随机坐标置换下线性码的插入-删除距离

Hengfeng Liu、Chunming Tang 和 Zhengchun Zhou

arXiv:2611.32240v1 [cs.IT] 24 Nov 2026

摘要

我们研究了通过置换固定$[n,k]_q$线性码的坐标所能达到的最大插入-删除（insdel）距离。利用公共子序列与非交叉匹配之间的联系，我们在大字母表上的随机坐标置换下建立了概率下界。特别地，若非恒定码字中的最大符号重数为 $O(n/q)$，则对于充分大的 q，均匀随机坐标置换以高概率产生至少为 $2n - O(n/\sqrt{q})$ 的插入删除距离。此外，对于满足相同符号重数条件的射影线性码，我们证明了所能达到的最大插入删除距离为 $2n - \Theta(n/\sqrt{q})$。作为应用，对于包含全一向量的 $[n,k]_q$ MDS 码，只要 $(k-1)(\log q)^2 = o(n)$，随机坐标置换以高概率产生至少为 $2n - O(\sqrt{kn})$ 的插入删除距离。特别地，对于全长 Reed–Solomon（RS）码，当 $k = o(q/(\log q)^2)$ 时，我们以高概率得到下界 $2q - O(\sqrt{kq})$，在允许更宽维数范围的同时改进了此前已知的下界。最后，我们给出了平方阶有限域上全长 RS 码的显式构造，达到了阶最优的插入删除距离 $2q - \Theta(\sqrt{q})$。

索引词

线性码、坐标置换、插入-删除距离、MDS码、Reed–Solomon码

I. 引言

通信系统和存储系统中的同步错误可能导致符号被插入传输序列或从传输序列中删除。此时，接收序列将不再与传输序列保持同步，这些错误被称为插入和删除错误（简称插入-删除错误）。近年来，由于高密度存储系统和无线通信系统中出现了插入-删除错误，插入-删除码受到了广泛关注，例如DNA存储、语言处理和磁道存储器错误纠正[1], [2]。

插入-删除码理论始于Levenshtein的开创性工作 [3], ，他证明了Varshamov– Tenengolts（VT）码 [4] 能够纠正一次插入或删除。对于固定数量的同步错误，一个具有挑战性的问题是构造冗余度接近理论最小值的显式码。大量在纠正固定数量的删除方面已取得进展，尤其是在双删除情形下 [5]–[8]。然而，纠正固定数量删除所需的最优冗余通常仍未可知。尽管许多已知的插入-删除码都是非线性的，线性码因其表示紧凑、编码高效而备受关注。不过，它们的插入-删除纠错能力受到更严格的限制。 [9] 已证明，能够纠正哪怕一次删除的线性码，其码率至多为 $1/2$。

对于向量 $\mathbf{u}$、$\mathbf{v} \in \mathbb{F}_q^n$，汉明距离 $d_{\mathrm{H}}(\mathbf{u},\mathbf{v})$ 计算它们取值不同的坐标数。它们的插入删除距离 $d_{\mathrm{I}}(\mathbf{u},\mathbf{v})$ 是将一个向量转换为另一个向量所需的最少插入和删除次数。对于固定的线性码，在汉明度量下，坐标置换会产生等价的码，但它们的插入删除距离可能不同。事实上，即使对于较短的 Reed–Solomon码，插入删除距离也可能取多个不同的值。例如，在 $\mathbb{F}_{13}$ 上，取赋值向量 $\boldsymbol{\alpha} = (0, 1, \ldots, 12)$，并考虑三维全长Reed–Solomon码 $\mathrm{RS}^{13,3}(\boldsymbol{\alpha} = \{(f(\alpha_i))_{i=1}^{13} : f \in \mathbb{F}_{13}[x], \deg f < 3\}$。该码的汉明距离为 11，但其坐标的 13! $= 6\,227\,020\,800$ 种置换所对应的插入删除距离分布如表 I 所示。

Hengfeng Liu 任职于中国成都西南交通大学数学学院（电子邮箱：hengfengliu@163.com）。Chunming Tang 和 ZhengchunZhou 任职于中国成都西南交通大学信息科学与技术学院（电子邮箱：tangchunmingmath@163.com；zzc@swjtu.edu.cn）。

表 I 在 $\mathbb{F}_{13}$ 上 $\mathrm{RS}_{13,3}(\boldsymbol{\alpha})$ 的所有坐标置换下的插入删除距离分布，其中 $\boldsymbol{\alpha}=(0、1、\ldots\ldots、12)$。

d_{I}	置换	百分比 (%)
2	382 668	0.00614528
4	42 183 492	0.67742655
6	1 422 576 324	22.84521555
8	4 728 796 020	75.93994258
10	33 082 296	0.53127004

近年来，人们已经建立了线性码插入删除距离的若干上界。对于一个 $[n,k]_q$线性码 $\mathcal{C}$，Cheng 等人 [10] 提出了半 Singleton 界，而 Chen [11]给出了其非渐近版本$d_{\mathrm{I}}(\mathcal{C}) \leq \max\{2(n-2k+2) 2\}$。随后，Ji 等人 [12] 将其加强为 $d_{\mathrm{I}}(\mathcal{C}) \leq \max\{2(n-2k+1) 2\}$，当 $\mathbf{1}\notin\mathcal{C}$ 时，该界称为严格半 Singleton 界。文献 [11], [13], [14] 得到了对任意坐标排序均成立的更多上界。除了建立上界之外，Chen [11] 还指出了为精心构造的二元线性码的插入删除距离获得良好下界所面临的挑战。在这方面，Qiu 等人 [15] 为二元一阶 Reed–Muller 码和单纯形码的某些子码建立了下界。

尽管确定一般线性码的插入删除距离颇具挑战，但针对 Reed–Solomon（RS）码的研究已经取得了 substantial progress。人们已经针对 RS 码和广义 RS 码提出了若干界限与构造 [13], [16], [17], ，尤其关注于达到半Singleton界。对于一般维数，研究已经证明：在充分大的域上存在能够达到该界限的 RS 码，并给出了确定性构造 [18], [19]。对于二维情形，达到该界限所需的最小域大小已被确定为 $\Theta(n^3)$ [18], [20]。最近，Con 等人 [21] 证明，在使用线性规模字母表时，随机 RS 码以高概率逼近半Singleton界。

全长 RS 码为研究坐标置换对插入删除距离的影响提供了一个自然场景。其求值集合为 $\mathbb{F}_q$，因此选择一个码 $\mathrm{RS}_{q,\ k}(\boldsymbol{\alpha})$就等同于对这些求值点进行排序。最近，Beelen 等人 [22] 开启了这项研究。对于一般的 $k\geq 2$，他们证明，如果 $q\geq 10ek^3e^{6k}$，则存在一种对 $\mathbb{F}_q$ 的排序 $\boldsymbol{\alpha}$，使得相应的全长 RS 码$\mathrm{RS}_{q,\ k}(\boldsymbol{\alpha})$的插入删除距离至少为 $q/(5k)$。对于 $k=2$，他们证明，当 q充分大时，均匀随机排列以高概率使得插入删除距离至少为$2(1-\delta)q$，且对每个固定的 $\delta\in(0,1)$均成立。然而，即使对于二维全长 RS 码，要找到具有相当大插入删除距离的显式排序仍然颇具挑战，正如 [22]中的开放问题所指出的那样。对于一般线性码而言，由于缺乏 RS 码的特殊结构，相应的问题则更加困难。

我们将通过置换大字母表上线性码的坐标所能达到的最大插删距离 $d_{\mathrm{I}}^{\max}(\mathcal{C})$记为。在本文中，我们探讨以下两个自然问题：

问题 *1*。对于固定的线性码 $\mathcal{C}$， $d_{\mathrm{I}}^{\max}(\mathcal{C})$可以有多大？随机坐标置换能够达到怎样的插删距离？

问题 *2*。当随机坐标置换以高概率产生较大的插删距离时，我们能否显式构造出达到相近距离的排列？

我们的主要结果如下。

- **随机坐标置换以高概率产生较大的插入删除距离**。我们根据最小汉明距离和最大符号重数，建立了随机坐标置换下线性码插入删除距离的概率下界，并给出了失败概率的显式指数估计（定理 12）。当 q 足够大时，若非恒定码字中的最大符号重数为 $O(n/q)$，则均匀随机的坐标置换以高概率使插入删除距离至少为 $2n-O(n/\sqrt{q})$（推论 13）。我们的证明利用了这样的表示：码字之间的公共子序列对应于其关系图中的非交叉匹配。作为应用，对于包含全一向量的 MDS 码，我们以高概率得到下界 $2n-2(k-1)-(2e+o(1))\sqrt{(k-1)n}$，只要$(k-1(\log q)^2=o(n)$。特别地，对于全长 RS 码，取 $n=q$ 可得插入删除

距离至少为 $2q - O(\sqrt{kq})$，只要 $k = o(q/(\log q)^2)$。对于 $k \geq 3$，这改进了 Beelen 等人所建立的下界 $q/(5k)$。[22, 定理 2] 在 $q \geq 10ek^3e^{6k}$ 条件下成立，并适用于更广的维数范围。对于 $k = 2$，首项系数为 $2e$ 的下界也可通过改进他们的计数论证得到（注 16）。

- **阶的最大插删距离** $2n - \Theta(n/\sqrt{q})$。我们证明了射影线性码的一个依赖字母表的上界，该上界与坐标排序无关。当非恒定码字中的最大符号重数为 $O(n/q)$ 时，该上界渐近地优于表 II 中比较的已有上界：此前已知的上界距离 $2n$的亏损为 $O(n/q)$，而我们的上界的亏损为$\Theta(n/\sqrt{q})$。结合我们的概率下界，这在相同条件下确定了 $d_{\mathrm{maxI}}(\mathcal{C}) = 2n - \Theta(n/\sqrt{q})$。

- **显式排序对于全长 RS 码可实现插删距离** $2q - \Theta(\sqrt{q})$。我们构造了全长 RS 码的确定性排序，使其在平方阶有限域上的距离为 $2q - \Theta(\sqrt{q})$。该构造将求值点按子域上的多项式水平集分组，并利用代数曲线的贝祖定理控制公共子序列。它适用于特征大于$k+1$的每个固定维数 $k \geq 2$。一种改进后的构造在任意特征下都适用于二维。对于 q，这些排序可在多项式时间内计算，并回答了平方阶有限域上的显式排序问题。

论文结构

第二节介绍后续证明所需的工具和辅助结果。第三节建立概率下界，并将其应用于 MDS 码、全长 RS 码和两种非 MDS 求值码；我们还证明了射影线性码的一个上界，并将其与下界结合起来，确定符号重数满足某一条件时可达到的最大插入删除距离。第四节给出平方阶有限域上全长 RS 码的显式排序。最后，我们在第五节提出开放问题。

II. 预备知识

本节回顾插入删除距离的基本事实、用不交叉匹配表示公共子序列的方法，以及显式构造所需的函数域和平面曲线相关结果。

A. 线性码的插入删除距离

一个序列 $(a_1, \ldots, a_r)$ 是公共子序列$\mathbf{u}$、$\mathbf{v} \in \mathbb{F}_q^n$的公共子序列，当且仅当存在指标 $1 \leq i_1 < \cdots < i_r \leq n$ 和$1 \leq j_1 < \cdots < j_r \leq n$，使得对于每个 $1 \leq t \leq r$，都有 $u_{i_t} = v_{j_t} = a_t$。用 $\mathrm{LCS}(\mathbf{u}, \mathbf{v})$ 表示它们之间最长公共子序列的长度。文献 [23] 已证明插入删除距离与最长公共子序列之间存在以下这一众所周知的关系。

引理 1（[23]）**.** 对于任意 $\mathbf{u},\mathbf{v} \in \mathbb{F}_q^n$,

$$d_{\mathrm{I}}(\mathbf{u}, \mathbf{v}) = 2\big(n - \mathrm{LCS}(\mathbf{u}, \mathbf{v})\big). \tag{1}$$

因此，对于码 $\mathcal{C} \subseteq \mathbb{F}_q^n$，有

$$d_{\mathrm{I}}(\mathcal{C}) = 2\big(n - \mathrm{LCS}(\mathcal{C})\big),$$

其中 最长公共子序列$(\mathcal{C}) :=$ 最大值
$\mathbf{u},\mathbf{v}\in\mathcal{C}$ 最长公共子序列$(\mathbf{u},\mathbf{v}).\mathbf{u}\neq\mathbf{v}$

对于固定长度的码，纠正至多 t 次删除等价于纠正至多 t 次插入 [3]。以下判据用最长公共子序列 [22], 刻画这种纠错能力，这可直接由引理1得出。

命题 2（[22]）**.** 设 $\mathcal{C} \subseteq \mathbb{F}_q^n$ 为一个至少包含两个码字的码，且 t 为非负整数。那么$\mathcal{C}$ 能够纠正 t 次插入和删除，当且仅当

$$\mathrm{LCS}(\mathcal{C}) \leq n - t - 1,$$

或等价地，$d_{\mathrm{I}}(\mathcal{C}) \geq 2t + 2$。

B. 单词的公共子序列与二部图中的非交叉匹配

对于 $\mathbf{u}$，$\mathbf{v} \in \mathbb{F}_q^n$，长度为 r 的公共子序列可以表示为二分图 中的一个大小为 r 的不交叉匹配 [24]。我们先回顾匹配中的相关概念，更多细节请参阅 [25] 。

下文所考虑的有向图均为有限图。对于有向图 $D = (V, E)$，有向匹配是一个有向边集合，其中各边的尾端顶点两两不同，头端两两不同。令 V^- 和 V^+ 为 V的两个不相交副本。二部表示 $\mathcal{B}(D)$ 是一个二分图，其顶点类为 V^- 和 V^+；当且仅当 $(i, j) \in E(D)$ [25]时， i^- 和 j^+相邻。反过来，这两个副本上的每个二分图都确定了 V上的一个有向图。这种对应关系在 $\mathcal{B}(D)$中的匹配与D中的有向匹配之间建立了一一对应。对角边 i^-i^+ 对应于 D中的自环 (i, i)，因此删除所有对角边相当于移除所有自环。全文中，我们用 (i, j) 表示边 i^-j^+ 及其对应的有向边 D。

固定 $\prec$ 上的一个全序，该全序在 V^- 和 V^+上诱导出相同的顺序。如果一个匹配 $M = \{(i_s, j_s) : 1 \le s \le r\}$ 的边可以编号为

$$i_1 \prec \cdots \prec i_r, \qquad j_1 \prec \cdots \prec j_r.$$

D 中的有向匹配关于 $\prec$ 是非交叉的，当且仅当其在 $\mathcal{B}(D)$ 中的对应匹配是非交叉的。对于任一种表示 G，令 $L_\prec(G)$ 表示关于 $\prec$的非交叉匹配的最大大小。

现在我们可以将公共子序列与非交叉匹配联系起来。设 $V = [n]$。对于 $\mathbf{u}$、$\mathbf{v} \in \mathbb{F}_q^n$，定义其关系图为二分图 $G_{\mathbf{u},\mathbf{v}} = (V^- \cup V^+, E_{\mathbf{u},\mathbf{v}})$，其中

$$E_{\mathbf{u},\mathbf{v}} = \{(i, j) \in V \times V : u_i = v_j\},$$

~，并将相应的 $G_{\mathbf{u},\mathbf{v}}$ 记为 V上的无环有向图，其边集为 $E_{\mathbf{u},\mathbf{v}} \setminus \{(i, i) : i \in V\}$。对于 $[n]$ 上的全序 $\prec$ （等价地，对于 V），令 $\mathbf{u}_\prec$ 和$\mathbf{v}_\prec$ 分别表示将$\mathbf{u}$ 和$\mathbf{v}$ 的坐标按照 $\prec$排列后得到的序列。$G_{\mathbf{u},\mathbf{v}}$ 中的一个非交叉匹配，其边为 (i_1, j_1) ..... (i_r, j_r)，确定了公共子序列

$$(u_{i_1}, \ldots, u_{i_r}) = (v_{j_1}, \ldots, v_{j_r}).$$

因此，我们有

$$\mathrm{LCS}(\mathbf{u}_\prec, \mathbf{v}_\prec) = L_\prec(G_{\mathbf{u},\mathbf{v}}). \tag{2}$$

以下示例说明顶点序如何影响不交叉匹配以及 $L_\prec(G_{\mathbf{u},\mathbf{v}})$的取值。

示例 3. 令 $\mathbf{u} = (0, 0, 1, 0, 1)$，$\mathbf{v} = (1, 1, 0, 1, 0)$，并将关系图写作 $G = G_{\mathbf{u},\mathbf{v}}$。在自然序 $1 \prec 2 \prec 3 \prec 4 \prec 5$ 下，我们有 $\mathrm{LCS}(\mathbf{u}, \mathbf{v}) = L_\prec(G) = 3$。在公共序 $1 \prec^* 3 \prec^* 2 \prec^* 5 \prec^* 4$ 下，重排后的词为 $\mathbf{u}_{\prec^*} = (0, 1, 0, 1, 0)$ 和 $\mathbf{v}_{\prec^*} = (1, 0, 1, 0, 1)$。它们的公共子序列 $(0, 1, 0, 1)$ 由以下位置对实现

$$M = \{(1, 3), (3, 2), (2, 5), (5, 4)\}.$$

图1展示了这两种坐标顺序下的同一匹配。在 $\prec$ 下它是交叉的，在 $\prec^*$ 下则是非交叉的。公共子序列至多包含两个零和两个一，因此在任何坐标顺序下，其长度都至多为4 。因此

$$L_{\prec^*}(G) = \max_{\prec} L_\prec(G) = 4.$$

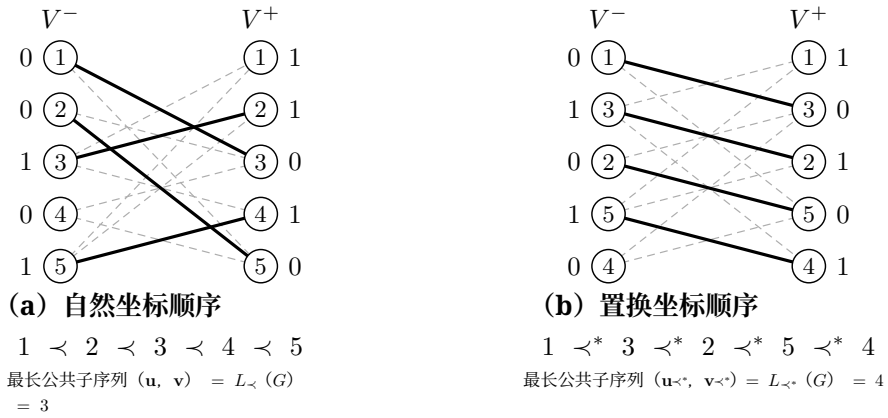


图 1. 在两种常见坐标顺序下，$\mathbf{u} = (0, 0, 1, 0, 1)$ 和 $\mathbf{v} = (1, 1, 0, 1, 0)$ 的关系图。两个面板中均突出显示了大小为 4 的同一匹配 M。

根据引理 1 和公式（2），这两个有序向量之间的插入删除距离可写为

$$d_{\mathrm{I}}(\mathbf{u}_{\prec}, \mathbf{v}_{\prec}) = 2\big(n - L_{\prec}(G_{\mathbf{u},\mathbf{v}})\big).$$

对于一个码 $\mathcal{C} \subseteq \mathbb{F}_q^n$，记 $\mathcal{C}_{\prec} := \{\mathbf{u}_{\prec} : \mathbf{u} \in \mathcal{C}\}$ 为按照 $\prec$ 排列每个码字的坐标所得到的码。对不同码字取最小值，得到

$$d_{\mathrm{I}}(\mathcal{C}_{\prec}) = 2\left(n - \max_{\mathbf{u}\neq\mathbf{v}\in\mathcal{C}} L_{\prec}(G_{\mathbf{u},\mathbf{v}})\right).$$

因此，在所有坐标置换下确定最大插删距离 $d_{\mathrm{I}}^{\max}(\mathcal{C})$，等价于在不同码字的关系图中，将最大的非交叉匹配最小化。

关系图 $G_{\mathbf{u},\mathbf{v}}$ 的边数和顶点次数由符号重数确定。对于每个$a \in \mathbb{F}_q$，定义 $I_a = \{i \in V\colon u_i = a\}$ 和 $J_a = \{j \in V\colon v_j = a\}$。连接 a 各次出现位置的边构成 $I_a \times J_a$。由于 $G_{\mathbf{u},\mathbf{v}}$ 的边被划分为集合 $I_a \times J_a$，我们有

$$|E(G_{\mathbf{u},\mathbf{v}})| = \sum_{a\in\mathbb{F}_q} |I_a|\,|J_a|. \tag{3}$$

此外，$i^- \in V^-$ 和 $j^+ \in V^+$ 的次数分别为 $|J_{u_i}|$ 和 $|I_{v_j}|$。对角边对应于 $\mathbf{u}$ 和 $\mathbf{v}$ 相同的坐标，因此其数量为 $n - d_{\mathrm{H}}(\mathbf{u},\mathbf{v})$。因此

$$\begin{aligned}|E(\widetilde{G}_{\mathbf{u},\mathbf{v}})| &= |\{(i,j) \in V \times V : i \neq j,\ u_i = v_j\}| \\ &= \sum_{a\in\mathbb{F}_q} |I_a|\,|J_a| - n + d_{\mathrm{H}}(\mathbf{u},\mathbf{v}).\end{aligned}$$

~从 $G_{\mathbf{u},\mathbf{v}}$ 中的一个非交叉匹配移除对角边后，得到 $G_{\mathbf{u},\mathbf{v}}$ 中的一个非交叉有向匹配。因此，

$$\mathrm{LCS}(\mathbf{u}_{\prec}, \mathbf{v}_{\prec}) \leq n - d_{\mathrm{H}}(\mathbf{u},\mathbf{v}) + L_{\prec}(\widetilde{G}_{\mathbf{u},\mathbf{v}}). \tag{4}$$

~对于一个有向匹配 M在 $G_{\mathbf{u}\text{中，}\mathbf{v}}$考虑有向子图 (V, M) 其顶点集为 V，边集为 M。由于 M的边两两具有不同的尾端顶点，且头端顶点也两两不同，图 (V, M) 中每个顶点的出度和入度都至多为一。更一般地，我们有以下有用的观察（见 [26]）。这里，$d_G^+(v)$ 和$d_G^-(v)$ 分别表示有向图 G中 v 的出度和入度。 G的弱连通分量是忽略边的方向后得到的连通分支；若一个分量包含一条边，则称其为非平凡的。

命题 4（ [26]）。 设 G 是一个有限无环有向图，对于每个 $v \in V(G)$，均有 $d_G^+(v) \leq 1$ 且 $d_G^-(v) \leq 1$。那么，G 的每个非平凡弱连通分量都是一条有向路径或一个有向环。

特别地，$(V,\ M)$ 的非平凡弱连通分支是有向路径或有向环。在第三节中，我们利用这一分解来估计在均匀随机顶点序下，固定有向匹配为非交叉的概率。

C. 代数函数域与平面曲线

我们回顾第四节中用到的函数域和平面曲线相关事实，遵循 [27], [28], ，并简要证明下文所需的辅助引理。

设 K 是一个代数闭域，并设 $\mathcal{X}$ 是定义在 K上的非奇异射影不可约曲线，其函数域为 $F = K(\mathcal{X})$。 F 的赋值点与 $\mathcal{X}$的点一一对应。此外，非奇异射影曲线的函数域之间的 K-同构会诱导相应曲线之间的同构 [28, 第 7.1 节，推论 3 和 4]。

设 $\mathbb{P}_F$表示 F的赋值点集合，并设 v_P是 $P \in \mathbb{P}_F$处的归一化离散赋值。除子$D = \sum_P n_P P$是赋值点的有限整数和；若所有 $n_P \geq 0$，则记作 $D \geq 0$。由于 K 是代数闭域，其次数为 次数 $D = \sum_P n_P$。对于 $z \in F^\times$，其主除子 和极点除子 定义如下:

$$(z) = \sum_{P \in \mathbb{P}_F} v_P(z)P, \qquad (z)_\infty = \sum_{\substack{P \in \mathbb{P}_F \\ v_P(z)<0}} -v_P(z)P.$$

对于每个 $z \in F \setminus K$，我们有 [27, 定理 1.4。11]

$$[F : K(z)] = \deg(z)_\infty. \tag{5}$$

与除子 D 相关的黎曼–罗赫空间为

$$\mathcal{L}(D) = \{z \in F^\times : (z) + D \geq 0\} \cup \{0\}.$$

对于整数 $m \geq 0$， $\mathcal{L}(mP)$ 由除 P之外没有极点、且在P处的极点阶至多为 m 的函数，以及零函数组成。对于稍后使用的层曲线 $X^r + Y^s = c$，这些空间具有以下描述。

引理5。 设 $r > s \geq 2$ 为与特征 $K \nmid rs$互素的整数，并设 $c \in K^\times$。记 $F_c = K(x, y)$ 为 $x^r + y^s = c$的函数域。函数 x 和 y 具有相同的唯一极点 P_∞，且

$$(x)_\infty = sP_\infty, \qquad (y)_\infty = rP_\infty.$$

对于每个整数 $m \geq 0$,

$$\mathcal{L}(mP_\infty) = \left\langle x^i y^j : i \geq 0,\ 0 \leq j < s,\ is + jr \leq m \right\rangle_K. \tag{6}$$

证明。 根据 [27, 命题3.1， $F_c/K(x)$的扩张次数为 s，并且在无穷远点处完全分歧，这是因为 $v_\infty(x^r - c) = -r$ 且最大公约数$(r, s) = 1$。记 P_∞ 为无穷远点上方的唯一位置，则关系式 $y^s = c - x^r$ 给出$(x)_\infty = sP_\infty$以及$(y)_\infty = rP_\infty$。

由于该仿射曲线是非奇异的，且 y 在 $K[x]$上整，故其坐标环 $K[x, y]$ 是 F_c中 $K[x]$ 的整闭包，因此根据 [27, 定理 3.2.6]，它也是在 P_∞ 之外正则的函数环。每个这样的函数都有唯一的展开式 $\sum_{j=0}^{s-1} a_j(x)y^j$，其中 $a_j(x) \in K[x]$，其单项式 $x^i y^j$ 的极点阶数两两不同$is + jr$，因此属于 $\mathcal{L}(mP_\infty)$ 等价于每个非零项都满足 $is + jr \leq m$，从而证明了式 (6)。 □

我们还将用到以下不可约性陈述，该陈述不要求对特征或 c施加限制。

引理 6. 设 $r > s \geq 2$ 为互素整数。那么，对于每个 $c \in K$，多项式 $X^r + Y^s - c$ 在 $K[X, Y]$中不可约。

证明。 将多项式视为 $Y^s + (X^r - c)$ 上的多项式，其中 $K(X)$。在 X的极点处，其常数系数的赋值为$-r$，与 s互素，且所有中间系数均为零。因此， [27, 命题 3.1.15] 给出了它在 $K(X)[Y]$中的不可约性。于是高斯引理给出了它在 $K[X, Y]$中的不可约性。 □

对于 $A \subseteq K^2$，其 *Zariski*闭包 A^Z是 K^2中的公共零点集，其中包含 $K[U, V]$ 中所有在A [28,第 6 节. 1]上消失的多项式。

引理 7。 设 u， $v \in F$ 不同时为常数，并令 $S \subseteq \mathcal{X}$ 为一个包含它们所有极点的有限集。*S et*

$$A = \{(u(P), v(P)) : P \in \mathcal{X} \setminus S\}.$$

那么，*Zariski*闭包 $\Gamma = A^Z$是一条不可约平面曲线，而 $\Gamma \setminus A$ 是有限的。代换 $U \mapsto u$ 和$V \mapsto v$将 $K(\Gamma)$ 与 $K(u,v) \subseteq F$对应起来。

证明。设 $\rho\colon K[U,V] \to F$将 U映到 u，并将 V映到 v。在 $\mathcal{X} \setminus S$上消失的有理函数必为零，因此在 A 上消失的多项式所构成的理想是 核ρ。因此，

$$K[\Gamma] \cong K[u,v].$$

这是一个在 K 上超越次数为一的整环。因此， Γ 是一条不可约平面曲线，并且在所示的对应关系下， $K(\Gamma) = K(u,\ v)$。

设 $\Gamma \subseteq \mathbb{P}^2$为 Γ的射影闭包。有理映射 $P \mapsto [1\colon u(P)\colon v(P)]$延拓为一个态射$\Phi\colon \mathcal{X} \to \Gamma$，因为 $\mathcal{X}$ 是非奇异的 [28, 第 7.1 节，推论 1]。此态射非常值，因而是满射。因此， $\Gamma \setminus A \subseteq \Phi(S)$是有限集。 □

最后，取射影闭包，即可由 Bézout 定理的射影版本推出下述仿射形式的 Bézout 定理 [28, 第 5.3 节，推论 3]。

定理 8。 在 K 上，两条次数分别为 d_1 和 d_2 、且没有公共不可约分支的平面曲线，至多有 $d_1 d_2$ 个不同的公共仿射点。

III. 界　　坐标置换下线性码的插入删除距离

我们建立了坐标置换下可达到的插入删除距离的下界和上界。首先，我们推导了共同随机顶点序下非交叉有向匹配的尾部估计。将这些估计应用于码字对的关系图，可以得到以汉明距离和符号重数表示的高概率下界，并应用于 MDS 码、全长 RS 码以及两种非 MDS 求值码。随后，我们为射影线性码建立了一个对任意坐标顺序都成立的上界；结合概率下界，该上界确定了满足符号重数条件时的最优渐近距离。

*A. M*共同随机顶点序下关系图的最大非交叉匹配　　*er*

由于关系图的两个顶点类共享相同的坐标顺序，我们首先估计在均匀随机顶点序下，固定有向匹配为非交叉的概率。

引理 9. 设 M 是无环有向图 $G = (V,E)$中一个大小为 $r \geq 1$ 的固定有向匹配，并考虑有向子图 (V,M)。如果 (V,M) 包含有向环，那么不存在任何使 M非交叉的 V全序。否则，其非平凡弱连通分支均为有向路径。如果 (V,M) 有 c个非平凡弱连通分支，那么 V 的均匀随机全序满足

$$\mathbf{P}\left(M \textit{ is noncrossing}\right) \leq \frac{1}{c!(r-c)!} \leq \frac{2^r}{r!}.$$

证明。根据命题 4，(V,M) 的每个非平凡弱连通分支都是有向路径或有向环。假设 M 相对于顶点序 $\prec$ 是非交叉的，且包含有向环 C。令 V_0 为 C的顶点集。C 的边定义了一个置换 $\sigma\colon V_0 \to V_0$，其中 $(v,\sigma(v))$ 是离开 v的边。由于M 是非交叉的， $u \prec v$ 意味着对所有 $u,v \in V_0$， $\sigma(u) \prec \sigma(v)$，但有限全序集上的每个递增置换都是恒等置换。因此，对每个 $v \in V_0$， $\sigma(v) = v$，这与 G 无环的假设矛盾。

现在假设(V,M) 的非平凡弱连通分支都是 c 有向路径。由于这些路径总共有 r 条边， M的支撑有 $r+c$ 个顶点。将该支撑上诱导出的顺序中的位置标记为 $[r+c]$，并令 D 和 R 分别为尾端和头端所占据位置的集合。补集 D^c和 R^c 分别由终端顶点和初始顶点组成。它们是不相交的集合，且大小均为 c，因此至多有

$$\binom{r+c}{c}\binom{r}{c}$$

个可能的有序对(D^c, R^c)。

写出 $D = \{d_1 < \cdots < d_r\}$和 $R = \{s_1 < \cdots < s_r\}$。非交叉匹配必须对每个 i将 d_i与 s_i相连。因此， D和 R确定了有序支撑上的边。(V,M) 的 c个路径分量可以分配给

最终得到的路径分量在 $[r+c]$ in 中至多有 $c!$ 种排列方式。每种排列都确定了路径上所有顶点的位置。由于支撑的相对顺序是均匀分布的，

$$\mathbf{P}\left(M \text{ is noncrossing}\right) \leq \frac{\binom{r+c}{c}\binom{r}{c}\, c!}{(r+c)!} = \frac{1}{c!(r-c)!} = \frac{\binom{r}{c}}{r!} \leq \frac{2^r}{r!}.$$

□

接下来我们将证明，随机全序允许一个较大的不交叉有向匹配的概率，可以用边数来界定；当最大入度和最大出度也受到控制时，还可以得到更精确的估计。

命题 10。 设 $G=(V,\ E)$ 是一个无环有向图，至多含有 m 条边，其中 $m \geq 1$ 是一个整数。均匀随机地选取 V 的一个全序 $\prec$ 。对于每个整数 $1 \leq r \leq m$,

$$\mathbf{P}\left(L_{\prec}(G) \geq r\right) \leq \binom{m}{r}\frac{2^r}{r!}. \tag{7}$$

如果每个顶点的入度和出度都至多为 Δ，则

$$\mathbf{P}\left(L_{\prec}(G) \geq r\right) \leq \frac{m^r}{(r!)^2}\exp\left(2\sqrt{\frac{\Delta r^3}{m}}\right). \tag{8}$$

证明。 事件 $L_{\prec}(G)\ \geq r$ 要求一个含有 r条边的有向匹配是不交叉的。这样的匹配至多有 $\binom{m}{r}$ 个，因此由引理 9及并合界可得式 (7)。

对于第二个估计，引理 9允许我们将注意力限制在满足以下条件的有向匹配 M 上：$(V,\ M)$ 的每个非平凡弱连通分支都是一条有向路径。固定这些路径分量的数量 c，并记为 $a_v = d_G^+(v)$。它们的初始边的选择数至多为

$$\sum_{\substack{S\subseteq V\\ |S|=c}}\prod_{v\in S} a_v \leq \frac{1}{c!}\left(\sum_v a_v\right)^c \leq \frac{m^c}{c!}.$$

按照其尾端的固定辅助序，为选定的初始边标记。路径长度构成一个和为 r 的正整数向量，共有 $\binom{r-1}{c-1}$ 个可能的长度向量。一旦这些数据确定，每条剩余的边至多可以用 Δ 种方式扩展一条路径。允许重复顶点只会增加计数，因此至多有

$$\frac{m^c}{c!}\binom{r-1}{c-1}\Delta^{r-c}$$

有向匹配 M 条边且有 c 个非平凡路径分量的有向匹配于 （$V,\ M$）。现在由引理 9 可得

$$\mathbf{P}\left(L_{\prec}(G) \geq r\right) \leq \sum_{c=1}^{r}\frac{m^c\Delta^{r-c}}{(c!)^2(r-c)!}\binom{r-1}{c-1}.$$

令 $j=r-c$，并提出因子 $m^r/$（$r!$ ^{2}。 j 次项中的剩余因子为

$$\frac{1}{(j!)^2}\left(\frac{\Delta}{m}\right)^j\left(\frac{r!}{(r-j)!}\right)^2\frac{(r-1)!}{(r-1-j)!} \leq \frac{1}{(j!)^2}\left(\frac{\Delta r^3}{m}\right)^j.$$

对于 $x \geq 0$，我们有

$$\sum_{j\geq 0}\frac{x^j}{(j!)^2} \leq \left(\sum_{j\geq 0}\frac{x^{j/2}}{j!}\right)^2 = e^{2\sqrt{x}}.$$

取 $x=\Delta r^3/m$ 即可证明式（8）。 □

B. 随机坐标置换下线性码的插入删除距离

对于线性码 $\mathcal{C}$，令 $\mathcal{C}^+ = \mathcal{C} + \langle \mathbf{1} \rangle$。非恒定码字的最大符号重数满足以下显然的恒等式。

引理 11. 对于包含非恒定码字的线性码 $\mathcal{C}$，

$$\max_{\substack{\mathbf{u} \in \mathcal{C} \setminus \langle \mathbf{1} \rangle \\ a \in \mathbb{F}_q}} |\{i : u_i = a\}| = n - d_{\mathrm{H}}(\mathcal{C}^+).$$

下文中，我们用 $\mu(\mathcal{C}) := n - d_{\mathrm{H}}(\mathcal{C}^+)$ 表示 $\mathcal{C}$ 中非恒定码字的最大符号重数。结合引理 11 和命题 10，我们得出以下关于插入删除距离的高概率下界。

定理 12. 设 $\mathcal{C}$ 是一个包含非恒定码字的 $[n,k]_q$ 线性码序列。假设 $\mu(\mathcal{C}) = o(n)$，且当 $n \to \infty$ 时，$k \log q = o(\sqrt{n\mu(\mathcal{C})})$。对于任意固定的 $\varepsilon > 0$，当 n 足够大时，均匀随机的坐标顺序满足

$$\mathbf{P}\left(d_{\mathrm{I}}(\mathcal{C}_\prec) \geq 2d_{\mathrm{H}}(\mathcal{C}) - (2e + \varepsilon)\sqrt{n\mu(\mathcal{C})} \right) \geq 1 - \exp\left(-\frac{\varepsilon}{2}\sqrt{n\mu(\mathcal{C})} \right) \longrightarrow 1.$$

证明。我们首先为任意包含非恒定码字的 $[n,k]_q$ 线性码给出有限估计。对于不同的非恒定码字，引理 11 和第二节中的式 (3) 给出 $|E(\widetilde{G}_{\mathbf{u},\mathbf{v}})| \leq n\mu(\mathcal{C})$ 以及 $d^+_{\widetilde{G}_{\mathbf{u},\mathbf{v}}}(i)\ d^-_{\widetilde{G}_{\mathbf{u},\mathbf{v}}}(i) \leq \mu(\mathcal{C})$。对于恒定码字 $a\mathbf{1}$ 和非恒定码字 $\mathbf{u}$，引理 11 给出

$$\mathrm{LCS}(a\mathbf{1}, \mathbf{u}_\prec) = |\{i : u_i = a\}| \leq \mu(\mathcal{C}).$$

因此，任何最长公共子序列长度大于 $\mu(\mathcal{C})$ 的码字对，都必须由两个非恒定码字组成。

对于一对不同的非恒定码字，其关系图有 $n - d_{\mathrm{H}}(\mathbf{u},\mathbf{v}) \leq n - d_{\mathrm{H}}(\mathcal{C})$ 条对角边。根据第 II 节中的式 (4)，长度大于 $n - d_{\mathrm{H}}(\mathcal{C}) + r$ 的公共子序列会在 $G_{\mathbf{u},\mathbf{v}}$中产生一个至少有 r 条边的非交叉有向匹配。应用命题 10，取 $m = n\mu(\mathcal{C})$ 和 $\Delta = \mu(\mathcal{C})$，并对少于 q^{2k} 对不同非恒定码字构成的有序对取并合界可得，对于每个整数 $1 \leq r \leq n$,

$$\mathbf{P}\left(\mathrm{LCS}(\mathcal{C}_\prec) > \max\{\mu(\mathcal{C}), n - d_{\mathrm{H}}(\mathcal{C}) + r\}\right) \leq q^{2k}\binom{n\mu(\mathcal{C})}{r}\frac{2^r}{r!} \leq q^{2k}\left(\frac{2e^2 n\mu(\mathcal{C})}{r^2}\right)^r$$

and

$$\mathbf{P}\left(\mathrm{LCS}(\mathcal{C}_\prec) > \max\{\mu(\mathcal{C}), n - d_{\mathrm{H}}(\mathcal{C}) + r\}\right) \leq q^{2k}\frac{(n\mu(\mathcal{C}))^r}{(r!)^2}\exp\left(2\sqrt{\frac{r^3}{n}}\right). \tag{9}$$

现在假设 $\mu(\mathcal{C}) = o(n)$ 且 k 对数 $q = o(\sqrt{n\mu(\mathcal{C})})$。固定 $\varepsilon > 0$ 并令 $r = \left\lfloor (e + \frac{\varepsilon}{2})\sqrt{n\mu(\mathcal{C})} \right\rfloor$。当 n充分大时，我们有 $1 \leq r \leq n$ 且 $\mu(\mathcal{C}) \leq r$。由于 $2r \leq (2e + \varepsilon)\sqrt{n\mu(\mathcal{C})}$，式 (9) 给出

$$\begin{aligned}&\mathbf{P}\left(d_{\mathrm{I}}(\mathcal{C}_\prec) < 2d_{\mathrm{H}}(\mathcal{C}) - (2e + \varepsilon)\sqrt{n\mu(\mathcal{C})} \right) \\ &\quad \leq \exp\left(2k\log q - 2r\log\frac{r}{e\sqrt{n\mu(\mathcal{C})}} + 2r\sqrt{\frac{r}{n}} \right).\end{aligned}$$

不等式 $x\log x \geq x - 1$ 对 $x > 0$ 成立，意味着

$$r\log\frac{r}{e\sqrt{n\mu(\mathcal{C})}} \geq r - e\sqrt{n\mu(\mathcal{C})}.$$

因此，指数至多为

$$-\varepsilon\sqrt{n\mu(\mathcal{C})} + 2 + 2k\log q + 2r\sqrt{\frac{r}{n}}.$$

最后三项为 $o(\sqrt{n\mu(\mathcal{C})})$。因此，当 n 足够大时，失败概率至多为 $\exp(-\varepsilon\sqrt{n\mu(\mathcal{C})}/2)$。

□

当符号重数至多为其平均值 n/q 的某个固定倍数时，由 Singleton界可得维数条件。所得的距离下界形式为 $2n - O(n/\sqrt{q})$。

推论 13. 设 $\mathcal{C}$ 是一个包含非恒定字的 $[n,k]_q$ 线性码序列，其中 $q \to \infty$。假设

$$n - d_{\mathrm{H}}(\mathcal{C}^+) \le K\frac{n}{q}$$

对于一个固定常数 $K > 0$。对于每个固定的 $\varepsilon > 0$ 以及所有充分大的 q，均匀随机的坐标顺序满足

$$\mathbf{P}\left(d_{\mathrm{I}}(\mathcal{C}_\prec) \ge 2n - (2e+\varepsilon)\sqrt{K}\,\frac{n}{\sqrt{q}}\right) \ge 1 - \exp\left(-\frac{\varepsilon n}{4\sqrt{q}}\right) \longrightarrow 1.$$

证明. 由引理 11 和 Singleton界可得

$$\frac{n}{q} \le \mu(\mathcal{C}) \le K\frac{n}{q}, \qquad k \le \dim \mathcal{C}^+ \le \mu(\mathcal{C}) + 1 \le 2\mu(\mathcal{C}).$$

特别地，$\mu(\mathcal{C}) \ge 1$ 且 $n \ge q/K$。因此 $\mu(\mathcal{C}) = o(n)$，并且

$$\frac{k\log q}{\sqrt{n\mu(\mathcal{C})}} \le 2\sqrt{K}\,\frac{\log q}{\sqrt{q}} \longrightarrow 0.$$

将 $\varepsilon/2$ 代入定理 12。由于 $d_{\mathrm{H}}(\mathcal{C}) \ge n - \mu(\mathcal{C})$，其距离下界至少为

$$2n - 2\mu(\mathcal{C}) - \left(2e + \frac{\varepsilon}{2}\right)\sqrt{n\mu(\mathcal{C})} \ge 2n - (2e+\varepsilon)\sqrt{K}\,\frac{n}{\sqrt{q}}$$

对于充分大的 q。这里 $\sqrt{n\mu(\mathcal{C})} \le \sqrt{K}\,n/\sqrt{q}$ 且 $2\mu(\mathcal{C}) = o(n/\sqrt{q})$。此外，$\sqrt{n\mu(\mathcal{C})} \ge n/\sqrt{q}$；因此失败概率至多为 $\exp(-\varepsilon n/(4\sqrt{q}))$。最后，$n \ge q/K$ 推出 $n/\sqrt{q} \to \infty$。 □

我们首先将定理 12 应用于包含全一向量的 MDS 码。

命题 14。 令 $\mathcal{C}$ 为包含 $\mathbf{1}$ 的 $[n,k]_q$ *MDS* 码序列，其中 $2 \le k < n$ 和 $(k-1)(\text{对数 } q)^2 = o(n)$ 为 $n \to \infty$。对于每个固定的 $\varepsilon > 0$ 以及所有充分大的 n，均匀随机的坐标顺序 $\prec$ 满足

$$\begin{aligned}\mathbf{P}\left(d_{\mathrm{I}}(\mathcal{C}_\prec) \ge 2(n-k+1) - (2e+\varepsilon)\sqrt{(k-1)n}\right)\\ \ge 1 - \exp\left(-\frac{\varepsilon}{2}\sqrt{(k-1)n}\right) \longrightarrow 1.\end{aligned}$$

证明。由于 $\mathbf{1} \in \mathcal{C}$，我们有 $\mathcal{C}^+ = \mathcal{C}$ 以及 $\mu(\mathcal{C}) = n - d_{\mathrm{H}}(\mathcal{C}) = k - 1$。由定理 12 可得结论，因为 $(k-1)(\text{对数 } q)^2 = o(n)$ 意味着 $\mu(\mathcal{C}) = o(n)$ 以及 k 对数 $q = o(\sqrt{n\mu(\mathcal{C})})$。 □

将命题 14 专门应用于全长 RS 码，可得到以下推论 .

推论 15。 令 $k \ge 2$ 满足 $k = o(q/(\text{对数 } q)^2)$ 为 $q \to \infty$，并令 $\boldsymbol{\alpha}$ 为 $\mathbb{F}_q$的均匀随机排列。对于每个固定的 $\varepsilon > 0$ 以及所有充分大的 q，

$$\begin{aligned}\mathbf{P}\left(d_{\mathrm{I}}(\mathrm{RS}_{q,k}(\boldsymbol{\alpha})) \ge 2(q-k+1) - (2e+\varepsilon)\sqrt{(k-1)q}\right)\\ \ge 1 - \exp\left(-\frac{\varepsilon}{2}\sqrt{(k-1)q}\right) \longrightarrow 1.\end{aligned}$$

证明。由于全长 RS 码是包含 $\mathbf{1}$ 的 MDS 码，因此由命题 14 和$n = q$可得结论。

□

注 16。 令 $\mathcal{C} = \mathrm{RS}_{q,\ 2}(\boldsymbol{\alpha})$，其中 $\boldsymbol{\alpha}$ 是 $\mathbb{F}_q$的均匀随机排列。将 $\binom{r+c}{r}^2$ 在 [22, 命题 2] 的证明 $\binom{r+c}{r}\binom{r}{c}$ 中对长度为 r 且并集大小固定为 $r+c$ 的递增索引序列对的计数替换为精确计数 $3 \le r < q$，

$$\mathbf{P}\big(\mathrm{LCS}(\mathcal{C}) \ge r\big) \le q^2 \sum_{c=1}^{\min\{r,q-r\}} \frac{q^c}{(r-c)!(c!)^2} \le q^2\frac{q^r}{(r!)^2}\exp\left(\frac{r^2}{q}\right).$$

最后一个不等式通过令 $j = r - c$ 并使用 $r!\ /(r-j)!\ \le r^j$得到。对于每个固定的 $\varepsilon > 0$，取$r = \lceil (e+\varepsilon/2)\sqrt{q} \rceil$ 并使用 $r!\ \ge (r/e)^r$，可得 $d_{\mathrm{I}}(\mathcal{C}) \ge 2q-(2e+\varepsilon)\sqrt{q}$，其概率至少为 $1-\exp(-\Omega_\varepsilon(\sqrt{q}))$，对所有充分大的 q均成立。

接下来，我们将推论 13 应用于通过对仿射函数求值而得到的两个非 MDS 码。

例 17. 设 $q \ge 3$ 是一个素数幂。该码

$$\mathcal{C} = \mathrm{RM}_q(1,2) = \{(a+bx+cy)_{(x,y)\in\mathbb{F}_q^2} : a,b,c \in \mathbb{F}_q\}$$

是定义在 $\mathbb{F}_q^2$ [29]上的一阶 Reed–Muller 码。由于每个非常值仿射函数都恰好取每个值 q次， $\mathcal{C}$的参数为 $[q^2, 3, q^2-q]_q$，且不是 MDS 码。它包含 $\mathbf{1}$ 并满足

$$\mu(\mathcal{C}) = n - d_{\mathrm{H}}(\mathcal{C}^+) = q = \frac{n}{q}.$$

对于每个固定的 $\varepsilon > 0$ 以及所有充分大的 q，由推论 13 取 $K = 1$ 可得

$$\begin{aligned}\mathbf{P}\left(d_{\mathrm{I}}(\mathcal{C}_\prec) \ge 2q^2 - (2e+\varepsilon)q^{3/2}\right)\\ \ge 1 - \exp\left(-\frac{\varepsilon}{4}q^{3/2}\right) \longrightarrow 1\end{aligned}$$

对于均匀随机的坐标顺序 $\prec$。

第二个例子是通过.....得到的 通过在厄米特曲线的仿射有理点上对仿射函数求值而得到。

例 18. 设 q 是一个素数幂，且 S 是厄米特曲线的仿射 $\mathbb{F}_{q^2}$-有理点集合

$$y^q + y = x^{q+1}.$$

厄米特曲线码

$$\mathcal{C} = \{(a+bx+cy)_{(x,y)\in S} : a,b,c \in \mathbb{F}_{q^2}\}$$

的参数为 $[q^3, 3, q^3-q-1]_{q^2}$，且不是 MDS [27, 第 8 节。3]。由于 $\mathbf{1} \in \mathcal{C}$，我们有

$$\mu(\mathcal{C}) = n - d_{\mathrm{H}}(\mathcal{C}^+) = q+1 \le 2\frac{n}{q^2}, \qquad n = q^3.$$

对于每个固定的 $\varepsilon > 0$ 以及所有充分大的 q，由字母表大小为 q^2 且 $K = 2$ 的推论 13 可得

$$\begin{aligned}\mathbf{P}\left(d_{\mathrm{I}}(\mathcal{C}_\prec) \ge 2q^3 - (2e+\varepsilon)\sqrt{2}\,q^2\right)\\ \ge 1 - \exp\left(-\frac{\varepsilon}{4}q^2\right) \longrightarrow 1\end{aligned}$$

对于均匀随机的坐标顺序 $\prec$。

C. 射影线性码插入删除距离的一个上界

现在，我们用一个依赖字母表的上界来补充前述概率下界。该上界适用于任意坐标顺序，限制了射影线性码的插入删除距离。结合推论 13 给出的高概率下界，当非恒定码字的符号重数为 $O(n/q)$时，可确定可达到的最大插入删除距离为 $2n - \Theta(n/\sqrt{q})$。

如果生成矩阵的任意两列都线性无关，则该线性码称为射影线性码。这保证了两个独立随机码字的坐标匹配指示变量成对独立。随后，对连续块应用二阶矩论证，即可得到一个较长的公共子序列。

定理 19. 设 $\mathcal{C}$ 是一个满足 $k \ge 2$ 的射影 $[n,k]_q$ 线性码。那么

$$d_{\mathrm{I}}^{\max}(\mathcal{C}) \le 2n - \left(1 - \frac{1}{\lfloor\sqrt{q}\rfloor}\right)\left\lfloor\frac{n}{\lfloor\sqrt{q}\rfloor}\right\rfloor. \tag{10}$$

证明. 固定一个坐标顺序，并令 $g_1, \dots, g_n \in \mathbb{F}_q^k$ 为生成元列。设 $b = \lfloor \sqrt{q} \rfloor$ 和 $M = \lfloor n/b \rfloor$，并取 M 个互不相交、长度为 b 的连续块。独立且均匀地选择 $U, V \in \mathbb{F}_q^k$。对于每个块B，令

$$X_B = \sum_{i,j \in B} \mathbf{1}_{\{U \cdot g_i = V \cdot g_j\}}.$$

由于生成元列均非零且两两不成比例，该和式中的不同等式对$(U,\ V)$施加相互独立的线性条件。因此，这些指示变量两两独立，且每个变量的均值为 $1/q$，并且

$$\mathbb{E}[X_B] = \frac{b^2}{q}, \qquad \mathbb{E}[X_B^2] = \frac{b^2}{q} + \frac{b^2(b^2-1)}{q^2}.$$

根据柯西–施瓦茨不等式,

$$\mathbf{P}(X_B > 0) \geq \frac{\mathbb{E}[X_B]^2}{\mathbb{E}[X_B^2]} = \frac{b^2}{q + b^2 - 1}.$$

令 Y 统计满足 $X_B > 0$ 的块数。则

$$\mathbb{E}[Y] \geq M \frac{b^2}{q + b^2 - 1}.$$

由于当 $U = V$ 时 $Y = M$，且 $\mathbf{P}(U = V) = q^{-k}$,

$$\mathbb{E}[Y \mid U \neq V] \geq M \frac{b^2/(q + b^2 - 1) - q^{-k}}{1 - q^{-k}}.$$

由于 $q \leq b^2 + 2b$ 且 $q^k \geq (b+1)^2$,, 我们有

$$\frac{b^2/(q + b^2 - 1) - q^{-k}}{1 - q^{-k}} \geq \frac{b/(2(b+1)) - q^{-k}}{1 - q^{-k}} \geq \frac{b-1}{2b}.$$

因此,

$$\mathbb{E}[Y \mid U \neq V] \geq \frac{b-1}{2b} M.$$

因此，存在 $U \neq V$，使得 $Y \geq (b-1)M/(2b)$。对应的码字互不相同，并且从这些块中各取一对匹配项，就能得到长度至少为 Y 的公共子序列。因此,

$$\begin{aligned} d_{\mathrm{I}}(\mathcal{C}) &\leq 2n - 2Y \\ &\leq 2n - \left(1 - \frac{1}{b}\right) M \\ &= 2n - \left(1 - \frac{1}{\lfloor \sqrt{q} \rfloor}\right) \left\lfloor \frac{n}{\lfloor \sqrt{q} \rfloor} \right\rfloor. \end{aligned}$$

□

注20。 我们将定理 19 与此前关于 $[n,k]_q$ 码的上界进行比较，其中 $k \geq 2$ 为 $q \to \infty$，并假设 $\mu(\mathcal{C}) = n - d_{\mathrm{H}}(\mathcal{C}^{+}) = O(n/q)$。表 II 直接给出了任意坐标顺序下码的插入删除距离上界。表中的早期上界右端均为 $2n - O(n/q)$ 的形式，而定理 19 给出 $2n - \Theta(n/\sqrt{q})$。结合推论 13，可得 $d_{\mathrm{I}}^{\max}(\mathcal{C}) = 2n - \Theta(n/\sqrt{q})$，均匀随机的坐标顺序以高概率达到相同的渐近距离。

严格半Singleton 界仅在 $\mathbf{1} \notin \mathcal{C}$ [12]时适用。对于严格直接界，当 q 足够大时， [12]的推论6.2适用于此处:根据假设，射影性成立，而 $d_{\mathrm{H}}(\mathcal{C}) \geq n - \mu(\mathcal{C}) > (n+1)/2$。Singleton界和 $\mathcal{C} \subseteq \mathcal{C}^{+}$ 给出 $1 \leq k - 1 \leq n - d_{\mathrm{H}}(\mathcal{C}) \leq \mu(\mathcal{C}) = O(n/q)$，因此 $n/q = \Omega(1)$。在 Chen 的界 [11, 定理 2.1]中， $h_{\prec}$ 是所选线性无关生成元列的数量，$z_{\prec}$ 是所选位置之前第一个位置之前或排序 $\prec$中最后一个位置之后，所选非零码字的零坐标数量。因此，对每个 $\prec$，都有 $h_{\prec} \leq k$ 和$z_{\prec} \leq n - d_{\mathrm{H}}(\mathcal{C})$，从而 $h_{\prec} + z_{\prec} - 1 = O(n/q)$ 在 $\prec$ 中一致成立。在 Xie等人的界 [14]中，L 是两个线性无关码字的汉明重量的共同下界。由于$L \leq n$，我们有$\lceil L/(q-1) \rceil = O(n/q)$。

表 II 任意坐标顺序 $\prec$ 下投影 $[n,k]_q$ 码的 $\mu(\mathcal{C}) = O(n/q)$ 的上界。

界	$d_{\mathrm{I}}(\mathcal{C}_\prec)$ 的上界	界的渐近形式
半 Singleton 界 [10]	$2n - (4k-4)$	$2n - O(n/q)$
严格半 Singleton 界 [12]	$2n - (4k-2)$	$2n - O(n/q)$
直接界	$2d_{\mathrm{H}}(\mathcal{C})$	$2n - O(n/q)$
严格直接界 [12]	$2(d_{\mathrm{H}}(\mathcal{C}) - 1)$	$2n - O(n/q)$
Chen [11, 定理 2.1]	$2n - 2(h_\prec + z_\prec - 1)$	$2n - O(n/q)$
谢等人 [14,定理 2。1]	$2n - 2\lceil L/(q-1) \rceil$	$2n - O(n/q)$
定理 19	$2n - \left(1 - \frac{1}{\lfloor\sqrt{q}\rfloor}\right)\left\lfloor\frac{n}{\lfloor\sqrt{q}\rfloor}\right\rfloor$	$2n - \Theta(n/\sqrt{q})$

IV. 具有较大插入删除距离的全长 Reed–Solomon 码的显式排序

现在回到全长 RS 码，并寻求对于固定维数具有距离 $2q - \Theta(\sqrt{q})$ 的显式排序。推论 15 和定理 19 表明，随机排序可以达到这一渐近距离。在平方阶有限域上，我们通过将求值点划分为连续块，并利用代数交集界来控制不同块之间的公共值，构造出这样的排序。

引理 21。 设 G 为一个二分图，其顶点类为 V^- 和 V^+，二者都是有限集 V的副本。设 $V = V_1 \sqcup \cdots \sqcup V_m$ 为 V的一个划分，其中各子集非空；固定 $\prec$上的一个排序 V，使得这些子集按照 V_1、.....、V_m的顺序作为连续块出现。每个块内部的顺序任意，并且两个顶点类使用相同的顺序。对于每个 i，令 V_i^-和 V_i^+分别表示 V_i在 V^- 和 V^+中的副本，并用 $G[V_i^-, V_j^+]$ 表示由 $V_i^- \cup V_j^+$诱导的子图。假设对于每个 $1 \le i,\ j \le m,\ G[V_i^-, V_j^+]$ 的普通匹配数至多为 b。则

$$L_\prec(G) \le (2m-1)b.$$

证明。 设 M 是一个非空的不交叉匹配，并按左端点递增的顺序列出其边。它们的右端点也按相同顺序排列。因此，对应的块索引在两个坐标上均单调不减。移除连续重复项后，设访问到的块对为 $(i_1,\ j_1)\ldots,$ $(i_t,\ j_t)$。由于每一步至少有一个索引增加，

$$t - 1 \le (i_t - i_1) + (j_t - j_1) \le 2(m-1).$$

M 中属于任意一个块对的边构成一个大小至多为 b 的普通匹配。因此，$|M| \le tb \le (2m-1)b$。

□

在引理 21 的证明中，沿着一个不交叉匹配的块索引 $(i,\ j)$ 在两个坐标上均单调不减。下图展示了 $m = 3$ 的一种可能序列，其中包含五个不同的块对，达到最大值 5。

	V_1^+	V_2^+	V_3^+
V_1^-	1	2	
V_2^-		3	4
V_3^-			5

图 2. $m = 3$ 的五个连续块对。

对于一般维数 k，我们给出如下全长 RS 码构造。对于每个固定的 k，所得码的距离为 $2q - \Theta(\sqrt{q})$，因此该构造达到了阶最优。

定理 22。 设 $k \geq 2$，且令 $q = q_0^2$ 为特征为 $p > k+1$ 的素数幂的平方。选取 $\theta \in \mathbb{F}_q \setminus \mathbb{F}_{q_0}$。对于每个 $c \in \mathbb{F}_{q_0}$，定义

$$S_c = \{x + \theta y : x, y \in \mathbb{F}_{q_0},\ x^{k+1} + y^k = c\}.$$

将非空集合 S_c 依次列出，以获得 $\boldsymbol{\alpha}$ 的一个排序 $\mathbb{F}_q$。集合的排列顺序以及集合内部元素的顺序均可任意选择。于是

$$d_{\mathrm{I}}(\mathrm{RS}_{q,k}(\boldsymbol{\alpha})) \geq 2q - 4\big(1 + (k+1)^2(k-1)^3\big)\sqrt{q}. \tag{11}$$

对于固定的 k，可以在关于 q的多项式时间内确定性地构造出这样的排序。

为应用引理 21，我们通过平面曲线的交来界定非零块之间的匹配。在 $K = \mathbb{F}_q$上进行处理，并令 $\sigma\colon a \mapsto a^{q_0}$ 为 $\mathbb{F}_q/\mathbb{F}_{q_0}$ 的弗罗贝尼乌斯自同构。对于 $f \in \mathbb{F}_q[T]$，记 f^σ 为将 σ 逐系数应用于 f 所得的多项式。对于 $c \in \mathbb{F}_{q_0}^*$，令 F_c 为引理5中的函数域，其中（r，s）=（$k+1$，k），并令 $\mathcal{X}_c$为其非奇异射影模型。我们将 $\mathcal{X}_c \setminus \{P_\infty\}$ 与仿射曲线$X^{k+1} + Y^k = c$等同。对于一个非常数多项式 $f \in \mathbb{F}_q[T]$，令 $\Gamma_{f,\,c}$ 为 K^2 中 $\mathcal{X}_c \setminus \{P_\infty\}$ 在该映射下的像的 Zariski闭包

$$(x, y) \longmapsto \big(f(x + \theta y), f^\sigma(x + \theta^{q_0} y)\big).$$

第一坐标函数在 P_∞ 处的极点阶数为$(k+1)$ 次 f。因此，由引理 7 可知，结合 $S = \{P_\infty\}$，$\Gamma_{f,\,c}$ 是一条不可约平面曲线。

引理 23。设 $c,\ \ d \in \mathbb{F}_{q_0}^*$，并设 $f,\ h \in \mathbb{F}_q[T]$ 是次数小于 k 的非恒定多项式。 n

$$\deg \Gamma_{f,c} \leq (k+1)\deg f. \tag{12}$$

此外，当 $f \neq h$ 时，$\Gamma_{f,\,c} \neq \Gamma_{h,\,d}$。

证明。设 $m =$次 f，由引理 7 可知，一条一般仿射直线 ℓ 与 $\Gamma_{f,\,c}$ 相交于 次$\Gamma_{f,\,c}$ 个互不相同的点，这些点都位于该映射的像中。将该映射的坐标函数代入 ℓ 的方程，得到一条次数至多为 m 的平面曲线，它与 $X^{k+1} + Y^k = c$ 相交于至少 次$\Gamma_{f,\,c}$ 个互不相同的点。由贝祖定理（定理 8）可得 (12)。

根据引理 7，我们可以确定

$$K(\Gamma_{f,c}) = K\big(f(x + \theta y), f^\sigma(x + \theta^{q_0} y)\big) \subseteq F_c.$$

令 a_m 为 f 的首项系数，并令 $N = m$（$k+1$）。两个生成元的极点除子均为 NP_∞。考虑归一化差

$$w = \frac{f(x + \theta y)}{a_m \theta^m} - \frac{f^\sigma(x + \theta^{q_0} y)}{a_m^{q_0} \theta^{m q_0}} \in K(\Gamma_{f,c}).$$

y^m 次项相消，而 xy^{m-1} 的系数为 $m\ (\theta^{-1} - \theta^{-q_0}) \neq 0$，因为 $m < p$ 和 $\theta \notin \mathbb{F}_{q_0}$。单项式 $x^i y^j$ 的极点阶为 $ki +$（$k+1$）j。因此，xy^{m-1} 的极点阶为 $N-1$，而其他 m 次项的极点阶至多为 $N-2$，低次数项的极点阶至多为（$m-1$）（$k+1$）。由于 w 没有有限极点，（w）$_\infty =$（$N-1$）P_∞。因此，由式 (5) 得到

$$[F_c : K(f(x + \theta y))] = N, \qquad [F_c : K(w)] = N - 1.$$

两个子域均包含于 K（$\Gamma_{f,\,c}$）中，因此塔定律表明 $[F_c$：K（$\Gamma_{f,\,c}$）$]$ 同时整除 N 和 $N-1$。这些整数互素，因此 K（$\Gamma_{f,\,c}$）$= F_c$。

假设 $\Gamma_{f,\,c} = \Gamma_{h,\,d}$。前述计算将 F_c 和 F_d 识别为其公共像的函数域。令 x'、$y' \in F_c$ 为 F_d的生成元的像。像的第一坐标在每条曲线上都有一个唯一极点。因此，它们的无穷远处赋值点相对应，比较其极点阶数可得 $\deg h = m$。由引理5，

$$\mathcal{L}(kP_\infty) = \langle 1, x\rangle_K, \qquad \mathcal{L}((k+1)P_\infty) = \langle 1, x, y\rangle_K.$$

由于 x' 和 y' 的极点阶数分别恰为 k 和 $k+1$，因此它们具有以下形式

$$x' = ax + b, \qquad y' = ux + vy + e, \qquad av \neq 0.$$

公共像上的坐标函数相等，从而得到

$$f(x+\theta y)=h(x'+\theta y'),$$
$$f^{\sigma}(x+\theta^{q_0}y)=h^{\sigma}(x'+\theta^{q_0}y').$$

代入这些表达式，并将 x,y 替换为 X,Y 后，两边之差是次数至多为 $m<k+1$ 、且在 $X^{k+1}+Y^k=c$ 上消失的多项式。该不可约曲线的理想由 $X^{k+1}+Y^k-c$ 生成。因此，每个差都能被这个多项式整除，而次数上界迫使两个差都是零多项式。比较 X^m 和 $X^{m-1}Y$ 的系数，得到

$$a+\theta u=v=a+\theta^{q_0}u.$$

因此， $u=0$ 和 $v=a$。关系 $(x')^{k+1}+(y')^k=d$ 现在表明，$(aX+b)^{k+1}+(aY+e)^k-d$ 是 $X^{k+1}+Y^k-c$ 的一个标量倍。比较 X^{k+1} 和 Y^k 的系数，得到 $a^{k+1}=a^k$，从而得到 $a=1$。由于$p>k+1$， X^k 和 Y^{k-1} 的系数给出 $b=e=0$。由此可得 $c=d$，而第一坐标恒等式化简为 $f(x+\theta y)=h(x+\theta y)$。由于 $x+\theta y$ 不是常数，这意味着 $f=h$。

□

现在我们可以证明定理 22。

定理 22 的证明。设 f、$h\in\mathbb{F}_q[T]$ 是次数小于 k 的两个不同非恒定多项式。它们的关系图中，块 S_c 与 S_d 之间的边对应于多项式值相同的求值点对。对于满足 x 的 $x+\theta y\in S_c$、$y\in\mathbb{F}_{q_0}$，$\Gamma_{f,\,c}$ 上对应的点为 $(f(x+\theta y),\ f(x+\theta y)^{q_0})$。因此，当 c、$d\neq 0$ 时， $f(S_c)\cap h(S_d)$ 中不同的公共值对应于 $\Gamma_{f,\,c}\cap\Gamma_{h,\,d}$ 上不同的点。根据引理 23，这些曲线互不相同，因此贝祖定理给出公共值的数量至多为 $(k+1)^2\deg f\deg h$。每个值对匹配的贡献至多为 $\min\{\deg f,\deg h\}$ 条边，因为非恒定多项式取任意给定值的次数至多等于其次数。因此，任意两个非零块之间的普通匹配数至多为

$$(k+1)^2\deg f\,\deg h\min\{\deg f,\deg h\}\le(k+1)^2(k-1)^3.$$

我们声称 $|S_0|=q_0$。除 $(0,\ 0)$ 外， $x^{k+1}+y^k=0$ 在 $\mathbb{F}_{q_0}$ 上的解可由 b 参数化 y

$$(x,y)=(-t^k,-t^{k+1}),\qquad t\in\mathbb{F}_{q_0}^{\times},$$

其中 t 可唯一恢复为 y/x，因此恰有 q_0-1 个非零解。在不交叉匹配的任一侧删除所有与 S_0相邻的边，至多会移除 $2q_0$ 条边，剩下的匹配在每一侧至多支撑于 q_0-1 个连续块上，且前述匹配界适用于每一对剩余的块。引理 21 给出

$$\begin{aligned}\mathrm{LCS}(f(\boldsymbol{\alpha}),h(\boldsymbol{\alpha}))&\le 2q_0+(2q_0-3)(k+1)^2(k-1)^3\\&\le 2\big(1+(k+1)^2(k-1)^3\big)q_0.\end{aligned}$$

如果恰有一个多项式为常数，则最长公共子序列至多为 $k-1$；若两个常数不同，则最长公共子序列长度为零。证明完毕。 □

在二维情形下，我们证明，水平集多项式的其他选取方式在特征 2 和3下会得到相同的渐近距离界。

命题24。 设 $q=q_0^2$ 的特征为 $p\in\{2,3\}$。取 $\theta\in\mathbb{F}_q\setminus\mathbb{F}_{q_0}$，并令

$$(r,s)=\begin{cases}(5,2), & p=3,\\(5,3), & p=2.\end{cases}$$

对于每个 $c\in\mathbb{F}_{q_0}$，定义

$$S_c=\{x+\theta y:x,y\in\mathbb{F}_{q_0},\ x^r+y^s=c\}.$$

将非空集合 S_c 依次列出，得到 $\mathbb{F}_q$的一个排列 $\boldsymbol{\alpha}$ 。集合的排列顺序以及每个集合内部的顺序均可任意。于是

$$d_{\mathrm{I}}(\mathrm{RS}_{q,2}(\boldsymbol{\alpha}))\ge 2q-100\sqrt{q}+50. \tag{13}$$

证明。 这些指数满足 最大公约数$(r, s) = 1$, $p \nmid rs$, and $r - s = p$. 设 $\theta^2 = \tau\theta + \nu$，其中 $\tau, \nu \in \mathbb{F}_{q_0}$和 $\nu \neq 0$。对于每个非恒等仿射映射 $\gamma(z) = az + b$，满足 $a \neq 0$，我们断言

$$|S_c \cap \gamma^{-1}(S_d)| \leq 25. \tag{14}$$

写出 $a = a_0 + a_1\theta$和 $b = b_0 + b_1\theta$。公式（14）中计数的点对应于.....的公共仿射点

$$X^r + Y^s = c$$

and

$$(a_0X + \nu a_1Y + b_0)^r + \big(a_1X + (a_0 + \tau a_1)Y + b_1\big)^s = d.$$

根据引理 6，且由于乘以 a 是可逆的，这两条曲线都是次数为 $r = 5$ 的绝对不可约曲线。

若这两条曲线有公共分支，则它们的定义多项式相差一个非零标量因子 λ。比较次数为 r 的齐次项可得

$$(a_0X + \nu a_1Y)^r = \lambda X^r.$$

因此 $a_1 = 0$ 和 $\lambda = a_0^r$。于是， X^{r-1} 和 Y^{s-1} 的系数给出 $b_0 = b_1 = 0$，因为 $p \nmid rs$。比较 Y^s 的系数可得 $a_0^{r-s} = 1$。由于 $r - s = p$，这迫使 $a_0 = 1$，与 $\gamma \neq \mathrm{id}$矛盾。然后，B´ezout 定理证明公式（14）。

对于不同的非恒定仿射多项式 f， h，关系 $f(x) = h(y)$是关于非恒等仿射映射 $\gamma = h^{-1} \circ f$ 的 $y = \gamma(x)$。其关系图是一个匹配，每个块对至多有 25 条边。引理 21 给出

$$\mathrm{LCS}(f(\boldsymbol{\alpha}), h(\boldsymbol{\alpha})) \leq 25(2q_0 - 1).$$

涉及常数多项式的不同二元组，其最长公共子序列长度至多为一。证明完毕。 □

结合针对 $k = 2$ 的定理 22，此命题给出了在每个平方阶有限域上距离为 $2q - \Theta(\sqrt{q})$ 的确定性排序。

V. 结论

本文研究了坐标排序如何影响大字母表上线性码的插入删除性能。公共子序列与非交叉匹配之间的联系，为理解随机排列所能达到的距离以及显式排序的设计提供了共同视角。对于符号重数受到适当控制的射影线性码，我们的结果表明，随机排序能够达到距离与 $2n$ 之间差距的最优渐近阶 $\Theta(n/\sqrt{q})$ 。这一认识也导出了固定维数的全长 Reed–Solomon 码在平方阶有限域上的显式排序，其差距达到了相同的平方根阶。

对可达到的最大距离进行更精确的理解仍值得研究，尤其是与 $2n$ 之间差距为 $n/\sqrt{q}$ 阶时的常数。具体而言，对于全长 Reed–Solomon 码，确定固定维数下 $\sqrt{q}$ 项的系数，并理解维数随域大小增长时最大距离如何变化，都很有意思。将显式构造推广到任意有限域，也将是后续研究的一个自然方向。

参考文献

[1] M. Cheraghchi 和 J. Ribeiro，“同步信道容量结果概述”，*IEEE Trans. Inf. Theory*，第 67 卷，第 6 期，第 3207–3232 页，2021 年 6 月。[2] B. Haeupler 和 A. Shahrasbi，“用于插入和删除的同步字符串与编码——综述”，*IEEE Trans. Inf. Theory*，第 67 卷，第 6 期，第 3190–3206 页，2021 年 6 月。[3] V. I. Levenshtein，“能够纠正删除、插入和反转的二进制码”，《苏联物理学文献》，第 10 卷，第 8 期，第 707–710 页，1966 年 2 月。[4] R. R. Varshamov 和 G. M. Tenengolts，“能够纠正单个非对称错误的码”，《自动化与遥控》，第 26 卷，第 2 期，第 286–290 页，1965 年。[5] J. Brakensiek、V. Guruswami 和 S. Zbarsky，“用于纠正多次删除的高效低冗余码”，*IEEE Trans. Inf. Theory*，第 64 卷，第 5 期，第 3403–3410 页，2018 年 5 月。[6] J. Sima 和 J. Bruck，“关于最优 k-删除纠错码”，*IEEE Trans. Inf. Theory*，第 67 卷，第 6 期，第 3360–3375 页，2021 年 6 月。

[7] V. Guruswami 和 J. H˚astad，“冗余度达到存在性界的显式双删除码”，IEEE Trans. Inf. Theory，第 67 卷，第 10 期，第 6384–6394 页，2021 年 10 月。[8] S. Liu、I. Tjuawinata 和 C. Xing，“具有低冗余度的显式 q元双删除纠错码构造”，IEEE Trans. Inf. Theory，第 70 卷，第 6 期，第 4093–4101 页，2024 年 6 月。[9] K. A. S. Abdel-Ghaffar、H. C. Ferreira 和 L. Cheng，“使用线性码和循环码纠正删除”，IEEE Trans. Inf. Theory，第 56 卷，第 10 期，第 5223–5234 页，2010 年 10 月。[10] K. Cheng、V. Guruswami、B. Haeupler 和 X. Li，“用于纠正插入和删除的高效线性码与仿射码”，SIAM J. Discrete Math.，第 37 卷，第 2 期，第 748–778 页，2023 年 6 月。[11] H. Chen，“线性插入删除码的无坐标排序上界”，IEEE Trans. Inf. Theory，第 68 卷，第 8 期，第 5126– 5132 页，2022 年 8 月。[12] Q. Ji、D. Zheng、H. Chen 和 X. Wang，“线性插入删除码的严格半 Singleton 界、严格直接上界与最优码”，IEEE Trans. Inf. Theory，第 69 卷，第 5 期，第 2900–2910 页，2023 年 5 月。[13] B. Chen 和 G. Zhang，“插入删除码的改进 Singleton界与最优构造”，IEEE Trans. Inf. Theory，第 68 卷，第 5 期，第 3028–3033 页，2022 年 5 月。[14] C. Xie、H. Chen、L. Qu 和 L. Liu，“线性插入删除码新的与维数无关的上界”，Adv. Math. Commun.，第 18 卷，第 6 期，第 1575–1589 页，2024 年 12 月。[15] R. Qiu 、 S. Zhu 、 Y. Li 和 Z. Sun，“ 关于二元 Reed–Muller 码和单纯形码的插入删除纠错容量”，IEEE Trans. Inf. Theory，第 72 卷，第 4 期，第 2099–2111 页，2026 年 4 月。[16] T. D. Duc、S. Liu、I. Tjuawinata 和 C. Xing，“高插入删除噪声区间内二维Reed–Solomon码的显式构造”，IEEE Trans. Inf. Theory，第 67 卷，第 5 期，第 2808–2820 页，2021 年 5 月。[17] S. Liu 和 I. Tjuawinata，“关于具有最优渐近纠错能力的二维插入删除Reed–Solomon码”，Finite Fields Appl.，第 73 卷，2021 年 8 月，文章编号 101841。[18] R. Con、A. Shpilka 和 I. Tamo，“针对对抗性插入和删除的 Reed–Solomon 码”，IEEE Trans. Inf. Theory，第 69 卷，第 5 期，第 2991–3000 页，2023 年 5 月。[19] J. Liu，“针对对抗性插入和删除的最优 RS 码与 GRS码及最优构造”，IEEE Trans. Inf. Theory，第 70 卷，第 9 期，第 6269–6279 页，2024 年 9 月。[20] R. Con、A. Shpilka 和 I. Tamo，“纠正插入和删除的最优二维Reed–Solomon码”，IEEE Trans. Inf. Theory，第 70 卷，第 7 期，第 5012–5016 页，2024 年 7 月。[21] R. Con、Z. Guo、R. Li 和 Z. Zhang，“随机Reed–Solomon码在大规模线性字母表上达到插入删除的半 Singleton 界”，载于第 52 届国际自动机、语言与程序设计专题讨论会（ICALP）论文集，系列：莱布尼茨国际信息学会议录（LIPIcs），第 334 卷，2025 年，第 60:1–60:21 页。[22] P. Beelen、R. Con、A. Gruica、M. Montanucci 和 E. Yaakobi，“针对插入和删除的 Reed–Solomon 码：全长度码与码率为1/2的码”，IEEE Trans. Inf. Theory，第 72 卷，第 2 期，第 1149–1160 页，2026 年 2 月。[23] B. Haeupler 和 A. Shahrasbi，“同步字符串：逼近 Singleton界的插入删除码”，载于第 49 届 ACM SIGACT 计算理论年会（STOC）论文集，2017 年 6 月，第 33–46 页。[24] M. Kiwi、M. Loebl 和 J. Matou˘sek，“大字母表的最长公共子序列的期望长度”，Adv. Math.，第 197 卷，第 2 期，第 480–498 页，2005 年 11 月。[25] L. Lov´asz 和 M. D. Plummer，《匹配理论》。美国罗德岛州普罗维登斯：AMS Chelsea Publishing，2009 年。[26] B. E. Sagan，《组合学：计数的艺术》，系列：数学研究生教材，第 210 卷。美国罗德岛州普罗维登斯：美国数学学会，2020 年。[27] H. Stichtenoth，《代数函数域与编码》，第 2 版。德国柏林：Springer，2009 年。[28] W. Fulton，《代数曲线：代数几何导论》。美国加利福尼亚州雷德伍德城：Addison-Wesley，1989 年。[29] P. Delsarte、J.-M. Goethals 和 F. J. MacWilliams，“关于广义 Reed–Muller码及其相关码”，Inf. Control.，第 16 卷，第 5 期，第 403–442 页，1970 年 7 月。