International Symposium on Information Theory and Its Applications
Honolulu, Hawaii, U.S.A., November 5–8, 2000

Codes and Iterative Decoding on Algebraic Expander Graphs

John Lafferty†
School of Computer Science
Carnegie Mellon University
5000 Forbes Avenue
Pittsburgh, PA 15213 USA
lafferty@cs.cmu.edu

Dan Rockmore‡
Departments of Mathematics
and Computer Science
Dartmouth College
Hanover, NH 03755 USA
rockmore@cs.dartmouth.edu

Abstract

The notion of graph expansion was introduced as a tool in coding theory by Sipser and Spielman, who used it to bound the minimum distance of a class of low-density codes, as well as the performance of various iterative decoding algorithms for these codes. In spite of its usefulness in establishing theoretical bounds on iterative decoding, graph expansion has not been widely used to design codes. Instead, random graphs are the primary means used to obtain graphs for codes, raising the question of whether comparable performance can be achieved using explicit constructions. In this paper we investigate the use of explicit algebraic expander graphs and algebraic subcodes, and show that the resulting coding schemes achieve excellent performance, competitive with standard low-density parity-check codes over a wide range of block lengths. Since the code constructions are based on graphs of groups, the Fourier transform can be used to obtain fast encoding algorithms for these codes.

1. Introduction

Gallager's low-density parity-check codes [3] have been the focus of a great deal of recent work in coding theory. Low-density parity-check codes with regular bit-degree sequences have been shown experimentally to perform very well [7], and variations that make use of irregular degree sequences and non-binary alphabets achieve outstanding performance [8] on the Gaussian channel. The work presented here is motivated by the fact that most low-density code techniques are based on random constructions and average case analysis over an ensemble of random codes. This raises the question of whether comparable performance can be achieved using explicit constructions—ideally, one would like to have *specific* codes with provable properties and good experimental performance. An additional motivation for the current work is that while capacity has been shown to be effectively achieved for long codes, the relative performance of coding schemes for shorter block lengths is not well understood. Our approach is inspired by Tanner's formulation of codes defined hierarchically on algebraically constructed graphs [11], and by Sipser and Spielman's analysis [10] showing the importance of graph expansion.

In this paper we study three families of explicit expander codes. The first family of codes is closely related to the explicit asymptotically good family of expander codes constructed in [10]. These codes are built on Cayley graphs of the non-abelian group $PSL_2(\mathbb{F}_q)$, using Hamming subcodes as constraints. The resulting codes achieve excellent performance; however, the constructions are somewhat elaborate, and the block sizes and parameters of the codes are limited. We also consider much simpler constructions in terms of expander graphs for the dihedral groups D_n, and cyclic groups C_n. In each case the constructions perform well when decoded using the sum-product algorithm, with lower bit-error rates than *regular* low-density parity-check codes with similar parameters on the Gaussian channel. While lower error rates can be obtained by using random graphs with irregular degree sequences and non-binary alphabets, the results presented here demonstrate that explicit, algebraically defined low-density codes can be competitive with random constructions for a range of block lengths.

2. Preliminaries on Expander Graphs

Graph expansion is a measure of the degree of connectivity of a graph—in a good expander every subset of vertices has a large number of neighbors that are not in the subset. More precisely, the following definitions are standard.

Definition 1. *A graph $\Gamma = (V; E)$ with n vertices*

†Research supported in part by NSF grant CCR-9805366.
‡Research supported in part by the National Science Foundation.

is said to be an ε-*expander if for any vertex subset* $S \subset V$ *with* $|S| \leq n/2$, $|\partial(S)| \geq \varepsilon|S|$, *where* $\partial(S) = \{v \in V/S \mid (s,t) \in E \text{ for some } s \in S\}$.

An error correcting code can be viewed as a bipartite graph by associating the left nodes with variables, and the right nodes with constraints (linear or nonlinear). Intuitively, if this graph is a good expander, then iterative decoding may work well on the graph because a small group of erroneous variables will give rise to a large number of violated constraints. Some of these constraints will be able to correct some of their neighboring bits, decreasing the number of erroneous bits in each round of iterative decoding. The Sipser and Spielman results give a precise statement and justification of this intuition, and show the *necessity* of expansion for asymptotically good codes. Unfortunately, the bounds established in the associated spectral analysis are quite weak, and do not fully explain the spectacular performance of the best low-density parity-check coding schemes. For example, using an intricate analysis based on spectral bounds, it is shown in [10] that expander codes on $PSL_2(\mathbb{F}_q)$ can be designed with minimum distance bounded away from zero, but falling well below the Gilbert-Varshamov bound, as shown in Figure 1. However, as we show below, these same codes in fact perform better than low-density parity-check codes.

A useful bound on expansion is given in terms of the *spectral gap* of the underlying graph [1].

Definition 2. *A connected* k-*regular graph is said to be a Ramanujan graph if* $\mu_1 \leq 2\sqrt{k-1}$, *where* μ_1 *is the second-largest absolute value of any eigenvalue of the graph's adjacency matrix. This is asymptotically the best spectral gap possible, as a consequence of the inequality* $\liminf_{n\to\infty} \mu_1(X_{n,k}) \geq 2\sqrt{k-1}$ *where the infimum is taken over all* k-*regular graphs with* n *vertices [1].*

In this paper all of the codes that we consider are built on Ramanujan graphs. In particular, we use Cayley graphs of abelian and non-abelian groups: given a group G and a set of generators $\{s_i\}$, we form the graph $\Gamma(G, \{s_i\})$ with vertices labeled by G, and with edges given by $(s, s_i g)$. By a *Cayley code* [5] we mean a code constructed in terms of a Cayley graph, where the bits (variables) are placed on either the vertices or the edges of the graph. If the adjacency matrix of the graph is sparse (equivalently, the chosen set of generators of the group is small), then we say that the code is *low-density*. The bits labeling the neighbors of a given vertex are required to form a codeword in some subcode, the simplest example being a parity check. A Cayley code inherits symmetries from the group. All of the classical cyclic codes are Cayley codes, with the bits on the vertices, where the group G is $\mathbb{Z}/n\mathbb{Z}$, and the subcode is a parity check. Fourier analysis can be used to obtain efficient encoding algorithms for these codes in general [5].

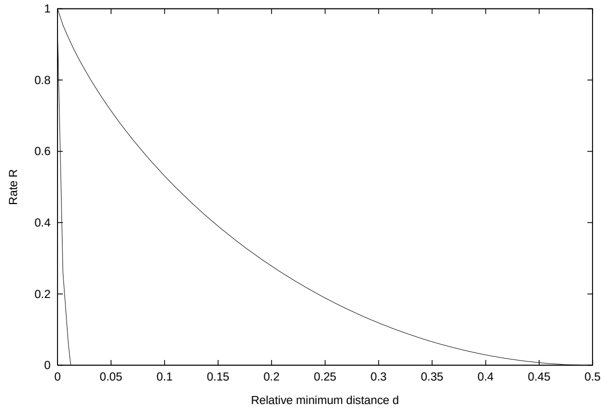

Figure 1: The provable minimum distance [10] of expander codes on $PSL_2(\mathbb{F}_q)$, compared with the Gilbert-Varshamov bound. The empirical performance of these codes (see Figure 2) shows that this lower bound is quite weak.

3. Expander Codes on $PSL_2(\mathbb{F}_q)$

The Lubotzky-Philips-Sarnak graphs [6] are Cayley graphs of the group $PSL_2(\mathbb{F}_q)$. Recall that $SL_2(\mathbb{F}_q)$ is the *special linear group* of 2×2 matrices of determinant one having entries from the finite field $\mathbb{F}$ of q elements. The *projective special linear group* $PSL_2(\mathbb{F}_q)$ is obtained by dividing $SL_2(\mathbb{F}_q)$ by its center, $\{\pm I\}$ where I is the 2×2 identity matrix. It is a simple finite group of Lie type for $q \geq 5$.

We suppose that $q \equiv 1 \ (mod\ 4)$. Let p be another prime which is a quadratic residue $(mod\ q)$. For $p > 3$ the LPS graphs $X^{p,q}$ are $(p+1)$-regular graphs on $PSL_2(\mathbb{F}_q)$ determined by the generating matrices

$$\frac{1}{\sqrt{p}} \begin{pmatrix} a_0 + i\, a_1 & a_2 + i\, a_3 \\ -a_2 + i\, a_3 & a_0 - i\, a_1 \end{pmatrix} \tag{1}$$

where $i = \sqrt{-1}\ (mod\ q)$, and (a_0, a_1, a_2, a_3) vary over the $p+1$ integral solutions to the equation $a_0^2 + a_1^2 + a_2^2 + a_3^2 = p$ having $a_0 > 0$ and odd and a_1, a_2, a_3 even.

Not only are the LPS graphs good expanders, they also have large girth: $girth(X^{p,q}) \geq 2\log_p q = \Omega(\log n)$. This is important for iterative decoding, since the independence assumption that the sum-product algorithm makes is invalid after a number of iterations equal to the girth of the graph.

Taking $p = 13$, we obtain 14-regular graphs on the group $G = PSL_2(\mathbb{F}_q)$ with $q(q^2-1)/2$ vertices. To construct a code with rate $\frac{1}{2}$ on these 14-regular graphs, we use the following procedure. By puncturing the $[15, 11, 3]$ Hamming code we obtain a $[14, 11]$ code $\mathcal{S}_1$. Expurgating the even codewords of $\mathcal{S}_1$ yields a $[14, 10]$ code $\mathcal{S}_2$. Assigning $\mathcal{S}_1$ to a proportion $0 \le \alpha \le 1$ of the constraint nodes, and $\mathcal{S}_2$ to the remaining proportion of $1 - \alpha$ constraint nodes, we obtain a code with overall rate of *at least* $1 - \frac{3\alpha + 4(1-\alpha)}{7} = \frac{3+\alpha}{7}$. Choosing $\alpha = \frac{1}{2}$ yields a code of rate at least $\frac{1}{2}$ and block length $\frac{7}{2} q(q^2 - 1)$. A sample simulation is shown in Figure 2 with $q = 17$ and $q = 29$, giving codes of block length 17,136 and 85,260. These expander codes achieve a coding gain of approximately 0.25 dB over $(3, 6)$-regular low-density parity-check codes with equal block length.

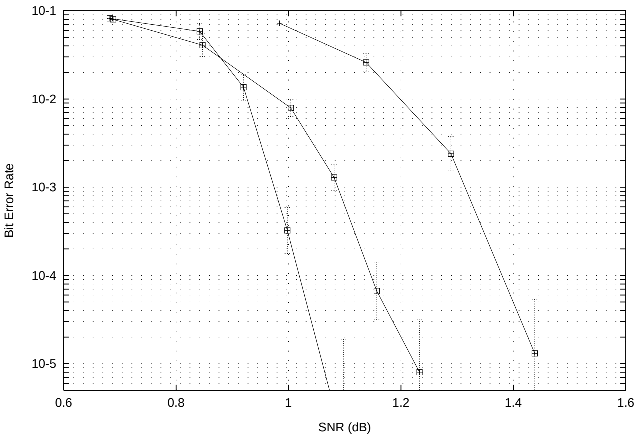

Figure 2: The leftmost and center waterfall curves are for LPS graphs for $PSL_2(\mathbb{F}_{29})$ and $PSL_2(\mathbb{F}_{17})$ respectively, each using modified Hamming codes as subcodes, to obtain a rate of (at least) $\frac{1}{2}$. The rightmost curve is for a Gallager code with degrees (3,6) and block length n =17,136, equal to the length of the $PSL_2(\mathbb{F}_{17})$ code.

4. Expander Codes on C_n and D_n

The primary technique used in finding expanders on cyclic groups is the bounding of character sums. This method also applies to dihedral groups, yielding a family of Ramanujan graphs that we use to design codes. In contrast to the LPS graphs, these graphs are asymptotically of unbounded degree; however, they are still useful for building codes with moderate block lengths, in the range of 500–50,000 bits.

If $C_n = \mathbb{Z}/nZ$ is the cyclic group of order n, and $\mathcal{S} = \{a_1, \ldots, a_k\}$ is a symmetric set of generators, then the eigenvalues of the Cayley graph $\Gamma(C_n, \mathcal{S})$ are given by $\sum_{i=1}^{k} \theta^{a_i}$, where θ ranges over the n-th roots of unity. Letting $n = p^2 - 1$, where p is prime, we have that $\mathbb{F}_{p^2}^* \cong \mathbb{F}_p(\omega)$, where ω is the root of an irreducible quadratic polynomial in $\mathbb{F}_p[x]$. The key technical fact is a theorem of Katz on character sums: if ψ is any multiplicative character on $\mathbb{F}_p(\omega)$, then

$$\sum_{i \in \mathbb{F}_p} \psi(\omega + i) \le \sqrt{p}\,.$$

Using this result, Chung [2] observes that the following generating set yields a Ramanujan graph. Let g be a generator for $\mathbb{F}_{p^2}^*$, and define $a_i = \log_g(\omega + i)$. Then the set $\mathcal{S} = \{\pm a_i\}_{i=0}^{p-1}$ generates $\mathbb{Z}/n\mathbb{Z}$, and has size $2p - 2 \le |\mathcal{S}| \le 2p$. Katz's theorem, applied to the character $\psi(g) = \theta$, shows that the Cayley graph $\Gamma(C_n, \mathcal{S})$ is Ramanujan.

Schellwat [9], modifies this construction to build Ramanujan graphs on the dihedral group D_{p^2-1}. The dihedral group D_n is made up of two copies of C_n pasted together by an involution: a natural presentation is $D_n = \{\langle r, s\rangle \mid r^n = s^2 = (sr)^2 = 1\}$. Each element can be written either as r^k or sr^k for suitable k. Let g, ω, and a_i be as given above. Then the set $\mathcal{D} = \{sr^{a_i}\}$ generates D_n, and furthermore, is symmetric, since $(sr^k)^2 = 1$. By Katz's theorem, this gives a family of bipartite Ramanujan graphs $\Gamma(D_{p^2-1}, \mathcal{D})$ of degree p. Numerical calculation of the spectrum, as well as our code simulations, indicates that they are much better expanders than the Chung graphs.

A simple variant of Katz's theorem allows us to add or delete small numbers of generators and still obtain Ramanujan graphs. For example, a simple use of the triangle inequality shows that the addition or deletion of an arbitrary element b to $\mathcal{D}$ implies that

$$\Big|\sum_{i \in \mathbb{F}_p} \theta^{a_i} \pm \theta^b\Big| \le \sqrt{p} + 1 < 2\sqrt{(p \pm 1) - 1}$$

for $p > 3$. Taking $p = 17$, this allows us to delete an element and use an extended Hamming code as a subcode.

Figure 3 shows the simulation of a code on D_{13}, with a block length of $n = 2{,}184$ bits, and constraints of size 13, for which we use punctured Hamming codes, similar to those used for $PSL_2(\mathbb{F}_q)$. A $(3, 6)$-regular Gallager code is shown for comparison. Figure 4 shows a simulation of the code on D_{17}, where we delete one of the generators to obtain constraints of size 16, for which we use an extended Hamming code. A comparable rate-$\frac{1}{2}$ Gallager code is also shown.

5. Decoding Complexity

The expander codes described above are decoded using the sum-product algorithm, where in each iteration the BCJR algorithm is carried out on the minimal

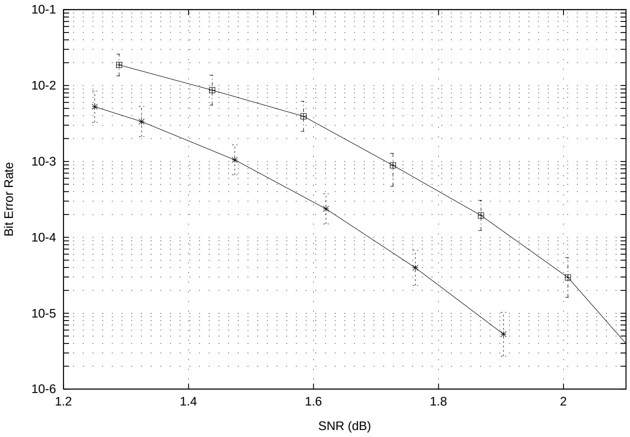

Figure 3: Expander code on D_{13^2-1}, having block length $n = 2{,}184$ and rate $\frac{1}{2}$ (lower curve) compared to a Gallager code with similar parameters.

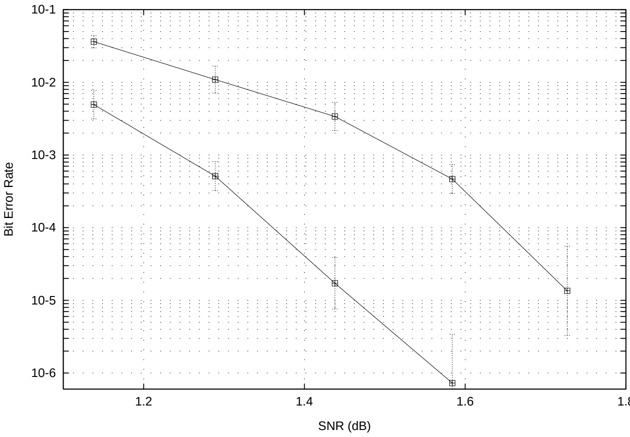

Figure 4: Expander code on D_{17^2-1}, with a generating set of size 16, resulting in a code of block length $n = 4{,}608$, using an extended Hamming code as subcode (lower curve), compared to a $(3,6)$-regular Gallager code with the same block length and rate.

trellis assigned to each constraint node. While the trellises we use are more complex than those needed for simple parity checks, there are fewer constraint nodes than in a Gallager code with the same rate and block length; however, there is certainly greater complexity overall.

To quantify the additional complexity, we note that on a trellis with E edges and V vertices, the BCJR algorithm requires $2E - V + 1$ floating point operations—we'll refer to these as trellis operations, or "tops." In terms of this measure of trellis complexity, one iteration of the sum-product algorithm on a rate-$\frac{1}{2}$ Gallager code requires 14.5 tops/bit. Using punctured and expurgated Hamming subcodes for graphs on $PSL_2(\mathbb{F}_q)$ requires 32.9 tops/bit. The use of an extended Hamming subcode for the dihedral group codes requires 57.4 tops/bit.

6. Conclusions

Recent work on low-density, parallel concatenated codes and iterative decoding methods has relied almost exclusively on random graphs. While graph expansion—a notion that arose from work in theoretical computer science—has been useful as an analytical tool, the results in this paper indicate that it can also be useful for designing explicit low-density codes. All of the algebraic constructions based on expanders that we have investigated are competitive with standard low-density parity-check codes. Though variations such as irregular degree sequences and non-binary fields give comparable gains, there are many further possibilities in designing codes within the framework of hierarchical codes defined on algebraic expander graphs.

References

[1] N. Alon, "Eigenvalues and expanders," *Combinatorica*, Vol. 6, 1986, pp. 83–96.

[2] F. Chung, "Diameters and eigenvalues," J. Amer. Math. Soc., Vol. 2, 1989, pp. 187–196.

[3] R. G. Gallager, *Low-Density Parity-Check Codes*, MIT Press, Cambridge, MA, 1963.

[4] J. Lafferty and D. Rockmore, "Fast Fourier analysis for SL_2 over a finite field and related numerical experiments," *Experimental Mathematics*, Vol. 1, No. 2, 1992, pp. 115–139.

[5] J. Lafferty and D. Rockmore, "Spectral techniques for expander codes," in Proceedings of *ACM Symposium on Theory of Computing (STOC)*, 1997, pp. 160–167.

[6] A. Lubotzky, R. Phillips, and P. Sarnak, "Ramanujan graphs," *Combinatorica*, Vol. 8, No. 3, 1988, pp. 261–277.

[7] D. MacKay, "Good error-correcting codes based on very sparse matrices," *IEEE Trans. Inform. Theory*, Vol. 45, No. 2, March 1999, pp. 399–431.

[8] T. Richardson, A. Shokrollahi and R. Urbanke, "Design of provably good low-density parity-check codes," submitted to *IEEE Trans. on Information Theory*.

[9] H. Schellwat, "Highly expanding graphs obtained from dihedral groups," In *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, Vol. 10, 1993, pp. 117–123.

[10] M. Sipser and D. Spielman, "Expander codes," *IEEE Trans. on Information Theory*, Vol. 42, No. 6, November 1996, pp. 1710–1722.

[11] R. M. Tanner, "A recursive approach to low complexity codes," *IEEE Trans. on Information Theory*, Vol. 27, No. 5, September, 1981, pp. 533–547.