Algebraic Expander Codes

Swastik Kopparty* Itzhak Tamo†

Abstract

Expander codes are a family of error-correcting codes based on expander graphs achieving constant rate and constant relative distance, while also admitting rich local structure. They are constructed by imposing a local code of rate r on the neighborhoods of an expanding graph. The standard constraint-counting argument then gives the global-rate lower bound $R \geq 2r - 1$, and therefore gives no positive-rate guarantee when $r \leq 1/2$. In particular, expander codes with positive global rate (from this argument) must have local relative distance at most $1/2$.

We construct explicit expander codes, based on algebraic ideas, that bypass this barrier. For every fixed $r \in (0, 1)$, including $r \leq 1/2$, our codes have local Reed–Solomon codes of rate r and relative distance $1 - r$, constant global rate, and constant global relative distance. The codes are obtained by evaluating a structured space of polynomials on an orbit of a noncommutative subgroup of $\mathrm{AGL}(1, \mathbb{F})$. The corresponding local constraints form a one-dimensional coset geometry whose incidence graph is a spectral expander with N vertices and degree $O(N^\epsilon)$ for any constant $\epsilon > 0$.

Our analysis of the rate does not use constraint counting. Instead, we explicitly exhibit a large space of codewords satisfying both local Reed–Solomon constraints, using a new notion of degree of a polynomial and a polytope-volume count. The family also satisfies a Schur-product closure property (i.e., these codes are "multiplication codes"), yielding an algebraic decoder of Berlekamp-Welch type, in addition to the standard iterative expander-code decoder.

1 Introduction

Expander/Tanner codes [15, 18] are one of the cleanest local-to-global constructions in coding theory: a long code is defined by imposing a short local code on the neighborhoods of a graph. When the graph has strong expansion and the local constraints are not too restrictive, one obtains constant rate, constant relative distance distance, and efficient iterative decoding (see, e.g., [15, 19, 8]).

The usual rate argument, however, has a sharp limitation. For an expander code with local rate r, the standard constraint-counting argument only guarantees global rate $R \geq 2r - 1$ [18]. Thus, below local rate $1/2$, the standard argument gives no positive-rate guarantee, but it does not preclude the existence of positive-rate codes. In particular, since the local relative distance is always at most $1 - r$, the usual rate argument cannot give codes with both (1) constant global rate, and (2) local relative distance $> 1/2$.

The regime $r < 1/2$ is especially interesting because of the potential for using local codes with the "multiplication property". Specifically, consider taking the local codes to be Reed-Solomon codes of rate r. The coordinate-wise product (formally called the *Schur-product*) of two Reed–Solomon codewords of rate r lie in a Reed–Solomon code of rate roughly $2r$, so the product local code has

*Department of Mathematics and Department of Computer Science, University of Toronto. Email: swastik.kopparty@utoronto.ca. Research supported by an NSERC Discovery Grant.

†Department of Electrical Engineering, Tel Aviv University. Email: tamo@tauex.tau.ac.il. Supported by the European Research Council (ERC) under Grant 852953.

nontrivial distance only if $r < 1/2$. This multiplicative structure is important in computational applications such as PCPs, cryptography, and recent quantum-code constructions [13, 6].

In this work, we give an explicit algebraic construction of a special expander graph and carefully chosen local codes that bypasses this constraint-counting barrier. For every fixed local rate $r \in (0, 1)$, including all $r \leq 1/2$, our codes have positive global rate and linear minimum distance. Crucially, the local constraints are Reed–Solomon codes, so the resulting family retains the multiplication property while its distance is proved through the spectral expansion of an underlying coset graph [2].

Main theorem (informal). *Fix any constant local rate $r \in (0, 1)$ and an integer $m > 0$. For infinitely many block lengths n we give an explicit linear code $\mathcal{C} = \mathcal{C}_{r,m,n} \subseteq \mathbb{F}^n$ such that: (i) $\mathcal{C}$ is the expander code arising from an explicit bipartite spectral expander graph with n edges, and all vertices having degree $O(n^{1/m})$. (ii) the local code on each vertex is a Reed–Solomon code of rate at most r; (iii) $\mathcal{C}$ has constant rate and constant relative distance (depending only on r and m); (iv) for any $0 < r_1, r_2 < 1$ with $r_1 + r_2 < 1$, we have the Schur-product containment: $\mathcal{C}_{r_1,m,n} * \mathcal{C}_{r_2,m,n} \subseteq \mathcal{C}_{r_1+r_2,m,n}$.*

In the standard "expander graph + local code" paradigm, one chooses a d-regular expanding graph and a length-d local code independently and combines them using the expander/Tanner code construction. Our code construction is most naturally viewed in a different way: as an algebraic code, evaluating a certain family of univariate polynomials at an additively and multiplicatively structured collection of points in $\mathbb{F}$. The expander graph then emerges naturally out of the pseudorandom interaction between the additive and multiplicative structures (as a *coset geometry* in the language of Kaufmann-Oppenheim [9]), and the local Reed-Solomon codes arise from the choice of the family of univariate polynomials. We term these codes *Algebraic Expander Codes* to emphasize this distinction.

This dual structure of the code, as both an expander code and an algebraic code, also has algorithmic consequences for decoding. Viewed as an expander code, the construction inherits the standard iterative expander-code decoders [15, 19]. On the other hand, viewed as an algebraic orbit-evaluation code that satisfies the multiplication property gives a different global decoder via error-correcting pairs [12], in the spirit of the Berlekamp-Welch algorithm. The first decoder has the better quantitative decoding radius, but the second highlights that the multiplication code property is algorithmically useful.

1.1 Connection to Locally Recoverable Codes and Graph Density

Our construction builds upon the algebraic framework of Locally Recoverable Codes (LRCs) [17]. In that framework, a code is defined by evaluating a special subspace of polynomials over some orbits of a subgroup G of the affine group $\mathrm{AGL}(1, \mathbb{F})$. The special subspace of polynomials guarantees that the restriction of any codeword to an orbit of G corresponds to a univariate polynomial of bounded degree, ensuring the local view is an RS code.

In [17], this algebraic framework was extended to construct codes with multiple recovering sets, where each coordinate participates in multiple (typically two) local codes. To achieve this, the evaluation set is partitioned into orbits under the action of two distinct subgroups of $\mathrm{AGL}(1, \mathbb{F})$, denoted as G and H. However, the constructions in [17] were restricted to subgroups of the *same algebraic type*, either two additive subgroups representing translations in distinct directions, or two multiplicative subgroups representing two different scalings.

We note that this approach had two primary limitations. First, it did not support arbitrary local rates, requiring $r > 1/2$ to ensure a non-trivial rate, which mirrors the bounds of standard

expander codes. Second, and more critically, using subgroups of the same type imposes inherent structural limitations on the code's underlying graph, which we explain next.

These codes admit a natural bipartite graph representation: code coordinates correspond to edges, while local codes (the orbits of G or H) correspond to vertices. Specifically, an edge connects the unique G-orbit and H-orbit containing that coordinate. The degree of any vertex is simply the size of its corresponding orbit.

When G and H share the same algebraic type, the corresponding actions commute, and the coset intersection graph associated with two commuting subgroups is complete bipartite (see [9, Proposition 4.3]). Thus the graph has excellent spectral expansion, but it is far too dense for the expander-code viewpoint: the degree is linear in the number of vertices.

This is where non-commutativity matters in this work. By taking one subgroup of translations and one subgroup of scalings, the generated group $\langle G, H\rangle$ becomes much larger relative to $|G|$ and $|H|$. The same coset-incidence construction therefore gives a substantially sparser graph.

1.2 Overview of Results

Our main contribution is an explicit construction of expander codes that overcome the limitations of commutative algebraic constructions. We summarize our results.

Main Results. Fix a prime p_0, a target local rate $r \in (0,1)$, and a sparsity parameter $m \geq 2$. Then for infinitely many powers $p = p_0^{\ell}$ with $p \to \infty$, we give an explicit linear code $\mathcal{C} = \mathcal{C}_{r,m,n} \subseteq \mathbb{F}^n$, where $n = p^{\Theta(m^2)}$ and $\mathbb{F}$ is a field of characteristic p_0, with the following properties:

1. **Graph Structure:** $\mathcal{C}$ is a Tanner code $\mathrm{Tanner}(\Gamma, (\mathcal{C}_v)_{v\in V})$, where:

   - $\Gamma = (V, E)$ is a bipartite *spectral expander graph* with $|E| = n$, with left and right degrees both $O(p^{m+1})$, and with normalized second largest singular value at most $O(1/\sqrt{p})$,
   - for each vertex $v \in V$, there is a code $\mathcal{C}_v \subseteq \mathbb{F}^{E(v)}$, where $E(v)$ denotes the set of edges incident on v.

   and thus $\mathcal{C}$ is precisely the set of vectors $f \in \mathbb{F}^E$ such that for all $v \in V$:

   $$f|_{E(v)} \in \mathcal{C}_v.$$

2. **Algebraic Structure:** $\mathcal{C}$ is defined by evaluating an explicit $\mathbb{F}$-linear space of polynomials on an orbit $\Omega \subseteq \mathbb{F}$ of a group $A = \langle G, H\rangle \leq \mathrm{AGL}(1, \mathbb{F})$.

3. **RS Locality:** The codes $\mathcal{C}_v$ are Reed–Solomon codes of length $|E(v)|$ and dimension at most $\lceil r|E(v)|\rceil$. In particular, each coordinate participates in two local RS constraints of local rate $r + o(1)$ and local relative distance at least $1 - r - o(1)$.

4. **Multiplication property:** The family satisfies a Schur-product closure property,

   $$\mathcal{C}_{r_1,m,n} * \mathcal{C}_{r_2,m,n} \subseteq \mathcal{C}_{r_1+r_2,m,n},$$

   whenever $r_1 + r_2 < 1$.

5. **Linear Global Rate:** The global rate R is bounded below by an explicit positive constant depending on r and m, even when the local rate $r \leq 1/2$. This bypasses the $2r - 1$ constraint-counting barrier for this algebraic code family.

6. **Linear Distance:** The global relative distance satisfies $\delta \geq (1-r)^2 - o(1)$. This comes from the Tanner code structure and the spectral expansion of the underlying graph.

7. **Decoding.** The construction admits two natural decoding algorithms, reflecting its two complementary descriptions. From the Tanner code description, the standard iterative expander-code decoder corrects a $(1-r)^2/2 - o(1)$ fraction of adversarial errors in time $\widetilde{O}(n^{1+1/(m+1)})$. From the algebraic description, the multiplication property gives a second global decoder via error-correcting pairs. Thus the same code family supports both local expander-style decoding and global algebraic decoding.

Remark 1.1 (Fixed-characteristic)**.** Throughout the two explicit instantiations, the characteristic is a fixed prime p_0. The parameter p denotes a power $p = p_0^\ell$, and all asymptotics $p \to \infty$ are taken by letting $\ell \to \infty$ while keeping p_0 and m fixed.

Limitations. We note two limitations of this construction. First, the alphabet size is not constant, but instead grows polynomially with the block length. Second, while the graph is sparse compared to the commutative case, its degree is still polynomial in n, scaling as $O(n^{1/(m+1)})$ rather than being constant. Overcoming these limitations remains an open question.

Remark 1.2. Our results are in fact more general by incorporating an additional global degree constraint on the space of evaluated polynomials (see Section 2.3). This gives rise to a family of algebraic subcodes of the previously described codes, while providing improved guarantees on the minimum distance. Importantly, these subcodes preserve the same local Reed–Solomon structure.

1.3 Techniques

The most significant challenge lies in proving that the global rate remains positive even when the local rate satisfies $r \leq 1/2$. In this regime, the standard dimension-counting argument for expander codes fails. To address this, we introduce a new notion of polynomial degree, explained next.

Establishing the Rate via u-Base Degree. To lower bound the dimension of the message space, we first introduce the necessary algebraic definitions. Let $g(X)$ and $h(X)$ be polynomials invariant under the actions of G and H, respectively. That is, $g(\phi(x)) = g(x)$ for any affine transformation $\phi \in G$ and all $x \in \mathbb{F}$, and similarly for $h(X)$ with respect to H.

We define the concept of the *u-base degree* as follows: Let $u(X)$ be a polynomial of positive degree d. Any polynomial $f(X)$ admits a unique u-adic expansion $f(X) = \sum_i c_i(X)u(X)^i$, where $\deg(c_i) < d$ for all i. The u-base degree of f is defined as the maximum degree among its coefficient polynomials, i.e., $\max_i\{\deg(c_i)\}$.

The task of lower bounding the code dimension is then reduced to bounding the dimension of the space of polynomials that simultaneously possess small g-base degree and small h-base degree.

A key technical contribution is proving that the u-base degree is sub-additive. This property allows us to translate the conditions on the message polynomials into a set of linear constraints on their h-base degrees. We then estimate the dimension of the message space by approximating the volume of the resulting high-dimensional polytope.

Spectral Expansion via Character Sums. To establish the graph expansion, we analyze the spectral gap (equivalently, the second largest singular value of the bi-adjacency matrix) by studying the two-step random walk operator. Using standard Fourier analysis techniques, we express the eigenvalues of this operator in terms of sums of non-trivial additive characters over a multiplicative

	Instantiation I (Sec. 4)	**Instantiation II (Sec. 5)**
Additive subgroup G	roots of $X^{p^m} + X^p + X$	$\mathbb{F}_{p^m}$
Multiplicative subgroup H	$\mathbb{F}_{p^m}^\times$	subgroup of $\mathbb{F}_{p^{m+1}}^\times$
Left/right degrees	$\lvert G\rvert = p^m$, $\lvert H\rvert = p^m - 1$	$\lvert G\rvert = p^m$, $\lvert H\rvert = \gamma(p^{m+1} - 1)$, $0 < \gamma < 1$
Character sum bound	Orthogonality of additive characters	Gauss-sum

Table 1: Two instantiations of the algebraic expander-code framework.

subgroup [1]. By applying classical bounds on these sums (e.g., bounds on Gauss sums), we bound the second singular value of the graph by $O(1/\sqrt{p})$, where p is the characteristic of the field. This confirms that the graph is a strong spectral expander.

Graph Sparsity via Mixed-Type Subgroups. As discussed in Section 1.1, the sparsity comes from using subgroups of different algebraic types: translations and scalings. Their noncommutativity makes $A = \langle G, H\rangle$ much larger relative to $|G|$ and $|H|$, and hence makes the corresponding coset graph much sparser than in the commuting case.

Two explicit instantiations. We instantiate the general framework in two complementary ways. The first instantiation (Section 4) takes the simplest multiplicative group $H = \mathbb{F}_{p^m}^\times$ (the entire multiplicative group of a field), yielding a clean spectral analysis that follows from the orthogonality of additive characters. This choice also makes the two local lengths (in other words, the degree of the vertices on each side) as balanced as possible (they differ by one). However, the construction requires working over a field large enough to realize the chosen additive subgroup G. See Section 4.1 for the definition of the group G and more details.

The second instantiation (Section 5) keeps the left local code algebraically simple by taking $G = \mathbb{F}_{p^m}$ and chooses H as a subgroup of $\mathbb{F}_{p^{m+1}}^\times$ of constant index. This introduces a tunable parameter $0 < \gamma < 1$ that governs the right degree, $|H| = \gamma(p^{m+1} - 1)$. The strong expansion follows via standard Gauss sum bounds. Unlike the first instantiation, this construction yields a graph with asymmetric degrees. This demonstrates the versatility of the general construction, showing that it can accommodate different group structures and rely on alternative character sum bounds (e.g., Gauss sums vs. orthogonality).

Table 1 provides a comparison between the two code instantiations.

1.4 Related Work

Expander/Tanner codes and expander-based coding. Expander codes originate in the work of Sipser and Spielman [15], building on the Tanner-code viewpoint [18]. A well-developed body of work relates global distance and decoding to graph expansion (see, e.g., [8, 19, 2]).

Unlike classical expander codes, our construction is not a black-box construction that works for an arbitrary expander graph and an arbitrary local code. This is deliberate. For Tanner codes with two local constraints per coordinate and local rate below 1/2, the standard constraint-counting argument provides no positive-rate guarantee. More importantly, a generic choice of graph and local constraints offers no apparent mechanism by which the large number of local constraints could become highly linearly dependent. Our construction addresses this by giving all local constraints a common algebraic origin. As a result, many of the local constraints are necessarily dependent, allowing the global code to have positive rate even below the 1/2 threshold. The graph, the local Reed–Solomon constraints, and the global polynomial space are therefore co-designed. This loss of genericity is the price paid for obtaining both positive rate in the low-local-rate regime and the multiplication property.

Algebraic locality and multiplication. On the algebraic side, our construction builds on the orbit-evaluation framework for locally recoverable codes introduced in [17], where subgroup invariants are used to enforce RS local views. A central motivation for using RS locality is the Schur-product (multiplication) behavior of RS codes: the coordinatewise product of two RS codewords of rate roughly r is itself a codeword in a related RS code of rate roughly $2r$ [13]. Such multiplicative structure has become increasingly important in coding theory and its applications, including PCP constructions and cryptography [11, 14, 3, 4], recent quantum-code constructions based on products of algebraic codes [6], and recent HDX-based code constructions that explicitly exploit multiplication [5]. More broadly, expander-based outer structures have also been used to combine the algebraic properties of Reed–Solomon codes with sparse combinatorial constructions for efficient coding algorithms [7].

Coset geometries. The graph underlying our construction is a one-dimensional instance of the standard coset-geometry framework of Kaufman–Oppenheim [9]. In particular, the face-coset description, and the complete-bipartite behavior in the commuting-subgroups case are all part of this general framework. Our contribution is the specific instantiation of this framework, i.e., the specific non-commuting subgroups inside $\mathrm{AGL}(1, \mathbb{F})$, together with the orbit-evaluation code placed on the resulting graph.

Comparison with Dinur–Liu–Zhang. The closest prior work to ours is the family of HDX codes of Dinur, Liu, and Zhang [5]. Their construction can be described as Tanner codes on a two-dimensional coset complex whose coordinates are triangles and whose local views around edges are RS codewords. Thus, their work and ours share the same broad coset-locality template: functions on a group-like set with RS constraints on subgroup cosets.

The main difference lies in the dimension and the resulting parameters. In the two-dimensional construction of [5], each coordinate participates in three local RS constraints, so the standard constraint-counting argument guarantees positive global rate whenever the local rate exceeds $2/3$. They obtain constant alphabet, constant locality, and local testability in a low-local-rate regime, however with exponentially small global rate. On the other hand, if they let the alphabet size grow, they lose the constant locality and testability, but obtain a polynomially small global rate. In contrast, our construction is one-dimensional: each coordinate participates in only two local RS constraints, so the corresponding constraint-counting threshold drops to local rate greater than $1/2$. Our main result, however, establishes constant global rate and constant relative distance even below this threshold. The price we pay is that the alphabet size and graph degree grow polynomially with the block length, and we do not prove local testability.

Organization of the paper. Section 2 introduces the general algebraic framework, defines the expander code, and establishes the associated bipartite graph structure. In Section 3, we analyze the spectral expansion of this graph by reducing the bound on the second eigenvalue to bounds on additive character sums. The two code instantiations are presented in Section 4 and Section 5. Section 6 provides the two decoding algorithms for these codes. Finally, Section 7 offers concluding remarks and discusses open problems.

2 General Code Construction

2.1 Algebraic Setup and Evaluation Domain

Fix a prime p_0, and let $p = p_0^\ell$ be a power of p_0. Let $\mathbb{F}$ be a finite extension field of $\mathbb{F}_p$; in particular, $\mathbb{F}$ has characteristic p_0. Throughout, we use X to denote a formal indeterminate in polynomial rings (e.g., $\mathbb{F}[X]$), and we use $x \in \mathbb{F}$ to denote a field element (an evaluation point) on which affine transformations act.

Let

$$\mathrm{AGL}(1, \mathbb{F}) = \{\, x \mapsto ax + b : a \in \mathbb{F}^\times,\ b \in \mathbb{F} \,\}$$

denote the group of affine transformations acting on $\mathbb{F}$. Let $G, H \leq \mathrm{AGL}(1, \mathbb{F})$ be subgroups, and let

$$A = \langle G, H \rangle$$

be the subgroup they generate. In our main instantiations we will take G to be a (pure) translation subgroup and H to be a (pure) scaling subgroup. See definitions ahead.

We assume without loss of generality that $|\mathbb{F}| \geq |A|$ (replacing $\mathbb{F}$ by a sufficiently large extension field if necessary). Under this assumption, there exists an element $\alpha \in \mathbb{F}$ whose stabilizer in A is trivial, i.e., $\phi(\alpha) \neq \alpha$ for all $\phi \in A \setminus \{\mathrm{id}\}$. Indeed, any non-identity affine map $\phi(x) = ax + b$ has at most one fixed point in $\mathbb{F}$. Hence, the set of elements of $\mathbb{F}$ fixed by some nontrivial element of A has cardinality at most $|A| - 1$, and since $|\mathbb{F}| \geq |A|$, we can choose α outside this set.

Let Ω denote the orbit of α under the action of A:

$$\Omega = \{\phi(\alpha) : \phi \in A\} \subseteq \mathbb{F}.$$

By construction, the stabilizer of α is only id, and therefore the action of A on Ω is free. Since Ω is an A-orbit, the action is also transitive, and hence it is regular.

We will construct an evaluation code over $\mathbb{F}$ by evaluating at the points of Ω. Consequently, the block length of the code is $n := |\Omega| = |A|$.

We will next need the notion of invariant polynomials.

Invariant polynomials. Let $G \leq \mathrm{AGL}(1, \mathbb{F})$, and let $f(X) \in \mathbb{F}[X]$. We say that f is *G-invariant* if

$$f(g(x)) = f(x) \qquad \text{for all } g \in G \text{ and all } x \in \mathbb{F}.$$

In particular, any G-invariant polynomial is constant on G-orbits in $\mathbb{F}$. Indeed, for every $a \in \mathbb{F}$ and $g \in G$,

$$f(g(a)) = f(a).$$

Clearly, every constant polynomial is G-invariant, but there are also nontrivial examples. For instance, it is easy to verify that

$$\prod_{g \in G} g(X), \tag{1}$$

is G-invariant.

We next consider two special cases: when G is a pure translation group and when G is a pure scaling group.

Pure translations. Suppose

$$G = \{\, x \mapsto x + u : u \in U \,\}$$

for some additive subgroup $U \leq (\mathbb{F}, +)$. Then (1) becomes

$$\prod_{g \in G} g(X) \;=\; \prod_{u \in U} (X + u) \;=\; \prod_{u \in U} (X - u), \tag{2}$$

where the last equality holds since U is closed under negation. The right-hand side of (2) is precisely the *annihilator polynomial* (also known as the *subspace polynomial*) of the additive subgroup U.

Pure scalings. Suppose

$$G = \{\, x \mapsto ux : u \in U \,\}$$

for some multiplicative subgroup $U \leq (\mathbb{F}^\times, \cdot)$. Then (1) becomes

$$\prod_{g \in G} g(X) \;=\; \prod_{u \in U} (uX) \;=\; X^{|U|} \prod_{u \in U} u. \tag{3}$$

Since $\prod_{u \in U} u \in \mathbb{F}^\times$ is a constant, (3) implies that the monomial $X^{|U|}$ is constant on the G-orbits of $\mathbb{F}$ under the scaling action.

In the sequel, we take G to be a pure translation group and define its associated G-invariant polynomial by

$$g(X) \;:=\; \prod_{u \in U} (X - u), \tag{4}$$

i.e., the right-hand side of (2). We will freely identify the translation subgroup $G \leq \mathrm{AGL}(1, \mathbb{F})$ with its set of translation parameters $U \leq (\mathbb{F}, +)$. Likewise, we take H to be a pure scaling group and define its associated H-invariant polynomial by

$$h(X) \;:=\; X^{|H|}. \tag{5}$$

We will repeatedly use that $g(X)$ and $h(X)$ are constant on the corresponding G- and H-orbits in $\mathbb{F}$.

2.2 Base u degree of a polynomial

We need the following notion of the degree of a polynomial.

Definition 2.1 (Base-u Degree)**.** Let $u(X) \in \mathbb{F}[X]$ be a *non-constant* polynomial. Any polynomial $f(X) \in \mathbb{F}[X]$ can be uniquely expanded in base $u(X)$ as:

$$f(X) = \sum_i c_i(X) u(X)^i,$$

where $\deg(c_i) < \deg(u)$ for all i. We define the *u-base degree* of f, denoted $\deg_u(f)$, as the maximum degree of its coefficients:

$$\deg_u(f) := \max_i \{\deg(c_i)\},$$

with the convention $\deg(0) := -\infty$.

Then, clearly, $\deg_u(f) \in \{-\infty, 0, \ldots, \deg(u) - 1\}$ and $\deg_u(f + g) \leq \max\{\deg_u(f), \deg_u(g)\}$. The following lemma shows that the u-base degree is sub-additive.

Lemma 2.2 (Subadditivity of the u-base degree)**.** *Let $u(X) \in \mathbb{F}[X]$ be a nonconstant polynomial. Then for all $f, g \in \mathbb{F}[X]$,*

$$\deg_u(fg) \;\leq\; \deg_u(f) + \deg_u(g).$$

Proof. Let $d := \deg(u)$. If $\deg_u(f) + \deg_u(g) \geq d$, then the result is trivially true (since $\deg_u$ of any polynomial is at most $d-1$).

Thus assume that $\deg_u(f) + \deg_u(g) < d$, and write

$$f = \sum_i c_i u^i, \qquad g = \sum_j d_j u^j,$$

with $\deg(c_i), \deg(d_j) < d$. Then

$$fg = \sum_{k \geq 0} e_k \, u^k, \tag{6}$$

where $e_k := \sum_{i+j=k} c_i d_j \in \mathbb{F}[X]$. Now, for each k, we have

$$\deg(e_k) \;\leq\; \max_{i+j=k} \big(\deg(c_i) + \deg(d_j) \big) \;\leq\; \deg_u(f) + \deg_u(g) < d,$$

and thus (6) is the base u expansion of fg, and the result follows. □

2.3 Message Space and Code Definition

In this section we give the definition of the message space and the code. This is specified by a local rate parameter $r \in (0,1)$, subgroups G and H of $\mathrm{AGL}(1,\mathbb{F})$, and a global degree parameter $D \leq n$.

Remark 2.3. The primary case of interest in this paper is $D = n$, in which the code coincides with a Tanner code defined on the coset graph. The more general regime $D \leq n$ yields a natural family of algebraic subcodes of this Tanner code.

For a first reading, we therefore recommend focusing on the case $D = n$, as it simplifies the construction while capturing all the main results described in Section 1.2.

For the groups G and H, let $g(X)$ and $h(X)$ be the associated G- and H-invariant polynomials, respectively.

We define the message space of polynomials $\mathcal{W}_{r,D}$ by

$$\mathcal{W}_{r,D} := \Big\{ f(X) \in \mathbb{F}[X] \;\Big|\; \deg(f) < D, \;\; \deg_g(f) < r|G|, \;\; \deg_h(f) < r|H| \Big\}^{\,[1]}. \tag{7}$$

The evaluation code $\mathcal{C}$ is defined as the image of the message space under evaluation over Ω:

$$\mathcal{C} = \mathcal{C}(G, H; r, D) := \big\{ (f(\beta))_{\beta \in \Omega} : f \in \mathcal{W}_{r,D} \big\}. \tag{8}$$

We collect a few fundamental properties of $\mathcal{C}$. The code length is n. Since every message polynomial has degree strictly less than $D \leq n$, the evaluation map is injective. Thus, $\dim_{\mathbb{F}}(\mathcal{C}) = \dim_{\mathbb{F}}(\mathcal{W}_{r,D})$. Furthermore, since $\mathcal{C}$ is a linear subcode of the Reed–Solomon code of length n and degree D, its minimum distance is at least $n - D + 1$.

Proposition 2.4. *The code $\mathcal{C}$ is a linear block code with parameters* $[\, n, \; \dim_{\mathbb{F}}(\mathcal{W}_{r,D}), \; \geq n - D + 1 \,]_{\mathbb{F}}$.

Our construction naturally admits two complementary decoding viewpoints. In the principal case $D = n$, the code is a Tanner/expander code and therefore inherits the standard iterative (combinatorial) decoding paradigm of expander codes. In the more general regime $D \leq n$, the code remains a subcode of the $D = n$ construction and thus still supports iterative decoding; however,

[1] It is straightforward to verify that $\mathcal{W}_{r,D}$ is an $\mathbb{F}$-linear subspace of $\mathbb{F}[X]$.

it additionally admits an algebraic decoding algorithm, as it can be viewed as a subcode of a Reed–Solomon code.

Thus, the parameter D interpolates between two structural perspectives on the same construction: a combinatorial one, based on expander decoding, and an algebraic one, based on global polynomial constraints.

Remark 2.5. We emphasize that we are primarily interested in the regime where the local rate satisfies $r \leq 1/2$. In this case, establishing a lower bound on the global rate is non-trivial. In contrast, if the local rate satisfies $r > 1/2$, a standard dimension counting argument suffices to show that the global rate of the code is at least $(2r-1)D/n - o(1)$.

Indeed, it suffices to show that the dimension of the subspace $\mathcal{W}_{r,D}$ is at least $(2r-1)D - 2$. It is easy to verify that $\mathcal{W}_{r,D} = U \cap V$, where

$$U = \mathrm{Span}\left(X^i g(X)^j : \deg(X^i g(X)^j) < D,\ i < r\,|G|\right),$$
$$V = \mathrm{Span}\left(X^i h(X)^j : \deg(X^i h(X)^j) < D,\ i < r\,|H|\right),$$

and $\dim(U), \dim(V) \geq \lfloor Dr \rfloor \geq Dr - 1$. Therefore,

$$D \geq \dim(U+V) = \dim(U) + \dim(V) - \dim(U \cap V) \geq 2rD - 2 - \dim(\mathcal{W}_{r,D}).$$

By rearranging the claim follows.

2.4 Local Codes

Since Ω is A-invariant, it is partitioned into orbits under the actions of both G and H. We will show that the restriction of any codeword to each such orbit lies in a Reed–Solomon code of small rate. Equivalently, every code coordinate participates in two local Reed–Solomon constraints (one induced by a G-orbit and one induced by an H-orbit), each of small rate.

Let $\beta \in \Omega$. The orbit of β under G, denoted β^G, has size $|G|$. As noted in Section 2.1, $g(x)$ takes a constant value, say $\lambda = g(\beta)$, on β^G. Let $f \in \mathcal{W}_{r,D}$ be a message polynomial. Expanding f in g-base, we have $f(X) = \sum_i c_i(X) g(X)^i$. Restricting f to the domain β^G, this becomes:

$$f|_{\beta^G}(x) = \sum_i c_i(x) \lambda^i.$$

The condition $\deg_g(f) < r|G|$ implies that $\deg(c_i) < r|G|$ for all i. Therefore, the restriction of the codeword to β^G lies in a Reed–Solomon code of length $|G|$ and dimension at most $\lceil r|G| \rceil$. Equivalently, its local rate is at most $\lceil r|G| \rceil / |G| = r + o(1)$, and its local relative distance is at least $1 - r - o(1)$.

A symmetric argument applies to the subgroup H. Since $\deg_h(f) < r|H|$ and $h(x)$ is constant on β^H, the restriction of the codeword to any orbit β^H belongs to a Reed–Solomon code of length $|H|$ and rate at most $r + o(1)$.

2.5 Multiplication Property

The polynomial definition of the code also gives the following Schur-product behavior, called the multiplication property. That is, the Schur-product of two algebraic expander codes with small enough local rates is contained in another algebraic expander code.

Write $\mathcal{C}_{r,D} := \mathcal{C}(G, H; r, D)$. If $a, b \in \mathbb{F}^\Omega$, let $a * b = (a(\omega)b(\omega))_{\omega \in \Omega}$ denote their Schur product. For two linear codes $\mathcal{C}_1, \mathcal{C}_2 \subseteq \mathbb{F}^\Omega$, define

$$\mathcal{C}_1 * \mathcal{C}_2 := \mathrm{Span}_{\mathbb{F}}\{a * b : a \in \mathcal{C}_1,\ b \in \mathcal{C}_2\}.$$

We will use the following simple observation about reducing polynomials modulo the vanishing polynomial of the evaluation set. Let $Z_\Omega(X) := \prod_{\beta\in\Omega}(X-\beta)$. Since Ω is a union of G-orbits and also a union of H-orbits, and since $g(X)$ and $h(X)$ are the corresponding invariant polynomials, we have $Z_\Omega \in \mathbb{F}[g] \cap \mathbb{F}[h]$. Equivalently,

$$\deg_g(Z_\Omega) = \deg_h(Z_\Omega) = 0.$$

Lemma 2.6. *Let u be either g or h. Then reduction modulo Z_Ω does not increase the u-base degree, that is, for any $f \in \mathbb{F}[X]$*

$$\deg_u(f \bmod Z_\Omega) \le \deg_u(f).$$

Proof. Since $Z_\Omega \in \mathbb{F}[u]$, write $Z_\Omega = Q(u)$ for some polynomial Q. Let $d = \deg(u)$, and write f in the form

$$f = \sum_{j=0}^{d-1} X^j A_j(u),$$

which is just the u-base expansion grouped according to the powers of X. Reducing modulo $Z_\Omega = Q(u)$ only replaces each polynomial $A_j(u)$ by its remainder modulo $Q(u)$. In particular, no term with a larger power of X is introduced. Hence the largest j that appears, which is exactly the u-base degree, cannot increase. □

Proposition 2.7 (Multiplication property)**.** *Let $r_1, r_2 \in (0,1)$ with $r_1 + r_2 < 1$, and let $D_1, D_2 \le n$. Then*

$$\mathcal{C}_{r_1,D_1} * \mathcal{C}_{r_2,D_2} \subseteq \mathcal{C}_{r_1+r_2,\ \min\{n,D_1+D_2\}}.$$

Proof. Let $c_1 = (f_1(\beta))_{\beta\in\Omega} \in \mathcal{C}_{r_1,D_1}$ and $c_2 = (f_2(\beta))_{\beta\in\Omega} \in \mathcal{C}_{r_2,D_2}$, where $f_i \in \mathcal{W}_{r_i,D_i}$. Set $D' = \min\{n, D_1 + D_2\}$ and $f' = f_1 f_2 \bmod Z_\Omega$. Then clearly $deg(f') < D'$ and the product word $c_1 * c_2$ is the evaluation of f' on Ω.

It remains to check the two base-degree conditions. By Lemma 2.6 and Lemma 2.2,

$$\deg_g(f') \le \deg_g(f_1 f_2) \le \deg_g(f_1) + \deg_g(f_2) < (r_1 + r_2)|G|,$$

and similarly $\deg_h(f') < (r_1 + r_2)|H|$. Thus $f' \in \mathcal{W}_{r_1+r_2,D'}$, and therefore $c_1 * c_2 \in \mathcal{C}_{r_1+r_2,D'}$. □

Remark 2.8. In particular, for $r < 1/2$ we get

$$\mathcal{C}_{r,D} * \mathcal{C}_{r,D} \subseteq \mathcal{C}_{2r,\ \min\{n,2D\}}.$$

Thus the product of two codewords lands in a related code whose local Reed–Solomon constraints still have nontrivial distance and therefore the code has nontrivial global distance. This is the main reason the low-local-rate regime $r < 1/2$ is relevant for multiplication.

2.6 Bipartite Expander-Code Viewpoint

The locality structure described above induces an underlying graph topology. In this section, we formalize this viewpoint by identifying the code coordinates (the elements of Ω) with the edges of a bipartite graph, and the local constraints with its vertices.

More specifically, each local constraint corresponds to the restriction of a codeword to an orbit of Ω under the action of G or H. We associate a vertex with each such orbit, and we associate an edge with each code coordinate in Ω. An edge (an element of Ω) is incident to the two vertices

corresponding to the unique G-orbit and H-orbit that contain it. This naturally yields a bipartite graph whose left and right vertex sets are the G-orbits and H-orbits in Ω, respectively.

Next, we view this graph as a coset graph of the subgroups G and H within A. Since the action of A on Ω is regular, we may identify the code coordinates (and hence the graph edges) with the elements of A via the bijection $\iota : A \to \Omega$ given by

$$\iota(\phi) = \phi(\alpha). \tag{9}$$

To determine the vertex sets, we identify which subsets of coordinates correspond to single local constraints. Fix an affine map $\phi \in A$ and let $P = \phi(\alpha)$. The G-orbit of P is

$$\mathrm{Orb}_G(P) \;=\; \{\, g(P) : g \in G \,\} \;=\; \{\, g(\phi(\alpha)) : g \in G \,\}.$$

Since the group operation in A is composition, $g(\phi(\alpha)) = (g \circ \phi)(\alpha)$. Hence, under the identification ι, the G-orbit corresponds to the set $G\phi \;:=\; \{\, g\phi : g \in G \,\}$, namely the right coset of G in A containing ϕ. Consequently, the vertices corresponding to G-orbits are naturally indexed by the set of right cosets

$$A/G \;=\; \{\, G\phi : \phi \in A \,\}.$$

An identical argument shows that the vertices corresponding to H-orbits are indexed by the set of right cosets A/H.

This leads to the following graph definition.

Graph definition. We define the bipartite graph $\mathsf{B}(G, H)$ using the coset geometry of A as follows:

- **Left vertices (V_L):** the set of right cosets of G in A, i.e., $V_L := \{\, G\phi : \phi \in A \,\}$.
- **Right vertices (V_R):** the set of right cosets of H in A, i.e., $V_R := \{\, H\phi : \phi \in A \,\}$.
- **Edges:** the edge set is identified with the group A. For each $\phi \in A$, we add an edge connecting the left vertex $G\phi$ to the right vertex $H\phi$.

This is precisely the one-dimensional coset complex associated with the group A and the two subgroups G, H, up to our convention of using right cosets. In the usual coset-complex language, vertices are subgroup cosets and incidences are given by nonempty intersections. More generally, faces of a coset complex can be identified with cosets of intersections of the defining subgroups; for standard facts about the coset-geometry framework see [9]. In our main instantiations $G \cap H = \{\mathrm{id}\}$, so the edges are naturally identified with the elements of A.

The graph is biregular. Indeed, each left vertex has degree $|G|$ and each right vertex has degree $|H|$. Equivalently, $\mathsf{B}(G, H)$ is the standard coset incidence graph between A/G and A/H.

When $D = n$ the code $\mathcal{C}$ is the Tanner code on $\mathsf{B}(G, H)$ with the following local constraints:

- For every left vertex $u = G\phi \in V_L$, the incident edges are indexed by the coset $G\phi$, and the restriction of a codeword to these edges must belong to a Reed–Solomon code of length $|G|$ and dimension at most $\lceil r|G| \rceil$.
- For every right vertex $v = H\phi \in V_R$, the incident edges are indexed by the coset $H\phi$, and the restriction of a codeword to these edges must belong to a Reed–Solomon code of length $|H|$ and dimension at most $\lceil r|H| \rceil$.

For general $D \leq n$, the code $\mathcal{C}$ is a subcode of this Tanner code.

This representation enables the use of standard expander-code tools: bounds on the minimum distance and efficient decoding can be derived from spectral expansion of $\mathsf{B}(G, H)$, quantified by the second-largest singular value of its (normalized) adjacency operator.

Remark 2.9. Note that the graph is not a multigraph (i.e., it is a simple graph) if and only if G and H intersect only trivially. Indeed, consider the two cosets Ha, Ga that contain the element $a \in A$. Then the number of edges between the corresponding vertices Ha and Ga is equal to $|Ha \cap Ga| = |H \cap G|$ and the claim follows.

3 Spectral Expansion of the Coset Graph

To prove that the constructed code is indeed an expander code, it suffices to show that the associated coset graph is expanding.

In this section we analyze the spectral properties of the coset (bipartite) graph $\mathsf{B}(G, H)$ arising from the construction. In particular, we derive an upper bound on the second largest singular value of its normalized bi-adjacency operator (equivalently, the second largest eigenvalue in absolute value of the normalized adjacency matrix of the bipartite graph). This spectral bound quantifies the expansion of $\mathsf{B}(G, H)$, and will allow us to lower bound the relative minimum distance of the global code in terms of the expansion of the graph and the minimum distances of the local codes.

By abuse of notation, we let B also denote the rectangular bi-adjacency matrix indexed by the set of left vertices A/G and right vertices A/H, where the entry (Ga, Hb) equals the number of edges connecting the cosets, i.e., $|Ga \cap Hb|$.

We consider the *normalized* adjacency operator $T : \ell_2(A/H) \to \ell_2(A/G)$ defined by:

$$T := \frac{1}{\sqrt{|G||H|}}\mathsf{B}.$$

$\sigma_2(T)$ is the second largest singular value of the normalized bi-adjacency operator, and equals the second largest eigenvalue in absolute value of the normalized adjacency matrix $\begin{pmatrix} 0 & T \\ T^* & 0 \end{pmatrix}$. It is well known that the largest singular value is $\sigma_1(T) = 1$, corresponding to the constant eigenfunction $\mathbf{1}$. Thus, $\sigma_2(T)$ satisfies:

$$\sigma_2(T) = \sup_{\substack{f \perp \mathbf{1} \\ \|f\|=1}} \|Tf\|.$$

To bound $\sigma_2(T)$, we analyze the eigenvalues of the self-adjoint operator T^*T, where it holds that $\sigma_2(T)^2 = \lambda_2(T^*T)$.

3.1 Spectral Analysis via Random Walks

The operator T^*T acts on the space $\ell_2(A/H)$. The entry $(T^*T)_{Hj,Hj'}$ corresponds to the probability weight of a path of length 2 starting at Hj and ending at Hj' in the bipartite graph. Specifically,

$$(T^*T)_{Hj,Hj'} = \frac{1}{|G||H|} \sum_{Gi \in A/G} |Gi \cap Hj| \cdot |Gi \cap Hj'|.$$

For a function $f : A/H \to \mathbb{C}$, the action of the operator is given by:

$$(T^*Tf)(Hj) = \sum_{Hj' \in A/H} (T^*T)_{Hj,Hj'} f(Hj') = \sum_{Hj' \in A/H} \frac{1}{|G||H|} \sum_{Gi \in A/G} |Gi \cap Hj| \cdot |Gi \cap Hj'| f(Hj'). \tag{10}$$

This summation can be interpreted as a random walk on the graph. Starting from a right coset Hj, the walk proceeds as follows:

1. Choose a random edge incident to Hj. This corresponds to picking a random $h \in H$ and moving to the element $x = hj$.
2. Traverse to the left vertex containing x, which is $G(hj)$.
3. Choose a random edge incident to $G(hj)$. This corresponds to picking a random $g \in G$ and moving to the element $y = g(hj)$.
4. Traverse to the right vertex containing y, which is $H(ghj)$.

Thus, the operator (10) acts as an averaging operator over the H-cosets:

$$(T^*Tf)(Hj) = \mathbb{E}_{\substack{g\in G \\ h\in H}}\big[f(Hghj)\big]. \tag{11}$$

3.2 Analysis for Translation and Scaling Subgroups

In this section, we analyze the operator (11) for specific groups H, G, as follows.

- Let $H \subseteq \mathbb{F}^\times$ be a multiplicative subgroup. We identify H with the group of scaling maps $\{x \mapsto hx \mid h \in H\}$.
- Let $G \subseteq \mathbb{F}$ be an $\mathbb{F}_p$-vector subspace, viewed as a subgroup of $(\mathbb{F}, +)$. We identify G with the group of translations $\{x \mapsto x + g : g \in G\}$.

Before proceeding to characterize the group $A = \langle G, H\rangle$ generated by G and H, we note that for our choice of G and H the resulting graph is simple. Indeed, G and H intersect trivially, $G \cap H = \{\mathrm{id}\}$, since the identity map is the only affine transformation that is simultaneously a translation and a scaling.

Theorem 3.1. *Let S be the additive closure of G under the action of H, defined as,*

$$S = \sum_{h\in H} hG = \left\{ \sum_k h_k g_k \;\middle|\; h_k \in H, g_k \in G \right\} \subseteq \mathbb{F}.$$

Then the group A is isomorphic to the semidirect product $S \rtimes H$. Specifically:

1. *The elements of A are exactly the affine maps $\{x \mapsto hx + s \mid h \in H, s \in S\}$.*
2. *The group operation corresponds to $(s_1, h_1) \cdot (s_2, h_2) = (s_1 + h_1 s_2, h_1 h_2)$.*

Proof. See Appendix A. □

The following lemma gives another characterization of the set S that will be useful later.

Lemma 3.2. *For H and G as above, the set $S = \sum_{h\in H} hG$ satisfies $S = \mathrm{Span}_{\mathbb{F}_p(H)}(G)$*

Proof. See Appendix A. □

Following Theorem 3.1, we identify the right cosets of H in A with the elements of the subspace S. Specifically, the coset containing $(s,1)$ is $H(s,1) = \{(0,h)(s,1) : h \in H\} = \{(hs,h) : h \in H\}$. This identification allows us to view the operator T^*T as acting on the space $\mathbb{C}^S$. Given $f \in \ell_2(A/H)$, we define $f' \in \mathbb{C}^S$ by $f'(s) := f(H(s,1))$.

We trace the random walk in (11) starting from the coset $H(s,1)$:

$$H(s,1) \xrightarrow{x \in H(s,1)} Gx \xrightarrow{y \in Gx} Hy.$$

A generic element in $H(s,1)$ is (hs,h) for a random $h \in H$. This lies in the right coset $G(hs,h) = \{(g+hs,h) : g \in G\}$. A generic element in this right coset is $(g+hs,h)$. We must identify the right coset $H(s',1)$ containing $(g+hs,h)$. Since $(g+hs,h) = (0,h)(h^{-1}(g+hs),1) = (0,h)(h^{-1}g+s,1)$, the new state is $s' = s + h^{-1}g$. Thus, the operator acts on $\mathbb{C}^S$ as:

$$(T^*Tf')(s) = \mathbb{E}_{\substack{g \in G \\ h \in H}} \left[f'(s + h^{-1}g) \right]. \tag{12}$$

Next, for an arbitrary subset $M \subseteq \mathbb{F}$, define its trace-dual subspace by

$$M^\perp \;:=\; \{\, a \in \mathbb{F} : \mathrm{Tr}(am) = 0 \text{ for all } m \in M \,\},$$

where $\mathrm{Tr} = \mathrm{Tr}_{\mathbb{F}/\mathbb{F}_p}$ denotes the field trace from $\mathbb{F}$ to the subfield $\mathbb{F}_p$. Fix a nontrivial additive character $\psi_p : \mathbb{F}_p \to \mathbb{C}^\times$, for instance

$$\psi_p(z) = \exp\left(\frac{2\pi i}{p_0} \mathrm{Tr}_{\mathbb{F}_p/\mathbb{F}_{p_0}}(z) \right).$$

For an $\mathbb{F}_p$-subspace $S \leq (\mathbb{F},+)$, the additive characters of S are indexed by cosets in $\mathbb{F}/S^\perp$. For $a \in \mathbb{F}$ define

$$\chi_a(s) \;:=\; \psi_p(\mathrm{Tr}(as)) \qquad \text{for all } s \in S.$$

Note that χ_a depends only on the coset $a + S^\perp$, since if $a' \equiv a \pmod{S^\perp}$ then $\mathrm{Tr}((a-a')s) = 0$ for all $s \in S$, and hence $\chi_{a'}(s) = \chi_a(s)$.

These characters form an eigenbasis for the operator T^*T. Indeed, fixing a character χ_a and substituting into (12), we obtain

$$(T^*T\,\chi_a)(s) = \mathbb{E}_{g,h}\big[\chi_a(s + h^{-1}g)\big] = \chi_a(s) \cdot \mathbb{E}_{g,h}\big[\chi_a(h^{-1}g)\big],$$

where we used the homomorphism property $\chi_a(x+y) = \chi_a(x)\chi_a(y)$. This implies that the eigenvalue of χ_a is $\lambda_a = \mathbb{E}_{h \in H}\mathbb{E}_{g \in G}[\chi_a(h^{-1}g)]$. The inner expectation is 1 if $h^{-1}a \in G^\perp$ and 0 otherwise. Thus the eigenvalue λ_a equals

$$\lambda_a = \Pr_{h \in H}(h^{-1}a \in G^\perp).$$

The second largest singular value $\sigma_2(T)$ corresponds to the square root of the second largest eigenvalue. We conclude that,

$$\sigma_2(T) = \sqrt{\max_{a \in \mathbb{F} \setminus S^\perp} \Pr_{h \in H}(h^{-1}a \in G^\perp)}. \tag{13}$$

3.3 Explicit Bounds via Character Sums

In this section we bound $\sigma_2(T)$ via character-sum estimates. More specifically, we derive an explicit upper bound on $\sigma_2(T)$ in terms of additive character sums over the multiplicative subgroup H.

Proposition 3.3. *Let*

$$M := \max_{a \in \mathbb{F} \setminus H^\perp} \left| \sum_{h \in H} \chi_a(h) \right|,$$

Then, the second largest singular value of the graph satisfies

$$\sigma_2(T) \leq \sqrt{\frac{1}{p} + \frac{M}{|H|}}.$$

Proof. By (13) we seek to bound $\lambda_a = \frac{1}{|H|}|\{h \in H : h^{-1}a \in G^\perp\}|$ for $a \in \mathbb{F} \setminus S^\perp$. By character orthogonality on the finite additive group G, we have

$$\mathbf{1}_{G^\perp}(y) = \frac{1}{|G|} \sum_{g \in G} \chi_g(y).$$

Indeed, if $y \in G^\perp$, then every term in the sum equals 1. Otherwise, $g \mapsto \chi_g(y)$ is a nontrivial additive character of G, and its sum over G is zero.

Then,

$$\lambda_a = \frac{1}{|H|} \sum_{h \in H} \mathbf{1}_{G^\perp}(h^{-1}a) = \frac{1}{|H|} \sum_{h \in H} \frac{1}{|G|} \sum_{g \in G} \chi_g(h^{-1}a) = \frac{1}{|G|} \sum_{g \in G} \left(\frac{1}{|H|} \sum_{h \in H} \chi_g(h^{-1}a) \right). \tag{14}$$

Note that $G \cap (H \cdot a)^\perp$ is a proper $\mathbb{F}_p$ subspace of G. Otherwise $G \subseteq (H \cdot a)^\perp$, which would imply that $a \in S^\perp$. For each $g \in G \cap (H \cdot a)^\perp$ the inner sum is 1, and there are at most $|G|/p$ such g's, therefore, by the triangle inequality, the RHS of (14) is at most

$$\frac{1}{p} + \frac{1}{|G|} \sum_{g \in G \setminus (H \cdot a)^\perp} \frac{1}{|H|} \left| \sum_{h \in H} \chi_g(h^{-1}a) \right| = \frac{1}{p} + \frac{1}{|G|} \sum_{ga \in G \cdot a \setminus H^\perp} \frac{1}{|H|} \left| \sum_{h \in H} \chi_{ga}(h^{-1}) \right| \leq \frac{1}{p} + \frac{M}{|H|}.$$

Since $\sigma_2(T) = \sqrt{\max_{a \in \mathbb{F} \setminus S^\perp} \lambda_a}$, the result follows. □

4 First Code Instantiation

In this section, we present our first construction of algebraic expander codes as a concrete instantiation of the general framework established in Section 2. We show that for any desired target local rate, this instantiation yields an infinite family of asymptotically good codes with strong spectral expansion, whose local constraints are given by relatively short Reed–Solomon codes.

However, for fixed m and fixed characteristic p_0, the alphabet size grows polynomially with the block length, and the maximum degree of the underlying graph grows polynomially with the number of vertices.

4.1 Code Instantiation

Let $m \geq 2$ be a fixed integer, and let $p = p_0^\ell$ be a power of the fixed prime p_0. We consider the asymptotic regime where m and p_0 are fixed and $p \to \infty$.

We apply the general construction $\mathcal{C}(G, H; r, D)$ given in Section 2.3 with the following parameters.

- **Additive Subgroup (G):** Let $g(X) = X^{p^m} + X^p + X$. We define G to be the set of roots of $g(X)$ in its splitting field. Since $g(X)$ is a p-linearized polynomial with $\gcd(g, g') = 1$, G is an $\mathbb{F}_p$-linear subspace of order $|G| = p^m$. We identify G with the group of translations $\{x \mapsto x + a : a \in G\}$, and then by (4) its associated G-invariant polynomial is $g(X)$.
- **Multiplicative Subgroup (H):** Let $H = \mathbb{F}_{p^m}^\times$ be the multiplicative group of order $p^m - 1$. We identify H with the group of scaling maps $\{x \mapsto hx : h \in H\}$, and then by (5) its associated H-invariant polynomial is $h(X) = X^{|H|}$.

Structure and Code Length. Let $\mathbb{F}$ be a finite extension of $\mathbb{F}_p$ that contains the splitting field of $g(X)$, contains $\mathbb{F}_{p^m}$, and satisfies $|\mathbb{F}| \geq |A|$. Then $A = \langle G, H \rangle$ is a subgroup of $\mathrm{AGL}(1, \mathbb{F})$.

By Theorem 3.1, the group A is isomorphic to $S \rtimes H$, where $S = \sum_{h \in H} hG$. The following lemma shows that $|S| = p^{m^2}$, and therefore the code length is

$$n = |A| = |S| \cdot |H| = p^{m^2}(p^m - 1) = p^{m^2+m}(1 - o(1)),$$

where $o(1)$ tends to zero as $p \to \infty$.

Lemma 4.1. *Let G and H, as defined above, be additive and multiplicative subgroups of the field $\mathbb{F}$, respectively. Then the set $S = \sum_{h \in H} hG$ has size p^{m^2}.*

Proof. See Appendix B. □

Remark 4.2 (Field size)**.** For concreteness, we give a crude bound on the size of the field over which the code can be defined.

The ambient field must contain the splitting field of $g(X) = X^{p^m} + X^p + X$, must contain $\mathbb{F}_{p^m}$, and must have size at least $|A|$.

Let e be the smallest positive integer such that $g(X)$ divides $X^{p^e} - X$. Hence, $\mathbb{F}_{p^e}$ is the splitting field of $g(X)$. The p-associate polynomials of $g(X)$ and $X^{p^e} - X$ are

$$f(Y) = Y^m + Y + 1 \qquad \text{and} \qquad Y^e - 1,$$

respectively. It is well known that, for p-linearized polynomials with coefficients in $\mathbb{F}_p$,

$$g(X) \mid X^{p^e} - X \qquad \Longleftrightarrow \qquad f(Y) \mid Y^e - 1.$$

Thus, e is the order of the polynomial Y in the ring $\mathbb{F}_{p_0}[Y]/(f(Y))$. Indeed, Y is a unit in this ring since it is coprime to $f(Y)$. Since the ring has p_0^m elements, the order of Y is at most p_0^m. Hence $e \leq p_0^m$.

Consequently, if we take $\mathbb{F} = \mathbb{F}_{p^{m(m+1)e}}$, then $\mathbb{F}$ contains the splitting field of $g(X)$, contains $\mathbb{F}_{p^m}$, and satisfies $|\mathbb{F}| \geq |A|$. Since $n = p^{m^2}(p^m - 1) = p^{m(m+1)}(1 - o(1))$, this gives the crude alphabet-size bound

$$|\mathbb{F}| = p^{m(m+1)e} = n^{e + o(1)} \leq n^{p_0^m + o(1)}.$$

Thus, for fixed m and fixed characteristic p_0, the alphabet size is polynomial in the block length.

4.2 Code Parameters

In this section we analyze the rate and relative distance of the code. Recall that $m \geq 2$ is fixed. To simplify the analysis, we work in the asymptotic regime where $p \to \infty$.

Fix $D \leq n$ and $r \in (0,1)$, and let $\mathcal{C} = \mathcal{C}(G, H; r, D)$ as in (8). Define the normalized degree bound $\rho = \lim_{p\to\infty} D/n$, where clearly $\rho \in [0,1]$.

We first provide the parameters of the code in the important case of $D = n$, i.e., $\rho = 1$.

Theorem 4.3. *Let $m \geq 2$ and $r \in (0,1)$ be fixed. The code $\mathcal{C}$ is an explicit family of expander codes of length n which, as $p \to \infty$, has:*

1. *Left and right local code lengths $|G| = p^m$ and $|H| = p^m - 1$ (both scaling as $\Theta(n^{\frac{1}{m+1}})$), with local rate at most $r + o(1)$ and local relative distance at least $1 - r - o(1)$.*
2. *Global relative distance $\delta \geq (1-r)^2$.*
3. *Global rate*
$$R \geq \frac{r^{2m+1}}{(2m+1)!}. \tag{15}$$

The proof of Theorem 4.3 follows from the next theorem by setting $\rho = 1$.

Next, we give the parameters of the code in the general case of $D \leq n$.

Theorem 4.4. *Let $m \geq 2$ be fixed. Let $r \in (0,1)$ and $\rho \in (0,1]$. The code $\mathcal{C}$ is an explicit family of expander codes of length n which, as $p \to \infty$, has:*

1. *Left and right local code lengths $|G| = p^m$ and $|H| = p^m - 1$ (both scaling as $\Theta(n^{\frac{1}{m+1}})$), with local rate at most $r + o(1)$ and local relative distance at least $1 - r - o(1)$.*
2. *Global relative distance $\delta \geq \max(1-\rho, (1-r)^2)$.*
3. *Global rate*
$$R \geq \frac{1}{(C+1)!}\left(r^{C+1} - \max(0, r-\rho)^{C+1}\right),$$
where $C = 2m$.

Proof. The claims about the local rate and relative distance follow from the code construction and the fact that the local codes are RS codes. We proceed to establish the bounds on the global rate and relative distance.

Distance Bound. We utilize both the algebraic definition of the code and the combinatorial properties of its underlying graph.

Algebraic Bound: The global degree constraint ensures that $\deg(f) < D \leq n$ for all $f \in \mathcal{W}(r, D)$. Then, since the evaluation map is injective, the relative distance is bounded by $\delta \geq 1 - D/n$. Asymptotically, this implies that $\delta \geq 1 - \rho$.

Combinatorial Bound: View the code as an expander code with local relative distance $\delta_0 = 1 - r$. In Theorem 4.8 (below) we show that the coset graph $\mathsf{B}(G, H)$ is a strong spectral expander; in particular, the normalized second largest singular value of its bi-adjacency matrix satisfies $\lambda(\mathsf{B}) = O(1/\sqrt{p})$. Hence, the global relative distance of the code satisfies

$$\delta \geq \delta_0(\delta_0 - \lambda(\mathsf{B})).$$

As $p \to \infty$, $\lambda(\mathsf{B}) \to 0$, yielding the asymptotic bound $\delta \geq (1-r)^2$.

Combining these yields $\delta \geq \max(1-\rho, (1-r)^2)$.

Rate Bound. The rate analysis is slightly more technical, as we need to lower bound the dimension of the subspace $\mathcal{W}_{r,D}$. Towards this end we show that many polynomials of the form $g(X)^i X^j$ have small h-base degree.

We begin first with the analysis of the degree of $g(X)$ in h-base . Note that $g(X) = Xh(X) + X^p + X$, and therefore $\deg_h(g) = p$. Next, we establish the h-base degree of g under the Frobenius automorphism.

Lemma 4.5. *Let $k \geq 0$ be an integer. Then,*

$$W_k := \deg_h(g^{p^k}) = \begin{cases} p^{(k \bmod m)+1} & k \bmod m \neq m-1, \\ p^{m-1} & k \bmod m = m-1. \end{cases}$$

Proof. Write $k = am + b$, where $0 \leq b < m$. Let $Q = \frac{p^{am}-1}{p^m-1}$. We use the identity $X^{p^{am}} = X \cdot h(X)^Q$. Thus, $X^{p^k} = (X^{p^{am}})^{p^b} = X^{p^b} h(X)^{Qp^b}$.

We compute $g(X)^{p^k} = (Xh(X) + X^p + X)^{p^k} = X^{p^k} h(X)^{p^k} + X^{p^{k+1}} + X^{p^k}$. Substituting the expressions for X^{p^k}:

$$\begin{aligned} g(X)^{p^k} &= \left(X^{p^b} h(X)^{Qp^b}\right) h(X)^{p^k} + \left(X^{p^{b+1}} h(X)^{Qp^{b+1}}\right) + \left(X^{p^b} h(X)^{Qp^b}\right) \\ &= X^{p^b} \left(h(X)^{Qp^b+p^k} + h(X)^{Qp^b}\right) + X^{p^{b+1}} h(X)^{Qp^{b+1}}. \end{aligned}$$

We analyze the degrees of the coefficients in the h-base expansion.

Case 1: $b < m-1$. The degrees of the coefficients are p^b and p^{b+1}. Since $p^{b+1} < p^m = \deg(h)+1$, these coefficients are already reduced. The maximum degree is $p^{b+1} = p^{(k \bmod m)+1}$.

Case 2: $b = m-1$. The degrees are p^{m-1} and p^m. We must reduce the term X^{p^m} using $X^{p^m} = Xh(X)$ we get

$$g(X)^{p^k} = X^{p^{m-1}} \left(h(X)^{Qp^{m-1}+p^k} + h(X)^{Qp^{m-1}}\right) + (Xh(X))h(X)^{Qp^m}.$$

The degrees of the coefficients are p^{m-1} and 1. The maximum degree is p^{m-1}. □

Note that the weights W_k defined in Lemma 4.5 for $k \geq 0$ form a periodic sequence of period length length m: $p^1, \ldots, p^{m-2}, p^{m-1}, p^{m-1}$. Also, the dominant weight is p^{m-1}.

Next, note that any polynomial of the form $g(X)^i X^j$ such that

1. $0 \leq j < r \cdot \deg(g) = rp^m$
2. $\deg_h\big(g(X)^i X^j\big) < r \cdot \deg(h) = r(p^m - 1)$
3. $\deg\big(g(X)^i X^j\big) = ip^m + j < D$

belongs to the subspace $\mathcal{W}_{r,D}$ (7). Next, we would like to bound the number of such polynomials. Note that all such polynomials are of distinct degree, and therefore linearly independent, which implies a lower bound on the dimension of $\mathcal{W}_{r,D}$.

Next, we define a simplified set of constraints

1. **Local Constraint:** $\sum_k i_k W_k + j < r(p^m - 1)$, where $i = \sum_k i_k p^k$ is the base-p expansion of i.
2. **Global Constraint:** $\deg(g^i X^j) = i \cdot p^m + j = \sum_k i_k p^{k+m} + j < D$.

Note that any polynomial $g(X)^i X^j$ that satisfies the local and global constraints also satisfies the three constraints $(1)-(3)$, above, and therefore they belong to $\mathcal{W}(r,D)$. Indeed, the global constraint is identical to the last constraint (3). Furthermore, if it satisfies the local constraint, then necessarily $j < rp^m$. Lastly,

$$\deg_h(g(X)^i X^j) = \deg_h\left(\prod_k \left(g(X)^{p^k}\right)^{i_k} X^j\right) \le \sum_k i_k W_k + j < r(p^m-1),$$

where the first inequality follows from the subadditivity of the h-base degree (Lemma 2.2) and Lemma 4.5.

Next, we lower bound the number of monomials that satisfy the local and global constraints. The global degree constraint implies $i < D/p^m \approx \rho p^{m^2}$, restricting the base-$p$ expansion of i to indices $0 \le k < m^2$.

We define the normalized variables $x_k := i_k/p \in [0,1)$ and $z := j/p^m \ge 0$, and analyze the constraints asymptotically as $p \to \infty$.

1. **Asymptotic local constraint:** Let I_{dom} be the set of indices corresponding to the dominant weight p^{m-1}, namely

$$I_{\text{dom}} := \{\, 0 \le k < m^2 : W_k = p^{m-1} \,\}, \qquad \text{and } |I_{\text{dom}}| = 2m.$$

   Let $C := 2m$. Dividing the local constraint by p^m and taking the limit $p \to \infty$ yields

$$\sum_{k \in I_{\text{dom}}} x_k + z < r. \tag{16}$$

2. **Asymptotic global constraint:** Let $k_{\max} = m^2 - 1$. Dividing the global constraint by p^{m^2+m} and taking the limit $p \to \infty$ gives

$$x_{k_{\max}} < \rho. \tag{17}$$

Note that $k_{\max} \in I_{\text{dom}}$, since by Lemma 4.5 we have $W_{k_{\max}} = W_{m^2-1} = p^{m-1}$. Thus, the asymptotic global constraint (17) restricts a variable that also appears in the asymptotic local constraint (16).

To lower bound the dimension of $\mathcal{W}_{r,D}$, we count the number of valid discrete tuples $(j, \{i_k\})$. As $p \to \infty$, this discrete counting problem can be approximated by a continuous volume via a standard multidimensional Riemann-sum argument. The $m^2 - C$ subdominant variables x_k (for $k \notin I_{\text{dom}}$) vanish from the asymptotic constraints (16) and (17). Thus, they are asymptotically unconstrained and contribute a factor of 1 to the normalized volume. The fraction of valid tuples therefore converges exactly to the continuous volume of the polytope $\mathcal{P}$ defined by the dominant variables $\{x_k\}_{k \in I_{\text{dom}}}$ and z. Since the total number of unconstrained tuples is $p^m \cdot p^{m^2} = p^{m^2+m}$, the number of admissible monomials is asymptotically $p^{m^2+m}\text{Vol}(\mathcal{P})$.

The next lemma computes the volume of $\mathcal{P}$.

Lemma 4.6. *The volume of the polytope $\mathcal{P}$ defined by constraints (16) and (17) is*

$$\mathit{Vol}(\mathcal{P}) = \frac{1}{(C+1)!}\left(r^{C+1} - \max(0, r-\rho)^{C+1}\right).$$

Proof. See Appendix C. □

Since the code length is $n = p^{m^2+m}(1-o(1))$, dividing the asymptotically number of admissible monomials $p^{m^2+m}\mathrm{Vol}(\mathcal{P})$ by n shows that the global rate R as $p \to \infty$ is bounded below by $\mathrm{Vol}(\mathcal{P})$. Hence, by Lemma 4.6, we obtain that

$$R \geq \frac{1}{(C+1)!}\left(r^{C+1} - \max(0, r-\rho)^{C+1}\right).$$

□

From Theorem 4.4 we identify three regions where the parameters of the code behave differently, depending on the relation between r and ρ. We summarize it in the following corollary.

Corollary 4.7. *For a fixed $m \geq 2$, a local rate $r \in (0,1)$ and $\rho \in (0,1)$, the code $\mathcal{C}$ satisfies*

$$R \geq \begin{cases} \frac{r^{C+1}-(r-\rho)^{C+1}}{(C+1)!} & \rho \leq r, \\ \frac{r^{C+1}}{(C+1)!} & r < \rho. \end{cases}$$

$$\delta \geq \begin{cases} 1-\rho & \rho \leq 2r - r^2, \\ (1-r)^2 & 2r - r^2 < \rho. \end{cases}$$

By inspecting Corollary 4.7, we observe an unsatisfying behavior of our rate bound. Intuitively, one expects that increasing ρ should increase the global rate R of the code. However, once ρ exceeds r, our lower bound on the rate no longer improves. This raises the question of whether the bound is simply not tight, and in fact the true rate is increasing in ρ throughout the admissible range, or whether the global rate indeed does not increase beyond this point.

4.3 Spectral Expansion

In this section we establish the final ingredient needed to conclude that our code is an expander code, namely that the coset graph associated with the instantiation of Section 4.1 is expanding. In particular, we prove that this graph is a strong spectral expander, with second singular value on the order of $O(1/\sqrt{p})$.

Theorem 4.8. *The second largest singular value of the normalized adjacency operator of the coset graph associated with the code instantiation in Section 4.1 satisfies,*

$$\sigma_2(T) \leq \sqrt{\frac{1}{p} + \frac{1}{p^m - 1}}.$$

Proof. The result will follow by Proposition 3.3, and by showing that

$$\max_{a \in \mathbb{F} \setminus H^\perp} \left| \sum_{h \in H} \chi_a(h) \right| = 1.$$

Indeed, for any $a \in \mathbb{F} \setminus H^\perp$, orthogonality of additive characters over $\mathbb{F}_{p^m}$ gives

$$\sum_{h \in \mathbb{F}_{p^m}} \chi_a(h) = 0,$$

so

$$\left| \sum_{h \in H} \chi_a(h) \right| = \left| \sum_{h \in \mathbb{F}_{p^m}} \chi_a(h) - \chi_a(0) \right| = 1.$$

□

5 Second Code Instantiation

In this section, we give another instantiation of the code construction from Section 2. Recall that in the first instantiation (Section 4.1) the two degrees in the graph are as close as possible (they differ by one). Moreover, it exploits the choice $H = \mathbb{F}_{p^m}^\times$, which yields an especially clean and strong character-sum bound and hence excellent spectral expansion. However, this choice is relatively rigid, as it essentially fixes the right degree to be of the form $|H| = p^m - 1$, since H must be the full multiplicative group of a field.

Furthermore, the field over which the code is defined must contain the splitting field of the linearized polynomial $g(X)$ that defines G, which can be large.

In this section we present a second instantiation that addresses these two caveats, thereby exemplifying the flexibility of the general construction. In particular, in this instantiation we take $G = \mathbb{F}_{p^m}$ (hence algebraically simpler than in the first instantiation) and choose H to be a large cyclic subgroup of $\mathbb{F}_{p^{m+1}}^\times$ of constant index $1/\gamma$. This allows us to tune the right degree $|H| = \gamma(p^{m+1} - 1)$ while still obtaining strong spectral expansion via standard Gauss-sum estimates for multiplicative subgroups. The result is an explicit expander-code family with the same qualitative distance guarantees, together with a parameter γ that controls the locality/degree tradeoff on the H-side.

We proceed with the code instantiation.

5.1 Code Instantiation

Fix an integer $m \geq 2$ and let $p = p_0^\ell$ tend to infinity along the sequence specified below. Let $q := p^{m+1}$. We apply the general construction $\mathcal{C}(G, H; r, D)$ given in Section 2.3 with the following parameters.

- **Additive Subgroup:** Let $G := \mathbb{F}_{p^m}$.
- **Multiplicative Subgroup:** Let $\gamma = 1/a \in (0, 1)$, where a is a fixed positive integer with $\gcd(a, p_0) = 1$. We restrict to powers $p = p_0^\ell$ such that $a \mid p^{m+1} - 1$[2]. Let $H \leq \mathbb{F}_q^\times$ be the unique cyclic subgroup of order

$$|H| = \gamma \cdot (p^{m+1} - 1).$$

  The divisibility condition ensures that $|H|$ is an integer.

As before, we identify G with the translation group $\{x \mapsto x + g : g \in G\}$ and H with the scaling group $\{x \mapsto hx : h \in H\}$. The associated invariant polynomials are therefore

$$g(X) = \prod_{a \in \mathbb{F}_{p^m}} (X - a) = X^{p^m} - X, \text{ and } h(X) = X^{|H|}.$$

Structure and Code Length. Let $\mathbb{F}$ be a large enough field extension of $\mathbb{F}_p$ such that the group $A = \langle G, H \rangle$ is a subgroup of $\mathrm{AGL}(1, \mathbb{F})$, and $|\mathbb{F}| \geq |A|$. Let $S := \sum_{h \in H} hG$ be the additive closure of G under the action of H. The following lemma characterizes S.

Lemma 5.1. *The set S satisfies*

$$S = \mathbb{F}_{p^{m(m+1)}}.$$

[2]There are infinitely many such powers. Indeed, if t is the multiplicative order of p_0 modulo a, then every ℓ divisible by t gives $p = p_0^\ell \equiv 1 \pmod{a}$, and hence $a \mid p^{m+1} - 1$.

Proof. For all sufficiently large p, the subgroup H is not contained in any proper subfield of $\mathbb{F}_{p^{m+1}}$. Indeed, every proper subfield of $\mathbb{F}_{p^{m+1}}$ has size at most p^m, whereas

$$|H| = \gamma(p^{m+1} - 1) > p^m - 1$$

for all sufficiently large p. Hence $\mathbb{F}_p(H) = \mathbb{F}_{p^{m+1}}$. Then, by Lemma 3.2

$$S = \mathrm{Span}_{\mathbb{F}_p(H)}(G) = \mathrm{Span}_{\mathbb{F}_{p^{m+1}}}(G) = \mathrm{Span}_{\mathbb{F}_{p^{m+1}}}(\mathbb{F}_{p^m}) = \mathbb{F}_{p^{m(m+1)}},$$

where the last equality follows since $\gcd(m, m+1) = 1$, $\mathbb{F}_{p^m} \cap \mathbb{F}_{p^{m+1}} = \mathbb{F}_p$. Hence any $\mathbb{F}_p$-basis of $\mathbb{F}_{p^m}$ is $\mathbb{F}_{p^{m+1}}$-linearly independent, so $\dim_{\mathbb{F}_{p^{m+1}}}\left(\mathrm{Span}_{\mathbb{F}_{p^{m+1}}}(\mathbb{F}_{p^m})\right) = m$. But $[\mathbb{F}_{p^{m(m+1)}} : \mathbb{F}_{p^{m+1}}] = m$, so this span is an m-dimensional $\mathbb{F}_{p^{m+1}}$-subspace of $\mathbb{F}_{p^{m(m+1)}}$, and therefore equals the whole field. □

Recall that the code length $n = |A| = |S||H|$, where $A = \langle G, H\rangle \cong S \rtimes H$. Hence, by Lemma 5.1

$$n = |S||H| = p^{m(m+1)} \cdot \gamma \cdot (p^{m+1} - 1) = \gamma \cdot p^{(m+1)^2} \cdot (1 - o(1)),$$

where $o(1)$ tends to zero as $p \to \infty$.

Remark 5.2 (Field size)**.** For concreteness, we show that the field over which the code is defined can be taken to be $\mathbb{F} = \mathbb{F}_{p^{2m(m+1)}}$.

Indeed, since $S = \mathbb{F}_{p^{m(m+1)}}$ must be contained in $\mathbb{F}$, the extension degree $[\mathbb{F} : \mathbb{F}_p]$ must be a multiple of $m(m+1)$. The choice $\mathbb{F} = \mathbb{F}_{p^{m(m+1)}}$ is too small, because it has size $|S| < |A| = |S||H|$. The next possible extension degree is therefore $2m(m+1)$, and $p^{2m(m+1)} \geq |A|$ for all sufficiently large p. Thus the field $\mathbb{F}_{p^{2m(m+1)}}$ satisfyies both requirements: it contains S and it has enough elements to choose an orbit of size $|A|$.

We conclude that the resulting code has length $n = \Theta\left(p^{(m+1)^2}\right)$, over a field of size $|\mathbb{F}| = p^{2m(m+1)} = O\left(n^{2m/(m+1)}\right)$.

Choose an element $\alpha \in \mathbb{F}$ that is not stabilized by any nontrivial affine transformation in A, and let

$$\Omega := \{a(\alpha) : a \in A\} \subseteq \mathbb{F},$$

be the orbit of α under the action of A. We take this orbit as the evaluation set of the code.

5.2 Code Parameters

Fix $r \in (0, 1)$ and $D \leq n$, and let $\rho = \lim_{p\to\infty} D/n \in (0, 1]$. Let $\mathcal{C} = \mathcal{C}(G, H; r, D)$ be the code from (8).

We first provide the parameters of the code in the important case of $D = n$, i.e., $\rho = 1$.

Theorem 5.3. *The code $\mathcal{C}$ is an explicit family of expander codes of length n which, as $p \to \infty$, has:*

1. *Left local code length $|G| = p^m = \Theta(n^{\frac{m}{(m+1)^2}})$ and right local code length $|H| = \gamma(p^{m+1} - 1) = \Theta(n^{\frac{1}{m+1}})$, with local rate at most $r + o(1)$ and local relative distance at least $1 - r - o(1)$.*
2. *Global relative distance $\delta \geq (1 - r)^2$.*
3. *Global rate bounded below by*

$$R \geq \frac{\gamma^{2m}\, r^{2m+2}}{(2m+1)!}. \tag{18}$$

The proof of Theorem 5.3 follows from Theorem 5.4 below by setting $\rho = 1$.

The following theorem summarizes the parameters of the code in the general case of $D \leq n$.

Theorem 5.4. *The code $\mathcal{C}$ is an explicit family of expander codes of length n which, as $p \to \infty$, has:*

1. *Left local code length $|G| = p^m = \Theta(n^{\frac{m}{(m+1)^2}})$ and right local code length $|H| = \gamma(p^{m+1} - 1) = \Theta(n^{\frac{1}{m+1}})$, with local rate at most $r + o(1)$ and local relative distance at least $1 - r - o(1)$.*
2. *Global relative distance $\delta \geq \max(1 - \rho, (1 - r)^2)$.*
3. *Global rate bounded below by*

$$R \geq \frac{\gamma^{2m} r}{(2m+1)!} \left(r^{2m+1} - \max(0, r - \rho)^{2m+1}\right).$$

Proof. Items (1) and the RS local distance bound are immediate from the locality analysis in Section 2.4.

Distance Bound. For (2), we combine the algebraic bound $\delta \geq 1 - \rho$ (since $\deg(f) < D$ for $f \in \mathcal{W}_{r,D}$ and evaluation is injective) with the standard expander-code distance bound applied to the underlying coset graph. By Theorem 5.8 below, the graph satisfies $\sigma_2(T) = o(1)$ as $p \to \infty$, and hence asymptotically $\delta \geq (1 - r)^2$. Combining the two bounds gives $\delta \geq \max(1 - \rho, (1 - r)^2)$.

Rate Bound. For (3), we lower bound $\dim(\mathcal{W}_{r,D})$ by exhibiting many polynomials of the form $g(X)^i X^j$ that are in $\mathcal{W}_{r,D}$ and they all have distinct degrees.

Consider polynomials $g(X)^i X^j$ with $0 \leq j < rp^m$ and $ip^m + j < D$. Then, clearly $\deg_g(g^i X^j) = j < r|G|$. Write $i = \sum_{k=0}^{k_{\max}} i_k p^k$ with $0 \leq i_k < p$, where $k_{\max} = m(m+1)$ since $i < D/p^m \leq p^{m^2+m+1}$.

For $k \geq 0$ define the weight

$$W_k := \deg_h\left(g^{p^k}\right),$$

to be the h-base degree of the polynomial g^{p^k}. The following lemma gives a precise description of the weights.

Lemma 5.5. *For a nonnegative integer k*

$$W_k = \begin{cases} p^m & k \bmod (m+1) = 0, \\ p^{k \bmod (m+1)} & \textit{else.} \end{cases}$$

By the subadditivity of the h-base degree (Lemma 2.2),

$$\deg_h(g^i X^j) \;\leq\; j + \sum_k i_k W_k.$$

Thus it suffices to consider the simplified constraint

$$j + \sum_k i_k W_k < r|H|.$$

We conclude that any polynomial of the form $g(X)^i X^j$ such that

1. $j < rp^m$

2. **Global Constraint:** $ip^m + j < D$

3. **Local Constraint:** $j + \sum_k i_k W_k < r|H|$,

is in $\mathcal{W}_{r,D}$. Next, we give a lower bound on the number of such polynomials.

Let $I_{\text{dom}} := \{0 \le k \le k_{\max} : k \bmod (m+1) \in \{0, m\}\}$. By Lemma 5.5, $W_k = p^m$ for $k \in I_{\text{dom}}$ and $W_k \le p^{m-1}$ otherwise. Clearly, $|I_{\text{dom}}| = 2m + 1$.

We define the normalized variables $x_k := i_k/p \in [0, 1)$ for $k \in I_{\text{dom}}$, and

$$z := j/p^m \in [0, r), \tag{19}$$

and analyze the constraints asymptotically as $p \to \infty$.

By dividing the local constraint by p^{m+1} and letting $p \to \infty$ we get the asymptotic local constraint (as subdominant terms vanish)

$$\sum_{k \in I_{\text{dom}}} x_k \;<\; r\gamma, \tag{20}$$

since $|H|/p^{m+1} \to \gamma$ and $j/p^{m+1} \to 0$.

Similarly, by dividing the global constraint by p^{m^2+2m+1} and letting $p \to \infty$ we get the asymptotic global constraint

$$x_{k_{\max}} < \rho\gamma. \tag{21}$$

As before, we approximate the discrete count of valid tuples via a standard Riemann-sum argument. The $m^2 - m$ subdominant variables x_k vanish from the asymptotic limits and therefore are unconstrained in the interval $[0, 1)$, contributing a factor of 1 to the volume. Thus, the fraction of valid tuples converges to the geometric volume of the polytope $\mathcal{P}$ defined by the constraints (19), (20) and (21) on the dominant variables $\{x_k\}_{k \in I_{\text{dom}}}$ and z. Since the total number of unconstrained tuples is $p^m \cdot p^{m^2+m+1} = p^{m^2+2m+1}$, the asymptotic dimension of $\mathcal{W}_{r,D}$ is $p^{m^2+2m+1}\text{Vol}(\mathcal{P})$.

The next lemma computes the volume of $\mathcal{P}$.

Lemma 5.6. *The volume of the polytope $\mathcal{P}$ defined by the constraints (19), (20) and (21) is*

$$\text{Vol}(\mathcal{P}) = \frac{\gamma^{2m+1} r \left(r^{2m+1} - \max(0, r-\rho)^{2m+1}\right)}{(2m+1)!}.$$

Proof. See Appendix D. □

Since the block length is $n \sim \gamma p^{m^2+2m+1}$, dividing the asymptotic dimension $p^{m^2+2m+1}\text{Vol}(\mathcal{P})$ by n yields the stated lower bound on the global rate $R \ge \frac{1}{\gamma}\text{Vol}(\mathcal{P})$. □

Next, we prove Lemma 5.5.

Proof of Lemma 5.5. Write $k = a(m+1) + t$, where $t = k \bmod (m+1)$. Since $|H|$ divides $p^{m+1} - 1$, we have the identity

$$X^{p^{m+1}} = X^{p^{m+1}-1} \cdot X = X^{|H| \cdot \frac{p^{m+1}-1}{|H|}} \cdot X = X \cdot h(X)^{\frac{p^{m+1}-1}{|H|}} = X \cdot h(X)^{\frac{1}{\gamma}}.$$

Then by induction on a, for any positive integer a we have

$$X^{p^{a(m+1)}} = X \cdot h(X)^{(\text{some integer})}.$$

Raising both sides to p^t shows that X^{p^k} can be written in base h as

$$X^{p^k} = X^{p^t} \cdot h(X)^{(\text{some integer})},$$

with coefficient degree $\deg(X^{p^t}) = p^t \le p^m < \deg(h)$.

Now

$$W_k = \deg_h\left(g^{p^k}\right) = \deg_h\left(X^{p^{m+k}} - X^{p^k}\right). \tag{22}$$

If $t = 0$, then $k = a(m+1)$ and $m + k = a(m+1) + m$, so the RHS of (22) becomes

$$\deg_h\left(X^{p^m} \cdot h(X)^{(\text{some integer})} - X \cdot h(X)^{(\text{some integer})}\right) = p^m.$$

If $1 \le t \le m$, then $m + k = a(m+1) + m + t = (a+1)(m+1) + (t-1)$, so the RHS of (22) becomes

$$\deg_h\left(X^{p^{(t-1)}} \cdot h(X)^{(\text{some integer})} - X^{p^t} \cdot h(X)^{(\text{some integer})}\right) = p^t.$$

□

5.3 Spectral Expansion via Gauss Sums

In this section we show that the coset graph underlying the code instantiation of Section 5.1 is a strong spectral expander, with second singular value on the order of $O(1/\sqrt{p})$. This expansion bound follows by applying Proposition 3.3.

To this end, it suffices to bound the additive character sums

$$\left|\sum_{h \in H} \chi_a(h)\right| \qquad \text{for } a \notin H^{\perp}.$$

We obtain the required estimate via a standard application of Gauss sums, as we now explain.

Lemma 5.7. *Let $H \le \mathbb{F}_q^{\times}$ be a multiplicative subgroup. Then for every nontrivial additive character ψ of $\mathbb{F}_q$,*

$$\left|\sum_{h \in H} \psi(h)\right| \;\le\; \sqrt{q}.$$

Proof. Let $s = (q-1)/|H|$ be the index of H, and fix a multiplicative character χ of $\mathbb{F}_q^{\times}$ of exact order s. Then, the indicator of H satisfies

$$\mathbf{1}_H(x) = \frac{1}{s}\sum_{j=0}^{s-1} \chi(x)^j \qquad (x \in \mathbb{F}_q^{\times}).$$

Hence

$$\sum_{h \in H} \psi(h) = \sum_{x \in \mathbb{F}_q^{\times}} \mathbf{1}_H(x)\psi(x) = \frac{1}{s}\sum_{j=0}^{s-1} \sum_{x \in \mathbb{F}_q^{\times}} \chi(x)^j \psi(x).$$

The $j = 0$ term equals $\sum_{x \in \mathbb{F}_q^{\times}} \psi(x) = -1$ since ψ is nontrivial. For $j \ne 0$ the inner sum is a Gauss sum and has magnitude $\sqrt{q}$. Thus

$$\left|\sum_{h \in H} \psi(h)\right| \le \frac{1}{s}\Big(1 + (s-1)\sqrt{q}\Big) \le \sqrt{q}.$$

□

Theorem 5.8. *Let T be the normalized bi-adjacency operator of the coset graph $\mathsf{B}(G, H)$. Then*

$$\sigma_2(T) \;\le\; \sqrt{\frac{1}{p} + \frac{1+o(1)}{\gamma\, p^{(m+1)/2}}} \;=\; \frac{1+o(1)}{\sqrt{p}}.$$

Proof. Let χ_a be an additive character of $\mathbb{F}$ with $a \in \mathbb{F} \setminus H^{\perp}$. Therefore, χ_a is a nontrivial character since $a \notin H^{\perp}$. Hence, by Lemma 5.7 we have

$$\left|\sum_{h\in H} \chi_a(h)\right| \;\le\; \sqrt{q} \Rightarrow M = \max_{a\in\mathbb{F}\setminus H^{\perp}} \left|\sum_{h\in H} \chi_a(h)\right| \;\le\; \sqrt{q}.$$

Plugging into Proposition 3.3 gives

$$\sigma_2(T) \le \sqrt{\frac{1}{p} + \frac{M}{|H|}} \le \sqrt{\frac{1}{p} + \frac{\sqrt{q}}{|H|}}.$$

Finally use $|H| = \gamma \cdot (q-1)$ to obtain the asymptotic form. □

6 Decoding

In this section we state decoding guarantees for the Tanner-code regime $D = n$. The subcodes with $D < n$ inherit the same local Tanner structure and also have additional global algebraic structure; we briefly discuss the resulting decoding possibilities in Remark 6.6.

Recall that the construction admits two complementary descriptions: a combinatorial Tanner-code description and an algebraic orbit-evaluation description. Each viewpoint naturally gives rise to a different decoding algorithm.

First, let $\mathcal{C}_r = \mathcal{C}(G, H, r, n)$ denote the Tanner code with local rate r on the coset graph $B(G, H)$, whose local constraints are RS codes on the G- and H-orbits. As a Tanner code on an expander graph, $\mathcal{C}_r$ can be decoded using the standard iterative decoding algorithm for expander codes [19]. We refer to this as the *iterative combinatorial decoder*.

Second, the family $\{\mathcal{C}_r\}_{0<r<1}$ possesses additional algebraic structure beyond its Tanner-code description. Specifically, it satisfies the multiplication property $\mathcal{C}_r * \mathcal{C}_s \subseteq \mathcal{C}_{r+s}$, (see Proposition 2.7) which gives rise to a different decoding algorithm via the framework of error-correcting pairs introduced by Pellikaan [12]. We refer to this as the *global algebraic decoder*. Conceptually, it can be viewed as a Berlekamp–Welch-type decoder for the hierarchy of algebraic expander codes $\{\mathcal{C}_r\}_{0<r<1}$.

The iterative combinatorial decoder has better performance, both in terms of running time and decoding radius, and is therefore the decoder of choice. Nevertheless, we include the global algebraic decoder because it demonstrates that the multiplication property is not only structural, but also has direct algorithmic consequences. Furthermore, the picture is less clear for the subcodes obtained by imposing an additional global degree constraint $D < n$ (see Remark 6.6).

Throughout, saying that a decoder has relative decoding radius τ means that it can correct η fraction of errors for $\eta < \tau$.

6.1 Combinatorial decoding via the expander structure

We first use the Tanner-code structure of $\mathcal{C}_r$ and apply the standard GMD iterative decoding algorithm for expander codes.

Theorem 6.1 (GMD expander-code decoding; see [20, 16, 10])**.** *Let $\mathcal{C}$ be a Tanner code on a biregular graph with n edges, normalized second singular value λ, and maximum degree Δ. Suppose that all local codes have relative distance at least δ_0, and let $\delta = \delta_0(\delta_0 - \lambda)$.*

Then $\mathcal{C}$ has an iterative decoder with decoding radius $\delta/2$. Moreover, if the local codes are RS codes, then the decoder can be implemented in time $\widetilde{O}(n\Delta)$ field operations.

By applying Theorem 6.1 directly we get the following result.

Theorem 6.2 (Iterative decoding)**.** *The algebraic expander code $\mathcal{C}_r$ from Theorem 4.3 or Theorem 5.3 has an iterative decoder with decoding radius $(1-r)^2/2$ and running time $\widetilde{O}\left(n^{1+\frac{1}{m+1}}\right)$ field operations.*

Proof. The local codes of $\mathcal{C}_r$ are RS codes of relative distance at least $1 - r - o(1)$ and normalized second singular value $\lambda = o(1)$. Therefore, by Theorem 6.1 the decoding radius is $\delta/2 = (1-r)^2/2$.

In both code instantiations $\Delta = \Theta(n^{1/(m+1)})$ therefore the running time is $\widetilde{O}(n\Delta) = \widetilde{O}\left(n^{1+\frac{1}{m+1}}\right)$. □

6.2 Algebraic decoding via error-correcting pairs

We now explain how the multiplication property gives a decoding algorithm using the framework of error-correcting pairs by Pellikaan [12].

For a linear code $\mathcal{C}$ let $\dim(\mathcal{C})$ and $d(\mathcal{C})$ denote its dimension and minimum distance, respectively. We will need the following result.

Theorem 6.3. *[12, Theorem 2.14] Let $\mathcal{A}, \mathcal{B}, \mathcal{C} \subseteq \mathbb{F}^n$ be linear codes and $t > 0$ an integer such that*

$$\mathcal{A} * \mathcal{B} \subseteq \mathcal{C}^\perp, \qquad t < \min\{\dim(\mathcal{A}), d(\mathcal{B}^\perp)\}, \qquad d(\mathcal{A}) + d(\mathcal{C}) > n.$$

Then any t errors in $\mathcal{C}$ can be corrected with complexity $O(n^3)$ field operations.

We apply Theorem 6.3 to the code $\mathcal{C}_r$ and obtain the following result.

Theorem 6.4. *The algebraic expander code $\mathcal{C}_r$ from Theorem 4.3 or Theorem 5.3 can be decoded from a decoding radius*

$$\tau := \sup_{0<s<1-\sqrt{2r-r^2}} \min\left\{\underline{R}(s), (1-r-s)^2\right\},$$

with complexity $O(n^3)$, where $\underline{R}(s) := \max\{R(s), 2s-1\}$ and $R(s)$ is the code rate lower bound given in (15) or (18).

Proof. Note first that $\underline{R}(s)$ is indeed a lower bound on the rate of $\mathcal{C}_s$, since the standard constraint-counting argument also yields the lower bound $2s-1$ (see Remark 2.5). We now apply Theorem 6.3.

Fix $s > 0$ with $r + s < 1$, and define

$$\mathcal{A} = \mathcal{C}_s, \qquad \mathcal{B} = \mathcal{C}_{r+s}^\perp, \qquad \mathcal{C} = \mathcal{C}_r.$$

We verify the hypotheses of Theorem 6.3.

Let $a \in \mathcal{C}_s$, $b \in \mathcal{C}_{r+s}^\perp$, and $c \in \mathcal{C}_r$. Then

$$\langle a * b, c\rangle = \langle b, a * c\rangle = 0,$$

where the last equality follows from the multiplication property (Proposition 2.7), since $a * c \in \mathcal{C}_s * \mathcal{C}_r \subseteq \mathcal{C}_{r+s}$. Hence $a * b \in \mathcal{C}_r^\perp$, and therefore $\mathcal{A} * \mathcal{B} \subseteq \mathcal{C}^\perp$.

Now let $t = \lfloor \eta n \rfloor$ for some $\eta < \tau$. The condition

$$t < \min\{\dim(\mathcal{A}), d(\mathcal{B}^\perp)\} = \min\{\dim(\mathcal{C}_s), d(\mathcal{C}_{r+s})\}$$

holds because the rate of $\mathcal{C}_s$ is at least $\underline{R}(s) - o(1)$, while the relative distance of $\mathcal{C}_{r+s}$ is at least $(1-r-s)^2 - o(1)$. It remains to verify that $d(\mathcal{A}) + d(\mathcal{C}) > n$. Since $\mathcal{A} = \mathcal{C}_s$ and $\mathcal{C} = \mathcal{C}_r$,

$$\frac{d(\mathcal{A}) + d(\mathcal{C})}{n} \geq (1-s)^2 + (1-r)^2 - o(1).$$

Thus $d(\mathcal{A}) + d(\mathcal{C}) > n$ whenever $(1-s)^2 + (1-r)^2 > 1$, or equivalently, $s < 1 - \sqrt{2r - r^2}$.

Therefore, Theorem 6.3 implies that, for every such s, one can correct t errors. Optimizing over s yields the claimed decoding radius. □

Remark 6.5. The iterative combinatorial decoder is stronger. It corrects up to essentially $\frac{(1-r)^2}{2}$ fraction of errors, while the algebraic decoder is limited by the dimension of $\mathcal{C}_s$ and the distance of $\mathcal{C}_{r+s}$. For example, in the very low-rate regime $r \to 0$, the optimized algebraic decoding radius tends to $3 - 2\sqrt{2} \approx 0.171$, whereas the iterative decoder tends to radius $1/2$.

Remark 6.6. For $D < n$, the code $\mathcal{C}_{r,D}$ is a subcode of the Tanner code $\mathcal{C}_{r,n}$, so the iterative expander decoder still applies without modification. At the same time, the additional degree constraint gives extra algebraic structure. In particular, the multiplication property continues to hold in the form

$$\mathcal{C}_{r_1,D_1} * \mathcal{C}_{r_2,D_2} \subseteq \mathcal{C}_{r_1+r_2,\min\{n,D_1+D_2\}},$$

and therefore one can also build error-correcting-pair decoders using auxiliary error-correcting pairs $\mathcal{C}_{s,D_A}$ as in Theorem 6.3 . Optimizing over the additional degree parameter D_A, and comparing the resulting algebraic decoder with the purely iterative decoder and the ambient Reed–Solomon decoder, leads to a two-parameter decoding problem. We do not pursue this optimization here, and focus instead on the cleaner Tanner-code case $D = n$.

7 Concluding Remarks and Open Questions

In this work, we presented a general framework for constructing algebraic expander codes utilizing the action of subgroups of the affine group $\mathrm{AGL}(1, \mathbb{F})$. By instantiating this framework with non-commuting translation and scaling subgroups, we derived explicit families of codes that circumvent the classical rate barrier for expander codes. Specifically, our constructions achieve a positive global rate even when the local code rate satisfies $r \leq 1/2$, a regime where generic expander constructions typically fail.

We conclude by highlighting several open problems and directions for future research arising from this construction.

1. **Tightness of the Rate Bound.** The lower bound on the global rate $R(r, m, \rho)$ derived in Theorems 4.4 and 5.4 establishes the existence of asymptotically good codes, but it exhibits asymptotic behaviors that suggest it is far from tight.

    - *Saturation:* The bound saturates once the normalized global degree ρ exceeds the local rate r. Intuitively, increasing the allowed global degree D (and thus ρ) should strictly increase the dimension of the message space $\mathcal{W}_{r,D}$, yet our polytope volume argument remains constant in this regime.

- *Vanishing with Sparsity:* As the sparsity parameter m increases, our lower bound vanishes factorially (scaling roughly as $1/(2m)!$). However, in the regime where the local rate satisfies $r > 1/2$, the standard constraint-counting bound for expander codes (Remark 2.5) guarantees a rate of at least $2r - 1$. This standard bound is independent of the graph degree (and thus independent of m). The fact that our algebraic bound approaches zero for large m, even when $r > 1/2$, indicates that our analysis, which relies on the simultaneous satisfaction of g-base and h-base degree constraints, likely underestimates the true dimension of the code. Bridging the gap between our bound and the constant $2r - 1$ baseline is a key challenge.

2. **Optimizing Code Parameters (Degree and Expansion).** The coset graphs constructed in this work have a degree that grows polynomially with the block length (at most $n^{1/(m+1)}$ in the instantiations above). While this improves upon the linear density of commuting-subgroup constructions, achieving *constant* degree (or polylogarithmic degree) remains a major goal for algebraic expander codes.

   - Is it possible to instantiate this framework with different subgroups of $\mathrm{AGL}(1, \mathbb{F})$ (or potentially $\mathrm{AGL}(k, \mathbb{F})$) to yield constant-degree graphs?
   - Furthermore, while we proved that the constructed graphs are strong spectral expanders, determining whether they can achieve the optimal Ramanujan bound is of interest.

3. **Local Testability.** The codes constructed here share structural similarities with affine-invariant codes and lifted codes, which are known to support local testing. Given that these codes are defined by local Reed–Solomon constraints on an expander graph, it is natural to ask whether they are also locally testable.

4. **High-Dimensional Expanders (HDX).** Our construction utilizes the interaction between two subgroups, G and H, to define a bipartite graph (a 1-dimensional simplicial complex) where edges correspond to group elements and vertices correspond to cosets. A natural generalization is to consider the geometry generated by *three* subgroups $G_1, G_2, G_3 \leq \mathrm{AGL}(1, \mathbb{F})$. Can such a construction define a 2-dimensional simplicial complex (e.g., where faces correspond to elements of $\langle G_1, G_2, G_3 \rangle$ and vertices to cosets of the G_i's) that exhibits high-dimensional expansion? If so, placing local codes on the faces of this complex could lead to new algebraic constructions of HDX codes.

A Proofs of Section 3

Proof of Theorem 3.1. We begin by showing that S is an H-invariant subspace. By definition, S is an additive group. Next, we show that S is closed under multiplication by H. For any $s = \sum h_k g_k \in S$ and any $h' \in H$, we have $h's = \sum (h'h_k)g_k$. Since H is a group, $h'h_k \in H$, so $h's$ is a sum of elements in $H \cdot G$.

Next, let $\mathcal{A} = \{\phi_{s,h} : x \mapsto hx + s \mid h \in H, s \in S\}$. We will show that $\mathcal{A}$ is a group and that $\mathcal{A} = A$.

$\mathcal{A}$ is a Group. Consider the composition of $\phi_{s_1,h_1}, \phi_{s_2,h_2} \in \mathcal{A}$:

$$(\phi_{s_1,h_1} \circ \phi_{s_2,h_2})(x) = h_1(h_2x + s_2) + s_1 = (h_1h_2)x + (h_1s_2 + s_1).$$

Since H is a group, $h_1h_2 \in H$. Since S is an additive group closed under H, $h_1s_2 \in S$, and thus $s' = s_1 + h_1s_2 \in S$. The inverse of $\phi_{s,h}$ is $\phi_{-h^{-1}s,h^{-1}}$, which is also in $\mathcal{A}$. Thus $\mathcal{A} \leq \mathrm{AGL}(1, \mathbb{F})$.

$A = \mathcal{A}$. Clearly $G \subset \mathcal{A}$ (via $h = 1$) and $H \subset \mathcal{A}$ (via $s = 0$). Thus $A \subseteq \mathcal{A}$. Conversely, let $\phi_{s,h} \in \mathcal{A}$. Since $\phi_{s,h} = \phi_{0,h} \circ \phi_{h^{-1}s,1}$, where $\phi_{0,h} \in H$ and $h^{-1}s \in S$, it suffices to show that for any $s \in S$, $\phi_{s,1} \in A$. Next, since an $s \in S$ is a sum of terms hg, it suffices to show that the translation $x \mapsto x + hg \in A$. Indeed, it can be realized as $\phi_{0,h} \circ \phi_{g,1} \circ \phi_{0,h^{-1}}$, where each transformation is either in G or H. Thus $\mathcal{A} \subseteq A$.

Semidirect Product. The mapping $(s, h) \mapsto \phi_{s,h}$ is clearly a bijection between $S \rtimes H$ and A. The group operation of $\mathcal{A}$ matches the standard semidirect product definition. □

Proof of Lemma 3.2. Let

$$K \;:=\; \mathrm{Span}_{\mathbb{F}_p}(H) \;=\; \left\{ \sum_{i=1}^{\ell} c_i h_i \;:\; \ell \geq 0,\; c_i \in \mathbb{F}_p,\; h_i \in H \right\} \subseteq \mathbb{F}.$$

We first claim that K is a subfield of $\mathbb{F}$ and hence $K = \mathbb{F}_p(H)$.

Indeed, K is closed under addition by definition and contains 1. It is also closed under multiplication. Indeed, if $x = \sum_i c_i h_i$ and $y = \sum_j d_j h'_j$, then

$$xy = \sum_{i,j} (c_i d_j)\,(h_i h'_j),$$

and since $c_i d_j \in \mathbb{F}_p$ and $h_i h'_j \in H$, we have $xy \in K$. Thus K is a finite subring of the field $\mathbb{F}$. Being a subring of a field, K has no zero divisors, hence K is a finite integral domain, and therefore a field. Since K contains $\mathbb{F}_p$ and H, minimality of $\mathbb{F}_p(H)$ gives $\mathbb{F}_p(H) \subseteq K$. For the other direction, $K \subseteq \mathbb{F}_p(H)$ because $\mathbb{F}_p(H)$ is closed under $\mathbb{F}_p$-linear combinations. Hence $K = \mathbb{F}_p(H)$.

Next, we prove the desired equality. First, note that every hg with $h \in H$ and $g \in G$ lies in $\mathrm{Span}_{\mathbb{F}_p(H)}(G)$ because $h \in \mathbb{F}_p(H)$. Since $\mathrm{Span}_{\mathbb{F}_p(H)}(G)$ is an additive subgroup, it follows that

$$S = \sum_{h \in H} hG \;\subseteq\; \mathrm{Span}_{\mathbb{F}_p(H)}(G).$$

For the reverse inclusion, we show that S is closed under multiplication by scalars from $\mathbb{F}_p(H) = K$. By construction, S is an additive subgroup closed under multiplication by $h \in H$ and multiplication by $c \in \mathbb{F}_p$ (since G is an $\mathbb{F}_p$-subspace). Therefore, for any $a = \sum_i c_i h_i \in K$ and any $s \in S$,

$$as = \left(\sum_i c_i h_i \right) s = \sum_i c_i (h_i s) \;\in\; S,$$

so S is a K-vector space. Since $G \subseteq S$, we conclude that $\mathrm{Span}_K(G) \;\subseteq\; S$. Combining both inclusions gives $S = \mathrm{Span}_{\mathbb{F}_p(H)}(G)$. □

B Proof of Lemma 4.1

Proof. By Lemma 3.2, we have $S = \mathrm{Span}_{\mathbb{F}_{p^m}}(G)$. Next, let $(v_1, \ldots, v_m)$ be a basis of G over $\mathbb{F}_p$. We claim that this basis is also linearly independent over $\mathbb{F}_{p^m}$, which implies $\left|\mathrm{Span}_{\mathbb{F}_{p^m}}(G)\right| = p^{m^2}$.

Suppose, for the sake of contradiction, that $v_1, \ldots, v_m$ are linearly dependent over $\mathbb{F}_{p^m}$. Let $t \geq 2$ be the minimal number of vectors in a dependent subset, and assume without loss of generality

that $v_1, \ldots, v_t$ are such a set. Then there exist coefficients $\alpha_i \in \mathbb{F}_{p^m}$ with

$$\sum_{i=1}^{t} \alpha_i v_i = 0, \quad \text{and} \quad \alpha_1 = 1.$$

Applying the polynomial g (which is $\mathbb{F}_p$-linear) to both sides yields

$$0 = g\left(\sum_{i=1}^{t} \alpha_i v_i\right) = \sum_{i=1}^{t} \left(\alpha_i^{p^m} v_i^{p^m} + \alpha_i^p v_i^p + \alpha_i v_i\right).$$

Since each v_i is a root of g, we have $v_i^{p^m} + v_i^p + v_i = 0$. Subtracting $\sum_{i=1}^{t} \alpha_i \left(v_i^{p^m} + v_i^p + v_i\right)$ from the above and using $\alpha_i^{p^m} = \alpha_i$ (because $\alpha_i \in \mathbb{F}_{p^m}$) gives

$$\sum_{i=1}^{t} \left(\alpha_i^p - \alpha_i\right) v_i^p = 0.$$

The coefficient of v_1^p is zero since $\alpha_1 = 1$, so we obtain

$$\sum_{i=2}^{t} \left(\alpha_i^p - \alpha_i\right) v_i^p = 0. \tag{23}$$

Since all the v_i lie in some finite field $\mathbb{F}_{p^s}$, raising (23) to the p^{s-1}-th power maps v_i^p to v_i and yields

$$\sum_{i=2}^{t} \left(\alpha_i^p - \alpha_i\right)^{p^{s-1}} v_i = 0.$$

By the minimality of t, all coefficients in this relation must be zero, i.e.,

$$\alpha_i^p - \alpha_i = 0 \quad \text{for all } i \geq 2.$$

Thus $\alpha_i \in \mathbb{F}_p$ for all i, including $\alpha_1 = 1 \in \mathbb{F}_p$.

Therefore, the original dependence relation is an $\mathbb{F}_p$-linear dependence among $v_1, \ldots, v_t$, contradicting the fact that they are part of an $\mathbb{F}_p$-basis. It follows that $v_1, \ldots, v_m$ are linearly independent over $\mathbb{F}_{p^m}$. □

C Proof of Lemma 4.6

Proof. We integrate over $y = x_{k_{\max}}$. The range is $0 \leq y < \min(r, \rho)$. For a fixed y, the volume of the remaining polytope on the remaining C variables is $\frac{(r-y)^C}{C!}$. We prove it by induction on C. For $C = 1$ the region is the interval $[0, r - y)$ of length $r - y$. Assume the claim holds for $C - 1$. For C variables, fix $x_1 = t \in [0, r - y]$, then the volume of the remaining polytope on the remaining $C - 1$ variables is by the induction hypothesis $\frac{(r-y-t)^{C-1}}{(C-1)!}$. Integrating t gives

$$\int_0^{r-y} \frac{(r - y - t)^{C-1}}{(C-1)!}\, dt = \frac{(r-y)^C}{C!}.$$

Lastly, the volume of the polytope $\mathcal{P}$ equals

$$\mathrm{Vol}(\mathcal{P}) = \int_0^{\min(r,\rho)} \frac{(r-y)^C}{C!} dy = \left[-\frac{(r-y)^{C+1}}{(C+1)!}\right]_0^{\min(r,\rho)}.$$

Since $r - \min(r, \rho) = \max(0, r - \rho)$, the result follows. □

D Proof of Lemma 5.6

Proof. Recall that $\mathcal{P}$ is defined over the variables $\{x_k\}_{k\in I_{\mathrm{dom}}}$ and z by

$$0 \le z < r, \qquad x_k \ge 0 \ (k \in I_{\mathrm{dom}}), \qquad \sum_{k\in I_{\mathrm{dom}}} x_k < r\gamma, \qquad x_{k_{\max}} < \rho\gamma,$$

where $|I_{\mathrm{dom}}| = 2m+1$ and $k_{\max} \in I_{\mathrm{dom}}$.

First note that z is completely decoupled from the other constraints, and contributes a multiplicative factor of $\mathrm{Vol}(\{z : 0 \le z < r\}) = r$. Hence, $\mathrm{Vol}(\mathcal{P}) = r \cdot \mathrm{Vol}(\mathcal{Q})$ where $\mathcal{Q}$ is the polytope in the $2m+1$ variables $\{x_k\}_{k\in I_{\mathrm{dom}}}$. For a fixed $y := x_{k_{\max}} \in (0, \min(r\gamma, \rho\gamma)]$, the volume of the polytope defined by the remaining $2m$ variables equals $\frac{(r\gamma - y)^{2m}}{(2m)!}$. The proof follows by induction as in the proof of Lemma 4.6.

Therefore,

$$\begin{aligned}
\mathrm{Vol}(\mathcal{Q}) &= \int_0^{\min(r\gamma,\rho\gamma)} \frac{(r\gamma - y)^{2m}}{(2m)!}\, dy = \left[-\frac{(r\gamma - y)^{2m+1}}{(2m+1)!}\right]_0^{\min(r\gamma,\rho\gamma)} \\
&= \frac{(r\gamma)^{2m+1} - (r\gamma - \min(r\gamma, \rho\gamma))^{2m+1}}{(2m+1)!} \\
&= \frac{\gamma^{2m+1}\left(r^{2m+1} - \max(0, r-\rho)^{2m+1}\right)}{(2m+1)!},
\end{aligned}$$

where the last equality follows since

$$r\gamma - \min(r\gamma, \rho\gamma) = \max(0, r\gamma - \rho\gamma) = \gamma \max(0, r - \rho).$$

Multiplying by the factor r coming from z yields the result. □

References

[1] N. Alon and J. Bourgain. Additive patterns in multiplicative subgroups. *Geometric and Functional Analysis*, 24(3):721–739, June 2014. 5

[2] A. Barg and G. Zémor. Distance properties of expander codes. *IEEE Transactions on Information Theory*, 52(1):78–90, Jan. 2006. 2, 5

[3] I. Cascudo, H. Chen, R. Cramer, and C. Xing. Asymptotically good ideal linear secret sharing with strong multiplication over any fixed finite field. In *Proceedings of the 29th Annual International Cryptology Conference on Advances in Cryptology*, CRYPTO '09, page 466–486, Berlin, Heidelberg, 2009. Springer-Verlag. 6

[4] R. Cramer, I. Damgård, and U. Maurer. General secure multi-party computation from any linear secret-sharing scheme. In B. Preneel, editor, *Advances in Cryptology–EUROCRYPT 2000: International Conference on the Theory and Application of Cryptographic Techniques*, volume 1807 of *Lecture Notes in Computer Science*, pages 316–334, Bruges, Belgium, May 2000. Springer. 6

[5] I. Dinur, S. Liu, and R. Y. Zhang. New Codes on High Dimensional Expanders. In S. Srinivasan, editor, *40th Computational Complexity Conference (CCC 2025)*, volume 339 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 27:1–27:42, Dagstuhl, Germany, 2025. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. 6

[6] L. Golowich and T.-C. Lin. Quantum ldpc codes with transversal non-clifford gates via products of algebraic codes, 2024. 2, 6

[7] V. Guruswami and A. Rudra. Explicit codes achieving list decoding capacity: Error-correction with optimal redundancy. Electronic Colloquium on Computational Complexity (ECCC), 2005. Report 2005/133. 6

[8] S. Hoory, N. Linial, and A. Wigderson. Expander graphs and their applications. *Bulletin of the American Mathematical Society*, 43:439–561, 2006. 1, 5

[9] T. Kaufman and I. Oppenheim. High dimensional expanders and coset geometries. *European Journal of Combinatorics*, 111:103696, 2023. 40th Anniversary Edition. 2, 3, 6, 12

[10] S. Kim. Generalized minimum distance iterative decoding of tanner codes. *IEEE Communications Letters*, 9(8):738–740, 2005. 28

[11] O. Meir. Ip = pspace using error-correcting codes. *SIAM Journal on Computing*, 42(1):380–403, 2013. 6

[12] R. Pellikaan. On decoding by error location and dependent sets of error positions. *Discrete Mathematics*, 106-107:369–381, 1992. 2, 27, 28

[13] H. Randriambololona. On products and powers of linear codes under componentwise multiplication. In *Algorithmic Arithmetic, Geometry, and Coding Theory*, volume 637 of *Contemporary Mathematics*, pages 3–78. American Mathematical Society, 2015. 2, 6

[14] A. Shamir. Ip = pspace. *J. ACM*, 39(4):869–877, Oct. 1992. 6

[15] M. Sipser and D. A. Spielman. Expander codes. *IEEE Transactions on Information Theory*, 42(6):1710–1722, 1996. 1, 2, 5

[16] V. Skachek and R. Roth. Generalized minimum distance iterative decoding of expander codes. In *Proceedings 2003 IEEE Information Theory Workshop (Cat. No.03EX674)*, pages 245–248, 2003. 28

[17] I. Tamo and A. Barg. A family of optimal locally recoverable codes. *IEEE Transactions on Information Theory*, 60(8):4661–4676, 2014. 2, 6

[18] R. Tanner. A recursive approach to low complexity codes. *IEEE Transactions on Information Theory*, 27(5):533–547, 1981. 1, 5

[19] G. Zemor. On expander codes. *IEEE Transactions on Information Theory*, 47(2):835–837, 2001. 1, 2, 5, 27

[20] G. Zémor. On expander codes. *IEEE Transactions on Information Theory*, 47(2):835–837, 2001. 28