Almost Ramanujan Expanders from Arbitrary Expanders via Operator Amplification

Fernando Granha Jeronimo* Tushant Mittal Sourya Roy
Avi Wigderson†

WORKING DRAFT
Please do not distribute

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 quantified by the second largest singular value of the 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 the quality of the sample improves with minimizing the spectral expansion 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 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 SL=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.*

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

[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.

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 graphs for these.

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 answer our original question, i.e., the case of general expander graphs. A result of König that says that the adjacency matrix of an arbitrary regular graph can

[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 $\Theta(\lambda^{2+o(1)})$.

[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].

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. For example, Harrow [Har07] gave a 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].

We give the first explicit family of almost Ramanujan quantum expanders.

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.*

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 a similar approach as used for Theorem 1.1, we express it as a sum of permutation matrices and amplify their expansion obtaining the following result.

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 $\lambda(X_i) \leq \lambda$.*

Remark 1.5. 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 move 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.

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

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.

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 and an improvement to the construction of *dimension expanders* in [LZ08].

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.*

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 a 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

[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.

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[7] is an operator valued function, $\rho : G \to \mathsf{M}_\ell(\mathbb{C})$, 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 $G = \mathbb{Z}_2^k$ has been extensively studied, e.g., [ABN+92, TS17].

Rozenman and Wigderson introduced the following "scalar amplification" technique using walks on an (auxiliary) expander graph, 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 $\mathsf{f} \colon S \to \{\pm 1\}$ be *any* function. A mapping T_W from functions on the space $\{\pm 1\}^S$ to the space $\{\pm 1\}^W$ can be defined as

$$\mathsf{T}_W(\mathsf{f})(w) \coloneqq \mathsf{f}(s_1) \cdots \mathsf{f}(s_t) \qquad \forall\, w = (s_1, \ldots, s_t) \in W\,.$$

In words, the value of each walk is given by the product of the values f assigns to its vertices. For a sufficiently "pseudorandom"[8] collection W and any function f satisfying $\mathrm{bias}(\mathsf{f}) \leq \varepsilon_0$, the bias of the amplified function, $\mathsf{T}_W(\mathsf{f})$, decreases exponentially (roughly) as $\mathrm{bias}(\mathsf{T}_W(\mathsf{f})) \leq \varepsilon_0^{t/2}$. Note that, when f is a character ρ (later we will consider more general representations), we can use the homomorphism property to write

$$\mathsf{T}_W(\rho)(w) = \rho(s_1 \cdots s_t) \qquad \forall\, w = (s_1, \ldots, s_t) \in W\,.$$

This allow us to interpret $\mathsf{T}_W(\rho)$ as a function with domain on the multiset $S' = \{s_1 \cdots s_t \mid w = (s_1, \ldots, s_t) \in W\}$, 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 1.6 (Scalar Amplification)**.** *For every $\lambda > 0$ and constant $\lambda_0 > 0$, there exists a deterministic polynomial time algorithm to construct $W \subseteq S^t$ of size $|W| \leq O(|S|/\lambda^{2+o(1)})$ such that for every function $\mathsf{f} : S \to \mathbb{C}$ with $|\mathbb{E}_{s\sim S}[\mathsf{f}(s)]| \leq \lambda_0$ and $\|\mathsf{f}\|_\infty \leq 1$, we have $|\mathbb{E}_{w\sim W}[\mathsf{T}_W(\mathsf{f})(w)]| \leq \lambda$.*

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

[7] The representation theory of finite groups over $\mathbb{C}$ is well-understood (e.g., [SS96]). We will need little of it here.

[8] This amounts to X being sufficiently expanding.

for Cayley graphs of general groups, but the resulting degree $O(|S|/\lambda^{11})$ to achieve final expansion λ is sub-optimal.

Our Results To extend the above approach to non-Abelian groups, it is necessary to work with operator valued functions, $\mathsf{f}\colon S \to \mathsf{M}_\ell(\mathbb{C})$, 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 $\mathsf{M}_\ell(\mathbb{C})^S$ to $\mathsf{M}_\ell(\mathbb{C})^W$.

Theorem 1.7 (Operator Amplification (this work))**.** *For every $\lambda > 0$ and constant $\lambda_0 > 0$, there exists a deterministic polynomial time algorithm to construct $W \subseteq S^t$ of size $|W| \leq O(|S|/\lambda^{2+o(1)})$ such that for every function $\mathsf{f} : S \to \mathsf{M}_\ell(\mathbb{C})$ with $\|\mathbb{E}_{s\sim S}[\mathsf{f}(s)]\|_{\mathrm{op}} \leq \lambda_0$ and $\max_s \|\mathsf{f}(s)\|_{\mathrm{op}} \leq 1$, we have $\|\mathbb{E}_{w\sim W}[\mathsf{T}_W(\mathsf{f})(w)]\|_{\mathrm{op}} \leq \lambda$.*

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

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

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

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

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 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 $\mathsf{f} : S \to \mathcal{L}(\mathcal{H})$ where $\mathcal{L}(\mathcal{H})$ is the space of bounded linear operators on an arbitrary Hilbert space, $\mathcal{H}$. This is useful if the underlying group is not finite but finitely generated by S.

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 of operators on some Hilbert space, of unit norm, such that $\|\mathbb{E}_{s\in S}[s]\|_{\mathrm{op}} \leq \lambda_0$. For any $\lambda > 0$ take $k = c\log(1/\lambda)$ (for appropriate c), and the set S_k of all k-wise products of operators in S^k. Clearly, $\|\mathbb{E}_{r\in S_k}[r]\|_{\mathrm{op}} \leq (\|\mathbb{E}_{s\in S}[s]\|_{\mathrm{op}})^k \leq \lambda$. Our main result is an explicit construction of a (pseudorandom)

subset $S' \in S_k$, of size only $|S'| = O(|S|/\lambda^{2+o(1)})$, with a similar guarantee, namely $\|\mathbb{E}_{s' \in S'}[s']\|_{\mathrm{op}} \leq \lambda$.

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^k, 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 combinatroial construction, 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 the set S above. When the elements of S are the (operator) values of some non-trivial representation of a group, on a set of expanding generators of it, the trivial amplification S_k, and the near-optimal amplification S' are simply the values of k-products of these generators. The analysis of the expected norm above directly yields the required expansion bound, in a way that has no dependence on the group or the representation. This *local* structure of the set S is of course what underlies the versatility of our spectral amplification, and its ability to preserve the structure of whatever expander whose expansion is amplified.

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 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 some parameters to some algebraic ones, yield new results in group theory.

Iterated Pseudorandomness Another curious aspect of our result is the following. Recall that expanders are pseudorandom objects for many purposes. One important purpose is sampling - rather than sample k independent random elements in some set V, one may sample k points along a random path in an expander on V - this affords significant savings in random bits spent (a nontrivial result of [Gil93]). For this result, any expander would do. What happens in this paper is an iterated use of expanders as samplers. The choice of the set S' of pseudorandom k-paths inside S^k again uses walks on expanders. This iterated use improves the quality of the original expander to being near-optimal spectrally. However, now the choice of which expander is used for the selection of paths seems critical, and (as in Ta-Shma's paper) is chosen to be the s-wide zig-zag graph of [BATS08].

Group Theory For us, the most surprising consequence of our results is that "weak" simple groups, especially the symmetric group,[9] 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). The symmetric group is weak (due to a celebrated result of Kassabov [Kas07] who found expanding generators for it), 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 for all we have almost Ramanujan generators.

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 to bootstrap the more involved construction based on s-wide replacement product which is the subject of Section 4. Here, we prove the main amplification result 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 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[10] 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. *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[11] on $\mathcal{H}$. We denote by $\mathsf{U}_{\mathcal{H}}$ the unitary group of operators acting on $\mathcal{H}$.

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

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

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

Definition 2.2 (Unitary Group Representation)**.** For a group G, a unitary representation is a pair $(\rho, \mathcal{H})$ where $\rho : G \to \mathrm{U}_{\mathcal{H}}$ 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 no (non-trivial[12]) subspace of $\mathcal{H}$, is invariant under the action of $\rho(G)$.

Every group has two special representations, which are,

1. (Trivial representation) - $(\rho, \mathbb{C})$ where for every g, $\rho(g) = 1$.
2. ((left) Regular representation) - $(\rho_{reg}, \mathcal{V}_{reg})$ where, $\mathcal{V}_{reg} = \mathbb{C}[G]$ is a vector space with the elements of G being an orthonormal basis, and $\rho_{reg}(g) : h \mapsto g \cdot h$.

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

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

where $\mathrm{Irrep}(G)$ 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 $\mathsf{A} = \mathbb{E}_{s \sim S}[\rho_{reg}(s)]$. 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,

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

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

where for any bounded linear operator, T, between (non-empty) Hilbert spaces,

$$\|\mathsf{T}\|_{\mathrm{op}} = \sup_{\|v\|=1} \|\mathsf{T}v\| = \sup_{\|v\|=\|w\|=1} |\langle \mathsf{T}v, w \rangle| .$$

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 $\varepsilon_0 \in (0, 1)$. We say that a multiset S of elements of a group G is ε_0-biased if for every non-trivial irreducible representation ρ, we have $\|\mathbb{E}_{s \sim S}[\rho(s)]\|_{\mathrm{op}} \leq \varepsilon_0$. We sometimes use the shorthand $\mathrm{bias}(S) \leq \varepsilon_0$.

Irreducible representations of Abelian groups, called *characters*, have dimension 1. Thus, this definition coincides with the usual one in which a biased distribution is one that *fools* non-trivial characters.

[12]The entire space $\mathcal{H}$ and $\{0\}$ are called trivial spaces.

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 $\mathsf{A}_X, \mathsf{A}_Y$.
- 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 $\mathcal{X}_{\mathcal{H}} \coloneqq \mathcal{H} \otimes \mathbb{C}[V_X]$, 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 $\mathcal{XY}_{\mathcal{H}} \coloneqq \mathcal{X}_{\mathcal{H}} \otimes \mathbb{C}[V_Y]$.
- Orthogonal decomposition: $\mathcal{X}_{\mathcal{H}} = \mathcal{X}_{\mathcal{H}}^{\parallel} \oplus \mathcal{X}_{\mathcal{H}}^{\perp}$ where $\mathcal{X}_{\mathcal{H}}^{\parallel} \coloneqq \mathrm{span}\{v \otimes \vec{1} \mid v \in \mathcal{H}\}$. Here, $\vec{1}$ denotes the un-normalized all-ones vector. Similarly, $\mathcal{XY}_{\mathcal{H}} = \mathcal{XY}_{\mathcal{H}}^{\parallel} \oplus \mathcal{XY}_{\mathcal{H}}^{\perp}$, where $\mathcal{XY}_{\mathcal{H}}^{\parallel} \coloneqq \mathrm{span}\{z \otimes \vec{1} \mid z \in \mathcal{X}_{\mathcal{H}}\}$.
- The operator $\overset{\circ}{\mathsf{A}}$ 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 $\overset{\circ}{\mathsf{A}}_X = \mathsf{I}_{\mathcal{H}} \otimes \mathsf{A}_X$ but if we were working on $\mathcal{XY}_{\mathcal{H}}$, it would be $\overset{\circ}{\mathsf{A}}_X = \mathsf{I}_{\mathcal{H}} \otimes \mathsf{A}_X \otimes \mathsf{I}_Y$. [13]
- For any operator valued function $\mathsf{f} : V_X \to \mathcal{L}(\mathcal{H})$, we will write f_x instead of $\mathsf{f}(x)$ and assume that it is normalized, $\max_x \|\mathsf{f}_x\|_{\mathrm{op}} \leq 1$.
- Given such a map f, the generalized bias operator is defined as[14],

$$\Pi_{\mathsf{f}} \colon \mathcal{X}_{\mathcal{H}} \to \mathcal{X}_{\mathcal{H}}, \quad v \otimes x \mapsto \mathsf{f}_x v \otimes x.$$

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 prove the following operator analogue.

Theorem 3.1 (Operator Amplification via Expander Walks)**.** *Let X be a $\lambda(X)$-spectral expander and let f be such that $\|\mathbb{E}_{x \in V_X}[\mathsf{f}_x]\|_{\mathrm{op}} \leq \lambda_0$ and $\max_{x \in V_X} \|\mathsf{f}_x\|_{\mathrm{op}} \leq 1$. Then,*

$$\left\| \Pi_{\mathsf{f}} \left(\overset{\circ}{\mathsf{A}}_X \Pi_{\mathsf{f}} \right)^t \right\|_{\mathrm{op}} \leq \left(2\lambda(X) + \lambda_0 \right)^{\lfloor t/2 \rfloor}.$$

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

[14] An equivalent matrix definition is $\Pi_{\mathsf{f}} \coloneqq \sum_{x \in V_X} \mathsf{f}_x \otimes \mathsf{E}_{x,x}$ where $\mathsf{E}_{x,x} \in \mathbb{C}^{V_X \times V_X}$ 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 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 the structure of 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 (see Theorem 3.2). We have the following theorem.

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 S. 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 $\mathcal{L}(\mathcal{H}) = \mathsf{M}_\ell(\mathbb{C})$. Let A_X be the normalized adjacency matrix of X and $\mathring{\mathsf{A}}_X = \mathsf{I}_{\mathcal{H}} \otimes \mathsf{A}_X$ be its extension to $\mathcal{X}_{\mathcal{H}} := \mathcal{H} \otimes \mathbb{C}[V_X]$. Let $\mathsf{f} \colon S \to \mathcal{L}(\mathcal{H})$ be *any* operator valued function and define Π_{f} to be the generalized bias operator acting as $\Pi_{\mathsf{f}}(v \otimes x) = \mathsf{f}_x v \otimes x$. In the scalar case, since $\mathcal{H} = \mathbb{C}$, we had an implicit identification of $\mathbb{C} \otimes \mathbb{C}[V_X] \cong \mathbb{C}[V_X]$. This identification no longer is suitable when f is operator valued. 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 $W_t \subset S^t$ be the collection of all length t walks on the graph X. Formally, we have

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

This can be shown easily via an induction on t and we refer to Lemma 4.9 for a formal proof. A minor technicality is that the operators in the image of f act on $\mathcal{H}$ whereas Π_{f} acts on the space $\mathcal{X}_{\mathcal{H}}$. Thus, we use projection and lifting maps to move between the spaces

$$\mathsf{P}_{\mathcal{H}}(w \otimes x) = w, \;\; \mathsf{L}_{\mathcal{H}}(v) = \mathop{\mathbb{E}}_{x \in X} [v \otimes x] \,.$$

It follows directly from the definition that $\|\mathsf{L}_{\mathcal{H}}\|_{\mathrm{op}} = \frac{1}{\sqrt{|V_X|}}$ and we can use Cauchy-Schwarz to get that $\|\mathsf{P}_{\mathcal{H}}\|_{\mathrm{op}} = \sqrt{|V_X|}$. Now, we put this together to get

$$\begin{aligned}
\left\| \mathop{\mathbb{E}}_{w \in W_t} [\mathsf{T}_{W_t}(\mathsf{f})(w)] \right\|_{\mathrm{op}} &= \left\| \mathop{\mathbb{E}}_{(s_0, \ldots s_t) \in W_t} [\mathsf{f}_{s_t} \cdots \mathsf{f}_{s_0}] \right\|_{\mathrm{op}} && (2) \\
&= \sup_{\|v\|=1} \left\| \mathop{\mathbb{E}}_{(s_0, \cdots, s_t) \in W_t} [\mathsf{f}_{s_t} \cdots \mathsf{f}_{s_0}]\, v \right\|_2 && (3) \\
&= \sup_{\|v\|=1} \left\| \mathsf{P}_{\mathcal{H}} \Big(\mathop{\mathbb{E}}_{(s_0, \cdots, s_t) \in W_t} [\mathsf{f}_{s_t} \cdots \mathsf{f}_{s_0}]\, v \otimes s_t \Big) \right\|_2 && (4)
\end{aligned}$$

$$= \sup_{\|v\|=1} \left\| \mathsf{P}_{\mathcal{H}} \Pi_{\mathsf{f}} \left(\mathring{\mathsf{A}}_X \Pi_{\mathsf{f}} \right)^t \mathop{\mathbb{E}}_{s \in S} [v \otimes s] \right\|_2 \tag{5}$$

$$= \sup_{\|v\|=1} \left\| \mathsf{P}_{\mathcal{H}} \Pi_{\mathsf{f}} \left(\mathring{\mathsf{A}}_X \Pi_{\mathsf{f}} \right)^t \mathsf{L}_{\mathcal{H}} v \right\|_2 \tag{6}$$

$$\leq \left\| \Pi_{\mathsf{f}} \left(\mathring{\mathsf{A}}_X \Pi_{\mathsf{f}} \right)^t \right\|_{\mathrm{op}} \left\| \mathsf{P}_{\mathcal{H}} \right\|_{\mathrm{op}} \left\| \mathsf{L}_{\mathcal{H}} \right\|_{\mathrm{op}} \tag{7}$$

$$\leq \left\| \Pi_{\mathsf{f}} \left(\mathring{\mathsf{A}}_X \Pi_{\mathsf{f}} \right)^t \right\|_{\mathrm{op}} . \tag{8}$$

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 $\left\| \mathbb{E}_{s \sim S} [\rho(s)] \right\|_{\mathrm{op}} \leq \lambda_0$, then we have $\left\| \mathbb{E}_{s \sim S'} [\rho(s)] \right\|_{\mathrm{op}} \leq \lambda \ll \lambda_0$. 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 $\mathsf{f}_s = \rho(s)$ in this case. Let

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

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

$$\mathop{\mathbb{E}}_{w \in W_t} \left[\mathsf{T}_{W_t}(\mathsf{f})(w) \right] = \mathop{\mathbb{E}}_{(s_0, \ldots, s_t) \in W_t} \left[\rho(s_t) \cdots \rho(s_0) \right] = \mathop{\mathbb{E}}_{s' \in S'} \left[\rho(s') \right] ,$$

and the multiset S' is the new biased set of the construction.

3.1 Bounding the Operator Norm Decay

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 $\mathcal{X}_{\mathcal{H}}^{\|} := \mathrm{span}\{ v \otimes \vec{1} \mid v \in \mathcal{H} \}$. 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 3.3 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 3.3. *For $z \in \mathcal{X}_{\mathcal{H}}^{\|}$, we have*

$$\left\| (\Pi_{\mathsf{f}} z)^{\|} \right\|_2 \leq \left\| \mathop{\mathbb{E}}_{x \in V_X} [\mathsf{f}_x] \right\|_{\mathrm{op}} \cdot \|z\|_2 .$$

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,

$$\left\| \left(\Pi_{\mathsf{f}} \left(u \otimes \vec{1} \right) \right)^{\|} \right\|_2 = \sup_{w \in \mathcal{H} \colon \left\| w \otimes \vec{1} \right\|_2 = 1} \left| \left\langle w \otimes \vec{1}, \Pi_{\mathsf{f}} \left(u \otimes \vec{1} \right) \right\rangle \right|$$

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

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

Lemma 3.4. *Let X be a $\lambda(X)$-spectral expander and let f be such that $\left\|\mathbb{E}_{x\in V_X}[\mathsf{f}_x]\right\|_{\mathrm{op}} \leq \lambda_0$ and $\max_{x\in V_X}\left\|\mathsf{f}_x\right\|_{\mathrm{op}} \leq 1$. Then,*

$$
\left\|\left(\mathring{\mathsf{A}}_X\Pi_{\mathsf{f}}\right)^2\right\|_{\mathrm{op}} \leq 2\lambda(X)+\lambda_0\,.
$$

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

$$
\begin{aligned}
\left\|\left(\mathring{\mathsf{A}}_X\Pi_{\mathsf{f}}\mathring{\mathsf{A}}_X\right)\left(z^{\perp}+z^{\parallel}\right)\right\|_2 &\leq \lambda(X)+\left\|\left(\mathring{\mathsf{A}}_X\Pi_{\mathsf{f}}\mathring{\mathsf{A}}_X\right)z^{\parallel}\right\|_2 \\
&\leq \lambda(X)+\left\|\mathring{\mathsf{A}}_X\left(\left(\Pi_{\mathsf{f}}z^{\parallel}\right)^{\perp}+\left(\Pi_{\mathsf{f}}z^{\parallel}\right)^{\parallel}\right)\right\|_2 \\
&\leq \lambda(X)+\left\|\mathring{\mathsf{A}}_X\left(\Pi_{\mathsf{f}}z^{\parallel}\right)^{\perp}\right\|_2+\left\|\left(\Pi_{\mathsf{f}}z^{\parallel}\right)^{\parallel}\right\|_2 \\
&\leq 2\lambda(X)+\left\|\left(\Pi_{\mathsf{f}}z^{\parallel}\right)^{\parallel}\right\|_2 \\
&\leq 2\lambda(X)+\lambda_0. \qquad \text{(By Claim 3.3)}
\end{aligned}
$$

$\blacksquare$

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 to perform walks on. We obtain this from a recent construction of Alon.

Theorem 3.5 (Corollary of [Alo21, Thm. 1.3])**.** *For any n, λ, we have an explicit construction of a graph X on $m_\lambda n$ vertices where with degree $9/\lambda^2$ such that $\lambda(X) \le \lambda$. Here, $m_\lambda \in \mathbb{N}$.*

We now establish the key amplification lemma.

Lemma 3.6. *Let $S \subseteq G$ such that $\mathrm{bias}(S) = \lambda_0 < 1$. Then, for any $\lambda > 0$, we can explicitly compute S' such that $\mathrm{bias}(S') \le \lambda$ 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 3.5 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

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

Let ρ be any irreducible representation. From Eq. (8) and Theorem 3.1, we get,

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

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 3.6 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 3.6 again with S_1 as the initial set and the final expansion as λ to obtain S'. This time we fix $\varepsilon_0 = 1/2$ in the proof of Lemma 3.6 and by our choice of λ', we have $\delta(\lambda') \le \beta$. Thus, the final size is $|S'| \le O_{\lambda'}\left(\frac{|S_1|}{\lambda^{4+\delta(\lambda')}}\right) \le O_{\lambda_0,\beta}\left(\frac{|S|}{\lambda^{4+\beta}}\right)$. $\blacksquare$

3.3 Derandomized Powering from Any Bias

We now show that amplification occurs whenever $\|\mathbb{E}_{x\in V_X}[\mathsf{f}_x]\|_{\mathrm{op}} < 1$ and the auxiliary graph X has positive spectral gap. This establishes that expander walks can be used to derandomize powers of an opeartor, 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 3.7 (Operator Amplification via Expander Walks (strengthening of Theorem 3.1))**.** *Let X be a $\lambda(X)$-spectral expander and let f be such that $\|\mathbb{E}_{x\in V_X}[\mathsf{f}_x]\|_{\mathrm{op}} \le \lambda_0$ and $\max_{x\in V_X}\|\mathsf{f}_x\|_{\mathrm{op}} \le 1$. Then,*

$$\left\|\Pi_{\mathsf{f}}\left(\mathring{\mathsf{A}}_X\Pi_{\mathsf{f}}\right)^t\right\|_{\mathrm{op}} \le \left(1-(1-\lambda(X))^2(1-\lambda_0)\right)^{\lfloor t/2\rfloor}.$$

This above amplification follows from the following improved version of Lemma 3.4. The proof explores the structural syntactically similarity between the operator amplification and known zig-zag analysis [RVW02, Rei05, TSD18].

Lemma 3.8. *Let X be a $\lambda(X)$-spectral expander and let f be such that $\|\mathbb{E}_{x\in V_X}[\mathsf{f}_x]\|_{\mathrm{op}} \leq \lambda_0$ and $\max_{x\in V_X}\|\mathsf{f}_x\|_{\mathrm{op}} \leq 1$. Then,*

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

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

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

By Lemma 3.4 and the fact that $\lambda(\mathsf{A}_J) = 0$, we obtain

$$\left\|\mathring{\mathsf{A}}_J\Pi_{\mathsf{f}}\mathring{\mathsf{A}}_J\right\|_{\mathrm{op}} \leq 2\lambda(\mathsf{A}_J) + \lambda_0 = \lambda_0\,,$$

Recall that $\|\Pi_{\mathsf{f}}\|_{\mathrm{op}} \leq 1$ since $\max_x\|\mathsf{f}_x\|_{\mathrm{op}} \leq 1$, and we also have $\|\mathsf{E}\|_{\mathrm{op}}, \|\mathsf{A}_J\|_{\mathrm{op}} \leq 1$. Then,

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

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 3.11). 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 $\mathrm{SL}_2(p)$ is the group of 2×2 invertible matrices over $\mathbb{F}_p$ with determinant 1. We obtain a base generating set for $\mathrm{SL}_2(p)$ via the following result.

Theorem 3.9 ([Lub11])**.** *There exists an explicit generating set S (of constant size) for $\mathrm{SL}_2(p)$ for any $p > 17$ such that $\lambda\left(Cay(\mathrm{SL}_2(p), S)\right) \leq \lambda_0$ for some absolute constant $\lambda_0 < 1$.*

Lemma 3.10 ([Che10])**.** *For every $n \geq 2^{3\cdot 2^{15}}$, there exists a prime in $[n, n+4n^{2/3}]$.*

Corollary 3.11. *For any $n > 2^{9\cdot 2^{15}}, \lambda > 0$, there is a deterministic polynomial time algorithm to construct an (n', d, λ)-graph $Cay(\mathrm{SL}_2(p), S)$, 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 Lemma 3.10, via brute-force search. Since, $\mathrm{SL}_2(p)$ is a group of order $(p^2 - 1)p$, we have $n \leq |\mathrm{SL}_2(p)| \leq n + O(n^{8/9})$. We use the constant-sized generating set S from Theorem 3.9 and amplify using Theorem 3.2. ∎

4 Operator Bias Reduction via the s-wide Replacement Walk

This section establishes our key technical result which states that given any initial set having a constant bias < 1 with respect to any *operator valued function*, we can construct a biased distribution with almost optimal size-bias trade-off. 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])**.** *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 $2\lambda(X) + \|\mathbb{E}_{x \in V_X}[\mathsf{f}_x]\|_{\mathrm{op}} \leq \lambda(Y)^2$ and $\max_{x \in V_X} \|\mathsf{f}_x\|_{\mathrm{op}} \leq 1$, we have*

$$\left\| \mathop{\mathbb{E}}_{w \in W_t} [\mathsf{T}_{W_t}(\mathsf{f})(w)] \right\|_{\mathrm{op}} \leq \left(\lambda(Y)^s + s \cdot \lambda(Y)^{s-1} + s^2 \cdot \lambda(Y)^{s-3} \right)^{\lfloor t/s \rfloor}.$$

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). Therefore, we need a sparser distribution (set) while retaining the same bias (expansion) guarantee. 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].

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 G^s over any group of appropriate size, 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 4.9 which is identical in spirit to Eq. (1) . We then finish the analysis in a manner similar to [TS17]. In Section 4.4, we provide details about instantializing the setup by explicitly constructing the graphs we need.

4.1 The s-wide Replacement Product

The standard replacement product takes an *outer* graph X, which is d_1-regular, and replaces each vertex of X with a "cloud" which is a copy of an *inner* graph on d_1 vertices, say, Y. 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). 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 $\mathsf{X}_0, \mathsf{X}_1, \ldots, \mathsf{X}_{s-1}$ 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 $\mathsf{I} \otimes \mathsf{A}_Y$, where A_Y is the normalized adjacency matrix of Y. The inter-cloud substeps 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 $t-1$ steps on the s-wide replacement product is

$$\prod_{i=t-2}^{0} \mathsf{X}_{i \bmod s}(\mathsf{I} \otimes \mathsf{A}_Y).$$

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 $(t-1)$-step walks on the s-wide replacement product is,

$$|V_X| \cdot |V_Y| \cdot \deg(Y)^{t-1} = n \cdot \deg(X)^s \cdot \deg(Y)^{t-1} \ll n \cdot \deg(X)^{t-1},$$

which therefore gives us an even sparser set of walks. Thus the s-wide replacement product will be used to simulate random walks on X while requiring a reduced amount of randomness (of course 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 4.2 (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 4.3 (Cayley Graphs are Locally Invertible)**.** Let G be a group and $A \subseteq G$ where the set A is closed under inversion. Label the neighbours 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 4.4 (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 $Y = ([d_1]^s, E')$.

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 $t-1$ steps of the s-wide replacement product are given by the operator

$$\mathsf{X}_{t-2 \bmod s} \overset{\circ}{\mathsf{A}}_Y \cdots \mathsf{X}_{0 \bmod s} \overset{\circ}{\mathsf{A}}_Y .$$

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, 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. (8)). 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 decay of bias.

Definition 4.5 (Operators and Distributions)**.** Given a tuple of random walk operators $\mathsf{B} = (\mathsf{B}_0, \cdots, \mathsf{B}_{t-1})$ on $\mathbb{C}[V_X] \otimes \mathbb{C}[V_Y]$ and a starting vertex x_0, we can define a distribution induced by the walk using these operators. More precisely, $\mathcal{D}(\mathsf{B}, x_0)$ is the distribution on on $(V_X \times V_Y)^{t+1}$ such that for every $1 \leq \ell \leq t$,

$$(\mathsf{B}_{\ell-1} \cdots \mathsf{B}_0) \left(x_0 \otimes \frac{1}{|V_Y|} \vec{1} \right) = \mathbb{E}_{(\vec{x}, \vec{y}) \sim \mathcal{D}(\mathsf{B})} x_\ell \otimes y_\ell. \tag{9}$$

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

Using this definition, we define the operators for the distributions we wish to study.

Uniform Distribution We define B_U where for each i, $\mathsf{B}_i = \mathsf{A}_X \otimes \mathsf{I}_Y$ for every i. Then, for any ℓ, $(\mathsf{A}_X \otimes \mathsf{I}_Y)^\ell = \mathsf{A}_X^\ell \otimes \mathsf{I}_Y$. Therefore, we obtain that $\mathcal{D}_X(\mathsf{B}_U)$ 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

$$\mathsf{B}[a,b] = \left(\mathsf{X}_a \overset{\circ}{\mathsf{A}}_Y, \mathsf{X}_{a+1}\overset{\circ}{\mathsf{A}}_Y, \cdots, X_b \overset{\circ}{\mathsf{A}}_Y\right).$$

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 neighbours according to which we will perform the walk in Y. To formalize this, let $\mathsf{A}_Y = (1/d_2)\sum_{j=1}^{d_2} \mathsf{P}_j$ 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

$$\mathsf{B}[a,b,J] = \left(\mathsf{X}_a \overset{\circ}{\mathsf{P}}_{j_0}, \mathsf{X}_{a+1}\overset{\circ}{\mathsf{P}}_{j_1}, \cdots, \mathsf{X}_b \overset{\circ}{\mathsf{P}}_{j_{b-a}}\right).$$

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 4.6 (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[15]

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

Remark 4.7. 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.

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

Lemma 4.8 ([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 it is a Cayley graph, we can think of $J \in S^t$ and the permutation matrices as $\mathsf{P}_g = \rho_{reg}(g)$. Recall that for any $y \in G^s$,

$$\mathsf{P}_g y = gy = (g_1 r_1, \cdots, g_s r_s),\ \mathsf{X}_i y = (r_1, \cdots, r_{i-1}, \varphi(r_i), r_{i+1}, \cdots, r_s)$$

Let $y = (r_1, \cdots, r_s) \sim G^s$ be the initial vector which is uniform. Since g_i and φ are fixed, the above operators don't change the 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, $x_i = x_{i-i}[\tau_{a+i-1,i}(r_{a+i-1})]$ and we take at most s steps and therefore, we use r_i for distinct $i \in [a,b]$ which are all independent.Thus, $x_i \sim \mathsf{A}_X^i x_0$. ∎

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 the spaces are slightly more involved, we formalize the computation more explicitly. Recall that $\overset{\circ}{\Pi}_{\mathsf{f}}(v \otimes x \otimes y) = \mathsf{f}_x v \otimes x \otimes y$.

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

Lemma 4.9 (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*

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

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

$$\begin{aligned}\left(\mathring{\mathsf{B}}_0\mathring{\Pi}_{\mathsf{f}}\right)\left(v \otimes x_0 \otimes \frac{1}{|V_Y|}\vec{1}\right) &= \mathring{\mathsf{B}}_0\left(\mathsf{f}_{x_0} v \otimes x_0 \otimes \frac{1}{|V_Y|}\vec{1}\right)\\ &= \mathop{\mathbb{E}}_{(\vec{x},\vec{y})\sim D(\mathsf{B})}\left[\mathsf{f}_{x_0} v \otimes x_1 \otimes y_1\right] \qquad \text{(Using Eq. (9) for } \ell = 1)\end{aligned}$$

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

$$\begin{aligned}\left(\mathring{\mathsf{B}}_{t-1}\mathring{\Pi}_{\mathsf{f}} \cdots \mathring{\mathsf{B}}_0\mathring{\Pi}_{\mathsf{f}}\right)(v \otimes x_0 \otimes y_0) &= \mathring{\mathsf{B}}_{t-1}\mathring{\Pi}_{\mathsf{f}} \cdot \prod_{i=t-2}^{0}\left(\mathring{\mathsf{B}}_i\mathring{\Pi}_{\mathsf{f}}\right)(v \otimes x_0 \otimes y_0)\\ &= \mathring{\mathsf{B}}_{t-1}\mathring{\Pi}_{\mathsf{f}} \mathop{\mathbb{E}}_{(\vec{x},\vec{y})\sim D(\mathsf{B})}\left[\mathsf{f}_{x_{t-2}} \cdots \mathsf{f}_{x_0} v \otimes x_{t-1} \otimes y_{t-1}\right]\\ &= \mathring{\mathsf{B}}_{t-1} \mathop{\mathbb{E}}_{(\vec{x},\vec{y})\sim D(\mathsf{B})}\left[\mathsf{f}_{x_{t-1}}\mathsf{f}_{x_{t-2}} \cdots \mathsf{f}_{x_0} v \otimes x_{t-1} \otimes y_{t-1}\right]\\ &= \mathop{\mathbb{E}}_{(\vec{x},\vec{y})\sim D(\mathsf{B})}\left[\mathsf{f}_{x_{t-1}} \cdots \mathsf{f}_{x_0} v \otimes x_t \otimes y_t\right].\end{aligned}$$

The second equality uses the inductive hypothesis and the third uses the fact that $\mathring{\Pi}_{\mathsf{f}}$ acts on the tensor space diagonally. Last two equalities use Eq. (9) for $\ell = t-1$ and $\ell = t$ respectively. ∎

4.3 The s-wide Norm Decay

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

Lemma 4.10 (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,*

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

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 $t = s_2 - s_1 + 1$.

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

Therefore we can fix $J = (j_{s_1}, \cdots, j_{s_2}) \in [d_2]^t$ and prove it for that. Applying Lemma 4.9 to $\mathsf{B}[s_1, s_2, J]$, we get,

$$\prod_{i=s_1}^{s_2}\left(\mathring{\mathsf{X}}_i\mathring{\mathsf{P}}_{j_i}\mathring{\Pi}_{\mathsf{f}}\right)\left(v \otimes x_0 \otimes \frac{1}{|V_Y|}\vec{1}\right) = \mathop{\mathbb{E}}_{(\vec{x},\vec{y})\sim D(\mathsf{B}[s_1,s_2,J])}\left[\mathsf{f}_{x_{t-1}} \cdots \mathsf{f}_{x_0} v \otimes x_t \otimes y_t\right]$$

$$= \sum_{\vec{x}\in V_X^t} \mathop{\mathbb{E}}_{y_0\sim V_Y} [\mathsf{f}_{\vec{x}} v \otimes x_t \otimes y_t]\, \mathbb{I}[y_0 \text{ gives rise to } \vec{x}],$$

where $\mathsf{f}_{\vec{x}} = \mathsf{f}_{x_t} \cdots \mathsf{f}_{x_0}$. 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 $x_i = x_{i-i}[\tau_{s_1+i-1,i}(r_{a+i-1})]$ 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

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

We now use[16] Lemma 4.9 for B_U and take inner product to get,

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

From Lemma 4.8, we know that Y is compatible and thus, $\mathcal{D}_X(\mathsf{B}[s_1, s_2, J]) = \mathcal{D}_X(\mathsf{B}_U)$. Thus, the right hand side of these two expressions 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 (c.f. Lemma 3.6), this amplification via the s-wide walks will be analyzed by bounding the norm decay for steps of length s using Lemma 4.10 similarly to [BATS08, TS17]. We will use the shorthand $\mathsf{L}_i := \mathring{\mathsf{X}}_i \mathring{\Pi}_{\mathsf{f}} \mathring{\mathsf{A}}_Y$. The goal is to bound $\|\mathsf{L}_{s-1} \cdots \mathsf{L}_0\|_{\mathrm{op}}$ which controls the bias of the set obtained by s-wide walks (c.f. proof of Eq. (8)). Equivalently, we will bound $\langle \prod_i \mathsf{L}_i \mathbf{v}_0, \mathbf{w}_s \rangle$ for any unit vectors[17] $v_0, w_s \in \mathcal{XY}_{\mathcal{H}}$. We will use the orthogonal decomposition for the space $\mathcal{XY}_{\mathcal{H}} = \mathcal{X}_{\mathcal{H}} \otimes \mathbb{C}[V_Y]$ as $\mathcal{XY}_{\mathcal{H}} = \mathcal{XY}_{\mathcal{H}}^{\|} \oplus \mathcal{XY}_{\mathcal{H}}^{\perp}$ where $\mathcal{XY}_{\mathcal{H}}^{\|} := \mathrm{span}\{z \otimes \vec{1} \mid z \in \mathcal{X}_{\mathcal{H}}\}$.

For $i \geq 1$, we inductively define the vectors v_i, w_i, z_i and bound their norms[18],

$$v_i = \mathsf{L}_{i-1} v_{i-1}^{\perp}, \qquad z_{s-i} = \left(\mathring{\mathsf{X}}_{s-i} \mathring{\Pi}_{\mathsf{f}}\right)^* w_{s-i+1}, \qquad w_{s-i} = \left(\mathring{\mathsf{A}}_Y\right)^* z_{s-i}^{\perp} \tag{10}$$

$$\|v_i\| \leq \lambda(Y)^i, \qquad \|z_{s-i}\| \leq \lambda(Y)^{i-1}, \qquad \|w_{s-1}\| \leq \lambda(Y)^i. \tag{11}$$

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

[16] 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 $\mathcal{D}_X(\mathsf{B}_U)$ is the same as the random walk distribution on X.

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

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

Lemma 4.11. *For any v_0, w_s and $0 \le r \le s-2$ we have,*

$$\begin{aligned}
\mathsf{L}_{s-1}\cdots\mathsf{L}_0 v_0 &= v_s + \sum_{i=0}^{s-1} \mathsf{L}_{s-1}\cdots\mathsf{L}_i v_i^{\|} \\
\mathsf{L}_{s-1}^* w_s &= w_{s-1} + z_{s-1}^{\|} \\
\mathsf{L}_r^*\cdots\mathsf{L}_{s-1}^* w_s &= w_r + z_r^{\|} + \sum_{i=r+1}^{s-1} \mathsf{L}_r^*\cdots\mathsf{L}_{i-1}^* z_i^{\|}
\end{aligned}$$

Theorem 4.1 (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 $2\lambda(X) + \|\mathbb{E}_{x\in V_X}[\mathsf{f}_x]\|_{\mathrm{op}} \le \lambda(Y)^2$ and $\max_{x\in V_X}\|\mathsf{f}_x\|_{\mathrm{op}} \le 1$, we have*

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

Proof. As discussed earlier, $\left\|\mathbb{E}_{(x_0,\cdots,x_t)\in W_t}[\mathsf{f}_{x_t}\cdots\mathsf{f}_{x_0}]\right\|_{\mathrm{op}} \le \|\mathsf{L}_t\cdots\mathsf{L}_0\|_{\mathrm{op}} \le \|\mathsf{L}_{s-1}\cdots\mathsf{L}_0\|_{\mathrm{op}}^{\lfloor t/s\rfloor}$.

$$\begin{aligned}
\langle \mathsf{L}_{s-1}\cdots\mathsf{L}_0 v_0, w_s\rangle &= \langle v_s, w_0\rangle + \sum_{r=0}^{s-1}\left\langle \mathsf{L}_{s-1}\cdots\mathsf{L}_r v_r^{\|}, w_s\right\rangle \\
&= \langle v_s, w_s\rangle + \sum_{r=0}^{s-1}\left\langle v_r^{\|}, \mathsf{L}_r^*\cdots\mathsf{L}_{s-1}^* w_s\right\rangle \\
&= \langle v_s, w_s\rangle + \sum_{i=0}^{s-1}\left\langle v_r^{\|}, w_r + z_r^{\|}\right\rangle + \sum_{r=0}^{s-2}\sum_{i=r+1}^{s-1}\left\langle v_r^{\|}, \mathsf{L}_r^*\cdots\mathsf{L}_{i-1}^* z_i^{\|}\right\rangle \\
&= \langle v_s, w_s\rangle + \sum_{i=0}^{s-1}\left\langle v_r^{\|}, z_r^{\|}\right\rangle + \sum_{r=0}^{s-2}\sum_{i=r+1}^{s-1}\left\langle v_r^{\|}, \mathsf{L}_r^*\cdots\mathsf{L}_{i-1}^* z_i^{\|}\right\rangle.
\end{aligned}$$

The last step uses $\left\langle v_r^{\|}, w_r\right\rangle = \left\langle \mathring{\mathsf{A}}_Y v_r^{\|}, z_r^{\perp}\right\rangle = 0$. Using Eq. (11), we get $\left\langle v_r^{\|}, z_r^{\|}\right\rangle \le \lambda(Y)^{s-1}$. To bound the last term, we finally use Lemma 4.10. Let $v_r^{\|} = v_r' \otimes \vec{1}$, and $z_i^{\|} = z_i' \otimes \frac{1}{|V_Y|}\vec{1}$. Then,

$$\begin{aligned}
\left\langle v_r^{\|}, \mathsf{L}_r^*\cdots\mathsf{L}_{i-1}^* z_i^{\|}\right\rangle &= \left\langle v_r', \left(\mathring{\mathsf{A}}_X\Pi_{\mathsf{f}}\right)^{i-r} z_i'\right\rangle && \text{(Using Lemma 4.10)} \\
&\le \left\|\left(\mathring{\mathsf{A}}_X\Pi_{\mathsf{f}}\right)^{i-r}\right\|_{\mathrm{op}} \|z_i'\|\,\|v_r'\| \\
&\le \lambda(Y)^{2\lfloor\frac{i-r}{2}\rfloor}\lambda(Y)^{r+s-i-1} \le \lambda(Y)^{s-3},
\end{aligned}$$

where the penultimate inequality uses Theorem 3.1 and plugs in the assumption that $2\lambda(X) + \|\mathbb{E}_{x\in V_X}[\mathsf{f}_x]\|_{\mathrm{op}} \le \lambda(Y)^2$. 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[19] S. The final set is obtained by multiplying elements along each $(t-1)$-length walks on the s-wide replacemement 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 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 $\mathrm{SL}_2(p)$ constructed using Corollary 3.11. By Example 4.3, 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 $\mathbb{Z}_2^n$ and therefore by Lemma 4.8, it is pseudorandom. 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[20].

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

Every other parameter is a function of s.

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

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

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

Now, we mention the central claim that we need from our choice of parameters.

Claim 4.12. *The selection of the parameters above implies the following bounds on t,*

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

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

Lemma 4.13. *The number of walks of length $t-1$ on the s-wide replacement product of X and Y is $O(|S|\,/\lambda^{2+\beta})$.*

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

$$|V(X)|\,|V(Y)| \cdot d_2^{(t-1)} \;=\; n' \cdot d_1^s \cdot d_2^{(t-1)} = n' \cdot d_2^{(t-1)+5s}$$

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

[20] **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

$$= \Theta\left(|S| \cdot d_2^{(t-1)+5s+5}\right)$$
$$= O\left(|S| \cdot d_2^{(1+5\alpha)(t-1)}\right).$$

which from Claim 4.12 (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 4.14. *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. Let ρ be any non trivial irreducible representation of a group G.

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

Theorem 4.15 (Almost Ramanujan Expanders I)**.** *Let $\mathrm{Cay}(G,S)$ be λ_0-expander with constant $\lambda_0 \in (0,1)$. For every function*[21] *$\beta(\lambda) > 0$, 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 $|S|\,d_2^5$.

The s-wide walk Obtain an (n', d_1, λ_1) Cayley graph X as explained before. We add $n' - n = O(n^{8/9})$ copies of the identity to S_1 to obtain S_2. By Lemma 4.14 and the assumption that $s \ge 2^{10}$, S_2 is a $\lambda_2^2/3 + O(n^{-1/9}) \le 2\lambda_2^2/3$-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

$$O(t\log(d_2)) \le O((1+10\alpha)\log(1/\lambda)) \qquad \text{(Using Claim 4.12 [ii])}$$

[21] For a first reading, it can 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.

$$\leq O(\log(1/\lambda)) \qquad \text{(By the assumption that } \alpha \leq 1/128\text{)}$$

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 4.1 to the function ρ on $S_2 \leftrightarrow V(X)$. The condition that $2\lambda(X) + \left\|\mathbb{E}_{g\sim S_2}[\rho(g)]\right\|_{\mathrm{op}} \leq \lambda(Y)^2$ 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,

$$\begin{aligned}
\left\|\mathop{\mathbb{E}}_{g\in S'}[\rho(g)]\right\|_{\mathrm{op}} &:= \left(\lambda_2^s + s\cdot\lambda_2^{s-1} + s^2\cdot\lambda_2^{s-3}\right)^{\lfloor (t-1)/s\rfloor} \\
&\leq \left(3s^2\lambda_2^{s-3}\right)^{((t-1)/s)-1} && \left(\text{Using } \lambda_2 = \frac{20s^2\log s}{s^{2s^2}} \leq \frac{1}{3s^2}\right) \\
&\leq \left(\lambda_2^{s-4}\right)^{(t-1-s)/s} \\
&\leq \lambda_2^{(1-5/s)(1-s/(t-1))(t-1)} \\
&\leq \lambda_2^{(1-5\alpha)(1-\alpha)(t-1)} && \text{(Using Claim 4.12 [i])} \\
&= \lambda_2^{(1-5\alpha)(1-\alpha)(t-1)} \leq \lambda, && \text{(From the choice of } t\text{)} \quad \blacksquare
\end{aligned}$$

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 (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 $\mathsf{P} = \rho_{\mathrm{def}}(\sigma)$ 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 allows us amplify Kassabov's generating set to a close to optimum bound showing that the symmetric group has explicit almost Ramanujan Cayley expanders.

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 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[22] 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 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* - $(\rho_{\mathrm{def}}(\sigma), \mathbb{C}^n)$ for Sym_n is defined as the representation that maps a permutation to the matrix defining it More formally, $\rho_{\mathrm{def}}(\sigma)e_i = e_{\sigma(i)}$. It is a fact that $\mathcal{V}_{\mathrm{def}} = \mathcal{V}_{triv} \oplus \mathcal{V}_{\mathrm{standard}}$ where $\mathcal{V}_{\mathrm{standard}}$ is an irreducible representation. Therefore, if we are given a set $\{\mathsf{P}_1, \cdots, \mathsf{P}_r\}$ of permutation matrices, we can identify a set $S = \{\sigma_1, \cdots, \sigma_r\} \subseteq \mathrm{Sym}_n$ such that $\rho_{\mathrm{def}}(\sigma_i) = \mathsf{P}_i$.

Corollary 5.1 (Permutation Amplification)**.** *Let $\mathsf{P} = \{\mathsf{P}_1, \cdots, \mathsf{P}_r\}$ be a collection of permutation matrices such that $\lambda(\mathbb{E}_{i\sim[r]}[\mathsf{P}_i]) \le \lambda_0$. Then, for any $\lambda \in (0,1)$, we can explicitly construct a collection P' such that*

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

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

Proof. Let $\mathsf{P}_i = \sigma_i$. Applying Theorem 4.1 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 $k = O_{\lambda_0}\left(\log(1/\lambda)\right)$. By the decomposition of the defining representation, we have that

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

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 $\mathcal{V}_{\mathrm{standard}}$. ∎

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* $\mathsf{P}_1, \ldots, \mathsf{P}_d \in \mathbb{R}^{n\times n}$ *such that*

$$\mathsf{A}_X = \frac{1}{d}\sum_{j=1}^{d} \mathsf{P}_j.$$

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*-expander 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

$$\mathsf{A}_{X_i} = \frac{1}{d_0} \sum_{j=1}^{d_0} \mathsf{P}_j,$$

where $\mathsf{P}_1, \ldots, \mathsf{P}_{d_0}$ are permutation matrices, then the adjacency matrix of X'_i is

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

where each P'_j is the product of at most $k = O_{\lambda_0}(\log(1/\lambda))$ permutation matrices among $\mathsf{P}_1, \ldots, \mathsf{P}_{d_0}$.

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 an usual degree-d expander graph $X = (V, E)$ is given by d permutation matrices acting on a vector space $\mathbb{R}[V]$, 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 [23] 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 construction cannot be efficiently implemented and thus used in applications like [AHL+14] which rely on existing explicit constructions 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[24] $\sum_{i=1}^{d} B_i^\dagger B_i = \mathsf{I}$.
- ("Expansion") The second largest eigenvalue[25] of Φ as a linear map is $\leq \lambda$.

[23] In the Hilbert–Schmidt inner product.

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

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

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 expander 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 [26]) such that the edges set E is the following disjoint union

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

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 it constant). We stress that our almost Ramanujan bound is with respect to the usual notion of 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 $\lambda(X) = \max\{|\lambda_2(\mathsf{A}_X)|, |\lambda_n(\mathsf{A}_X)|\}$.

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,

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

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

$$\mathsf{A}_Y = \sum_{i=1}^{d_0} \mathsf{P}_i \,,$$

where $\mathsf{P}_i = \mathsf{M}_{f_i} + \mathsf{M}_{\overline{f_i}}$ and $(\mathsf{M}_{f_i})_{x,y} = \mathbb{1}\left[f_i(x) = y\right]$.

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 $\lambda(X_i) \leq \lambda$.*

Proof. Let $\{X'_n\}_{n\in\mathbb{N}}$ be the family in Corollary 5.12. Fix $X = X'_n$ and let $\mathsf{P}_1, \ldots, \mathsf{P}_{d_0} \in \mathbb{R}^{n\times n}$ be the permutation matrices guaranteed by Corollary 5.12, where each $\mathsf{P}_i = \mathsf{M}_{f_i} + \mathsf{M}_{\overline{f}_i}$. 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 $\mathsf{P}_1, \ldots, \mathsf{P}_{d_0}$ and so we obtain

$$\begin{aligned}\mathsf{P}_{i_1} \cdots \mathsf{P}_{i_k} &= \sum_{g_i \in \{f_i, \overline{f}_i\}} \mathsf{M}_{g_{i_1}} \cdots \mathsf{M}_{g_{i_k}} \\ &= \sum_{g_i \in \{f_i, \overline{f}_i\}} M_{g_{i_1} \circ g_{i_2} \circ \cdots g_{i_k}} \,,\end{aligned}$$

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 $\mathsf{P}_{i_1} \cdots \mathsf{P}_{i_k} \neq 0$). 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 suitable notion of "spectral gap" for discrete groups (and so it is related to bias). These groups can have infinitely many irreducible representations

on more general Hilbert spaces. 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.

Let G be a discrete group and S a finite set 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

$$\begin{aligned}
\overline{\mathcal{K}}(G,S) &:= \inf\{\overline{\mathcal{K}}(G,S,\rho) \mid (\rho,\mathcal{H}) \text{ irreducible and non-trivial}\} \\
\overline{\mathcal{K}}(G,S,\rho) &:= \inf_{v\in\mathcal{H}:\ \|v\|_2=1} \frac{1}{|S|}\sum_{g\in S} \|\rho(g)v - v\|_2^2 \\
&= \inf_{v\in\mathcal{H}:\ \|v\|_2=1} \frac{1}{|S|}\sum_{g\in S} 2 - 2\langle \rho(g)v, v\rangle \\
&= \inf_{v\in\mathcal{H}:\ \|v\|_2=1} 2 - 2\left\langle \mathop{\mathbb{E}}_{g\sim S}[\rho(g)]\, v, v\right\rangle \\
&= 2\left(1 - \left\| \mathop{\mathbb{E}}_{g\sim S}[\rho(g)]\right\|_{\mathrm{op}}\right).
\end{aligned}$$

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. We remark that the above amplification can also similarly improve the constant of property (τ) (the latter being a weaker version of property (T)).

We now 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

The idea of dimension expanders arose in [BISW01] motivated by applications in theoretical computer science. A conjectured construction of dimension expanders based on irreducible representations was suggested by Wigderson to hold over every field. It was subsequently established by Lubotzky and Zelmanov [LZ08] for fields of characteristic zero.

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

For an irreducible unitary representation ρ, there exists an associated irreducible representation[27] adj_ρ. The construction in [LZ08] relates expansion with the Kazhdan constant,

Proposition 5.17 (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 $\dim(\mathcal{W})$ 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 $\mathcal{W}$) 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 5.18. *Let $\mathcal{W}, \mathcal{W}' \subseteq \mathbb{C}^d$ be two vector spaces. Let P, P' be orthogonal projectors onto $\mathcal{W}, \mathcal{W}'$, respectively. Then*

$$\mathrm{Re}\,\mathrm{Tr}(\mathsf{P}\mathsf{P}') = \mathrm{Tr}(\mathsf{P}\mathsf{P}') \geq \dim(\mathcal{W} \cap \mathcal{W}').$$

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

Remark 5.19. 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].

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 and Seress conjecture [BS88]) or with respect to some constant size generating set as in [BKL89]. Here, we explore the former 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.

[27] Let $\mathfrak{sl}_n(\mathbb{C}) = \{\mathrm{tr}(\mathsf{A}) = 0 \mid \mathsf{A} \in \mathsf{M}_n(\mathbb{C})\}$. Equip the space with the Frobenius inner product defined as $\langle \mathsf{A}, \mathsf{B}\rangle = \mathrm{tr}(\mathsf{A}^\dagger\mathsf{B})$ where $\mathsf{A}^\dagger$ 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 $\mathrm{adj}_\rho(g)\cdot\mathsf{A} = \rho(g)\cdot\mathsf{A}\cdot\rho(g)^{-1}$. Since conjugation preserves the trace, $\mathfrak{sl}_n$ is closed under the representation. Moreover, it is unitary as

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

Lemma 5.20. *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 $\lambda, \beta > 0$. 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

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

This implies that $\mathsf{A}_i e_g$ 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] is to apply a suitable version of operator expander mixing lemma, where they obtain a dependence factor $1/\lambda^{11}$ in the degree. We show that this approach 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 case [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. A similar result was first derived in [CMR13].

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

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

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, the projection map $\mathsf{P}_{\mathcal{H}} := \mathcal{X}_{\mathcal{H}} \to \mathcal{H}$ defined via $w \otimes x \mapsto w$ and a lifting map $\mathsf{L}_{\mathcal{H}} := \mathcal{H} \to \mathcal{X}_{\mathcal{H}}$ defined via $w \mapsto \frac{1}{|V_X|} w \otimes \vec{1}$. We will need again that $\|\mathsf{P}_{\mathcal{H}}\|_{\mathrm{op}}\|\mathsf{L}_{\mathcal{H}}\|_{\mathrm{op}} = 1$.

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.*

$$\mathop{\mathbb{E}}_{(x,x')\in E}[\mathsf{f}_{x'}\cdot\mathsf{f}_x] = \mathsf{P}_{\mathcal{H}}\Pi_z\mathring{\mathsf{A}}_X\Pi_z\mathsf{L}_{\mathcal{H}}.$$

$$\mathop{\mathbb{E}}_{x,x'\in V}\left[\mathsf{f}_{x'}\cdot\mathsf{f}_x\right] \;=\; \mathsf{P}_{\mathcal{H}}\Pi_z\mathring{\mathsf{J}}_X\Pi_z\mathsf{L}_{\mathcal{H}}.$$

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

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

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

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

concluding the proof. ∎

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

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

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.

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

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

Let $X_0 = (V_0, E_0)$ be an explicit expander via Theorem 3.5, 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,

$$\lambda(G, S_1) \leq \lambda_0^2 + \varepsilon_0 \leq \lambda_0(1 - \gamma_0), \quad |S_0| = 9\,|S|\,/\varepsilon_0^2 = O(|S|)$$

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 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 3.5, 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 $\lambda_i \leq 2^{-(a-1)(2^i)}$ 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,

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

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

[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. 24, 41

[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. 26, 29

[Alo21] Noga Alon. Explicit expanders of every degree and size. *Combinatorica*, February 2021. doi:10.1007/s00493-020-4429-x. 12, 15, 16

[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. 8

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

[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, 26, 29

[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, 6, 26, 29

[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, 8, 17, 22

[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, 27, 29

[BISW01] B. Barak, R. Impagliazzo, A. Shpilka, and A. Wigderson. Dimension expanders. unpublished, 2001. 4, 32

[BKL89] László Babai, William M. Kantor, and A. Lubotsky. Small-diameter cayley graphs for finite simple groups. *Eur. J. Comb.*, 10, 1989. 33

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

[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. 33

[BY13] Jean Bourgain and Amir Yehudayoff. Expansion in sl 2(ℝ)and monotone expanders. *Geometric and Functional Analysis*, 23(1), 2013. 4, 30

[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. 16

[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. 6, 7, 9, 10, 34

[DS09] Zeev Dvir and Amir Shpilka. Towards dimension expanders over finite fields. *Combinatorica*, 31(3), sep 2009. 4, 30, 33

[DW10] Zeev Dvir and Avi Wigderson. Monotone expanders: Constructions and applications. *Theory of Computing*, 6(12), 2010. 30, 33

[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. 33

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

[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, 8

[Har07] Aram W. Harrow. Quantum expanders from any classical cayley graph expander. *Quantum Information & Computation*, 2007. 4, 26, 29, 30, 33

[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, 26, 29

[Has07b] M. B. Hastings. Entropy and entanglement in quantum ground states. *Phys. Rev. B*, 2007. 26, 29

[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, 29

[HH09] M. B. Hastings and A. W. Harrow. Classical and quantum tensor product expanders. *Quantum Info. Comput.*, 2009. 26, 29

[HLW06] Shlomo Hoory, Nathan Linial, and Avi Wigderson. Expander graphs and their applications. *Bull. Amer. Math. Soc.*, 43(04):439–562, August 2006. 1, 27

[JM21] Akhil Jalan and Dana Moshkovitz. Near-optimal cayley expanders for abelian groups, 2021. arXiv:2105.01149. 6, 7

[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. 40

[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, 9, 26, 30

[KKN21] Marek Kaluba, Dawid Kielak, and Piotr W. Nowak. On property (T) for $\mathrm{Aut}(F_n)$ and $\mathrm{SL}_n(\mathbb{Z})$. *Annals of Mathematics*, 193(2):539 – 562, 2021. doi:10.4007/annals.2021.193.2.3. 27

[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. 27

[LPS88] Alexander Lubotzky, R. Phillips, and Peter Sarnak. Ramanujan graphs. *Combinatorica*, 8:261–277, 1988. 1, 16

[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. 16

[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. 5, 27, 32, 33

[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, 27

[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. 1, 2

[MW04] Roy Meshulam and Avi Wigderson. Expanders in group algebras. *Combinatorica*, 24, 2004. 8

[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

[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. 32

[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, 16

[RSW06] Eyal Rozenman, Aner Shalev, and Avi Wigderson. Iterative construction of cayley expander graphs. *Theory of Computing*, 2(5):91–120, 2006. 3, 8, 9

[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, 6, 8, 17

[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. 16

[SS96] L. L. Scott and J. P. Serre. *Linear Representations of Finite Groups*. Graduate Texts in Mathematics. Springer New York, 1996. 6

[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, 11, 16, 17, 19, 20, 21, 22, 23, 34, 40, 41

[TSD18] Amnon Ta-Shma and Dean Doron. Combinatorial constructions of expanders. the zig-zag product. Lecture notes, 2018. 16

[Vad12] Salil P. Vadhan. *Pseudorandomness*. Now Publishers Inc., 2012. 31

[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 3.11 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 $\frac{c}{\lambda_1^{2*4.1}} \leq \frac{cd_2^{4.1}}{b_2^{8.2}} \leq d_2^5$. We increase its degree to d_2^5 by taking multiple copies of the generating set which doesn't change bias[28]. Thus, we obtain a (n', d_1, λ_1)-graph where $n' = n + O(n^{8/9})$.

[28] 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.

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[29]. The set A of generators comes from a small bias code derived 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.
$Y : (n_2, d_2, \lambda_2), \quad n_2 = d_2^{5s}, \quad d_2 = s^{4s}, \quad \lambda_2 = \frac{b_2}{\sqrt{d_2}}, \quad b_2 = 5s\log d_2$
$X : (n', d_1, \lambda_1), \quad n' \approx n = O(\lvert S\rvert\, d_2^5), \quad d_1 = d_2^5, \quad \lambda_1 = \frac{\lambda_2^2}{10}$
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 4.12. *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,

$$\lambda_2^{(1-5\alpha)(1-\alpha)^2(t-1)} = \lambda_2^{(1-5\alpha)(1-\alpha)(t-2)\left(\frac{(1-\alpha)}{1-1/(t-1)}\right)} \tag{12}$$
$$\geq \lambda_2^{(1-5\alpha)(1-\alpha)(t-2)} \qquad (t-1 \geq s = 1/\alpha) \tag{13}$$
$$\geq \lambda \qquad \text{(From the choice of minimal } t\text{)} \tag{14}$$

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

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

[29] 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.

We obtain (ii)

$$\begin{aligned}
(d_2)^{(t-1)} &\leq \lambda_2^{\frac{-(t-1)}{1-2\alpha}} \\
&\leq \lambda^{\frac{-2}{(1-2\alpha)(1-5\alpha)(1-\alpha)^2}} && \text{(Using Eq. (14))} \\
&\leq \lambda^{-2(1+10\alpha)}. && \blacksquare
\end{aligned}$$