Curve-lifted codes for local recovery using lines

Gretchen L. Matthews*[1], Travis Morrison†[2], and Aidan W. Murphy‡[3]

[1,3]Department of Mathematics, Virginia Tech
[2]Johns Hopkins Applied Physics Laboratory

Abstract

In this paper, we introduce curve-lifted codes over fields of arbitrary characteristic, inspired by Hermitian-lifted codes over $\mathbb{F}_{2^r}$. These codes are designed for locality and availability, and their particular parameters depend on the choice of curve and its properties. Due to the construction, the numbers of rational points of intersection between curves and lines play a key role. To demonstrate that and generate new families of locally recoverable codes (LRCs) with high availabilty, we focus on norm-trace-lifted codes. In some cases, they are easier to define than their Hermitian counterparts and consequently have a better asymptotic bound on the code rate.

1 Introduction

Algebraic geometry codes were introduced in the 1980s [7] and quickly received attention due to the existence of sequences with parameters exceeding the Gilbert-Varshamov bound [18]. These codes are defined by evaluating functions on a smooth projective curve over a finite field at rational points, a construction that is quite flexible allowing for customizations or variants that achieve particular goals. In [2], it was demonstrated how coverings of curves yield locally recoverable codes (also known as LRCs or codes with locality). A code C of length n is said to be **locally recoverable** if there exist recovery sets $R_1, \ldots, R_n$ such that for every codeword coordinate i, the i^{th} coordinate c_i of any codeword $c \in C$ may be recovered from the codeword symbols $c\mid_{R_i}$. The cardinality of the largest R_i is called the locality of C. A code has availability t if each coordinate has t disjoint recovery sets and is called a code with availability if $t > 1$.

*gmatthews@vt.edu

†tmo@vt.edu

‡Aidan.Murphy@jhuapl.edu

The work of the first and second authors is supported in part by the Commonwealth Cyber Initiative. The work of the first author is supported by NSF DMS-2201075. This work was performed while the third author was at Virginia Tech. Some results on binary norm-trace codes were presented at ISIT 2021 and in the third author's dissertation.

Locally recoverable codes, whose study originated in [6, 16], are studied due to their applicability in setting such as distributed storage where vast amounts of data are stored across many servers which may be temporarily offline (and is modeled as an erasure). To limit network traffic involved in recovery, it is desirable that each coordinate (or server) can be recovered using information from a small subset all other coordinates (or the rest of the network). Moreover, it is useful to have multiple ways in which information can be recovered, so that if a coordinate is lost, the ability to recover does not depend on the availability of some other coordinate which might also be lost.

Algebraic geometry codes can be adapted to define locally recoverable codes, as noted in [2, 8]. In this paper, we define curve-lifted codes, a generalization of the Hermitian-lifted codes [10], for a projective curve $\mathcal{X}$ over a finite field $\mathbb{F}_q$. Curve-lifted codes are evaluation codes in which codewords are determined by evaluating particular functions at affine points on the curve. The functions to evaluate depend on what we will refer to as an intersection number: given a curve $\mathcal{X}$ and a collection of lines $\mathbb{L}$, the intersection number is

$$\min \left\{ \mid \left(L \cap \mathcal{X} \right) \left(\mathbb{F}_q \right) \mid : L \in \mathbb{L} \right\}.$$

Codewords arise from the evaluation of rational functions f on X that restrict to low-degree polynomials on $\left(L \cap \mathcal{X} \right) \left(\mathbb{F}_q \right)$. Recovery sets for a position associated with a point P consist of collections of other points of intersection between a line through P and $\mathcal{X}$. The parameters of such codes depend on the collection $\mathbb{L}$ of lines selected and the number of $\mathbb{F}_q$-rational points that lie on both the curve $\mathcal{X}$ and a line $L \in \mathbb{L}$. When $\mathbb{L}$ is the set of all lines in affine space over $\mathbb{F}_q$, we refer to such a number as an intersection number.

Intersection numbers for particular curves can be challenging to determine. We demonstrate progress in this direction for the **norm-trace curve** which is defined by

$$\mathcal{X}_{q,r} : y^{q^{r-1}} + y^{q^{r-2}} + \cdots + y^q + y = x^{\frac{q^r-1}{q-1}}$$

over the field $\mathbb{F}_{q^r}$ with q^r elements. Taking $r = 2$ gives the **Hermitian curve** $H_q : y^q + y = x^{q+1}$ over $\mathbb{F}_{q^2}$. This allows us to give important instances of the curve-lifted construction, resulting in norm-trace-lifted codes.

Curve-lifted codes and norm-trace-lifted codes are strongly inspired by Hermitian-lifted codes. We note some key distinctions between the binary Hermitian case ($r = 2$) and the more general norm-trace one in which $r > 2$. The functions which give rise to codewords can be described explicitly for the case $r > 2$, while much of the effort in studying the Hermitian case is devoted to finding enough functions to get a positive rate. It should be mentioned that lifted codes are not traditional algebraic geometry codes; while the codewords are of a similar form, obtained by evaluating functions at rational points, the space of functions is not a Riemann-Roch space.

Determining the functions to evaluate to define a curve-lifted code is a crucial task. As we will see in the norm-trace case when $r > 2$, we must first determine the appropriate locality, which depends on the number of rational points of intersection between the curve $\mathcal{X}_{q,r}$ and a line $\mathbb{F}_{q^r}$. That is a primary contribution of this work. It pays off in that once this is determined, we can immediate specify enough functions which yield codewords in the norm-trace-lifted code to prove that this family of codes has positive rate even as the code length grows, for a fixed characteristic. This is in contrast to the Hermitian case (meaning

$r = 2$) where it is an open problem to explicitly describe all functions that give rise to codewords.

This paper is organized as follows. Section 2 reviews the necessary terminology and the Hermitian-lifted codes. We define curve-lifted codes in Section 3 and give examples of them. The cardinalities of intersections between the norm-trace curves and lines are found in Section 4, laying a foundation to define an array of norm-trace-lifted codes over arbitrary fields in Section 5. The paper ends with a conclusion in Section 6.

2 Preliminaries

In this section, we review notation to be used throughout the paper as well as the background on other curve-lifted codes in the literature. Throughout, we let q be a power of a prime and $r \geq 2$ be an integer. The finite field with q elements is denoted $\mathbb{F}_q$, and the multiplicative group of its nonzero elements is denoted $\mathbb{F}_q^\times$. The set of nonnegative integers is denoted by $\mathbb{N}$, and for a positive integer n, $[n] := \{1, \ldots, n\}$.

An $[n, k, d]$ **linear code** C over a finite field $\mathbb{F}_q$ is a k-dimensional $\mathbb{F}_q$-subspace of $\mathbb{F}_q^n$ in which any two distinct elements (called codewords) differ in at least d coordinates. Such a code has **length** n, **dimension** k, and **minimum distance** d. Any $d - 1$ erasures in a received word may be recovered by accessing the remaining $n - d + 1$ coordinates. We are interested in recovering erasures by accessing a small number of coordinates. As is standard in the erasure recovery model, we assume a received word w have the form $w \in \mathbb{F}_q^n \cup \{?\}$ where there exists a codeword $c \in C$ such that $w_i \in \{c_i, ?\}$ for all $i \in [n]$. All codes considered in this paper are linear, so we use the terms code and linear code interchangeably.

A code C of length n over $\mathbb{F}_q$ has **locality** s if and only if for all $i \in \{1, \ldots, n\}$, there exists

$$R_i \subseteq [n] \backslash \{i\}$$

where $|R_i| \leq s$ and for all codewords $c \in C$,

$$c_i = \varphi_i(c\,|_{R_i})$$

for some function $\varphi_i : \mathbb{F}_q^s \to \mathbb{F}_q$. Note that a code with locality s clearly has locality s' for all $s' \geq s$. However, we typically use the term to refer to $\max\{|R_i| : i \in [n]\}$. The set R_i is called a **recovery set** for i. The set $\overline{R_i} := R_i \cup \{i\}$ is called a **repair group** for i. We say that C has **availability** t if there exist recovery sets

$$R_{i,1}, \ldots, R_{i,t} \subseteq [n] \backslash \{i\}$$

for each index $i \in [n]$, such that

$$R_{i,j} \cap R_{i,j'} = \emptyset$$

for all $j \neq j'$.

The codes considered in this paper are reminiscent of algebraic geometry codes, in that they are defined using rational points and functions on curves over finite fields. An **algebraic geometry code** $C(D, G)$ of length n over a finite field $\mathbb{F}$ is defined by fixing a curve $\mathcal{X}$ over $\mathbb{F}$ along with divisors G and D on $\mathcal{X}$ so that their supports are disjoint and D is the sum of

$\mathbb{F}$-rational points $P_1, \ldots, P_n$. Each codeword is of the form $(f(P_1), \ldots, f(P_n))$ where f is a function in the Riemann-Roch space of G. If D is not specified, it is taken to be the sum of all $\mathbb{F}$-rational points other than those in the support of G. For example, Reed-Solomon codes are algebraic geometry codes on projective lines. The next most commonly studied algebraic geometry codes are Hermitian codes. Hermitian codes are special cases of norm-trace codes, meaning algebraic geometry codes defined on norm-trace curves. Norm-trace codes were first defined by Geil [5]. They are defined by evaluating functions from Riemann-Roch spaces at rational points on

$$\mathcal{X}_{q,r} : N(x) = Tr(y)$$

over the field $\mathbb{F}_{q^r}$ with q^r elements; here and throughout, the norm and trace will be considered with respect to the extension $\mathbb{F}_{q^r}/\mathbb{F}_q$, so that for any $a \in \mathbb{F}_{q^r}$, they are

$$\text{Norm}(a) := a^{\frac{q^r-1}{q-1}}, \text{Tr}(a) := \sum_{i=0}^{r-1} a^{q^i} \in \mathbb{F}_q.$$

Note that $\mathcal{X}_{q,r}$ has genus

$$g = \frac{1}{2}\left(\frac{q^r-1}{q-1} - 1\right)\left(q^{r-1} - 1\right).$$

Let $\mathcal{X}_{q,r}(\mathbb{F}_{q^r})$ denote the set of $\mathbb{F}_{q^r}$-rational point on the curve $\mathcal{X}_{q,r}$. For each $a \in \mathbb{F}_{q^r}$, there are q^{r-1} elements $b \in \mathbb{F}_{q^r}$ such that $\text{Norm}(a) = \text{Tr}(b)$, each giving rise to a rational point $(a, b) \in \mathcal{X}_{q,r}(\mathbb{F}_{q^r})$, referred to as an affine point of $\mathcal{X}_{q,r}$. In addition, there is a unique point at infinity, $P_\infty \in \mathcal{X}_{q,r}(\mathbb{F}_{q^r})$. Hence, $|\mathcal{X}_{q,r}(\mathbb{F}_{q^r})| = q^r q^{r-1} + 1 = q^{2r-1} + 1$. Consider the vector space of functions $V \subseteq \cup_{m=0}^{\infty} \mathcal{L}(mP_\infty)$ with no poles other than at P_∞, where $P_\infty := (0 : 1 : 0)$ is the unique point at infinity on $\mathcal{X}_{q,r}$. It is worth noting that the Riemann-Roch space of the divisor mP_∞ is

$$\mathcal{L}(mP_\infty) = \left\langle x^i y^j : 0 \leq i \leq \frac{q^r-1}{q-1} - 1, iq^{r-1} + j\frac{q^r-1}{q-1} \leq m \right\rangle \subseteq \mathbb{F}_{q^r}[x, y]. \tag{2.1}$$

We denote the set of polynomials with coefficients in $\mathbb{F}_{q^r}$ and indeterminates x, y of total degree at most k by $\mathbb{F}_{q^r}[x, y]_{\leq k}$. A Hermitian code is an algebraic geometry code over the Hermitian curve. For convenience, we will identify rational functions of the form $\frac{f(x,y)}{z^{\deg f}}$ where $f(x, y) \in \mathbb{F}[x, y]$ with the polynomial $f(x, y) \in \mathbb{F}[x, y]$.

A Hermitian code with $m \leq q^2-1$ over $\mathbb{F}_{q^2}$ has locality q and availability q^2-1 [10]. To see this, note that each affine $\mathbb{F}_{q^2}$-rational point P on the Hermitian curve $\mathcal{H}_q$ has the property that any non-tangent line to $\mathcal{H}_q$ intersects the curve in $q + 1$ $\mathbb{F}_{q^2}$-rational points and there are $q^2 - 1$ such lines. Because any function f in the Riemann-Roch space $\mathcal{L}(mP_\infty)$ may be expressed as $f(x, y) = \sum_{i=0}^{q-1} \sum_{j=0}^{\lfloor \frac{m-iq}{q+1} \rfloor} x^i y^j$, such a function restricted to a non-tangent line can be viewed as a univariate polynomial of degree at most $q - 1$. To recover an erasure at point P_{ab}, one may treat the word corresponding to the $\mathbb{F}_{q^2}$-rational points on any line through P_{ab} as a Reed-Solomon codeword. Thus, the set of $\mathbb{F}_{q^2}$-rational points on any non-tangent line through P_{ab}, other than P_{ab} itself, form a recovery set for the coordinate corresponding to P_{ab}, demonstrating that the Hermitian code $C(mP_\infty, D)$ over $\mathbb{F}_{q^2}$ has locality q and availability $q^2 - 1$. Unfortunately, the rate of these codes approaches 0 as q goes to infinity.

Lifting is a mechanism introduced to increase the rate of codes while maintaining desirable properties. Hermitian-lifted codes over binary fields were introduced in [10] and yield a family of codes with a positive lower bound on the rate as q goes to infinity. More recent work with an improved bound on the rate of Hermitian-lifted codes appears in [1].

3 Curve-lifted codes

In this section, we present curve-lifted codes, which are evaluation codes whose codewords arise from functions that restrict to low degree polynomials on a collection of lines through the curve. In the first subsection, we develop the notion in this section and consider some examples. As we will see, intersection numbers will be a necessary ingredient, and they will be further explored in later sections. In the second subsection, we drill down to better describe the defining sets of functions for particular curves.

3.1 Construction

Consider a projective curve $\mathcal{X}$ given by $F(x, y) = 0$ over the finite field $\mathbb{F}_q$. Take an $\mathbb{F}_q$-rational point $P \in \mathcal{X}(\mathbb{F}_q)$ and enumerate the other $\mathbb{F}_q$-rational points on $\mathcal{X}$: $P_1, \ldots, P_n$. Set $D = P_1 + \cdots + P_n$, and let $V \subseteq \mathbb{F}_q(\mathcal{X})$ be a set of rational function on $\mathcal{X}$ with no poles among $P_1, \ldots, P_n$. Consider the map

$$\begin{array}{rccc} \mathrm{ev}_D\colon & V & \to & \mathbb{F}_q^n \\ & f & \mapsto & (f(P_1), \ldots, f(P_n)). \end{array}.$$

We define an **evaluation code** $C(V, D)$ to be the image of the map ev_D. Its properties will depend on V and $\{P_1, \ldots, P_n\}$.

We will be interested in non-horizontal lines

$$L_{\alpha,\beta}(x, y) : y = \alpha x + \beta$$

with $\alpha, \beta \in \mathbb{F}_q$, so $\alpha \neq 0$. Let

$$\mathbb{L}_q := \{L_{\alpha,\beta} : \alpha, \beta \in \mathbb{F}_q, \alpha \neq 0\}.$$

We consider which rational functions $f \in \mathbb{F}_q(\mathcal{X})$ restrict to certain polynomials on the intersections of particular lines and the curve $\mathcal{X}$. To do so, let

$$m_{\alpha,\beta,F}(x) := F(x, \alpha x + \beta).$$

When F is clear from the context, we may write $m_{\alpha,\beta}(x)$ for $m_{\alpha,\beta,F}(x)$. It is relevant to consider functions f on $\mathcal{X}$ modulo the polynomial $m_{\alpha,\beta}$. More precisely, we use the following notion.

Definition 3.1. For a polynomial $f(t) \in \mathbb{F}_q[t]$, define $\bar{f}_{\alpha,\beta}(t)$ to be the remainder resulting upon division of $f(t)$ by $m_{\alpha,\beta}(t)$; that is,

$$\bar{f}_{\alpha,\beta}(t) := f(t) \bmod m_{\alpha,\beta}(t).$$

Set

$$\deg_{\alpha,\beta}(f) := \deg(\bar{f}_{\alpha,\beta}(t)).$$

Note that

$$\deg\left(\bar{f}_{\alpha,\beta}(t)\right) \leq \deg\left(m_{\alpha,\beta}\right) - 1$$

for all $f \in \mathbb{F}_{q^r}[t]$. We also write $f \equiv_{\alpha,\beta} h$ to mean $f \equiv h \mod m_{\alpha,\beta}$, omitting subscripts if they are clear from the context.

Consider for each point P_i, $i \in [n]$, the set of lines $\mathbb{L}_i$ containing the point P_i. Let

$$B \leq \min\left\{\mid (L \cap \mathcal{X})(\mathbb{F}_q) \mid : L \in \mathbb{L}_i, i \in [n]\right\}.$$

Notice that any line L through P_i intersects $\mathcal{X}$ at at least $B-1$ other $\mathbb{F}_q$-points.

Definition 3.2. Given a curve $\mathcal{X}$ over $\mathbb{F}_q$ with divisor $D := P_1 + \cdots + P_n$ supported by n distinct $\mathbb{F}_q$-rational points, a collection of lines $\mathbb{L} \subseteq \mathbb{L}_q$, and an integer B, the associated curve-lifted code is $C(D, \mathcal{F}_{\mathbb{L},B})$ where

$$\mathcal{F}_{\mathbb{L},B} := \left\{f \in \mathbb{F}_q(\mathcal{X}) : \exists g \in \mathbb{F}_q[t]_{\leq B-2} \text{ with } f \circ L \equiv g \ \forall L \in \mathbb{L}\right\}. \tag{3.1}$$

According to Definition 3.2, the codewords in a curve-lifted code are obtained by evaluating at each point in the support of D functions which restrict on all lines in $\mathbb{L}$ to low-degree polynomials. In the next result, we see that this construction provides locality and availability.

Proposition 3.3. *The curve-lifted code $C(D, \mathcal{F}_{\mathbb{L},B})$ is a code of length $n \leq |\mathcal{X}(\mathbb{F}_q)|$ over $\mathbb{F}_q$ with locality $B-1$ and availability $q-1$.*

Proof. For $i \in [n]$, consider the set of lines $\mathbb{L}_i$ through the $\mathbb{F}_q$-rational point point P_i on $\mathcal{X}$. By definition, any line $L \in \mathbb{L}$ that contains P_i intersects $\mathcal{X}$ in at least $B-1$ other points among the $P_j, j \in [n] \setminus \{i\}$. Let $R_{i,L} \subseteq (L \cap \mathcal{X})(\mathbb{F}_q) \setminus \{P_i\}$ such that $|\mathcal{R}_{i,L}| = B-1$.

Consider a received word w resulting from $\langle()\, f)$ in which there is an erasure in the coordinate corresponding P_i. We claim that $R_{i,L}$ is a recovery set for position i, for all $L \in \mathbb{L}_i$. To demonstrate this fact, we must determine from $R_{i,L}$ the value $f(P_i)$. Observe that for each of the points in the set $R_{i,L} := \left\{P_{j_1}, \ldots P_{j_{B-1}}\right\}$, the value $f(P_{j_t})$ is known. Since $f\mid_L = g$, the values $g\left(P_{j_1}\right), \ldots, g\left(P_{j_{B-1}}\right)$ are known. Because $\deg g \leq B-2$, the polynomial g may be found by interpolation using the $B-1$ values $g\left(P_{j_1}\right), \ldots, g\left(P_{j_{B-1}}\right)$. Then $f(P_i) = g(P_i)$. Hence $R_{i,L}$ is a recovery set for i. Moreover, the intersection of any two such lines L and L' satisfies

$$L \cap L' = \{P_i\}.$$

Thus, the sets $R_{i,L}$, $L \in \mathbb{L}_i$ are disjoint recovery sets indicating that C has availabilty $q-1$. □

Example 3.4. Consider taking $\mathcal{X} = \mathcal{H}_q$, the Hermitian curve over $\mathbb{F}_{q^2}$ where q is even. According to [10], $B = q+1$. Hence, Proposition 3.3 states such codes have length q^3, locality $B-1 = q$, and availability q^2-1. These are precisely the Hermitian-lifted codes considered in [10].

Example 3.5. In this example, we consider norm-trace-lifted codes over fields $\mathbb{F}_{3^r}$ for small values of r. When $r = 3$, we have the curve

$$\mathcal{X}_{3,3} : y^9 + y^3 + y = x^{13}$$

which has genus 48 and 243 $\mathbb{F}_{27}$-rational points other than P_∞. Using [3], we see that each line in affine space over $\mathbb{F}_{27}$ intersects the curve in either 7, 10, or 13 $\mathbb{F}_{27}$-rational points. Taking

$$B = \min\{7, 10, 13\} = 7$$

and $\mathbb{L} = \mathbb{L}_{27}$ gives

$$\mathcal{F}_{\mathbb{L}_{27},7} \coloneqq \{f \in \mathbb{F}_{27}\left(\mathcal{X}_{3,3}\right) : \exists g \in \mathbb{F}_{27}[t]_{\leq 5} \text{ with } f \circ L \equiv g \ \forall L \in \mathbb{L}\}$$

and the code $C(D, \mathcal{F}_{\mathbb{L},7})$ of length 243 with locality 6 and availability $3^3 - 1 = 26$. Codewords are of the form $\mathrm{ev}_D(f)$ where $\deg f \mid_{(L\cap\mathcal{X}_{3,3})(\mathbb{F}_{27})} \leq 5$ for all lines L over $\mathbb{F}_{27}$. As a result, an erasure can be recovered by utilizing only 6 of the other 242 coordinates and in 26 different (disjoint) ways, as each non-horizontal line forms a repair group for each point of intersection with $\mathcal{X}_{3,3}$. We may also note that

$$\left\langle x^a y^b : a + b \leq 5 \right\rangle \subseteq \mathcal{F}_{\mathbb{L}_{27},7},$$

so

$$\left\langle \mathrm{ev}_D(x^a y^b) : a + b \leq 5 \right\rangle \subseteq C(D, \mathcal{F}_{\mathbb{L}_{27},7}).$$

The containment may be strict as in the Hermitian case; it will be further explored in Section 5.

We may also consider taking a proper subset of lines. Calculating intersection numbers [3], we see that for every point $(a, b) \in X_{3,3}(\mathbb{F}_{27})$ with $a \neq 0$, there are

a. 6 lines through (a, b) meeting $\mathcal{X}_{3,3}$ in 13 points

b. 10 lines through (a, b) meeting $\mathcal{X}_{3,3}$ in 7 points

c. 10 lines through (a, b) meeting $\mathcal{X}_{3,3}$ in 10 points.

For the points $(0, b) \in X_{3,3}(\mathbb{F}_{27})$, there are

d. 13 lines through $(0, b)$ meeting $\mathcal{X}_{3,3}$ in 13 points

e. 13 lines through $(0, b)$ meeting $\mathcal{X}_{3,3}$ in 7 points.

Taking $\mathbb{L}$ to be the sets of lines in a. and d. above and setting $B = 13$, Proposition 3.3 applies to give a code $C(D, \mathcal{F}_{\mathcal{L},13})$ which has length 243, locality 12, and availability $\min\{6, 13\} = 6$. Notice that the code $C(D, \mathcal{F}_{\mathcal{L},13})$ includes codewords defined by functions that become polynomials of degree at most 11 when restricted to the lines in a. and d. In particular,

$$\left\langle \mathrm{ev}_D(x^a y^b) : a + b \leq 11 \right\rangle \subseteq C(D, \mathcal{F}_{\mathbb{L}_{27},13}).$$

This suggests that taking only lines which intersect the curve in more points (13 opposed to 7) yields codes of larger dimension, a fact that will be considered in Section 5.

Taking instead $r = 4$, we have the curve

$$\mathcal{X}_{3,4} : y^{27} + y^9 + y^3 + y = x^{40}$$

over $\mathbb{F}_{81}$ which has genus 507 and 2187 affine $\mathbb{F}_{81}$-rational points. Computations [3] indicate that each line in affine space of $\mathbb{F}_{81}$ intersects $\mathcal{X}_{3,4}$ in 22, 28, or 31 $\mathbb{F}_{81}$-rational points. Thus, we take

$$B = \min\{22, 28, 31\} = 22$$

and note that

$$\left\langle x^a y^b : a + b \leq 20 \right\rangle \subseteq \mathcal{F}_{\mathbb{L}_{81},22} = \left\{ f \in \mathbb{F}_{81}\left(\mathcal{X}_{3,4}\right) : \exists g \in \mathbb{F}_{81}[t]_{\leq 20} \text{ with } f \circ L \equiv g \ \forall L \in \mathbb{L} \right\}.$$

Hence, $C(D, \mathcal{F}_{\mathbb{L}_{81},22})$ is a code over $\mathbb{F}_{81}$ of length 2187, locality 21, and availability 80. It follows that an erasure can be recovered by utilizing only 20 of the 2186 other coordinates and in 80 different (disjoint) ways.

Example 3.6. In this example, we consider the curve $\mathcal{X}$ given by $y^8 + y = x^3$ over $\mathbb{F}_{64}$, from the first family of non-classical curves described by Schmidt [17]. One may note that $\mathcal{X}$ is maximal [4] and a so-called Castle curve [14]. Note that $\mathcal{X}$ which has genus 7 and 176 $\mathbb{F}_{64}$-rational points other than P_∞. Using [3], we see that each line in affine space over $\mathbb{F}_{64}$ intersects the curve in either 1, 2, 4, or 7 $\mathbb{F}_{64}$-rational points. Taking

$$B = \min\{1, 2, 4, 7\} = 1$$

and $\mathbb{L} = \mathbb{L}_{64}$ gives

$$\mathcal{F}_{\mathbb{L}_{64},1} = \emptyset.$$

Thus, to obtain curve-lifted codes on $\mathcal{X}$, we must take a proper subset of lines in $\mathbb{L}_{64}$. Calculating intersection numbers [3], we see that there are 168 points $P_1, \ldots, P_{168}$ such that for each P_i, $i \in [168]$, there are

a. 3 lines whose only point of intersection with the curve $\mathcal{X}$ in P_i,

b. 12 lines through P_i meeting $\mathcal{X}$ in 2 points,

c. 11 lines through P_i meeting $\mathcal{X}$ in 3 points,

d. 30 lines through P_i meeting $\mathcal{X}$ in 4 points, and

e. 7 lines through P_i meeting $\mathcal{X}$ in 7 points.

In addition, there are 8 points $P_{169}, \ldots, P_{176}$ such that for each P_i, $i \in \{169, \ldots, 176\}$, there are

f. 21 lines through P_i meeting $\mathcal{X}$ in 3 points and

g. 42 lines meeting the curve $\mathcal{X}$ in 4 points.

To design a curve-lifted code on $\mathcal{X}$ with locality 3, one could take the set $\mathbb{L}$ to consist of those lines in d. and g. above. In this case, the code $C(P_1 + \cdots + P_{176}, \mathcal{F}_{\mathbb{L},4})$ has length 176 over $\mathbb{F}_{64}$, locality 3, and availability $\min\{30, 42\} = 30$. It is formed by taking those functions $f \in \mathbb{F}_{64}(\mathcal{X})$ that reduce to quadratics on the intersection of each line $L \in \mathbb{L}$ and the curve $\mathcal{X}$.

To increase the dimension, we might consider increasing the locality. However, there are no lines which contain any of the points P_i, $i \in \{169, \ldots, 176\}$, and intersect the curve in more than 4 points. Consequently, Definition 3.2 does not support the design of a length 176 code over $\mathbb{F}_{64}$ which has locality greater than 3. Even so, we note that each of the points P_i, $i \in [168]$, lies on 7 lines that intersect the curve in 7 points as specified in e. above. The Proposition 3.3 implies that $C(P_1 + \cdots + P_{168}, \mathcal{F}_{\mathbb{L},7})$ has length 168 over $\mathbb{F}_{64}$, locality 5, and availability 7.

As these examples show, a key component of these curve-lifted constructions is the determination of the number of points of intersection between the curve and lines. In Examples 3.5 and 3.6, computational tools were used to determine them for particular curves over specific fields. In order to determine infinite families of such codes, we need more sophisticated theoretical tools. We will demonstrate that in the next section where we bound intersection numbers of sets of lines on the norm-trace curve.

One may also note that the traditional code parameters dimension and minimum distance are absent in Proposition 3.3. Because locally recoverable codes are designed for erasure recovery using small sets of other coordinates, the minimum distance is not as relevant as in standard error correction or recovery of collections of erasures using the entire received word. However, rates for families of LRCs provide a useful gauge of their capabilities. To examine code rates, we will also need the more information on particular curves. We make some headway on this front in the next subsection.

Definition 3.7. A monomial $M_{a,b}(x, y)$ is said to be **good** for $\mathcal{F}_{\mathbb{L},B}$ if for all lines $L_{\alpha,\beta} \in \mathbb{L}$,

$$\deg_{\alpha,\beta}(M_{a,b} \circ L_{\alpha,\beta}) \leq B - 2.$$

We may simply say that a monomial is good if the set of lines $\mathbb{L}$ and integer B are clear from the context. Some monomials are good regardless of the choice of $\mathbb{L}$. For instance, $M_{a,b}(x, y)$ is good for $\mathcal{F}_{\mathbb{L},B}$ for each $(a, b) \in \mathbb{N}^2$ with $a + b \leq B - 2$, for all $\mathbb{L} \subseteq \mathbb{L}_q$

3.2 Sporadic monomials

In this subsection, we study monomials $x^a y^b$ which are good but not simply because $a + b$ is small enough (as mentioned above). This notion is made precise in the following definition.

Definition 3.8. A monomial $x^a y^b$ is called **sporadic** for $\mathcal{F}_{\mathbb{L},B}$ if it is good for $\mathcal{F}_{\mathbb{L},B}$ and $a + b \geq B - 1$. A monomial good for $\mathcal{F}_{\mathbb{L},B}$ is called **typical** if it is not sporadic.

Example 3.9. Recall that Hermitian-lifted codes have locality q and we may think of them as $C(D, \mathcal{F}_{\mathbb{L}_{q^2},q+1})$ where $D = P_1 + \cdots + P_{q^3}$ is supported by the $\mathbb{F}_{q^2}$-rational points on $X_{q,2} : y^q + y = x^{q+1}$. The set of typical monomials is

$$\left\{x^a y^b : a + b \leq q - 1\right\}.$$

Loosely speaking, Hermitian-lifted codes are defined with two sets of monomials $x^a y^b$: those with $a + b \leq q - 1$, meaning degree less than the locality which are always good, and some with $a + b \geq q$ that happen to reduce to those of degree less than locality on all lines. The monomials in the latter set were called sporadic, since their behavior is not yet fully understood, meaning to date, only some of them have been described explicitly. For instance, according to [10, Theorem 10],

$$\left\{x^a y^b : \begin{array}{c} a \leq q-1, b \leq q^2-1, a+b \geq q, \exists l, i \in [l], 0 \leq s \leq i-1 \text{ so that } b = wq + b' \\ \text{with } b' < 2^{l-1}, 2^i \mid w, a < 2^{l-1}, \text{ and no } 2^s \text{ term in binary expansions of } a \text{ and } b' \end{array}\right\}$$

is a subset of sporadic monomials.

In the Hermitian-lifted case, accounting for the sporadic monomials is necessary to determine that the codes have a rate bounded away from 0 as the code length grows, as demonstrated in [10]. We will see that the $r > 2$ case for the norm-trace-lifted codes is quite different. While there may be sporadic monomials, due to the larger locality (or B value used), enough typical monomials may be found to demonstrate an even better bound on the asymptotic rate. To better understand the functions that define codewords of curve-lifted codes, we will use the following observation.

Lemma 3.10. *Let $f(x) = x^m + g(x)$ where $\deg g < m$. Assume $d \geq m$. Then the remainder of x^d after division by f has degree at least $\deg g$.*

Proof. Consider the set of integers $A = \{k \geq 1 : d + (k-1)(\deg g - m) \geq m\}$. This set contains $k = 1$, so it is nonempty. Since $\deg g - m < 0$, the sequence $d + (k-1)(\deg g - m)$, $k = 1, 2, \ldots$ is a strictly decreasing sequence of integers, so A contains a maximal element. Let k be a maximal element of A. If we write $x^d = q(x)f(x) + r(x)$ with $0 \leq \deg r < \deg f$, then

$$q(x) = x^{d-m} - x^{d-2m}g + x^{d-3m}g^2 - \cdots + (-1)^{k-1}x^{d-km}g^{k-1}$$

and $r(x) = (-1)^k x^{d-km} g^k$. Since k is the smallest integer such that $\deg r = d + k(\deg g - m) < m$, we see that $\deg r$ must be the unique integer in the interval $[\deg g, m)$ congruent to d modulo $m - \deg g$. In particular, $\deg r \geq \deg g$. □

The next result provides insight into when we might expect to find sporadic monomials; a variant of which may be found in [15],

Proposition 3.11. *Consider a curve $\mathcal{X}$ given by $F(x, y) = 0$ over a finite field $\mathbb{F}_q$ and a collection $\mathbb{L} \subseteq \mathbb{L}_q$ of lines containing $L_{1,0}$. Let d denote the degree of the second highest degree term of F. Suppose a curve-lifted code $C(D, \mathcal{F}_{,B})$ is defined on $\mathcal{X}$ for some integer B, selected so that each line in $\mathbb{L}$ intersects $\mathcal{X}$ in at least B $\mathbb{F}_q$-rational rational points. If $B - 1 \leq d$, then there are no sporadic good monomials for $\mathcal{F}$.*

Proof. Consider a good monomial $M_{a,b}$. Note that the line $L_{1,0}(x) = (x, x)$ gives $(M_{a,b} \circ L_{1,0})(x) = x^{a+b}$ and $L_{1,0} \in \mathbb{L}$. The intersection of the line defined by $x = y$ with $\mathcal{X}$ is cut out by the equation $F(x, x) = 0$, i.e.

$$m_{1,0}(x) = F(x, x) = 0.$$

We are interested in the degree of $\overline{x^{a+b}} = x^{a+b} \mod m_{1,0}$. Assume $a + b \geq B - 1$, meaning $M_{a,b}$ is sporadic. We consider two cases, depending on the value $a + b$. First, suppose $a + b < \deg F$. Then

$$\overline{x^{a+b}} = x^{a+b},$$

since $\deg x^{a+b} = a + b < \deg F$. However, it then follows that

$$\deg_{1,0}(x^{a+b}) = a + b \geq B - 1 > B - 2$$

which shows $M_{a,b}$ is not good for $\mathcal{F}$. Hence, there are no sporadic monomials x^{a+b} good for $\mathcal{F}$ where $a + b < \deg F$.

Now suppose $a + b \geq \deg F$. Then the remainder of x^{a+b} upon division by $m_{1,0}(x)$ has degree at least d by Lemma 3.10. Consequently,

$$\deg_{1,0}(x^{a+b}) \geq d \geq B - 1 > B - 2,$$

so there are no sporadic monomials x^{a+b} good for $\mathcal{F}_{\mathbb{L},B}$. □

4 Intersection numbers of norm-trace curves

In this section, we determine the number of points in $\mathcal{X}_{q,r}(\mathbb{F}_{q^r})$ on an intersection of a line $L_{\alpha,\beta}$ where $\alpha \neq 0$ and $\alpha, \beta \in \mathbb{F}_{q^r}$ with the norm-trace curve $\mathcal{X}_{q,r}$. We loosely refer to the number of such points as an intersection number. Intersection numbers will be applied in Section 5 to construct norm-trace-lifted codes over fields of arbitrary characteristic. In particular, they will be used to set the value B as in Proposition 3.3 and a degree bound which will define an appropriate set of functions to support local recovery with high availability and positive rate. In particular, we will find an integer B, depending only on q and r, such that for all $L_{\alpha,\beta} \in \mathbb{L}_{q^r}$,

$$B \leq \# \left(L_{\alpha,\beta} \cap \mathcal{X}_{q,r}\right)\left(\mathbb{F}_{q^r}\right)$$

which will ultimately play a role in the locality of the norm-trace-lifted codes.

Given $\alpha, \beta \in \mathbb{F}_{q^r}$, it will be useful to consider the polynomial

$$m_{\alpha,\beta,q,r}(x) \coloneqq x^{(q^r-1)/(q-1)} - \mathrm{Tr}(\beta) - \sum_{i=0}^{r-1} (\alpha x)^{q^i} \in \mathbb{F}_{q^r}[x]_{\leq \frac{q^r-1}{q-1}},$$

or $m_{\alpha,\beta}$ for short. Define

$$n_{q,r}(\alpha, \beta) \coloneqq \#(L_{\alpha,\beta} \cap \mathcal{X}_{q,r})(\mathbb{F}_{q^r}).$$

First, we observe that if α and α' are nonzero elements with the same norm and β and β' have the same trace, then $n_{q,r}(\alpha, \beta) = n_{q,r}(\alpha', \beta')$.

Lemma 4.1. *For any $\alpha \in \mathbb{F}_{q^r}^{\times}$ and $\beta \in \mathbb{F}_{q^r}$, $n_{q,r}(\alpha, \beta)$ depends only on* $\mathrm{Tr}(\alpha)$ *and* $\mathrm{Norm}(\beta)$.

Proof. First, plug in $y = \alpha x + \beta$ into the equation for $\mathcal{X}_{q,r}$ to obtain

$$x^{(q^r-1)/(q-1)} = \sum_{i=0}^{r-1} (\alpha x + \beta)^{q^i} = \sum_{i=0}^{r-1} (\alpha x)^{q^i} + \beta^{q^i} = \mathrm{Tr}(\beta) + \sum_{i=0}^{r-1} (\alpha x)^{q^i}.$$

Then notice that $n_{q,r}(\alpha,\beta)$ is the number of zeros in $\mathbb{F}_{q^r}$ of the polynomial $m_{\alpha,\beta}(x)$. Making the substitution $t = \alpha^{-1}x$ yields the polynomial

$$\begin{aligned} g(t) &= m_{\alpha,\beta}(\alpha^{-1}x) \\ &= (\alpha t)^{(q^r-1)/(q-1)} - \mathrm{Tr}(\beta) - \sum_{i=0}^{r-1} t^{q^i} \\ &= \mathrm{Norm}(\alpha) t^{(q^r-1)/(q-1)} - \mathrm{Tr}(\beta) - \sum_{i=0}^{r-1} t^{q^i}. \end{aligned}$$

The number of zeros in $\mathbb{F}_{q^r}$ of g and $m_{\alpha,\beta}$ are the same since the map $t \mapsto \alpha^{-1}x$ is a bijection sending zeros of $m_{\alpha,\beta}$ to zeros of g, and the number of zeros of g depends only on $\mathrm{Tr}(\beta)$ and $\mathrm{Norm}(\alpha)$. □

Next, we apply Lemma 4.1 to the case where $q = 2$; see also [11, Lemma 1].

Proposition 4.2. *Let $r \geq 2$ and $\alpha, \beta \in \mathbb{F}_{2^r}$ with $\alpha \neq 0$. If $\mathrm{Tr}(\beta) \neq 0$, then*

$$n_{2,r}(\alpha,\beta) = 2^{r-1} - 1.$$

If $\mathrm{Tr}(\beta) = 0$, then

$$n_{2,r}(\alpha,\beta) = 2^{r-1} + 1.$$

Thus, for any line $L_{\alpha,\beta} \in \mathbb{L}_{2,r}$, the cardinality of its intersection with the norm-trace curve $\mathcal{X}_{2,r}$ over $\mathbb{F}_{2^r}$ is

$$|L_{\alpha,\beta} \cap \mathcal{X}_{2,r}(\mathbb{F}_{2^r})| = 2^{r-1} \pm 1.$$

Proof. To determine $n_{2,r}(\alpha,\beta)$, according to Lemma 4.1, we need only consider the two cases, depending on $\mathrm{Tr}(\beta) = 0$ or $\mathrm{Tr}(\beta) = 1$.

Notice that points in the intersection $L_{\alpha,\beta} \cap \mathcal{X}_{2,r}(\mathbb{F}_{2^r})$ correspond to roots of the polynomial

$$x^{2^r-1} - (\alpha x + \beta)^{2^{r-1}} + \cdots + (\alpha x + \beta)^2 + (\alpha x + \beta)$$

which are also elements of $\mathbb{F}_{2^r}$. Because the roots of $x^{2^r-1} - x$ are precisely the $\gamma \in \mathbb{F}_{2^r}$, the problem of determining $n_{2,r}$ reduces to finding the degree of $h(x) = \gcd(m_{\alpha,\beta}(x), x^{2^r} - x)$ by Freshman's Dream.

In the case $\mathrm{Tr}(\beta) = 0$, the Euclidean Algorithm reveals

$$\gcd(m_{\alpha,\beta}(x), x^{2^r} - x) = \alpha^{2^{r-1}} x^{2^{r-1}+1} + \cdots + \alpha x^2 + x$$

which has degree $2^{r-1} + 1$. In the case $\mathrm{Tr}(\beta) = 1$,

$$\gcd(m_{\alpha,\beta}(t), x^{2^r} - x) = \alpha^{2^{r-1}} x^{2^{r-1}-1} + \cdots + \alpha,$$

which has degree $2^{r-1} - 1$. Therefore, $n_{2,r}(\alpha,\beta) = 2^{r-1} \pm 1$. □

For $q \neq 2$, we observe more intricate behavior. We will use results of Moisio and Moisio-Wan, who build on work of Katz [9], to establish lower and upper bounds on $n_{q,r}(\alpha, \beta)$. For an integer $r \geq 2$ and elements $a, b \in \mathbb{F}_q$, let

$$N_{q,r}(a,b) = \#\{\alpha \in \mathbb{F}_{q^r} : \mathrm{Tr}(\alpha) = a, \mathrm{Norm}(\alpha) = b\}.$$

In [9], Katz proves the following result on counting elements of $\mathbb{F}_{q^r}$ with prescribed norm and trace:

Lemma 4.3. *[9, Theorem 4] If $r \geq 2$ and $a, b \in \mathbb{F}_q^\times$, then*

$$\left| N_{q,r}(a,b) - \frac{q^r - 1}{q(q-1)} \right| \leq r q^{(r-2)/2}.$$

We will use the following improvement to Katz' result when $a \neq 0$, due to Moisio and Wan [13].

Lemma 4.4. *[13, Theorem 1.2] Let $a, b \in \mathbb{F}_q^\times$ and $r \geq 2$. Then*

$$\left| N_{q,r}(a,b) - \frac{q^{r-1} - 1}{q-1} \right| \leq (r-1) q^{(r-2)/2}.$$

The previous bound is complemented by the following result of Moisio [12], which provides a bound $N_{q,r}(a,b)$ when $a = 0$.

Lemma 4.5. *[12] Let $r \geq 2$, $b \in \mathbb{F}_q^\times$, and $d = \gcd(r, q-1)$. We have*

$$\left| N_{q,r}(0,b) - \frac{q^{r-1} - 1}{q-1} \right| \leq (d-1) q^{(r-2)/2}.$$

We now use [12, 13] to produce lower bounds on $n_{q,r}(\alpha, \beta)$.

Theorem 4.6. *Let $r \geq 2$, $\alpha, \beta \in \mathbb{F}_{q^r}$ with $\alpha \neq 0$, and $d = \gcd(r, q-1)$. If $\mathrm{Tr}(\beta) \neq 0$, then*

$$n_{q,r}(\alpha, \beta) \geq q^{r-1} - (d - 1 + (r-1)(q-2)) q^{(r-2)/2} - 1.$$

If $\mathrm{Tr}(\beta) = 0$, then

$$n_{q,r}(\alpha, \beta) \geq q^{r-1} - (r-1)(q-1) q^{(r-2)/2}.$$

Proof. Let $a = \mathrm{Norm}(\alpha)^{-1}$ and $b = \mathrm{Tr}(\beta)$. Let $g(t) = a^{-1} t^{(q^r-1)/(q-1)} - b - \sum_{i=0}^{r-1} t^{q^i}$. We have that $n_{q,r}(\alpha, \beta)$ is the number of roots of g in $\mathbb{F}_{q^r}$. Note that the roots of g are in bijection with the set

$$\bigcup_{t \in \mathbb{F}_q} \{\gamma \in \mathbb{F}_{q^r} : \mathrm{Tr}(\gamma) = t \text{ and } \mathrm{Norm}(\gamma) = at + ab\}.$$

Thus the number of roots of g (and hence $n_{q,r}(\alpha, \beta)$) is given by

$$\begin{aligned} n_{q,r}(\alpha, \beta) &= \#\{\gamma \in \mathbb{F}_{q^r} : g(\gamma) = 0\} \\ &= \sum_{t \in \mathbb{F}_q} \#\{\gamma \in \mathbb{F}_{q^r} : \mathrm{Tr}(\gamma) = t \text{ and } \mathrm{Norm}(\gamma) = at + ab\} \\ &= \sum_{t \in \mathbb{F}_q} N_r(t, at + ab). \end{aligned}$$

First, assume $b \neq 0$. When $t = -b$, we have $N_{q,r}(t, at + ab) = N_{q,r}(-b, 0) = 0$. If $t \neq -b$, then $at + ab \neq 0$ so we may apply Lemma 4.4 to bound $N_{q,r}(t, at + ab)$. Since $\alpha \neq 0$, we may apply Lemma 4.5 to bound $N_{q,r}(0, ab)$. We therefore obtain the desired lower bound for $n_{q,r}(\alpha, \beta)$:

$$\begin{aligned}
n_{q,r}(\alpha,\beta) &= N_{q,r}(0,ab) + N_{q,r}(-b,0) + \sum_{\substack{t\in\mathbb{F}_q \\ t\neq 0,-b}} N_r(t, at+ab) \\
&= N_{q,r}(0,ab) + \sum_{\substack{t\in\mathbb{F}_q \\ t\neq 0,-b}} N_r(t, at+ab) \\
&\geq \frac{q^{r-1}-1}{q-1} - (d-1)q^{(r-2)/2} + \sum_{\substack{t\in\mathbb{F}_q \\ t\neq 0,-b}} \frac{q^{r-1}-1}{q-1} - (r-1)q^{(r-2)/2} \\
&= \frac{q^{r-1}-1}{q-1} - (d-1)q^{(r-2)/2} + (q-2)\left(\frac{q^{r-1}-1}{q-1} - (r-1)q^{(r-2)/2}\right) \\
&= \frac{q^{r-1}-1}{q-1}(1+q-2) - q^{(r-2)/2}(d-1+(q-2)(r-1)) \\
&= q^{r-1} - (d-1+(q-2)(r-1))q^{(r-2)/2} - 1.
\end{aligned}$$

Now assume $b = 0$. Then applying Lemma 4.4 to each $N_{q,r}(t, at)$ for $t \neq 0$ we get

$$\begin{aligned}
n_{q,r}(\alpha,\beta) &= N_{q,r}(0,0) + \sum_{\substack{t\in\mathbb{F}_q \\ t\neq 0}} N_r(t, at) \\
&\geq 1 + \sum_{\substack{t\in\mathbb{F}_q \\ t\neq 0}} \frac{q^{r-1}-1}{q-1} - (r-1)q^{(r-2)/2} \\
&= 1 + (q-1)\left(\frac{q^{r-1}-1}{q-1} - (r-1)q^{(r-2)/2}\right) \\
&= q^{r-1} - (r-1)(q-1)q^{(r-2)/2}.
\end{aligned}$$

□

Corollary 4.7. *For any line $L_{\alpha,\beta} \in \mathbb{L}_{q^r}$, the cardinality of its intersection with the norm-trace curve $\mathcal{X}_{q,r}$ satisfies*

$$|L_{\alpha,\beta} \cap \mathcal{X}_{q,r}(\mathbb{F}_{q,r})| \geq q^{r-1} - (r-1)(q-1)q^{(r-2)/2} - 1.$$

There are $q^r - 1$ lines in $\mathbb{L}_{q^r}$. Let $d = \gcd(r, q-1)$. For those lines $L_{\alpha,\beta} \in \mathbb{L}_{q^r}$ with $Tr(\beta) \neq 0$, the cardinality of the intersection of $L_{\alpha,\beta}$ with the norm-trace curve $\mathcal{X}_{q,r}$ satisfies

$$|L_{\alpha,\beta} \cap \mathcal{X}_{q,r}(\mathbb{F}_{q,r})| \geq q^{r-1} - (r-1)(q-1)q^{(r-2)/2} - 1 + q^{\frac{r-2}{2}}$$

if $r \neq d$. There are $q^r - q^{r-1} = q^{r-1}(q-1)$ such lines.

Proof. Notice that $d = \gcd(r, q-1) \leq r$. For convenience, let $N = q^{r-1} - (q-1)(r-1)q^{(r-2)/2} - 1$, which is the right-hand side of the lower bound on the intersection number for lines $L_{\alpha,\beta}$ with $\text{Tr}(\beta) \neq 0$, and $Z = q^{r-1} - (d - 1 + (q-2)(r-1))q^{(r-2)/2} - 1$, which is the right-hand side of the lower bound on the intersection number for lines $L_{\alpha,\beta}$ with $\text{Tr}(\beta) = 0$. Then for all $\alpha, \beta \in \mathbb{F}_{q^r}$, $n_{q,r}(\alpha, \beta) \geq \min\{N, Z\}$. Calculating the difference between the values given in the two lower bounds, we see that

$$N - Z = (r-d)q^{\frac{r-2}{2}} - 1 = \begin{cases} -1 & \text{if } d = r \\ q^{\frac{r-2}{2}} - 1 + t, \text{ for some } t \in \mathbb{N} & \text{otherwise.} \end{cases}$$

Therefore,

$$N = \begin{cases} Z - 1 & \text{if } d = r \\ Z - 1 + q^{\frac{r-2}{2}} + t, \text{ for some } t \in \mathbb{N} & \text{otherwise.} \end{cases}$$

Consequently, every line $L_{\alpha,\beta} \in \mathbb{L}$ satisfies

$$n_{q,r}(\alpha, \beta) \geq \min\{N, Z\} \geq Z - 1.$$

□

We will use the bound in the previous corollary to establish families of norm-trace-lifted codes in arbitrary characteristic.

Remark 4.8. Notice Theorem 4.6 and Corollary 4.7 provide little to no information in the case $r = 2$. Indeed, $q^{r-1} - (r-1)(q-1)q^{(r-2)/2} - 1 = q - (q-1) = 1$. Moreover, in the case that q is even, $d = \gcd(2, q-1) = 1$ and $q^{r-1} - (d - 1 + (r-1)(q-2))\, q^{(r-2)/2} - 1 = q - (1 - 1 + (2-1)(q-2))\, q^{(2-2)/2} = 2$. Furthermore, if q is odd, then $d = 2$ and $q^{r-1} - (d - 1 + (r-1)(q-2))\, q^{(r-2)/2} - 1 = q - (2 - 1 + (2-1)((q-2))\, q^{(2-2)/2} = 1$. Hence, these lower bounds are quite poor, since it is known that the actual value for $n_{q,2}(\alpha, \beta) = q+1$.

5 Norm-trace-lifted codes

In this section, we combine the results from Sections 3 and 4 to define and study norm-trace-lifted codes defined using the norm-trace curve $X_{q,r} : Tr(y) = N(x)$ over $\mathbb{F}_{q^r}$ where q is any prime power. In light of Remark 4.8 and [10], we restrict our attention to $r > 2$. Now, having found a lower bound to take for B in Proposition 3.3 on intersection numbers as in Corollary 4.7, we may now consider codes defined by sets of rational functions that reduce on all lines to polynomials of degree at most $B - 2$. Also, because the curve itself is specified by a particular equation, we can obtain bounds on the code rates, including some that exceed the comparable Hermitian cases.

5.1 Constructions with highest availability

To obtain codes from the norm-trace curve with highest availability, we use points on every line through a point to form a repair group so as to obtain the largest number of disjoint recovery sets. With that in mind, consider taking $B = B_{q,r}$ where

$$B_{q,r} := \begin{cases} q^{r-1} - 1 & \text{if } q = 2 \\ q^{r-1} - (r-1)(q-1)q^{\frac{r-2}{2}} - 1 & \text{otherwise} \end{cases}$$

with the goal of forming a repair group $R_{ab,L}$ for an affine point $P_{ab} := (a,b)$ from each line $L \in \mathbb{L}_{q^r}$ through P_{ab}. We will use the shorthand notation $\mathcal{F}_{q,r} := \mathcal{F}_{\mathbb{L}_{q^r},B_{q,r}}$ so that

$$\mathcal{F}_{q,r} = \left\{ f \in \mathbb{F}_{q^r}[x,y] : \exists g \in \mathbb{F}_{q^r}[t]_{\leq B_{q,r}-2} \text{ with } f \circ L_{\alpha,\beta} \equiv g \text{ for all } L_{\alpha,\beta} \in \mathbb{L}_{q^r} \right\}.$$

To do so, we will take a subset $R_{ab,L} \subseteq L \cap \mathcal{X}_{q,r}(\mathbb{F}_{q^r})$ such that $P_{ab} \in R_{ab,L}$ and $|R_{ab,L}| = B_{q,r}$. Such a subset exists by Corollary 4.7. Recall that

$$m_{\alpha,\beta}(t) := t^{(q^r-1)/(q-1)} - \operatorname{Tr}(\beta) - \sum_{i=0}^{r-1} (\alpha t)^{q^i} \in \mathbb{F}_{q^r}[t]_{\leq \frac{q^r-1}{q-1}}.$$

Note that

$$\deg\left(\bar{f}_{\alpha,\beta}(t)\right) \leq q^{r-1} + q^{r-2} + \cdots + q$$

for all $f \in \mathbb{F}_{q^r}[t]$, as $\deg(m_{\alpha,\beta}(t)) = \frac{q^r-1}{q-1}$.

According to (2.1), rational functions on $\mathcal{X}_{q,r}$ with no poles at any of the affine points are elements of $\mathbb{F}_{q^r}[x,y]$, since $\cup_{m\in\mathbb{N}}\mathcal{L}(mP_\infty) \subseteq \mathbb{F}_{q^r}[x,y]$.

Definition 5.1. The **norm-trace-lifted code** defined over $\mathbb{F}_{q^r}$ is $C(D, \mathcal{F}_{q,r})$, the image of $\mathcal{F}_{q,r}$ under the evaluation map ev; that is,

$$C(D, \mathcal{F}_{q,r}) := \{\operatorname{ev}_D(f) : f \in \mathcal{F}_{q,r}\} \subseteq \mathbb{F}_{q^r}^n.$$

Clearly, $C(D, \mathcal{F}_{q,r})$ is a code of length n. To ascertain its dimension, we set out to determine the functions in $\mathcal{F}_{q,r}$. Based on the definition of $\mathcal{F}_{q,r}$, we are interested in polynomials $f(x,y)$ such that

$$\deg_{\alpha,\beta}(f \circ L_{\alpha,\beta}) \leq B_{q,r} - 2$$

for all $L_{\alpha,\beta} \in \mathbb{L}_{q^r}$. Recall that $M_{a,b}(x,y) := x^a y^b$ where $a, b \in \mathbb{N}$.

The number of $(a,b) \in \mathbb{N}^2$ with $a + b \leq B$ for some specified positive integer B is $\sum_{a=0}^{B} B - a + 1 = \frac{1}{2}(B+1)(B+2)$. To ensure that the set of such monomials gives rise to an independent set of codewords, we appeal to a result recorded in [15].

Lemma 5.2. *[15, Lemma 2.14] The set of vectors*

$$\left\{ \operatorname{ev}(M_{a,b}(x,y)) : 0 \leq a \leq \frac{q^r-1}{q-1} - 1, 0 \leq b \leq q^{r-1} - 1 \right\}$$

are linearly independent.

Proof. Drawing inspiration from the proof of Proposition 5 of [10], we observe that the kernel of the evaluation map ev is generated by $x^{\frac{q^r-1}{q-1}} - y^{q^{r-1}} - \cdots - y^q - y$, $x^{q^r} - x$, and $y^{q^r} - y$. Under monomial orderings with $x^{\frac{q^r-1}{q-1}} < y^{q^{r-1}}$,

$$\left\{ x^{\frac{q^r-1}{q-1}} - y^{q^{r-1}} - \cdots - y^q - y, x^{q^r} - x \right\}$$

is a Gröbner basis for the kernel of the evaluation map, and so the evaluations of $M_{a,b}$ cannot contain any element from the kernel of the evaluation map. Thus, the evaluations of $M_{a,b}$ are linearly independent. □

Theorem 5.3. *The norm-trace-lifted code $C(D, \mathcal{F}_{q,r})$ over $\mathbb{F}_{q^r}$ is a code of length q^{2r-1}, dimension at least*

$$\frac{1}{2}q^{r/2-1}\left(q^{r-1}-(q-1)(r-1)q^{r/2-1}+1\right)\left(q^{r/2}-qr+q+r-1\right),$$

locality $q^{r-1}-(r-1)(q-1)q^{\frac{r-2}{2}}-2$, and availability q^r-1 with rate approaching $\frac{1}{2q}$ as the code length grows for a fixed characteristic.

Proof. Given an erasure in the coordinate corresponding to an affine point P_{ab} on $\mathcal{X}_{q,r}$, consider the set of lines $\mathcal{L}$ through P_{ab}. According to Corollary 4.7, any line $L \in \mathcal{L}$ intersects $\mathcal{X}_{q,r}$ in at least $B_{q,r}-1$ other affine points. Moreover, any function $f \in \mathcal{F}_{q,r}$ has the property that $f_{|L} \equiv g$ where $\deg g \leq B_{q,r}-2$. We claim that the $B_{q,r}-1$ affine points in the intersection $L \cap X_{q,r}$ other than P_{ab} may be used to interpolate and find g; that is, we claim that $L \cap X_{q,r} \setminus \{P_{ab}\}$ is a recovery set for the coordinate associated with P_{ab}. Evaluating $g(P_{ab})$ allows for recovery of the erased coordinate using $B_{q,r}-1$ points. Hence, the locality of $C(D, \mathcal{F}_{q,r})$ with this choice of recovery sets is $B_{q,r}-1 = q^{r-1}-(r-1)(q-1)q^{\frac{r-2}{2}}-2$. Because there are q^r-1 such lines L, $C(D, \mathcal{F}_{q,r})$ has availability q^r-1.

Notice that the set $S := \left\{x^a y^b : a+b \leq B_{q,r}-2\right\}$ is a set of monomials good for $\mathcal{F}_{q,r}$ of cardinality

$$\frac{1}{2}\left(q^{r-1}-(r-1)(q-1)q^{\frac{r-2}{2}}\right)\left(q^{r-1}-(r-1)(q-1)q^{\frac{r-2}{2}}+1\right).$$

According to Lemma 5.2, S is linearly independent, demonstrating that $C(D, \mathcal{F}_{q,r})$ has dimension at least

$$\frac{1}{2}\left(q^{r-1}-(r-1)(q-1)q^{\frac{r-2}{2}}\right)\left(q^{r-1}-(r-1)(q-1)q^{\frac{r-2}{2}}+1\right).$$

and rate

$$\frac{\frac{1}{2}\left(q^{r-1}-(r-1)(q-1)q^{\frac{r-2}{2}}\right)\left(q^{r-1}-(r-1)(q-1)q^{\frac{r-2}{2}}+1\right)}{q^{2r-1}} \to \frac{1}{2q} > 0$$

as $r \to \infty$. □

Note that the rate is bounded away from 0 for fixed characteristic as the degree of the extension grows. Hence, the codes over fields of small characteristic provide the best asymptotic rates. We also observe that taking functions that reduce to low degree polynomials on all lines (rather than just some) provides the largest availability given by the lifted construction, since the lines are the recovery sets. We may also consider fewer lines with more refined intersection numbers, as in the next subsection.

5.2 Constructions with high availability and larger dimension

As suggested by Proposition 3.3, we may adapt the collections of lines $\mathcal{L}$ used in Definition 5.1 and the value B to a more refined bound on the intersection numbers. Consider

$$B'_{q,r} := q^{r-1} - (\gcd(r, q-1) - 1 + (r-1)(q-2))\, q^{(r-2)/2} - 1$$

and the collection of lines

$$\mathbb{L}'_{q^r} = \{L_{\alpha,\beta} : \mathrm{Tr}(\beta) \neq 0\}.$$

Then set $\mathcal{F}'_{q,r} := \mathcal{F}_{\mathbb{L}'_{q^r}, B'_{q,r}}$ so that

$$\mathcal{F}'_{q,r} = \left\{ f \in \mathbb{F}_{q^r}[x,y] : \exists g \in \mathbb{F}_{q^r}[t]_{\leq B'_{q,r}-2} \text{ with } f \circ L_{\alpha,\beta} \equiv g \text{ for all } L_{\alpha,\beta} \in \mathbb{L}'_{q^r}, \right\}.$$

We say the **refined norm-trace-lifted code** is

$$C\left(D, \mathcal{F}_{\mathbb{L}'_{q^r}, B'_{q,r}}\right) = \{\mathrm{ev}(f) : f \in \mathcal{F}'_{q,r}\} \subseteq \mathbb{F}_{q^r}^n.$$

We say that $M_{a,b}(x,y)$ is good for $\mathcal{F}'_{q,r}$ if for all lines $L_{\alpha,\beta} \in \mathbb{L}_{q,r}$, $\deg_{\alpha,\beta}(M_{a,b} \circ L_{\alpha,\beta}) \leq B'_{q,r}$.

Notice that $M_{a,b}(x,y)$ is good for $\mathcal{F}'_{q,r}$ for each $(a,b) \in \mathbb{N}^2$ with $a+b \leq B'_{q,r}$. Consequently, we have the following result.

Proposition 5.4. *The refined norm-trace-lifted code $C'_{q,r} = \mathrm{ev}\left(\mathcal{F}'_{q,r}\right)$ is a code over $\mathbb{F}_{q^r}$ of length q^{2r-1}, locality $q^{r-1} - (\gcd(r,q-1) - 1 + (r-1)(q-2))\, q^{\frac{r-2}{2}} - 2$, and availability $q^r - q^{r-1} - 2$ with rate bounded away from 0 as the code length grows for a fixed characteristic.*

Proof. The proof is similar to that of Theorem 5.3. □

In the next subsection, we draw some comparisons between these codes and other families.

5.3 Comparisons

Recall that the norm-trace-lifted codes defined over fields of arbitrary characteristic arise from the set of functions $\mathcal{F}_{q,r}$ which depend on the degree bounds given in Theorem 4.6. These bounds hold for any q. Tighter bounds may give rise to codes with better parameters.

We begin this subsection with a demonstration of this.

Example 5.5. Consider a curve-lifted code on the norm-trace curve $\mathcal{X}_{2,r}$ over $\mathbb{F}_{2^r}$. First consider $C(D, (\mathcal{F}_{\mathbb{L}_{2^r}, 2^{r-1}-(r-1)2^{(r-2)/2}}))$. Recall that Theorem 4.6 gives

$$n_{2,r} \geq 2^{r-1} - (r-1)2^{(r-2)/2}$$

whereas

$$n_{2,r}(\alpha,\beta) = 2^{r-1} \pm 1,$$

as shown in Proposition 4.2. Hence, a code with larger dimension is obtained by considering the binary norm-trace-lifted code $C(D, (\mathcal{F}_{\mathbb{L}_{2^r}, B_{2,r}}))$, which was also studied in [11]. We see that $[2^{2r-1}, (0.25 - \varepsilon_r) \cdot 2^{2r-1}, \geq 2^r]$ code with locality $2^{r-1} - 2$, availability $2^r - 1$, and asymptotic rate 0.25 [11, Theorem 3].

To increase the dimension further, we may take $\mathbb{L}''_{2^r} := \{L_{\alpha,\beta} : Tr(\beta) = 0\}$. In doing so, according to the proof of Proposition 4.2, we obtain a code $C(D, \mathcal{F}_{\mathbb{L}''_{2^r}, 2^{r-1}+1})$. We will see that

$$\left\langle \mathrm{ev}_B(x^a y^b) : a + b = 2^{r-1} - 1, 2^{r-1} - 2 \right\rangle \in C(D, \mathcal{F}_{\mathbb{L}''_{2^r}, 2^{r-1}+1}) \setminus C(D, \left(\mathcal{F}_{\mathbb{L}_{2^r}, B_{2,r}}\right)).$$

Further comparisons are captured in Table 1.

	1-pt norm-trace	HLC	NTLC	RNTLC
locality	$2^{r-1} - 2$	$2^{r/2}$	$2^{r-1} - 2$	2^{r-1}
availability	$2^r - 1$	$2^r - 1$	$2^r - 1$	$2^{r-1} - 1$
length	2^{2r-1}	$2^{3r/2}$	2^{2r-1}	2^{2r-1}
dimension	$\leq 2^{2r-4} - 1$	$\geq 0.007 \cdot 2^{3r/2}$	$(0.25 - \varepsilon_r) \cdot 2^{2r-1}$	$(0.25 - \varepsilon_r) \cdot 2^{2r-1}$
asymptotic rate	$\leq \frac{1}{8} - \frac{1}{2^{2r-1}}$	≥ 0.007	0.25	0.25

Table 1: Parameters of one-point norm-trace, Hermitian-lifted, binary norm-trace-lifted, and refined binary norm-trace-lifted codes over $\mathbb{F}_{2^r}$; see also [15]

Remark 5.6. We recognize that while the binary norm-trace-lifted codes $C(D, \left(\mathcal{F}_{\mathbb{L}_{2^r}, B_{2,r}}\right))$ have dimension exceeding their counterparts $C(D, \left(\mathcal{F}_{\mathbb{L}_{2^r}, 2^{r-1}-(r-1)2^{(r-2)/2}}\right))$ defined from the same curve, their asymptotic rates behave similarly. Moreover, there is a tradeoff in locality which is larger for the codes $\mathrm{ev}\left(\mathcal{F}_{b,2,r}\right)$ than $\mathrm{ev}\left(\mathcal{F}_{2,r}\right)$, meaning more symbols are needed in order to perform recovery.

It is also worth comparing the codes introduced in this paper with their Hermitian counterparts. A crucial distinction arises if we wish to consider the codes concretely in terms of bases or generator matrices. According to the next result, these are readily available for norm-trace-lifted codes with $r > 2$.

Corollary 5.7. *For $r > 2$, the norm-trace-lifted code $C(D, \mathcal{F}_{q,r})$ over $\mathbb{F}_{q^r}$ is defined exclusively by typical monomials, meaning*

$$\left\{ \mathrm{ev}\left(x^a y^b\right) : a + b \leq B_{q,r} - 2 \right\}$$

is a basis for $C(D, \mathcal{F}_{q,r})$.

Proof. The result follows immediately from Proposition 3.11, since $q^{r-1} \geq B_{q,r}$. □

Notice that Proposition 3.11 does not apply in the Hermitian case, because $r = 2$. In particular, $q \leq B = q + 1$.

Remark 5.8. Consider the code $C(D, \mathcal{F}_{\mathbb{L}_{q^r}, n_{q,r}})$ on the norm-trace curve $\mathcal{X}_{q,r}$ over $\mathbb{F}_{q,r}$, where

$$n_{q,r} := \min\left\{ n_{q,r}(\alpha, \beta) : \alpha, \beta \in \mathbb{F}_{q^r} \right\}.$$

While (to our knowledge) at present there is not a closed form expression for $n_{q,r}$, values can be computed as in Table 2. Using the precise intersection numbers for odd q produces higher dimensional codes.

We now provide examples to illustrate this fact.

Example 5.9. Consider the curve $\mathcal{X}_{7,3}$ over $\mathbb{F}_{343}$. According to Table 2, $n_{q,r} = 43$. The curve-lifted code $C(D, \mathcal{F}_{\mathbb{L}_{343},43})$ includes codewords given by evaluating $x^a y^b$, $a + b \leq 41$, whereas the norm-trace-lifted code given by the bound $n_{7,3} \geq 16$ only considers those with $x^a y^b$, $a + b \leq 14$, according to Corollary 5.7.

p	r	$\#(L_{\alpha,\beta} \cap \mathcal{X}_{p,r})(\mathbb{F}_{p^r})$	$B_{p,r}$	$B'_{p,r}$
3	2	1, 4	0	0
3	3	13, 7, 10	1	4
3	4	22, 28, 31	8	14
3	5	73, 76, 85, 91	38	59
3	6	229, 244, 256	152	188
3	7	703, 715, 742, 757	540	634
5	2	1, 6	0	0
5	3	21, 26,31	6	10
5	4	111, 121, 126 141	64	64
5	5	561, 611, 621, 626, 641, 681	445	489
5	6	3056, 3106, 3126, 3131, 3206	2624	2724
5	7	15631, 15751, 15731, 15501, 15681, 15456	14282	14617
7	2	1, 8	0	0
7	3	43, 50, 57	16	16
7	4	351, 358, 316, 379, 337	216	230
7	5	2451, 2325, 2465, 2381, 2437, 2395, 2353, 2402	1955	2029
7	6	16773, 16738, 17053, 16843, 16801, 16633, 16808	15336	15336
7	7	117433, 117615, 118693, 117580, 118063, 116173, 117895, 117853, 117685, 117643, 117475	112980	113758

Table 2: Collection of actual intersection numbers $n_{q,r}$ for all α, β and bounds

6 Conclusion

In this paper, we defined curve-lifted codes which allow for local recovery by taking as repair groups the points of intersection of the curve with lines through the evaluation points. While inspired by Hermitian-lifted codes, they may exhibit different behaviors depending on the particular defining curve. We demonstrate when such codes have an explicit basis arising from typical monomial, unlike the Hermitian case where sporadic monomials are needed. In addition, we determined bounds on the number of affine points of intersection between a norm-trace curve and a line. We then used them to define norm-trace-lifted codes over fields of arbitrary characteristic. These new codes have high availability and positive rate, bounded away from zero as the code length goes to infinity. We note that codes over fields of small characteristic provide the best asymptotic rates. The opportunity to obtain greater rate by evaluating more functions may motivate one to consider more tailored bounds for the intersection numbers.

References

[1] Austin Allen, Eric Pabón-Cancel, Fernando Piñero-González, and Lesley Polanco. Improving the dimension bound of Hermitian lifted codes, 2023.

[2] Alexander Barg, Itzhak Tamo, and Serge Vlăduţ. Locally recoverable codes on algebraic curves. *IEEE Transactions on Information Theory*, 63(8):4928–4939, 2017.

[3] Wieb Bosma, John Cannon, and Catherine Playoust. The Magma algebra system. I. The user language. *J. Symbolic Comput.*, 24(3-4):235–265, 1997. Computational algebra and number theory (London, 1993).

[4] A. Garcia and P. Viana. Weierstrass points on certain non-classical curves. *Arch. Math*, 46:315–322, 1986.

[5] Olav Geil. On codes from norm–trace curves. *Finite Fields and Their Applications*, 9(3):351–371, 2003.

[6] Parikshit Gopalan, Cheng Huang, Huseyin Simitci, and Sergey Yekhanin. On the locality of codeword symbols. *IEEE Transactions on Information Theory*, 58(11):6925–6934, 2012.

[7] V. D. Goppa. Algebraico-Geometric Codes. *Izvestiya: Mathematics*, 21(1):75–91, February 1983.

[8] Kathryn Haymaker, Beth Malmskog, and Gretchen L. Matthews. Locally recoverable codes with availability $t \geq 2$ from fiber products of curves. *Advances in Mathematics of Communications*, 12(2):317–336, 2018.

[9] Nicholas M. Katz. Estimates for Soto-Andrade sums. *J. Reine Angew. Math.*, 438:143–161, 1993.

[10] Hiram H. López, Beth Malmskog, Gretchen L. Matthews, Fernando Piñero González, and Mary Wootters. Hermitian-lifted codes. *Designs, Codes and Cryptography*, 89:497–515, 2021.

[11] Gretchen L. Matthews and Aidan W. Murphy. Norm-trace-lifted codes over binary fields. In *2022 IEEE International Symposium on Information Theory (ISIT)*, pages 3079–3084, 2022.

[12] Marko Moisio. Kloosterman sums, elliptic curves, and irreducible polynomials with prescribed trace and norm. *Acta Arith.*, 132(4):329–350, 2008.

[13] Marko Moisio and Daqing Wan. On Katz's bound for the number of elements with given trace and norm. *J. Reine Angew. Math.*, 638:69–74, 2010.

[14] Carlos Munuera, Alonso Sepúlveda, and Fernando Torres. Algebraic geometry codes from castle curves. In Ángela Barbero, editor, *Coding Theory and Applications*, pages 117–127, Berlin, Heidelberg, 2008. Springer Berlin Heidelberg.

[15] Aidan W. Murphy. *Codes from norm-trace curves: local recovery and fractional decoding.* PhD thesis, Virginia Tech, 2022.

[16] Dimitris S. Papailiopoulos and Alexandros G. Dimakis. Locally repairable codes. In *2012 IEEE International Symposium on Information Theory Proceedings*, pages 2771–2775, 2012.

[17] F. K. Schmidt. Zur arithmetischen theorie der algebraischen funktionen. ii. allgemeine theorie der weierstraßpunkte. *Mathematische Zeitschrift*, 45:75–96, 1939.

[18] M. A. Tsfasman, S. G. Vlădutx, and Th. Zink. Modular curves, Shimura curves, and Goppa codes, better than the Varshamov-Gilbert bound. *Mathematische Nachrichten*, 109(1):21–28, 1982.