Poisson 度序列分布函数构造的复损码通常也称做旋风码 (Tornado Code)，下图给出了旋风码和正则复损码的性能比较。

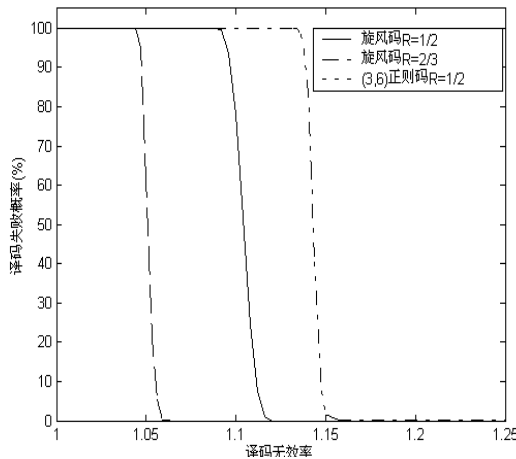


图 3.6 基于正则序列和 Tornado 序列的低密度删码的性能比较

3.4.2 右边正则度序列分布^[75]

不同于 Heavy-Tail/Poisson 度序列分布，文[75]提出了另一种能够任意逼近删除信道容量限的度序列分布函数，其构造如下：

给定正整数 $a_r \geq 3$ 和 $N \geq 2$ ，令

$$\rho(x) = x^{a_r-1} \quad (3-60)$$

$$\lambda(x) = \left(\sum_{k=1}^{N-1} \binom{\theta}{k} (-1)^{k+1} x^k \right) / \left(1 - \frac{N}{\theta} \binom{\theta}{N} (-1)^{N+1} \right) \quad (3-61)$$

其中 $\theta = 1/(a_r - 1)$, $\binom{\theta}{k}$ 定义如下:

$$\binom{\theta}{k} := \begin{cases} 1 & k = 0 \\ \frac{\theta(\theta-1)\cdots(\theta-k+1)}{k!} = (-1)^{k-1} \frac{\theta}{k} \left(1 - \frac{\theta}{k-1}\right) \cdots \left(1 - \frac{\theta}{2}\right) (1-\theta) & k > 0 \end{cases} \quad (3-62)$$

注意到下面函数的泰勒展开式为:

$$h(x) = 1 - (1-x)^\theta = \sum_{k=1}^{+\infty} \binom{\theta}{k} (-1)^{k+1} x^k \quad (3-63)$$

同时用数学归纳法可知有下面等式成立,

$$\sum_{k=1}^{N-1} \binom{\theta}{k} (-1)^{k+1} = \frac{1}{\theta} \left(\theta - N \binom{\theta}{N} (-1)^{N+1} \right) \quad (3-64)$$

因此度分布函数 $\lambda(x)$ 实际上就是函数 $h(x)$ 泰勒展开式的前有限项并做归一化处理。

再注意到下面等式:

$$\sum_{k=1}^{N-1} \binom{\theta}{k} \frac{(-1)^{k+1}}{k+1} = \frac{\theta - \binom{\theta}{N} (-1)^{N+1}}{\theta+1} \quad (3-65)$$

可以计算出左边信息比特节点的平均度数为:

$$\begin{aligned} a_l &= \frac{1}{\int_{x=0}^1 \lambda(x) dx} = \frac{\theta - N \binom{\theta}{N} (-1)^{N+1}}{\theta \sum_{k=1}^{N-1} \binom{\theta}{k} \frac{(-1)^{k+1}}{k+1}} \\ &= \frac{\theta - N \binom{\theta}{N} (-1)^{N+1}}{\theta \frac{\theta - \binom{\theta}{N} (-1)^{N+1}}{\theta+1}} = \frac{\theta+1}{\theta} \frac{\theta - N \binom{\theta}{N} (-1)^{N+1}}{\theta - \binom{\theta}{N} (-1)^{N+1}} \end{aligned} \quad (3-66)$$

因此, 码的设计码率为:

$$R = 1 - \frac{a_l}{a_r} = 1 - \frac{\frac{\theta+1}{\theta} \frac{\theta - N \binom{\theta}{N} (-1)^{N+1}}{\theta - \binom{\theta}{N} (-1)^{N+1}}}{1 + \frac{1}{\theta}} = 1 - \frac{\theta - N \binom{\theta}{N} (-1)^{N+1}}{\theta - \binom{\theta}{N} (-1)^{N+1}} \quad (3-67)$$

文[75]证明了通过增大右边校验比特节点度数 a_r 和参数 N ，具有该度序列分布的 LDPC 码能够以任意接近删除信道容量限的速率实现可靠传输，下表给出了不同设计码率时各设计参数和性能参数的取值。

表 3.2 满足右边正则度分布的码参数（设计码率分别为 2/3、1/2 和 1/3）

N	a_r	$1-R$	$p^*(\lambda, \rho)$	$p^*(\lambda, \rho)/(1-R)$	$\bar{p}^*(\lambda, \rho)$	$p^*(\lambda, \rho)/\bar{p}^*(\lambda, \rho)$
2	6	0.33333	0.20000	0.60000	0.29099	0.68731
3	7	0.31677	0.23611	0.74537	0.29714	0.82230
6	8	0.32886	0.28994	0.88166	0.31243	0.92801
11	9	0.33645	0.31551	0.93777	0.32690	0.96514
17	10	0.33357	0.32024	0.96001	0.32724	0.97860
27	11	0.33392	0.32558	0.97502	0.32984	0.98711
42	12	0.33381	0.32847	0.98401	0.33113	0.99197
64	13	0.33312	0.32963	0.98953	0.33134	0.99484
13	6	0.50090	0.48090	0.96007	0.49232	0.97679
29	7	0.50164	0.49287	0.98251	0.49759	0.99052
60	8	0.49965	0.49545	0.99159	0.49762	0.99563
125	9	0.49985	0.49784	0.99598	0.49885	0.99797
257	10	0.50000	0.49903	0.99805	0.49951	0.99904
523	11	0.50002	0.49954	0.99904	0.49977	0.99953
1058	12	0.49999	0.49975	0.99953	0.49986	0.99977
111	6	0.66677	0.66475	0.99698	0.66584	0.99837
349	7	0.66667	0.66603	0.99904	0.66636	0.99950
1077	8	0.66663	0.66642	0.99969	0.66653	0.99984
3298	9	0.66669	0.66662	0.99990	0.66665	0.99995

表中的码率并不严格等于设计码率，实际上在(3-67)式中， $\theta = 1/(a_r - 1)$ ，而 a_r 和 N 均为正整数，即该码率的取值不连续。故基于右边正则序列所构造的低密度删码的码率不能严格等于给定码率，只能通过增大 a_r 使所构造的删码的码率逼近给定码率；然而这种调整同样会增加基于级联稀疏二部图的低密度删码的译码复杂度。

3.5 基于改进型的右边正则度序列设计

3.5.1 改进型右边正则度序列设计

为了克服由右边正则度分布函数构造的复损码的码率不能够任意逼近设计码率的缺陷，我们提出了改进型的右边正则度分布函数。对所有实数 $x \in [0, 1]$ 和 $s \in [2, \infty)$ 令

$$f(x) = \sum_{k \geq 1} f_k x^k = \sum_{k \geq 1} \binom{\theta}{k} (-1)^{k+1} x^k \quad (3-69)$$

$$\tilde{f}(x, s) = \sum_{k=1}^{\lfloor s \rfloor - 1} f_k x^k + (s - \lfloor s \rfloor) f_{\lfloor s \rfloor} x^{\lfloor s \rfloor} = \sum_{k=1}^{\lfloor s \rfloor - 1} \binom{\theta}{k} (-1)^{k+1} x^k + (s - \lfloor s \rfloor) \binom{\theta}{\lfloor s \rfloor} (-1)^{\lfloor s \rfloor + 1} x^{\lfloor s \rfloor} \quad (3-70)$$

其中 $\lfloor s \rfloor$ 表示不超过 s 的最大整数， θ 的定义与前面相同。

取改进型右边正则序列为

$$\lambda(x) = \tilde{f}(x, s) / \tilde{f}(1, s) \quad (3-71)$$

$$\rho(x) = x^{a_r-1} \quad (3-72)$$

易见 $\lambda(0) = 0$ 和 $\lambda(1) = 1$ 。将(3-70)式代入(3-71)式有

$$\lambda(x) = \frac{\sum_{k=1}^{\lfloor s \rfloor - 1} \binom{\theta}{k} (-1)^{k+1} x^k + (s - \lfloor s \rfloor) \binom{\theta}{\lfloor s \rfloor} (-1)^{\lfloor s \rfloor + 1} x^{\lfloor s \rfloor}}{1 - \frac{\lfloor s \rfloor}{\theta} \binom{\theta}{\lfloor s \rfloor} (-1)^{\lfloor s \rfloor + 1} + (s - \lfloor s \rfloor) \binom{\theta}{\lfloor s \rfloor} (-1)^{\lfloor s \rfloor + 1}} \quad (3-73)$$

容易计算出，具有改进型右边正则度分布的 LDPC 码的码率 R 为

$$\begin{aligned} R &= 1 - \int_0^1 \rho(x) dx / \int_0^1 \lambda(x) dx \\ &= 1 - \frac{1 - \frac{\lfloor s \rfloor}{\theta} \binom{\theta}{\lfloor s \rfloor} (-1)^{\lfloor s \rfloor + 1} + (s - \lfloor s \rfloor) \binom{\theta}{\lfloor s \rfloor} (-1)^{\lfloor s \rfloor + 1}}{1 - \frac{1}{\theta} \binom{\theta}{\lfloor s \rfloor} (-1)^{\lfloor s \rfloor + 1} + \frac{\theta + 1}{\theta} (s - \lfloor s \rfloor) \binom{\theta}{\lfloor s \rfloor} (-1)^{\lfloor s \rfloor + 1}} \frac{1}{\lfloor s \rfloor + 1} \end{aligned} \quad (3-74)$$

实际上，将(3-73)式同(3-61)式进行比较可以发现，当 s 不是整数时，新的右边度序列分布函数比原来多了一个余项并重新做了归一化处理，当 s 取整数即 $s = N$ 时，余项恰好为 0，改进型度序列分布函数就退化为最初的右边正则度序列分布函数。通过调整 s 的具体取值，就可以使对应码的码率严格等于设计码率。

3.5.2 改进型右边正则度序列的性能分析

为了说明改进型右边正则序列是渐近似最优的，令

$$R_f(x, s) = \tilde{f}(x, s) / \tilde{f}(1, s) \quad (3-75)$$

$$G_f(s) = \int_0^1 \tilde{f}(x, s) dx \quad (3-76)$$

$$H_f(s) = \int_0^1 R_f(x, s) dx = G_f(s) / \tilde{f}(1, s) \quad (3-77)$$

引理 3.6: 设 $f(x)$ 、 $R_f(x, s)$ 、 $G_f(s)$ 和 $H_f(s)$ 如上定义，则有(a)、对 $\forall x \in [0, 1]$ ， $R_f(x, s)$ 关于 s 为 $[2, \infty)$ 上的单调递减函数；(b)、 $H_f(s)$ 为 $[2, \infty)$ 上单调递减连续函数。

证明: (a)、对 $\forall x \in [0, 1]$ 和 $\forall s_1, s_2 \in [2, \infty)$ ，若 $s_2 \leq s_1$ ，则有

$$\begin{aligned} R_f(x, s_1) - R_f(x, s_2) &= \tilde{f}(x, s_1) / \tilde{f}(1, s_1) - \tilde{f}(x, s_2) / \tilde{f}(1, s_2) \\ &\leq \frac{1}{\tilde{f}(1, s_1) \tilde{f}(1, s_2)} \left[\sum_{i=2}^{\lfloor s_1 \rfloor - \lfloor s_2 \rfloor} \sum_{k=1}^{\lfloor s_2 \rfloor} f_k f_{\lfloor s_2 \rfloor + i} (x^{\lfloor s_2 \rfloor + i} - x^k) \right. \\ &\quad + (s_2 - \lfloor s_2 \rfloor) \sum_{i=2}^{\lfloor s_1 \rfloor - \lfloor s_2 \rfloor} f_{\lfloor s_2 \rfloor + 1} f_{\lfloor s_2 \rfloor + i} (x^{\lfloor s_2 \rfloor + i} - x^{\lfloor s_2 \rfloor + 1}) \\ &\quad \left. + (s_1 - \lfloor s_1 \rfloor) \sum_{k=1}^{\lfloor s_2 \rfloor} f_k f_{\lfloor s_1 \rfloor + 1} (x^{\lfloor s_1 \rfloor + 1} - x^k) \right] \end{aligned}$$

$$+ (s_1 - \lfloor s_1 \rfloor)(s_2 - \lfloor s_2 \rfloor) f_{\lfloor s_1 \rfloor + 1} f_{\lfloor s_2 \rfloor + 1} (x^{\lfloor s_1 \rfloor + 1} - x^{\lfloor s_2 \rfloor + 1}) \quad (3-78)$$

对 $\forall x \in [0, 1]$ 和任意正整数 k_1 和 k_2 , 若 $k_1 \geq k_2$, 则必有 $x^{k_1} - x^{k_2} \leq 0$ 。结合 $f_k = \binom{\alpha}{k} (-1)^{k+1} > 0$ ($k=1, 2, \dots$)、 $\tilde{f}(1, s_1) > 0$ 和 $\tilde{f}(1, s_2) > 0$ 得 $R_f(x, s_1) - R_f(x, s_2) \leq 0$, 即 $R_f(x, s)$ 关于 s 为 $s \in [2, \infty)$ 上的单调递减函数;

(b)、由(a)易知 $H_f(s)$ 为 $[2, \infty)$ 上的单调递减函数。下面证明函数

$$G_f(s) = \int_0^1 \tilde{f}(x, s) dx = \sum_{k=1}^{\lfloor s \rfloor - 1} \frac{f_k}{k+1} + (s - \lfloor s \rfloor) \frac{f_{\lfloor s \rfloor}}{\lfloor s \rfloor + 1} \quad (3-79)$$

为 $[2, \infty)$ 上的连续函数。易见对实数 $s \in [1, \infty)$ 和 $s \neq \lfloor s \rfloor$, $G_f(s)$ 为 $[2, \infty)$ 上的连续函数。

对于充分小的 $\delta > 0$, 可知

$$\lim_{\delta \rightarrow 0^+} (G_f(\lfloor s \rfloor + \delta) - G_f(\lfloor s \rfloor)) = \lim_{\delta \rightarrow 0^+} \delta \frac{f_{\lfloor s \rfloor}}{\lfloor s \rfloor + 1} = 0 \quad (3-80)$$

$$\lim_{\delta \rightarrow 0^+} (G_f(\lfloor s \rfloor) - G_f(\lfloor s \rfloor - \delta)) = \lim_{\delta \rightarrow 0^+} \delta \frac{f_{\lfloor s \rfloor - 1}}{\lfloor s \rfloor} = 0 \quad (3-81)$$

于是, $G_f(s)$ 在 $s = \lfloor s \rfloor$ 为左连续和右连续的, 而且左极限和右极限相等。因此 $G_f(s)$ 在 $s = \lfloor s \rfloor$ 连续。故 $G_f(s)$ 为 $[2, \infty)$ 上的连续函数。类似可证 $\tilde{f}(1, s)$ 关于 s 为 $[2, \infty)$ 上的连续函数。所以, $H_f(s)$ 为 $[2, \infty)$ 上的连续函数。#

引理 3.7^[75]: 对 $\forall 0 < \theta \leq 1/2$ 和整数 $N \geq 2$, 存在正常数 c , 满足

$$\frac{c\theta}{N^{\theta+1}} \leq \binom{\theta}{N} (-1)^{N+1} \leq \frac{\theta}{N^{\theta+1}} \quad (3-82)$$

定理 3.7: 改进型的右边正则序列为是渐近似最优的。

证明: 对于满足(3-71)式和(3-72)式的改进型右边正则序列 $\lambda(x)$ 和 $\rho(x)$, 首先计算对所有 $x \in (0, p)$ 使得 $p\lambda(1 - \rho(1 - x)) < x$ 成立的最大的 p 。令

$$\Delta = 1 - \frac{\lfloor s \rfloor}{\theta} \binom{\theta}{\lfloor s \rfloor} (-1)^{\lfloor s \rfloor + 1} + (s - \lfloor s \rfloor) \binom{\theta}{\lfloor s \rfloor} (-1)^{\lfloor s \rfloor + 1} \quad (3-83)$$

注意到对于所有实数 $s \in [2, \infty)$ 有

$$\binom{\theta}{\lfloor s \rfloor} (-1)^{\lfloor s \rfloor + 1} > 0, \quad \theta = 1/(a_r - 1) \quad (3-84)$$

$$1 - (1 - x)^\theta = \sum_{k=1}^{\infty} \binom{\theta}{k} (-1)^{k+1} x^k \quad (3-85)$$

从而对所有 $x \in (0, p)$ 有

$$p\lambda(1 - \rho(1 - x)) \leq \frac{p}{\Delta} \left[\sum_{k=1}^{\lfloor s \rfloor - 1} \binom{\theta}{k} (-1)^{k+1} [1 - \rho(1 - x)]^k + \binom{\theta}{\lfloor s \rfloor} (-1)^{\lfloor s \rfloor + 1} [1 - \rho(1 - x)]^{\lfloor s \rfloor} \right]$$

$$\begin{aligned} &\leq \frac{p}{\Delta} \sum_{k=1}^{\infty} \binom{\theta}{k} (-1)^{k+1} (1-\rho(1-x))^k = \frac{p}{\Delta} \left\{ 1 - [1 - (1-\rho(1-x))]^{\theta} \right\} \\ &= \frac{p}{\Delta} \left\{ 1 - [(1-x)^{1/\theta}]^{\theta} \right\} = \frac{p}{\Delta} x \end{aligned} \quad (3-86)$$

考虑到 $p\lambda(1-\rho(1-x)) < x$, 有

$$p \leq \frac{1}{1 - \frac{\lfloor s \rfloor}{\theta} \binom{\theta}{\lfloor s \rfloor} (-1)^{\lfloor s \rfloor+1} + (s - \lfloor s \rfloor) \binom{\theta}{\lfloor s \rfloor} (-1)^{\lfloor s \rfloor+1}} \quad (3-87)$$

结合引理 3.7 可得

$$\frac{p}{1-R} = 1 - \frac{1}{\theta} \binom{\theta}{\lfloor s \rfloor} (-1)^{\lfloor s \rfloor+1} + \frac{\theta+1}{\theta} (s - \lfloor s \rfloor) \binom{\theta}{\lfloor s \rfloor} (-1)^{\lfloor s \rfloor+1} \frac{1}{\lfloor s \rfloor+1} \geq 1 - \frac{1}{(\lfloor s \rfloor)^{\theta+1}} \quad (3-88)$$

注意到 $\theta > 0$, 即 $\theta+1 > 1$, 故当 $s \rightarrow +\infty$ 时(3-88)式变为

$$p/(1-R) \geq 1 \quad (3-89)$$

而根据信道编码定理可知, 码率的上限即为信道容量, 于是有

$$R \leq 1-p \quad (3-90)$$

整理后得到

$$p/(1-R) \leq 1 \quad (3-89)$$

结合(3-89)式和(3-91)式可知:

$$\lim_{s \rightarrow +\infty} p/(1-R) = 1 \quad (3-92)$$

这就证明了改进型右边正则序列是渐近似最优的。#

下表给出了不同设计码率时改进型右边正则序列的各设计参数和性能参数的取值。

表 3.3 满足改进型右边正则度分布的码参数 (设计码率分别为 2/3、1/2 和 1/3)

a_r	s	$1-R$	$p^*(\lambda, \rho)$	$p^*(\lambda, \rho)/(1-R)$
6	110.833	2/3	0.665260	0.997890
7	348.979	2/3	0.666188	0.999282
8	1077.836	2/3	0.666504	0.999756
9	3296.383	2/3	0.666611	0.999917
6	12.873	1/2	0.488574	0.977148
7	28.397	1/2	0.494072	0.988144
8	60.311	1/2	0.497021	0.994042
9	125.315	1/2	0.498499	0.996998
10	256.982	1/2	0.499242	0.998484
11	522.797	1/2	0.499617	0.999234
12	1058.302	1/2	0.499806	0.999612
6	2.0	1/3	0.280000	0.840000
7	3.698	1/3	0.308179	0.924537
8	6.377	1/3	0.313220	0.939660
9	10.537	1/3	0.320161	0.960483
10	16.939	1/3	0.324701	0.974103
11	26.744	1/3	0.327520	0.982560
12	41.656	1/3	0.329447	0.988341

下面两图分别为采用改进型右边正则序列构造的低密度删码与采用正则序列、Tornado 序列和原右边正则序列构造的低密度删码的性能比较。

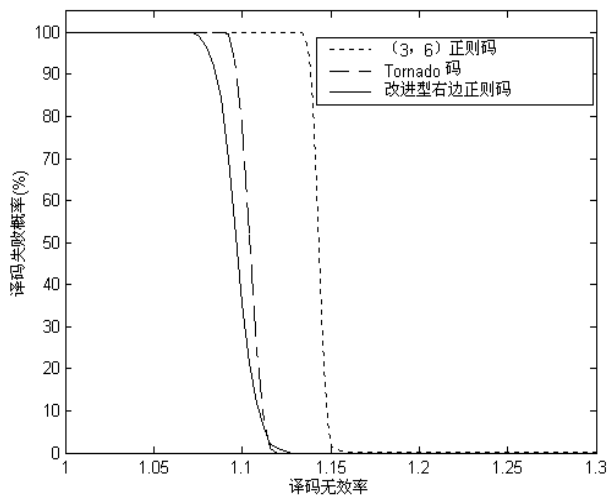


图 3.7 基于(3,6)正则序列、Tornado 序列和改进型右边正则序列的低密度删码的性能比较

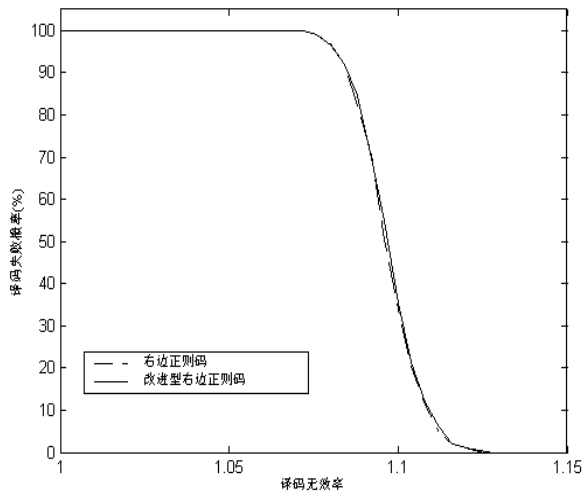


图 3.8 基于右边正则序列和改进型右边正则序列的低密度删码的性能比较

注：图 3.4~图 3.8 中的低密度删码均指级联型低密度删码

3.6 本章小结

本章介绍了删码的基本概念及其发展；分析了正则 LDPC 码在删除信道下的理论性能；证明了 $(d, 2d)$ -正则码都不是渐进最优码并进一步直观地解释了所有的正则 LDPC 码不是渐进最优码的原因；介绍了几种非正则的度序列分布函数并对相应 LDPC 码在删除信道下的性能进行了仿真；基于右边正则度序列分布函数提出了一种改进型的右边正则度序列分布函数并从理论上证明了用改进的度分布函数构造的低密度删码也是渐进最优的，仿真结果表明，改进型的右边正则度序列分布函数克服了原度分布函数不能设计任意码率的低密度删码的缺点，同时复杂度并没有任何增加，在性能上也没有任何损失。

第四章 LDPC 码的围长研究

本章首先介绍了几种现有的具有较大围长的 LDPC 码设计方法, 在此基础上, 提出了一种新的构造具有较大围长 LDPC 码的方法, 经检测和仿真表明, 采用该方法设计的 LDPC 码具有较大的围长, 并且在高斯信道下的纠错性能不差于相同参数、随机构造的 LDPC 码, 在高信噪比时甚至优于相同参数的随机码; 最后提出了一种 PEG 算法的修正算法, 给出了本章总结。

4.1 现有的几种围长设计方法

第二章已经给出 LDPC 码围长的定义, 分析了围长对码的性能的影响并推导了一个简单的正则 LDPC 码围长的上限。为了尽量减小 Tanner 图中环的存在对相应 LDPC 码在和积译码算法下所得到性能的影响, 人们在进行编码设计时往往希望对应的 Tanner 图中不存在长度较小的环, 同时一些研究学者基于代数方法和启发式搜索方法, 也提出了一些具有较大围长的 LDPC 码构造方法。研究表明通过增大 LDPC 码的围长在一定程度上可以改善码的纠错性能。下面介绍几种现有的码构造方法。

4.1.1 一种基于RS码的代数构造方法^[46]

令 C 为有限域 $GF(q)$ 上的一个 $(n, 2)$ RS 码, 由 RS 码的性质可知该码集合中共有 q^2 个码字且最小汉明距离为 $n-1$ 。下面根据码 C 构造一个具有 q^2 个变量节点和 qn 个校验节点的 Tanner 图。

令 Tanner 图中的每一个变量节点对应码 C 中的一个码字, 将 qn 个校验节点平均分为 n 组, 每组中的 q 个节点对应有限域 $GF(q)$ 中的 q 个元素; 节点间之边的连接规则如下: 设一个变量节点对应的码字为 (x_1, x_2, \dots, x_n) , 则该变量节点共与 n 个校验节点相连, 依次为校验节点集的第 1 组中对应于元素 x_1 的校验节点、第 2 组中对应于 x_2 的校验节点、……第 n 组中对应于 x_n 的校验节点。这样, 根据该 Tanner 图构造就得到一个长度为 q^2 的 (n, q) 正则 LDPC 码。

下面考虑 Tanner 图中环的情况。假设其中存在一个长度为 4 的环, 这就意味着 $(n, 2)$ RS 码中有两个码字 (对应于长度为 4 的环经过的两个变量节点) 的某两个分量 (对应于环经过的两个校验节点所在组的编号) 是相同的, 于是这两个码字之间的汉明距离至多为 $n-2$, 与该 RS 码的最小汉明距离为 $n-1$ 矛盾, 因此该 Tanner 图对应的 LDPC 码的围长至少是 6。

设所构造 LDPC 码的围长为 $2l$, 当 $l=3$ 时, 应用第二章给出的围长上限可得到:

$$1+n(q-1) \leq q^2 \quad (4-1)$$

整理后得到

$$n \leq q+1 \quad (4-2)$$

于是当选择 $n = q+1$ 时就得到最优的 Tanner 图，此时围长为 6。

当 $n < q+1$ 时，所得到的 Tanner 图未必是最优的，其围长有可能大于 6。很容易可以看出，对 $n \neq 2$ 的情况，所得 Tanner 图的围长就是 6，而当 $n = 2$ 时，围长为 8。因为 Tanner 图中存在一个长度为 6 的环意味着码 C 中有三个码字两两存在某个分量相同，而这对长度为 2 的码是不可能的。

4.1.2 一种基于矩阵分裂的代数构造方法^[50]

文[50]基于矩阵分裂的思想，提出了一种具有较大围长的正则 LDPC 码校验矩阵的构造方法，具体方法如下。

设 H 为一个 (λ, ρ) 正则 LDPC 码的校验矩阵，维数为 $\lambda p \times \rho p$ ，其中 p 为一个素数。将 H 分裂为 $\lambda \times \rho$ 个维数为 $p \times p$ 的小方阵，每一个方阵为一个单位阵或者单位阵的行循环移位，设第 i 行、第 j 列的方阵是由一个单位阵经过各行循环左移 $s_{i,j}$ 位得到的，则各方阵循环移位参数的确定过程为：令 a 和 b 为有限域 $GF(p)$ 上的两个非零元素， $a \neq b$ ，且 a 和 b 关于有限域中乘法运算的级数 (Order) 为 λ 和 ρ ，于是 λ 和 ρ 必为 $p-1$ 的因子，则第 i 行、第 j 列的方阵的循环移位参数由下面等式确定。

$$s_{i,j} = b^{(i-1)} a^{(j-1)} \quad (4-3)$$

这样得到的校验矩阵每列均有 λ 个“1”，每行均有 ρ 个“1”，相应的码就是一个长度为 $n = \rho p$ 的 (λ, ρ) 正则 LDPC 码。容易看出这样构造的校验矩阵的秩至多为 $\lambda p - \lambda - 1$ ，故相应 LDPC 码的码率 R 也不能严格等于设计码率 $1 - \lambda/\rho$ ，而要略高些。

根据有限域的相关理论可知这样得到的 LDPC 码具有较大的围长，下表给出了采用该方法构造的几个 (3,5) 正则 LDPC 码的围长。

表 4.1 (3,5) 正则 LDPC 码的围长

码长 $n = 5p$	p	R	围长 g
155	31	0.4129	8
305	61	0.4066	10
755	151	0.4424	10
905	181	0.4022	12
1055	211	0.4019	12
1205	241	0.4017	12
1355	271	0.4015	12
2105	421	0.4010	12
3305	661	0.4006	12
5105	1021	0.4004	12
6455	1291	0.4003	12
11555	2311	0.4001	12

可以看出，采用该方法构造的 LDPC 码完全避免了长度为 4 和 6 的小环的出现，但遗憾的是这样得到的码的围长至多为 12，不能随着码长的增加而增加。

4.1.3 一种启发式搜索构造方法——PEG构造方法^[60]

前面介绍的两种方法均为基于代数的构造方法，而且只能适用于正则 LDPC 码的构造，文[60]采用“步步最优”的策略，提出了一种有效的构造具有较大围长的 LDPC 码的方法——渐进边增长（PEG, Progressive Edge-Growth）算法。该算法通过依次在已有 Tanner 图上添加边来构造最终的 Tanner 图，每次添加边时都尽可能减少对已有 Tanner 图的围长的影响。它不但适用于正则 LDPC 码的构造，也适用于非正则 LDPC 码的构造。为便于描述，用 $\mathbf{H} = \{h_{ij}\}$ ($0 \leq i < m$, $0 \leq j < n$) 表示码的校验矩阵，用 (\mathbf{V}, \mathbf{E}) 表示所构造的 Tanner 图。其中 $\mathbf{V} = \mathbf{V}_c \cup \mathbf{V}_s$ 为节点集， $\mathbf{V}_c = \{c_0, c_1, \dots, c_{m-1}\}$ 为校验节点集， m 为校验节点个数； $\mathbf{V}_s = \{s_0, s_1, \dots, s_{n-1}\}$ 为变量节点集， n 为变量节点个数； $\mathbf{E} = \mathbf{V}_c \times \mathbf{V}_s$ 表示 Tanner 图中边的集合，边 $(c_i, s_j) \in \mathbf{E}$ 当且仅当 $h_{ij} \neq 0$, $0 \leq i < m$, $0 \leq j < n$ 。

若已知边的度分布函数 $\lambda(x) = \sum_i \lambda_i x^{i-1}$ 和 $\rho(x) = \sum_i \rho_i x^{i-1}$ ，即可求出变量节点和校验节点的度分布函数 $\hat{\lambda}(x) = \sum_i \hat{\lambda}_i x^i$ 和 $\hat{\rho}(x) = \sum_i \hat{\rho}_i x^i$ ，并按照该分布随机给各个变量节点和校验节点分配度数，记为 $D_s = \{d_{s_0}, d_{s_1}, \dots, d_{s_{n-1}}\}$ 和 $D_c = \{d_{c_0}, d_{c_1}, \dots, d_{c_{m-1}}\}$ ，其中 d_{s_j} (d_{c_i}) 表示变量节点 s_j (校验节点 c_i) 的度数，通常集合 D_s 和 D_c 按照升序排列，即有 $d_{s_0} \leq d_{s_1} \leq \dots \leq d_{s_{n-1}}$ 和 $d_{c_0} \leq d_{c_1} \leq \dots \leq d_{c_{m-1}}$ ；同时，将边的集合根据变量节点集 \mathbf{V}_s 表示为 $\mathbf{E} = \mathbf{E}_{s_0} \cup \mathbf{E}_{s_1} \cup \dots \cup \mathbf{E}_{s_{n-1}}$ ，其中 $\mathbf{E}_{s_j} = \{E_{s_j}^k, 0 \leq k \leq d_{s_j} - 1\}$ 表示所有与变量节点 s_j 相连的边构成的集合， $E_{s_j}^k$ 为与变量节点 s_j 相连的第 k 条边。定义 $\mathbf{N}_{s_j}^l$ 为当前 Tanner 中所有与变量节点 s_j 之间的最短路径长度（所经过的边的个数）不超过 $2l+1$ 的校验节点构成的集合（如图 4.1 所示），并用 $\overline{\mathbf{N}}_{s_j}^l = \mathbf{V}_c \setminus \mathbf{N}_{s_j}^l$ 表示校验节点集中除去 $\mathbf{N}_{s_j}^l$ 后剩下的集合，则整个 PEG 算法可以描述如下：

```

for (j=0; j<n; j++)
{
  for(k=0; k<dsj; k++)
  {
    if(k==0)
    {
      添加边  $E_{s_j}^0 \rightarrow (c_i, s_j)$ ，其中  $c_i$  为当前 Tanner 图
      中度数最小的校验节点
    }
    else
    {
      添加  $E_{s_j}^k \rightarrow (c_i, s_j)$ ，若对  $\forall l \in N$ ，恒有
       $\overline{\mathbf{N}}_{s_j}^l \neq \Phi$ ，则  $c_i \in \overline{\mathbf{N}}_{s_j}^{l_{\max}}$ ；若  $\exists l \in N$ ，使得
       $\overline{\mathbf{N}}_{s_j}^l \neq \Phi$  而  $\overline{\mathbf{N}}_{s_j}^{l+1} = \Phi$ ， $c_i$  为集合  $\overline{\mathbf{N}}_{s_j}^{l+1}$  中
      度数最小的校验节点。
    }
  }
}

```

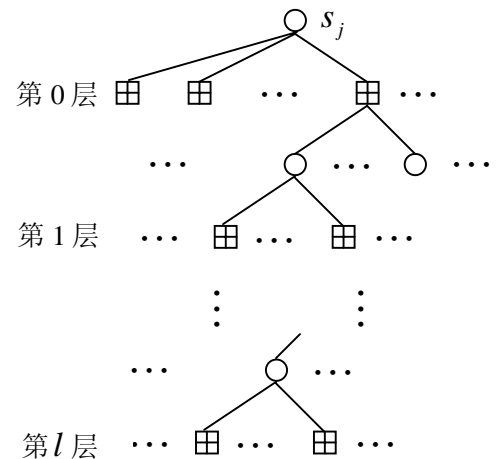


图 4.1 节点集 $\mathbf{N}_{s_j}^l$ 的结构示意图

文[60]对采用 PEG 算法构造的 LDPC 码的性能进行了分析, 给出了所构造码的围长上下限和最小距离下限, 仿真结果表明, PEG 算法具有较好的实用性, 采用该算法可以构造出围长为 8、码长为 1008 的 (3,6) 正则 LDPC 码。

4.2 一种具有较大围长的正则 LDPC 码构造方法

考虑下面长为 16 的 (2, 4) 正则 LDPC 码对应的 Tanner 图。

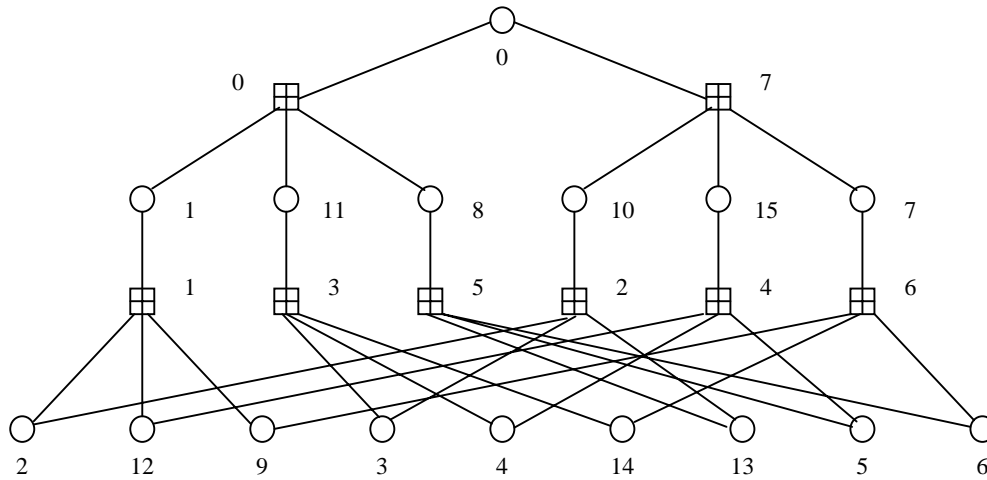


图 4.2 (16,2,4)LDPC 码的 Tanner 图表示

显然该 Tanner 图中所有环的最小长度为 8, 因此对应 LDPC 码的围长也为 8。按图 4.2 将其中的变量结点和校验结点依次编号, 可以得到对应 LDPC 码的校验矩阵如下:

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

图 4.3 (16,2,4)LDPC 码的校验矩阵表示

上面矩阵很有规律, 可以看作是由两个行重为 2、维数为 8×8 的循环方阵拼接而成。因此可以猜想, 采用某些有规律的矩阵合并成校验矩阵, 这样生成的 LDPC 码很可能会具有较大的围长。或者说, 将 LDPC 码的校验矩阵分裂为若干个子矩阵, 然后每个子矩阵再按照某种规律构造, 就有可能避免小环的出现。

类似于 4.1.2 节所述方法，这里也采用矩阵分裂的思想。设要构造一个长为 n ($n = \rho U$ $U \in N$) 的 (λ, ρ) 正则 LDPC 码，我们将该码的校验矩阵分裂为 (λ, ρ) 个 $U \times U$ 的子矩阵，即：

$$\mathbf{H} = \begin{pmatrix} \mathbf{H}_0 \\ \mathbf{H}_1 \\ \vdots \\ \mathbf{H}_{\lambda-1} \end{pmatrix} = \begin{pmatrix} \mathbf{H}_{0,0} & \mathbf{H}_{0,1} & \cdots & \mathbf{H}_{0,\rho-1} \\ \mathbf{H}_{1,0} & \mathbf{H}_{1,1} & \cdots & \mathbf{H}_{1,\rho-1} \\ \vdots & \vdots & \cdots & \vdots \\ \mathbf{H}_{\lambda-1,0} & \mathbf{H}_{\lambda-1,1} & \cdots & \mathbf{H}_{\lambda-1,\rho-1} \end{pmatrix} \quad (4-4)$$

其中每个子矩阵 $\mathbf{H}_{i,j} = \mathbf{I}(s_{i,j})$ ($0 \leq i < \lambda$, $0 \leq j < \rho$) 均为一个单位阵或者单位阵的循环移位， $s_{i,j}$ 表示该单位阵的各行循环右移的位数。显然，这样构造的校验矩阵也不可能为满秩，至多为 $\lambda\rho - \lambda - 1$ 。

为便于描述，用 $\mathbf{S} = (s_{i,j})$ 表示由各个子方阵的循环移位参数组成的矩阵，用 (I, J, i, j) 表示校验矩阵中的元素，其中 (I, J) 为该元素所属的子矩阵的坐标， (i, j) 为该元素在它所属的子矩阵中的坐标。

称 Tanner 图中每个变量结点参与的所有环的最小长度为该变量结点的环长，则显然相应 LDPC 码的围长就等于各个变量结点的环长的最小值。将 n 个变量结点分为 ρ 组，每一组变量结点对应一列子矩阵，则考虑到各个子矩阵的循环特性，有如下定理成立。

定理 4.1: 属于同组的变量结点具有相同的环长。

证明: 设任意两个同组的变量结点 x 和 y ，分别对应一列子矩阵的第 x 列和第 y 列，且 $y - x = d_{\text{mod } U}$ ，其环长分别为 $C(x)$ 和 $C(y)$ ，并设变量结点 x 的最小环的路径为：

$$\begin{array}{ccc} (I_1, J_1, i_1, j_1) & \leftrightarrow & (I_1, J_2, i_1, j_2) \\ \updownarrow & & \updownarrow \\ & & (I_2, J_2, i_2, j_2) \leftrightarrow (I_2, J_3, i_2, j_3) \\ & & \vdots \\ (I_{C(x)/2}, J_1, i_{C(x)/2}, j_1) & \longleftrightarrow & (I_{C(x)/2}, J_{C(x)/2}, i_{C(x)/2}, j_{C(x)/2}) \end{array}$$

图 4.4 变量节点 x 的环路示意图

根据各个子矩阵的循环特性，可以找到另一个环的路径为：

$$\begin{array}{ccc} (I_1, J_1, (i_1 + d)_{\text{mod } U}, (j_1 + d)_{\text{mod } U}) & \leftrightarrow & (I_1, J_2, (i_1 + d)_{\text{mod } U}, (j_2 + d)_{\text{mod } U}) \\ \updownarrow & & \vdots \\ (I_{C(x)/2}, J_1, (i_{C(x)/2} + d)_{\text{mod } U}, (j_1 + d)_{\text{mod } U}) & \leftrightarrow & (I_{C(x)/2}, J_{C(x)/2}, (i_{C(x)/2} + d)_{\text{mod } U}, (j_{C(x)/2} + d)_{\text{mod } U}) \end{array}$$

图 4.5 变量节点 y 的环路示意图

显然该环路长度为 $C(x)$ 且经过变量结点 y ，故有

$$C(x) \geq C(y) \quad (4-5)$$

同理可得

$$C(x) \leq C(y) \quad (4-6)$$

综合上面两式，有

$$C(x) = C(y) \quad (4-7)$$

即对任意两个同组的变量结点，它们的环长均相同，证毕。#

由定理4.1可知，按照上述方法构造的校验矩阵所对应的LDPC码，所有变量结点的环长至多有 ρ 种情况，因此对这样构造的矩阵只需要分别从各组中抽取一个变量结点，然后只对这 ρ 个变量接点进行检测，即可确定整个码的围长。

首先来考虑4环的情况。如果一个LDPC码含有4环，则它所对应的校验矩阵中必然有4个“1”处于某个矩形的四个顶点，该环路路径可表示为：

$$\begin{array}{ccc} (I_1, J_1, i, (i+s_{I_1 J_1})_{\text{mod}U}) & \leftrightarrow & (I_1, J_2, i, (i+s_{I_1 J_2})_{\text{mod}U}) \\ \updownarrow & & \updownarrow \\ (I_2, J_1, (i+s_{I_1 J_2} - s_{I_2 J_2})_{\text{mod}U}, (i+s_{I_1 J_2} - s_{I_2 J_2} + s_{I_2 J_1})_{\text{mod}U}) & \leftrightarrow & (I_2, J_2, (i+s_{I_1 J_2} - s_{I_2 J_2})_{\text{mod}U}, (i+s_{I_1 J_2})_{\text{mod}U}) \end{array}$$

图 4.5 4 环的环路示意图

显然，应有

$$i + s_{I_1 J_1} = i + s_{I_1 J_2} - s_{I_2 J_2} + s_{I_2 J_1} \pmod{U} \quad (4-8)$$

化简后即得：

$$s_{I_1 J_1} + s_{I_2 J_2} - s_{I_1 J_2} - s_{I_2 J_1} = 0 \pmod{U} \quad (4-9)$$

至此，可以得到如下定理。

定理4.2：按照(4-4)式所示的矩阵分裂方法构造的矩阵所对应的LDPC码不含有长为4的环的充要条件为下式成立：

$$s_{I_1 J_1} + s_{I_2 J_2} - s_{I_1 J_2} - s_{I_2 J_1} \neq 0 \pmod{U} \quad (I_1 \neq I_2 \in \{0, 1, \lambda-1\}, J_1 \neq J_2 \in \{0, \dots, \rho-1\}) \quad (4-10)$$

该定理的正确性从前面的描述中即可得知，这里不再赘述。

由定理4.2很容易可以得到下面推论。

推论4.1：按照(4-4)式所示的矩阵分裂方法构造的矩阵所对应的LDPC码不含有长为 $2l$ 的环的充要条件为下式成立：

$$\sum_{k=0}^{l-1} (s_{I_k J_k} - s_{I_{(k+1)\text{mod}l}}) \neq 0_{\text{mod}U} \quad (I_k \neq I_{(k+1)\text{mod}l} \in \{0, 1, \lambda-1\}, J_k \neq J_{(k+1)\text{mod}l} \in \{0, \dots, \rho-1\}) \quad (4-11)$$

在编码设计时，可以首先确定所构造LDPC码的设计围长，然后根据上面的定理和推论列出相应的不等约束，进而寻找满足这些不等约束的参数即可。

考虑到对校验矩阵只进行行交换和列交换并不会改变它所对应的因子图的拓扑结构，可以通过交换校验矩阵的各行来将某一列子矩阵变为单位阵，也可以通过交换校验矩阵的各列来将某一行子矩阵变为单位阵。从这个角度讲，可以直接将某一行和某

一列（这里选择第 0 行和第 0 列）子矩阵的循环移位参数设定为 0，而对剩余的 $(\lambda-1)(\rho-1)$ 个子矩阵的循环移位参数的选取进行求解，这样并不影响结果的正确性，而且能使所要确定的变量从 $\lambda \times \rho$ 个减为 $(\lambda-1)(\rho-1)$ 个。

在进行参数选择时，可以根据上面的分析和设计的围长列出各参数所应满足的约束方程，然后再寻找满足这些约束方程的参数取值。然而，由于这些约束方程均为不等约束，因而无法采用一般的方程组求解法；而如果采用穷举的方法去遍历各个参数的所有可能组合，继而从中找出满足约束的一组，搜索的范围将有 $U^{(\lambda-1)(\rho-1)}$ ，这样即使 U 的取值较小（如 10^2 ），总的搜索范围也将很大（对 (3,6) 正则码为 10^{20} ），因而无法实现。

为了实现参数的快速选取，可以采用如下的逐参试探算法：

- (1)、令 $s_{i,0} = 0 (i = 0, 1, \dots, \lambda-1)$ 及 $s_{0,j} = 0 (j = 1, \dots, \rho-1)$ ；
- (2)、随机在 $\{0, \dots, U-1\}$ 中选取 $s_{1,1}$ 的取值，然后判断 $s_{1,1}$ 是否满足给定的不等约束（不包括含有未定参数的约束方程），若满足则确定 $s_{1,1}$ 的取值，否则重新执行(2)；
- (3)、按照(2)的方法依次确定剩余子矩阵的循环移位参数。

按照上面算法，每个参数至多需要 U 次试探，这样总共的试探次数至多为 $(\lambda-1)(\rho-1)U$ ，远远小于整个搜索空间 $U^{(\lambda-1)(\rho-1)}$ 。采用计算机编程试探可以很快实现。

显然，这样确定参数的方法并没有遍历各个参数的所有可能组合，因而不能保证在任何存在满足约束的参数组合的情况下必然找到该参数组合。下面对该算法的实用性和可行性进行分析，不失一般性，这里以 (3,6) 正则码为例。

由于该算法采用逐个确定参数的方法，显然最后确定的参数受到的约束是最多的，定义 $N(l)$ 为考虑消除 Tanner 图中长度为 $2l$ 的环时最后一个参数受到的约束方程个数，则有：

$$N(2) = C_5^1 \times C_2^1 = 10 \quad (4-12)$$

$$N(3) = C_5^1 \times C_2^1 \times C_4^1 = 40 \quad (4-13)$$

$$N(4) = C_5^1 \times C_2^1 \times C_5^1 \times C_2^1 \times C_4^1 = 400 \quad (4-14)$$

由于各个约束方程均为不等约束，每个不等约束只能限制参数不能取某个（在取模的情况下可能有两个）特定的值，因此所有不等约束限制参数所不能取的值的个数至多为约束方程数目的两倍。考虑到我们所要构造的 LDPC 码的码长， U 的取值一般在 100 左右，因此消除六环一般都是可行的。

取 $U = 168$ ，按照上面方法构造长度为 1008 的 (3,6) 正则 LDPC 码，通过计算机搜索，得到个子方阵的循环参数为：

$$\mathbf{S} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 69 & 12 & 29 & 19 & 80 \\ 0 & 95 & 54 & 6 & 58 & 59 \end{pmatrix} \quad (4-15)$$

通过检测发现，该LDPC码的围长为10，为了保证所构造码的码率严格等于1/2，可以从生成的校验矩阵中删去2个“1”。该码在AWGN信道下的纠错性能如下图所示，图中的另外两条曲线分别为相同长度、随机构造、不消除4环和消除4环的(3,6)正则LDPC码的性能曲线，其中girth表示围长，ave表示所有变量节点的平均环长。

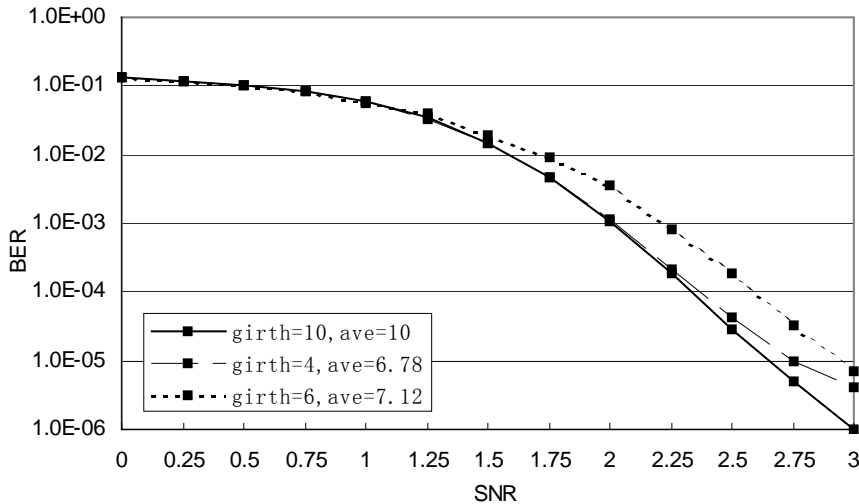


图 4.6 长为 1008 的(3,6)正则 LDPC 码的性能图

对于长度为1008的(3,6)正则LDPC码， $U = 168$ 不是素数，不能采用4.1.2节所述的方法构造；文[60]采用PEG算法所构造的长度为1008的(3,6)正则LDPC码的围长为8，平均环长为9.66，稍劣于上面构造的LDPC码，因此该方法用于正则LDPC码的构造时要优于其它的构造方法。

通过分析发现，采用该方法构造的正则LDPC码也和4.1.2节所述方法一样，其围长存在一个上限，下面进行详细介绍。考虑一个维数为 $2U \times 3U$ 的矩阵，将其分裂为6个维数为 $U \times U$ 的子方阵，每个方阵均为单位阵或单位阵的行循环移位，则可以得到一个行重为3、列重为2的矩阵。不失一般性，令第一行和第一列子方阵均为单位阵，其余两个子方阵的行右循环移位参数分别为 $s_{1,1}$ 和 $s_{1,2}$ ，则不论 $s_{1,1}$ 、 $s_{1,2}$ 如何取值，该矩阵始终存在下图所示的12环。

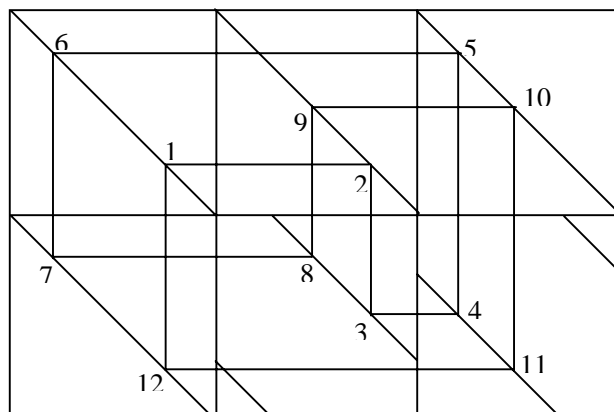


图 4.7 12环在矩阵上的环路示意图

将上图环上各个的非零元素依次编号，并令编号为 1 的元素坐标为 $(0,0,x,x)$ ，则环上各节点的坐标如表 4.2 所示。

表 4.2 12 环上非零元素的坐标

编号	坐标	编号	坐标
1	$(0,0,x,x)$	7	$(1,0,x-s_{1,1}+s_{1,2},x-s_{1,1}+s_{1,2})$
2	$(0,1,x,x)$	8	$(1,1,x-s_{1,1}+s_{1,2},x+s_{1,2})$
3	$(1,1,x-s_{1,1},x)$	9	$(0,1,x-s_{1,1}+s_{1,2},x+s_{1,2})$
4	$(1,2,x-s_{1,1},x-s_{1,1}+s_{1,2})$	10	$(0,2,x+s_{1,2},x+s_{1,2})$
5	$(0,2,x-s_{1,1}+s_{1,2},x-s_{1,1}+s_{1,2})$	11	$(1,2,x,x+s_{1,2})$
6	$(0,0,x-s_{1,1}+s_{1,2},x-s_{1,1}+s_{1,2})$	12	$(1,0,x,x)$

因此，若采用上面所述方法构造 (λ, ρ) 正则 LDPC 码，只要满足 $\lambda \geq 2$ ， $\rho \geq 2$ 且 $\lambda + \rho \geq 5$ ，相应的校验矩阵中也就必然包含图 4.7 所示的子矩阵或其转置矩阵，于是所得到的 LDPC 码的围长也就必然不可能超过 12。

4.3 PEG 算法研究

4.1.3 节介绍的 PEG 构造算法中，所有变量节点是根据其度数按照升序依次连边的，但文献中没有给出变量节点必须根据度数按照升序依次进行连边的依据。本文对该问题进行了研究，取码长 $n = 1008$ ，度分布函数如(4-15)和(4-16)式所示：

$$\lambda(x) = 0.21x^4 + 0.25x^3 + 0.25x^2 + 0.29x \quad (4-16)$$

$$\rho(x) = x^5 \quad (4-17)$$

在运用 PEG 算法进行该码的构造过程中，分别按照各变量节点度数的升序和降序来确定其连边次序，得到 LDPC 码的平均环长分别为 8.13 和 9.59，这说明变量节点根据其度数按照降序依次连边能够获得更大的平均环长。

对给定码长 n 和度分布函数 $\lambda(x)$ 和 $\rho(x)$ ，应用该算法可构造出具有较大围长的 LDPC 码，但所构造的码仅能满足变量节点的度分布。虽然目前已发现优于现有非正则度分布的右边正则度分布，这并不能说明最优的度分布函数一定是右边正则的，而且在实际中可能需要校验节点也满足一定度分布。这里拟将已有算法加以修正，使所构造码能严格满足给定度分布。令向量 $\mathbf{K} = (\kappa_0, \kappa_1, \dots, \kappa_{\max})$ 表示校验节点的度分布，其中 κ_i 表示度数不小于 i （包括 i ）的校验节点个数， \max 为校验节点最大度数，并设与 \mathbf{K} 相同维数的向量 $\mathbf{T} = (t_0, t_1, \dots, t_{\max})$ ，将其初始化为 $\mathbf{T} = (m, 0, \dots, 0)$ ，则改进后的 PEG 算法流程如下：

```

for (j=0; j<n; j++)
{ for(k=0; k<dsj; k++)
  { if(k==0)
    {
      添加边  $E_{s_j}^0 \rightarrow (c_i, s_j)$ , 其中  $c_i$  为当前 Tanner 图中任意一个校验节点, 且满足
       $t_{d+1} < \kappa_{d+1}$ ,  $d$  为该校验节点的度数;
       $t_{d+1}$  增加 1;
    }else
    {
      添加  $E_{s_j}^k \rightarrow (c_i, s_j)$ , 若对  $\forall l \in N$ , 恒有  $\overline{\mathbf{N}}_{s_j}^l \neq \Phi$ , 则  $c_i \in \overline{\mathbf{N}}_{s_j}^{l_{\max}}$ , 然后令
       $d = 0$ ; 若  $\exists l \in N$ , 使得  $\overline{\mathbf{N}}_{s_j}^l \neq \Phi$  而  $\overline{\mathbf{N}}_{s_j}^{l+1} = \Phi$ ,  $c_i$  为集合  $\overline{\mathbf{N}}_{s_j}^{l-1}$  中的校验节
      点且满足  $t_{d+1} < \kappa_{d+1}$ ,  $d$  为该校验节点的度数; 若  $\overline{\mathbf{N}}_{s_j}^{l-1}$  中不存在这样的校验节
      点, 则将  $l$  自减 1, 继续寻找满足条件的校验节点, 直至找到为止;
       $t_{d+1}$  增加 1;
    }
  }
}

```

4.4 本章小结

本章介绍了几种现有的构造具有较大围长的 LDPC 码方法, 并提出了一种新的具有较大围长的正则 LDPC 码构造方法。仿真表明用该方法构造的码在 AWGN 信道下要优于随机构造的码, 在有些参数取值情况下得到的码甚至还可能大于采用 PEG 算法构造码的平均环长; 同时本章还对 PEG 算法进行了研究, 优化了变量节点的连边次序, 并给出了一种改进的算法, 使改进后的 PEG 算法能够构造具有较大围长、同时完全满足给定度序列分布的 LDPC 码。

第五章 LDPC 码的快速编码研究

本章首先简要给出了 LDPC 码快速编码的原理及其实现,接着介绍了两种能够达到线性编码的码类,指出了它们与可快速编码的 LDPC 码之间的关系,然后根据快速编码的原理提出了两类能够达到线性编码的 LDPC 码构造方法并对其在 AWGN 信道下的纠错性能进行了仿真,最后给出了对 LDPC 码线性编码的几点设想并对本章进行了总结。

5.1 LDPC 码的快速编码

第二章已经介绍了 LDPC 码的译码,在 Tanner 图上描述时,LDPC 码的译码实质上就是在变量节点和校验节点之间传递信息,所传递信息的量与整个 Tanner 图中边的总数呈正比,而由于 LDPC 码的 Tanner 图是和其低密度的校验矩阵一一对应的,很容易可以看出 LDPC 码的译码具有线性的复杂度,因此相对来讲,LDPC 码的编码反而成为其实用化过程中的瓶颈。

由于 LDPC 码属于线性分组码,它的编码过程通常采用线性分组码的通用编码方法。设 LDPC 码的校验矩阵为 $\mathbf{H}_{(n-k) \times n}$, 对应的生成矩阵为 $\mathbf{G}_{k \times n}$, 则对任意的输入信息序列 $\mathbf{m}_{1 \times k}$, 相应的码字序列 $\mathbf{x}_{1 \times n}$ 应为:

$$\mathbf{x} = \mathbf{m} \cdot \mathbf{G} = \mathbf{m} \cdot (\mathbf{g}_0, \mathbf{g}_1, \dots, \mathbf{g}_{n-1}) \quad (5-1)$$

因此整个编码的复杂度取决于(5-1)式的运算复杂度,对于二元的情况,其中主要的操作包括与运算和异或运算(模二加),设生成矩阵 \mathbf{G} 的平均列重为 m , 则整个编码过程中大约需要 mn 次与运算, $(m-1)n$ 次异或运算。尽管 LDPC 码的校验矩阵是非常稀疏的,但它的生成矩阵却并不稀疏,通常 m 与 n 的比值是 $[0,1]$ 之间不可忽略的数,这使得其编码复杂度往往与其码长的平方呈正比。因此,相对于 LDPC 码的译码来说,它的编码反而具有较高的复杂度,这一点与 Turbo 码不同(其编译码复杂度均为线性)。

文[76][77]对 LDPC 码的快速编码问题进行了研究,指出虽然 LDPC 码的编码复杂度与其码长的平方呈正比,但采用适当的编码算法,相应的系数可以取得很小。文[77]给出实现 LDPC 码的快速编码的方法,即通过行列的置换将码的校验矩阵变换成下三角或近似下三角形式,如下图所示。

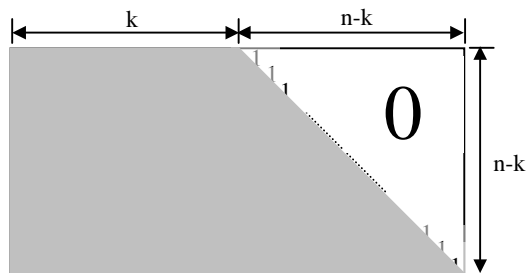


图 5.1 LDPC 码校验矩阵的下三角形式

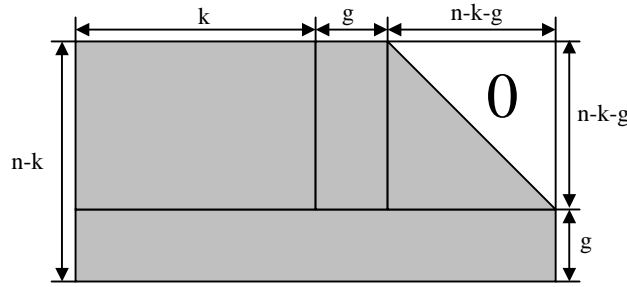


图 5.2 LDPC 码校验矩阵的近似下三角形式

若 LDPC 码的校验矩阵具有如图 5.1 所示的下三角形式，在图中所示的斜线上为全“1”，而其余的“1”均在斜线的左边，则可以采用迭代的方式进行编码。设码字向量为 $\mathbf{x} \in F^n$ ，将其分为两部分，即信息位向量 $\mathbf{s} \in F^k$ 和校验位向量 $\mathbf{p} \in F^{n-k}$ ，亦即 $\mathbf{x} = (\mathbf{s}, \mathbf{p})$ ，则该码的编码过程可具体描述为：

1)、直接将信息比特的值赋给信息位向量 \mathbf{s} ；

2)、采用后向迭代确定所有校验位的值，即对所有的 $l \in [0, n-k-1]$ ，从小到大依次计算下式，

$$p_l = \sum_{i=0}^{k-1} h_{i,j} \cdot s_j + \sum_{i=0}^{l-1} h_{i,j+k} \cdot p_j \quad (5-2)$$

其中 $h_{i,j}$ 表示校验矩阵第 i 行、第 j 列上的元素。

实际上，该编码过程就是在从上到下依次利用校验矩阵的各行校验约束关系。对于每一个校验约束关系，其中涉及的变量除斜线上的“1”所对应的校验位外，其余的变量要么是信息位，要么就是前面已经求出的校验位，也就是说，该校验约束关系中只有一个未知变量，因此可以很容易求得相应校验位的值。

设校验矩阵经过变换成为下三角形式后的平均行重为 m ，则整个编码约需要 $m(n-k)$ 次与运算， $(m-1)(n-k)$ 次异或运算，当 m 相对于 n 可以看作很小的常数时，该编码方法就具有线性的复杂度。

若 LDPC 码的校验矩阵具有如图 5.2 所示的近似下三角形式，在图中所示的斜线上仍为全“1”，其余的“1”均在斜线的左边和下边，则该码的校验位向量可分为两部分，即 $\mathbf{p} = (\mathbf{p}_a, \mathbf{p}_b) \in F^{n-k}$ ，其中 $\mathbf{p}_a \in F^g$ ， $\mathbf{p}_b \in F^{n-k-g}$ ，对于 \mathbf{p}_a 部分，需要采用其它方式进行求解，复杂度一般为 $O(n^2)$ ，而对于 \mathbf{p}_b 部分，仍然可以采用上面介绍的迭代方式进行编码，当校验矩阵足够稀疏时，可以看作只有线性的复杂度。

使校验矩阵具有图 5.1 所示的下三角形式或图 5.2 所示的近似下三角形式只能使编码可以直接在校验矩阵上完成，而真正保证编码具有线性复杂度的是经过整形变换之后的校验矩阵的密度。一般事先给定的码校验矩阵是具有低密度的，所以人们希望在将校验矩阵变换成下三角形式或近似下三角形式的过程中仍能保持该矩阵的低密度，而显然采用高斯消元法对矩阵进行行变换是做不到这一点的，因此通常校验矩阵的下三角化或近似下三角化都是只通过行列的置换实现的。文[77]指出只通过行列的置换将一个矩阵

变换成下三角形式是一件非常困难的事，有时甚至是不可实现的；同时，文中还给出了几种能够将已知校验矩阵通过行列置换变为近似下三角形式的贪婪算法，对 LDPC 码的快速编码实现有一定的实际意义。

5.2 两类可快速编码的码类

5.2.1 Tornado 码及其编码

文[28]提出的 Tornado 码也是基于稀疏矩阵的码，而且它和普通的 LDPC 码具有非常相似的图结构。两者的不同之处在于 Tornado 码对应的二部图中的校验结点不再表示校验约束关系，而表示码字中的校验比特位，原先的变量结点也不再表示码字中的所有比特位，而仅表示信息比特位，图 5.3 给出了一个 Tornado 码的二部图表示。可以看出，图 5.3 所示的 Tornado 码和图 5.4 所示的(8,2,4)LDPC 码的图结构非常相似，但它的码长已变成了 12，其中 x_1, x_2, \dots, x_8 为信息位， y_1, y_2, \dots, y_4 为校验位，校验位的值由所有与它相连的信息位的值模二相加得到。由于矩阵 H 的稀疏性，Tornado 码为计算校验位的值所需要的操作数与信息位的个数呈正比，因而可以达到编码的线性复杂度。

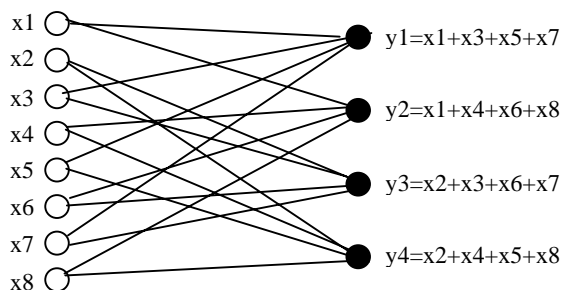


图 5.3 Tornado 码二部图表示

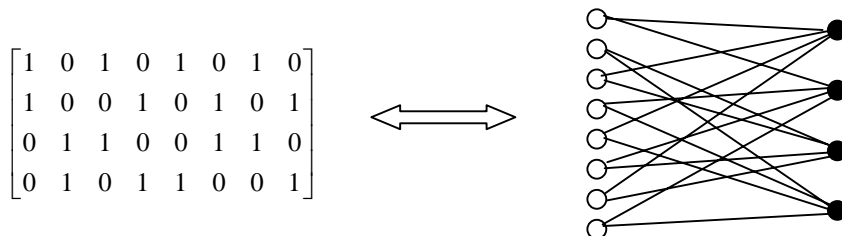


图 5.4 (8,2,4) LDPC 码的校验矩阵表示及二部图表示

设图 5.4 所示的(8,2,4)LDPC 码的校验矩阵为 H ，很容易可以发现，图 5.3 所示 Tornado 码的校验矩阵实际上可以写成 $H_T = [H|I]$ ，显然，它也是个稀疏矩阵，且具有下三角形式，同时在矩阵中斜线的正下方没有“1”。

然存在重量为 1 的列(矩阵的最后一列), 亦即相应的二部图中必然存在度为 1 的变量结点, 这一点可能会与给定的度序列要求相矛盾, 但考虑到图 5.1 所示的校验矩阵只有最后一列的列重必须为 1, 可以忽略不计, 这样构造既基本满足给定度序列分布, 又能够实现线性编码的 LDPC 码还是可行的。

我们采用直接构造法生成了一个长度为 1008 的 LDPC 码, 该码的校验矩阵的行重基本上均为 6, 列重基本上均为 3, 因而可以近似看作是一个(3, 6)的正则 LDPC 码。我们对其在高斯信道下对其性能进行了仿真, 并将它与一个随机生成、同样长度的(3, 6)正则 LDPC 码进行比较, 结果如图 5.7 所示。

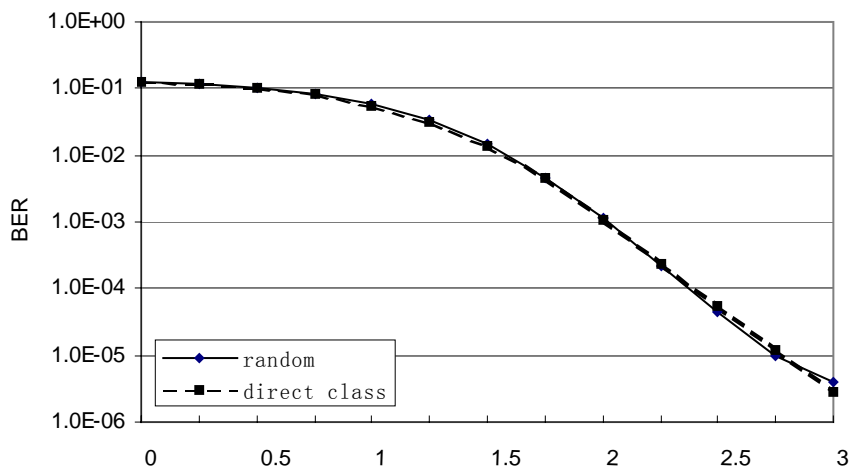


图 5.7 直接构造法生成的 LDPC 码的性能

从图 5.7 可以看出, 两个 LDPC 码在高斯信道下的纠错性能非常接近, 这表明具有下三角形校验矩阵的 LDPC 码作为整个 LDPC 码集合中的一个子集, 其中任意一个码的纠错性能都以很大概率不差于整个 LDPC 码集合的平均纠错性能, 而能够实现线性编码则使其在实际应用中更加具有优势。因此, 在实际应用中使用既具有线性的编译码复杂度, 又具有较强的纠错能力的 LDPC 码是可能的。经检测, 图 5.7 中的两个 LDPC 码的围长均为 4, 这也是直接构造法生成的 LDPC 码在纠错性能上不能超过相同参数、随机生成的 LDPC 码的主要原因。

5.3.2 删除构造法

LDPC 码的和积译码算法(SPA)是基于译码过程中所传递的消息间的独立性假设推导出来的, 而在实际应用中该假设往往是不能成立的, LDPC 码对应的图结构中围长的大小和环长分布对该码的纠错性能有着重要的影响, 为此, 人们基于代数构造和启发式搜索提出了许多具有较大围长和较好环长分布的 LDPC 码的设计方法, 以减弱消息的非独立性对译码性能的影响。因此如果采用直接构造法生成能够线性编码的 LDPC

码，这些现有的消除小环的码设计方法往往不能借用，即直接构造法无法用已有方法来生成既能实现线性编码，又能消除小环的 LDPC 码。

为此，可以先利用现有的其它方法生成具有较大围长的 LDPC 码，然后在对其校验矩阵进行行列置换的同时，删去适当位置上的“1”，从而将该校验矩阵变换成为图 3 所示的形式。由于行列的置换不会影响相应的二部图的拓扑结构，而删“1”操作也只能使图中已有的环的环长增大，所以用删除构造法得到的 LDPC 码一定可以具有不小于原 LDPC 码的围长，从而得到既能实现线性编码，又具有较大围长的 LDPC 码。在变换过程中，为尽量保持 LDPC 码原有的度序列分布，应使删去的“1”尽可能的少，而这一点的最优化实现是非常困难的，下面给出一种近优的实现算法——**删除算法**。

删除算法的具体实现步骤如下：

1)、按照其它某种方法先生成满足给定度序列分布和环长要求的 LDPC 码的校验矩阵；

2)、设置变量 row 和 col 表示矩阵中某一个“1”所在的行号和列号，均赋初值为 0，并设置变量 weight_row 和 weight_col 表示第 row 行和第 col 列的重量，分别赋初值为 0 和 MAX(一个足够大的正整数)；

3)、遍历矩阵所有的列，如果当前列的列重小于 weight_col，则执行 4)，若当前列的列重等于 weight_col，则执行 5)，否则什么都不执行；

4)、将 col 更新为当前列的列号，将 weight_col 更新为第 col 列的列重，然后在该列中所有“1”所在的行中寻找重量最大的行，记为 row，并将 weight_row 更新为第 row 行的行重，返回 3)继续遍历下一列；

5)、观察当前列中所有的“1”所在行的行重，如果某一个“1”所在行的行重大于 weight_row，则将 row 和 col 更新为该“1”所在的行号和列号，并将 weight_row 更新为该“1”所在行的行重，返回 3)继续遍历下一列；

6)、如果所有的列均遍历完毕，则通过行列置换将第 row 行移到矩阵的最后一行，将第 col 列置换到矩阵的最后一列，并将最后一列中除最后一行上的“1”以外的“1”全部删去；

7)、将去掉最后一行和最后一列的矩阵作为新的矩阵，重新执行 2)至 6)步直至新的矩阵为空，需要注意的是统计行重和列重时仅计算当前的新的矩阵的行重和列重，进行行列置换时则应对原先的整个校验矩阵进行行列置换。

可以发现，该算法在每一次从某一列中删去多余的“1”时都选取重量最小的列，使每次被删去的“1”最少；同时该列中保留的“1”所在行的行重又尽可能的大，使下一步新的矩阵中的“1”的数量尽可能的少，从而使以后各步中“1”被删除的概率尽可能的小，通过这种“步步最优”的策略使最终整个校验矩阵被删除的“1”的个数尽可能少。

对第四章中得到的围长为 10、长度为 1008 的(3, 6)正则 LDPC 码，这里用删除算法对其校验矩阵进行一定的变换，最终得到具有下三角形式的矩阵，即新的 LDPC 码

的校验矩阵。经过统计，在删除算法的实现过程中删去的“1”共有 40 个，与矩阵中原有的“1”的总数 3024 相比，仅为 1.3%，因此可以保证新得到的 LDPC 码与原码具有近似的度序列分布。新的 LDPC 码在 AWGN 信道下的纠错性能如图 5.8 所示。

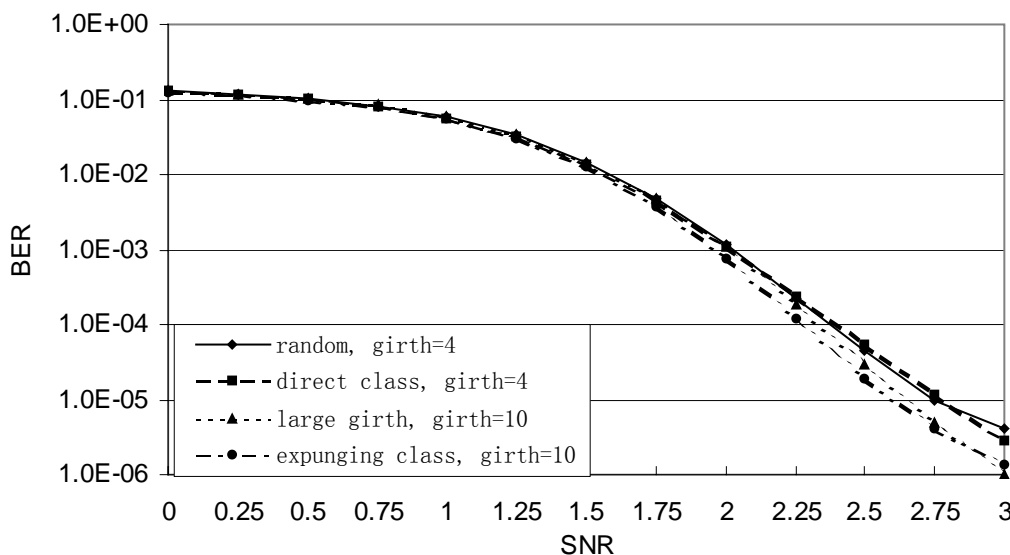


图 5.8 删除构造法生成的 LDPC 码的性能

从图中可以看出，应用删除构造法得到的新码不但具有线性的编码复杂度，而且在性能上与原码非常接近，均优于含有长度较小的环的 LDPC 码。这说明既具有线性的编译码复杂度，同时具有强大的纠错能力的 LDPC 码是存在的。

5.4 关于线性编码的几点设想

回忆第二章介绍的 LDPC 码在删除信道下的译码算法，整个译码过程上就是一个在码所对应的 Tanner 图上去边的过程。通过分析可以发现，其本质就是利用校验矩阵所限定的比特间约束关系，根据正确接收的比特值来求解被删除比特值的过程，而上面提到的 LDPC 码快速编码方法也是利用校验矩阵所限定的比特间约束关系，根据信息位比特的值来求解校验位比特的值，两者非常相似。由此可以猜想，对于删除信道下的某一个 LDPC 码，若某一次传输过程中，被删除的比特位恰好全部是校验位，则相应的译码过程就相当于一个编码过程，也就是说，在对系统形式的 LDPC 码进行编码时，把所有的校验位看作删除的比特位，就可以采用删除信道下的译码操作来实现编码，或者说 LDPC 码的快速编码实际上基于其线性复杂度的译码算法实现的。问题的关键在于对这样特殊的错误图样，整个过程能否成功进行直至所有校验位都被求出。

在删除信道下 LDPC 码的译码过程中，每一步操作都需要在剩余二部图上寻找一个度数为 1 的校验结点，当剩余二部图上不存在度数为 1 的校验结点时，译码就宣告失

败，经过分析发现，LDPC 码的截止集（Stopping Set）是影响其在删除信道下性能的一个主要因素。所谓截止集就是这样一些变量节点组成的集合，在它们和所有与其相连的校验节点构成的二部图子图中，不存在度唯一的校验节点。集合中变量节点的个数成为该截止集的阶数（Size）。下图给出了几个截止集的例子。

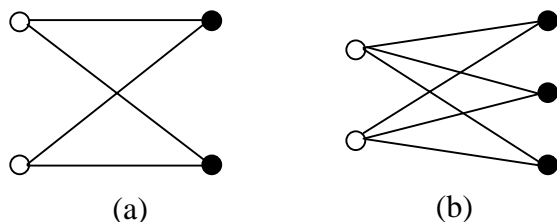


图 5.9 截止集示意图

因此，为了保证采用 LDPC 码在删除信道下的译码算法能够成功实现线性编码，必须保证由所有校验位对应的变量节点集中不存在截止集，而将校验矩阵限制为下三角形式就恰好可以确保这一点。

综上，实现 LDPC 码线性编码的一个途径就是利用其在删除信道下的译码算法，而删除信道下 LDPC 码的译码算法不能保证对任意 LDPC 码都能成功实现，因此需要将码的形式加以限制（限制其校验矩阵满足下三角形式）。进一步可以猜想，除了在删除信道下，LDPC 码在 BSC 信道和 AWGN 信道下也有相应的译码算法，并且它们也都具有线性的复杂度，因此也可以采用这些译码算法来实现 LDPC 码的线性编码。考虑到 AWGN 信道下的和积译码算法需要软输出，涉及到实数运算，因此较为复杂，不适于 LDPC 码的编码应用；而 BSC 信道下的译码算法属于硬判决译码，实现较为简单，因此尽管 LDPC 码在 BSC 信道下的译码算法也不能保证译码的必然成功，考虑如何将 LDPC 码的校验矩阵加以限制，使其采用 BSC 信道下的译码算法能够成功进行校验位的求解，或许是实现 LDPC 码线性编码的另一个途径。

5.5 本章小结

本章介绍了 LDPC 码的快速编码及其实现，指出了 Tornado 码和重复累积码能够实现线性编码的原因，然后提出了两类能够实现线性编码的 LDPC 码类设计方法，对其在 AWGN 信道下的纠错性能进行了仿真，最后给出了作者关于 LDPC 码线性编码的几点设想。

结 束 语

信道编码理论及技术作为现代通信系统必不可少的关键技术，近几十年在 Shannon 信道编码定理的指引下已经经历了飞速的发展并取得了大量的研究成果。目前，基于图模型的低密度校验码，使人们以低复杂度的删除错误恢复算法实现了逼近删除信道容量和可靠通信的理想，它的研究不仅具有重要的学术价值和理论指导意义，更具有强烈的应用背景和非常显著的经济效益。

作者结合国家自然科学基金(编号 69972035)、重庆市/信息产业部移动通信技术重点实验室开放课题基金和“适合中国的 4G 及后 3G 中的关键技术研究”(与三星公司合作项目)等相关课题的科研项目，对 LDPC 码进行了前期研究，在此基础上主要围绕 LDPC 码的编码设计对相关理论进行了深入的研究，提出了一些新的理论观点，获得了一些研究成果。本文所做的贡献主要包括：

1、系统地分析和总结了 LDPC 码基于图模型的编译码思想。推导了 LDPC 码在不同信道下的译码算法，对影响 LDPC 码性能的两个关键性因素——度序列分布函数和围长——进行了深入分析，指出了优化的方向；

2、分析了 LDPC 码在删除信道下的理论性能，推导了正则度分布的阈值，从理论上证明了基于 $(d, 2d)$ -正则度序列的低密度删删码都不是渐近最优码 $(d \geq 3)$ ，研究了非正则 LDPC 码的度序列分布，基于对右边正则序列的详细分析提出了一种改进型右边正则序列，并证明了此序列为渐近似最优的，同时对基于正则序列、Heavy-Tail/Poisson 序列、右边正则序列和改进型右边正则序列的级联型低密度删删码进行了模拟仿真及性能分析。

3、通过对 LDPC 码的围长分析和对现有构造具有较大围长的 LDPC 码方法的研究，提出了一种新的具有较大围长的正则 LDPC 码构造方法，并对用该方法构造的 LDPC 码在 AWGN 信道下的纠错性能进行了仿真，提出了一种 PEG 算法的修正算法，使改进后的 PEG 算法能够构造出具有较大围长且严格满足给定度序列分布的 LDPC 码；

4、通过对 LDPC 码快速编码的研究现状的深入分析，指出了 Tornado 码和重复累积码能够达到线性编码的原因及其与可线性编码的 LDPC 码之间的关系，提出了两种能够实现线性编码的 LDPC 码构造方法并对其在 AWGN 信道下的纠错性能进行了仿真，提出了关于 LDPC 码线性编码实现的几点设想。

图模型上 LDPC 码的研究内容和涉及学科范围相当广泛，还有许多理论和实际应用问题没有得到解决。本文的研究只涉及了其中个别问题，着重于围绕 LDPC 码的编

码设计及其相关的理论问题，这些研究成果还有待于做进一步的扩展研究，主要包括：

1. 从理论和实用的角度，如何利用图论和最优化理论的知识寻找获得最优度分布序列方法，从而为构造良好性能的 LDPC 码奠定基础；
2. 是否存在一种确定的二部图使得基于此二部图的级联型低密度删码的性能最好；
3. 如何构造具有更大围长、且简单实用的 LDPC 码仍然是一个值得研究的问题；
4. 是否存在一种 LDPC 码的快速编码方法，对任意的 LDPC 码都适用；
5. LDPC 码的和积译码算法只有在无环图上才能获得最优的译码性能，该算法是否存在优化的余地，使其能够在有环图上也能获得最优或近似更优的译码性能，如果存在，如何优化；
6. 仿真表明，LDPC 码对码字中各个比特的保护能力并不相同，如何优化设计 LDPC 码的码结构，以增强对信息比特的保护；
7. LDPC 码的分析理论大多都是考虑当码长趋于无穷时的渐进性能，如何构造长度有限的 LDPC 码使其具有最优的纠错性能；
8. LDPC 码的硬件实现优化，以推进其实用化的进程。

致 谢

首先,我想把我最诚挚的谢意和深深的祝福献给我的导师王新梅教授和师母吉箴琴老师。能在王老师的悉心指导和深切关怀下顺利完成硕士学业,我感到非常荣幸。论文中每个研究成果都是在王老师的亲自指导下获得的。王老师严谨的治学态度、孜孜不倦的钻研精神、深厚的学术造诣以及平易近人的品格,将使我终生受益。王老师倡导了一个民主、和谐的学术气氛,经常给我们提出问题,并给予可能解决问题的分析方法;也总能够认真听取学生的想法,而后积极参与讨论并提出宝贵的指导意见,鼓励我们不断创新。

非常感谢马文平老师、白宝明老师、贺玉成老师、刘东苏老师、李颖老师和郭旭东师兄在我硕士期间所给予的指点和帮助,他们严谨的学风和渊博的知识永远值得我学习。

特别感谢孙韶辉博士、慕建君博士、童胜博士、王单博士和刘斌硕士,我们经常在一起讨论、相互启发和解决难点问题,本文中的许多成果都是我们共同努力的结晶。孙韶辉博士最早带领我接触了 LDPC 码的一些知识;我和慕建君博士一起分析了第三章中正则低密度纠删码的性能和基于改进型右边正则度序列的低密度纠删码的性能等方面的问题,并共同建立了基于这两类度序列的低密度纠删码的仿真模型和仿真程序;我同童胜博士、王单博士和刘斌硕士共同讨论了第四章中 LDPC 码的围长设计和第五章 LDPC 码的快速编码方面的问题,提出了围长设计方案和可快速编码的 LDPC 码设计方案并进行了仿真实现。

感谢香港城市大学的李坪老师、马啸博士、吴克颖博士、刘力海博士和李毅博士,美国加州大学的毛用才博士后,西安电子科技大学的冶继民老师、王玉华老师、周幸妮老师、曹丽娜老师和张卓奎老师,上海交通大学的刘杰博士和李强硕士,清华大学的吕永强博士和章洪灿博士,中山大学的李小玮博士;感谢大唐公司的陈恭华先生和夫人王艳萍女士。

感谢信道编码研究室的兄弟姐妹们,他们是李鸿培博士、任亚安博士、孙永兴博士、韦宝典博士、王辉球博士、郑荣博士、余斌霄博士、张卫党博士、徐朝军博士、肖鸿博士、孙蓉博士、刘景伟博士、许卫东博士、程相国博士、周秦英博士、孙岳博士、池凯凯博士、车书玲博士、刘景美博士、邓浩硕士、秦玉荣硕士、路成业硕士、李云鹏硕士、惠俊红硕士、袁聪硕士、吕继强硕士、张鸿雁硕士、韩锦蓉硕士、王向晖硕士、刘隽硕士、涂世龙硕士、张丰翼硕士、刘晓寒硕士、马允龙硕士、韩亚辉硕士、高延玲硕士、张庆志硕士、李亚汉硕士、苏加军硕士、黄建忠硕士、邓双成硕士、杨晨硕士和鲁慧硕士等,与他们朝夕相处,如

同生活在一个大家庭里。

感谢一次又一次，不厌其烦地和我一起验证“保皇分布”和“双扣定理”的陈鹏硕士、赵鹏硕士、李天保硕士、顾芳明硕士和李刚强硕士，感谢顾华熙博士、杨国亮博士、孙曦博士、田海波博博士、王磊博士、高军涛博士、孙友炜硕士、郑欣硕士、史小斌硕士、樊宁波硕士、倪源硕士、石鹏硕士、张骞硕士、苑雪硕士、黄新芳硕士、范晓鹏硕士、崔树鹏硕士、魏乐硕士、冯涛硕士、聂远飞硕士、博列峰硕士、尤国强硕士、张东波硕士、徐晓宁硕士、朱文凯硕士和宋锐硕士等，谢谢你们在两年多的硕士生活中所给予我的帮助和照顾。

感谢 1-9711 班的所有同学，非常怀念和你们一起度过的大学四年，认识你们，是我一生的财富。

感谢好友周豪杰、谢非、魏刚、翟斌、石维、董云志、宋亮亮、刘敏、彭勇、段宏亮、王中馗、谢丹、杨磊、齐红波、张柯、张国昭、常玉、赵文克、茅涛、李军坡、靳鹏飞、何雪慧、李哲、司海涛、滑向峰、杜华、付朝霞、左米前、左米凡、孙晓存、李峰伟、孟治江等。

特别感谢我的家人，尤其是我的父母和外婆，感谢你们对我无私的付出和对我的理解与支持，没有你们，我不可能有现在的学习机会，在以后的人生道路上，我将尽我所能，以不辜负你们对我的期望。

感谢所有给过我帮助和关爱的人，祝你们永远平安、幸福！

参考文献

- [1] A.R. Calderbank. The art of signaling: Fifty years of coding theory. *IEEE Trans. Inf. Theory*, 44(6):2561-2595, Oct. 1998.
- [2] C.E. Shannon. A mathematical theory of communication. *Bell Syst. Tech. J.*, 27:379-423/623-656. July/Oct. 1948.
- [3] S. Verdu, "Fifty years of Shannon theory," *IEEE Trans. Inform. Theory*, vol. 44, no. 6, pp. 2057-2078, Oct. 1998.
- [4] J.G. Proakis. *Digital Communications*, 3rd ed. New York: McGraw-Hill, 1995.
- [5] 王新梅, 肖国镇. 纠错码原理与方法. 西安电子科技大学出版社, 1991.
- [6] 王育民, 梁传甲. 信息与编码理论. 西北电讯工程学院出版社, 1986.
- [7] T.M. Cover and J.A. Thomas, *Elements of Information Theory*. New York, Wiley 1991.
- [8] R. E. Blahut, *Principles and Practice of Information Theory*. Reading, MA: Addison-wesley, 1987.
- [9] C. Berrou and A. Glavieux. Near optimum error correcting coding and decoding: turbo-codes. *IEEE Trans. Comm.*, 44(10):1261-1271, Oct. 1996.
- [10] C. Berrou, A. Glavieux, and P. Thitimajshima. Near Shannon limit error-correcting coding and decoding: turbo-codes. In *Proc. of ICC'93*, pp.1064-1070, May 1993.
- [11] R.G. Gallager. *Low density parity-check codes*. Cambridge, MA:MIT Press, 1963.
- [12] R.G. Gallager. Low density parity-check codes. *IRE Trans. Inf. Theory*, 8(1):21-28, Jan. 1962.
- [13] G. D. Forney, Jr, "Codes on graphs: New and views," In *Proc. 2nd International Symposium on Turbo Codes and Related Topics*, Brest France, pp. 9-16, Sept. 1988.
- [14] G. D. Forney, Jr, "Codes on graphs: normal realizations," *IEEE Trans. Inform. Theory*, vol47, no2, pp. 520-548, Feb. 2001.
- [15] R. J. McEliece, "Are Turbo-like codes effective on nonstandard channels?" *IEEE Inform. Theory Society Newsletter*, vol.51, no.4, Dec. 2001.
- [16] D.J.C. MacKay. Good error correcting codes based on very sparse matrices. *IEEE Trans. Inf. Theory*, 45(2): 399-431, 1999.
- [17] N. Wiberg. *Codes and decoding on general graphs*. Ph.D. dissertation. Linköping University, Linköping, Sweden, 1996.
- [18] D.J.C. MacKay. Turbo codes are low-density parity-check codes. [Online], available: <http://www.cs.toronto.edu/~mackay/abstracts/-turbo-ldpc-html>, 1998.
- [19] N. Wiberg, H.-A. Loeliger, and R. Köetter. Codes and iterative decoding on general graphs. *European Trans. Telecomm.*, 6:513-525, 1995.
- [20] R.M. Tanner. A recursive approach to low complexity codes. *IEEE Trans. Inf. Theory*, 27: 533-547, Sept. 1981.
- [21] D.J.C. MacKay and R.M. Neal. Near Shannon limit performance of low-density parity check codes. *Electron. Lett.*, 32:1645-1646, Aug. 1996.
- [22] M.C. Davey. *Error-correction using low-density parity-check codes*. Ph.D. dissertation, University Cambridge, Cambridge, UK, Dec. 1999.

- [23] M.C. Davey and D.J.C. Mackay. Low density parity check codes GF(q). *IEEE Comm. Letters*, 2(6):165-167, June 1998.
- [24] D. A. Spielman, Computationally efficient error-correcting codes and holographic proofs. Ph. D dissertation. MIT, 1995
- [25] D.A. Spielman. Linear-time encodeable and decodable error-correcting codes. *IEEE Trans. Inf. Theory*, 42(6):1723-1731, Nov. 1996.
- [26] M. Sipser and D.A. Spielman. Expander codes. *IEEE Trans. Inf. Theory*, 42(6):1710-1722, Nov. 1996.
- [27] D.A. Spielman. "Constructing error-correcting codes from expander graphs," Available at <http://www-math.mit.edu/~spielman>.
- [28] M.G. Luby, M. Mitzenmacher, M.A. Shokrollahi, D.A.Spielman, and V. Stemann. Practical loss-resilient codes. In *Proc. 29th Annu. Symp. Theory of Computing*, pp.150-159, 1997.
- [29] J. W. Byers, M. Luby, M. Mitzenmacher, and A. Rege. A digital fountain approach to reliable distribution of bulk data. [Online], available at <http://www.icsi.berkeley.edu/~luby/>, 1998.
- [30] S.-Y. Chung, G.D. Forney, Jr., T.J. Richardson, and R. Urbanke. On the design of low-density parity-check codes within 0.0045dB of the Shannon limit. *IEEE Comm. Letters*, 5(2):58-60, Feb. 2001.
- [31] M.G. Luby, M. Mitzenmacher, M.A. Shokrollahi, and D.A. Spielman. Analysis of low-density codes and improved designs using irregular graphs. In *Proc. 30th Annu. Symp. Theory of Computing*, pp.249-258, 1998.
- [32] T.J. Richardson, M.A. Shokrollahi, and R.L. Urbanke. Design of capacity-approaching irregular low-density parity-check codes. *IEEE Trans. Inf. Theory*, 47(2):619-637, Feb. 2001.
- [33] S.-Y. Chung, T.J. Richardson, and R.L. Urbanke. Analysis of sum-product decoding of low-density parity-check codes using a Gaussian approximation. *IEEE Trans. Inf. Theory*, 47(2):657-670, Feb. 2001.
- [34] M.G. Luby, M. Mitzenmacher, and M.A. Shokrollahi. Analysis of random processes via and-or tree evaluation. In *Proc. 9th Annu. ACM-SIAM Symp. Discrete Algorithms*, pp.364-373, 1998.
- [35] G.D. Forney, Jr. Codes on graphs: normal realizations. *IEEE Trans. Inf. Theory*, 47(2):520-548, 2001.
- [36] T.J. Richardson and R.L. Urbanke. The capacity of low-density parity-check codes under message-passing decoding. *IEEE Trans. Inf. Theory*, 47(2): 599-618, Feb. 2001.
- [37] Luby, M. Mitzenmacher, M.A. Shokrollahi, and D.A. Spielman. Improved low-density parity-check codes using irregular graphs and belief propagation. In *Proc. of 1998 IEEE ISIT*, Cambridge, MA, pp.117, Aug. 1998.
- [38] M. A. Shokrollahi and R. Storn, "Design of efficient erasure codes with differential evolution," In Proc. of 2000 IEEE ISIT, Sorrento, Italy, June 2000.
- [39] M. A. Shokrollahi, "New sequences of linear time erasure codes approaching the channel capacity," In Pro. AAEECC' 13, Lecture Notes in Computer Science 1719, pp. 65-76. 1999.
- [40] M. A. Shokrollahi, "Codes and graphs," available at <http://www.shokrollahi.com/amain.pyb.html>.
- [41] M. A. Shokrollahi, "Capacity-achieving sequences," available at <http://www.shokrollahi.com/amain.pyb.html>.
- [42] 贺玉成. 基于图模型的低密度校验码理论及应用研究. 西安电子科技大学博士论文, 2002

- [43] 孙韶辉. 低密度校验码中几个关键问题的研究. 西安电子科技大学博士论文, 2002.
- [44] 慕建君. 低密度纠错码和网格复杂度的研究. 西安电子科技大学博士论文, 2002
- [45] P. Li and W.K. Leung. Decoding low density parity check codes with finite quantization bits. *IEEE Comm. Letters*, 4(2):62-64, Feb. 2000.M.G.
- [46] Amin Shokrollahi. "LDPC Codes: An Introduction", available at <http://www.ipm.ac.ir/ipm/homepage>.
- [47] Xiao-Yu Hu, et.al. "Efficient Implementations of the Sum-Product Algorithm for Decoding LDPC Codes," available at <http://www.zurich.ibm.com/pdf/sys/storage>.
- [48] S.-Y. Chung. *On the construction of some capacity-approaching coding schemes*. Ph.D. Dissertation, Massachusetts Institute of Technology, 2000.
- [49] Xian-ping Ge, David Eppstein, Padhraic Smyth, "The Distribution of Cycle Lengths in Graphical Models for Iterative Decoding", Technical Report UCI-ICS 99-10, UCI, 1999.
- [50] R. M. Tanner, "A Class of Group-Structured LDPC Codes", ICSTA 2001 Proceedings, Ambleside, England.
- [51] S. J. Johnson and S. R. Weller, "Construction of Low-Density Parity-Check Code from Kirkman Triple Systems", Proc. IEEE Globecom Conf. San Antonio, TX, 2002.
- [52] S. J. Johnson and S. R. Weller, "Regular Low-Density Parity-Check Codes from Combinatorial Designs".
- [53] Jon-Lark Kim, "Explicit Construction of Families of LDPC Codes with Girth at Least Six",
- [54] Yu Kou, Shu Lin, 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, Nov. 2001.
- [55] Hongwei Song, Jingfeng Liu and B. V. K. Vijaya Kumar, "Low Complexity LDPC Codes for Partial Response Channels", Globecom'2002, Taipei, Taiwan, November 2002.
- [56] Tong Zhang and Keshab K. Parhi, "A Class of Efficient-Encoding Generalized Low-Density Parity-Check Codes", Proc. of 2001 IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP), vol. 4, pp. 2477-2480, Salt Lake City, Utah, May 2001.
- [57] Tong Zhang and Keshab K. Parhi, "Joint Code and Decoder Design for Implementation-Oriented (3,k)-Regular LDPC Codes".
- [58] S. Hirst and B. Honary, "Decoding of Generalized Low-Density Parity-Check Codes Using Weighted Bit-flip Voting".
- [59] M. M. Mansour and N. R. Shanbhag, "Low-Power VLSI Decoder Architectures for LDPC Codes", International Symposium on Low Power Electronics and Design, 2002, Monterey, California, USA 284 - 289.
- [60] Xiao-Yu Hu, et al. "Progressive Edge-Growth Tanner Graphs", IEEE Trans. Inform. Theory. Feb. 2001, pp.995-1001.
- [61] J.C. Lin, S. Paul. RMTP: A reliable Multicast Transport Protocol. *IEEE INFOCOM'96*, pp.1414-1424, Mar. 1996.
- [62] S. Deering. *Multicast Routing in a Datagram Internetwork*. Ph.D. dissertation. Stanford University, Dec. 1991.
- [63] D. Rubenstein, J. Kurose, and D. Towsley. Real-time reliable multicast using proactive forward error correction. [Online], available at <http://citeseer.nj.nec.com/110093.html>.
- [64] S. Floyd, V. Jacobson, C.G. Liu, S. McCanne, L. Zhang. A reliable multicast framework for light-weight sessions and application level framing. In *ACM SIGCOMM'95*, pp. 342-356,

- Aug. 1995.
- [65] S. McCanne, V. Jacobson, and M. Vetterli. Receiver-driven layered multicast. In *Proc. of ACM SIGCOMM'96*, pp. 117-130, 1996.
 - [66] C.K. Miller. Reliable multicast protocols: A practical View. In *Proc. of the 22nd Annual Conference on Local Computer networks(LCN'97)*, 1997.
 - [67] R. Yavatkar, J. Griffioen, and M. Sudan. A reliable dissemination protocol for interactive collaborative applications. In *Proc. of ACM Multimedia'95*, San Francisco, pp.333-344, 1995.
 - [68] L. Rizzo. Effective erasure codes for reliable computer communication protocols. In *Computer communication Review*. April 1997.
 - [69] L. Rizzo, and I. Vicisano. A reliable multicast data distribution protocol based on software FEC Techniques. In *Proc. of HPCS'97*, Greece, June 1997.
 - [70] L. Vicisano, I. Rizzo, and J. Crowcroft. TCP-like congestion control for layered multicast data transfer. In *Proc. of INFOCOM'98*, CA, April 1998.
 - [71] L. Rizzo. On the feasibility of software FEC. *DEIT technical Report LR-970131*.
 - [72] F.J. MacWilliams and N.J.A. Sloane. *Theory of Error-Correcting Codes*. Amsterdam. The Netherlands:North- Holland. 1977.
 - [73] J. Blömer, M. Mitzenmacher, and A. Shokrollahi. An XOR-based erasure-resilient coding scheme ICSI Technical Report , No. TR-95048, Aug. 1995. Available at <http://www.icsi.berkeley.edu/~luby/>.
 - [74] M. Luby, et al. Habilitationsschrift. [Online], available at <http://citeseer.nj.nec.com/419720.html>, 2001.
 - [75] M.A. Shokrollahi. New sequences of linear time erasure codes approaching the channel capacity. In *Proc. 13th Conf. Applied Algebra, Error Correcting codes, and Cryptography* (Lecture Notes in Computer Science), Berlin, Germany: Springer Verlag, pp.65-76, 1999.
 - [76] D.J. C. Mackay, S. T. Wilson, and M. C. Davey, “Comparison of constructions of irregular Gallager codes”, in Proc. 36th Allerton Conf. Communication, Control. And Computing, Sept. 1998.
 - [77] T. J. Richardson and R. L. Urbanke, “Efficient Encoding of Low-Density Parity-Check Codes”, IEEE Transactions on Information Theory, Vol 47, No. 2, Feb. 2001.
 - [78] D.Divsalar, H. Jin, and R. J. McEliece. “Coding theorems for ‘turbo-like’ codes”. In Proc. 36th Allerton Conf. On Communication, Control, and Computing, Allerton, Illinois, pp. 201-210, Sept. 1998.
 - [79] H. Jin, A. Khandekar, and R. J. McEliece. “Irregular Repeat-Accumulate Codes.” In Proc. 2nd International Symposium on Turbo Codes and Related Topics, Brest France, pp.1-8, Sept. 2000.
 - [80] R.J. McEliece, D.J.C. MacKay, J.-F. Cheng. Turbo decoding as an instance of Pearl’s “belief propagation” algorithm. *IEEE J. Select. Areas Comm.*, 16(2): 140-152, Feb. 1998.
 - [81] Mu Jian-jun, He Yu-cheng, Wang Xin-mei, “Design Method of Sequences of Degree Distribution for Low-Density Erasure Codes”, *Chinese Journal of Electronics*, 2002, 11(3), pp.372-376.
 - [82] 慕建君, 孙韶辉, 王新梅, “关于线性时间复损码的研究”, 电子学报, 2002, 第 30 卷, 第 1 期, pp.122-125.
 - [83] 慕建君, 路成业, 王新梅, “关于纠删码的研究与进展”, 电子信息学报, 2002, 第 24 卷, 第 9 期, pp.1276-1281.

硕士期间完成的论文和科研工作

一、完成论文

- [P1] 慕建君, 王鹏, 王新梅, “正则低密度纠错码的性能分析”, 西安电子科技大学学报, 2003 年第 4 期, p469-472.
- [P2] 慕建君, 王鹏, 王新梅, “基于改进型右边正则度分布序列的低密度纠错码”, 计算机学报, 2003 年第 12 期, 1734-1738.
- [P3] 王鹏, 王新梅, “LDPC 码的线性编码研究”, 西安电子科技大学学报, 已录用.
- [P4] 王鹏, 王单, 童胜, 王新梅, “一种消除小环的正则 LDPC 码构造方法”, 现代通信理论与信号处理进展——2003 年通信理论与信号处理年会论文集, 电子工业出版社, 2003. 10. 第一版, p76-82.
- [P5] 王鹏, 王单, 童胜, 王新梅, “一种不含小环的正则 LDPC 码构造方法”, 电子学报, 已投.
- [P6] Wang Peng, Tong Sheng, Wang Xinmei, “Construction of LDPC codes with large girth”, submitted to ISIT2004.
- [P7] 童胜, 王鹏, 王单, 王新梅, “LDPC 码和积译码算法的量化实现”, 西安电子科技大学学报, 已录用.
- [P8] 童胜, 王鹏, 王单, 王新梅, “LDPC 码量化和积译码算法实现”, 现代通信理论与信号处理进展——2003 年通信理论与信号处理年会论文集, 电子工业出版社, 2003. 10. 第一版, p88-95.
- [P9] 童胜, 王鹏, 王单, 王新梅, “LDPC 码量化和积译码算法的高效实现”, 2003 年西安电子科技大学研究生学术年会.
- [P10] 王鹏, 王新梅, “LDPC 码的快速编码研究”, 2003 年西安电子科技大学研究生学术年会.
- [P11] 王鹏, 王单, 童胜, 王新梅, “一种具有较大围长的正则 LDPC 码构造方法”, 2003 年西安电子科技大学研究生学术年会.
- [P12] 刘隽, 王鹏, 李颖, 王新梅, “差分空时码与 LDPC 码级联系统研究”, 2003 年西安电子科技大学研究生学术年会.
- [P13] 王向辉, 王鹏, 王新梅, “复合测距伪随机码的选择”, 通信技术, 2003 年第 11 期, p3-5.
- [P14] Zhang Weidang, Wang Peng, Wang Xinmei, “A simple method to increase the minimum weight of Turbo codes”, submitted to IEE Electronics Letter.

二、主要科研

1. 国家自然科学基金项目，“基于图模型的低密度校验码编码理论研究”，项目编号：60272057；
2. 重庆市/信息产业部移动通信技术重点实验室开放课题基金项目，“通信网中低密度纠错码的研究”，项目编号：69972035；
3. 韩国三星综合技术院资助项目，“中国 Beyond IMT-2000 信道编码技术”；
4. 国家自然科学基金委员会和香港科技局联合资助项目，“级联空时码研究”，项目编号：60131160742；
5. 陕西省自然科学基金，“计算机网络中低密度纠错码的研究”，项目编号：NO.2002F01。