



# Algebraic hierarchical locally recoverable codes with nested affine subspace recovery

Kathryn Haymaker<sup>1</sup> · Beth Malmskog<sup>2</sup> · Gretchen Matthews<sup>3</sup>

Received: 20 February 2024 / Revised: 16 August 2024 / Accepted: 1 October 2024  
© The Author(s) 2024

## Abstract

Codes with locality, also known as locally recoverable codes, allow for recovery of erasures using proper subsets of other coordinates. These subsets are typically of small cardinality to promote recovery using limited network traffic and other resources. Hierarchical locally recoverable codes allow for recovery of erasures using sets of other symbols whose sizes increase as needed to allow for recovery of more symbols. In this paper, we describe a hierarchical recovery structure arising from geometry in Reed–Muller codes and codes with availability from fiber products of curves. We demonstrate how the fiber product hierarchical codes can be viewed as punctured subcodes of Reed–Muller codes, uniting the two constructions. This point of view provides natural structures for local recovery with availability at each level in the hierarchy.

**Keywords** Locality · Hierarchy · Availability · Fiber product codes · Reed–Muller codes

**Mathematics Subject Classification** 94B27 · 11T71 · 14H05

---

Communicated by C. Li.

---

Beth Malmskog and Gretchen Matthews have contributed equally to this work.

---

✉ Kathryn Haymaker  
kathryn.haymaker@villanova.edu

Beth Malmskog  
bmalmskog@coloradocollege.edu

Gretchen Matthews  
gmatthews@vt.edu

<sup>1</sup> Department of Mathematics and Statistics, Villanova University, 800 E. Lancaster Ave, Villanova, PA 19085, USA

<sup>2</sup> Department of Mathematics & Computer Science, Colorado College, 819 N. Tejon St., Colorado Springs, CO 80903, USA

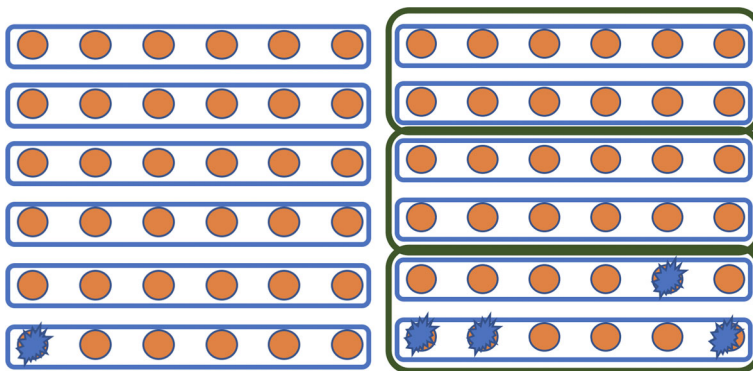
<sup>3</sup> Department of Mathematics, Virginia Tech, 225 Stanger Street, Blacksburg, VA 24061-1026, USA

## 1 Introduction

Error-correcting codes provide ways of encoding information into vectors for storage or communication with redundancy included, so that errors can be detected and recovered by the retriever or recipient. Not all errors are created equal, however, nor are all errors equally likely to occur across different applications. For example, symbol erasure (where a vector symbol is erased, leaving a blank coordinate) is easier to detect (and potentially to correct) than symbol changing, because the existence and position of the erasure is known. In cloud storage, a regular event and major concern is that one or more servers will fail catastrophically or be overloaded with requests for data so that they are unable to satisfy additional queries, effectively erasing all data on the server. Locally recoverable codes are motivated by the desire to facilitate erasure recovery in this setting. The idea is to determine a small set of helper positions for each position  $i$  so that position  $i$  may be recovered using only the helper set. A linear code is said to have locality  $r$  if for each coordinate  $i$  of a codeword, there is a set of  $r$  other helper coordinates, called a recovery or helper set, so that in any codeword, the symbol in position  $i$  can be recovered from the symbols in the helper set.

Two structural generalizations that handle multiple simultaneous erasures are availability and hierarchy [3, 23, 28]. A locally recoverable code is said to have availability  $t$ , referred to as an LRC( $t$ ), if each coordinate has  $t$  independent recovery sets. A locally recoverable code is said to have hierarchical recovery and is referred to as an H-LRC if the recovery set for each coordinate is contained in a larger recovery set that can recover additional erasures beyond what the smaller recovery set can, as depicted in Fig. 1.

This feature can be generalized to several levels of hierarchy. These two structural generalizations can be combined, giving rise to hierarchical LRC( $t$ )s. Aside from [1], combined hierarchy and availability is a relatively unstudied area, yet many geometric code constructions naturally possess both availability and hierarchy. Furthermore, geometric constructions often offer great flexibility, allowing hierarchical recovery in many different ways. In this paper, we examine how hierarchy and availability arise naturally from underlying structure in two related types of geometric codes: Reed–Muller codes and codes from fiber products. In [18], the authors construct LRC( $t$ )s from fiber products of  $t$  curves. We relate these fiber product codes to Reed–Muller codes and demonstrate how to perform hierarchical recovery for all such codes. Though Reed–Muller codes have been well-studied, and the hierarchy



**Fig. 1** This is a cartoon illustrating a code with 2-level hierarchy. Small repair groups of size 6 can recover a single erasure, while larger repair groups of size 12 can recover up to 4 erasures

arises quite naturally, this approach to hierarchical locality for Reed–Muller codes has not appeared before in the literature.

There has been extensive and interesting work on codes that are optimal with respect to availability bounds. In one fairly recent example, the authors of [4] construct optimal codes with availability capable of repairing  $\delta - 1$  erased symbols in the case where each repair set contains one check symbol. In [5] the authors prove a new bound for locally repairable codes with multiple disjoint repair sets (availability) and present an optimal construction of codes that meet the bound. A class of locally recoverable codes with availability that meets the generalized Singleton bound for LRCs is presented in [13]. In [24], new classes of LRCs with small locality ( $r = 2$  or  $3$ ) and high availability are presented, and for many cases, the authors show the codes are alphabet-optimal or Greisner-optimal.

There has also been significant work on H-LRCs, including the following examples, which do not include availability at each level. The authors of [9] use matrix product codes, concatenated codes, and cyclic codes to construct optimal H-LRCs with lengths either  $q^2 + q$ ,  $n$  dividing  $q - 1$ , or unbounded. In [10], the authors give an improved bound for H-LRCs for a special set of parameters and present optimal constructions that meet this bound for those parameters. Optimal H-LRCs via generalized Reed–Solomon codes are constructed in [30]. The complete hierarchical locality of the punctured Simplex codes is given in [16]. Note that the codes discussed in [16] also have availability but combined hierarchy and availability is not the topic of that paper.

There are few examples in the literature of LRCs with combined hierarchy and availability. Constructions in [1] of codes with  $H = 2$  and availability at each level are given in subsections VIII.D and VIII.E. The difference between the codes in [1] and the ones here are highlighted in later remarks. In [12], the authors give an overview of the connections between matroid theory and linear LRCs with hierarchical availability via examples. They conclude that a level-specific Singleton-type bound is needed, and that remains an open question in this area.

The general constructions in this paper give multiple choices of recovery set at each level of hierarchy (and typically high-availability, all arising naturally from geometry). A major goal of this paper is to begin the exploration and discussion of codes with structure beyond what has been discussed in the literature—in particular, codes that have availability, hierarchy, and many choices of hierarchical structure (this could be thought of as “availability of hierarchy”). Instead of devising new codes that have this structure, a major contribution of this paper is to describe how these properties arise naturally in existing, well-studied codes. The codes in this paper are generally not optimal with respect to existing bounds for codes with availability or hierarchy (e.g. in [1, 23]) because these bounds do not take into account the additional structure we describe.

However, we show that these generally non-optimal codes offer intrinsic versatility in recovery beyond that offered by codes optimal with respect to bounds for either availability or hierarchy. It remains an open question to establish bounds that fully capture this level of flexibility in recovery.

An important contribution of this work is the illumination of nested affine subspaces within existing code families—specifically Reed–Muller codes and codes from fiber products of curves—that can be used for the purpose of natural hierarchical recovery with high availability. This work is the first time that the combined properties of hierarchy and availability are being exploited in full generality in these code families. Moreover, since many algebraic geometry codes are punctured Reed–Muller codes, the ideas here can be generalized to other structures, resulting in more flexible recovery methods for other existing code families.

This paper is organized as follows. Coding theory preliminaries are contained in Sect. 2. The main families of codes considered in this paper, Reed–Muller codes and fiber product codes, are reviewed here. Section 3 contains results on their localities and availability [18], connects these codes concretely and conceptually, and describes how to obtain basic (2-level) hierarchy from each of these constructions. Section 4 extends this construction to  $H$ -level hierarchy and demonstrates that these constructions also inherently have availability at each level of recovery. Section 5 uncovers the combined availability and hierarchy inherent in the geometry of Reed–Muller and fiber product codes. The conclusion is found in Sect. 6.

## 2 Locality, availability, and code constructions

In this section, we review the notions of locality and availability and set notation to be used throughout the paper. This section contains definitions of terminology of algebraic curves. We also give an overview of the main code constructions that will be needed in later sections.

### 2.1 Background and notation

We use standard notation from coding theory. The finite field with  $q$  elements is denoted by  $\mathbb{F}_q$ . An  $[n, k, d]$  code  $C$  over  $\mathbb{F}_q$  is a  $k$ -dimensional  $\mathbb{F}_q$ -subspace of  $\mathbb{F}_q^n$  in which any elements differ in at least  $d$  coordinates. The set of positive integers is denoted  $\mathbb{Z}^+$ . Given an  $[n, k, d]$  code  $C$  and  $I := \{i_1, \dots, i_s\} \subseteq \{1, \dots, n\}$ , the corresponding punctured code is  $C|_I := \{(c_{i_1}, \dots, c_{i_s}) : (c_1, \dots, c_n) \in C\}$ .

The notion of locality was introduced by Gopalan et al. in 2012 [14]. It was generalized to consider multiple failures shortly thereafter, as captured in the definition below.

**Definition 1** [22] Let  $n, r, \rho \in \mathbb{Z}^+$ . A linear code  $C$  of length  $n$  over  $\mathbb{F}_q$  is  $(r, \rho)$ -**locally recoverable** if for each  $i \in \{1, \dots, n\}$  there exists a punctured code  $C_i$  whose support contains  $i$ , of length at most  $r + \rho - 1$  and minimum distance at least  $\rho$ . The support of  $C_i$  is called a **repair group** for position  $i$ , and  $I_i := \text{Supp } C_i \setminus \{i\}$  is called a **recovery set** for position  $i$ . We sometimes say that  $C$  has  $(r, \rho)$  locality.

Notice that  $\text{Supp } C_i$  is a repair group for all  $j \in \text{Supp } C_i$  not just  $i$  itself.

Given any  $c \in C$  where  $C$  has locality  $(r, \rho)$ , the symbol  $c_i$  in position  $i$  may be recovered using at most  $r$  symbols indexed by elements of  $I_i$ , assuming no more than  $\rho - 1$  erasures have occurred. If  $r$  symbols are not available due to erasures, the codeword symbols indexed by  $\{1, \dots, n\} \setminus \{i\}$  may be used to recover  $c_i$ . This idea could be seen as foreshadowing the notion of hierarchical recovery: if too many erasures prevent recovery by a smaller set ( $I_i$  in this case), recovery is attempted using the larger set  $\{1, \dots, n\} \setminus \{i\}$ . Of course, we are interested in the setting in which there is another set slightly larger than  $I_i$  that can be used for recovery rather than immediately considering all remaining coordinates indexed by  $\{1, \dots, n\} \setminus \{i\}$ . This is known as hierarchical locality, and the formal definition will be given in Sect. 3.

If too many positions within a recovery set become unavailable, local recovery may not be possible. This leads to what is known as the availability problem. One way to address this problem is to construct multiple recovery sets for each position of the codeword. A code is said to have availability  $t$  if each coordinate has  $t$  independent recovery sets. The concept of availability was introduced in [28]. More formally, we have the following definition provided by Ballentine, Barg, and Vlăduț.

**Definition 2** [1] Let  $n, t, r_j, \rho_j \in \mathbb{Z}^+$  for all  $j \in \{1, \dots, t\}$ . A linear code of length  $n$  over  $\mathbb{F}_q$  is  $(r_j, \rho_j)_{1 \leq j \leq t}$ -**locally recoverable with availability**  $t$  if, for each  $i \in \{1, \dots, n\}$  and each  $j \in \{1, \dots, t\}$ , there exists a punctured code  $C_{i,j}$  with support positions  $\{i\} \cup I_{i,j}$  such that  $\dim(C_{i,j}) \leq r_j$ , there are at least  $\dim(C_{i,j})$  linearly independent positions in  $I_{i,j} \setminus \cup_{k \in [t] \setminus \{j\}} I_{i,k}$ , and  $d(C_{i,j}) \geq \rho_j$ ; that is,  $C_{i,j}$  is a  $(r_j, \rho_j)$ -locally recoverable code. The set  $I_{i,j}$  is known as the  $j$ -th recovery set for position  $i$ .

The following definitions related to algebraic curves over finite fields are from Chapter 13 of [20], and we refer the reader to that reference as well as [19, 26] for additional details and examples. For a thorough treatment of algebraic curves and their associated function fields, please see Appendix B of [25].

Let  $\mathbb{F}$  be a field. The  $n$ -dimensional affine space over  $\mathbb{F}$ , sometimes denoted  $\mathbb{A}^n(\mathbb{F})$ , is the  $n$ -dimensional vector space  $\mathbb{F}^n$ , where the elements of  $\mathbb{F}^n$  are called *points*. The  $n$ -dimensional projective space  $\mathbb{P}^n(\mathbb{F})$  is the set of all equivalence classes

$$(\{[x_1 : x_2 : \dots : x_{n+1}] : x_i \in \mathbb{F} \text{ for all } i\} \setminus \{[0 : 0 : \dots : 0]\}) / \sim,$$

where  $[x_1, x_2, \dots, x_{n+1}] \sim [y_1, y_2, \dots, y_{n+1}]$  if and only if  $x_i = \lambda y_i$  for all  $i$ , where  $\lambda \in \mathbb{F} \setminus \{0\}$ . We generally call  $P = [x_1 : x_2 : \dots : x_{n+1}] \in \mathbb{P}^n(\mathbb{F})$  a *point at infinity* if  $x_{n+1} = 0$ . *Affine points* in  $\mathbb{P}^n(\mathbb{F})$  are those points not at infinity. Each affine point in  $\mathbb{P}^n(\mathbb{F})$  can be uniquely represented as  $[x_1 : x_2 : \dots : x_n : 1]$ , giving a natural identification with the points of  $\mathbb{F}^n$ . The *projective line* over  $\mathbb{F}$  is  $\mathbb{P}^1(\mathbb{F})$  and the *projective plane* over  $\mathbb{F}$  is  $\mathbb{P}^2(\mathbb{F})$ . Let  $X_1, X_2, \dots, X_n$  be indeterminates. A polynomial  $f$  in  $\mathbb{F}[X_1, X_2, \dots, X_n]$  is *homogeneous of degree*  $d$  if every term of  $f$  is of degree  $d$ . A polynomial can be made homogeneous by introducing a new variable  $X_{n+1}$  and multiplying each term by an appropriate power of  $X_{n+1}$  to make all terms have degree  $d$ . We denote by  $f^H$  the homogenized polynomial of  $f$ . This process gives a one-to-one correspondence between polynomials in  $n$  variables of degree at most  $d$  and homogeneous polynomials of degree  $d$  in  $n + 1$  variables. The  $\mathbb{F}$ -points of an *affine plane curve*  $\mathcal{X}_f$  defined by  $f(x, y) \in \mathbb{F}[X, Y]$  a non-constant polynomial is

$$\mathcal{X}_f(\mathbb{F}) = \{(x, y) \in \mathbb{F}^2 : f(x, y) = 0\}.$$

The associated function field  $\mathbb{F}(\mathcal{X}_f)$  is the field of rational functions  $\mathbb{F}(x, y)$ , where the relation between  $x$  and  $y$  is given by  $f(x, y) = 0$ . A *projective plane curve*  $\mathcal{X}_f$  is the set of projective points  $[x : y : z] \in \mathbb{P}^2(\mathbb{F})$  such that  $f(x, y, z) = 0$ , where  $f(x, y, z)$  is a non-constant homogeneous polynomial. If a projective curve is defined by a polynomial  $f^H$ , it is called the *projective closure* of  $\mathcal{X}_f(\mathbb{F})$ . Curves defined over  $\mathbb{F}$  can also have points over an extension field  $\mathbb{E}$ , in which case the notation  $\mathcal{X}_f(\mathbb{E})$  is used. The partial derivative of a function is defined as in calculus, and a point  $(y_1, y_2, \dots, y_n)$  on an affine (resp., projective) curve  $\mathcal{X}_f(\mathbb{F})$  is called *singular* if every partial derivative of  $f$  equals 0 at the point. A curve that has no singular points is called *smooth*.

To define an evaluation code with evaluation points coming from a curve, we require a vector space of functions that can be evaluated at the points. Let  $p$  be a homogeneous polynomial of positive degree that defines a projective plane curve  $\mathcal{X}_p$  over  $\mathbb{F}$ . Define the *field of rational functions on*  $\mathcal{X}_p$  over  $\mathbb{F}$ , denoted by  $\mathbb{F}(\mathcal{X}_p)$ , as the collection of equivalence classes of rational functions  $\frac{g}{h}$ , where  $g$  and  $h$  are homogeneous of equal degree, with  $p \nmid h$ , and where  $\frac{g}{h} \sim \frac{g+ap}{h+bp}$  for any  $a, b$  homogeneous polynomials of appropriate degree.

In [2], the authors construct locally recoverable codes with availability  $t = 2$  based on fiber products of curves and propose a group-theoretic perspective on the construction, whereby a curve can sometimes be expressed as a fiber product of its quotient curves by certain subgroups of the automorphism group of the curve; see [3] for an extended version.

In [18], the authors give a closely related construction of codes with availability  $t$  recovery sets for any  $t \geq 2$  based on fiber products of  $t$  curves, including a different method of designing the LRC( $t$ )s, bounding the parameters for these codes, and constructing examples based on the generalized Giulietti–Korchmaros curves, the Suzuki curves, and fiber products of Artin-Schreier curves proposed by van der Geer and van der Vlugt. For an additional reference on codes from fiber products of curves, we recommend [7]. In the next subsection, we review these constructions in the context of the framework above in preparation for taking a hierarchical perspective.

## 2.2 Reed–Muller and fiber product codes

In this section, we introduce Reed–Muller codes and codes from fiber products of curves, then describe a framework that unites these codes as locally recoverable codes with availability. In the next section, we will see that it unifies the hierarchical recovery structure of these two families as well.

Recall that Reed–Muller codes are evaluation codes, defined as follows.

**Definition 3** The ( $q$ -ary) **Reed–Muller code** is a linear code over  $\mathbb{F}_q$  formed by evaluations of polynomials in  $\mathbb{F}_q[x_1, \dots, x_m]$  of total degree at most  $v$  at points in  $\mathbb{F}_q^m$ , and is denoted by  $\mathcal{RM}_q(v, m)$ ; that is,

$$\mathcal{RM}_q(v, m) = \{(f(P_1), \dots, f(P_n)) : f \in \mathbb{F}_q[x_1, \dots, x_m]_{\leq v}\}$$

where  $n = q^m$  and  $\mathbb{F}_q^m = \{P_1, \dots, P_n\}$ .

In thinking about the geometry of  $\mathbb{F}_q^m$ , we will make use of linear subspaces and their translations. The following definition makes these objects precise.

**Definition 4** Viewing  $V = \mathbb{F}_q^m$  as a vector space over  $\mathbb{F}_q$ , define an **affine subspace**  $A$  of  $V$  to be a coset of a subspace  $S$  of  $V$ ,

$$A = \vec{v} + S.$$

When  $S$  has dimension  $r$ , the set  $A$  is also called an  **$r$ -flat**.

Local properties of Reed–Muller codes were considered by Yekhanin in [29] for the purpose of local decodability. If  $v < q - 1$ , the code  $\mathcal{RM}_q(v, m)$  has local recovery with  $r = q - 1$  and availability  $\frac{q^m - 1}{q - 1}$ . To see this, note that the value of any polynomial of total degree at most  $v$  at a point  $P_i$  in  $\mathbb{F}_q^m$  can be determined by the value of that polynomial on the other  $q - 1$  points on any of the  $\frac{q^m - 1}{q - 1}$  distinct  $\mathbb{F}_q$ -rational lines through  $P_i$ . It is important to note that because two lines intersect in at most one point, and in this case that point is  $P_i$ , these lines with  $P_i$  removed satisfy the disjoint repair group property. In this way, we see the inherent geometric structure giving rise to LRCs with availability.

Now, we turn our attention to another code design in which the underlying geometry gives rise to availability. In what follows, we build on Sect. 2.1, but do use some more general and technical definitions. These objects are defined carefully in the references [19, 20, 25, 26]. However, we will attempt to provide some intuition on the more general and technical terms as we proceed so that the main ideas may be clearer to the reader. Further, the codes considered here are defined in full generality in [18] and discussed at length in [7], where space is devoted to developing a more concrete view.

Intuitively, an algebraic curve is the one-dimensional vanishing set of some polynomials defined over a given field. The curve is smooth if it does not have any self-crossings or cusps, i.e. it has a one-dimensional tangent space at each point. In this context, projective means that we include any “points at infinity” on each curve, which can be made algebraically precise by homogenizing the defining equations for the curve and looking for solutions in projective space. A rational map of curves is a map defined by rational functions. Let  $\mathcal{Y}, \mathcal{Y}_1, \mathcal{Y}_2, \dots, \mathcal{Y}_t$  be smooth, projective algebraic curves over  $\mathbb{F}_q$  each with a rational, separable map  $h_j : \mathcal{Y}_j \rightarrow \mathcal{Y}$  of degree  $d_{h_j}$  for each  $j \in \{1, 2, \dots, t\}$ . We note that separability is a more technical concept, but that a map is certainly separable if the characteristic of  $\mathbb{F}_q$  does not divide the degree of the map.

Each projective algebraic curve  $\mathcal{X}$  has an associated field of rational functions defined over  $\mathbb{F}_q$ , denoted  $\mathbb{F}_q(\mathcal{X})$ . We assume two further technical conditions on our curves, namely that  $\mathbb{F}_q$  is the full field of constants within each of the associated function fields  $\mathbb{F}_q(\mathcal{Y}_i)$ , and that the extensions  $\mathbb{F}_q(\mathcal{Y}_i)/\mathbb{F}_q(\mathcal{Y})$  are linearly disjoint (see [7] for more on this). Let  $\mathcal{X}$  be the fiber product  $\mathcal{Y}_1 \times_{\mathcal{Y}} \mathcal{Y}_2 \times_{\mathcal{Y}} \dots \times_{\mathcal{Y}} \mathcal{Y}_t$ . For a full discussion of fiber products, see [15]. Intuitively, the points of  $\mathcal{X}$  can be thought of as  $t$ -tuples of points, one from each of the curves  $\mathcal{Y}_i$ , where each point in the tuple maps to the same point on  $\mathcal{Y}$ . That is, points of  $\mathcal{X}$  are of the form  $(P_1, P_2, \dots, P_t)$  where  $h_1(P_1) = h_2(P_2) = \dots = h_t(P_t)$ . Though it is not immediately obvious, this construction results in a curve  $\mathcal{X}$ . Let the natural projection map  $g_j : \mathcal{X} \rightarrow \mathcal{Y}_j$  have degree  $d_{g_j}$  for each  $j, 1 \leq j \leq t$ , and define the rational, separable map  $g = g_j \circ h_j : \mathcal{X} \rightarrow \mathcal{Y}$  (for any  $j$ ) of degree  $d_g$ . For each  $j$ , define the curve

$$\tilde{\mathcal{Y}}_j = \mathcal{Y}_1 \times_{\mathcal{Y}} \dots \times_{\mathcal{Y}} \mathcal{Y}_{j-1} \times_{\mathcal{Y}} \mathcal{Y}_{j+1} \times_{\mathcal{Y}} \dots \times_{\mathcal{Y}} \mathcal{Y}_t.$$

More generally (for later use) we define  $\tilde{\mathcal{Y}}_A$  for any set  $A \subseteq \{1, \dots, t\}$  to be the fiber product over  $\mathcal{Y}$  of all  $\mathcal{Y}_i$  with  $i \notin A$ . Then  $\mathcal{X} = \mathcal{Y}_j \times_{\mathcal{Y}} \tilde{\mathcal{Y}}_j$ . Denote the associated natural maps by

$$\tilde{g}_j : \mathcal{X} \rightarrow \tilde{\mathcal{Y}}_j \quad \text{and} \quad \tilde{h}_j : \tilde{\mathcal{Y}}_j \rightarrow \mathcal{Y}.$$

The degree of  $\tilde{g}_j$  must be equal to  $d_{h_j}$ . Intuitively, the map  $\tilde{g}_j$  “forgets” the information coming from the curve  $\mathcal{Y}_j$  while retaining the data of the fiber product that come from the other factors, meaning the curves  $\mathcal{Y}_i$  with  $i \neq j$ .

The function field  $\mathbb{F}_q(\mathcal{X})$  is isomorphic to the compositum of the function fields  $\mathbb{F}_q(\mathcal{Y}_i)$ , where the function field  $\mathbb{F}_q(\mathcal{Y})$  is embedded into each  $\mathbb{F}_q(\mathcal{Y}_i)$  as induced by the map  $h_i$ . For ease of exposition, we identify each function field with its image inside  $\mathbb{F}_q(\mathcal{X})$ , so for each  $i$ ,

$$\mathbb{F}_q(\mathcal{Y}) \subseteq \mathbb{F}_q(\mathcal{Y}_i) \subseteq \mathbb{F}_q(\mathcal{X}).$$

This framework sets the stage for a code design reminiscent of algebraic geometry codes, though harnessing the fiber product structure to facilitate local recovery with high (or bespoke) availability. Let  $y_j \in \mathbb{F}_q(\mathcal{Y}_j)$  be a primitive element of  $\mathbb{F}_q(\mathcal{Y}_j)/\mathbb{F}_q(\mathcal{Y})$  where  $y_j$  is the root of a degree  $d_{h_j}$  polynomial with coefficients in  $\mathbb{F}_q(\mathcal{Y})$ . This yields

$$\mathbb{F}_q(\mathcal{Y})(y_1, y_2, \dots, y_t) \cong \mathbb{F}_q(\mathcal{X}).$$

A divisor on a curve is a formal sum of points on the curve. Let  $D_j$  be the principal divisor of the function  $y_j \in \mathbb{F}_q(\mathcal{X})$ . Let  $D_j = D_{j,+} - D_{j,-}$ , where  $D_{j,+} = (y_j)_0$  is the zero divisor of  $y_j$ , and  $D_{j,-} = (y_j)_\infty$  is the pole divisor of  $y_j$ . Let  $d_{y_j}$  be the degree of  $y_j : \mathcal{Y}_j \rightarrow \mathbb{P}^1$ . Then, if  $y_j$  is viewed as a function  $y_j : \mathcal{X} \rightarrow \mathbb{P}^1$ , its degree is  $d_{g_j} d_{y_j}$ , which is equal to  $\deg(D_{j,-})$ . Now choose simultaneously a divisor  $D$  on  $\mathcal{Y}(\mathbb{F}_q)$  and a set  $S$  of points in  $\mathcal{Y}(\mathbb{F}_q)$  as follows. Let  $\tilde{D} = \sum_{j=1}^t g(D_{j,-})$ , so  $\text{Supp}(\tilde{D})$  consists of all points of  $\mathcal{Y}$  which, for some

$j$ , are the image under  $g$  of a point on  $\mathcal{X}$  at which the function  $y_j$  has a pole. Then choose  $D$  an effective divisor on  $\mathcal{Y}(\mathbb{F}_q)$  of degree  $\deg(D) = l$  and  $S = \{P_1, \dots, P_s\} \subseteq \mathcal{Y}(\mathbb{F}_q)$  so that the following conditions are satisfied:

- $|g^{-1}(P_i) \cap \mathcal{X}(\mathbb{F}_q)| = d_g$ , for all  $i \in [s]$ ,
- $S \cap \text{Supp}(\tilde{D}) = \emptyset$ ,
- $S \cap \text{Supp}(D) = \emptyset$ ,
- $l < s$ .

Let  $\mathcal{L}(\mathcal{Y}, D)$  be the Riemann–Roch space of the divisor  $D$  on the curve  $\mathcal{Y}$ . See [25] for a discussion of Riemann–Roch spaces and their properties. Let  $\ell(D)$  be the dimension of  $\mathcal{L}(\mathcal{Y}, D)$  as an  $\mathbb{F}_q$ -vector space. Let  $\{f_1, f_2, \dots, f_m\}$  be a basis for  $\mathcal{L}(\mathcal{Y}, D)$ , so  $m = \ell(D)$ . These functions are naturally in  $\mathbb{F}(\mathcal{Y})$ , and we also consider them to be functions in  $\mathbb{F}(\mathcal{X})$  through the natural inclusion. Note that since  $l < s$ , each non-zero function in  $\mathcal{L}(\mathcal{Y}, D)$  will be non-zero when evaluated at some point in  $S$ . Then set

$$B = g^{-1}(S)$$

so that  $n := |B| = d_g s$ . Order the points in  $B$  and denote them as  $\{Q_1, Q_2, \dots, Q_n\}$ . Let

$$V = \text{Span} \left\{ f_k y_1^{e_1} \cdots y_t^{e_t} : \begin{array}{l} e_i \in \mathbb{Z}, \\ 0 \leq e_j \leq d_{h_j} - \rho_j \forall j \in \{1, \dots, t\}, k \in [m] \end{array} \right\}.$$

Define  $C(D, B) := \text{Im}(ev_B)$  where

$$\begin{aligned} ev_B : V &\rightarrow \mathbb{F}_q^n \\ f &\mapsto (f(Q_i))_{i \in \{1, \dots, n\}}. \end{aligned}$$

The following is a restatement of results from [18], using the above terminology, in preparation for an investigation of hierarchy.

**Theorem 2.1** [18] *Let curves  $\{\mathcal{Y}_j\}_{j \in \{1, \dots, t\}}$ ,  $\mathcal{Y}$ , maps  $\{h_j : \mathcal{Y}_j \rightarrow \mathcal{Y}\}_{j \in \{1, \dots, t\}}$ , a divisor  $D$  on  $\mathcal{Y}(\mathbb{F}_q)$ , and sets  $S \subseteq \mathcal{Y}(\mathbb{F}_q)$  and  $B = g^{-1}(S)$  be all as described above, where  $l = \deg(D) \leq |S|$  and the quantity  $d$  below is positive. For  $i \in \{1, 2, \dots, n\}$  and  $j \in \{1, 2, \dots, t\}$ , set  $C_{i,j}$  to be the punctured code within  $C(D, B)$  with support  $\{k : Q_k \in \tilde{g}_j^{-1}(\tilde{g}_j(Q_i))\}$ . Then the code  $C(D, B)$  is an LRC( $t$ ) with*

- length  $n = |B| = d_g |S|$ ,
- dimension  $k = \ell(D)(d_{h_1} - \rho_1 + 1)(d_{h_2} - \rho_2 + 1) \cdots (d_{h_t} - \rho_t + 1)$
- minimum distance  $d \geq n - ld_g - \sum_{j=1}^t (d_{h_j} - \rho_j)(d_g d_{y_j})$ , and
- locality  $r_j = d_{h_j} - \rho_j + 1$  for  $1 \leq j \leq t$ .

Next, we provide a unified framework that will aid the investigation of hierarchical locally recoverable codes from Reed–Muller and fiber product codes.

### 2.3 Relating fiber product codes and Reed–Muller codes

Locally recoverable codes with many recovery sets arising from fiber product constructions were presented in great generality in [18]. Though this generality is valuable, many examples of these codes can be understood much more simply. In these cases, the availability and hierarchical recovery scheme that will be outlined is essentially the same as the natural geometric availability and hierarchical recovery scheme that we will present for generalized Reed–Muller codes. We will see that this perspective orients fiber product codes in the larger



landscape of techniques to increase the rate of Reed–Muller codes and punctured Reed–Muller codes while still retaining availability.

Let all notation be as in Theorem 2.1. Assume that

$$\mathcal{L}(\mathcal{Y}, D) \subseteq \mathbb{F}_q[z_1, z_2, \dots, z_\mu]_{\leq \psi}$$

so that the maximum total degree of all functions in  $\mathcal{L}(\mathcal{Y}, D)$  is at most  $\psi$ . Then take

$$V \subseteq \left\langle z_1^{a_1} \cdots z_\mu^{a_\mu} y_1^{e_1} \cdots y_t^{e_t} : \begin{array}{l} a_i, e_i \in \mathbb{Z}, 0 \leq e_j \leq d_{h_j} - \rho_j \forall j \in \{1, \dots, t\}, \\ a_1 + \cdots + a_\mu \leq \psi \end{array} \right\rangle.$$

The curve  $\mathcal{X}$  abstractly exists in a space which is the product of many projective spaces. Consider a collection  $S$  of affine  $\mathbb{F}_q$ -rational points on  $\mathcal{Y}$ , meaning  $S \subseteq \mathcal{Y}(\mathbb{F}_q)$  may be thought of as

$$S \subseteq \mathbb{F}_q^\mu.$$

Take  $B = g^{-1}(S) \subseteq \mathcal{X}(\mathbb{F}_q)$  which we identify with points in  $\mathbb{F}_q^{\mu+t}$  so that the evaluation points for the code  $C(B, D)$  satisfy

$$B \subseteq \mathbb{F}_q^{\mu+t};$$

more precisely, the function

$$\phi : \mathcal{X} \rightarrow \mathbb{F}_q^{\mu+t}$$

gives

$$(z_1, z_2, \dots, z_\mu, y_1, y_2, \dots, y_t) \mapsto \mathbb{F}_q^{\mu+t}.$$

The codewords of  $C(B, D)$  are then the evaluations of polynomials in  $\mathbb{F}_q[z_1, z_2, \dots, z_\mu, y_1, y_2, \dots, y_t]$  on points  $S$  of the fiber product curve  $\mathcal{X}$ . Thus the fiber product code  $C(D, B)$  is a subcode of the Reed–Muller code  $\mathcal{RM}_q(v, \mu + t)$  for some  $v$ , punctured to the set  $\phi(g^{-1}(S))$  where  $v \leq \psi + \sum_{j=1}^t (d_{h_j} - \rho_j)$ :

$$C(B, D) \subseteq \mathcal{RM}_q(v, \mu + t) |_{\phi(g^{-1}(S))}.$$

When viewed from this perspective, the  $t$  recovery sets of a position in a  $t$ -fold fiber product code correspond to the intersections of lines parallel to the  $y_1, \dots, y_t$  coordinate axes with the curve. In a Reed–Muller code, each line through a point gives a recovery set for the position of that point; in the fiber product case we limit ourselves to particular lines yielding known-cardinality intersections with the evaluation set. By doing this, we are able to use functions of larger total degree than would be possible when using a standard Reed–Muller code. In particular, we are able to let each  $y_k$  have degree up to  $d_{h_k} - 2$ , without any limit on the total degree of the function. Consequently, the more general fiber product codes yield larger rates than would be possible for the punctured Reed–Muller code.

This approach is one of several ways to define codes with larger rates. One way of increasing rate is lifting, pioneered by Guo et al. in [17]. Their lifted Reed–Solomon codes make use of the observation that there are generally monomials of total degree  $\geq q - 1$  in  $\mathbb{F}_q[x_1, \dots, x_m]$  that reduce to degree  $< q - 1$  univariate polynomials on every line in  $\mathbb{F}_q^m$ . Adding these monomials to a lower degree Reed–Muller code greatly increases the rate without losing the very high availability. In [11], the authors create partially lifted codes, which increase the rate further by limiting the degree condition to a subset of lines as well as including non-monomial functions that meet the degree condition on lines. In [21], the authors consider lifted codes

where the evaluation points are  $\mathbb{F}_{q^2}$ -points of the Hermitian curve  $\mathcal{H}_q$  in the plane. All non-tangent lines with  $q + 1$  points of intersection with the affine patch of  $\mathcal{H}_q$  yield recovery sets for this code, and the degree allowed for the univariate polynomials on lines is restricted by the size of the intersection with the curve.

In fiber product codes, we use a subset of curve points as the evaluation sets, but the lines corresponding to recovery sets have been severely limited to those parallel to many coordinate axes. The degree allowed for the restricted polynomials on these lines is again limited by the cardinality of the intersections of the lines with the curve, which in the case of the fiber product construction is the degree of the map  $h_k$  for some  $1 \leq k \leq t$ . This allows for many more monomials than bounding total degree as is the case in the Reed–Muller code.

In the next sections, we demonstrate how these constructions and modifications naturally give rise to hierarchical local recovery with availability.

### 3 Two-level hierarchical recovery of Reed–Muller and generalized fiber product constructions

In the previous section, we observed how lines and fiber products of curves can provide many disjoint recovery sets for a codeword coordinate. Another concept for addressing the availability problem has been advanced in the form of hierarchical locality [1, 10, 12, 16, 23]. The idea is that a position and its recovery set could form a locally recoverable code with smaller locality and local minimum distance, offering two nested recovery sets. The larger recovery set would be used if the smaller set was insufficient to recover the given erasures. This notion captured in the next definition.

**Definition 5** [23] Let  $n, n_1, n_2, s_1, s_2, \delta_1, \delta_2 \in \mathbb{Z}^+$  with  $n_2 < n_1$ ,  $s_2 \leq s_1$ , and  $\delta_2 < \delta_1$ . A linear code  $C$  of length  $n$  is said to have **hierarchical locality** with parameters  $((n_1, s_1, \delta_1), (n_2, s_2, \delta_2))$  if for each  $i \in \{1, \dots, n\}$  there exists a punctured code  $C_i$  of length  $n_1$ ,  $\dim(C_i) \leq s_1$ , and minimum distance  $d(C_i) \geq \delta_1$  with  $i$  in the support of  $C_i$  such that  $C_i$  is an  $(s_2, \delta_2)$ -locally recoverable code, where each local repair group has size  $n_2$ .

Note that up to  $\delta_2 - 1$  erasures among the support of any local repair group within  $C_i$  can be locally recovered using the local recovery process for  $C_i$ . Further erasures up to  $\delta_1 - 1$  total can be recovered using all the coordinates of  $C_i$ . Hence, there are two levels of hierarchy present:

1. one level using at most  $n_2 - \delta_2 + 1$  symbols from a recovery set for the symbol within  $C_i$  if there are fewer than  $\delta_2$  erasures and
2. another larger one using at most  $n_1 - \delta_1 + 1$  symbols of  $C_i$  if there are between  $\delta_2$  and  $\delta_1 - 1$  erasures.

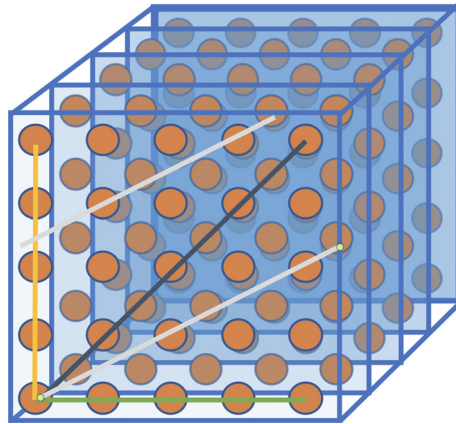
#### 3.1 Hierarchical recovery of Reed–Muller codes

The geometric definitions of Reed–Muller codes give them built-in nested structure in the form of intersections of the evaluation set with lines, planes, and hyperplanes within the ambient space, as depicted in Fig. 2.

We see how to use this structure for basic (2-level) hierarchical recovery in both settings.

**Theorem 3.1** *The Reed–Muller code  $\mathcal{RM}_q(v, m)$  has hierarchical locality with parameters*

$$\left( \left( q^2, \binom{v+2}{2}, q^2 - qv \right), (q, v+1, q-v) \right).$$



**Fig. 2** This cartoon illustrates how some hierarchical repair groups naturally arise in the evaluation points of  $RM_5(v, 3)$  for  $v \leq 3$ . The positions corresponding to points of any  $\mathbb{F}_5$ -rational line passing through the given point can act as a repair group for local recovery, with a middle code arising from any plane containing the line. Here we see one way to partition the space into planes, the points of which correspond to the positions in the support of middle codes. For each point in these planes, there are six lines in the plane passing through the point (four are illustrated here), any of which can be chosen to give the positions in the support of the lower code for the chosen position

**Proof** Recall that the code  $C = \mathcal{RM}_q(v, m)$ , where  $v \leq q - 1$ , has the following parameters:

$$n = q^m, k = \binom{v + m}{m}, d = (q - v)q^{m-1}.$$

Consider the H-LRC formed by taking planes and lines, respectively, through a given position. We have that any plane containing  $P_i$  can be the support for a middle code  $C_i$ , with local recovery inside the plane given by any partition of the plane into parallel lines. If  $v \leq q - \delta_2$ , then  $C_i$  is a  $(q^2, q - \delta_2 + 1, \delta_2)$ -locally recoverable code, and itself has dimension  $\dim(C_i) \leq \binom{q - \delta_2 + 2}{2}$  and minimum distance  $d(C_i) \geq q\delta_2$ .

Put in terms of  $v$ , the code  $C = \mathcal{RM}_q(v, m)$  has hierarchical locality with parameters  $((n_1, s_1, \delta_1), (n_2, s_2, \delta_2))$ , where

$$n_1 = q^2, s_1 = \binom{v + 2}{2}, \delta_1 = q^2 - qv$$

$$n_2 = q, s_2 = v + 1, \delta_2 = q - v.$$

Indeed, the middle code for a fixed position  $c_i = (x_1, \dots, x_m)_i$  consists of evaluations of all points in a plane containing the point  $c_i$ . Therefore  $s_1 \leq q^2$ , the number of points on a plane. The evaluations of the points on a plane form an LRC with parameters  $(v + 1, q - v)$ , since there is a restriction to evaluations of a line contained in the plane. The middle code is isomorphic to  $\mathcal{RM}_q(2, q - 2)$ , and the bottom code is isomorphic to  $\mathcal{RM}(1, q - 2)$ .  $\square$

**Example 1** (Two-level hierarchical recovery using a Reed–Muller code) Consider  $\mathcal{RM}_7(5, 3)$ , with  $H = 2$ . Then  $C_2 \subseteq C_1 \subseteq C$ , where the parameters of  $C$  are [343, 56, 98], and the H-LRC parameters are [49, 21, 14], and [7, 6, 2], respectively.

**Remark 1** Variations of the above two-level hierarchy construction above can be obtained using a code  $\mathcal{RM}(v, m)$  where the supports used for  $n_1, n_2$  are defined by nested  $\ell_1$ - and

$\ell_2$ -flats in the geometry, with  $m - 1 > \ell_1 > \ell_2 > 1$ . The hierarchical parameters are

$$\left( \left( q^{\ell_1}, \binom{v + \ell_1}{\ell_1} \right), (q - v)q^{\ell_1 - 1} \right), \left( q^{\ell_2}, \binom{v + \ell_2}{\ell_2} \right), (q - v)q^{\ell_2 - 1} \right)$$

giving added flexibility in the construction.

### 3.2 Hierarchical recovery of fiber product codes

In this subsection, we return to fiber product codes, examining the family’s ability to support local recovery. In the fiber product construction of [18], we see that the  $j$ -th recovery set for position  $i$  consists of the positions corresponding to points on the curve that all share the same images under the function  $g$  and each  $y_k$  except  $k = j$ . To provide the appropriate setting to discuss hierarchical recovery, consider a fiber product code  $C(D, B)$  and define  $C_{i,j}$  to be the punctured code of  $C(D, B)$  with support  $\{k : Q_k \in \tilde{g}_j^{-1}(\tilde{g}_j(Q_i))\}$ . In other words, the  $i$ -th coordinate corresponds to a point  $Q_i$  on the fiber product curve. The  $j$ -th recovery set corresponds to a factor curve  $\mathcal{Y}_j$  of the fiber product, and element  $y_j$  in the function field  $\mathbb{F}_q(\mathcal{X})$ . The support of  $C_{i,j}$  will contain all coordinates corresponding to points  $Q$  such that  $y_k(Q) = y_k(Q_i)$  for all  $k \neq j$  and  $g(Q) = g(Q_i)$ . Local recovery of coordinate  $i$  with  $C_{i,j}$  is accomplished by observing that on the points with indices in the support of  $C_{i,j}$ , any function in  $V$  restricts to a polynomial of degree at most  $d_{h_j} - \rho_j$  in  $y_j$ , which can be interpolated at any missing points given the value of the polynomial on at least  $d_{h_j} - \rho_j + 1$  points. Thus each  $C_{i,j}$  is a subcode of a punctured Reed–Solomon code. Any  $\rho_j - 1$  erasures within the support of  $C_{i,j}$  are recoverable using  $C_{i,j}$ . We formalize the  $C(D, B)$  as a two-level H-LRC and record its parameters in the following theorem.

**Theorem 3.2** *Let  $t$  be a positive integer with  $t \geq 2$ . Let  $C = C(D, B)$  be an LRC( $t$ ) constructed from a fiber product of curves as in Theorem 2.1. Choose any  $j, k \in \{1, 2, \dots, t\}$  with  $j \neq k$ . Then  $C$  has hierarchical locality with  $n_1 = d_{h_j}d_{h_k}$ ,  $n_2 = d_{h_j}$ ,  $s_1 = (d_{h_j} - \rho_j + 1)(d_{h_k} - \rho_k + 1)$ ,  $s_2 = d_{h_j} - \rho_j + 1$ ,  $\delta_1 = \rho_j\rho_k$  and  $\delta_2 = \rho_j$ .*

**Proof** We define the code  $C_{i,(j,k)}$  to be the punctured code from  $C(D, B)$  with support  $\{l : Q_l \in \tilde{g}_k^{-1}(\tilde{g}_k(\tilde{g}_j^{-1}(\tilde{g}_j(Q_i))))\}$ . This consists of positions corresponding to all points  $Q$  in  $B$  so that for each  $l \in \{1, 2, \dots, t\} \setminus \{j, k\}$ ,  $y_l(Q) = y_l(Q_i)$ , and  $g(Q) = g(Q_i)$ . We now have that on the set of points corresponding to the positions in the support of  $C_{i,(j,k)}$ , any function in  $V$  restricts to a polynomial in  $y_j$  and  $y_k$  with bounded degree in each coordinate. Thus  $C_{i,(j,k)}$  is essentially a subcode of a punctured Reed–Muller code on two variables. We will take the middle code  $C_{l_i} = C_{i,(j,k)}$  for each  $i \in \{1, 2, \dots, n\}$ . We observe that

$$\text{Supp}(C_{i,(j,k)}) = \bigcup_{l \in \text{Supp}(C_{i,j})} \text{Supp}(C_{l,k}) = \bigcup_{l \in \text{Supp}(C_{i,k})} \text{Supp}(C_{l,j}).$$

Further,

$$\text{Supp}(C_{l_1,k}) \cap \text{Supp}(C_{l_2,k}) = \emptyset$$

for  $l_1 \neq l_2, l_1, l_2 \in \text{Supp}(C_{i,j})$ . This is because if  $l_1$  and  $l_2$  are both in  $\text{Supp}(C_{i,j})$ , then for the corresponding points  $Q_{l_1}$  and  $Q_{l_2}$  in the evaluation set  $B$ , we have  $y_j(Q_{l_1}) \neq y_j(Q_{l_2})$ . If  $Q$  is a point with index in the intersection  $\text{Supp}(C_{l_1,k}) \cap \text{Supp}(C_{l_2,k})$ , then  $y_j(Q_{l_1}) = y_j(Q) = y_j(Q_{l_2})$ , a contradiction.

Thus every coordinate in  $\text{Supp}(C_{i,(j,k)})$  has two disjoint recovery sets in  $C_{i,(j,k)}$ , namely  $C_{i,j}$  and  $C_{i,k}$ , a stronger property than is required for hierarchical recovery. However, for the

purpose of understanding the hierarchical recovery process, we will consider local recovery through the  $j$ -th recovery set corresponding to  $C_{i,j}$  to determine the parameters. This yields  $s_2 = r_j = \deg(h_j) - \rho_j + 1$  and  $\delta_2 = \rho_j$ , where  $r_j$  and  $\rho_j$  are the local recovery parameters and  $d_{h_j}$  is the degree of the covering map  $h_j$  from the construction in Theorem 2.1.

The length of  $C_{i,(j,k)}$  is  $d_{h_j}d_{h_k}$ . The dimension of  $C_{i,(j,k)}$  is at most  $s_1 = (d_{h_j} - \rho_j + 1)(d_{h_k} - \rho_k + 1)$  based on the maximum degrees in  $y_j$  and  $y_k$  of functions in  $V$ . The minimum distance of  $C_{i,(j,k)}$  can be easily bounded below as follows. Suppose that there are fewer than  $\rho_j\rho_k$  erasures in a codeword of  $C_{i,(j,k)}$ . Then, there must be some  $i'$  with fewer than  $\rho_k$  erasures in the indices corresponding to  $C_{i',k}$  or fewer than  $\rho_j$  erasures in the indices corresponding to  $C_{i',j}$ . Thus the value of the  $i'$ -th position can be locally recovered by one of  $C_{i',j}$  or  $C_{i',k}$ . By repeated application, this implies that all erasures can be recovered, so the minimum distance of  $C_{i,(j,k)}$  must be at least  $\rho_j\rho_k$ . □

**Remark 2** In [1], the authors construct codes with hierarchical locality by a natural construction using towers of curves  $\mathcal{X} \rightarrow \mathcal{Y} \rightarrow \mathcal{Z}$ . Our hierarchical codes can be viewed in this light, where the tower we construct is  $\mathcal{X} \rightarrow \tilde{Y}_j \rightarrow \tilde{Y}_{\{j,k\}}$ . The construction in [1, Proposition IV.1] requires  $\delta_2 = 2$ , while the fiber product construction here allows for a choice of  $\delta_2 = \rho_j \geq 2$ . Due to the particularly nice arrangement of recovery groups in this construction, the method of recovery using the middle code that we describe here gives a different lower bound  $\delta_1$  for the minimum distance of the middle code.

### 3.2.1 Hierarchical recovery of LRC(t) from a fiber product of Artin–Schreier curves

Let  $p$  be a prime,  $h, t \in \mathbb{Z}^+$  with  $t \leq h$ , and  $q = p^h$ . In [6, 18], the authors apply the LRC( $t$ ) construction to create codes defined over  $\mathbb{F}_{q^2}$  on  $\mathcal{A}_{q,t}$ , a fiber product of  $t$  Artin–Schreier curves studied in [27]. When  $t = h$ ,  $\mathcal{A}_{q,t}$  is isomorphic to  $\mathcal{H}_q$ . We let  $A = \{a \in \mathbb{F}_{q^2} : a^q + a = 0\}$ , a  $\mathbb{F}_p$ -linear space generated over  $\mathbb{F}_p$  by  $\{a_1, a_2, \dots, a_n\}$ . Let  $\mathcal{Y}$  be the projective line in coordinate  $y_0$ . For  $1 \leq i \leq t$ , let  $\mathcal{Y}_i$  be defined by  $y_i - y_i^p = a_i y_0^{q+1}$ . Let  $h_i : \mathcal{Y}_i \rightarrow \mathcal{Y}$  be the map given by projection onto the  $y_0$  coordinate. We may then define  $\mathcal{X} = \mathcal{A}_{q,t}$  to be the fiber product of these curves  $\mathcal{Y}_i$  over  $\mathcal{Y}$ ; i.e.,

$$\mathcal{A}_{q,t} = \mathcal{Y}_1 \times_{\mathcal{Y}} \mathcal{Y}_2 \times_{\mathcal{Y}} \dots \times_{\mathcal{Y}} \mathcal{Y}_t.$$

As described in [6], we may identify  $\mathcal{A}_{q,t}$  with its image in  $\mathbb{P}^{t+1}$ , where the affine points of  $\mathcal{A}_{q,t}$  are given by

$$B = \{(y_0, y_1, y_2, \dots, y_t) \in (\mathbb{F}_{q^2})^{t+1} : y_i^p - y_i = a_i y_0^{q+1} \text{ for all } 1 \leq i \leq t\}. \tag{1}$$

Let  $B$  be the set of  $p^t q^2$  affine points in  $\mathcal{X}(\mathbb{F}_{q^2})$ . Let

$$P_i = (\alpha, \beta_1, \beta_2, \dots, \beta_t) \in B.$$

Then  $B_{i,j}$ , the  $j$ -th recovery set for the position corresponding to  $P_i$ , is the set of positions corresponding to the points in  $\{(\alpha, y_1, y_2, \dots, y_t) \in B : y_k = \beta_k \forall k \neq j\}$ . We then have  $|B_{i,j}| = p$ . On points corresponding to the positions in  $B_{i,j}$ , any function in  $V$  varies as a polynomial in  $y_j$  of degree at most  $(p - 2)$  and can therefore be interpolated by knowing its values on any  $p - 1$  points.

Given  $h, t$  as above, choose  $l \leq q^2 - \frac{t(p-2)(q+1)p^{t-1}+1}{p^t}$ ; this will ensure an appropriate value of  $d$ . Let  $P_\infty$  denote the unique point at infinity on  $\mathcal{Y}$ , and let  $D = lP_\infty$ . Then  $\mathcal{L}(D)$  is the set of polynomials in  $y_0$  of degree at most  $l$ , a vector space of dimension  $\ell(D) = l + 1$ .

Applying the hierarchical recovery construction in Theorem 3.2 to the codes constructed in Theorem 3.3, we immediately obtain an H-LRCs over  $\mathbb{F}_{q^2}$ .

**Theorem 3.3** Consider  $C_{\mathcal{A}_{q,t,l}} := C(V, B)$  where  $\mathcal{X} = \mathcal{A}_{q,t}$  is the fiber product of the specified Artin–Schreier curves, with  $B$  and  $l$  as above and  $D = l\infty_{\mathcal{Y}}$ , and  $V$  as in Theorem 2.1. Then the  $[p^t q^2, (l+1)(p-1)^t, n-lp^t-t(p-2)(q+1)p^{t-1}]$  locally recoverable code  $C_{\mathcal{A}_{q,t,l}}$  over  $\mathbb{F}_{q^2}$  with availability  $t$  and locality  $(p-1, p-1, \dots, p-1)$  is a 2-level H-LRC with hierarchical parameters  $n_2 = p$ ,  $n_1 = p^2$ ,  $s_2 = p-1$ ,  $s_1 = (p-1)^2$ ,  $\delta_2 = 2$ ,  $\delta_1 = 4$ .

**Proof** The fact that  $C_{\mathcal{A}_{q,t,l}}$  is an LRC( $t$ ) with the parameter shown follows from [6, 18]. To verify that it is an H-LRC as stated, set  $j, k \in \{1, \dots, t\}$ ,  $j \neq k$ . Since the  $t$  factor curves are all isomorphic, with degree  $p$  maps to the base curve, we have  $d_{h_j} = d_{h_k} = p$ ,  $r_j = r_k = p-1$ , and  $\rho_j = \rho_k = 2$  for any choice of  $j, k$ . Directly applying Theorem 3.2 gives the required hierarchical parameters.  $\square$

### 3.2.2 Hierarchical recovery of LRC(2) from the Hermitian curve as a fiber product

Codes on the Hermitian curve have been extensively studied, including in the context of locally recoverable codes with availability. Let  $\mathcal{H}_q$  be the Hermitian curve, i.e. the projective curve defined over  $\mathbb{F}_{q^2}$  by the affine equation  $x^q + x = y^{q+1}$ . The Hermitian curve can be constructed as a fiber product as follows. Let  $\mathcal{Y} = \mathbb{P}^1$ ,  $\mathcal{Y}_1: u = y^{q+1}$ , and  $\mathcal{Y}_2: u = x^q + x$ . Let  $h_j: \mathcal{Y}_j \rightarrow \mathbb{P}^1$  be projection onto  $u$  for  $j = 1, 2$ . Note that these maps have coprime degrees. Thus, the corresponding function field extensions are linearly disjoint. Then the fiber product  $\mathcal{X} = \mathcal{Y}_1 \times_{\mathcal{Y}} \mathcal{Y}_2$  is isomorphic to the curve  $\mathcal{H}_q: x^q + x = y^{q+1}$ . Indeed, the affine points of  $\mathcal{X}(\mathbb{F}_{q^2})$  are given by

$$\{(y, u), (x, u) : x, y, u \in \mathbb{F}_{q^2}, y^{q+1} = u = x^q + x\} \subseteq \mathbb{P}^2 \times \mathbb{P}^2.$$

Hence, this fiber product is isomorphic (by the natural map) to the intersection of the two hypersurfaces in  $\mathbb{P}^3$  with affine equations  $u = x^q + x$  and  $u = y^{q+1}$  and also to the curve  $\mathcal{H}_q$  defined in  $\mathbb{P}^2$  by affine equation  $y^{q+1} = x^q + x$ . Let  $C_{\mathcal{H}_q}$  be the LRC(2) presented in Proposition 5.1 of [2], as well as Theorem 7 of [6]. For this code, we take the curve  $\mathcal{H}_q$  with evaluation set  $B_{\mathcal{H}_q} = \{P \in \mathcal{H}_q(\mathbb{F}_{q^2}) : y(P) \neq 0\}$ . We can check that  $|B_{\mathcal{H}_q}| = q^3 - q$ . Let  $V_{\mathcal{H}_q}$  be the space of functions with basis  $\{x^{e_1}y^{e_2} : 0 \leq e_1 \leq q-2, 0 \leq e_2 \leq q-1\}$ . This choice of functions corresponds to choosing a zero divisor  $D = 0$  on the base curve  $\mathcal{Y}$  in the construction from Theorem 2.1, yielding  $\ell(D) = 1$ , and  $\rho_1 = \rho_2 = 2$ . Let  $C_{\mathcal{H}_q} = C(V_{\mathcal{H}_q}, B_{\mathcal{H}_q})$ .

**Theorem 3.4** [2, 6] The  $[q^3 - q, q^2 - q, q^3 - 2q^2 + q + 2]$ ,  $((q-2, 2), (q-1, 2))$  LRC(2)  $C_{\mathcal{H}_q}$  is a 2-level H-LRC with  $n_1 = q(q+1)$ ,  $s_1 = q(q-1)$ ,  $\delta_1 = 4$ ,  $n_2 = q+1$ ,  $s_2 = q-1$ ,  $\delta_2 = 2$ .

**Proof** The supports of the two punctured codes giving recovery sets for the position corresponding to a point  $P \in B_{\mathcal{H}_q}$  consist of the positions corresponding to points  $Q \in B_{\mathcal{H}_q}$ ,  $Q \neq P$  sharing the same  $x$ -coordinate as  $P$  and those sharing the same  $y$ -coordinate value as  $P$ , respectively. Concretely, if  $P_i = (x_0, y_0)$  and  $\alpha_i \in \mathbb{F}_{q^2}$  such that  $x_0^q + x = \alpha_i = y_0^{q+1}$ , then  $C_{i,1}$  has support  $\{j : P_j = (x_0, y) \in B_{\mathcal{H}_q}, y^{q+1} = \alpha_i\}$  and  $C_{i,2}$  has support  $\{j : P_j = (x, y_0) \in B_{\mathcal{H}_q}, \alpha_i = x^q + x\}$ .

Directly applying the construction in the proof of Theorem 3.2 to the fiber product code  $C_{\mathcal{H}_q}$ , we obtain an H-LRC with the given hierarchical parameters. Define  $C_{1i} = C_{i,(2,1)}$  to

be the code with support  $\{j : P_j = (x, y) \in B_{\mathcal{H}_q}, y^{q+1} = \alpha_i = x^q + x\}$ . The length of  $C_{1_i}$  is  $q(q + 1)$ ; the dimension of  $C_{1_i}$  is at most  $q(q - 1) = s_1$ ; and the minimum distance of  $C_{1_i}$  is at least  $4 = \delta_1$ . The parameters of the local recovery sets are  $s_2 = r_2 = q - 2$  and  $\delta_2 = \rho_2 = 2$ . □

**Remark 3** 1. In [1, Example VII.3], the authors create an H-LRC from the Hermitian curve.

We obtain a different H-LRC by our fiber product-based construction.

2. Notice that by simply by considering the other local recovery set, we also have an H-LRC with the less favorable  $s_2 = r_1 = q - 2$ . We have chosen the better parameter in the theorem.
3. In comparing to the standard hierarchy bound of Sasidharan et al. [23], Theorem 3.4 shows a gap between the bound and the code parameters because the code also has natural availability described in Sect. 5 that is not accounted for in the bound.

We observe that the dimension of the middle code  $C_i$  is equal to the dimension of  $C_{\mathcal{H}_q}$ . This will occur in the general construction when  $t = 2$  and  $\ell(D) = 1$ . It is desirable to recover the required erasures by accessing fewer symbols in the middle code than we would have to access in the full code. Given that the length  $n_1$  of  $C_{1_i}$  is smaller than the length of  $C_{\mathcal{H}_q}$ , giving a higher rate for  $C_{1_i}$ , one might wonder whether  $C_{\mathcal{H}_q}$  offers any advantages at all. However,  $C_{\mathcal{H}_q}$  has much larger minimum distance than  $C_{1_i}$  promoting greater global error correction or erasure repair.

This situation motivates discerning additional levels of hierarchy in constructions, the primary topic of the next section.

### 4 Codes with H-level hierarchy from Reed–Muller and generalized fiber product construction

We now present the natural extension of H-LRCs to multiple levels of hierarchy. This generalization was mentioned in [23] and precisely defined in [8, 16]. Adapted to the notation above, we have the following definition.

**Definition 6** Let  $H, n_j, s_j, \delta_j \in \mathbb{Z}^+$  for all  $j \in [H]$  with  $H \geq 2, n_1 > n_2 > \dots > n_H, s_1 \geq s_2 \geq \dots \geq s_H$ , and  $\delta_1 > \delta_2 > \dots > \delta_H$ . An  $[n, k, d]$  linear code  $C$  has  $H$ -level hierarchical locality with parameters  $[(n_1, s_1, \delta_1), (n_2, s_2, \delta_2), \dots, (n_H, s_H, \delta_H)]$  if, for each  $i \in \{1, \dots, n\}$ , there is a set of  $H$  punctured codes  $\{(C_j)_i : j \in [H]\}$ , each code of length  $n_j$ , so that  $i \in \text{Supp}((C_j)_i), I_{j_i} = \text{Supp}((C_j)_i) \setminus \{i\}$ , so that

- $(I_H)_i \subseteq (I_{H-1})_i \subseteq \dots \subseteq (I_1)_i$ ,
- $\dim((C_j)_i) \leq s_j$  for all  $j$ ,
- $(C_j)_i$  has minimum distance at least  $\delta_j$  for all  $j$ ,
- $(C_j)_i$  has  $(H - j)$ -level hierarchical locality through the set of codes  $\{(C_k)_i : j < k \leq H\}$  for all  $k < H - 1$ .

The convention is that  $n_1 < n$ , but that is not required. While one could take  $C_1 = C$ , we typically do not do so to highlight the local nature of recovery meaning using a proper subset of coordinates of the code are employed at each level.

We think of the codes  $(C_j)_1, \dots, (C_j)_n$  as middle codes for the  $j$ -th level in the hierarchy for positions  $i = 1, \dots, n$ . Observe that a code with  $((n_1, s_1, \delta_1), (n_2, s_2, \delta_2))$  hierarchical locality as in Definition 5 has  $H$ -level hierarchy with  $H = 2$ .

As noted in the previous section, some codes with 2-level hierarchy can be further examined to produced additional tiers for recovery. In this section, we demonstrate that, noting how these recovery sets arise naturally from embedded structures.

### 4.1 H-level hierarchy from Reed–Muller codes

Recall that a Reed–Muller code consists of evaluations of points in  $\mathbb{F}_q^m$ , points which form the finite affine geometry of dimension  $m$ . Substructures in this geometry have dimension  $m - 1, m - 2, \dots, 1$ , where the substructures of dimension 1 are lines, those of dimension 2 are planes, and those of dimension  $d \geq 3$  are  $d$ -flats. An  $(m - 1)$ -flat is called a hyperplane. Starting with a codeword position corresponding to a point and choosing a line, then plane, a 3-flat, and so on as the nested recovery sets for that position, the result is an  $H$ -level H-LRC with  $H = m - 1$  and hierarchical parameters as described in the next result.

**Theorem 4.1** *The Reed–Muller code  $\mathcal{RM}_q(v, m)$  is an  $(m - 1)$ -level  $H$ -LRC with parameters*

- $n_j = q^{H-j+1} = q^{m-j}$
- $s_j = \binom{v+m-j}{m-j}$
- $\delta_j = (q - v)q^{m-j-1}$

for all  $j \in \{1, \dots, m - 1\}$ .

**Proof** Consider a point  $P \in \mathbb{F}_q^m$ . To recover an erasure in the coordinate indexed  $i$  by  $P$ , we may consider codes with increasing supports

$$(I_{m-1})_i \subseteq (I_{m-2})_i \subseteq \dots \subseteq (I_2)_i \subseteq (I_1)_i$$

where  $I_j \cup \{P\}$  is a  $(m - j)$ -dimensional subspace of  $\mathbb{F}_q^m$ . Then for each  $j \in \{1, \dots, m - 1\}$ , the punctured code  $(C_j)_i$  is isometric to the Reed–Muller code  $\mathcal{RM}_q(v, m - j)$ , which completes the proof. □

### 4.2 H-level hierarchy from fiber product codes

An  $LRC(t)$  arising from a  $t$ -fold fiber product naturally gives rise to up to  $t$ -level hierarchy. Figure 3 is a diagram of the curve covers that give this hierarchical structure.

**Theorem 4.2** *Let  $t \in \mathbb{Z}^+, t \geq 2$ . Let  $\mathcal{X}$  be a fiber product satisfying the hypotheses of and with all notation as in Theorem 2.1. Let  $C = C_{\mathcal{X}}$  be the  $[n, k, d], [(r_j = d_{h_j} - \rho_j + 1, \rho_j)]_{j \in \{1, \dots, t\}}$  LRC( $t$ ) constructed in the proof of Theorem 2.1. For  $H \in \{1, \dots, t\}, 2 \leq H, C$  is an  $H$ -LRC with  $H$ -level hierarchy, and hierarchical parameters:*

- $n_j = \prod_{k=j}^H d_{h_k}$ ,
- $s_j = \prod_{k=j}^H (d_{h_k} - \rho_k + 1)$ ,
- $\delta_j = \prod_{k=j}^H \rho_k$ .

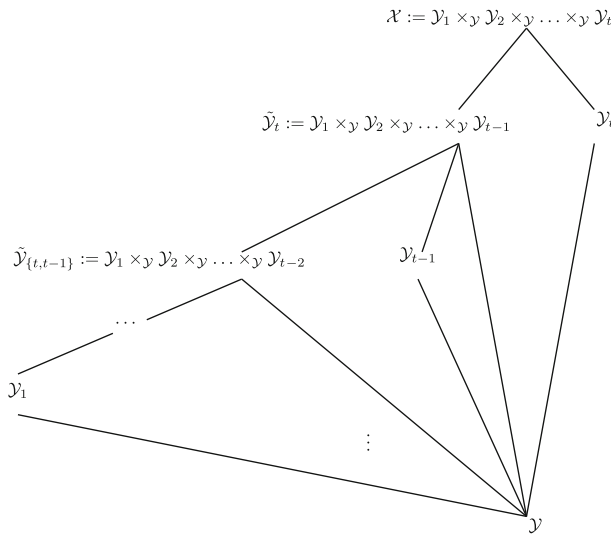
**Proof** Recall that  $S \subseteq \mathcal{X}(\mathbb{F}_q)$  is the set of evaluation points of  $C$ . We define the following two functions to simplify exposition. Given any set  $T \subseteq S$  and  $k \in [H]$ , let

$$F_k(T) = \{\tilde{g}_k^{-1}(\tilde{g}_k(Q)) : Q \in T\},$$

and

$$I(T) = \{i \in \{1, \dots, n\} : Q_i \in T\}.$$





**Fig. 3** A diagram of covers in the iterated fiber product construction. The tower of covers leading to  $t$ -level hierarchy is the far left sequence of covers

For any  $i \in \{1, \dots, n\}$ , let  $Q_i \in S$  be the corresponding evaluation point. We then take

$$(S_H)_i = F_H(\{Q_i\}).$$

This is the set of all evaluation points that share a value of  $g$  and all coordinate functions with  $Q_i$  except for the  $y_H$  coordinate. For  $j \in [H - 1]$ , let

$$(S_{H-j})_i = F_{H-j}((S_{H-j+1})_i).$$

Similarly, the points in set  $(S_j)_i$  will share a value of  $g$  and all coordinate functions with  $Q_i$  except  $y_j, y_{j+1}, \dots, y_H$ .

We then let  $(C_j)_i$  be the puncture of the code  $C$  to indices in  $I((S_j)_i)$ . This gives  $(I_j)_i = I((S_j)_i) \setminus \{i\}$ . The parameters of these codes are now obtainable from simple counting. The length  $n_j$  of the code  $(C_j)_i$  will be the product of the degrees of the corresponding maps  $h_k$  for  $j \leq k \leq H$ , because the points of  $S$  have been required to fully split in all the corresponding extensions. Since the evaluation points of the code  $(C_j)_i$  share a value of  $g$  and all coordinate functions except  $y_j, y_{j+1}, \dots, y_H$ , any function in the original evaluation set  $V$  will restrict to a function in the span of the set  $\{y_j^{e_j} y_{j+1}^{e_{j+1}} \dots y_H^{e_H} : 0 \leq e_k \leq d_{h_k} - \rho_k \text{ for all } k\}$ . This yields the given upper bound on the dimension  $s_j$  of the code  $(C_j)_i$ .

To bound the minimum distance of the code  $(C_j)_i$ , consider that, by a straight-forward extension of the argument in the proof of Theorem 3.2, each position  $i'$  in the support of  $(C_j)_i$  has  $H - j + 1$  disjoint recovery sets in  $(C_j)_i$  and the intersection possibilities of any recovery sets for any distinct positions in the support of  $(C_j)_i$  is extremely constrained. Thus we are able to conclude that if there are fewer than  $\rho_j \rho_{j+1} \dots \rho_H$  erasures in a codeword of  $(C_j)_i$ , there must be some position  $i'$  with fewer than  $\rho_k$  erasures in its  $k$ -th recovery set for some  $k \in \{j, j + 1, \dots, H\}$ . Therefore position  $i'$  can be recovered through its  $k$ -th recovery set; by repeated application any pattern of fewer than  $\rho_j \rho_{j+1} \dots \rho_H$  erasures can be recovered.  $\square$

**Remark 4** In the construction of Theorem 4.2, we have arbitrarily chosen to use the coordinates  $y_1, \dots, y_H$  and fixed an order on these so that the smallest code has evaluation points varying in  $y_H$ , intermediate middle codes each incorporate variability in previous coordinates, and the largest middle code has evaluation points varying in  $y_1$  through  $y_H$ . Given any choice of  $H$  values in  $\{1, \dots, t\}$  and any order of these values, we could use the fact that the fiber product construction is (up to isomorphism) commutative and associative to reorder the factors of the fiber product so that our chosen factors are in the positions 1 through  $H$  in the fiber product. Thus, this arbitrary choice of coordinates does not result in any loss of generality.

**Remark 5** In [1, Section VIII Part D], the authors create codes with hierarchical locality from coverings of algebraic curves, including a construction involving fiber products which is substantially different than that described here. The fiber product construction of Ballentine, Barg, and Vladut uses curves  $\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \mathcal{C}$ , where there exist covering maps  $\phi_2 : \mathcal{X} \rightarrow \mathcal{Y}$  and  $\phi_1 : \mathcal{Y} \rightarrow \mathcal{Z}$  with  $\deg(\phi_2) = ab$  and  $\deg(\phi_1) = b$ , and both  $\mathcal{X} \times_{\mathcal{Z}} \mathcal{C}$  and  $\mathcal{Y} \times_{\mathcal{Z}} \mathcal{C}$  are smooth, absolutely irreducible curves. Thus the construction starts with a tower of curves  $\mathcal{X} \rightarrow \mathcal{Y} \rightarrow \mathcal{Z}$  and then takes the fiber products (over  $\mathcal{Z}$ ) of each of these curves with another curve  $\mathcal{C}$  that is also a cover of  $\mathcal{Z}$ . The code is then defined over  $\mathcal{X} \times_{\mathcal{Z}} \mathcal{C}$ , with hierarchy resulting from covers in the tower  $\mathcal{X} \times_{\mathcal{Z}} \mathcal{C} \rightarrow \mathcal{Y} \times_{\mathcal{Z}} \mathcal{C} \rightarrow \mathcal{Z} \times_{\mathcal{Z}} \mathcal{C} \cong \mathcal{C}$ . In this work, we pursue a more general perspective on fiber product codes based on the construction in [18]. One major difference is that the iterated fiber product in [18] starts with some collection of curves  $\mathcal{X}_1, \mathcal{X}_2, \dots, \mathcal{X}_t$ , with no covering relationships among them but with each having a map to a shared curve  $\mathcal{Y}$ . We consider a code defined on

$$\mathcal{X} = \mathcal{X}_1 \times_{\mathcal{Y}} \mathcal{X}_2 \times_{\mathcal{Y}} \cdots \times_{\mathcal{Y}} \mathcal{X}_t.$$

The underlying fiber product of Ballentine et al. in [1] may coincide with this construction in some very special cases, but most situations of this paper and [18] are not covered by [1]. First, our construction does not require the strong restriction on the map degrees in [1]. Further, the highest level of hierarchy in our construction arises from the pullback of the map  $\mathcal{X} \rightarrow \mathcal{Y}$ , where  $\mathcal{Y}$  is the base curve of every fiber product. The base curve of the fiber product is not represented in the final tower of [1]. Also, the construction in [1] begins with a two-covering tower of curves and uses fiber products to obtain a related two-covering tower of curves. The construction in [18] that is continued here begins with many single-level covers of curves and uses each of these to create an additional level of hierarchy, with  $H$  levels of hierarchy arising from  $H$  single-level covers. Though not explicitly stated, it seems that the fiber product construction in [1] would need to start with an  $H$ -level tower of curves to generalize to  $H$ -level hierarchy after the fiber product is applied.

We return to the curve  $\mathcal{A}_{q,t}$  described in Subsect. 3.2.1. Let  $p$  be an odd prime,  $t, h \in \mathbb{Z}^+$ ,  $q = p^h$ ,  $t \leq h$ . Choose  $l \in \mathbb{Z}$  as before to define an LRC( $t$ )  $C_{\mathcal{A}_{q,t},l}$  over  $\mathbb{F}_{q^2}$  with parameters given in Theorem 3.3. This code has  $H$ -level hierarchy for any  $H \leq t$ . To display the full range of hierarchy, we choose  $H = t$ .

**Theorem 4.3** *The code  $C_{\mathcal{A}_{q,t},l}$  constructed in Theorem 3.3 is a  $t$ -level  $H$ -LRC with hierarchical parameters*

- $n_j = p^{t+1-j}$ ,
- $s_j = (p-1)^{t+1-j}$ , and
- $\delta_j = 2^{t+1-j}$ ,

for  $1 \leq j \leq t$ .

**Proof** The parameters are a direct application of Theorem 4.2, where our construction gives  $\rho_k = 2$  and  $d_{hk} = p$  for all  $k$ . □

### 5 Combined hierarchy and availability

Finally, we incorporate the notion of availability into our study of hierarchical locality. The following definition is a generalization of those given by Freij-Hollanti et al. in [12] and Ballentine et al. in [1].

**Definition 7** Let  $H, n_j, t_j, s_j, \delta_j, r_{j,k}, \rho_{j,k} \in \mathbb{Z}^+$  for all  $j \in [H]$  and  $k \in [t_j]$  with  $H \geq 2, s_1 \geq s_2 \geq \dots \geq s_H$  and  $\delta_1 > \delta_2 > \dots > \delta_H$ . Let  $C$  be an  $[n, k, d]$  linear code with  $H$ -level hierarchical locality with parameters  $(n_j, s_j, \delta_j)_{j \in [H]}$ . For every  $i \in \{1, \dots, n\}, j \in [H]$ , let  $(C_j)_i$  be the  $j$ -th level middle code for position  $i$ . If, for all  $i, (C_j)_i$  is a locally recoverable code with availability  $t_j$  and local parameters  $((r_{j,k}, \rho_{j,k})_{k \in [t_j]})$ , we say that  $C$  has  $H$ -level hierarchy with availability  $(t_j)_{j \in [H]}$ .

The notion of H-LRCs with availability is useful in capturing the additional flexibility of the geometric constructions we present here. Intuitively, if a middle code has availability, it potentially offers the ability to recover many erasure patterns in the middle code using local recovery. There may be another erasure recovery algorithm for the middle code, but having availability means we have additional local strategies at our disposal.

We now discuss the parameters of our example families when viewed in the more structured framework as codes with both  $H$ -level hierarchy and availability. Definition 7 describes codes with an  $H$ -level hierarchical recovery structure and  $t_j$  linearly independent recovery sets at level  $j, j \in [H]$ .

#### 5.1 Reed–Muller codes with hierarchy and availability

At each hierarchy level of the Reed–Muller code, there is natural availability using the underlying geometry of the affine spaces of  $\mathbb{F}_q^m$ .

**Corollary 1** Let  $C$  be the hierarchical locally recoverable code in Theorem 4.1. Then  $C$  is also an  $H = (m - 1)$ -LRC with availability parameters  $t_j = \frac{q^{m+1-j}-1}{q-1}$ , for  $j = 1, \dots, m - 1$ , and  $(r_{j,k}, \rho_{j,k}) = (r_k, \rho_k)$  for  $k \in [t_j]$  and  $j \in [H]$ .

**Remark 6** For a fixed point  $P \in \mathbb{F}_q^m$  that corresponds to the codeword position of  $C$  that we seek to recover, any pair of  $(m - j)$ -flats that contain  $P$  must overlap in an affine subspace of dimension  $(m - j - 1)$ . Suppose  $A_1, A_2$  are two affine subspaces of dimension  $(m - j)$  with  $P \in A_1, P \in A_2$ ; then  $B = A_1 \cap A_2$  is an affine subspace of dimension  $(m - j - 1)$ . We note that the conditions of Definition 7 are still satisfied by the code  $C$  since there is a recovery method using only the points in  $A_1 \setminus B$  (resp.,  $A_2 \setminus B$ ), as follows. To recover the codeword value at the point  $P$ , consider a line  $L$  through  $P$  such that  $L \setminus P \subseteq A_1 \setminus B$ . Notice that  $|(L \setminus P) \cap (A_1 \setminus B)| = q - 1$ , since otherwise the entire line would be contained in the affine subspace  $B$ . Thus, since  $v < q$  the value of a degree  $v$  polynomial can be recovered from the evaluation of the polynomial on the points on  $L \setminus P$ .

#### 5.2 Fiber product codes with hierarchy and availability

For codes from fiber products, we get natural availability at each level.

**Corollary 2** Let  $t \in \mathbb{Z}^+$ ,  $t \geq 2$ . Let  $\mathcal{X}$  be a fiber product satisfying the hypotheses of and with all notation as in Theorem 2.1. Let  $C = C_{\mathcal{X}}$  be the  $[n, k, d]$ ,  $[(r_j = d_{h_j} - \rho_j + 1, \rho_j)]_{j \in \{1, \dots, t\}}$  LRC( $t$ ) constructed in the proof of Theorem 2.1. Let  $H \in \{1, \dots, t\}$ ,  $2 \leq H$ , and consider the code  $C$  described as an  $H$ -LRC in Theorem 4.2. Then  $C$  is an  $H$ -LRC with  $H$ -level hierarchy and hierarchical parameters as given in the theorem, with availability parameters  $t_j = H + 1 - j$  for  $j \in [H]$  and  $(r_{j,k}, \rho_{j,k}) = (r_k, \rho_k)$  for  $k \in [t_j]$  and  $j \in [H]$ .

**Remark 7** As in the case of simple hierarchy, the authors of [1] devise  $H$ -LRCs with availability from a certain 2-level fiber product (see Fig. 2 in the cited paper). Their construction is a specialization of their main construction of hierarchy from a two-cover tower to a situation where each cover can be decomposed such that the corresponding covering curve is expressed as a fiber product over the covered curve. Our construction is not captured by [1]. Our construction uses fiber products differently. One very concrete difference is that our construction does not ever involve a fiber product using another fiber product as a base curve. Our hierarchy comes from the fibers of maps in a tower of fiber products with an increasing number of factors, with all fiber products using  $\mathcal{Y}$  as a base curve, as in Fig. 3.

## 6 Conclusion

In this paper, we harnessed the underlying structure of Reed–Muller and fiber product codes to provide hierarchical local recovery of erasures.  $H$ -LRCs have the advantage that they make use of smaller recovery sets for larger numbers of erasures than LRCs without tiered recovery sets of varying sizes. The constructions considered here are especially useful as their properties rely on the underlying geometry and immediately satisfy the disjoint or linearly independent repair group property. It remains to consider hierarchical recovery when disjoint or linearly independent repair groups are not required and how that may yield additional recovery sets at each level.

**Acknowledgements** The National Science Foundation partially supported the second author (DMS-2137661) and the third author (DMS-2201075). The third author is also partially supported by the Commonwealth Cyber Initiative.

**Author contributions** These authors contributed equally to this work.

**Data availability** No datasets were generated or analysed during the current study.

## Declarations

**Competing interests** The authors declare no competing interests.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. Ballentine S., Barg A., Vlăduț S.: Codes with hierarchical locality from covering maps of curves. *IEEE Trans. Inf. Theory* **65**(10), 6056–6071 (2019).
2. Barg A., Tamo I., Vlăduț S.: Locally recoverable codes on algebraic curves. In: 2015 IEEE International Symposium on Information Theory (ISIT), pp. 1252–1256 (2015).
3. Barg A., Tamo I., Vlăduț S.: Locally recoverable codes on algebraic curves. *IEEE Trans. Inf. Theory* **63**(8), 4928–4939 (2017).
4. Cai H., Cheng M., Fan C., Tang X.: Optimal locally repairable systematic codes based on packings. *IEEE Trans. Commun.* **67**(1), 39–49 (2019).
5. Cai H., Miao Y., Schwartz M., Tang X.: On optimal locally repairable codes with multiple disjoint repair sets. *IEEE Trans. Inf. Theory* **66**(4), 2402–2416 (2019).
6. Chara M., Kottler S., Malmskog B., Thompson B., West M.: Minimum distance and parameter ranges of locally recoverable codes with availability from fiber products of curves. arXiv preprint [arXiv:2204.03755](https://arxiv.org/abs/2204.03755) (2022).
7. Chara M., Kottler S., Malmskog B., Thompson B., West M.: Minimum distance and parameter ranges of locally recoverable codes with availability from fiber products of curves. *Des. Codes Cryptogr.* **91**(5), 2077–2105 (2023).
8. Chen Z., Barg A.: Cyclic LRC codes with hierarchy and availability. In: 2020 IEEE International Symposium on Information Theory (ISIT), pp. 616–621 (2020).
9. Chen B., Zhang G., Li W.: New optimal linear codes with hierarchical locality. *IEEE Trans. Inf. Theory* **69**(3), 1544–1550 (2022).
10. Dukes A., Micheli G., Lavorante V.P.: Optimal locally recoverable codes with hierarchy from nested F-adic expansions. *IEEE Trans. Inf. Theory* **69**(11), 6981–6988 (2023). <https://doi.org/10.1109/TIT.2023.3298401>.
11. Frank-Fischer S., Guruswami V., Wootters M.: Locality via partially lifted codes. In: Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2017), pp. 43–14317 (2017).
12. Freij-Hollanti R., Westerbäck T., Hollanti C.: Locally repairable codes with availability and hierarchy: matroid theory via examples. In: International Zurich Seminar on Communications-Proceedings, pp. 45–49 (2016).
13. Garrison C., Micheli G., Nott L., Lavorante V.P., Waitkevich P.: On a class of optimal locally recoverable codes with availability. In: 2023 IEEE International Symposium on Information Theory (ISIT), pp. 2021–2026. *IEEE* (2023).
14. Gopalan P., Huang C., Simitci H., Yekhanin S.: On the locality of codeword symbols. *IEEE Trans. Inf. Theory* **58**(11), 6925–6934 (2012).
15. Görtz U., Wedhorn T.: *Fiber Products*, pp. 93–117. Vieweg+Teubner, Wiesbaden (2010). [https://doi.org/10.1007/978-3-8348-9722-0\\_5](https://doi.org/10.1007/978-3-8348-9722-0_5).
16. Grezet M., Hollanti C.: The complete hierarchical locality of the punctured simplex code. *Des. Codes Cryptogr.* **89**(1), 105–125 (2021).
17. Guo A., Kopparty S., Sudan M.: New affine-invariant codes from lifting. In: Proceedings of the 4th Conference on Innovations in Theoretical Computer Science, pp. 529–540 (2013).
18. Haymaker K., Malmskog B., Matthews G.L.: Locally recoverable codes with availability  $t \geq 2$  from fiber products of curves. *Adv. Math. Commun.* **12**(2), 317–336 (2018).
19. Høholdt T., Van Lint J.H., Pellikaan R.: Algebraic geometry codes. *Handb. Coding Theory* **1**(Part 1), 871–961 (1998).
20. Huffman W.C., Pless V.: *Fundamentals of Error-Correcting Codes*. Cambridge University Press, Cambridge (2010).
21. López H.H., Malmskog B., Matthews G.L., Piñero-González F., Wootters M.: Hermitian-lifted codes. *Des. Codes Cryptogr.* **89**, 497–515 (2021).
22. Prakash N., Kamath G.M., Lalitha V., Kumar P.V.: Optimal linear codes with a local-error-correction property. In: 2012 IEEE International Symposium on Information Theory Proceedings, pp. 2776–2780 (2012).
23. Sasidharan B., Agarwal G.K., Kumar P.V.: Codes with hierarchical locality. In: 2015 IEEE International Symposium on Information Theory (ISIT), pp. 1257–1261 (2015).
24. Silberstein N., Zeh A.: Anticode-based locally repairable codes with high availability. *Des. Codes Cryptogr.* **86**, 419–445 (2018).
25. Stichtenoth H.: *Algebraic Function Fields and Codes*. Graduate Texts in Mathematics, vol. 254. Springer Berlin, Heidelberg (2009).

26. Tsfasman M.A., Vlăduț S.G., Nogin D.: Algebraic Geometric Codes: Basic Notions: Basic Notions, vol. 1. American Mathematical Society, Providence (2007).
27. van der Geer G., van der Vlugt M.: How to construct curves over finite fields with many points. In: Arithmetic Geometry, Symposia Mathematica XXXVII, pp. 169–189 (1997).
28. Wang A., Zhang Z.: Repair locality with multiple erasure tolerance. *IEEE Trans. Inf. Theory* **60**(11), 6979–6987 (2014). <https://doi.org/10.1109/TIT.2014.2351404>.
29. Yekhanin S.: Locally decodable codes. *Found. Trends@ Theor. Computer. Sci.* **6**(3), 139–255 (2012).
30. Zhang G., Liu H.: Constructions of optimal codes with hierarchical locality. *IEEE Trans. Inf. Theory* **66**(12), 7333–7340 (2020).

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.