构造小域上的最优局部修复码

献给万哲先教授 90 华诞

张志芳*, 徐敬可, 刘木兰

中国科学院数学与系统科学研究院, 北京 100190

E-mail: zfz@amss.ac.cn, xujingke14@mails.ucas.ac.cn, mlliu@amss.ac.cn

收稿日期: 2016-11-03; 接受日期: 2017-01-18; 网络出版日期: 2017-03-16; * 通信作者

摘要 局部修复码是近几年分布式存储编码领域一个非常热门的研究方向. 满足局部修复性 r 即要求码字的一位能够被其他至多 r 位恢复. 这种性质对于提高分布式存储系统中失效节点的修复效率非常重要. 本文主要考虑了域规模小于 n 时最优局部修复码的构造. 具体地, 本文构造了有限域 $\mathbb{F}_q$ ($q = \frac{r}{r+1}n + 1$) 上满足所有位局部修复性 $r = 2$ 、 $d = 6$ 和 $d = r + 1$ 的最优局部修复码.

关键词 局部修复码 最优极小距离 分布式存储系统

MSC (2010) 主题分类 11T71, 68P20

1 引言

在大数据时代, 海量数据的存储离不开分布式存储系统. 分布式存储就是将原始数据经过编码后分散地存储在多个节点上, 使得只用连接其中部分节点就可以获取所需要的数据. 与传统的集中式存储方式相比, 分布式存储在降低存储负荷、提高存储可靠性和疏散访问交通等方面具有明显的优势. 根据实际需求选用特定的编码是分布式存储的一项关键技术, 其中用到的局部修复码 (locally repairable code) 是近几年非常热门的一个研究方向. 它最早在文献 [1] 中正式提出并得到一些经典的结果, 随后大批学者围绕局部修复码展开了深入和系统的研究.

局部修复码的定义可以有以下简单叙述: 设 $\mathcal{C}$ 是有限域 $\mathbb{F}_q$ 上一个 $[n, k]$ 线性码, 如果 $\mathcal{C}$ 的第 i 位是其他至多 r 位在 $\mathbb{F}_q$ 上的一个线性组合, 则称 $\mathcal{C}$ 的第 i 位具有局部性 r . 它的一个实际背景就是在分布式存储系统中, 一个失效的存储节点可以通过连接其他至多 r 个节点来实现数据恢复. 在实际使用中, r 越小则节点修复的效率越高. 特别地, 如果码 $\mathcal{C}$ 的每一位都具有局部性 r , 则称为所有位局部修复码; 如果弱化一点, 只要求 k 个信息位都具有局部性 r , 则称为信息位局部修复码. 文献 [1] 的一个经典结果是, 证明了对于信息位局部修复码而言, 它的极小距离 d 满足

$$d \leq n - k + 1 - \left(\left\lceil \frac{k}{r} \right\rceil - 1 \right), \quad (1.1)$$

这个界很像编码理论里经典的 Singleton 界. 通过 Pyramid 码 [2] 可以构造出一般参数下达到界 (1.1)

英文引用格式: Zhang Z F, Xu J K, Liu M L. Constructions of optimal locally repairable codes over small fields (in Chinese). Sci Sin Math, 2017, 47: 1607–1614, doi: 10.1360/N012016-00206

的信息位局部修复码, 而所有位局部修复码虽然也满足界 (1.1), 但却不是处处可以达到的. 由于所有位的局部修复性较信息位局部修复性在实际使用中更加方便, 本文只关注前者, 所提到的局部性如不作特殊说明都认为是所有位局部修复性.

最早, Gopalan 等^[1]指出, 当 $(r+1) \nmid n$ 并且 $r \mid k$ 时, 所有位局部修复码达不到界 (1.1), 而对于可以达到界 (1.1) 的情形, 他们只给出了 $(r+1) \mid n$ 时在充分大的有限域上的存在性结果. 随后, 文献 [3,4] 中界 (1.1) 分别被推广到了向量码和非线性码的情形, 对于界的可达性条件并没有给出一般的构造. 文献 [5] 得到了一个与有限域的大小 q 有关的极小距离上界, 由于依赖于编码理论中尚未确定的参数, 目前只能在个别 q 的取值下给出较界 (1.1) 更紧的上界. 最近, 文献 [6] 在一定参数范围内得到了极小距离紧的上界, 这是对界 (1.1) 的一个改进结果. 但是, 界 (1.1) 目前仍然是局部修复码里最受关注的一个经典的界.

作为刻画最优极小距离的一个方面, 不少工作专注于在可达性条件下如何构造达到界 (1.1) 的局部修复码, 这对于实际应用也具有相当重要的意义. 为简单起见, 我们称达到界 (1.1) 的局部修复码为最优局部修复码. 文献 [7,8] 首先分别用 Reed-Solomon 码和 Gabidulin 码构造了最优的局部修复码, 他们用到的有限域大小关于码长 n 是指数的. 2014 年, Tamo 和 Barg^[9] 利用特殊多项式插值, 在与 n 差不多大的有限域上构造了最优的局部修复码, 他们的这一工作获得了 2015 年 IEEE 信息论领域的最佳论文奖. 这些最优的构造都是在参数条件 $(r+1) \mid n$ 时给出的, 其中文献 [8] 中的构造在满足 $n \bmod (r+1) > k \bmod r > 0$ 的条件下也是最优的. 在一些更小的有限域 (如二元域上) 也有一些局部修复码的构造^[10-12], 不过一般没有达到界 (1.1), 而是达到文献 [5] 中给出的限定在二元域上的极小距离上界.

对于 Singleton 界, MDS 猜想^[13]指出, 当 $k \leq q$ 时, 有限域 $\mathbb{F}_q$ 上达到 Singleton 界的 $[n, k]$ 线性码的码长至多是 $q+2$. 那么, 对于界 (1.1), 是否也有类似的码长与有限域大小之间的制约关系? 由于目前已有在与 n 差不多大的有限域上最优局部修复码的构造, 因此, 考虑能否在更小的有限域上一般地构造最优的局部修复码是一个很有意义的问题. 文献 [14] 在有限域 $\mathbb{F}_q$ ($q \geq r+1$) 上构造了一个极小距离为 4 的最优局部修复码. 能否构造其他较一般参数的小域上的最优局部修复码是本文考虑的主要问题. 注意, 这里所称的“小域”是相对于文献 [9] 中的构造而言, 即我们要求 $q < n$; “最优”是指达到界 (1.1) 的码.

本文的结构如下: 第 2 节给出特殊参数 $n = 3l$, $k = 2l - 3$ 和 $r = 2$ 下局部修复码的构造, 极小距离 $d = 6$ 达到最优, 用到的有限域大小为 $q = 2l + 1$, 其中 l 是使得 $2l + 1$ 为某个素数幂的一个正整数. 第 3 节将这一构造进行推广, 得到了较为一般的参数条件 $(r+1) \mid n$ 并且 $d = r+1$ 时最优局部修复码的构造, 用到的有限域大小为 $\frac{r}{r+1}n + 1$. 第 4 节是本文的结论.

2 例子: $d = 6$

首先介绍一下常用记号, 对于正整数 n , 记 $[n] = \{1, \dots, n\}$. 对于向量 $\mathbf{c} = (c_1, \dots, c_n) \in \mathbb{F}_q^n$, 定义

$$\text{supp}(\mathbf{c}) = \{i \in [n] \mid c_i \neq 0\}.$$

令 $\text{Mat}_q(m, n)$ 表示 $\mathbb{F}_q$ 上 $m \times n$ 阶矩阵的全体.

设 $\mathcal{C}$ 是一个 $\mathbb{F}_q$ 上的 $[n, k]$ 线性码, 易知 $\mathcal{C}$ 的第 i ($1 \leq i \leq n$) 位具有局部修复性 r , 当且仅当 $\mathcal{C}$ 的对偶码存在码字 $\mathbf{c}$ 满足 $i \in \text{supp}(\mathbf{c})$ 并且 $|\text{supp}(\mathbf{c})| \leq r+1$. 根据这一事实, 我们从校验矩阵的角度构造局部修复码.

设 $q = 2l + 1$ 是某个素数的方幂, 其中 $l \geq 3$. 设 β 是 $\mathbb{F}_q$ 的一个本原元. 令

$$H = \begin{pmatrix} 1 & & x_1 & 1 & & & & \\ & 1 & & & x_2 & 1 & & \\ & & \ddots & & & & \ddots & \\ & & & 1 & & & x_l & 1 \\ & & & & 1 & \beta & \beta^2 & \beta^3 \dots \beta^{2l-2} & \beta^{2l-1} \\ & & & & 1 & \beta^2 & \beta^4 & \beta^6 \dots \beta^{4l-4} & \beta^{4l-2} \\ & & & & 1 & \beta^3 & \beta^6 & \beta^9 \dots \beta^{6l-6} & \beta^{6l-3} \end{pmatrix}, \quad (2.1)$$

其中

$$x_j \in \mathbb{F}_q^*, \quad x_j \neq \frac{(\beta^{2i-2}, \beta^{2i-1}, \beta^{2j-2})}{(\beta^{2i-2}, \beta^{2i-1}, \beta^{2j-1})}, \quad 1 \leq i \neq j \leq l, \quad (2.2)$$

这里 $(a, b, c) \triangleq abc(c-b)(c-a)(b-a)$, 矩阵中的空白位置都表示 0. 易见, $H \in \text{Mat}_q(l+3, 3l)$. 条件 (2.2) 说明 x_j 在 F^* 中至多排除了 $l-1$ 个取值, 由于 $q-1 = 2l > l-1$, 因此, 这样的 x_j 总是可以在 F^* 中取到, 即我们定义的 H 总是存在的. 另外, 在条件 (2.2) 中, x_j 的取值只是从 $\mathbb{F}_q^*$ 中至多排除了 $l-1$ 个值, 由于 $q-1 = 2l > l-1$, 因此, x_j ($1 \leq j \leq l$) 都可以在 $\mathbb{F}_q^*$ 中取到.

定理 1 以 (2.1) 中的矩阵 H 作为检验阵, 生成 $\mathbb{F}_q$ 上的一个 $[n = 3l, k = 2l - 3, d = 6, r = 2]$ 的局部修复码, 且极小距离最优.

证明 码长 $n = 3l$ 是显然的. 矩阵 H 的前 l 行说明了该码的所有位满足局部修复性 $r = 2$. H 的后 3 行的非零部分是一个 Vandermonde 矩阵, 显然线性无关, 再有 H 的左上部分是一个 l 阶单位矩阵, 从而容易看出 H 是行满秩的, 即作为检验阵生成的线性码维数 $k = 3l - (l + 3) = 2l - 3$.

下面证明极小距离 $d = 6$. 首先, 我们说明 H 的任意 5 列是线性无关的. 由于 Vandermonde 矩阵的性质, H 右边的 $2l$ 列中的任意 3 列是线性无关的, 并且与左边的 l 列都是线性无关的, 所以只需要考虑从 H 右边的 $2l$ 列中任取超过 3 列的情形, 即以下两种情形:

(1) 从 H 右边的 $2l$ 列中任取 5 列. 根据 H 的前 l 行的特点可知, 此时可能出现线性相关的只有如下情形:

$$M_{i,j} = \begin{pmatrix} x_i & 1 & & & \\ & & x_j & 1 & \\ \beta^{2i-2} & \beta^{2i-1} & \beta^{2j-2} & \beta^{2j-1} & \\ \beta^{4i-4} & \beta^{4i-2} & \beta^{4j-4} & \beta^{4j-2} & \\ \beta^{6i-6} & \beta^{6i-3} & \beta^{6j-6} & \beta^{6j-3} & \end{pmatrix}, \quad 1 \leq i \neq j \leq l,$$

这里的 $M_{i,j}$ 是其中的 4 列去掉 0 位置得到的子矩阵. 我们将说明 $M_{i,j}$ 去掉第一行后得到一个 4×4 的可逆子阵, 从而, $M_{i,j}$ 的列秩为 4. 事实上, 由 x_j 的取法可知, 所得子阵的行列式

$$x_j(\beta^{2i-2}, \beta^{2i-1}, \beta^{2j-1}) - (\beta^{2i-2}, \beta^{2i-1}, \beta^{2j-2}) \neq 0,$$

因此, 这样的 4 列线性无关, 从而, 在该情形下的任意 5 列是线性无关的.

(2) 从 H 右边的 $2l$ 列中任取 4 列, 再从 H 的前 l 列里任取 1 列. 注意到此时所取 5 列构成的子阵中, 若存在两行的 Hamming 重量均为 1, 则这 5 列线性无关. 因此可能线性相关的情形, 删除所有的 0 行, 得到子矩阵

$$M_{i,j} = \begin{pmatrix} 1 & x_i & 1 & & \\ & & & x_j & 1 \\ \beta^{2i-2} & \beta^{2i-1} & \beta^{2j-2} & \beta^{2j-1} & \\ \beta^{4i-4} & \beta^{4i-2} & \beta^{4j-4} & \beta^{4j-2} & \\ \beta^{6i-6} & \beta^{6i-3} & \beta^{6j-6} & \beta^{6j-3} & \end{pmatrix}.$$

同理可证后 4 行线性无关, 第一行显然与后 4 行线性无关.

从 H 的任意 5 列线性无关可知, 码的极小距离 $d > 5$. 另一方面, 由界 (1.1) 知, $d \leq n - k + 2 - \lceil \frac{k}{r} \rceil = 3l - (2l - 3) + 2 - \lceil \frac{2l-3}{2} \rceil = 6$, 因此, 码的极小距离 $d = 6$ 并且达到界 (1.1), 是最优的. $\square$

下面给出这个构造的两个具体例子.

例 1 设 $l = 3, q = 7, 3$ 是 $\mathbb{F}_7$ 的一个本原元. 令

$$H = \begin{pmatrix} 1 & & & 4 & 1 & & & \\ & 1 & & & & 4 & 1 & \\ & & 1 & & & & 4 & 1 \\ & & & 1 & 3 & 2 & 6 & 4 & 5 \\ & & & & 1 & 2 & 4 & 1 & 2 & 4 \\ & & & & & 1 & 6 & 1 & 6 & 1 & 6 \end{pmatrix}.$$

为码 $\mathcal{C}$ 的校验阵, 则 $\mathcal{C}$ 是 $[9, 3, 6]$ 的线性码, 满足所有位局部修复性 $r = 2$. 可以验证

$$\begin{aligned} x_1 &= 4 \notin \left\{ 0, \frac{(2, 6, 1)}{(2, 6, 3)} = 3, \frac{(4, 5, 1)}{(4, 5, 3)} = 1 \right\}, \\ x_2 &= 4 \notin \left\{ 0, \frac{(1, 3, 2)}{(1, 3, 6)} = 6, \frac{(4, 5, 2)}{(4, 5, 6)} = 3 \right\}, \\ x_3 &= 4 \notin \left\{ 0, \frac{(1, 3, 4)}{(1, 3, 5)} = 3, \frac{(2, 6, 4)}{(2, 6, 5)} = 6 \right\}. \end{aligned}$$

并且 $d = 6$ 达到界 (1.1), 因而是最优的.

例 2 设 $l = 4, q = 9 = 3^2, \beta \in \mathbb{F}_9$ 是一个本原元, 满足 $\beta^2 = \beta + 1$. 令

$$H = \begin{pmatrix} 1 & & & & 1 & 1 & & & \\ & 1 & & & & & 1 & 1 & \\ & & 1 & & & & & 1 & 1 \\ & & & 1 & & & & & 1 & 1 \\ & & & & 1 & \beta & \beta^2 & \beta^3 & 2 & 2\beta & 2\beta^2 & 2\beta^3 \\ & & & & & 1 & \beta^2 & \beta^4 & \beta^6 & 1 & \beta^2 & 2 \\ & & & & & & 1 & \beta^3 & \beta^6 & \beta & 2 & 2\beta^3 & \beta^2 & 2\beta \end{pmatrix}.$$

为码 $\mathcal{C}$ 的校验阵, 则 $\mathcal{C}$ 是 $[12, 5, 6]$ 的线性码, 满足所有位局部修复性 $r = 2$. 可以验证

$$\begin{aligned} x_1 = 1 &\notin \left\{ 0, \frac{(\beta^2, \beta^3, 1)}{(\beta^2, \beta^3, \beta)} = \beta^3, \frac{(\beta^4, \beta^5, 1)}{(\beta^4, \beta^5, \beta)} = 2\beta^2, \frac{(2\beta^2, 2\beta^3, 1)}{(2\beta^2, 2\beta^3, \beta)} = \beta \right\}, \\ x_2 = 1 &\notin \left\{ 0, \frac{(1, \beta, \beta^2)}{(1, \beta, \beta^3)} = \beta, \frac{(\beta^4, \beta^5, \beta^2)}{(\beta^4, \beta^5, \beta^3)} = \beta^3, \frac{(2\beta^2, 2\beta^3, \beta^2)}{(2\beta^2, 2\beta^3, \beta^3)} = \beta^3 \right\}, \\ x_3 = 1 &\notin \left\{ 0, \frac{(1, \beta, \beta^4)}{(1, \beta, \beta^5)} = \beta^7, \frac{(\beta^2, \beta^3, \beta^4)}{(\beta^2, \beta^3, \beta^5)} = \beta, \frac{(2\beta^2, 2\beta^3, \beta^4)}{(2\beta^2, 2\beta^3, \beta^5)} = \beta^3 \right\}, \\ x_4 = 1 &\notin \left\{ 0, \frac{(1, \beta, \beta^6)}{(1, \beta, \beta^7)} = \beta^3, \frac{(\beta^2, \beta^3, \beta^6)}{(\beta^2, \beta^3, \beta^7)} = \beta, \frac{(\beta^4, \beta^5, \beta^6)}{(\beta^4, \beta^5, \beta^7)} = \beta \right\}. \end{aligned}$$

并且 $d = 6$ 达到界 (1.1), 因而是最优的.

3 一般构造: $d = r + 1$

在上一节的构造中, 对参数一个较强的限制条件是 $r = 2$, 本节将去掉这一限制条件, 构造方法还是从设计码的校验矩阵入手.

设 $q = rl + 1$ 是一个素数幂, 其中 $l \geq 3, r \geq 2$. 设 β 是 $\mathbb{F}_q$ 的一个本原元. 对于 $1 \leq i \leq l$, 记

$$\begin{aligned} H_{i,1} &= \begin{pmatrix} 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} \rightarrow \text{ith} \in \text{Mat}_q(l, r), \\ H_{i,2} &= \begin{pmatrix} \beta_i & & & \\ & \beta_i^2 & & \\ & & \ddots & \\ & & & \beta_i^{(r-1)} \end{pmatrix} \begin{pmatrix} 1 & \beta & \cdots & \beta^{r-1} \\ 1 & \beta^2 & \cdots & \beta^{2r-2} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \beta^{r-1} & \cdots & \beta^{(r-1)^2} \end{pmatrix} \in \text{Mat}_q(r-1, r), \end{aligned}$$

其中

$$\beta_i = \beta^{(i-1)r},$$

$H_{i,1}$ 的第 i 行为全 1 行, 其余行为 0 行. 令

$$H = \begin{pmatrix} I_l & H_{1,1} & H_{2,1} & \cdots & H_{l,1} \\ O & H_{1,2} & H_{2,2} & \cdots & H_{l,2} \end{pmatrix}, \quad (3.1)$$

其中 I_l 表示 l 阶单位矩阵, O 表示一个全零矩阵. 我们有下面的结论.

定理 2 以 (3.1) 中的矩阵 H 作为检验阵, 生成 $\mathbb{F}_q$ 上的一个

$$[n = (r+1)l, k = r(l-1) + 1, d = r+1, r]$$

的局部修复码, 且极小距离最优.

证明 由于 H 的后 $r-1$ 行的非零部分是一个 Vandermonde 矩阵, 所以, 这 $r-1$ 行线性无关, 并且其的中任意 $r-1$ 列也是线性无关的. 容易得到, H 的秩为 $l+r-1$, 因此,

$$\begin{aligned} k &= n - (l + r - 1) \\ &= (r + 1)l - (l + r - 1) \\ &= r(l - 1) + 1. \end{aligned}$$

另外, 所有位满足局部修复性 r 也是显然的.

下面说明 H 的任意 r 列线性无关. 由于 H 的后 $r-1$ 行的 Vandermonde 结构, H 的后 rl 列中的任意 $r-1$ 列都是线性无关的, 并且与前 l 列也是线性无关的, 所以只需考虑从后 rl 列中任取 r 列的情形. 这时, 如果把它们的前 $l-1$ 行都加到第 l 行, 则得到一个全 1 行, 并且与剩下的 $r-1$ 行合起来构成一个 $r \times r$ 的 Vandermonde 矩阵, 因此, 行秩为 r , 即这 r 列线性无关.

至此我们已说明了极小距离 $d \geq r+1$. 接下来, 由界 (1.1) 知,

$$\begin{aligned} d &\leq n - k + 2 - \left\lceil \frac{k}{r} \right\rceil \\ &= l + r - 1 + 2 - \left\lceil \frac{r(l-1) + 1}{r} \right\rceil \\ &= r + 1. \end{aligned}$$

因此, 极小距离达到界 (1.1), 是最优的. $\square$

可以看到, 定理 2 的证明比定理 1 的证明反而要简单一些, 这是因为我们在后者的构造中作了一定的简化, 即将矩阵 $H_{i,1}$ ($1 \leq i \leq l$) 的非零位置全部设定为 1, 但是这样带来的一个损失是极小距离只做到了 $r+1$. 如果仿照定理 1 中的构造, 在矩阵 $H_{i,1}$ 的非零位置引入更多 $\mathbb{F}_q^*$ 中的元素, 那么我们可以进一步提高极小距离. 当然, 此时 $H_{i,2}$ 的行数是少于 $r-1$ 行的. 另外, 类似于文献 [14, 15] 中的构造, 可以适当删去检验矩阵 H 后面的一些行, 虽然会导致极小距离减少, 但仍然可以达到界 (1.1), 即也是最优的局部修复码, 并且此时的信息率会比定理 2 中的码高.

下面表 1 给出本文中一般性构造与一些已有结果的比较. 表 1 中, 局部修复码的码长均为

$$n = l(r + 1),$$

表 1 几类最优局部修码的参数比较

	域的大小	维数	极小距离
本文的构造	$q = \frac{n}{r+1}r + 1$	$(l-1)r + 1$	$r + 1$
文献 [7, 构造 1]	$q \geq (\frac{nr}{r+1})^{\frac{nr}{r+1}+2}$	$(l-1)r + 1$	$r + 1$
文献 [9, 构造 1]	$q \geq n$	$(l-1)r + 1$	$r + 1$
文献 [14, 构造 2]	$q \geq r + 1$	$n - l - 2$	4
文献 [14, 构造 3]	$q \geq 2n + 1$	$n - l - 3$	5
文献 [15, 引理 3.2]	$q \geq n$	$(l-1)r + 1$	$r + 1$
文献 [16, 定理 15]	$q \geq \binom{n}{k-1}$	$(l-1)r + 1$	$r + 1$

r 为局部修复性. 由表 1 可以看到, 我们构造的局部修复码需要的有限域的规模比其他的构造都要小, 除了文献 [14, 构造 2]. 但事实上, 当 $r = 3$ 时, 在我们的构造中取 $\beta_i = 1$ ($i \in \{1, \dots, l\}$), 则也可以在有限域 $\mathbb{F}_q$ ($q = r + 1$) 上得到最优的局部修复码, 这时的构造与文献 [14, 构造 2] 一致.

4 结论

本文主要考虑了域规模小于 n 时最优局部修复码的构造. 具体地, 我们给出了有限域 $\mathbb{F}_q$ ($q = \frac{r}{r+1}n + 1$) 上满足所有位局部修复性 $r = 2$ 、 $d = 6$ 和 $d = r + 1$ 的最优局部修复码. 能否在更加小的有限域上一般地构造最优的局部修复码, 并且, 如何刻画最优局部修复码 (或者说界 (1.1)) 对于域大小的限制, 是值得进一步研究的问题.

参考文献

- 1 Gopalan P, Huang C, Simitci H, et al. On the locality of codeword symbols. *IEEE Trans Inform Theory*, 2012, 58: 6925–6934
- 2 Huang C, Chen M, Li J. Pyramid codes: Flexible schemes to trade space for access efficiency in reliable data storage systems. In: *Proceedings of 6th IEEE International Symposium on Network Computing and Applications*. Cambridge: IEEE, 2007, 79–86
- 3 Papailiopoulos D S, Dimakis A G. Locally repairable codes. In: *Proceedings of IEEE International Symposium on Information Theory (ISIT)*. Cambridge: IEEE, 2012, 2771–2775
- 4 Forbes M, Yekhanin S. On the locality of codeword symbols in non-linear codes. *Discrete Math*, 2014, 324: 78–84
- 5 Cadambe V, Mazumdar A. An upper bound on the size of locally recoverable codes. In: *IEEE International Symposium on Network Coding (NetCod)*. Calgary: IEEE, 2013, 1–5
- 6 Wang A, Zhang Z. An integer programming-based bound for locally repairable codes. *IEEE Trans Inform Theory*, 2015, 61: 5280–5294
- 7 Tamo I, Papailiopoulos D S, Dimakis A G. Optimal locally repairable codes and connections to matroid theory. In: *Proceedings of IEEE International Symposium on Information Theory (ISIT)*. Istanbul: IEEE, 2013, 1814–1818
- 8 Silberstein N, Rawat A S, Koyluoglu O O. Optimal locally repairable codes via rank-metric codes. In: *Proceedings of IEEE International Symposium on Information Theory (ISIT)*. Istanbul: IEEE, 2013, 1819–1823
- 9 Tamo I, Barg A. A family of optimal locally recoverable codes. *IEEE Trans Inform Theory*, 2014, 60: 4661–4676
- 10 Huang P, Yaakobi E, Uchikawa H, et al. Cyclic linear binary locally repairable codes. In: *IEEE Information Theory Workshop (ITW)*. Jerusalem: IEEE, 2015, 1–5
- 11 Silberstein N, Zeh A. Optimal binary locally repairable codes via anticodes. In: *IEEE International Symposium on Information Theory (ISIT)*. Hong Kong: IEEE, 2015, 1247–1251
- 12 Zeh A, Yaakobi E. Optimal linear and cyclic locally repairable codes over small fields. In: *IEEE Information Theory Workshop (ITW)*. Jerusalem: IEEE, 2015, 1–5
- 13 Ron M. *Introduction to Coding Theory*. Cambridge: Cambridge University Press, 2006
- 14 Hao J, Xia S T. Bounds and constructions of locally repairable codes: Parity-check matrix approach. *ArXiv:1601.05595*, 2016
- 15 Blaum M. On locally recoverable (LRC) codes. *ArXiv:1512.06161*, 2015
- 16 Song W, Dau S, Yuen C, et al. Optimal locally repairable linear codes. *IEEE J Sel Areas Commun*, 2014, 32: 1019–1036

Constructions of optimal locally repairable codes over small fields

ZHANG ZhiFang, XU JingKe & LIU MuLan

Abstract Locally repairable codes is an attractive research object in recent years. A code coordinate has locality r if the value of this coordinate can be recovered as a function of the values of at most r other coordinates.

This is an important property for reducing the repair complexity in distributed storage systems. In this paper, we consider the constructions of optimal locally repairable codes over finite fields with size smaller than the code length n . Specifically, over the finite field $\mathbb{F}_q$ with $q = \frac{r}{r+1}n + 1$, we construct two classes of optimal locally repairable codes with $r = 2, d = 6$ and $d = r + 1$ respectively.

Keywords locally repairable code, optimal minimum distance, distributed storage system

MSC(2010) 11T71, 68P20

doi: 10.1360/N012016-00206