Near-Optimal Cayley Expanders for Abelian Groups

Akhil Jalan[1]

Department of Computer Science, University of Texas at Austin, USA

akhil@cs.utexas.edu

Dana Moshkovitz

Department of Computer Science, University of Texas at Austin, USA

danama@cs.utexas.edu

Abstract

We give an efficient deterministic algorithm that outputs an expanding generating set for any finite abelian group. The size of the generating set is close to the randomized construction of Alon and Roichman [9], improving upon various deterministic constructions in both the dependence on the dimension and the spectral gap. By obtaining optimal dependence on the dimension we resolve a conjecture of Azar, Motwani, and Naor [14] in the affirmative. Our technique is an extension of the bias amplification technique of Ta-Shma [40], who used random walks on expanders to obtain expanding generating sets over the additive group of $\mathbb{F}_2^n$. As a consequence, we obtain (i) randomness-efficient constructions of almost k-wise independent variables, (ii) a faster deterministic algorithm for the Remote Point Problem, (iii) randomness-efficient low-degree tests, and (iv) randomness-efficient verification of matrix multiplication.

2012 ACM Subject Classification Theory of computation → Pseudorandomness and derandomization; Theory of computation → Error-correcting codes

Keywords and phrases Cayley graphs, Expander walks, Epsilon-biased sets, Derandomization

Digital Object Identifier 10.4230/LIPIcs.FSTTCS.2021.40

Related Version The full version of this paper is [29].
Full Version: `https://arxiv.org/abs/2105.01149`

Funding This material is based upon work supported by the National Science Foundation under grants number 1218547 and 1678712.

1 Our Contributions

1.1 Main Result

A graph is an expander if there exists a constant $\alpha > 0$ such that the spectral gap of its adjacency matrix (namely, the difference between its top eigenvalue and its second eigenvalue) is at least α. Such graphs are very well-connected in the sense that they lack sparse cuts. Expanders that are additionally sparse are immensely important in computer science and mathematics (see, e.g. the survey [28]).

Cayley graphs are an important class of graphs built from groups. Given a group G and a generating set $S \subset G$, the graph $\mathrm{Cay}(G, S)$ has vertex set G and edges $(g, g \cdot s)$ for all $g \in G$, $s \in S$. In addition to describing various well-known graphs such as the hypercube and the torus, Cayley graphs of (non-abelian) groups gave the first explicit constructions of near-optimal expander graphs [34]. Moreover, their algebraic structure makes Cayley

[1] Corresponding author.

graphs easier to analyze. In particular, the eigenvectors and eigenvalues of a Cayley graph are well-understood through the Fourier transform on the group.

When is a Cayley graph an expander? Alon and Roichman showed that given a group G, integer $n \geq 1$, and $\epsilon > 0$, taking a uniformly random subset $S \subset G^n$ of size $O(\frac{n\log(|G|)}{\epsilon^2})$ gives an expander with spectral gap $1 - \epsilon$, with high probability [9]. They also proved a nearly matching lower bound of $|S| = \Omega((\frac{n\log(|G|)}{\epsilon^2})^{1-o(1)})$ when G is abelian. When $G = \mathbb{F}_2$ the lower bound is $\Omega(\frac{n}{\epsilon^2\log(1/\epsilon)})$ [6] [2].

An explicit construction with parameters matching the Alon-Roichman bound has remained elusive, despite being widely studied in the pseudorandomness literature [32, 35, 6, 36, 1, 7, 26, 14, 23, 12, 17, 11].

The best known results achieve $O((\log(|G|) + \frac{n^2}{\epsilon^2})^5)$ for arbitrary abelian G [12], $O(\frac{n^2}{\epsilon^2})$ for abelian G where $|G| \leq \log(\frac{n^2}{\epsilon^2})^{O(1)}$, and $O(\frac{n\log(|G|)^{O(1)}}{\epsilon^{11}})$ for general G [23]. For solvable subgroups of permutation groups one can improve this to $O(\frac{n^2}{\epsilon^8})$ [11].

In this paper we give an explicit construction of expanding generating sets for abelian groups whose size is near the Alon-Roichman bound.

▶ **Theorem 1.** *There is a deterministic, polynomial-time algorithm which, given a generating set of an abelian group G, integer $n \geq 1$, and $\epsilon > 0$, outputs a generating set $S \subset G^n$ of size $O(\frac{n\log(|G|)^{O(1)}}{\epsilon^{2+o(1)}})$ such that $\mathrm{Cay}(G^n, S)$ has spectral gap $1 - \epsilon$.*

Our construction immediately improves parameters in several applications - see Section 1.3 for details. We remark that in most settings, one fixes a group G while $n \to \infty$ and $\epsilon \to 0$. In this regime, since $|G|$ is a constant, the size of the generating set in Theorem 1 is optimal up to an $\epsilon^{-o(1)}$ factor. The $o(1)$ term in the exponent approaches 0 as $\epsilon \to 0$.

Expanding Cayley graphs are equivalent to pseudorandom objects called ϵ-biased sets. These were originally defined over $\mathbb{F}_2^n$ by Naor and Naor [35]. A set $S \subseteq \mathbb{F}_2^n$ is said to be ϵ-biased if for every non-empty $T \subseteq [n]$, we have $\mathbb{E}_{x\in S}[\bigoplus_{i\in T} x_i] = 1/2 \pm \epsilon$.

Naor and Naor initiated a long line of work culminating in a recent breakthrough result by Ta-Shma, that achieves $|S| = O(\frac{n}{\epsilon^{2+o(1)}})$ [40]. This construction approaches the Alon-Roichman bound as $\epsilon \to 0$.

Ta-Shma's construction follows previous work in using a 2-step "bias amplification" approach. First, identify an explicit set $S_0 \subset \mathbb{F}_2^n$ with constant bias, usually through algebraic methods. Second, amplify the bias of S_0 to any $\epsilon > 0$ by performing a random walk on an expander graph. While this general method was already known, it could only achieve $|S| = O(\frac{n}{\epsilon^{4+o(1)}})$. To break this barrier, Ta-Shma identified a graph structure obtained from a "wide replacement product", which was more effective for the bias amplification step and resulted in $|S| = O(\frac{n}{\epsilon^{2+o(1)}})$.

Our main contribution is to show that the wide replacement walk is a near-optimal "character sampler," and therefore also amplifies bias well for abelian Cayley graphs.

1.2 Wide Replacement Walks are Near-Optimal Character Samplers

Random walks on expander graphs are useful for a variety of algorithmic purposes. A classical fact is that expander walks are good approximate samplers, in the sense that a sufficiently long random walk on an expander will visit sets of density δ for approximately a δ fraction

[2] It is possible that this lower bound is tight. A candidate construction based on algebraic-geometric codes could achieve this lower bound [17].

of the steps. This is called the "expander Chernoff bound" and one can characterize this as the property that expander walks fool a suitable test function.

Ta-Shma observed that expander walks fool the much more sensitive class of parity functions on $\{0,1\}^n$ as well. Parity functions are sensitive to input perturbations - flipping a single bit in the input can change the output. The classical expander Chernoff bound is not fine-grained enough to prove that t-step expander walks fool parity functions. The fact that they nevertheless do fool parity functions is therefore surprising, and Ta-Shma referred to this fact as "expanders are good parity samplers" [40].

Since parity functions are just the characters of $\mathbb{F}_2^n$, we can ask: do expander walks also fool the characters of more general classes of groups? We show that this is indeed true, and therefore "expander walks are good character samplers." Moreover, just as in the $\mathbb{F}_2$ case, a random walk on a wide replacement product of expander graphs is a near-optimal type of character sampler.

Character sampling explained: Let us precisely explain what we mean by "character sampling." A character of an abelian group is a homomorphism $\chi : G \to \mathbb{C}^*$, where $\mathbb{C}^*$ is the multiplicative group of complex numbers. The eigenvalues of an abelian Cayley graph $\mathrm{Cay}(G,S)$ are given by $|\mathbb{E}_{x\sim S}\,\chi(x)|$ for all characters χ. Note that the constant function that maps all values to 1 is a character, and the eigenvalue associated with it is the top eigenvalue. Therefore, we are interested in generating sets S such that $|\mathbb{E}_{x\sim S}\,\chi(x)| \le \epsilon$ for all non-constant χ.

For simplicity, consider the case $G = \mathbb{Z}_d$ for some $d \ge 2$. Let $\omega_d := \exp(\frac{2\pi i}{d})$. In this case the characters are just the maps $x \mapsto \omega_d^{x\cdot j}$ for $j = 0, 1, \ldots, d-1$.

Now, suppose we have some ϵ_0-biased set $G_0 \subset G$, where $\epsilon_0 < 1$ is a constant. First, observe that taking t *independent* samples from G_0 and outputting their sum obtains a distribution with bias $(\epsilon_0)^t$. However, since independent sampling also results in a distribution with support size $|G_0|^t$, there is no improvement in size as a function of bias.

The idea of the random walk approach is to derandomize independent sampling by taking *correlated* samples. Specifically, identify G_0 with the vertices of some degree-regular expander graph Γ. We need to show that taking a random walk of length t on Γ and then summing the elements in the path gives a distribution with lower bias than G_0.

A t-step walk on Γ gives a sequence of group elements $(x_0, \ldots, x_t) \in G_0^{t+1}$. We are interested in the bias of the random group element $\sum_i x_i$. In general, we cannot hope that $(\sum_i x_i)$ is close to the uniform distribution in *statistical distance*. However, if Γ is an expander with second eigenvalue λ, then for every non-constant character χ the quantity $|\mathbb{E}[\chi(\sum_i x_i)]|$ is at most $(\epsilon_0 + \lambda)^{\lfloor t/2 \rfloor}$, where the expectation is over paths $(x_0, \ldots, x_t)$ in the graph. Notice that $\mathbb{E}_{x\in G}[\chi(x)] = 0$, so the random element $(\sum_i x_i)$ is close to uniform in the weaker sense of fooling characters. Therefore, the expander walk is a good "character sampler."

Why expanders are character samplers: We express the bias of the random walk distribution algebraically in terms of matrix norms corresponding to the random walk.

Abusing notation, let Γ denote the random walk matrix of the graph Γ. Let the character $\chi^* : \mathbb{Z}_d \to \mathbb{C}$ be the worst-case character for the random-walk distribution. Partition G_0 into $S_0, \ldots, S_{d-1}$ depending on their values with respect to χ^*, so that $x \in S_k \iff \chi^*(x) = \omega_d^k$.

We need to track how often the walk enters $S_0, S_1, \ldots, S_{d-1} \subset V(\Gamma)$. Identify each S_i with an $|S_i|$-dimensional subspace of $\mathbb{C}^{V(\Gamma)}$. For $i \in \mathbb{Z}_d$ let $\Pi_i : \mathbb{C}^{V(\Gamma)} \to \mathbb{C}^{V(\Gamma)}$ be the projection onto this subspace. Finally, let $\Pi = \sum_{y\in\mathbb{Z}_d} \omega_d^y \Pi_y$ be the weighted projection matrix.

Given some initial distribution $\vec{u}$ on the vertices, the vector $\Gamma^t\vec{u}$ tracks the distribution after taking a t-step walk on the graph. The matrix Π tracks how often the walk enters the sets $S_0, \ldots, S_{d-1}$, and so the bias of the random walk distribution can be bounded by the norm of $(\Pi\Gamma)^t$.

Let $V^{\parallel}$ denote the subspace spanned by the all-ones vector $\vec{1}$, and $V^{\perp} = (V^{\parallel})^{\perp}$. For a vector $v \in V^{\parallel} \oplus V^{\perp}$, let $v^{\parallel}$ and $v^{\perp}$ denote the projections onto $V^{\parallel}, V^{\perp}$ respectively.

While $\|\Pi\Gamma\| = 1$ since $\|\Pi\Gamma\vec{1}\| = \|\Pi\vec{1}\| = 1$, it turns out that $\|(\Pi\Gamma)^2\| \leq bias(G_0) + 2\lambda(\Gamma)$, where $\lambda(\Gamma)$ is the second eigenvalue of Γ in absolute value.

To see this, notice that if $\vec{v} \in V^{\perp}$ is a unit vector, then $\|\Pi\Gamma\Pi\Gamma\vec{v}\| \leq \|\Pi\Gamma\Pi\|\lambda(\Gamma)\|\vec{v}\| \leq \lambda(\Gamma)$. Therefore, the "bad" case is when $\vec{v} \in V^{\parallel}$. Let $u = \frac{1}{\sqrt{|V(\Gamma)|}}\vec{1}$. Using the fact that $\|\Pi\| = 1$,

$$\begin{aligned}
\|\Pi\Gamma\Pi\Gamma u\| &= \|\Pi\Gamma\Pi u\| \\
&\leq \|\Pi\Gamma(\Pi u)^{\parallel}\| + \|\Pi\Gamma(\Pi u)^{\perp}\| \\
&\leq \|\Pi(\Pi u)^{\parallel}\| + \lambda(\Gamma)\|\Pi(\Pi u)^{\perp}\| \\
&\leq \|\Pi(\Pi u)^{\parallel}\| + \lambda(\Gamma)
\end{aligned}$$

It remains to show that $\|\Pi(\Pi u)^{\parallel}\| \leq bias(G_0)$. To see this, notice that Π is a diagonal matrix and u is just $\vec{1}$ scaled by a constant. Further, Π is a block-diagonal matrix of the form

$$\Pi = \begin{bmatrix} I_{|S_0|} & & & \\ & \omega_d I_{|S_1|} & & \\ & & \ddots & \\ & & & \omega_d^{d-1} I_{|S_{d-1}|} \end{bmatrix}$$

Note that we have reordered the vertices of the graph in order of S_0, S_1 and so on.

If the blocks are exactly the same size, then $\Pi u \in V^{\perp}$, because $\sum_{y \in \mathbb{Z}_d} \omega_d^y = 0$. In general the blocks have different dimensions, but they are the same size up to the bias of G_0. Therefore $\|(\Pi u)^{\parallel}\| \leq bias(G_0)$.

It follows that a random walk on Γ is a good character sampler. However, this approach can never amplify bias fast enough to achieve a generating set smaller than $O(\frac{|G_0|}{\epsilon^{4+o(1)}})$. The reason is because while we can bound $\|(\Pi\Gamma)^2\|$, we cannot bound $\|\Pi\Gamma\|$ below 1. Therefore, we effectively only gain from one in every two steps.

Wide Replacement Walks are Near-Optimal Character Samplers: To circumvent the "2-step barrier" of expander walks outlined above, Ta-Shma used the *wide replacement walk* on a product of two expander graphs [40]. The idea of the wide replacement walk is to take the product of a D_1-regular graph Γ as before with an "inner graph" H on D_1^s vertices, for some $s \geq 2$. The product graph replaces every vertex of Γ with a copy of H (called a "cloud") and then connects clouds to other clouds according to the edge structure of Γ.

Analyzing the bias of the walk involves bounding the matrix norm of $\dot{\Pi}\dot{\Gamma}\dot{H}$, where $\dot{\Gamma}$ and $\dot{H}$ are random walk matrices on the product corresponding to Γ, H.

Let $V^{\parallel}$ denote the subspace of vectors which are constant on the H-component of the product, and let $V^{\perp} = (V^{\parallel})^{\perp}$.

Similar to the above case, one can show that $\dot{\Pi}\dot{\Gamma}\dot{H}$ shrinks the norm of any $v \in V^{\perp}$ by a factor of $\lambda(H)$. The difficult case is when $v \in V^{\parallel}$. Here we arrive at the core idea of the replacement product: if the inner graph H is *pseudorandom* with respect to Γ, then when the walk is in $V^{\parallel}$, the next s steps approximate the ordinary random walk on Γ.

This is enough to circumvent the "2-step barrier" since in even the "bad case" where the walk is stuck in $V^{\parallel}$, we can shrink the bias as though it were taking an ordinary walk on Γ. As we showed above, this shrinks the bias from some ϵ_0 to $(\epsilon_0 + 2\lambda(\Gamma))^{\lfloor s/2 \rfloor}$ every s steps. If we select Γ, H such that $\epsilon_0 + 2\lambda(\Gamma) \le \lambda(H)^2$, then we conclude that we shrink the bias by a factor of $\lambda(H)^{s-O_s(1)}$ every s steps. So we gain from $s - O(1)$ out of every s steps.

Going from the $\mathbb{F}_2$-case to the case of general abelian groups simply requires a more careful analysis of characters. We defer the full proof to Appendix 2.2.

Morally speaking, the only difference in the analysis is that the projection matrix Π which tracks how often the walk enters each S_i is different. This does not change the overall argument much; in particular, we can use almost identical graphs Γ, H as in [40].

We conclude that a wide replacement walk allows us to amplify bias of a constant-biased subset $G_0 \subset G^n$ of size $O(n\log(|G|)^{O(1)})$ (e.g. the construction of [11]) to an ϵ-biased set of size $O(\frac{n\log(|G|)^{O(1)}}{\epsilon^{2+o(1)}})$, nearly matching the Alon-Roichman bound. For explicit parameters of the construction, see Appendix C.

1.3 Applications

Explicit constructions of expander graphs are an essential component of algorithms, especially for derandomization. Here we are interested in the setting of constructing an expanding Cayley graph from a given abelian group G. Our construction achieves a near-optimal degree, which improves parameters in various applications. We defer precise statements of these results and the full proofs to the full version.

Almost k-wise independence: A distribution $D \sim G^n$ is (ϵ, k)-wise independent if for every index set $I \subset [n]$ of size k, the restriction of D to I is ϵ-close to uniform in statistical distance. Almost k-wise independent distributions are a fundamental object in and of themselves. They also have a variety of applications in derandomization, including load balancing [24], derandomization of Monte-Carlo simulations [24], derandomization of CSP approximation algorithms [21], and pseudorandom generators [22]. We note that certain applications (e.g. quantum t-designs [10]) really require almost k-wise independent distributions over *arbitrary* alphabet size rather than just the binary alphabet, which motivates our study of ϵ-biased sets over arbitrary abelian groups.

Vazirani's XOR Lemma asserts that an ϵ-biased distribution D is also $(\epsilon\sqrt{|G|^k}, k)$-wise indepdent for all $k \le n$. Therefore, by constructing an ϵ'-biased distribution where $\epsilon' = \frac{\epsilon}{\sqrt{|G|^k}}$, we also obtain explicit constructions of (ϵ, k)-wise independent random variables on G^n.

▶ **Proposition 2** (Almost k-wise independent sets over abelian groups)**.** *Let G be a finite abelian group given by some generating set. For any $\epsilon > 0$ and $n \ge k \ge 1$ there exists a deterministic, polynomial-time algorithm whose output is an (ϵ, k)-wise independent distribution over G^n. The support size is $O(\frac{n\cdot|G|^{k+o(1)}}{\epsilon^{2+o(1)}})$.*

Remote Point Problem: A matrix $A \in \mathbb{F}_2^{m\times n}$ is (k, d)-rigid iff for all rank-k matrices $R \in \mathbb{F}_2^{m\times n}$, the matrix $A - R$ has a row with at least d nonzero entries. Valiant initiated the study of rigid matrices in circuit complexity, proving that an explicit construction of an $(\Omega(n), n^{\Omega(1)})$-rigid matrix for $m = O(n)$ would imply superlinear circuit lower bounds [43]. After more than four decades of research, state of the art constructions have yet to meet this goal [19].

The Remote Point Problem was introduced by Alon, Panigrahy, and Yekhanin as an intermediate problem in the overall program of rigid matrix constructions [8]. Arvind and Srinivasan generalized the problem to any group [12].

Let G be a group, $n \ge 1$, and $H \le G^n$ a subgroup given by some generating set. For a given G, H and integer $r > 0$, the Remote Point Problem is to find a point $x \in G^n$ such that x has Hamming distance greater than r from all $h \in H$, or else reject. In the case of

$G^n = \mathbb{F}_2^n$, this is a relaxation of the matrix rigidity problem, since rather than finding m vectors $x_1, \ldots, x_m \in \mathbb{F}_2^n$ whose linear span is far from all low-dimensional subspaces, we are given a single subspace and must find just a single point far from it.

To find a remote point, existing algorithms first construct a collection of subgroups $H_1, \ldots, H_m \leq G^m$ whose union covers all points of distance at most r from H. In the $\mathbb{F}_2$ case, [8] find a point $x \notin \bigcup_i H_i$ by the method of pessimistic estimators. In the general case, [12] instead prove that any generating set $S \subset G^n$ such that $\mathrm{Cay}(G^n, S)$ has sufficiently good expansion must contain a point outside of $\bigcup_i H_i$. They find this remote point by first constructing an expanding generating set S, and then exhaustively searching it. Their argument implicitly uses the fact that small-bias sets correspond to rigid matrices, albeit with weak parameters - this connection was developed further in [5].

The construction of [12] for small-bias sets over abelian groups has size $O((\log(|G|) + \frac{n^2}{\epsilon^2})^5)$ in general, and for $\log(|G|) \leq \log(\frac{n^2}{\epsilon^2})^{O(1)}$ this is improved to $O(\frac{n^2}{\epsilon^2})$. Our algorithm improves the dependence on n from n^2 to n.

Randomness-Efficient Low-Degree Testing: Let $\mathbb{F}_q$ be the finite field on q elements. Low-degree testing is a property testing problem in which, when given query access to a function $f : \mathbb{F}_q^n \to \mathbb{F}_q$ and $d \geq 1$, one must decide whether f is a degree d polynomial or far (in Hamming distance) from all degree d polynomials. These tests are a key ingredient in constructions of Locally Testable Codes (LTCs) and Probabilistically Checkable Proofs (PCPs) [18].

To test whether f is a degree-d polynomial, a natural test is to sample $x, y \sim \mathbb{F}^n$ and check whether $f(x)$ agrees with the unique (degree-d, univariate) polynomial obtained by Lagrange interpolation along $d+1$ points on the line $\{x + ty : t \in \mathbb{F}_q\}$.

Rubinfeld and Sudan introduced a low-degree test using this idea [38]. It is given query access to the function f, along with a *line oracle* function g. Let $\mathbb{L}$ denote all lines $\{\vec{a} + t\vec{b} : t \in \mathbb{F}_q\} \subset \mathbb{F}_q^n$, where $\vec{a}, \vec{b} \in \mathbb{F}^n$. Given a description of a line, the line oracle g returns a univariate polynomial of degree d defined on that line. Hence we write $g : \mathbb{L} \to \mathbb{F}_q[t]$, where the image of g is understood to only contain degree-d polynomials.

If f is indeed a degree-d polynomial, then one can set $g(\ell) = f|_\ell$ for all $\ell \in \mathbb{L}$, and the following two-query test clearly accepts.

(i) Select $x, y \in \mathbb{F}^n$ independently, uniformly at random.

(ii) Let ℓ be the line determined by $\{x + ty : t \in \mathbb{F}\}$. Accept iff $f(x)$ agrees with $g(\ell)(x)$.

They also showed this test is sound: when f is far from degree-d polynomials, the test rejects with high probability.

Ben-Sasson et al derandomized this test by replacing the second uniform sample y with a sample from an ϵ-biased set [18]. This modification improves the randomness efficiency of the tests, and therefore the length of the resulting LTC and PCP constructions. Moreover, they showed that the soundness guarantees of low-degree tests are almost unchanged due to the expansion properties of the Cayley graph on $\mathbb{F}_q^n$.

Our constructions of small-bias sets immediately imply improved randomness-efficiency of this low-degree test.

► **Proposition 3** (Improved [18] Theorem 4.1). *Let $\mathbb{F}_q$ be the finite field of q elements, $n \geq 1$, $f : \mathbb{F}_q^n \to \mathbb{F}_q$ a function, and $g : \mathbb{L} \to \mathbb{F}_q[t]$ a line oracle. There exists a degree-d test which has sample space size $O(q^n \cdot \frac{n \log(q)^{O(1)}}{\epsilon^{2+o(1)}})$. For $d \leq q/3$ and sufficiently small $\delta > 0$, if the test accepts with probability $\geq 1 - \delta$ then f has Hamming distance at most 4δ from a degree d polynomial.*

Randomness-Efficient Verification of Matrix Multiplication: Let R denote some finite field $\mathbb{F}_q$ or cyclic group $\mathbb{Z}_q$ for $q \geq 2$. Given $A, B, C \in R^{n\times n}$, the matrix multiplication verification problem asks whether $AB = C$.

Naively, one could multiply A, B and then check whether $AB = C$ entry-wise in $O(n^\omega)$ time, where $\omega \approx 2.373$ [2]. A classical result of Freivalds suggests the following much simpler quadratic-time randomized algorithm: Sample $x \in R^n$ and check whether $ABx = Cx$ [27].

Observe that the entries of ABx and Cx are linear functions of x. Therefore, sampling x from a small-bias set gives a randomness-efficient version of Freivalds' algorithm, at the cost of slightly higher error. Our construction therefore gives the following randomness efficient algorithm for verification of matrix multiplication.

▶ **Proposition 4.** *Let R denote a finite field $\mathbb{F}_q$ or cyclic group $\mathbb{Z}/q\mathbb{Z}$. Given matrices $A, B, C \in R^{n\times n}$ and ϵ-biased set $S \subset R^n$, there exists randomized algorithm to decide whether $AB = C$ with one-sided error $(\frac{1}{q}+\epsilon)$. Its runtime is $O(n^2)$ and it uses $\log(\frac{n\log(q)^{O(1)}}{\epsilon^{2+o(1)}})$ random bits.*

We note that if $R = \mathbb{Z}$, there exists a deterministic $O(n^2)$ time algorithm to verify matrix multiplication [33]. However, this result relies on the fact that $\mathbb{Z}$ has characteristic zero. For the analysis to hold in the case of $\mathbb{Z}_q$, we would need a very strong bound on the entries of A, B, C - namely, that $\max_{i,j}\{|A_{i,j}|, |B_{i,j}|, |C_{i,j}|\} \leq q^{\frac{1}{n-1}}$.

1.4 Related Work

Explicit Constructions: Explicit constructions of expanding generating sets for Cayley graphs have been mostly studied in the pseudorandomness literature in the context of small-bias sets for derandomization. Naor and Naor gave a combinatorial construction over $\mathbb{F}_2^n$ of size $O(\frac{n}{\epsilon^3})$ [35]. Alon, Goldreich, Hastad, and Peralta used algebraic arguments to give constructions over finite fields $\mathbb{F}^n$ of size $O(\frac{n^2}{\epsilon^2})$, assuming the field size is bounded as $\log(|\mathbb{F}|) < \frac{n}{\log(n)+\log(1/\epsilon)}$ [6].

Resarchers in various communities have obtained constructions that achieve size $O(\text{poly}(\frac{n\log(|G|)}{\epsilon}))$, but suboptimal exponents. In number theory and additive combinatorics researchers studying the case of $n = 1$ gave constructions over $\mathbb{Z}_d$ of size $O((\frac{\log(d)}{\epsilon})^{O(1)})$ [36], $O(\frac{\log(d)^{O(1)}}{\epsilon^2})$ [32], and $O(\frac{d}{\epsilon^{O(\log^*(d))}})$ [1].

Other constructions equivalent to small-bias sets include $O(\frac{(n-1)^2}{\epsilon^2})$-sized ϵ-discrepancy sets over finite fields of prime order p when $n \leq p$ [7], and ϵ-balanced codes over finite fields, corresponding to small-bias sets over $\mathbb{F}_q^n$ of size $O(n \cdot q)$ with constant bias [31].

Ta-Shma's tour de force gave the first explicit construction of expanding generating sets of size $O(\frac{n\log(|G|)}{\epsilon^{2+o(1)}})$, nearly attaining the Alon-Roichman bound, but only for the special case of $G = \mathbb{F}_2$ [40]. Our work is an extension of Ta-Shma's bias amplification technique to the more general setting of arbitrary abelian groups.

Azar, Motwani, and Naor generalized the study of small-bias sets to finite abelian groups [14]. Over $\mathbb{Z}_d^n$ they used character sum estimates to give a construction of size $O((d + \frac{n^2}{\epsilon^2})^C)$, where $C \leq 5$ is Linnik's constant [45]. Assuming the Extended Riemann Hypothesis, $C \leq 2 + o(1)$ [15]. When $\log(d) \leq \log(\frac{n^2}{\epsilon^2})^{O(C)}$ they improve the size to $O((1 + o(1))\frac{n^2}{\epsilon^2})$.

Arvind and Srinivasan proved that one can project small-bias sets over $\mathbb{Z}_d^n$ to any abelian group G^n when d is the largest invariant factor of G. Therefore, using the construction from [14] they obtain small-bias sets over G^n with the same bias and size as [14], with $d = O(\log(|G|))$ [12].

The most general setting is to consider Cayley graphs over non-abelian groups. Wigderson and Xiao derandomized the Alon-Roichman construction using the method of pessimistic estimators [44]. Arvind, Mukhopadhyay, and Nimbhorkhar later gave a derandomization for both directed and undirected Cayley graphs using Erdos-Renyi sequences [13]. However, both algorithms require the entire group table of G^n as input, rather than just a generating set. Since generating sets are of size $O(n\log(|G|))$, these algorithms are exponentially slower, running in time $O(\text{poly}(|G|^n))$ rather than $O(\text{poly}(n\log(|G|))$. Nevertheless, they have applications to settings such as homomorphism testing [39], which Wigderson and Xiao derandomized using their construction of expanding generating sets [44].

Chen, Moore, and Russell obtained generating sets of size $O(\frac{n\log(|G|)^{O(1)}}{\epsilon^{11}})$ over arbitrary groups G^n when $|G|$ is a constant [23] . Like Ta-Shma, their technique is to use bias amplification via expander graphs; specifically, they amplify bias via an iterated application of a 1-step random walk on an expander graph. Alon in 1993, and later Rozenman and Wigderson in 2004, had already noted that this technique amplifies bias for $G = \mathbb{F}_2$ [25]. Chen, Moore, and Russell generalized this analysis to all groups, using techniques from harmonic analysis and random matrix theory [23].

Existing work seems far from obtanining constructions for non-abelian groups near the Alon-Roichman bound. Known work tends to concentrate on special classes of non-abelian groups with some useful algebraic structure. Chen, Moore, and Russell constructed generating sets of size $O(\frac{(n\log(|G|))^{1+o(1)}}{\epsilon^{O(1)}})$ for smoothly solvable groups with constant-exponent abelian quotients [23]. Their analysis exploits the structure of solvable groups via Clifford theory. It also hinges on the assumption that the quotients in the derived series have constant exponent.

Arvind et al later gave a construction of size $\tilde{O}(\frac{\log(|G|)^{2-o(1)}}{\epsilon^8})$ for solvable subgroups G of permutation groups [11]. Their construction recursively generates expanding generating sets for quotients in the derived series of the group, and uses the thin sets construction of [1] as a base set. Unlike [23] they do not require successive quotients of the derived series to be small; however, their argument does rely on an $O(\log(n))$ upper bound on the length of the derived series for any solvable $G \leq S_n$, which is not true for solvable groups in general.

Lower Bounds: Alon and Roichman gave a randomized upper bound of $O(\frac{n\log(|G|)}{\epsilon^2})$ on the size of a generating set for any finite G^n with spectral gap $(1-\epsilon)$ [9]. In the same paper, they gave a nearly matching lower bound when G is abelian, of $\Omega((\frac{n\log(|G|)}{\epsilon^2})^{1-o(1)})$. This is a sharper version of the folklore result that an abelian group G^n requires $O(n\log(|G|))$ generators for its Cayley graph to be connected.

For non-abelian groups, the existence of sparse expanders means the best lower bound in general is the Alon-Boppana bound. This removes the dependence on $|G|$ and n, only requiring a generating set of size $\Omega(\frac{1}{\epsilon^2})$ [3] to achieve spectral gap of $1-\epsilon$. Indeed, explicit constructions of Ramanujan graphs can be built from Cayley graphs of non-abelian groups [34], and therefore attain this bound.

Expander Walks: Random walks on expander graphs are an essential tool in computer science. Rather than surveying the vast literature, we refer the reader to the surveys [28, 42]. Two remarks are in order.

First, our use of wide replacement walks is essentially a way of building expander graphs from other expander graphs. This is thematic of several previous works, such as the zig-zag product [37]. Note that the zig-zag product is just a modification of the replacement product; indeed, the (wide) replacement product itself can be used to give explicit, combinatorial constructions of Ramanujan graphs [16]. Ta-Shma used wide replacement walks to amplify spectral gaps of Cayley graphs on $\mathbb{F}_2^n$ [40]; this construction relied on previous constructions of expander graphs, although the expander graphs were not required to be Cayley graphs themselves.

Second, the fact that "expanders are good character samplers" is surprising given that characters are sensitive to input perturbations. A recent work of Cohen, Peri, and Ta-Shma uses Fourier-analytic techniques to classify a large class of Boolean functions which can be fooled by expander walks, including all symmetric Boolean functions [25].

1.5 Open Problems

In this work, we gave an efficient deterministic algorithm to compute an expanding generating set of an abelian group. Our construction achieves optimal dependence on dimension and near-optimal dependence on error, resulting in improvements in various applications. Here, we discuss some natural open questions raised by our work.

Expanding generating sets of optimal size: The Alon-Roichman theorem proves that every group G^n has an expanding generating set $S \subset G^n$ of size $|S| = O(\frac{\log(|G|)}{\epsilon^2})$ [9]. This construction has not been fully derandomized for any group; even in the case of $G^n = \mathbb{F}_2^n$, Ta-Shma's construction only asympotically approaches a size of $O(\frac{n}{\epsilon^2})$ as $\epsilon \to 0$. The actual size of the generating set is $O(\frac{n}{\epsilon^{2+o(1)}})$, and this $o(1)$ term is seemingly unavoidable when using expander walks [40].

Similarly, our algorithm gives an expanding generating $S \subset G^n$ of size $O(\frac{n \log(|G|)^{O(1)}}{\epsilon^{2+o(1)}})$, for finite abelian G. The additional $\operatorname{poly}\log(|G|)$ factor comes from the bounds on constant-bias subsets of abelian groups; any construction of a constant-bias set $S \subset G^n$ of size $O(n \log(|G|))$ would immediately give expanding generating sets of size $O(\frac{n \log(|G|)}{\epsilon^{2+o(1)}})$. To our knowledge, not even a candidate construction exists which would give constant-bias subsets of size $O(n \log(|G|))$ for abelian groups; this is an interesting and potentially easier open problem, since it requires none of the expander walks machinery that we need to get arbitrarily small ϵ.

There is a candidate construction that could beat the Alon-Roichman bound for $G = \mathbb{F}_2$, based on algebraic-geometric codes [17]. The code construction would give an ϵ-biased set $S \subset \mathbb{F}_2^n$ of size $|S| = O(\frac{n}{\epsilon^2 \log(1/\epsilon)})$, assuming a conjecture in algebraic geometry. The authors themselves note that they have "no idea" whether this conjecture is valid [17].

Expanding generating sets of non-abelian groups: While wide replacement walks amplify bias quite naturally for abelian groups, it is unclear whether they can do so for general groups. Dealing with matrix-valued irreducible representations, rather than scalar-valued characters, makes the analysis of bias amplification considerably more involved; hence even the analysis of the 1-step walk is nontrivial [23]. It would be very interesting to see whether one can place algebraic conditions on a group that are weaker than commutativity, but still ensure that the wide replacement walk amplifies bias.

Existing works on expanding generating sets for non-abelian groups have studied solvable groups, which generalize abelian groups [23, 11]. However, if we restrict the algorithm to input instances which are all non-abelian groups, then existence results suggest that one should be able to *beat* the Alon-Roichman bound.

For example, it is known that for every finite *simple* non-abelian group G^n, there exists a generating set $S \subset G^n$ such that $\operatorname{Cay}(G^n, S)$ has spectral gap $1 - \epsilon$, and $|S|$ is independent of n [20]. Therefore, restricting input instances to simple groups seems too easy, while an algorithm for all groups seems too hard. Is there some natural natural class of non-abelian, non-simple groups for which algorithms can efficiently find expanding generating sets near (or even below) the Alon-Roichman bound?

Decoding over any finite field: A recent work of Jeronimo et al gives a decoding algorithm for a modified version of Ta-Shma's codes [30]. Since our work gives ϵ-balanced codes over any finite field, it would be interesting to extend both the modification of the codes and the decoding algorithm of [30] to this general setting.

Classifying the power of expander walks on groups: So far we have discussed how random walks on expanders are good samplers in various ways, such as the expander Chernoff bound, parity sampling, and character sampling. Cohen, Peri, and Ta-Shma study the class of all Boolean functions that expander walks fool [25]. It would be very interesting to extend their results to functions on groups, perhaps using similar tools from harmonic analysis and representation theory. For example, for which groups G besides $\mathbb{F}_2$ do expander walks fool all symmetric functions on G^n?

1.6 Organization

The rest of this paper is organized as follows. In Section 2 we prove that our wide replacement walk construction gives an expanding generating set over any finite abelian group with near-optimal degree. Due to space constraints we defer some proofs to the full version of the paper.

Appendix C contains the precise parameters of the construction. Appendices A and B contain technical preliminaries on Cayley graphs and wide replacement walks, respectively.

2 Expanding Generating Sets for Abelian Groups

Throughout this section, let G be a finite abelian group and $n \geq 1$. In this section, we will describe an efficient deterministic algorithm to construct a generating set $S \subset G^n$ such that the Cayley graph $Cay(G^n, S)$ has second eigenvalue at most ϵ. The degree is $|S| = O(\frac{n \log(|G|)^{O(1)}}{\epsilon^{2+o(1)}})$.

The inputs to our algorithm are a generating set $G' \subset G$, integer $n \geq 1$, and desired expansion $\epsilon > 0$. The algorithm proceeds as follows:

(i) Construct an ϵ_0-biased set $S_0 \subset G^n$ with support size $O(n \log(|G|)^{O(1)})$ for a constant $\epsilon_0 < 1$.

(ii) Perform a wide replacement walk to amplify the bias of S_0 to ϵ. Specifically, we identify S_0 with the vertices of an outer graph Γ, and then choose an inner graph H in a manner described later. We emphasize that while Γ is an expander graph whose vertex set is S_0, it is not required to be a Cayley graph on S_0. For the purposes of this step, the group structure of G is irrelevant.

Let $t \geq 1$ be the walk length, to be chosen later. The output ϵ-biased set $S \subset G^n$ corresponds to length-t walks on the wide replacement product of Γ and H. Given a sequence of vertices $(x_0, ..., x_t) \in V(\Gamma) \times V(H)$, we add up the components corresponding to $V(\Gamma)$, which are just elements of S_0, to obtain some element of G^n. This gives the elements of S.

Next, let us informally describe parameter choices (precise choices are in section C). Let D_2 be the degree of H. At every step in the wide replacement walk we need to specify some $i \in [D_2]$ to take a step. It follows that $S \subset G^n$ has a size of $O(n \log(|G|)^{O(1)} \cdot D_2^t)$. We must choose t large enough to shrink the bias to ϵ. The choice t (walk length) and D_2 (degree of the inner graph) will determine the overall size of the output generating set.

These choices hinge on the bias amplification bound of the wide replacement walk. We show that the s-wide replacement walk shrinks the bias by a factor of $O(s^2 \cdot \lambda(H)^{s-3})$ every s steps. However, the size of the walk distribution grows by a factor of $O(D_2^s)$ every s steps. This imperfect bias amplification is why we cannot get optimal dependence on ϵ, as that would require that the bias shrinks by exactly $O(\lambda(H)^s)$ every s steps.

Therefore we cannot choose H to be an optimal spectral expander with $\lambda(H) = \Theta(\frac{1}{\sqrt{D_2}})$. Instead, optimizing for the size of the output distribution, we set $s = \Theta(\frac{\log(1/\epsilon)^{1/3}}{\log\log(1/\epsilon)^{1/3}})$, second eigenvalue $\lambda(H) = \Theta(\frac{s\cdot\log(D_2)}{\sqrt{D_2}})$, and the walk length $t = \Theta(\frac{\log(1/\epsilon)}{\log(1/\lambda(H))} \cdot \frac{s^2}{s^2-5s+1}) = \Theta((\frac{\log(1/\epsilon)}{\log(1/\lambda(H))})^{1+o(1)})$. This is exactly the reason our output set has a dependence of $O(\frac{1}{\epsilon^{2+o(1)}})$ rather than exactly $O(\frac{1}{\epsilon^2})$, and the same is true for [41].

This section is organized as follows. In section 2.1, we describe how one can identify the elements S_0 with the vertices of an expander graph, and then perform the ordinary random walk on the graph to amplify the bias of S_0, albeit suboptimally. In section 2.2 we show how to express the bias of a wide replacement walk algebraically. In section 2.3 we prove an upper bound on this algebraic expression, therefore proving the bias amplification bound of the wide replacement walk. Finally, in section C we describe the details and exact parameters for the wide replacement walk, as well as the ϵ_0-biased subset of G^n.

2.1 The ordinary expander walk

Let G be a finite abelian group. For ease of notation, we will refer to G rather than G^n until section C, when we need to discuss parameters. Since H^n is a finite abelian group for all abelian H, there is no loss of generality.

In this section we will show how to amplify the bias of a small-bias set in G by performing a random walk on an expander. This will be a lemma in the analysis of our actual construction, which involves a *wide replacement walk*.

To state the bias amplification theorem, we need some notation.

Let $G = \mathbb{Z}_{d_1} \oplus \cdots \oplus \mathbb{Z}_{d_k}$ be the invariant factor decomposition of G. Notice that $d_i|d_j$ for any $i < j$. In particular, all d_i divide d_k. For $x \in G$ write $x = (x_1, ..., x_k)$, so that $x_i \in \mathbb{Z}_{d_i}$ for each i.

Fix a nontrivial character $\chi : G \to \mathbb{C}^*$ corresponding to a group element $a \in G$. Let $a = (a_1, ..., a_k)$. Then for a given $x \in G$, $\chi(g) = \omega_{d_1}^{a_1\cdot x} \cdots \omega_{d_k}^{a_k\cdot x}$. Since all d_i divide d_k, we can write this as

$$\chi(g) = \omega_{d_k}^{\sum_{i=1}^{k}(\frac{d_k}{d_i}a_i\cdot x_i) \mod d_k}$$

Now, let $S_{init} \subset G$ have bias ϵ_0. Identify S_{init} with the vertices of some degree-regular expander graph Γ. We write $V := V(\Gamma) = S_{init}$. In order to understand the bias of a random walk on Γ with respect to χ, we have to track how often the walk enters vertices which map to $\omega_{d_k}, \omega_{d_k}^2$, and so on.

We will partition S_{init} as follows. For $y \in \mathbb{Z}_{d_k}$, let S_y be the elements of S_{init} which are mapped to $\omega_{d_k}^y$ by χ. Formally, $S_y = \{x \in S_{init} : y = (\sum_{i=1}^{k} \frac{d_k}{d_i} x_i \cdot a_i) \mod d_k\}$. Observe that $\{S_y : y \in \mathbb{Z}_{d_k}\}$ is a partition of S_{init}.

Next, let $t > 0$ be the walk length. We will partition all length-$(t+1)$ sequences in S_{init} according to their sum. For $y \in \mathbb{Z}_{d_k}$, let $T_y = \{b \in \mathbb{Z}_{d_k}^{t+1} : (\sum_i b_i) \mod d_k = y\}$. Again, notice that $\{T_y : y \in \mathbb{Z}_{d_k}\}$ is a partition of $\mathbb{Z}_{d_k}^{t+1}$.

Finally, fix $y \in \mathbb{Z}_{d_k}$. The set S_y corresponds to some subset of the vertices of Γ. Therefore we can identify S_y with an $|S_y|$-dimensional subspace of $\mathbb{C}^V$. Let $\Pi_y : \mathbb{C}^V \to \mathbb{C}^V$ be the projection matrix onto this subspace. Let $\Pi = \sum_{y\in\mathbb{Z}_{d_k}} \omega_{d_k}^y \Pi_y$. We write $\Pi = \Pi(\chi)$ to indicate the dependence on choice of χ.

We can now state the bias amplification theorem for ordinary expander walks.

▶ **Theorem 5** (Ordinary t-step expander walk). *Let $S_{init} \subset G$ have bias ϵ_0 and let $\Gamma = (S_{init}, E)$ be a d-regular expander graph with $\lambda(\Gamma) = \lambda < 1$. Suppose $D \sim G$ is the distribution induced by beginning at a uniform vertex and taking a t-step random walk $(x^{(0)}, ..., x^{(t)})$ and then adding the results of the walk to get an element $(\sum_i x^{(i)}) \in G$.*

Let $\chi^* : G \to \mathbb{C}^*$ be the nontrivial character which maximizes the bias of D. Let $\Pi = \Pi(\chi^*)$, and $\|\cdot\|$ be the matrix operator norm. Finally, abusing notation, let Γ be the random walk matrix of Γ. Then,

$$bias(D) = bias(\chi^*) \leq \|(\Pi\Gamma)^t\Pi\|$$

Proof. Let $u = \frac{1}{\sqrt{|V(\Gamma)|}}\vec{1}$ be the normalized all-ones vector. Let $a^* \in G$ be the element corresponding to χ^*. Let $(a_1^*, ..., a_k^*) \in \mathbb{Z}_{d_1} \oplus \cdots \oplus \mathbb{Z}_{d_k}$ denote a^* written in the invariant factor decomposition.

Let $W \sim V^{t+1}$ denote the distribution of all t-step walks on Γ. Let $(x^{(0)}, ..., x^{(t)}) \sim W$ be some sequence of random walk steps. So $x^{(0)} \sim S_{init}$ (since the walk begins at a uniformly random vertex) $x^{(i+1)}$ is a uniformly random neighbor of $x^{(i)}$. If $\vec{v}^{(i)} \in \mathbb{C}^V$ is the distribution at step i, then $\vec{v}^{(i+1)} = \Gamma\vec{v}^{(i)}$.

Recall that we use subscripts to denote invariant factors, so $x = (x_1, ..., x_k) \in \bigoplus_{i=1}^{k} \mathbb{Z}_{d_i}$.

$$\begin{aligned}
Bias(D) &= Bias_D(\chi^*) \\
&= \left| \mathop{\mathbb{E}}_{(x^{(0)},...,x^{(t)})\sim W} \prod_{i=1}^{k} \omega_{d_i}^{x_i \cdot a_i^*} \right| \\
&= \left| \mathop{\mathbb{E}}_{(x^{(0)},...,x^{(t)})\sim W} \omega_{d_k}^{\sum_{i=1}^{k} \frac{d_k}{d_i} x_i \cdot a_i^*} \right| \\
&= \left| \sum_{y \in \mathbb{Z}_{d_k}} \omega_{d_k}^{y} \mathop{\mathbb{P}}_{(x^{(0)},...,x^{(t)})\sim W} [y = (\sum_{j=0}^{t} \sum_{i=1}^{k} \frac{d_k}{d_i} x_i^{(j)} \cdot a_i^*) \mod d_k] \right| \\
&= \left| \sum_{y \in \mathbb{Z}_{d_k}} \sum_{b \in T_y} \omega_{d_k}^{y} \mathop{\mathbb{P}}_{(x^{(0)},...,x^{(t)})\sim W} [\bigwedge_{j=0}^{t} (x^{(j)} \in S_{b_j})] \right| \\
&= \left| \sum_{y \in \mathbb{Z}_{d_k}} \omega_{d_k}^{y} (u^T \sum_{b \in T_y} \Pi_{b_t}\Gamma \cdots \Pi_{b_1}\Gamma\Pi_{b_0} u) \right| \\
&= \left| u^T (\sum_{b \in \mathbb{Z}_{d_k}^{t+1}} \omega_{d_k}^{\sum_j b_j} \Pi_{b_t}\Gamma \cdots \Pi_{b_1}\Gamma\Pi_{b_0}) u \right| \\
&= \left| u^T (\sum_{b_t \in \mathbb{Z}_{d_k}} \omega_{d_k}^{b_t} \Pi_{b_t})\Gamma \cdots (\sum_{b_1 \in \mathbb{Z}_{d_k}} \omega_{d_k}^{b_1} \Pi_{b_1})\Gamma(\sum_{b_0 \in \mathbb{Z}_{d_k}} \omega_{d_k}^{b_0} \Pi_{b_0}) u \right| \\
&= \left| u^T (\Pi\Gamma)^t \Pi u \right| \\
&\leq \|(\Pi\Gamma)^t\Pi\|
\end{aligned}$$

◀

We have thus obtained an algebraic expression for the bias of the walk distribution, which we will now upper-bound. We defer the proof to the full version.

▶ **Theorem 6** (Matrix norm bounds). *Let* Π, Γ *be as before.*
(i) $\|\Pi\| = 1$.
(ii) $\|(\Pi\Gamma)^2\| \leq \epsilon_0 + 2\lambda$
It follows that $\|(\Pi\Gamma)^t\Pi\| \leq (\epsilon_0 + 2\lambda)^{\lfloor t/2 \rfloor}$.

Combining the two results in this section, it follows that a t-step walk amplifies the bias to $(\epsilon_0 + 2\lambda)^{\lfloor t/2 \rfloor}$.

2.2 The wide replacement walk

In this section and the subsequent one, we will show how the wide replacement walk amplifies bias more efficiently than an ordinary expander walk. We will proceed in a similar manner to the last section, by first obtaining an algebraic expression for the bias of the random walk distribution, and then upper-bounding the algebraic expression in section 2.3.

2.2.1 Setup

Let $\Gamma = (S_{init}, E)$ be a graph whose vertices are some constant-bias set $S_{init} \subset G$ as before. Suppose Γ is D_1-regular. Let $\phi_\Gamma : [D_1] \to [D_1]$ be the local inversion function of Γ.

Let $s > 0$ be an integer, and let H be a D_2-regular expander graph on $[D_1]^s$ vertices. We will abuse notation and use Γ, H to denote the random walk matrices of Γ, H respectively.

Let $V^1 = \mathbb{C}^{S_{init}} = \mathbb{C}^{V(\Gamma)}$ and $V^2 = \mathbb{C}^{D_1^s} = \mathbb{C}^{V(H)}$. We define three operators on $V^1 \otimes V^2$ that we need to describe the bias of the wide replacement walk. Let $v^1 \otimes v^2 \in V^1 \otimes V^2$.

For $i \in [s]$ define the projection matrix $P_i : V^2 \to \mathbb{C}^{D_1}$ as follows. Notice $V^2 = \mathbb{C}^{V(H)} \cong \mathbb{C}^{D_1^s}$. Identifying $V(H)$ with $\mathbb{Z}_{D_1}^s$, let $Z_i \subset V(H)$ correspond to $\{(0, ..., 0, a_i, 0, ..., 0) \in \mathbb{Z}_{D_1}^s : a_i \in \mathbb{Z}_{D_1}\}$. So we can identify $Z_i \subset V(H)$ with a D_1-dimensional subspace of $\mathbb{C}^{V(H)}$. Then let $P_i : V^2 \to \mathbb{C}^{D_1}$ be the projection onto this subspace.

Given some $v^1 \in V^1$ and $j \in [D_1]$, the vector $v^1[j] \in V^1$ is a permutation of the coordinates of v^1 based on the mapping of each vertex to its j^{th} neighbor in Γ [3]. This corresponds to taking a step in Γ, by moving along the edge numbered j incident to the current vertex. For $w \in \mathbb{C}^{D_1}$, let $v^1[w] = \sum_{j=1}^{D_1} w_j \cdot v_1[j]$.

Finally, given the local inversion function $\phi_\Gamma : [D_1] \to [D_1]$ of Γ and $i \in [s]$, define $\psi_\Gamma^{(i)} : [D_1]^s \to [D_1]^s$ as the function which applies ϕ_Γ to the i^{th} coordinate and leaves other coordinates unchanged. Since ϕ_Γ is a permutation on $[D_1]$, $\psi_\Gamma^{(i)}$ is a permutation on $[D_1]^s$. Abusing notation, let $\psi_\Gamma^{(i)} : \mathbb{C}^{D_1^s} \to \mathbb{C}^{D_1^s}$ denote the permutation matrix which permutes coordinates according to $\psi_\Gamma^{(i)}$.

We are ready to define the three operators which describe the bias of the wide replacement walk.

$$\dot{H}(v^1 \otimes v^2) = v^1 \otimes H(v^2)$$
$$\forall \chi \in \hat{G}, y \in \mathbb{Z}_d : \dot{\Pi}_y(\chi)(v^1 \otimes v^2) = \Pi_y(\chi)(v^1) \otimes v^2$$

[3] This is well-defined as long as the graph Γ is d-regular, since its adjacency matrix is then just a sum of d permutation matrices.

$$\forall \ell \in \{0, 1, ..., s-1\} : \dot{\Gamma}_\ell(v^1 \otimes v^2) = v^1[P_\ell(v^2)] \otimes \psi_\Gamma^{(\ell)}(v^2)$$

Note that each of these operators is a tensor product of operators on V^1, V^2, and hence preserves tensor products.

Moreover, notice $\dot{H}, \dot{\Gamma}_{t \mod s}$ are precisely the transition matrices of the H-step and Γ-step in the wide replacement walk at time t.

For a character $\chi : G \to \mathbb{C}^*$ let $\dot{\Pi}(\chi) = \sum_{y \in \mathbb{Z}_{d_k}} \omega_{d_k}^y \dot{\Pi}_y(\chi)$. $\dot{\Pi}$ plays the role of Π from the analysis of the ordinary expander walk.

For notational convenience,

$$\dot{L}_j(\chi) := \dot{\Pi}(\chi)\dot{\Gamma}_j\dot{H}$$

2.2.2 Algebraic Expression for the Bias

In this section we will express the bias of the wide replacement walk distribution in terms of the matrix norms of $\dot{L}_0, ..., \dot{L}_{s-1}$.

► **Proposition 7** (t-step s-wide replacement product walk)**.** *Let G be a finite abelian group. Let $S_{init} \subset G$ have bias ϵ_0 and let $\Gamma = (S_{init}, E)$ be a D_1-regular expander graph. Let H be a D_2 regular expander on $[D_1]^s$ vertices for some integer $s \geq 1$.*

Let $D_{walk} \sim G$ be the t-step s-wide replacement product walk distribution. It is defined by beginning at a uniform vertex and performing an t-step wide replacement wide on $V(\Gamma) \times V(H)$. Given a sequence of vertices $((a_0, b_0), ..., (a_t, b_t)) \in V(\Gamma) \times V(H)$ obtained from a walk, we output $(\sum_i a_i) \in G$. Then $D_{walk} \sim G$ is the distribution induced by taking all such t-step walks.

We claim that if $\chi^* : G \to \mathbb{C}^*$ is the nontrivial character which maximizes the bias of D_{walk}, and $\dot{\Pi} = \dot{\Pi}(\chi^*)$, then using the notation from above,

$$bias(D_{walk}) = bias(D_{walk}, \chi^*) \leq \|\dot{L}_{s-1}(\chi^*) \cdots \dot{L}_0(\chi^*)\|^{\lfloor t/s \rfloor}$$

The proof is similar to that of Theorem 5. See the full version.

It remains to be shown that this matrix norm is indeed bounded. To show that the wide replacement walk gains from $s - O(1)$ out of every s steps, we need to show that $\|\dot{L}_{s-1} \cdots \dot{L}_0\| \leq \lambda(H)^{s-O(1)}$.

2.3 Bounding the matrix norm

In the previous section we showed that the bound the bias of the wide-replacement walk distribution, it suffices to bound the operator norm of the following matrix, defined with respect to the worst-case character χ^* of the walk distribution:

$$\dot{L}_{s-1} \cdots \dot{L}_0$$

This is almost exactly the same matrix as the one analyzed in [41]. The difference is that the operator $\dot{\Pi}$, instead of tracking how often the walk enters the sets in a bipartition of S_{init}, now tracks how often the walk enters the sets in a d_k-way partition of S_{init}. Here $d_k = \Omega(\log(|G|))$ is the largest invariant factor of G.

As a consequence, the diagonal entries of $\dot{\Pi}$ now come from the ${d_k}^{th}$ roots of unity, rather than $\{\pm 1\}$. The analysis of the matrix bound from [41] mostly carries through, although working over $\mathbb{C}^{V_1} \otimes \mathbb{C}^{V^2}$ rather than the reals will require some care.

As in [41], our argument will proceed by considering arbitrary vectors v, w and analyazing $\langle v, \dot{L}_{s_1} \cdots \dot{L}_0 w\rangle$. We will repeatedly decompose the vectors into their parallel and perpendicular components. Let $V^{\parallel} = V^1 \otimes \vec{1}$ denote vectors whose H-component is a scalar multiple of $\vec{1}$ ("parallel vectors"), and $V^{\perp} = (V^{\parallel})^{\perp}$ ("perpendicular vectors").

Because of the spectral expansion of H, every time a vector is in $V^{\perp}$ we can show it shrinks by a factor of $\lambda(H)$. The hard case is when vectors are in $V^{\parallel}$. Here, we will prove a technical lemma which is a straightforward generalization of the core lemma in [41]. The lemma shows if the walk distribution is in $V^{\parallel}$, then any *sequence* of s steps imitates a random walk of s steps on the outer graph Γ. This allows us to argue that the bias is amplified as though taking the ordinary random walk on Γ. If the bias so far is α, then this scales the bias by $\alpha \mapsto (\alpha + 2\lambda(\Gamma))^{s/2}$ after s steps.

This turns out to be enough. Let $\epsilon_0 = bias(S_{init})$ be the bias of the initial set $S_{init} \subset G$. Since ϵ_0 is a constant, we can select graphs Γ, H such that $\epsilon_0 + 2\lambda(\Gamma) \leq \lambda(H)^2$. Therefore, while we do not gain a factor of $(\lambda(\Gamma))^s$ every s steps, we will gain according to a factor of $(\lambda(H))^{s-O(1)}$.

Therefore, whether in the $V^{\perp}$ or $V^{\parallel}$ case, we shrink the bias by a factor of $\lambda(H)^{s-O(1)}$ for every s steps.

We begin by proving the technical lemma about parallel vectors. We will frequently use the following fact.

▶ **Proposition 8** (Operator-Averaging, [41] Claim 14). *Let Ω be a finite set and P, Q probability distributions on Ω. Let $\|P - Q\|_1$ denote the difference of the distributions in the 1-norm. Further, let $\{T_x\}_{x\in\Omega}$ be a family of linear operators on $\mathbb{C}^n$ indexed by Ω, such that for all $x \in \Omega$, $\|T_x\| \leq 1$. Let $A = \mathbb{E}_{x\sim P}[T_x]$ and $B = \mathbb{E}_{x\sim Q}[T_x]$. We claim that for all $v, w \in \mathbb{C}^n$ that*

$$|\langle Av, w\rangle - \langle Bv, w\rangle| \leq \|P - Q\|_1 \|v\| \|w\|$$

Next, we need to formalize the notion of the wide replacement walk "imitating" the ordinary random walk on the outer graph, which we do via the notion of a pseudorandom inner graph.

▶ **Definition 9.** *(Pseudorandom inner graph) Let Γ be a D_1-regular graph with local inversion function $\phi_\Gamma : [D_1] \to [D_1]$. Let H be a D_2-regular graph on D_1^s vertices. Let $\zeta \geq 0$. We say H is ζ-pseudorandom with respect to Γ if for all s-step sequences in the s-wide replacement walk, the corresponding V^1-instructions are ζ-close to $Unif([D_1]^s)$ in ℓ_1-norm.*

Formally, let the adjacency matrix of H be $H = \frac{1}{D_2}\sum_{i=1}^{D_2} \Xi_i$, where each Ξ_i is a permutation matrix [4]*. Let $\xi_i : V(H) \to V(H)$ be the permutation map corresponding to Ξ_i. For $0 \leq k < s$, let $\psi_k : [D_1]^s \to [D_1]^s$ be $\psi_k(a_0, ..., a_{s-1}) = (a_0, ..., a_{k-1}, \phi_\Gamma(a_k), a_{k+1}, ..., a_{s-1})$.*

Fix $(j_0, ..., j_{s-1}) \in [D_2]^s$. For some $(u^1, u^2) \in V(\Gamma) \times V(H)$ let $\sigma_{j_0}(u^2) = \gamma_{j_0}(u^2)$. For $\ell > 0$, let

$$\sigma_{j_\ell,...,j_0}(u^2) = \gamma_{j_\ell}(\psi_{\ell-1}(\sigma_{j_{\ell-1},...,j_0}(u^2)))$$

We say $(j_0, ..., j_{s-1}) \in [D_2]^s$ is ζ-pseudorandom with respect to Γ if

[4] By the Birkhoff-von Neumann Theorem, the adjacency matrix of a d-regular graph is a sum of d permutation matrices.

$$\|(\pi_0(\sigma_{j_0}(Unif([D_1]))), ..., \pi_{s-1}(\sigma_{j_{s-1},...,j_0}(Unif([D_1])))) - Unif([D_1]^s)\|_1 \le \zeta$$

We say the inner graph H is ζ-pseudorandom with respect to the outer graph Γ if for all $(j_0, ..., j_{s-1}) \in [D_2]^s$, $(j_0, ..., j_{s-1})$ is ζ-pseudorandom with respect to Γ.

If we unravel the definition, this is simply requiring that H is compatible with the edge labeling of Γ in precisely the way that we want. Pseudorandomness is a strong condition on H which, by definition, guarantees the wide-replacement walk imitates the ordinary walk on Γ in a suitable sense.

With this definition we can return to proving the lemma. We will begin by proving the pseudorandomness claim for the case where $D_2 = 1$; the general case where $D_2 \neq 1$ follows from another application of operator averaging, viewing the matrix H as an average of D_2 permutation matrices. We defer the proofs to the full version.

▶ **Proposition 10** (Action on parallel vectors). *Let $\ell \le s$. Suppose that the sequence $(j_0, ..., j_{\ell-1}) \in [D_2]^s$ is ζ-pseudoranom with respect to the local inversion function $\phi : [D_1] \to [D_1]$. Let $\tilde{\Xi}_{j_0}, ..., \tilde{\Xi}_{j_{\ell-1}}$ denote the operators on $V^1 \otimes V^2$ corresponding to the permutations $\xi_{j_0}, ..., \xi_{j_{\ell-1}}$ on $V(H)$. Let $1_{V(H)}$ denote the normalized all-ones vector of length $|V(H)|$.*

For any $\tau = \tau^1 \otimes 1_{V(H)}$ and $\upsilon = \upsilon^1 \otimes 1_{V(H)}$,

$$\left|\langle \dot{\Pi}\dot{\Gamma}_{\ell-1}\tilde{\Xi}_{j_{\ell-1}} \cdots \dot{\Pi}\dot{\Gamma}_0\tilde{\Xi}_{j_0}\tau, \upsilon\rangle - \langle (\pi\Gamma)^\ell \tau^1, \upsilon^1\rangle\right| \le \zeta\|\tau\|\|\upsilon\|$$

▶ **Corollary 11** (Generalized action on parallel vectors ([41] Theorem 27)). *Suppose that H is ζ-pseudorandom with respect to the local inversion function ϕ_Γ of Γ. For every $i_1, i_2 \in \{0, 1, ..., s-1\}$, and every $\tau, \upsilon \in V^\parallel$,*

$$\left|\langle \dot{L}_{i_2} \cdots \dot{L}_{i_1}\tau, \upsilon\rangle - \langle (\Pi\Gamma)^{i_2-i_1+1}\tau^1, \upsilon^1\rangle\right| \le \zeta\|\tau\|\|\upsilon\|$$

Now we are ready to prove bound the matrix norm of $\dot{L}_{s_1} \cdots \dot{L}_0$, which expresses the bias of the wide replacement walk. Our argument will proceed by considering the quadratic form $\langle v, \dot{L}_{s_1} \cdots \dot{L}_0 w\rangle$ for arbitrary v, w and then repeatedly decomposing v, w into their $V^\parallel$ and $V^\perp$ components. Because of the spectral expansion of H, every time a vector is in $V^\perp$ we can show it shrinks by a factor of $\lambda_2 = \lambda(H)$.

The hard case is when vectors are in $V^\parallel$. Here, we will use Corollary 11 to argue that any *sequence* of s steps imitates a random walk on the outer graph Γ. This allows us to argue that the bias is amplified as though taking the ordinary random walk on Γ. This scales the bias by $(\epsilon_0 + 2\lambda_1)^{s/2}$ at every s steps.

This is enough, as we can assume that $\epsilon_0 + 2\lambda_1 \le \lambda_2^2$. Therefore, while we do not gain a factor of $(\lambda_1)^s$ every s steps, we will gain according to a factor of $(\lambda_2)^s$. Since $\lambda_2 < 1$, the difference between gaining according to λ_2 or λ_1 does not matter asymptotically.

▶ **Theorem 12** (Bounding algebraic expression for bias). *Suppose that:*

i) H is ζ-pseudorandom with respect to ϕ_Γ

ii) $\epsilon_0 + 2\lambda(\Gamma) \le \lambda(H)^2$

Then we obtain the following bound for the bias of the walk after s steps.

$$\|\dot{L}_{s-1} \cdots \dot{L}_0\| \le \lambda(H)^s + s\lambda(H)^{s-1} + s^2(\lambda(H)^{s-2} + \zeta)$$

We defer the proof to the full version.

References

1 Miklós Ajtai, Henryk Iwaniec, János Komlós, János Pintz, and Endre Szemerédi. Construction of a thin set with small Fourier coefficients. *Bull. London Math. Soc.*, 22(6):583–590, 1990. doi:10.1112/blms/22.6.583.

2 Josh Alman and Virginia Vassilevska Williams. A refined laser method and faster matrix multiplication. In *Proceedings of the 2021 ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 522–539. SIAM, 2021.

3 Noga Alon. Eigenvalues and expanders. *Combinatorica*, 6(2):83–96, 1986.

4 Noga Alon. Explicit expanders of every degree and size. *Combinatorica*, pages 1–17, 2021.

5 Noga Alon and Gil Cohen. On rigid matrices and u-polynomials. In *2013 IEEE Conference on Computational Complexity*, pages 197–206. IEEE, 2013.

6 Noga Alon, Oded Goldreich, Johan Håstad, and René Peralta. Simple constructions of almost k-wise independent random variables. *Random Structures & Algorithms*, 3(3):289–304, 1992.

7 Noga Alon and Yishay Mansour. ϵ-discrepancy sets and their application for interpolation of sparse polynomials. *Inform. Process. Lett.*, 54(6):337–342, 1995. doi:10.1016/0020-0190(95)00032-8.

8 Noga Alon, Rina Panigrahy, and Sergey Yekhanin. Deterministic approximation algorithms for the nearest codeword problem. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, pages 339–351. Springer, 2009.

9 Noga Alon and Yuval Roichman. Random Cayley graphs and expanders. *Random Structures Algorithms*, 5(2):271–284, 1994. doi:10.1002/rsa.3240050203.

10 Andris Ambainis and Joseph Emerson. Quantum t-designs: t-wise independence in the quantum world. In *Twenty-Second Annual IEEE Conference on Computational Complexity (CCC'07)*, pages 129–140. IEEE, 2007.

11 V. Arvind, Partha Mukhopadhyay, Prajakta Nimbhorkar, and Yadu Vasudev. Expanding generating sets for solvable permutation groups. *SIAM J. Discrete Math.*, 32(3):1721–1740, 2018. doi:10.1137/17M1148979.

12 V. Arvind and Srikanth Srinivasan. The remote point problem, small bias spaces, and expanding generator sets. In *STACS 2010: 27th International Symposium on Theoretical Aspects of Computer Science*, volume 5 of *LIPIcs. Leibniz Int. Proc. Inform.*, pages 59–70. Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2010.

13 Vikraman Arvind, Partha Mukhopadhyay, and Prajakta Nimbhorkar. Erdős-rényi sequences and deterministic construction of expanding cayley graphs. In *Latin American Symposium on Theoretical Informatics*, pages 37–48. Springer, 2012.

14 Yossi Azar, Rajeev Motwani, and Joseph Naor. Approximating probability distributions using small sample spaces. *Combinatorica*, 18(2):151–171, 1998. doi:10.1007/PL00009813.

15 Eric Bach and Jonathan Sorenson. Explicit bounds for primes in residue classes. *Mathematics of Computation*, 65(216):1717–1735, 1996.

16 Avraham Ben-Aroya and Amnon Ta-Shma. A combinatorial construction of almost-ramanujan graphs using the zig-zag product. *SIAM Journal on Computing*, 40(2):267–290, 2011.

17 Avraham Ben-Aroya and Amnon Ta-Shma. Constructing small-bias sets from algebraic-geometric codes. *Theory Comput.*, 9:253–272, 2013. doi:10.4086/toc.2013.v009a005.

18 Eli Ben-Sasson, Madhu Sudan, Salil Vadhan, and Avi Wigderson. Randomness-efficient low degree tests and short PCPs via epsilon-biased sets. In *Proceedings of the Thirty-Fifth Annual ACM Symposium on Theory of Computing*, pages 612–621. ACM, New York, 2003. doi:10.1145/780542.780631.

19 Amey Bhangale, Prahladh Harsha, Orr Paradise, and Avishay Tal. Rigid matrices from rectangular pcps or: Hard claims have complex proofs. In *2020 IEEE 61st Annual Symposium on Foundations of Computer Science (FOCS)*, pages 858–869. IEEE, 2020.

20 Emmanuel Breuillard and Alexander Lubotzky. Expansion in simple groups. *arXiv preprint arXiv:1807.03879*, 2018.

21 Moses Charikar, Konstantin Makarychev, and Yury Makarychev. Near-optimal algorithms for maximum constraint satisfaction problems. *ACM Trans. Algorithms*, 5(3):Art. 32, 14, 2009. doi:10.1145/1541885.1541893.

22 Eshan Chattopadhyay, Pooya Hatami, Kaave Hosseini, and Shachar Lovett. Pseudorandom generators from polarizing random walks. *Theory Comput.*, 15:Paper No. 10, 26, 2019. doi:10.4086/toc.2019.v015a010.

23 Sixia Chen, Cristopher Moore, and Alexander Russell. Small-bias sets for nonabelian groups: derandomizations of the Alon-Roichman theorem. In *Approximation, randomization, and combinatorial optimization*, volume 8096 of *Lecture Notes in Comput. Sci.*, pages 436–451. Springer, Heidelberg, 2013. doi:10.1007/978-3-642-40328-6_31.

24 Tobias Christiani and Rasmus Pagh. Generating k-independent variables in constant time. In *2014 IEEE 55th Annual Symposium on Foundations of Computer Science*, pages 196–205. IEEE, 2014.

25 Gil Cohen, Noam Peri, and Amnon Ta-Shma. Expander random walks: A fourier-analytic approach. In *Electron. Colloquium Comput. Complex*, volume 27, page 6, 2020.

26 Guy Even, Oded Goldreich, Michael Luby, Noam Nisan, and Boban Veličković. Efficient approximation of product distributions. *Random Structures Algorithms*, 13(1):1–16, 1998. doi:10.1002/(SICI)1098-2418(199808)13:1<1::AID-RSA1>3.0.CO;2-W.

27 Rusins Freivalds. Probabilistic machines can use less running time. In *IFIP congress*, volume 839, page 842, 1977.

28 Shlomo Hoory, Nathan Linial, and Avi Wigderson. Expander graphs and their applications. *Bulletin of the American Mathematical Society*, 43(4):439–561, 2006.

29 Akhil Jalan and Dana Moshkovitz. Near-optimal cayley expanders for abelian groups. *arXiv preprint arXiv:2105.01149*, 2021.

30 Fernando Granha Jeronimo, Dylan Quintana, Shashank Srivastava, and Madhur Tulsiani. Unique decoding of explicit epsilon-balanced codes near the gilbert-varshamov bound. In *2020 IEEE 61st Annual Symposium on Foundations of Computer Science (FOCS)*, pages 434–445. IEEE, 2020.

31 Jørn Justesen. Class of constructive asymptotically good algebraic codes. *IEEE Transactions on Information Theory*, 18(5):652–656, 1972.

32 Nicholas M. Katz. An estimate for character sums. *J. Amer. Math. Soc.*, 2(2):197–200, 1989. doi:10.2307/1990974.

33 Ivan Korec and Jiří Wiedermann. Deterministic verification of integer matrix multiplication in quadratic time. In *International Conference on Current Trends in Theory and Practice of Informatics*, pages 375–382. Springer, 2014.

34 Alexander Lubotzky, Ralph Phillips, and Peter Sarnak. Ramanujan graphs. *Combinatorica*, 8(3):261–277, 1988.

35 Joseph Naor and Moni Naor. Small-bias probability spaces: efficient constructions and applications. *SIAM J. Comput.*, 22(4):838–856, 1993. doi:10.1137/0222053.

36 A. Razborov, E. Szemerédi, and A. Wigderson. Constructing small sets that are uniform in arithmetic progressions. *Combin. Probab. Comput.*, 2(4):513–518, 1993. doi:10.1017/S0963548300000870.

37 Omer Reingold, Salil Vadhan, and Avi Wigderson. Entropy waves, the zig-zag graph product, and new constant-degree expanders and extractors. In *Proceedings 41st Annual Symposium on Foundations of Computer Science*, pages 3–13. IEEE, 2000.

38 Ronitt Rubinfeld and Madhu Sudan. Robust characterizations of polynomials with applications to program testing. *SIAM Journal on Computing*, 25(2):252–271, 1996.

39 Amir Shpilka and Avi Wigderson. Derandomizing homomorphism testing in general groups. *SIAM Journal on Computing*, 36(4):1215–1230, 2006.

40 Amnon Ta-Shma. Explicit, almost optimal, epsilon-balanced codes. In *STOC'17—Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*, pages 238–251. ACM, New York, 2017. doi:10.1145/3055399.3055408.

41 Amnon Ta-Shma. Explicit, almost optimal, epsilon-balanced codes. In *TR 17-041*. Electronic Colloqium on Computational Complexity, 2017.
42 Salil Vadhan. *Pseudorandomness*, volume 7. Now Delft, 2012.
43 Leslie G Valiant. Graph-theoretic arguments in low-level complexity. In *International Symposium on Mathematical Foundations of Computer Science*, pages 162–176. Springer, 1977.
44 Avi Wigderson and David Xiao. Derandomizing the ahlswede-winter matrix-valued chernoff bound using pessimistic estimators, and applications. *Theory of Computing*, 4(1):53–76, 2008.
45 Triantafyllos Xylouris. *Über die Nullstellen der Dirichletschen L-Funktionen und die kleinste Primzahl in einer arithmetischen Progression*, volume 404 of *Bonner Mathematische Schriften [Bonn Mathematical Publications]*. Universität Bonn, Mathematisches Institut, Bonn, 2011. Dissertation for the degree of Doctor of Mathematics and Natural Sciences at the University of Bonn, Bonn, 2011.

A Cayley Graphs and Expanders

We begin with some preliminaries on graphs and group theory.

▶ **Definition 13** (Spectral expander graph)**.** *Let $G = ([n], E, w)$ be a weighted, d-regular undirected graph. By d-regular we mean that for all $u \in V$, $\sum_{v \in V} w(\{u, v\}) = d$.*

Let $A \in \mathbb{C}^{n \times n}$ be the (weighted) adjacency operator of G, and let $M = \frac{1}{d}A$ be the normalized adjacency operator, also known as the random walk matrix. Let the eigenvalues of M be denoted $\lambda_n \leq ... \leq \lambda_2 \leq \lambda_1 = 1$, counting multiplicity. Then G is a one-sided spectral expander if $\lambda_2 < 1-\Omega(1)$, and G is a two-sided spectral expander if $\max\{|\lambda_n|, |\lambda_2|\} < 1-\Omega(1)$.

Let $\lambda(G) := \max\{|\lambda_n|, |\lambda_2|\}$. The two-sided spectral gap of G is $1 - \lambda(G)$.

Next, we define Cayley graphs.

▶ **Definition 14.** *(Symmetric generating set) Let G be a group and $S \subset G$. We say that S is symmetric if for all $s \in S$, $s^{-1} \in S$. Further, S is a generating set if for all $g \in G$ there exist $s_1, ..., s_k \in S$ (possibly repeated) such that $s_k \cdots s_1 = g$.*

We write $\langle S \rangle = G$.

▶ **Definition 15.** *(Cayley Graph) Let G be a group and $S \subset G$ be a symmetric generatring set, and $w : S \to \mathbb{R}_{\geq 0}$ a weight function. The Cayley graph $Cay(G, S, w)$ is the graph with vertex set G and edge set $\{\{g, g \cdot s\} : g \in G, s \in S\}$. The weight of an edge $\{g, g \cdot s\}$ is $w(s)$.*

We will require the total weight of S to be normalized to $|S|$ by convention. Notice that since S is symmetric, we can consider the graph $\mathrm{Cay}(G, S)$ to be an undirected and weighted $|S|$-regular multigraph.

The eigenvectors of abelian Cayley graphs are described by their group characters.

▶ **Definition 16** (Characters of abelian group)**.** *Let $\mathbb{C}^*$ be the multiplicative group of nonzero complex numbers. For any finite abelian group G, the characters of G, denoted $\hat{G}$, are the set of all homomorphisms $\chi : G \to \mathbb{C}^*$.*

▶ **Proposition 17.** *Let G be a finite abelian group and $S \subset G$ a symmetric generating set. Then the eigenvalues of $Cay(G, S)$ are given by*

$$\{|\mathop{\mathbb{E}}_{x \sim S}[\chi(x)]| : \chi \in \hat{G}\}$$

Notice that any group has a *trivial character* $\chi : G \to \mathbb{C}^*$ such that $\chi(g) = 1$ for all g. The eigenvalue corresponding to the trivial character is always 1. Therefore, for a Cayley graph to be an expander we need bounds on all of its nontrivial characters.

▶ **Definition 18** (Small-bias distributions for abelian groups)**.** *Let G be a finite abelian group and $D \sim G$ a random variable. For any character χ of G, the bias of D with respect to χ is*

$$Bias_{\chi}(D) := |\mathop{\mathbb{E}}_{x\sim D}[\chi(x)]|$$

Let χ_0 denote the trivial character. The bias of D is its maximum bias with respect to nontrivial characters.

$$Bias(D) := \max_{\chi\neq\chi_0} Bias_{\chi}(D)$$

If $S \subset G$, then $bias(S)$ is the bias of the uniform distribution on S. If S is a symmetric generating set, $\lambda(Cay(G,S)) = Bias(S)$.

Notice that if S is non-negatively weighted, we can normalize weights to sum to 1 and obtain a (not necessarily uniform) distribution on S. Then the bias of S is just the bias of this distribution.

Finally, we will need a few more facts about characters of abelian groups.

▶ **Proposition 19.** *(Characters of cyclic groups) Let $\mathbb{Z}_d$ be the cyclic group on $d \geq 2$ elements. Let $\omega_d := exp(\frac{2\pi i}{d})$. The characters of $\mathbb{Z}_d$ are the maps $\chi_j(x) = \omega_d^{j\cdot x}$ for $j = 0, 1, ..., d-1$.*

▶ **Definition 20.** *(Direct sum of groups) Let A, B be abelian groups. The direct sum $A \oplus B$ is the abelian group whose elements belong to the Cartesian product $A \times B$. For $(a_1, b_1), (a_2, b_2) \in A \times B$, the group operation is $(a_1, b_1) + (a_2, b_2) = (a_1 + a_2, b_1 + b_2)$.*

Notice that the direct sum is associative.

▶ **Proposition 21.** *(Fundamental theorem of finite abelian groups) Let G be a finite abelian group. Then G is isomorphic to a direct sum of cyclic groups. That is, there exist $d_1, ..., d_k \geq 2$ such that*

$$G \cong \mathbb{Z}_{d_1} \oplus \cdots \oplus \mathbb{Z}_{d_k}$$

Moreover, $d_i | d_j$ for all $i < j$.

We refer to $\mathbb{Z}_{d_1} \oplus \cdots \oplus \mathbb{Z}_{d_k}$ as the invariant factor decomposition of G. The integers $d_1, ..., d_k$ are the invariant factors.

From the above propositions one can show that the characters of a finite abelian group are products of maps of the form $x \mapsto \omega_{d_i}^{j\cdot x}$. This structure is crucial to our overall argument.

B Wide Replacement Walks

In this section we define what it means to take a wide replacement walk.

Let G be a D_1-regular graph on N_1 vertices and H be a D_2-regular graph on D_1 vertices. The *replacement product* $G \textcircled{r} H$ is a $(D_2 + 1)$-regular graph on $N_1 \cdot D_1$ vertices. Each vertex of G (the "outer graph") is replaced by a copy of H (the "inner graph"). We call these copies *clouds.*

The intra-cloud edges in each cloud of $G \textcircled{r} H$ are just the edges from H. However, $G \textcircled{r} H$ also has *inter*-cloud edges which arise by identifying the D_1 vertices of H with the D_1 incident edges of a vertex $v \in V(G)$. This identification requires that we number the edges of every vertex in G. We formalize this with the concept of a rotation map.

▶ **Definition 22.** *(Rotation map) Let G be a D-reguluar graph such that the edges incident to every $v \in V(G)$ are numbered $1, ..., D$. Formally there is a function $N : V \times [D] \to V$ such that $N(v, i) = w$ iff w is the i^{th} neighbor of v.*

Then a rotation map is a function $Rot : V \times [D] \to V \times [D]$ such that for all $v, w \in V$ and $i, j \in [D]$, $Rot(v, i) = (w, j)$ iff the i^{th} neighbor of v is w and the j^{th} neighbor of w is v.

For technical reasons, we need a special kind of rotation map called a local inversion function. This is a rotation map where if (v, i) maps to (w, j) then j only depends on i.

▶ **Definition 23.** *(Local inversion function) Let G be a D-regular graph with a rotation map $Rot : V \times [D] \to V \times [D]$. A local inversion function $\phi_G : [D] \to [D]$ is a permutation on $[D]$ such that for all $v \in V, i \in [D]$,*

$$Rot(v, i) = (N(v, i), \phi_G(i))$$

We are ready to define the wide replacement product walk. Instead of the usual inner graph H we use a "wide" inner graph on D_1^s vertices for some integer $s \geq 1$. The vertices of H correspond to s-tuples that define s local inversion functions. The walk cycles through them.

To take a step in the usual replacement product walk, we start at some vertex $v \in G \textcircled{r} H$ then compose two steps: an intra-cloud step which changes the H-component, and an inter-cloud step which changes the G-component. Every vertex in $G \textcircled{r} H$ is incident to a unique inter-cloud edge; therefore, there is only one choice of neighboring cloud, and so the position after the intra-cloud step determines the entire step.

The s-wide replacement walk modifies the inter-cloud step so that there are s choices during inter-cloud step. If G is D_1-regular, then a vertex of H corresponds to some vector $(a_0, ..., a_{s-1}) \in [D_1]^s$. The wide replacement walk maintains a clock which tracks how many steps have been taken. At time step t, the clock is set to $\ell = t \mod s$, and the inter-cloud step moves to a neighboring cloud according to the value of $a_\ell \in [D_1]$.

After deciding which neighboring cloud to move to, the choice of which vertex in the cloud to land in is also determined by a_ℓ. The walk updates the H-component by feeding the ℓ^{th} coordinate to the local inversion function $\phi_G : [D_1] \to [D_1]$ of G, and leaving all other coordinates unchanged. So $(a_0, ..., a_{s-1}) \in [D_1]^s$ is mapped to $(a_0, ..., a_{\ell-1}, \phi_G(a_\ell), a_{\ell+1}, ..., a_{s-1})$. This completes the inter-cloud step.

The utility of the wide replacement walk is that the H-component of a vertex now stores $O(s \log(D_1))$ bits of information, rather than just $O(\log(D_1))$ bits. As we discussed in the introduction, the barrier to bias amplification is when the walk distribution is uniform within clouds.

Now, the values of the H-component are precisely the instructions for the inter-cloud steps of the walk; therefore, the fact that the H-component is uniform is no longer bad news, since it means that the inter-cloud steps of the replacement walk imitate the truly random walk on the outer graph for the next s steps.

▶ **Definition 24.** *Let G be a D_1-regular graph with local inversion function $\phi_G : [D_1] \to [D_1]$. Let H be a D_2-regular graph on D_1^s vertices, for integer $s \geq 1$. A random step in the wide replacement product is determined as follows.*

Let $(v^{(1)}, v^{(2)}) \in V(G) \times V(H)$ be the current state of the walk at time $t \in \mathbb{N}$. Sample random $i \in [D_2]$. Then the time-t step according to i, denoted $Step_{i,t}(v^{(1)}, v^{(2)})$ is given by the composition of two steps:

(i) Intra-cloud step: Leave the G-component $v^{(1)}$ unchaged. Move the $v^{(2)}$ component to its i^{th} neighbor in H. Formally, set

$$w^{(1)} = v^{(1)}$$
$$w^{(2)} = v^{(2)}[i]$$

(ii) Inter-cloud step: Identifying $V(H)$ with $[D_1]^s$, let $\pi_j : [D_1]^s \to [D_1]$ be projection onto the j^{th} coordinate. Write $w^{(2)} \in V(H)$ as $w^{(2)} = (\pi_0(w^{(2)}), ..., \pi_{s-1}(w^{(2)})) \in [D_1]^s$.

Let $\ell = t \mod s$. Move to the neighbor of $w^{(1)}$ in G that is numbered by $\pi_\ell(w^{(2)}) \in D_1$. Then, update the ℓ^{th} coordinate of H-component $w^{(2)}$ by the local inversion function $\phi_G : [D_1] \to [D_1]$ and leave other coordinates unchaged. Formally, let $\psi_\ell : [D_1]^s \to [D_1]^s$ be

$$\psi_\ell(a_0, ..., a_{s-1}) = (a_0, ..., a_{\ell-1}, \phi_G(a_\ell), a_{\ell+1}, ..., a_{s-1})$$

Set

$$Step_{i,t}(v^{(1)}, v^{(2)}) = (w^{(1)}[\pi_\ell(w^{(2)})], \psi_\ell(w^{(2)}))$$

A few remarks are in order. First, notice that the number of random bits needed to specify a random step is only $O(\log(D_2))$, despite the fact that we are moving on a graph with $V(G) \times V(H)$ vertices. This will be crucial in the analysis of the tradeoff between bias amplification and size increase of the small-bias set.

Second, once a value of t is fixed, so the clock is set to $\ell = t \mod s$, the wide replacement walk can be regarded as taking a usual step in the usual replacement walk. The intra-cloud step is unchaged, and the inter-cloud step depends only on the ℓ^{th} coordinate of the H-component.

Since we have specified what it means to take a random step, this is sufficient to describe the walk. We simply initialize at a uniform vertex of $V(G) \times V(H)$ and then take some number of steps, to be chosen later.

C Parameters of the Construction

In this section we describe how to optimize parameters such that the wide replacement walk construction achieves our desired support size. Our construction and hence the parameters we choose are almost identical to those discussed in Section 5 of [41].

The algorithm is given integer $n \geq 1$, desired second eigenvalue $\epsilon > 0$, and an arbitrary generating set for a group G.

It first generates an ϵ_0-biased set $S_{init} \subset G^n$ of size $O(\frac{n \log(|G|)^{O(1)}}{poly(\epsilon_0)})$ for a constant ϵ_0. For concreteness we set $\epsilon_0 = 0.1$.

▶ **Proposition 25.** *There exists a deterministic, polynomial time algorithm which, given a generating set for an abelian group G and integer $n \geq 1$, outputs a generating set $S_{init} \subset G^n$ of size $O(n(\log(|G|))^{O(1)})$ such that the Cayley graph has second eigenvalue at most 0.1.*

Proof. First, by Theorem 4 of [23], we can construct a generating set $S \subset G$ with second eigenvalue $(1 - \frac{C}{\log\log(|G|)} + \beta)$ for a parameter β and universal constant C. Its size will be $|S| = O(\frac{n \log(|G|)}{\beta^{O(1)}}) = O(n \log(|G|)^2)$. Setting $\beta = \frac{C}{2 \log\log(|G|)}$, we obtain second eigenvalue $(1 - \frac{C}{2\log\log(|G|)})$.

Next, we can amplify the bias of S to 0.1 by taking a t-step ordinary expander walk. By the results of section 3.1, if we take a walk on a D-regular expander graph with second eigenvalue

λ and $D = O(1)$, then the t-step walk will amplify the bias to $((1 - \frac{C}{2\log\log(|G|)}) + 2\lambda)^{\lfloor t/2 \rfloor}$. For this quantity to be at most 0.1, it suffices to set $t > \frac{\log\log(|G|)}{C}(1+2\lambda) = \Theta(\log\log(|G|))$.

Therefore, after t steps we obtain a generating set $S_0 \subset G^n$ with bias 0.1, whose size is $|S_0| \cdot D^t = O(\frac{n\log(|G|)^2}{(0.1)^{O(1)}} \cdot 2^{\Theta(\log\log(|G|))}) = O(n(\log(|G|))^{O(1)})$. ◀

Next, the algorithm performs a wide replacement walk. We must specify the inner and outer graphs as well as the number of steps. Our parameters are almost identical to [41].

Let $\alpha = \Theta((\frac{\log\log(\frac{1}{\epsilon})}{\log(\frac{1}{\epsilon})})^{1/3})$. We will show that the wide replacement walk amplifies bias to ϵ and produces a generating set of size $O(\frac{n\log(|G|)^{O(1)}}{\epsilon^{2+O(\alpha)}}) = O(\frac{n\log(|G|)^{O(1)}}{\epsilon^{2+o(1)}})$.

Let the "width" $s = \frac{1}{\alpha}$.

Inner Graph: Let D_2 be the least power of two such that $D_2 \geq s^{4s}$. Let $b_2 = 4s\sqrt{2}\log(D_2)$. Let $D_1 = D_2^4$. Let $m = \log(D_1)$.

Let $H = Cay(\mathbb{Z}_2^{ms}, A)$ for a generating set of size $|A| = D_2$ (found, e.g via [41]) such that the second eigenvalue is $\lambda(H) = \frac{b_2}{\sqrt{D_2}}$.

Outer graph: Let $D_1 = D_2^4$. Find a D_1-regular expander graph Γ with $\lambda(\Gamma) = \Theta(\frac{1}{\sqrt{D_1}})$ (using, e.g. [4]). Identify its vertices with the ϵ_0-biased set S_{init}.

Walk length: Finally, set t to be the least integer such that $\lambda(H)^{(1-4\alpha)(1-\alpha)t} \leq \epsilon$ and $t \geq \frac{s}{\alpha}$.

▶ **Proposition 26.** *The t-step wide replacement walk distribution is ϵ-biased.*

Proof. The bias after t steps is given by $(\lambda(H)^s + s\lambda(H)^{s-1} + s^2\lambda(H)^{s-2})^{\lfloor t/s \rfloor}$. Therefore,

$$
\begin{aligned}
(\lambda(H)^s + s\lambda(H)^{s-1} + s^2\lambda(H)^{s-2})^{\lfloor t/s \rfloor} &\leq (2s^2\lambda(H)^{s-3})^{\lfloor t/s \rfloor} \\
&\leq (2s^2\lambda(H)^{s-3})^{t/s-1} \\
&\leq (\lambda(H)^{s-4})^{t/s-1} \\
&= \lambda(H)^{\frac{s-4}{s}(t-s)} \\
&= \lambda(H)^{(1-\frac{4}{s})(1-\frac{s}{t})t} \\
&\leq \lambda(H)^{(1-4\alpha)(1-\alpha)t} \\
&\leq \epsilon
\end{aligned}
$$

The last step follows by assumption on t. ◀

▶ **Proposition 27.** *The support size of the wide replacement walk distribution is $O(|S_{init}| \cdot \frac{1}{\epsilon^{2+O(\alpha)}})$, where S_{init} is the initial constant-bias set.*

Proof. Recall that we identify our initial 0.1-biased distribution with the vertices of the outer graph Γ. Therefore $N_1 = |V(\Gamma)| = O(\frac{n\log(|G|)^{O(1)}}{\epsilon_0^c})$ for constant $\epsilon_0, c > 0$. Since ϵ_0 is constant we can assume $D_2 \geq \epsilon_0^{-1}$. The walk begins at a uniform vertex of the replacement product, so the initial support size is N_1N_2. After t steps it increases by a factor of D_2^t. Therefore

$$
\begin{aligned}
N_1N_2D_2^t &= O(\frac{n\log(|G|)^{O(1)}}{\epsilon_0^c}N_2D_2^t) \\
&= O(\frac{n\log(|G|)^{O(1)}}{\epsilon_0^c}D_2^{4s}D_2^t)
\end{aligned}
$$

$$= O(n\log(|G|)^{O(1)} \cdot D_2^{4s+t+c})$$
$$\leq O(n\log(|G|)^{O(1)} \cdot D_2^{4\alpha t+t+c})$$
$$\leq O(n\log(|G|)^{O(1)} \cdot D_2^{t(1+5\alpha)})$$

Next, notice $b_2 = 4\sqrt{2}s\log(D_2) = 4\sqrt{2}\cdot 4s^2\log(s) \leq s^4$ for sufficiently large s (equivalently, small enough ϵ). Therefore, $D_2 \geq (s^4)^s \geq b_2^s = b_2^{1/\alpha}$. Therefore $D_2^{1/2-\alpha} \leq \lambda(H)^{-1} = \frac{\sqrt{D_2}}{b_2}$.

It follows that for small enough α (equivalently, small enough ϵ), that

$$D_2^t \leq (\lambda(H)^{-1})^{\frac{t}{1/2-\alpha}} = (\lambda(H)^{-1})^{\frac{2t}{1-2\alpha}} = (\epsilon^{-1})^{\frac{1}{(1-4\alpha)(1-\alpha)t}\frac{2t}{1-2\alpha}} \leq (\epsilon^{-1})^{2(1+8\alpha)}$$

Finally, $D_2^{t(1+5\alpha)} \leq (\epsilon^{-1})^{2(1+8\alpha)(1+5\alpha)} \leq (\epsilon^{-1})^{2(1+14\alpha)}$.

Therefore, the overall size of the generating set is $O(\frac{n\log(|G|)^{O(1)}}{\epsilon^{2+O(\alpha)}})$. In particular, since $\alpha \to 0$ as $\epsilon \to 0$, the size is $O(\frac{n\log(|G|)^{O(1)}}{\epsilon^{2+o(1)}})$. ◀