Mathematical proof note

Binary Cyclic Codes with Simultaneously Near-Linear Minimum and Dual Distances

Yang Shifu
National Mobile Communications Research Laboratory and the Frontiers Science Center for Mobile Information Communication and Security, Southeast University, Nanjing 211189, China

Version 1.0 — October 16, 2026

Result overview

For every odd integer $m \geq 3$, put $n = 2^m - 1$. We give a self-contained probabilistic proof that ordinary binary cyclic codes of dimension $(n-1)/2$ can have both minimum distance and dual distance at least $(1/2 - o(1))n/(m \log_2 m)$. A finite-length sufficient condition is $H_2(\delta) + \delta \leq 1/(2m)$. The proof counts only low-weight words whose zeroth coordinate is one. Shift symmetry contributes a factor δ to the count, and a first-moment estimate for random reciprocal-factor choices completes the argument. A probability-tending-to-one version has the same leading distance constant.

1 Finite-Length Statement

Identify $\mathbb{F}_2^n$ with $R_n = \mathbb{F}_2[X]/(X^n - 1)$ by the coefficient vector. An ordinary binary cyclic code is an ideal of R_n. The dual is taken with respect to the usual binary dot product, and $d(A)$ denotes the minimum Hamming weight of a nonzero word of A. Write

$$H_2(x) = -x \log_2 x - (1-x) \log_2 (1-x), \qquad 0 < x < 1.$$

Theorem 1.1. *Let $m \geq 3$ be odd and $n = 2^m - 1$. If $0 < \delta < 1/2$ and*

$$H_2(\delta) + \delta \leq \frac{1}{2m}, \tag{1}$$

then there is an ordinary binary cyclic code C satisfying

$$\dim C = \frac{n-1}{2}, \qquad C \subset C^{\perp}, \qquad \min\{d(C), d(C^{\perp})\} > \delta n. \tag{2}$$

Condition (1) is sufficient; no necessity assertion is made. In particular, the distance guarantee is the integer bound $\lfloor \delta n \rfloor + 1$.

2 A First-Coordinate Counting Lemma

For $f \in R_n$, define its cyclic rank by

$$r(f) = \dim_{\mathbb{F}_2}(R_n f) = n - \deg \gcd(f, X^n - 1). \tag{3}$$

For $1 \le r \le n$, set

$$\mathcal{S}_r = \{f : r(f) = r,\ 0 < \mathrm{wt}(f) \le \delta n\}, \qquad \mathcal{S}_r^{(1)} = \{f \in \mathcal{S}_r : f_0 = 1\}.$$

Lemma 2.1. *For $0 < \delta < 1/2$,*

$$|\mathcal{S}_r^{(1)}| \le \delta\, 2^{2rH_2(\delta)}. \tag{4}$$

Proof. The first k coordinates form an information set for every k-dimensional binary cyclic code. Indeed, if g is its monic generator, every word has a unique expression ug with $\deg u < k$. Since $g(0) = 1$, the first k coefficients of ug determine those of u by a triangular system with diagonal entries one.

If $r(f), r(h) \le r$, then $f + h$ belongs to the cyclic code $R_n f + R_n h$, of dimension at most $2r$. Hence projection onto the first $\ell = \min(2r, n)$ coordinates is injective on the union of all words of cyclic rank at most r.

Assume $\mathcal{S}_r \neq \varnothing$ and let F be uniform on $\mathcal{S}_r$. This set is shift invariant, so every coordinate has the same probability

$$p = \Pr(F_i = 1) = \frac{\mathbb{E}\,\mathrm{wt}(F)}{n} \le \delta.$$

Prefix injectivity and entropy subadditivity give

$$\log_2 |\mathcal{S}_r| = H(F_0, \dots, F_{\ell-1}) \le \ell H_2(p) \le 2rH_2(\delta).$$

Finally, $|\mathcal{S}_r^{(1)}| = p|\mathcal{S}_r|$, which proves (4). The empty case is immediate. □

Only $\mathcal{S}_r$, not $\mathcal{S}_r^{(1)}$, is asserted to be shift invariant. The argument uses no independence among word coordinates.

3 Reciprocal Factors and One First Moment

Proof of Theorem 1.1. For a polynomial with constant term one, let $P^*(X) = X^{\deg P} P(X^{-1})$ denote its reciprocal. Since $n = 2^m - 1$ is odd, $X^n - 1$ is square-free and splits over $\mathbb{F}_{2^m}$. Its factorization over $\mathbb{F}_2$ has the form

$$X^n - 1 = (X + 1) \prod_{i=1}^{t} P_i P_i^*, \qquad P_i \neq P_i^*, \quad \deg P_i \le m. \tag{5}$$

To check the distinct reciprocal pairing, suppose a nontrivial root α and α^{-1} were Frobenius conjugate. If $s > 1$ is the order of α, then $2^j \equiv -1 \pmod{s}$ for some j. But the order of 2 modulo s divides the odd integer m, whereas -1 has order two, a contradiction.

Independently and uniformly choose $Q_i \in \{P_i, P_i^*\}$ and put

$$g = \prod_i Q_i, \qquad D = \langle g \rangle, \qquad C = \langle (X+1)g \rangle.$$

The dual generator formula yields

$$g_{D^\perp} = \left(\frac{X^n - 1}{g} \right)^* = (X+1)g.$$

Thus $C = D^\perp \subset D$, $\dim C = (n-1)/2$, and

$$\min\{d(C), d(C^\perp)\} = d(D). \tag{6}$$

Fix $f \neq 0$. If neither factor in some pair P_i, P_i^* divides f, then $\Pr(f \in D) = 0$. Otherwise, let b be the number of pairs in which exactly one factor divides f. Each such pair forces one independent choice, so $\Pr(f \in D) = 2^{-b}$. By (3), the cyclic rank is the sum of the degrees of the irreducible factors that do not divide f. The factor $X+1$ contributes at most one, and therefore $r(f) \leq 1 + mb$. In all cases,

$$\Pr(f \in D) \leq 2^{-(r(f)-1)/m}. \tag{7}$$

Count only the low-weight words with zeroth coordinate one:

$$Z_\delta(D) = |\{f \in D : f_0 = 1,\ \mathrm{wt}(f) \leq \delta n\}|.$$

Any nonzero word can be shifted to such a coordinate position. Hence the exact equivalence is

$$d(D) > \delta n \quad \Longleftrightarrow \quad Z_\delta(D) = 0. \tag{8}$$

Let $a = 1/m - 2H_2(\delta)$. If $a > 0$, Lemma 2.1, (7), and linearity of expectation imply

$$\begin{aligned} \Pr\{d(D) \leq \delta n\} &\leq \mathbb{E} Z_\delta(D) \\ &\leq \sum_{r=1}^{n} \delta\, 2^{2rH_2(\delta)} 2^{-(r-1)/m} \\ &\leq \frac{\delta\, 2^{1/m}}{2^a - 1}. \end{aligned} \tag{9}$$

Under (1), $a \geq 2\delta$. Since $2^{2\delta} - 1 > 2\delta \ln 2$,

$$\mathbb{E} Z_\delta(D) \leq \frac{\delta\, 2^{1/m}}{2^{2\delta} - 1} < \frac{2^{1/m}}{2 \ln 2} \leq \frac{2^{1/3}}{2 \ln 2} < 1. \tag{10}$$

The nonnegative integer $Z_\delta(D)$ is therefore zero for at least one choice. Equations (6) and (8) finish the proof. □

4 Asymptotic Distance and Success Probability

Corollary 4.1. *As $m \to \infty$ through odd integers, there are ordinary binary cyclic codes C_m of length $n = 2^m - 1$ and dimension $(n-1)/2$ such that*

$$\min\{d(C_m), d(C_m^{\perp})\} \geq \left(\frac{1}{2} - o(1)\right) \frac{n}{m \log_2 m}. \tag{11}$$

Proof. Let $\delta_m \in (0, 1/2)$ be the unique solution of

$$H_2(\delta_m) + \delta_m = \frac{1}{2m}.$$

The left side is strictly increasing on $(0, 1/2)$. Since

$$H_2(x) + x = x \log_2(1/x) + O(x) \qquad (x \downarrow 0),$$

its inverse satisfies $\delta_m = (1 + o(1))/(2m \log_2 m)$. Apply Theorem 1.1. □

The random choices themselves can be good with probability tending to one at the same leading distance constant. Put

$$\ell = \log_2 m, \qquad \epsilon_m = \ell^{-1/2}, \qquad H_2(\widetilde{\delta}_m) = \frac{1 - \epsilon_m}{2m}.$$

Then $\widetilde{\delta}_m = (1 + o(1))/(2m\ell)$ and $a = \epsilon_m/m$. From (9),

$$\Pr\{d(D) \leq n\widetilde{\delta}_m\} \leq \frac{2^{1/m} m \widetilde{\delta}_m}{(\ln 2)\epsilon_m} = O(\ell^{-1/2}) = o(1). \tag{12}$$

This version uses (9) directly and does not require (1).

Scope of the conclusion

The equivalence (8) is exact. The first-moment condition $\mathbb{E}Z_\delta < 1$ and the scalar condition (1) are only sufficient. The result is a probabilistic existence theorem, not a deterministic efficient procedure for choosing a good generator. The relative distance guaranteed by (11) tends to zero; no positive constant relative distance is proved. No orbit-size estimate, rank–weight inequality, nested all-rate construction, self-dual lifting, or channel-capacity theorem is needed for this proof.