Almost Ramanujan Expanders from Arbitrary Expanders via Operator Amplification

Fernando Granha Jeronimo* Tushant Mittal Sourya Roy
Avi Wigderson†

We give an efficient algorithm that transforms any bounded degree expander graph into another that achieves almost optimal (namely, near-quadratic, $d \leq 1/\lambda^{2+o(1)}$) trade-off between (any desired) spectral expansion λ and degree d. Furthermore, the algorithm is *local*: every vertex can compute its new neighbors as a subset of its original neighborhood of radius $O(\log(1/\lambda))$. The optimal quadratic trade-off is known as the Ramanujan bound, so our construction gives almost Ramanujan expanders from arbitrary expanders.

The locality of the transformation preserves structural properties of the original graph, and thus has many consequences. Applied to Cayley graphs, our transformation shows that *any* expanding finite group has almost Ramanujan expanding generators. Similarly, one can obtain almost optimal explicit constructions of quantum expanders, dimension expanders, monotone expanders, etc., from existing (suboptimal) constructions of such objects. Another consequence is a "derandomized" random walk on the original (suboptimal) expander with almost optimal convergence rate. Our transformation also applies when the degree is not bounded or the expansion is not constant.

We obtain our results by a generalization of Ta-Shma's technique in his breakthrough paper [STOC 2017], used to obtain explicit almost optimal binary codes. Specifically, our spectral amplification extends Ta-Shma's analysis of bias amplification from scalars to matrices of arbitrary dimension in a very natural way. Curiously, while Ta-Shma's explicit bias amplification derandomizes a well-known probabilistic argument (underlying the Gilbert–Varshamov bound), there seems to be no known probabilistic (or other existential) way of achieving our explicit ("high-dimensional") spectral amplification.

*This material is based upon work supported by the NSF grant CCF-1900460. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the NSF.

†This work was partially supported by NSF grant CCF-1900460.

Contents

1 Introduction

1.1 Background

Expander graphs are fundamental objects in computer science and mathematics, possessing a variety of applications in both fields [HLW06, Lub12]. Indeed, expanders (and expansion) play a central role in numerous algorithmic advances, cryptographic schemes, circuit and proof complexity lower bounds, derandomization and pseudorandom generators, error correcting codes, ... and are central to structural results in group theory, algebra, number theory, geometry, combinatorics. In light of this wealth, a central question is

Which graphs are expanders?

A central *quality* measure of expansion of an infinite family of d-regular graphs $\{X_i\}_{i\in\mathbb{N}}$ is the second largest singular value of its normalized adjacency matrix, which we denote by $\lambda(X_i) \in [0,1]$. We say that a family $\{X_i\}_{i\in\mathbb{N}}$ is λ-expanding, for some fixed $\lambda < 1$, if $\lambda(X_i) \leq \lambda$ for every member X_i of the family. The smaller is the expansion parameter λ, the more spectrally expanding is the family. (For simplicity, we will sometimes discuss single graphs rather than families, and say that X is a (d,λ)-expander if it is d-regular and satisfies $\lambda(X) \leq \lambda$.)

A random d-regular graph with $d \geq 3$ is easily shown [Pin73] to be $.99$-expanding with high probability, giving rise to the existence of expanding families. The quest to explicitly construct bounded degree expanders started with Margulis' paper [Mar73], and has been an extremely active research area in the past half century. Today we have a large arsenal of constructions and tools to establish expansion which are quite different in nature, algebraic, analytic, combinatorial, and mixtures of these (for a short survey of this wealth see [Wig18, Sec 8.7]), and we will discuss a few of them below.

Returning to the main discussion, all different constructions above yield d-regular λ-expanding families with *some* specific constants d and λ. Now, a large variety of structural and algorithmic applications call for optimizing both parameters, and understanding the best trade-off between them. One example which is directly related to this paper is the study of random walks on expanders sometimes used for randomness-efficient error-reduction of probabilistic algorithms, and also in the construction of randomness extractors. The surprising *expander Chernoff bound* of Gillman [Gil93] informally says that a sequence of *highly correlated* k vertices along a random walk in a (d,λ)-expander, is almost as good a sampler as a sequence of k *independent* vertices. Saving randomness calls for minimizing the degree d, while improving the quality of the sample requires minimizing the expansion parameter λ.

However, for any choice of degree d, the spectral expansion λ cannot be made arbitrarily small. The Alon–Boppana bound [Nil91] shows that $\lambda(X_i) \geq 2\sqrt{d-1}/d - o(1)$. It intuitively says that the *infinite* d-regular tree is the best possible spectral expander, raising the challenge of achieving it by *finite graphs*. This challenge was first met, by the (independent) seminal papers of [LPS88, Mar88]; they constructed optimal spectrally expanding families, dubbed *Ramanujan graphs*, satisfying the (Ramanujan bound) $\lambda(X_i) \leq 2\sqrt{d-1}/d$. The investigation of expanding families near or achieving the optimal Ramanujan bound

has received much attention. However since then, only one essentially different construction of Ramanujan graphs was found, 30 years later, by [MSS15].

The quest towards almost optimal trade-offs can be summarized as a sharpening of our original major question above:

Which graphs are (almost) Ramanujan expanders?

A study of almost Ramanujan expanders, in which the bound above is nearly matched, has received much attention as well. Friedman [Fri03] greatly strengthened Pinsker's bound above [Pin73], showing that with high probability, a random d-regular graph X satisfies $\lambda(X) \leq 2\sqrt{d-1}/d + o(1)$. Thus, for random regular graphs, expansion and (near) optimal expansion occur "together". For explicit constructions, an approach towards such a bound, which is central for this paper, follows from the *zig-zag product* of [RVW00]. They showed that their basic zig-zag construction achieves an explicit family of expanders with $d \leq 1/\lambda^4$, they further derandomize the basic zig-zag product to achieve $d \leq 1/\lambda^3$, and ask if further derandomization can decrease the exponent to (the optimal) quadratic bound. Ben-Aroya and Ta-Shma [BATS08] in their ingenious "s-wide zig-zag product", nearly matched the optimal quadratic bound[1], achieving $d \leq 1/\lambda^{2+o(1)}$. Their "higher-order" version of zig-zag [BATS08] will be central in our work. A different path to explicitly construct almost Ramanujan graphs was the *lifting method* of Bilu–Linial [BL06], which achieves $d \leq \widetilde{O}(1/\lambda^2)$, and famously led to the (exact) Ramanujan expanders of [MSS15] mentioned above.

It is important to note that while for some applications and structural results, *any* family of expanders would suffice, for many others, the graphs are externally given to us (as e.g. is the case for understanding the expansion of Cayley graphs of groups). Moreover, seeking different constructions and analysis tools has led to surprising applications beyond those intended (e.g., the resolution of the Kadison–Singer conjecture by [MSS14] and the proof of $\mathrm{SL} = \mathrm{L}$ by Reingold [Rei05]).

When is it possible for a family of expanders to get close to the Ramanujan bound?

We show that this is always possible: *any* expander family can be *locally and efficiently* converted into an almost Ramanujan family. More precisely, starting from any family of bounded degree expanders, it is possible to obtain, for any desired target expansion $\lambda > 0$, a new family of λ-expanders close to the Ramanujan bound.

1.2 Main Results

Our main result for general families of expander graphs is as follows.

Theorem 1.1 (Main I - Informal). *Let $\{X_i\}_{i\in\mathbb{N}}$ be a family of (d_0, λ_0)-expanders where $\lambda_0 < 1$ is a constant. For any (target) $\lambda \in (0,1)$ and X_i, we can explicitly construct a (d, λ)-expander, X'_i, on the same vertex set, where $d = O(d_0/\lambda^{2+o(1)})$. Moreover, the construction is local in the sense that edges in X'_i correspond to short walks in X_i.*

[1] We call any such bound near-optimal or almost Ramanujan. Of course, reducing the $o(1)$ slack in the exponent is clearly of much interest.

We obtain our results by considering the seemingly more specialized case of Cayley expanders, which are based on group theory and represent a prominent way of constructing expanders. Recall that a Cayley graph $\mathrm{Cay}(G, S)$ on a finite group G is specified by a symmetric set of generators $S \subseteq G$, where vertices are elements of G and $g, g' \in G$ are adjacent if and only if $g'g^{-1}$ belongs to S.

While many groups admit Cayley expanders, most of these are far from the Ramanujan bound. This is true, in particular, in the case of non-Abelian finite simple groups which includes the symmetric group. Breuillard and Lubotzky [BL18] ask whether it is possible to have near-Ramanujan expanders for all families of finite simple groups. More generally,

Which groups admit expanding Cayley graphs close to the Ramanujan bound?

An equivalent viewpoint arising from the theory of pseudorandomness, is that of *biased distributions*. Here we work with a definition (formalized in Definition 2.4) for operators which naturally generalizes the one for scalars. The equivalence is quite direct – a set $S \subseteq G$ is a a λ-biased distribution if and only if $\mathrm{Cay}(G, S)$ is a λ-expander.

Our key result is that any group that admits a Cayley expander also admits one that is almost Ramanujan.

Theorem 1.2 (Main II)**.** *Let G be a finite group and S be such that $\mathrm{Cay}(G, S)$ is a λ_0-expander, for some constant $\lambda_0 \in (0, 1)$. For every $\lambda \in (0, 1)$, there exists S' such that*

- *$\mathrm{Cay}(G, S')$ is a λ-expander. Equivalently, S' is an λ-biased distribution.*
- *$|S'| = O\left(|S| / \lambda^{2+o(1)}\right)$, and*
- *S' can be computed deterministically in $\mathrm{poly}(|S| / \lambda)$-time assuming an oracle for group operations.*

Furthermore, if $\mathrm{Cay}(G, S)$ is strongly explicit[2], then so is $\mathrm{Cay}(G, S')$.

Remark 1.3. The breakthrough construction of explicit almost optimal binary codes of Ta-Shma [TS17] close to the Gilbert–Varshamov [Gil52, Var57] bound can be viewed as a particular case of Theorem 1.2 applied to a specific family of Abelian groups[3].

Since expanding families of Cayley graphs are known for non-Abelian finite simple groups [BL18, Theorem 3.1], this result makes substantial progress towards the question asked therein (the $o(1)$ term needs to be removed to resolve it completely). Moreover, these are strongly explicit (except for the Suzuki group). Thus, our result yields strongly explicit almost Ramanujan Cayley graphs for these these groups, which notably includes the symmetric group!

[2] Neighbors of a vertex can be computed in polytime in the *description length* of a vertex.

[3] A linear λ_0-balanced code over $\mathbb{F}_2^{n_0}$ of dimension k is equivalent to a Cayley λ_0-expander over $G = \mathbb{F}_2^k$ of degree n_0. Let $S \subseteq G$ be the rows of a generator matrix of a good λ_0-balanced code (good means k/n_0 and $\lambda_0 < 1$ are constants). Applying Theorem 1.2 above to S with final expansion parameter $\lambda > 0$, we obtain a generating set $S' \subseteq G$ of a Cayley λ-expander with degree $O(k/\lambda^{2+o(1)})$, or equivalently, we obtain a λ-balanced code of rate $\Theta(\lambda^{2+o(1)})$.

Corollary 1.4 (Explicit almost Ramanujan Cayley Expanders)**.** *For every non-Abelian finite* simple[4] *group* G *and* $\lambda > 0$*, we can explicitly construct almost-Ramanujan* (d, λ)*-Cayley multigraphs on* G *where* $d \leq O(1/\lambda^{2+o(1)})$.

We can now move from Cayley graphs back to general graphs and answer our original question. A result of König that says that the adjacency matrix of an arbitrary regular graph can be written as a sum of permutation matrices which can be interpreted as elements of the symmetric group. Using this set of permutations as our base set, we can amplify it close to the optimum bound (essentially[5]) using Theorem 1.2. Thus, we obtain Theorem 1.1.

1.3 Applications

We will now discuss some applications of this operator amplification technique which allows us to improve other pseudorandom objects. All the "pseudorandom" objects below are expanders (with various stronger properties). For all, we amplify their spectral bound to almost Ramanujan. We stress that our amplification preserves the underlying structure, and so produces another object with the same properties. Precise definitions of these objects will be given in Section 5.

Quantum Expanders Roughly speaking, a quantum expander is an operator defined by d complex matrices, whose (linear) action on quantum states has a constant spectral gap. Quantum expanders were defined in [AS04, BASTS08, Has07a], and Hastings [Has07c] showed that the Ramanujan bound also applies to them. Existing explicit constructions are far from the Ramanujan bound. In [Har07], Harrow gave a generic construction using expanding Cayley graphs which is explicit if the group has a large irreducible representation and admits efficient Quantum Fourier Transform (QFT). Both these conditions are satisfied by the symmetric group Sym_n using the generating family by Kassabov [Kas07] and the QFT algorithm by Beals [Bea97].

By amplifying the expansion of the generators of [Kas07], we give the first explicit family of almost Ramanujan quantum expanders.

Corollary 1.5 (Explicit Almost Ramanujan Quantum Expanders)**.** *For every* $\lambda \in (0, 1)$*, there is an explicit infinite family of (efficient)* $(O(1/\lambda^{2+o(1)}), \lambda)$*-quantum expanders.*

Monotone Expanders Monotone expanders are expanders, whose edge set can be decomposed into a constant number of *monotone* maps on $[n]$. Bourgain and Yehudayoff [BY13] gave the only known explicit construction of monotone expanders with *constant* degree. By an approach similar to that used for Theorem 1.1, we express it as a sum of permutation matrices and amplify their expansion obtaining the following result.

Corollary 1.6 (Almost Ramanujan Monotone Expanders)**.** *For every* $\lambda > 0$*, there is an explicit family* $\{X_i\}_{i \in \mathbb{N}}$ *of (vertex)* d*-regular* $d^{O(1)}$*-monotone expanders with* $d = O(1/\lambda^{2+o(1)})$ *and* $\lambda(X_i) \leq \lambda$.

[4] This holds for other groups as well, as long as they have expanding generators. One non-simple example is the Cayley expanders of Rozenman, Shalev and Wigderson [RSW06].

[5] Actually, we only consider the *standard* representation in this amplification.

Remark 1.7. There are two natural notions of degree for a monotone expander. The usual vertex degree and the number of monotone maps. Our almost Ramanujan trade-off is with respect to the vertex degree (and the monotone degree is polynomial in the vertex degree). It would be really interesting to obtain an almost Ramanujan trade-off with respect to the monotone degree.

Dimension Expanders Loosely speaking, dimension expanders (over any field $\mathbb{F}$) are a linear algebraic extension of expanders: a collection of d linear maps on $\mathbb{F}^n$, which significantly *expands* (the span of) any vector space of dimension below $n/2$. They were defined by Barak et al. in [BISW01]. Over the complex numbers, any quantum expander is a dimension expander. More generally, Dvir and Shpilka [DS09] proved that a monotone expander directly yields a dimension expander over every field. We give spectral almost Ramanujan expanders that have the additional property of being dimension expanders. Additionally, if the starting dimension is small enough then we achieve almost doubling of the starting dimension.

Kazhdan Constant We can also amplify operators in *infinite dimensional* Hilbert spaces. This allows us to obtain improved (average) Kazhdan constants of groups with "Property (T)", which is an analogue of expansion for discrete groups. This implies better bounds for the *product replacement algorithm* to sample group elements.

Corollary 1.8 (Amplifying Average Kazhdan Constant)**.** *Let* G *be a discrete group and* S *a finite set of generators such that the average Kazhdan constant* $\overline{\mathcal{K}}(G,S)$ *is equal to* $2\cdot(1-\lambda_0)$ *for some constant* $\lambda_0 \in (0,1)$*. For every* $\lambda \in (0,1)$*, there is a set* $S' \subseteq G$ *such that*

1. $\overline{\mathcal{K}}(G,S') \geq 2\cdot(1-\lambda)$*, and thus,* $\mathcal{K}(G,S') \geq 2\cdot(1-\lambda)$*.*
2. $|S'| = O_{\lambda_0}(|S|/\lambda^{2+o(1)})$*, and*
3. S' *can be found in time* $\mathrm{poly}(|S|/\lambda)$ *assuming an oracle for group operations on* G*.*

Randomness-efficient Walks An immediate consequence of being able to achieve an almost optimum degree versus expansion trade-off in this generic way is that we obtain randomness-efficient random walks.

1.4 Techniques

We consider the main contribution of this work to be the broad applicability of the near-optimal operator amplification to *any* family of expanders. For instance, the existence of almost Ramanujan expanders for all expanding groups, including the symmetric group, is quite surprising to us. On the technical side, we view our main contribution as the identification of appropriate natural linear algebraic extensions to Ta-Shma's amplification framework [TS17] that accommodate amplification of operators as described above. This extension will be so natural that it may almost feel that we are replacing absolute values in the original scalar analysis [TS17] by operator norms. However, appropriate generalizations and care are needed in such an extension to operators.

We first recall the problem and see why it is non-trivial. Let G be a finite group and S be a symmetric multiset such that $\mathrm{Cay}(G, S)$ is a λ_0-expander for some $\lambda_0 \in (0, 1)$. Assume that $\mathrm{Cay}(G, S)$ is far from being Ramanujan, e.g., $|S| = 1/\lambda_0^{100}$. Our goal is to construct a new generating set S' such that $\mathrm{Cay}(G, S')$ is a λ-spectral expander with an almost optimal final degree, say, $|S'| = O(1/\lambda^{2.001})$.

A first approach would be to take S' to be the power S^t with $t \approx \log_{\lambda_0}(\lambda)$. However, now the degree, $|S|^t = O(1/\lambda^{100})$, has also increased and the trade-off remains the same. Thus, we want to efficiently compute a sparse subset of S^t that retains the expansion. Since we know what degree we are aiming for, we could try take a sparse random sample $S' \subseteq S^t$ of size $d = O(1/\lambda^{2.001})$ and hope that some form of matrix concentration ensures that $\mathrm{Cay}(G, S')$ is λ'-spectral expander with $\lambda' \approx \lambda$. Unfortunately, it is not clear how to show even the existence of a *single* sparse subset S' that achieves the required expansion[6]. Standard probabilistic techniques, such as the matrix Chernoff, have a forbidding dependence on the dimension of the matrices for this application.

Switching to the bias distribution viewpoint, a subset $S \subseteq G$ is said to be ε-biased if it *fools* all non-trivial irreducible representations, i.e., for every non-trivial irreducible *representation*, ρ, of G, we have $\|\mathbb{E}_{s\sim S}[\rho(s)]\|_{\mathrm{op}} \leq \varepsilon$. Here, a representation of a group is an operator valued function, $\rho : G \to M_\ell(\mathbb{C})$, that is multiplicative, i.e., for every two group elements g_1, g_2 we have $\rho(g_1 g_2) = \rho(g_1)\rho(g_2)$. As mentioned earlier, $\mathrm{Cay}(G, S)$ is λ-expanding if and only if S is λ-biased set. Thus, the problem of constructing optimal Cayley expander can be reformulated as construction of small biased distribution with optimal support size. In fact, we will see that the techniques work for general matrix value functions (not just representations).

Earlier Work Much of the earlier work has focused on the case of Abelian groups. It is well-known that the irreducible representations of these groups are 1-dimensional, i.e., scalar valued functions called *characters*. The special case of ε-biased distributions over $G = \mathbb{Z}_2^k$ introduced in the pioneering work of Naor and Naor [NN90]. One of their constructions of ε-biased distribution uses walks on expander graphs. These distributions have found numerous applications (e.g., [ABN+92, Vad12, TS17]).

Rozenman and Wigderson (see analysis in [TS17]) introduced the following "scalar amplification" technique using walks on an (auxiliary) expander graph X, whose vertices are identified with elements of S. Let $W \subseteq S^t$ be the collection of all walks of length $(t-1)$ on X. Let $f : S \to \{\pm 1\}$ be *any* function. The collection W naturally gives rise to a map T_W that lifts $\{\pm 1\}$-valued functions on S to $\{\pm 1\}$-valued functions on W by multiplication as follows

$$T_W(f)(w) = f(w) := f(s_0) \cdots f(s_{t-1}) \qquad \forall\, w = (s_0, \ldots, s_{t-1}) \in W\,.$$

In words, the value of each walk is given by the product of the values f assigns to its vertices. For a sufficiently "pseudorandom"[7] collection W and any function f satisfying

[6] To some extent this difficulty is also present in the proof of the Alon–Roichman theorem [AR94] and the reason why even for non-Abelian groups the only generic upper bound known uses $\Omega(\log(|G|))$ random generators to obtain an expander. Recall that matrix Chernoff bounds deteriorate with the dimension of the matrices, and we have no fixed bound on their dimension here.

[7] This amounts to X being sufficiently expanding.

$\mathsf{bias}(f) \le \varepsilon_0$, they argue that the bias of the amplified function, $T_W(f)$, decreases exponentially (roughly) as $\mathsf{bias}(T_W(f)) \le \varepsilon_0^{t/2}$. Note that, when f is a character ρ (later we will consider more general representations), we can use the homomorphism property to write

$$T_W(\rho)(w) = \rho(s_0 \cdots s_{t-1}) \qquad \forall\, w = (s_0, \ldots, s_{t-1}) \in W\,.$$

This allow us to interpret $T_W(\rho)$ as a function with domain on the multiset $S' = \{s_0 \cdots s_{t-1} \mid w = (s_0, \ldots, s_{t-1}) \in W\}$, our new biased set. This technique gives an ε-biased distribution with support size $O(1/\varepsilon^{4+o(1)})$ (*cf.*, [TS17]), which is quite good but still sub-optimal.

Ta-Shma [TS17] managed to close the gap almost optimally using the *s-wide replacement product* to derandomize the amplification of Rozenman and Wigderson. Recall that the s-wide replacement product of Ben-Aroya and Ta-Shma [BASTS08] is a higher-order version of the zig-zag product [RVW00]. Using the collection of walks on the s-wide replacement product allows for a much smaller collection $W \subseteq S^t$ with nearly optimal size. This scalar technique was later applied to the more general case of arbitrary Abelian groups by Jalan and Moshkowitz [JM21]. These results can be encapsulated in the following statement.

Theorem 1.9 (Scalar Amplification)**.** *Let S be a finite set and $\lambda_0 \in (0,1)$ be a constant. For every $\lambda > 0$, there exists a deterministic polynomial time algorithm to construct $W \subseteq S^t$ of size $|W| \le O(|S|/\lambda^{2+o(1)})$ such that for every function $f : S \to \mathbb{C}$ with $|\mathbb{E}_{s\sim S}[f(s)]| \le \lambda_0$ and $\|f\|_\infty \le 1$, we have $|\mathbb{E}_{w\sim W}[f(w)]| \le \lambda$.*

Our Results To extend Ta-Shma's approach to non-Abelian groups, it is necessary to work with operator valued functions, $f\colon S \to M_\ell(\mathbb{C})$, as the irreducible representations are no longer of dimension one. In fact, the amplification applies to any operator valued function. Our main technical result is a *dimension independent* generalization of the scalar amplification result to operator valued functions. Note that the definition of T_W extends naturally to a mapping from $M_\ell(\mathbb{C})^S$ to $M_\ell(\mathbb{C})^W$.

Theorem 1.10 (Operator Amplification (this work))**.** *Let S be a finite set and $\lambda_0 \in (0,1)$ be a constant. For every $\lambda > 0$, there exists a deterministic polynomial time algorithm to construct $W \subseteq S^t$ of size $|W| \le O(|S|/\lambda^{2+o(1)})$ such that for every function $f : S \to M_\ell(\mathbb{C})$ with $\|\mathbb{E}_{s\sim S}[f(s)]\|_{\mathrm{op}} \le \lambda_0$ and $\max_s \|f(s)\|_{\mathrm{op}} \le 1$, we have $\|\mathbb{E}_{w\sim W}[f(w)]\|_{\mathrm{op}} \le \lambda$.*

To establish the operator valued generalization, we make a simple and yet extremely useful change in the bias operator (Π_f) defined by Ta-Shma which is a key object in the analysis of both [TS17] and [JM21]. In both these cases, f is scalar, and one defines

$$\Pi_f : \mathbb{C}[S] \to \mathbb{C}[S] \;\text{ where } \Pi_f \cdot s = f(s) \cdot s\,.$$

However, this approach is not readily generalizable to operators and the view we take is that if $f : S \to M_\ell(\mathbb{C})$, then, Π_f is actually an operator on $\mathbb{C}^\ell \otimes \mathbb{C}[S]$ defined as

$$\Pi_f : \mathbb{C}^\ell \otimes \mathbb{C}[S] \to \mathbb{C}^\ell \otimes \mathbb{C}[S] \;\text{ where } \Pi_f(v \otimes s) = f(s)\, v \otimes s\,.$$

Clearly, in the Abelian case, we have $\ell = 1$ and this is isomorphic to the setup by Ta-Shma. This generalization is very natural and we show that not only does the older

machinery gel well with this, but the proof remains intuitive with the different spaces neatly delineated. More precisely, we first establish an *operator version* of the Rozenman and Wigderson amplification, and then we derandomize it using (a suitable version of) the s-wide replacement product. Furthermore, since the result does not depend on the dimension, ℓ, we can use it even for functions $f : S \to \mathcal{L}(\mathcal{H})$ where $\mathcal{L}(\mathcal{H})$ is the space of bounded linear operators on an arbitrary Hilbert space, $\mathcal{H}$, possibly infinite dimensional. This is useful if the underlying group is not finite but finitely generated by S.

To the best of our knowledge, only one general result was known for general groups. Chen, Moore and Russell [CMR13] analyzed the usual expander walk construction using a matrix version of the expander mixing lemma. This gives an amplification procedure for Cayley graphs of general groups, but the resulting degree $O(|S|/\lambda^{11})$ to achieve final expansion λ is sub-optimal. Analogous to the (folklore results of the) scalar case, we show that the analysis in [CMR13] of the amplification via (iterated applications of) expander mixing lemma can be improved to get $O(|S|/\lambda^{4+o(1)})$ achieving similar parameters to the expander walk approach.

1.5 Discussion

The results of this paper have some curious features, which we would like to elaborate on. For most of them, we will use the following "bare bones" description of our main spectral amplification result. Namely, let S be a finite set and $\mathcal{H}$ a Hilbert space. Let f be a function mapping elements of S to operators on $\mathcal{H}$ of unit norm, such that $\|\mathbb{E}_{s\in S}[f(s)]\|_{op} \le \lambda_0$. For any $\lambda > 0$ take $t = c\log(1/\lambda)$ (for appropriate c). We extend f from S to S^t by defining $f(s_1,\ldots,s_t) = f(s_1)\cdots f(s_t)$. Clearly, $\|\mathbb{E}_{r\in S^t}[f(r)]\|_{op} \le (\|\mathbb{E}_{s\in S}[f(s)]\|_{op})^t \le \lambda$. Our main result is an explicit construction of a (pseudorandom) subset $S' \subseteq S^t$, of size only $|S'| = O(|S|/\lambda^{2+o(1)})$, with a similar guarantee, namely $\|\mathbb{E}_{s'\in S'}[f(s')]\|_{op} \le \lambda$.

Dimension Independence Note that if the operators in S are 1-dimensional, namely scalars, then the *existence* of a set S' of this size (which is best possible even in this 1-dimensional case) follows directly from the Chernoff bound. Indeed, Ta-Shma's construction [TS17] may be viewed as derandomizing this result, producing an explicit such S'.

One may try to do the same for operators in a higher dimension, say ℓ, by appealing to the Matrix Chernoff bounds of Ahlswede–Winter [AW02] (see also Tropp [Tro15]). However, these concentration inequalities pay a factor of ℓ in the tail bound, resulting in a set S' of size $\Omega(\log(\ell))$. As the dimension ℓ is arbitrary (indeed, may be infinite), such a bound is useless.

Thus, our explicit construction has no known probabilistic (or other existential) analog! What is curious is that our dimension-independent analysis follows very closely that of Ta-Shma for 1-dimension, roughly speaking, replacing scalar absolute values by the operator norm in any dimension. We feel that it would be extremely interesting to find a matrix concentration inequality for sampling product sets like S^t, which is dimension independent.

Algebraic vs. Combinatorial Expander Constructions Our explicit construction of the pseudorandom set S' above uses expanders obtained from the s-wide zig-zag product of [BATS08]. This is a combinatorial construction, a refinement of the original zig-zag product construction of [RVW00]. Nonetheless, it has significant consequences to algebraic expander constructions which use group theory, namely to the expansion of Cayley graphs. This is possible due to the abstraction of how elements of S are mapped to operators via some function f. We can take S to be expanding generating set of a group and f to be some non-trivial irreducible representation ρ. Instead of using the t-product of elements of S^t to obtain a new amplified generating set S', a much sparser subset is chosen using the s-wide zig-zag construction. The analysis of the norm amplification discussed above yields the required expansion bound, in a way that has no dependence on the group or the representation. The flexibility in mapping element of S to operators underlies the versatility of our spectral amplification. It allow us to preserve some of the structure of the expanders whose expansion are being amplified. In this case, both the starting expander and the amplified expander are Cayley graph over the same group.

It is interesting to note that this is a recurring phenomenon. In [ALW01], it was discovered that the zig-zag product may be viewed as a combinatorial generalization of the algebraic semi-direct product of groups. This connection made possible the construction of new expanding Cayley graphs in groups that are far from being simple, e.g., in [MW04, RSW06]. It is rewarding to see again how new combinatorial constructions, sometimes inferior in certain parameters to some algebraic ones, yield new results in group theory.

Iterated Pseudorandomness Another interesting aspect of our result is the following. Recall that expanders are pseudorandom objects for many purposes. One important purpose is sampling - rather than sampling t independent random elements in some set S, one may sample t points along a random walk on an expander on the vertex set S and a Chernoff type bound still holds (a nontrivial result of [Gil93])- this affords significant savings in the number of random bits spent. For this result, any expander would do. What happens in this paper is an iterated use of expanders as samplers as follows. We first choose a sparse pseudorandom set of t-walks inside S^t using expanders walks. Then, we choose a yet sparser pseudorandom set inside it, again using walks on an additional expander. This repeated use of expanders improves the trade-off between quality of spectral amplification and the size of the final pseudorandom set to near-optimal. Now the construction of this iterated selection of walks seems critical, and (as in Ta-Shma's paper) is chosen to come from the s-wide zig-zag product of two expanders [BATS08].

Group Theory For us, the most surprising consequence of our results is that "weak" simple groups, especially the symmetric group,[8] can have near-Ramanujan generators. The question of which groups are expanding, and just how expanding they are, is an old quest of group theory. One dichotomy is whether *every* finite set of generators of the group is expanding (these are "strongly expanding" groups), or if some are and some aren't (these are "weakly expanding" groups). For the symmetric group, many finite non-expanding

[8] See also the groups in [RSW06], which are iterated wreath products of symmetric groups.

generating sets of constant size were long known, and Kassabov's breakthrough construction [Kas07] designed a constant size *expanding* generating set. The symmetric group is then a weakly expanding group, while, e.g., simple groups of Lie type (namely, matrix groups) are believed, and in some cases known, to be strongly expanding. Nonetheless, our construction works equally well for all, and we have almost Ramanujan generators for all expanding groups.

Semigroups and universality Perhaps the most general way to view our main result is the following abstraction. Let S be any finite set (which may be best viewed as an alphabet) of some size $|S| = n$, and for every integer t we consider subsets $W \subseteq S^t$ of words of length t. We call such a subset, W, λ-universal if, informally, W amplifies the bias of any linear operator valued function on S. More precisely, if for every Hilbert space $\mathcal{H}$ and every function $f : S \to \mathcal{L}(H)$ satisfying $\|f(v)\|_{op} \leq 1$ for all $v \in S$, and $\|\mathbb{E}_S[f(v)]\|_{op} \leq \frac{1}{2}$ say,[9] we have $\|\mathbb{E}_W[f(w)]\|_{op} \leq \lambda$, where for $w = (v_0, v_1, \ldots, v_{t-1}) \in S^t$, $f(w)$ is a shorthand for $f(w) = f(v_0)f(v_1)\ldots f(v_{t-1})$.

This semigroup viewpoint of words stresses the non-commutativity of composing the operators. It is easy to see how to derive the results for groups directly from the result above: take S to be a set of expanding[10] elements in some group G, and f is some irreducible representation of G. In this nice case, W will itself be a subset of G, and so an almost Ramanujan expanding set of generators (and as f is a homomorphism in this case, $f(w)$ with w interpreted as a group element will actually match the definition of $f(w)$ when w is interpreted as a word over S).

Closeness to the Ramanujan Bound As mentioned above, a family of d-regular graphs is called Ramanujan if its spectral expansion parameter λ is at most $2\sqrt{d-1}/d$. This terminology was introduced in the seminal work of Lubotzky, Phillips and Sarnak [LPS88], and it designates the optimal degree versus expansion trade-off that a family of bounded degree expanders can achieve. Several notions of closeness to the Ramanujan bound were investigated, e.g., $(2\sqrt{d-1}+\varepsilon)/d$ (with $\varepsilon > 0$ small or vanishing) in [Fri03, MOP20, JMO+22], $O(1/\sqrt{d})$ in [ACKM19], $\mathrm{polylog}(d)/\sqrt{d}$ in[BL06, JMO+22] and more generally $d^{o_d(1)}/d^{1/2}$.

In this work, we obtain a bound of the form $\lambda \leq O(2^{\log(d)^c}/d^{1/2})$ for some constant $c \in (0,1)$, which we refer to as an almost Ramanujan bound. Rephrasing in terms of the expansion parameter, we achieve expansion λ with degree $O(1/\lambda^{2+\beta})$, where $\beta = O(1/\log(1/\lambda))^{c'}$ for some $c' \in (0,1)$. We stress that the nomenclatures almost Ramanujan, near Ramanujan and etc, may vary depending on the author. Improving the results in this work to achieve trade-offs even closer to the Ramanujan bound (if possible) is of great interest. We suspect that new ideas may be required to substantially improve the bound to, say, expansion λ versus degree $O(\mathrm{polylog}(1/\lambda)/\lambda^2)$.

[9] The constant $\frac{1}{2}$ is chosen for simplicity - in general we will have an initial bias parameter λ_0.
[10] With second eigenvalue $1/2$.

1.6 Outline

We start in Section 2 by summarizing basic definitions and the notation used throughout the paper. In Section 3, we generalize the simpler construction of Ta-Shma based on expander walks. Apart from serving as a nice warm-up to the more-involved construction, it will be used as a bootstrap for the more involved construction based on s-wide replacement product which is the subject of Section 4. Here, we prove the main amplification result (a formal version of Theorem 1.10 above) and instantiate using known constructions and those obtained from Section 3 which establishes Theorem 1.2. Section 5 discusses the permutation amplification trick and formally completes the proof of Theorem 1.1. It also discusses the other applications in more detail. Finally Section 6 gives an operator version of the expander mixing lemma which improves the analysis of [CMR13].

2 Preliminaries

Let $X = (V, E)$ be an n-vertex d-regular multigraph for some $d \geq 1$. We denote by A_X the normalized adjacency matrix of X, i.e., the adjacency matrix divided by d.

Definition 2.1 (λ-spectral Expander)**.** Let the eigenvalues of the matrix A_X, denoted as $\mathrm{Spec}(A_X)$, be $\{1 = \lambda_1 \geq \cdots \geq \lambda_n\}$ and define $\lambda(X) = \max\{|\lambda_2|, |\lambda_n|\}$. We say that X is a λ-spectral expander if $\lambda(X) \leq \lambda$.

We denote by G a finite group (except in Section 5.5 where we only need it to be finitely generated). For a multiset $S \subseteq G$, $\mathrm{Cay}(G, S)$ denotes a multigraph[11] with the vertex set being G and edges $\{(g, sg) \mid g \in G,\ s \in S\}$.

Group Representations In order to study the expansion of a Cayley graph, we will use the notion of a group representation[12]. *Weyl's unitary trick*, says that for a large family of groups (which includes all finite groups), every representation can be made unitary and thus, we can restrict to studying these.

Let $\mathcal{H}$ be a complex Hilbert space and denote by $\mathcal{L}(\mathcal{H})$ the algebra of bounded linear operators[13] on $\mathcal{H}$. We denote by $U_{\mathcal{H}}$ the unitary group of operators acting on $\mathcal{H}$.

Definition 2.2 (Unitary Group Representation)**.** For a group G, a unitary representation is a pair $(\rho, \mathcal{H})$ where $\rho : G \to U_{\mathcal{H}}$ is a group homomorphism, i.e., for every $g_1, g_2 \in G$, we have $\rho(g_1 g_2) = \rho(g_1)\rho(g_2)$. A representation is *irreducible* if the only subspaces of $\mathcal{H}$ that are invariant under the action of $\rho(G)$ are the empty space, $\{0\}$, and the entire space, $\mathcal{H}$.

Every group has two special representations, which are,

1. (Trivial representation) - $(\rho, \mathbb{C})$ where for every g, $\rho(g) = 1$.

[11] Note that unless $S = S^{-1}$, the graph $\mathrm{Cay}(G, S)$ is a directed multigraph.

[12] Additional background on representation theory of finite groups can be found in [SS96].

[13] For most applications, one can think of $\mathcal{H} = \mathbb{C}^n$ for some n, and $\mathcal{L}(\mathcal{H}) = M_n(\mathbb{C})$, the space of $n \times n$ complex matrices. However, we will need the generality in Section 5.5.

2. ((left) Regular representation) - $(\rho_{\mathrm{reg}}, \mathcal{V}_{\mathrm{reg}})$ where, $\mathcal{V}_{\mathrm{reg}} = \mathbb{C}[G]$ is a vector space with the elements of G being an orthonormal basis, and $\rho_{\mathrm{reg}}(g) : h \mapsto g \cdot h$.

Fact 2.3. Let G be a finite group and let $\mathcal{V}_{\mathrm{reg}}$ be the regular representation over $\mathbb{C}$. We have

$$\mathcal{V}_{\mathrm{reg}} \cong \bigoplus_{(\rho, V_\rho) \in \mathrm{Irrep}(G)} \dim(\rho) \cdot \mathcal{V}_\rho,$$

where Irrep(G) denotes the set of irreducible unitary representations of G.

Expanders and Biased Distributions It follows from definitions that the normalized adjacency matrix of $\mathrm{Cay}(G, S)$ is given by $A = \mathbb{E}_{s \sim S}[\rho_{\mathrm{reg}}(s)]$. Moreover, the copy of the trivial representation is the space spanned by the all-ones vector. Fact 2.3 implies that this can be block diagonalized and therefore,

$$\mathrm{Spec}(A) = \bigcup_{\rho \in \mathrm{Irrep}(G)} \mathrm{Spec}(\mathop{\mathbb{E}}_{s \sim S}[\rho(s)]), \quad \text{and thus,}$$

$$\lambda(\mathrm{Cay}(G, S)) = \max_{\substack{\rho \in \mathrm{Irrep}(G) \\ \rho \text{ is non-trivial}}} \left\| \mathop{\mathbb{E}}_{s \sim S}[\rho(s)] \right\|_{\mathrm{op}}.$$

Recall that for any bounded linear operator, $T : \mathcal{H} \to \mathcal{H}'$, between (non-empty) Hilbert spaces, we have

$$\|T\|_{\mathrm{op}} = \sup_{v \in \mathcal{H} : \|v\| = 1} \|Tv\| = \sup_{v \in \mathcal{H}, w \in \mathcal{H}' : \|v\| = \|w\| = 1} |\langle Tv, w \rangle| ,$$

where $\|v\| = \sqrt{\langle v, v \rangle_{\mathcal{H}}}$ and $\|w\| = \sqrt{\langle w, w \rangle_{\mathcal{H}'}}$.

Given this equivalence, we will find it convenient to work with the operator norm version referred to as *bias* in the literature [CMR13].

Definition 2.4 (Biased Distribution on G)**.** Let $\varepsilon \in (0, 1)$. We say that a multiset S of elements of a group G is ε-biased if for every non-trivial irreducible representation ρ, we have $\|\mathbb{E}_{s \sim S}[\rho(s)]\|_{\mathrm{op}} \leq \varepsilon$. We sometimes use the shorthand $\mathrm{bias}(S) \leq \varepsilon$, where $\mathrm{bias}(S) = \lambda(\mathrm{Cay}(G, S))$.

Irreducible representations of Abelian groups, called *characters*, have dimension 1. Thus, this definition coincides with the usual one of ε-biased distribution *fooling* non-trivial characters [NN90, AGHP92]. These pseudorandom distributions were introduced in the pioneering work of Naor and Naor where several applications to derandomization were given [NN90].

Notation

Since we deal with various vector spaces and graphs, we will find it useful to establish some convenient notation. While we recall these in the relevant section, the following is a summary for ready reference.

- The main multigraphs we study will be X and Y with vertices V_X, V_Y and normalized adjacency operators A_X, A_Y.
- We denote vertices of X, Y by x, y and an ordered tuple of vertices by $\vec{x} = (x_0, \cdots, x_t)$.
- We use u, v, w to denote arbitrary vectors in $\mathcal{H}$ and x, y for basis vectors of $\mathbb{C}[V_X], \mathbb{C}[V_Y]$ where $\mathbb{C}[V_X]$ is the complex vector space with the elements of V_X being a orthonormal basis.
- The tensored vector spaces have an induced inner product. For $\mathcal{X}_\mathcal{H} := \mathcal{H} \otimes \mathbb{C}[V_X]$, it is $\langle v \otimes x, w \otimes x' \rangle = \langle v, w \rangle_\mathcal{H} \langle x, x' \rangle$. Similarly, we have one on $\mathcal{X}\mathcal{Y}_\mathcal{H} := \mathcal{X}_\mathcal{H} \otimes \mathbb{C}[V_Y]$.
- Orthogonal decomposition: $\mathcal{X}_\mathcal{H} = \mathcal{X}_\mathcal{H}^{\|} \oplus \mathcal{X}_\mathcal{H}^{\perp}$ where $\mathcal{X}_\mathcal{H}^{\|} := \text{span}\{v \otimes \vec{1} \mid v \in \mathcal{H}\}$. Here, $\vec{1}$ denotes the un-normalized all-ones vector. Similarly, $\mathcal{X}\mathcal{Y}_\mathcal{H} = \mathcal{X}\mathcal{Y}_\mathcal{H}^{\|} \oplus \mathcal{X}\mathcal{Y}_\mathcal{H}^{\perp}$, where $\mathcal{X}\mathcal{Y}_\mathcal{H}^{\|} := \text{span}\{z \otimes \vec{1} \mid z \in \mathcal{X}_\mathcal{H}\}$.
- The operator $\mathring{A}$ denotes the extension of operator A to a tensor product of spaces where it acts as identity on the other spaces. For example, A_X acts on $\mathbb{C}[V_X]$ and its extension to $\mathcal{X}_\mathcal{H}$ is $\mathring{A}_X = I_\mathcal{H} \otimes A_X$. However, if we were working on $\mathcal{X}\mathcal{Y}_\mathcal{H}$, it would be $\mathring{A}_X = I_\mathcal{H} \otimes A_X \otimes I_Y$ instead[14].
- Given an operator valued function $f : V_X \to \mathcal{L}(\mathcal{H})$, the generalized *bias operator* is defined as[15],

$$\Pi_f : \mathcal{X}_\mathcal{H} \to \mathcal{X}_\mathcal{H}, \quad v \otimes x \mapsto f(x)\, v \otimes x.$$

3 Operator Bias Reduction via Expander Walks

In this section, we establish a new *operator* analogue of the (expander walk based) bias amplification procedure for *scalars* due to Rozenman and Wigderson. An analysis of this scalar amplification was given by Ta-Shma in [TS17]. More precisely, we first prove the following operator analogue for constant bias (Theorem 3.1) and later generalize it to any bias (Theorem 3.8) in Section 3.3.

Theorem 3.1 (Operator Amplification via Expander Walks)**.** *Let X be a $\lambda(X)$-spectral expander and let $\mathcal{W}_t$ be the collection of walks obtained from walks of length t on X. Then for any operator valued function f such that $\|\mathbb{E}_{x \in V_X}[f(x)]\|_{op} \le \lambda_0$ and $\max_{x \in V_X} \|f(x)\|_{op} \le 1$, we have*

$$\left\| \mathop{\mathbb{E}}_{(s_0, \ldots s_t) \in \mathcal{W}_t} [f(s_t) \cdots f(s_0)] \right\|_{op} \le (2\lambda(X) + \lambda_0)^{\lfloor t/2 \rfloor}.$$

We remark that a precursor of these techniques, in the simpler setting of Abelian groups, appears in the pioneering work of Naor and Naor introducing ε-biased distributions over the group $\mathbb{Z}_2^m$. There, they also amplify bias using expander walks on an auxiliary expander [NN90].

[14] The spaces will be self-evident and the use of the same notation should not be confusing.

[15] An equivalent matrix definition is $\Pi_f := \sum_{x \in V_X} f(x) \otimes E_{x,x}$ where $E_{x,x} \in \mathbb{C}^{V_X \times V_X}$ is the diagonal matrix with exactly one non-zero entry of value 1 in the row and column indexed by the vertex x.

This simpler amplification of Theorem 3.1 will be crucially used in the full almost optimal amplification (which derandomizes it) and also to bootstrap it. Moreover, it yields a construction of expanding Cayley graphs of small sizes which will be required later.

This bias reduction procedure uses walks on an auxiliary expander graph. Here, we only use its expansion property (as opposed to later when we rely on its structure for the s-wide construction). With this it is already possible to obtain $1/\lambda^{4+o(1)}$ dependence on the final degree of an λ-expander.

Theorem 3.2. *Let $S \subseteq G$ such that $\lambda(\mathrm{Cay}(G,S)) = \lambda_0 < 1$. For every $\lambda \in (0,1)$ and constant $\beta \in (0,1)$, we can find $S' \subseteq G$ in time $\mathrm{poly}(|S|, 1/\lambda_0, 1/\lambda)$ such that $\lambda(\mathrm{Cay}(G,S')) \leq \lambda$ and $|S'| = O_{\lambda_0}(|S|/\lambda^{4+\beta})$.*

Towards this, we first formalize the connection between bias of a special subset of a group and the operator norm of a certain operator. The subset is obtained by taking random walks over an expander graph as mentioned above. We then proceed to bound this operator norm. Finally, we instantiate our construction with an explicit expander graph due to [Alo21].

The Analysis Let S be any finite set and let X be a graph on the vertex set $V_X = S$ with A_X being its normalized adjacency matrix. Let $\mathcal{H}$ be a complex Hilbert space and $\mathcal{L}(\mathcal{H})$ be the (bounded) operators on $\mathcal{H}$; an important example will be $\mathcal{L}(\mathcal{H}) = M_\ell(\mathbb{C})$. For *any* operator valued function, $f : S \to \mathcal{L}(\mathcal{H})$, we define the generalized bias operator as

$$\Pi_f : \mathcal{H} \otimes \mathbb{C}[V_X] \mapsto \mathcal{H} \otimes \mathbb{C}[V_X], \quad \Pi_f(v \otimes x) = f(x)\, v \otimes x\,.$$

In the scalar case, since $\mathcal{H} = \mathbb{C}$, earlier works [TS17, JM21] used the implicit identification $\mathbb{C} \otimes \mathbb{C}[V_X] \cong \mathbb{C}[V_X]$ and defined Π_f as a diagonal matrix. This identification no longer is suitable when f is operator valued in dimension > 1. However, a simple yet crucial observation is that merely decoupling the spaces allows us to collect the terms as we proceed along the walk.

Let $\mathcal{W}_t \subseteq S^{t+1}$ be the collection of all length t walks on the graph X and we define $\mathring{A}_X = I_{\mathcal{H}} \otimes A_X$. Then, we have

Lemma 3.3.

$$\Pi_f \left(\mathring{A}_X\, \Pi_f\right)^t \mathop{\mathbb{E}}_{s \in S} [v \otimes s] \;=\; \mathop{\mathbb{E}}_{(s_t,\cdots,s_0)\in \mathcal{W}_t} [f(s_t)\cdots f(s_0)]\, v \otimes s_t\,. \tag{1}$$

This can be shown easily via an induction on t and we refer to Lemma 4.7 for a formal proof of a more general statement. A minor technicality is that the operators in the image of f act on $\mathcal{H}$ whereas Π_f acts on the space $\mathcal{X}_{\mathcal{H}} := \mathcal{H} \otimes \mathbb{C}[V_X]$. We use projection and lifting maps to move between the spaces $\mathcal{X}_{\mathcal{H}}$ and $\mathcal{H}$. Define $P_{\mathcal{H}} : \mathcal{X}_{\mathcal{H}} \to \mathcal{H}$ and $L_{\mathcal{H}} : \mathcal{H} \to \mathcal{X}_{\mathcal{H}}$, as,

$$P_{\mathcal{H}}(w \otimes x) = w, \quad L_{\mathcal{H}}(v) = \mathop{\mathbb{E}}_{x \in V_X} [v \otimes x]\,.$$

It follows directly from the definition that $\|L_{\mathcal{H}}\|_{op} = 1/\sqrt{|V_X|}$ and we can use Cauchy-Schwarz to get that $\|P_{\mathcal{H}}\|_{op} = \sqrt{|V_X|}$. Now, we put this together to obtain a simple expres-

sion on the quantity we need to bound

$$\begin{aligned}
\left\| \mathop{\mathbb{E}}_{(s_0,\dots s_t)\in\mathcal{W}_t} [f(s_t)\cdots f(s_0)] \right\|_{op} &= \sup_{\|v\|=1} \left\| \mathop{\mathbb{E}}_{(s_0,\cdots,s_t)\in\mathcal{W}_t} [f(s_t)\cdots f(s_0)]\, v \right\|_2 \\
&= \sup_{\|v\|=1} \left\| P_{\mathcal{H}} \left(\mathop{\mathbb{E}}_{(s_0,\cdots,s_t)\in\mathcal{W}_t} [f(s_t)\cdots f(s_0)]\, v \otimes s_t \right) \right\|_2 \\
&= \sup_{\|v\|=1} \left\| P_{\mathcal{H}} \Pi_f \left(\mathring{A}_X\, \Pi_f \right)^t \mathop{\mathbb{E}}_{s\in S} [v\otimes s] \right\|_2 \\
&= \sup_{\|v\|=1} \left\| P_{\mathcal{H}} \Pi_f \left(\mathring{A}_X\, \Pi_f \right)^t L_{\mathcal{H}} v \right\|_2 \\
&\le \left\| \Pi_f \left(\mathring{A}_X\, \Pi_f \right)^t \right\|_{op} \|P_{\mathcal{H}}\|_{op}\, \|L_{\mathcal{H}}\|_{op} \\
&\le \left\| \Pi_f \left(\mathring{A}_X\, \Pi_f \right)^t \right\|_{op} .
\end{aligned}$$

The Construction of Amplified Biased Sets The particular case of $S \subseteq G$ (for some group G) and the function f being a unitary representation ρ on $\mathcal{H}$ leads to the amplification of biased sets. We will construct a new multiset $S' \subseteq G$ such if $\|\mathbb{E}_{s\sim S}[\rho(s)]\|_{op} \le \lambda_0$, then we have $\|\mathbb{E}_{s\sim S'}[\rho(s)]\|_{op} \le \lambda \ll \lambda_0$. Note here that the construction of S' is agnostic to ρ, and thus we can reduce the bias of all irreducible representations simultaneously! Assume that we have a graph X on the vertex set S. For $s \in S$, we have $f(s) = \rho(s)$ in this case. Let

$$S' = \{ s_t s_{t-1} \cdots s_0 \mid (s_0, s_1, \cdots s_t) \in \mathcal{W}_t \},$$

which will be our new amplified biased set. Using the homomorphism property of ρ, we have the following simplification

$$\mathop{\mathbb{E}}_{w=(s_0,\dots s_t)\in\mathcal{W}_t} [f(s_t)\cdots f(s_0)] = \mathop{\mathbb{E}}_{(s_0,\dots,s_t)\in\mathcal{W}_t} [\rho(s_t)\cdots\rho(s_0)] = \mathop{\mathbb{E}}_{s'\in S'} [\rho(s')] , \tag{2}$$

$$\text{and thus, } \operatorname{bias}(S') \le \left\| \Pi_f \left(\mathring{A}_X\, \Pi_f \right)^t \right\|_{op} \tag{3}$$

where S' is the new biased multiset of the construction and the second inequality follows from the preceding calculation when $\mathcal{W}_t$ is a collection of walks on X.

3.1 Operator Norm Decay from Constant Bias

Now that we have reduced the problem to studying the operator norm, we will study how the norm decays as we take walks. We use the decomposition, $\mathcal{X}_{\mathcal{H}} = \mathcal{X}_{\mathcal{H}}^{\|} \oplus \mathcal{X}_{\mathcal{H}}^{\perp}$ where $\mathcal{X}_{\mathcal{H}}^{\|} := \operatorname{span}\{v\otimes\vec{1} \mid v \in \mathcal{H}\}$. The decay comes from two sources. For $z \in \mathcal{X}_{\mathcal{H}}^{\perp}$, we get a decay by $\lambda(X)$ by the definition of X being an expander. Claim 3.4 shows that for $z \in \mathcal{X}_{\mathcal{H}}^{\|}$, we get a decay from Π_f, equal to the initial bias. We put this together in Theorem 3.1 to obtain the desired exponential decay.

Claim 3.4. *For $z \in \mathcal{X}_{\mathcal{H}'}^{\|}$, we have*

$$\left\|(\Pi_f\, z)^{\|}\right\|_2 \;\le\; \left\|\mathop{\mathbb{E}}_{x\in V_X}[f(x)]\right\|_{op} \cdot \|z\|_2 \,.$$

Proof. The equation trivially holds when $z = 0$, so assume $z \neq 0$ and scale it so that $\|z\|_2 = 1$. From definition of $\mathcal{X}_{\mathcal{H}'}^{\|}$, we can assume that $z = u \otimes \vec{1}$. Computing we have,

$$\begin{aligned}
\left\|\left(\Pi_f\left(u\otimes\vec{1}\right)\right)^{\|}\right\|_2 &= \sup_{w\in\mathcal{H}:\ \|w\otimes\vec{1}\|_2=1}\left|\left\langle w\otimes\vec{1}, \Pi_f\left(u\otimes\vec{1}\right)\right\rangle\right| \\
&= \sup_{w\in\mathcal{H}:\ \|w\otimes\vec{1}\|_2=1}\left|\left\langle w\otimes\vec{1}, \Pi_f\left(u\otimes\sum_{x\in V_X} x\right)\right\rangle\right| \\
&= \sup_{w\in\mathcal{H}:\ \|w\otimes\vec{1}\|_2=1}\left|\left\langle w\otimes\vec{1}, \sum_{x\in V_X}(f(x)\,u\otimes x)\right\rangle\right| \\
&= \sup_{w\in\mathcal{H}:\ \|w\otimes\vec{1}\|_2=1}\left|\sum_{x\in V_X}\langle w, f(x)\,u\rangle\left\langle\vec{1}, x\right\rangle\right| \\
&= \sup_{w\in\mathcal{H}:\ \|w\otimes\vec{1}\|_2=1}\left|\left\langle w, |V_X|\left(\mathop{\mathbb{E}}_{x\in V_X}[f(x)]\right)u\right\rangle\right| \\
&\le \left\|\mathop{\mathbb{E}}_{x\in V_X}[f(x)]\right\|_{op}|V_X|\,\|w\|\,\|u\| = \left\|\mathop{\mathbb{E}}_{x\in V_X}[f(x)]\right\|_{op}. \qquad \blacksquare
\end{aligned}$$

We show that for every two[16] steps of the walk, the norm of the (associated) operator decays as follows.

Lemma 3.5. *Let X be a $\lambda(X)$-spectral expander and let f be such that $\|\mathbb{E}_{x\in V_X}[f(x)]\|_{op} \le \lambda_0$ and $\max_{x\in V_X}\|f(x)\|_{op} \le 1$. Then,*

$$\left\|\left(\mathring{A}_X\,\Pi_f\right)^2\right\|_{op} \;\le\; 2\lambda(X) + \lambda_0\,.$$

Proof. Since $\|\Pi_f\|_{op} = \max_{x\in V_X}\|f(x)\|_{op} \le 1$, it is enough to bound $\left\|\mathring{A}_X\,\Pi_f\,\mathring{A}_X\right\|_{op}$. Let $z \in \mathcal{X}_{\mathcal{H}}$ be a unit vector which is decomposed as $z = z^{\|} + z^{\perp}$. We have

$$\begin{aligned}
\left\|\left(\mathring{A}_X\,\Pi_f\,\mathring{A}_X\right)\left(z^{\perp}+z^{\|}\right)\right\|_2 &\le \lambda(X) + \left\|\left(\mathring{A}_X\,\Pi_f\,\mathring{A}_X\right)z^{\|}\right\|_2 \\
&\le \lambda(X) + \left\|\mathring{A}_X\left(\left(\Pi_f z^{\|}\right)^{\perp} + \left(\Pi_f z^{\|}\right)^{\|}\right)\right\|_2 \\
&\le \lambda(X) + \left\|\mathring{A}_X\left(\Pi_f z^{\|}\right)^{\perp}\right\|_2 + \left\|\left(\Pi_f z^{\|}\right)^{\|}\right\|_2
\end{aligned}$$

[16] This is the source of loss of a factor of 2 in the exponent (which leads to degree $O(|S|/\lambda^{4+o(1)})$ rather than the desired degree of $O(|S|/\lambda^{2+o(1)})$ we will later achieve. Note that the same loss occurs in the original zig-zag analysis of [RVW00], which was later remedied by the s-wise zig-zag of [BATS08].

$$\leq 2\lambda(X) + \left\| \left(\Pi_f z^{\|} \right)^{\|} \right\|_2$$
$$\leq 2\lambda(X) + \lambda_0. \qquad \text{(By Claim 3.4)}$$

■

Theorem 3.1 now follows from the lemma above and the submultiplicativity of the operator norm.

3.2 Instantiating the Construction

To construct S', our construction requires an auxiliary expander graph X to perform walks on. One convenient source (among several) is a recent construction of Alon.

Theorem 3.6 (Corollary of [Alo21, Thm. 1.3])**.** *For every $n \in \mathbb{N}$, $\lambda \in (0,1)$, there exists a positive integer m_λ and an explicit construction of a graph X on $m_\lambda n$ vertices with degree at most $9/\lambda^2$ and $\lambda(X) \leq \lambda$.*

We now establish the key amplification lemma.

Lemma 3.7. *Let $S \subseteq G$ such that $\mathsf{bias}(S) = \lambda_0 < 1$. Then, for any $\lambda > 0$, we can explicitly compute S' such that $\mathsf{bias}(S') \leq \lambda$ and $|S'| = O_{\lambda_0}\left(\frac{|S|}{\lambda^{4+\delta(\lambda_0)}}\right)$.*

Proof. Pick a constant ε_0 such that $\lambda_1 := (1 + 2\varepsilon_0)\lambda_0 < 1$ and use Theorem 3.6 to obtain an explicit $(m|S|, d, \varepsilon_0\lambda_0)$-graph X. Let S_1 be the multiset consisting of m copies of S. The bias remains the same and now, $|V(X)| = |S_1|$. We construct S' by multiplying elements of t-length walks on X where $t = \lceil 2(1 + \log_{\lambda_1}(\lambda)) \rceil$. The size of S' is

$$\begin{aligned} |S'| = (m|S|) \cdot d^t &= O_{\lambda_0}(|S|) \cdot \left(\frac{3}{\varepsilon_0\lambda_0}\right)^{4\log_{\lambda_1}\lambda} \\ &= O_{\lambda_0}(|S|) \cdot \lambda^{\frac{-4\log\left(\frac{3}{\varepsilon_0\lambda_0}\right)}{\log(1/\lambda_1)}} \\ &\leq O_{\lambda_0}(|S|) \cdot \lambda^{-4\left(1+\frac{\log\left(\frac{3+6\varepsilon_0}{\varepsilon_0}\right)}{\log(1/\lambda_0)}\right)}. \end{aligned}$$

Let ρ be any irreducible representation. From Eq. (3) and Theorem 3.1, we get,

$$\left\| \mathop{\mathbb{E}}_{s_0 \cdots s_t \in S'} [\rho(s_t \cdots s_0)] \right\|_{op} \leq (2\lambda(X) + \mathsf{bias}(S))^{t/2-1} \leq (\lambda_1)^{t/2-1} \leq \lambda. \qquad ■$$

Using the amplification above, we now derive our first simplified explicit construction.

Proof of Theorem 3.2. Pick a *constant* $\lambda' < \min\left(\frac{1}{2}, \left(\frac{3}{4}\right)^{4\beta}\right)$. Use Lemma 3.7 with the target expansion $\lambda = \lambda'$ to obtain a set S_1 with size $|S_1| = O_{\lambda_0,\beta}(|S|)$ as λ' is a constant. Now use Lemma 3.7 again with S_1 as the initial set and the final expansion as λ to obtain S'. This time we fix $\varepsilon_0 = \frac{1}{2}$ in the proof of Lemma 3.7 and by our choice of λ', we have $\delta(\lambda') \leq \beta$. Thus, the final size is $|S'| \leq O_{\lambda'}\left(\frac{|S_1|}{\lambda^{4+\delta(\lambda')}}\right) \leq O_{\lambda_0,\beta}\left(\frac{|S|}{\lambda^{4+\beta}}\right)$. ■

3.3 Operator Norm Decay from any Bias

The amplification guarantee of Theorem 3.1 trivializes if $2\lambda(X) + \lambda_0 \geq 1$. Nonetheless, we now show that amplification does occur under much weaker conditions, namely, whenever $\|\mathbb{E}_{x \in V_X}[f(x)]\|_{op} < 1$ and the auxiliary graph X has expansion $\lambda(X) < 1$. This establishes that expander walks can be used to derandomize powers of an operator, itself given by an average of bounded operators, in the general case. In this derandomization, we still have an exponential norm decay, but we only "pay additional randomness" proportional to the degree of the auxiliary expander regardless of the number of operators.

Theorem 3.8 (Operator Amplification via Expander Walks (strengthening of Theorem 3.1))**.** *Let X be a $\lambda(X)$-spectral expander and let $\mathcal{W}_t$ be the collection of walks obtained from walks of length t on X. Then for any operator valued function f such that $\|\mathbb{E}_{x \in V_X}[f(x)]\|_{op} \leq \lambda_0$ and $\max_{x \in V_X} \|f(x)\|_{op} \leq 1$, we have*

$$\left\| \mathop{\mathbb{E}}_{(s_0,\ldots s_t) \in \mathcal{W}_t} [f(s_t) \cdots f(s_0)] \right\|_{op} \leq \left[1 - (1-\lambda(X))^2(1-\lambda_0)\right]^{\lfloor t/2 \rfloor} .$$

The above amplification follows from the following improved version of Lemma 3.5. The proof explores the structural *syntactic* similarity between the operator amplification and known zig-zag analysis [RVW02, Rei05, TSD18]. This regime of bias amplification was instrumental in the breakthrough SL=L result of Reingold [Rei05].

Lemma 3.9. *Let X be a $\lambda(X)$-spectral expander and let f be such that $\|\mathbb{E}_{x \in V_X}[f(x)]\|_{op} \leq \lambda_0$ and $\max_{x \in V_X} \|f(x)\|_{op} \leq 1$. Then,*

$$\left\| \left(\mathring{A}_X \, \Pi_f \right)^2 \right\|_{op} \leq 1 - (1-\lambda(X))^2(1-\lambda_0) .$$

Proof. Let $A_J = J/|V(X)|$, where J is the $|V(X)| \times |V(X)|$ all ones matrix. We can write $A_X = (1-\lambda)A_J + \lambda E$, where $\lambda = \lambda(X)$ and $\|E\|_{op} \leq 1$. Then

$$\begin{aligned} \left\| \mathring{A}_X \, \Pi_f \, \mathring{A}_X \right\|_{op} \leq\ & (1-\lambda)^2 \left\| \mathring{A}_J \, \Pi_f \, \mathring{A}_J \right\|_{op} + \lambda(1-\lambda) \left\| \mathring{E} \, \Pi_f \, \mathring{A}_J \right\|_{op} \\ & + (1-\lambda)\lambda \left\| \mathring{A}_J \, \Pi_f \, \mathring{E} \right\|_{op} + \lambda^2 \left\| \mathring{E} \, \Pi_f \, \mathring{E} \right\|_{op} . \end{aligned}$$

By Lemma 3.5 and the fact that $\lambda(A_J) = 0$, we obtain

$$\left\| \mathring{A}_J \, \Pi_f \, \mathring{A}_J \right\|_{op} \leq 2\lambda(A_J) + \lambda_0 = \lambda_0 ,$$

Recall that $\|\Pi_f\|_{op} \leq 1$ since $\max_x \|f(x)\|_{op} \leq 1$, and we also have $\|E\|_{op}, \|A_J\|_{op} \leq 1$. Then,

$$\begin{aligned} \left\| \mathring{A}_X \, \Pi_f \, \mathring{A}_X \right\|_{op} &\leq (1-\lambda)^2\lambda_0 + 2\lambda(1-\lambda) + \lambda^2 \\ &= (1-\lambda)^2\lambda_0 + 1 - (1-\lambda)^2, \\ &= 1 - (1-\lambda)^2(1-\lambda_0) , \end{aligned}$$

concluding the proof. ∎

3.4 Explicit Expanders of Small Sizes

As an application of Theorem 3.2, we demonstrate an construction of explicit Cayley expanders of sizes close to any desired n (as in Corollary 3.12). While a recent work of Alon [Alo21] gives a construction for every n, it does not have a Cayley graph structure which is convenient for us to prove Theorem 5.4. Moreover, the construction of Cayley graph as in [TS17] based on [LPS88] does not suffice for us as they only work in the regime when n is very large.

Recall that $SL_2(p)$ is the group of 2×2 invertible matrices over $\mathbb{F}_p$ with determinant 1. We obtain a base generating set for $SL_2(p)$ via the following result.

Theorem 3.10 ([Lub11])**.** *There exists an explicit generating set S (of constant size) for $SL_2(p)$ for any prime $p > 17$ such that $\lambda(\mathrm{Cay}(SL_2(p), S)) \le \lambda_0$ for some absolute constant $\lambda_0 < 1$.*

Theorem 3.11 ([Che10])**.** *For every $n \ge 2^{3 \cdot 2^{15}}$, there exists a prime in $[n, n + 4n^{2/3}]$.*

Corollary 3.12. *For any $n > 2^{9 \cdot 2^{15}}, \lambda > 0$, there is a deterministic polynomial time algorithm to construct an (n', d, λ)-graph $\mathrm{Cay}(SL_2(p), S)$, where $n' = n + O(n^{8/9})$ and $d = O(\lambda^{-4.1})$.*

Proof. Find a prime $p \in [n^{1/3} + 1, n^{1/3} + O(n^{2/9})]$, which exists due to Theorem 3.11, via brute-force search. Since, $SL_2(p)$ is a group of order $(p^2 - 1)p$, we have $n \le |SL_2(p)| \le n + O(n^{8/9})$. We use the constant-sized generating set S from Theorem 3.10 and amplify using Theorem 3.2. ∎

4 Operator Bias Reduction via the s-wide Replacement Walk

We have seen in Section 3 that bias reduction via random walks on an expander X is sub-optimal (by a factor of 2 in the exponent). We will derandomize this random walk construction to achieve an almost optimal bias reduction. The idea is to introduce a new graph Y which has a much smaller degree, and to "simulate" a random walk on X via a walk on Y. This is realized by a higher-order version of the zig-zag product [RVW00] called the *s-wide replacement product* defined by Ben-Aroya and Ta-Shma [BATS08] (see Definition 4.5).

This section establishes our key technical result which states that given any initial *operator valued function* of constant bias < 1, we amplify the bias in an almost optimal way. This generalizes the analysis of Ta-Shma [TS17] from *scalar* valued functions to *operator* valued functions.

Theorem 4.1 (Operator Generalization of Theorem 24 [TS17])**.** *Fix integers $t \ge s \ge 1$. Let X be any d_1-regular graph[17] and Y be any d_2-regular Cayley graph on $\mathbb{F}_2^{s \log d_1}$. Let $\mathcal{W}_t$ be the collection of length t walks on the s-wide replacement product of X and Y. Let $\mathcal{H}$ be a Hilbert space. For any operator valued function $f : V_X \to \mathcal{L}(\mathcal{H})$, satisfying $\max_{x \in V_X} \|f(x)\|_{op} \le 1$ and $\|\mathbb{E}_{x \in V_X}[f(x)]\|_{op} := \lambda_0 \le \lambda(Y)^2 - 2\lambda(X)$, we have*

$$\left\| \mathop{\mathbb{E}}_{(s_0, \cdots, s_t) \in \mathcal{W}_t} [f(s_0) \cdots f(s_t)] \right\|_{op} \le \left(\lambda(Y)^s + s \cdot \lambda(Y)^{s-1} + s^2 \cdot \lambda(Y)^{s-3}\right)^{\lfloor t/s \rfloor} \le O_s\left(\lambda(Y)\right)^{(1 - o_s(1))t}.$$

[17] With d_1 a power of 2.

Furthermore, the size of the collection is $|\mathcal{W}_t| = |X| \cdot d_1^s \cdot d_2^t$.

Remark 4.2. Note that there is an inherent trade-off between the spectral bound amplification (on the operator norm), and the degree bound (on the number of walks), which causes the suboptimality in how close this technique lets us approach the Ramanujan bound. As in [TS17], the $o(1)$ term we obtain from the bound above is $(1/\log(1/\lambda))^c$ for some $c > 0$ (see Theorem 4.17 for the precise computation).

The rest of this section is planned as follows. In Section 4.1, we recall the s-wide replacement product and describe random walks on it. Then, in Section 4.2, we formalize the distributions we work with and reprove the result that if Y is a Cayley graph over any product group of appropriate size[18] then it is *compatible*, i.e., it enables the transfer of pseudorandomness from Y to X. The key generalization to operator valued functions is established in Lemma 4.7 which is identical in spirit to Eq. (1). In Section 4.3, we then finish the amplification analysis in a manner similar to [TS17]. In Section 4.4, we provide details about instantiating the setup by explicitly constructing the graphs we need.

4.1 The s-wide Replacement Product and its Walks

To describe the sparse derandomized subset of walks on X from the s-wide product, we give an informal description, and then move to a formal description. Before doing so, we first recall the standard replacement product of graphs. This product takes an *outer* graph X on n vertices, which is d_1-regular, and replaces each vertex of X with a "cloud" which is a copy of an *inner* d_2-regular graph Y on the vertex set $[d_1]$. The edges within each cloud are determined by Y while the edges between clouds are based on the edges of X (and a rotation map). By taking $d_2 \ll d_1$, the replacement product yields a new graph that derandomizes the degree of X.

The s-wide replacement product generalizes this to allow $V_Y = [d_1]^s$ for any positive integer s. We will now need s rotation maps given by the operators $X_0, X_1, \ldots, X_{s-1}$ which we describe now.

The i-th operator X_i specifies one inter-cloud edge for each vertex $(v, (a_0, \ldots, a_{s-1})) \in V_X \times V_Y$, which goes to the cloud whose X component is $v_X[a_i]$, the a_i-th neighbor of v in X indexed by the i-th coordinate of the Y component. (We will discuss what happens to the Y component after taking such a step momentarily.)

Walks on the s-wide replacement product consist of steps with two different parts: an intra-cloud part followed by an inter-cloud part. All of the intra-cloud steps simply move to a random neighbor in the current cloud, which corresponds to applying the operator $I \otimes A_Y$, where A_Y is the normalized adjacency matrix of Y. The inter-cloud steps are all deterministic, with the first moving according to X_0, the second according to X_1, and so on, returning to X_0 for step number s. The operator for such a walk taking t steps on the

[18] Any product group of the form G^s with $|G| = d_1$ can be used in the s-wide construction and it satisfies this *compatible* property. Note that in Theorem 4.13 we used $G = \mathbb{F}_2^{\log_2(d_1)}$.

s-wide replacement product is

$$\prod_{i=t-1}^{0} X_{i \bmod s}(I \otimes A_Y)\,.$$

Observe that a walk on the s-wide replacement product yields a walk on the outer graph X by recording the X component after each step of the walk. Since a walk is completely determined by its intra-cloud steps, the number of t-step walks on the s-wide replacement product is,

$$|V_X| \cdot |V_Y| \cdot d_2^t = n \cdot d_1^s \cdot d_2^t \ll n \cdot d_2^t\,,$$

which therefore gives us a very sparse subset of all t-walks on X. Thus the s-wide replacement product will be used to simulate random walks on X while requiring a reduced amount of randomness (as we shall see this simulation is only possible under special conditions, namely, when we are uniformly distributed on each cloud).

We now formally define the s-wide replacement product and consider the labeling of neighbors in X more carefully. Suppose X is a d_1-regular graph. For each $x \in V_X$ and $j \in [d_1]$, let $x[j]$ be the j-th neighbor of x in X.

Definition 4.3 (Locally Invertible Rotation Map)**.** X admits a locally invertible rotation map if there exists a bijection $\varphi\colon [d_1] \to [d_1]$ such that for every $(x, j) \in V_X \times [d_1]$,

$$\text{if } x' = x[j], \text{ then, } x'[\varphi(j)] = x\,.$$

Example 4.4 (Cayley Graphs are Locally Invertible)**.** Let G be a group and $A \subseteq G$ where the set A is closed under inversion. Label the neighbors of vertices in $\mathrm{Cay}(G, A)$, by elements of A such that $g[a] = a \cdot g$. Then, $\mathrm{Cay}(G, A)$ is locally invertible as the map $\varphi\colon A \to A$ defined as $\varphi(a) = a^{-1}$ clearly satisfies the criteria,

$$\text{if } g' = g[a] = a \cdot g, \text{ then, } g'[\varphi(a)] = a^{-1} \cdot g' = g\,,$$

for every $g \in G$, $a \in A$.

Definition 4.5 (s-wide Replacement Product)**.** Suppose we are given the following:

- A d_1-regular graph X with a bijection $\varphi : [d_1] \to [d_1]$ which defines a locally invertible rotation map.
- A d_2-regular graph Y on the vertex set $[d_1]^s$.

And we define:

- For $i \in \{0, 1, \ldots, s-1\}$, we define $\mathrm{Rot}_i\colon V_X \times V_Y \to V_X \times V_Y$ such that,

  $$\mathrm{Rot}_i((x, (a_0, \ldots, a_{s-1}))) \coloneqq (x[a_i], (a_0, \ldots, a_{i-1}, \varphi(a_i), a_{i+1}, \ldots, a_{s-1}))\,,$$

  for every $x \in V_X$ and $(a_0, \ldots, a_{s-1}) \in V_Y = [d_1]^s$. (Note that the Y component of the rotation map depends only on a vertex's Y component, not its X component.)
- Denote by X_i the operator on $\mathbb{C}[V_X \times V_Y]$ which acts on the natural basis via the permutation Rot_i and let A_Y be the normalized random walk operator of Y.

Then t steps of the s-wide replacement product are given by the operator

$$X_{t-1 \bmod s} \mathbin{\mathring{A}_Y} \cdots X_{1 \bmod s} \mathbin{\mathring{A}_Y} X_{0 \bmod s} \mathbin{\mathring{A}_Y}\,.$$

4.2 The Collection of Derandomized Walks

We now describe the distribution obtained by the walks on the s-wide replacement product using the language of operators.

Recall that, in the expander walk case discussed in Section 3, we first relate the set of walks W_t to the action of the t-step walk operator (Eq. (1)) and then obtain that the task of bounding the bias reduces to bounding the operator norm (Eq. (2)). Similarly for s-wide case, we express a t-step walk in terms of a s-wide operator that act on the extended space $\mathbb{C}[V_X] \otimes \mathbb{C}[V_Y]$. Then we prove a core lemma that intuitively says: the action of t-step s-wide operator is same as the action of t-step random walk operator in an appropriate sense, whenever $t \leq s$. The scalar version of this lemma is present (Lemma 26) in [TS17] and we generalize it for operator valued functions. This generalization requires some care and appropriate notational setup. Finally, we use this lemma to show bias decay for any value of t.

Definition 4.6 (Operators and Distributions)**.** Given a tuple of random walk operators[19] $B = (B_0, \cdots, B_{t-1})$ on $\mathbb{C}[V_X] \otimes \mathbb{C}[V_Y]$ and a starting vertex $x_0 \in V_X$, we can define a distribution induced by the walk using these operators. More precisely, $\mathcal{D}(B, x_0)$ is the distribution on on $(V_X \times V_Y)^{t+1}$ such that for every $1 \leq \ell \leq t$,

$$(B_{\ell-1} \cdots B_0)\left(x_0 \otimes \frac{1}{|V_Y|}\vec{1}\right) = \mathbb{E}_{(\vec{x},\vec{y})\sim\mathcal{D}(B)} x_\ell \otimes y_\ell. \tag{4}$$

We typically suppress x_0 as it will not matter and denote $\mathcal{D}(B) = (\mathcal{D}_X(B), \mathcal{D}_Y(B))$ to specify the projections to V_X, V_Y.

The next lemma is a generalization of Eq. (1) which we need for the s-wide replacement walk. This can also be specialized to prove Eq. (1) by letting Y be a graph with one vertex (and thus $\mathcal{X}_\mathcal{H} \cong \mathcal{X}\mathcal{Y}_\mathcal{H}$). Since we now work with the tensor products of three spaces (one for the graph X, one for the graph Y, and one for the operator valued function f), we formalize the computation more explicitly. Recall that $\mathring{\Pi}_f(v \otimes x \otimes y) = f(x)\, v \otimes x \otimes y$.

Lemma 4.7 (Operator Generalization)**.** *For any tuple of random walk operators B, any operator valued f, and any $v \in \mathcal{H}$, $x_0 \in V_X$, we have*

$$\left(\mathring{B}_{t-1}\,\mathring{\Pi}_f \cdots \mathring{B}_0\,\mathring{\Pi}_f\right)\left(v \otimes x_0 \otimes \frac{1}{|V_Y|}\vec{1}\right) = \mathop{\mathbb{E}}_{(\vec{x},\vec{y})\sim D(B)}\left[f(x_{t-1}) \cdots f(x_0)\, v \otimes x_t \otimes y_t\right].$$

Proof. We prove the computation via induction on t. The base case is when $t = 1$

$$\begin{aligned}\left(\mathring{B}_0\,\mathring{\Pi}_f\right)\left(v \otimes x_0 \otimes \frac{1}{|V_Y|}\vec{1}\right) &= \mathring{B}_0\left(f(x_0)\, v \otimes x_0 \otimes \frac{1}{|V_Y|}\vec{1}\right) \\ &= \mathop{\mathbb{E}}_{(\vec{x},\vec{y})\sim D(B)}\left[f(x_0)\, v \otimes x_1 \otimes y_1\right] \quad \text{(Using Eq. (4) for } \ell = 1)\end{aligned}$$

Let $y_0 = \frac{1}{|V_Y|}\vec{1}$ and assume the statement holds for $t-1$. Then,

$$\left(\mathring{B}_{t-1}\,\mathring{\Pi}_f \cdots \mathring{B}_0\,\mathring{\Pi}_f\right)(v \otimes x_0 \otimes y_0) = \mathring{B}_{t-1}\,\mathring{\Pi}_f \cdot \prod_{i=t-2}^{0}\left(\mathring{B}_i\,\mathring{\Pi}_f\right)(v \otimes x_0 \otimes y_0)$$

[19] Markov chain operators on $V_X \times V_Y$.

$$
\begin{aligned}
&= \mathring{B}_{t-1} \mathring{\Pi}_f \mathop{\mathbb{E}}_{(\vec{x},\vec{y})\sim D(B)} [f(x_{t-2}) \cdots f(x_0) v \otimes x_{t-1} \otimes y_{t-1}] \\
&= \mathring{B}_{t-1} \mathop{\mathbb{E}}_{(\vec{x},\vec{y})\sim D(B)} [f(x_{t-1}) f(x_{t-2}) \cdots f(x_0) v \otimes x_{t-1} \otimes y_{t-1}] \\
&= \mathop{\mathbb{E}}_{(\vec{x},\vec{y})\sim D(B)} [f(x_{t-1}) \cdots f(x_0) v \otimes x_t \otimes y_t] \,.
\end{aligned}
$$

The second equality uses the inductive hypothesis and the third uses the fact that $\mathring{\Pi}_f$ acts on the tensor space diagonally. Last two equalities use Eq. (4) for $\ell = t-1$ and $\ell = t$ respectively. ■

Using Definition 4.6, we further define the operators for the distributions we wish to study.

Uniform Distribution Let us first capture using this notation the uniform distribution on walks on X starting from $x_0 \in V_X$. We define B_U where for each i, $B_i = A_X \otimes I_Y$ for every i. Then, for any ℓ, $(A_X \otimes I_Y)^\ell = A_X^\ell \otimes I_Y$. Therefore, we obtain that $\mathcal{D}_X(B_U)$ is the t-step random walk distribution on X i.e., $x_i \sim A_X^i x_0$.

The s-wide Distribution This is the distribution obtained by the s-wide walks as described in the earlier section. For $0 \le a \le b \le s$, we define

$$
B[a,b] = \left(X_a \mathring{A}_Y, X_{a+1} \mathring{A}_Y, \cdots, X_b \mathring{A}_Y \right) .
$$

We can view this random walk as occurring in two steps. The first being picking an initial vertex $y_0 \in Y$ and then, picking the sequence of neighbors according to which we will perform the walk in Y. To formalize this, let $A_Y = (1/d_2) \sum_{j=1}^{d_2} P_j$ where P_j are permutation matrices and let $J = (j_0, \cdots, j_{b-a}) \in [d_2]^{b-a+1}$. The conditional distribution, is defined by

$$
B[a,b,J] = \left(X_a \mathring{P}_{j_0}, X_{a+1} \mathring{P}_{j_1}, \cdots, X_b \mathring{P}_{j_{b-a}} \right) .
$$

We would like these two distributions to be the same and a graph Y is said to be *compatible* with respect to (X, φ), if for any fixed sequence, J, of a walk of length $\ell \le s$, the distribution obtained on X via the uniform sampling of y_0, is the same as the usual ℓ-length walk on X from any fixed initial vertex, x_0. Thus, the randomness of sampling a vertex from Y is effectively *transferred* to a random walk on X.

Definition 4.8 (Compatible)**.** A graph Y is *compatible* with respect to (X, φ) if for every $0 \le a \le b \le s$, $J \in [d_2]^{b-a+1}$ and $x_0 \in V_X$, we have[20]

$$
\mathcal{D}_X(B[a,b,J], x_0) = \mathcal{D}_X(B_U, x_0) = A_X^{b-a+1} x_0 \,.
$$

Remark 4.9. This *compatible* property is the same as 0-pseudorandom property in [TS17]. We rename it as it is more of a structural compatibility property than a pseudorandomness one.

[20] It is important to note that $\mathcal{D}_Y(B[a,b,J]) \neq \mathcal{D}_Y(B_U)$.

We now prove, for the sake of completeness, that Cayley graphs are compatible with every locally invertible graph.

Lemma 4.10 ([TS17, Lemma 29])**.** *Let* $Y = \mathrm{Cay}(G^s, T)$ *where* $|G| = d_1$*. Then,* Y *is compatible with respect to any* X, φ*.*

Proof. Since Y is a Cayley graph, we can think of $J \in S^t$ and the permutation matrices as $P_g = \rho_{\mathrm{reg}}(g)$. Recall that for any $y = (r_1, \ldots r_s) \in G^s$,

$$P_g y = gy = (g_1 r_1, \cdots, g_s r_s), \quad \text{and} \quad X_i y = (r_1, \cdots, r_{i-1}, \varphi(r_i), r_{i+1}, \cdots, r_s)\,.$$

Suppose $y = (r_1, \cdots, r_s) \sim G^s$ is now sampled uniformly. Since g_i and φ are fixed, the above operators P_g and X_i preserve the uniform distribution. Moreover, r_i is independent of r_j as $r_i \mapsto \tau_{i,k}(r_i)$ after k steps for some fixed permutation $\tau_{i,k}$ depending only on J and φ.

By definition, $x_i = x_{i-1}[\tau_{a+i-1,i}(r_{a+i-1})]$ and we take at most s steps and therefore, we use r_i for distinct $i \in [a, b]$ which are all independent. Thus, $x_i \sim A_X^i x_0$. ∎

4.3 The s-wide Operator Norm Decay

We are now ready to establish the key technical lemma in the analysis of the s-wide replacement.

Lemma 4.11 (Simulation Lemma (generalization of Lemma 26 from [TS17]))**.** *Let* $0 \le s_1 \le s_2 < s$*. For every pair of vectors* $z, z' \in \mathcal{X}_{\mathcal{H}}$*, we have,*

$$\left\langle \prod_{i=s_1}^{s_2} \left(\mathring{X}_i \mathring{A}_Y \mathring{\Pi}_f \right) \left(z \otimes \frac{1}{|V_Y|} \vec{1} \right), z' \otimes \vec{1} \right\rangle \;=\; \left\langle \left(\mathring{A}_X \Pi_f \right)^{s_2 - s_1 + 1} z, z' \right\rangle .$$

Proof. Let $z = \sum_x v_x \otimes x$ and $z' = \sum_x w_x \otimes x$. Since the expression is bilinear, it suffices to prove the equation for $v \otimes x,\ w \otimes x'$ for an arbitrary pair (x, x'). Let $t = s_2 - s_1 + 1$.

$$\prod_{i=s_1}^{s_2} \left(\mathring{X}_i \mathring{A}_Y \mathring{\Pi}_f \right) \;=\; \mathop{\mathbb{E}}_{(j_{s_1}, \cdots, j_{s_2}) \sim [d_2]^t} \left[\prod_{i=s_1}^{s_2} \left(\mathring{X}_i \mathring{P}_{j_i} \mathring{\Pi}_f \right) \right]$$

Therefore, we can fix $J = (j_{s_1}, \cdots, j_{s_2}) \in [d_2]^t$ and prove it for that. Applying Lemma 4.7 to $B[s_1, s_2, J]$, we get,

$$\begin{aligned} \prod_{i=s_1}^{s_2} \left(\mathring{X}_i \mathring{P}_{j_i} \mathring{\Pi}_f \right) \left(v \otimes x_0 \otimes \frac{1}{|V_Y|} \vec{1} \right) \;&=\; \mathop{\mathbb{E}}_{(\vec{x}, \vec{y}) \sim D(B[s_1, s_2, J])} [f(x_{t-1}) \cdots f(x_0)\, v \otimes x_t \otimes y_t] \\ &= \sum_{\vec{x} \in V_X^t} \mathop{\mathbb{E}}_{y_0 \sim V_Y} [f(\vec{x})\, v \otimes x_t \otimes y_t]\, \mathbb{I}[y_0 \text{ gives rise to } \vec{x}], \end{aligned}$$

where $f(\vec{x}) = f(x_{t-1}) \cdots f(x_0)$. The second equality uses the fact that J is fixed and we only pick the starting vertex uniformly at random which determines the entire sequence $\vec{x}, \vec{y}$. For each given $\vec{x} = (x_0, \cdots, x_t)$, there are exactly d_1^{s-t} starting vertices $y_0 = (r_1, \cdots, r_s)$

that give rise to $\vec{x}$. This is because, the only requirement is that each of the t constraints $x_i = x_{i-1}[\tau_{s_1+i-1,i}(r_{a+i-1})]$ is satisfied where $\tau_{s_1+i-1,i}$ is a fixed permutation (for a given J). Each of these equations determine one of the r_i's and therefore we have d_1^{s-t} free choices. Therefore, the conditioning on y doesn't change the distribution $\mathcal{D}_X$ and when we take inner products, we obtain

$$\left\langle \prod_{i=s_1}^{s_2} \left(\mathring{X}_i \mathring{P}_{j_i} \mathring{\Pi}_f \right) \left(v \otimes x_0 \otimes \frac{1}{|V_Y|} \vec{1} \right), w \otimes x' \otimes \vec{1} \right\rangle = \frac{d_1^{s-t}}{d_1^s} \sum_{\vec{x} \in V_X^t} \langle x_t, x' \rangle \langle f(x_{t-1}) \cdots f(x_0) v, w \rangle$$
$$= \mathop{\mathbb{E}}_{\vec{x} \sim D_X(B[s_1,s_2,J])} [\langle x_t, x' \rangle \langle f(x_{t-1}) \cdots f(x_0) v, w \rangle] .$$

We now use[21] Lemma 4.7 for B_U and take inner product to get,

$$\left\langle \left(\mathring{A}_X \Pi_f \right)^{s_2-s_1+1} (v \otimes x_0), w \otimes x' \right\rangle = \mathop{\mathbb{E}}_{\vec{x} \sim D_X(B_U)} [\langle x_t, x' \rangle \langle f(x_{t-1}) \cdots f(x_0) v, w \rangle] .$$

From Lemma 4.10, we know that Y is compatible and thus, $\mathcal{D}_X(B[s_1, s_2, J]) = \mathcal{D}_X(B_U)$. Thus, the right hand side of these two equations above are equal. ∎

The s-step Decay Just like the amplification in Section 3 was analyzed by studying the norm decay obtained in every two steps (*cf.*,Lemma 3.7), this amplification via the s-wide walks will be analyzed by bounding the norm decay for steps of length s using Lemma 4.11 similarly to [BATS08, TS17]. We will use the shorthand $L_i := \mathring{X}_i \mathring{\Pi}_f \mathring{A}_Y$.

The goal is to bound $\|L_{s-1} \cdots L_0\|_{op}$ which controls the bias of the set obtained by s-long, s-wide walks (*cf.*,proof of Eq. (2)). Equivalently, we will bound $\langle (\prod_i L_i) v_0, w_s \rangle$ for any unit vectors[22] $v_0, w_s \in \mathcal{XY}_{\mathcal{H}}$. We will use the orthogonal decomposition,

$$\mathcal{XY}_{\mathcal{H}} := \mathcal{X}_{\mathcal{H}} \otimes \mathbb{C}[V_Y] = \mathcal{XY}_{\mathcal{H}}^{\|} \oplus \mathcal{XY}_{\mathcal{H}}^{\perp} \text{ where } \mathcal{XY}_{\mathcal{H}}^{\|} := \text{span}\{z \otimes \vec{1} \mid z \in \mathcal{X}_{\mathcal{H}}\} .$$

For $i \geq 1$, we inductively define the vectors v_i, w_i, z_i and bound their norms[23],

$$v_i = L_{i-1} v_{i-1}^{\perp}, \qquad z_{s-i} = \left(\mathring{X}_{s-i} \mathring{\Pi}_f \right)^* w_{s-i+1}, \qquad w_{s-i} = \left(\mathring{A}_Y \right)^* z_{s-i}^{\perp} \quad (5)$$
$$\|v_i\| \leq \lambda(Y)^i, \qquad \|z_{s-i}\| \leq \lambda(Y)^{i-1}, \qquad \|w_{s-i}\| \leq \lambda(Y)^i . \quad (6)$$

The following lemma follows readily from a calculation and we omit its proof.

Lemma 4.12. *For any v_0, w_s and $0 \leq r \leq s-2$ we have,*

$$L_{s-1} \cdots L_0 v_0 = v_s + \sum_{i=0}^{s-1} L_{s-1} \cdots L_i v_i^{\|}$$

[21] As we only want to work with the space $\mathcal{X}_{\mathcal{H}}$ here, we can assume in the application of the lemma that $|V_Y| = 1$. Else, one could directly apply Eq. (1) and use the observation that $\mathcal{D}_X(B_U)$ is the same as the random walk distribution on X.

[22] Here we deviate from our notation and use v, w for vectors in $\mathcal{XY}_{\mathcal{H}}$.

[23] By definition $\|v_i\| \leq \left\| \mathring{A}_Y v_{i-1}^{\perp} \right\| \leq \lambda(Y) \|v_{i-1}\|$ .The computation is similar for w and z.

$$L_{s-1}^* w_s = w_{s-1} + z_{s-1}^{\|}$$
$$L_r^* \cdots L_{s-1}^* w_s = w_r + z_r^{\|} + \sum_{i=r+1}^{s-1} L_r^* \cdots L_{i-1}^* z_i^{\|}$$

Theorem 4.13 (Operator Generalization of Theorem 24 [TS17]). *Let X be any d_1-regular graph and Y be a Cayley graph on $\mathbb{F}_2^{s \log d_1}$. Let W_t be the collection of t-length s-wide walks, on the s-wide replacement product on X and Y. For any operator valued function f on V_X, such that $\max_{x \in V_X} \|f(x)\|_{op} \le 1$ and $\|\mathbb{E}_{x \in V_X}[f(x)]\|_{op} := \lambda_0 \le \lambda(Y)^2 - 2\lambda(X)$,*

$$\left\| \mathop{\mathbb{E}}_{(s_0,\cdots,s_t) \in W_t} [f(s_0) \cdots f(s_t)] \right\|_{op} \le \left(\lambda(Y)^s + s \cdot \lambda(Y)^{s-1} + s^2 \cdot \lambda(Y)^{s-3}\right)^{\lfloor t/s \rfloor}.$$

Proof. Using Lemma 4.7, we can repeat the proof of Eq. (2) to see that,

$$\left\| \mathop{\mathbb{E}}_{(x_0,\cdots,x_t) \in W_t} [f(s)x_t \cdots f(s)x_0] \right\|_{op} \le \|L_t \cdots L_0\|_{op} \le \|L_{s-1} \cdots L_0\|_{op}^{\lfloor t/s \rfloor}.$$

$$\begin{aligned}
\langle L_{s-1} \cdots L_0 v_0, w_s \rangle &= \langle v_s, w_0 \rangle + \sum_{r=0}^{s-1} \left\langle L_{s-1} \cdots L_r v_r^{\|}, w_s \right\rangle \\
&= \langle v_s, w_s \rangle + \sum_{r=0}^{s-1} \left\langle v_r^{\|}, L_r^* \cdots L_{s-1}^* w_s \right\rangle \\
&= \langle v_s, w_s \rangle + \sum_{i=0}^{s-1} \left\langle v_r^{\|}, w_r + z_r^{\|} \right\rangle + \sum_{r=0}^{s-2} \sum_{i=r+1}^{s-1} \left\langle v_r^{\|}, L_r^* \cdots L_{i-1}^* z_i^{\|} \right\rangle \\
&= \langle v_s, w_s \rangle + \sum_{i=0}^{s-1} \left\langle v_r^{\|}, z_r^{\|} \right\rangle + \sum_{r=0}^{s-2} \sum_{i=r+1}^{s-1} \left\langle v_r^{\|}, L_r^* \cdots L_{i-1}^* z_i^{\|} \right\rangle.
\end{aligned}$$

The last step uses $\left\langle v_r^{\|}, w_r \right\rangle = \left\langle \mathring{A}_Y v_r^{\|}, z_r^{\perp} \right\rangle = 0$. Using Eq. (6), we get $\left\langle v_r^{\|}, z_r^{\|} \right\rangle \le \lambda(Y)^{s-1}$. To bound the last term, we finally use Lemma 4.11. Let $v_r^{\|} = v_r' \otimes \vec{1}$, and $z_i^{\|} = z_i' \otimes \frac{1}{|V_Y|}\vec{1}$. Then,

$$\begin{aligned}
\left\langle v_r^{\|}, L_r^* \cdots L_{i-1}^* z_i^{\|} \right\rangle &= \left\langle v_r', \left(\mathring{A}_X \Pi_f\right)^{i-r} z_i' \right\rangle && \text{(Using Lemma 4.11)} \\
&\le \left\| \left(\mathring{A}_X \Pi_f\right)^{i-r} \right\|_{op} \left\| z_i' \right\| \|v_r'\| \\
&\le \lambda(Y)^{2\lfloor \frac{i-r}{2} \rfloor} \lambda(Y)^{r+s-i-1} \le \lambda(Y)^{s-3},
\end{aligned}$$

where the penultimate inequality uses Theorem 3.1 and plugs in the assumption that $2\lambda(X) + \|\mathbb{E}_{x \in V_X}[f(x)]\|_{op} \le \lambda(Y)^2$. Substituting this back in our expression above gives us the result. ∎

4.4 Instantiating the s-wide Replacement Product

Overview

The goal of this section is to explicitly construct the graphs X and Y, in order to finish the proof of Theorem 1.2. Once we obtain the graphs, we identify the vertices of X, i.e., V_X with the initial generating set[24] S. The final set is obtained by multiplying elements along each $(t-1)$-length walks on the s-wide replacement product of X and Y. We will only summarize the construction here and show that the choice of the parameters does in fact yield our main result. Detailed computation and verification is present in Appendix A.

The construction Recall that a graph is said to be an (n, d, λ)-graph if it has n vertices, is d-regular, and has second largest singular value of its normalized adjacency matrix at most λ.

- The outer graph X will be an (n', d_1, λ_1)-graph which is a Cayley graph on $SL_2(p)$ constructed using Corollary 3.12. By Example 4.4, it is locally invertible.
- The inner graph Y will be a (d_1^s, d_2, λ_2)-graph which is a Cayley graph on $\mathbb{Z}_2^m$ and therefore by Lemma 4.10, it is compatible. For this, we use the construction of Alon et al. [AGHP92], and the analysis of Ta-Shma Lemma A.1.

The parameters $n', d_1, d_2, \lambda_1, \lambda_2$ and s are chosen as follows for a fixed $\beta(\lambda)$. [25]

s is the smallest power of 2 such that $\frac{32}{\beta} \le 2^{10} \le s \le \left(\frac{\log(1/\lambda)}{4\log\log(1/\lambda)}\right)^{1/3}$

Every other parameter is a function of s.

$Y : (n_2, d_2, \lambda_2), \quad n_2 = d_2^{5s}, \quad d_2 = s^{4s}, \quad \lambda_2 \le \frac{b_2}{\sqrt{d_2}}, \quad b_2 = 5s\log d_2$

$X : (n', d_1, \lambda_1), \quad n' \approx n = O(|S|\, d_2^5), \quad d_1 = d_2^5, \quad \lambda_1 = \frac{\lambda_2^2}{10}$

t : smallest integer such that $(\lambda_2)^{(1-5\alpha)(1-\alpha)(t-1)} \le \lambda, \ ; \ $ where $\alpha = 1/s$

Now, we mention the central claim that we need from our choice of parameters. We defer its proof to Appendix A.

Claim 4.14. *The selection of the parameters above implies the following bounds on* t,

i $t - 1 \ge 2s^2$

ii $(d_2)^{(t-1)} \le \lambda^{-2(1+10\alpha)}$,

Lemma 4.15. *The number of walks of length* $t-1$ *on the* s*-wide replacement product of* X *and* Y *is* $O(|S|/\lambda^{2+\beta})$.

[24] More precisely, a slightly modified set S', obtained by duplicating and adding identities

[25] **Note:** While we let β be a function of λ, it might be instructive to make the simplifying assumption that it is an arbitrarily small constant. We will deonte it simply as β from now on.

Proof. Since each step of the walk has d_2 options, the number of walks is

$$\begin{aligned} |V(X)|\,|V(Y)| \cdot d_2^{(t-1)} &= n' \cdot d_1^s \cdot d_2^{(t-1)} = n' \cdot d_2^{(t-1)+5s} \\ &= \Theta\left(|S| \cdot d_2^{(t-1)+5s+5}\right) \\ &= O\left(|S| \cdot d_2^{(1+5\alpha)(t-1)}\right). \end{aligned}$$

which from Claim 4.14 (ii), implies a size of

$$O\left(|S| \cdot d_2^{(1+5\alpha)(t-1)}\right) = O\left(\frac{|S|}{\lambda^{2(1+10\alpha)(1+5\alpha)}}\right) = O\left(\frac{|S|}{\lambda^{2+32\alpha}}\right) = O\left(\frac{|S|}{\lambda^{2+\beta}}\right). \qquad \blacksquare$$

Before we prove the main result, we need the following simple observation.

Lemma 4.16. *Let S be an ε-biased set of a group G. And let S' be obtained by adding $\theta\,|S|$ many identity elements. Then, S' is an $(\varepsilon+\theta)$-biased set.*

Proof. Denote by e the identity element of G. Let ρ be any non trivial irreducible representation of a group G. Computing we have

$$\begin{aligned} \left\|\mathbb{E}_{s\in S'}\rho(s)\right\|_{op} &= \frac{1}{1+\theta}\left\|\mathbb{E}_{s\in S}\rho(s) + \theta\cdot\mathbb{E}_{s\in S\setminus S'}\rho(e)\right\|_{op} \\ &\le \left\|\mathbb{E}_{s\in S}\rho(s)\right\|_{op} + \theta && (\|\rho(e)\|_{op} = 1) \\ &\le \varepsilon + \theta && (S \text{ is } \varepsilon\text{- biased}) \qquad \blacksquare \end{aligned}$$

Theorem 4.17 (Almost Ramanujan Expanders I)**.** *Let $\mathrm{Cay}(G,S)$ be λ_0-expander with constant $\lambda_0 \in (0,1)$. For every function*[26] *$\beta(\lambda) > 0$, and for any $\lambda > 0$, sufficiently small such that*

$$\frac{32}{\beta(\lambda)} \le \left(\frac{\log(1/\lambda)}{4\log\log(1/\lambda)}\right)^{1/3},$$

there exists a deterministic polynomial time algorithm to construct S' such that $\mathrm{Cay}(G,S')$ is a λ-expander with degree $|S'| = O_{\lambda_0}(|S|/\lambda^{2+\beta})$.

Furthermore, each element in S' is the product of $O(\log(1/\lambda))$ elements of S.

Proof. We can assume that $s \ge 2^{10}$ since otherwise λ is a constant and we can just use Theorem 3.2.

Initial Boost We first boost the expansion from λ_0 to $1/d_2 \le \lambda_2^2/3$. Using Theorem 3.2 (with its parameter β equal to 1), we can find a new set of generators, S_1, such that $\mathrm{Cay}(G,S_1)$ is $1/d_2$-spectral expander and $|S_1| = O(|S|\,d_2^5)$. Moreover, we also know that, each element in S_1 is a multiple of at most $\log\left(d_2^5\right)$ elements in S. We add multiple copies of the entire set to make the size $|S|\,d_2^5$.

[26] For a first reading, it may be helpful to assume that β is a very small but fixed constant not depending on λ. Since each of the parameters depend on β, they all become constants under this assumption.

The s-wide walk Obtain an (n', d_1, λ_1) Cayley graph X from Corollary 3.12 as explained before. We add $n' - n = O(n^{8/9})$ copies of the identity to S_1 to obtain S_2. By Lemma 4.16 and the assumption that $s \geq 2^{10}$, S_2 is a $\lambda_2^2/3 + O(n^{-1/9}) \leq 2\lambda_2^2/3$-biased set. We denote by S' the final set of generators obtained by t steps of the s-wide replacement product of X and Y. By definition, each element in S' is a product of t elements in S_2 which has the same elements as S_1. Thus, each element in S' is a product of at most

$$\begin{aligned} O(t \log(d_2)) &\leq O((1 + 10\alpha)\log(1/\lambda)) && \text{(Using Claim 4.14 [ii])} \\ &\leq O(\log(1/\lambda)) && \text{(By the assumption that } \alpha \leq 1/128) \end{aligned}$$

elements of S. The only thing that remains is to prove expansion of $\mathrm{Cay}(G, S')$. We pick any irreducible representation ρ and apply Theorem 4.13 to the function ρ on $S_2 \leftrightarrow V(X)$. The condition that $2\lambda(X) + \left\|\mathbb{E}_{g\sim S_2}[\rho(g)]\right\|_{op} \leq \lambda(Y)^2$ translates to $\lambda_1 \leq \lambda_2^2/6$ which is satisfied by our choice of λ_1. Thus, the final expansion is given by,

$$\begin{aligned} \left\|\mathop{\mathbb{E}}_{g\in S'}[\rho(g)]\right\|_{op} &:= \left(\lambda_2^s + s\cdot\lambda_2^{s-1} + s^2\cdot\lambda_2^{s-3}\right)^{\lfloor (t-1)/s\rfloor} \\ &\leq \left(3s^2\lambda_2^{s-3}\right)^{((t-1)/s)-1} && \left(\text{Using } \lambda_2 \leq \frac{20s^2\log s}{s^{2s^2}} \leq \frac{1}{3s^2}\right) \\ &\leq \left(\lambda_2^{s-4}\right)^{(t-1-s)/s} \\ &\leq \lambda_2^{(1-5/s)(1-s/(t-1))(t-1)} \\ &\leq \lambda_2^{(1-5\alpha)(1-\alpha)(t-1)} && \text{(Using Claim 4.14 [i])} \\ &= \lambda_2^{(1-5\alpha)(1-\alpha)(t-1)} \leq \lambda, && \text{(From the choice of } t) \quad \blacksquare \end{aligned}$$

5 Some Applications

Our operator amplification leads to almost optimal explicit constructions of many pseudorandom objects (from existing suboptimal ones): transforming arbitrary expander graphs into almost-Ramanujan expanders (Section 5.2), quantum expanders (Section 5.3), monotone expanders (Section 5.4), to generating sets with improved (average) Kazhdan constants (Section 5.5) and to dimension expanders (Section 5.6). These pseudorandom objects embody various notions of expansion.

Permutation Amplification The key to these applications is observing that the adjacency matrix of an arbitrary graph and that of a monotone expander can be written as a sum of permutation matrices which can be interpreted as $P_\sigma = \rho_{\mathrm{def}}(\sigma)$ for the *defining* (or *natural*) representation ρ_{def}. We plug in the collection of these permutations $\{\sigma\}$ in our amplification machinery to obtain almost optimal spectral expanders and monotone expanders.

Almost Ramanujan Expanders for the Symmetric Group Constructing constant size expanding generating sets for the symmetric group was quite challenging (even non-explicitly). In a breakthrough work [Kas07], Kassabov provided the first family of such

expanding generators which was also explicit. However, this family was not close to the Ramanujan bound and no such generating set was known. Theorem 1.2 lets us amplify Kassabov's generating set to one close to optimum bound showing that the symmetric group has explicit almost Ramanujan Cayley expanders. The same obviously holds for every expanding group.

Quantum Expanders A quantum expander is a generalization of an expander graph having many applications in quantum information theory [AS04, BASTS08, Has07b, Has07a, HH09, AHL+14]. Harrow [Har07] showed that Cayley graphs can be used to construct quantum expanders inheriting the expansion of the starting Cayley graph. However, the construction is only explicit if the group admits an efficient quantum Fourier transform (QFT). Since we can now obtain almost Ramanujan Cayley graphs for the symmetric group which has a known efficient QFT [Bea97], we obtain the first explicit almost Ramanujan quantum expanders.

Improving the Kazhdan Constant The *Kazhdan constant* $\mathcal{K}(G, S)$ of a finitely generated group G, with respect to a generating set S, is a quantitative version of Property (T) which has been used to construct explicit expanders (e.g., Margulis [Mar88]). We show that this can be amplified by considering a slightly different version called the *average Kazhdan constant* which directly relates to the bias of the set S. This is interesting as typically the bound on the Kazhdan constant is used to construct expanders but here we construct expanding generating sets to improve the constant! The improved constants and the generating sets have algorithmic implications and we mention two of them.

- *Dimension expanders* - Lubotzky and Zelmanov [LZ08] showed that the image of a generating set of a group under an irreducible representation gives a dimension expander and its expansion is controlled by its Kazhdan constant.
- *Product replacement algorithm* - uses random walks on k-tuples of groups elements. Lubotzky and Pak [LP00] showed that the mixing time of the algorithm relates to the Kazhdan constant (assuming Property (T)) of certain lattice groups like $\mathrm{SL}_n(\mathbb{Z})$. This crucial assumption was proven in complete generality[27] recently by Kaluba, Kielak and Nowak [KKN21]. In particular, we have a mixing time bound of $\frac{4 \log |G|}{\mathcal{K}(G,S)^2}$.

Using our amplified generating set (Corollary 5.14), we can improve both these results.

Sampling Group elements Another application of having almost optimal Ramanujan Cayley graphs is to sample random group elements efficiently. Given a Cayley graph, $\mathrm{Cay}(G, S)$, one can consider a random walk on G which starts at an arbitrary vertex g and at each step moves to a random neighbor $g \to sg$. Spectral expansion guarantees that walks mix quickly, i.e., in at most $O_\lambda(\log |G|)$ steps (See [HLW06]). The amount of randomness used in each step is $\log d$ and since the degree versus expansion trade-off is now almost

[27] In general, we have quotients of $\mathrm{Aut}(F_n)$, the automorphism group of the free group generated by n elements and [KKN21] proves that $\mathrm{Aut}(F_n)$ has Property (T).

optimal, we can achieve the same convergence guarantee using a smaller degree and thus the random walk is more efficient in terms of randomness.

5.1 Permutation Amplification

The *defining representation* - $(\rho_{\text{def}}(\sigma), \mathbb{C}^n)$ for Sym_n is defined as the representation that maps a permutation to the matrix defining it. More formally, $\rho_{\text{def}}(\sigma)e_i = e_{\sigma(i)}$ for every unit basis vector e_i of $\mathbb{C}^n$. It is a fact that $\mathcal{V}_{\text{def}} = \mathcal{V}_{\text{triv}} \oplus \mathcal{V}_{\text{standard}}$ where $\mathcal{V}_{\text{standard}}$ is an irreducible non-trivial representation. Note that if we are given a set $\{P_1, \cdots, P_r\}$ of permutation matrices acting on $\mathbb{C}^n$, we can identify a set $S = \{\sigma_1, \cdots, \sigma_r\} \subseteq \mathrm{Sym}_n$ such that $\rho_{\text{def}}(\sigma_i) = P_i$.

Corollary 5.1 (Permutation Amplification)**.** *Let $P = \{P_1, \cdots, P_r\}$ be a collection of permutation matrices such that $\lambda(\mathbb{E}_{i\sim[r]}[P_i]) \leq \lambda_0$. Then, for any $\lambda \in (0,1)$, we can explicitly construct a collection P' such that*

1. *$\lambda(\mathbb{E}_{M\sim P'}[M]) \leq \lambda$,*
2. *$|P'| \leq O\left(|P|/\lambda^{2+o(1)}\right)$ and*
3. *each $P'_i \in P'$ is a product of at most $O_{\lambda_0}(\log(1/\lambda))$ many matrices from P.*

Proof. Let $P_i = \sigma_i$. Applying Theorem 4.13 to the set $S = \{\sigma_i\}$ we get a larger set of permutations, S' of the form $\sigma' = \sigma_{i_1} \circ \cdots \circ \sigma_{i_k}$ where $k = O_{\lambda_0}(\log(1/\lambda))$. By the decomposition of the defining representation, we have that

$$\begin{aligned}\mathrm{Spec}\left(\mathop{\mathbb{E}}_{M\sim P'}[M]\right) &= \mathrm{Spec}\left(\mathop{\mathbb{E}}_{\sigma'\sim S'}[\rho_{\text{def}}(\sigma')]\right) \\ &= \{1\} \cup \mathrm{Spec}\left(\mathop{\mathbb{E}}_{\sigma'\sim S'}[\rho_{\text{standard}}(\sigma')]\right).\end{aligned}$$

where the 1 corresponds to the eigenvalue from the trivial representation. Since the operator amplification reduces the bias of every non-trivial irreducible representation, it also does so for $\mathcal{V}_{\text{standard}}$. ∎

5.2 Arbitrary Expanders via Permutation Amplification

We can make any family of bounded degree expander graphs into an almost Ramanujan family while preserving their adjacency structure. First, we recall König's theorem that says that the adjacency matrix of a d-regular graph can be expressed in terms of permutation matrices.

Theorem 5.2 (König)**.** *Let A_X be normalized adjacency matrix of a d-regular n-vertex simple graph X. Then, there exists d permutation matrices $P_1, \ldots, P_d \in \mathbb{R}^{n\times n}$ such that*

$$A_X = \frac{1}{d}\sum_{j=1}^{d} P_j.$$

It is also efficient to find permutation matrices as above.

Claim 5.3. *The permutations in Theorem 5.2 can be found in time* $\mathrm{poly}(n)$.

Proof. We view A_X as encoding the adjacency relation of a bipartite graph with vertex bipartition $(A = V(X), B = V(X))$. This bipartite graph is d-regular so it has at least one perfect matching M, which can be found in $\mathrm{poly}(n)$ time. We remove this matching M obtaining a $(d-1)$-regular graph and we repeat till the resulting graph is empty. ∎

Our general transformation into an almost Ramanujan bound follows by using Claim 5.3 to obtain an initial set of permutation matrices which are amplified using Corollary 5.1.

Theorem 5.4 (Main I (Formal version of Theorem 1.1))**.** *Let* $\{X_i\}_{i\in\mathbb{N}}$ *be a family of* d_0*-regular* λ_0*-expanders with constant* $\lambda_0 < 1$*. For any* $\lambda \in (0,1)$ *and any expander* X_i*, we can deterministically compute a* d*-regular* λ*-expander* X'_i *with* $d = O_{\lambda_0}(d_0/\lambda^{2+o(1)})$ *in time* $\mathrm{poly}(|V(X_i)|)$*. Moreover, the construction is local in the sense that edges in* X'_i *correspond to short walks in* X_i*. More precisely, if the adjacency matrix of* X_i *is*

$$A_{X_i} = \frac{1}{d_0}\sum_{j=1}^{d_0} P_j,$$

where $P_1, \ldots, P_{d_0}$ *are permutation matrices, then the adjacency matrix of* X'_i *is*

$$A_{X'_i} = \frac{1}{d}\sum_{j=1}^{d} P'_j,$$

where each P'_j *is the product of at most* $k = O_{\lambda_0}(\log(1/\lambda))$ *permutation matrices among* $P_1, \ldots, P_{d_0}$.

5.3 Explicit Almost Ramanujan Quantum Expanders

Quantum expanders were defined in [AS04, BASTS08, Has07a] and have found many applications in quantum information theory. For instance, they can be used in the construction of designs and gates sets [HH09], in quantum statistical zero knowledge (QSZK) [BASTS08], in detecting EPR pairs [AHL+14] and in the study of *entanglement* [Has07b].

Roughly speaking, a quantum expander is a generalization of an expander graph (see Definition 5.5 for precise details). While a usual degree-d expander graph $X = (V, E)$ is given by d permutation matrices acting on a vector space $\mathbb{C}[V]$, a quantum expander is given by d (suitable) linear operators acting on quantum states (i.e., PSD matrices of trace 1). The normalized adjacency matrix of a λ-expander shrinks the ℓ_2-norm of vectors orthogonal the all ones function by a factor of λ. Similarly, a quantum expander shrinks the Frobenius norm of PSD matrices orthogonal [28] to the identity matrix (the quantum analogue of the all ones function) by a factor of λ (the quantum expansion parameter).

In [Has07c], Hastings showed that the Ramanujan bound also applies to quantum expanders and that d random unitaries get arbitrarily close to the bound. However, such a

[28] With respect to the Hilbert–Schmidt inner product.

construction cannot be efficiently implemented and thus used in applications like [AHL+14] which rely on existing explicit constructions (e.g., [BASTS08, Har07]) that are far from the Ramanujan bound and thus give sub-optimal results.

We deduce the existence of explicit families of almost Ramanujan quantum expanders by applying our amplification of Cayley graphs together with a result of Harrow [Har07]. For this, it is important that we can efficiently construct almost Ramanujan Cayley expanders on the symmetric group Sym_n, for which efficient Quantum Fourier Transform (QFT) is known [Bea97].

Definition 5.5 (Quantum Expander [AHL+14])**.** The (super) operator $\Phi : \mathbb{C}^{N\times N} \to \mathbb{C}^{N\times N}$ is an (N, d, λ) quantum expander if

- ("Degree") The operator Φ can be expressed as a sum of d linear operators as follows, $\Phi(\rho) = \sum_{i=1}^{d} B_i \rho B_i^\dagger$ where[29] $\sum_{i=1}^{d} B_i^\dagger B_i = I_N$.
- ("Expansion") The second largest eigenvalue[30] of Φ as a linear map is $\leq \lambda$.

Theorem 5.6 (Harrow [Har07])**.** *Let G be a group and $S \subseteq G$ be a multiset such that $\mathrm{Cay}(G, S)$ is a λ-spectral expander. Let V^μ be an irreducible representation of G of dimension N. Then, there exists an $(|S|, \lambda)$-quantum expander of dimension N. Furthermore, if the group G admits an efficient QFT and $\log N = \Omega(\log |G|)$, then the quantum expander is explicit.*

As a corollary of Harrow's result and our explicit family of almost Ramanujan Cayley expanders over the symmetric group obtained from the expanding family of Kassabov [Kas07], we deduce the following corollary.

Corollary 5.7 (Explicit Almost Ramanujan Quantum Expanders)**.** *For every $\lambda \in (0, 1)$, there is an explicit infinite family of (efficient) $(O(1/\lambda^{2+o(1)}), \lambda)$-quantum expanders.*

5.4 Explicit Almost Ramanujan Monotone Expander

We now show how to obtain almost Ramanujan monotone expanders starting from the explicit construction in Bourgain and Yehudayoff [BY13]. Monotone expanders are dimension expanders over any field as observed by Dvir and Shpilka [DS09, DW10]. First, we recall the definition of a monotone graph.

Definition 5.8 (Monotone Graph)**.** A bipartite graph $X = ([n]_A \sqcup [n]_B, E)$ is a d-monotone graph if there are d partial monotone maps $f_1, \ldots, f_d$ with domain and images in $[n]$ (as an ordered set [31]) such that the edges set E is the following disjoint union

$$E = \bigsqcup_{i=1}^{d} \{(v_A, f_i(v)_B) \mid v \in \mathrm{Domain}(f_i)\}.$$

[29] A useful special case is when each B_i is a (normalized) unitary.

[30] If ρ satisfies $\mathrm{Tr}(\rho) = 0$, then $\|\Phi(\rho)\|_2 \leq \lambda \|\rho\|_2$, where $\|\rho\|_2 := \sqrt{\mathrm{Tr}(\rho^\dagger \rho)}$.

[31] Under the natural order, i.e., $1 \leq 2 \leq \cdots \leq n$.

We observe that there are two notions of degree of a monotone graph: the usual vertex degree and the number of monotone functions. Clearly, if a graph is d-monotone, all vertex degrees are at most d. The converse is not necessarily true (e.g., every bipartite graph $X = (V, E)$ is $|E|$-monotone – it is important to keep this parameter constant). We stress that our almost Ramanujan bound is with respect to the usual notion of vertex degree (and keeps the number of monotone maps polynomial in the vertex degree).

Definition 5.9 (Monotone Vertex Expander)**.** We say that $X = (A = [n]_A \sqcup B = [n]_B, E)$ is a d-monotone expander if it is a d-monotone graph and there exists $\delta > 0$ such that for all $A' \subseteq A$ with $|A| \leq n/2$, we have $|\partial(A')| \geq (1+\delta)|A'|$, where $\partial(A')$ is the set of vertices in B adjacent to A'.

Theorem 5.10 (Bourgain and Yehudayoff [BY13])**.** *There is an explicit family $\{X_n\}_{n\in\mathbb{N}}$ of d-monotone vertex expanders with $d = \Theta(1)$.*

We will work with a spectral definition of monotone expander.

Definition 5.11 (Spectral Monotone Expander)**.** Let $X = (A = [n]_A \sqcup B = [n]_B, E)$ be a d-monotone graph. We define A_X to be the adjacency matrix of X when the two vertex partitions are identified (as $x_A = x_B$ for $x \in [n]$) and define $\lambda(X) = \max\{|\lambda_2(A_X)|, |\lambda_n(A_X)|\}$.

It is well-known that starting from a monotone expander (not necessarily a vertex regular graph), we can add partial monotone functions to obtain a monotone graph of regular (vertex) degree that is still expanding. We use this to establish the following,

Corollary 5.12. *There is explicit family $\{X_n\}_{n\in\mathbb{N}}$ of d_0-regular $2d_0$-monotone expanders with $\lambda(X_n) \leq \lambda_0 < 1$ and $d_0 = \Theta(1)$. Furthermore, the unormalized adjacency matrix of X_n can be written as a sum of d_0 permutation matrices each corresponding to two monotone maps.*

Proof Sketch: Let $\{X'_n\}_{n\in\mathbb{N}}$ be the family in Theorem 5.10. Let $X = X'_n$ be a fixed d_0-regular monotone expander with the maps $\{f_i\}$.

For each monotone function f_i, we define its "complement", $\overline{f}_i$, as the (unique) partial monotone function $\overline{f}_i$ such that $f_i \cup \overline{f}_i$ is a total function. Let Y be the $2d_0$-monotone graph corresponding to the maps $\{f_i, \overline{f_i}\}$. Then, its adjacency can be written as as follows

$$A_Y = \sum_{i=1}^{d_0} P_i\,,$$

where $P_i = M_{f_i} + M_{\overline{f_i}}$ and $(M_{f_i})_{x,y} = \mathbb{1}[f_i(x) = y]$.

Each matrix P_i is a permutation matrix as $f_i \cup \overline{f}_i$ is a total function. Adding more maps preserves the constant vertex expansion parameter which (together with having constant vertex degree) implies constant spectral expansion bounded away from 1 (see [Vad12, Theorem 4.19]). Thus, $\{Y_n\}_{n\in\mathbb{N}}$ is the required family. □

In the amplification process, we will be multiplying permutation matrices rather than just composing monotone maps since the latter operation can result in a map with empty domain. We now establish the derandomized spectral amplification of monotone expanders.

Corollary 5.13 (Almost Ramanujan Monotone Expanders)**.** *For every $\lambda > 0$, there is an explicit family $\{X_i\}_{i\in\mathbb{N}}$ of (vertex) d-regular $d^{O(1)}$-monotone expanders with $d = O(1/\lambda^{2+o(1)})$ and $\lambda(X_i) \le \lambda$.*

Proof. Let $\{X'_n\}_{n\in\mathbb{N}}$ be the family in Corollary 5.12. Fix $X = X'_n$ and let $P_1, \dots, P_{d_0} \in \mathbb{R}^{n\times n}$ be the permutation matrices guaranteed by Corollary 5.12, where each $P_i = M_{f_i} + M_{\overline{f}_i}$. Use Corollary 5.1 to obtain a collection of $d := O(1/\lambda^{2+\beta})$ permutation matrices each of which is a product of k permutation matrices from $P_1, \dots, P_{d_0}$ and so we obtain

$$\begin{aligned} P_{i_1} \cdots P_{i_k} &= \sum_{g_i \in \{f_i, \overline{f_i}\}} M_{g_{i_1}} \cdots M_{g_{i_k}} \\ &= \sum_{g_i \in \{f_i, \overline{f_i}\}} M_{g_{i_1} \circ g_{i_2} \circ \cdots \circ g_{i_k}}, \end{aligned}$$

where $g_{i_1} \circ g_{i_2} \circ \cdots g_{i_k}$ is the composed map which is monotone (possibly with empty domain). This means that we can have at most 2^k monotone maps (and at least one since $P_{i_1} \cdots P_{i_k} \neq 0$). Therefore, the total number of maps is at most $d \cdot 2^k = d^{O(1)}$ as $k = O_{\lambda_0}(\log(1/\lambda))$. This can be made undirected by adding f^{-1} for each f and thereby doubling the degree. ∎

5.5 Amplifying the Average Kazhdan Constant

The *Kazhdan constant* is a notion of "spectral gap" (and so it is related to bias) for discrete groups which predates and was central to the study of expansion in finite groups and graphs. These groups can have infinitely many irreducible representations on more general Hilbert spaces, possibly of infinite dimension. Nonetheless, we can still apply our operator version of Ta-Shma's amplification procedure as it is independent of dimension and works for any unitary representation ρ. Therefore, we amplify the average Kazhdan constant which also amplifies the Kazhdan constant. We now define these two parameters formally.

Let G be a discrete group generated by a finite set S of generators. The Kazhdan constant of G with respect to generators S is defined as

$$\mathcal{K}(G, S) := \inf\{\mathcal{K}(G, S, \rho) \mid (\rho, \mathcal{H}) \text{ irreducible and non-trivial}\},$$

where $\mathcal{K}(G, S, \rho) := \inf_{v\in\mathcal{H}:\, \|v\|_2=1} \max_{g\in S} \|\rho(g)v - v\|_2^2$.

Analogously, an average version of the Kazhdan constant, as in the work of Pak and Zuk [PZ01], can be defined as

$$\begin{aligned} \overline{\mathcal{K}}(G, S) &:= \inf\{\overline{\mathcal{K}}(G, S, \rho) \mid (\rho, \mathcal{H}) \text{ irreducible and non-trivial}\} \\ \overline{\mathcal{K}}(G, S, \rho) &:= \inf_{v\in\mathcal{H}:\, \|v\|_2=1} \frac{1}{|S|} \sum_{g\in S} \|\rho(g)v - v\|_2^2 \\ &= \inf_{v\in\mathcal{H}:\, \|v\|_2=1} \frac{1}{|S|} \sum_{g\in S} 2 - 2\langle \rho(g)v, v\rangle \end{aligned}$$

$$= \inf_{v \in \mathcal{H}:\, \|v\|_2 = 1} 2 - 2 \left\langle \mathop{\mathbb{E}}_{g \sim S} [\rho(g)]\, v, v \right\rangle$$
$$= 2 \left(1 - \left\| \mathop{\mathbb{E}}_{g \sim S} [\rho(g)] \right\|_{op} \right).$$

Theorem 1.2 gives an improved generating set in this more general setting.

Corollary 5.14 (Amplifying Average Kazhdan Constant)**.** *Let* G *be a discrete group and* S *a finite set of generators such that the average Kazhdan constant* $\overline{\mathcal{K}}(G, S)$ *is equal to* $2 \cdot (1 - \lambda_0)$ *for some constant* $\lambda_0 \in (0, 1)$*. For every* $\lambda \in (0, 1)$*, there is a set* $S' \subseteq G$ *such that*

1. $\overline{\mathcal{K}}(G, S') \geq 2 \cdot (1 - \lambda)$*, and thus,* $\mathcal{K}(G, S') \geq 2 \cdot (1 - \lambda)$*.*
2. $|S'| = O_{\lambda_0}(|S| / \lambda^{2+o(1)})$*, and*
3. S' *can be found in time* $\mathrm{poly}(|S| / \lambda)$ *assuming an oracle for group operations on* G*.*

Remark 5.15. Note that the above amplification for $\overline{\mathcal{K}}$ immediately implies the same amplification for $\mathcal{K}$ (since the maximum is at least the average). Moreover, we remark that the above amplification can also similarly improve the constant of Lubotzky's property (τ) (the latter being a weaker version of property (T)), so it is more general and applies to expansion in many more discrete groups [RL10].

In Section 5.6, we will apply this corollary to a specific family of representations which will give a simple improvement to the bounds on the dimension expander constructed in [LZ08].

5.6 Explicit Almost Ramanujan Dimension Expanders

Dimension expanders were defined in [BISW01] motivated by applications in theoretical computer science. A conjectured construction based on irreducible representations was suggested by Wigderson to hold over every field. The conjecture was subsequently established by Lubotzky and Zelmanov [LZ08] for fields of characteristic zero. We now define dimension expanders, explain the [LZ08] proof, and our amplification in this setting.

Definition 5.16 (Dimension Expander [LZ08])**.** Let $\mathbb{F}$ be a field, $d \in \mathbb{N}$, $\varepsilon > 0$, V be a vector space of dimension n and $T_1, \ldots, T_d : V \to V$ be linear transformations. We say that $(V, \{T_i\}_{i \in [d]})$ is an ε-dimension expander if for every subspace $W \subseteq V$ of dimension at most $n/2$, we have $\dim(W + \sum_{i=1}^{d} T_i(W)) \geq (1 + \varepsilon) \cdot \dim(W)$.

Remark 5.17. Observe that if the maps T_i are restricted to being permutation matrices, and the expansion condition is restricted only to subspaces W generated by elementary basis vectors, then one obtains the usual definition of vertex expansion of graphs. Thus dimension expanders may be viewed as a linear-algebraic extension of expander graphs.

For an irreducible unitary representation ρ, there exists an associated representation[32] adj_ρ. The construction in [LZ08] relates dimension expansion with the Kazhdan constant as follows.

[32] Let $\mathfrak{sl}_n(\mathbb{C}) = \{\mathrm{tr}(A) = 0 \mid A \in M_n(\mathbb{C})\}$. Equip the space with the Frobenius inner product defined as $\langle A, B \rangle = \mathrm{tr}(A^\dagger B)$ where $A^\dagger$ is the conjugate transpose. For any finite dimensional unitary representation $\rho : G \to \mathbb{U}_n$, we have an adjoint representation $(\mathrm{adj}_\rho, \mathfrak{sl}_n)$ where the action is by conjugation $\mathrm{adj}_\rho(g) \cdot A =$

Proposition 5.18 (Adapted from [LZ08]). Let $\rho\colon G \to \mathbb{U}_{\mathbb{C}^n}$ be a unitary irreducible representation. Then $(\mathbb{C}^n, \{\rho(g)\}_{g\in S})$ is ε-expander, where $\varepsilon = (1/2 - o(1)) \cdot \mathcal{K}(G, S, \mathrm{adj}_\rho)$ (if we additionally assume $\dim(W)$ is sufficiently small).

By definition, $\mathcal{K}(G, S, \mathrm{adj}_\rho) \geq \mathcal{K}(G, S)$ and therefore for a group G which satisfies the condition of Corollary 5.14, we obtain a set S' (at the expense of restricting the dimension of W) such that $\mathcal{K}(G, S', \mathrm{adj}_\rho) \geq 2(1-\varepsilon)$ for any $\varepsilon > 0$. Which we can get ε arbitrarily close to 1 in the definition of dimension expander. In fact, we need another simple improvement to a computation in [LZ08] which we state without proof.

Claim 5.19. *Let $W, W' \subseteq \mathbb{C}^d$ be two vector spaces. Let P, P' be orthogonal projectors onto W, W', respectively. Then,*

$$\mathrm{Re}\left(\mathrm{Tr}(PP')\right) = \mathrm{Tr}(PP') \geq \dim(W \cap W').$$

With the above claim and the analysis in [LZ08], we obtain stronger dimension expansion for small dimensional spaces.

Remark 5.20. Forbes and Guruswami [FG15] point out that the quantum expander construction of Harrow [Har07] also yields a dimension expander (with a similar construction of the dimension expanders from [LZ08]). As mentioned earlier, monotone expanders are dimension expanders over any field [DS09, DW10]. Moreover, the Bourgain and Yehudayoff [BY13] construction of monotone expanders with constant generating set yields such dimension expanders with constant generating set!

5.7 Diameter of Finite Groups

The study of the diameter of Cayley graphs can take many forms, e.g., it can be with respect to every generating set (as in the celebrated Babai–Seress conjecture [BS88]) or with respect to some constant size generating set as in [BKL89]. Here, we explore the latter case.

First, recall that any n-vertex degree-d graph has diameter at least $\log_{d-1}(n)$. On the other hand, it is well-known that expansion directly implies diameter at most $C \cdot \log_{d-1}(n)$ for some constant $C \geq 1$ (depending on the expansion).

Using the operator amplification, we deduce that any expanding group G has a constant degree-d Cayley expander of diameter $\approx 2 \cdot \log_{d-1}(|G|)$. More precisely, we have the following.

Lemma 5.21. *Suppose $\{\mathrm{Cay}(G_i, S_i)\}_{i\in\mathbb{N}}$ is a family of bounded degree Cayley expanders. Then, there exists a family $\{\mathrm{Cay}(G_i, S_i')\}_{i\in\mathbb{N}}$ of constant degree-d Cayley expanders with diameter at most $(2 + o_d(1)) \cdot \log_{d-1}(G_i)$.*

Proof. We apply Theorem 1.2 to the family $\{\mathrm{Cay}(G_i, S_i)\}_{i\in\mathbb{N}}$ obtaining a new family of $\{\mathrm{Cay}(G_i, S_i')\}_{i\in\mathbb{N}}$ of (d, λ)-expanders with $d = 1/\lambda^{2+\beta}$ for some sufficiently small constants

$\rho(g)\cdot A\cdot\rho(g)^{-1}$. Since conjugation by unitary matrices preserves the trace, $\mathfrak{sl}_n$ is closed under the representation. Moreover, it is unitary as

$$\left\langle \mathrm{adj}_\rho(g)A, \mathrm{adj}_\rho(g)B \right\rangle = \mathrm{tr}(\rho(g)A^\dagger\rho(g)^\dagger\rho(g)B\rho(g)^{-1}) = \langle A, B\rangle .$$

$\lambda, \beta > 0$. Let A_i be the normalized adjacency matrix of $\mathrm{Cay}(G_i, S_i')$ and $n_i = |G_i|$. Let e_g be the indicator vector of some fixed $g \in G_i$. Note that

$$\left\|(A_i - J/n_i)^t e_g\right\|_2 \leq \lambda^t = d^{-t/(2+\beta)} < 1/|G_i|\,,$$
$$\text{for } t = (2+2\beta)\cdot \log_d(|G_i|) = (2 + o_{d,\beta}(1))\cdot \log_{d-1}(|G_i|).$$

This implies that $A_i e_g$ is supported on all elements of G_i, and thus the diameter of G_i is at most t. ∎

6 Operator Expander Mixing Lemma

In Section 3, we showed an operator amplification based on walks on an auxiliary expander. An alternative approach due to Chen, Moore and Russell [CMR13] proves an operator version of the expander mixing lemma (EML) and applies it in an iterated way (using different auxiliary graphs) for bias amplification. They obtain a dependence factor $1/\lambda^{11}$ in the degree. We show that this approach [CMR13] can achieve a dependence factor of $1/\lambda^{4+o(1)}$ which is similar to the expander walk approach Theorem 3.2 (also follows similar trade-offs to the scalar amplification via random walks [TS17]). We formally prove the following result.

Theorem 6.1 (Iterated Operator EML)**.** *Let $S \subseteq G$. Suppose $\lambda(\mathrm{Cay}(G,S)) = \lambda_0 < 1$, where $\lambda_0 \in (0,1)$. For every $\lambda \in (0,1)$, we can find $S' \subseteq G$ such that,*

1. *$\lambda(\mathrm{Cay}(G,S')) \leq \lambda$ and $|S'| = O_{\lambda_0}(|S|/\lambda^{4+o(1)})$, and*
2. *the running time is $\mathrm{poly}(|S|, 1/\lambda_0, 1/\lambda)$.*

We now show an operator version of the expander mixing lemma for completeness. As we mentioned above, a similar result was first derived in [CMR13]. While a simple generalization of EML, it is of the same nature of the generalizations of this paper, and is of independent interest.

Lemma 6.2 (Matrix EML [CMR13])**.** *Let $X = (V, E)$ be a $\lambda(X)$-spectral expander and let $f : V \to \mathcal{L}(\mathcal{H})$. Then,*

$$\left\| \mathop{\mathbb{E}}_{(x',x)\in E}[f(x')\cdot f(x)] - \left(\mathop{\mathbb{E}}_{x\in V_X}[f(x)]\right)^2 \right\|_{\mathrm{op}} \leq \lambda(X)\cdot \max_{x\in V_X} \|f(x)\|_{\mathrm{op}}^2 .$$

We start with a simple claim describing an operator form the process of sampling according to the edges of an expander and sampling according to pairs of vertices. Recall the following maps from Section 3, $P_{\mathcal{H}} : \mathcal{X}_{\mathcal{H}} \to \mathcal{H}$ and $L_{\mathcal{H}} : \mathcal{H} \to \mathcal{X}_{\mathcal{H}}$,

$$P_{\mathcal{H}}(w \otimes x) = w, \quad L_{\mathcal{H}}(v) = \mathop{\mathbb{E}}_{x\in V_X}[v \otimes x]\,.$$

We will need again that $\|P_{\mathcal{H}}\|_{\mathrm{op}} \|L_{\mathcal{H}}\|_{\mathrm{op}} = 1$.

Claim 6.3. *Let A_X be the normalized adjacency matrix of a d-regular graph X and let J_X be the normalized $|V_X| \times |V_X|$ all-ones matrix.*

$$\mathop{\mathbb{E}}_{(x,x')\in E}[f(x') \cdot f(x)] = P_{\mathcal{H}} \Pi_z \mathring{A}_X \Pi_z L_{\mathcal{H}}.$$
$$\mathop{\mathbb{E}}_{x,x'\in V}[f(x') \cdot f(x)] = P_{\mathcal{H}} \Pi_z \mathring{J}_X \Pi_z L_{\mathcal{H}}.$$

Proof. The proof is identical for both so we prove just the first one. For any $w \in \mathcal{H}$, we have

$$\begin{aligned}
P_{\mathcal{H}} \Pi_z \mathring{A}_X \Pi_z L_{\mathcal{H}} w &= \frac{1}{|V_X|} P_{\mathcal{H}} \Pi_z \mathring{A}_X \Pi_z \left(\sum_{x\in V_X} x \otimes w \right) \\
&= \frac{1}{|V_X|} P_{\mathcal{H}} \Pi_z \mathring{A}_X \left(\sum_{x\in V_X} x \otimes f(x) w \right). \\
&= \frac{1}{d|V_X|} P_{\mathcal{H}} \Pi_z \left(\sum_{x\in V_X} \sum_{x'\sim x} x' \otimes f(x) w \right). \\
&= \frac{1}{|E|} P_{\mathcal{H}} \left(\sum_{x\in V_X} \sum_{x'\sim x} x' \otimes f(x') f(x) w \right). \\
&= \frac{1}{|E|} \sum_{x\sim x'} f(x') f(x) w = \mathop{\mathbb{E}}_{(x',x)\in E}[f(x') \cdot f(x)] w.
\end{aligned}$$

as claimed. ∎

We now prove the operator mixing lemma above.

Proof of Lemma 6.2. By Claim 6.3, it is enough to bound the operator norm

$$\begin{aligned}
\left\| P_{\mathcal{H}} \Pi_z \left(\mathring{A}_X - \mathring{J}_X \right) \Pi_z L_{\mathcal{H}} \right\|_{op} &\le \|P_{\mathcal{H}}\|_{op} \|\Pi_z\|_{op}^2 \left\| \left(\mathring{A}_X - \mathring{J}_X \right) \right\|_{op} \|L_{\mathcal{H}}\|_{op} \\
&\le \lambda(X) \cdot \|\Pi_z\|_{op}^2 = \lambda(X) \cdot \max_{x\in V_X} \|f(x)\|_{op}^2,
\end{aligned}$$

concluding the proof. ∎

Corollary 6.4 (Non-abelian EML)**.** *Let $X = (V, E)$ be a $\lambda(X)$-spectral expander, $\rho : G \to U_{\mathcal{H}}$ be an unitary representation and $(g_v)_{v\in V} \in G^V$. Then*

$$\left\| \mathop{\mathbb{E}}_{(u,v)\in E} [\rho(g_u) \cdot \rho(g_v)] - \left(\mathop{\mathbb{E}}_{u\in V} [\rho(g_u)] \right)^2 \right\|_{op} \le \lambda(X).$$

Proof. Follows immediately from Lemma 6.2 and the fact that unitary operators have operator norm bounded by 1. ∎

We now prove the main result of this section. This iterated amplification also appears in the derandomized squaring of Rozenman and Vadhan [RV05] used to give an alternative proof of the SL = L result of Reingold [Rei04].

Proof of Theorem 6.1. We amplify the expansion in two phases. The first phase amplifies the initial expansion of S from λ_0 to a *constant* expansion $\lambda_0'' = 1/4$. This phase increases the size of the generator set by a constant factor.

(First Phase) Let ε_0, γ_0 be constants such that

$$\varepsilon_0 = \lambda_0(1-\lambda_0)/2, \quad 0 < \gamma_0 \leq (1-\lambda_0)/2 < 1$$

Let $X_0 = (V_0, E_0)$ be an explicit expander via Theorem 3.6, with $\lambda(X_0) \leq \varepsilon_0$, degree $O(1/\varepsilon_0^2)$ and with the number of vertices $|V_0| = m|S|$ with $m = O(1)$. Replicate each element of S m times and still call the resulting multiset S (observe that expansion remains λ_0). For every edge $(u, v) \in E_0$, add $g_u g_v$ to S_0. By Corollary 6.4,

$$\lambda(G, S_0) \leq \lambda_0^2 + \varepsilon_0 \leq \lambda_0(1-\gamma_0), \quad |S_0| = 9|S|/\varepsilon_0^2 = O(|S|)$$

Repeat this procedure $\log_{1-\gamma_0} 1/4\lambda_0$ times which ensures that the expansion is $\lambda_0'' = 1/4$. Let S_0 be this final set.

(Second Phase) We will amplify the bias inductively using a stronger (i.e., more expanding) auxiliary expander graph X_i at each step. As mentioned, this inductive amplification is similar to the derandomized squaring of Rozenman and Vadhan [RV05]. We start with S_0 and expansion $\lambda_0'' = 2^{-2}$ as in the first phase. At each step assume that you have a set S_{i-1} with expansion λ_{i-1}. Use Theorem 3.6, to construct X_{i-1} to have expansion λ_{i-1}^2 and degree at most $9/\lambda_{i-1}^4$. Then, S_i is obtained via edges of X_i as before and we have $\lambda_i \leq 2\lambda_{i-1}^2$. It is easy to check that the recurrence yields $\lambda_i \leq 2^{-(2^i)}$ for $i \geq 1$. Assume for convenience that $\log \lambda = -2^r$. Clearly, then we need to iterate this r times. In each iteration, the size grows by a factor of the degree which is $9/\lambda_{i-1}^4$ and thus the final size of S' can be bounded as,

$$|S'| = |S_0| \prod_{i=0}^{r-1} \frac{9}{\lambda_i^4} \leq |S_0| \cdot 9^r 2^{4+4(2^0+\cdots+2^{r-1})} = \frac{|S_0|}{\lambda^4} \cdot \left(\frac{1}{\log \lambda}\right)^{\log 9} \leq O_{\lambda_0}\left(\frac{|S|}{\lambda^{4+o(1)}}\right).$$

concluding the proof. ∎

Acknowledgement

We thank Alexander Lubotzky for stimulating and enlightening discussions in the initial phase of this project.

References

[ABN+92] N. Alon, J. Bruck, J. Naor, M. Naor, and R. Roth. Construction of asymptotically good, low-rate error-correcting codes through pseudo-random graphs. *IEEE Transactions on Information Theory*, 28:509–516, 1992. 6

[ACKM19] Naman Agarwal, Karthekeyan Chandrasekaran, Alexandra Kolla, and Vivek Madan. On the expansion of group-based lifts. *SIAM J. Discret. Math.*, 33(3):1338–1373, 2019. doi:10.1137/17M1141047. 10

[AGHP92] N. Alon, O. Goldreich, J. Håstad, and R. Peralta. Simple constructions of almost k-wise independent random variables. *Random Structures and Algorithms*, 3(3):289–304, 1992. 12, 27, 46

[AHL+14] Dorit Aharonov, Aram W. Harrow, Zeph Landau, Daniel Nagaj, Mario Szegedy, and Umesh V. Vazirani. Local tests of global entanglement and a counterexample to the generalized area law. In *Proceedings of the 55th IEEE Symposium on Foundations of Computer Science*, 2014. arXiv:1410.0951, doi:10.1109/FOCS.2014.34. 30, 32, 33

[Alo21] Noga Alon. Explicit expanders of every degree and size. *Combinatorica*, February 2021. doi:10.1007/s00493-020-4429-x. 14, 17, 19

[ALW01] N. Alon, A. Lubotzky, and A. Wigderson. Semi-direct product in groups and zig-zag product in graphs: connections and applications. In *Proceedings of the 42nd IEEE Symposium on Foundations of Computer Science*, 2001. 9

[AR94] Noga Alon and Yuval Roichman. Random cayley graphs and expanders. *Random Struct. Algorithms*, 5(2):271–285, 1994. doi:10.1002/rsa.3240050203. 6

[AS04] Andris Ambainis and Adam D. Smith. Small pseudo-random families of matrices: Derandomizing approximate quantum encryption. In Klaus Jansen, Sanjeev Khanna, José D. P. Rolim, and Dana Ron, editors, *APPROX-RANDOM 2004, Cambridge, MA, USA, August 22-24, 2004, Proceedings*, volume 3122 of *Lecture Notes in Computer Science*, pages 249–260. Springer, 2004. arXiv:0404075, doi:10.1007/978-3-540-27821-4_23. 4, 30, 32

[AW02] R. Ahlswede and A. Winter. Strong converse for identification via quantum channels. *IEEE Transactions on Information Theory*, 48(3), 2002. 8

[BASTS08] Avraham Ben-Aroya, Oded Schwartz, and Amnon Ta-Shma. Quantum expanders: Motivation and constructions. In *Proceedings of the 23rd IEEE Conference on Computational Complexity*, pages 292–303. IEEE Computer Society, 2008. doi:10.1109/CCC.2008.23. 4, 7, 30, 32, 33

[BATS08] Avraham Ben-Aroya and Amnon Ta-Shma. A combinatorial construction of almost-ramanujan graphs using the zig-zag product. In *Proceedings of the 40th ACM Symposium on Theory of Computing*, page 325–334, 2008. 2, 9, 16, 19, 25

[Bea97] Robert Beals. Quantum computation of fourier transforms over symmetric groups. In *Proceedings of the 29th ACM Symposium on Theory of Computing*, STOC '97, page 48–53, 1997. 4, 30, 33

[BISW01] B. Barak, R. Impagliazzo, A. Shpilka, and A. Wigderson. Dimension expanders. unpublished, 2001. 5, 36

[BKL89] László Babai, William M. Kantor, and A. Lubotzky. Small-diameter cayley graphs for finite simple groups. *Eur. J. Comb.*, 10, 1989. 37

[BL06] Yonatan Bilu and Nathan Linial. Lifts, discrepancy and nearly optimal spectral gap. *Combinatorica*, 26(5):495–519, October 2006. 2, 10

[BL18] Emmanuel Breuillard and Alexander Lubotzky. Expansion in simple groups, 2018. arXiv:1807.03879. 3

[BS88] László Babai and Akos Seress. On the diameter of cayley graphs of the symmetric group. *Journal of Combinatorial Theory, Series A*, 49(1), 1988. 37

[BY13] Jean Bourgain and Amir Yehudayoff. Expansion in sl 2(ℝ)and monotone expanders. *Geometric and Functional Analysis*, 23(1), 2013. 4, 33, 34, 37

[Che10] Yuan-You Fu-Rui Cheng. Explicit estimate on primes between consecutive cubes. *Rocky Mountain Journal of Mathematics*, 40(1), February 2010. arXiv:0810.2113, doi:10.1216/rmj-2010-40-1-117. 19

[CMR13] Sixia Chen, Cristopher Moore, and Alexander Russell. Small-bias sets for nonabelian groups - derandomizations of the Alon–Roichman theorem. In *APPROX-RANDOM*, volume 8096 of *Lecture Notes in Computer Science*, pages 436–451, 2013. 8, 11, 12, 38

[DS09] Zeev Dvir and Amir Shpilka. Towards dimension expanders over finite fields. *Combinatorica*, 31(3), sep 2009. 5, 33, 37

[DW10] Zeev Dvir and Avi Wigderson. Monotone expanders: Constructions and applications. *Theory of Computing*, 6(12), 2010. 33, 37

[FG15] Michael A. Forbes and Venkatesan Guruswami. Dimension Expanders via Rank Condensers. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2015)*, volume 40, pages 800–814, 2015. 37

[Fri03] Joel Friedman. A proof of alon's second eigenvalue conjecture. In *Proceedings of the 35th ACM Symposium on Theory of Computing*, 2003. 2, 10

[Gil52] E.N. Gilbert. A comparison of signalling alphabets. *Bell System Technical Journal*, 31:504–522, 1952. 3

[Gil93] D. Gillman. A Chernoff bound for random walks on expander graphs. In *Proceedings of the 34th IEEE Symposium on Foundations of Computer Science*, pages 680–691, 1993. 1, 9

[Har07] Aram W. Harrow. Quantum expanders from any classical cayley graph expander. *Quantum Information & Computation*, 2007. 4, 30, 33, 37

[Has07a] M. B. Hastings. Entropy and entanglement in quantum ground states. *Physical Review B*, 76(3), jul 2007. arXiv:0701055, doi:10.1103/physrevb.76.035114. 4, 30, 32

[Has07b] M. B. Hastings. Entropy and entanglement in quantum ground states. *Phys. Rev. B*, 2007. 30, 32

[Has07c] M. B. Hastings. Random unitaries give quantum expanders. *Phys. Rev. A*, 76:032315, Sep 2007. URL: https://link.aps.org/doi/10.1103/PhysRevA.76.032315, arXiv:0706.0556, doi:10.1103/PhysRevA.76.032315. 4, 32

[HH09] M. B. Hastings and A. W. Harrow. Classical and quantum tensor product expanders. *Quantum Info. Comput.*, 2009. 30, 32

[HLW06] Shlomo Hoory, Nathan Linial, and Avi Wigderson. Expander graphs and their applications. *Bull. Amer. Math. Soc.*, 43(04):439–562, August 2006. 1, 30

[JM21] Akhil Jalan and Dana Moshkovitz. Near-optimal cayley expanders for abelian groups, 2021. arXiv:2105.01149. 7, 14

[JMO+22] Fernando Granha Jeronimo, Tushant Mittal, Ryan O'Donnell, Pedro Paredes, and Madhur Tulsiani. Explicit abelian lifts and quantum ldpc codes. In *ITCS 2022*, 2022. 10

[JQST20] Fernando Granha Jeronimo, Dylan Quintana, Shashank Srivastava, and Madhur Tulsiani. Unique decoding of explicit ε-balanced codes near the Gilbert–Varshamov bound. In *Proceedings of the 61st IEEE Symposium on Foundations of Computer Science*, 2020. 45

[Kas07] Martin Kassabov. Symmetric groups and expander graphs. *Inventiones mathematicae*, 170(2):327–354, August 2007. arXiv:0505624, doi:10.1007/s00222-007-0065-y. 4, 10, 29, 33

[KKN21] Marek Kaluba, Dawid Kielak, and Piotr W. Nowak. On property (T) for $\mathrm{Aut}(\mathbb{F}_n)$ and $\mathrm{SL}_n(\mathbb{Z})$. *Annals of Mathematics*, 193(2):539 – 562, 2021. doi:10.4007/annals.2021.193.2.3. 30

[LP00] Alexander Lubotzky and Igor Pak. The product replacement algorithm and kazhdan's property (t). *Journal of the American Mathematical Society*, 14(2):347–363, October 2000. doi:10.1090/s0894-0347-00-00356-8. 30

[LPS88] Alexander Lubotzky, R. Phillips, and Peter Sarnak. Ramanujan graphs. *Combinatorica*, 8:261–277, 1988. 1, 10, 19

[Lub11] Alexander Lubotzky. Finite simple groups of Lie type as expanders. *Journal of the European Mathematical Society*, pages 1331–1341, 2011. arXiv:0904.3411, doi:10.4171/JEMS/282. 19

[Lub12] Alexander Lubotzky. Expander graphs in pure and applied mathematics. *Bull. Amer. Math. Soc.*, 2012. 1

[LZ08] Alexander Lubotzky and Efim Zelmanov. Dimension expanders. *Journal of Algebra*, 319(2):730–738, 2008. 30, 36, 37

[Mar73] G. A. Margulis. Explicit constructions of concentrators. *Probl. Peredachi Inf.*, 9, 1973. 1

[Mar88] G. A. Margulis. Explicit group-theoretical constructions of combinatorial schemes and their application to the design of expanders and concentrators. 1988. 1, 30

[MOP20] Sidhanth Mohanty, Ryan O'Donnell, and Pedro Paredes. Explicit near-ramanujan graphs of every degree. In *Proceedings of the 52nd ACM Symposium on Theory of Computing*, pages 510–523. ACM, 2020. 10

[MSS14] Adam Marcus, Daniel Spielman, and Nikhil Srivastava. Interlacing families ii: Mixed characteristic polynomials and the kadison–singer problem. *Annals of Mathematics*, 2014. 2

[MSS15] Adam Marcus, Daniel Spielman, and Nikhil Srivastava. Interlacing families i: Bipartite Ramanujan graphs of all degrees. *Annals of Mathematics*, 2015. 2

[MW04] Roy Meshulam and Avi Wigderson. Expanders in group algebras. *Combinatorica*, 24, 2004. 9

[Nil91] Alon Nilli. On the second eigenvalue of a graph. *Discrete Mathematics*, 91(2):207–210, 1991. doi:10.1016/0012-365X(91)90112-F. 1

[NN90] J. Naor and M. Naor. Small-bias probability spaces: efficient constructions and applications. In *Proceedings of the 22nd ACM Symposium on Theory of Computing*, pages 213–223, 1990. 6, 12, 13

[Pin73] Mark S. Pinsker. On the complexity of a concentrator. In *7th International Teletraffic Conference*, 1973. 1, 2

[PZ01] Igor Pak and Andrzej Zuk. Two Kazhdan constants and mixing of random walks. Technical report, Int. Math. Res. Not. 2002, 2001. 35

[Rei04] Omer Reingold. Undirected st-connectivity in log-space. Technical Report TR04-094, Electronic Colloquium on Computational Complexity, 2004. 39

[Rei05] Omer Reingold. Undirected ST-connectivity in log-space. In *Proceedings of the 37th ACM Symposium on Theory of Computing*, pages 376–385, 2005. 2, 18

[RL10] J.D. Rogawski and A. Lubotzky. *Discrete Groups, Expanding Graphs and Invariant Measures*. Modern Birkhäuser Classics. Birkhäuser Basel, 2010. 36

[RSW06] Eyal Rozenman, Aner Shalev, and Avi Wigderson. Iterative construction of cayley expander graphs. *Theory of Computing*, 2(5):91–120, 2006. 4, 9

[RV05] Eyal Rozenman and Salil Vadhan. Derandomized squaring of graphs. In *Proceedings of RANDOM'05*, pages 436–447. Springer-Verlag, 2005. 39, 40

[RVW00] O. Reingold, S. Vadhan, and A. Wigderson. Entropy waves, the zig-zag graph product, and new constant-degree expanders and extractors. In *Proceedings of the 41st IEEE Symposium on Foundations of Computer Science*, 2000. 2, 7, 9, 16, 19

[RVW02] Omer Reingold, Salil Vadhan, and Avi Wigderson. Entropy waves, the zig-zag graph product, and new constant-degree expanders. *Annals of Mathematics*, 155(1):157–187, 2002. 18

[SS96] L. L. Scott and J. P. Serre. *Linear Representations of Finite Groups*. Graduate Texts in Mathematics. Springer New York, 1996. 11

[Tro15] Joel A. Tropp. An introduction to matrix concentration inequalities. *Found. Trends Mach. Learn.*, 2015. 8

[TS17] Amnon Ta-Shma. Explicit, almost optimal, epsilon-balanced codes. In *Proceedings of the 49th ACM Symposium on Theory of Computing*, STOC 2017, pages 238–251, New York, NY, USA, 2017. ACM. 3, 5, 6, 7, 8, 13, 14, 19, 20, 22, 23, 24, 25, 26, 38, 45, 46

[TSD18] Amnon Ta-Shma and Dean Doron. Combinatorial constructions of expanders. the zig-zag product. Lecture notes, 2018. 18

[Vad12] Salil P. Vadhan. *Pseudorandomness*. Now Publishers Inc., 2012. 6, 34

[Var57] R.R. Varshamov. Estimate of the number of signals in error correcting codes. *Doklady Akademii Nauk SSSR*, 117:739–741, 1957. 3

[Wig18] Avi Wigderson. Mathematics and computation. Book draft at https://www.math.ias.edu/files/mathandcomp.pdf, 2018. 1

A Explicit Structures and their Parameters

The way we choose parameters and objects for it borrows heavily from Ta-Shma's arguments in [TS17]. The analysis follows an analogous structure of [JQST20] for binary codes, which in turn builds on the original analysis of Ta-Shma [TS17].

Given as input $|S|, \lambda$ and a slowly growing function $\beta(\lambda)$, we construct the graphs X, Y as described below with the following parameters which is similar (but not identical) to Ta-Shma's choice. Let s be the smallest power of 2 greater than $\frac{32}{\beta}$ and let $d_2 = s^{4s}$.

The outer graph X**.** We use our construction of expander from Corollary 3.12 to construct a graph on $n' \approx n$ vertices with expansion $\lambda_1 = \frac{\lambda_2^2}{10}$. The condition on the size is satisfied as $n = 2|S|d_2^5 \geq d_2^5 \geq 2^{2^{17}}$ by the assumption that $s \geq 2^{10}$. Moreover, the degree is $\frac{c}{\lambda_1^{2*4.1}} \leq \frac{c\, d_2^{4.1}}{b_2^{8.2}} \leq d_2^5$. We increase its degree to d_2^5 by taking multiple copies of the generating set which doesn't change bias[33]. Thus, we obtain a (n', d_1, λ_1)-graph where $n' = n + O(n^{8/9})$.

The inner graph Y**.** We obtain a Cayley graph $Y = \mathrm{Cay}(\mathbb{Z}_2^{\log(n_2)}, A)$ such that Y is an $(n_2 = d_2^{5s}, d_2, \lambda_2)$ graph[34]. The set A of generators comes from a small bias code derived

[33] This is wasteful but we do it to ensure that $V(Y) = d_1^s$ and that d_1^s is a power of 2.

[34] Notice that since s (and therefore d_2, n_2) is chosen to be a power of 2, the conditions of Lemma A.1 are satisfied.

from a construction of Alon et al. [AGHP92], but we will rely on Ta-Shma's analysis.

Lemma A.1 (Based on Lemma 6 [TS17]). *For every $m \in \mathbb{N}^+$ and $d = 2^{2k} \leq 2^m$, there exists a fully explicit set $A \subseteq \mathbb{Z}_2^m$ such that the graph $\mathrm{Cay}(\mathbb{Z}_2^m, A)$ is a $(2^m, d, \lambda = \frac{m}{\sqrt{d}})$-expander graph.*

We summarize the construction and the choice of parameters here -

s is the smallest power of 2 such that $\frac{32}{\beta} \leq s \leq \left(\frac{\log(1/\lambda)}{4\log\log(1/\lambda)}\right)^{1/3}$
Every other parameter is a function of s.
$Y : (n_2, d_2, \lambda_2), \quad n_2 = d_2^{5s}, \quad d_2 = s^{4s}, \quad \lambda_2 \leq \frac{b_2}{\sqrt{d_2}}, \quad b_2 = 5s\log d_2$
$X : (n', d_1, \lambda_1), \quad n' \approx n = O(\lvert S\rvert\, d_2^5), \quad d_1 = d_2^5, \quad \lambda_1 = \frac{\lambda_2^2}{10}$
t : smallest integer such that $(\lambda_2)^{(1-5\alpha)(1-\alpha)(t-1)} \leq \lambda$, ; where $\alpha = 1/s$

Note: We can assume that $s \geq 2^{10}$ since otherwise λ is a constant and we can just use Theorem 3.2.

Claim 4.14. *The selection of the parameters above implies the following bounds on t,*

i $t - 1 \geq 2s^2$

ii $(d_2)^{(t-1)} \leq \lambda^{-2(1+10\alpha)}$,

Proof. Proof of (i) Using $d_2 = s^{4s}$ and the upper bound on s, we have

$$\left(\frac{1}{\lambda_2}\right)^{(1-5\alpha)(1-\alpha)2s^2} \leq \left(\frac{1}{\lambda_2}\right)^{2s^2} = \left(\frac{d_2}{b_2^2}\right)^{s^2} \leq (d_2)^{s^2} = s^{4s^3}$$
$$= 2^{4s^3\log_2(s)} \leq 2^{\log_2(1/\lambda)} = \frac{1}{\lambda}.$$

Hence, $(\lambda_2)^{(1-5\alpha)(1-\alpha)s/\alpha} \geq \lambda$ and thus $t-1$ must be at least $2s^2$. Also observe that,

$$\lambda_2^{(1-5\alpha)(1-\alpha)^2(t-1)} = \lambda_2^{(1-5\alpha)(1-\alpha)(t-2)\left(\frac{(1-\alpha)}{1-1/(t-1)}\right)} \tag{7}$$
$$\geq \lambda_2^{(1-5\alpha)(1-\alpha)(t-2)} \qquad (t-1 \geq s = 1/\alpha) \tag{8}$$
$$\geq \lambda \qquad (\text{From the choice of minimal } t) \tag{9}$$

Since $b_2 = 5s\log_2(d_2) = 20s^2\log_2(s) \leq s^4$ (recall that $s = 1/\alpha \geq 2^{10}$),

$$d_2^{1-2\alpha} = \frac{d_2}{d_2^{2\alpha}} = \frac{d_2}{s^8} \leq \frac{d_2}{b_2^2} = \frac{1}{\lambda_2}.$$

We obtain (ii)

$$(d_2)^{(t-1)} \leq \lambda_2^{\frac{-(t-1)}{1-2\alpha}}$$
$$\leq \lambda^{\frac{-2}{(1-2\alpha)(1-5\alpha)(1-\alpha)^2}} \qquad (\text{Using Eq. (9)})$$
$$\leq \lambda^{-2(1+10\alpha)}. \qquad \blacksquare$$