arXiv:2301.01668v2 [math.CO] 5 Jan 2023

三角形自由图上码率趋近于一的存储码构造

Hexiang Huang QingXiang*

2022年12月29日

摘要

考虑将比特分配给连通图 $\Gamma(V,E)$ 的各个顶点，并满足每个顶点的取值都是其邻居取值的函数。这样一组分配称为 Γ 上长度为 $|V|$ 的存储码。本文在无三角形图上构造了一个无限族的二进制线性存储码，其码率可以任意接近一。

关键词。 存储码、凯莱图、群代数、湮灭子

1 引言

设 Γ 是一个具有 n 个顶点的简单图， $\mathcal{C}$ 是一个长度为 n 的二元码，其坐标由 Γ 的顶点索引。如果对于每个码字 $c\in\mathcal{C}$， c 的每个坐标都能由其邻居的取值唯一确定，则称 $\mathcal{C}$ 是 Γ 上的存储码。

给定一个图 Γ，我们可以在其上定义一个存储码：假设 Γ 有 n 个顶点，记为 v_1、v_2、.....、v_n。令 $A(\Gamma)$ 为 Γ 的邻接矩阵，其行和列分别以 v_1、v_2、.....、v_n 标记。令 $H=A(\Gamma)+I$，并令 $\mathcal{C}$ 为定义在 $\mathbb{F}_2$ 上、校验矩阵为 H 的线性码。令 $c=(c_1,c_2,.....,c_n)\in\mathcal{C}$ 为一个码字。对于 c 的第 v_i^{th} 个分量，即 c_i， H 的第 v_i 行给出一个线性方程，$c_i=\sum_{v_j\in N(v_i)}c_j$，其中 $N(v_i)$ 是 v_i 在 Γ 中的邻居集合；因此，可以根据其邻居的取值恢复 c_i。Mazumdar [1, 2] 以及 Shanmugam 和 Dimakis [3] 引入了图上的存储码。 [4] 和 [5] 的作者也提出了图上存储码的概念，不过表述方式有所不同。我们只考虑二元线性存储码。存储码 $\mathcal{C}$ 的码率记为 $R(\mathcal{C})$，即其维数与环境空间维数之比。若有一族存储码 $\mathcal{C}_r$，其中 r 是一个参数，并假设 $\lim_{r\to\infty}R(\mathcal{C}_r)$ 存在，则该极限称为这一族的码率。

构造码率接近 1 的存储码很容易：令 Γ 为有 n个顶点的完全图， $\mathcal{C}$ 为由单个校验方程 $\sum_{i=1}^{n}x_i=0$定义的二元线性码。则$\mathcal{C}$ 是定义在 Γ 上的存储码，码率为 $1-1/n$。

在上述示例中，用于构造码率接近 1 的存储码的图非常稠密（事实上，稠密程度已达可能的最大值），并且包含大量团。因此，很自然地

*本研究部分受国家自然科学基金委员会项目 12071206、12131011、12150710510 以及中德流动计划 M-0157 资助

会考虑这样一个问题：在不含团的图上，存储码可达到的最大码率是多少，也就是说，在无三角形的图上，存储码的最大码率是多少。

在这类图上构造高码率的存储码是一项挑战。考虑最稠密的无三角形图，即完全二部图$K_{m,n}$以及其上的存储码 $\mathcal{C}$ 。其中有两个独立顶点集$K_{m,n}$。对于每个顶点，我们都可以根据另一个（顶点）独立集中的顶点恢复该顶点上的消息，因此 $R(\mathcal{C}) \leq 1/2$。在早期研究中 [5], ，作者曾猜想，对于无三角形图， $R = 1/2$ 是可达到的最大码率。后来，[4] 一些孤立的例子推翻了这一猜想。

最近，Barg 和 Z´emor [6] 提出了四个定义在无三角形图上的存储码无限族。他们使用了凯莱图方法：取 S 为 $\mathbb{F}_2^r$ 的子集，使得 $0 \notin S$，且 S 中任意三个向量都线性无关。于是 $\Gamma = \mathrm{Cay}(\mathbb{F}_2^r, S)$ 是无三角形的。令 $H = A(\Gamma) + I$，并利用 H 作为校验矩阵定义二元线性码 $\mathcal{C}$ 。由此，我们得到定义在无三角形图 Γ 上的存储码 $\mathcal{C}$ 。通过这种方法，子集 $S \subseteq \mathbb{F}_2^r$ 将生成一个无三角形图及其上的存储码。

Barg 和 Z´emor 的构造 [6]

- Clebsch族：这是一个定义在广义Clebsch图上的无限存储码族，其码率高于 $1/2$，并逐渐趋近于 $1/2$。
- 汉明族：要构造这个族，首先取长度为 2^{r-1} 的扩展二元Hamming码的校验矩阵，为其增添一列全零列，然后再添加一行权重为 2 的行，其中首项和末项均为 1。将所得的 $(r+1) \times n$ 矩阵记为 M_r。例如，对于 $r = 4$，我们得到矩阵

$$M_4 = \left[\begin{array}{cccccccc|c} 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ \hline 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right].$$

  令 S 为 M_r的列向量集合。该码族的码率逐渐趋近于$3/4$。我们将在定理 2.3 中给出这一事实的新证明，然后推广这个码族。
- BCH族：设 B_s 为BCH码的校验矩阵， S 为B_s的列集合。这个码族的码率能否达到 1，仍是一个开放问题。然而，计算机计算结果表明，当 s 从 4 变化到 8 时，码率会提高，尤其是当 $s = 8$ 时，码率为 0.8196。这是截至[6]撰写时，已知的无三角形图上存储码的最高码率。

- Reed-Muller族：设 M 是一个矩阵，其各行由线性和二次布尔多项式函数的求值向量组成。设 S 为 M的列集。这个族完全满足码率高所需的必要条件；但我们不知道这个族的码率。

前两个族的码率是通过矩阵方法 [6]计算出来的。除了给出这些构造外， [6] 的作者还提出了码率较高所需的必要条件。

定理 1.1（[6]）。设 $\Gamma = \mathrm{Cay}(\mathbb{F}_2^r, S)$，其中 S不包含零向量。设 M为一个矩阵，其列由 S 中的元素组成，并且 $\mathcal{C}$ 为校验矩阵 $H = A(\Gamma) + I$的二元码。那么

(i) 若 $R(\mathcal{C}) > 1/2$，则 n 为奇数，且 M 的所有行的重量均为偶数；(ii) 若 $R(\mathcal{C}) > (2^k-1)/2^k, k = 2, 3, \ldots,$ 则 M 的任意 k 行的交集均为偶数。

2 码率为 1 的无三角形图上的一族存储码

定义 2.1。设 G 是一个乘法记号的有限群，其单位元为 e，并设 S是 G 的一个子集，使得 $e \notin S$ 且 $S = S^{-1}$，其中 $S^{-1} = \{g^{-1} \mid g \in S\}$。以 S为连接集的 G 上的凯莱图，记作 $\Gamma = \mathrm{Cay}(G, S)$，是一个以 G 中的元素为顶点的图；两个顶点 g_1、$g_2 \in G$ 相邻，当且仅当 $g_1 g_2^{-1} \in S$。

本文将仅使用凯莱图方法来构造存储码。回顾一些记号： S 是 $\mathbb{F}_2^r$ 的一个子集，其中 $0 \notin S$，$\Gamma = \mathrm{Cay}(\mathbb{F}_2^r, S)$，且 $H = A(\Gamma) + I$。存储码 $\mathcal{C}$ 是 $\mathbb{F}_2$ 上的线性码，其校验矩阵为 H。注意，在 Γ 中，每个顶点都表示 $\mathbb{F}_2^r$ 中的一个向量，而 H 的行和列由 $\mathbb{F}_2^r$ 中的向量标记。若 Γ 中存在一个三角形，即存在三个向量 v_1、v_2、v_3，使得 $v_1 - v_2$、$v_2 - v_3$、$v_3 - v_1 \in S$，那么就存在三个向量 w_1、w_2、$w_3 \in S$，使得 $w_1 + w_2 + w_3 = 0$。因此，若 S 中任意三个向量之和都不为零，则 Γ 是无三角形图。

如果我们有一个 $S \subseteq \mathbb{F}_2^r$，根据凯莱图的定义，我们有

$$A(\Gamma) = (a_{uv}), \quad a_{uv} = \begin{cases} 1, \ if \ v - u \in S, \\ 0, \ otherwise. \end{cases}$$

由于 $H = A(\Gamma) + I$，我们有

$$H = (h_{uv}), \quad h_{uv} = \begin{cases} 1, \ if \ v - u \in S', \\ 0, \ otherwise, \end{cases} \tag{1}$$

其中 $S' = S \cup \{0\}$。我们称 H 为 陪集矩阵 S。考虑 S更为方便，它包含零向量。由于 H 是 $\mathcal{C}$的校验矩阵，$\mathcal{C}$ 是满足 c 的行向量集合 $H \cdot c^{\intercal} = 0$。注意，在（1）中， $v - u = v + u$ 因为我们是在 $\mathbb{F}_2$上进行运算，所以 H 是对称的。因此 $\mathcal{C} = \{c \in \mathbb{F}_2^{2^r} \mid c \cdot H = 0\}$，我们记 $\mathcal{C} = \mathrm{Null}(H)$，即 H的零空间。

我们的目的是构造一个 $S \subseteq \mathbb{F}_2^r$，它满足 $0 \in S$，且 S 中任意三个非零向量之和都不为零，同时使码率 $R(\mathrm{Null}(H))$尽可能高。我们发现，如果将汉明族中的 S 转化到更具代数性的框架中，问题便会变得简洁明了，也更容易推广。

$\mathcal{C}$ 的环境空间是 $\mathbb{F}_2^{2^r}$。通过交替使用 $c = (c_v \mid v \in \mathbb{F}_2^r)$和$\sum_{v \in \mathbb{F}_2^r} c_v v$，我们可以将 $\mathbb{F}_2^{2^r}$ 视为群代数 $\mathbb{F}_2[\mathbb{F}_2^r]$。将向量空间 $\mathbb{F}_2^r$ 中的加法记为“$\oplus$”，则$\mathbb{F}_2[\mathbb{F}_2^r]$ 是一个环，其加法为“$+$”，乘法为“$\oplus$”。

注意到， H 的每一行都是 $\mathbb{F}_2^r$某个子集的示性向量。通过对这两个概念稍作混用，我们可以看出， H 的第 v 行是$(\mathbb{F}_2^r, \oplus)$中的陪集 $v \oplus S$。正如循环码可以视为某个群环中的主理想，我们将证明，$\mathrm{rowspan}(H) \subseteq \mathbb{F}_2[\mathbb{F}_2^r]$ 与相应环中的一个主理想同构，而零空间 $\mathcal{C} = \mathrm{Null}(H) \subseteq \mathbb{F}_2[\mathbb{F}_2^r]$与该主理想的湮灭子同构。

设 $G = \langle x_1, x_2, \ldots, x_r | x_1^2 = x_2^2 = \cdots = x_r^2 = 1 \rangle$ 为大小为 2^r的初等阿贝尔 2-群。为避免使用两个加法符号带来的麻烦，我们将使用以下群同构

$$\sigma : (\mathbb{F}_2^r, \oplus) \to (G, \cdot)$$
$$(a_1, a_2, \ldots, a_r) \mapsto x_1^{a_1} x_2^{a_2} \cdots x_r^{a_r}$$

注意， σ 在群代数之间诱导出一个环同构

$$\phi : (\mathbb{F}_2[\mathbb{F}_2^r], +, \oplus) \to (\mathbb{F}_2[G], +, \cdot) \simeq P_r$$
$$\sum_{v \in \mathbb{F}_2^r} k_v v \mapsto \sum_{v \in \mathbb{F}_2^r} k_v \sigma(v)$$

其中 $P_r := \mathbb{F}_2[x_1, x_2, \ldots, x_r] / \langle x_1^2 - 1, x_2^2 - 1, \ldots, x_r^2 - 1 \rangle$。

设 z 是某个交换环 A 中的一个元素。回顾一下，主理想 $\langle z \rangle = \{a \cdot z \mid a \in A\}$，湮灭子 $\mathrm{Ann}(z) = \{a \in A \mid a \cdot z = 0\}$。

假设 $S \subseteq \mathbb{F}_2^r$ 包含零向量。令 $f_S = \sum_{w \in S} \sigma(w) \in P_r$，且 H 为 S 的陪集矩阵。于是我们有以下两个命题。

命题 2.1. 对偶码 $\mathcal{C}^\perp =$ 的行空间(H) 同构于主理想 $\langle f_S \rangle \subseteq P_r$。证明. 将 H 中由 $v \oplus S$ 索引的行记为 $h_v, v \in \mathbb{F}_2^r$。注意， h_v 是 $\mathbb{F}_2^{2^r}$ 中的一个向量，因此也是 $\mathbb{F}_2[\mathbb{F}_2^r]$ 中的一个元素；此外， $h_0 = \sum_{w \in S} w$ 和 $h_v = \sum_{w \in S} v \oplus w = v \oplus h_0$。因此

$$\mathrm{rowspan}(H) = \left\{ \sum_{v \in \mathbb{F}_2^r} k_v h_v \,|\, k_v \in \mathbb{F}_2 \right\} = \left\{ \sum_{v \in \mathbb{F}_2^r} k_v (v \oplus h_0) \,|\, k_v \in \mathbb{F}_2 \right\}$$
$$= \left\{ \left(\sum_{v \in \mathbb{F}_2^r} k_v v \right) \oplus h_0 \,|\, k_v \in \mathbb{F}_2 \right\} = \{a \oplus h_0 \,|\, a \in \mathbb{F}_2[\mathbb{F}_2^r]\}.$$

对于每个 $a \oplus h_0 \in$ 行空间(H)，我们定义

$$\phi'(a \oplus h_0) = \phi(a \oplus h_0) = \phi(a) \cdot \phi(h_0)$$
$$= \phi(a) \cdot \phi\left(\sum_{w \in S} w \right) = \phi(a) \cdot \sum_{w \in S} \sigma(w) = \phi(a) \cdot f_S.$$

由于 $\phi'(a \oplus h_0) \in \langle f_S \rangle$，映射 ϕ' 是良定义的。此外，映射 ϕ' 是线性且单射的，因为它是 ϕ 在 $\mathcal{C}^\perp$ 上的限制。它也是满射的： ϕ 是双射，因此存在逆映射。于是对于每个 $g \in P_r$，我们有 $\phi'(\phi^{-1}(g) \oplus h_0) = \phi(\phi^{-1}(g)) \cdot f_S = g \cdot f_S$，因此 $g \cdot f_S$ 在 $\phi^{-1}(g) \oplus h_0 \in$ 行空间(H) 中有一个原像。由此可得， ϕ' 是线性双射。

□

命题 2.2. 码 $\mathcal{C} =$ 零空间(H) 与 湮灭子$(f_S) \subseteq P_r$ 同构。

证明。对于每个 $a \in$零空间(H)，我们定义 $\phi''(a) = \phi(a)$。注意，

$$\mathrm{Null}(H) = \left\{ a \in \mathbb{F}_2^{2^r} \mid \sum_{v \in \mathbb{F}_2^r} a_v h_v = 0 \right\} = \left\{ a \in \mathbb{F}_2^{2^r} \mid \sum_{v \in \mathbb{F}_2^r} a_v (v \oplus h_0) = 0 \right\}$$
$$= \left\{ a \in \mathbb{F}_2^{2^r} \mid \left(\sum_{v \in \mathbb{F}_2^r} a_v v \right) \oplus h_0 = 0 \right\} = \{a \in \mathbb{F}_2[\mathbb{F}_2^r] \,|\, a \oplus h_0 = 0\}.$$

注意　在最后一个等式中，我们从线性空间 $\mathbb{F}_2^{2^r}$ 转到群代数 $\mathbb{F}_2[\mathbb{F}^r_2]$。我们可以验证，$\phi''$ 是良定义的：对于每个 $a \in \mathsf{Null}(H)$ $\phi''(a) \cdot f_S = \phi(a) \cdot \phi(h_0) = \phi(a \oplus h_0) = 0$，因此 $\phi''(a) \in \mathsf{Ann}(f_S)$。$\phi''$ 是线性且单射的，因为它是 ϕ 在 $\mathsf{Null}(H)$ 上的限制。它也是满射的：对于每个 $g = \sum_{m\in G} k_m m \in \mathsf{Ann}(f_S)$，我们都能在 $\mathsf{Null}(H)$ 中找到其原像 $\sum_{m\in G} k_m \sigma^{-1}(m)$。因此，$\phi''$ 是线性双射。 □

现在我们可以在商环 $P_r = \mathbb{F}_2[x_1, x_2, \ldots, x_r]/\langle x_1^2 - 1, x_2^2 - 1, \ldots, x_r^2 - 1\rangle$ 中进行运算。假设 $S \subseteq \mathbb{F}_2^r$ 包含零向量，且 S 中任意三个非零向量的和都不为零。通常，我们将存储码视为 $f_S \in P_r$ 的湮灭子，而不明确提及环境图 $\mathsf{Cay}(\mathbb{F}_2^r, S \setminus \{0\})$。下面列出 P_r 的一些性质：

(i) $(x_i + 1)^2 = 0$ 属于 P_r，对 $i = 1, \ldots, r$；(ii) $R(\langle x_1 + 1\rangle + \langle x_2 + 1\rangle + \cdots + \langle x_k + 1\rangle) = 1 - \frac{1}{2^k}$；(iii) $R\left(\prod_{i=1}^n \sum_{j=1}^{k_i} \langle x_{ij} + 1\rangle\right) = \prod_{i=1}^n R\left(\sum_{j=1}^{k_i} \langle x_{ij} + 1\rangle\right)$。

为证明第一个性质，注意到 $(x_i+1)^2 = x_i^2+1$ 属于理想 $\langle x_1^2 - 1, x_2^2 - 1, \ldots, x_r^2 - 1\rangle$，因此它在商环 P_r 中为零。其余两个性质的证明见第 3节；第二个性质将在定理 3.6 中证明，第三个性质将在推论3.9中证明。

利用这些性质，我们可以简单地证明汉明族的码率为 $3/4$。

定理 2.3（汉明族）。汉明族中的存储码具有 $\mathsf{Ann}(f_r) \subseteq P_{r+1}$ 的形式，其中

$$f_r = (x_r + 1)(x_{r+1} + 1) + x_r \cdot (x_1 + 1)(x_2 + 1)\cdots(x_{r-1} + 1).$$

关联　图是无三角形的，并且码率 $R(\mathsf{Ann}(f_r))$ 在 r 趋于无穷大时趋近于 $3/4$。

证明。令 M_r 为汉明族中构造的矩阵。例如，当 $r = 4$时，我们有

$$M_4 = \left[\begin{array}{cccccccc|c} 0&0&0&0&1&1&1&1&0\\ 0&0&1&1&0&0&1&1&0\\ 0&1&0&1&0&1&0&1&0\\ 1&1&1&1&1&1&1&1&0\\ \hline 1&0&0&0&0&0&0&0&1 \end{array}\right]$$

显然，M_r 的任意三个列向量都线性无关，因此环境图是无三角形的。请记住，我们需要一个包含零向量的 $S \subseteq \mathbb{F}_2^{r+1}$，所以在 M_r 中增添一列全零列，得到 M_r'，并令 $f_r \in P_{r+1}$ 为对应于 M_r' 的元素。于是，存储码为湮灭子 f_r。我们有

$$\begin{aligned} f_r &= x_{r+1}x_r + x_r \cdot \left(\sum_{(s_1,\ldots,s_{r-1})\neq(0,\ldots,0)} x_1^{s_1} x_2^{s_2} \cdots x_{r-1}^{s_{r-1}}\right) + x_{r+1} + 1 \\ &= x_{r+1}x_r + x_r \cdot \left(\sum x_1^{s_1} x_2^{s_2} \cdots x_{r-1}^{s_{r-1}}\right) + x_r + x_{r+1} + 1 \\ &= (x_r + 1)(x_{r+1} + 1) + x_r \cdot (x_1 + 1)(x_2 + 1)\cdots(x_{r-1} + 1). \end{aligned}$$

利用性质 (i)， $x_i+1 \in$ 湮灭子(x_i+1)，我们便得到一个包含关系

$$\mathrm{Ann}(f_r) \supseteq (\langle x_r+1\rangle + \langle x_{r+1}+1\rangle)(\langle x_1+1\rangle + \langle x_2+1\rangle + \cdots + \langle x_{r-1}+1\rangle).$$

利用性质(ii) 和(iii)，我们得到

$$\begin{aligned} R(\mathrm{Ann}(f_r)) &\geq R((\langle x_r+1\rangle + \langle x_{r+1}+1\rangle)(\langle x_1+1\rangle + \langle x_2+1\rangle + \cdots + \langle x_{r-1}+1\rangle)) \\ &= R(\langle x_r+1\rangle + \langle x_{r+1}+1\rangle) \cdot R(\langle x_1+1\rangle + \langle x_2+1\rangle + \cdots + \langle x_{r-1}+1\rangle) \\ &= \frac{3}{4} \cdot \left(1 - \frac{1}{2^{r-1}}\right). \end{aligned}$$

注意到 M_r 的最后两行恰好在一个位置相交，因此 R(湮灭子(f_r)) $\leq 3/4$ 由定理 1.1可知。因此，码率的极限 R(湮灭子(f_r)) 为 $3/4$。 □

注记。我们注意到，定理 f_r 2.3中构造的元素x_r+1是一个“短”项（$x_{r+1}+1$）（$x_r\cdot$）与一个“长”项 x_1+1（x_2+1）（$\cdots$）（$x_{r-1}+1$）之和。注意，根据性质(ii)

$$R(\mathrm{Ann}((x_r+1)(x_{r+1}+1))) \geq R(\langle x_r+1\rangle + \langle x_{r+1}+1\rangle) = 3/4.$$

因此，仅“短”项就已经达到了 $3/4$ 的码率。然而，存在一些单项式 x_r、 x_{r+1}、 x_rx_{r+1}（在 $(x_r+1)($ $x_{r+1}+1$) 的展开式中），它们表示 S 中三个线性相关向量，因此 $(x_r+1)(x_{r+1}+1)$ 会产生一个含有三角形的凯莱图。为消除线性相关，需要“长”项；如果“长”项中的额外变量数量 $r-1$ 很大，那么码率值就不会降低太多。

考虑 $f=(x_1+1)(x_2+1)(x_3+1) \in P_3$。我们有 Ann($f$) $\supseteq \langle x_1+1\rangle + \langle x_2+1\rangle + \langle x_3+1\rangle$，由性质 (i) 可知；而 $R(\mathrm{Ann}(f)) \geq 7/8$，由性质 (ii)可知。因此很容易构造高码率存储码；然而，此处的环境凯莱图并非无三角形图，因为存在许多“线性相关”的三元组，例如 x_1、 x_2、 x_1x_2。

定理 2.4. 设 $\mathcal{C}_k$ =Ann(f_k) $\subseteq P_{3k+3}$，其中

$$\begin{aligned} f_k =& (x_{3k+1}+1)(x_{3k+2}+1)(x_{3k+3}+1) + x_{3k+1}x_{3k+2}(x_1+1)\cdots(x_k+1) \\ &+ x_{3k+1}x_{3k+3}(x_{k+1}+1)\cdots(x_{2k}+1) + x_{3k+2}x_{3k+3}(x_{2k+1}+1)\cdots(x_{3k}+1). \end{aligned}$$

那么， $\mathcal{C}_k$ 是无三角形图上的存储码。该族的码率为 $7/8$。

证明。设 $S \subseteq \mathbb{F}_2^{3k+3}$为与完全展开的$f_k$中的单项式相对应的向量子集。我们需要验证 $0 \in S$，并且 S 中任意三个非零向量之和不为零。

f_k的展开式中有一个 1，因此 $0 \in S$。注意， S 中任意三个非零向量之和不为零，等价于 f_k+1 展开式中任意三个单项式的乘积不为 $1 \in P_{3k+3}$。三个单项式的选取有以下几种情况：

(i) 三个单项式均选自“短”项（$x_{3k+1}+1$）（$x_{3k+2}+1$）（$x_{3k+3}+1$）中保留下来的单项式。注意，在 f_k+1的展开式中，偶次数项 1、 $x_{3k+1}x_{3k+2}$、 $x_{3k+1}x_{3k+3}$、 $x_{3k+2}x_{3k+3}$均被抵消，因此只有奇次数项，即 x_{3k+1}、 x_{3k+2}、x_{3k+3}和 $x_{3k+1}x_{3k+2}x_{3k+3}$ 保留下来，所以其中任意三个的乘积都不可能为 1。

(ii) 从某个“长”项展开式中保留下来的单项式项里任选三个，例如从 $x_{3k+1}x_{3k+2}(x_1+1)\cdots($ x_k+1) 的展开式中选取。则这三个单项式项的乘积会含有因子 $x_{3k+1}^3x_{3k+2}^3 = x_{3k+1}x_{3k+2}$，因此不等于$1 \in P_{3k+3}$。

(iii) 还有另外三种情况：从“短”项中选两个、从某个“长”项中选一个；从“短”项中选一个、从“长”项中选两个；从多个“长”项中选三个。对于所有这些情况，注意到乘积必定在某个变量$x_i, i = 1, \ldots, 3k$ 中具有 1 次度，因此不等于 $1 \in P_{3k+3}$。

因此，凯莱图是无三角形的。与定理 2.3的证明一样，我们有包含关系

$$\begin{aligned}\mathrm{Ann}(f_k) \supseteq (\langle x_{3k}+1\rangle + \langle x_{3k+2}+1\rangle + \langle x_{3k+3}+1\rangle) \cdot (\langle x_1+1\rangle + \cdots + \langle x_k+1\rangle) \\ \cdot (\langle x_{k+1}+1\rangle + \cdots + \langle x_{2k}+1\rangle) \cdot (\langle x_{2k+1}+1\rangle + \cdots + \langle x_{3k}+1\rangle).\end{aligned}$$

应用性质 (ii) 和 (iii)，于是得到 $R(\text{湮灭子}(f_k)) \geq \frac{7}{8} \cdot \left(1 - \frac{1}{2^k}\right)^3$。

注意，在 $f_k + 1$的展开式中，只有一项含有 $x_{3k+1}x_{3k+2}x_{3k+3}$，即 $x_{3k+1}x_{3k+2}x_{3k+3}$ 本身。因此由定理 1.1， $R(\text{湮灭子}(f_k)) \leq 7/8$。因此这是一个码率为 $7/8$的族。

□

注。这种构造的主要思路是利用长项消去 $(x_{3k+1}+1)(x_{3k+2}+1)(x_{3k+3}+1)$ 中的偶次单项式，使剩下的项全都是奇数次，因此其中任意三项的乘积都不可能为 1。我们还可以进一步推进这一思路。

定理 2.5（广义汉明族）。设 r, k 为正整数，且 $m = 2^{r-1} - 1$。那么定义 $\mathcal{C}_{r,\ k} = \text{湮灭子}(f_{r,\ k}) \subseteq P_{mk+r}$，其中

$$f_{r,k} = (x_{mk+1}+1)(x_{mk+2}+1)\cdots(x_{mk+r}+1) + \sum_{i=0}^{m-1} h_i \prod_{j=1}^{k}(x_{ik+j}+1),$$

and

$$\{h_0, h_1, \ldots, h_{m-1}\} = \left\{ x_{mk+1}^{s_1} x_{mk+2}^{s_2} \cdots x_{mk+r}^{s_r} \mid (0,0,\ldots,0) \neq (s_1, s_2, \ldots, s_r) \in \{0,1\}^r \text{ and } \sum_{i=1}^{r} s_i \text{ is even} \right\}.$$

这是一个定义在无三角形图上的存储码族，当 k 趋于无穷时，其码率为 $1 - \frac{1}{2^r}$。

证明。采用与定理 2.4 证明中类似的论证，可以证明该凯莱图是无三角形的。再由定理 1.1，项 $y_1 y_2 \cdots y_r$ 的存在意味着 $R(\mathcal{C}_{r,\ k}) \leq 1 - \frac{1}{2^r}$。另一方面，包含关系

$$\mathrm{Ann}(f_{r,k}) \supseteq (\langle x_{mk+1}+1\rangle + \langle x_{mk+2}+1\rangle + \cdots + \langle x_{mk+r}+1\rangle) \cdot \prod_{i=0}^{m-1} \sum_{j=1}^{k} \langle x_{ik+j}+1\rangle$$

意味着 $R(\text{湮灭子}(f_{r,\ k})) \geq (1 - \frac{1}{2^r}) \cdot \left(1 - \frac{1}{2^k}\right)^m$。。随着 k 趋于无穷， $R(\mathcal{C}_{r,\ k})$ 趋于 $1 - \frac{1}{2^r}$。因此，只要选择足够大的 r 和 k，该码族的码率就可以任意接近 1。 □

所构造图的稀疏性

考虑一族简单连通图 $\Gamma(V, E)$，其中 V 是顶点集， E 是边集。设 $N = |V|$。假设 $|E|$ 是 N的函数。注意，由于 Γ 是连通的， $|E| \geq N - 1$；又由于 Γ是简单的， $|E| \leq N(N-1)/2$；因此 $|E| = O(N^t)$，其中 $1 \leq t \leq 2$。

我们希望求出定理 2.5 中构造的凯莱图的边数。设 r 为一个固定的正整数。令 $n = r + k(2^{r-1} - 1)$，并令 S 为对应于 $f_{r,k} + 1$ 的 $\mathbb{F}_2^n$ 子集。那么，存储码 $\mathcal{C}$ 的环境图为 $\Gamma = \mathsf{Cay}(\mathbb{F}_2^n, S)$。因此，顶点数为 $N = 2^n = 2^r \times 2^{k(2^{r-1}-1)}$，于是

$$k = \frac{\log_2 N - r}{2^{r-1} - 1}. \tag{2}$$

注意，Γ 是一个正则图，并且每个顶点的度为 $|S| = 2^{r-1} + (2^{r-1} - 1)2^k$。利用 (2)，我们得到

$$\begin{aligned} |S| &= 2^{r-1} + (2^{r-1} - 1)2^{\frac{\log_2 N - r}{2^{r-1}-1}} \\ &= 2^{r-1} + (2^{r-1} - 1)\left(\frac{N}{2^r}\right)^{\frac{1}{2^{r-1}-1}}. \end{aligned}$$

设 E 为 Γ的边集。那么 E 的大小为

$$\begin{aligned} |E| = \frac{N|S|}{2} &= \frac{N\left(2^{r-1} + (2^{r-1} - 1)\left(\frac{N}{2^r}\right)^{\frac{1}{2^{r-1}-1}}\right)}{2} \\ &\approx \frac{2^{r-1} - 1}{2^{1-r/(2^{r-1}-1)}} N^{1+\frac{1}{2^{r-1}-1}} = O(N^{1+\frac{1}{2^{r-1}-1}}). \end{aligned}$$

因此，$|E|$的大小为 $N^{1+\frac{1}{2^{r-1}-1}}$，其指数 $1 + \frac{1}{2^{r-1}-1}$在 r趋于无穷大时趋近于 1。因此，该图族实际上相当稀疏。

3 一些经典理想的码率

回顾一下，在第 2 节中，我们构造了一个存储码族，其码率主要利用性质 (ii) 和 (iii)计算得出。本节将证明这些性质。

按照前面的约定，当环和理想都视为向量空间时，理想的码率是理想的维数与其环境环的维数之比。为了标明我们使用的变量，通常用 $P[x_1, \ldots, x_n]$ 作为$P_n = \mathbb{F}_2[x_1, x_2, \ldots, x_n]/\langle x_1^2 - 1$、$x_2^2 - 1$、.....、 $x_n^2 - 1\rangle$的缩写。

注意， P_n 是一个多项式商环。因此， P_n 中的每个元素都可以表示为关于某些变量的多项式。设有两个元素 f、 $g \in P_n$。若 $f = f(x_1, \ldots, x_k)$且 $g = g(x_{k+1}, \ldots, x_n)$，则称 f和 g是关于互不相交的变量的。

引理 3.1. 将 P_n 视为 $\mathbb{F}_2$上的向量空间，我们有 P_n的两个基:

- $B_1 = \{x_1^{s_1} \cdots x_n^{s_n} \mid s_i \in \{0, 1\}\}$;
- $B_2 = \{(x_1 + 1)^{s_1} \cdots (x_n + 1)^{s_n} \mid s_i \in \{0, 1\}\}$.

证明。在群代数 P_n中，第一个基是显然的，因为它们恰好就是群元素。注意， B_2 与 B_1的大小相同。因此，只需证明 B_1 中的每个元素都可以由 B_2线性表示。例如， $x_1 = (x_1 + 1) + 1$, $x_1x_2 = (x_1 + 1)(x_2 + 1) + (x_1 + 1) + (x_2 + 1) + 1$。 □

引理 3.2. 如果 $f \neq 0$ 和 $g \neq 0$ 是互不相交的变量，那么 $fg \neq 0$。

证明. 假设环境环为 $P[x_1,\ldots,x_n]$，且

$$f=\sum a_{s_1,\ldots,s_k}x_1^{s_1}\cdots x_k^{s_k},g=\sum b_{s_{k+1},\ldots,s_n}x_{k+1}^{s_{k+1}}\cdots x_n^{s_n},$$

其中系数 $a_{s_1,\ldots,s_k},b_{s_{k+1},\ldots,s_n}\in\mathbb{F}_2$ 并非全为零。必定存在某个乘积 $a_{s_1,\ldots,s_k}b_{s_{k+1},\ldots,s_n}\neq 0$。因此 $fg=\sum a_{s_1,\ldots,s_k}b_{s_{k+1},\ldots,s_n}x_1^{s_1}\cdots x_k^{s_k}x_{k+1}^{s_{k+1}}\cdots x_n^{s_n}\cdots x_n^{s_n}\neq 0$，因为 $\{x_1^{s_1}\cdots x_n^{s_n}\mid s_i\in\{0,1\}\}$ 在 $\mathbb{F}_2$ 上线性无关。

□

引理 3.3。对于所有 $i=1,\ldots,n$，$\mathrm{Ann}(x_i+1)=\langle x_i+1\rangle$。

证明。我们只考虑 $i=1$ 的情况。其他情况也可用相同方法处理。包含关系“⊇”显然成立。对于 $f\in P_n$，我们可以写成 $f=\sum a_{s_1,\ldots,s_n}(x_1+1)^{s_1}\cdots(x_n+1)^{s_n}$，由引理 3.1。若 $f\in\mathrm{Ann}(x_1+1)$，则

$$\begin{aligned}(x_1+1)f&=(x_1+1)\sum a_{s_1,\ldots,s_n}(x_1+1)^{s_1}(x_2+1)^{s_2}\cdots(x_n+1)^{s_n}\\&=(x_1+1)\sum_{s_1=0} a_{s_1,\ldots,s_n}(x_1+1)^{s_1}(x_2+1)^{s_2}\cdots(x_n+1)^{s_n}\\&=(x_1+1)\sum a_{0,s_2,\ldots,s_n}(x_2+1)^{s_2}\cdots(x_n+1)^{s_n}=0.\end{aligned}$$

由引理 3.2 可知 $\sum a_{0,\ s_2,\ \ldots,\ s_n}(x_2+1)^{s_2}\cdots(x_n+1)^{s_n}=0$。由引理 3.1，每个系数 $a_{0,\ s_2,\ \ldots,\ s_n}=0$。因此 $f=\sum a_{1,\ s_2,\ \ldots,\ s_n}(x_1+1)(x_2+1)^{s_2}\cdots(x_n+1)^{s_n}$，从而得到 $f\in\langle x_1+1\rangle$。

□

命题 3.4. 理想 $\langle x_1+1\rangle\langle x_2+1\rangle\cdots\langle x_k+1\rangle$ 的码率为 $1/2^k$，其中 $1\leq k\leq n$。
证明。设 $f\in P_n$，即 $f=\sum a_{s_1,\ \ldots,\ s_n}(x_1+1)^{s_1}\cdots(x_n+1)^{s_n}$. 若 $f\in\langle x_1+1\rangle\langle x_2+1\rangle\cdots\langle x_k+1\rangle$，则 $f\in\mathrm{Ann}(x_i+1)$，因此 $a_{s_1,\ \ldots,\ s_{i-1},\ 0,\ s_{i+1},\ \ldots,\ s_n}=0$，如引理 3.3 关于 $i=1$、…、k 的证明所示。于是 $f=\sum a_{1,\ \ldots,\ 1,\ s_{k+1},\ \ldots,\ s_n}(x_1+1)\cdots(x_k+1)(x_{k+1}+1)^{s_{k+1}}\cdots(x_n+1)^{s_n}$。另一方面，每个这样的 f 都属于 $\langle x_1+1\rangle\langle x_2+1\rangle\cdots\langle x_k+1\rangle$。因此该理想具有如下基

$$\{(x_1+1)\cdots(x_k+1)(x_{k+1}+1)^{s_{k+1}}\cdots(x_n+1)^{s_n}\mid s_i\in\{0,1\}\}.$$

该理想的维数为 2^{n-k}，因此码率为 $2^{n-k}/2^n=1/2^k$。

□

命题 3.5. $\langle x_1+1\rangle\cdots\langle x_k+1\rangle\cap\langle x_{k+1}+1\rangle=\langle x_1+1\rangle\cdots\langle x_k+1\rangle\langle x_{k+1}+1\rangle$，对于 $1\leq k\leq n-1$。

证明. 包含关系“⊇”显然。设 $f\in\langle x_1+1\rangle\cdots\langle x_k+1\rangle\cap\langle x_{k+1}+1\rangle$。则我们可以写成 $f=(x_1+1)\cdots(x_k+1)g$，其中 $g\in P_n$。利用这一事实 $(x_i+1)x_i=x_i+1$，可得 $(x_1+1)\cdots(x_k+1)g=(x_1+1)\cdots(x_k+1)g(x_1=1,\cdots,x_k=1)$，因此不失一般性，我们可以假设 g 只涉及变量 $x_{k+1},\ldots,x_n$。由于 $f\in\langle x_{k+1}+1\rangle$，我们有 $(x_{k+1}+1)f=(x_1+1)\cdots(x_k+1)(x_{k+1}+1)g=0$。因此，根据引理 3.2 $(x_{k+1}+1)g=0$，并由引理 3.3 $g\in\langle x_{k+1}+1\rangle$。证毕。 □

定理 3.6。理想 $\langle x_1+1\rangle+\langle x_2+1\rangle+\cdots+\langle x_k+1\rangle$ 的码率为 $1-1/2^k$，对于 $1\leq k\leq n$。

证明。回顾一下，对于 P_n的两个向量子空间 V和 W，我们有 $\dim(V+W) = \dim V + \dim W - \dim V \cap W$。应用此公式，我们得到

$$\begin{aligned}\dim \sum_{i=1}^{k} \langle x_i + 1\rangle &= \sum_{s=1}^{k} (-1)^{s+1} \sum_{i_1<\cdots<i_s} \dim\langle x_{i_1}+1\rangle \cap \cdots \cap \langle x_{i_s}+1\rangle \\ &= \sum_{s=1}^{k} (-1)^{s+1} \sum_{i_1<\cdots<i_s} \dim\langle x_{i_1}+1\rangle \cdots \langle x_{i_s}+1\rangle,\end{aligned}$$

其中最后一个等式用到了命题3.5 。将两边除以环境环的维数，并应用命题3.4，得到

$$R\left(\sum_{i=1}^{k}\langle x_i+1\rangle\right) = \sum_{s=1}^{k}(-1)^{s+1}\binom{k}{s}\frac{1}{2^s} = 1 - \frac{1}{2^k}.$$

□

引理 3.7. 设 $\{f_i(x_1,\ldots,x_r)\}_{i=1,\ldots,k}, \{g_j(y_1,\ldots,y_s)\}_{j=1,\ldots,\ell}$ 是 $P[x_1,\ldots,x_r,y_1,\ldots,y_s]$ 中由互不相交的变量构成的两组元素。若这两组元素均线性无关，则集合 $\{f_i g_j \mid i=1,\ldots,k, j=1,\ldots,\ell\}$ 线性无关。

证明。为方便起见，我们记

$$\{v_1,\ldots,v_n\} = \left\{x_1^{t_1}\cdots x_r^{t_r} \mid t_1,\ldots,t_r \in \{0,1\}\right\},$$

$$\{w_1,\ldots,w_m\} = \left\{y_1^{t_1}\cdots y_s^{t_s} \mid t_1,\ldots,t_s \in \{0,1\}\right\}.$$

于是，对于 $i=1,\ldots,k$ 和 $j=1,\ldots,\ell$，有 $f_i = \sum_{c=1}^{n} a_{ic}v_c$ 和 $g_j = \sum_{e=1}^{m} b_{je}w_e$，其中 a_{ic}，$b_{je} \in \mathbb{F}_2$。设存在某个 $\lambda_{ij} \in \mathbb{F}_2$，使得零线性组合 $\sum_{i,j}\lambda_{ij}f_ig_j = 0$ 成立。那么

$$0 = \sum_{i,j}\lambda_{ij}f_ig_j = \sum_{i,j}\lambda_{ij}\sum_{c=1}^{n}a_{ic}v_c\sum_{e=1}^{m}b_{je}w_e = \sum_{c,e}\sum_{i,j}\lambda_{ij}a_{ic}b_{je}\cdot v_cw_e.$$

No注意到 $\{v_cw_e \mid c=1,\ldots,n, e=1,\ldots,m\}$ 是引理 3.1中的第一个基，因此我们有

$$\sum_{i,j}\lambda_{ij}a_{ic}b_{je} = \sum_{i}a_{ic}\sum_{j}\lambda_{ij}b_{je} = 0,\ c=1,2,\ldots,n,\ e=1,2,\ldots,m.$$

由于 $\{f_i\}$ 线性无关，系数向量 $\{[a_{i1}\ a_{i2}\ \cdots\ a_{in}] \mid i=1,\ldots,k\}$ 也线性无关。因此，对于所有 i 和 e，都有 $\sum_j \lambda_{ij}b_{je} = 0$；又由于 $\{g_j\}$ 线性无关，这进一步意味着对于所有 i 和 j，都有 $\lambda_{ij} = 0$。证毕。□

定理 3.8. 设 $\{f_i(x_1,\ldots,x_r)\}_{i=1,\ldots,k}, \{g_j(y_1,\ldots,y_s)\}_{j=1,\ldots,\ell}$ 是 $P[x_1,\ldots,x_r,y_1,\ldots,y_s]$ 中两个定义在不相交变量上的元素集。那么我们有率公式 $R(\alpha\beta) = R(\alpha)R(\beta)$，其中$\alpha = \langle f_1,\ldots,f_k\rangle$ 和 $\beta = \langle g_1,\ldots,g_\ell\rangle$。

证明。 设 $V = \alpha \cap P[x_1, \ldots, x_r]$ 和 $W = \beta \cap P[y_1, \ldots, y_s]$。那么 V和 W是向量空间，并且我们假设 $\{v_1, v_2, \ldots, v_n\}$ 和 $\{w_1, w_2, \ldots, w_m\}$ 分别是它们的基。定义张量积 $V \otimes W$为 $\mathbb{F}_2$-张成的子空间 $\{v_c w_e \mid c = 1, \ldots, n, e = 1, \ldots, m\}$。根据引理 3.7 $\{v_c w_e\}$线性无关，因此 $\dim V \otimes W = \dim V \times \dim W = n \times m$。

我们断言 $\alpha\beta = V \otimes W$。包含关系“$\supseteq$”显然成立。令 $h \in \alpha\beta$，即 $h = \sum_{ij} d_{ij} \cdot f_i g_j$，其中 $d_{ij} \in P_{r+s}$。对于每个 d_{ij}，我们可以写成 $d_{ij} = \sum_k a_{ijk} b_{ijk}$，其中 $a_{ijk} \in P[x_1, \ldots, x_r], b_{ijk} \in P[y_1, \cdots, y_s]$。因此

$$h = \sum_{ij} \sum_{k} a_{ijk} b_{ijk} \cdot f_i g_j = \sum_{k} \sum_{ij} a_{ijk} f_i \cdot b_{ijk} g_j.$$

注意到 $a_{ijk} f_i \in V$和 $b_{ijk} g_j \in W$。于是 $a_{ijk} f_i \cdot b_{ijk} g_j \in V \otimes W$也成立， h亦然，因此 $\alpha\beta = V \otimes W$。故 $\dim \alpha\beta = n \times m$。

同理，我们有 $\alpha = V \otimes P[y_1, \ldots, y_s]$，从而 $\dim\ \alpha = \dim\ V \times \dim\ P[y_1, \ldots, y_s] = n \times 2^s$，得到 $R(\alpha) = \dim \alpha / 2^{r+s} = n/2^r$。类似地， $R(\beta) = m/2^s$。最后，我们得到

$$R(\alpha\beta) = \frac{\dim \alpha\beta}{2^{r+s}} = \frac{n \times m}{2^r \times 2^s} = R(\alpha) R(\beta).$$

□

推论 3.9. $R\left(\prod_{i=1}^{n} \sum_{j=1}^{k_i} \langle x_{ij} + 1 \rangle\right) = \prod_{i=1}^{n} R\left(\sum_{j=1}^{k_i} \langle x_{ij} + 1 \rangle\right)$。

证明。 这由定理3.8 通过归纳法得出。 □

致谢。在我们完成本文写作之际，我们得知了 [7], ，其中作者构造了一族码率渐近为 1 的存储码。由于本文中的构造和 [7] 中的构造都是对汉明族的推广，它们看起来很相似。然而，用于计算存储码码率的方法截然不同。

参考文献

[1] Mazumdar, A.《可修复网络的存储容量》。IEEE 信息论汇刊 61 (2015)，第 11 期，5810-5821。[2] Mazumdar, A., McGregor, A. 和 Vorotnikova, S.《作为信息论顶点覆盖的存储容量与索引编码率》，IEEE 信息论汇刊 65 (2019)，第 9 期，5580–5591。[3] Shanmugam, K. 和 Dimakis, A. G.《通过索引编码和局部可修复码界定多重单播》。2014 年 IEEE 国际信息论研讨会 (2014)：296-300。[4] Cameron, P. J., Dang, A. N. 和 Riis, S.《无三角形图上的猜测游戏》。电子组合学杂志 23 (2014)。

[5] Christofides, D. 和 Markstr¨om, K.《无向图的猜测数》。电子组合学杂志 18 (2011)。[6] Barg, A. 和 Z´emor, G.《无三角形图上的高码率存储码》。IEEE 信息论汇刊 68 (2021)，第 12 期，7787–7797。[7] Barg, A., Schwartz 和 Yohananov, L.《渐近单位码率的无三角形图存储码》。arXiv:2212.12117v1数学系和深圳国家应用数学中心，南方科技大学，中国深圳 518055。电子邮箱：hhxiang1999@foxmail.com 数学系和深圳国际数学中心，南方科技大学，中国深圳 518055。电子邮箱：xiangq@sustech.edu.cn