分类号：_____ 单位代码：10335

学　　号：11235062

浙江大学

博士学位论文

中文论文题目：极值组合方法在几类信息问题中的应用

英文论文题目：**The application of the extremal combinatorial methods in several problems of information science**

申请人姓名：汪馨

指导教师：葛根年　教授

专业名称：应用数学

研究方向：组合数学

所在学院：数学科学学院

论文提交日期　2017年5月20日

极值组合方法在几类信息问题中的应用

论文作者签名： _____

指导教师签名： _____

论文评阅人1：_____

评阅人2：_____

评阅人3：_____

评阅人4：_____

评阅人5：_____

答辩委员会主席：范更华　教授　福州大学

委员1：范更华　教授　福州大学

委员2：宗传明　教授　天津大学

委员3：符方伟　教授　南开大学

委员4：吴佃华　教授　广西师范大学

委员5：葛根年　教授　浙江大学

答辩日期：2017年5月19日

The application of the extremal combinatorial methods

in several problems of information science

Author's signature: _____

Supervisor's signature: _____

External Reviewers: _____

_____

_____

_____

_____

Examining Committee Chairperson:

Prof.Genghua Fan, Fuzhou University

Examining Committee Members:

Prof.Genghua Fan, Fuzhou University

Prof.Chuanming Zong, Tianjin University

Prof.Fangwei Fu, Nankai University

Prof.Dianhua Wu, Guangxi Normal University

Prof.Gennian Ge, Zhejiang University

Date of oral defence: May 19th, 2017

浙江大学研究生学位论文独创性声明

本人声明所呈交的学位论文是本人在导师指导下进行的研究工作及取得的研究成果。除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含为获得浙江大学或其他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谢意。

学位论文作者签名：　　签字日期： 2017 年 5 月 10 日

学位论文版权使用授权书

本学位论文作者完全了解浙江大学有权保留并向国家有关部门或机构送交本论文的复印件和磁盘，允许论文被查阅和借阅。本人授权浙江大学可以将学位论文的全部或部分内容编入有关数据库进行检索和传播，可以采用影印、缩印或扫描等复制手段保存、汇编学位论文。

（保密的学位论文在解密后适用本授权书）

学位论文作者签名：　　导师签名：

签字日期： 2017 年 5 月 10 日　　签字日期： 2017 年 5 月 10 日

致　　谢

首先我要衷心地感谢我的导师葛根年教授。葛老师带我走进了组合数学的殿堂，他渊博的知识、严谨的治学态度、认真细致的做事风格都令我十分钦佩，这将是我受益终生的宝贵财富。在这五年的博士生生涯中，葛老师在科研、生活与为人处世等方面给予了我很多的教导和建议。葛老师教会了我科学的思维方式、正确的研究方法、前沿的思想理念，他每一次的提点都让我获益匪浅。

我还要感谢这五年中在学习和生活上给予过我指导的各位老师，特别是，日本筑波大学缪莹教授，美国 Delaware 大学向青教授，苏州大学季利均教授，浙江大学冯涛研究员，首都师范大学张俊老师，同济大学的杨亦挺老师等。在与他们的交流中，我得以开拓研究视野，体会到科研的乐趣。也感谢他们对我的种种建议与鼓励。

感谢在一起学习与研究的同门：张会师姐、高斐师兄、朱明志师兄、魏恒嘉师兄、胡思煌师兄、李抒行师兄、林浩师兄、张一炜师兄、上官冲、张韬、顾玉杰、陈寿长、马景学、丁报昆、钱昺辰、孔祥梁、林灯、韩雪娇、徐子翔等。

在这共同学习和生活的岁月中我们一起留下了许多美好的回忆。尤其是同门中魏恒嘉师兄带我跨过了科研中的第一道坎，张一炜师兄不厌其烦地和我探讨各种问题。

还要感谢我亲爱的朋友们：俞星、钱皓磊、俞阳阳、傅盼等。感谢你们的关怀和鼓励。

最后，谨以此文献给我挚爱的父母和长辈，他们在背后的默默支持是我前进的动力。在此，祝愿他们身体健康，心情愉快！

由于作者水平有限，加之时间和篇幅所限，文中难免有谬误和不详之处，敬请各位专家学者不吝批评指正！

摘　　要

组合数学和信息科学有着密不可分的联系，一方面组合数学的强大理论工具为信息科学的研究提供强有力的支撑，另一方面信息科学中产生的各种与组合数学相关的问题进一步刺激了组合数学的发展。极值组合学是近几十年来组合数学中发展最为迅猛的一个分支，同时它与信息科学的交叉最为紧密。本学位论文主要应用极值组合学的基本方法，对相关的问题进行了研究，并取得了一定的进展。

在第 1 章绪论部分，我们将简要介绍所研究问题的背景和本文的主要贡献。

在第 2 章中，我们考虑循环压缩感知矩阵的构造。压缩感知矩阵的构造一直是信号处理领域最关心的问题之一，我们利用部分指数和所给出的循环矩阵，既渐近地达到了理论界，并且具有存储量小和运算速度快的优点。

在第 3 章中，我们考虑多重常重码的理论研究。多重常重码是物理不可克隆函数和编码理论的桥梁。我们借助球面码的上界给出了多重常重码的新上界，从而改进了第三型 Johnson 界；利用图分解的工具完全决定了两类多重常重码的最大码字容量。

在第 4 章中，我们考虑极值集合论中的 L-相交系问题，利用线性代数方法，包括关联矩阵秩的估计和多线性多项式，对 Alon-Babai-Suzuki 不等式进行了改进。

在第 5 章中，我们考虑私人信息检索中的 PIR 阵列码的构造问题。从应用的角度出发，构造只需少量服务器的最优 PIR 过程是非常迫切的。我们利用组合设计的思想，在 $t > d^2 - d$ 的情形下，设计了服务器数量达到最少的最优 PIR 阵列码，同时还给出了新上界的刻画。

第 6 章对本人博士期间的其他工作进行了总结。

关键词： 压缩感知，部分指数和，多重常重码，球面码，图分解，L-相交系，线性代数方法，私人信息检索，组合设计

Abstract

Combinatorics has a close connection with information science. On one hand, combinatorics provides the theoretical supports for the study of information science. On the other hand, a variety of problems arising from information science stimulate the rapid development of combinatorics. Extremal combinatorics is one of the most active branches of combinatorics in recent decades. Simultaneously it has the closest interaction with information science. This dissertation aims to use the basic methods of extremal combinatorics to study several problems arising from information science and give some improvements.

In Chapter 1, we will briefly introduce the backgrounds of the problems concerned, and summarize our main contributions to these problems.

In Chapter 2, we focus on the construction of the circulant compressed sensing matrix. The construction of the compressed sensing matrix is one of the most important problems in signal processing. We will use the estimation of the partial exponential sum to give the construction of the circulant matrices, which not only achieve the theoretic bound but also perform well in storage and computation.

In Chapter 3, we focus on the theory of multiply constant-weight codes. Multiply constant-weight codes establish the connection between the design of the Loop physically unclonable functions and coding theory. We give a new upper bound for multiply constant-weight codes through the sphere codes, which improves the Johnson bound. By the tool of the graph decomposition, we show the existence of two classes of optimal multiply constant-weight codes.

In Chapter 4, we consider the problem about L-intersection from extremal set system. We use the linear algebra method, including incidence matrix and multilinear polynomial, to improve the estimation of Alon-Babai-Suzuki inequality.

In Chapter 5, we focus on the construction of the PIR array codes. A scheme with a small number of servers is of great interest due to applicable reasons. We use the idea from combinatorial designs to construct the optimal PIR array code with the minimal number of servers. Meanwhile,

we give a new upper bound for PIR array codes.

In Chapter 6, we summarize the other works while pursuing the doctor degree.

Keywords: compressed sensing, partial exponential sum, multiply constant-weight codes, sphere codes, graph decomposition, L-intersection, linear algebra method, private information retrieval, combinatorial designs

插　　图

2-1　基于 Zadoff-Chu 序列的 29×840 随机、确定部分循环矩阵和 29×840 部分傅里叶矩阵的实验结果..... 18
2-2　基于 m 序列的 29×840 随机和确定部分循环矩阵的实验结果..... 18

表　格

3-1　小参数下的 MCWC$(2, n_1; 2, n_2; 6)$，$3 \leq n_1 \leq 9$..... 40

目　　次

致谢 I

摘要 III

Abstract V

插图 VII

表格 IX

目次

1　绪论 1

1.1　循环压缩感知矩阵 1

1.2　多重常重码 2

1.3　*L*-相交系 4

1.4　私人信息检索 5

2　循环压缩感知矩阵 7

2.1　介绍 7

2.2　预备工作 8

2.2.1　RIP 和 MIP 的定义 8

2.2.2　最坏情形和平均情形 9

2.2.3　部分指数和 9

2.3　矩阵的构造 12

2.3.1　基于 Zadoff-Chu 序列族的构造 12

2.3.2　基于 m 序列的构造 14

2.4　实验结果 16

2.5　小结 17

3　多重常重码 19

3.1　介绍 19

3.2　定义和记号 20

3.2.1 多重常重码 20
3.2.2 图分解 21
3.3 多重常重码上下界的改进 22
3.3.1 由球面码导出的上界 22
3.3.2 多重常重码的 Plotkin 界 24
3.3.3 多重常重码的线性规划界 26
3.3.4 多重常重码的 GV 界 28
3.4 两类最优码 31
3.4.1 极小距离为 $2\sum_{i=1}^{m} w_i - 2$ 的最优多重常重码 31
3.4.2 极小距离为 $2mw - 2w$ 的最优多重常重码 33
3.5 重量为 4、极小距离为 6 的多重常重码 35
3.6 小结 43
4 L-相交系 45
4.1 介绍 45
4.2 定理 4.14 的证明 48
4.3 定理 4.15 的证明 54
4.4 小结 59
5 私人信息检索 61
5.1 介绍 61
5.2 $1 < s \leq 2$: 最优 PIR 阵列码的构造 63
5.2.1 $t \geq d^2$ 64
5.2.2 $d^2 - d < t < d^2$ 65
5.3 $s > 2$: $g(s,t)$ 的上下界研究 68
5.3.1 $g(s,t)$ 的新上界 68
5.3.2 PIR 阵列码的一般构造 71
5.4 小结 75
6 其它研究工作 77
6.1 可逆 2×2 子矩阵比例问题 77
6.2 置换码 77
6.3 防诬陷码 78
6.4 光正交签名码 78
6.5 关于集合差的集族问题 78

参考文献 81
作者简历 89
攻读博士学位期间主要研究成果 91

1 绪论

信息科学与组合数学有着密不可分的联系。

信息科学所研究的对象大都是离散结构，具有很强的组合性质。正如通信理论的奠基人 Shannon 在《通信的数学理论》中所提及，数学模型是最能简单有效地指出通信技术发展方向和趋势的。21 世纪是信息产业飞速发展的全新时代。随着传输数据量激增以及用户对多媒体通信的质量与速度的需求，通信传输技术必须向高性能、低成本和智能化等方向发展，寻求新的信息快速传输算法与处理方式和物理实现是现代通信领域面临的一大挑战。与此同时，随着数据流量的增大和信息的争夺，通信容易被干扰和窃听，存在严重的安全隐患。在这样的大背景下，组合数学有了更广阔的应用舞台。

极值组合学是近几十年来组合数学中发展最为蓬勃的一个分支。自从天才数学家 Erdös 在组合学的研究中引入概率方法以来，在 Alon，Bollobás，Lovász，Rödl 和 Szemerédi 等一批数学家的推动下，极值组合学已经发展成为组合数学乃至整个数学领域中最为重要的分支之一。同时，伴随着计算机科学、通信技术、互联网等信息科学的发展，产生了大量的极值组合问题，极值组合学中的概率方法和代数方法为信息科学的研究提供了强大的理论支撑，从另一方面促进了极值组合学的发展。

本文将利用组合学中的常用方法，对一些从信息科学中产生和源于组合学本身的问题进行研究和思考。具体包括：信号采样中的压缩感知矩阵、多重常重码、L-相交系和私人信息检索中的阵列码等问题。下面将介绍各研究课题的背景意义，并概述本文在各研究课题上所做的工作。

1.1 循环压缩感知矩阵

压缩感知理论[10,12,27,61] 利用信号的稀疏特性，在远远小于奈奎斯特采样率的条件下，通过有效快速的算法来重构原始的信号。这个发现在信号处理领域有许多潜在的应用。其中的采样过程我们可以利用一个测量矩阵 $\Phi \in C^{m\times N}$ 来描述，其中 $m \ll N$。给定一个长

度为 N 的 k-稀疏信号 x，整个压缩感知的过程可以利用如下线性方程组来表达：

$$y = \Phi x + e,$$

其中 y 表示 m 长的采样向量， e 为系统的噪声。

如何构造一个好的压缩感知矩阵是数学和信息学的交叉研究课题。本文所研究的循环压缩感知矩阵不仅在稀疏信道估计[40] 、傅里叶光学[66] 和雷达图像[66] 中有着重要的应用，同时具有存储量小和运算速度快等优点，因而得到广泛的重视。

受限等距性（RIP）[11,12] 是最基本的衡量矩阵是否具有好的稀疏信号恢复能力的度量，随机矩阵的理论表明，包括高斯矩阵、伯努利矩阵在内的许多随机矩阵[5,12,29,63] 都具有好的 RIP 性质。相对的，之前关于确定性矩阵构造的结果[26,51–53,77]，大都集中在矩阵相干性（MIP）的讨论上:

$$\mu(\Phi) = \max_{1\le i\ne j\le N} \frac{|\ \langle \Phi_i, \Phi_j\rangle\ |}{\|\ \Phi_i\ \|_2 \|\ \Phi_j\ \|_2},$$

其中 Φ_i 表示矩阵 Φ 中的第 i 列。

在本文之前，所有的关于循环压缩感知矩阵的研究中都需要借助于随机矩阵的理论：

1. 确定采样和随机矩阵：在文献[47,61] 中证明了当 $m \ge O(k\log^4 N)$ 时矩阵 Φ 满足 RIP 性质；

2. 随机采样和随机矩阵：Romberg 在[66] 中证明了当 $m \ge O(k\log^5 N)$ 时矩阵 Φ 满足 RIP 性质；

3. 随机采样和确定性矩阵：在文献[50] 中，李等人构造了一些确定性的矩阵，再利用随机采样的方法证明了在 $m \ge O(k\log^4 N)$ 的条件下这些矩阵也具有 RIP 性质。

本文的主要贡献在于提出新的方案来实现完全确定的循环压缩感知矩阵的构造。我们利用部分指数和的估计构造了两类循环矩阵，从 MIP 的角度出发它们都与之前的确定性矩阵相当，从统计的角度出发我们所得到的矩阵达到了理论界 $m \ge O(k\log N)$，同时实验的结果也支撑了我们的理论。这部分的工作已发表在《Finite Fields and Their Application》。

1.2 多重常重码

现代密码体系严重依赖于单向函数的应用，但是通常的单向函数总是基于还未被证明的猜想来保证它的安全性。由 Pappu 等人在文献[58] 中提出的物理不可克隆函数（PUF）提供了一种认证消耗低和抵抗物理攻击的新选择。近年来，物理不可克隆函数的研究已经成

为无线电频率识别和智能卡领域[18,37,58,72]的一种潮流。在文献[14,19] 中，Chee 等人提出了多重常重码（MCWC），从而为设计环形物理不可克隆函数和编码理论建立了桥梁。设 X 可以被划分成 $X = X_1 \cup X_2 \cup \cdots \cup X_m$，其中 $|X_i| = n_i$，$i = 1, 2, \ldots, m$。如果每个码字都满足在由 X_1 标记的坐标上重量恰为 w_1，由 X_2 标记的坐标上重量恰为 w_2，以此类推，我们称这个 (N, d) 码 $\mathcal{C} \subseteq \mathbb{Z}_2^X$ 是多重常重的。

与经典的编码理论相似，多重常重码的研究也主要关心它的上下界问题。在文献[14] 中，Chee 等人推广了 Johnson 关于常重码的相关研究[44] 给出了多重常重码的一些上下界。他们进一步还证明了在相差一个常数因子的意义下，他们的界是渐近紧的。在文献[16] 中，Chee 等人给出了几类新的最优多重常重码的构造。进一步地，在文献[15] 中，他们在极小距离为 $2mw - 2$ 时，证明了 Johnson 界是渐近精确的。

一般参数意义下的码容量问题，一直是编码领域的公开问题。而球面码作为欧式距离下的几何对象，在很大的范围内它的码容量都已经被完全决定。本文通过多重常重码与球面码的联系，利用球面码的上界改进了多重常重码的第三型 Johnson 界。同时，我们还得到了一般意义下多重常重码的 Gilbert-Varshamov 界，渐近意义下改进了文献[14] 中由级联得到的下界。

记有限集合 X 中的所有有序对为 $\overline{\binom{X}{2}}$，用三元组 $G = (V, C, E)$ 来表示边着色的有向图，其中 V 表示顶点集合，C 表示颜色集合，E 是集合 $\overline{\binom{X}{2}} \times C$ 的一个子集（E 一般称为边集）。如果 G 中的每条边都恰好属于集族 $\mathcal{F}$ 的一个元素，称集族 $\mathcal{F}$ 是图 G 的分解。图分解可以看成是区组设计的一种自然推广。Lamken 和 Wilson 最早在文献[48] 中对这类问题进行了系统地研究。本文利用图分解的工具，完全确定极小距离为 $2\sum_{i=1}^{m} w_i - 2$ 和 $2mw - 2w$ 这两类多重常重码的最大码字个数：

定理 1.1. 设 $w_1 \geq w_2 \geq \cdots \geq w_m$，$w = \sum_{i=1}^{m} w_i$，存在正常数 n_0 使得在 $w_1 > w_2$ 时满足 $n - 1 \equiv 0 \pmod{w_1(w_1 - 1)}$；否则，满足 $n - 1 \equiv 0 \pmod{w_1^2}$，对任意 $n \geq n_0$，都有：

$$T(w_1, n; \ldots; w_m, n; 2w - 2) = \begin{cases} \frac{n(n-1)}{w_1(w_1-1)}, & \text{如果 } w_1 > w_2; \\ \frac{n(n-1)}{w_1^2}, & \text{如果 } w_1 = w_2. \end{cases}$$

定理 1.2. 给定正整数 k 和 w，满足 $k \mid w$，存在一个正常数 $m_0 = m_0(k, w)$ 使得对任意 $m \geq m_0$，

$$M(m, n, 2(mw - w), w) = m(k - 1) + 1$$

其中 $n = w(m(k-1)+1)/k$。

这部分的工作已经发表在《IEEE Transactions on Information Theory》。

1.3 L-相交系

如果集族 $\mathcal{A}$ 中的任何一对集合 $A_i, A_j \in \mathcal{A}$ 都有非空的交，我们称 $[n] = \{1, 2, \ldots, n\}$ 中的集族 $\mathcal{A}$ 是相交的。令 L 是由 s 个非负整数所构成的集合。如果集族 $\mathcal{A}$ 中的任何一对集合 A_i, A_j 满足 $|A_i \cap A_j| \in L$，我们称 $\mathcal{A}$ 为 L-相交系。L-相交系的研究与常重码的上界问题有着密不可分的联系，同时 L-相交系的构造推动着私人信息检索方案的设计。关于 L- 相交系的系统研究，可以追溯到 1975 年，Ray-Chaudhuri 和 Wilson[64] 通过分析 Wilson 矩阵秩的方法，对于均匀 L-相交系得到了如下重要的进展：

定理 1.3. 设 $\mathcal{A}$ 是 $[n]$ 中的 k-均匀 L-相交系，那么 $|\mathcal{A}| \leq \binom{n}{s}$。

1991 年，Alon，Babai 和 Suzuki 在文献[3] 中成功地借鉴了 Seidel 等人在研究欧式空间上 2-距离集中引进的多项式方法，对上面的问题进行了重新的研究，并得到如下推广形式的结论：

定理 1.4. 设 p 为一素数，$K = \{k_1, k_2, \ldots, k_r\}$ 和 $L = \{l_1, l_2, \ldots, l_s\}$ 是 $\{0, 1, \ldots, p-1\}$ 中两个互不相交的子集。如果集族 $\mathcal{A}$ 满足 $|A_i| \pmod p \in K$ 对任意 $A_i \in \mathcal{A}$，$|A_i \cap A_j| \pmod p \in L$ 对任意 $i \neq j$，当 $r(s-r+1) \leq p-1$ 和 $n \geq s + \max_{1\leq i\leq r} k_i$ 时，有 $|\mathcal{A}| \leq \binom{n}{s} + \binom{n}{s-1} + \cdots + \binom{n}{s-r+1}$ 成立。

本文综合地运用了线性代数方法改进了上面的结果：

定理 1.5. 设 p 为一素数，$K = \{k_1, k_2, \ldots, k_r\}$ 和 $L = \{l_1, l_2, \ldots, l_s\}$ 是 $\{0, 1, \ldots, p-1\}$ 中两个互不相交的子集。如果集族 $\mathcal{A}$ 满足 $|A_i| \pmod p \in K$ 对任意 $A_i \in \mathcal{A}$，$|A_i \cap A_j| \pmod p \in L$ 对任意 $i \neq j$，当 $n \geq 2s - 2r + 1$ 时，有 $|\mathcal{A}| \leq \binom{n-1}{s} + \binom{n-1}{s-1} + \cdots + \binom{n-1}{s-2r+1}$ 成立。

定理 1.6. 设 p 为一素数，$K = \{k_1, k_2, \ldots, k_r\}$ 和 $L = \{l_1, l_2, \ldots, l_s\}$ 是 $\{0, 1, \ldots, p-1\}$ 中两个互不相交的子集。如果集族 $\mathcal{A}$ 满足 $|A_i| \pmod p \in K$ 对任意 $A_i \in \mathcal{A}$，$|A_i \cap A_j| \pmod p \in L$ 对任意 $i \neq j$，当 $n \geq s + \max_{1\leq i\leq r} k_i$ 时，有 $|\mathcal{A}| \leq \binom{n-1}{s} + \binom{n-1}{s-1} + \cdots + \binom{n-1}{s-2r+1}$ 成立。

这部分工作已经投稿至《Discrete Mathematics》。

1.4 私人信息检索

私人信息检索（PIR）最早在文献[20]中提出：假设有 n 比特的数据和 k 个服务器，每个服务器都包含有全部的信息，因此它的总存储量为 nk，一个基于 k 个服务器的 PIR 过程允许用户检索到需要的信息，但是服务器却无法得知用户的需求。比如说，假设数据集为 $\mathbf{x}=(x_1,x_2,\ldots,x_n)$，某个用户想要知道第 i 位的信息 x_i。在一个基于 2 个服务器的 PIR 过程中，用户可以随机地选取一个向量 $\mathbf{v}\in\{0,1\}^n$。第一个服务器接收到问询 $\mathbf{v}$，向用户反馈 $\mathbf{v}\cdot\mathbf{x}$。第二个服务器收到问询 $\mathbf{v}+\mathbf{e_i}$，向用户反馈 $(\mathbf{v}+\mathbf{e_i})\cdot\mathbf{x}$。那么用户可以通过 $x_i=(\mathbf{v}+\mathbf{e_i})\cdot\mathbf{x}-\mathbf{v}\cdot\mathbf{x}$ 检索到自己想要的信息。但是由于 $\mathbf{v}$ 是随机选取的，每个服务器都得不到任何关于用户的有用信息。

近来，PIR 过程与分布式存储的思想产生了联系[4,13,33,69,73]，将原有的在每个服务器上存储全部的信息，改变成服务器上都只存储部分信息（即采用编码的方案）。在开创性的工作[34,35]中，Fazeli 等人证明了利用 m 个服务器（其中 $m>k$），可以大大地降低服务器的总存储量。这个问题最终被转化为 PIR 阵列码的设计问题。

这个问题得到了 Blackburn 和 Etzion 的关注[6]，他们希望去构造达到最优码率的阵列码。在给定 $t\geq 2$，$1\leq d\leq t$ 且 $s=1+\frac{d}{t}$ 时，他们已经给出了最优阵列码的构造，但是在他们的构造中列数惊人的大 $m=\binom{t+d}{t}\frac{v}{d}+\binom{t+d}{d+1}\frac{v}{t}$，其中 v 是 d 和 t 的最小公倍数。这就意味着他们的设计方案需要海量的服务器才能实现。从应用角度出发，构造只需少量服务器的方案是非常迫切的。我们将考虑如下问题：如何利用最少的服务器数量 m 使得码率 k/m 达到最优。在本文中，我们利用区组设计的思想在 $s=1+\frac{d}{t}, t>d^2-d$ 的情形下给出了上面问题的完整回答。但是在 $s>2$ 时，我们甚至连最优码率究竟是多少也不能给出直接的回答，只能利用细致的组合分析得到了部分结果，从而改进了文献[6]中的上界。这部分的工作已经投稿至《IEEE Transactions on Information Theory》。

2 循环压缩感知矩阵

2.1 介绍

压缩感知理论[10,12,27,61]利用信号的稀疏特性，在远远小于奈奎斯特采样率的条件下，通过有效快速的算法来重构原始的信号。这个发现在信号处理领域有许多潜在的应用。其中的采样过程我们可以利用一个测量矩阵 $\Phi \in C^{m\times N}$ 来描述，其中 $m \ll N$。给定一个长度为 N 的信号 x，如果存在一组基 Ψ，使得这个信号 x 有 k-稀疏表示，也就是说，存在一个向量 f 只有 $k \ll N$ 个坐标非零，使得 $x = \Psi f$。整个压缩感知的过程可以利用如下线性方程组来表达：

$$y = \Phi x + e = \Phi\Psi f + e,$$

其中 y 表示 m 长的采样向量， e 为系统的噪声。为了考虑问题的方便起见，在本章中我们总是考虑无噪声的情形（$e = 0$）。

我们的目标是去设计好的压缩感知矩阵。受限等距性（RIP）[11,12]是研究测量矩阵是否具有好的采样稀疏信号能力的最基本度量，同时它在信号恢复算法分析中也扮演着举足轻重的作用[10,27,56,57,74]。随机矩阵的理论表明，包括高斯矩阵、伯努利矩阵在内[5,12,29,63]的随机矩阵都具有好的 RIP 性质。另一个用来衡量矩阵优劣的度量是由 Donoho 和 Huo 提出的矩阵相干性（MIP）。几乎所有的确定性矩阵的构造都依赖于 MIP。

尽管随机矩阵在理论上拥有很多优势，但是它们在实现过程中却有着需要海量的存储容量和矩阵运算时间等不可克服的缺点。拿核磁共振为例，我们从一个傅里叶变换矩阵中随机的抽取足够多的行来作为测量矩阵，那么我们在矩阵乘法中就可以利用快速傅里叶变换（FFT）来大大减少运算时间，已经有理论表明这样的部分傅里叶矩阵同时拥有良好的 RIP[12,67] 和 MIP[77] 性质。基于此，许多科研工作者都致力于带有结构的压缩感知矩阵的研究[30,40,62,66]。

本章研究的循环压缩感知矩阵在稀疏信道估计[40] 、傅里叶光学[66] 和雷达图像[66] 中均有着重要的应用。部分循环矩阵 Φ 可以用如下方式来生成：

$$\Phi = \frac{1}{\sqrt{m}} R_\Omega A,$$

其中 A 是一个循环矩阵：

$$A = \begin{pmatrix} a_0 & a_1 & \cdots & a_{N-1} \\ a_{N-1} & a_0 & \cdots & a_{N-2} \\ \vdots & \vdots & \ddots & \vdots \\ a_1 & a_2 & \cdots & a_0 \end{pmatrix},$$

R_Ω 为矩阵 A 的采样算子，其中 Ω 来标记采样的位置。

众所周知一个循环矩阵可以通过傅里叶变换来进行对角化：

$$A = \frac{1}{\sqrt{N}} F \Sigma F^H,$$

其中 $\Sigma = diag(\sigma) = diag(\sigma_0, \sigma_1, \cdots, \sigma_{N-1})$，矩阵 F 和 F^H 分别表示离散傅里叶变换矩阵及其逆矩阵。因此循环矩阵可以通过 FFT 来进行乘法运算。

我们对与本章相关的工作做如下总结：

1. 确定采样和随机矩阵：在文献[40,62] 中，向量 a 随机生成，确定采样的过程可由 $\{1, 2, \cdots, N\}$ 的任意大小为 $|\Omega| = m$ 的子集所确定。在文献[47,61] 中证明了当 $m \geq O(k \log^4 N)$ 时矩阵 Φ 满足 RIP 性质；

2. 随机采样和随机矩阵： Romberg[66] 证明了在这种情形下当 $m \geq O(k \log^5 N)$ 时矩阵 Φ 满足 RIP 性质；

3. 随机采样和确定性矩阵：在文献[50] 中，李等人构造了一批确定的矩阵 A，再利用随机采样的方法证明了在 $m \geq O(k \log^4 N)$ 的条件下这些矩阵也具有 RIP 性质。

本章的主要贡献在于提出一种新方案来实现完全确定的循环压缩感知矩阵的构造。本章的结构如下：第 2.2 节介绍必要的定义和定理。矩阵的构造将在第 2.3 节中给出。在第 2.4 节中我们对构造的矩阵进行了实验模拟。

2.2 预备工作

2.2.1 RIP 和 MIP 的定义

对一个 $m \times N$ 的矩阵 Φ，任意 k-稀疏向量 x，我们将满足如下事实的最小正实数 δ_k 定义为受限等距常数（RIC），

$$(1 - \delta_k)\|x\|_2 \leq \|\Phi x\|_2 \leq (1 + \delta_k)\|x\|_2.$$

从定义我们可以看出 RIC 的计算需要遍历所有的 k-稀疏向量。

定义1. 我们将矩阵 Φ 中任意两列的内积绝对值的最大值 $\mu(\Phi)$ 定义为这个矩阵的相干数：

$$\mu(\Phi) = \max_{1\leq i\neq j\leq N} \frac{|\langle \Phi_i, \Phi_j\rangle|}{\| \Phi_i \|_2 \| \Phi_j \|_2},$$

其中 Φ_i 表示矩阵 Φ 中的第 i 列。

从稀疏恢复的角度出发，我们自然希望一个矩阵的相干性越小越好。但事实上，它会被矩阵的行数和列数所控制 $\mu \geq \sqrt{\frac{N-m}{(N-1)m}}$，这也就是我们熟知的 Welch 界[76] 。

2.2.2 最坏情形和平均情形

这一节主要为我们所构造的矩阵提供理论保障。下面的两个定理都刻画了关于矩阵最大的稀疏程度 k 使得一个 k-稀疏的向量 x 可以完全（或者几乎完全）从 Φx 中被恢复出来。

当考虑最坏的情形时，我们希望任意 k-稀疏的信号都可以被恢复出来。在过往的研究中，有很多的工作都集中在这个方向[26,51–53,77]。

定理 2.1. (最坏情形)[28] 令 Φ 为 $m \times N$ 的矩阵，它的相干数是 μ，那么在 $k = O(\mu^{-1})$ 的条件下，任意 k-稀疏的信号都可以被恢复。

但是在很多实际应用中，我们并不要求可以恢复所有的 k-稀疏向量[9]。因此在考虑平均的意义下，我们的目标转化为可以恢复尽可能多的 k-稀疏向量，和一般的数学研究一样，我们自然期望：当矩阵的规模趋向于无穷大时，能够恢复的 k-稀疏向量的比例趋向于 1。在这种情形下，稀疏信号一般通过随机选取支撑集后，再随机选取非零元素来实现。

定理 2.2. (平均情形)[9,75] 令 Φ 为 $m \times N$ 的矩阵，它的相干数满足 $\mu(\Phi) \leq A_0(\log N)^{-1}$，其中 A_0 是一个正常数。若 $k \leq \Omega(N/(\|\Phi\|^2 \log N))$ 成立，其中 $||\Phi||$ 表示矩阵 Φ 的算子范数，则随机选取的稀疏信号被恢复的概率大于 $1 - 1/N$。

2.2.3 部分指数和

设 G 是一个大小为 $|G|$ 的有限阿贝尔群，它的单位元记作 1_G。G 的特征 χ 是 G 到由绝对值为 1 所有复数够成的乘法群 U 的一个同态。我们称一个特征是平凡的，如果它将 G 中的所有元素都映成 1。

设 $\mathbb{F}_q$ 是阶为 q 的有限域，其中 $q=p^n$ 是一个素数幂，p 是域 $\mathbb{F}_q$ 的特征。现在，我们去考虑 $\mathbb{F}_q$ 的加法群。设 $\mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_p}:\mathbb{F}_q\to\mathbb{F}_p$ 是 $\mathbb{F}_q$ 到 $\mathbb{F}_p$ 的绝对迹函数，由

$$\mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_p}(\alpha)=\alpha+\alpha^p+\cdots+\alpha^{p^{n-1}}$$

所定义，对任意的 $\alpha\in\mathbb{F}_q$。函数

$$\chi_1(c)=e^{\frac{2\pi i\mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_p}(c)}{p}},\quad \text{对所有 } c\in\mathbb{F}_q$$

是 $\mathbb{F}_q$ 加法群上的特征，其中 $i=\sqrt{-1}$。对 $b\in\mathbb{F}_q$，函数 χ_b 满足 $\chi_b(c)=\chi_1(bc)$ 对任意的 $c\in\mathbb{F}_q$，因此它也是 $\mathbb{F}_q$ 的一个加法特征。因为所有的加法特征都可以由 χ_1 表示出来，我们称 χ_1 为 $\mathbb{F}_q$ 中的本原加法特征。

考虑 $\mathbb{F}_q^*$ 中的乘法循环群，g 是它的生成元，$\mathbb{F}_q$ 上的乘法特征 ψ_j 可以按如下方式定义：

$$\psi_j(g^k)=e^{\frac{2\pi ijk}{q-1}},\quad \text{对 } k=0,\cdots,q-2.$$

对一个非平凡的乘法特征，Katz 在文献[45] 中特别地考虑了子域 $\mathbb{F}_q$ 陪集上的特征和，得到了如下令人振奋的结果：

$$\left|\sum_{x\in\mathbb{F}_q}\psi(x-\alpha)\right|\le(n-1)\sqrt{q},$$

其中 α 使得 $\mathbb{F}_{q^n}=\mathbb{F}_q(\alpha)$。

事实上，更早一些时间，Bombieri 关于加法特征得到了如下更一般的结论：

引理 2.3. [7] 令 $g(x)$ 为 $\mathbb{F}_p$ 上的非常值有理多项式，ψ_p 为素域 $\mathbb{F}_p$ 上的非平凡加法特征。那么我们有如下估计：

$$\left|\sum_{x\in\mathbb{F}_p}\psi_p(g(x))\right|\le(t+\deg(g)_\infty-2)\sqrt{p},$$

其中 $\deg(g)_\infty$ 表示 $g(x)$ 的极点除子的度，t 为 $g(x)$ 中不同极点的个数。

命题 2.4. 令 p 为素数且 $q=p^n$，$\gamma\in\mathbb{F}_q$ 是 $\mathbb{F}_{p^n}$ 的一个生成元，α 为 $\mathbb{F}_q$ 中的任一非零元。当 $n<p$ 时，有理多项式 $\frac{\alpha}{x-\gamma}$ 与任意形如 $h(x)^p-h(x)+c\in\mathbb{F}_q[x]$ 的多项式在 $\mathbb{F}_p$ 上均不等价。在这里如果两个有理多项式在 $\mathbb{F}_p$ 上取值相同，我们称它们是等价的。

证明. 我们不妨设 $h(x)=a_0x^d+a_1x^{d-1}+\cdots+a_{d-1}x+a_d\in\mathbb{F}_q[x]$，存在 $c\in\mathbb{F}_q$ 满足 $\frac{\alpha}{x-\gamma}$ 与 $h(x)^p-h(x)+c\in\mathbb{F}_q[x]$ 在 $\mathbb{F}_p$ 上等价。也就是说，

$$\frac{\alpha}{x-\gamma}\equiv h(x)^p-h(x)+c\pmod{x^p-x}.$$

不失一般性，我们可以假设 $d \leq p-1$，因此我们得到：

$$\begin{aligned} h(x)^p &= a_0^p x^{pd} + a_1^p x^{p(d-1)} + \cdots + a_{d-1}^p x^p + a_d^p \\ &\equiv a_0^p x^d + a_1^p x^{d-1} + \cdots + a_{d-1}^p x + a_d^p \pmod{x^p - x}. \end{aligned}$$

通过整理我们有

$$\begin{aligned} &(x-\gamma)((a_0^p - a_0)x^d + (a_1^p - a_1)x^{d-1} + \cdots \\ &+(a_{d-1}^p - a_{d-1})x + (a_d^p - a_d) + c) \equiv \alpha \pmod{x^p - x}. \end{aligned}$$

Case 1: 当 $d < p-1$ 时，多项式 $(x-\gamma)((a_0^p - a_0)x^d + (a_1^p - a_1)x^{d-1} + \cdots + (a_{d-1}^p - a_{d-1})x + (a_d^p - a_d) + c)$ 的度 $\leq p-1$。因此上面的同余式可以化简为：

$$\begin{aligned} &(x-\gamma)((a_0^p - a_0)x^d + (a_1^p - a_1)x^{d-1} + \cdots \\ &+(a_{d-1}^p - a_{d-1})x + (a_d^p - a_d) + c) = \alpha. \end{aligned}$$

那么最高次项的系数 $a_0^p - a_0$ 为 0。我们得到 $a_0 \in \mathbb{F}_p$。以此类推，可以依次得到 $a_1 \in \mathbb{F}_p, \cdots, a_{d-1} \in \mathbb{F}_p$。因此上面的等式可以约简为 $(x-\gamma)((a_d^p - a_d) + c) = \alpha$，这是不可能的。

Case 2: 当 $d = p-1$ 时，令 $b_i = a_i^p - a_i$ 对 $i = 0, 1, \cdots, d-1$，$b_d = (a_d^p - a_d) + c$。由同余关系我们得到：

$$\begin{aligned} b_1 &= \gamma b_0, \\ b_2 &= \gamma b_1, \\ &\cdots, \\ b_{d-1} &= \gamma b_{d-2}, \\ b_0 + b_d &= \gamma b_{d-1}, \\ -\gamma b_d &= \alpha. \end{aligned}$$

通过求解以上的方程组，我们得到：

$$\begin{aligned} b_0 &= \frac{\alpha}{\gamma - \gamma^p}, \\ b_1 &= \frac{\gamma\alpha}{\gamma - \gamma^p}, \\ &\cdots, \\ b_{d-1} &= \frac{\gamma^{d-1}\alpha}{\gamma - \gamma^p}. \end{aligned}$$

由于 $n < p$，我们由 $n-1 < d$ 可以得到 $\{\frac{\alpha}{\gamma-\gamma^p}, \frac{\gamma\alpha}{\gamma-\gamma^p}, \cdots, \frac{\gamma^{n-1}\alpha}{\gamma-\gamma^p}\}$ 是迹映射 $\mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_p} : \mathbb{F}_q \to \mathbb{F}_p$ 核的一个子集。显然 $\frac{\alpha}{\gamma-\gamma^p}, \frac{\gamma\alpha}{\gamma-\gamma^p}, \cdots, \frac{\gamma^{n-1}\alpha}{\gamma-\gamma^p}$ 在 $\mathbb{F}_p$ 上是线性无关的, 因此它们可以张成整个域 $\mathbb{F}_q$。我们得到迹映射是零映射，这是不可能的。 □

2.3 矩阵的构造

在如下讨论中，我们令 $A_0, A_1, \ldots, A_{N-1}$ 为 $N \times N$ 循环矩阵 A 的行，也就是说

$$A = \begin{pmatrix} A_0 \\ A_1 \\ \vdots \\ A_{N-1} \end{pmatrix},$$

其中 A_0=$(a_0, a_1, \ldots, a_{N-1})$，$A(s,t)$ 记做 A 中第 s 行、第 t 列的元素。对 $\{0, 1, \ldots, N-1\}$ 中的任意子集 $M = \{m_0, m_1, \ldots, m_r\}$，我们定义由 M 所决定的部分循环矩阵如下所示：

$$A_M := \begin{pmatrix} A_{m_0} \\ A_{m_1} \\ \vdots \\ A_{m_r} \end{pmatrix}.$$

2.3.1 基于 Zadoff-Chu 序列族的构造

设 N, γ 是两个正整数，γ 与 N 互素。那么 Zadoff-Chu 序列族中的第 γ 个序列可以表示为：

$$s_\gamma(k) = \begin{cases} e^{-\frac{i\pi\gamma k(k+2g)}{N}}, & N\text{是偶数;} \\ e^{-\frac{i\pi\gamma k(k+1+2g)}{N}}, & N\text{是奇数,} \end{cases}$$

对 $k = 0, 1, \ldots, N-1$，其中 g 是一个整数。在本节中，我们只需去考虑最简单的情形，即 $g = 0$，$\gamma = 1$，N 是偶数。

正如上文所说的一个循环矩阵完全由它的第一行所决定。在第一个构造中，第一行取自 Zadoff-Chu 序列族。令 $a_k := e^{\frac{-i\pi k^2}{N}}$，$N = q^n - 1$ 对 $k = 0, 1, \ldots, N-1$。容易看出，整个矩阵可以由下式表示：$A(s,t) = e^{\frac{-i\pi(s-t)^2}{N}}$ 对任意 $s, t \in \{0, 1, \cdots, N-1\}$。

定理 2.5. 令 q 为素数幂，$\alpha \in \mathbb{F}_{q^n}$ 使得 $\mathbb{F}_{q^n} = \mathbb{F}_q(\alpha)$。设 g 是循环群 $\mathbb{F}_{q^n}^*$ 的一个生成元，n 为大于 1 的正整数。取 $N = q^n - 1$，和

$$M = \{m = \log_g(t - \alpha) : t \in \mathbb{F}_q\}.$$

那么取出的 $q \times N$ 矩阵 $\Phi = \frac{1}{\sqrt{q}} A_M$ 的相干数满足：

$$\mu \leq \frac{n-1}{\sqrt{q}}.$$

证明. 对 $0 \leq i \leq N-1$，记 Φ_i 为矩阵 Φ 的第 i 列。对任意 $0 \leq j, k \leq N-1$，$j \neq k$，第 j 列和第 k 列的内积为：

$$\begin{aligned} |\langle \Phi_j, \Phi_k \rangle| &= \left| \frac{1}{q} \sum_{r=0}^{q-1} e^{\frac{-i\pi(m_r - j)^2}{N}} e^{\frac{i\pi(m_r - k)^2}{N}} \right| \\ &= \left| \frac{1}{q} e^{\frac{i\pi(k^2 - j^2)}{N}} \sum_{r=0}^{q-1} e^{\frac{2i\pi(j-k)m_r}{N}} \right| \\ &= \left| \frac{1}{q} \sum_{r=0}^{q-1} e^{\frac{2i\pi(j-k)m_r}{N}} \right|. \end{aligned}$$

利用 Katz 关于乘法特征和的估计式可得：

$$\mu = \max_{j \neq k} |\langle \Phi_j, \Phi_k \rangle| \leq \frac{n-1}{\sqrt{q}}.$$

□

注2. 这里我们需要指出的是上面构造的技巧是借鉴了[77] 中部分傅里叶矩阵的构造。事实上文献 [77] 中关于改进 *Katz* 估计的结果，也可以类似的应用到本文中。限于篇幅原因，我们就不再详述了。定理 *2.5* 中所构造的矩阵可以完全恢复 k-稀疏信号，当 k 满足如下式子：

$$k < \frac{1}{2}(1 + \frac{\sqrt{q}}{n-1}).$$

这个界与 *DeVore* 关于二元压缩感知矩阵[26] 和 *Xu* 与 *Xu* 关于部分傅里叶矩阵[77] 的理论界是一致的，也就是说从理论上这三类构造的性能是一样的。在后文的实验模拟部分我们甚至可以发现本节中所构造的矩阵在恢复能力可以稍优于前者。

然而，上面得到的理论界与随机矩阵由 RIP 分析得到的界还有很大的差距。但是从实验结果可以看出，我们以上得到的理论界还是过于悲观。下面我们将对平均情形进行分析来合理地解释最后的实验结果。

命题 2.6. *定理 2.5 中构造的矩阵的算子范数为* $\sqrt{\frac{N}{q}}$。

证明. Zadoff-Chu 序列的自相关数为

$$\begin{aligned}
&\sum_{t=0}^{N-1} A(s_1,t)\overline{A(s_2,t)} \\
&= \sum_{t=0}^{N-1} e^{\frac{-i\pi(s_1-t)^2}{N}} e^{\frac{i\pi(s_2-t)^2}{N}} \\
&= e^{\frac{i\pi(s_2-s_1)(s_2+s_1)}{N}} \sum_{t=0}^{N-1} e^{\frac{-2\pi i(s_2-s_1)t}{N}} \\
&= 0.
\end{aligned}$$

因此 $\Phi\Phi^H = \frac{N}{q} I_{q\times q}$，这可以导出 $||\Phi|| = \sqrt{\max_i |\lambda_i(\Phi\Phi^H)|}$，其中 $\lambda_i(A)$ 记为矩阵 A 的第 i 大的特征值。 □

定理 2.7. *当* $k = O(\frac{q}{\log N})$ *时，定理 2.5 中所构造的矩阵可以几乎完全恢复所有的* k*-稀疏向量。*

证明. 这个定理可以很容易地从命题 2.6，定理 2.5 和定理 2.2 中得到。 □

2.3.2 基于 m 序列的构造

令 $N = p^n - 1$，$a_k := \chi(g^k)$ 对 $k = 0, \ldots, N-1$，其中 χ 是 $\mathbb{F}_{p^n}$ 中的本原加法特征，g 为循环群 $\mathbb{F}_{p^n}^*$ 的生成元。整个矩阵可以由下式表示：$A(s,t) = \chi(g^{(t-s)})$。

由引理 2.3 和性质 2.4，我们可以得到如下的界：

定理 2.8. *令* p *为一个奇素数，*$q = p^n$ *其中* $n \le \sqrt{p}$*。设* $\gamma \in \mathbb{F}_q$ *是* $\mathbb{F}_{p^n}$ *的生成元，也就是说* $\mathbb{F}_{p^n} = \mathbb{F}_p(\gamma)$*；*$\alpha$ *为* $\mathbb{F}_q$*中的任意非零元。那么*

$$\left| \sum_{x\in\mathbb{F}_p} \psi_p(\mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_p}(\frac{\alpha}{\mathrm{x}-\gamma})) \right| \le (n-1)\sqrt{p}.$$

证明. 函数

$$\mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_p}(\frac{\alpha}{\mathrm{x}-\gamma}) = \sum_{\mathrm{i}=0}^{\mathrm{n}-1} \frac{\alpha^{\mathrm{p^i}}}{\mathrm{x^{p^i}} - \gamma^{\mathrm{p^i}}}$$

在 $\mathbb{F}_p$ 上等价于

$$\sum_{i=0}^{n-1} \frac{\alpha^{p^i}}{x-\gamma^{p^i}}.$$

函数 $\sum_{i=0}^{n-1} \frac{\alpha^{p^i}}{x-\gamma^{p^i}}$ 有一个极点除子。因为 $\{\gamma, \gamma^p, \cdots, \gamma^{p^{n-1}}\}$ 在 Frobenius 的作用下是封闭的，它的度在 $\mathbb{F}_p$ 上 $\leq n$。由命题 2.4，$\mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_p}(\frac{\alpha}{x-\gamma})$ 在 $\mathbb{F}_p$ 上不是常值函数。因此由引理 2.3，我们得到如下估计：

$$\left| \sum_{x\in\mathbb{F}_p} \psi_p(\mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_p}(\frac{\alpha}{x-\gamma})) \right| \leq (n-1)\sqrt{p}.$$

□

定理 2.9. 令 p 为奇素数，n 为正整数使得 $1 < n \leq \sqrt{p}$。设 $N = p^n - 1$，g 是 $\mathbb{F}_{p^n}$ 中的本原元，γ 是 $\mathbb{F}_{p^n}$ 中的生成元。采样子集由下式所决定：

$$M = \{m = \log_g(t-\gamma) : t \in \mathbb{F}_p\},$$

那么上面所得到的矩阵的相干数为：

$$\mu \leq \frac{n-1}{\sqrt{p}}.$$

证明. 对任意 $0 \leq j, k \leq N-1$，$j \neq k$，矩阵 Φ 的第 j 和第 k 列的内积为：

$$\begin{aligned}
|\langle \Phi_j, \Phi_k \rangle| &= \left| \frac{1}{p} \sum_{r=0}^{p-1} \chi(g^{(j-m_r)}) \overline{\chi}(g^{(k-m_r)}) \right| \\
&= \left| \frac{1}{p} \sum_{r=0}^{p-1} \chi(g^{-m_r}(g^j - g^k)) \right| \\
&= \left| \frac{1}{p} \sum_{r=0}^{p-1} \chi_{(g^j-g^k)}(g^{-m_r}) \right| \\
&= \left| \frac{1}{p} \sum_{t\in\mathbb{F}_p} \chi_{(g^j-g^k)}(\frac{1}{t-\gamma}) \right|.
\end{aligned}$$

特别的，我们在定理 2.8 中取 ψ_p 为 $\mathbb{F}_p$ 上的本原加法特征，$\alpha = g^j - g^k$，可以得到如下的估计：

$$|\langle \Phi_j, \Phi_k \rangle| = \left| \frac{1}{p} \sum_{t\in\mathbb{F}_p} \chi_{(g^j-g^k)}(\frac{1}{t-\gamma}) \right| \leq \frac{n-1}{\sqrt{p}}.$$

□

和上文中类似，我们同样也去考察这个矩阵在平均情形下的恢复能力。

命题 2.10. *定理 2.9 中得到的矩阵的算子范数为* $\sqrt{\frac{N+1}{p}}$。

证明. m 序列的自相关数为：

$$\begin{aligned}
&\sum_{t=0}^{N-1} A(s_1,t)\overline{A(s_2,t)} \\
&= \sum_{t=0}^{N-1} \chi(g^{(t-s_1)})\overline{\chi(g^{(t-s_2)})} \\
&= \sum_{t=0}^{N-1} \chi(g^{t-s_1} - g^{t-s_2}) \\
&= \sum_{t=0}^{N-1} \chi(g^t(g^{-s_1} - g^{-s_2})) \\
&= \sum_{t=0}^{N-1} \chi_{g^{-s_1}-g^{-s_2}}(g^t) \\
&= -\chi_{g^{-s_1}-g^{-s_2}}(0) = -1.
\end{aligned}$$

因此 $\Phi\Phi^H = \frac{1}{p}(NI_{p\times p} - J_{p\times p})$，这可以导出我们最后的结论 $\max_i |\lambda_i(\Phi\Phi^H)| = \frac{N+1}{p}$。

□

定理 2.11. *当* $k = O(\frac{p}{\log N})$ *时，定理 2.9 中得到的矩阵可以几乎完全恢复所有的* k*-稀疏向量。*

2.4 实验结果

在本节中，我们主要将上文中所得到的矩阵和随机矩阵的恢复能力进行比较。为了保证参数的一致性，随机矩阵将从同样的循环矩阵中随机的选择相同行数来形成。为了计算的方便，我们统一采用了正交匹配追踪算法（OMP）[74] 来进行实验。我们对稀疏度从 1 到 20 都进行了 100 次实验，最后计算出它们的恢复成功率。

例3. *(Zadoff-Chu 序列)* 取 $q = 29$，$N = q^2 - 1 = 840$。在有限域 $\mathbb{F}_{29^2}$ 中，α 为其中的本原

元且满足 $\alpha^2+24\alpha+2=0$。我们有：

$$\begin{aligned}M=&\{\log_\alpha(x-\alpha):x\in\mathbb{F}_{29}\}\\=&\{421,547,324,576,323,29,647,312,663,94,\\&524,51,602,821,455,488,310,285,170,292,175,\\&709,238,219,496,626,327,228,703\}.\end{aligned}$$

由定理 *2.5* 我们可以得到一个 29×840 的矩阵。

图 1 为相应的实验结果。

例4. *(m 序列)* 取 $q=29$，$N=q^2-1=840$。在有限域 $\mathbb{F}_{29^2}$ 中，α 为其中的本原元且满足 $\alpha^2+24\alpha+2=0$。我们有：

$$\begin{aligned}M=&\{\log_\alpha(x-\alpha):x\in\mathbb{F}_{29}\}\\=&\{421,547,324,576,323,29,647,312,663,94,\\&524,51,602,821,455,488,310,285,170,292,175,\\&709,238,219,496,626,327,228,703\}.\end{aligned}$$

图 2 为相应的实验结果。

2.5　小结

本章主要利用了部分指数和的估计，给出了两类循环压缩感知矩阵的构造。从理论角度我们证明了在最坏情形和平均情形下，它们从阶上都是最优的。同时实验结果也支撑了我们的理论证明。遗憾的是，因为我们的构造过度依赖于部分指数和的结果，尤其是有限域的结构，因此在参数的选择上会有非常大的局限。在后续的研究中，如何将代数的结构从构造出隐去，更多的引入一些伪随机的思想来构造相应的矩阵从而突破当前的限制是一个可以考虑的研究方向。

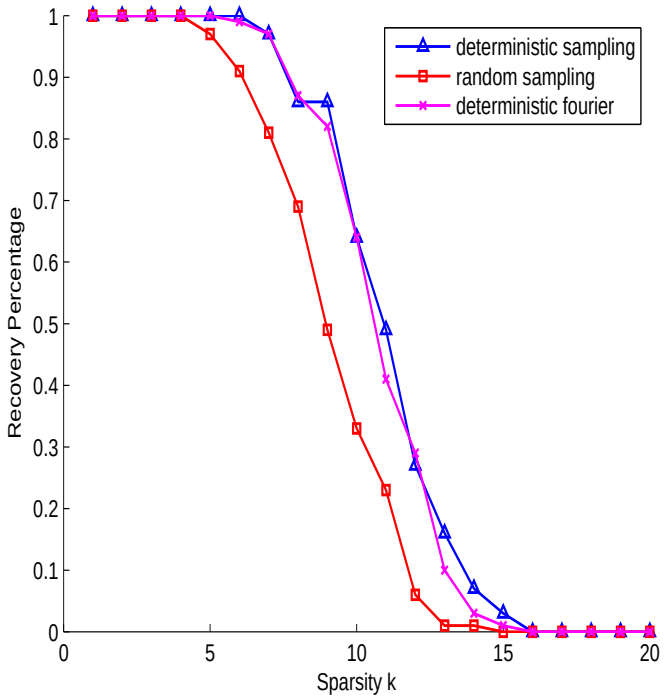


图 2-1　基于 Zadoff-Chu 序列的 29×840 随机、确定部分循环矩阵和 29×840 部分傅里叶矩阵的实验结果

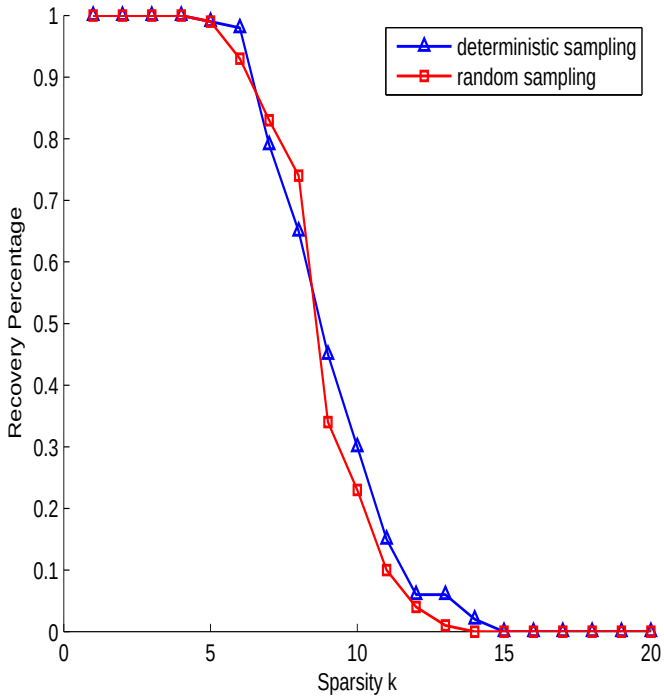


图 2-2　基于 m 序列的 29×840 随机和确定部分循环矩阵的实验结果

3 多重常重码

3.1 介绍

现代密码体系主要依赖于单向函数的应用，但是通常的单向函数总是基于还未被证明的猜想来保证它的安全性。由 Pappu 等人在文献[58] 中提出的物理不可克隆函数（PUF）提供了一种认证消耗低和抵抗物理攻击的新选择。近年来，物理不可克隆函数的研究已经成为在无线电频率识别和智能卡领域[18,37,58,72] 的一种潮流。在文献[19] 中，Chee 等人提出了多重常重码（MCWC），从而为设计环形物理不可克隆函数和编码理论建立起了桥梁。在一个多重常重码中，每一个码字是一个长为 mn 的二元向量，它可以被划分成 m 个相同大小的部分，且每部分的重量都恰好是 w。这个定义事实上是常重码（CWC）（$m = 1$）和双重常重码（$m = 2$）的自然推广[44,49]。

关于多重常重码的研究目前还处在初始阶段。在文献[14] 中，Chee 等人推广了 Johnson 关于常重码的相关研究[44] 给出了多重常重码的一些基本事实。他们进一步还证明了在相差一个常数因子的意义下，他们的界是渐近紧的。在文献[16] 中，Chee 等人给出了几类新的最优多重常重码的构造。进一步地，在文献[15] 中，他们在极小距离为 $2mw - 2$ 时，证明了 Johnson 界是渐近精确的。

在本章中，我们将继续关于多重常重码上下界的研究，我们的主要贡献如下所述：

- 我们将推广 Agrell 等人在文献[2] 中的方法，从而改进文献[14] 中导出的第三型 Johnson 界。同时我们得到了多重常重码的 Gilbert-Varshamov 界，并且证明了在渐近意义下可以改进文献[14] 中通过级联的方法得到的下界。

- 我们得到了两类最优多重常重码的渐近存在性。其中的第一类我们推广了文献[15] 中的结果，考虑不同的部分拥有不同的重量。另一类我们考虑了极小距离为 $2mw - 2w$ 的多重常重码的 Johnson 界是紧的。

本章的结构如下：在第 3.2 节中，我们给出一些基本的定义和符号，并且总结之前的工作。在第 3.3 节中，我们主要讨论多重常重码的上下界。在第 3.4 节中，利用图分解的

工具来证明两类最优多重常重码的渐近存在性。在第 3.5 节中，我们考虑了重量为 4、极小距离为 6 的多重常重码。

3.2 定义和记号

3.2.1 多重常重码

令 m，N 为正整数，X 是大小为 N 的集合。设 X 可以被分成 $X = X_1 \cup X_2 \cup \cdots \cup X_m$，其中 $|X_i| = n_i$，$i = 1, 2, \ldots, m$。如果每个码字在由 X_1 标记的坐标上重量恰为 w_1，由 X_2 标记的坐标上重量恰为 w_2，以此类推，我们称一个 (N, d) 码 $\mathcal{C} \subseteq \mathbb{Z}_2^X$ 是多重常重的，并且记为 MCWC$(w_1, n_1; w_2, n_2; \cdots; w_m, n_m; d)$。特别地，当 $w_1 = w_2 = \cdots = w_m = w$ 且 $n_1 = n_2 = \cdots = n_m = n$，我们可以将 $N = mn$ 长的多重常重码简记为 MCWC(m, n, d, w)。当 $m = 1$，这个码就是常重码，记为 CWC(n, d, w)。

记一个码 $(n, d)_q$ 的最大容量为 $A_q(n, d)$。当 $q = 2$ 时，简记为 $A(n, d)$。一个多重常重码 MCWC$(w_1, n_1; w_2, n_2; \ldots; w_m, n_m; d)$ 的最大容量记为 $T(w_1, n_1; w_2, n_2; \ldots; w_m, n_m; d)$。特别地，用 $M(m, n, d, w)$ 和 $A(n, d, w)$ 来分别表示 MCWC(m, n, d, w) 和 CWC(n, d, w) 的最大容量。如果一个码可以达到最大容量，那么我们称之为最优的。

下面我们将不加证明的给出一些多重常重码的基本结果，详细的证明过程可以参考文献[14]。

命题 3.1. *[14]* 设 $q \leq A(n, d_1, w)$，我们有

$$M(m, n, d_1 d_2, w) \geq A_q(m, d_2).$$

特别地，$M(m, qw, 2d, w) \geq A_q(mw, d)$。

由于多重常重码作为常重码的一种自然推广，著名的 Johnson 界可以做如下形式的推广：

命题 3.2. *[14]*

$$T(w_1, n_1; w_2, n_2; \ldots; w_m, n_m; d)$$
$$\leq \lfloor \frac{n_i}{w_i} T(w_1, n_1; \ldots; w_i - 1, n_i - 1; \ldots; w_m, n_m; d) \rfloor, \tag{3-1}$$
$$T(w_1, n_1; w_2, n_2; \ldots; w_m, n_m; d)$$
$$\leq \lfloor \frac{n_i}{n_i - w_i} T(w_1, n_1; \ldots; w_i, n_i - 1; \ldots; w_m, n_m; d) \rfloor, \tag{3-2}$$

$$T(w_1, n_1; w_2, n_2; \ldots; w_m, n_m; d)$$

$$\leq \lfloor \frac{u}{w_1^2/n_1 + w_2^2/n_2 + \cdots + w_m^2/n_m - \lambda} \rfloor, \tag{3-3}$$

其中 $d = 2u$，$\lambda = w_1 + w_2 + \cdots + w_m - u$。

命题 3.3. *[14]*

$$M(m, n, d, w) \leq \lfloor \frac{n^m}{w^m} M(m, n-1, d, w-1) \rfloor, \tag{3-4}$$

$$M(m, n, d, w) \leq \lfloor \frac{n^m}{(n-w)^m} M(m, n-1, d, w) \rfloor, \tag{3-5}$$

$$M(m, n, d, w) \leq \lfloor \frac{d/2}{d/2 + mw^2/n - mw} \rfloor. \tag{3-6}$$

3.2.2 图分解

我们记有限集合 X 中的所有有序对为 $\overline{\binom{X}{2}}$，用三元组 $G = (V, C, E)$ 来表示一张边着色的有向图，其中 V 表示顶点集合，C 表示颜色集合，E 是集合 $\overline{\binom{X}{2}} \times C$ 的一个子集（E 称为边集）。我们用 $K_n^{(r)}$ 来表示 n 个顶点、r 着色的完全有向图，也就是说 $|V| = n$，$|C| = r$，$E = \overline{\binom{X}{2}} \times C$。

我们称集族 $\mathcal{F}$ 是一张边着色的有向图 K 的分解，如果 K 中的每条边都恰好属于集族 $\mathcal{F}$ 的一个元素。给定一族边着色的有向图 $\mathcal{G}$，如果集族 $\mathcal{F}$ 中的每一张边着色的有向图都与 $G \in \mathcal{G}$ 同构，我们称集族 $\mathcal{F}$ 是一张边着色的有向图 K 的 $\mathcal{G}$-分解。在文献[48] 中，Lamken 和 Wilson 证明了对一族给定的有向图，图 $K_n^{(r)}$ 分解的渐近存在性。为了更好地介绍他们的相关结果，我们还需要引进一些记号。

设 $G = (V, C, E)$ 为一张边着色的有向图，其中 $|C| = r$。

令 $((u, v), c) \in E$ 表示一条从 u 指向 v 并且染了颜色 c 的边。对任意顶点 u 和颜色 c，我们将打入和离开顶点 u 且染色为 c 的边的数目分别记为顶点 u 关于颜色 c 的入度和出度。因此，我们可以对顶点 u 定义一个 $2r$ 长的向量 $\tau(u, G)$ 来反映它在图 G 中的出度和入度，其中 $\tau(u, G) = (\mathrm{in}_1(u, G), \mathrm{out}_1(u, G), \ldots, \mathrm{in}_r(u, G), \mathrm{out}_r(u, G))$。定义 $\alpha(\mathcal{G})$ 为整数 t 的最大公因子使得 $2r$ 长向量 $(t, t, \ldots, t)$ 是顶点 u 所对应的向量 $\tau(u, G)$ 的非负整线性组合。

对任意图 $G = (V, C, E) \in \mathcal{G}$，令 $\mu(G)$ 是由 $\mu(G) = (m_1(G), m_2(G), \ldots, m_r(G))$ 定义的一个 r 长的向量，其中 $m_i(G)$ 表示图 G 中颜色为 i 的边的数目。类似上文，定义 $\beta(\mathcal{G})$ 为整数 m 的最大公因子使得 r 长向量 $(m, m, \ldots, m)$ 是图 G 所对应的的向量 $\mu(G)$ 的非负整线性组合。如果 $(1, 1, \ldots, 1)$ 可以由 $\mu(G)$ 的正有理线性组合表出，我们称集族 $\mathcal{G}$ 是可容许的。

定理 3.4. *(Lamken，Wilson[48])* 如果集族 $\mathcal{G}$ 是可容许的，那么存在常数 $n_0 = n_0(\mathcal{G})$ 使得在 $n(n-1) \equiv 0 \pmod{\beta(\mathcal{G})}$ 和 $n-1 \equiv 0 \pmod{\alpha(\mathcal{G})}$ 成立时，对任意 $n \geq n_0$，$K_n^{(r)}$ 均存在 $\mathcal{G}$-分解。

在同一篇论文中，他们还将上面的定理推广到多重的情形：对图 $K_n^{[\lambda_1,\lambda_2,\ldots,\lambda_r]}$ 进行 $\mathcal{G}$-分解，其中 $K_n^{[\lambda_1,\lambda_2,\ldots,\lambda_r]}$ 表示定义在 n 个顶点上的有向图，对任意的有向点对 (x,y) 和颜色 i 都恰有 λ_i 条边。

令 $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_r)$ 是由正整数构成的向量。同上，令 $\alpha(\mathcal{G};\boldsymbol{\lambda})$ 定义为最小的正整数 t 使得常向量 $t\boldsymbol{\lambda}$ 是顶点 u 对应的向量 $\tau(u,G)$ 的非负整线性组合，$\beta(\mathcal{G};\boldsymbol{\lambda})$ 定义为最小的正整数 m 使得常向量 $m\boldsymbol{\lambda}$ 是图 G 对应的向量 $\mu(G)$ 的非负整线性组合。如果 $\boldsymbol{\lambda}$ 可以由 $\mu(G)$ 的正有理线性组合表出，我们称集族 $\mathcal{G}$ 是 $\boldsymbol{\lambda}$-可容许的。

定理 3.5. *(Lamken，Wilson[48])* 如果集族 $\mathcal{G}$ 是 $\boldsymbol{\lambda}$-可容许的，其中 $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_r)$，那么存在常数 $n_0 = n_0(\mathcal{G}, \boldsymbol{\lambda})$ 使得在 $n(n-1) \equiv 0 \pmod{\beta(\mathcal{G};\boldsymbol{\lambda})}$ 和 $n-1 \equiv 0 \pmod{\alpha(\mathcal{G};\boldsymbol{\lambda})}$ 成立时，对任意 $n \geq n_0$，$K_n^{[\lambda_1,\lambda_2,\ldots,\lambda_r]}$ 均存在 $\mathcal{G}$-分解。

3.3 多重常重码上下界的改进

3.3.1 由球面码导出的上界

本节我们从球面码的定义开始说起，不同于一般的经典码，球面码定义在欧式空间上。球面码是 $S(n)$ 上的一个有限集合，其中 $S(n) := \{\mathbf{x} \in R^n : \|\mathbf{x}\| = 1\}$。这里 $\|\cdot\|$ 就是我们通常意义下的欧式范数。我们用欧式距离 $d_E(\mathbf{c}_1, \mathbf{c}_2) := \|\mathbf{c}_1 - \mathbf{c}_2\|$ 来定义两个码字之间的距离。事实上，由于每个码字都取自单位球面，我们可以用向量之间的夹角 ϕ 或者三角函数值 s 来等价刻画两者之间的距离，通过简单的计算可以得到如下关系式：

$$s = \cos\phi = 1 - \frac{d_E^2}{2}.$$

在球面码中，我们一般用参数 s 来衡量码字之间的距离。同样的，我们定义距离为 s 的 n-维球面码的最大码字容量为 $A_S(n,s)$。

与一般的编码理论不同，当 $s \leq 0$ 时，$A_S(n,s)$ 的确切值早在六十年代就由一批匈牙

利数学家所完全决定[1,23,31,60,68]：

$$\begin{aligned} &A_S(n,s) = \lfloor 1 - \tfrac{1}{s} \rfloor, \quad \text{如果 s} \le -\tfrac{1}{\text{n}}; \\ &A_S(n,s) = n+1, \qquad \text{如果 } -\tfrac{1}{\text{n}} \le \text{s} < 0; \\ &A_S(n,0) = 2n. \end{aligned}$$

在继续后面的讨论之前，我们先注意到一个非常简单但很实用的观察：我们可以将一个二元码在适当的映射下变成一个球面码。因此球面码的上界可以导出二元码的一个上界。令人惊讶的是，在一些情况下，这样给出的上界可以改进之前的结果。

定义

$$\mathcal{H}(n) = \{0,1\}^n,$$

$$\mathcal{M}(m,n,w) = \{\mathbf{x} \in \mathcal{H}(mn) : \mathbf{x} \cdot \mathbf{u}_i = w, \text{ 对 } 1 \le \text{i} \le \text{m}\},$$

其中 $\mathbf{u}_i = \mathbf{e}_i \otimes \mathbf{j}_n$，$\mathbf{e}_i$ 是 m-维单位向量，$\mathbf{j}_n$ 是 n-维全 1 向量。那么 $\mathcal{H}(n) = \{0,1\}^n$ 的任意子集是一个二元码，$\mathcal{M}(m,n,w)$ 的子集是多重常重码。

令 $\Omega(*)$ 是一个从二元汉明空间到欧式空间的映射，它将 $0 \to 1$，$1 \to -1$。$\Omega(*)$ 将 $\mathcal{M}(m,n,w)$ 映射成：

$$\Omega(\mathcal{M}(m,n,w)) = \{\mathbf{x} \in \Omega(\mathcal{H}(mn)) : \mathbf{x} \cdot \mathbf{u}_i = n - 2w \text{ 对 } 1 \le \text{i} \le \text{m}\}.$$

对任意 $\mathbf{x} \in \mathcal{M}(m,n,w)$，$\mathbf{x}$ 满足 $(\Omega(\mathbf{x}) - \mathbf{x}_0) \cdot \mathbf{u}_i = 0$ 和 $\|\Omega(\mathbf{x}) - \mathbf{x}_0\| = r$，其中

$$\mathbf{x}_0 = (1 - \frac{2w}{n})\mathbf{j}_{mn},$$

$$r = 2\sqrt{\frac{mw(n-w)}{n}}.$$

因此 $\Omega(\mathcal{M}(m,n,w))$ 是以 $\mathbf{x}_0$ 为中心的半径为 r 的 $(nm-m)$-维超球面。由上面的分析我们可以得到下面的界：

定理 3.6.

$$\begin{aligned} &M(m,n,2d,w) \le \lfloor \tfrac{d}{b} \rfloor, \qquad \text{如果 b} \ge \tfrac{\text{d}}{\text{nm}-\text{m}+1}, \\ &M(m,n,2d,w) \le m(n-1)+1, \quad \text{如果 } 0 < \text{b} < \tfrac{\text{d}}{\text{n}}, \end{aligned}$$

其中

$$b = d - \frac{mw(n-w)}{n}.$$

证明. 设 $\mathcal{C}$ 是一个 MCWC$(m,n,2d,w)$。对 $\Omega(\mathcal{C})$ 先平移 $-\mathbf{x}_0$，再伸缩 $1/r$。通过上面的分析，得到一个极小距离为 $s=1-\frac{dn}{mw(n-w)}$ 的 $(nm-m)$- 维球面码。

$$\begin{cases} M(m,n,2d,w) \leq A_S(m(n-1),s), & \text{如果 s} \geq -1; \\ M(m,n,2d,w)=1, & \text{如果 s} < -1. \end{cases}$$

□

注5. *定理 3.6 的第一个上界恰好与 Johnson 界 (3-3) 是一致的。当 $0<b<\frac{d}{n}$ 时，第二个上界可以稍微改进命题 3.2 中的 Johnson 界。*

3.3.2 多重常重码的 Plotkin 界

为了文章的完整性，我们仍然给出下面这个众所周知的结果的证明。

命题 3.7. *[2] 设码 $\mathcal{C}$ 是一个 (n,d) 码，那么*

$$|\mathcal{C}| \leq \frac{d/2}{d/2-\sum_{i=1}^{n} f_i(1-f_i)}$$

其中分母必须是正的，f_i 表示整个码中 1 在第 i 位所占的比例。

证明. 证明可以由计数的方法得到，一方面，

$$d_{av}=\frac{1}{M(M-1)}\sum_{c_1,c_2\in C} d(c_1,c_2) \geq d,$$

其中 $M=|\mathcal{C}|$。另一方面，

$$d_{av}=\frac{2M}{M-1}\sum_{i=1}^{n} f_i(1-f_i).$$

那么

$$\frac{2M}{M-1}\sum_{i=1}^{n} f_i(1-f_i) \geq d.$$

□

对多重常重码来说，我们对 f_i 将有更多的限制，因此我们也希望可以得到更好的上界。

定理 3.8.

$$M(m,n,2d,w) \leq \max\{\frac{d}{d-\sum_{i=1}^{mn} f_i(1-f_i)}\} \tag{3-7}$$

其中最大值遍历所有满足以下条件的 f_i $(1 \le i \le mn)$：

$$\begin{aligned} f_1 + f_2 + \cdots + f_n &= w \\ f_{n+1} + f_{n+2} + \cdots + f_{2n} &= w \\ &\vdots \\ f_{(m-1)n+1} + f_{(m-1)n+2} + \cdots + f_{mn} &= w. \end{aligned}$$

证明. 定理的证明可以由多重常重码的定义和命题 3.7 导出。 □

推论 3.9.

$$M(m, n, 2d, w) \le \lfloor \frac{d}{b} \rfloor, \tag{3-8}$$

其中

$$b = d - \frac{mw(n-w)}{n}.$$

证明. 为了得到多重常重码的上界，我们只需去决定当 $f_1 + f_2 + \cdots + f_n = w$ 时，$\sum_{i=1}^{n} f_i^2$ 的最小值。利用拉格朗日乘子法，设 γ 为辅助变量。我们考虑如下函数：

$$g(f_1, f_2, \ldots, f_n, \gamma) = \sum_{i=1}^{n} f_i^2 + \gamma(f_1 + f_2 + \cdots + f_n - w).$$

那么

$$\frac{\partial g}{\partial f_i} = 2f_i + \gamma = 0,$$

$$\frac{\partial g}{\partial \gamma} = \sum_{i=1}^{n} f_i - w = 0.$$

因此当 $f_i = \frac{w}{n}$ 时，原函数达到最小值。将 f_i 以 $\frac{w}{n}$ 代入到 (3-7) 中，我们可以得到 (3-8)。 □

注6. *(3-8) 与命题 3.3中的 (3-6) 是不谋而合的，但是我们注意到事实上 f_i 只能是 $1/M$ 的倍数，因此我们可以将问题从连续的 $[0,1]$ 区间替换成离散的 $\{0, 1/M, 2/M, \ldots, 1\}$，与上面的讨论类似，我们可以得到如下的隐式上界。*

推论 3.10. 如果 $b > 0$，那么

$$M(m, n, 2d, w) \le \lfloor d/b \rfloor,$$

其中

$$b = d - \frac{mw(n-w)}{n} + \frac{nm}{M^2}\{Mw/n\}\{M(n-w)/n\},$$
$$M = M(m, n, 2d, w),$$
$$\{x\} = x - \lfloor x \rfloor.$$

3.3.3 多重常重码的线性规划界

设 X 是一个至少由两个元素组成的有限集合，对整数 $n \geq 1$，令 $\mathcal{R} = \{R_0, R_1, \ldots, R_n\}$ 是 $n+1$ 个 X 中的等价关系。我们将 $(X, \mathcal{R})$ 称为 n 类的结合方案，如果下面的三个条件同时成立：

1. 集族 $\mathcal{R}$ 是 X^2 的一个划分，其中 $R_0 = \{(x, x)|x \in X\}$；
2. 对 $i = 0, 1, \ldots, n$，R_i 的逆 $R_i^{-1} = \{(y, x)|(x, y) \in R_i\}$ 也属于 $\mathcal{R}$；
3. 对任意三个整数 $i, j, k = 0, 1, \ldots, n$，存在 $p_{i,j}^{(k)} = p_{j,i}^{(k)}$ 使得对任意 $(x, y) \in R_k$：

$$|\{z \in X|(x, z) \in R_i, (z, y) \in R_j\}| = p_{i,j}^{(k)}. \tag{3-9}$$

其中 $p_{i,j}^{(k)}$ 称为 $(X, \mathcal{R})$ 的相交数。

任意的等价关系 R_i 可以由邻接矩阵 $D_i \in \mathbb{C}(X, X)$ 来表示：

$$D_i(x, y) = \begin{cases} 1, & (x, y) \in R_i, \\ 0, & (x, y) \notin R_i. \end{cases}$$

我们称线性空间

$$A = \{\sum_{i=0}^{n} \alpha_i D_i | \alpha_i \in \mathbb{C}\}$$

是结合方案 $(X, \mathcal{R})$ 上的 Bose-Mesner 代数。同时，存在一组相互正交的幂等矩阵 $J_0, J_1, \ldots, J_n$ 构成 Bose-Mesner 代数的另一组基。

给定 Bose-Mesner 代数的两组基 $\{D_k\}$ 和 $\{J_k\}$，他们之间的相互表示如下：

$$D_k = \sum_{i=0}^{n} P_k(i) J_i, \quad k = 0, 1, \ldots, n.$$

我们可以定义阶为 $n+1$ 的矩阵 P，它的 (i, k) 位置上的元素是 $P_k(i)$：

$$P = [P_k(i) : 0 \leq i, k \leq n].$$

由于 P 是非奇异的，一定存在唯一的矩阵 Q 使得：

$$PQ = QP = |X|I.$$

我们将 P 和 Q 一起称为结合方案的特征矩阵。

设 $\mathcal{R} = \{R_0, R_1, \ldots, R_n\}$ 是定义在 X 上的一个结合方案。对 X 中的一个非空子集 Y，让我们来定义 Y 关于 $\mathcal{R}$ 的内分布为 $\alpha = (\alpha_0, \alpha_1, \ldots, \alpha_n)$，其中 α_i 由 $\alpha_i = |Y|^{-1}|R_i \cap Y^2|$ 给出。

在文献[25] 中，Delsarte 给出了关于内分布和特征矩阵 Q 的如下重要联系：

定理 3.11. *[25]* αQ 中的分量 αQ_k 都是非负的。

设 w 和 n 都是整数，且 $1 \leq w \leq n$。在 n 维的 Hamming 空间中，我们考虑如下集合：

$$X = \{x \in \mathbb{F}^n | w_H(x) = w\},$$

并且可以定义它们之间的距离关系 $R_0, R_1, \ldots, R_w$：

$$R_i = \{(x, y) \in X^2 | d(x, y) = 2i\}.$$

对给定的 n 和 w，$1 \leq w \leq n/2$，我们称 $(X, \mathcal{R})$ 是一个 Johnson 方案 $J(w, n)$，事实上就是一个二元常重码。

给定整数 k，$0 \leq k \leq w$，我们定义关于变量 u 的 Eberlein 多项式 $E_k(u)$：

$$E_k(u) = \sum_{i=0}^{k} (-1)^i \binom{u}{i} \binom{w-u}{k-i} \binom{n-w-u}{k-i}.$$

定理 3.12. *[25] Johnson 方案* $J(w, n)$ 的特征矩阵 P 和 Q 分别是：

$$P_k(i) = E_k(i),$$

$$Q_i(k) = \frac{\mu_i E_k(i)}{\binom{w}{i}\binom{n-w}{i}},$$

其中 $\mu_i = \frac{n-2i+1}{n-i+1}\binom{n}{i}$。

设 $\mathcal{C}$ 是一个 MCWC$(m, n, 2d, w)$。码 $\mathcal{C}$ 的距离分布为：

$$A_{2i_1, 2i_2, \ldots, 2i_m} := \frac{1}{|\mathcal{C}|} \sum_{\mathbf{c} \in \mathcal{C}} A_{2i_1, 2i_2, \ldots, 2i_m}(\mathbf{c}),$$

其中 $A_{2i_1, 2i_2, \ldots, 2i_m}(\mathbf{c}) := |\{\mathbf{c}_1 \in \mathcal{C} : (\mathbf{c}_1 \oplus \mathbf{c}) \cdot \mathbf{u}_j = 2i_j\}|$, $\mathbf{u}_j := \mathbf{e}_j \otimes \mathbf{j}_n$，$\mathbf{e}_j$ 是 m 维单位向量，$\mathbf{j}_n$ 是 n 维全 1 向量。

推论 3.13. *设码 $\mathcal{C}$ 是一个 MCWC$(m,n,2d,w)$，那么*

$$\sum_{i_1=0}^{w}\sum_{i_2=0}^{w}\cdots\sum_{i_m=0}^{w}Q_{k_1}(i_1)Q_{k_2}(i_2)\cdots Q_{k_m}(i_m)A_{2i_1,2i_2,\ldots,2i_m}\geq 0.$$

证明. 对 $v=1,2,\ldots,m$，如果 $(X^{(v)};R_0^{(v)},\cdots,R_w^{(v)})$ 是一个结合方案，它的相交数是$p_{ijk}^{(v)}$，邻接矩阵是 $D_i^{(v)}$，幂等矩阵是 $J_i^{(v)}$，特征值为 $P_k^{(v)}(i)$ 和 $Q_k^{(v)}(i)$。那么它们的笛卡尔积 $(X^{(1)}\times X^{(2)}\times\cdots\times X^{(m)};R_{i_1\ldots i_m}=R_{i_1}^{(1)}\times\cdots\times R_{i_m}^{(m)},0\leq i_j\leq m$ 对 $1\leq j\leq m)$ 也是一个结合方案，它的特征值是 $Q_{k_1}^{(1)}(i_1)Q_{k_2}^{(2)}(i_2)\cdots Q_{k_m}^{(m)}(i_m)$。而码 $\mathcal{C}$ 是 m 个 Jonhnson 方案的笛卡尔积。结论得证。 □

定理 3.14.

$$M(m,n,2d,w)\leq 1+\lfloor\max\sum_{i_1=0}^{w}\sum_{i_2=0}^{w}\cdots\sum_{i_m=0}^{w}A_{2i_1,\ldots,2i_m}\rfloor,$$

其中

$$A_{2i_1,\ldots,2i_m}\geq 0,$$

$$A_{2i_1,\ldots,2i_m}=0,\ \text{对}\ \sum_{\mathrm{j}=1}^{\mathrm{m}}\mathrm{i_j}<\mathrm{d},$$

$$\sum_{i_1=0}^{w}\sum_{i_2=0}^{w}\cdots\sum_{i_m=0}^{w}Q_{k_1}(i_1)Q_{k_2}(i_2)\cdots Q_{k_m}(i_m)A_{2i_1,2i_2,\ldots,2i_m}\geq 0. \tag{3-10}$$

3.3.4 多重常重码的 GV 界

在本节中，我们考虑当 m 趋于无穷大时，n 是关于 m 的函数的情况下，$M(m,n,d,w)$ 的渐近码率，其中 $d=\lfloor\delta mn\rfloor$，$w=\lfloor\omega n\rfloor$ 对 $0<\delta,\omega<1$。我们按如下方式定义 $\mu(\delta,\omega)$：

$$\mu(\delta,\omega):=\limsup_{m\to\infty}\frac{\log_2 M(m,n,\lfloor\delta mn\rfloor,\lfloor\omega n\rfloor)}{mn}.$$

在文献[14] 中，Chee 等人利用级联的技巧给出了多重常重码的渐近下界。

命题 3.15. *[14] 对 $\delta\leq 1/2$，我们有*

$$\mu(\delta,1/2)\geq 1-H(\delta),$$

其中 $H(x)$ 表示二元熵函数

$$H(x):=-x\log_2 x-(1-x)\log_2(1-x).$$

在本节中，我们首先将推广命题 3.15 中的结果，同时我们也会导出多重常重码的 GV 下界，并且证明后者总是优于前者的。

我们先选取一个可以达到 GV 界的 q 元码。为了方便起见，我们不妨假设 $\frac{1}{\omega}$ 和 δmn 都是整数。

定理 3.16. 对 $\omega \leq 1/2$ 和 $\delta \leq \max\{1/2, 2\omega\}$，我们有

$$\mu_c(\delta,\omega) \geq \omega \log_2(\frac{1}{\omega})(1 - H_{\frac{1}{\omega}}(\frac{\delta}{2\omega})),$$

其中 $H_q(x) := x\log_q(q-1) - x\log_q x - (1-x)\log_q(1-x)$ 对 $0 < x \leq \frac{q-1}{q}$。

证明. 利用级联的技巧我们可以得到 $M(m,n,\delta mn,\omega n) \geq A_{\frac{1}{\omega}}(mwn, \frac{\delta mn}{2})$。由于 $A_q(n,d) \geq q^{(1-H_q(d/n))n}$，那么

$$M(m,n,\delta mn,\omega n) \geq (\frac{1}{\omega})^{(1-H_{\frac{1}{\omega}}(\frac{\delta}{2\omega}))mwn},$$

因此

$$\mu_c(\delta,\omega) \geq \omega \log_2(\frac{1}{\omega})(1 - H_{\frac{1}{\omega}}(\frac{\delta}{2\omega})).$$

□

注7. 事实上，有比达到 *GV* 界更优的代数几何码存在。但是从命题的结果来看，利用代数几何码并不能带来实质性的改进。因此我们从讨论的方便起见，仍然只选取了达到 *GV* 界的码。

对一个 MCWC$(m,n,2d,w)$，它的半径为 $2d-1$ 的 Hamming 球的大小为

$$\sum_{i_1+i_2+\ldots+i_m \leq d-1} \binom{w}{i_1}\binom{n-w}{i_1}\cdots\binom{w}{i_m}\binom{n-w}{i_m}.$$

定理 3.17. 对 $\omega \leq 1/2$ 和 $\delta \leq \max\{1/2, 2\omega\}$，我们有

$$\mu_{GV}(\delta,\omega) \geq H_2(\omega) - \omega H_2(\frac{\delta}{2\omega}) - (1-\omega)H_2(\frac{\delta}{2(1-\omega)}).$$

证明. 由于

$$M(m,n,\delta mn,\omega n) \geq \frac{\binom{n}{\omega n}^m}{\sum_{i_1+i_2+\ldots+i_m \leq \frac{\delta mn}{2}-1}\binom{\omega n}{i_1}\binom{(1-\omega)n}{i_1}\cdots\binom{\omega n}{i_m}\binom{(1-\omega)n}{i_m}}$$

$$\geq \frac{\binom{n}{\omega n}^m}{\sum_{0\leq i\leq \frac{\delta mn}{2}} \binom{\omega mn}{i}\binom{(1-\omega)mn}{i}},$$

我们有

$$\mu_{GV}(\delta,\omega) \geq \frac{\log_2 \frac{2^{nmH_2(\omega)}}{2^{\omega nmH_2(\frac{\delta}{2\omega})}2^{(1-\omega)mnH_2(\frac{\delta}{2(1-\omega)})}}}{mn}$$
$$\geq H_2(\omega) - \omega H_2(\frac{\delta}{2\omega}) - (1-\omega)H_2(\frac{\delta}{2(1-\omega)}).$$

□

在这一节的最后，我们给出上述两个界的比较。

定理 3.18.

$$\mu_{GV}(\delta,\omega) \geq \mu_c(\delta,\omega),$$

等号成立当且仅当 $w=\frac{1}{2}$ 或者 $\delta = 2(\omega-\omega^2)$。

证明. 令

$$f(\delta,\omega) = \mu_{GV}(\delta,\omega) - \mu_c(\delta,\omega)$$
$$= H_2(\omega) - (1-\omega)H_2(\frac{\delta}{2(1-\omega)})$$
$$+\frac{\delta}{2}\log_2(\frac{1}{\omega}-1) - (1-\omega)\log_2(1-\omega).$$

为了方便起见，我们进行变量替换 $x=\frac{\delta}{2}$，我们可以得到

$$f(x,\omega) = -(2-2\omega-x)\log_2(1-\omega) + x\log_2(\tfrac{x}{\omega})$$
$$+(1-\omega-x)\log_2(1-\omega-x).$$

我们将通过两种情况的讨论完成整个证明：$\omega \leq \frac{1}{4}, x \leq \omega$ 和 $\frac{1}{4} < \omega \leq \frac{1}{2}, x \leq \frac{1}{4}$。

(a) $\omega \leq \frac{1}{4}, x \leq \omega$.

当 $x=0$ 时，$f(0,\omega) = -(1-\omega)\log_2(1-\omega) > 0$。

当 $x=\omega$ 时，$f(\omega,\omega) = (3\omega-2)\log_2(1-\omega) - (2\omega-1)\log_2(1-2\omega)$。我们想要去证 $f(\omega,\omega)\geq 0$。由于 $f(0,0)=0$ 和 $f(\frac{1}{4},\frac{1}{4}) = 2-\frac{5}{4}\log_2 3 > 0$，我们需要去说明 $g(\omega)=f(\omega,\omega)$ 是单调递增的。

$$g'(\omega) = 3\log_2(1-\omega) - 2\log_2(1-2\omega) + \frac{\omega}{\omega-1},$$
$$g''(\omega) = \frac{\omega(3-2\omega)}{(\omega-1)^2(1-2\omega)} > 0.$$

由于$g^{'}(0)=0$，$g^{'}(\frac{1}{4})=\frac{5}{3}+3\log_2(\frac{3}{4})>0$，我们得到$g^{'}(\omega)\geq 0$，因此$f(\omega,\omega)\geq 0$。

特别地，$\frac{\partial f(x,\omega)}{\partial x}=\log_2\frac{x(1-\omega)}{\omega(1-\omega-x)}=0$，我们得到$x=\omega-\omega^2$。由于$f(\omega-\omega^2,\omega)=0$，通过上面的分析，最终证明了$f(\delta,\omega)\geq 0$。

(b) $\frac{1}{4}<\omega\leq\frac{1}{2}, x\leq\frac{1}{4}$。

当$x=0$时，$f(0,\omega)=-(1-\omega)\log_2(1-\omega)>0$。

当$x=\frac{1}{4}$时，$f(\frac{1}{4},\omega)=-(\frac{7}{4}-2\omega)\log_2(1-\omega)+\frac{1}{4}\log_2(\frac{1}{4\omega})+(\frac{3}{4}-\omega)\log_2(\frac{3}{4}-\omega)$。我们想要去说明$f(\frac{1}{4},\omega)\geq 0$。由于$f(\frac{1}{4},\frac{1}{4})=-\frac{5}{4}\log_2(\frac{3}{4})-\frac{1}{2}>0$ 和 $f(\frac{1}{4},\frac{1}{2})=0$，我们将证明$f(\frac{1}{4},\omega)$是单调递减的。

$$f^{'}(\frac{1}{4},\omega)=\frac{1}{\ln 2}(2\ln(1-\omega)-\ln(\frac{3}{4}-\omega)+1-\frac{1}{4(1-\omega)}-\frac{1}{4\omega}),$$

$$f^{''}(\frac{1}{4},\omega)=\frac{1}{\ln 2}(\ +\frac{1-2\omega}{4\omega^2(1-\omega)^2})\geq 0.$$

由于$f^{'}(\frac{1}{4},\frac{1}{4})=\frac{1}{\ln 2}(\ln(\frac{9}{8})-\frac{1}{3})<0$ 和 $f^{'}(\frac{1}{4},\frac{1}{2})=0$，我们得到$f^{'}(\frac{1}{4},\omega)\leq 0$，因此$f(\frac{1}{4},\omega)\geq 0$。

等号成立的证明部分与情形 1 是类似的。 □

3.4 两类最优码

在文献[15]中，Chee 等人证明了几类常重复合码、非二元常重码和多重常重码的 Johnson 界是渐近精确的。特别地，他们证明了在极小距离为$2mw-2$时，上界 (3-1) 是渐近精确的。

定理 3.19. *(Chee 等人[15])* 对固定的m和w，总是存在一个正整数n_0使得当$n-1\equiv 0 \pmod{w^2}$时，对任意$n\geq n_0$均有

$$M(m,n,2mw-2,w)=\frac{n(n-1)}{w^2}.$$

在本节中，我们将定理 3.19 作如下两方面的推广：1、w_i不完全相同；2、极小距离为$2mw-2w$。

3.4.1 极小距离为$2\sum_{i=1}^{m}w_i-2$的最优多重常重码

不妨设$w_1\geq w_2\geq\cdots\geq w_m$都是非负整数。令$w=\sum_{i=1}^{m}w_i$。由 Johnson 界 (3-1) 可以

得到

$$T(w_1,n;w_2,n;\ldots;w_m,n;2w-2)\leq\begin{cases}\frac{n(n-1)}{w_1(w_1-1)}, & \text{如果 } w_1>w_2;\\ \frac{n(n-1)}{w_1^2}, & \text{如果 } w_1=w_2.\end{cases}$$

我们将证明这个上界是渐近紧的。为了利用定理 3.4，我们首先定义集族 $\mathcal{G}$，用 $[m]$ 中的 m^2 个有序点对来标记颜色，定义 $\overline{w}=[w_1,w_2\ldots,w_m]$。令 $G(\overline{w})$ 是顶点如下的有向图：

$$V(G(\overline{w}))=W_1\cup W_2\cup\cdots\cup W_m \tag{3-11}$$

其中 W_i 互不相交的顶点集合，并且 $|W_i|=w_i$。这里对任意 $x,y\in V(G(\overline{w}))$，定义由 x 到 y 这条边上的颜色为 (i,j)，其中 $x\in W_i$，$y\in W_j$。因此在 $G(\overline{w})$ 中，当 $i\neq j$ 时，总共有 w_iw_j 条着色为 (i,j) 的边；有 $w_i(w_i-1)$ 条着色为 (i,i)。对任意 $i,j\in[m]$，定义 G_{ij} 为一张由两个顶点和一条着色为 (i,j) 的边所得到的有向图。我们到了来正式定义 $\mathcal{G}(\overline{w})$ 的时候，我们依据 $w_1=w_2$ 是否成立，分两种情况来讨论：

1. 当 $w_1>w_2$ 时，我们有 $w_1(w_1-1)\geq w_1w_2$。令 r 为最大的正整数使得 $w_1-1=w_2=\cdots=w_r$ 成立。那么 $\mathcal{G}(\overline{w})=\{G(\overline{w})\}\cup\{G_{ij}:(i,j)\in([m]\times[m])\backslash\{(1,i),(i,1):1\leq i\leq r\}\}$。

2. 当 $w_1=w_2$ 时，我们有 $w_1w_2>w_1(w_1-1)$。令 r 为最大的正整数使得 $w_1=\cdots=w_r$。那么 $\mathcal{G}(\overline{w})=\{G(\overline{w})\}\cup\{G_{ij}:(i,j)\in([m]\times[m])\backslash\overline{\binom{[r]}{2}}\}$。

命题 3.20. *如果 $K_n^{(m^2)}$ 存在 $\mathcal{G}(\overline{w})$-分解，那么*

$$T(w_1,n;\ldots;w_m,n;2w-2)=\begin{cases}\frac{n(n-1)}{w_1(w_1-1)}, & \text{如果 } w_1>w_2;\\ \frac{n(n-1)}{w_1^2}, & \text{如果 } w_1=w_2.\end{cases}$$

证明. 令 V 为 $K_n^{(m^2)}$ 的顶点集，$\mathcal{F}$ 为 $K_n^{(m^2)}$ 的 $\mathcal{G}(\overline{w})$-分解。令 $X=\{1,2,\ldots,m\}\times V$。我们按照如下方式来生成每一个码字。对任意 $F\in\mathcal{F}$ 与 $G(\overline{w})$ 同构的，存在唯一的顶点分解 $V(F)=\cup_{i=1}^m S_i$ 使得当 $x\in S_i$，$y\in S_j$ 时，F 中的每条由 x 指向 y 的边着 (i,j) 色。构造一个码字 $\mathbf{u}$ 使得当 $x\in S_i$ 时，$\mathbf{u}_{(i,x)}=1$；否则 $\mathbf{u}_{(i,x)}=0$。由于 $|S_i|=w_i$，这个码确实是一个参数为 $(w_1,n;\ldots;w_m,n;d)$ 的多重线性码。又注意到每一条着色的有向边都至多在 $\mathcal{F}$ 中出现一次，任意两个码字 $\mathbf{u}$ 和 $\mathbf{v}$ 的支撑集满足 $|\mathrm{supp}(\mathbf{u})\cap\mathrm{supp}(\mathbf{v})|\leq 1$。因此这个码的极小距离至少为 $2w-2$。

最后我们计算 $\mathcal{F}$ 中与 $G(\overline{w})$ 同构的有向图的数目，容易得到：当 $w_1 > w_2$ 时，$m = \frac{n(n-1)}{w_1(w_1-1)}$；否则 $m = \frac{n(n-1)}{w_1^2}$。 □

注意到 $m_{(i,j)}(G(\overline{w})) = w_i w_j$，$i \neq j, m_{(i,i)}(G(\overline{w})) = w_i(w_i-1)$，$m_{(i,j)}(G_{ij}) = 1$，我们有

$$\beta(\mathcal{G}(\overline{w})) = \begin{cases} w_1(w_1-1), & \text{如果 } w_1 > w_2; \\ w_1^2, & \text{如果 } w_1 = w_2. \end{cases}$$

又由于当 $i \neq j$ 时，$\mathrm{in}_{(i,j)}(G(\overline{w})) = w_j, \mathrm{out}_{(i,j)}(G(\overline{w})) = w_i$；$\mathrm{in}_{(i,i)}(G(\overline{w})) = \mathrm{out}_{(i,i)}(G(\overline{w})) = w_i - 1$。容易计算得到

$$\alpha(\mathcal{G}(\overline{w})) = \begin{cases} w_1(w_1-1), & \text{如果 } w_1 > w_2; \\ w_1, & \text{如果 } w_1 = w_2. \end{cases}$$

最后我们利用定理 3.4，我们得到如下的结论：

定理 3.21. 设 $w_1 \geq w_2 \geq \cdots \geq w_m$，$w = \sum_{i=1}^{m} w_i$，存在正整数 n_0 使得在 $w_1 > w_2$ 时满足 $n-1 \equiv 0 \pmod{w_1(w_1-1)}$；否则，满足 $n-1 \equiv 0 \pmod{w_1^2}$ 时，对任意 $n \geq n_0$，都有：

$$T(w_1, n; \ldots; w_m, n; 2w-2) = \begin{cases} \frac{n(n-1)}{w_1(w_1-1)}, & \text{如果 } w_1 > w_2; \\ \frac{n(n-1)}{w_1^2}, & \text{如果 } w_1 = w_2. \end{cases}$$

3.4.2 极小距离为 $2mw - 2w$ 的最优多重常重码

我们首先将建立 α-可分解平衡不完全区组设计与最优多重常重码之间的联系。

命题 3.22. 如果 α-可分解平衡不完全区组设计 (v, k, λ) 存在，那么 $M(m, n, d, w) = v$，其中 $m = \frac{\lambda(v-1)}{\alpha(k-1)}$，$n = \frac{\alpha v}{k}$，$d = 2\lambda\frac{v-k}{k-1}$，$w = \alpha$。

证明. 由 Johnson 界 (3-3) 可以得到 $M(m, n, d, w) \leq v$，其中 $m = \frac{\lambda(v-1)}{\alpha(k-1)}$，$n = \frac{\alpha v}{k}$，$d = 2\lambda\frac{v-k}{k-1}$，$w = \alpha$。

设 $(X, \mathcal{B})$ 是 α-可分解平衡不完全区组设计 (v, k, λ)。由于在 $\mathcal{B}$ 中共有 $\frac{\lambda(v-1)}{\alpha(k-1)}$ 个 α-平行类，每个平行类恰包含 $\frac{\alpha v}{k}$ 个区组，我们可以将这些区组排进一个 $m \times n$ 的阵列中，其中 $m = \frac{\lambda(v-1)}{\alpha(k-1)}$，$n = \frac{\alpha v}{k}$，使得每一行中的区组恰好构成一个 α-平行类。按照如下方式构造对应的码：对任意 $x \in X$，构造一个码字 $\mathbf{u}$，其中当在 (i, j) 上的区组包含 x 时令 $\mathbf{u}_{(i,j)} = 1$，

否则令 $\mathbf{u}_{(i,j)}=0$。由于每个点在每一行中都出现 α 次，上面我们构造的码是一个码字个数为 v 的多重常重码 (m,n,d,α)。由于 X 中任意两个不同的点恰好同时出现在 λ 个区组中，任意两个码字的支撑集恰好相交于 λ 位，因此码的极小距离为 $d=2(mw-\lambda)=2\lambda\frac{v-k}{k-1}$。□

在下面的行文中，我们只需要说明对满足 $v\equiv 1 \pmod{k-1}$ 充分大的 v，当 $\alpha=\lambda$ 和 $k\mid\alpha$ 时，α-可分解平衡不完全区组设计 (v,k,λ) 存在。

我们首先定义 $r=k^2-k$ 着色的有向图集合 $\mathcal{G}$。我们用 $[k-1]$ 中的单点集 (i)，$i=1,2,...,k-1$ 和 $(k-1)^2$ 个有序点对来标记颜色。令 $\boldsymbol{\lambda}$ 为一个 k^2-k 长的取值为 λ 的常值向量。对任意非负整数 $(k-1)$-组 $\mathbf{t}=(t_1,t_2\ldots,t_{k-1})$，其和恰好为 k，令 $G(\mathbf{t})$ 为定义在 $k+1$ 个顶点上的有向图：

$$V(G(\mathbf{t}))=\{w\}\cup T_1\cup T_2\cup\cdots\cup T_{k-1} \tag{3-12}$$

其中 T_i 为互不相交的顶点集，满足 $|T_i|=t_i$，w 是 T_i 之外的另一个顶点。这里对任意不相同的顶点 $x,y\in V(G(\mathbf{t}))$，由 x 指向 y 的边着 (i,j) 色，其中 $x\in T_i$，$y\in T_j$；由 w 指向 $x\in T_i$ 的边着 (i)。令 $\mathcal{G}$ 为所有满足条件的图 $G(\mathbf{t})$ 的集合。

命题 3.23. *设 $r=k^2-k$，$m=\frac{v-1}{k-1}$。如果 r 着色的有向图 $K_m^{[\lambda,\lambda,\ldots,\lambda]}$ 存在 $\mathcal{G}$-分解，那么 λ-可分解平衡不完全区组设计 $(m(k-1)+1,k,\lambda)$ 也存在。*

证明. 令 V 是 $K_m^{[\lambda,\lambda,\ldots,\lambda]}$ 的顶点集合，$X=\{\infty\}\cup(V\times[k-1])$，$B_x=\{\infty\}\cup(\{x\}\times\{1,2,\ldots,k-1\})$，$\mathcal{B}=\{B_x:x\in V\}$。用 V 中的元素来标记 λ-平行类，记作 $\mathcal{P}_x$，$x\in V$。对每个 $F\in\mathcal{F}$，$k+1$ 个顶点 $V(F)\subset V$ 存在唯一的划分

$$V(F)=\{w\}\cup S_1\cup S_2\cup\cdots\cup S_{k-1}.$$

令

$$A_F=\cup_{i=1}^{k-1}S_i\times\{i\};$$

取 λ 份拷贝在对应的平行类 $\mathcal{P}_w$ 中。令 $\mathcal{A}=\{A_F:F\in\mathcal{F}\}$，$\mathcal{B}^\lambda$ 表示 $\mathcal{B}$ 中的每个元素都重复 λ 次的多重集。容易验证 $(X,\mathcal{A}\cup\mathcal{B}^\lambda)$ 是 $((k-1)m+1,k,\lambda)$-平衡不完全区组设计，每个 $\mathcal{P}_w$ 恰好是一个 λ-平行类。比如说，在 $\mathcal{P}_w$中包含 $(y,i),y\neq w$ 的 λ 个区组是那些 A_F，其中 F 是包含由 w 指向 y 的着 (i) 色边的图。□

通过与文献[48] 定理 10.1 相似的讨论，我们可以证明 $m(m-1)(\lambda,\lambda,\ldots,\lambda)$ 是 $\mu(G(\mathbf{t}))$ 的整线性组合，$(m-1)(\lambda,\lambda,\ldots,\lambda)$ 是 $\tau(x,G(\mathbf{t}))$ 的整线性组合。因此定理 3.5 的条件都满足，我们得到如下结论：

定理 3.24. *给定正整数 k 和 λ，满足 $k \mid \lambda$，存在一个常数 $m_0 = m_0(k,\lambda)$ 使得对任意 $m \geq m_0$，λ-可分解平衡不完全区组设计 $(m(k-1)+1,k,\lambda)$ 存在。*

结合命题 3.22 和定理 3.24，我们最终确定了极小距离为 $2mw-2w$ 的多重常重码的最大容量：

定理 3.25. *给定正整数 k 和 w，满足 $k \mid w$，存在一个常数 $m_0 = m_0(k,w)$ 使得对任意 $m \geq m_0$，*

$$M(m,n,2(mw-w),w) = m(k-1)+1$$

其中 $n = w(m(k-1)+1)/k$。

3.5 重量为 4、极小距离为 6 的多重常重码

在文献[16] 中，作者留下了是否可以决定参数为 $m=2$，$w_1=w_2=2$，$d=6$，$n_1 \leq n_2 \leq 2n_1-1$ 且 n_1 和 n_2 都是奇数的最优多重常重码的公开问题。在本节中，我们将去考虑解决这个问题。

引理 3.26. *设 n_1,n_2 是两个奇数，$0<n_1 \leq n_2 \leq 2n_1-1$。那么 $T(2,n_1;2,n_2;6) \leq \lfloor \frac{n_2(n_1-1)}{4} \rfloor$。*

我们将证明这个上界在绝大部分的情形下都可以达到。首先我们先去定义一个新的组合构型，以及建立它与最优多重常重码的联系。

设 V 是一个 v 元集，S 是一个 s 元集。我们将斜几乎可解方简记为 SAS(s,v)，它是一个 $s \times s$ 的阵列，它的行和列都由 S 中的元素所标记，每个单元可以是空集或者包含 V 中的点对，使得：

1. 对任意的 (i,j) 和 (j,i)，$i \neq j$，至多有一个被填充；
2. 对角线的元素都是空集；
3. 任意 V 中的点对至多出现一次；

4. 对任意 $i \in S$，第 i 行和第 i 列出现的点对是 $V\backslash\{x\}$ 的一个划分，对某个 $x \in V$。

命题 3.27. 设 $v \equiv 1 \pmod 4$，$s \equiv 1 \pmod 2$，且 $v \le s \le 2v-1$。存在码字个数为 $\frac{s(v-1)}{4}$ 的 *MCWC*$(2, v; 2, s; 6)$ 当且仅当 *SAS*(s, v) 存在。

证明. 设 A 是由 SAS(s, v) 所对应的阵列。不失一般性，我们假定 V 和 S 是不同的。令 $X = V \cup S$。码按照如下方式来生成。对 A 中每个被填充的单元 (i, j)，$A(i, j) = \{a, b\}$，构造码字 $\mathbf{u}$，其中 $\mathbf{u}_x = 1$，如果 $x \in \{a, b, i, j\}$；否则 $\mathbf{u}_x = 0$。那么我们得到了一个 MCWC$(2, v; 2, s; d)$。而 1)3)4) 可以保证任一点对至多出现在一个码字的支撑集中，因此任意两个不同码字 $\mathbf{u}$ 和 $\mathbf{v}$ 的支撑集至多交于一个位置，那么码的极小距离为 6。根据 4)，对每个 $i \in S$，存在 $\frac{v-1}{2}$ 个单元在第 i 行和 i 列被填充。因此我们总共有 $\frac{s(v-1)}{4}$ 个单元被填充，码字的个数为 $\frac{s(v-1)}{4}$。

相反的，令 $X = X_1 \cup X_2$，$|X_1| = v$，$|X_2| = s$。令 $\mathcal{C}$ 是一个大小为 $\frac{s(v-1)}{4}$ 的 MCWC$(2, v; 2, s; 6)$。下面我们来构造一个 $s \times s$ 的阵列。对每个码字 $\mathbf{u} \in \mathcal{C}$，$\mathrm{supp}(\mathbf{u}) = \{a, b, i, j\}$，$a, b \in X_1$，$i, j \in X_2$，在单元 (i, j) 中填充点对 $\{a, b\}$。容易验证这的确是一个 SAS(s, v)。 □

在上面 SAS 的定义中，如果我们将条件 4) 替换成如下条件，我们可以得到 SAS$^*(s, v)$ 的定义。

4)' 存在 $i_0 \in S$ 使得对每个 $i \in S\backslash\{i_0\}$，第 i 行和第 i 列的点对是 $V\backslash\{x\}$ 的一个划分，对某个 $x \in V$。而第 i_0 行和第 i_0 列的点对是 $V\backslash\{x, y, z\}$ 的一个划分，对某些不同的 $x, y, z \in V$。

相似的，我们有如下的关于 SAS$^*(s, v)$ 和最优多重常重码的联系，它的证明与命题 3.27 是一致的。

命题 3.28. 设 $v \equiv 3 \pmod 4$ 和 $s \equiv 1 \pmod 2$，$v \le s \le 2v-1$。存在大小为 $\lfloor \frac{s(v-1)}{4} \rfloor$ 的 *MCWC*$(2, v; 2, s; 6)$ 当且仅当 *SAS*$^*(s, v)$ 存在。

下面我们将介绍一个有用的技巧来帮助我们去构造关于 SAS 和 SAS* 的无穷类。

设 V 是一个 v 元集，S 是一个 s 元集。设 $\{H_1, H_2, \ldots, H_n\}$ 是 V 的一个划分，$|H_i| = h_i$。设 $\{S_1, S_2, \ldots, S_n\}$ 是 S 的一个划分，$|S_i| = s_i$。一个型为 $\{(s_i, h_i) : 1 \le i \le n\}$ 的斜

frame 可解方（SFS）是一个 $s \times s$ 的阵列，其中行和列由 S 中的元素来标记，每个单元可以是空集或者是 V 中的点对，使得：

1. 对任意的 (i,j) 和 (j,i)，$i \neq j$，至多有一个被填充；

2. 由 $S_i \times S_i$ 所标记的子矩阵是空的，称之为洞；

3. 任意 V 中的点对至多出现一次；

4. H_i 中的点对都不出现；

5. 对任意 $l \in S_i$，第 l 行和第 l 列中的点对是 $V\backslash H_i$ 的一个划分。

我们将利用指数记号 $(s_1,g_1)^{n_1}\cdots(s_n,g_n)^{n_t}$ 来表示在划分中 (s_i,g_i) 出现 n_i 次。

我们将利用可分组设计（GDD）给出 SFS 的递归构造。

构造8. 设 $(X,\mathcal{G},\mathcal{B})$ 是一个可分组设计，$s,v : X \to \mathbb{Z}^+ \cup \{0\}$ 是 X 上的两个赋权函数。如果对每一个区组 $B \in \mathcal{B}$，都有型 $\{(s(x),v(x)) : x \in B\}$ 的 *SFS* 存在。那么型 $\{(\sum_{x\in G} s(x), \sum_{x\in G} v(x)) : G \in \mathcal{G}\}$ 的 *SFS* 也存在。

证明. 对任意 $x \in X$，设 $S(x)$ 是 $s(x)$ 个元素的指标集，对任意 $x \neq y \in X$，$S(x)$ 和 $S(y)$ 互不相交。对每个 $B \in \mathcal{B}$，我们在 $\cup_{x\in B}(\{x\}\times\{1,2,\ldots,v(x)\})$ 上构造型 $\{(s(x),v(x)) : x \in B\}$ 的 SFS $\mathcal{A}_B$，用 $\cup_{x\in\mathcal{B}}S(x)$ 来标记它的行和列。

记 $S = \cup_{x\in X}S(x)$，$V = \cup_{x\in X}(\{x\} \times \{1,2,\ldots,v(x)\})$。我们按如下方式在 V 上构造所需的 SFS $\mathcal{A}$，用 S 来标记它的行和列：对 $\mathcal{A}$ 中的每个单元用 (α,β) 来标记，如果 $\alpha \in S(x)$，$\beta \in S(y)$，$x \neq y$ 且存在一个区组 $B \in \mathcal{B}$ 包含 x,y，那么我们将单元中的元素替换成由 (α,β) 标记的 $\mathcal{A}_B$；否则单元为空集。

对任意 $G_i \in \mathcal{G}$，记 $S_i = \cup_{x\in G_i}S(x)$ 和 $H_i = \cup_{x\in G_i}(\{x\} \times \{1,2,\ldots,v(x)\})$。容易验证 SFS 定义中的 1) – 4) 均满足。现在，对每个 $\alpha \in S_i$，我们考虑行 α 和列 α 中的点对。不妨设 $\alpha \in S(x)$ 对某个 $x \in G_i$。由于对每个 $y \notin G_i$，存在唯一的区组同时包含 x 和 y，$\{B\backslash\{x\} : x \in B \in \mathcal{B}\}$ 是 $X\backslash G_i$ 的一个划分。注意到对每个 $\mathcal{A}_B$，$x \in B$，行 α 和列 α 中的点对是 $\cup_{y\in B,y\neq x}(y \times \{1,2,\ldots,v(y)\})$ 的一个划分。那么行 α 和列 α 中的点对在 $\mathcal{A}$ 上是

$$\bigcup_{x\in B,B\in\mathcal{B}}\left(\bigcup_{y\in B,y\neq x}(y\times\{1,2,\ldots,v(y)\})\right) = \bigcup_{y\in X\backslash G_i}(y\times\{1,2,\ldots,v(y)\}) = V\backslash H_i$$

的一个划分。

因此我们证明了 $\mathcal{A}$ 是型 $\{(\sum_{x\in G} s(x), \sum_{x\in G} v(x)) : G \in \mathcal{G}\}$ 的 SFS。 □

设 V 是一个 v 元集合，S 是一个 s 元集合。设 W 是 V 的一个子集，$|W| = w$，T 是 S 的一个子集，$|T| = t$。一个带洞的斜几乎可解方，记为 HSAS$(s, v; t, w)$，是一个 $s \times s$ 的阵列，其中行和列都由 S 来标记，每个单元可以是空集或者是 V 中的点对，使得：

1. 对任意的 (i, j) 和 (j, i)，$i \neq j$，至多有一个被填充；
2. 由 $T \times T$ 所标记的子矩阵是空的，称之为洞；
3. 任意 V 中的点对至多出现一次；
4. W 中的点对都不出现；
5. 对每个 $t \in T$，行 t 和列 t 中的点对是 $V\backslash W$ 的一个划分；
6. 对每个 $l \in S\backslash T$，行 l 和列 l 中的点对是 $V\backslash\{x\}$ 的一个划分，对某个 $x \in V$。

下面的事实简单却很有效。

命题 3.29. *如果 HSAS$(s, v; t, w)$ 和 SAS(t, w) 都存在，那么 SAS(s, v) 存在。*

下面，我们将通过 SFS 来构造 SAS。

构造9. *设型 $\{(s_i, h_i) : 1 \leq i \leq n\}$ 的 SFS 存在。设 $s = \sum_{i=1}^{n} s_i$，$v = \sum_{i=1}^{n} h_i$。如果对每个 $1 \leq i \leq n-1$，都存在 HSAS$(s_i + e, h_i + w; e, w)$，并且*

(1) 如果存在 HSAS$(s_n + e, h_n + w; e, w)$，那么 HSAS$(s + e, v + w; e, w)$ 存在；

(2) 如果存在 SAS$(s_n + e, h_n + w)$，那么 SAS$(s + e, v + w)$ 存在；

(3) 如果存在 SAS$^*(s_n + e, h_n + w)$，那么 SAS$^*(s + e, v + w)$ 存在。

证明. 设 A 是在 $V = \cup_{i=1}^{s} H_i$ 上型 $\{(s_i, h_i) : 1 \leq i \leq n\}$ 的 SFS。设 W 是大小为 w 的集合，与 V 互不相交，我们将新的点集取做 $V \cup W$。现在，添加新的 e 行和 e 列。对任意 $1 \leq i \leq n-1$，用 HSAS$(s_i + e, h_i + w; e, w)$ 将 $s_i \times s_i$ 的子方和新增的 e 行 e 列填满，使得新增行和列的交形成一个洞。然后用 HSAS$(s_n + e, h_n + w; e, w)$ 将 $s_n \times s_n$ 的子方和新增的 e行 e 列填满。容易去验证得到的是一个 HSAS$(s + e, v + w; e, w)$。剩下几种情形的证明也是类似的。 □

如果 $\mathbf{u} \in \mathbb{Z}_2^X$ 是 MCWC$(w_1, n_1; w_2, n_2; d)$ 的一个码字。我们可以用一个四元组 $\langle a_1, a_2, a_3, a_4 \rangle \in X^4$ 来表示 $\mathbf{u}$，其中 $\mathbf{u}_{a_1} = \mathbf{u}_{a_2} = \mathbf{u}_{a_3} = \mathbf{u}_{a_4} = 1$。在本节中，我们将一直应用这种记号来表示一个码字。

引理 3.30. 设 $n_1 \in \{3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 25, 29, 33, 37\}$，$n_1 \le n_2 \le 2n_1 - 1$，$n_2$ 是奇数，那么

1. $T(2, n_1; 2, n_2; 6) = \lfloor \frac{n_2(n_1-1)}{4} \rfloor$，除了 $(n_1, n_2) = (5, 7)$；
2. $T(2, 5; 2, 7; 6) = 6$。

证明. T$(2, 5; 2, 7; 6) \le 6$ 的证明可以在[2] 中找到。码的构造如下所示。

对 $3 \le n_1 \le 9$，设 $X = \{0, 1, 2, \ldots, n_1 + n_2 - 1\}$。$X$ 可以被划分为 $X = X_1 \cup X_2$，$X_1 = \{0, 1, \ldots, n_1 - 1\}$ 和 $X_2 = \{n_1, n_1 + 1, \ldots, n_1 + n_2 - 1\}$。在表 3-1 中列出了所需构造的码字。

对 $n_1 \in \{13, 17, 21, 25, 29, 33, 37\}$，码的具体构造可以参考原文的附录。

对 $n_1 \in \{11, 15, 19\}$，$n_1 \le n_2 \le 2n_1 - 3$，从原文的附录中取出 HSAS$(n_2, n_1; 3, 3)$，用 SAS$^*(3, 3)$（等价于 MCWC$(2, 3; 2, 3; 6)$）将它的洞填上，可以得到 SAS$^*(n_2, n_1)$。由命题 3.28，这等价于 MCWC$(2, n_1; 2, n_2; 6)$ 的码字数目为 $\lfloor \frac{n_2(n_1-1)}{4} \rfloor$。对 $n_1 \in \{11, 15, 19\}$ 和 $n_2 = 2n_1 - 1$，我们进行类似的操作；从原文的附录中取出 HSAS$(n_2, n_1; 5, 3)$ 用 SAS$^*(5, 3)$（等价于 MCWC$(2, 5; 2, 3; 6)$）将它的洞填上。 □

引理 3.31. 设 t 是一个正整数，$2t + 1 \ge 21$，$2t + 1 \notin \{23, 27, 29, 33, 39, 43, 51, 59, 75, 83, 87, 95, 99, 107, 139, 179\}$。设 $n_1 = 4t + 1$ 或者 $4t + 3$，$n_1 \le n_2 \le 2n_1 - 1$，n_2 是奇数。那么 $T(2, n_1; 2, n_2; 6) = \lfloor \frac{n_2(n_1-1)}{4} \rfloor$。

证明. 由命题 3.27 和 3.28，我们只需去构造相应的 SAS(n_2, n_1)，$n_1 \equiv 1 \pmod 4$；或者 SAS$^*(n_2, n_1)$，$n_1 \equiv 3 \pmod 4$。

对每个给定的 t 且 $2t+1 \notin \{71, 111, 113, 115, 119\}$，取出 $(2t+1, \{5, 7, 9\}, 1)$-PBD，删去一个点得到型 $4^i 6^j 8^k$ 的 $\{5, 7, 9\}$-GDD，$4i + 6j + 8k = 2t$。将重量 $(4, 2)$ 或者 $(2, 2)$ 分配到每个点上，应用构造 8；以原文的附录中的型 $(4, 2)^a (2, 2)^b$ 的 SFS 作为输入，$a + b \in \{5, 7, 9\}$。那么我们可以得到的 SFS 的型是

$$(8, 8)^{i_8} (10, 8)^{i_{10}} \cdots (16, 8)^{i_{16}} (12, 12)^{j_{12}} \cdots (24, 12)^{j_{24}} (16, 16)^{k_{16}} \cdots (32, 16)^{k_{32}},$$

表 3-1　小参数下的 MCWC$(2, n_1; 2, n_2; 6)$，$3 \le n_1 \le 9$

(n_1, n_2)	码字
$(3,3)$	$\langle 0,1,3,4\rangle$
$(3,5)$	$\langle 0,1,3,4\rangle$ $\langle 1,2,5,6\rangle$
$(5,5)$	$\langle 0,1,5,6\rangle$ $\langle 0,2,7,8\rangle$ $\langle 1,3,7,9\rangle$ $\langle 2,4,5,9\rangle$ $\langle 3,4,6,8\rangle$
$(5,7)$	$\langle 0,1,5,6\rangle$ $\langle 0,2,7,8\rangle$ $\langle 0,3,9,10\rangle$ $\langle 1,2,9,11\rangle$ $\langle 1,4,7,10\rangle$ $\langle 3,4,5,8\rangle$
$(5,9)$	$\langle 0,3,10,9\rangle$ $\langle 2,3,5,13\rangle$ $\langle 0,2,8,7\rangle$ $\langle 0,4,11,12\rangle$ $\langle 1,2,9,11\rangle$ $\langle 1,3,7,12\rangle$ $\langle 0,1,6,5\rangle$ $\langle 1,4,8,13\rangle$
$(7,7)$	$\langle 0,1,7,8\rangle$ $\langle 0,2,9,10\rangle$ $\langle 0,3,11,12\rangle$ $\langle 1,2,11,13\rangle$ $\langle 1,4,9,12\rangle$ $\langle 2,5,7,12\rangle$ $\langle 3,4,7,10\rangle$ $\langle 3,5,8,9\rangle$ $\langle 4,6,8,11\rangle$ $\langle 5,6,10,13\rangle$
$(7,9)$	$\langle 0,1,7,8\rangle$ $\langle 0,2,9,10\rangle$ $\langle 0,3,11,12\rangle$ $\langle 0,4,13,14\rangle$ $\langle 1,2,11,13\rangle$ $\langle 1,3,9,14\rangle$ $\langle 1,4,10,12\rangle$ $\langle 2,3,7,15\rangle$ $\langle 2,5,8,12\rangle$ $\langle 3,5,10,13\rangle$ $\langle 4,5,7,9\rangle$ $\langle 4,6,8,11\rangle$ $\langle 5,6,14,15\rangle$
$(7,11)$	$\langle 0,1,7,8\rangle$ $\langle 0,2,9,10\rangle$ $\langle 1,5,11,13\rangle$ $\langle 0,4,13,14\rangle$ $\langle 0,5,15,16\rangle$ $\langle 3,6,16,13\rangle$ $\langle 1,3,9,14\rangle$ $\langle 0,3,11,17\rangle$ $\langle 1,6,15,17\rangle$ $\langle 2,3,7,15\rangle$ $\langle 2,4,8,16\rangle$ $\langle 2,5,12,14\rangle$ $\langle 3,5,8,10\rangle$ $\langle 1,4,12,10\rangle$ $\langle 4,5,7,17\rangle$ $\langle 4,6,9,11\rangle$
$(7,13)$	$\langle 0,1,7,8\rangle$ $\langle 0,2,9,10\rangle$ $\langle 0,3,11,12\rangle$ $\langle 1,4,12,10\rangle$ $\langle 5,4,7,9\rangle$ $\langle 0,6,17,18\rangle$ $\langle 1,2,11,13\rangle$ $\langle 1,3,9,14\rangle$ $\langle 3,5,10,8\rangle$ $\langle 1,5,17,19\rangle$ $\langle 3,4,19,18\rangle$ $\langle 2,4,8,16\rangle$ $\langle 5,0,16,15\rangle$ $\langle 0,4,14,13\rangle$ $\langle 2,3,7,17\rangle$ $\langle 3,6,13,16\rangle$ $\langle 4,6,11,15\rangle$ $\langle 2,5,12,18\rangle$ $\langle 2,6,14,19\rangle$
$(9,9)$	$\langle 7,3,15,12\rangle$ $\langle 2,1,16,11\rangle$ $\langle 4,8,9,15\rangle$ $\langle 0,3,11,10\rangle$ $\langle 2,8,10,13\rangle$ $\langle 2,6,9,12\rangle$ $\langle 4,5,11,12\rangle$ $\langle 1,0,12,17\rangle$ $\langle 6,7,13,17\rangle$ $\langle 6,0,14,15\rangle$ $\langle 6,4,10,16\rangle$ $\langle 1,4,14,13\rangle$ $\langle 7,1,10,9\rangle$ $\langle 7,8,14,11\rangle$ $\langle 3,8,17,16\rangle$ $\langle 5,2,15,17\rangle$ $\langle 3,5,14,9\rangle$ $\langle 0,5,13,16\rangle$
$(9,11)$	$\langle 4,8,13,11\rangle$ $\langle 3,0,14,10\rangle$ $\langle 6,5,11,19\rangle$ $\langle 3,1,16,11\rangle$ $\langle 0,8,16,15\rangle$ $\langle 8,2,9,17\rangle$ $\langle 6,2,14,13\rangle$ $\langle 6,8,12,10\rangle$ $\langle 4,3,17,15\rangle$ $\langle 6,3,9,18\rangle$ $\langle 4,1,12,18\rangle$ $\langle 3,5,13,12\rangle$ $\langle 8,1,19,14\rangle$ $\langle 7,0,11,12\rangle$ $\langle 1,7,9,13\rangle$ $\langle 0,5,17,18\rangle$ $\langle 6,7,17,16\rangle$ $\langle 2,7,19,18\rangle$ $\langle 7,5,14,15\rangle$ $\langle 0,4,9,19\rangle$ $\langle 1,2,10,15\rangle$ $\langle 4,5,10,16\rangle$
$(9,13)$	$\langle 6,4,13,17\rangle$ $\langle 3,2,17,9\rangle$ $\langle 0,1,9,10\rangle$ $\langle 6,3,18,15\rangle$ $\langle 5,0,16,17\rangle$ $\langle 2,5,18,13\rangle$ $\langle 7,1,18,20\rangle$ $\langle 2,1,12,15\rangle$ $\langle 2,4,16,14\rangle$ $\langle 1,5,19,21\rangle$ $\langle 6,7,9,12\rangle$ $\langle 8,7,13,16\rangle$ $\langle 8,2,20,19\rangle$ $\langle 0,3,19,13\rangle$ $\langle 4,5,20,9\rangle$ $\langle 8,0,18,12\rangle$ $\langle 7,3,14,21\rangle$ $\langle 7,5,15,10\rangle$ $\langle 7,4,19,11\rangle$ $\langle 1,3,16,11\rangle$ $\langle 6,8,10,11\rangle$ $\langle 6,0,20,14\rangle$ $\langle 3,4,12,10\rangle$ $\langle 1,8,14,17\rangle$ $\langle 0,2,11,21\rangle$ $\langle 4,8,15,21\rangle$

对任意非负整数 $i_8, i_{10}, \ldots, i_{16}, j_{12}, \ldots, k_{32}$，

$$i_8 + i_{10} + \cdots + i_{16} = i$$
$$j_{12} + j_{14} + \cdots + j_{24} = j$$
$$k_{16} + k_{18} + \cdots + k_{32} = k.$$

现在我们可以通过下面的三种方式来填 SFS 的洞：

1. 新增一行一列，利用构造 9 (2)，‘$e = 1$’ 和 ‘$w = 1$’；输入的 HSAS$(r, v; 1, 1)$（i.e. SAS(r, v)）可以由引理 3.30 得到，$v \in \{9, 13, 17\}$，$v \le r \le 2v - 1$。最终我们得到了一个 SAS$(s, 4t + 1; 1, 1)$，其中 $4t + 1 \le s \le 8t + 1$。

2. 新增三行三列，利用构造 9 (1)，‘$e = 3$’ 和 ‘$w = 3$’；输入的 HSAS$(r, v; 3, 3)$ 的构造在原文的附录中给出，$v \in \{11, 15, 19\}$，$v \le r \le 2v - 3$。我们得到 HSAS$(s, 4t + 3; 3, 3)$，$4t + 3 \le s \le 8t + 3$。然后用引理 3.30 中得到的 SAS$(3, 3)$ 来填洞，最终得到 SAS$^*(s, 4t + 3)$，$4t + 3 \le s \le 8t + 3$。

3. 当 SFS 的型是 $(16, 8)^i(24, 12)^j(32, 16)^k$ 时，新增五行和五列，应用构造 9 (1)，‘$e = 5$’ 和 ‘$w = 3$’；输入的 HSAS$(r, v; 5, 3)$ 的构造在原文的附录中给出，$(r, v) \in \{(21, 11), (29, 15), (37, 19)\}$。我们得到 HSAS$(8t + 5, 4t + 3; 5, 3)$。然后利用引理 3.30 中构造的 SAS$^*(5, 3)$ 来填洞，最终得到 SAS$^*(8t + 5, 4t + 3)$。

对 $2t + 1 = 71$，取 TD$(9, 8)$，从某个组中截短六个点，可以得到一个型 $8^8 6^1$ 的 $\{8, 9\}$-GDD。与上面的方法类似，我们可以得到 SAS$(s, 4t + 1)$ 和 SAS$^*(s, 4t + 3)$。这里输入的型 $(4, 2)^a(2, 2)^{8-a}$ 的 SFS 也在原文的附录中给出，$0 \le a \le 8$。

对 $2t + 1 \in \{111, 113, 115, 119\}$，取型 8^{15} 的 $\{7, 9\}$-GDD[21]，截短最后两个组得到型 $8^{13}6^1$，8^{14}，$8^{13}6^14^1$，$8^{14}6^1$ 的 $\{5, 6, 7, 8, 9\}$-GDD。同理我们可以得到想要的 SAS 和 SAS*。输入的型 $(4, 2)^a(2, 2)^{6-a}$ 的 SFS 在原文的附录中给出。 □

引理 3.32. *设 t 是一个正整数，$2t + 1 \in \{39, 43, 51, 59, 75, 99\}$。设 $n_1 = 4t + 1$ 或者 $4t + 3$，$n_1 \le n_2 \le 2n_1 - 1$，n_2 是奇数。那么 $T(2, n_1; 2, n_2; 6) = \lfloor \frac{n_2(n_1-1)}{4} \rfloor$。*

证明. 对 $2t + 1 = 39$，取一个型 $6^6 2^1$ 的 $\{5, 7\}$-GDD[21]，注意到 $6 \times 6 + 2 = 2t$。那么我们可以得到型 $(12, 12)^{i_{12}}(14, 12)^{i_{14}} \cdots (24, 12)^{i_{24}}(4, 4)^{j_4}(8, 4)^{j_8}$ 的 SFS，对任意非负整数 $i_{12}, i_{14}, \ldots, i_{24}, j_4, j_8$，$i_{12} + i_{14} + \cdots + i_{24} = 6$ 且 $j_4 + j_8 = 1$。现在我们可以通过下面的三种方式来填 SFS 的洞：

1. 新添一行一列，利用构造 9 (2)，‘$e = 1$’ 和 ‘$w = 1$’；输入的 HSAS$(r, 13; 1, 1)$ （i.e. SAS$(r, 13)$），$13 \le r \le 25$，SAS$(5, 5)$ 和 SAS$(9, 5)$可以从引理 3.30 中得到。最终我们得到了一个 SAS$(s, 77)$，其中 $77 \le s \le 153$。

2. 新增三行三列，利用构造 9 (3)，‘$e = 3$’ 和 ‘$w = 3$’；输入的 HSAS$(r, 15; 3, 3)$ 的构造在原文的附录中给出，$15 \le r \le 27$ ，SAS$^*(7, 7)$ 和 SAS$^*(11, 7)$ 从引理 3.30 中得到。最终得到 SAS$^*(s, 79)$，$79 \le s \le 155$。

3. 当 SFS 的型为 $(24, 12)^6(8, 4)^1$ 时，新增五行和五列，应用构造 9 (3)，‘$e = 5$’ 和 ‘$w = 3$’；输入的 HSAS$(29, 15; 5, 3)$ 的构造在原文的附录中给出，SAS$^*(13, 7)$ 由引理 3.30 给出。我们得到 SAS$^*(157, 79)$。

对 $2t + 1 \in \{43, 51, 59, 75, 99\}$，我们分别从型 8^52^1、8^62^1、8^72^1、8^92^1、$8^{12}2^1$ 的 $\{5, 6, 7, 8, 9\}$-GDD 出发，来构造我么所需的 SAS 和 SAS*；这里我们应用 SAS$(r, 17)$（引理 3.30）、HSAS$(r, 19; 3, 3)$（附录）、HSAS$(37, 19; 5, 3)$（附录）来填 SFS 的洞。 □

引理 3.33. *设 t 是一个正整数，$2t + 1 \in \{107, 139, 179\}$。设 $n_1 = 4t + 1$ 或者 $4t + 3$，$n_1 \le n_2 \le 2n_1 - 1$，n_2 是奇数。那么 $T(2, n_1; 2, n_2; 6) = \lfloor \frac{n_2(n_1-1)}{4} \rfloor$。*

证明. 对 $2t + 1 = 107$，取 TD$(6, 20)$，从它的最后一个组中截短六个点，可以得到一个型 20^56^1 的 $\{5, 6\}$-GDD。那么我们可以得到型 $(40, 40)^{i_{40}}(42, 40)^{i_{42}} \cdots (80, 40)^{i_{80}}(12, 12)^{j_{12}}(14, 12)^{j_{14}} \cdots (24, 12)^{j_{24}}$ 的 SFS，对任意的非负整数 $i_{40}, i_{42}, \ldots, i_{80}, j_{12}, \ldots, j_{24}$，$i_{40} + i_{42} + \cdots + i_{80} = 5$，$j_{12} + j_{14} + \ldots + j_{24} = 1$。现在我们可以通过下面的三种方式来填 SFS 的洞：

1. 新添一行一列，利用构造 9 (2)，‘$e = 1$’ 和 ‘$w = 1$’；输入的 HSAS 和 SAS 可以从引理 3.30–3.31 中得到。最终我们得到了一个 SAS$(s, 213)$，其中 $213 \le s \le 425$。

2. 新增三行三列，利用构造 9 (3)，‘$e = 3$’ 和 ‘$w = 3$’；输入的 HSAS$(r, 43; 3, 3)$，$43 \le r \le 83$，和 SAS$^*(r, 15)$可以从引理 3.30–3.31 中得到。最终得到 SAS$^*(s, 215)$，$215 \le s \le 427$。

3. 当 SFS 的型是 $(80, 40)^5(24, 12)^1$ 时，新增五行和五列，应用构造 9 (3)，‘$e = 5$’ 和 ‘$w = 3$’；输入的 HSAS$(85, 43; 5, 3)$ 和 SAS$^*(29, 15)$ 由引理 3.30–3.31 给出。我们得到 SAS$^*(429, 215)$。

对 $2t + 1 = 139$ 或者 179，取 TD$(8, 24)$，截短它的最后三个组，可以得到型 24^56^3、$24^76^14^1$ 的 $\{5, 6, 7, 8, 9\}$-GDD。然后我们可以用类似的方法得到想要的 SAS 和 SAS*。 □

同理，我们可以得到如下的结论：

引理 3.34. 设 t 是一个正整数，$2t+1 \in \{83, 87, 95\}$。设 $n_1 = 4t+1$，$n_1 \leq n_2 \leq 2n_1 - 1$，$n_2$ 是奇数。那么 $T(2, n_1; 2, n_2; 6) = \frac{n_2(n_1-1)}{4}$.

综上所述，我们可以得到以下的结论

定理 3.35. 设 n_1, n_2 是两个奇数，$0 < n_1 \leq n_2 \leq 2n_1 - 1$。除了 $(n_1, n_2) = (5, 7)$ 的情况和 $n_1 \in \{23, 27, 31, 35, 39, 45, 47, 53, 55, 57, 59, 65, 67, 165, 175, 191\}$ 这些可能情况之外，都有 $T(2, n_1; 2, n_2; 6) = \lfloor \frac{n_2(n_1-1)}{4} \rfloor$。

3.6 小结

本章主要介绍了对多重常重码的研究。一方面在一些特定的参数下，我们利用球面码的上界导出的新的关于多重常重码的上界，这可以用来改进第三型 Johnson 界，同时我们得到了多重常重码的一般 GV 界，这可以改进由级联得到的下界。另一方面，我们几乎完全确定了极小距离为 $2\sum_{i=1}^{m} w_i - 2$ 和 $2mw - 2w$ 这两类码的最大码字数目。最后，我们彻底解决了重量为 4、极小距离为 6 的最优多重常重码。

事实上，常重码的研究要比一般的二元码更加困难，常重码的渐近上界和对 GV 界的一般改进目前还没有系统的结果。而多重常重码作为一种特殊的常重码，它的研究不仅对物理不可克隆函数的设计具有重大意义，并且对常重码的理论研究具有很大的参考意义。在未来的研究中，我们希望可以得到更多的最优多重常重码的例子，从而为常重码的研究提供理论依据。

4 L-相交系

4.1 介绍

我们称 $[n] = \{1, 2, \ldots, n\}$ 中的集族 $\mathcal{A}$ 是相交的，如果集族 $\mathcal{A}$ 中的任何一对集合 $A_i, A_j \in \mathcal{A}$ 都有非空的交。令 L 是由 s 个非负整数所构成的集合。如果集族 $\mathcal{A}$ 中的任何一对集合 A_i, A_j 都满足 $|A_i \cap A_j| \in L$，我们称之为 L-相交系。如果集族中的每个集合大小都相同，则我们称这个集族是均匀的。极值集合论的一个主要研究对象就是确定相交集族和 L-相交系的大小。其中最著名的两个定理是 EKR 定理和 RW 定理，它们分别刻画了相交集族和 L-相交系的最大可能值，并确定了其极值情形。

下面是 de Bruijin 和 Erdös 在 1948 年得到的关于相交系的结果。

定理 4.1. *[24]* 如果 $[n]$ 中的集族 $\mathcal{A}$ 满足 $|A_i \cap A_i| = 1$ 对任意的 $A_i, A_j \in \mathcal{A}$，那么 $|\mathcal{A}| \leq n$。

一年之后，在文献[8] 中 Bose 对上面的结果进行了一般性的推广。

定理 4.2. *[8]* 如果 $[n]$ 中的集族 $\mathcal{A}$ 满足 $|A_i \cap A_i| = \lambda$（$\lambda \neq 0$）对任意的 $A_i, A_j \in \mathcal{A}$，那么 $|\mathcal{A}| \leq n$。

1961 年，Erdös，Ko 和 Rado 正式发表了他们经典的论文[32]，证明了如下关于 k-均匀相交集族的结果。

定理 4.3. 设 $n > 2k$，$\mathcal{A}$ 为 $[n]$ 中的 k-均匀相交集族，那么 $|\mathcal{A}| \leq \binom{n-1}{k-1}$。等号成立当且尽当集族 $\mathcal{A}$ 的所有 k 元子集都包含一个共同的元素。

之后在 1975 年，Ray-Chaudhuri 和 Wilson[64] 在均匀 L-相交系方面得到了如下重要的进展：

定理 4.4. 设 $\mathcal{A}$ 是 $[n]$ 中的 k-均匀 L-相交系，那么 $|\mathcal{A}| \leq \binom{n}{s}$。

在一些特定的参数下，这个不等式是最优的（当我们考虑 $L=\{0,1,\ldots,s-1\}$，集族 $\mathcal{A}$ 由所有的 s 元子集所构成）。

1981 年，Frankl 和 Wilson[36] 对上面的结果进行了一般化地推广：

定理 4.5. *设 $\mathcal{A}$ 是 $[n]$ 中的 L-相交系，那么 $\mathcal{A}\leq\binom{n}{s}+\binom{n}{s-1}+\cdots+\binom{n}{0}$。*

当我们考虑的集族由所有大小至多为 s 的子集所构成时，上述的界也是不能被改进的。

在同一篇论文中他们还得到了模意义下的 RW 定理。

定理 4.6. *设 p 为任意素数，$\mathcal{A}$ 是 $[n]$ 中的 k-均匀集族满足 $k \pmod p\notin L$ 和 $|A_i\cap A_j| \pmod p\in L$ 对所有 $i\neq j$，那么 $|\mathcal{A}|\leq\binom{n}{s}$。*

在 1991 年，Alon，Babai 和 Suzuki[3] 将定理 4.6 中的均匀条件替换成 $\mathcal{A}$ 中的集合在模 p 下有 r 种不同的大小，得到了如下形式的推广：

定理 4.7. *设 $K=\{k_1,k_2,\ldots,k_r\}$ 和 $L=\{l_1,l_2,\ldots,l_s\}$ 为 $\{0,1,\ldots,p-1\}$ 中两个互不相交的子集，其中 p 为任一素数。若集族 $\mathcal{A}$ 满足 $|A_i| \pmod p\in K$ 对任意 $A_i\in\mathcal{A}$，$|A_i\cap A_j| \pmod p\in L$ 对 $i\neq j$，当 $r(s-r+1)\leq p-1$ 和 $n\geq s+\max_{1\leq i\leq r}k_i$ 时，那么 $|\mathcal{A}|\leq\binom{n}{s}+\binom{n}{s-1}+\cdots+\binom{n}{s-r+1}$。*

在定理 4.7 的证明中，Alon，Babai 和 Suzuki 使用了非常强力的线性代数方法以及一个很有技巧性的引理 3.6。遗憾的是定理中的条件 $r(s-r+1)\leq p-1$ 和 $n\geq s+\max_{1\leq i\leq r}k_i$ 看起来并不是非常自然，因此他们提出了条件 $r(s-r+1)\leq p-1$ 是否可以被去掉的猜想。Snevily 最早对这个问题进行了探索，在 1994 年文献[70] 中得到了当 n 充分大时，ABS 猜想成立。

定理 4.8. *设 p 是一个素数，K,L 是 $\{0,1,\ldots,p-1\}$ 中的两个互不相交的子集。设 $|L|=s$，$\mathcal{A}$ 是 $[n]$ 中的集族，满足 $|A_i| \pmod p\in K$ 对任意 $A_i\in\mathcal{A}$ 且 $|A_i\cap A_j| \pmod p\in L$ 对任意 $i\neq j$，那么 $|\mathcal{A}|\leq\binom{n-1}{s}+\binom{n-1}{s-1}+\cdots+\binom{n-1}{0}$。*

2000 年 Qian 和 Ray-Chaudhuri[59] 从另一个角度利用线性代数方法，将 4.7 中的条件

$r(s-r+1) \le p-1$ 和 $n \ge s + \max_{1\le i\le r} k_i$ 替换成了一个非常简洁的条件 $n \ge 2s-r$。

定理 4.9. 设 p 为一素数，$K=\{k_1,k_2,\ldots,k_r\}$ 和 $L=\{l_1,l_2,\ldots,l_s\}$ 是 $\{0,1,\ldots,p-1\}$ 中两个互不相交的子集。如果 $n \ge 2s-r$，集族 $\mathcal{A}$ 满足 $|A_i| \pmod p \in K$ 对任意 $A_i \in \mathcal{A}$ 和 $|A_i \cap A_j| \pmod p \in L$ 对任意 $i \ne j$，那么 $|\mathcal{A}| \le \binom{n}{s} + \binom{n}{s-1} + \cdots + \binom{n}{s-r+1}$。

直到最近，ABS 猜想才由 Hwang 和 Kim 完全解决[42]。

定理 4.10. 设 p 为一素数，$K=\{k_1,k_2,\ldots,k_r\}$ 和 $L=\{l_1,l_2,\ldots,l_s\}$ 是 $\{0,1,\ldots,p-1\}$ 中两个互不相交的子集。如果集族 $\mathcal{A}$ 满足 $|A_i| \pmod p \in K$ 对任意 $A_i \in \mathcal{A}$ 和 $|A_i \cap A_j| \pmod p \in L$ 对任意 $i \ne j$，当 $n \ge s+\max_{1\le i\le r} k_i$ 时，那么 $|\mathcal{A}| \le \binom{n}{s}+\binom{n}{s-1}+\cdots+\binom{n}{s-r+1}$。

在文献[17] 中，Chen 和 Liu 在条件 $\min\{k_i\} > \max\{l_i\}$ 下改进了定理 4.8 中的上界。

定理 4.11. 设 p 是一个素数，$L=\{l_1,l_2,\ldots,l_s\}$ 和 $K=\{k_1,k_2,\ldots,k_r\}$ 是 $\{0,1,\ldots,p-1\}$ 中互不相交的两个子集，使得 $\min\{k_i\} > \max\{l_i\}$。如果 $\mathcal{A}$ 是 $[n]$ 中的集族，满足 $|A_i| \pmod p \in K$ 对任意 $A_i \in \mathcal{A}$，且 $|A_i \cap A_j| \pmod p \in L$ 对任意 $i \ne j$，那么 $|\mathcal{A}| \le \binom{n-1}{s} + \binom{n-1}{s-1} + \cdots + \binom{n-1}{s-2r+1}$。

在文献[55] 中，Liu 和 Yang 将定理 4.11 中的结论在条件 $k_i > s-r$ 下进行了推广。

定理 4.12. 设 p 是一个素数，$L=\{l_1,l_2,\ldots,l_s\}$ 和 $K=\{k_1,k_2,\ldots,k_r\}$ 是 $\{0,1,\ldots,p-1\}$ 中互不相交的两个子集，使得 $k_i > s-r$ 对任意 i。如果 $\mathcal{A}$ 是 $[n]$ 中的集族，满足 $|A_i| \pmod p \in K$ 对任意 $A_i \in \mathcal{A}$，且 $|A_i \cap A_j| \pmod p \in L$ 对任意 $i \ne j$，那么 $|\mathcal{A}| \le \binom{n-1}{s} + \binom{n-1}{s-1} + \cdots + \binom{n-1}{s-2r+1}$。

在同一篇文章中，作者还在定理 4.7 的条件下，得到了相同的结论。

定理 4.13. 设 p 是一个素数，$L=\{l_1,l_2,\ldots,l_s\}$ 和 $K=\{k_1,k_2,\ldots,k_r\}$ 是 $\{0,1,\ldots,p-1\}$ 中互不相交的两个子集，使得 $r(s-r+1) \le p-1$，$n \ge s+\max_{1\le i\le r} k_i$。如果 $\mathcal{A}$ 是 $[n]$ 中的集族，满足 $|A_i| \pmod p \in K$ 对任意 $A_i \in \mathcal{A}$，且 $|A_i \cap A_j| \pmod p \in L$ 对任意 $i \ne j$，那么 $|\mathcal{A}| \le \binom{n-1}{s} + \binom{n-1}{s-1} + \cdots + \binom{n-1}{s-2r+1}$。

在本章中，我们将对以上的结果进行显著的改进。我们的主要结论如下：

定理 4.14. 设 p 为一素数，$K=\{k_1,k_2,\ldots,k_r\}$ 和 $L=\{l_1,l_2,\ldots,l_s\}$ 是 $\{0,1,\ldots,p-1\}$ 中两个互不相交的子集。如果 $n\geq 2s-2r+1$，集族 $\mathcal{A}$ 满足 $|A_i|\pmod p\in K$ 对任意 $A_i\in\mathcal{A}$ 和 $|A_i\cap A_j|\pmod p\in L$ 对任意 $i\neq j$，那么 $|\mathcal{A}|\leq\binom{n-1}{s}+\binom{n-1}{s-1}+\cdots+\binom{n-1}{s-2r+1}$。

定理 4.15. 设 p 为一素数，$K=\{k_1,k_2,\ldots,k_r\}$ 和 $L=\{l_1,l_2,\ldots,l_s\}$ 是 $\{0,1,\ldots,p-1\}$ 中两个互不相交的子集。如果集族 $\mathcal{A}$ 满足 $|A_i|\pmod p\in K$ 对任意 $A_i\in\mathcal{A}$ 和 $|A_i\cap A_j|\pmod p\in L$ 对任意 $i\neq j$，当 $n\geq s+\max_{1\leq i\leq r}k_i$ 时，那么 $|\mathcal{A}|\leq\binom{n-1}{s}+\binom{n-1}{s-1}+\cdots+\binom{n-1}{s-2r+1}$。

注意到如下事实：当 $n\geq 2s-2$时，对 $1\leq i\leq r-1$ 有 $\binom{n-1}{s}+\binom{n-1}{s-1}+\cdots+\binom{n-1}{s-2r+1}=\binom{n}{s}+\binom{n}{s-2}+\cdots+\binom{n}{s-2(r-1)}$ 和 $\binom{n}{s-2i}<\binom{n}{s-i}$。定理 4.14 改进了 Qian 和 Ray-Chaudhuri 定理 4.9 中的结果；而当 $n\geq 2s-2$ 时，定理 4.15 加强了定理 4.10 中的结果。

定理 4.7，4.9，4.12，4.13 都已经在文献[38,55] 中被推广到了 k 个集合相交的情形。利用相似的想法，我们也可以将定理 4.14 和 4.15 进行如下推广。

定理 4.16. 设 p 是一个素数，$k\geq 2$。设 $L=\{l_1,l_2,\ldots,l_s\}$ 和 $K=\{k_1,k_2,\ldots,k_r\}$ 是 $\{0,1,\ldots,p-1\}$ 中的两个互不相交的子集。如果 $[n]$ 中的集族 $\mathcal{A}$ 满足 $|A_i|\pmod p\in K$ 对任意 $A_i\in\mathcal{A}$，且 $|A_{i_1}\cap A_{i_2}\cap\ldots\cap A_{i_k}|\pmod p\in L$ 对$\mathcal{A}$ 中任意 k 个不同的子集。如果 $n\geq 2s-2r+1$ 或者 $n\geq s+\max_{1\leq i\leq r}k_i$，那么 $|\mathcal{A}|\leq(k-1)\left[\binom{n-1}{s}+\binom{n-1}{s-1}+\cdots+\binom{n-1}{s-2r+1}\right]$。

4.2 定理 4.14 的证明

本节我们进行定理 4.14 的证明。

在本节中我们对记号做如下规定：设 $X=[n-1]=\{1,2,\ldots,n-1\}$ 是由 $(n-1)$ 个元素构成的集合，p 为一素数，$L=\{l_1,l_2,\ldots,l_s\}$ 和 $K=\{k_1,k_2,\ldots,k_r\}$ 是 $\{0,1,\ldots,p-1\}$ 中的两个互不相交的子集。设 $\mathcal{A}=\{A_1,A_2,\ldots,A_m\}$ 是 $[n]$ 中的集族满足：(1) $|A_i|\pmod p\in K$ 对任意 $1\leq i\leq m$，(2) $|A_i\cap A_j|\pmod p\in L$ 对 $i\neq j$。不失一般性，我们设总是存在一个正常数 t，使得 $n\notin A_i$ 对 $1\leq i\leq t$，$n\in A_i$ 对 $i\geq t+1$。记

$$\mathbb{P}_i(X)=\{S|S\subset X\ ,\ |S|=i\}.$$

我们将变量 x_i 与集合 $A_i\in\mathcal{A}$ 相关联并且令 $x=(x_1,x_2,\ldots,x_m)$。对任意子集 $I\subset X$，定

义

$$L_I = \sum_{i:I\subset A_i\in\mathcal{A}} x_i.$$

在 $\mathbb{F}_p$ 上考虑如下线性方程组：

$$\{L_I = 0,\ \text{其中 } I \text{ 遍历 } \cup_{i=0}^{s}\mathbb{P}_i(X)\}. \tag{4-1}$$

命题 4.17. *如果集族 $\mathcal{A}$ 满足(1)(2)，那么上面的线性方程组只有零解。*

证明. 我们不妨设 $v=(v_1,v_2,\ldots,v_m)$ 是方程组 (4-1) 的一个解。我们要证明在 $\mathbb{F}_p$ 上 v 只能是零向量。定义如下两个辅助多项式：

$$g(x)=\prod_{j=1}^{s}(x-l_j)$$

和

$$h(x)=g(x+1)=\prod_{j=1}^{s}(x+1-l_j).$$

由于 $\binom{x}{0},\binom{x}{1},\ldots,\binom{x}{s}$ 在次数至多为 s 的多项式空间 $\mathbb{F}_p[x]$ 中形成一组基，分别存在常数 $a_0,a_1,\ldots,a_s\in\mathbb{F}_p$ 和 $b_0,b_1,\ldots,b_s\in\mathbb{F}_p$ 使得

$$g(x)=\sum_{i=0}^{s}a_i\binom{x}{i}$$

和

$$h(x)=\sum_{i=0}^{s}b_i\binom{x}{i}$$

成立。设 A_{i_0} 是 $v_{i_0}\neq 0$ 所对应的集合。下面我们证明两个恒等式：

如果 $n\notin A_{i_0}$，那么

$$\sum_{i=0}^{s}a_i\sum_{I\in\mathbb{P}_i(X),I\subset A_{i_0}}L_I=\sum_{A_i\in\mathcal{A}}g(|A_i\cap A_{i_0}|)x_i; \tag{4-2}$$

如果 $n\in A_{i_0}$，那么

$$\sum_{i=0}^{s}b_i\sum_{I\in\mathbb{P}_i(X),I\subset A_{i_0}}L_I=\sum_{i=1}^{t}h(|A_i\cap A_{i_0}|)x_i+\sum_{i\geq t+1}h(|A_i\cap A_{i_0}|-1)x_i. \tag{4-3}$$

我们通过比较两边的系数来证明它们。对任意 $A_i\in\mathcal{A}$，在式子 (4-2) 中 x_i 在左边的系数是

$$\sum_{i=0}^{s}a_i|\{I\in\mathbb{P}_i(X):I\subset A_{i_0},I\subset A_i\}|=\sum_{i=0}^{s}a_i\binom{|A_i\cap A_{i_0}|}{i},$$

由 a_i 的定义可以得到它与 $g(|A_i \cap A_{i_0}|)$ 相等。这就证明了等式 (4-2)。

对任意 $i \le t$，在式子 (4-3) 中 x_i 在左边的系数是

$$\sum_{i=0}^{s} b_i |\{I \in \mathbb{P}_i(X) : I \subset A_{i_0}, I \subset A_i\}| = \sum_{i=0}^{s} b_i \binom{|A_i \cap A_{i_0}|}{i},$$

对任意 $i \ge t+1$，在式子 (4-3) 中 x_i 在左边的系数是

$$\sum_{i=0}^{s} b_i |\{I \in \mathbb{P}_i(X) : I \subset A_{i_0}, I \subset A_i\}| = \sum_{i=0}^{s} b_i \binom{|A_i \cap A_{i_0}| - 1}{i}.$$

这就证明了等式 (4-3)。

如果 $n \notin A_{i_0}$，将 v_i 代入到等式 (4-2) 中，我们得到

$$\sum_{i=0}^{s} a_i \sum_{I \in \mathbb{P}_i(X), I \subset A_{i_0}} L_I(v) = \sum_{A_i \in \mathcal{A}} g(|A_i \cap A_{i_0}|) v_i.$$

由于 v 是方程组 (4-1) 的一个解，显然等式左边等于 0。当 $i \ne i_0$，对 $A_i \in \mathcal{A}$，有 $|A_i \cap A_{i_0}| \pmod p \in L$，因此 $g(|A_i \cap A_{i_0}|) = 0$。由上面得到等式的右边等于 $g(|A_{i_0}|)v_{i_0}$。综上我们有 $g(|A_{i_0}|)v_{i_0} = 0$。由于 $L \cap K = \emptyset$，我们可以通过 $g(|A_{i_0}|) \ne 0$，得到 $v_{i_0} = 0$。这与 v 是一个非零向量这一假定是矛盾的。

如果 $n \in A_{i_0}$，将 v_i 代入到等式 (4-3) 中，我们得到

$$\begin{aligned}\sum_{i=0}^{s} b_i \sum_{I \in \mathbb{P}_i(X), I \subset A_{i_0}} L_I(v) &= \sum_{i=1}^{t} h(|A_i \cap A_{i_0}|) v_i + \sum_{i \ge t+1} h(|A_i \cap A_{i_0}| - 1) v_i \\ &= \sum_{i \ge t+1} h(|A_i \cap A_{i_0}| - 1) v_i. \quad \text{(由于对任意 } i \le t,\ v_i = 0)\end{aligned}$$

由于 $h(|A_i \cap A_{i_0}| - 1) = g(|A_i \cap A_{i_0}|)$，我们可以通过和上文类似的讨论而得出矛盾。 □

由上述的命题我们可以得到：

$$|\mathcal{A}| \le \dim(\{L_I : I \in \cup_{i=0}^{s} \mathbb{P}_i(X)\}),$$

其中我们记 $\dim(\{L_I : I \in \cup_{i=0}^{s} \mathbb{P}_i(X)\})$ 为向量空间 $\{L_I : I \in \cup_{i=0}^{s} \mathbb{P}_i(X)\}$ 的维数。本节余下的部分我们将估计这个空间的维数。

引理 4.18. *对任意 $i \in \{0, 1, \ldots, s-2r+1\}$，$I \in \mathbb{P}_i(X)$，在 $\mathbb{F}_p$ 上线性形*

$$\sum_{H \in \mathbb{P}_{i+2r}(X), I \subset H} L_H$$

可以被线性形 $\{L_H : i \le |H| \le i+2r-1, H \subset X\}$ 所表出。

证明. 定义

$$f(x)=\left(\prod_{j=1}^{r}(x-(k_j-i))\right)\times\left(\prod_{j=1}^{r}(x-(k_j-1-i))\right).$$

我们分两种情况进行讨论：

(a) 对所有 i 均有 $i \pmod p \notin K$ 和 $i+1 \pmod p \notin K$。在这种情形下，对任意 $k_j \in K$，在 $\mathbb{F}_p$ 中 $k_j-i\neq 0$ 且 $k_j-i-1\neq 0$，因此 $c=(k_1-i)(k_2-i)\cdots(k_r-i)(k_1-i-1)\cdots(k_r-i-1)\neq 0$。显然存在常数 $a_1,a_2,\ldots,a_{2r-1}\in\mathbb{F}_p$，$a_{2r}=(2r)!\in\mathbb{F}_p-\{0\}$ 使得

$$a_1\binom{x}{1}+a_2\binom{x}{2}+\cdots+a_{2r}\binom{x}{2r}=f(x)-c$$

成立。

下面我们将证明：

$$\sum_{j=1}^{2r}a_j\sum_{H\in\mathbb{P}_{i+j}(X),I\subset H}L_H=-cL_I. \tag{4-4}$$

事实上式子的两边都是关于 x_A 的线性形，x_A 在左边的系数等于 $\sum_{j=1}^{2r}a_j|\{H|I\subset H\subset A,n\notin H,|H|=i+j\}|$。分情况可以得到如下事实：

$$\begin{cases}0,\ \text{如果 } I\not\subset A;\\ a_1\binom{|A|-i}{1}+a_2\binom{|A|-i}{2}+\cdots+a_{2r}\binom{|A|-i}{2r},\ \text{如果 } I\subset A \text{ 和 } n\notin A;\\ a_1\binom{|A|-i-1}{1}+a_2\binom{|A|-i-1}{2}+\cdots+a_{2r}\binom{|A|-i-1}{2r},\ \text{如果 } I\subset A \text{ 和 } n\in A.\end{cases}$$

由上面的等式我们可以得到：

$$\sum_{j=1}^{2r}a_j\binom{|A|-i}{j}=f(|A|-i)-c=-c \text{ 由于 } |A| \pmod p\in K;$$

$$\sum_{j=1}^{2r}a_j\binom{|A|-i-1}{j}=f(|A|-i-1)-c=-c \text{ 由于 } |A| \pmod p\in K.$$

x_A 在右边的系数显然是相等的。这就证明了 (4-4)。

我们将 (4-4) 用另一种方式表达：

$$\sum_{H\in\mathbb{P}_{i+2r}(X),I\subset H}L_H=-\frac{1}{(2r)!}(cL_I+\sum_{j=1}^{2r-1}a_j\sum_{H\in\mathbb{P}_{i+j}(X),I\subset H}L_H).$$

(b) 存在 i 使得 $i \pmod p \in K$ 或者 $i+1 \pmod p \in K$ 成立。在这种情形下，式子 $(x-(k_1-i))(x-(k_2-i))\cdots(x-(k_r-i))(x-(k_1-i-1))\cdots(x-(k_r-i-1))$ 的常数项在 $\mathbb{F}_p$ 上是 0。因此存在常数 $a_1, a_2, \ldots, a_{2r-1} \in \mathbb{F}_p$，$a_{2r} = (2r)! \in \mathbb{F}_p - \{0\}$ 使得 $a_1\binom{x}{1} + a_2\binom{x}{2} + \cdots + a_{2r}\binom{x}{2r} = f(x)$ 成立。我们得到

$$\sum_{j=1}^{2r} a_j \sum_{H\in\mathbb{P}_{i+j}(X), I\subset H} L_H = 0 \ \ \forall I \in \mathbb{P}_i(X),$$

也就说

$$\sum_{H\in\mathbb{P}_{i+2r}(X), I\subset H} L_H = -\frac{1}{(2r)!}\Big(\sum_{j=1}^{2r-1} a_j \sum_{H\in\mathbb{P}_{i+j}(X), I\subset H} L_H\Big).$$

这就完成了引理的证明。 □

推论 4.19. *在与引理 4.18 相同的条件下我们有*

$$\begin{aligned}&\langle L_H : H \in \cup_{j=i}^{i+2r-1}\mathbb{P}_j(X)\rangle \\ &= \big\langle L_H : H \in \cup_{j=i}^{i+2r-1}\mathbb{P}_j(X)\big\rangle + \left\langle \sum_{H\in\mathbb{P}_{i+2r}(X), I\subset H} L_H : I \in \mathbb{P}_i(X)\right\rangle\end{aligned}$$

其中 $\langle L_H : H \in \cup_{j=i}^{i+2r-1}\mathbb{P}_j(X)\rangle$ 是由 $\{L_H : H \in \cup_{j=i}^{i+2r-1}\mathbb{P}_j(X)\}$ 所张成的向量空间。

引理 4.20. *对任意的正整数 u, v 满足 $u < v < p$ 且 $u+v \le n-1$，我们有*

$$\dim\left(\frac{\langle L_J : J \in \mathbb{P}_v(X)\rangle}{\langle \sum_{J\in\mathbb{P}_v(X), I\subset J} L_J : I \in \mathbb{P}_u(X)\rangle}\right) \le \binom{n-1}{v} - \binom{n-1}{u}.$$

其中 $\frac{A}{B}$ 是两个向量空间 A 和 B 的商空间。

引理 4.21. *对任意 $i \in \{0, 1, \ldots, s-2r+1\}$,*

$$\begin{aligned}&\binom{n-1}{i} + \binom{n-1}{i+1} + \cdots + \binom{n-1}{i+2r-1} + \dim\left(\frac{\langle L_H : H \in \cup_{j=i}^{s}\mathbb{P}_j(X)\rangle}{\langle L_H : H \in \cup_{j=i}^{i+2r-1}\mathbb{P}_j(X)\rangle}\right) \\ &\le \binom{n-1}{s-2r+1} + \binom{n-1}{s-2r+2} + \cdots + \binom{n-1}{s}.\end{aligned}$$

证明. 我们对 $s-2r+1-i$ 进行归纳。当 $s-2r+1-i=0$ 时，结论是显然成立的。不妨设结论在 $s-2r+1-i<l$ 时成立。我们只需证明结论在 $s-2r+1-i=l$ 时也同样成立

即可。由定理的条件我们可以得到 $i+i+2r \leq (s-2r)+(s-2r)+2r \leq n-1$。结合推论 4.19 和引理 4.20，我们有

$$\begin{aligned}
&\dim\left(\frac{\langle L_H : H \in \cup_{j=i}^{i+2r}\mathbb{P}_j(X)\rangle}{\langle L_H : H \in \cup_{j=i}^{i+2r-1}\mathbb{P}_j(X)\rangle}\right)\\
&\quad=\dim\left(\frac{\langle L_H : H \in \cup_{j=i}^{i+2r-1}\mathbb{P}_j(X)\rangle + \langle L_H : H \in \mathbb{P}_{i+2r}(X)\rangle}{\langle L_H : H \in \cup_{j=i}^{i+2r-1}\mathbb{P}_j(X)\rangle + \langle \sum_{H\in\mathbb{P}_{i+2r}(X), I\subset H} L_H : I \in \mathbb{P}_i(X)\rangle}\right)\\
&\quad\leq\dim\left(\frac{L_H : H \in \mathbb{P}_{i+2r}(X)}{\sum_{H\in\mathbb{P}_{i+2r}(X), I\subset H} L_H : I \in \mathbb{P}_i(X)}\right)\\
&\quad\leq\binom{n-1}{i+2r}-\binom{n-1}{i}.
\end{aligned}$$

现在我们可以来证明这个引理，

$$\begin{aligned}
&\binom{n-1}{i}+\binom{n-1}{i+1}+\cdots+\binom{n-1}{i+2r-1}+\dim\left(\frac{\langle L_H : H \in \cup_{j=i}^{s}\mathbb{P}_j(X)\rangle}{\langle L_H : H \in \cup_{j=i}^{i+2r-1}\mathbb{P}_j(X)\rangle}\right)\\
&\quad=\binom{n-1}{i}+\binom{n-1}{i+1}+\cdots+\binom{n-1}{i+2r-1}+\dim\left(\frac{\langle L_H : H \in \cup_{j=i}^{i+2r}\mathbb{P}_j(X)\rangle}{\langle L_H : H \in \cup_{j=i}^{i+2r-1}\mathbb{P}_j(X)\rangle}\right)\\
&\qquad+\dim\left(\frac{\langle L_H : H \in \cup_{j=i}^{s}\mathbb{P}_j(X)\rangle}{\langle L_H : H \in \cup_{j=i}^{i+2r}\mathbb{P}_j(X)\rangle}\right)\\
&\quad=\binom{n-1}{i}+\binom{n-1}{i+1}+\cdots+\binom{n-1}{i+2r-1}+\dim\left(\frac{\langle L_H : H \in \cup_{j=i}^{i+2r}\mathbb{P}_j(X)\rangle}{\langle L_H : H \in \cup_{j=i}^{i+2r-1}\mathbb{P}_j(X)\rangle}\right)\\
&\qquad+\dim\left(\frac{\langle L_H : H \in \mathbb{P}_i(X)\rangle + \langle L_H : H \in \cup_{j=i+1}^{s}\mathbb{P}_j(X)\rangle}{\langle L_H : H \in \mathbb{P}_i(X)\rangle + \langle L_H : H \in \cup_{j=i+1}^{i+2r}\mathbb{P}_j(X)\rangle}\right)\\
&\quad\leq\binom{n-1}{i}+\binom{n-1}{i+1}+\cdots+\binom{n-1}{i+2r-1}+\dim\left(\frac{\langle L_H : H \in \cup_{j=i}^{i+2r}\mathbb{P}_j(X)\rangle}{\langle L_H : H \in \cup_{j=i}^{i+2r-1}\mathbb{P}_j(X)\rangle}\right)\\
&\qquad+\dim\left(\frac{\langle L_H : H \in \cup_{j=i+1}^{s}\mathbb{P}_j(X)\rangle}{\langle L_H : H \in \cup_{j=i+1}^{i+2r}\mathbb{P}_j(X)\rangle}\right)\\
&\quad\leq\binom{n-1}{i}+\binom{n-1}{i+1}+\cdots+\binom{n-1}{i+2r-1}+\binom{n-1}{i+2r}-\binom{n-1}{i}\\
&\qquad+\dim\left(\frac{\langle L_H : H \in \cup_{j=i+1}^{s}\mathbb{P}_j(X)\rangle}{\langle L_H : H \in \cup_{j=i+1}^{i+2r}\mathbb{P}_j(X)\rangle}\right)\\
&\quad=\binom{n-1}{i+1}+\cdots+\binom{n}{i+2r}+\dim\left(\frac{\langle L_H : H \in \cup_{j=i+1}^{s}\mathbb{P}_j(X)\rangle}{\langle L_H : H \in \cup_{j=i+1}^{i+2r}\mathbb{P}_j(X)\rangle}\right)\\
&\quad\leq\binom{n-1}{s-2r+1}+\cdots+\binom{n-1}{s},
\end{aligned}$$

其中最后一步来自于对 $s-2r+1-(i+1)<l$ 的归纳假设。 □

现在我们到了来证明定理 4.14 的时候。

定理 4.14 的证明.

$$\begin{aligned}
|\mathcal{A}| \leq & \dim(\langle L_H : H \in \cup_{i=0}^{s}\mathbb{P}_i(X)\rangle) \\
\leq & \dim(\langle L_H : H \in \cup_{i=0}^{2r-1}\mathbb{P}_i(X)\rangle) + \dim\left(\frac{\langle L_H : H \in \cup_{i=0}^{s}\mathbb{P}_j(X)\rangle}{\langle L_H : H \in \cup_{i=0}^{2r-1}\mathbb{P}_j(X)\rangle}\right) \\
\leq & \binom{n-1}{0} + \binom{n-1}{1} + \cdots + \binom{n-1}{2r-1} + \dim\left(\frac{\langle L_H : H \in \cup_{i=0}^{s}\mathbb{P}_j(X)\rangle}{\langle L_H : H \in \cup_{i=0}^{2r-1}\mathbb{P}_j(X)\rangle}\right) \\
\leq & \binom{n-1}{s-2r+1} + \binom{n-1}{s-2r+2} + \cdots + \binom{n-1}{s} \text{ 在引理 4.21 中取 } i = 0.
\end{aligned}$$

□

4.3 定理 4.15 的证明

在本节中，我们规定 p 是一素数，$x = (x_1, x_2, \ldots, x_n)$ 表示取值在 0、1 上的 n 维向量。我们称一个多项式 $f(x)$ 是多线性的，如果任一变量 x_i 的次数在每个单项式中至多是 1。显然地，当 x_i 的取值只能是 0 或 1 时，任意一个多项式我们都可以看成是多线性的。对 $[n]$ 中的一个子集 A，我们用 $v = (v_1, v_2, \ldots, v_n)$，$v_i = 1$ 当 $i \in A$ 时，否则 $v_i = 0$ 来定义它所对应的特征向量 v_A。

设 $L = \{l_1, l_2, \ldots, l_s\}$ 和 $K = \{k_1, k_2, \ldots, k_r\}$ 是 $\{0, 1, \ldots, p-1\}$ 中的两个互不相交的子集，其中 K 中的元素我们用升序来排列。不失一般性，我们假设存在整数 t，使得 $n \in A_j$ 当 $j \geq t+1$ 时，而当 $1 \leq j \leq t$ 时 $n \notin A_j$。

对任意集合 $A_j \in \mathcal{A}$，我们定义其对应的多项式

$$f_{A_j}(x) = \prod_{i=1}^{s}(v_{A_j} \cdot x - l_i).$$

容易看出 $f_{A_j}(x)$ 是一个次数至多为 s 的多线性多项式。

设 Q 为 $[n-1]$ 中所有大小至多为 $s-1$ 的集合所构成的集族，那么 $|Q| = \sum_{i=0}^{s-1}\binom{n-1}{i}$。对每个子集 $L \in Q$，定义

$$q_L(x) = (1 - x_n)\prod_{i \in L} x_i.$$

$q_L(x)$ 也是一个次数至多为 s 的多线性多项式。

记 $K - 1 = \{k_i - 1 | k_i \in K\}$，容易看出 $|K \cup (K-1)| \leq 2r$。令

$$g(x) = \prod_{h \in K \cup (K-1)}\left(\sum_{i=1}^{n-1} x_i - h\right).$$

设 Q 为 $[n-1]$ 中所有大小至多为 $s-2r$ 的集合所构成的集族，那么 $|W|=\sum_{i=0}^{s-2r}\binom{n-1}{i}$。对每个子集 $I\in W$，定义

$$g_I(x)=g(x)\prod_{i\in I}x_i.$$

$g_I(x)$ 也是一个次数至多为 s 的多线性多项式。

我们希望证明这些多项式

$$\{f_{A_i(x)}|1\le i\le m\}\cup\{q_L(x)|L\in Q\}\cup\{g_I(x)|I\in W\}$$

在域 $\mathbb{F}_p$ 上是线性无关的。如果在域 $\mathbb{F}_p$ 上有一个线性组合使得这些多项式的和为 0：

$$\sum_{i=1}^{m}a_if_{A_i}(x)+\sum_{L\in Q}b_Lq_L(x)+\sum_{I\in W}u_Ig_I(x)=0. \tag{4-5}$$

断言1. 对任意 i 满足 $n\in A_i$，均有 $a_i=0$。

如果存在 i_0，使得 $n\in A_{i_0}$ 且 $a_{i_0}\neq 0$。由于 $n\in A_{i_0}$，$q_L(v_{A_{i_0}})=0$ 对任意 $L\in Q$，在考虑到 $f_{A_j}(v_{i_0})=0$ 对 $j\neq i_0$ 并且 $g(v_{i_0})=0$。我们将 $x=v_{A_{i_0}}$ 代入到等式 (4-5) 中，可以得到 $a_{i_0}f_{A_{i_0}}(v_{A_{i_0}})\equiv 0\pmod p$。又因为 $f_{A_{i_0}}(v_{A_{i_0}})\neq 0$，我们得到 $a_{i_0}=0$，这是与假设矛盾的。

断言2. 对任意 i 满足 $n\notin A_i$，均有 $a_i=0$。

利用断言 1，我们可以得到下式：

$$\sum_{i=1}^{t}a_if_{A_i}(x)+\sum_{L\in Q}b_Lq_L(x)+\sum_{I\in W}u_Ig_I(x)=0. \tag{4-6}$$

如果存在 i_0，使得 $n\notin A_{i_0}$ 且 $a_{i_0}\neq 0$。令 $v'_{i_0}=v_{i_0}+(0,0,\ldots,0,1)$，那么对任意 $L\in Q$，都有 $q_L(v'_{i_0})=0$。注意到对满足 $n\notin A_j$ 的 j 都有 $f_{A_j}(v'_{i_0})=f_{A_j}(v_{i_0})$，并且 $g(v'_{i_0})=0$。将 $x=v'_{i_0}$ 代入到等式 (4-6) 中，我们有 $a_{i_0}f_{A_{i_0}}(v'_{i_0})=a_{i_0}f_{A_{i_0}}(v_{i_0})\equiv 0\pmod p$。这意味着 $a_{i_0}=0$，与假设相矛盾。

断言3. 对任意 $L\in Q$，都有 $b_L=0$。

通过断言 1 和 2，我么有

$$\sum_{L\in Q}b_Lq_L(x)+\sum_{I\in W}u_Ig_I(x)=0. \tag{4-7}$$

将 $x_n=0$ 代入到等式 (4-7) 中，得到

$$\sum_{L\in Q}b_L\prod_{i\in L}x_i+\sum_{I\in W}u_Ig_I(x)=0.$$

我们将上面的式子从等式 (4-7) 中减去，

$$\sum_{L\in Q} b_L \left(x_n \prod_{i\in L} x_i \right) = 0.$$

令上式中 $x_n = 1$，

$$\sum_{L\in Q} b_L \prod_{i\in L} x_i = 0.$$

不难验证 $\prod_{i\in L} x_i$（$L \in Q$）这些多项式是线性无关的。因此我们得到 $b_L = 0$。

由断言 $1-3$，我们将原式化简为

$$\sum_{I\in W} u_I g_I(x) = 0.$$

因此我们只需证明 g_I 是线性无关的。

设 N 是一个正整数，$H = \{h_1, h_2, \ldots, h_u\}$ 是 $[N]$ 中的用升序排列的子集。如果 $h_1 \geq g-1$，$N - h_u \geq g-1$，或者存在 i（$1 \leq i \leq u-1$）使得 $h_{i+1} - h_i \geq g$，那么我们称 H 有一个大小为 g 的间隙。下面的引理由 Alon，Babai 和 Suzuki 在文献[3] 中给出了严格的证明。

引理 4.22. *设 H 是 $\{0, 1, \ldots, p-1\}$ 中的子集，将多项式 $\prod_{h\in H}(x_1 + x_2 + \cdots + x_N - h)$ 记为 $p(x)$。如果集合 $(H+p\mathbb{Z})\cap[N]$ 的间隙大于等于 $g+1$，那么多项式 $\{p_I(x) : |I| \leq g-1, I \in N\}$ 在 $\mathbb{F}_p$ 上是线性无关的，其中 $p_I(x) = p(x)\prod_{i\in I} x_i$。*

为了应用引理 4.22，我们定义：$H = (K \cup (K-1) + p\mathbb{Z}) \cap [n-1]$，再将 $n-1$ 分成如下四种情形进行讨论：

1. $s + k_r - 1 \leq n - 1 < p + k_1 - 1$;
2. $s + k_r - 1 < p + k_1 - 1 \leq n - 1$;
3. $(s - 2r + 1) + k_r < p + k_1 - 1 \leq s + k_r - 1 \leq n - 1$;
4. $p + k_1 - 1 \leq (s - 2r + 1) + k_r \leq s + k_r - 1 \leq n - 1$.

Case 1: $s + k_r - 1 \leq n - 1 < p + k_1 - 1$。

由于 $n - 1 < p + k_1 - 1$，H 仅由 $\{k_1 - 1, k_1, \ldots, k_r\}$ 所构成。又由 $s + k_r - 1 \leq n - 1$，我们得到 $n - 1 - k_r \geq s - 1 \geq s - 2r + 1$，集合 H 的间隙大于等于 $s - 2r + 2$。

Case 2: $s+k_r-1<p+k_1-1\leq n-1$。

由于 $n-1\geq p+k_1-1$，H 至少包含如下的元素：$\{k_1-1,k_1,\ldots,k_r,p+k_1-1\}$。由 $s+k_r-1<p+k_1-1$，我们可以导出 $(p+k_1-1)-k_r\geq s\geq s-2r+2$，因此集合 H 的间隙大于等于 $s-2r+2$。

Case 3: $(s-2r+1)+k_r<p+k_1-1\leq s+k_r-1\leq n-1$。

由于 $n-1\geq p+k_1-1$，H 至少包含如下的元素：$\{k_1-1,k_1,\ldots,k_r,p+k_1-1\}$。由 $(s-2r+1)+k_r<p+k_1-1$，我们有 $(p+k_1-1)-k_r>s-2r+1$，因此集合 H 的间隙大于等于 $s-2r+2$。

利用引理 4.22，我们可以得到在 1–3 的情形下，多项式

$$\{f_{A_i(x)}|1\leq i\leq m\}\cup\{q_L(x)|L\in Q\}\cup\{g_I(x)|I\in W\}$$

在 $\mathbb{F}_p$ 上是线性无关的。由于所有次数至多为 s 的单项式构成次数至多为 s 的多线性多项式空间的一组基，

$$|\mathcal{A}|+\sum_{i=0}^{s-1}\binom{n-1}{i}+\sum_{i=0}^{s-2r}\binom{n-1}{i}\leq\sum_{i=0}^{s}\binom{n}{i},$$

这意味着

$$|\mathcal{A}|\leq\binom{n-1}{s}+\binom{n-1}{s-1}+\cdots+\binom{n-1}{s-2r+1}.$$

至此我们完成了 1–3 情形的证明。

由于在定理 4.14 中，我们已经证明了在 $n\geq 2s-2r+1$ 的条件下，结论是正确的。最后我们只需证明在情形 4 和 $n\leq 2s-2r$ 时，也有相同的结论。在这部分的证明中我们将利用 Hwang 和 Kim 在证明 ABS 猜想中所采用的相关技巧。

由于 $p+k_1-1\leq(s-2r+1)+k_r\leq s+k_r-1\leq n-1\leq 2s-2r-1$，我们可以依次得到 $k_r\leq s-2r$，$r+s\leq p\leq s-2r+2+k_r-k_1\leq 2s-4r+2$，$s\geq 5r-2$。由于 $n\leq 2s-2r<2p$，存在 $1\leq c\leq r$ 使得 $|A_i|\in(K+p\mathbb{Z})\cap[n]=\{k_1,k_2,\ldots,k_r,p+k_1,\ldots,p+k_c\}$ 成立。

$$|\mathcal{A}|\leq\binom{n}{k_1}+\binom{n}{k_2}+\cdots+\binom{n}{k_r}+\binom{n}{p+k_1}+\cdots+\binom{n}{p+k_c}.$$

我们要证明不等式右边小于等于 $\binom{n-1}{s}+\binom{n-1}{s-1}+\ldots+\binom{n-1}{s-2r+1}=\binom{n}{s}+\binom{n}{s-2}+\ldots+\binom{n}{s-2r+2}$。由于 $s+r+k_1-1\leq p+k_1-1\leq(s-2r+1)+k_r$，我们有 $k_r\geq 3r-2+k_1$。由于 $2s-2r\geq n\geq s+k_r\geq s+3r-2+k_1$，令 $n=2s-2r-\delta$，其中 $0\leq\delta\leq s-5r+2$。又序列 $\{\binom{n}{k}\}$ 是单峰的并且关于 $n/2$ 是对称的，$|s-n/2|=r+\delta/2>r-\delta/2-2=|n/2-(s-2r+2)|$。

综上我们得到：

$$\min\left[\binom{n}{s},\binom{n}{s-2},\ldots,\binom{n}{s-2r+2}\right]=\binom{n}{s}. \tag{4-8}$$

由于 $n=2s-2r-\delta\geq p+k_c\geq r+s+k_c$，我们有 $k_c\leq s-3r-\delta$。对 $1\leq i\leq c$，我们将 k_i 用 $k_i=s-3r-\delta-a_i$ 来表示，其中 $0\leq a_i\leq s-3r-\delta$。因此，我们有 $p+k_i\geq r+s+k_i=2s-2r-\delta-a_i$，其中 $1\leq i\leq c$。由于 $2s-2r-\delta-a_i\geq s+r>n/2$，可得:

$$\sum_{i=1}^{c}\left(\binom{n}{k_i}+\binom{n}{p+k_i}\right)\leq\sum_{i=1}^{c}\left(\binom{n}{s-3r-\delta-a_i}+\binom{n}{2s-2r-\delta-a_i}\right).$$

对于 $c+1\leq i\leq r$，我们可以导出 $k_i\leq k_r<s-2r-\delta<n/2$。注意到 $|s-n/2|=r+\delta/2=|n/2-(s-2r-\delta)|$，对任意 $c+1\leq i\leq r$，我们得到 $\binom{n}{k_i}\leq\binom{n}{s}$。

$$\begin{aligned}|\mathcal{A}|&\leq\sum_{i=1}^{c}\left(\binom{n}{k_i}+\binom{n}{p+k_i}\right)+\sum_{i=c+1}^{r}\binom{n}{k_i}\\&\leq\sum_{i=1}^{c}\left(\binom{n}{s-3r-\delta-a_i}+\binom{n}{2s-2r-\delta-a_i}\right)+(r-c)\binom{n}{s}.\end{aligned}$$

最后我们利用下面的引理来完成整个证明。

引理 4.23. [42] 对任意 $0\leq c<k\leq n/2$，均有

$$\binom{n}{k-1-c}+\binom{n}{c}\leq\binom{n}{k}.$$

令 $k=n-s=s-2r-\delta<n/2$，利用引理 4.23，对任意 $0\leq a\leq s-3r-\delta<k$，我们有

$$\begin{aligned}&\binom{n}{s-3r-\delta-a}+\binom{n}{2s-2r-\delta-a}\\&=\binom{n}{n-s-r-a}+\binom{n}{n-a}\\&=\binom{n}{k-r-a}+\binom{n}{a}\\&\leq\binom{n}{k-1-a}+\binom{n}{a}\\&\leq\binom{n}{k}=\binom{n}{s}.\end{aligned}$$

我们来完成在情形 4 下定理 4.15 的证明。

$$|\mathcal{A}|\leq\sum_{i=1}^{c}\left(\binom{n}{s-3r-\delta-a_i}+\binom{n}{2s-2r-\delta-a_i}\right)+(r-c)\binom{n}{s}\leq r\binom{n}{s}.$$

由 (4-8)，我们有

$$|\mathcal{A}|\leq\binom{n}{s}+\binom{n}{s-2}+\cdots+\binom{n}{s-2r+2}=\binom{n-1}{s}+\binom{n-1}{s-1}+\cdots+\binom{n-1}{s-2r+1}.$$

4.4 小结

L-相交系的研究已经有了三十多年的历史，但是在模意义下的结果（除了在均匀的情形下）至今还得不到上下界完全一致的结论。在本章中，我们综合应用了两种线性代数方法改进了 Alon-Babai-Suzuki 不等式，事实上我们仅仅将次主项从 $\binom{n}{s-1}$ 提高到 $\binom{n}{s-2}$，离 Snevily 在 90 年代关于次主项为 0 的猜想还有非常大的距离。我们认为可能单纯地使用线性代数方法很难去完全解决 Snevily 猜想，应该要在L-相交系的研究中引入新的方法和技巧。

5 私人信息检索

5.1 介绍

私人信息检索（PIR）最早在文献[20]中提出：假设我们有 n 比特的数据和 k 个服务器，每个服务器都包含有全部的信息，因此它的总存储量为 nk，一个基于 k 个服务器的 PIR 过程允许用户检索到需要的信息，但是服务器无法得知用户的需求。比如说，假设数据集为 $\mathbf{x} = (x_1, x_2, \ldots, x_n)$，某个用户想要知道第 i 位的信息 x_i。在一个基于 2 个服务器的 PIR 过程中，用户可以随机的选取一个向量 $\mathbf{v} \in \{0,1\}^n$。第一个服务器接收到问询 $\mathbf{v}$，向用户反馈 $\mathbf{v} \cdot \mathbf{x}$。第二个服务器收到问询 $\mathbf{v} + \mathbf{e_i}$，向用户反馈 $(\mathbf{v} + \mathbf{e_i}) \cdot \mathbf{x}$。那么用户可以通过 $x_i = (\mathbf{v} + \mathbf{e_i}) \cdot \mathbf{x} - \mathbf{v} \cdot \mathbf{x}$ 检索到自己想要的信息。但是由于 $\mathbf{v}$ 是随机选取的，每个服务器都得不到任何关于用户的有用信息。

近来，PIR 过程与分布式存储的思想产生了联系[4,13,33,69,73]，将原有的在每个服务器上存储全部的信息，改变成服务器上都只存储部分信息。在开创性的工作[34,35]，Fazeli 等人证明了利用 m 个服务器（其中 $m > k$），可以大大地降低整个服务器的总存储量。我们继续以上面的问题为例，设三个服务器分别存储：$\mathbf{x}' = (x_1, \ldots, x_{n/2})$，$\mathbf{x}'' = (x_{n/2+1}, \ldots, x_n)$ 和 $\mathbf{x}' + \mathbf{x}''$。不失一般性，我们假设用户希望检索信息 x_i。用户可以随机的选取向量 $\mathbf{u} \in \{0,1\}^{n/2}$，向三个服务器分别问询：$\mathbf{u}$，$\mathbf{u} + \mathbf{e_i}$ 和 $\mathbf{u} + \mathbf{e_i}$。那么通过 $x_i = -\mathbf{u} \cdot \mathbf{x}' - (\mathbf{u} + \mathbf{e_i}) \cdot \mathbf{x}'' + (\mathbf{u} + \mathbf{e_i}) \cdot (\mathbf{x}' + \mathbf{x}'')$，用户可以成功的检索到想要的信息 x_i。相比于原有的方案，它将总存储量从 $2n$ 降到了 $\frac{3n}{2}$。

在文献[35]中，设计一个 PIR 的过程最终可以归结为设计一个相对应的 PIR 阵列码。给定正整数 t，m，p 和 k，一个 $[t \times m, p]$ *阵列码* 是一个 $t \times m$ 的矩阵，其中的每个元素都是 $\{x_1, \ldots, x_p\}$ 的线性组合。如果对任意 $i \in \{1, 2, \ldots, p\}$，都有 k 个互不相交的集合 $S_1, S_2, \ldots, S_k$，使得每个集合 S_j 可以线性地张成 x_i，我们称这个阵列码具有 *k-PIR 性质*。进一步，我们称这样的阵列码是参数为 $[t \times m, p]$ k-PIR 阵列码。下面我们给出 $[3 \times 6, 6]$ 4-PIR 阵列码的一个例子：

x_1	x_2	x_3	x_4	x_5	x_6
x_2	x_3	x_4	x_5	x_6	x_1
x_3+x_4	x_4+x_5	x_5+x_6	x_6+x_1	x_1+x_2	x_2+x_3

我们可以非常容易的去验证这个矩阵满足条件，比方说，x_1 可以由 $S_1=\{1\}$，$S_2=\{6\}$，$S_3=\{2,5\}$ 和 $S_4=\{3,4\}$ 所生成。

一个 $[t\times m,p]$ k-PIR 阵列码和 PIR 过程的关系如下：将 n 个比特的数据集划分成 p 部分 $\{x_1,x_2,\ldots,x_p\}$，每一部分用有限域里的一个元素来表示。阵列中的每一列表示一个服务器。每个服务器存储 t 个 $\{x_1,x_2,\ldots,x_p\}$ 的线性组合。这个方案的总存储量是 ntm/p，当 $tm/p<k$ 时，它比原有方案的 nk 要好。令 s 是每个服务器的存储量和总数据量的比例，也就是说 $s=\frac{n}{nt/p}=p/t$。我们的目标是总的存储负荷最低，因此我们希望 $\frac{nk}{ntm/p}=s\frac{k}{m}$ 越大越好。这个阵列码的码率可以定义为 k/m。现在问题可以归结如下：给定 s 和 t，我们希望设计一个 $[t\times m,p]$ k-PIR 阵列码使得码率 k/m 达到最大，记为 $g(s,t)$。更一般的，我们可以去考虑分析 $g(s)=\overline{\lim}_{t\to\infty}g(s,t)$。

最近，这个问题得到了 Blackburn 和 Etzion 的关注，他们希望去构造达到最优码率的阵列码。下面我们将列出已知的结果，详细地证明可以去参考文献[6]。

定理 5.1. *对任意给定的正整数 s，$g(s,1)\le\frac{2^{s-1}}{2^s-1}$ 等号成立当且仅当 k 可以被 2^{s-1} 整除。*

定理 5.2. *对任意整数 $s\ge 3$，我们有 $g(s,s-1)\ge\frac{s}{2s-1}$。*

定理 5.3. *对任意有理数 $s>1$，我们有 $g(s)\le\frac{s+1}{2s}$，并且不存在 t 使得 $g(s,t)=\frac{s+1}{2s}$ 成立。*

定理 5.4. *对任意整数 $t\ge 2$ 和正整数 d，$s=1+\frac{d}{t}$ 和 $p=t+d$，我们有*

$$g(1+\frac{d}{t},t)\le\frac{(2d+1)t+d^2}{(t+d)(2d+1)}=1-\frac{d^2+d}{(t+d)(2d+1)}.$$

特别地，当 $1<s\le 2$，这个上界是紧的，$g(2,t)=\frac{3t+1}{4t+2}$；对 $t\ge 2$，$1\le d\le t-1$，$g(1+\frac{d}{t},t)=\frac{(2d+1)t+d^2}{(t+d)(2d+1)}$。

定理 5.5. *存在如下参数的 PIR 阵列码：*

1. *令 $s=\frac{rt-(r-2)r-1}{t}$，$p=rt-(r-2)r-1$，其中 $3\le r\le t$。那么 $g(s,t)\ge\frac{1}{2}+\frac{t-r+1}{2(rt-(r-2)r-1)}$。*

2. 令 $s = r + d/t$，$p = rt + d$，其中 $r \geq 2$ 是一个整数，$t \geq r$，$1 \leq d \leq t-1$。那么 $g(s,t) \geq 1 - \frac{(rt+d-t+r)(rt+d-t)}{(rt+d)(2rt+2d-2t+r)}$。

3. 令 $s > 2$ 是一个整数，$t \geq s$，那么 $g(s,t) \geq \frac{st+t+1}{s(2t+1)}$。

4. 令 $s > 2$ 是一个整数，$(s-1)t = lb$，$t \geq l + b$，其中 l 和 b 是两个正整数，那么 $g(s,t) \geq \frac{s+1}{2s} - \frac{l}{2st}$。

给定 $t \geq 2$，$1 \leq d \leq t$ 且 $s = 1 + \frac{d}{t}$，尽管达到上界的 PIR 阵列码的构造已经由 Blackburn 和 Etzion 给出，但是在他们给出的构造中，阵列的列数大的惊人 $m = \binom{t+d}{t}\frac{v}{d} + \binom{t+d}{d+1}\frac{v}{t}$，其中 v 是 d 和 t 的最小公倍数。这就意味着在他们的设计方案中，需要海量的服务器才能达到最优的码率。从应用角度出发，构造只需少量服务器的方案是非常迫切的。因此我们将考虑如下问题：最小的服务器数量 m 使得码率 k/m 达到最优。在文献[6]中，$d = 1$ 的情形已经被解决。在本章中，我们将给出 $t > d^2 - d$ 情形下的一个完整回答。另一方面，在 $s > 2$ 的情形下，我们导出了 PIR 阵列码的全新上界改进了定理 5.4 的结果。

5.2 $1 < s \leq 2$: 最优 PIR 阵列码的构造

令 $s = 1 + d/t$，其中 $1 \leq d \leq t$，那么 $1 < s \leq 2$ 且 $p = ts = t + d$。关于码率的上界已经在定理 5.4 中给出：$1 - \frac{d^2+d}{p(2d+1)}$。令 ω 是 $d^2 + d$ 和 $p(2d+1)$ 的最大公因子，因此一个最优的 PIR 阵列码的最小可能服务器数目为 $p(2d+1)/\omega$。在本节中，我们将在某个区域内构造达到最小服务器数目的最优 PIR 阵列码。

由于 $\omega | d^2 + d$，我们可以将 ω 分解成 $\omega = \omega_1\omega_2$，其中 $\omega_1 | d$，$\omega_2 | (d+1)$。此外，由于 d 和 $d+1$ 是互素的，ω_1 和 ω_2 也是互素的。记 $d = \omega_1 d_1$，$d + 1 = \omega_2 d_2$。又由于 $\gcd(d, 2d+1) = 1$，$\gcd(d+1, 2d+1) = 1$，我们可以导出 $\omega | p$。记 $p = \mu\omega = \mu\omega_1\omega_2$，那么我们所希望的服务器数目可以写成 $m = \frac{p(2d+1)}{\omega} = \mu(2d+1)$。

首先我们定义两种形式的服务器。第一种形式的服务器只包含单个元素 $\{x_1, x_2, \ldots, x_p\}$，称之为单点服务器。这样的服务器包含 t 个单点元素，比如说 $\{y_1, y_2, \ldots, y_t\}$，我们将它记为 $\overline{A}$，其中 $A = \{x_1, x_2, \ldots, x_p\} \backslash \{y_1, y_2, \ldots, y_t\}$。第二种形式的服务器包含 $t-1$ 个单点元素，比如说 $\{z_1, z_2, \ldots, z_{t-1}\}$，而余下的一个位置是除了 $\{z_1, z_2, \ldots, z_{t-1}\}$ 这些元素的加和。我们称之为 Σ 型服务器，记为 ΣB，其中 $B = \{x_1, x_2, \ldots, x_p\} \backslash \{z_1, z_2, \ldots, z_{t-1}\}$。我们将要构造的 PIR 阵列码只应用到了上述两种简单形式的服务器。在本节中，如果不加另外的说

明，所有的指标都是在模 p 意义下的。

构造（给定 t，d 满足 $t > d^2 - d$）：

1. 我们有如下的 $\mu(d+1)$ 个单点服务器。对 $0 \le j \le \mu\omega_2 - 1$，定义 $A_j = \{x_{j+\alpha+\beta\mu\omega_2} : 0 \le \alpha \le d_1 - 1, 0 \le \beta \le \omega_1 - 1\}$。由于 $d_1\omega_1 = d < p = \mu\omega_1\omega_2$，我们有 $d_1 < \mu\omega_2$，集合 A_j 中没有重复的元素。因此 A_j 的大小恰好是 $d_1\omega_1 = d$。我们得到 $\mu(d+1) = \mu\omega_2 d_2$ 个单点服务器分别是 $\overline{A_0}, \overline{A_1}, \ldots, \overline{A_{\mu\omega_2-1}}$，每个恰好出现 d_2 次。

2. 我们有如下的 μd 个 Σ 形服务器。对 $0 \le j \le \mu\omega_1 - 1$，定义 $B_j = \{x_{j+\gamma d_1+\lambda\mu\omega_1} : 0 \le \gamma \le d_2 - 1, 0 \le \lambda \le \omega_2 - 1\}$。由于 $t > d^2 - d$，我们有 $d_1\omega_1(d_2-1)\omega_2 \le d^2 < p = \mu\omega_1\omega_2$，因此 $d_1(d_2 - 1) < \mu$， B_j 中没有重复的元素。B_j 的大小恰好是 $d_2\omega_2 = d + 1$。我们得到 $\mu d = \mu\omega_1 d_1$ 个 Σ 型服务器分别是 $\Sigma B_0, \Sigma B_1, \ldots, \Sigma B_{\mu\omega_1-1}$，每个恰好出现 d_1 次。

下面我们将分成 $t \ge d^2$ 和 $d^2 - d < t < d^2$ 分别说明上面的构造确实是最优的 PIR 阵列码。

5.2.1 $t \ge d^2$

定理 5.6. *对 $t \ge d^2$，存在码率 $\frac{k}{m}$ 达到 $g(s,t) = 1 - \frac{d^2+d}{p(2d+1)}$ 包含 $m = \mu(2d+1)$ 个服务器的 k-PIR 阵列码。*

证明. 我们先证明当 $t \ge d^2$ 时，对任意的 A_{j_1} 和 B_{j_2}，都有 $|A_{j_1} \bigcap B_{j_2}| \le 1$。否则，至少存在两个不同的元素在 $A_{j_1} \bigcap B_{j_2}$，这意味着存在 $\alpha_1, \alpha_2, \beta_1, \beta_2, \gamma_1, \gamma_2, \lambda_1, \lambda_2$ 使得

$$j_1 + \alpha_1 + \beta_1\mu\omega_2 \equiv j_2 + \gamma_1 d_1 + \lambda_1\mu\omega_1 \pmod{p},$$

$$j_1 + \alpha_2 + \beta_2\mu\omega_2 \equiv j_2 + \gamma_2 d_1 + \lambda_2\mu\omega_1 \pmod{p}.$$

上面两式相减可以得到

$$(\alpha_2 - \alpha_1) + (\beta_2 - \beta_1)\mu\omega_2 \equiv (\gamma_2 - \gamma_1)d_1 + (\lambda_2 - \lambda_1)\mu\omega_1 \pmod{p}.$$

那么我们有

$$(\alpha_2 - \alpha_1) \equiv (\gamma_2 - \gamma_1)d_1 \pmod{\mu}.$$

当 $t \ge d^2$ 时，$\mu = \frac{p}{\omega_1\omega_2} \ge \frac{d(d+1)}{\omega_1\omega_2} = d_1 d_2$。由于 $1 - d_1 \le \alpha_2 - \alpha_1 \le d_1 - 1$ 和 $1 - d_2 \le \gamma_2 - \gamma_1 \le d_2 - 1$，因此 $(\alpha_2 - \alpha_1) \equiv (\gamma_2 - \gamma_1)d_1 \pmod{\mu}$ 成立当且仅当 $\alpha_2 - \alpha_1 = \gamma_2 - \gamma_1 = 0$。那么

我们有 $(\beta_2-\beta_1)\mu\omega_2 \equiv (\lambda_2-\lambda_1)\mu\omega_1 \pmod p$ 和

$$(\beta_2-\beta_1)\omega_2 \equiv (\lambda_2-\lambda_1)\omega_1 \pmod{\omega_1\omega_2}.$$

由于 ω_1 和 ω_2 是互素的，$1-\omega_1 \le \beta_2-\beta_1 \le \omega_1-1$ 且 $1-\omega_2 \le \lambda_2-\lambda_1 \le \omega_2-1$，那么 $(\beta_2-\beta_1)\omega_2 \equiv (\lambda_2-\lambda_1)\omega_1 \pmod{\omega_1\omega_2}$ 成立当且仅当 $\beta_2-\beta_1=\lambda_2-\lambda_1=0$。这样我们就可以导出矛盾。因此，$|A_{j_1} \bigcap B_{j_2}| \le 1$。

要说明 k-PIR 性质，由于构造的对称性我们只需要去考虑 x_0 的生成情况。在 $\mu(d+1)$ 个单点服务器上总共有 $t\mu(d+1)$ 个单元，其中 $\frac{t\mu(d+1)}{p}$ 个包含单点 x_0。因此总共 $\frac{t\mu(d+1)}{p}=\frac{td_2}{\omega_1}$ 个单点服务器包含 x_0，而其余的 $\frac{d\mu(d+1)}{p}=d_1d_2$ 则不含有 x_0。在 μd 个 Σ 型服务器中，总共有 $(t-1)\mu d$ 个单元，其中 $\frac{(t-1)\mu d}{p}$ 个包含单点 x_0。因此 $\frac{(t-1)\mu d}{p}=\frac{(t-1)d_1}{\omega_2}$ 个 Σ 型服务器包含单点 x_0，剩下的 $\frac{(d+1)\mu d}{p}=d_1d_2$ 个不包含 x_0。任意地从不含 x_0 的单点服务器中取出一个，它所对应的集合 $\overline{A_{j_1}}$，其中 $x_0 \in A_{j_1}$。任意地从不含 x_0 的 Σ 型服务器中选取一个，它对应的集合为 ΣB_{j_2}，其中 $x_0 \in B_{j_2}$。由于我们在之前已经证明了 $|A_{j_1} \bigcap B_{j_2}| \le 1$，因此服务器 $\overline{A_{j_1}}$ 知道除了 x_0 之外的所有服务器 ΣB_{j_2} 中的线性加和项，因此它们可以共同生成 x_0。

最终我们可以得到 $k=\frac{td_2}{\omega_1}+\frac{(t-1)d_1}{\omega_2}+d_1d_2=\frac{d^2+2td+t}{\omega_1\omega_2}$。因此上面构造的阵列码的码率为 $k/m=\frac{d^2+2td+t}{\mu(2d+1)\omega_1\omega_2}=1-\frac{d^2+d}{p(2d+1)}$，可以达到上界。 □

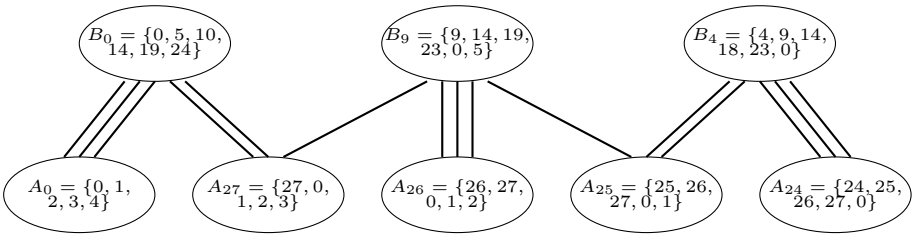
注10. 事实上，我们可以在不包含 x_0 的单点服务器和 Σ 型服务器之间构造一张二部图，其中的边集由两个点可以生成 x_0 所构成。在上面的证明过程中，我们证明了当 $t \ge d^2$ 时，这是一张完全二部图。而要说明方案确实是合理的，我们只需要在图上找到一个完美匹配就可以了。基于这一点我们可以将定理 *5.6* 适当推广。

5.2.2 $d^2-d<t<d^2$

在复杂的分析之前，我们先通过一个例子来说明证明的基本思路。

例11 ($d=5,t=23,p=28$)**.** 可以计算出对应的参数为 $\omega=2$，$\mu=14$，$\omega_1=1$，$d_1=5$，$\omega_2=2$ 和 $d_2=3$。我们有 84 个单点服务器：$\overline{A_j}$，$0 \le j \le 27$，每个出现三次，其中 $A_j=\{x_{j+\alpha}:0\le\alpha\le 4\}$。另外我们有 70 个 Σ 型服务器：ΣB_j，$0 \le j \le 13$，每个均出现五次，其中 $B_j=\{x_{j+5\gamma+14\lambda:0\le\gamma\le 2,0\le\lambda\le 1}\}$。不包含 x_0 的服务器是：ΣB_0，ΣB_9 和 ΣB_4，每个出现五次；$\overline{A_0}$，$\overline{A_{27}}$，$\overline{A_{26}}$，$\overline{A_{25}}$ 和 $\overline{A_{24}}$，每个出现三次。注意到由于 $B_0 \bigcap A_{24}=\{0,24\}$，

ΣB_0 与 $\overline{A_{24}}$ 是不相连的。同理 ΣB_4 与 $\overline{A_0}$ 也是不相连的。但是我们还是可以在对应的图中找到如下的完美匹配。



引理 5.7. 对 $d^2 - d < t < d^2$，如果 $|A_{j_1} \bigcap B_{j_2}| > 1$ 或者 $0 \in A_{j_1} \bigcap B_{j_2}$，那么 $j_2 = 0$ 或者 $j_2 = \mu\omega_1 - d_1(d_2 - 1)$。

证明. 存在 $\alpha_1, \alpha_2, \beta_1, \beta_2, \gamma_1, \gamma_2, \lambda_1, \lambda_2$ 使得

$$j_1 + \alpha_1 + \beta_1\mu\omega_2 \equiv j_2 + \gamma_1 d_1 + \lambda_1\mu\omega_1 \equiv 0 \pmod{p},$$

$$j_1 + \alpha_2 + \beta_2\mu\omega_2 \equiv j_2 + \gamma_2 d_1 + \lambda_2\mu\omega_1 \pmod{p}.$$

通过上面的两式相减，我们可以得到

$$(\alpha_2 - \alpha_1) + (\beta_2 - \beta_1)\mu\omega_2 \equiv (\gamma_2 - \gamma_1)d_1 + (\lambda_2 - \lambda_1)\mu\omega_1 \pmod{p}.$$

那么我们有 $(\alpha_2 - \alpha_1) \equiv (\gamma_2 - \gamma_1)d_1 \pmod{\mu}$。当 $t > d^2 - d$ 时，$\mu = \frac{p}{\omega_1\omega_2} > \frac{d^2}{\omega_1\omega_2} \geq d_1(d_2 - 1)$。由于 $1 - d_1 \leq \alpha_2 - \alpha_1 \leq d_1 - 1$ 和 $1 - d_2 \leq \gamma_2 - \gamma_1 \leq d_2 - 1$，因此 $(\alpha_2 - \alpha_1) \equiv (\gamma_2 - \gamma_1)d_1 \pmod{\mu}$ 成立当且仅当下面的实事成立：

• Case I. $\alpha_2 - \alpha_1 = \gamma_2 - \gamma_1 = 0$。那么我们可以通过和定理 5.6 中相同的分析，从而导出矛盾。

• Case II. $\gamma_2 - \gamma_1 = d_2 - 1$，因此 $\gamma_2 = d_2 - 1$，$\gamma_1 = 0$。那么我们有 $j_2 + \lambda_1\mu\omega_1 \equiv 0 \pmod{p}$。由于 $0 \leq j_2 \leq \mu\omega_1 - 1$ 和 $0 \leq \lambda_1 \leq \omega_2 - 1$，我们可以导出一定有 $j_2 = 0$。

• Case III. $\gamma_2 - \gamma_1 = 1 - d_2$，因此 $\gamma_1 = d_2 - 1$，$\gamma_2 = 0$。那么我们有 $j_2 + d_1(d_2 - 1) + \lambda_1\mu\omega_1 \equiv 0 \pmod{p}$。由于 $0 \leq j_2 \leq \mu\omega_1 - 1$，$0 \leq \lambda_1 \leq \omega_2 - 1$ 和 $d_1(d_2 - 1) < \mu \leq \mu\omega_1$，我们可以导出一定有 $j_2 = \mu\omega_1 - d_1(d_2 - 1)$。 □

因此我们只需要考察两个特出的 Σ 型服务器：ΣB_0，$\Sigma B_{\mu\omega_1 - d_1(d_2-1)}$。

引理 5.8. $A_{j_1} \bigcap B_0 = \{x_0\}$ 当且仅当 $j_1 = 0$ 或者 $\mu\omega_2 - \mu + d_1(d_2 - 1) + 1 \leq j_1 \leq \mu\omega_2 - 1$。

证明. $j_1+\alpha_1+\beta_1\mu\omega_2\equiv 0 \pmod p$ 成立当且仅当

$$j_1=\alpha_1=\beta_1=0$$

或者

$$\beta_1=\omega_1-1，\ j_1=\mu\omega_2-\alpha_1.$$

因此满足 $x_0\in A_{j_1}$ 的 j_1 的可能集合为 $j_1\in\{0\}\bigcup[\mu\omega_2-d_1+1,\mu\omega_2-1]$。

我们还需要排除那些使得 $|A_{j_1}\bigcap B_0|>1$ 的 j_1。继续引理 5.7 中 Case II 的讨论，$\gamma_2-\gamma_1=d_2-1$，那么 $\alpha_2-\alpha_1=d_1(d_2-1)-\mu$。因此 $\alpha_1\in[0,d_1-1]\bigcap[\mu-d_1(d_2-1),\mu-d_1(d_2-1)+d_1-1]=[\mu-d_1(d_2-1),d_1-1]$。我们有满足 $|A_{j_1}\bigcap B_0|>1$ 的 j_1 的可能集合为 $j_1\in[\mu\omega_2-d_1+1,\mu\omega_2-\mu+d_1(d_2-1)]$。因此，将这些 j_1 都排除在之外，我们可以最终得到 $A_{j_1}\bigcap B_0=\{x_0\}$ 当且仅当 $j_1=0$ 或者 $\mu\omega_2-\mu+d_1(d_2-1)+1\le j_1\le\mu\omega_2-1$。 □

引理 5.9. $A_{j_1}\bigcap B_{\mu\omega_1-d_1(d_2-1)}=\{x_0\}$ 当且仅当 $\mu\omega_2-d_1+1\le j_1\le\mu\omega_2-d_1d_2+\mu$。

证明. $j_1+\alpha_1+\beta_1\mu\omega_2\equiv 0 \pmod p$ 成立当且仅当

$$j_1=\alpha_1=\beta_1=0$$

或者

$$\beta_1=\omega_1-1 \text{ 且 } j_1=\mu\omega_2-\alpha_1.$$

因此满足 $x_0\in A_{j_1}$ 的j_1的可能集合为 $j_1\in\{0\}\bigcup[\mu\omega_2-d_1+1,\mu\omega_2-1]$。

我们还需要排除那些使得 $|A_{j_1}\bigcap B_{\mu\omega_1-d_1(d_2-1)}|>1$ 的 j_1。继续引理 5.7 中 Case III 的讨论，$\gamma_2-\gamma_1=1-d_2$，那么 $\alpha_2-\alpha_1=\mu-d_1(d_2-1)$。因此 $\alpha_1\in[0,d_1-1]\bigcap[d_1(d_2-1)-\mu,d_1(d_2-1)-\mu+d_1-1]=[0,d_1(d_2-1)-\mu+d_1-1]$。我们有满足 $|A_{j_1}\bigcap B_0|>1$ 的 j_1 的可能集合为 $j_1\in[\mu\omega_2-d_1(d_2-1)+\mu-d_1+1,\mu\omega_2-1]\bigcup\{0\}$。因此，将这些 j_1 都排除在之外，我们可以最终得到 $A_{j_1}\bigcap B_{\mu\omega_1-d_1(d_2-1)}=\{x_0\}$ 当且仅当 $\mu\omega_2-d_1+1\le j_1\le\mu\omega_2-d_1d_2+\mu$。

□

引理 5.10. $d_2(\mu-d_1(d_2-1))\ge d_1$。

证明. 这个不等式等价于 $\frac{d+1}{\omega_2}(\frac{p}{\omega_1\omega_2}-\frac{d}{\omega_1}(\frac{d+1}{\omega_2}-1))\ge\frac{d}{\omega_1}$。将它整理和化简，我们可以得到 $d\omega_2(d+1-\omega_2)\ge(d+1)(d^2+d-p)$。由于 $\omega_2|d+1$ 和 $d^2<p<d^2+d$，因此 p 不是 $d+1$ 的倍数，$\omega_2\ne d+1$。我们有 $1\le\omega_2\le\frac{d+1}{2}$，左式至少是 d^2。由于 $d^2<p$，我们有 $d^2+d-p\le d-1$，右式至多为 d^2-1。因此不等式得证。 □

结合上面的引理，我们最终可以得到如下结论。

定理 5.11. *对 $d^2-d<t<d^2$，存在码率 $\frac{k}{m}$ 达到 $g(s,t)=1-\frac{d^2+d}{p(2d+1)}$ 服务器数目为 $m=\mu(2d+1)$ 的最优 k-PIR 阵列码。*

证明. 我们只需证明由那些不包含 x_0 的服务器所诱导的二部图中存在一个完美匹配。对 d_1 个名为 ΣB_0 的服务器，由引理 5.8 我们知道这样的服务器与满足 $j\in\mathcal{S}\triangleq[\mu\omega_2-\mu+d_1(d_2-1)+1,\mu\omega_2-1]\bigcup\{0\}$ 的名为 $\overline{A_j}$ 的服务器相连。由于每个 A_j 出现 d_2 次，因此总共有 $d_2(\mu-d_1(d_2-1))$ 个这样的服务器。

相似的，对 d_1 个名为 $\Sigma B_{\mu\omega_1-d_1(d_2-1)}$ 的服务器，由引理 5.9 我们知道这样的服务器只与满足 $j\in\mathcal{T}\triangleq[\mu\omega_2-d_1+1,\mu\omega_2-d_1d_2+\mu]$ 的名为 $\overline{A_j}$ 的服务器相连。由于每个 A_j 出现 d_2 次，因此总共有 $d_2(\mu-d_1(d_2-1))$ 个这样的服务器。

对任意除了 ΣB_0，$\Sigma B_{\mu\omega_1-d_1(d_2-1)}$ 的其他 Σ 型服务器，引理 5.7 告诉我们它们与所有不包含 x_0 的单点服务器所相连。因此，为了找到一个完美匹配，我们只需合理地找到 ΣB_0 和 $\Sigma B_{\mu\omega_1-d_1(d_2-1)}$ 的边，而剩下的就可以随意选取。由引理 5.10，$d_2|\mathcal{S}|=d_2|\mathcal{T}|=d_2(\mu-d_1(d_2-1))\geq d_1$。此外 $d_2|\mathcal{S}\bigcup\mathcal{T}|=d_1d_2\geq 2d_1$，其中 $d_2\geq 2$ 由引理 5.10 中的 $\omega_2\neq d+1$ 导出。因此我们可以找到所有 ΣB_0 和 $\Sigma B_{\mu\omega_1-d_1(d_2-1)}$ 的边，结论得证。 □

5.3 $s>2$: $g(s,t)$ 的上下界研究

5.3.1 $g(s,t)$ 的新上界

在本节中，我们将尝试着去诱导出 $g(s,t)$ 在 $s>2$ 的情形下（等价于原文[6] 的 $d>t$ 和 $p=d+t>2t$）的全新上界，试图去改进定理 5.4 中的原有结果。

对任意给定的 PIR 阵列码，我们首先将所有的服务器分成四类。第一类包含所有的单点服务器，它的数目记为 l。第二类服务器包含 $t-1$ 个单点元素，剩下的一个位置是余下 $p-t+1$ 个元素中 η 个之和，其中 $2\leq\eta\leq t+1$，它的数目记为 r。第三类服务器包含 $t-1$ 个单点元素，剩下的一个位置是余下 $p-t+1$ 个元素中 λ 个之和，其中 $t+1<\lambda\leq p-t+1$，它的数目记为 u。我们将剩余的服务器都归到第四类，它们的数目记为 w。那么我们有 $l+r+u+w=m$。

定理 5.12. 对任意正整数 $t \geq 2$，$d > t$，我们有

$$g(1+\frac{d}{t},t) \leq \frac{d^2+2t^2+3td+2t}{2(t+d)(d+t+1)}.$$

证明. 假设我们有一个 $[t \times m, p]$ k-PIR 阵列码。对任意 $i \in \{1,2,\ldots,p\}$，设 $S_1^i, S_2^i, \ldots, S_{k_i}^i$ 是那些可以张成元素 x_i 的互不相交的集合，其中 k_i 尽可能取得越大越好。为了得到 k/m 的上界，我们只需证明：

$$\sum_{i=1}^{p} k_i \leq \frac{d^2+2t^2+3td+2t}{2(d+t+1)} m.$$

在上面的四类服务器中，不失一般性，任意包含 x_i 的服务器显然是可以拿来作为 $S_1^i, S_2^i, \ldots, S_{k_i}^i$ 中的元素的。设在四类服务器中这样的服务器的数量分别是：l_i，r_i，u_i 和 w_i。设 f_i 是那些在 $S_1^i, S_2^i, \ldots, S_{k_i}^i$ 中由一个单点服务器和一个非单点服务器所构成集合的数目。设 g_i 是那些 $S_1^i, S_2^i, \ldots, S_{k_i}^i$ 中由至少两个单点服务器和一个非单点服务器所构成集合的数目。而在 $S_1^i, S_2^i, \ldots, S_{k_i}^i$ 中剩余的集合都至少包含两个非单点服务器。因此我们可以得到如下的不等式：

$$k_i \leq l_i + r_i + u_i + w_i + f_i + g_i + \frac{r - r_i + u - u_i + w - w_i - f_i - g_i}{2}.$$

下面我们将用两种方式去估计 $\sum k_i$。首先，我们从计数单点服务器出发 $f_i + 2g_i \leq l - l_i$。因此我们有

$$\begin{aligned}
k_i &\leq l_i + r_i + u_i + w_i + f_i + g_i + \frac{r - r_i + u - u_i + w - w_i - f_i - g_i}{2} \\
&= l_i + \frac{r + r_i + u + u_i + w + w_i}{2} + \frac{f_i}{2} + \frac{g_i}{2} \\
&\leq l_i + \frac{r + r_i + u + u_i + w + w_i}{2} + \frac{f_i}{2} + g_i \\
&\leq l_i + \frac{r + r_i + u + u_i + w + w_i}{2} + \frac{l - l_i}{2} \\
&= \frac{l + l_i + r + r_i + u + u_i + w + w_i}{2}.
\end{aligned}$$

通过计数四类服务器中的单点集数目，我们有$\sum l_i = lt$，$\sum r_i = r(t-1)$，$\sum u_i = u(t-1)$ 和 $\sum w_i \leq w(t-2)$。这可以导出

$$\begin{aligned}
\sum k_i &\leq \frac{p(l+r+u+w)}{2} + \frac{lt + r(t-1) + u(t-1) + w(t-2)}{2} \\
&= l\frac{p+t}{2} + r\frac{p+t-1}{2} + u\frac{p+t-1}{2} + w\frac{p+t-2}{2}.
\end{aligned} \tag{5-1}$$

关于第二种估计我们去分析 f_i。注意到非单点服务器不可能来自第三类，这是由于那些不包含 x_i 的第三类服务器，它的非单点元素的形式只能是 $x_i + \sum_{j=1}^{\lambda-1} y_j$，其中 $\lambda - 1 > t$。

任意不包含单点 x_i 的单点服务器，至多提供 t 个其他单点。因此它们之间是不可能通过合作生成 x_i 的。最后，我们有 $f_i \le r - r_i + w - w_i$ 和 $\sum f_i \le pr - r(t-1) + pw - \sum w_i$。

然而上面的分析仍然无法得到我们想要的结论，我们将通过下面的重要观察来最终得到证明。对任意第二类的非单点服务器，它的非单点元素至多由 $t+1$ 元素的和所构成。因此它对 $\sum f_i$ 的贡献至多是 $t+1$。我们得到比 $\sum f_i \le pr - r(t-1) + pw - \sum w_i$ 更好的估计 $\sum f_i \le r(t+1) + pw - \sum w_i$。我们有

$$\begin{aligned} k_i &\le l_i + r_i + u_i + w_i + f_i + g_i + \frac{r - r_i + u - u_i + w - w_i - f_i - g_i}{2} \\ &= l_i + \frac{r + r_i + u + u_i + w + w_i}{2} + \frac{f_i}{4} + \frac{g_i}{2} + \frac{f_i}{4} \\ &\le l_i + \frac{r + r_i + u + u_i + w + w_i}{2} + \frac{l - l_i}{4} + \frac{f_i}{4} \\ &= \frac{l + 3l_i + 2r + 2r_i + 2u + 2u_i + 2w + 2w_i}{4} + \frac{f_i}{4}, \end{aligned}$$

$$\begin{aligned} \sum k_i &\le \sum \frac{l + 3l_i + 2r + 2r_i + 2u + 2u_i + 2w + 2w_i}{4} + \frac{\sum f_i}{4} \\ &= \frac{lp + 3lt + 2rp + 2r(t-1) + 2up + 2u(t-1) + 2wp + 2\sum w_i}{4} + \frac{r(t+1) + wp - \sum w_i}{4} \\ &= \frac{lp + 3lt + 2rp + 2r(t-1) + 2up + 2u(t-1) + 2wp + r(t+1) + wp}{4} + \frac{\sum w_i}{4} \\ &\le l\frac{p+3t}{4} + r\frac{2p+3t-1}{4} + u\frac{p+t-1}{2} + w\frac{3p+t-2}{4}. \end{aligned} \tag{5-2}$$

现在我们已经得到了 $\sum k_i$ 的两种估计 (1) 和 (2)。记作

$$F(l,r,u,w) = l\frac{p+t}{2} + r\frac{p+t-1}{2} + u\frac{p+t-1}{2} + w\frac{p+t-2}{2},$$

$$G(l,r,u,w) = l\frac{p+3t}{4} + r\frac{2p+3t-1}{4} + u\frac{p+t-1}{2} + w\frac{3p+t-2}{4},$$

那么 $\sum k_i \le \min\{F,G\}$。为了估计出 $\sum k_i$ 的上界，我们需要去确定 $\min\{F,G\}$ 的最大值。不妨设最大值在 $(\widetilde{l},\widetilde{r},\widetilde{u},\widetilde{w})$ 处取到。容易验证得到 $F(\widetilde{l},\widetilde{r},\widetilde{u},\widetilde{w}) \le F(\widetilde{l},\widetilde{r}+\widetilde{u},0,\widetilde{w})$ 和 $G(\widetilde{l},\widetilde{r},\widetilde{u},\widetilde{w}) \le G(\widetilde{l},\widetilde{r}+\widetilde{u},0,\widetilde{w})$。因此我们有 $\widetilde{u}=0$。这样问题就可以约简为

$$\begin{aligned} \max \quad & \min\left\{ l\frac{p+t}{2} + r\frac{p+t-1}{2} + w\frac{p+t-2}{2},\ l\frac{p+3t}{4} + r\frac{2p+3t-1}{4} + w\frac{3p+t-2}{4} \right\}, \\ s.t. \quad & l + r + w = m,\ l,r,w \in \mathbb{N} \end{aligned}$$

为了解决上面的问题，我们首先将 w 固定。那么当 $l = m\frac{t+1}{p+1} + w\frac{p-2t+1}{p+1}$ 时，有 $m\frac{p^2+tp+2t}{2(p+1)} - w\frac{t}{(p+1)}$。而为了得到上面取值的最大值，我们只能让 $w = 0$。总而言之，当 $w = 0$，$l = m\frac{t+1}{p+1}$ 和 $r = m\frac{p-t}{p+1}$ 时取到最优解 $\frac{d^2+2t^2+3td+2t}{2(d+t+1)}m$。 □

最后，我们非常容易验证当 $s > 2$ 时（即 $d > t$时），$\frac{d^2+2t^2+3td+2t}{2(t+d)(d+t+1)} < \frac{(2d+1)t+d^2}{(t+d)(2d+1)}$ 成立，因此我们得到的上界要优于定理 5.4。

5.3.2 PIR 阵列码的一般构造

在接下来的部分，我们将给出在 $s > 2$ 时，PIR 阵列码的一个全新构造。在这之前我们先简单地去回顾一下 Blackburn 和 Etzion 在[6] Section 4 中的构造，我们在后面的行文中将记为 B-E 构造。这里，我们只介绍 s 是整数的情形，事实上 s 不是整数的情况也是可以相似的去给出的。

对一个服务器包含 $t-1$ 个单点元素和另外一个在余下的 $st-t+1$ 中的 j 这元素之和，我们将称这个服务器是 j 型的，对 $1 \le j \le st-t+1$。那么 1 型的服务器就是我们上文中一直提及的单点服务器。对 $1 \le r \le s$，设 T_r 是由所有型 $(r-1)t+1$ 服务器所构成的集合。也就是说，$|T_1| = \binom{st}{t}$，$|T_r| = \binom{st}{t-1}\binom{st-t+1}{(r-1)t+1}$ 对 $2 \le r \le s$。B-E 构造包含所有服务器集合 T_r，其中 $1 \le r \le s$，并且每个 T_r 都出现 η_r 次。对任意给定的元素 x_i，我们可以将那些不含有单点 x_i 的服务器放到 $s-1$ 个二部图中。η_r 的选择是用来保证对任意给定的元素 x_i，$s-1$ 个二部图可以找到一个完美匹配，等价于说，所有不包含单点 x_i 的服务器都可以被两两配对，从而去生成 x_i。

对任意给定的 x_i，$s-1$ 个二部图如下所示。二部图 G_r，$1 \le r \le s-1$，由两边所构成。其中第一部分由所有 T_r 中的服务器所构成，每个都恰好出现 η_r，其中 x_i 在所有的点中都没有出现。第二部分由所有 T_{r+1} 中的服务器所构成，每个恰好出现 η_{r+1}，其中 x_i 都出现在加和中。在第一部分的顶点 v 和第二部分的顶点 u 之间有边相连，当且仅当在 v 中出现的 $t-1$ 个单点集和 $(r-1)t+1$ 个元素之和都恰好是 u 中 rt 个加和项，除了给定的 x_i。可以非常容易地计算得到，为了要保证 G_r 中有一个完美匹配，参数必须去满足：$\eta_1 : \eta_2 = \binom{p-t-1}{t-1} : 1$ 和 $\eta_r : \eta_{r+1} = \binom{p-rt-1}{t-1} : \binom{rt}{t-1}$ 对 $2 \le r \le s-1$。

例12 ($s = 4$)**.** 服务器 T_1, T_2, T_3, T_4 分别出现 $\eta_1, \eta_2, \eta_3, \eta_4$ 次。下面的比例是必须的。$\eta_1 : \eta_2 = \binom{3t-1}{t-1} : 1$，$\eta_2 : \eta_3 = (t+1) : 2t$ 和 $\eta_3 : \eta_4 = 1 : \binom{3t}{t-1}$。所以我们选择 $\eta_1 = (t+1)\binom{3t-1}{t-1}$，$\eta_2 = t+1$, $\eta_3 = 2t$ 和 $\eta_4 = 2t\binom{3t}{t-1}$。

相比定理 5.5 中的其他 PIR 阵列码的构造，B-E 构造是一个一般性的构造，它可以去适应一切的参数。

下面我们将在 B-E 构造的基础上，进行适当地改造，得到一个新的 PIR 阵列码的构造。我们将说明这个构造在码率上比定理 5.5 中的构造都要好。

构造(给定 $t, s > 2$ 和 $p = ts$):

1. 选取所有的那些 $\binom{p}{t}$ 单点服务器，每个都出现 δ 次，其中 $\delta = \binom{p-t-1}{t-1}$。

2. 对那些只有 $t-1$ 个单点的服务器，其他的一个元素由剩余的 $p-t+1$ 中的 j 个元素之和，我们称这些为型 j 的服务器，$2 \le j \le p-t+1$。取出所有型 $t+1, t+2, \ldots, p-t+1$ 的服务器，每种都只出现一次。

我们容易观察得到这个构造关于所有元素 $\{x_1, x_2, \ldots, x_p\}$ 都是对称的。为了去张成某个元素 x_i，所有包含 x_i 的服务器是毫无疑问的，我们期望那些不包含 x_i 的服务器可以两两进行配对来生成 x_i。下面的定理就用来说明上面的想法是可以被实现的。

定理 5.13. *上面的构造给出的 PIR 阵列码的参数是：* $m = \binom{p}{t}\binom{p-t-1}{t-1} + \binom{p}{t-1}\sum_{t+1\le j\le p-t+1}\binom{p-t+1}{j}$，$k = \frac{p+t}{2p}\binom{p}{t}\binom{p-t-1}{t-1} + \frac{p+t-1}{2p}\binom{p}{t-1}\sum_{t+1\le j\le p-t+1}\binom{p-t+1}{j}$。

证明. 总共有 $\binom{p}{t} \times \delta$ 个单点服务器和 $\binom{p}{t-1}\binom{p-t+1}{j}$ 个型 j 的服务器，加起来一共

$$m = \binom{p}{t}\binom{p-t-1}{t-1} + \binom{p}{t-1}\sum_{t+1\le j\le p-t+1}\binom{p-t+1}{j}.$$

在所有的单点服务器中，包含 x_i 的单点服务器恰好有 $\frac{t}{p}\binom{p}{t}\binom{p-t-1}{t-1}$ 个。在所有的非单点服务器中，包含 x_i 的恰好有 $\frac{t-1}{p}\binom{p}{t-1}\sum_{t+1\le j\le p-t+1}\binom{p-t+1}{j}$ 个。

对那些不包含 x_i 的单点服务器，我们不妨假设它包含 $\{y_1, y_2, \ldots, y_t\}$。我们将它与一个型 $t+1$ 的服务器相配对，它的非单点元素恰为 $x_i + \sum_{j=1}^{t} y_j$。这两个服务器恰好可以去张成 x_i。存储 $\{y_1, y_2, \ldots, y_t\}$ 的服务器有 $\delta = \binom{p-t-1}{t-1}$，同时型 $t+1$ 的包含 $x_i + \sum_{j=1}^{t} y_j$ 的服务器也有 $\binom{p-t-1}{t-1}$ 个。因此我们可以将这两部分的服务器进行一一配对。

下面我们讨论满足如下条件的型 j 的服务器：1) 它不包含单点集 x_i；2) x_i 也没有出现在它的加和项中，其中 $t+1 \le j < p-t+1$。我们不妨假设这种服务器包含 $z_1, z_2, \ldots, z_{t-1}$ 和 $\omega_1 + \omega_2 + \cdots + w_j$。那么我们可以将他和那些包含 $z_1, z_2, \ldots, z_{t-1}$ 和 $x_i + \omega_1 + \omega_2 + \cdots + w_j$ 的型 $j+1$ 的服务器相配对。显然它们两者可以联合去导出 x_i。

应用上面的方式，所有不含有 x_i 的服务器都可以被两两配对起来，因此我们有

$$\begin{aligned} k &= \frac{t}{p}\binom{p}{t}\binom{p-t-1}{t-1} + \frac{t-1}{p}\binom{p}{t-1}\sum_{t+1\le j\le p-t+1}\binom{p-t+1}{j} \\ &\quad + \frac{m - \frac{t}{p}\binom{p}{t}\binom{p-t-1}{t-1} - \frac{t-1}{p}\binom{p}{t-1}\sum_{t+1\le j\le p-t+1}\binom{p-t+1}{j}}{2} \\ &= \frac{p+t}{2p}\binom{p}{t}\binom{p-t-1}{t-1} + \frac{p+t-1}{2p}\binom{p}{t-1}\sum_{t+1\le j\le p-t+1}\binom{p-t+1}{j}. \end{aligned}$$

□

例13 ($s = 4$)**.** 我们有所有的 $\binom{4t}{t}$ 单点服务器，每个都出现 $\binom{3t-1}{t-1}$ 次，并且型 $t+1, t+2, \ldots, 3t+1$ 的服务器，总共 $\binom{4t}{t-1}\sum_{t+1\le j\le 3t+1}\binom{3t+1}{j}$ 个。容易地去计算看出所需的服务器要比例 *12* 中的要少得多。

要精确的计算出我们所构造的阵列码的码率非常地繁琐，但是我们注意到我们的构造中事实上只有两类情况发生：包含 x_i 的服务器；不包含 x_i 的服务器均可以两两配对。因此，如果我们有 α 个单点服务器和 β 个其他服务器。所有的单点服务器总共包含有 $t\alpha$ 个元素。有构造的对称性可以看出，x_i 总共出现了 $t\alpha/p$ 次。同理，在所有其他服务器中 x_i 总共出现了 $(t-1)\beta/p$。因此，我们有 $m = \alpha+\beta$ 和 $k = t\alpha/p + (t-1)\beta/p + \frac{m - t\alpha/p - (t-1)\beta/p}{2}$。最终的码率为

$$k/m = \frac{t+p}{2p}\frac{\alpha}{\alpha+\beta} + \frac{t+p-1}{2p}\frac{\beta}{\alpha+\beta},$$

是 $\frac{t+p}{2p}$ 和 $\frac{t+p-1}{2p}$ 的附权平均，严格大于 $\frac{t+p-1}{2p} = \frac{ts+t-1}{2ts}$。

从上面的观察可以得到比例 $\frac{\alpha}{\alpha+\beta}$ 越大，那么码率也越大。下面我们将说明我们的构造比定理 5.5 中的构造在码率上都要更大。

由于我们所构造的阵列码的码率严格大于 $\frac{t+p-1}{2p} = \frac{ts+t-1}{2ts}$，我们首先用这个值 $\frac{ts+t-1}{2ts}$ 去和现有的构造比较。尽管在这种情况下，我们可以完成大部分的说明，但还有一些零碎的情形需要我们更仔细的去比较。

- 与定理 5.5 的比较，第一个实例（构造 7 和定理 10[6]）：

设 $3 \le r \le t$，$s = r - \frac{(r-2)+1}{t}$。在这个实例中 $p = ts = tr - t^2 + 2r - 1$。容易验证当 $r \ge 3$ 时，$\frac{ts+t-1}{2ts} > \frac{1}{2} + \frac{t-r+1}{2(rt-(r-2)r-1)}$ 成立。

• 与定理 5.5 的比较，第二个实例（构造 8 和定理 11[6]）：

设 $s=r+\frac{d}{t}$，$p=rt+d$，其中 $r\geq 2$ 是一个整数，$t\geq r$，$1\leq d\leq t-1$。我们要去说明 $\frac{ts+t-1}{2ts}=\frac{rt+d+t-1}{2(rt+d)}$ 严格大于 $1-\frac{(rt+d-t+r)(rt+d-t)}{(rt+d)(2rt+2d-2t+r)}$，最终可以被约简为 $(r-2)(rt+d-t)>r$。这显然在 $r\geq 3$ 是始终成立的。

剩余的情况是当 $r=2$,$p=2t+d$ 时，我们进行仔细地比较。在我们的构造中单点服务器的数量为 $A=\binom{2t+d}{t}\binom{t+d-1}{t-1}$，其他服务器的数量为 $B=\binom{2t+d}{t-1}\sum_{i=0}^{d}\binom{t+d+1}{i}$，因此最终的码率是：

$$k/m=\frac{A\frac{3t+d}{4t+2d}+B\frac{3t+d-1}{4t+2d}}{A+B}=\frac{3t+d}{4t+2d}-\frac{1}{(4t+2d)(\frac{A}{B}+1)}.$$

关于 $\frac{A}{B}$ 比例的下界可以按如下方式导出，其中 $\sum_{i=0}^{d}\binom{t+d+1}{i}\leq\frac{d(t+d+1)!}{d!(t+1)!}$ 可以导出第三个不等式，$d\leq t-1$ 可以导出第四个不等式：

$$\frac{A}{B}=\frac{\binom{2t+d}{t}\binom{t+d-1}{t-1}}{\binom{2t+d}{t-1}\sum_{i=0}^{d}\binom{t+d+1}{i}}=\frac{(t+d+1)(t+d-1)!}{t(t-1)!d!\sum_{i=0}^{d}\binom{t+d+1}{i}}\geq\frac{t+1}{d(t+d)}>\frac{1}{t+d}.$$

因此码率的估计为 $\frac{k}{m}\geq\frac{3t+d}{4t+2d}-\frac{1}{(4t+2d)(\frac{1}{t+d}+1)}$，可以容易验证这比 $1-\frac{(t+d+2)(t+d)}{(2t+d)(2t+2d+2)}$ 要大。

• 与定理 5.5 的比较，第二个实例（构造 9 和定理 12[6]）：

设 $s>2$ 是一个整数，$t\geq s$。不等式 $\frac{ts+t-1}{2ts}>\frac{ts+t+1}{s(2t+1)}$ 等价于 $ts>3t+1$，这在 $s\geq 4$ 时是成立的。剩余 $s=3$ 的情形我们进行如下的分析。

当 $s=3$ 时，那么 $p=3t$。在我们的构造中 $m=\binom{3t}{t}\binom{2t-1}{t-1}+\binom{3t}{t-1}2^{2t}$，$k=\frac{2}{3}\binom{3t}{t}\binom{2t-1}{t-1}+\frac{4t-1}{6t}\binom{3t}{t-1}2^{2t}$。下面我们来证明码率 k/m 是严格大于 $\frac{4t+1}{6t+3}$：

$$\begin{aligned}&\frac{\frac{2}{3}\binom{3t}{t}\binom{2t-1}{t-1}+\frac{4t-1}{6t}\binom{3t}{t-1}2^{2t}}{\binom{3t}{t}\binom{2t-1}{t-1}+\binom{3t}{t-1}2^{2t}}>\frac{4t+1}{6t+3}\\ \Longleftrightarrow\ &\frac{2t+1}{t}\binom{2t-1}{t-1}\cdot\frac{1}{6t+3}>2^{2t}(\frac{4t+1}{6t+3}-\frac{4t_1}{6t})\\ \Longleftrightarrow\ &(4t+2)\binom{2t-1}{t-1}>2^{2t}\end{aligned}$$

其中最后一个不等式可以通过归纳给出：当 $t=1$ 时，可以导出 $6>4$；归纳假设的步骤如下 $\frac{(4t+6)\binom{2t+1}{t}}{(4t+2)\binom{2t-1}{t-1}}=\frac{4t+6}{t+1}>4$。

- 与定理 5.5 的比较，第二个实例（构造 10 和定理 13[6]）：

设 $s > 2$ 是一个整数。设 $(s-1)t = lb$ 和 $t \geq l + b$，其中 l 和 b 是正整数。显然我们要有 l 大于 1。因此 $\frac{ts+t-1}{2ts} = \frac{s+1}{2s} - \frac{1}{2st} > \frac{s+1}{2s} - \frac{l}{2st}$ 成立。

综上所述，我们证明了我们得到的 PIR 阵列码的码率是要优于定理 5.5 中的构造的。

5.4　小结

在本章中，我们考虑了最优 PIR 阵列码的构造问题。在情形 $1 < s \leq 2$ 下，我们完全确定了当 $t > d^2 - d$ 时，最优 PIR 阵列码所需的最小服务器数量。我们相信在其他情形下应该也会有与本章相似的结论。在情形 $s > 2$ 下，我们得到的上界事实上说明定理 5.4 的结果肯定不是紧的。一个自然的问题是我们得到的上界是不是紧的，从证明的过程来分析，我们相信这个界也不是最优的，如何去得到这个问题的最优答案是一个十分具有挑战的研究课题。

6 其它研究工作

本章将简要介绍本人攻读博士期间的其它研究成果，限于篇幅原因，我只对这些课题进行简要介绍。

6.1 可逆 2×2 子矩阵比例问题

AONT 变换由 RSA 密码体制的创始人 Rivest 提出[65]，用来作为分组密码的预处理过程之一。它可以转化为子矩阵的可逆性问题。然而，在二元域上完全符合 AONT 变换要求的矩阵是不存在的。Stinson 等人[22] 提出了一个相关的问题：一个可逆的二元矩阵中有多大比例的可逆 2×2 子矩阵? 这可以看做是一个有信息背景的纯组合问题，Stinson 等人给出了它的上下界分别为 0.492 和 0.625，并通过组合设计和分圆等方法得到了一些构造，猜测这个最大比例值的极限是存在的。我们完整解决了此问题，证明此最大的比例值为 0.5。通过理论的分析，我们得到要达到理论的上界，矩阵中的非零数目的比例是一个无理数。而采用一般的代数构造显然是很难完成这个目标的，因此我们转而通过矩阵的随机构造，利用二阶矩方法，证明了下界确实是 0.5。这部分的工作已经发表在《Discrete Mathematics》。

6.2 置换码

设 S_n 为 n 个元素上的置换群。置换码是置换群 S_n 中的在某种度量意义下的一个子集。依据不同的应用背景，可以定义多种不同的度量。我们主要考虑的是 Hamming 和 Kendall's τ 这两种距离。前者与电力线技术的应用密切相关，后者则是闪存中的排序调制体系[43] 中的必需要求。

与经典码的研究类似，我们也希望在给定置换群 S_n 和给定极小距离 d 的情况下，计算最大的码字数目并得到相应的构造。在 Hamming 距离下，甚至连 Gilbert-Varshamov 型的平凡下界都没法在一般意义下进行改进。我们将研究码字数目的问题转化到研究图的独立集的问题，给出了对应图的一个恰当的染色，利用了图独立数与染色数之间的制约关

系，得到了在 d 给定，而 n 趋于无穷大意义下新的下界，比 Gilbert-Varshamov 型的平凡下界提高了 n 倍。在 Kendall's τ-距离下，我们也用类似的染色方法，得到了新的下界，进一步缩小了上下界之间的距离。这部分的工作已经被《Designs, Codes and Cryptography》所录用。

6.3 防诬陷码

在当今法制社会下，数字产品的版权和知识产权的保护越来越得到重视。数字产品的生产者在其产品中加上一些额外的信息作为数字指纹，这就可以在发现盗版时，通过数字指纹而追踪到源头。然而，在多个盗版者进行合作后，他们可以对数字产品进行一些组合，从而使数字指纹也产生变形，大大加大了追踪的难度。为了解决由这种合作而带来的追踪困难，近二十年来一大批密码学家和组合学家设计出一系列合谋安全码，如防诬陷码、追踪码、可分离码等。防诬陷码的研究与经典的编码问题类似，主要关注它的上下界问题。概率方法是研究离散对象存在性的强有力工具，但是有时候我们直接从整个概率空间中选取，往往不能得到想要的结构（但是非常接近），这时候就需要利用修正方法（alteration），对已经选好的对象进行修补，最终完成整个结构的存在性证明。我们利用这种方法对防诬陷码的下界进行了研究，在很大的范围内改进了文献[71] 中的结果，同时我们也应用计数的技巧在字母表较小的情况下改进了原有的上界。这部分的工作已经投稿至《IEEE Transactions on Information Theory》。

6.4 光正交签名码

光正交签名码（OOSPC）本质上是一族具有良好的自相关性和互相关性的二元矩阵。光正交签名码在光 CDMA 网络上有着非常重要的应用[39,46,78]。我们主要用多项式所对应的关联矩阵给出了三类渐近意义下达到最优的光正交签名码的构造。同时，r-简单矩阵在各类光正交码的递归构造中都起到了不可替代的作用，改进 r-简单矩阵的参数就可以丰富各类光正交码的构造。我们利用代数几何码的想法，给出新参数的 r-简单矩阵的构造。这部分的工作已经投稿至《IEEE Transactions on Information Theory》。

6.5 关于集合差的集族问题

在 L-相交系的研究中，我们主要考虑两个集合之间的相交关系。相对的，在这里我们考虑两个集合之间的差的关系，回答了[54] 中提出的公开问题，并改进了文献[41] 的上界。

具体结果如下：

1. 设 $L=\{l_1,l_2,\ldots,l_s\}$ 和 $K=\{k_1,k_2,\ldots,k_r\}$ 是两个非负整数集合。若 $\mathcal{F}=\{F_1,F_2,\ldots,F_m\}$ 是 $[n]$ 中的集族满足对任意 i，$|F_i|\in K$，且对任意 $i\neq j$，$|F_i\cap F_j|\in L$。如果 $n\geq 2s-r$，那么 $|\mathcal{F}|\leq \binom{n}{s}+\binom{n}{s-1}+\cdots+\binom{n}{s-r+1}$。

2. 设 p 是一个素数，$L=\{l_1,l_2,\ldots,l_s\}$ 和 $K=\{k_1,k_2,\ldots,k_r\}$ 是 $\{1,2,\ldots,p-1\}$ 中的两个子集。若 $\mathcal{F}=\{F_1,F_2,\ldots,F_m\}$ 是 $[n]$ 中的集族满足对任意 i，$|F_i| \pmod p\in K$，且对任意 $i\neq j$，$|F_i-F_j| \pmod p\in L$。如果 $n\geq 2s-2r+1$ 或者 $n\geq s+\max_{1\leq i\leq r}k_i$，那么 $|\mathcal{F}|\leq \binom{n-1}{s}+\binom{n-1}{s-1}+\cdots+\binom{n-1}{s-2r+1}$。

这部分的工作已经投稿至《Journal of Combinatorial Designs》。

参考文献

[1] J. ACZÉL AND T. SZELE, *Solution 35*, Matematikai Lapok, 3 (1952), pp. 94–95.

[2] E. AGRELL, A. VARDY, AND K. ZEGER, *Upper bounds for constant-weight codes*, IEEE Trans. Inform. Theory, 46 (2000), pp. 2373–2395.

[3] N. ALON, L. BABAI, AND H. SUZUKI, *Multilinear polynomials and Frankl–Ray-Chaudhuri–Wilson type intersection theorems*, J. Combin. Theory Ser. A, 58 (1991), pp. 165–180.

[4] D. AUGOT, F. LEVY-DIT VEHEL, AND A. SHIKFA, *A storage-efficient and robust private information retrieval scheme allowing few servers*, in Cryptology and network security, vol. 8813 of Lecture Notes in Comput. Sci., Springer, Cham, 2014, pp. 222–239.

[5] R. BARANIUK, M. DAVENPORT, R. DEVORE, AND M. WAKIN, *A simple proof of the restricted isometry property for random matrices*, Constr. Approx., 28 (2008), pp. 253–263.

[6] S. BLACKBURN AND T. ETZION, *PIR array codes with optimal PIR rate*, arXiv preprint arXiv:1607.00235, (2016).

[7] E. BOMBIERI, *On exponential sums in finite fields*, Amer. J. Math., 88 (1966), pp. 71–105.

[8] R. C. BOSE, *A note on Fisher's inequality for balanced incomplete block designs*, Ann. Math. Statistics, 20 (1949), pp. 619–620.

[9] E. J. CANDÈS AND Y. PLAN, *Near-ideal model selection by ℓ_1 minimization*, Ann. Statist., 37 (2009), pp. 2145–2177.

[10] E. J. CANDÈS, J. ROMBERG, AND T. TAO, *Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information*, IEEE Trans. Inform. Theory, 52 (2006), pp. 489–509.

[11] E. J. CANDÈS, J. K. ROMBERG, AND T. TAO, *Stable signal recovery from incomplete and inaccurate measurements*, Comm. Pure Appl. Math., 59 (2006), pp. 1207–1223.

[12] E. J. CANDÈS AND T. TAO, *Near-optimal signal recovery from random projections: universal encoding strategies?*, IEEE Trans. Inform. Theory, 52 (2006), pp. 5406–5425.

[13] T. H. CHAN, S.-W. HO, AND H. YAMAMOTO, *Private information retrieval for coded storage*, in 2015 IEEE International Symposium on Information Theory (ISIT), IEEE, 2015, pp. 2842–2846.

[14] Y. CHEE, Z. CHERIF, J.-L. DANGER, S. GUILLEY, H. KIAH, J. KIM, P. SOLE, AND X. ZHANG, *Multiply constant-weight codes and the reliability of loop physically unclonable functions*, IEEE Trans. Inform. Theory, 60 (2014), pp. 7026–7034.

[15] Y. CHEE, F. GAO, H. KIAH, A. C. H. LING, H. ZHANG, AND X. ZHANG, *Decompositions of edge-colored digraphs: A new technique in the construction of constant-weight codes and related families*, arXiv:1401.3925, (2014).

[16] Y. CHEE, H. KIAH, H. ZHANG, AND X. ZHANG, *Constructions of optimal and near-optimal multiply constant-weight codes*, arXiv:1411.2513, (2014).

[17] W. Y. C. CHEN AND J. LIU, *Set systems with L-intersections modulo a prime number*, J. Combin. Theory Ser. A, 116 (2009), pp. 120–131.

[18] Z. CHERIF, J.-L. DANGER, S. GUILLEY, AND L. BOSSUET, *An easy-to-design puf based on a single oscillator: The loop puf*, in Proc. 15th Euromicro Conf. Digit. Syst. Design, Izmir, Turkey, Sep 2012, pp. 156–162.

[19] Z. CHERIF, J.-L. DANGER, S. GUILLEY, J.-L. KIM, AND P. SOLE, *Multiply constant weight codes*, in Proc. IEEE Int. Symp. Inf. Theory, Istanbul, Turkey, July 2013, pp. 306–310.

[20] B. CHOR, O. GOLDREICH, E. KUSHILEVITZ, AND M. SUDAN, *Private information retrieval*, J. ACM, 45 (1998), pp. 965–982.

[21] C. J. COLBOURN AND J. H. DINITZ, eds., *Handbook of combinatorial designs*, Chapman & Hall/CRC, second ed., 2007.

[22] P. D'ARCO, N. N. ESFAHANI, AND D. R. STINSON, *All or nothing at all*, arXiv:1510.03655, (2015).

[23] H. DAVENPORT AND G. HAJÓS, *Problem 35*, Matematikai Lapok, 2 (1951), p. 68.

[24] N. G. DE BRUIJN AND P. ERDÖS, *On a combinatorial problem*, Nederl. Akad. Wetensch., Proc., 51 (1948), pp. 1277–1279 = Indagationes Math. 10, 421–423 (1948).

[25] P. DELSARTE, *An algebraic approach to the association schemes of coding theory*, Philips Res. Repts. Suppl., (1973).

[26] R. A. DEVORE, *Deterministic constructions of compressed sensing matrices*, J. Complexity, 23 (2007), pp. 918–925.

[27] D. L. DONOHO, *Compressed sensing*, IEEE Trans. Inform. Theory, 52 (2006), pp. 1289–1306.

[28] D. L. DONOHO AND M. ELAD, *Optimally sparse representation in general (nonorthogonal) dictionaries via l^1 minimization*, Proc. Natl. Acad. Sci. USA, 100 (2003), pp. 2197–2202 (electronic).

[29] D. L. DONOHO AND J. TANNER, *Counting faces of randomly projected polytopes when the projection radically lowers dimension*, J. Amer. Math. Soc., 22 (2009), pp. 1–53.

[30] M. F. DUARTE AND Y. C. ELDAR, *Structured compressed sensing: from theory to applications*, IEEE Trans. Signal Process., 59 (2011), pp. 4053–4085.

[31] P. ERDÖS, *Problem 20*, Matematikai Lapok, 1 (1950), p. 226.

[32] P. ERDÖS, C. KO, AND R. RADO, *Intersection theorems for systems of finite sets*, Quart. J. Math. Oxford Ser. (2), 12 (1961), pp. 313–320.

[33] G. FANTI AND K. RAMCHANDRAN, *Efficient private information retrieval over unsynchronized databases*, IEEE J. Sel. Topics Signal Process., 9 (2015), pp. 1229–1239.

[34] A. FAZELI, A. VARDY, AND E. YAAKOBI, *Codes for distributed PIR with low storage overhead*, in 2015 IEEE International Symposium on Information Theory (ISIT), IEEE, 2015, pp. 2852–2856.

[35] ——, *PIR with low storage overhead: coding instead of replication*, arXiv preprint arXiv:1505.06241, (2015).

[36] P. FRANKL AND R. M. WILSON, *Intersection theorems with geometric consequences*, Combinatorica, 1 (1981), pp. 357–368.

[37] B. GASSEND, D. CLARKE, M. VAN DIJK, AND S. DEVADAS, *Silicon physical random functions*, in Proc. 9th ACM Conf. Comput. Commun. Secur., 2002, pp. 148–160.

[38] V. GROLMUSZ AND B. SUDAKOV, *On k-wise set-intersections and k-wise Hamming-distances*, J. Combin. Theory Ser. A, 99 (2002), pp. 180–190.

[39] A. A. HASSAN, J. E. HERSHEY, AND N. A. RIZA, *Spatial optical cdma*, IEEE J. Sel. Areas Commun, 13 (1995), pp. 609–613.

[40] J. HAUPT, W. U. BAJWA, G. RAZ, AND R. NOWAK, *Toeplitz compressed sensing matrices with applications to sparse channel estimation*, IEEE Trans. Inform. Theory, 56 (2010), pp. 5862–5875.

[41] K.-W. HWANG, T. KIM, L. C. JANG, P. KIM, AND G. SOHN, *Alon-Babai-Suzuki's conjecture related to binary codes in nonmodular version*, J. Inequal. Appl., (2010), pp. Art. ID 546015, 4.

[42] K.-W. HWANG AND Y. KIM, *A proof of Alon-Babai-Suzuki's conjecture and multilinear polynomials*, European J. Combin., 43 (2015), pp. 289–294.

[43] A. JIANG, R. MATEESCU, M. SCHWARTZ, AND J. BRUCK, *Rank modulation for flash memories*, IEEE Trans. Inform. Theory, 55 (2009), pp. 2659–2673.

[44] S. JOHNSON, *Upper bounds for constant weight error correcting codes*, Discrete Math, 3 (1972), pp. 109–124.

[45] N. M. KATZ, *An estimate for character sums*, J. Amer. Math. Soc., 2 (1989), pp. 197–200.

[46] K. KITAYAMA, *Novel spatial spread spectrum based fiber optic cdma networks for image transmission*, IEEE J. Sel. Areas Commun, 12 (1994), pp. 762–772.

[47] F. KRAHMER, S. MENDELSON, AND H. RAUHUT, *Suprema of chaos processes and the restricted isometry property*, Comm. Pure Appl. Math., 67 (2014), pp. 1877–1904.

[48] E. R. LAMKEN AND R. M. WILSON, *Decompositions of edge-colored complete graphs*, J. Combin. Theory Ser. A, 89 (2000), pp. 149–200.

[49] V. I. LEVENSHTEIN, *Upper-bound estimates for fixed-weight codes*, Probl. Pered. Inform., (1971), pp. 3–12.

[50] K. LI, L. GAN, AND C. LING, *Convolutional compressed sensing using deterministic sequences*, IEEE Trans. Signal Process., 61 (2013), pp. 740–752.

[51] S. LI, F. GAO, G. GE, AND S. ZHANG, *Deterministic construction of compressed sensing matrices via algebraic curves*, IEEE Trans. Inform. Theory, 58 (2012), pp. 5035–5041.

[52] S. LI AND G. GE, *Deterministic construction of sparse sensing matrices via finite geometry*, IEEE Trans. Signal Process, 62 (2014), pp. 2850–2859.

[53] ——, *Deterministic sensing matrices arising from near orthogonal systems*, IEEE Trans. Inform. Theory, 60 (2014), pp. 2291–2302.

[54] S. LI AND H. ZHANG, *Set systems with L-intersections and k-wise L-intersecting families*, J. Combin. Des., 24 (2016), pp. 514–529.

[55] J. LIU AND W. YANG, *Set systems with restricted k-wise L-intersections modulo a prime number*, European J. Combin., 36 (2014), pp. 707–719.

[56] D. NEEDELL AND J. A. TROPP, *CoSaMP: iterative signal recovery from incomplete and inaccurate samples*, Appl. Comput. Harmon. Anal., 26 (2009), pp. 301–321.

[57] D. NEEDELL AND R. VERSHYNIN, *Uniform uncertainty principle and signal recovery via regularized orthogonal matching pursuit*, Found. Comput. Math., 9 (2009), pp. 317–334.

[58] R. PAPPU, B. RECHT, J. TAYLOR, AND N. GERSHENFELD, *Physical one-way functions*, Science, 297 (2002), pp. 2026–2030.

[59] J. QIAN AND D. K. RAY-CHAUDHURI, *On mod-p Alon-Babai-Suzuki inequality*, J. Algebraic Combin., 12 (2000), pp. 85–93.

[60] R. A. RANKIN, *The closest packing of spherical caps in n dimensions*, Proc. Glasgow Math. Assoc., 2 (1955), pp. 139–144.

[61] H. RAUHUT, *Compressive sensing and structured random matrices*, in Theoretical foundations and numerical methods for sparse recovery, vol. 9 of Radon Ser. Comput. Appl. Math., Walter de Gruyter, Berlin, 2010, pp. 1–92.

[62] H. RAUHUT, J. ROMBERG, AND J. A. TROPP, *Restricted isometries for partial random circulant matrices*, Appl. Comput. Harmon. Anal., 32 (2012), pp. 242–254.

[63] H. RAUHUT, K. SCHNASS, AND P. VANDERGHEYNST, *Compressed sensing and redundant dictionaries*, IEEE Trans. Inform. Theory, 54 (2008), pp. 2210–2219.

[64] D. K. RAY-CHAUDHURI AND R. M. WILSON, *On t-designs*, Osaka J. Math., 12 (1975), pp. 737–744.

[65] R. L. RIVEST, *All-or-nothing encryption and the package transform*, in Fast Software Encryption, Springer, 1997, pp. 210–218.

[66] J. ROMBERG, *Compressive sensing by random convolution*, SIAM J. Imaging Sci., 2 (2009), pp. 1098–1128.

[67] M. RUDELSON AND R. VERSHYNIN, *On sparse reconstruction from Fourier and Gaussian measurements*, Comm. Pure Appl. Math., 61 (2008), pp. 1025–1045.

[68] K. SARKADI AND T. SZELE, *Solution 20*, Matematikai Lapok, 2 (1951), pp. 76–77.

[69] N. B. SHAH, K. RASHMI, AND K. RAMCHANDRAN, *One extra bit of download ensures perfectly private information retrieval*, in 2014 IEEE International Symposium on Information Theory, IEEE, 2014, pp. 856–860.

[70] H. S. SNEVILY, *On generalizations of the de Bruijn-Erdős theorem*, J. Combin. Theory Ser. A, 68 (1994), pp. 232–238.

[71] D. R. STINSON, R. WEI, AND K. CHEN, *On generalized separating hash families*, J. Combin. Theory Ser. A, 115 (2008), pp. 105–120.

[72] G. E. SUH AND S. DEVADAS, *Physical unclonable functions for device authentication and secret key generation*, in Proc. 44th ACM/IEEE Ann. Design Autom. Conf., June 2007, pp. 9–14.

[73] R. TAJEDDINE AND S. E. ROUAYHEB, *Private information retrieval from MDS coded data in distributed storage systems*, arXiv preprint arXiv:1602.01458, (2016).

[74] J. A. TROPP, *Greed is good: algorithmic results for sparse approximation*, IEEE Trans. Inform. Theory, 50 (2004), pp. 2231–2242.

[75] ——, *On the conditioning of random subdictionaries*, Appl. Comput. Harmon. Anal., 25 (2008), pp. 1–24.

[76] L. WELCH, *Lower bounds on the maximum cross correlation of signals (corresp.)*, IEEE Trans. Inform. Theory, 20 (1974), pp. 397–399.

[77] G. XU AND Z. XU, *Compressed sensing matrices from fourier matrices*, IEEE Trans. Inform. Theory, 61 (2015), pp. 469–478.

[78] G. C. YANG AND W. C. KWONG, *Two-dimensional spatial signature patterns*, IEEE Trans. Commun, 44 (1996), pp. 184–191.

作者简历

- 汪馨，男，浙江大学数学科学学院博士生，导师：葛根年.
- 通信地址：中国浙江省杭州市浙江大学玉泉校区数学科学学院，310027.
- 联系方式：(+86)18157303863，11235062@zju.edu.cn
- 教育经历：

  2008.9–2012.8，浙江大学数学系，数学与应用数学专业，理学学士.

  2012.9–至今，浙江大学数学科学学院，应用数学专业，理学博士，研究方向：组合数学与编码密码学.
- 研究兴趣：极值组合学，编码理论.

攻读博士学位期间主要研究成果

1. Xin Wang, Jun Zhang, Gennian Ge, “Determinstic Convolutional Compressed Sensing Matrices”, Finite Fields and Their Application, vol. 42, pp. 102-117, Nov. 2016.

2. Xin Wang, Hengjia Wei, Chong Shangguan, Gennian Ge, “New bounds and constructions for multiply constant-weight codes”, IEEE Transactions on Information Theory, vol. 62, no. 11, pp. 6315-6327, Nov. 2016.

3. Yiwei Zhang, Tao Zhang, Xin Wang, Gennian Ge, “Invertible binary matrix with maximum number of 2-by-2 invertible submatrices”, Discrete Mathematics, vol. 340, no. 2, pp. 201-208, Feb. 2017.

4. Xin Wang, Yiwei Zhang, Yiting Yang, Gennian Ge, “New bounds on permutation codes under Hamming metric and Kendall’s τ-metric”, Designs, Codes and Cryptography, to appear.

5. Xin Wang, Hengjia Wei, Gennian Ge, “A strengthened inequality of Alon-Babai-Suzuki’s conjecture in set systems with restricted intersections modulo p”, submitted.

6. Chong Shangguan, Xin Wang, Ying Miao, Gennian Ge, “New Bounds For Frameproof Codes”, submitted.

7. Yiwei Zhang, Xin Wang, Hengjia Wei, Gennian Ge, “On private information retrieval array codes”, submitted.

8. Lijun Ji, Baokun Ding, Xin Wang, Gennian Ge, “Asymptotically Optimal Optical Orthogonal Signature Pattern Codes”, submitted.

9. Xiangliang Kong, Xin Wang, Gennian Ge, “On set systems with restricted intersections and differences”, submitted.