Locally Testable Cyclic Codes

László Babai, Amir Shpilka, and Daniel Štefankovič

***Abstract*—Cyclic linear codes of block length n over a finite field $\mathbb{F}_q$ are linear subspaces of $\mathbb{F}_q^n$ that are invariant under a cyclic shift of their coordinates. A family of codes is *good* if all the codes in the family have constant rate and constant normalized distance (distance divided by block length). It is a long-standing open problem whether there exists a good family of cyclic linear codes.**

A code C is r-*testable* if there exists a randomized algorithm which, given a word $x \in \mathbb{F}_q^n$, adaptively selects r positions, checks the entries of x in the selected positions, and makes a decision (accept or reject x) based on the positions selected and the numbers found, such that

i) if $x \in C$ then x is surely accepted;
ii) if $\mathrm{dist}(x, C) \geq \epsilon n$ then x is probably rejected. ("dist" refers to Hamming distance.)

A family of codes is *locally testable* if all members of the family are r-testable for some constant r. This concept arose from holographic proofs/PCP's. Recently it was asked whether there exist good, locally testable families of codes.

In this paper the intersection of the two questions stated is addressed.

Theorem. There are no good, locally testable families of cyclic codes over any (fixed) finite field.

In fact the result is stronger in that it replaces condition ii) of local testability by the condition

ii') if $\mathrm{dist}(x, C) \geq \epsilon n$ then x has a positive chance of being rejected.

The proof involves methods from Galois theory, cyclotomy, and diophantine approximation.

***Index Terms*—Coding theory, cyclic codes, locally testable codes.**

I. Introduction

All codes discussed in this paper are **linear.** We study rate/distance tradeoffs for locally testable cyclic codes. A family of codes is *good* if the codes in the family have constant rate and constant normalized distance.

An example of a class of locally testable cyclic codes is the Hadamard codes [11]. However, these codes have logarithmic information rate and are therefore far from being good.

It is a classical open question whether there exist good families of cyclic codes (see [21, p. 270, Research Problem (9.2)]). Extending results by Berman [10], we prove the nonexistence of good families of cyclic codes for a large class of block lengths, including block lengths with a large "smooth" divisor (a smooth number has no large prime divisors) (Theorems 1.9 and 5.1).

Recently, Goldreich and Sudan asked whether there exist good locally testable codes [17]. Our main result is that locally testable *cyclic* codes cannot be good. In view of the general tradeoffs we establish for cyclic codes, the critical remaining case is when the block length has a large prime divisor. In this case, we use diophantine approximation to establish a tradeoff involving the testability parameter. Our results require only a considerably weaker property than local testability: words that are far from the code are only required to have a *positive* chance of rejection (the chance may not be bounded away from zero). We formalize this concept and state the main results in Section I-B.

A. Preliminaries

Throughout this paper we shall use the following notation. Let p be a prime and $\ell \geq 1$. Let $q = p^\ell$, the order of the field $\mathbb{F}_q$. A linear code C of length ("block length") n over $\mathbb{F}_q$ is a subspace $C \leq \mathbb{F}_q^n$. The dimension $\dim(C)$ is referred to as the number of information bits. The ratio $\mathrm{rate}(C) := \dim(C)/n$ is the *rate* of C. We say that a family of codes $C_i \leq \mathbb{F}_q^{n_i}$ has *constant rate* if $\mathrm{rate}(C_i) = \Omega(1)$, i.e., if $\dim(C_i) = \Omega(n_i)$. The *weight* $\mathrm{wt}(w)$ of a "word" $w \in \mathbb{F}_q^n$ is the number of nonzero entries of w. The *distance* of the code C is

$$\mathrm{dist}(C) = \min_{w \in C, w \neq 0} \mathrm{wt}(w),$$

the minimum weight of nonzero codewords. The *normalized distance* of C is the quotient $\mathrm{dist}(C)/n$.

We say that $\{C_i\}$ is a family of *good codes* if both the rates and the normalized distances of the C_i are bounded away from zero (i.e., these quantities are $\Omega(1)$).

A code is *cyclic* if it is invariant under the cyclic shift of the coordinates, i.e.,

$$(a_0, \ldots, a_{n-1}) \in C \Leftrightarrow (a_{n-1}, a_0, \ldots, a_{n-2}) \in C.$$

Cyclic codes have a voluminous literature; several of the well-known families of classical codes are cyclic (Bose–Chaudhuri–Hocquenghem (BCH) codes, Reed–Muller codes).

B. Weakly Locally Testable Codes: The Main Results

For $u, v \in \mathbb{F}_q^n$, we use $u \cdot v$ to denote $\sum_i u_i v_i \in \mathbb{F}_q$ (inner product over $\mathbb{F}_q$). For $C \leq \mathbb{F}_q^n$, the dual subspace is

$$C^\perp = \left\{ x \in \mathbb{F}_q^n \mid (\forall\, w \in C)(x \cdot w = 0) \right\}.$$

Definition 1.1: An *r-tester* for a linear code on input x randomly selects r positions $i_1, \dots, i_r$ $(1 \le i_j \le n)$ and checks the corresponding entries in x. (i_j depends only on the pairs (i_k, x_{i_k}), $1 \le k \le j-1$, and on the random bits.) Then it chooses a Boolean function $f : \mathbb{F}_q^r \to \{0,1\}$ at random from a probability distribution that may depend on the sequence (i_k, x_{i_k}), and *accepts* x if $f(x_{i_1}, \dots, x_{i_r}) = 1$.

i) An r-tester is *complete* if it surely accepts each $w \in C$.
ii) An r-tester is *weakly sound* if for all $w \in \mathbb{F}_q^n$, if w is surely accepted then $\mathrm{dist}(w, C) < \mathrm{dist}(C)/3$.

A linear code C is *weakly* r-testable if C has a complete and weakly sound r-tester.

Remark 1.2: As Ben Sasson *et al.* [8] point out, in the case of linear codes we may assume that the test functions f are linear, i.e., $f(x) = 1$ $(x \in \mathbb{F}_q^r)$ iff $x \cdot \alpha = 0$ for a given vector $\alpha = \alpha(f) \in \mathbb{F}_q^r$. We shall not use this observation.

Definition 1.3: A family of codes is weakly locally testable if for some constant r, all codes in the family are weakly r-testable.

Remark 1.4: This is clearly a weaker condition than local testability: we do not require that words distant from C be rejected with constant probability, only with positive probability.

Goldreich and Sudan ask whether there exist good, locally testable codes [17]. We give a partial answer to this question; our result may support the view that some of the complications [17] go through (repeated concatenation steps, the use of probabilistic checkable proofs (PCPs) to build their (nearly linear) locally testable codes may be inevitable.

Theorem 1.5: Let $\mathbb{F}_q$ be a finite field. There are no good, weakly locally testable cyclic codes over $\mathbb{F}_q$.

This result is an immediate consequence of the following tradeoff. Recall that for a *good code*, both $\dim(C)$ and $\mathrm{dist}(C)$ must be $\Omega(n)$.

Theorem 1.6: Let q be a prime power. There is a constant $c = c(q)$ such that for any weakly r-testable cyclic code C, of length n, over $\mathbb{F}_q$ either

- $\dim(C) \le cn/(\log n \log\log n)^{\frac{1}{2(r-1)}}$ or
- $\mathrm{dist}(C) \le cn/(\log n \log\log n)^{1/2}$.

Under a widely accepted number-theoretic hypothesis (about "Wieferich primes") we obtain a better tradeoff for binary codes.

Conjecture 1.7: For all primes p

$$2^{p-1} \not\equiv 1 (\mathrm{mod}\ p^3).$$

This conjecture has been verified for all primes $\le 10^{12}$.

Theorem 1.8: Let C be a weakly r-testable binary cyclic code. If Conjecture 1.7 holds then either

- $\dim(C) \le cn/(\log n)^{\frac{1}{r-1}}$; or
- $\mathrm{dist}(C) \le cn/\log n$

where c is an absolute constant.

C. Cyclic Codes

Cyclic codes were first defined by Prange [22] in 1957. Since then many families of cyclic codes have been found and bounds on the rate and distance of cyclic codes were proved. In particular several of the well-known classical codes are cyclic (BCH codes, Reed-Muller codes). The monograph [21] is a good source on cyclic codes; the question of existence of good families of cyclic codes was also formalized there ([21, p. 270)]. The problem is still open; very little seems to have happened in this area since an important 1967 paper by S. D. Berman [10]. Berman deals with the semisimple case (when p and n are relatively prime) and assumes that all primes dividing n are bounded. We extend Berman's result in several directions: we drop the semisimplicity assumption; permit the primes to grow slowly while obtaining explicit tradeoffs; and allow n to have a large nonsmooth divisor m (up to $m \le n^{1/2-\varepsilon}$). However, in spite of these extensions, the basic ingredients from cyclotomy in our proof are not significantly different from Berman's.

The following is our main tradeoff for cyclic codes (without testability assumption).

Theorem 1.9: Let $B \ge 2$. Let $q = p^\ell$ be a prime power. Let $n = p_1^{\alpha_1} \dots p_k^{\alpha_k} m$ where $p_i \le B$. Let C be a cyclic code of length n over $\mathbb{F}_q$. Then

$$\mathrm{dist}(C)\dim(C) \le \ell \cdot n^{1+1.1/\ln\ln n} m^2 p^{B^2/\ln B}.$$

Note that $p_i = p$ is permitted in this result.

For bounded primes, our proof yields the following result.

Theorem 1.10: Let $p_1, \dots, p_k$ be fixed primes (not necessarily different from p). For every ρ, $0 < \rho < 1$, there exist c_ρ such that if C is a cyclic code, over $\mathbb{F}_q$, of length $n = m \cdot \prod_{i=1}^k p_i{}^{\alpha_i}$ with $\dim(C) \ge \rho n$ then $\mathrm{dist}(C) \le c_\rho m^2$.

Berman's result is the case $m = 1$, $p_i \ne p$. Theorems 1.9 and 1.10 will follow from Theorem 5.1 (Section V).

Another by-product of our proof is the following explicit tradeoff related to the powers of p dividing n.

Theorem 1.11: Let $q = p^\ell$ be a prime power. Let $n = p^s m$. Let C be a cyclic code of length n over $\mathbb{F}_q$. Then

$$\dim(C) \cdot \mathrm{dist}(C) \le pmn.$$

We give the one-paragraph proof in Section III. The following corollary to this result (with a rather more complicated proof) appeared in 1991 in Castagnioli *et al.* [12].

Corollary 1.12 [12, Theorem 3]: Let $\{C_i\}$ be a family of codes over $\mathbb{F}_q$ $(q = p^\ell)$. Let C_i have block length $n_i = p^{s_i}$. If the s_i are unbounded then the family $\{C_i\}$ is not good.

Proof: By Theorem 1.11

$$\dim(C_i) \cdot \mathrm{dist}(C_i) \le pm_i n_i = n_i^2/p^{s_i-1} \ne \Omega\left(n_i^2\right)$$

as would be the case for a good family. □

Remark 1.13: In view of these results, the search for good cyclic codes should focus on prime block length. Mersenne primes n would seem like prime candidates over $\mathbb{F}_2$ because for them $x^n - 1$ has the largest number of factors.

D. Locally Testable Codes

A code C is *r-testable* if there exists a randomized algorithm which, given a word $x \in \mathbb{F}_q^n$, adaptively selects r positions, checks the entries of x in the selected positions, and makes a decision (accept or reject x) based on the numbers found on the positions selected, such that

i) if $x \in C$ then x is surely accepted;
ii) if $\operatorname{dist}(x, C) \geq \epsilon n$ then x is rejected with a constant probability ("dist" refers to Hamming distance).

A family of codes is *locally testable* if there exists a constant r such that all members of the family are r-testable.

r-testable codes for very small values of r first arose from the analysis of the codes used as the "exterior hulls" of holographic proofs. Thus, as pointed out in [17], suitable (nontrivial) modification of [5] yields codes that are polylog-testable and have quasi-polynomial length as a function of the number b of information bits. The length of the code was reduced to nearly linear ($n = b^{1+\epsilon}$) in [6] while retaining its polylog-testability. ([6] has a correctable error; the proof as stated there yields nearly quadratic length.) The seminal PCP paper by Arora *et al.* [2] can be adapted to yield constant-testable codes of polynomial length.

The testing length was reduced to 3 bits by Blum, Luby, and Rubinfeld [11] at the cost of exponential length (Hadamard code).

Friedl and Sudan [15] were the first to formally define locally testable codes. In [15], they also gave a family of locally testable codes of nearly quadratic length over a large alphabet.

The code of [15] is constructed as follows. The messages are polynomials of degree d in m variables over the field $\mathbb{F}_q$; both m and q bounded by $d^{O(1)}$. A codeword is the set of restrictions of the message to each affine line in $\mathbb{F}_q^m$, that is, for every affine line in $\mathbb{F}_q^m$, we have a coordinate in the codeword, in which we write the degree d polynomial which is the restriction of the message to that line. So this code uses an alphabet of size q^{d+1} (one "letter" for each univariate polynomial of degree $\leq d$ over $\mathbb{F}_q$).

Starting from the codes in [15], Goldreich and Sudan [17] give two constructions of locally testable *binary* codes. In the first construction they first restrict the code of [15] a random subset of coordinates (which corresponds to taking a random subset of affine lines); then they concatenate the code with a Reed–Muller-type code, and finally, concatenate it with the Hadamard code. In this construction, the length is $n = b^{1+\epsilon}$ (any $\epsilon > 0$ can be achieved). In the second construction, Goldreich and Sudan again restrict the code of [15] a random subset of coordinates, but now instead of concatenating the code, they reduce the alphabet size by using PCPs. This construction gives a locally testable code of nearly linear length ($n = b^{1+o(1)}$). Both constructions are randomized. These constructions were derandomized by Ben Sasson *et al.* [9].

Goldreich and Sudan [17] also note, that, while the codes constituting the outer layer of known holographic proofs/probabilistically checkable proofs are not in themselves locally testable, they can be modified (nontrivially) to yield locally testable codes. Finally, Goldreich and Sudan raise the problem we partially address here: do *good* locally testable codes exist?

In a recent paper, Ben Sasson *et al.* [8] study locally testable codes with two queries. They show that any *linear* locally testable code with two queries, over a finite alphabet, that has a linear distance, contains at most a *constant* number of codewords. This result also holds for nonlinear binary codes. However, if the alphabet size is larger than 2 then there exist nonlinear locally testable codes with two queries of nearly linear length.

Local decodability, a strengthening of the concept of local testability, has been studied in [16], [18], [14], [19]. Stronger tradeoffs than ours are obtained by these authors under this stronger assumption; as shown in [17], such strong tradeoffs do not hold for locally testable codes.

In addition to PCPs, local testability and local decodability arise in several contexts in computational complexity and cryptography, including self-testing/correcting computations, pseudorandom generators, private information retrieval, fault-tolerant data storage. For further literature on these connections we refer to the bibliographies of [18], [14].

E. Organization of the Paper

In Section II, we give the basic tools for dealing with cyclic codes. In Section III, we prove the easy case where n is divisible by a high power of the characteristic of the field. This proof gives some intuition for the later proofs. In Section IV, we give some information about cyclotomic polynomials over finite fields that we will use in Section V for proving our main result on cyclic codes (without local testability assumptions). In Section VI, we show how to improve our tradeoff for binary codes under a number-theoretic conjecture. The main contributions of the paper follow in Sections VII and VIII where we use the local testability assumption through a diophantine approximation argument to conclude the proof of our main tradeoff result.

II. Cyclic Codes and Polynomials

In this section, we review the polynomial ring machinery used to handle cyclic codes. For more details we refer to [21].

Let $q = p^\ell$ be a prime power. Let $R = \mathbb{F}_q[x]/(x^n - 1)$. It is known that R is a principal ideal domain. In other words, every ideal I in R is generated by some element $I = (g)$. With the vector $a = (a_0, \ldots, a_{n-1}) \in \mathbb{F}_q^n$ we associate the polynomials

$$f_a(x) = \sum_{i=0}^{n-1} a_i x^i \in R$$

and

$$f_a^*(x) = \sum_{i=0}^{n-1} a_i x^{n-i-1} \in R.$$

Let $C \leq \mathbb{F}_q^n$ be a cyclic code. Let $I_C = \{f_a : a \in C\}$ and $J_C = \left\{f_a^* : a \in C^\perp\right\}$. The cyclicity of C implies that I_C and J_C are ideals in R. As R is a principal ideal domain, we have that $I_C = (g)$ where $g \mid x^n - 1$; we call g the *generator polynomial* of C. So $a \in C \Leftrightarrow g \mid f_a$. Let $h = (x^n - 1)/g$; then $J_C = (h)$ and $a \in C \Leftrightarrow x^n - 1 \mid (h \cdot f_a)$. We call h the parity-check polynomial for C.

Clearly, $\dim(C) = \deg(h)$. So C has constant rate if and only if $\deg(h) = \Omega(n)$.

For $f \in \mathbb{F}_q[x]$, let $\mathrm{wt}(f)$ be the number of nonzero terms in f. The standard representative of $f \in \mathbb{F}_q[x]$ in $\mathbb{F}_q[x]/(x^n-1)$ is $f \bmod (x^n-1)$.

We shall use the following properties of the weight function.

Observation 2.1: For any $f, g \in \mathbb{F}_q[x]$

- $f(x)^p = f(x^p)$;
- $\mathrm{wt}(f^p) = \mathrm{wt}(f)$;
- $\mathrm{wt}(fg) \le \mathrm{wt}(f)\mathrm{wt}(g)$;
- $\mathrm{wt}(f \bmod (x^n-1)) \le \mathrm{wt}(f)$.

The distance of C can be expressed as follows:

$$\begin{aligned}\mathrm{dist}(C) &= \min\{\mathrm{wt}(f) : f \not\equiv 0,\ hf \equiv 0 \pmod{x^n-1}\}\\ &= \min\{\mathrm{wt}(gf) : f \not\equiv 0 \pmod{x^n-1},\ \deg(f) < \deg(h)\}.\end{aligned}$$

The *testing weight* of C, $\mathrm{test}(C)$, is the smallest r such that there exist polynomials $h_1, \ldots, h_\ell$ such that $\mathrm{wt}(h_i) \le r$ and $h_1, \ldots, h_\ell$ generate the same ideal as h in R. We use $\mathrm{test}(h)$ to denote $\mathrm{test}(C)$.

III. High Powers of p

In this section, we prove Theorem 1.11, the tradeoff for the case when n is divisible by a high power of p. No local testability assumption will be made in this section. The idea is to use the factorization of h, the parity-check polynomial of C, to construct a codeword of small weight. The fact that $n = mp^s$ makes this analysis very easy. A similar argument will appear in later sections, but with more complications.

Proof of Theorem 1.11: Let $h \in \mathbb{F}_q[x]$ be the parity-check polynomial of C. Recall that $h \mid x^n - 1$. As p is the characteristic of the field, we have

$$h \mid x^n - 1 = (x^m - 1)^{p^s}.$$

Let k be the smallest integer such that $h \mid (x^m-1)^k$. Note that $\dim(C) = \deg(h) \le mk$. Let $f_1 \in \mathbb{F}_q[x]$ be a factor of $x^m - 1$ such that f_1^k divides h. Let $j = \lfloor \log_p k \rfloor$. Note that

$$f = \frac{x^n-1}{f_1^{p^j}} = (x^m-1)^{(p-1)(p^{s-1}+\cdots+p^j)} \left(\frac{x^m-1}{f_1}\right)^{p^j}$$

is a codeword as $(x^n-1) \mid f \cdot h$. Moreover, we can write $f(x) = f'(x^{p^j})$, for some polynomial $f' \in \mathbb{F}_q[x]$, so

$$\begin{aligned}\mathrm{dist}(C) &\le \mathrm{wt}(f) = \mathrm{wt}(f') \le n/p^j\\ &= p^{s-j}m \le p^{s+1}m/k = pn/k.\end{aligned}$$

Combining our two inequalities we obtain $\dim(C)\mathrm{dist}(C) \le mk \cdot pn/k = pmn$. □

In order to prove a more general result we need to consider the factorization of $x^n - 1$ to cyclotomic polynomials, and their irreducible factors.

IV. Cyclotomic Polynomials

In this section, we collect basic facts about cyclotomic polynomials over finite fields. We refer to [20] for proofs and further information. We continue to work over the field $\mathbb{F}_q$, $q = p^\ell$.

Let $\widetilde{\Phi}_n$ be the nth cyclotomic polynomial over the rationals. These polynomials can be calculated inductively from the identity $x^n - 1 = \prod_{d|n} \widetilde{\Phi}_d$. The degree of $\widetilde{\Phi}_n$ is $\varphi(n)$,[1] and $\widetilde{\Phi}_n$ is a polynomial with integer coefficients; let Φ_n be the same polynomial $\bmod\ p$. Then Φ_n is the nth cyclotomic polynomial over $\mathbb{F}_q$, and $x^n - 1 = \prod_{d|n} \Phi_d$ (equality $\bmod\ p$). For $p \nmid d$, the roots of Φ_d are the primitive dth roots of unity in the algebraic closure of $\mathbb{F}_q$. For $d = m \cdot p^s$, where $p \nmid m$, $\Phi_d = (\Phi_m)^{p^{s-1}(p-1)}$.

Lemma 4.1: Let m_1, m_2 be integers such that every prime dividing m_1 divides m_2. Then

$$\Phi_{m_1 m_2}(x) = \Phi_{m_2}(x^{m_1}).$$

Proof: The proof is a repeated application of Exercise 2.57(b) from [20], where this fact is stated for the case when m_1 is a prime. □

For n coprime to a let $\mathrm{ord}_n(a)$ denote the order of a in $\mathbb{Z}_n^\times$, the multiplicative group of the ring $\mathbb{Z}/n\mathbb{Z}$ (the integers modulo n). Note that for $q = p^\ell$ we have

$$\mathrm{ord}_n(q) = \mathrm{ord}_n(p)/\gcd(\mathrm{ord}_n(p), \ell) = \mathrm{lcm}(\mathrm{ord}_n(p), \ell)/\ell$$

($\mathrm{lcm}(a,b)$ is the least common multiple of a and b).

Lemma 4.2: ([20, p. 65]) For n coprime to q, all the irreducible factors of Φ_n over $\mathbb{F}_q$ have degree $\mathrm{ord}_n(q)$.

Let v be a prime different than p. When v is odd we define $t_{v,p}$ to be the smallest exponent such that $p^{v-1} \not\equiv 1 \pmod{v^{t_{v,p}+1}}$. When $v = 2$ we define $t_{v,p}$ to be the exponent satisfying $p^2 = 1 + 2^t \pmod{2^{t+1}}$. Note that in both cases $1 \le t_{v,p} \le v/\log_p v$.

We will use $t_{v,p}$ to determine the order of p modulo v^k for large values of k.

Lemma 4.3: Let p, v be different primes. For $k \le t_{v,p}$ we have

$$\mathrm{ord}_{v^k}(p) = \begin{cases} \mathrm{ord}_v(p), & v \text{ odd} \\ 1 \text{ or } 2, & v = 2. \end{cases}$$

For $k \ge t_{v,p}$

$$\mathrm{ord}_{v^k}(p) = \begin{cases} v^{k-t_{v,p}} \cdot \mathrm{ord}_v(p), & v \text{ odd} \\ v^{k-t_{v,p}} \cdot 2, & v = 2. \end{cases}$$

Proof: We begin with the first claim. Assume that $v \ne 2$. Let $t = t_{v,p}$ and let $g = \mathrm{ord}_v(p)$. Clearly, $g \mid v - 1$. Let e be such that $p^g = 1 + av^e \pmod{v^{e+1}}$, where a is not divisible by v. We will show that $e = t$, which implies that for every $k \le t$

$$p^g = 1 + av^t \pmod{v^{t+1}} = 1 \pmod{v^k}.$$

Thus, $\mathrm{ord}_{v^k}(p) \le g = \mathrm{ord}_v(p)$. As $\mathrm{ord}_v(p) \le \mathrm{ord}_{v^k}(p)$ the claim follows. Indeed, let $p^g = 1 + av^e \pmod{v^{e+1}}$. We have that

$$\begin{aligned}p^{v-1} &= (p^g)^{\frac{v-1}{g}} = (1+av^e)^{\frac{v-1}{g}} \pmod{v^{e+1}}\\ &= 1 + \frac{v-1}{g} av^e \pmod{v^{e+1}}.\end{aligned}$$

As $\frac{v-1}{g}a$ is not divisible by v we get that $t = e$.

[1] $\varphi(n) = \#\{m \le n \mid \gcd(m,n) = 1\}$.

Assume now that $v = 2$. By definition of $t_{2,p}$ we have that $p^2 = 1 + 2^t \pmod{2^{t+1}}$. Thus, $\mathrm{ord}_{2^k}(p) \le 2$, and the claim follows.

We now prove the claim for the case $k \ge t_{v,p}$. Let $t = t_{v,p}$, and denote $k = t + s$. We prove the claim for k by induction on s. From the proof of the first claim and the definition of $t_{2,p}$ we get that

$$p^{\mathrm{ord}_{v^t}(p)} = 1 + a \cdot v^t \pmod{v^{t+1}}$$

where a is not divisible by v. This proves the claim for the case $s = 0$. For the inductive step we need the following claim.

Claim 4.4: Let $m \equiv 1 + av^r \pmod{v^{r+1}}$ where a is not divisible by v. Assume that $r \ge 1$ for odd v and $r \ge 2$ for $v = 2$. Then $m^v \equiv 1 + av^{r+1} \pmod{v^{r+2}}$.

Proof: We have $m = 1 + av^r + bv^{r+1}$ for some integer b. Hence,

$$m^v = \sum_{i=0}^{v} \binom{v}{i} (av^r + bv^{r+1})^i.$$

When $r \ge 2$, we have that $2r \ge r + 2$ and, hence, only the first two terms in the sum can be nonzero modulo v^{r+2}. When $r = 1$ and $v \ge 3$ we have $3r \ge r + 2$ and, hence, only the first three terms in the sum can be nonzero modulo v^{r+2}; moreover, the third term is divisible by $\binom{v}{2} v^{2r}$ and, hence, is zero modulo v^{r+2}. Thus, only the first two summands are nonzero $\bmod\ v^{r+2}$ so we get that the sum is equivalent to

$$\begin{aligned} m^v &\equiv 1 + v(av^r + bv^{r+1}) \pmod{v^{r+2}} \\ &\equiv 1 + av^{r+1} \pmod{v^{r+2}}. \end{aligned}$$

□

Thus, by induction on s we obtain that

$$p^{\mathrm{ord}_{v^t}(p) \cdot v^s} \equiv 1 + av^{t+s} \pmod{v^{t+s+1}}$$

for every $s \ge 0$. Hence, the order of p in $\mathbb{Z}^{\times}_{v^{t+s+1}}$ divides $\mathrm{ord}_{v^t}(p) \cdot v^{s+1}$, is divisible by the $\mathrm{ord}_{v^t}(p) \cdot v^s$ (because of the induction hypothesis), and is not $\mathrm{ord}_{v^t}(p) \cdot v^s$. Hence, the claim follows. □

We are now ready to prove the main lemma of this section. Let $m_1 = p_1^{\gamma_1} \cdots p_k^{\gamma_k}$. The following lemma shows that the irreducible factors of Φ_{m_1} over the field $\mathbb{F}_q$ are exactly the irreducible factors of Φ_{m_2} evaluated at x^{m_1/m_2}, for $m_2 = p_1^{\delta_1} \cdots p_k^{\delta_k}$, where δ_i is determined by q, γ_i, and $p_1, \ldots, p_k$.

Lemma 4.5: Let $p, p_1, \ldots, p_k$ be different primes where $q = p^\ell$. Let c_i be the largest integer such that $p_i^{c_i}$ divides ℓ. Let c_i' be the largest integer t such that there exists $j \ne i$ for which p_i^t divides $\mathrm{ord}_{p_j}(p)$. Let $b_i = \max(c_i, c_i')$. Let $m_1 = p_1^{\gamma_1} \cdots p_k^{\gamma_k}$. Let $m_2 = p_1^{\delta_1} \cdots p_k^{\delta_k}$ where $\delta_i = \min\{\gamma_i, t_{p_i,p} + b_i\}$. Let $f_1, \ldots, f_d$ be the irreducible factors of Φ_{m_2} over $\mathbb{F}_q$. Then the irreducible factors of Φ_{m_1} over $\mathbb{F}_q$ are $f_1(x^{m_1/m_2}), \ldots, f_d(x^{m_1/m_2})$.

Proof: The idea of the proof is simple. First, we use Lemma 4.1 to obtain

$$\Phi_{m_1}(x) = \Phi_{m_2}(x^{m_1/m_2}).$$

This implies that $f_i(x^{m_1/m_2})$ is a factor of $\Phi_{m_1}(x)$. Now we show that $\deg(f_i(x^{m_1/m_2})) = \mathrm{ord}_{m_1}(q)$ and using Lemma 4.2 we get that it is an irreducible factor.

Note that $p_i^{b_i}$ divides $\mathrm{ord}_{p_i{}^{t_{p_i,p}+b_i}}(p)$ and $p_i^{b_i+1}$ does not divide $\mathrm{ord}_{p_j^B}(p)$ for any $j \ne i$ and any B (follows from Lemma 4.3). Moreover, $p_i^{b_i+1}$ does not divide ℓ. Using Lemma 4.3 we obtain that

$$\begin{aligned} &\mathrm{ord}_{m_1}(p^\ell) \\ &= \mathrm{lcm}(\ell, \mathrm{ord}_{p_1{}^{\gamma_1}}(p), \ldots, \mathrm{ord}_{p_k{}^{\gamma_k}}(p))/\ell \\ &= \mathrm{lcm}(\ell, p_1^{\gamma_1-\delta_1}\mathrm{ord}_{p_1{}^{\delta_1}}(p), \ldots, p_k^{\gamma_k-\delta_k}\mathrm{ord}_{p_k{}^{\delta_k}}(p))/\ell \\ &=^{(*)} p_1^{\gamma_1-\delta_1} \ldots p_k^{\gamma_k-\delta_k}\mathrm{lcm}(\ell, \mathrm{ord}_{p_1{}^{\delta_1}}(p), \ldots, \mathrm{ord}_{p_k{}^{\delta_k}}(p))/\ell \\ &= (m_1/m_2)\mathrm{ord}_{m_2}(p^\ell). \end{aligned}$$

Where equality $(*)$ holds because the δ_i's were defined in such a way that if $\delta_i < \gamma_i$ then the power of p_i that divides each $\mathrm{ord}_{p_j^{\delta_j}}(p)$ is smaller or equal to the power of p_i that divides $\mathrm{ord}_{p_i^{\delta_i}}(p)$. By Lemma 4.2, we get that for every i, $\mathrm{ord}_{m_2}(p^\ell) = \deg(f_i(x))$. From this we see that for every i, $f_i(x^{m_1/m_2})$ is a factor of minimal degree of $\Phi_{m_1}(x)$, and hence it is irreducible. This completes the proof of Lemma 4.5. □

Lemma 4.6: Let m_2 be as defined in Lemma 4.5. Then

$$m_2 \le \ell \cdot (p_1 - 1) \cdot \cdots \cdot (p_k - 1) \prod_{i=1}^{k} \min\left\{p_i^{t_{p_i,p}}, p_i^{\alpha_i}\right\}.$$

Proof: From the definition of the b_i it follows that each $p_i^{b_i}$ divides $\ell(p_1 - 1) \ldots (p_k - 1)$. As the p_i are distinct, we obtain that

$$m_2 \le \ell \cdot (p_1 - 1) \cdot \cdots \cdot (p_k - 1) \prod_{i=1}^{k} \min\left\{p_i^{t_{p_i,p}}, p_i^{\alpha_i}\right\}.$$

□

V. Small Primes

In this section, we settle the case when a large part of n factors into powers of very small primes. No local testability assumption will be made in this section.

Notation: Let n be a positive integer. The sum of divisors of n smaller than A will be denoted $\sigma(n, A) = \sum_{d|n, d<A} d$. Let $0 < \rho < 1$. Let $A(n, \rho)$ be the maximal A such that $\sigma(n, A) < \rho n$.

We now prove a generalization of both Theorems 1.9 and 1.10. The theorems will follow from estimates on $A(n, \rho)$ and the general theorem. In Section VI, we improve Theorem 5.1 for binary codes under a plausible number theoretic conjecture.

Theorem 5.1: Let $q = p^\ell$ be a prime power. Let $n = p^s p_1^{\alpha_1} \cdots p_k^{\alpha_k}$ where $p, p_1, \ldots, p_k$ are different primes. Let C be a cyclic code of length n over $\mathbb{F}_q$. Let $\dim(C) = \rho n$. Then

$$\mathrm{dist}(C) \le p\ell \frac{n}{A(n, \rho)} \prod_{i=1}^{k} (p_i - 1) p_i^{\beta_i}$$

where $\beta_i = \min\{t_{p_i,p}, \alpha_i\}$.

Proof: We first sketch the idea of the proof. Let h be the parity-check polynomial of C. We shall show that h shares a factor g_{n_1} with a cyclotomic polynomial Φ_{n_1} where n_1 has a large smooth divisor. It follows by Lemma 4.5 that g_{n_1} can be written as a polynomial of a large power of x. This in turn implies, as in the proof of Theorem 1.11, that the polynomial $(x^n-1)/g_{n_1}$ represents a codeword of small weight.

We now begin the formal proof. Let h be the parity-check polynomial of the code C. Recall that $h \mid x^n-1=\prod_{d|n}\Phi_d$. Let $\rho=\deg(h)/n$ be the rate of the code. For d not divisible by p let $g_d=\gcd(h,\Phi_d)$. For $d=p^j e$ $(1\le j\le s)$, where e is not divisible by p, we define g_d recursively on j

$$g_d=\gcd\left(\frac{h}{\prod_{i=0}^{j-1} g_{p^i e}},\Phi_d\right).$$

Thus, g_d is the part that Φ_d "contributes" to h. This is made precise in the next claim.

Claim 5.2:

a) $\prod_{d|n} g_d=h$.

b) If $f^t \mid \Phi_{ep^j}$ and $f \mid g_{ep^{j+1}}$ then $f^t \mid g_{ep^j}$.

Proof: Let e be a divisor of n which is not divisible by p. We have that

$$\gcd\left(h,\prod_{j=0}^{s}\Phi_{p^j e}\right)=\prod_{j=0}^{s} g_{p^j e}.$$

As for any two integers $e\ne e'$, which are not divisible by p, we have that

$$\gcd\left(\prod_{i=0}^{s}\Phi_{p^i e},\prod_{i=0}^{s}\Phi_{p^i e'}\right)=1.$$

We get that

$$h=\gcd(h,x^n-1)=\gcd\left(h,\prod_{d|n}\Phi_d\right)$$
$$=\prod_{e|n,p\nmid e}\gcd\left(h,\prod_{i=0}^{s}\Phi_{p^i e}\right)=\prod_{e|n,p\nmid e}\prod_{i=0}^{s} g_{p^i e}=\prod_{d|n} g_d.$$

The second claim follows from the observation that if $r<t$ is the maximal power of f such that $f^r \mid g_{p^j e}$, then $\frac{h}{\prod_{i=0}^{j} g_{p^i e}}$ is not divisible by f, as otherwise f^{r+1} would be a common factor of $\frac{h}{\prod_{i=0}^{j-1} g_{p^i e}}$ and of $\Phi_{p^j e}$ and, hence, a factor of $g_{p^j e}$. This implies that f is not a factor of $g_{p^{j+1}e}$ in contradiction. ☐

From the fact $\prod_{d|n} g_d=h$ we infer $\sum_{d|n}\deg(g_d)=\rho n$. We have $\deg(g_d)\le\varphi(d)\le d$. Therefore,

$$\rho n=\sum_{d|n}\deg(g_d)\le\sum_{deg(g_d)\ne 0} d$$

and hence there is $n_1\ge A(n,\rho)$ such that $n_1 \mid n$ and $\deg(g_{n_1})>0$.

This completes the first step of the proof as outlined. Let m_1 be the part of n_1 not divisible by p, i.e., $n_1=p^{s_1}m_1$. Let g be an irreducible factor of g_{n_1}.

Claim 5.3: $g^{p^{s_1-1}} \mid g_{n_1}$.

Proof: As $g_{n_1} \mid \Phi_{n_1}=(\Phi_{m_1})^{p^{s_1-1}(p-1)}$ and g is an irreducible factor of g_{n_1}, we get that g is an irreducible factor of Φ_{m_1}. Therefore, for every $1\le j$ we have that $g^{p^{j-1}(p-1)} \mid \Phi_{m_1p^j}$. As $g \mid g_{m_1p^{s_1}}$ we get, by Observation 5.2, that for every $1\le j<s_1$, $g^{p^{j-1}(p-1)} \mid g_{m_1p^j}$. Since $\prod_{j=0}^{s_1} g_{m_1p^j} \mid h$ we get that

$$g^{2+\sum_{j=1}^{s_1-1}p^{j-1}(p-1)} \Big| h$$

where the 2 in the exponent comes from the contributions of g_{m_1} and g_{n_1}. Thus, for $Z=p^{s_1-1}$ we have that $g^Z \mid h$. ☐

As $g^Z \mid h$, we get that $f=(x^n-1)/g^Z$ is a codeword. We will show that f has a small weight. Since Z is a power of p we have that $f=((x^{n/Z}-1)/g)^Z$. From Observation 2.1 it follows that $\mathrm{wt}(f)=\mathrm{wt}((x^{n/Z}-1)/g)$. We now show that g can be written as a polynomial in x^r for some large r that divides n/Z.

Recall that m_1 is the part of n_1 which is not divisible by p, and that g is a factor of Φ_{m_1}. Denote

$$m_1=\prod_{i=1}^{k} p_i^{\gamma_i}.$$

Let $b_1,\dots,b_k,\delta_1,\dots,\delta_k,m_2$ be defined as in Lemma 4.5. We obtain that the irreducible factors of Φ_{m_1} are $f_1(x^{m_1/m_2}),\dots,f_d(x^{m_1/m_2})$ where $f_1,\dots,f_d$ are the irreducible factors of Φ_{m_2}. Hence, g has all exponents divisible by m_1/m_2. Since $\frac{m_1}{m_2} \mid \frac{n}{Z}$ we can view both $x^{n/Z}-1$ and g as polynomials in x^{m_1/m_2} and conclude that

$$\mathrm{wt}(f)=\mathrm{wt}((x^{n/Z}-1)/g)\le\frac{n/Z}{m_1/m_2}=\frac{pnm_2}{n_1}.$$

From Lemma 4.6 we get that

$$m_2\le\ell\cdot(p_1-1)\cdot\dots\cdot(p_k-1)\prod_{i=1}^{k}\min\left\{p_i^{t_{p_i,p}},p_i^{\alpha_i}\right\}.$$

Hence,

$$\mathrm{wt}(f)\le p\cdot\ell\frac{n}{A(n,\rho)}\prod_{i=1}^{k}(p_i-1)\cdot\min\left\{p_i^{t_{p_i,p}},p_i^{\alpha_i}\right\}.$$

This completes the proof of the Theorem 5.1. ☐

We now show how to prove Theorems 1.9, 1.10.

A. *Proof of Theorem 1.9*

We use the following estimate on $A(n,\rho)$.

Lemma 5.4: For any $0<\rho<1$

$$A(n,\rho)\ge\rho n^{1-1.1/\ln\ln n}.$$

Proof: It is known that the number of divisors of a number $n\ge 3$ is less than $n^{1.1/\ln\ln n}$ (see, e.g., [7, p. 234]), so the claim follows. ☐

Using this estimate we obtain the following result.

Corollary 5.5: Let $q = p^\ell$ be a prime power. Let $n = p^s p_1^{\alpha_1} \cdots p_k^{\alpha_k}$ where $p, p_1, \ldots, p_k$ are different primes. Let C be a cyclic code of length n over $\mathbb{F}_q$. Then

$$\dim(C)\mathrm{dist}(C) \le p\ell n^{1+1.1/\ln\ln n} \prod_{i=1}^{k} (p_i - 1) p_i^{\beta_i} \qquad (1)$$

where $\beta_i = \min\{t_{p_i,p}, \alpha_i\}$.

Recall that the quantity $t_{v,p}$ is defined before Lemma 4.3, where it is noted that $t_{v,p} \le v/\log_p v$. Therefore, $p^{\beta_i} \le \min\{p_i^{\alpha_i}, p^{p_i}\}$. The way to prove the theorem will be to divide the primes p_i to the set of primes smaller than B (for some parameter B), and the primes larger than B, and to estimate the contribution of each to the (1). We shall need the following explicit estimates.

Theorem 5.6 (see [7, p. 233]): For all $B \ge 2$, we have $\sum_{p \le B} \ln p \le 1.000081B$, where the sum ranges over all primes $p \le B$.

Theorem 5.7 (see [7, p. 233]): Let $j \ge 6$. The jth prime p_j satisfies

$$j \ln j < p_j < j(\ln j + \ln\ln j).$$

Corollary 5.8: For all $B \ge 2$ we have

$$\sum_{p \le B} (p + \log_2(p-1)) \le (B^2/\ln B) - 1$$

where the sum ranges over all primes $p \le B$.

Proof: For $B \le 1000$, the corollary is verified by a computer. Now let $B \ge 1000$. We have $\ln\ln B \le 0.28 \ln B$. Let t be the largest integer such that $t \ln t \le B$. Note that $t \le B/(\ln B - \ln\ln B) \le B/(0.72 \ln B)$. Using the monotonicity of $x(\ln x + \ln\ln x)$ for $x > 1$ we obtain

$$\begin{aligned}\sum_{p \le B} p &\le 28 + \sum_{j=6}^{t} j(\ln j + \ln\ln j) \\ &\le 28 + \frac{t(t+1)}{2}(\ln t + \ln\ln t) \\ &\le 28 + \frac{B(t+1)}{2} + \frac{t(t+1)}{2}\ln\ln t \\ &\le \frac{B^2}{\ln B}(0.0002 + 0.70 + 0.28) \\ &= 0.9802\frac{B^2}{\ln B}.\end{aligned}$$

Hence,

$$\begin{aligned}\sum_{p \le B} (p + \log_2(p-1)) &\le 0.9802\frac{B^2}{\ln B} + \frac{1.000081}{\ln 2}B \\ &\le 0.9952\frac{B^2}{\ln B} \le \frac{B^2}{\ln B} - 1. \qquad \square\end{aligned}$$

Proof of Theorem 1.9: Let β_i be defined as in Theorem 5.1. For $p_i \le B$, we use $\beta_i \le t_{p_i,p} \le p_i/\log_p p_i$. From Corollary 5.8 we obtain

$$\prod_{i: p_i \le B} (p_i - 1) p_i^{\beta_i} \le q p^{B^2/\ln B - 1}.$$

For $p_i > B$, we use $\beta_i \le \alpha_i$ and hence,

$$\prod_{i: p_i > B} (p_i - 1) p_i^{\beta_i} \le \prod_{i: p_i > B} p_i^{2\alpha_i} = m^2.$$

Theorem 1.9 now follows from Corollary 5.5. $\square$

As a special case of Theorem 1.9, we obtain following result.

Corollary 5.9: Let q be a fixed prime power. Let $p(n)$ be the largest prime dividing n. For any $\varepsilon > 0$, there exists $c = c(\varepsilon, q) > 0$ such that if $p(n) \le c(\ln n \ln\ln n)^{1/2}$ then for cyclic codes over $\mathbb{F}_q$ of length n

$$\mathrm{dist}(C)\dim(C) \le n^{1+\varepsilon}.$$

B. *Proof of Theorem 1.10*

We shall need the following estimate on $A(n, \rho)$.

Lemma 5.10: Let $n = p_1^{\alpha_1} \cdots p_k^{\alpha_k}$ (p_i prime). Let $0 < \rho < 1$

$$c = (\max_i p_i) \cdot \prod_{i=1}^{k} \frac{p_i}{p_i - 1},$$

and $\delta = \min((3k)^{-4k}, (\rho/c)^2)$. Then $A(n, \rho) \ge \delta n$.

We defer the proof of this bound to the end of this section. By plugging the estimate on $A(n, \rho)$ to Theorem 5.1, we obtain the following corollary.

Corollary 5.11: Let $q = p^\ell$ be a prime power. Let $p_1, \ldots, p_k$ be a fixed set of primes different from p. Let $0 < \rho < 1$ be a constant. Let δ be as in Lemma 5.10. Let C be a cyclic code of length $n = p^s p_1^{\alpha_1} \cdots p_k^{\alpha_k}$ and dimension ρn. Then

$$\mathrm{dist}(C) \le \frac{p}{\delta} \cdot \ell \cdot \prod_{i=1}^{k} p_i \cdot \min\left\{p^{p_i}, p_i^{\alpha_i}\right\}.$$

The explicit estimate for Berman's result (Theorem 1.10) follows from the corollary and the fact that p and all the p_i's are fixed (and so $\prod_{i=1}^{k} p_i \cdot \min\{p^{p_i}, p_i^{\alpha_i}\}$ is also fixed).

We devote the rest of this section to proving Lemma 5.10.

Lemma 5.12: Let $n = p_1^{\alpha_1} \cdots p_k^{\alpha_k}$ (p_i prime). Then for any $0 < \delta < 1$

$$\sigma(n, \delta n) < \left(1 + \log\frac{1}{\delta}\right)^k \cdot \left(\max_i p_i\right) \cdot \prod_{i=1}^{k} \frac{p_i}{p_i - 1} \cdot \delta n.$$

Proof: Let

$$N(n, \delta) = \left\{d : d \mid n \text{ and } \delta n < d \le \delta n \cdot \max_i p_i\right\}.$$

Clearly, if $d \mid n$ and $d \le \delta n$ then there exist $\tilde{d}$ in $N(n, \delta)$ such that $d \mid \tilde{d}$. We thus obtain that

$$\sigma(n, \delta n) = \sum_{d \mid n,\ d < \delta n} d \le \sum_{\tilde{d} \in N(n,\delta)} \sum_{d \mid \tilde{d}} d.$$

Consider some $\tilde{d} \in N(n, \delta n)$. Denote $\tilde{d} = p_1^{\beta_1} \cdots p_k^{\beta_k}$. We get that

$$\sum_{d|\tilde{d}} d = \prod_{i=1}^{k} \left(\sum_{j=0}^{\beta_i} p_i^j \right) < \prod_{i=1}^{k} \frac{p_i^{\beta_i+1}}{p_i - 1}$$

$$= \tilde{d} \cdot \prod_{i=1}^{k} \frac{p_i}{p_i - 1} \leq \delta n \cdot \max_i p_i \cdot \prod_{i=1}^{k} \frac{p_i}{p_i - 1}.$$

Therefore,

$$\sigma(n, \delta n) < |N(n, \delta)| \cdot \delta n \cdot \prod_{i=1}^{k} \frac{p_i}{p_i - 1} \cdot \max_i p_i.$$

Claim 5.13: $|N(n, \delta)| \leq \left(1 + \lfloor \log \frac{1}{\delta} \rfloor\right)^k$.

Proof: Let $d \in N(n, \delta)$. Then $d \mid n$ and $\frac{n}{d} < \frac{1}{\delta}$. Denote $\frac{n}{d} = p_1^{\beta_1} \cdot \cdots \cdot p_k^{\beta_k}$. Clearly

$$\sum_{i=1}^{k} \beta_i \leq \left\lfloor \log \frac{1}{\delta} \right\rfloor.$$

Therefore, the number of such $(\beta_1, \ldots, \beta_k)$ is at most

$$\binom{k + \lfloor \log \frac{1}{\delta} \rfloor}{k} \leq \left(1 + \left\lfloor \log \frac{1}{\delta} \right\rfloor\right)^k. \qquad \square$$

This completes the proof of Lemma 5.12. $\square$

Proof of Lemma 5.10: We will show that for our choice of δ

$$\left(1 + \log \frac{1}{\delta}\right)^k \delta \leq \rho / c. \tag{2}$$

Note that $\delta^{1/2} \leq \rho/c$ and, hence, it is enough to show $(1 + \log 1/\delta)^{2k} \leq (1/\delta)$. Let $t = (1/\delta)^{1/2k}$. The inequality we want to prove becomes $1 + 2k \log t \leq t$, which is true for $t \geq (3k)^2$.

Using (2) and Lemma 5.12 we obtain

$$\sigma(n, \delta n) < \left(1 + \log \frac{1}{\delta}\right)^k \cdot c \cdot \delta n \leq \rho n$$

and hence $A(n, \rho) \geq \delta n$. $\square$

VI. Wieferich Primes

In this section, we improve our tradeoff under a widely accepted number-theoretic hypothesis. This section is only relevant to *binary* codes, i.e., codes over $\mathbb{F}_2$.

Primes such that $t_{p,2} > 1$, i.e., primes p satisfying $2^{p-1} \equiv 1 \pmod{p^2}$, are called Wieferich primes and have played an important role in certain cases of Fermat's Last Theorem [23]. There are only two Wieferich primes known (1093 and 3511). There are no other Wieferich primes less than $4 \cdot 10^{12}$ [13]. It is not known whether there are infinitely many Wieferich primes. For $p = 1093$ and 3511 we have $t_{p,2} = 2$ and for all other primes $p < 4 \cdot 10^{12}$ we have $t_{p,2} = 1$.

Conjecture 6.1: For every prime p, $t_{p,2} \leq 2$.

(This is a restatement of Conjecture 1.7.)

Using the fact that the product of primes less than x is $\exp(x + o(1))$, we obtain the following conditional consequence of Theorem 5.1.

Corollary 6.2: Let $p(n)$ be the largest prime dividing n. If Conjecture 6.1 holds then, for any $\varepsilon > 0$, there exists $c > 0$ such that if $p(n) \leq c \ln n$ then for binary cyclic codes of length n, $\operatorname{dist}(C) \dim(C) \leq n^{1+\varepsilon}$.

VII. r-Closed Codes

So far, our proofs worked for all cyclic codes. The remaining cases will require the assumption of local testability.

In this section, we show how to replace the algorithmic concept of r-testability with the algebraic concept of r-closure on which the proof in the next section will be based.

Definition 7.1: The *r-core* of a code C is the subspace of C spanned by the words of weight $\leq r$. The *r-closure* of C is the dual of the r-core of the dual. We say that C is *r-closed* if C is its own r-closure.

Observation 7.2: C is r-closed exactly if $C^\perp$ is spanned by its words of weight $\leq r$. Moreover, the r-closure of C is the smallest r-closed subspace containing C.

Let $[n] = \{1, \ldots, n\}$. For $A \subseteq [n]$ and $x \in \mathbb{F}_q^n$, let x_A denote the restriction of x to A (a string of $|A|$ numbers from $\mathbb{F}_q$). We shall need the following characterization of the r-closure.

Proposition 7.3: Let $x \in \mathbb{F}_q^n$. Then x belongs to the r-closure of C if and only if for all $A \subseteq [n]$, if $|A| \leq r$ then there exists $y \in C$ such that $x_A = y_A$.

Proof: The "if" part is trivial, so we prove the "only if" part (which is the part we shall need below). Let D denote the r-closure of C and let $x \in D$. Let $A \subset [n]$, $|A| \leq r$. We need to prove that $(\exists y \in C)(x_A = y_A)$. Let $C_A = \{y_A : y \in C\}$ (the projection of C obtained by restriction to A). We need to prove that $x_A \in C_A$. This will be accomplished by proving that $x_A \perp C_A^\perp$. Let $z_A \in C_A^\perp$. We need to show that $x_A \perp z_A$. Let $z = (0, z_A) \in \mathbb{F}_q^n$ be the word with zero entries outside A and restricting to z_A. Now $z \in C^\perp$ because for all $y \in C$, $z \cdot y = z_A \cdot y_A = 0$. But $\operatorname{wt}(z) \leq r$, so $z \in D^\perp$ and, therefore, $x \cdot z = 0$. Consequently, $x_A \cdot z_A = x \cdot z = 0$, as desired. $\square$

Proposition 7.4: Let $\mathcal{T}$ be a complete r-tester for C (surely accepts all words in C) and let D be the r-closure of C. Then $\mathcal{T}$ surely accepts all words in D.

Proof: Let $x \in D$. On input x, let $\mathcal{T}$ select $A \subseteq [n]$, $|A| = r$ (in some order). Now $(\exists y \in C)(x_A = y_A)$; therefore, on input y, $\mathcal{T}$ makes the exact same sequence of choices. But $\mathcal{T}$ must accept y; therefore, it will accept x. $\square$

The main result of this section follows. This result holds not only for cyclic codes but also for codes that are invariant under a transitive group action.

Definition 7.5: Let G be a group acting on $[n]$ as permutations. We say that G is *transitive* if for every $i, j \in [n]$ there exists $g \in G$ such that $g(i) = j$. Let $C \subset \mathbb{F}^n$ be a code. We say that C is *invariant* under the group G if for every $g \in G$ we have $(a_1, \dots, a_n) \in C \Leftrightarrow (a_{g(1)}, \dots, a_{g(n)}) \in C$.

Lemma 7.6: ([4]) Let $A, B \subseteq [n]$ and let G be a transitive permutation group acting on $[n]$. Then

$$E(|A \cap B^g|) = \frac{|A| \cdot |B|}{n}$$

where E denotes the expected value over uniform $g \in G$.

Proof: For $i \in [n]$, let T_i be the event that $i \in B^g$. Let $i, j \in [n]$ and let $h \in G$ be such that $h(i) = j$. Note that gh is uniformly random over G. We have $i \in B^g$ iff $j \in B^{gh}$ and, hence, $\Pr(T_i) = \Pr(T_j)$. Let X_i be the indicator random variable of T_i. The function $X_1 + \cdots + X_n$ is constantly equal to $|B|$ and, hence, $E(X_1 + \cdots + X_n) = |B|$. Using

$$E(X_i) = \Pr(T_i) = \Pr(T_j) = E(X_j)$$

we obtain $E(X_i) = |B|/n$ for any $i \in [n]$. Finally, we obtain

$$E(A \cap B^g) = E\left(\sum_{a \in A} X_a\right) = \sum_{a \in A} E(X_a) = \frac{|A| \cdot |B|}{n}. \quad \square$$

Corollary 7.7: Under the conditions of Lemma 7.6, the expected size of the symmetric difference is

$$E(|A \triangle B^g|) = |A| + |B| - 2\frac{|A| \cdot |B|}{n}.$$

Lemma 7.8: Let C be a weakly r-testable code. Assume C is invariant under the action of a transitive permutation group G. Then C is r-closed.

Proof: Let D denote the r-closure of C. Note that D is invariant under the action of G. Assume for a contradiction that $C \neq D$. Let $d = \mathrm{dist}(C)$ and let u be a word in $D \setminus C$. As D is accepted by $\mathcal{T}$, we must have that $\mathrm{dist}(u, C) < \frac{d}{3}$. Therefore, there is a nonzero word w in D such that $\mathrm{wt}(w) < \frac{d}{3}$. Let $v \in D$ be a word of weight $< 2d/3$; let $\mathrm{wt}(v)$ be maximum under this constraint.

Claim 7.9: $\mathrm{wt}(v) > d/3$.

Proof: Assume for a contradiction that $\mathrm{wt}(v) \le d/3$. Let w' be a random translate of w under the given transitive group action. (So $w' \in D$.)

Let A be the set of coordinates on which v is nonzero and let B be the set of coordinates on which w' is nonzero. Note that $|A| = \mathrm{wt}(v)$ and $|B| = \mathrm{wt}(w)$. By Corollary 7.7, the expected size of $A \triangle B$ is $\mathrm{wt}(v) + \mathrm{wt}(w) - 2\mathrm{wt}(v)\mathrm{wt}(w)/n$ and, hence, the expected weight of $v + w'$ is at least

$$\mathrm{wt}(v) + \mathrm{wt}(w) - 2\mathrm{wt}(v)\mathrm{wt}(w)/n \ge \mathrm{wt}(v) + \mathrm{wt}(w)/3 > \mathrm{wt}(v);$$

therefore, there exists w' (a translate of w) such that $\mathrm{wt}(v + w') > \mathrm{wt}(v)$. Therefore, by the maximality of $\mathrm{wt}(v)$, we have $\mathrm{wt}(v + w') \ge 2d/3$. But

$$\mathrm{wt}(v + w') \le \mathrm{wt}(v) + \mathrm{wt}(w') < 2d/3$$

which is a contradiction, proving Claim 7.9. $\square$

Now $d/3 < \mathrm{wt}(v) < 2d/3$, and therefore, $\mathrm{dist}(v, C) > d/3$. This is a contradiction because v is accepted by $\mathcal{T}$ according to Proposition 7.4. This completes the proof of Lemma 7.8. $\square$

VIII. Not All Primes Small

In this section, we settle the case when n is divisible by an unbounded prime ($p(n) = \omega(1)$, where $p(n)$ is the largest prime dividing n). This is the only case where local testability plays a role; by Lemma 7.8, we shall assume that our codes are r-closed.

Theorem 8.1: Let $v \neq p$ be a prime such that $v \mid n$. For every r-closed[2] cyclic code C of length n we have either $\dim(C) \le 4nv^{-1/(r-1)}$ or $\mathrm{dist}(C) \le n/v$.

We postpone the proof of Theorem 8.1 to the end of this section and first show hot to prove Theorem 1.6.

Proof of Theorem 1.6: Let $c = c(1/2, q)$, as defined in Corollary 5.9. We will show that we can choose

$$c_1 = \max\left\{1, 1/c, 4c^{-1/(r-1)}\right\}$$

in the role of the constant "c" appearing in Theorem 1.6. There are two cases. If the primes dividing n are bounded by $c(\ln n \ln\ln n)^{1/2}$ then, by Corollary 5.9, $\dim(C)\mathrm{dist}(C) \le n^{3/2}$ and, hence, either $\dim(C) \le n^{3/4}$ or $\mathrm{dist}(C) \le n^{3/4}$. Using $c_1 \ge 1$ we obtain

$$\begin{aligned} n^{3/4} &\le c_1 n/(\log n \log\log n)^{1/2} \\ &\le c_1 n/(\log n \log\log n)^{1/2(r-1)} \end{aligned}$$

and hence we proved Theorem 1.6 in the first case. In the case when the largest prime dividing n is at least $c(\ln n \ln\ln n)^{1/2}$ we use Theorem 8.1. $\square$

Similarly, Theorem 1.8 follows by combining Theorem 8.1 with Corollary 6.2. Let $c = c(1/2, 2)$, as defined in Corollary 5.9. Then we can choose $c_2 = \max\{4, 1/c, 4c^{-1/(r-1)}\}$ in the role of the constant "c" appearing in Theorem I.8.

Corollary 8.2: Let $\{C_i\}$ be a family of locally testable cyclic codes. Let C_i be of length n_i, and let $p(n_i)$ be the largest prime divisor of n_i. If $\sup p(n_i) = \infty$ then this family cannot be good.

We now turn to the proof of Theorem 8.1. We shall need the following classical result on simultaneous diophantine approximation.

Theorem 8.3 (Dirichlet): Let $\alpha_1, \dots, \alpha_r$ be real numbers and $L > 0$ an integer. Then there exist integers $P_1, \dots, P_n, Q$ such that $|\alpha_i Q - P_i| \le L^{-1/r}$ and $0 < Q \le L$.

Let $\overline{\mathbb{F}}_q$ denote the algebraic closure of $\mathbb{F}_q$. As before, we assume that C is a code over $\mathbb{F}_q$, where $q = p^\ell$.

Lemma 8.4: Let $v \neq p$ be a prime, $h \in \overline{\mathbb{F}}_q[x]$ and $b \in \overline{\mathbb{F}}_q$, $b \neq 0$. Assume that $\mathrm{wt}(h) \le r$. Then either $x^v - b$ divides h or $\deg(\gcd(h, x^v - b)) \le 4v^{1-1/(r-1)}$.

Proof: Let c be a root of $x^v - b$. Let $f_1(x) = h(xc)$. Clearly

$$\deg(\gcd(h, x^v - b)) = \deg(\gcd(f_1, x^v - 1)).$$

Let $a_0, \dots, a_{r-1}$ be the exponents of nonzero terms in f_1, i.e., $f_1(x) = \sum_{j=0}^{r-1} x^{a_j}$. Without loss of generality, we can assume that $a_0 = 0$ because zero is not a root of $x^v - 1$. Let

[2]We assume here that $r > 1$ because otherwise, since $C^\perp$ is cyclic, we would have $C^\perp = \mathbb{F}^n$, and so $C = 0$.

$\alpha_i = a_i/v, i = 0, \ldots, r-1$. Let $P_1, \ldots, P_{r-1}, Q$ be the best (in max-norm) simultaneous diophantine approximation of $\alpha_1, \ldots, \alpha_{r-1}$ with $Q \leq v-1$. Let $P_0 = 0$. By Dirichlet's Theorem VIII.3

$$|\alpha_i Q - P_i| \leq (v-1)^{-1/(r-1)} \leq 2v^{-1/(r-1)}.$$

Let $b_i = a_i Q - P_i v$. Let further $t = \max_{i=0}^{r-1} |b_i|$. We have $t \leq 2v^{1-1/(r-1)}$. Define

$$f_2(x) = x^t f_1(x^Q) \ (\mathrm{mod}\ x^v - 1).$$

We have

$$f_2(x) \equiv_{\ (\mathrm{mod}\ x^v-1)} x^t f_1(x^Q) = \sum_{j=o}^{r-1} x^t \cdot (x^Q)^{a_j}$$
$$= \sum_{j=0}^{r-1} x^{t+b_j+vP_j} \equiv_{\ (\mathrm{mod}\ x^v-1)} \sum_{j=0}^{r-1} x^{t+b_i}.$$

As $0 \leq t + b_i \leq 4v^{1-1/(r-1)}$, we get that

$$\deg(f_2(x)) \leq 4v^{1-1/(r-1)}.$$

Let R be the inverse of Q modulo v, that is, $R \cdot Q = 1 \ (\mathrm{mod}\ v)$. If $\omega \in \overline{\mathbb{F}}_q$ is a common root of f_1 and $x^v - 1$ then ω^R is a root of f_2. Indeed, as $w^v = 1$ we get that the following equalities hold:

$$f_2(w^R) = w^{Rt} f_1(w^{RQ}) = w^{Rt} f_1(w) = 0.$$

If f_2 is the zero polynomial then $x^v - 1$ divides $x^t f_1(x^Q)$, and hence, $x^v - 1$ divides $f_1(x^Q)$. As Q is invertible $\mathrm{mod}\ v$ this implies that x^v divides $f_1(x)$. To see this we note that $x^v - 1 \mid x^{Rv} - 1$ and $x^{Rv} - 1 \mid f_1(x^{RQ})$. As

$$f_1(x^{RQ}) \equiv_{\ (\mathrm{mod}\ x^v-1)} f_1(x)$$

we get that $x^v - 1 \mid f_1(x)$. Thus, if $f_2 \equiv 0$ we get that $x^v - b \mid h$.

If $f_2 \not\equiv 0$, then it has at most $4v^{1-1/(r-1)}$ roots, and so by the discussion above, f_1 and $x^v - 1$ have at most $4v^{1-1/(r-1)}$ common roots. Hence,

$$\deg(\gcd(h, x^v - b)) = \deg(\gcd(f_1, x^v - 1)) \leq 4v^{1-1/(r-1)}.$$

□

Proof of Theorem 8.1: Let $h_1, \ldots, h_t \in \mathbb{F}_q[x]$ be the test polynomials of the code. We have $\mathrm{wt}(h_i) \leq r$, $i = 1, \ldots, t$. Let h be the parity-check polynomial of C. Let $n = p^s m v$ where m is not divisible by p. We have

$$x^n - 1 = (x^{mv} - 1)^{p^s} = \left(\prod_b (x^v - b)\right)^{p^s}$$

where the product ranges over all mth roots of unity $b \in \overline{\mathbb{F}}_q$. If for each b we have $\deg(\gcd(h, x^v - b)) \leq 4v^{1-1/(r-1)}$ then

$$\dim(C) = \deg(h) = \deg(\gcd(h, x^n - 1))$$
$$\leq p^s m \cdot 4v^{1-1/(r-1)} = 4nv^{-1/(r-1)}.$$

Otherwise, there exists an mth root of unity b such that $\deg(\gcd(h, x^v - b)) > 4v^{1-1/(r-1)}$. This means that for every $i = 1, \ldots, t$

$$\deg(\gcd(h_i, x^v - b)) > 4v^{1-1/(r-1)}.$$

By Lemma 8.4, $x^v - b$ divides each h_i and hence also h (as the h_i's span h). Note that $x^v - b^{q^i}$ divides h for any $i \geq 0$ because $\sigma_i : a \mapsto a^{q^i}$ is a Frobenius automorphism of $\overline{\mathbb{F}}_q$, and the coefficients of h are in $\mathbb{F}_q$ and hence fixed by σ_i. Let $g \in \mathbb{F}_q[x]$ be the minimal polynomial of b. Then $g = \prod_{i=0}^{k-1}(x - b^{q^i})$, where k is the smallest positive integer such that $b^{q^k} = b$. Hence, $g(x^v)$ divides h. As b is an mth root of unity we get that $g(x) \mid x^m - 1$, and so $g(x^v)$ divides $x^{mv} - 1$, and hence it divides $x^n - 1$. Let $f := (x^n - 1)/g(x^v)$. Clearly, $f \in C$. Moreover, f has at most n/v terms because all the exponents are divisible by v. This concludes the proof of the Theorem 8.1. □

REFERENCES

[1] N. Alon and J. H. Spencer, *The Probabilistic Method*. New York: Wiley, 1992.

[2] S. Arora, C. Lund, R. Motwani, M. Sudan, and M. Szegedy, "Proof verification and intractability of approximation problems," *J. Assoc. Comp. Mach.*, vol. 45, pp. 501–555, 1998.

[3] S. Arora and M. Safra, "Probabilistically checkable proofs: A new characterization of NP," *J. Assoc. Comp. Mach.*, vol. 45, pp. 70–122, 1998.

[4] L. Babai and P. Erdös, "Representation of group elements as short products," in *Theory and Practice of Combinatorics*, A. Rosa, G. Sabidussi, and J. Turgeon, Eds., 1982, vol. 12, Annals of Discrete Math., pp. 27–30.

[5] L. Babai, L. Fortnow, and C. Lund, "Nondeterministic exponential time has two-prover interactive protocols," *Comput. Complexity*, vol. 1, pp. 3–40, 1991.

[6] L. Babai, L. Fortnow, L. A. Levin, and M. Szegedy, "Checking computations in polylogarithmic time," in *Proc. 23rd ACM Symp. Theory of Computing*, New Orleans, LA, May 1991, pp. 21–31.

[7] E. Bach and J. Shallit, *Algorithmic Number Theory*. Cambridge, MA: MIT Press, 1996.

[8] E. B. Sasson, O. Goldreich, and M. Sudan, "Bounds on 2-Query Codeword Testing," ECCC, Tech. Rep. TR03-019, 2003.

[9] E. Ben Sasson, M. Sudan, S. Vadhan, and A. Wigderson, "Randomness-efficient low degree tests and short PCP's via epsilon-biased sets," in *Proc. 35th ACM Symp. Theory of Computing*, San Diego, CA, Jun. 2003, pp. 612–621.

[10] S. D. Berman, "Semisimple cyclic and Abelian codes," *Cybernetics*, vol. 3, pp. 21–30, 1967.

[11] M. Blum, M. Luby, and R. Rubinfeld, "Self-testing/correcting with application to numerical problems," *J. Comp. Syst. Sci.*, vol. 47, no. 3, pp. 549–595, 1993.

[12] G. Castagnoli, J. L. Massey, P. A. Schoeller, and N. von Seeman, "On repeated-root cyclic codes," *IEEE Trans. Inf. Theory*, vol. 37, no. 2, pp. 337–342, Mar. 1991.

[13] G. Castagnoli, K. Dilcher, and C. Pomerance, "A search for Wieferich and Wilson primes," *Math. Comp.*, vol. 66, pp. 433–449, 1997.

[14] A. Deshpande, R. Jain, T. Kavitha, S. V. Lokam, and J. Radhakrishnan, "Better lower bounds for locally decodable codes," in *Proc. 17th IEEE Conf. Computational Complexity*, Montreal, QC, Canada, May 2002, pp. 152–161.

[15] K. Friedl and M. Sudan, "Some improvements to total degree tests," in *Proc. 3rd Ann. Israel Symp. Theory of Computing and Systems*, Tel-Aviv, Israel, Jan. 1995, pp. 190–198.

[16] O. Goldreich, H. Karloff, L. Schulman, and L. Trevisan, "Lower bounds for linear locally decodable codes and private information retrieval," in *Proc, 17th IEEE Conf. Computational Complexity*, Montreal, QC, Canada, May 2002, pp. 175–183.

[17] O. Goldreich and M. Sudan, "Locally testable codes and PCP's of almost linear length," in *Proc. 43rd Symp. Foundations of Computer Science*, Vancouver, BC, Canada, Nov. 2002, pp. 13–22.

[18] J. Katz and L. Trevisan, "On the efficiency of local decoding procedures for error-correcting codes," in *Proc. 32nd ACM Symp. Theory of Computing*, Portland, OR, May 2000, pp. 80–86.

[19] I. Kerenidis and R. de Wolf, "Exponential lower bound for 2-query locally decodable codes via a quantum argument," in *Proc. 35th ACM Symp. Theory of Computing*, San Diego, CA, Jun. 2003, pp. 106–115.

[20] R. Lidl and H. Niederreiter, "Finite fields," in *Encyclopedia of Mathematics and its Applications*. New York: Cambridge Univ. Press, 1997, vol. 20.

[21] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes*. Amsterdam, The Netherelands: Elsevier/North-Holland, 1977.

[22] E. Prange, "Cyclic Error-Correcting Codes in Two Symbols," Air Force Cambridge Research Center, Cambridge, MA, Tech. Rep. AFCRC-TN-57-103, 1957.

[23] A. Wieferich, "Zum letzten Fermat'schen Theorem," *J. Reine Angew. Math.*, vol. 136, pp. 293–302, 1909.