Locally Recoverable Codes on Surfaces

Cecília Salgado, Anthony Várilly-Alvarado, and José Felipe Voloch

***Abstract*— A linear error correcting code is a subspace of a finite-dimensional space over a finite field with a fixed coordinate system. Such a code is said to be locally recoverable with locality r if, for every coordinate, its value at a codeword can be deduced from the value of (certain) r other coordinates of the codeword. These codes have found many recent applications, e.g., to distributed cloud storage. We will discuss the problem of constructing good locally recoverable codes and present some constructions using algebraic surfaces that improve previous constructions and sometimes provide codes that are optimal in a precise sense. The main conceptual contribution of this paper is to consider surfaces fibered over a curve in such a way that each recovery set is constructed from points in a single fiber. This allows us to use the geometry of the fiber to guarantee the local recoverability and use the global geometry of the surface to get a hold on the standard parameters of our codes. We look in detail at situations where the fibers are rational or elliptic curves and provide many examples applying our methods.**

***Index Terms*— Error correcting codes, locally recoverable codes, algebraic surfaces.**

I. Introduction

MOTIVATED by applications to distributed cloud storage, Gopalan *et al.* [6] introduced a particular class of error correcting codes that efficiently correct erasures, known now as *locally recoverable codes*. The successful application of algebraic geometry to the construction of error-correcting codes [17] naturally prompted the search for locally recoverable codes using algebro-geometric methods [2], [3], [5], [9]–[12]. In particular, [3] gave a systematic way to produce optimal locally recoverable codes.

Manuscript received December 11, 2019; revised February 16, 2021; accepted June 10, 2021. Date of publication June 21, 2021; date of current version August 25, 2021. This work was supported in part by IMPA, in part by BIRS-Oaxaca, in part by MPIM, in part by IHP, in part by the University of Canterbury, and in part by the M. Stoll's Rational Points Workshop Series. The work of Cecília Salgado was supported in part by FAPERJ under Grant E-26/203.205/2016, in part by the Serrapilheira Institute under Grant Serra-1709-17759, in part by CNPq under Grant PQ2 310070/2017-1, and in part by the Capes-Humboldt Program. The work of Anthony Várilly-Alvarado was supported in part by the National Science Foundation (NSF) under Grant DMS-1352291 and Grant DMS-1902274. The work of José Felipe Voloch was supported in part by the Simons Foundation under Grant 234591 and in part by the Marsden Fund Council administered by the Royal Society of New Zealand. *(Corresponding author: José Felipe Voloch.)*

Cecília Salgado is with the Instituto de Matemática, Universidade Federal do Rio de Janeiro, Rio de Janeiro 21941-909, Brazil, and also with the Bernoulli Institute, University of Groningen, 9747 Groningen, The Netherlands (e-mail: c.salgado@rug.nl).

Anthony Várilly-Alvarado is with the Department of Mathematics, Rice University, Houston, TX 77005 USA (e-mail: av15@rice.edu).

José Felipe Voloch is with the School of Mathematics and Statistics, University of Canterbury, Christchurch 8140, New Zealand (e-mail: felipe.voloch@canterbury.ac.nz).

Communicated by G. Matthews, Associate Editor for Coding Theory.

Digital Object Identifier 10.1109/TIT.2021.3090939

Algebro-geometric codes are constructed from algebraic varieties, but the one-dimensional case of curves is the most amply studied, while surfaces and higher dimensional varieties have received less attention. The purpose of this article is to present new systematic constructions of locally recoverable codes using surfaces fibered over a curve in such a way that each recovery set is constructed from points in a single fiber. We use the geometry of the fiber to guarantee local recoverability the global geometry of the surface to get a hold on the standard parameters of our codes. We start by setting up a general framework for such constructions and showing how some previous constructions of locally recoverable codes fall into this framework. We then specialize our setup to consider codes constructed using ruled surfaces and elliptic surfaces. Some of the examples we produce are optimal (in the sense of achieving equality in inequality I.1) and are long in the sense that they have, for example, length $n = 4q$, where q is the size of the alphabet. These codes are longer than the other known explicit codes with same recoverability and dimension; however, they have bounded recoverability. We obtain the following theorem as a corollary of Theorem IV.5 (see Example IV.6).

Theorem I.1: For an integer d divisible by 4 and an integer $b \leq q$ such that $4b \geq d$, there exist optimal locally recoverable codes over $\mathbb{F}_q$ with parameters

$$(n, k, d, r) = \left(4b, 3b - \frac{3}{4}d + 1, d, 3\right).$$

For arbitrarily large recoverability, we construct, for every prime p, codes over $\mathbb{F}_{p^2}$ of recoverability p, length n about $2p^2$, distance d for any $d \leq n, (p+1) \mid d$, having dimension just shy of the optimal $p(n-d)/(p+1)$. The precise statement is as follows (Theorem VI.6)

Theorem I.2: For every odd prime (power) p and integer $d \leq 2(p+1)(p-2), (p+1) \mid d$, there exists a locally recoverable code $\mathcal{C}$ over $\mathbb{F}_{p^2}$ of recoverability p, length $n = 2(p+1)(p-2)$, minimum distance d and dimension

$$k = \frac{p(n-d)}{p+1} - \frac{p-1}{2}.$$

We believe that codes in this range are new.

Ultimately, the codes we construct are obtained by evaluating functions on a curve lying on a surface, and thus can be viewed as codes on curves. However, our proofs of the various properties these codes enjoy crucially rely on the internal geometry of the ambient surface. This point of view guided our work throughout, so we have kept the perspective it affords.

The work of [11] also uses curves embedded in higher dimensional varieties to construct locally recoverable codes. Their construction has some similarities and some differences to ours. We compare the two constructions, once we set up some terminology, in Section III-D

A. Locally Recoverable Codes

Let $\mathbb{F}_q$ be the finite field of q elements. A linear error correcting code is a subspace $\mathcal{C}$ of $\mathbb{F}_q^n$ for some n, which is called the length of $\mathcal{C}$. We denote by k the dimension of $\mathcal{C}$ as a $\mathbb{F}_q$-vector space and we denote by d the minimum distance of $\mathcal{C}$, defined as the minimum number of nonzero coordinates among the nonzero elements of $\mathcal{C}$.

The code $\mathcal{C}$ is said to be locally recoverable (LR) with locality r if, for each $i = 1, \ldots, n$, there is a subset $J_i \subset \{1, \ldots, n\} - \{i\}, \#J_i = r$ (called the recovery set), such that, if we know the values c_j for $j \in J_i$ of the coordinates of any $c \in \mathcal{C}$, then we can recover c_i. Codes with small locality can be used in distributed storage systems as they can reconstruct data erasures with smaller storage overhead than traditional back-ups. It is desirable to have codes with small locality, large dimension (equivalently, high information rate k/n) and large minimum distance for these applications. However, these parameters are not independent: they satisfy the basic constraint [6], [13]

$$d \leq n - k - \lceil k/r \rceil + 2, \tag{I.1}$$

and $\mathcal{C}$ is called an optimal LR code if equality holds. We write d_{opt} for the right hand side of (I.1).

An explicit construction of optimal LR codes with $n \leq q$ is given in [16]. There are known upper bounds for the length of LR codes and some general existence theorems [7]. One of the purposes of this paper is to explicitly construct longer optimal LR codes.

The LR codes we construct have the property that the sets $J_i \cup \{i\}$ form a partition of $\{1, \ldots, n\}$ but not every LR code has this property. We end this subsection by giving a simple proof of (I.1) for LR codes with this property.

Theorem I.3: Consider an [n,k,d]-LR code of locality r whose recovery sets J_i have the property that the union of the sets $J_i \cup \{i\}$ form a partition of $\{1, \ldots, n\}$. Then (I.1) holds.

Proof: Note that the recovery map for any coordinate on inputs all equal to 0 is 0, since the zero vector is a codeword. Now take $b = \lceil k/r \rceil - 1$ so $br < k$ and choose b disjoint sets of the form $J_i \cup \{i\}$ and set the r coordinates indexed by each J_i from this choice to 0. In addition, choose $k - 1 - br$ coordinates outside the union of the chosen $J_i \cup \{i\}$ and set them equal to 0 as well. Thus, a total of $k - 1$ conditions are imposed and there exists a non-zero codeword satisfying them all as our code has dimension k. But this non-zero codeword also has zero i-th coordinates for all of the chosen $J_i \cup \{i\}$. This gives us b additional zero coordinates. Hence the weight of this codeword is at most $n - (k-1) - b = n - k - \lceil k/r \rceil + 2$. □

B. Algebro-Geometric Codes

Let X be a quasi-projective [1] algebraic variety over a finite field $\mathbb{F}_q$. Concretely, this means that we select an open subset of affine or projective space where a collection of polynomials vanish.

The function field of X is the set of functions that can be expressed as quotients of polynomials in the coordinates of the ambient space modulo the equations defining X. Given a point P on X and an element σ of the function field of X, if the denominator of σ does not vanish at P, the function σ can be evaluated at P giving an element $\sigma(P)$ of $\mathbb{F}_q$.

Let $P_1, \ldots, P_n$ be a subset of the set $X(\mathbb{F}_q)$ of $\mathbb{F}_q$-rational points of X and V a finite-dimensional subspace of the function field of X. We assume that the evaluation, as above, of all elements of V at all the points $P_1, \ldots, P_n$ is defined and we can consider the image $\mathcal{C}$ of evaluation map, which is an error correcting code:

$$\begin{aligned} \text{ev}_V : V &\to (\mathbb{F}_q)^n \\ \sigma &\mapsto (\sigma(P_1), \ldots, \sigma(P_n)) . \end{aligned}$$

The length of the code is n. The dimension k of the code is

$$k = \dim_{\mathbb{F}_q}(\text{im}\, \text{ev}_V) = \dim V - \dim_{\mathbb{F}_q}(\ker \text{ev}_V)$$

which simplifies to $\dim V$ if ev_V is injective. The minimum distance d is the smallest Hamming distance between elements of $\mathcal{C}$. This is equal to n minus the largest number of $\mathbb{F}_q$-points of X vanishing on an element of $V \setminus \ker \text{ev}_V$.

For X a projective variety and D a divisor on X, we denote by $\mathcal{L}(X, D)$ the Riemann-Roch space of functions σ on X such that either $\sigma = 0$ or $(\sigma) + D$ is an effective divisor, where (σ) denotes the divisor of σ. The space $\mathcal{L}(X, D)$ is always finite-dimensional and we denote its dimension by $\ell(X, D)$. We will typically define our vector space V as above as a subspace of some $\mathcal{L}(X, D)$.

II. Baseline Codes From High-Dimensional Varieties

Let $\mathbb{A}^m$ denote affine m-dimensional space over a finite field $\mathbb{F}_q$. In this section we construct locally recoverable codes, with local recoverability parameter r from a projection morphism

$$\begin{aligned} \pi &\colon \mathbb{A}^{r-1} \times \mathbb{A}^1 \to \mathbb{A}^1, \\ (x_1, \ldots, x_{r-1}; t) &\mapsto t. \end{aligned}$$

We shall impose the smallest possible amount of structure on our choice of points for evaluation. This will give us a baseline to assess the parameters of other constructions.

[1] Many codes are naturally described as algebro-geometric (AG) codes in quasi-projective varieties that are not projective. Witness the classical Reed-Muller codes; they are AG codes in affine space $\mathbb{A}^n$. Every quasi-projective variety is an open subset of a projective variety, by definition, but the choice of a projective compactification is not unique, e.g., $\mathbb{A}^n$ can be embedded in projective space $\mathbb{P}^n$ or in the product $(\mathbb{P}^1)^n$, which are different. In this paper, we use specific choices of compatifications when determining parameters for our codes (e.g., when we consider Hirzebruch surfaces). In other circumstances, it is preferable not to, e.g., a curve can always be embedded in a unique projective curve without increasing the number of singular points. This is not true in higher dimensions.

Let M and N denote positive integers. We shall use the space of functions

$$V[M,N] := \{a_0(t) + \sum_{i=1}^{r-1} a_i(t)x_i : \deg a_0 \le M \text{ and } \deg a_i \le N \text{ for } i = 1,\dots,r-1\}$$

to construct an evaluation code (so $V[M,N]$ plays the rôle of the vector space V from § I-B). We pick, for some $b \le q$, some set of b distinct points on the target $\mathbb{A}^1$ of the morphism π and, in each of b fibers of π above these points, we pick $r+1$ points and take all these $b(r+1)$ points as the set of points where we evaluate the above functions. Thus, the length of the resulting code will be $n = b(r+1)$. The following lemma falls within the framework of [5, Proposition 4.2].

Lemma II.1: Fix $t = t_0 \in \mathbb{F}_q$, and let $P_1,\dots,P_{r+1}$ be $\mathbb{F}_q$-points in the fiber $\pi^{-1}(t_0)$, no r of which lie on a hyperplane. Let $\sigma \in V[M,N]$ be a function. Then the value of $\sigma(P_i)$ can be recovered from knowledge of the coordinates of $P_1,\dots,P_{r+1}$ and the r values $\sigma(P_1),\dots,\widehat{\sigma(P_i)},\dots,\sigma(P_{r+1})$.

Proof: Write $\sigma = a_0(t) + \sum_{i=1}^{r-1} a_i(t)x_i$. Let $a_i = a_i(t_0)$ for $i = 1,\dots,r+1$. Then we have the matrix equation

$$\begin{pmatrix} 1 & x_1(P_1) & \cdots x_{r-1}(P_1) \\ \vdots & & \\ \hat{1} & \widehat{x_1(P_i)} & \cdots \widehat{x_{r-1}(P_i)} \\ \vdots & & \\ 1 & x_1(P_{r+1}) & \cdots x_{r-1}(P_{r+1}) \end{pmatrix} \cdot \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_r \end{pmatrix} = \begin{pmatrix} \sigma(P_1) \\ \vdots \\ \widehat{\sigma(P_i)} \\ \vdots \\ \sigma(P_{r+1}) \end{pmatrix}. \tag{II.1}$$

Since no r of the points $P_1,\dots,P_{r+1}$ lie on a hyperplane, the $r \times r$ matrix in (II.1) is invertible, and hence we may compute $a_0,\dots,a_r$ from knowledge of the coordinates of $P_1,\dots,\widehat{P_i},\dots,P_{r+1}$ and the r values $\sigma(P_1),\dots,\widehat{\sigma(P_i)},\dots,\sigma(P_{r+1})$. We conclude that

$$\sigma(P_i) = a_0 + a_1x_1(P_i) + \cdots a_rx_{r-1}(P_i). \qquad \square$$

To construct what we will call a baseline code, let (as above)

$$\{t_1,\dots,t_b\} \subseteq \mathbb{A}^1(\mathbb{F}_q)$$

be b distinct points on the target $\mathbb{A}^1$ of the morphism π, and for each t_i, choose $r+1$ points $P_{i,1},\dots,P_{i,r+1}$ on the fiber $\pi^{-1}(t_i)$, no r of which lie on a hyperplane.

Proposition II.2: Suppose that $b - M, b-N \ge 1$. The baseline code

$$\mathcal{C} = \{(\sigma(P_{i,j}))_{1\le i\le b, 1\le j\le r+1} : \sigma \in V[M,N]\}.$$

has local recoverability r and its parameters satisfy

$$\begin{aligned} &n = b(r+1), \\ &k = (M+1) + (r-1)(N+1), \\ &d \le (r+1)\,(b-(N+1)) - (M-N) - \left\lceil \frac{M-N}{r} \right\rceil + 2, \\ &d \ge \min\{(b-M)(r+1), 2(b-N)\}. \end{aligned}$$

Proof: We have already discussed the length of $\mathcal{C}$. The dimension of the code is simply the $\mathbb{F}_q$-dimension of $V[M,N]$. The upper bound on the distance of the code is an application of (I.1). For the lower bound on d, we argue as follows: Suppose that $\sigma \in V[M,N]$ is a function with $a_i \equiv 0$ for $i = 1,\dots r-1$, i.e., $\sigma = a_0(t)$ for a polynomial $a_0(t)$ of degree $\le M$. Then at least $(b-M)$ of the values $a_0(t_1),\dots,a_0(t_b)$ are nonzero. The weight of the codeword associated to σ is thus at least $(b-M)(r+1)$. If, on the other hand, $\sigma \in S$ is a function where at least one $a_i \not\equiv 0$ for $i = 1,\dots,r-1$, then at least $(b-N)$ of the values $a_i(t_1),\dots,a_i(t_b)$ are nonzero. In the corresponding fibers of π, the function σ defines a hyperplane. The hypothesis that no r points on a fiber lie on a hyperplane ensures that σ takes on a nonzero value on at least two points in each of the $(b-N)$ fibers. Hence, $d \ge \min\{(b-M)(r+1), 2(b-N)\}$, as claimed.

Local recoverability of $\mathcal{C}$ follows from Lemma II.1. $\square$

Remark II.3: The proof of Proposition II.2 shows that if $\min\{(b-M)(r+1), 2(b-N)\} = (b-M)(r+1)$, then in fact $d = (b-M)(r+1)$. In addition, if

$$M + N > b \quad \text{and} \quad 2N > b \tag{II.2}$$

then it is always possible to construct a function σ whose associated code word has weight exactly $2(b-N)$. So under the conditions (II.2), the lower bound for d in Proposition II.2 is in fact sharp.

Example II.4: We specialize to the case where $r = 3$, $M = b-1$ and $N = b-2$. Then the upper and lower bounds for d meet and we have $d = 4$. This gives optimally recoverable codes with parameters

$$(n,k,d,r) = (4b, 3b-2, 4, 3).$$

Note that the information rate k/n is approximately 75%, and since $b \le q$, one can construct codes with $n = 4q$ and high information rate that are optimal locally recoverable. In particular, over any $\mathbb{F}_q$ with $q \ge 9$, we can construct a code with parameters $(n,k,d,r) = (32, 22, 4, 3)$.

Example II.5: If we now take $b \le q, r$ arbitrary and $M = N = b-1$, then the upper and lower bounds of Proposition II.2 also coincide and the code has distance $d = 2$.

The last two examples are the only cases where the upper and lower bounds of Proposition II.2 coincide and a baseline code with no additional properties is optimal (see Remark II.3).

To see this, let $\delta := M - N$; we consider two cases:

- $\min\{(b-M)(r+1), 2(b-N)\} = (b-M)(r+1)$: Then

$$(b-M)(r+1) = (r+1)\,(b-(N+1)) - (M-N) - \left\lceil \frac{M-N}{r} \right\rceil + 2,$$

from which one can conclude that

$$(r+1)(\delta-1) - \delta - \left\lceil \frac{\delta}{r} \right\rceil + 2 = 0. \tag{II.3}$$

Write $\lceil \frac{\delta}{r} \rceil = \frac{\delta}{r} + \epsilon, 0 \le \epsilon < 1$, then

$$\delta(r - 1/r) = r - 1 + \epsilon. \tag{II.4}$$

This implies in particular that $\delta > 0$. If $\delta \geq 2$ then, since $r \geq 3$, we have

$$\delta(r - 1/r) \geq 2r - 1 > r - 1 + \epsilon$$

and hence $\delta = 1$, since it is an integer.
If $\delta = 1$ then the hypothesis $2(b-N) \geq (b-M)(r+1)$ gives

$$b \leq M + \frac{2}{r-1} \leq M + 1 \text{ (whenever } r \geq 3),$$

from which we conclude that $b-M = 1$, and hence that $b-N = 2$. It follows that

$$(r+1) = (b-M)(r+1) = d = \min\{(b-M)(r+1), 2(b-N)\} = \min\{(r+1), 4\},$$

whence $r + 1 \leq 4$. Since we want codes with $r \geq 3$, we must have $r = 3$ and $d = 4$.

- $\min\{(b-M)(r+1), 2(b-N)\} = 2(b-N)$:
Then

$$b = N + 1 + \frac{\delta}{r-1} + \frac{1}{r-1}\left\lceil \frac{\delta}{r} \right\rceil. \tag{II.5}$$

On the other hand, $\min\{(b-M)(r+1), 2(b-N)\} = 2(b-N)$ gives

$$b \geq \frac{M(r+1) - 2N}{r-1}. \tag{II.6}$$

Substituting the value for b obtained in (II.5) into the inequality (II.6) we get

$$\delta\left(1 + \frac{1}{r-1}\right) \leq 1 + \frac{1}{r-1}\left\lceil \frac{\delta}{r} \right\rceil.$$

The latter implies that $\delta \leq 1$. If $\delta = 1$ then $M = N + 1$ and thus

$$2(b-N) \leq (b-(N+1))(r+1) \implies b \geq N + 1 + \frac{2}{r-1}.$$

We also have

$$\begin{aligned} b &= N + 1 + \frac{1}{r-1}\left\lceil \frac{1}{r} \right\rceil + \frac{1}{r-1} \text{ (by (II.5))} \\ &= N + 1 + \frac{2}{r-1} \\ &\leq N + 2. \end{aligned}$$

The distance is thus given by $d = 2(b-N) \leq 4$ and by our analysis, the inequality $2(b-N) \leq (b-(N+1))(r+1)$ is sharp, so $(b-M)(r+1) \leq 4$, which forces $r \leq 3$. Finally, since we assumed $r \geq 3$, we conclude that in fact $r = 3$ and $d = 4$.
If $\delta = 0$ then (II.5) gives $b = N + 1$, which implies that $d = 2$. If $\delta \leq -1$, we get $b \leq N$ which is not possible.

III. Codes From Ruled Surfaces: Affine Intimations

A. Tamo-Barg Codes

We present the construction of Tamo and Barg [16] of optimal LR codes of length at most q from the perspective of the last section, which we believe is new. We retain the notation of the previous section.

Let $g(x) \in \mathbb{F}_q[x]$ be a polynomial of degree $r+1$, viewed as a morphism $g\colon \mathbb{A}^1 \to \mathbb{A}^1$. Choose distinct $t_1, \ldots, t_b \in \mathbb{F}_q$ such that the fiber $g^{-1}(t_i)$ consists of $r+1$ distinct elements $x_{i,1}, \ldots, x_{i,r+1}$ of $\mathbb{F}_q$, for $i = 1, \ldots, b$. Note that the $x_{i,j}$ are therefore $n = b(r+1)$ distinct elements of $\mathbb{F}_q$. We define the points $P_{i,j} = (x_{i,j}, x_{i,j}^2, \ldots, x_{i,j}^{r-1}) \in \mathbb{A}^{r-1}(\mathbb{F}_q)$, and we consider the projection map

$$\pi\colon \mathbb{A}^{r-1} \times \mathbb{A}^1 \to \mathbb{A}^1, \quad (x_1, \ldots, x_{r-1}; t) \mapsto t.$$

For a fixed i, the fiber above t_i is an affine space $\mathbb{A}^{r-1}$ containing the points $P_{i,j}$ for $j = 1, \ldots, r+1$. Moreover, by their construction, these points lie on an *affine rational normal curve*, i.e., they lie on the image of the map

$$h\colon \mathbb{A}^1 \to \mathbb{A}^{r-1}, \quad x \mapsto (x, x^2, \ldots, x^{r-1}).$$

This guarantees that no r of them lie on a hyperplane. As in §II, we take the space of functions $V[M, N]$, but specialize to the case where $M = N$, and build a code $\mathcal{C}$. Lemma II.1 guarantees that $\mathcal{C}$ has local recoverability r. Put differently, the fact that the points $P_{i,j}$ lie on rational normal curves implies that the $r \times r$ matrix in (II.1) is a Vandermonde matrix, thus invertible.

The parameters n, k, and r for the code $\mathcal{C}$ are as before. However, in this special situation, we get a better lower bound for the minimum distance d as follows. Note that

$$\sigma(P_{i,j}) = a_0(g(x_{i,j})) + \sum_{\ell=1}^{r-1} a_\ell(g(x_{i,j}))x_{i,j}^\ell$$

is the value at $x = x_{i,j}$ of a polynomial of degree at most $N(r+1) + r - 1$ in x. This degree is an upper bound on the number of its zeros and thus $d \geq n - (N(r+1) + r - 1)$. On the other hand, as in the previous section, the upper bound (I.1) for d when $M = N$ is

$$(r+1)(b-(N+1)) - (M-N) - \left\lceil \frac{M-N}{r} \right\rceil + 2 = n - (N(r+1) + r - 1),$$

showing that these codes are optimal LR codes.

As mentioned above, these codes have $n \leq q$. To achieve n near q one needs to choose the polynomial $g(x)$ in such a way that the preimage of many values of $t \in \mathbb{F}_q$ under g consists of $r+1$ elements of $\mathbb{F}_q$. One such choice is $g(x) = x^{r+1}$ if $(r+1) \mid (q-1)$. For other choices and a full discussion, see [16].

B. Ruled Surfaces Perspective

An algebraic surface S over a field k is called a ruled surface if it is endowed with a morphism $\pi\colon S \to B$ to a base algebraic curve B such that for all but finitely many $b \in B(\bar{k})$, the fiber $\pi^{-1}(b)$ is a smooth rational curve, where $\bar{k}$ is a fixed algebraic closure of k. There is a ruled surface operating behind the scenes in our recasting of the Tamo–Barg codes [16], which we now describe.

Using the notation of §III-A, we let[2] $S = \mathbb{A}^1_x \times \mathbb{A}^1_t$, which maps to $\mathbb{A}^{r-1} \times \mathbb{A}^1_t$ via

$$h \times \mathrm{id}\colon (x,t) \mapsto (x, x^2, \ldots, x^{r-1}; t).$$

The variety S fits into the commutative diagram

$$\begin{array}{ccc} S & \xrightarrow{h\times \mathrm{id}} & \mathbb{A}^{r-1} \times \mathbb{A}^1_t \\ {\scriptstyle \pi'}\downarrow & & \downarrow{\scriptstyle \pi} \\ \mathbb{A}^1_x & \xrightarrow{\quad g \quad} & \mathbb{A}^1_t \end{array}$$

where the map $\pi'\colon S \to \mathbb{A}^1_x$ is projection onto the first coordinate. The variety S is our ruled surface, and the code constructed in §III-A can be described as an evaluation code on S, as follows. Given $t_1, \ldots, t_b$ outside the branch locus of the morphism $g\colon \mathbb{A}^1_x \to \mathbb{A}^1_t$, i.e., such that the fiber $g^{-1}(t_i)$ consists of b distinct points $x_{i,1}, \ldots, x_{i,r+1}$ in $\mathbb{A}^1_x(\mathbb{F}_q)$, we set

$$P_{i,j} = (x_{i,j}, t_i) \in S(\mathbb{F}_q) \quad \text{for } 1 \le i \le b, 1 \le j \le r+1,$$

so that the recovery set for the point $P_{i,j}$ is

$$J_{i,j} := \{P_{i,k} : 1 \le k \le r+1, k \ne j\}.$$

Then, letting

$$V[N] = \left\{ a_0(t) + \sum_{i=1}^{r-1} a_i(t) x^i : \deg a_i \le N \text{ for } i = 0, \ldots, r-1 \right\}$$

the Tamo–Barg codes are of the form

$$\mathcal{C} = \{(\sigma(P_{i,j}))_{1 \le i \le b, 1 \le j \le r+1} : \sigma \in V[N]\}.$$

C. Recasting and Extending Barg–Tamo–Vlăduţ Codes

Just as §§III-A–III-B gives a reinterpretation of the construction of [16], in this section we reinterpret the construction of [3] but here we go further and, aided by our geometric point of view, obtain better codes by a judicious choice of the space of functions to evaluate. Some of the codes we obtain are optimal.

In broad terms, we consider a *curve* C in the surface $S = \mathbb{A}^1_x \times \mathbb{A}^1_t$ and embed S (and consequently C) in $\mathbb{A}^{r-1} \times \mathbb{A}^1_t$ as above by $(x,t) \mapsto (x, x^2, \ldots, x^{r-1}, t)$. We choose C so that the projection in the t coordinate has degree $r+1$ and choose the values of $t \in \mathbb{F}_q$ to be those for which their preimage consists of $r+1$ rational points. Then, just as before, we can evaluate these points on a space of polynomials similar to the ones considered above to get an LR code with locality r.

[2]Keen readers will immediately note that $S = \mathbb{A}^2_{(x,t)}$. We prefer to use the product $\mathbb{A}^1_x \times \mathbb{A}^1_t$ because, as we shall see in §IV, the correct projective compactification of S to work with is $\mathbb{P}^1 \times \mathbb{P}^1$, and not $\mathbb{P}^2$.

In §III-B all the points in S used for the Tamo–Barg evaluation code lie on the curve $g(x) = t$. In this section, we instead consider the curve

$$C: \quad x^{r+1} = t^2 + 1, \tag{III.1}$$

which is a cyclic cover of $\mathbb{A}^1_{\mathbb{F}_q}$ via the map $(x,t) \mapsto t$. In order to have many fibers of cardinality $r+1$ over $\mathbb{F}_q$ we take $q \equiv 1 \bmod r+1$. Fix a positive integer $\mathfrak{d}$. The space of functions we use to define the code consists of functions of the form

$$\sigma = a_0(t) + a_1(t)x + \cdots + a_{r-1}(t)x^{r-1}, \tag{III.2}$$

where the $a_j(t)$ vary in the vector space defined by the inequalities

$$\deg a_j \le \frac{n - \mathfrak{d}}{r+1} - \epsilon_j$$

and

$$\epsilon_j = \begin{cases} 0 & \text{if } j = 0, \\ 1 & \text{if } 0 < j \le (r+1)/2, \\ 2 & \text{otherwise.} \end{cases}$$

The local recoverability with locality r of the resulting code follows, since for fixed t, with $r+1$ distinct values for x, the matrix determining the missing value is a Vandermonde matrix. The inequalities defining the space of functions to be evaluated ensure that the minimum distance of this code is at least $\mathfrak{d}$, because x has a pole of order 2 at infinity and t has a pole of order $r+1$ at infinity.

The space of functions at which we evaluate points of the curve has dimension, for r odd,

$$k = \frac{r}{r+1}(n - \mathfrak{d}) - \sum_{i=0}^{r-1} \epsilon_j + r = \frac{r}{r+1}(n - \mathfrak{d}) + \frac{5-r}{2}.$$

Note that the upper bound d_{opt} for the distance of this code is

$$\begin{aligned} n - k - \left\lceil \frac{k}{r} \right\rceil + 2 &= n - \frac{r}{r+1}(n - \mathfrak{d}) \\ &+ \frac{r-5}{2} - \left\lceil \frac{1}{r+1}(n - \mathfrak{d}) + \frac{5-r}{2r} \right\rceil + 2 \\ &= \mathfrak{d} + \frac{r-5}{2} + 2. \end{aligned}$$

The last equality holds for $r \ge 5$ whereas, for $r = 3$, we just get $\mathfrak{d}$. So the codes constructed this way are optimal for $r = 3$; for $r > 3$, these codes are further from the optimal bound the larger r gets.

For r even, a similar calculation gives $\mathfrak{d} + r/2$ as the upper bound for the distance when $r > 2$ and $\mathfrak{d}$ when $r = 2$. So the codes constructed are optimal for $r = 2$; for $r > 3$, these codes are further from the optimal bound the larger r gets.

We note again the similarity with the Tamo-Barg codes discussed above, which uses a space of functions of the same form as (III.2) but with $\deg a_j \le k/r - 1$ and a curve of the form $g(x) = t$ for a polynomial $g(x)$ in place of C. The length of their codes is at most q, whereas the codes above can be longer if the curve C in (III.1) has more than q affine points.

D. The Construction of Munuera and Tenório

We briefly describe the general construction of [7, Section 2.2]. Here t (to keep the notation of [11]) is a positive integer and not a variable as elsewhere in this section. They consider a map $(\phi_1, \ldots, \phi_t) : \mathbb{A}^m \to \mathbb{A}^t$ and another function $\phi_{t+1} : \mathbb{A}^m \to \mathbb{A}^1$. Their evaluation points lie in $\mathbb{A}^m$ but they use the map $(\phi_1, \phi_2, \ldots, \phi_{t+1})$ to view them in $\mathbb{A}^{t+1}$ and for the purpose of comparison it is enough to consider $\mathbb{A}^{t+1} = \mathbb{A}^t \times \mathbb{A}^1$ and the natural projections $\mathbb{A}^{t+1} \to \mathbb{A}^t$ and $\mathbb{A}^{t+1} \to \mathbb{A}^1$. Their $\mathbb{A}^1$ coordinate plays the role of the $\mathbb{A}^1_x$ coordinate in the previous section. In particular, they use the properties of the rational normal curve (under the guise of Lagrange interpolation) to get local recoverability. Their $\mathbb{A}^t$ plays the role of what we denote by $\mathbb{A}^1_t$ in the previous section.

When it comes to explicit constructions they consider an algebraic curve mapping to $\mathbb{A}^{t+1}$ (so the ϕ_i are functions on the curve) and take the evaluation points from the image of the curve. Their computation of the other parameters of the codes they construct use the intrinsic geometry of the curve and not the geometry of the curve within the ambient space, which is the viewpoint we will take in Section IV. This is where our construction and theirs diverge. Moreover, their examples lead to different code parameters which are not directly comparable to ours. Particularly, they mostly deal with values of the locality r different from those that we consider. In [7, Section 3.2, 3.3] they construct codes with $r = 2$ and [7, Section 3.4] they have codes with $r = q - 1$ over $\mathbb{F}_{q^2}$. Whereas we, in Theorems IV.5 and V.7, deal with r such that $(r + 1) \mid (q - 1)$ over $\mathbb{F}_q$ and, in Theorem VI.6, with codes with $r = q$ over $\mathbb{F}_{q^2}$. The one place where these intersect is the special case of $r = 2$ in Theorem IV.5 where we deal with an elliptic curve inside our surface. These codes are then very similar to those of [7, Section 3.3] that also use an elliptic curve. The ideas of [11] have been extended in [5] to construct (r, δ)-LRC codes, which is a direction we do not pursue here.

IV. Codes on Ruled Surfaces: $\mathbb{P}^1 \times \mathbb{P}^1$

In this section we add one more layer of geometry to the codes we constructed in §III by considering codes on the ruled surface $S = \mathbb{P}^1 \times \mathbb{P}^1$, which is a projective compactification of the surface $\mathbb{A}^1_x \times \mathbb{A}^1_t$. This extra layer of geometry affords important conceptual insights: a lower bound for the minimum distance of a code can be interpreted as an intersection number of two curves in S, and good lower bounds for a minimum distance can be achieved by forcing curves to intersect with high multiplicity at the point $(\infty, \infty) \in S$.

We begin with a toy model for our code, that is far from optimal, but which helps set ideas and notation. We let $S := \mathbb{P}^1_{(x:y)} \times \mathbb{P}^1_{(t:u)}$, where $(x : y)$ and $(t : u)$ are respective homogeneous coordinates for the factors of S.

A. A Coarse Construction

Let r be a positive odd integer, let $b \leq q$ be a positive integer, and set $n = b(r + 1)$. Choose an integer $\mathfrak{d}$ divisible by $r + 1$, so that

$$N := \frac{n - \mathfrak{d}}{r + 1}$$

is an integer, as well as a positive integer α. Consider a curve of the form

$$C : g(x, y; t, u) = 0$$

in S, where g is a bi-homogeneous polynomial of the bi-degree $(r + 1, \alpha)$. In other words, every monomial of g has total degree $r + 1$ in the variables x and y, and total degree α in the variables t and u. We say that C is of type $(r + 1, \alpha)$. Our code will be an evaluation code on the $\mathbb{F}_q$-vector space of functions of the form

$$\sigma = a_0(t, u)y^{r-1} + a_1(t, u)y^{r-2}x + \cdots a_{r-1}(t, u)x^{r-1}, \tag{IV.1}$$

where the $a_i(t, u)$ are homogeneous polynomials of degree N in t and u. We write $V_{r-1,N}$ for this vector space. Each function $\sigma \in V_{r-1,N}$ defines itself a curve in X given by $\sigma = 0$. We write (σ) for this curve[3]; it is a curve of type $(r - 1, N)$.

Write $p\colon S \to \mathbb{P}^1_{(t:u)}$ for the projection onto the second factor. To construct our code, we pick b points $(t_i : u_i) \in \mathbb{P}^1_{(t:u)}(\mathbb{F}_q)$ such that the fiber $p^{-1}((t_i, u_i)) \cap C$ consists of $r + 1$ distinct points

$$(x_{i,1} : y_{i,1}), \ldots, (x_{i,r+1} : y_{i,r+1}) \in \mathbb{P}^1_{(x:y)}(\mathbb{F}_q)$$

and set

$$P_{i,j} = ((x_{i,j} : y_{i,j}), (t_i : u_i)) \in S(\mathbb{F}_q).$$

Proposition IV.1: The code

$$\mathcal{C} := \{(\sigma(P_{i,j}))_{1 \leq i \leq b, 1 \leq j \leq r+1} : \sigma \in V_{r-1,N}\}$$

has parameters satisfying

$$\begin{aligned} n &= b(r + 1) \\ k &= r(N + 1) = \frac{r}{r + 1} \cdot (n - \mathfrak{d}) + r \\ d &\leq \mathfrak{d} - r + 1 \\ d &\geq \mathfrak{d} - \alpha(r - 1) \end{aligned}$$

Proof: The parameter k is simply the dimension of the $\mathbb{F}_q$-vector space $V_{r-1,N}$. The upper bound for the distance is the bound (I.1):

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

We have used here the divisibility relation $(r + 1) \mid (n - \mathfrak{d})$. For the lower bound on the distance, we note that the largest number of zeros in a code word in $\mathcal{C}$ is bounded above by

$$\max_{\sigma \in V_{r-1,N}} \#\left(C \cap (\sigma)\right),$$

[3]The notation (σ) is the usual notation in algebraic geometry for the divisor of zeroes of a global section of a line bundle; see §I-B.

i.e., the largest number of intersection points between C and the curve $(\sigma) \subset S$ given by $\sigma = 0$, as σ varies over the vector space $V_{r-1,N}$. The intersection theory of S shows that this number is independent of σ: indeed, the intersection of divisors on S of type (a, b) and (a', b') is $ab' + a'b$ [13, V, Example 1.4.3]. Since C is a curve of type $(r+1, \alpha)$ and (σ) is a curve of type $(r-1, N)$, we have

$$\begin{aligned}\#(C \cap (\sigma)) &= N(r+1) + \alpha(r-1) \\ &= n - \mathfrak{d} + \alpha(r-1).\end{aligned}$$

Hence, the lowest weight for a code word in $\mathcal{C}$ is

$$d \geq n - \#(C \cap (\sigma)) = \mathfrak{d} - \alpha(r-1),$$

as claimed. □

Remark IV.2: The codes in the above proposition have locality r. However, we defer the discussion of locality until after we refine the code in the next section.

Remark IV.3: The upper and lower bounds for d in Proposition IV.1 meet if and only if $\alpha = 1$; this is precisely the habitat for the Tamo–Barg codes. In the notation of §III-A, the affine curve $g(x) = t$ lies in the open set $\mathbb{A}^1_x \times \mathbb{A}^1_t = \{y, u \neq 0\}$ of S; its projective closure in S is given by $y^{r+1} g(x/y) u = t y^{r+1}$, which is a curve of type $(r+1, 1)$ in the notation of this section.

Remark IV.4: Let us compare the parameters in Proposition IV.1 with those of a base-line codes in Proposition II.2. The length n, dimension k, and upper bound for d coincide since we have specialized to the case where $M = N$ in Proposition IV.1. If $r \geq 3$, then the lower bound for d in Proposition II.2 is $2(b - N)$, while the bound for the codes just studied is

$$d \geq (r+1)(b-N) - \alpha(r-1)$$

The latter bound is better as long as $b > N + \alpha$, i.e., as long as $\mathfrak{d} > \alpha(r+1)$.

B. Refining the Construction

In this section, we show that one can narrow the gap between the upper and lower bounds for d in Proposition IV.1 by

1) choosing C judiciously,
2) using a particular proper subspace $V \subset V_{r-1,N}$ for the evaluation code,
3) using only points $P_{i,j} = ((x_{i,j} : y_{i,j}), (t_i : u_i))$ with $y_{i,j} = u_i = 1$.

Intuitively, our construction guarantees that the point

$$(\infty, \infty) := ((1:0), (1:0)) \in S(\mathbb{F}_q)$$

lies in the intersection $C \cap (\sigma)$ for all $\sigma \in V$ with high multiplicity. Note that $P_{i,j} \neq (\infty, \infty)$ for i and j by construction of $P_{i,j}$. This allows us certify the code $\mathcal{C}$ has minimum distance $d = \mathfrak{d}$.

Consider the curve

$$C : u^\alpha x^{r+1} - (t^\alpha + u^\alpha) y^{r+1} = 0,$$

which is a particular curve of type $(r+1, \alpha)$ in S. We shall use functions of the form (IV.1), but we constrain the degree in t of the polynomials $a_i(t, u)$, as follows:

$$\deg_t a_i(t, 1) \leq N - \left\lceil \frac{\alpha i}{r+1} \right\rceil.$$

This requires $N \geq \lceil \alpha(r-1)/(r+1) \rceil$, which we now assume. In other words, setting

$$\epsilon_i := \left\lceil \frac{\alpha i}{r+1} \right\rceil,$$

we assume that for each $0 \leq i \leq r-1$,

$$a_i(t, u) = u^{\epsilon_i} \cdot a'_i(t, u)$$

for a homogeneous polynomial $a'_i(t, u)$. When this is the case, the vector space of functions

$$\begin{aligned}V := \{\sigma \in V_{r-1,N} : \\ \sigma = a_0(t,u) y^{r-1} + u^{\epsilon_1} \cdot a_1(t,u) y^{r-2} x + \cdots + u^{\epsilon_{r-1}} a_{r-1}(t,u) x^{r-1}\}\end{aligned}$$

has dimension

$$k = r(N+1) - \sum_{i=0}^{r-1} \epsilon_i. \tag{IV.2}$$

The vector space V has the important property that $\sigma((\infty, \infty)) = 0$ for all $\sigma \in V$. This is key in improving our bounds for the minimum distance of the codes we define using the curve C and the space of functions V. We pick b points $(t_i : 1) \in \mathbb{P}^1_{(t:u)}(\mathbb{F}_q)$ such that the fiber $p^{-1}((t_i : 1)) \cap C$ consists of $r+1$ distinct points

$$(x_{i,1} : 1), \ldots, (x_{i,r+1} : 1) \in \mathbb{P}^1_{(x:y)}(\mathbb{F}_q).$$

Put

$$P_{i,j} = ((x_{i,j} : 1), (t_i : 1)) \in S(\mathbb{F}_q).$$

Theorem IV.5: Assume that $\alpha \mid (r+1)$ and $(r+1) \mid (q-1)$. The code

$$\mathcal{C} := \{(\sigma(P_{i,j}))_{1 \leq i \leq b, 1 \leq j \leq r+1} : \sigma \in V\}$$

has locality r and its parameters satisfy

$$\begin{aligned}&n = b(r+1), \\ &k = \begin{cases} r(N+1) - \frac{r(r-1)}{2}, \text{if } r+1 = \alpha, \text{ and} \\ r(N+1) + 2\alpha - \frac{(\alpha+1)(r+1)}{2}, \text{if } r+1 > \alpha, \end{cases} \\ &d \leq \mathfrak{d} + \frac{(\alpha-1)(r-3)}{2} - \left\lceil \frac{2\alpha}{r} - \frac{(\alpha+1)(r+1)}{2r} \right\rceil, \\ &d \geq \mathfrak{d}.\end{aligned}$$

In particular, the code $\mathcal{C}$ is an optimal LR code if $\alpha = 1$ or $r = 3$.

Example IV.6: Setting $\alpha = 2$, $r = 3$, and picking an integer d divisible by 4 such that $4b \geq d$, we obtain optimal LR codes with parameters

$$(n, k, d, r) = \left(4b, 3b - \frac{3}{4}d + 1, d, 3\right).$$

Since $b \leq q$, one can construct codes with $n = 4q$ with high information rate that are locally recoverable. Compare

this with the baseline codes from Example II.4, where a code with similar parameters is possible only when $d = 4$.

Proof: [Proof of Theorem IV.5] Assume that $r + 1 > \alpha$. By (IV.2), to establish the claim on $k = \dim_{\mathbb{F}_q} V$, it suffices to show that

$$\sum_{i=0}^{r-1} \epsilon_i = \frac{(\alpha+1)(r+1)}{2} - 2\alpha.$$

The sequence of integers $\epsilon_0, \ldots, \epsilon_{r-1}$ has the form

$$0, \underbrace{1, \ldots, 1}_{(r+1)/\alpha}, \underbrace{2, \ldots, 2}_{(r+1)/\alpha}, \underbrace{3, \ldots, 3}_{(r+1)/\alpha}, \ldots, \underbrace{\alpha-1, \ldots, \alpha-1}_{(r+1)/\alpha}, \underbrace{\alpha, \ldots, \alpha}_{(r+1)/\alpha-2}.$$

Hence

$$\begin{aligned}\sum_{i=0}^{r-1} \epsilon_i &= \sum_{l=1}^{\alpha-1} l \cdot \frac{r+1}{\alpha} + \alpha\left(\frac{r+1}{\alpha} - 2\right) \\ &= \frac{(\alpha-1)\alpha}{2} \cdot \frac{r+1}{\alpha} + (r+1) - 2\alpha \\ &= (\alpha-1)\frac{r+1}{2} + (r+1) - 2\alpha \\ &= \frac{(\alpha+1)(r+1)}{2} - 2\alpha.\end{aligned}$$

If $r + 1 = \alpha$, then $\epsilon_i = i$ and the result follows.

For the lower bound on the distance, note that the largest number of zeros in a code word in $\mathcal{C}$ is bounded above by

$$\max_{\sigma \in V} \#\left(C \cap (\sigma)\right),$$

just as in Proposition IV.1. We have already seen that

$$C \cdot (\sigma) = \alpha(r-1) + n - \mathfrak{d}.$$

However, for every $\sigma \in V$, the curves C and (σ) intersect at the point $(\infty, \infty) \in S(\mathbb{F}_q)$. We claim this happens with multiplicity *at least* $\alpha(r-1)$, and hence

$$\max_{\sigma \in V} \#\left(C \cap (\sigma)\right) \leq C \cdot (\sigma) - \alpha(r-1) = n - \mathfrak{d},$$

from which we deduce that

$$d \geq n - \max_{\sigma \in V} \#\left(C \cap (\sigma)\right) \geq \mathfrak{d}.$$

To establish the claim on the multiplicity of C and (σ) at (∞, ∞), note that the point (∞, ∞) is the origin of the affine patch $\mathbb{A}^2_{(y,u)}$ of S. In this patch, an affine equation for C is

$$C : u^\alpha = (1 + u^\alpha) y^{r+1},$$

which is in fact singular at the origin (this only helps increase the multiplicity of the intersection with the curve (σ)). In the complete local ring of C at the origin, the quantity $1 + u^\alpha$ has an α-th root. More precisely, let

$$A = k[y, u]/(u^\alpha - (1 + u^\alpha) y^{r+1})$$

be the affine coordinate ring of C, and let $\mathfrak{m} = (y, u)$ be the maximal ideal corresponding to the origin. Then in the completed local ring $\hat{A}_{\mathfrak{m}}$, the binomial expansion shows that

$$w = (1+u^\alpha)^{1/\alpha} = 1 + \binom{1/\alpha}{1} u^\alpha + \binom{1/\alpha}{2} u^{2\alpha} + \binom{1/\alpha}{3} u^{3\alpha} + \cdots$$

Let ζ denote an α-th root of unity in an algebraic closure of $\mathbb{F}_q$. Geometrically, C has α branches at the origin:

$$u = w y^{(r+1)/\alpha}, u = \zeta w y^{(r+1)/\alpha}, \ldots, u = \zeta^{\alpha-1} w y^{(r+1)/\alpha},$$

For each one of these branches, y is a uniformizer for the ideal $\mathfrak{m}$, and u has valuation $(r+1)/\alpha$ with respect to this uniformizer.[4] For $\sigma \in V$, a local equation for (σ) in the affine patch $\mathbb{A}^2_{(y,u)}$ is

$$a_0(1,u) y^{r-1} + u^{\epsilon_1} \cdot a_1(1,u) y^{r-2} + \cdots + u^{\epsilon_{r-1}} a_{r-1}(1,u) = 0$$

The monomial $u^{\epsilon_i} y^{r-1-i}$ has $\mathfrak{m}$-adic valuation

$$\left\lceil \frac{\alpha i}{r+1} \right\rceil \cdot \frac{r+1}{\alpha} + r - 1 - i.$$

As i ranges through $0, \ldots, r-1$, the *smallest* value of this quantity is $r - 1$. Hence, on each branch of C the minimal $\mathfrak{m}$-adic valuation of $\sigma \in V$ is $r - 1$, and therefore C and (σ) intersect at (∞, ∞) with multiplicity $\geq \alpha(r-1)$. This concludes the proof of the lower bound for d.

Next, we compute an upper bound for d using (I.1):

$$\begin{aligned}d &\leq n - k - \left\lceil \frac{k}{r} \right\rceil + 2 \\ &= n - r(N+1) - 2\alpha + \frac{(\alpha+1)(r+1)}{2} \\ &\quad - (N+1) - \left\lceil \frac{2\alpha}{r} - \frac{(\alpha+1)(r+1)}{2r} \right\rceil + 2 \\ &= \mathfrak{d} - (r+1) - 2\alpha + \frac{(\alpha+1)(r+1)}{2} \\ &\quad - \left\lceil \frac{2\alpha}{r} - \frac{(\alpha+1)(r+1)}{2r} \right\rceil + 2 \\ &= \mathfrak{d} + \frac{(\alpha-1)(r-3)}{2} \\ &\quad - \left\lceil \frac{2\alpha}{r} - \frac{(\alpha+1)(r+1)}{2r} \right\rceil.\end{aligned}$$

Finally, we discuss the locality of the code $\mathcal{C}$. Since all points $P_{i,j}$ used to construct $\mathcal{C}$ have $y_{i,j} = u_i = 1$, the set $\{P_{i,j}\}$ lies entirely in the affine patch $\mathbb{A}^1_x \times \mathbb{A}^1_t$ of S. Proceeding as in §III-B, we map this affine patch to $\mathbb{A}^{r-1} \times \mathbb{A}_t$ via

$$(x, t) \mapsto (x, x^2, \ldots, x^{r-1}; t).$$

The image of the points $\{P_{i,j}\}$ lie on a rational normal curve, so no r of them lie on a hyperplane, and hence Lemma II.1 shows the code $\mathcal{C}$ has locality r. □

Remark IV.7: Let us compare the parameters in Proposition IV.5 with those of a base-line codes in Proposition II.2. The length n is the same for both constructions. The dimension is smaller in Proposition IV.5; however, on the one hand, $\mathfrak{d} = (b - N)(r + 1)$, and on the other hand, when $M = N$ and $r \geq 3$ the lower bound for the distance in the base-line codes is $2(b - N)$. Hence, the lower bound $\mathfrak{d}$ represents an improvement on base-line codes of $(r - 1)(b - N)$. For a numerical example, take $r + 1 = \alpha = 5$ and $q = 16$. Then we can take $b = 10$, so $n = 50$ and $\mathfrak{d}$ can be any integer divisible by 5 with $\mathfrak{d} \leq 35$ and the parameters are given as in Proposition IV.5 with equality $d = \mathfrak{d}$.

[4]By this we mean: let $B = \bar{\mathbb{F}}_q[y, u]/(u - \zeta^i w y^{(r+1)/\alpha})$ be the geometric local coordinate ring of one of the branches of C. Then the $\mathfrak{m}$-adic completion $\hat{B}_{\mathfrak{m}}$ at the maximal ideal $\mathfrak{m} = (y, u)$ corresponding to the origin is a local discrete valuation ring. Hence the ideal $\mathfrak{m}\hat{B}_{\mathfrak{m}}$ is principal [14, Proposition 9.2]. The equation of the branch shows that y is a generator for this ideal, and that $u \in \mathfrak{m}^{(r+1)/\alpha} \setminus \mathfrak{m}^{(r+1)/\alpha - 1}$, which is to say that u has $\mathfrak{m}$-adic valuation $(r+1)/\alpha$.

V. Codes on Hirzebruch Surfaces

The ruled surface $\mathbb{P}^1 \times \mathbb{P}^1$ is an example of a Hirzebruch surface, which are ruled surfaces determined by a non-negative integer m. After recalling some of the geometry of these surfaces, we adapt the construction of codes in §IV to the setting of Hirzebruch surfaces.

A. Hirzebruch Surfaces $\mathbb{F}(m)$

Let $m \in \mathbb{Z}_{\geq 0}$; we let two copies of the multiplicative group $\mathbb{G}_m \times \mathbb{G}_m$ act on the product of two punctured affine planes $\mathbb{A}^2 \setminus \{(0,0)\} \times \mathbb{A}^2 \setminus \{(0,0)\}$ via

$$(\lambda, 1) : (x, y; t, u) \mapsto (\lambda^{-m}x, y; \lambda t, \lambda u)$$
$$(1, \mu) : (x, y; t, u) \mapsto (\mu x, \mu y; t, u).$$

The Hirzebruch surface $S = \mathbb{F}(m)$ is the quotient

$$\mathbb{A}^2 \setminus \{(0,0)\} \times \mathbb{A}^2 \setminus \{(0,0)\} \,/\mathbb{G}_m \times \mathbb{G}_m.$$

Such surfaces are endowed with a natural fibration $p\colon S \to \mathbb{P}^1_{(t:u)}$ given by

$$((x : y), (t : u)) \to (t : u). \tag{V.1}$$

Note that $\mathbb{P}^1 \times \mathbb{P}^1 = \mathbb{F}(0)$.

Lemma V.1: Let $S = \mathbb{F}(m)$ be as above. The following hold:

1) The Picard group $\mathrm{Pic}(S)$ is isomorphic to $\mathbb{Z}^2$, generated by the classes of the curves
$$A = \{t - u = 0\} \quad \text{and} \quad B = \{x = 0\},$$
which are, respectively, a fiber of (V.1) and the so-called negative section of S.
2) The intersection pairing on $\mathrm{Pic}(S)$ is determined by
$$A^2 = 0, \quad A \cdot B = 1 \quad \text{and} \quad B^2 = -m.$$
3) Let $M = mA + B \in \mathrm{Pic}(S)$. The canonical divisor K_S is linearly equivalent to $(m-2)A - 2M$.
4) For non-negative integers α, β satisfying $\alpha \geq m\beta - 1$, the Riemann–Roch space $\mathcal{L}(S, \alpha A + \beta B)$ has dimension
$$\ell(S, \alpha A + \beta B) = (\alpha + 1)(\beta + 1) - m\frac{\beta(\beta+1)}{2}.$$

Proof: For (1), (2) and (3) see [14], Sections B.2.9 and B 2.7. The Riemann–Roch theorem for surfaces gives the Euler characteristic of the class $\alpha A + \beta B$:

$$\frac{(\alpha A + \beta B) \cdot (\alpha A + \beta B - K_S)}{2} + 1 = (\alpha + 1)(\beta + 1) - n\frac{\beta(\beta+1)}{2}.$$

By, e.g., [16, Thm. 2.1.], the conditions $\beta \geq 0$ and $\alpha \geq m\beta - 1$ guarantee that this Euler characteristic coincides with the dimension of the Riemann–Roch space $\mathcal{L}(S, \alpha A + \beta B)$. □

Remark V.2: The morphism $\phi\colon X \to \bar{X} \subset \mathbb{P}^m$ defined by the sections generating the projectivized Riemann-Roch space $|M|$ is the natural resolution of the cone over the rational normal curve of degree n. The map ϕ contracts B to the vertex of the cone (see [15, B 2.9]).

B. Riemann-Roch Spaces for Codes

In this section, we give an explicit description of the elements of the Riemann–Roch spaces $V_{\beta,\alpha} := \mathcal{L}(S, \alpha A + \beta B)$ appearing in Lemma V.1. We assume throughout that α and β are non-negative integers.

Lemma V.3: Let $\alpha = \varepsilon + m\beta$ with $\varepsilon \geq 0$. The elements of $V_{\beta,\alpha}$ have the form

$$\sigma = a_0(t,u)y^\beta + a_1(t,u)y^{\beta-1}x + \cdots + a_\beta(t,u)x^\beta \tag{V.2}$$

where $a_i(t,u)$ is a homogeneous polynomial of degree $\varepsilon + im$ for $i = 0, \ldots, \beta$. We have

$$\dim V_{\beta,\alpha} = (\alpha+1)(\beta+1) - m\frac{\beta(\beta+1)}{2}.$$

Proof: Let σ be as in the statement of the lemma. First, we show that $\sigma \in V_{\beta,\alpha}$. Since A and B generate $\mathrm{Pic}(S)$, there are α' and β' such that $(\sigma) = \alpha' A + \beta' B$ as classes in $\mathrm{Pic}(S)$. To determine α' and β' we use the intersection pairing on $\mathrm{Pic}(S)$.

Since A is a curve defined by fixing the ratio t/u, we have that

$$(\sigma) \cdot A = \beta.$$

On the other, since $B = \{x = 0\}$, we see that

$$(\sigma) \cdot B = \varepsilon.$$

We obtain the system of equations

$$\beta = (\sigma) \cdot A = \alpha' \cdot A^2 + \beta' A \cdot B = \beta',$$
$$\varepsilon = (\sigma) \cdot B = \alpha' A \cdot B + \beta' B^2 = \alpha' - m\beta'.$$

Thus $\beta' = \beta$ and $\alpha' = \varepsilon + m\beta' = \alpha$ as claimed. Note that the condition that $a_i(t,u)$ is homogeneous of degree $\varepsilon + im$ ensures that the monomials are invariant under the action $(\lambda, 1) \in \mathbb{G}_m \times \mathbb{G}_m$.

The subspace of $V_{\beta,\alpha}$ generated by elements of the form (V.2) has dimension

$$\begin{aligned} k &= (\varepsilon+1) + (\varepsilon+1+m) + \cdots + (\varepsilon+1+\beta m) \\ &= \sum_{i=0}^{\beta} (\varepsilon+1) + im \\ &= (\beta+1)(\varepsilon+1) + m\frac{\beta(\beta+1)}{2} \\ &= (\alpha+1)(\beta+1) - m\frac{\beta(\beta+1)}{2} \end{aligned}$$

and hence must be equal to the entire vector space, by Lemma V.1(4). □

C. A Coarse Construction

Let r and $b \leq q$ be positive integers, and set $n = b(r+1)$. Choose an integer $\mathfrak{d}$, divisible by $r+1$, so that

$$N := \frac{n - \mathfrak{d}}{r+1}$$

is an integer, as well as a positive integer α. Set $\beta = r - 1$, and consider a curve of the form

$$C : g(x, y; t, u) = 0$$

in S, where g is an element of

$$V_{r+1,\alpha+m(r+1)} = \mathcal{L}\left((\alpha+m(r+1))A + (r+1)B\right).$$

We say C is of type $(r+1, \alpha+m(r+1))$. The fibration $p\colon S \to \mathbb{P}^1_{(t:u)}$ in (V.1) gives S the structure of a ruled surface. To construct evaluation codes using C, pick b points $(t_i : u_i) \in \mathbb{P}^1_{(t:u)}(\mathbb{F}_q)$ such that the fiber $p^{-1}((t_i : u_i)) \cap C$ consists of $r+1$ distinct points

$$(x_{i,1} : y_{i,1}), \ldots, (x_{i,r+1} : y_{i,r+1}).$$

Put

$$P_{i,j} = ((x_{i,j} : y_{i,j}), (t_i : u_i)) \in S(\mathbb{F}_q),$$

so that there are $n = b(r+1)$ points of the form $P_{i,j}$ in total. We shall use the vector space

$$V_{\beta,N+m\beta} = V_{r-1,N+m(r-1)}$$

to construct our evaluation codes.

Proposition V.4: The code

$$\mathcal{C} := \{(\sigma(P_{i,j}))_{1\le i\le b, 1\le j\le r+1} : \sigma \in V_{r-1,N+m(r-1)}\},$$

constructed using C, has locality r and its parameters satisfy

$$\begin{aligned} n &= b(r+1)\\ k &= (N+1)r + m\frac{r(r-1)}{2}\\ d &\le \mathfrak{d} - (r-1) - m\frac{(r^2-1)}{2}\\ d &\ge \mathfrak{d} - (r-1)(\alpha + m(r+1)). \end{aligned}$$

Proof: By Lemma V.2, we have

$$k = \dim V_{r-1,N+m(r-1)} = r(N+1) + m\frac{r(r-1)}{2}. \quad \text{(V.3)}$$

Next, if r is odd or m even, we have

$$\left\lceil \frac{k}{r} \right\rceil = N + 1 + m\frac{(r-1)}{2}.$$

Otherwise,

$$\left\lceil \frac{k}{r} \right\rceil = N + 1 + m\frac{(r-1)}{2} + \frac{1}{2} \ge N + 1 + m\frac{(r-1)}{2}.$$

Hence, an upper bound for d using (I.1) is

$$\begin{aligned} d &\le n - k - \left\lceil \frac{k}{r} \right\rceil + 2\\ &\le n - r(N+1) - m\frac{r(r-1)}{2} - (N+1) - m\frac{(r-1)}{2} + 2\\ &= n - (n - \mathfrak{d}) - (r+1) - m\frac{r(r-1)}{2} - m\frac{(r-1)}{2} + 2\\ &= \mathfrak{d} - (r-1) - m\frac{(r^2-1)}{2}. \end{aligned}$$

As in the proof of Proposition IV.5, a lower bound for the minimum distance of $\mathcal{C}$ is

$$\begin{aligned} d &\ge n - \max_{\sigma\in V} \#(C \cap (\sigma))\\ &\ge n - C\cdot(\sigma) \text{ for any } \sigma \in V_{r-1,N+m(r-1)} \end{aligned}$$

Since the equation for C is an element of $V_{r+1,\alpha+m(r+1)}$, we may use Lemma V.1(2) to compute

$$\begin{aligned} &C\cdot(\sigma)\\ &= ((\alpha+m(r+1))A + (r+1)B)\cdot\\ &((N+m(r-1))A + (r-1)B)\\ &= (r-1)(\alpha+m(r+1)) + (N+m(r-1))(r+1)\\ &- m(r^2-1)\\ &= (r-1)(\alpha+m(r+1)) + n - \mathfrak{d}, \end{aligned}$$

and hence

$$d \ge \mathfrak{d} - (r-1)(\alpha + m(r+1)).$$

as claimed. Finally, the locality is r by the same argument as in the end of the proof of Proposition IV.5. □

Remark V.5: When $m = 0$, we have $S = \mathbb{F}(0) = \mathbb{P}^1 \times \mathbb{P}^1$. In this case, the bounds on the distance for $\mathcal{C}$ coincide with the bounds of Proposition IV.1, as one would expect.

Remark V.6: The upper and lower bounds for the minimum distance in Proposition V.4 meet when

$$1 + m\frac{(r+1)}{2} = \alpha + m(r+1).$$

Since α, m and r are non-negative, we must have $m = 0$ (i.e., $S = \mathbb{P}^1 \times \mathbb{P}^1$) and $\alpha = 1$.

D. Refining the Construction

Consider the curve $C \subset S$ with affine model given by

$$C : x^{r+1} = t^\alpha + 1.$$

The projective closure of this curve in S is given by:

$$u^{\alpha+m(r+1)}x^{r+1} - (t^\alpha + u^\alpha)y^{r+1} = 0. \quad \text{(V.4)}$$

The left hand side of the above equation is an element of the vector space $V_{r+1,\alpha+m(r+1)}$.

To construct evaluation codes using C, as usual, pick b points $(t_i : u_i) \in \mathbb{P}^1_{(t:u)}(\mathbb{F}_q)$ such that the fiber $p^{-1}((t_i : u_i)) \cap C$ consists of $r+1$ distinct points

$$(x_{i,1} : y_{i,1}), \ldots, (x_{i,r+1} : y_{i,r+1}).$$

Put

$$P_{i,j} = ((x_{i,j} : y_{i,j}), (t_i : u_i)) \in S(\mathbb{F}_q),$$

so that there are $n = b(r+1)$ points of the form $P_{i,j}$ in total. For the vector space of function on which we evaluate the $P_{i,j}$, we constrain the degree in t of the polynomials $a_i(t,u)$, as follows:

$$\deg_t a_i(t,1) \le N + im - \left\lceil \frac{i(\alpha+m(r+1))}{r+1} \right\rceil.$$

Again, this requires $N \ge \lceil \alpha(r-1)/(r+1) \rceil$, which we now assume. In other words, setting

$$\epsilon_i := \left\lceil \frac{i(\alpha+m(r+1))}{r+1} \right\rceil,$$

we assume that for each $0 \le i \le r-1$,

$$a_i(t,u) = u^{\epsilon_i}\cdot a_i'(t,u)$$

for a homogeneous polynomial $a_i'(t,u)$. When this is the case, the calculation (V.3) shows that the vector space of functions

$$V := \{\sigma \in V_{r-1,N+m(r-1)} : \sigma = a_0(t,u)y^{r-1} + u^{\epsilon_1}\cdot a_1(t,u)y^{r-2}x + \cdots + u^{\epsilon_{r-1}}a_{r-1}(t,u)x^{r-1}\}$$

has dimension

$$k = r(N+1) + m\frac{r(r-1)}{2} - \sum_{i=0}^{r-1}\epsilon_i$$

If $\alpha = r+1$ then

$$\sum_{i=0}^{r-1}\epsilon_i = \sum_{i=0}^{r-1} i + im = (m+1)\frac{r(r-1)}{2}$$

Otherwise, if $r+1 > \alpha$ then

$$\sum_{i=0}^{r-1}\epsilon_i = \sum_{i=0}^{r-1}\left\lceil \frac{i\alpha}{r+1}\right\rceil + im = \frac{(\alpha+1)(r+1)}{2} - 2\alpha + m\frac{r(r-1)}{2}.$$

where the second equality follows by our work in the proof of Proposition IV.5. We conclude that

$$k = \begin{cases} r(N+1) - \frac{r(r-1)}{2}, \text{if } r+1=\alpha, \text{ and} \\ r(N+1)+2\alpha - \frac{(\alpha+1)(r+1)}{2} - m\frac{r(r-1)}{2}, \text{if } r+1>\alpha, \end{cases}$$

Theorem V.7: Assume that $\alpha \mid (r+1)$ and $(r+1) \mid (q-1)$. The code

$$\mathcal{C} := \{(\sigma(P_{i,j}))_{1\le i\le b, 1\le j\le r+1} : \sigma \in V\}$$

has locality r and its parameters satisfy

$$n = b(r+1),$$
$$k = \begin{cases} r(N+1) - \frac{r(r-1)}{2}, \text{if } r+1=\alpha, \text{ and} \\ r(N+1) + 2\alpha - \frac{(\alpha+1)(r+1)}{2} - m\frac{r(r-1)}{2}, \text{if } r+1>\alpha, \end{cases}$$
$$d \le \mathfrak{d} + \frac{(\alpha-1)(r-3)}{2} - \left\lceil \frac{2\alpha}{r} - \frac{(\alpha+1)(r+1)}{2r}\right\rceil + m\frac{(r^2-1)}{2},$$
$$d \ge \mathfrak{d}.$$

Proof: We have already discussed the values of n and k above. The upper bound for d is obtained from (I.1), proceeding as in the proof of Proposition IV.5.

For the lower bound on the distance, we note that, as before,

$$d \le \max_{\sigma\in V} \#\left(C \cap (\sigma)\right),$$

just as in Proposition IV.1. In the course of the proof of Proposition V.4, we saw that

$$C\cdot(\sigma) = (r-1)(\alpha + m(r+1)) + n - \mathfrak{d}.$$

However, for every $\sigma \in V$, the curves C and (σ) intersect at the point

$$[x,y;t,u] = [1,0;1,0] \in \mathbb{F}(m).$$

We claim this happens with multiplicity *at least* $(r-1)(\alpha + m(r+1))$, and hence

$$\max_{\sigma\in V} \#\left(C\cap(\sigma)\right) \le C\cdot(\sigma) - (r-1)(\alpha+m(r+1)) = n - \mathfrak{d},$$

from which we deduce that

$$d \ge n - \max_{\sigma\in V} \#\left(C\cap(\sigma)\right) \ge \mathfrak{d}.$$

The claim on the multiplicity is established as in the proof of Proposition IV.5: the point $[1,0;1,0] \in \mathbb{F}(m)$ is the origin of the affine patch of C given by

$$u^{\alpha+m(r+1)} = (1+u^\alpha)y^{r+1},$$

In the complete local ring of C at the origin, the quantity $1+u^\alpha$ has an $(\alpha + m(r+1))$-th root. Let ζ denote an $(\alpha+m(r+1))$-th root of unity in an algebraic closure of $\mathbb{F}_q$. Geometrically, C has $\alpha + m(r+1)$ branches at the origin:

$$u = wy^{(r+1)/(\alpha+m(r+1))}, u = \zeta wy^{(r+1)/(\alpha+m(r+1))}, \ldots, u = \zeta^{(\alpha+m(r+1))-1}wy^{(r+1)/(\alpha+m(r+1))},$$

For each one of these branches, y is a uniformizer for the maximal ideal at the origin of C, and u has valuation $(r+1)/(\alpha+m(r+1))$ with respect to this uniformizer (see the proof of Proposition IV.5 for more details). For $\sigma \in V$, a local equation for (σ) in the affine patch $\mathbb{A}^2_{(y,u)}$ is

$$a_0(1,u)y^{r-1} + u^{\epsilon_1}\cdot a_1(1,u)y^{r-2} + \cdots + u^{\epsilon_{r-1}}a_{r-1}(1,u) = 0$$

The monomial $u^{\epsilon_i}y^{r-1-i}$ has $\mathfrak{m}$-adic valuation

$$\left\lceil \frac{i(\alpha+m(r+1))}{r+1}\right\rceil \cdot \frac{r+1}{(\alpha+m(r+1))} + r - 1 - i.$$

As i ranges through $0,\ldots,r-1$, the *smallest* value of this quantity is $r-1$. Hence, on each branch of C the minimal valuation at the origin of $\sigma \in V$ is $r-1$, and therefore C and (σ) intersect at $[1,0;1,0]$ with multiplicity $\ge \alpha(r-1)(\alpha+m(r+1))$. This concludes the proof of the lower bound for d. □

When $m=0$, we recover Proposition IV.5. The parameters get slightly worse for $m>0$ but this more general construction might still be interesting.

VI. Locally Recoverable Codes From Elliptic Surfaces

A. Elliptic Surfaces

The definitions of this section hold over an arbitrary field k.

An algebraic surface $\mathcal{E}$ is called an elliptic surface if it is endowed with a morphism $\pi : \mathcal{E} \to B$ to a base algebraic curve B such that

i) for all but finitely many $t \in B(\bar{k})$, the fiber $\pi^{-1}(t)$ is a genus one curve, where $\bar{k}$ is a fixed algebraic closure of k.
ii) there is a section σ to π, i.e., a morphism $\sigma : B \to \mathcal{E}$ such that $\pi\circ\sigma = \mathrm{id}_B$.

The morphism π is called an elliptic fibration. Condition ii) implies that all but finitely many fibers of π are indeed elliptic curves.

Let $\pi : \mathcal{E} \to B$ be an elliptic fibration. A section $P : B \to \mathcal{E}$ is, by definition, a regular map such that $\pi\circ P$ is the identity on B. We denote by $\mathcal{O}$ the zero section and by abuse of notation also the zero element of any fiber. The set of sections of the

fibration π in the above sense can be made into an abelian group with identity $\mathcal{O}$ (in the same way one defines the group law on an elliptic curve). This group is called the Mordell-Weil group of $\mathcal{E}$ and it is finitely generated by the Néron-Severi-Mordell-Weil theorem.

We also have that $\mathcal{E}$ has a Weierstrass equation

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6$$

where $a_i \in k(B)$. We consider the divisor $D = n \cdot \infty + m \cdot \mathcal{O}$, where ∞ is the "fiber above ∞", and $\mathcal{O}$ is the zero section. A function on $\mathcal{E}$ whose polar divisor is bounded by D is of the form

$$\sum_{2i \leq m} \alpha_i x^i + \sum_{2i+3 \leq m} \beta_i x^i y,$$

where α_i and β_i are functions in the Riemann Roch space $\mathcal{L}(B, n \cdot \infty)$.

Each fiber E is embedded in $\mathbb{P}^{n-1}$ by the linear system $|n\mathcal{O}|$ (where $\mathcal{O}$ is the identity of E).

B. General Code Construction

Let $\pi : \mathcal{E} \to B$ be an elliptic fibration. We denote by $\mathcal{O}$ the zero section and by abuse of notation also the zero element of any fiber. We denote by $E_t = \pi^{-1}(t)$ the fiber above t and by $E_t[2]$ its subgroup of elements of order at most 2.

Lemma VI.1: Assume that for each t in a subset of $B(\mathbb{F}_q)$ such that the fiber E_t over t is an elliptic curve, we are given $\Gamma_t \subset E_t(\mathbb{F}_q) - E_t[2]$ all of same cardinality $r+1$ for some integer r with the property that $\sum_{P \in \Gamma_t} P \in E_t[2]$ in the group law of E_t.

Let $\Gamma = \bigcup_t \Gamma_t$ and V a finite-dimensional $\mathbb{F}_q$-vector space of functions on $\mathcal{E}$ such that the restriction of any element of V to a fiber above any t is in the Riemann-Roch space $\mathcal{L}(E_t, r\mathcal{O})$. We form a code $\mathcal{C}$ by evaluating the functions on V on the points of Γ. The code $\mathcal{C}$ is locally recoverable with locality r.

Proof: Given a function f and codeword $c = (f(P))_{P \in \Gamma}$ and suppose we need to recover $f(P_0)$. We have that $P_0 \in \Gamma_t$ for some t. Now, the restriction of f to E_t is a rational function f_t on E_t, which is an element of the Riemann-Roch space $\mathcal{L}(E_t, r\mathcal{O})$. We claim that $f_t(P_0)$ can be uniquely recovered from the values of $f_t(P), P \in \Gamma_t - \{P_0\}$. If there are two such functions with the same values, their difference vanishes at $\Gamma_t - \{P_0\}$ but has a pole of order at most r at $\mathcal{O}$. The only possibilty is that this function then has simple zeros at the points of $\Gamma_t - \{P_0\}$, a pole of order r at $\mathcal{O}$ and no other zeros or poles. That would imply, using Abel's theorem on E_t ([17, Corollary III 3.5]), that $\sum_{P \in \Gamma_t - \{P_0\}} P = \mathcal{O}$ and thus $P_0 \in E_t[2]$, which contradicts our hypothesis. This shows that the map $L(r\mathcal{O}) \to \mathbb{F}_q^r, h \mapsto (h(P))_{P \in \Gamma_t - \{P_0\}}$ is injective. As these spaces have the same dimension by Riemann-Roch, it is also surjective. □

A natural example is to take sections $P_i, i = 1, \ldots, r$ of the elliptic fibration $\pi : \mathcal{E} \to B$. If we let $P_{r+1} = -\sum_{i=1}^r P_i$ and $\Gamma_t = \{P_1(t), \ldots, P_{r+1}(t)\}$, we are in the above situation.

We can also use an irreducible curve C in $\mathcal{E}$. Then we have a map $C \to B$ and we assume that it has degree $r+1$ and take as Γ_t the fibers of this map above points that split completely. To ensure that the points of Γ_t add to zero we need to check the algebraic point defined by C has trace zero. Often the following lemma is useful.

Lemma VI.2: Let $\pi : \mathcal{E} \to B$ be an elliptic surface with finite Mordell-Weil group of order prime to the characteristic of k. Let C be an irreducible curve in $\mathcal{E}$ such that the map $C \to B$ is separable of degree $r+1$. If, for one $t \in B$ with $\pi^{-1}(t)$ an elliptic curve and whose preimage $\Gamma_t = (\pi|_C)^{-1}(t)$ in C has $r+1$ distinct points we have that $\sum_{P \in \Gamma_t} P = \mathcal{O}$, then for all other such t, we also have $\sum_{P \in \Gamma_t} P = \mathcal{O}$.

Proof: We can base change $\pi : \mathcal{E} \to B$ to $\pi' : \mathcal{E}' \to C$ via $C \to B$ and C itself pulls back to a section s of π' and we can then take the $C \to B$ trace of this section to get a section of π. Concretely, this section consists of adding the points on $(\pi|_C)^{-1}(t)$ and viewing that as a function of $t \in B$. By the assumption on the Mordell-Weil group, this section is of finite order. From [17, Proposition VII 3.1], for sections of finite order prime to the characteristic, the specialization map to a smooth fiber is injective. By assumption, for one such fiber, the specialization of s is zero. It follows that s itself is zero. □

Here are some explicit examples.

Example VI.3: Take $\mathcal{E}$ the Legendre family $y^2 = x(x-1)(x-t)$ and consider the curve $C : (u^2+t+1)^2 = u(u-1)(u-t)$ of genus 1 embedded in $\mathcal{E}$ by taking $x = u, y = u^2 + t + 1$, so $r = 3$. Lemma VI.2 applies with $t = -1$. If Γ has n points and $d < n, 4|(n-d)$, we consider functions of the form $f = a(t) + b(t)x + c(t)y$ with $\deg a \leq (n-d)/4, \deg b, \deg c < (n-d)/4$ and these restrict to C as a function of degree at most $n-d$, so the minimum distance is at least d. The dimension is $k = 3(n-d)/4+1$ and it follows that $d = n-k-\lceil k/3 \rceil+2$, i.e., the code is optimal, but typically not as long as the optimal codes from the previous sections.

Example VI.4: Let $\mathcal{E}$ be the elliptic surface $y^2 = x^3 + x - t^2 - 1$ over $\mathbb{F}_q$ and C the curve given by $x = y^2$ inside $\mathcal{E}$, which is $y^6 = t^2 + 1$. The elliptic surface has trivial Mordell Weil group over $\mathbb{F}_q(t)$ so the multisection corresponding to C automatically has trace zero. This leads to the same family of codes corresponding to the case $r = 5$ of subsection III-C by considering evaluation on functions of the form $f = a_0(t) + a_1(t)x + a_2(t)y + a_3(t)x^2 + a_4(t)xy$.

Example VI.5: We can also recover the case $r = 3$ of subsection III-C by taking $\mathcal{E}$ to be the elliptic surface $y^2+xy = x^3 + t^2 + 2$ over $\mathbb{F}_q$ and C the curve given by $x^2 = y = u$ inside $\mathcal{E}$, which is $u^4 = t^2 + 2$ and evaluation on functions of the form $f = a_0(t) + a_1(t)x + a_2(t)y$. We can take, for $q = 5, 13$ respectively, sets of size $b = 2, 4$ and get codes of length $n = 8, 16$.

Yet another example is a variant of the examples constructed by Ulmer [18] leading to the following theorem.

Theorem VI.6: For every odd prime (power) p and integer $d \leq 2(p+1)(p-2), (p+1)|d$, there exists a locally recoverable code $\mathcal{C}$ over $\mathbb{F}_{p^2}$ of recoverability p, length $n = 2(p+1)(p-2)$, minimum distance d and dimension

$$k = \frac{p(n-d)}{p+1} - \frac{p-1}{2}.$$

Proof: Consider the surface $\mathcal{E} : y^2 = x(x+1)(x+t^2+1)$ over $\mathbb{F}_{p^2}$, p odd and the curve C defined by $u^{p+1} = t^2 + 1$. Then C embeds in $\mathcal{E}$ by taking $x = u, y = u(u+1)^{(p+1)/2}$. The points on C on the fiber above $t = b$ are of the form $(c, c(c+1)^{(p+1)/2})$ for each c satisfying $c^{p+1} = b^2 + 1$. The function $y(x+1)^{(p-1)/2} - (x + b^2 + 1)$ has degree $p+2$ and vanishes on all these points and on the point $(-b^2 - 1, 0)$ of order 2. So lemma VI.1 applies once we exclude the points on C with $c = 0, c^{p+1} = 1$. Each allowed value of c gives two values of b since $c^{p+1} - 1 \in \mathbb{F}_p$ so has square roots in $\mathbb{F}_{p^2}$. So we have $n = 2(p+1)(p-2)$ points in C we can use to form Γ.

To construct a code we consider the following vector space, where $x_i = x^{(i+1)/2}, i$ odd and $x_i = yx^{(i-2)/2}, i$ even, $i > 0$.

$$V = \left\{ a_0(t) + \sum_{i=1}^{p-1} a_i(t) x_i : \deg a_i \leq N_i \text{ for } i = 0, \ldots, p-1 \right\}$$

where $N_0 = \frac{n-d}{p+1}$,

$$N_i = \frac{n-d}{p+1} - 1, \; i \text{ odd }, \qquad N_i = \frac{\text{n-d}}{p+1} - 2, \; i \text{ even }, i > 0. \tag{VI.1}$$

chosen so that the elements of V restrict to functions of degree $n-d$ on C and the codewords have weight at least d. The dimension k satisfies $k = \sum_{i=0}^{p-1}(N_i + 1)$ and the result follows. □

Remark VI.7: Note that, in the above theorem $d_{\text{opt}} = n - k - \lceil k/p \rceil + 2 = d + (p+3)/2$.

ACKNOWLEDGMENT

The authors would like to thank the following institutions for providing the opportunity for them to meet: IMPA, BIRS-Oaxaca, MPIM, IHP, University of Canterbury, and M. Stoll's Rational Points workshop series. José Felipe Voloch would also like to thank A. Dimakis for a helpful conversation. Finally, the authors would like to thank the referees for a careful reading and a number of suggestions to the manuscript.

REFERENCES

[1] M. F. Atiyah and I. G. Macdonald, *Introduction to Commutative Algebra*. Reading, MA, USA: Addison-Wesley, 1969.

[2] A. Barg, K. Haymaker, E. W. Howe, G. L. Matthews, and A. Várilly-Alvarado, *Locally Recoverable Codes From Algebraic Curves and Surfaces*, vol. 9. Cham, Switzerland: Springer, 2017, pp. 95–127.

[3] A. Barg, I. Tamo, and S. Vladut, "Locally recoverable codes on algebraic curves," *IEEE Trans. Inf. Theory*, vol. 63, no. 8, pp. 4928–4939, Aug. 2017.

[4] I. Coskun and J. Huizenga, "Brill-Noether theorems and globally generated vector bundles on Hirzebruch surfaces," *Nagoya Math. J.*, vol. 238, pp. 1–36, Jun. 2020.

[5] C. Galindo, F. Hernando, and C. Munuera, "Locally recoverable J-affine variety codes," *Finite Fields Their Appl.*, vol. 64, Jun. 2020, Art. no. 101661.

[6] P. Gopalan, C. Huang, H. Simitci, and S. Yekhanin, "On the locality of codeword symbols," *IEEE Trans. Inf. Theory*, vol. 58, no. 11, pp. 6925–6934, Nov. 2012.

[7] V. Guruswami, C. Xing, and C. Yuan, "How long can optimal locally repairable codes be?" *IEEE Trans. Inf. Theory*, vol. 65, no. 6, pp. 3662–3670, Jun. 2019.

[8] R. Hartshorne, "Algebraic geometry," in *Graduate Texts in Mathematics*, no. 52. New York, NY, USA: Springer, 1977.

[9] K. Haymaker, B. Malmskog, and G. L. Matthews, "Locally recoverable codes with availability $t \geq 2$ from fiber products of curves," *Adv. Math. Commun.*, vol. 12, no. 2, pp. 317–336, 2018.

[10] X. Li, L. Ma, and C. Xing, "Optimal locally repairable codes via elliptic curves," *IEEE Trans. Inf. Theory*, vol. 65, no. 1, pp. 108–117, Jan. 2019.

[11] C. Munuera and W. Tenório, "Locally recoverable codes from rational maps," *Finite Fields Their Appl.*, vol. 54, pp. 80–100, Nov. 2018.

[12] C. Munuera *et al.*, "Locally recoverable codes from algebraic curves with separated variables," *Adv. Math. Commun.*, vol. 14, no. 2, pp. 265–278, 2020.

[13] D. S. Papailiopoulos and A. G. Dimakis, "Locally repairable codes," *IEEE Trans. Inf. Theory*, vol. 60, no. 10, pp. 5843–5855, Oct. 2014.

[14] M. Reid, *Chapters on Algebraic Surfaces*, vol. 3. Providence, RI, USA: Amer. Math. Soc., 1997, pp. 3–159.

[15] J. Silverman, "The arithmetic of elliptic curves," in *Graduate Texts in Mathematics*, vol. 106, 2nd ed. Dordrecht, The Netherlands: Springer, 2009.

[16] I. Tamo and A. Barg, "A family of optimal locally recoverable codes," *IEEE Trans. Inf. Theory*, vol. 60, no. 8, pp. 4661–4676, Aug. 2014.

[17] M. Tsfasman, S. Vlăduţ, and D. Nogin, "Algebraic geometric codes: Basic notions," in *Mathematical Surveys and Monographs*, vol. 139. Providence, RI, USA: American Mathematical Society, 2007.

[18] D. Ulmer, "Explicit points on the Legendre curve," *J. Number Theory*, vol. 136, pp. 165–194, Mar. 2014.

Cecília Salgado received the Ph.D. degree in mathematics from the University of Denis Diderot (Paris VII), France, in 2009. She joined the Department of Mathematics, University of Leiden, The Netherlands, in 2009, as Post-Doctoral Researcher, where she stayed until 2011, when she became an Assistant Professor at the Federal University of Rio de Janeiro (UFRJ). Later, she became an Associate Professor at UFRJ, where she stayed until 2021. She is currently an Associate Professor at the Department of Mathematics, Bernoulli Institute, Groningen, The Netherlands. Her research interests include arithmetic algebraic geometry and coding theory.

Anthony Várilly-Alvarado received the A.B. degree in mathematics from Harvard University, Cambridge, MA, USA, in 2003, and the Ph.D. degree in mathematics from the University of California at Berkeley, Berkeley, CA, USA, in 2009. He joined the Department of Mathematics, Rice University, Houston, TX, USA, in 2009, as a G. C. Evans Instructor and was promoted to an Assistant Professor in 2012. He has been a Professor of mathematics at Rice University since 2019. His research interests include arithmetic algebraic geometry and coding theory. He is a fellow of the American Mathematical Society (class of 2021).

José Felipe Voloch received the Ph.D. degree in mathematics from the University of Cambridge, U.K., in 1985. Then, he took a position at IMPA, Rio de Janeiro, Brazil, where he stayed until 1992. After a year as a Visiting Scholar at the University of California at Berkeley, he took a position at the Department of Mathematics, The University of Texas at Austin, where he stayed until 2015. He is currently a Professor of mathematics at the School of Mathematics and Statistics, University of Canterbury. His research interests include number theory, algebraic geometry, cryptography, and coding theory.