Expansion properties of random Cayley graphs and vertex transitive graphs via matrix martingales

Demetres Christofides Klas Markström

Abstract

The Alon-Roichman theorem states that for every $\varepsilon > 0$ there is a constant $c(\varepsilon)$, such that the Cayley graph of a finite group G with respect to $c(\varepsilon)\log|G|$ elements of G, chosen independently and uniformly at random, has expected second largest eigenvalue less than ε. In particular, such a graph is an expander with high probability.

Landau and Russell, and independently Loh and Schulman, improved the bounds of the theorem. Following Landau and Russell we give a new proof of the result, improving the bounds even further. When considered for a general group G, our bounds are in a sense best possible. We also give a generalisation of the Alon-Roichman theorem to random coset graphs.

Our proof uses a Hoeffding-type result for operator valued random variables which we believe can be of independent interest.

1 Introduction

We say that a graph $X = (V, E)$ is an (n, d, ε)-***expander*** if it is a graph on n vertices, with maximum degree at most d, such that for every subset W of its vertices of size at most $n/2$, we have $|N(W) \setminus W| \geqslant \varepsilon|W|$, where $N(W)$ denotes the neighbourhood of W. The ***(vertex) expansion constant*** of X is the largest ε such that X is an (n, d, ε)-expander.

Expanders have many applications in computer science. For example in the construction of graphs with special connectivity properties and small number of edges, in the construction of graphs that are 'hard to pebble', in the construction of parallel sorting networks, in the establishment of lower bounds and time-space trade-offs for computing various functions, in complexity theory, in derandomization, in coding theory and in cryptography. References for the above applications can be found in [2, 3, 11, 14, 17].

The most useful families of expanders are ***linear expanders***: Families of graphs $\{X_i\}_{i\geqslant 1}$ such that each X_i is an (n_i, d, ε)-expander, with d and ε fixed, and n_i tending to infinity. Although it is not difficult to show via probabilistic methods that such families do exist, it has been quite hard to find explicit examples. The first such

construction was provided by Margulis [16]. For a more recent construction together with references to earlier constructions, the reader can look at [17].

Expanders whose degree is polylogarithmic in the number of vertices have also proved useful. Alon and Roichman [6] proved that 'random Cayley graphs are expanders'. (The degree needed for this result to hold turns out to be logarithmic in the number of vertices.) Recall that the ***Cayley diagram*** of a group G with respect to a multiset S of elements of G is the directed multigraph whose vertices are the elements of G and whose set of (directed) edges is the multiset of all ordered pairs (x, y) such that $y = sx$ for some $s \in S$. (I.e. if s appears k times in the multiset S, then there are k directed edges from x to sx.) Ignoring orientation but retaining multiple edges we get the ***Cayley graph*** $X(G, S)$ of G with respect to S. Note that this graph is $2|S|$-regular and it is connected if and only if S generates G.

One way to show that a graph has a given expansion property is via linear algebra. Recall that the ***adjacency matrix*** $A = A(X)$ of a graph X of order n is the $n \times n$ matrix (with rows and columns indexed by the vertices of X) defined by

$$A_{xy} = \begin{cases} 1 & \text{if } (x, y) \text{ is an edge;} \\ 0 & \text{otherwise.} \end{cases}$$

More generally, if X is a multigraph, then A_{xy} is defined to be the number of edges joining x and y, where we use the convention that A_{xx} is twice the number of loops at x. If the graph is d-regular we can also define the ***normalised adjacency matrix*** $T = T(X)$ of X, by $T = \frac{1}{d}A$. Note that T is a real symmetric matrix, so it has an orthonormal basis of real eigenvectors (and so all its eigenvalues are real). We'll write $\lambda_0 \geqslant \lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_{n-1}$ for its eigenvalues. (Warning: Some authors use the same notation for the eigenvalues of the adjacency matrix.) It can be easily checked that $\lambda_0 = 1$ and $|\lambda_i| \leqslant 1$ for every i. We will also denote by $\mu = \mu(G)$ the second largest element of the multiset of the absolute values of the eigenvalues of T. (I.e. $\mu = \max\{|\lambda_1|, |\lambda_{n-1}|\}$.) The relation between linear algebra and expansion properties of a graph comes from the following result of Alon and Milman (see [2, 4, 5]) which essentially says that if μ is bounded away from 1, then G is a good expander.

Lemma 1 (Alon-Milman [5])**.** *Let X be a d-regular graph on n vertices. Then X is an $(n, d, \frac{2-2\lambda_1}{3-2\lambda_1})$-expander. In particular it is an $(n, d, \frac{2-2\mu}{3-2\mu})$-expander.* □

Alon and Roichman used random walks on random Cayley graphs to prove the following:

Theorem 2 (Alon-Roichman [6])**.** *For every $\varepsilon > 0$ there is a $c(\varepsilon) > 0$, depending only on ε such that for every finite group G,*

$$\mathbb{E}(\mu(X(G, S))) \leqslant \varepsilon,$$

where S is a multiset of $c(\varepsilon) \log |G|$ elements of G chosen independently and uniformly at random.[1]

[1] All logarithms in this paper will be natural, unless otherwise stated.

Recently, Wigderson and Xiao [20], have obtained a polynomial time derandomization of this theorem.

It can be easily shown using martingales (see [6]) that μ is concentrated around its mean. Hence, by Theorem 2 we have:

Theorem 3. *For every $\delta > 0$, there is a $c'(\delta) > 0$ depending only on δ, such that for every finite group G, the Cayley graph $X(G, S)$ is an $(n, 2|S|, \delta)$-expander with high probability*[2] *as $|G| \to \infty$, where S is a multiset of $c'(\delta) \log |G|$ elements of G chosen independently and uniformly at random.*

Landau and Russell [12] and independently Loh and Schulman [13] gave new proofs of Theorem 2, replacing $\log |G|$ by $\log D(G)$, where $D(G)$ is the sum of the dimensions of the irreducible representations of the group G. This satisfies $\sqrt{|G|} < D(G) \leqslant |G|$. (See the next section for more details.) We will usually write D instead of $D(G)$ when the choice of G is clear. Landau and Russell also improved the constant $c(\varepsilon)$. Their proof used some Chernoff-type bounds on operator valued random variables from a recent paper of Ahlswede and Winter [1]. By proving a Hoeffding-type result for operator valued random variables we will deduce the following version of the Alon-Roichman theorem, improving the bounds of Landau and Russell even further.

Theorem 4. *For every $0 < \varepsilon < 1$, there is a function*

$$k = k(D) \leqslant \left(\frac{2}{\varepsilon^2} + o(1) \right) \log D$$

such that for every finite group G,

$$\mathbb{E}(\mu(X(G, S))) \leqslant \varepsilon,$$

where S is a multiset of k elements of G chosen independently and uniformly at random. The $o(1)$ term is with respect to increasing D.

To be more precise, the expression we get for k is

$$k = \left(\frac{1}{H_{1/2}\left(\frac{1+\varepsilon}{2}\right)} + o(1) \right) \log D,$$

where, H_p is the ***weighted entropy function***

$$H_p(x) = x \log \left(\frac{x}{p} \right) + (1 - x) \log \left(\frac{1 - x}{1 - p} \right).$$

In fact, the bounds in [1] used by Landau and Russell are not quite correct. (See Section 3 for more details.) Using a corrected version of the Ahlswede and Winter result, the Landau-Russell method also gives the same bounds as Theorem 4.

Theorem 4 is an easy corollary of the following theorem.

[2] Meaning that the probability tends to 1.

Theorem 5. *Let S be a multiset of k elements of a finite group G chosen independently and uniformly at random. Then, for every $0 < \varepsilon < 1$,*

$$\Pr(\mu(X(G,S)) \geqslant \varepsilon) \leqslant 2D \exp\left\{-kH_{1/2}\left(\frac{1+\varepsilon}{2}\right)\right\}.$$

Indeed, taking $k = \frac{1}{H_{1/2}((1+\varepsilon)/2)}[\log D + b + \log 2]$ we get $\Pr(\mu \geqslant \varepsilon) \leqslant e^{-b}$. Since μ takes values in $[0,1]$ we have $\mathbb{E}\mu \leqslant (1-e^{-b})\varepsilon + e^{-b} \leqslant \varepsilon + e^{-b}$. To deduce Theorem 4, replace ε by $\varepsilon' = \varepsilon(1-\delta)$ and b by $-\log(\varepsilon\delta)$ where $\delta = \delta(D)$ tends to 0 as D tends to infinity and $\log 1/\delta = o(\log D)$.

As we will see later, Theorem 5 is essentially the best possible theorem one could have for a general group G. To improve on the theorem, one possibly needs to know more about the structure of the specific group under consideration. It is worth mentioning that recently Kassabov, Lubotzky and Nikolov [10] have announced a proof that there exist a positive integer k and an $\varepsilon > 0$ such that every non-abelian finite simple group which is not a Suzuki group has a set S of k generators such that $X(G,S)$ is an ε-expander. It is believed that the result also holds for the Suzuki groups. On the other hand, as Lubotzky and Weiss [15] have shown, no such result can hold for a family of 'almost' abelian groups[3].

Recall that a graph X is called ***vertex transitive*** if for every two vertices x, y of X, there is an automorphism ϕ of X such that $\phi(x) = y$. Cayley graphs are vertex transitive graphs. However, there are vertex transitive graphs which are not Cayley graphs, the smallest such graph being the Petersen graph. As Sabidussi [18] observed, every vertex transitive graph can be obtained using the following construction based on cosets of a group:

Given a group G and a subgroup H of G, the ***coset diagram*** of G modulo H with respect to a multiset S of elements of G is defined as follows: Its vertices are the right cosets of H in G, and there is a directed edge from Hx to Hy if and only if $yx^{-1} \in HSH$. The multiplicity of each directed edge from Hx to Hy is defined to be the number of elements s of the multiset S such that $yx^{-1} \in HsH$. Note that the definition of the coset diagram is indeed well-defined, i.e. it does not depend on the choice of representatives for the cosets of H. We can now define the ***coset graph*** $X(G,H,S)$ of G modulo H with respect to the multiset S by retaining multiple edges and ignoring orientation. It is not difficult to check that $X(G,H,S)$ is a regular graph of degree $2\sum_{s\in S}\frac{|HsH|}{|H|}$. Note that if H is a normal subgroup of G, then the Cayley graph $X(G/H,S)$ is isomorphic to the coset graph $X(G,H,S)$.

As we will see in Section 5, using similar ideas as in the proof of Theorem 5 we can extend Theorem 5 to show that random coset graphs have good expansion properties.

For the reader's convenience, we recall in Section 2 all the results from representation theory that we will need. Proofs of the results can be found in many books, for

[3]Here 'almost' abelian refers to any family of soluble groups having bounded derived length.

example in Serre [19]. In Section 3 we will give the proof of our Hoeffding-type result for operator valued random variables. We believe that this result can be of independent interest in cases where one has to work with (operator valued) random variables which are not independently distributed. In Section 4 we give the proof of Theorem 5. In Section 5 we extend Theorem 5 to coset graphs. Finally in Section 6, we discuss some related results.

2 Some Representation Theory

A ***representation*** ρ of a finite group G is a homomorphism $\rho : G \to \mathrm{GL}(V)$, where V is a finite dimensional vector space over $\mathbb{C}$. The ***dimension*** d_ρ of ρ is simply the dimension of V. Equip V with an inner product $\langle , \rangle$. One can now replace $\langle , \rangle$ with a new inner product $\langle , \rangle'$ which is preserved under the action of ρ. Indeed define

$$\langle v, w\rangle' = \frac{1}{|G|} \sum_{g \in G} \langle \rho(g)v, \rho(g)w\rangle.$$

In this way, we may (and we will) consider ρ as a homomorphism into the unitary group of V.

We say that a subspace W of V is ***invariant*** if it is fixed by ρ. (I.e. $\rho(g)W \subseteq W$ for every $g \in G$.) It is then easily checked that the restriction $\rho_W : G \to \mathrm{GL}(W)$ is a representation. We say that ρ is ***irreducible*** if there is no non-trivial invariant subspace.

Theorem (Complete Reducibility)**.** *Any representation ρ can be decomposed into irreducible representations $\rho = \rho_1 \oplus \ldots \oplus \rho_k$. (Meaning $V = W_1 \oplus \ldots W_k$ where each W_i is invariant and $\rho_i = \rho_{W_i}$ is irreducible.)*

We say that two representations $\rho : G \to \mathrm{GL}(V)$ and $\rho' : G \to GL(V')$ are ***equivalent*** if there exists an isomorphism $\phi : V \to V'$ such that the following diagram commutes for every $g \in G$.

$$\begin{array}{ccc} V & \xrightarrow{\phi} & V' \\ {\scriptstyle \rho(g)}\downarrow & & \downarrow{\scriptstyle \rho'(g)} \\ V & \xrightarrow[\phi]{} & V' \end{array}$$

Theorem. *A finite group G has only a finite number of irreducible representations up to equivalence.*

Two important representations are the ***trivial representation*** $1 : G \to \mathrm{GL}(\mathbb{C})$; $g \mapsto id$ and the ***regular representation*** $R : G \to \mathrm{GL}(\mathbb{C}[G])$, where $\mathbb{C}[G]$ is the

vector space over $\mathbb{C}$ of formal sums of elements of G, defined by $(Rg)(\sum \alpha_h h) = \sum \alpha_h gh = \sum \alpha_{g^{-1}h} h$. It is an important fact that R decomposes as

$$R = \bigoplus \underbrace{\rho \oplus \cdots \oplus \rho}_{\dim \rho}$$

where the sum is over a complete set of inequivalent irreducible representations of G. In particular, since $\mathbb{C}[G]$ has dimension $|G|$, we deduce that $|G| = \sum (d_\rho)^2$. We define $D = D(G)$ by $D = \sum d_\rho$ and observe that $\sqrt{|G|} < D(G) \leqslant |G|$.

Given a subgroup H of a group G, by restricting to H, we can consider any representation ρ of G as a representation of H. It could be the case that ρ is an irreducible representation of G, but when restricted to H is not any more irreducible. We define $D(G, H)$ to be the sum of the dimensions of the irreducible representations of G which, when decomposed into irreducible representations of H, do not contain the trivial representation of H. For example, it is not hard to show that if H is normal in G, then $D(G/H) = D(G, H)$.

3 Operator valued random variables

Let V be a Hilbert space of dimension d, $A(V)$ be the set of self adjoint operators on V and $P(V)$ the cone of positive operators on V i.e.

$$P(V) = \{A \in A(V) : \text{all eigenvalues of } A \text{ are non-negative}\}.$$

This defines a partial order on $A(V)$ by $A \leqslant B \Leftrightarrow B - A \geqslant 0$. We denote by $[A, B]$ the set of all $C \in A(V)$ such that $A \leqslant C \leqslant B$. We also denote by $\|A\|$ the largest eigenvalue of A in absolute value.

In their proof of the Alon-Roichman theorem, Landau and Russell used the last part of Theorem 19 of [1] which is wrong for μ small enough. The proof of the theorem is correct apart from the very last line. For the case $\mu = 1/2$, which was the one used in the Landau-Russell proof, the proof of Theorem 19 in [1] gives the following better result.

Theorem 6. *Let V be a Hilbert space of dimension d and let $A, A_1, \ldots, A_k$ be independent identically distributed random variables taking values in $[0, I] \subseteq A(V)$ with $\mathbb{E}A = \frac{1}{2}I$. Then for all $\varepsilon \in [0, 1]$,*

$$\Pr\left(\frac{1}{k}\sum A_i \notin \left[\frac{1-\varepsilon}{2}, \frac{1+\varepsilon}{2}\right]\right) \leqslant 2d \exp\left\{-nH_{1/2}\left(\frac{1-\varepsilon}{2}\right)\right\}.$$

With this result, the Landau-Russell proof also yields Theorem 4.

The main difficulty in the proof of a Hoeffding-type result for operators, is the fact that the mapping $Y \mapsto e^Y$ is not operator convex. For example if $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ and

$B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ then clearly $A \geqslant B$ but it can be easily checked that $e^A \not\geqslant e^B$. This difficulty is resolved in the following lemma.

Lemma 7. *Let Y be a random variable taking values in $[-r, 1-r] \subseteq A(V)$ with $\mathbb{E}Y = 0$. Then, for $s \geqslant 0$ we have*

$$\mathbb{E}(e^{sY}) \leqslant \left((1-r)e^{-sr} + re^{s(1-r)}\right) I.$$

Proof. Firstly, by the convexity of the function $y \to e^y$, we deduce that for any $y \in \mathbb{R}$

$$e^{sy} \leqslant (1-r-y)e^{-sr} + (y+r)e^{s(1-r)}.$$

It follows that

$$e^{sY} \leqslant ((1-r)I - Y)e^{-sr} + (Y + rI)e^{s(1-r)}.$$

holds for any diagonal matrix Y. By diagonalising, we deduce that the same inequality holds for any $Y \in A(V)$. The result now follows by taking expectations. □

Before proving our Hoeffding-type result we include for completeness a proof of (a slightly weaker version of) the operator Markov inequality of [1].

Lemma 8 (Operator Markov [1])**.** *Let X be a random variable taking values in $P(V)$. Then*

$$\Pr(X \not\leqslant I) \leqslant \operatorname{Tr}(\mathbb{E}X).$$

Proof. We have

$$\mathbb{E}X = \sum_A \Pr(X = A)A \geqslant \sum_{A \not\leqslant I} \Pr(X = A)A.$$

The result follows by taking traces and noting that every $A \in P(V)$ with $A \not\leqslant I$ has $\operatorname{Tr}(A) \geqslant 1$. □

We are now ready to prove our Hoeffding-type result for operators. It will be used in the next section to give a proof of Theorem 5 and thus of Theorem 4.

Theorem 9 (Operator Hoeffding)**.** *Let V be a Hilbert space of dimension d and let $X_i = \mathbb{E}(X|\mathcal{F}_i)$ be a martingale, taking values in $A(V)$, whose difference sequence satisfies $Y_i \in [-r_i, 1-r_i]$. Let $r = \frac{1}{n}\sum_{i=1}^n r_i$ Then*

$$\Pr(X - \mathbb{E}X \not\leqslant nhI) \leqslant d \exp\{-nH_r(r+h)\}$$

and

$$\Pr(X - \mathbb{E}X \not\geqslant -nhI) \leqslant d \exp\{-nH_r(r-h)\}.$$

Proof. By Lemma 8, if $s > 0$, we have

$$\begin{aligned}
\Pr(X - \mathbb{E}X \not\leqslant hnI) &= \Pr\left(\sum Y_i \not\leqslant hnI\right) \\
&= \Pr\left(\exp\left\{s\sum Y_i\right\} \not\leqslant e^{shn} I\right) \\
&\leqslant e^{-shn} \operatorname{Tr} \mathbb{E} \exp\left\{s\sum Y_i\right\} \\
&= e^{-shn} \operatorname{Tr} \mathbb{E}\left(\mathbb{E}\left(\exp\left\{s\sum Y_i\right\}|\mathcal{F}_{n-1}\right)\right) \\
&= e^{-shn} \operatorname{Tr} \mathbb{E}\left(\exp\left\{s\sum_{i=1}^{n-1} Y_i\right\}\mathbb{E}\left(\exp\{sY_n\}|\mathcal{F}_{n-1}\right)\right),
\end{aligned}$$

where the second equality follows by diagonalising $\sum Y_i$. Applying Lemma 7 we deduce that

$$\Pr(X - \mathbb{E}X \not\leqslant hnI) \leqslant e^{-shn}[(1-r_n)e^{-sr_n} + r_n e^{s(1-r_n)}] \operatorname{Tr} \mathbb{E}\left(\exp\left\{s\sum_{i=1}^{n-1} Y_i\right\}\right).$$

By induction and the arithmetic-geometric mean inequality it follows that

$$\Pr(X - \mathbb{E}X \not\leqslant hnI) \leqslant d\left((1-r)e^{-s(h+r)} + re^{s(1-h-r)}\right)^n.$$

The right hand side is minimized by taking $s = \log\left(\frac{(h+r)(1-r)}{r(1-h-r)}\right)$. This gives the first inequality. The second inequality is obtained by considering $-X$. □

4 Proof of Theorem 5

Let $s_1, \ldots, s_k$ be elements of G chosen independently and uniformly at random. Let

$$s = \frac{1}{2k}\sum_{i=1}^{k}(s_i + s_i^{-1}) \in \mathbb{C}[G]$$

and observe that the matrix of the linear operator

$$R(s) = \frac{1}{2k}\sum_{i=1}^{k}(R(s_i) + R(s_i^{-1}))$$

with respect to the standard basis of $\mathbb{C}[G]$ (i.e. $\{g : g \in G\}$) is just the normalised adjacency matrix T of $X(G,S)$. Here, the homomorphism R, is extended linearly to $\mathbb{C}[G]$. But the eigenvalue 1 corresponds to the trivial representation, so (by the decomposition of R) we deduce that $\mu = \max_{\rho} \|\rho(s)\|$ where ρ runs over all irreducible non-trivial representations of G.

So we are left with estimating $\|\rho(s)\|$ for every non-trivial irreducible representation ρ. Fix a non-trivial irreducible representation ρ of G and let Y_i be the operator

$\frac{1}{2}[\rho(s_i) + \rho(s_i^{-1})]$. Then $X_i = Y_1 + \ldots + Y_i$ is a martingale satisfying the conditions of Theorem 9 with $d = d_\rho$ and $r_i = \frac{1}{2}$. Indeed, the matrix of the operator $\sum_{g \in G} R(g)$ with respect to the standard basis of $\mathbb{C}[G]$ is the 'all 1 matrix'. In particular it has rank 1. But so does the matrix of the trivial representation. Hence, by the decomposition of R, $\sum_{g \in G} \rho(g)$ is the zero operator and so $\mathbb{E}Y_i = 0$. It follows that

$$\begin{aligned}\Pr(\|\rho(s)\| \geqslant \varepsilon) &= \Pr\left(\left\|\frac{1}{k}\sum_{i=1}^{k} A_i\right\| \geqslant \frac{\varepsilon}{2}\right) \\ &= \Pr\left(\|X - \mathbb{E}X\| \geqslant \frac{\varepsilon k}{2}\right) \\ &\leqslant 2d_\rho \exp\left\{-kH_{1/2}\left(\frac{1+\varepsilon}{2}\right)\right\}.\end{aligned}$$

Summing over all irreducible non-trivial representations of G completes the proof of Theorem 5. □

The difference between our proof and the proofs in [12] and [13] are in the estimation of $\|\rho(s)\|$. Loh and Schulman proceeded via random walks as in the original Alon-Roichman proof while Landau and Russell used the operator Chernoff bound of [1].

5 Coset Graphs

Let G be a finite group, H a subgroup of G and S is a multiset of elements of G, chosen independently and uniformly at random. We want to show that if S is large enough, then $X = X(G, H, S)$ is a good expander. As before, the aim would be to deduce this from the spectral properties of X. However, due to problems with multiplicities, it seems that it is not easy to show that $\mu(X)$ is as small as one would hope for. The following simple lemma will enable us to deduce that X is a good expander by looking at a related graph Y for which $\mu(Y)$ can be computed more easily. Recall that a ***stochastic matrix*** is a matrix P with entries in $[0, 1]$ such that the elements in each row of P sum to 1.

Lemma 10. *Let X be a d-regular graph on n vertices. Let $\mu^* = \mu^*(X)$ be the infimum of the second largest eigenvalue in absolute value of P, where P runs over all $n \times n$ symmetric stochastic matrices satisfying $P_{ij} = 0$ whenever $T(X)_{ij} = 0$. Then X is an $(n, d, \frac{2-2\mu^*}{3-2\mu^*})$-expander.*

Proof. Immediate if we note that two graphs which have the same underlying simple graph, have the same expansion constant. Indeed, this observation says that the lemma is true if the entries of the matrix are rational. We can easily pass to matrices which also have irrational entries by taking limits. □

For further discussion of this lemma, see Section 6.4.

So, instead of looking at the spectrum of X, we can look at the spectrum of any graph Y which has the same underlying simple graph as X. We define Y as follows: Its vertex set is the set of right cosets of H in G, and the multiplicity of the edge (Hx, Hy) is the number of possible ways to express y as $h_1 s h_2 x$ where $h_1, h_2 \in H$ and $s \in S$ or $s^{-1} \in S$. Note that Y is $2|H||S|$-regular and it has indeed the same underlying simple graph as H.

We claim that every eigenvalue of Y, is also an eigenvalue of the Cayley graph Y' of G with respect to the multiset HSH. This follows from the fact that Y' is the 'blow-up' of Y. I.e. Y' is obtained from Y by replacing each vertex by $|H|$ new vertices corresponding to the elements of the coset, and replacing each edge by a $K_{|H|,|H|}$ on the corresponding vertices. Thus, the normalised adjacency matrix of Y' can be written up as a $\frac{|G|}{|H|} \times \frac{|G|}{|H|}$ matrix of $|H| \times |H|$ blocks, where each block is either the all 0's or all 1's matrix, depending on whether the corresponding entry in Y was a 0 or a 1. It then easily follows that if $v = (v_1, \ldots, v_{|G|/|H|})$ is an eigenvector of Y with eigenvalue λ, then $v' = (v_1, \ldots, v_1, \ldots, v_{|G|/|H|}, \ldots, v_{|G|/|H|})$ is an eigenvector of Y' with eigenvalue λ.

Thus, in order to show that X is a good expander, it is enough to show that $\mu(Y')$ is bounded away from 1.

Theorem 11. *With the above notation, we have*

$$\Pr(\mu(Y') \geqslant \varepsilon) \leqslant 2D(G,H) \exp\left\{-|S|H_{1/2}\left(\frac{1+\varepsilon}{2}\right)\right\}.$$

Proof. As in the proof of Theorem 5 we get

$$\begin{aligned}\Pr(\mu(Y') \geqslant \varepsilon) &\leqslant \sum_{\rho} \Pr\left(\left\|\rho\left(\frac{1}{2|S||H|^2}\sum_{\substack{h_1,h_2\in H\\ s\in S}}(h_1 s h_2 + h_1 s^{-1} h_2)\right)\right\| \geqslant \varepsilon\right)\\ &\leqslant \sum_{\|\sum_{h\in H}\rho(h)\|\neq 0} \Pr\left(\left\|\rho\left(\frac{1}{2|S|}\sum_{s\in S}(s+s^{-1})\right)\right\| \geqslant \frac{\varepsilon|H|^2}{\|\sum_{h\in H}\rho(h)\|^2}\right)\end{aligned}$$

The result will follow by showing that $\|\sum \rho(h)\|$ is equal to $|H|$ or 0, depending on whether ρ, as a representation of H, contains the trivial representation of H or not. But if ρ contains the trivial representation of H, then clearly $\|\sum \rho(h)\| = |H|$, while if it does not, then it decomposes into a sum of irreducible representations ρ_i of H such that each $\sum \rho_i(h)$ is the zero operator. This completes the proof. □

6 Further Comments and Results

6.1 Comparison with some lower bounds

Alon and Roichman also showed that if $G = \mathbb{Z}_2^m$, S is any subset of elements of G, and $\mu(X(G,S)) \leqslant 1-\delta$, then the following must hold:

$$\sum_{i<\frac{\delta|S|}{4}} \binom{|S|}{i} 2^m < 2^l.$$

In particular this implies that $|S| \geqslant \left(1+\Omega_m\left(\delta \log\left(\frac{1}{\delta}\right)\right)\right) m$.

The following result, proven by Alon and Roichman for abelian groups, shows that this lower bound is essentially sharp as an upper bound for every group.

Theorem 12. *For every sufficiently small $\delta > 0$, there is an $\varepsilon = O\left(\delta \log\left(\frac{1}{\delta}\right)\right)$ such that for every finite group G, $\mu(X(G,S)) \leqslant 1-\delta$ with high probability as $D \to \infty$, where S is a multiset of $(1+\varepsilon)\log_2 D$, chosen independently and uniformly at random.*

Proof. Taking

$$k = \frac{\log D + b + \log 2}{H_{1/2}(1-\delta/2)} = \frac{2(\log D + b + \log 2)}{(2-\delta)\log(2-\delta) + \delta \log \delta}$$

we deduce that $\Pr(\mu \geqslant 1-\delta) \leqslant e^{-b}$ and we can then proceed as before. The result follows by noting that as $\delta \to 0$ then

$$\frac{2}{(2-\delta)\log(2-\delta) + \delta\log\delta} \sim \frac{1}{\log 2} + \frac{\delta \log\left(\frac{1}{\delta}\right)}{2(\log 2)^2}.$$

□

Since Theorem 12 is a direct consequence of Theorem 5 and since it is essentially sharp for $G = \mathbb{Z}_2^m$, then in order to improve on Theorem 5 it seems we would have to take into consideration more of the structure of the group that we are concerned with.

6.2 Distribution assumptions in Theorem 5

Note that our operator Hoeffding result implies Theorem 6. In fact it implies the operator Chernoff result of [1][Theorem 19]. Moreover it has the advantage that the random variables need not be identically distributed nor independent. In our case, to prove the result of Theorem 5, we can give the s_i any distribution we like as long as $\mathbb{E}(\rho(s_i) + \rho(s_i^{-1})|s_1,\ldots,s_{i-1}) = 0$ for any non-trivial irreducible representation ρ. However it turns out that if we insist that $\Pr(s_i = g) = \Pr(s_i = g^{-1})$ for all g (this doesn't affect the final outcome) then the only possibility is choosing the s_i independently and uniformly at random. Indeed consider

$$x = \sum_{g\in G}(\Pr(s_i = g|s_1,\ldots,s_{i-1}) - \frac{1}{|G|})g \in \mathbb{C}[G]$$

and note that by our assumption

$$x = x' = \sum_{g \in G} (\Pr(s_i = g^{-1} | s_1, \ldots, s_{i-1}) - \frac{1}{|G|}) g.$$

Then for any irreducible representation ρ of G (trivial or not) extended linearly to $\mathbb{C}[G]$ we have $\rho(x + x') = 0$. Hence $\rho(x) = 0$ for any representation of G (irreducible or not). But this is also true for the regular representation R which satisfies $R(x) = 0$ if and only if $x = 0$ i.e. if and only if $\Pr(s_i = g | s_1, \ldots, s_{i-1}) = 1/|G|$. So essentially we can only pick the $s_i's$ independently and uniformly at random. However, if one is willing to lose a bit on $\mathbb{E}(\rho(s_i) + \rho(s_i^{-1}) | s_1, \ldots, s_{i-1}) = 0$ one might still be able to get a more general result than Theorem 5 with some loss in the constants. We don't pursue this line of thought any further.

6.3 Ramanujan-like Cayley graphs

Kahale [9] proved that the second largest eigenvalue of any d-regular graph on n vertices is at least

$$\frac{2\sqrt{d-1}}{d} \left(1 + O\left(\left(\frac{\log d}{\log n}\right)^2\right)\right).$$

This is essentially best possible; there are explicit constructions of ***Ramanujan graphs*** (see for example the book by Davidoff, Sarnak and Valette [8] and its references), d-regular graphs with second largest eigenvalue at most $\frac{2\sqrt{d-1}}{d}$. We can show that sufficiently dense random Cayley graphs are quite close to being Ramanujan.

Theorem 13. *For every $\delta > 0$ and every group G, if S is a set of k elements of G, chosen independently and uniformly at random, then $\mu(X(G,S)) \leqslant (\sqrt{2} + \delta)\sqrt{\frac{\log D}{k}}$ with high probability as $D \to \infty$.*

Proof. Immediate by Theorem 5 since

$$\Pr\left(\mu(X(G,S)) \geqslant (\sqrt{2} + \delta)\sqrt{\frac{\log D}{k}}\right) \leqslant 2D^{1-(1+\delta/\sqrt{2})^2}. \qquad \square$$

Taking $k = (\log D)^r$, for some fixed r, we deduce that with high probability, such a Cayley graph has second eigenvalue at most $\frac{\sqrt{2}+\delta}{\sqrt{k}} k^{1/2r}$.

6.4 Discussion of Lemma 10

In Section 5, we used Lemma 10 to pass to a matrix which had easily computable second largest eigenvalue. We now give two examples where Lemma 10 is used to improve the bound for the expansion constant.

Example. We begin with a connected 3-regular non-bipartite graph X. This guarantees that $\mu(X) < 1$. It is well known that each 3-regular bridgeless graph has a

perfect matching. We actually pick X such that it has a perfect matching M whose removal does not leave a Hamiltonian cycle. We now replace each edge of X not in M by n new edges to get new graph X_n. Clearly, $\mu^*(X_n) \leqslant \mu(X) < 1$. On the other hand, the normalised adjacency matrices of X_n tend to the normalised adjacency matrix of $X \setminus M$ in the operator norm, say. Since the eigenvalues vary continuously, we deduce that $\mu(X_n) \to \mu(X \setminus M) = 1$ as $n \to \infty$.

In the above example, we used Lemma 10 to get a better expansion constant for the multigraphs X_n than the one given by Lemma 1. However this was in a sense not a genuine example. All we did was to take a simple graph and produce some multigraphs with the same adjacencies, which had very bad spectral properties. So now we present an example of a simple graph X such that we can show that $\mu^*(X) < \mu(X)$.

Example. We take $X = K_4 \times C_{10}$. This can be viewed as a Cayley graph of $\mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_{10}$. Using the representation theoretic approach of Section 3, one can show that $\mu(X) = \frac{7+\sqrt{5}}{10} = 0.9236\ldots$. Consider the subgraph $Y = C_4 \times C_{10}$ of X. We give each edge of Y a weight of $\frac{6}{5}$ and each other edge a weight of $\frac{1}{5}$ to get a new (weighted) graph Z. Note that Z is a regular graph. It can be checked that $\mu^*(X) \leqslant \mu(Z) = \frac{23}{25} = 0.92 < \mu(X)$.

When can one use this method to improve the bound for the expansion constant? In the above example, we had a regular graph X which had some local dense parts together with some global other structure, and the global structure was responsible for the expansion of X. Picking a matching Y in each of these local parts and putting higher weights on the edges of $X \setminus Y$ we obtained a regular weighted graph Z. Why should we have $\mu(Z) < \mu(X)$? There is a relationship between $\mu(X)$ and random walks on the vertices of X. The smaller $\mu(X)$ is, the faster the random walk on the vertices of X will converge to the stationary distribution and vice versa (see e.g. [7, Chapter IX]). Intuitively, by having less weight in most edges of these dense local parts (but large weight in some of their edges) we would expect a random walk to spend less time in each local part, and hence to converge faster to the stationary distribution. We believe that in situations like this, where we have better local expansion properties than globally, one has a fairly good chance to be able to use Lemma 10 to deduce a better bound for the expansion properties for the graph X.

Acknowledgements

The first author is supported by grants from the Engineering and Physical Sciences Research Council and from the Cambridge Commonwealth Trust. The work was done while the second author was visiting the Department of Pure Mathematics and Mathematical Statistics of the University of Cambridge, and he would like to thank the department for its hospitality and also the Royal Physiographic Society in Lund for the grant, making this visit possible.

We would like to thank the referees for their constructive comments and remarks. We would also like to thank one of the referees who brought to our attention the paper

by Wigderson and Xiao [20].

References

[1] R. Ahlswede and A. Winter, Strong converse for identification via quantum channels, IEEE Trans. Inform. Theory **48** (2002), 569–579.

[2] N. Alon, Eigenvalues and expanders, Combinatorica **6** (1986), 83–96.

[3] N. Alon, Eigenvalues, geometric expanders, sorting in rounds, and Ramsey theory, Combinatorica **6** (1986), 207–219.

[4] N. Alon and V. D. Milman, λ_1, isoperimetric inequalities for graphs, and superconcentrators, J. Combin. Theory Ser. B **38** (1985), 73–88.

[5] N. Alon and V. D. Milman, Eigenvalues, expanders and superconcentrators, in *Twenty-fifth annual symposium on foundations of computer science*, Academic Press, Orlando, FL, 1986, 320–322.

[6] N. Alon and Y. Roichman, Random Cayley graphs and expanders, Random Structures Algorithms **5** (1994), 271–284.

[7] B. Bollobás, *Modern graph theory*, Springer, New York, 1998.

[8] G. Davidoff, P. Sarnak and A. Valette, *Elementary number theory, group theory, and Ramanujan graphs*, Cambridge Univ. Press, Cambridge, 2003.

[9] N. Kahale, On the second eigenvalue and linear expansion of regular graphs, in *Expanding graphs*, Amer. Math. Soc., Providence, RI, 1993, 49–62.

[10] M. Kassabov, A. Lubotzky and N. Nikolov, Finite simple groups as expanders, Proc. Natl. Acad. Sci. USA **103** (2006), 6116–6119.

[11] M. Klawe, Limitations on explicit constructions of expanding graphs, SIAM J. Comput. **13** (1984), 156–166.

[12] Z. Landau and A. Russell, Random Cayley graphs are expanders: a simple proof of the Alon-Roichman theorem, Electron. J. Combin. **11** (2004), Research Paper 62.

[13] P.-S. Loh and L. J. Schulman, Improved expansion of Random Cayley Graphs, Discrete Math. Theor. Comput. Sci. **6** (2004), 523–528.

[14] A. Lubotzky, *Discrete groups, expanding graphs and invariant measures*, Birkhäuser, Basel, 1994.

[15] A. Lubotzky and B. Weiss, Groups and expanders, in J. Friedman (ed.), *Expanding graphs*, Amer. Math. Soc., Providence, RI, 1993, 95–109.

[16] G. A. Margulis, Explicit constructions of expanders, Problemy Peredači Informacii **9** (1973), 71–80, (English translation in Problems Inform. Transmission **10** (1975), 325–332).

[17] O. Reingold, S. Vadhan and A. Wigderson, Entropy waves, the zig-zag graph product, and new constant-degree expanders, Ann. of Math. (2) **155** (2002), 157–187.

[18] G. Sabidussi, Vertex-transitive graphs, Monatsh. Math. **68** (1964), 426–438.

[19] J.-P. Serre, *Linear representations of finite groups*, Springer, New York, 1977.

[20] A. Wigderson and D. Xiao, Derandomizing the AW matrix-valued Chernoff bound using pessimistic estimators and applications, Electronic Colloquium on Computational Complexity, Report TR06-105, ISSN 1433-8092, 13th Year, 105th Report.

Demetres Christofides

University of Cambridge
Department of Pure Mathematics and Mathematical Statistics
Centre for Mathematical Sciences
Wilberforce Road
Cambridge CB3 0WB
United Kingdom

D.Christofides@dpmms.cam.ac.uk

Klas Markström

Department of Mathematics and Mathematical Statistics
Umeåuniversitet
SE-901 87
Umeå
Sweden

Klas.Markstrom@math.umu.se