Wedge-Lifted Codes

Jabari Hastings, Amy Kanne, Ray Li, Mary Wootters
Stanford University
Email: {jabarih, akanne, rayyli, marykw}@stanford.edu

***Abstract*—We define *wedge-lifted codes*, a variant of lifted codes, and we study their locality properties. We show that (taking the trace of) wedge-lifted codes yields binary codes with the t-disjoint repair property (t-DRGP). When $t = N^{1/2d}$, where N is the block length of the code and $d \geq 2$ is any integer, our codes give improved trade-offs between redundancy and locality among binary codes.**

A full version of this paper is accessible at: https://arxiv.org/pdf/2011.12435.pdf

I. Introduction

In this work, we define and study *Wedge-Lifted Codes*, and show they yield improved binary error correcting codes for a notion of locality known as the t-disjoint repair group property.

An *error correcting code* (or simply *code*) $\mathcal{C} \subset \Sigma^N$ is a set of strings of a fixed *length* N over an *alphabet* Σ. If $\Sigma = \{0,1\}$, $\mathcal{C}$ is called a *binary* code. We measure the quality of a code by the *redundancy*, denoted $K^\perp$, defined as $K^\perp = N - K$, where $K = \log_{|\Sigma|}|\mathcal{C}|$ is the *dimension* of the code. It is desirable for codes to be larger, or equivalently to have less redundancy.

In this work we are interested in constructing better (less redundant) binary codes with *locality*. There are several notions of locality in this literature, but informally a code exhibits locality if we can correct one or a small number of erasures by looking only *locally* at a few other symbols of the codeword. In this work we construct codes with a notion of locality known as the t-disjoint repair group property (t-DRGP).

Definition I.1. A code $\mathcal{C} \subseteq \Sigma^N$ has the $t-$*disjoint repair group property* (t-DRGP) if for every $i \in [N]$, there is a collection of t disjoint subsets $S_1, \ldots, S_t \subseteq [N] \setminus \{i\}$ and functions $f_1, \ldots, f_t$ so that for all $c \in \mathcal{C}$ and $j \in [t]$, $f_j\left(c|_{S_j}\right) = c_i$.

Codes with the t-DRGP are motivated by distributed storage, where one desires efficient recovery from a few erasures (see surveys [18], [21]). Codes with the t-DRGP are also relevant to *private information retrieval* (PIR) in cryptography, as all codes (with a systematic encoding) with the t-DRGP also form $(t+1)$-*PIR* codes [4]. Further, when $t = \Omega(N)$ is large, codes with the t-DRGP are equivalent to *locally correctable codes* [12], [26].

Previously the best constructions for codes with the disjoint repair group property were given by lifted multiplicity codes [10], [15] (see also [28]), but these codes have very large alphabet sizes, which is undesirable from the perspective of distributed storage and PIR. Hence, a natural question, explicitly asked in [15], is, what are the best constructions of *binary* codes with the t-DRGP? Our work makes progress on this question by giving new constructions of binary codes with the t-DRGP with the best known redundancy for some values of t.

Theorem I.2. *For positive integers* d *and infinitely many* N*, for* $t = N^{1/2d}$*, there exist binary codes of length* N *with redundancy* $t^{\log_2(2-2^{-d})}\sqrt{N}$ *that have the the* $(t-1)$*-DRGP.*

Theorem I.2 gives improved constructions of binary codes with the t-DRGP when $t = N^{1/2d}$ for integers $d \geq 2$. (see Figure 1 for a visual comparison and Section I-A for a more detailed comparison): When $d = 2$, Theorem I.2 improves over a construction of [5]. When $d \geq 3$, Theorem I.2 improves over the constructions of [4]. As all codes with the $(t-1)$-DRGP are also t-PIR codes, Theorem I.2 also gives improved constructions of t-PIR codes in the same parameter settings.

It is an interesting question whether the construction of [4] can be beaten for all $t \in (1, \sqrt{N})$. That is, for all $t = N^\alpha$ with $\alpha \in (0, 1/2)$, are there binary codes with the t-DRGP and redundancy $O(t^{1-\varepsilon}\sqrt{N})$ for some $\varepsilon > 0$ (possibly depending on α)? Our work shows this is true for $\alpha = 1/2d$ when d is a positive integer. Additionally, for a dense collection of $\alpha \in (0, 1/2)$, our binary codes essentially match the redundancy bound of $O(t\sqrt{N})$ from [4] (Theorem III.7). This is proved with a naive bound, so it is possible that, with a more refined analysis, our codes could achieve the improvement to $O(t^{1-\varepsilon}\sqrt{N})$ for all $\alpha \in (0, 1/2)$.

In the remainder of this section, we highlight some related work and our approach. In Section II, we state some preliminaries. In Section III, we define and analyze our construction of Wedge-Lifted Codes. In Section IV, we show demonstrate how to turn the codes in Section III, which are over a q-ary alphabet, into binary codes, proving Theorem I.2.

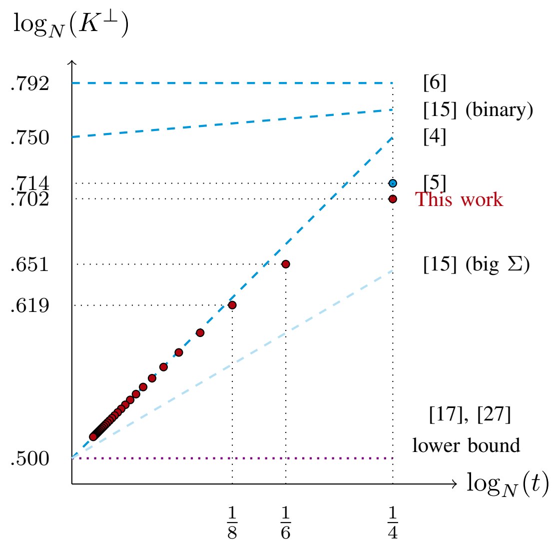

Fig. 1. The best trade-offs known between the number t of disjoint repair groups and the redundancy $K^\perp$, for $t \leq N^{1/4}$. Our results appear as the red dots.

A. Related work

a) Prior work on DRGP: We refer the reader to Figure 1 for a picture of the best known constructions of codes with the t-DRGP. For $t \leq \sqrt{N}$, [4] constructed codes with t-DRGP and $t\sqrt{N}$ bits of redundancy. We show (Corollary III.7) that there always exist Wedge-Lifted Codes that *at least* match the redundancy this construction. However, the construction of [4] is not optimal, at least when t is somewhat large. [6] showed that *Lifted Reed Solomon* codes have the $N^{1/2}$-DRGP with redundancy at most $N^{0.792} \ll t\sqrt{N}$. A later work [5] also improved [4] when $t = N^{1/4}$, giving codes with redundancy at most $N^{0.714}$. When $t = N^{1/4}$, Theorem I.2 improves on [5] by giving codes with redundancy roughly $N^{0.702}$. When $t = N^{1/2d}$ for all $d \geq 3$, Theorem I.2 improves [4] by giving codes with redundancy $t^{1-\varepsilon_d}\sqrt{N}$.

Additional results are known for larger values of t [1], [10]. We focus on the setting $t \leq N^{1/4}$ because we believe this parameter regime is already interesting, and because we believe that the improvements of Wedge-Lifted Codes over existing binary codes are the strongest in this parameter regime. We expect that generalizing our bivariate construction of Wedge-Lifted Codes to more than two variables would give improved binary codes with the t-DRGP for larger values of t.

b) Other notions of locality: The t-DRGP is closely related to other notions of locality, including PIR codes [1], [2], [4], batch codes [1], [9], [11], [20], *locally correctable codes* (LCCs) [6], [8], [13], [14] and *locally decodable codes* (LDCs) [3], [12], [26], and *LRCs with availability* [19], [23]–[25].

B. Our approach

The best known constructions of codes [5], [6], [10], [15] with the DRGP leverage an idea called *lifting* [6]. The basic idea of lifting is to improve algebraic codes like the Reed-Muller Code, which has some locality properties, by relaxing the construction so that the locality properties are preserved but so that the redundancy decreases. In this work, we propose a variation of the lifting technique which we call *wedge-lifts*. This method avoids the large alphabets of *lifted multiplicity codes* [10], [15], [16], [28], improves on the *partially lifted codes* of [5], and additionally gives the best constructions of codes with the t-DRGP for $t = N^{1/2d}$ for integers $d \geq 2$. Our wedge-lifted codes are not binary, but their alphabets are small enough that we can make them binary by taking the coordinate-wise trace of the code. For a full overview of our approach, see [7].

II. Preliminaries

In this section, we introduce the background and notation we use throughout the paper.

A. Notation and basic definitions

Let $\mathbb{F}_q$ denote the finite field of order q and let $\mathbb{F}_q^\times$ denote its multiplicative subgroup. We study linear codes $\mathcal{C} \subseteq \mathbb{F}_q^N$ of block length N over an alphabet of size q. Throughout this paper, we assume that $\mathbb{F}_q$ has characteristic 2 and write $q = 2^\ell$.

We need the following tools to reason about the binary representations of integers. If x and y are two non-negative integers with binary representations $x = \overline{x_{\ell-1} \cdots x_0}$ and $y = \overline{y_{\ell-1} \cdots y_0}$, then define the bitwise-OR $\vee$ and bitwise-AND $\wedge$ of x and y by

$$x \vee y = \overline{\max(x_{\ell-1}, y_{\ell-1}) \cdots \max(x_0, y_0)}$$
$$x \wedge y = \overline{\min(x_{\ell-1}, y_{\ell-1}) \cdots \min(x_0, y_0)}$$

Furthermore, we say that x *lies in the* 2*-shadow* of y, denoted $x \leq_2 y$, if $x_i \leq y_i$ for all $i \in \{0, \ldots, \ell - 1\}$. We are interested in 2-shadows because of the following corollary of Lucas's theorem.

Theorem II.1 (Follows from Lucas's Theorem)**.** *Let $x = \overline{x_{\ell-1}\cdots x_0}$ and $y = \overline{y_{\ell-1}\cdots y_0}$ be written in binary. Then, $\binom{x}{y} \equiv 1 \mod 2$ if and only if $y \leq_2 x$.*

The codes $\mathcal{C}$ we consider are polynomial evaluation codes. For a polynomial $P \in \mathbb{F}_q[X_1,\dots,X_m]$, we write its corresponding codeword as

$$\mathrm{eval}_q(P) = \langle P(x_1,\dots x_m)\rangle_{(x_1,\dots,x_m)\in\mathbb{F}_q^m}.$$

For the rest of the paper, take $m = 2$. We are concerned with the restriction of bivariate polynomials to lines. For a line $L : \mathbb{F}_q \to \mathbb{F}_q^2$ with $L(T) = (L_1(T), L_2(T))$ and a polynomial $P : \mathbb{F}_q^2 \to \mathbb{F}_q$, we define the restriction of P on L, denoted $P|_L$, to be the unique polynomial of degree at most $q-1$ so that $P|_L(T) = P(L_1(T), L_2(T))$.

B. Trace codes

Given a linear code $\mathcal{C} \subseteq \mathbb{F}_{p^\ell}^N$, it is sometimes desirable to construct another code $\mathcal{C}' \subseteq \mathbb{F}_p^N$ over a smaller alphabet that maintains some properties of $\mathcal{C}$. We briefly describe how trace functions give such codes.

Definition II.2. Let $\mathrm{tr}_p : \mathbb{F}_{p^\ell} \to \mathbb{F}_p$ be the trace function

$$\mathrm{tr}_p(\alpha) = \sum_{i=0}^{\ell-1} \alpha^{p^i}$$

We can extend $\mathrm{tr}_p : \mathbb{F}_{p^\ell}^N \to \mathbb{F}_p^N$ by defining

$$\mathrm{tr}_p(v) = (\mathrm{tr}_p(v_1),\dots,\mathrm{tr}_p(v_N)) \in \mathbb{F}_p^N$$

We can further extend it to codes $\mathcal{C} \subseteq \mathbb{F}_{p^\ell}^N$ by taking the the trace of every vector in the code:

$$\mathrm{tr}_p(\mathcal{C}) = \{\mathrm{tr}_p(v) : v \in \mathcal{C}\} \subseteq \mathbb{F}_p^N$$

Note that tr_p is a $\mathbb{F}_p$-linear map. Hence, if $\mathcal{C}$ is a linear code, then $\mathrm{tr}_p(\mathcal{C})$ is also a linear code. We can bound the rate of $\mathrm{tr}_p(\mathcal{C})$ in terms of the rate of $\mathcal{C}$ using the following corollary of Delsarte's theorem.

Theorem II.3 (Follows from Delsarte's theorem (see, e.g., [22]))**.** *For any $\mathbb{F}_{p^\ell}$-linear code $\mathcal{C} \subseteq \mathbb{F}_{p^\ell}^N$,*

$$\dim \mathcal{C} \leq \dim \mathrm{tr}_p(\mathcal{C}) \leq \ell \cdot \dim \mathcal{C}.$$

III. Wedge-lifted codes

In this section, we define and analyze wedge-lifted codes. As mentioned above, we focus on bivariate codes, although as remarked above we believe our work could be extended to more variables.

A. Definition of wedge-lifted codes

For a point $\vec{p} = (x, y) \in \mathbb{F}_q^2$ and a set $H \subseteq \mathbb{F}_q$, we define the *wedge* through $\vec{p}$ formed by H, denoted $W_{H,\vec{p}}$, as the set of affine lines passing through $\vec{p}$ whose slope is in H,

$$W_{H,\vec{p}} = \{L(T) = (T, \alpha(T - x) + y) : \alpha \in H\}.$$

For a wedge $W_{H,\vec{p}}$ and polynomial $P \in \mathbb{F}_q[X, Y]$, we define the *wedge restriction* of P to the wedge $W_{H,\vec{p}}$, denoted $P|_{W_{H,\vec{p}}}$, to be the sum of the restrictions of P to each line in the wedge,

$$P|_{W_{H,\vec{p}}} = \sum_{L\in W_{H,\vec{p}}} P|_L = \sum_{\alpha\in H}\sum_{T\in\mathbb{F}_q} P(T, \alpha(T - x) + y).$$

Note that when $W_{H,\vec{p}}$ consists of an odd number of lines and the field has characteristic 2, the wedge-restriction of P is equivalent to the sum of P's evaluations of each point in the wedge,

$$P|_{W_{H,\vec{p}}} = \sum_{(x,y)\in W_{H,\vec{p}}} P(x, y).$$

Definition III.1 (Wedge-lifted codes)**.** Let $\mathcal{H}$ be a collection of disjoint subsets of $\mathbb{F}_q$, with each $H \in \mathcal{H}$ having odd size. The $(\mathcal{H}, q)$ *wedge-lifted code* is a code $\mathcal{C}$ over alphabet $\Sigma = \mathbb{F}_q$ of length q^2 given by

$$\mathcal{C} = \left\{ \mathrm{eval}_q(P) \;\middle|\; \begin{array}{c} P \in \mathbb{F}_q[X, Y],\ P|_{W_{H,\vec{p}}} = 0 \\ \text{for all } H \in \mathcal{H} \text{ and } \vec{p} \in \mathbb{F}_q^2, \end{array} \right\}.$$

The number of disjoint repair groups for a wedge-lifted code follows from the code's definition (see full paper for proof [7]).

Proposition III.2. *The $(\mathcal{H}, q)$ wedge-lifted code has $|\mathcal{H}|$ disjoint repair groups.*

Following the approach of previous work [5], [6], [10], [15], [16], we show that wedge-lifted codes contain the evaluations of polynomials that lie in the span of "good" monomials. Informally, a monomial is $(\mathcal{H}, q)$-good if it restricts nicely to all the wedges given by $\mathcal{H}$.

Definition III.3 ($(\mathcal{H}, q)$-good monomials)**.** Let $\mathcal{H}$ be a collection of disjoint subsets of $\mathbb{F}_q$. We say that a monomial $P(X, Y) = X^aY^b$ with $0 \leq a, b \leq q - 1$ is $(\mathcal{H}, q)$-*good* if for every $H \in \mathcal{H}$ and every $\vec{p} \in \mathbb{F}_q^2$, $P|_{W_{H,\vec{p}}} = 0$, and say it is $(\mathcal{H}, q)$-*bad* otherwise.

By definition, the evaluations of all good monomials lie within our wedge-lifted codes. Furthermore, monomials X^aY^b with $a, b \leq q-1$ form a basis for polynomials of degree at most $q - 1$, which are in bijection with $\mathbb{F}_q^{q^2}$ through the eval_q map. Therefore, we can obtain a lower bound on the rate of our code by finding a lower bound on the number of good monomials.

Observation III.4. *For any $(\mathcal{H}, q)$ wedge-lifted code $\mathcal{C}$, the redundancy of $\mathcal{C}$ is at most the number of $(\mathcal{H}, q)$-bad monomials.*

B. Wedge-lifted codes via cosets

In this section we analyze wedge-lifted codes that arise when $\mathcal{H}$ is a collection of cosets. Using cosets allow us to leverage the following important fact.

Fact III.5. *Let $H \leq \mathbb{F}_q^\times$ be a subgroup. For a nonnegative integer n, the sum $\sum_{\alpha \in H} \alpha^n$ is $|H|$ if n is a multiple of $|H|$ and 0 otherwise.*

Lemma III.6. *Let $H \leq \mathbb{F}_q^\times$ be a subgroup, $\mathcal{H}$ be the collection of cosets gH of $\mathbb{F}_q^\times$, and a, b be integers such that $0 < a, b \leq q-1$. Then, a monomial X^aY^b is $(\mathcal{H}, q)$-bad if and only if both of the following conditions hold:*

1) *$a \vee b = q - 1$*
2) *There exists an $i \equiv b \mod |H|$ such that $i \leq_2 a \wedge b$.*

Proof. Say that $\vec{p} = (x, y)$ for any $x, y \in \mathbb{F}_q$, we aim to show that $P|_{W_{gH,\vec{p}}} = 0$ for any choice of x, y if and only if the two conditions hold. First, in the case that gH is a coset of a subgroup $H \leq F_q^\times$, we can simplify $P|_{W_{gH,\vec{p}}}$ to (see [7] for details)

$$P|_{W_{gH,\vec{p}}} = \sum_{\alpha \in H} \sum_{T \in \mathbb{F}_q} \sum_{i=0}^{b} \binom{b}{i} (g\alpha)^i (y - g\alpha x)^{b-i} T^{a+i}.$$

For brevity, we assume $a + b < 2q - 2$. The case $a = b = q - 1$ is similar, negligible, and handled in the full paper [7]. With this assumption and the fact that $\sum_{T \in \mathbb{F}_q} T^{a+i} = 1$ if $a + i = q - 1$ and 0 otherwise, we can simplify the above sum.

$$P|_{W_{gH,\vec{p}}} = \sum_{\alpha \in H} \binom{b}{q-1-a} (g\alpha)^{-a} (y - g\alpha x)^{b+a-(q-1)}$$

Theorem II.1 states that $\binom{b}{q-1-a} = 1$ if and only if $q - 1 - a \leq_2 b$, which is equivalent to $a \vee b = q - 1$. So, if $a \vee b \neq q - 1$, $P|_{W_{gH,\vec{p}}} = 0$, so P is good.

It remains to show that if $a \vee b = q-1$, then $P|_{W_{\alpha H,\vec{p}}} = 0$ if and only if there exists an $i \equiv b \mod |H|$ such that $i \leq_2 a \wedge b$. We then expand the binomial term $(y - g\alpha x)^{b+q-(q-1)}$ in the sum of the above expression and use Fact III.5 to get that $P|_{W_{gH,\vec{p}}}$ is

$$\sum_{\substack{0 \leq i \leq b+a-(q-1) \\ i \equiv b \mod |H|}}^{b+a-(q-1)} \binom{a+b-(q-1)}{i} y^i x^{a+b-(q-1)-i} g^{b-i}.$$

Then, if we view this as a bivariate polynomial in x, y, we see that it has degree at most $q-1$ in each variable, so it is 0 for every $x, y \in \mathbb{F}_q$ if and only if every coefficient of x and y is 0. Therefore, P is bad if and only if $g^{b-i}\binom{a+b-(q-1)}{i} \neq 0$ for some $i \equiv b \mod |H|$. Because $a \vee b = q - 1$ in this case, we can compute that $a + b - (q - 1) = a \wedge b$. Moreover, $g \neq 0$, so applying Theorem II.1 to $\binom{a \wedge b}{i}$ gives us the following equivalence.

$$g^{b-i}\binom{a+b-(q-1)}{i} \neq 0 \iff i \leq_2 a \wedge b,$$

which immediately yields the lemma statement. ☐

We can apply this result with a naive bound on the number of a, b such that those two conditions hold: essentially, there are $t = (q-1)/|H|$ possible values of $b-i$, and each choice of $b-i$ yields $\sqrt{N} = q$ possible bad pairs (a, b) (see [7] for details). This demonstrates that wedge-lifted codes are no worse than the non-algebraic constructions of [2], [4].

Corollary III.7. *Let $H \leq \mathbb{F}_q^\times$ be a subgroup and let $\mathcal{H}$ be the collection of cosets gH of $\mathbb{F}_q^\times$. Then, for $t = (q-1)/|H|$, the $(\mathcal{H}, q)$ wedge-lifted code has the t-DRGP and redundancy at most $t\sqrt{N}$.*

By standard arguments [7], in Corollary III.7, t can be taken to be $N^{\alpha+o(1)}$ for any $\alpha \in (0, 1/2)$ when q and H are appropriately chosen, so our construction indeed matches that of [2], [4] in the whole parameter regime $t = N^\alpha$ for $\alpha \in (0, 1/2)$.

C. Instantiations

We now give a better bound on the redundancy of wedge-lifted codes. We continue to assume that $\mathcal{H}$ is a collection of cosets gH, examining the special case where each coset in $\mathcal{H}$ has order $|H| = (q-1)/(q^{1/d}-1)$ where d is an integer and $\ell = \log_2(q)$ is a multiple of d. Under these conditions, we give a precise description of the monomials that fail to satisfy Lemma III.6. The key observation is that multiples of $(q-1)/(q^{1/d}-1)$ that are less than q have repeating bit patterns.

Observation III.8. *Let $q = 2^{\ell' d}$ for positive integers ℓ' and d. Then any nonnegative integer $a < q$ with binary representation $\overline{a_{d\ell'-1} \cdots a_0}$ is a multiple of $(q-1)/(q^{1/d}-1)$ if and only if $\overline{a_{\ell'-1} \cdots a_0} = \overline{a_{2\ell'-1} \cdots a_{\ell'}} = \cdots = \overline{a_{d\ell'-1} \cdots a_{(d-1)\ell'}}$.*

Combining this observation with Lemma III.6, we can prove the following characterization of bad monomials when $|H| = (q-1)/(q^{1/d}-1)$.

Lemma III.9. *Let $q = 2^{\ell' d}$ for some natural numbers ℓ' and d. Let $H \leq \mathbb{F}_q^\times$ be a subgroup of order $|H| = (q-1)/(q^{1/d}-1)$ and $\mathcal{H}$ be the collection of cosets gH. Then any monomial X^aY^b with $0 \leq a, b \leq q-1$ is $(\mathcal{H}, q)$-bad if and only if both of the following conditions hold:*

1) $a \vee b = q-1$
2) *There do not exist $j \in \{0, 1, \ldots, \ell'-1\}$ and $r, s \in \{0, 1 \ldots, d-1\}$ such that $b_{r\ell'+j} = a_{s\ell'+j} = 1$ and $a_{r\ell'+j} = b_{s\ell'+j} = 0$.*

Proof. We prove the forward direction and leave the reverse direction to the full version [7]. If X^aY^b is $(\mathcal{H}, q)$-bad, then there exists $i \equiv b \mod |H|$ such that $i \leq_2 a \wedge b$ by Lemma III.6. Since $i \leq_2 b$ and $b - i \equiv 0 \mod |H|$, if there exist $r, s \in \{0, 1 \ldots, d-1\}$ and $j \in \{0, 1, \ldots, \ell'-1\}$ such that $b_{r\ell'+j} = 1$ and $b_{s\ell'+j} = 0$, then $i_{r\ell'+j} = 1$ and $i_{s\ell'+j} = 0$. Note that $i \leq_2 a$ and $i_{r\ell'+j} = 1$ together imply that $a_{r\ell'+j} = 1$. Note also that $a \vee b = q-1$ and $b_{s\ell'+j} = 0$ together imply that $a_{s\ell'+j} = 1$. Hence, it is never the case that $b_{r\ell'+j} = a_{s\ell'+j} = 1$ and $a_{r\ell'+j} = b_{s\ell'+j} = 0$. □

With the description of $(\mathcal{H}, q)$-bad monomials in hand, we can give an exact count: essentially, for each values of $j \in \{0, \ldots, \ell'-1\}$, there are $2^{d+1}-1$ choices for the bits whose positions are $j \mod \ell'$ (bits $a_{r\ell'+j}, b_{r\ell'+j}$ for $r \in \{0, \ldots, d-1\}$) of a bad monomial X^aY^b.

Corollary III.10. *Let $q = 2^{\ell' d}$ for some natural numbers ℓ' and d. Let $H \leq \mathbb{F}_q^\times$ be a subgroup of order $|H| = (q-1)/(q^{1/d}-1)$ and $\mathcal{H}$ be the collection of cosets gH. Then, there are $(2^{d+1}-1)^{\ell'}$ $(\mathcal{H}, q)$-bad monomials.*

We summarize the properties of the wedge-lifted codes constructed from our special choice of $\mathcal{H}$.

Theorem III.11. *Let $q = 2^{\ell' d}$ for some natural numbers ℓ' and d. Let $H \leq \mathbb{F}_q^\times$ be a subgroup of order $|H| = (q-1)/(q^{1/d}-1)$ and $\mathcal{H}$ be the collection of cosets gH. Then, the $(\mathcal{H}, q)$ wedge-lifted code has*

- *length q^2.*
- *alphabet size q.*
- *redundancy at most $(2^{d+1}-1)^{\ell'}$.*
- *and the $(2^{\ell'}-1)$-disjoint repair group property.*

One can check that, for $t = 2^{\ell'}$, this gives the desired code for Theorem I.2, except that the code is not binary. To obtain a binary code, we take the coordinate-wise trace of the codewords, described in the next section.

IV. Trace of Wedge-Lifted Codes

In this section, we take the trace of our codes in Theorem III.11 and set parameters, in order to prove our main theorem, Theorem I.2.

The key lemma says that we can take the coordinate-wise trace of a code over a field of characteristic two without hurting the redundancy or locality.

Lemma IV.1. *Let $\mathcal{C} \subseteq \mathbb{F}_q^{q^2}$ be a $(\mathcal{H}, q)$-wedge-lifted code. Then, there exists a binary code $\mathcal{C}' \subseteq \mathbb{F}_2^{q^2}$ with the same or lower redundancy and the same number of disjoint repair groups.*

Proof. For any polynomial P such that $\mathrm{eval}(P) \in \mathcal{C}$, $H \in \mathcal{H}$, and $(x, y) = \vec{p} \in \mathbb{F}_q^2$, we know that $P|_{W_{H,\vec{p}}} = 0$. Therefore, using the fact that tr_2 commutes with addition,

$$\begin{aligned}\mathrm{tr}_2(0) &= \mathrm{tr}_2\left(\sum_{\alpha \in H}\sum_{T \in \mathbb{F}_q} P(T, \alpha(T-x)+y)\right) \\ &= \sum_{\alpha \in H}\sum_{T \in \mathbb{F}_q} \mathrm{tr}_2(P(T, \alpha(T-x)+y))\end{aligned}$$

Therefore, if we view the code $\mathrm{tr}_2(\mathcal{C})$ as a code over $\mathbb{F}_2^{q^2}$ with coordinates indexed by $\mathbb{F}_q^2$, any index $\vec{p}$ of a codeword $\mathrm{eval}(\mathrm{tr}_2 \circ P)$ of $\mathrm{tr}_2(\mathcal{C})$ can be repaired by summing over the lines whose slopes are in H. So, $\mathrm{tr}_2(\mathcal{C})$ and $\mathcal{C}$ both have $|\mathcal{H}|$ disjoint repair groups.

Furthermore, Theorem II.3 states that $\dim(\mathrm{tr}_2(\mathcal{C})) \geq \dim(\mathcal{C})$, so because these codes have the same length, $\mathrm{tr}_2(\mathcal{C})$ has the same or lower redundancy as $\mathcal{C}$. □

We obtain Theorem I.2 by applying Lemma IV.1 to Theorem III.11. As a corollary of Lemma IV.1, we can also make the codes of Corollary III.7, which match [4] for a dense collection of $\alpha \in (0, 1/2)$, into binary codes. See [7] for details.

V. Conclusion

In this paper, we introduced *wedge-lifted codes*, which give an improved construction of binary codes with the t-DRGP for several $t \leq \sqrt{N}$. We conclude with some open questions.

1) For all $\alpha \in (0, 1/2)$ with $t = N^\alpha$, are there binary codes with the t-DRGP and redundancy $O(t^{1-\varepsilon}\sqrt{N})$ for some $\varepsilon > 0$, possibly (but ideally not) depending on α? Our work shows this is true for $\alpha = 1/2d$ when d is any positive integer. The work of [15] showed this is true (with an absolute $\varepsilon = 0.425$) for nonbinary codes.
2) Can we improve [17], [27] to prove better lower bounds on the redudancy of t-DRGP codes?

References

[1] Hilal Asi and Eitan Yaakobi. Nearly optimal constructions of pir and batch codes. *IEEE Transactions on Information Theory*, 65(2):947–964, 2019.

[2] Simon R. Blackburn and Tuvi Etzion. PIR Array Codes with Optimal PIR Rate. *ArXiv e-prints*, July 2016.

[3] Klim Efremenko. 3-query locally decodable codes of subexponential length. *SIAM Journal on Computing*, 41(6):1694–1703, 2012.

[4] Arman Fazeli, Alexander Vardy, and Eitan Yaakobi. Codes for distributed pir with low storage overhead. In *2015 IEEE International Symposium on Information Theory (ISIT)*, pages 2852–2856. IEEE, 2015.

[5] S Luna Frank-Fischer, Venkatesan Guruswami, and Mary Wootters. Locality via partially lifted codes. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2017)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2017.

[6] Alan Guo, Swastik Kopparty, and Madhu Sudan. New affine-invariant codes from lifting. In *Innovations in Theoretical Computer Science, ITCS '13, Berkeley, CA, USA, January 9-12, 2013*, pages 529–540, 2013.

[7] Jabari Hastings, Amy Kanne, Ray Li, and Mary Wootters. Wedge-lifted codes. *arXiv preprint arXiv:2011.04453*, 2020.

[8] Brett Hemenway, Rafail Ostrovsky, and Mary Wootters. Local correctability of expander codes. *Information and Computation*, 243:178–190, 2015.

[9] Lukas Holzbaur, Rina Polyanskaya, Nikita Polyanskii, and Ilya Vorobyev. Lifted reed-solomon codes with application to batch codes. pages 634–639, 2020.

[10] Lukas Holzbaur, Rina Polyanskaya, Nikita Polyanskii, Ilya Vorobyev, and Eitan Yaakobi. Lifted multiplicity codes. *arXiv preprint arXiv:2008.04717*, 2020.

[11] Yuval Ishai, Eyal Kushilevitz, Rafail Ostrovsky, and Amit Sahai. Batch codes and their applications. In *Proceedings of the thirty-sixth annual ACM Symposium on the Theory of Computing*, STOC 2004, pages 262–271. ACM, 2004.

[12] Jonathan Katz and Luca Trevisan. On the efficiency of local decoding procedures for error-correcting codes. In *Proceedings of the 32nd symposium on Theory of Computing*, STOC 2000, pages 80–86, 2000.

[13] Swastik Kopparty, Or Meir, Noga Ron-Zewi, and Shubhangi Saraf. High-rate locally-correctable and locally-testable codes with sub-polynomial query complexity. In *Proceedings of the 48th Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2016, pages 202–215. ACM, 2016.

[14] Swastik Kopparty, Shubhangi Saraf, and Sergey Yekhanin. High-rate codes with sublinear-time decoding. *Journal of the ACM (JACM)*, 61(5):28, 2014.

[15] Ray Li and Mary Wootters. Lifted Multiplicity Codes and the Disjoint Repair Group Property. In Dimitris Achlioptas and László A. Végh, editors, *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2019)*, volume 145 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 38:1–38:18, Dagstuhl, Germany, 2019. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.

[16] N. Polyanskii and I. Vorobyev. Trivariate lifted codes with disjoint repair groups. In *2019 XVI International Symposium "Problems of Redundancy in Information and Control Systems" (REDUNDANCY)*, pages 64–68, 2019.

[17] Sankeerth Rao and Alexander Vardy. Lower bound on the redundancy of PIR codes. *arXiv preprint arXiv:1605.01869*, 2016.

[18] KV Rashmi, Nihar B Shah, Dikang Gu, Hairong Kuang, Dhruba Borthakur, and Kannan Ramchandran. A solution to the network challenges of data recovery in erasure-coded distributed storage systems: A study on the facebook warehouse cluster. In *5th {USENIX} Workshop on Hot Topics in Storage and File Systems (HotStorage 13)*, 2013.

[19] Ankit Singh Rawat, Dimitris S Papailiopoulos, Alexandros G Dimakis, and Sriram Vishwanath. Locality and availability in distributed storage. In *2014 IEEE International Symposium on Information Theory*, pages 681–685. IEEE, 2014.

[20] Ankit Singh Rawat, Zhao Song, Alexandros G Dimakis, and Anna Gál. Batch codes through dense graphs without short cycles. *IEEE Transactions on Information Theory*, 62(4):1592–1604, 2016.

[21] Vitaly Skachek. Batch and pir codes and their connections to locally repairable codes. In *Network Coding and Subspace Designs*, pages 427–442. Springer, 2018.

[22] Henning Stichtenoth. *Algebraic function fields and codes*, volume 254. Springer Science & Business Media, 2009.

[23] Itzhak Tamo and Alexander Barg. Bounds on locally recoverable codes with multiple recovering sets. In *2014 IEEE International Symposium on Information Theory*, pages 691–695. IEEE, 2014.

[24] Itzhak Tamo, Alexander Barg, and Alexey Frolov. Bounds on the parameters of locally recoverable codes. *IEEE Transactions on Information Theory*, 62(6):3070–3083, 2016.

[25] Anyu Wang and Zhifang Zhang. Repair locality with multiple erasure tolerance. *IEEE Transactions on Information Theory*, 60(11):6979–6987, 2014.

[26] David P. Woodruff. *A Quadratic Lower Bound for Three-Query Linear Locally Decodable Codes over Any Field*, pages 766–779. Springer Berlin Heidelberg, Berlin, Heidelberg, 2010.

[27] Mary Wootters. Linear codes with disjoint repair groups. Not intended for publication, available at https://sites.google.com/site/marywootters/disjoint_repair_groups.pdf, 2016.

[28] Liyasi Wu. Revisiting the multiplicity codes: A new class of high-rate locally correctable codes. In *2015 53rd Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pages 509–513. IEEE, 2015.