A Combinatorial Construction of Almost-Ramanujan Graphs Using the Zig-Zag Product

Avraham Ben-Aroya *
Department of Computer Science
Tel-Aviv University
Tel-Aviv 69978, Israel

Amnon Ta-Shma †
Department of Computer Science
Tel-Aviv University
Tel-Aviv 69978, Israel

ABSTRACT

Reingold, Vadhan and Wigderson [21] introduced the graph *zig-zag* product. This product combines a large graph and a small graph into one graph, such that the resulting graph inherits its size from the large graph, its degree from the small graph and its spectral gap from both. Using this product they gave the first fully-explicit combinatorial construction of expander graphs. They showed how to construct D–regular graphs having spectral gap $1-O(D^{-\frac{1}{3}})$. In the same paper, they posed the open problem of whether a similar graph product could be used to achieve the almost-optimal spectral gap $1-O(D^{-\frac{1}{2}})$.

In this paper we propose a generalization of the zig-zag product that combines a large graph and several small graphs. The new product gives a better relation between the degree and the spectral gap of the resulting graph. We use the new product to give a fully-explicit combinatorial construction of D–regular graphs having spectral gap $1-D^{-\frac{1}{2}+o(1)}$.

Categories and Subject Descriptors

F.m [**Theory of Computation**]: Miscellaneous;
G.2 [**Discrete Mathematics**]: Graph Theory

General Terms

Theory

Keywords

Expander graphs, zig-zag product

*abrhambe@tau.ac.il. Supported by the Adams Fellowship Program of the Israel Academy of Sciences and Humanities, and by the European Commission under the Integrated Project QAP funded by the IST directorate as Contract Number 015848.

†amnon@tau.ac.il. Supported by Israel Science Foundation grant 217/05 and by USA Israel BSF grant 2004390.

STOC'08, May 17–20, 2008, Victoria, British Columbia, Canada.

1. INTRODUCTION

Expander graphs are graphs of low-degree and high connectivity. There are several ways to measure the quality of expansion in a graph. One such way measures *set expansion*: given a not too large set S, it measures the size of the set $\Gamma(S)$ of neighbors of S, relative to $|S|$. Another way is *(Rényi) entropic expansion*: given a distribution π on the vertices of the graph, it measures the amount of (Rényi) entropy added in $\pi' = G\pi$. This is closely related to measuring the *algebraic expansion* given by the spectral gap of the adjacency matrix of the graph (see Section 2 for formal definitions, and [9] for an excellent survey).

Pinsker [19] was the first to observe that constant-degree random graphs have almost-optimal set expansion. Explicit graphs with algebraic expansion were constructed, e.g., in [14, 8, 11]. This line of research culminated by the works of Lubotzky, Philips and Sarnak [13], Margulis [15] and Morgenstern [17] who explicitly constructed Ramanujan graphs, i.e., D–regular graphs achieving spectral gap of $1-2\frac{\sqrt{D-1}}{D}$. Alon and Boppana (see [18]) showed that Ramanujan graphs achieve almost the best possible algebraic expansion, and Friedman [7] showed that random graphs are almost Ramanujan (we cite his result in Theorem 6). Several works [6, 3, 1, 12] showed intimate connections between set expansion and algebraic expansion. We refer the reader, again, to the excellent survey paper [9].

Despite the optimality of the constructions above, the search for new expander constructions is still going on. This is motivated, in part, by some intriguing remaining open questions. Another important motivation comes from the fact that expanders are a basic tool in complexity theory, with applications in many different areas. The above mentioned explicit constructions rely on deep mathematical results, while it seems natural to look for a purely combinatorial way of constructing and analyzing such objects. This goal was achieved recently by Reingold, Vadhan and Wigderson [21] who gave a *combinatorial* construction of algebraic expanders. Their construction has an intuitive analysis and is based on elementary linear algebra. The heart of the construction is a new graph product, named the zig-zag product, which we explain soon.

Following their work, Capalbo et. al. [5] used a variant of the zig-zag product to explicitly construct D–regular graphs with set expansion close to D (rather than $D/2$ that is guaranteed in Ramanujan graph constructions). Also, in a seemingly different setting, Reingold [20] gave a log-space algorithm for undirected connectivity, settling a long-standing

open problem, by taking advantage, among other things, of the simple combinatorial composition of the zig-zag product.

Several works studied different aspects of the zig-zag composition. Alon et. al [2] showed, somewhat surprisingly, an algebraic interpretation of the zig-zag product over non-Abelian Cayley graphs. This lead to new iterative constructions of Cayley expanders [16, 22], which were once again based on algebraic structures. While these constructions are not optimal, they contribute to our understanding of the power of the zig-zag product.

The expander construction presented in [21] has spectral gap $1 - O(D^{-\frac{1}{4}})$. As was noted in that paper, this is the best possible for the zig-zag product, because the zig-zag product takes two steps on a "small graph", and as we explain soon, one of these steps may be completely wasted. It is still possible, however, that a variant of the zig-zag product gives better expansion. Indeed, [5] modified the zig-zag product to get close to optimal *set* expansion. Also, [21] considered a "derandomized" variant of the zig-zag product, where one takes two steps on the small graph, one step on the large graph and then two more steps on the small graph, but where the first and last steps are correlated (in fact, identical). They showed this product has spectral gap $1 - O(D^{-\frac{1}{3}})$. They posed the open problem of finding a variant of the zig-zag product with almost-optimal spectral gap $1 - O(D^{-\frac{1}{2}})$. In fact, any combinatorial construction achieving the above spectral gap is yet unknown.

Our main result is a new variant of the zig-zag product, where instead of composing one large graph with one small graph, we compose one large graph with *several* small graphs. The new graph product we develop exhibits a better relationship between its degree and its spectral gap and retains most of the other properties of the standard zig-zag product. In particular, we use this product to construct an iterative family of D–regular expanders with spectral gap $1-D^{-\frac{1}{2}+o(1)}$, thus nearly resolving the open problem of [21].

Bilu and Linial [4] gave a different iterative construction of algebraic expanders that is based on 2-lifts. Their construction has close to optimal spectral gap $1-O((\log^{1.5} D)\cdot D^{-\frac{1}{2}})$. We mention, however, that their construction is only *mildly-explicit* (meaning that, given N, one can build a graph G_N on N vertices in poly(N) time). Our construction, as well as [21], is *fully-explicit* (meaning that given $v \in V = [N]$ and $i \in [D]$ one can output the i'th neighbor of v in poly(log(N)) time). This stronger notion of explicitness is crucial for some applications.

1.1 An intuitive explanation of the new product

1.1.1 The zig-zag product

Let us review the zig-zag product of [21]. We begin by first describing the replacement product between two graphs, where the degree, D_1, of the first graph G_1 equals the number of vertices, N_2, of the second graph H. In the resulting graph, every vertex v of G_1 is replaced with a cloud of D_1 vertices $\{(v,i)\}_{i\in[D_1]}$. We put an "inter-cloud" edge between (v,i) and (w,j) if $e=(v,w)$ is an edge in G_1 and e is the i'th edge leaving v, and the j'th edge leaving w. We also put copies of H on each of the clouds, i.e., for every v we put an edge between (v,i) and (v,j) if (i,j) is an edge in H.

The zig-zag product graph corresponds to 3-step walks on the replacement product graph, where the first and last steps are inner-cloud edges and the middle step is an inter-cloud edge. That is, the vertices of the zig-zag product graph are the same as the replacement product graph, and we put an edge between (v,i) and (w,j) if one can reach (w,j) from (v,i) by taking a 3-step walk: first an H edge on the cloud of v, then an inter-cloud edge to the cloud of w, and finally an H edge on the cloud of w. Roughly speaking, the resulting graph inherits its size from the large graph, its degree from the small graph, andits spectral gap from both.

Before we proceed, let us adopt a slightly more formal notation. We denote by V_1 the set of vertices of G_1 and its cardinality by N_1. Similarly, V_2 is the set of vertices of H and its cardinality is $N_2 = D_1$. The degree of H is denoted by D_2. We associate each of the graphs with its normalized adjacency matrix, and we let $\bar{\lambda}(\cdot)$ denote the second-largest eigenvalue of a given graph. We view G_1 as a linear operator on a dim-N_1 vector space. For a vertex $v \in V$, we denote by $\overrightarrow{v}$ the vector that its v–coordinate is 1 and all its other coordinates are 0. Next, we define an operator $\dot{G}_1$ on a vector space $\mathcal{V}$ of dimension $N_1 \cdot N_2$ that is the adjacency matrix of the inter-clouds edges (i.e., in the notation above, $\dot{G}_1(v,i) = (w,j)$). We also let $\tilde{H} = I \otimes H$, i.e., it is an H step on the cloud coordinates, unchanging the cloud itself. In this notation, the adjacency matrix of the zig-zag product is $\tilde{H}\dot{G}_1\tilde{H}$ and our task is to bound its second-largest eigenvalue. Notice that $\dot{G}_1$ is a permutation (in fact, a perfect matching).

Any distribution on $V = V_1 \times V_2$ can be thought of as giving each cloud some weight, and then distributing that weight within the cloud. Thus, the distribution has two components; the first corresponds to a cloud (a G_1 vertex) and the second corresponds to a position within a cloud (a H vertex). To give an intuition why $\tilde{H}\dot{G}_1\tilde{H}$ is an expander we analyze two extreme cases. In the first case, the distribution within each cloud is entropy-deficient (and hence far from uniform) and the first $\tilde{H}$ application already adds entropy. In the second case, the distribution within each cloud is uniform. In this case the first $\tilde{H}$ application does not change the distribution at all. However, as we are uniform on the clouds, applying $\dot{G}_1$ on the distribution propagates the entropy from the second component to the first one (this follows from the fact that G_1 is an expander). Any permutation, and $\dot{G}_1$ in particular, does not change the overall entropy of the distribution. Thus, we conclude that the entropy added to the first component was taken from the second component, and hence the second component is now entropy-deficient. Therefore, the second $\tilde{H}$ application adds entropy.

The formal analysis in [21] works by decomposing $\mathcal{V}$ into two subspaces: The first subspace, $\mathcal{V}^{||}$, includes all the vectors x that are uniform over clouds, i.e., all vectors of the form $x = x^{(1)} \otimes \mathbf{1}$, where $x^{(1)}$ is an arbitrary N_1-dimensional vector and $\mathbf{1}$ is the (normalized) all 1's vector. The second subspace, $\mathcal{V}^{\perp}$, is its orthogonal complement. Two observations are made. First, that $\left\langle x^{(1)} \otimes \mathbf{1}, \tilde{H}\dot{G}_1\tilde{H}(y^{(1)} \otimes \mathbf{1}) \right\rangle$ equals to $\left\langle x^{(1)}, G_1 y^{(1)} \right\rangle$ and therefore when $x, y \in \mathcal{V}^{||}$ and $x \perp \mathbf{1}$ we have that $\left\langle x, \tilde{H}\dot{G}_1\tilde{H}y \right\rangle \leq \bar{\lambda}(G_1)\,|\langle x,y\rangle|$. The second observation is that when either x or y belong to $\mathcal{V}^{\perp}$ we have that $\left\langle x, \tilde{H}\dot{G}_1\tilde{H}y \right\rangle \leq \bar{\lambda}(H)\,|\langle x,y\rangle|$. There-

fore, by linearity, we get that $\tilde{H}\dot{G}_1\tilde{H}$ maps vectors $x \perp \mathbf{1}$ in $\mathcal{V}$ to vectors with length smaller by a factor of at least $4 \cdot \min\ \bar{\lambda}(G_1), \bar{\lambda}(H)$. A more careful analysis yields a better bound.

The non-optimality of the zig-zag product comes from the following observation. The degree of the zig-zag graph is D_2^2 (where D_2 is the degree of H). However, when $x \in \mathcal{V}^{||}$ we have that $\tilde{H}\dot{G}_1\tilde{H}x = \tilde{H}\dot{G}_1 x$, and the operator $\tilde{H}\dot{G}_1$ corresponds to taking only *a single* step on H. Namely, we pay in the degree for two steps, but (on some vectors) we get the benefit of only one step. Therefore, the best we can hope for is getting the Ramanujan value for D_2, namely, $2\frac{\sqrt{D_2-1}}{D_2}$.

We would like to point out an interesting phenomena that occurs in the zig-zag product analysis. The analysis shows that $\left\langle x, \tilde{H}\dot{G}_1\tilde{H}y \right\rangle \leq \bar{\lambda}(G_1)\,|\langle x, y\rangle|$ for $x, y \in \mathcal{V}^{||}$ and $x \perp \mathbf{1}$. Thus, even though the degree of $\tilde{H}\dot{G}_1\tilde{H}$ is only $D_2^2 \ll D_1$, this part of the analysis gives us $\bar{\lambda}(G_1) \ll D_2^{-2}$. Saying it differently, when the operator acts on $x \in V^{||}$, it uses the entropy x has in each cloud, rather than the entropy that comes from the zig-zag graph degree.

1.1.2 The k-step zig-zag product

Now consider the variant of the zig-zag product where we take k steps on H rather than just 2. That is, we consider the graph whose adjacency matrix is $\tilde{H}\dot{G}_1\tilde{H}\ldots\tilde{H}\dot{G}_1\tilde{H}$ with k steps on H. How small is the second largest eigenvalue going to be? In particular, will it beat $\bar{\lambda}(H)^{k/2}$ that we get from sequential applications of the zig-zag product, or not? Obviously, the same argument as before shows that we must lose at least one $\tilde{H}$ application. Is it possible that this is indeed what we get and that the second-largest eigenvalue is of order $\bar{\lambda}(H)^{k-1}$?

The problem.

Let us consider what happens when we take three H steps. The operator we consider is therefore $\tilde{H}\dot{G}_1\tilde{H}\dot{G}_1\tilde{H}$. Given a distribution over the graph's vertices, we are asking how many of the $\tilde{H}$ applications add entropy. Suppose that the first $\tilde{H}$ application does not add entropy. This is immediately followed by $\dot{G}_1$, which (in this case) propagates entropy from the second component to the first one. Thus, the second $\tilde{H}$ application adds entropy. Now we apply $\dot{G}_1$ again. It is possible that at this stage the distribution on the second component is far from uniform. In this case $\dot{G}_1$ might cause the entropy to *propagate back* from the first component to the second component, possibly making the second component uniform again. If this happens, the third $\tilde{H}$ application does not add entropy at all. Thus, we have three H steps, but only one adds entropy.

We rephrase the problem in an algebraic language. Notice that in the zig-zag product we have just one application of $\dot{G}_1$, whereas in the new product we have $k-1$ such applications. G_1 is an operator that describes a stochastic process that randomly chooses one of D_1 possible neighbors. In contrast, $\dot{G}_1$ is a unitary operator, a permutation mapping one cloud element to another cloud element. In particular, it follows from the way $\dot{G}_1$ is defined that ${\dot{G}_1}^2 = I$. Therefore, it is possible, may be even plausible, that the second $\dot{G}_1$ step cancels the first $\dot{G}_1$ step. If that happens, we might end up with the second-largest eigenvalue of $\dot{G}_1\tilde{H}\dot{G}_1$ being a constant, completely independent of both $\bar{\lambda}(G)$ and $\bar{\lambda}(H)$.[1]

Thus, it seems that the only thing that can save us is the action of $\tilde{H}$ between two $\dot{G}_1$ steps. However, the prospects here do not look too bright, because $\dot{G}_1\tilde{H}\dot{G}_1$ is an operator acting on a large vector space of dimension N_1N_2 (recall that we think of N_2 as a constant and of N_1 as a growing parameter) while H should be a constant size graph. It seems highly unlikely that one can prove that there exists a good graph H, among the constant number of possible small graphs, such that on any vector of arbitrarily large dimension, the second application of $\dot{G}_1$ does not invert the first one.

The solution.

In order to gain more H steps we need to make sure that entropy does not flow in the wrong direction. This is achieved as follows. Whenever a $\tilde{H}$ application does not add entropy, we know that the distribution over the second component is uniform. We want to take advantage of this to make sure *all* the following $\dot{G}_1$ applications do not move entropy in the wrong direction. Thus, failure in a single $\tilde{H}$ application, guarantees success in all following $\tilde{H}$ applications.

When a $\tilde{H}$ application does not add entropy, the distribution over the second component is close to uniform. We make the second component large enough such that it can support k uniform G steps. For example, we can make the cloud size $|V_2|$ equal D_1^{4k}. The graph G_1 still has degree D_1, and we therefore need to specify how to translate a cloud vertex (from $[D_1]^{4k}$) to an edge-label (in $[D_1]$). For concreteness, let us assume we take the edge-label from the first $\log(D_1)$ bits of the cloud vertex. Now, all we need for the operator $\dot{G}_1$ to move entropy in the right direction is that the second component is uniform *only* on its first few bits.

Let us take a closer look at the situation. We start with a uniform distribution over the second component (because we are considering the case where $\tilde{H}$ fails) with about $4k\log(D_1)$ entropy. We apply $\dot{G}_1$ and up to $\log(D_1)$ entropy flows from the second component to the first one. Thus, there is still much entropy in the second component. We now apply $\tilde{H}$. Our goal is to guarantee that $\tilde{H}$ moves the entropy in the second component to the first $\log(D_1)$ bits. When this happens, the next $\dot{G}_1$ application moves more entropy from the second component to the first one, and entropy never flows in the wrong direction.

The problem is the condition we get on a "good" H seems to involve a large vector space $\mathcal{V} = \mathcal{V}_1 \otimes \mathcal{V}_2$ of dimension $N_1 \cdot D_1$, and there are only a constant number of possible graphs H on D_1^{4k} vertices (we think of D_1 and k as constants, and of N_1 as a growing parameter). The key observation here is that by enforcing an additional requirement on the graph G_1 that we soon describe, we can reduce the number of constraints, in particular making them independent of N_1.

---

[1] [23] also bound the expression $\left\langle \dot{G}_1\tilde{H}\dot{G}_1x, x \right\rangle$, when $x \in \mathcal{V}^{||}$ and $x \perp \mathbf{1}$. They express H as $H = (1-\lambda_2)J + \lambda_2 C$, where J is the normalized all one matrix and $\|C\| \leq 1$. This decomposition yields the bound $\lambda_1^2+\lambda_2$, which is useful when $\lambda_2 \ll \lambda_1$. In our case $\lambda_1 \ll \lambda_2$. Applying the decomposition on $\left\langle \tilde{H}\dot{G}_1\tilde{H}\dot{G}_1\tilde{H}x, x \right\rangle$, seems to give a bound that is larger than λ_2, which is not useful for us.

With this, the problem can be easily solved using standard probabilistic arguments.

A graph G_1 is *π-consistently labeled* [20] if for every edge $e = (v, w)$, if e is the i'th edge leaving v then e is the $\pi(i)$'th edge leaving w. In other words, we can reverse a step i by using the label $\pi(i)$.[2] We say a graph is *locally invertible* if it is π-consistently labeled for some π. That is, we can reverse a step i without knowing where we came from and where we are now. We show a natural condition guaranteeing that H is good for locally invertible G_1. The condition involves only edge labels and is therefore independent of N_1.

Armed with that we go back to the zig-zag analysis. As in [21], we decompose the vector space $\mathcal{V}$ to its parallel and perpendicular parts. However, because we have $k-1$ intermediate G_1 steps, we need to decompose not only the initial vectors, but also some intermediate vectors. Doing it carefully, we get that composing G_1 (of degree D_1 and second eigenvalue λ_1) with k graphs H_i (of degree D_2 and second eigenvalue λ_2 each) we get a new graph with degree D_2^k and second eigenvalue about $\lambda_2^{k-1} + \lambda_2^k + 2\lambda_1$. We can think of λ_1 as being arbitrarily small, as we can decrease it to any constant we wish without affecting the degree of the resulting graph. One can interpret the above result as saying that $k-1$ out of the k steps worked for us!

1.1.3 An almost-Ramanujan expander construction

We now go back to the iterative expander construction of [21] and replace the zig-zag component there with the k-step zig-zag product. Say, we wish to construct graphs of degree D, for D of the form $D = D_2^k$. Doing the iterative construction we get a degree D expander, with k steps over graphs $\{H_i\}$, each of degree D_2. Roughly speaking, the resulting eigenvalue is λ_2^{k-1} where λ_2 is the Ramanujan value for D_2, i.e., $\lambda_2 = \frac{2\sqrt{D_2-1}}{D_2}$. The optimal value we shoot for is the Ramanujan value for D, which is $\frac{2\sqrt{D-1}}{D}$. Our losses come from two different sources. First we lose one application of H out of the k applications, and this loss amounts to, roughly, $\sqrt{D_2}$ multiplicative factor. We also have a second loss of 2^{k-1} multiplicative factor emanating from the fact that $\lambda_{\mathrm{Ram}}(D_2)^k \approx 2^{k-1}\lambda_{\mathrm{Ram}}(D_2^k)$. This last loss corresponds to the fact that H^k is not Ramanujan even when H is. Balancing our losses gives:

THEOREM 1. *For every $D > 0$, there exists a fully-explicit family of graphs $\{G_i\}$, with an increasing number of vertices, such that each G_i is D–regular and $\bar{\lambda}(G_i) \leq D^{-\frac{1}{2}+O(\frac{1}{\sqrt{\log D}})}$.*

In this extended abstract we prove Theorem 1 only for degrees of a specific form. Proving Theorem 1 in its full generality is a bit more technical. This proof will appear in the full version of this paper.

1.2 Organization of the paper

In Section 2 we give preliminary definitions. Section 3 contains the formal definition of the k-step zig-zag product. Section 4 contains a proof that almost all graphs are good. Section 5 contains the analysis of the new product. Finally, in Section 6 we use the product to give an iterative construction of expanders.

[2]This should not be confused with the term *consistently labeled* (without a permutation π) which has a different meaning.

2. PRELIMINARIES

We associate a (directed or undirected) graph $G = (V, E)$ with its normalized adjacency matrix, also denoted by G, i.e., $G_{i,j} = \frac{1}{\deg(j)}$ if $(i, j) \in E$ and 0 otherwise. For a matrix G we denote by $s_i(G)$ the i'th largest singular value of G. If the graph G is regular (i.e., $\deg_{\mathrm{in}}(v) = \deg_{\mathrm{out}}(v) = D$ for all $v \in V$) then $s_1(G) = 1$. We also define $\bar{\lambda}(G) = s_2(G)$. We say a graph G is a (D, λ) graph, if it is D–regular and $\bar{\lambda}(G) \leq \lambda$. We also say G is a (N, D, λ) graph if it is a (D, λ) graph over N vertices. If G is an undirected graph then the matrix G is Hermitian, in which case there is an orthonormal eigenvector basis and the eigenvalues $\lambda_1 \geq \ldots \geq \lambda_N$ are real. In this case, $\bar{\lambda}(G) = s_2(G) = \max\{\lambda_2, |\lambda_N|\}$. We say a D–regular graph is *Ramanujan* if $\bar{\lambda}(G) \leq \lambda_{\mathrm{Ram}}(D) \stackrel{def}{=} \frac{2\sqrt{d-1}}{d}$.

We can convert a directed expander to an undirected expander simply by undirecting the edges. Say G is a (N, D, λ) directed graph. Then $U = \frac{1}{2}[G+G^\dagger]$ is an undirected graph. Also, $\mathbf{1} \stackrel{def}{=} \frac{1}{\sqrt{N}}(1, \ldots, 1)^t$ is an eigenvector of both G and $G^\dagger$ and so $s_2(U) = \frac{1}{2}s_2(G + G^\dagger) \leq s_2(G)$. It follows that U is a $(N, 2D, \lambda)$ graph.

To represent graphs, we use the rotation maps introduced in [21]. Let G be an undirected D–regular graph $G = (V, E)$. Assume that for every $v \in V$, its D outgoing edges are labeled by $[1..D]$. Let $v[i]$ denote the i'th neighbor of v in G. We define $\mathrm{Rot}_G : V \times [D] \to V \times [D]$ as follows. $\mathrm{Rot}_G(v, i) = (w, j)$ if $v[i] = w$ and $w[j] = v$. In words, the i'th neighbor of v is w, and the j'th neighbor of w goes back to v. Notice that if $\mathrm{Rot}_G(v, i) = (w, j)$ then $\mathrm{Rot}_G(w, j) = (v, i)$, i.e., Rot_G^2 is the identity mapping.

DEFINITION 1. *A graph G is* locally invertible *if its rotation map is of the form $\mathrm{Rot}_G(v, i) = (v[i], \phi(i))$ for some permutation $\phi : [d] \to [d]$. We say that ϕ is the local inversion function.*

For an n-dimensional vector x we let $|x|_1 = \sum_{i=1}^n |x_i|$ and $\|x\| = \sqrt{\langle x, x\rangle}$. We measure the distance between two distributions P, Q by $|P - Q|_1$. The operator norm of a linear operator L over a vector space is $\|L\|_\infty = \max_{x:\|x\|=1} \|Lx\|$.

We often use vectors coming from a tensor vector space $\mathcal{V} = \mathcal{V}_1 \otimes \mathcal{V}_2$, as well as vertices coming from a product vertex set $V = V_1 \times V_2$. In such cases we use superscripts to indicate the universe a certain object resides in. For example, we denote vectors from $\mathcal{V}_1$ by $x^{(1)}, y^{(1)}$ etc. In particular, when $x \in \mathcal{V}$ is a product vector then $x^{(1)}$ denotes the $\mathcal{V}_1$ component, $x^{(2)}$ denotes the $\mathcal{V}_2$ component and $x = x^{(1)} \otimes x^{(2)}$.

$\mathbb{S}_\Lambda$ represents the permutation group over Λ. $G_{N,D}$, for an even D, is the following distribution over D–regular, undirected graphs: First, uniformly choose $D/2$ permutations $\gamma_1, \ldots, \gamma_{D/2} \in \mathbb{S}_{[N]}$. Then, output the graph $G = (V = [N], E)$, whose edges are the *undirected* edges formed by the $D/2$ permutations.

3. THE K-STEP ZIG-ZAG PRODUCT

3.1 The product

The input to the product is:

- A possibly directed graph $G_1 = (V_1 = [N_1], E_1)$ that is a (D_1, λ_1) graph. We assume G_1 has a local inversion function $\phi = \phi_{G_1}$. That is, $\mathrm{Rot}_{G_1}(v^{(1)}, d_1) = (v^{(1)}[d_1], \phi_{G_1}(d_1))$.

- k undirected graphs $\bar{H} = (H_1, \ldots, H_k)$, where each H_i is a (N_2, D_2, λ_2) graph over the vertex set V_2.

In the replacement product (and also in the zig-zag product) the parameters are set such that the degree D_1 of G_1 equals the cardinality of V_2. An element $v_2 \in V_2$ is then interpreted as a label $d_1 \in [D_1]$. However, as explained in the introduction, we take larger graphs H_i, with $V_2 = [D_1]^{4k}$. That is, we have D_1^{4k} vertices in V_2 rather than D_1 in the replacement product. Therefore, we need to explain how to map a vertex $v^{(2)} \in V_2 = [D_1]^{4k}$ to a label $d_1 \in [D_1]$ of G_1. For that we use a map $f : V_2 \to [D_1]$ that is *regular*, i.e., every element of $[D_1]$ has the same number of f pre-images in V_2. For simplicity we fix one concrete such f – the function π_1 that takes the first $[D_1]$ coordinate of V_2. Namely, $\pi_1(v^{(2)}) = \pi_1(v_1^{(2)}, \ldots, v_{4k}^{(2)}) = v_1^{(2)}$.

The graph $G_{\text{new}} = G \textcircled{z} \bar{H}$ that we construct is related to a k–step walk over this new replacement product. The vertices of G_{new} are $V_1 \times V_2$. The degree of the graph is D_2^k and the edges are indexed by $\bar{i} = (i_1, \ldots, i_k) \in [D_2]^k$. We next define the rotation map $\text{Rot}_{G_{\text{new}}}$ of the new graph. For $v = (v^{(1)}, v^{(2)}) \in V_1 \times V_2$ and $\bar{i} = (i_1, \ldots, i_k) \in [D_2]^k$, $\text{Rot}_{G_{\text{new}}}(v, \bar{i})$ is defined as follows.

We start the walk at $(v_0^{(1)}, v_0^{(2)}) = v = (v^{(1)}, v^{(2)})$. For $j = 1, \ldots, 2k-1$, if j is odd, we set $t = \frac{j+1}{2}$ (and so $t = 1, \ldots, k$) and take one $H_t(\cdot, i_t)$ step on the second component. I.e., the first component is left untouched, $v_j^{(1)} = v_{j-1}^{(1)}$ and we set $(v_j^{(2)}, i'_t) = \text{Rot}_{H_t}(v_{j-1}^{(2)}, i_t)$. For even j, we take one step on G_1 with $\pi_1(v_{j-1}^{(2)})$ as the $[D_1]$ label to be used, i.e., $v_j^{(1)} = v_{j-1}^{(1)}[\pi_1(v_{j-1}^{(2)})]$. We set $v_j^{(2)} = \psi(v_{j-1}^{(2)})$, where

$$\psi(v^{(2)}) \;=\; (\phi_{G_1}(\pi_1(v^{(2)})), v_2^{(2)}, v_3^{(2)}, \ldots, v_{4k}^{(2)}). \qquad (1)$$

Namely, for the first $[D_1]$ coordinate of the second component we use the local inversion function of G_1, and all other coordinates are left unchanged. Finally, we specify $\text{Rot}_{G_{\text{new}}}(v, \bar{i}) = \;\; (v_{2k-1}^{(1)}, v_{2k-1}^{(2)}), (i'_k, \ldots, i'_1)$. It is straightforward to verify that $\text{Rot}_{G_{\text{new}}}$ is indeed a rotation map.

To summarize, we start with a D_1–regular graph over N_1 vertices (we think of D_1 as a constant and of $N_1 = |V_1|$ as a growing parameter) that is locally invertible. We replace each degree D_1 vertex with a "cloud" of D_1^{4k} vertices, and map a cloud vertex to a D_1 instruction using π_1. We then take a $(2k-1)$-step walk, with alternating H and G_1 steps, over the resulting graph.

3.2 A condition guaranteeing good algebraic expansion

We remind the reader of the discussion in Subsection 1.1.2 about "good" graphs H. We start with some $x \in \mathcal{V}$ that is uniform over clouds. We say the graphs $\bar{H} = (H_1, \ldots, H_k)$ are good if, for any $j > i$, applying $\tilde{H}_j \dot{G}_1 \tilde{H}_{j-1} \dot{G}_1 \ldots \tilde{H}_i \dot{G}_1$ on x always results in a vector that is uniform over the first $\log(D_1)$ bits of the cloud.

Each graph H_i is D_2–regular, and hence can be expressed as $H_i = \frac{1}{D_2} \sum_{j=1}^{D_2} \mathcal{H}_{i,j}$ where $\mathcal{H}_{i,j}$ is the transition matrix of a permutation $\gamma_{i,j} \in \mathbb{S}_{V_2}$. Instead of showing that $\bar{H}$ is good, we show that each sequence of permutations $\gamma_{1,j_1}, \ldots, \gamma_{k,j_k}$ is good in some sense that we define soon. Working with permutations is easier than working with $\bar{H}$ because a sequence of permutations induces a deterministic behavior while any $\tilde{H}_i$ is stochastic.

Assume we have a local inversion function on G_1 that is extended to a permutation $\psi : V_2 \to V_2$ as in Equation (1). We first determine the labels that are induced by replacing the H_i steps with the permutations $\gamma_1, \ldots, \gamma_k$:

DEFINITION 2. *Let $\psi, \gamma_1, \ldots, \gamma_{k-1} : V_2 \to V_2$ be permutations. Denote $\bar{\gamma} = (\gamma_1, \ldots, \gamma_{k-1})$. The permutation sequence $\bar{q} = (q_0, \ldots, q_{k-1})$* induced by *$(\bar{\gamma}, \psi)$ is defined as follows:*

- *$q_0(v^{(2)}) = v^{(2)}$,*
- *For $1 \le i < k$, $q_i(v^{(2)}) = \gamma_i(\psi(q_{i-1}(v^{(2)})))$.*

It can be checked that $q_j(v)$ is the V_2 value one reaches after taking a j-step walk starting at $v^{(2)}$ (and an arbitrary $v^{(1)}$) and taking each time a G_1 step followed by a γ_i permutation (for $i = 1, \ldots, j$).

We say $\bar{\gamma}$ is *ϵ–pseudorandom* with respect to ψ if the distribution of the first $\log(D_1)$ bits in each of the k labels we encounter is uniform. We define:

DEFINITION 3. *Let $q_0, \ldots, q_{k-1} : V_2 \to V_2$ be the permutations induced by $(\bar{\gamma} = (\gamma_1, \ldots, \gamma_{k-1}), \psi)$. We say $\bar{\gamma}$ is* ε–pseudorandom *with respect to ψ if*

$$\pi_1(q_0(U)) \circ \ldots \circ \pi_1(q_{k-1}(U)) - U_{[D_1]^k} \;_1 \le \varepsilon,$$

where $\pi_1(q_0(U)) \circ \ldots \circ \pi_1(q_{k-1}(U))$ is the distribution obtained by picking $v^{(2)} \in V_2$ uniformly at random and outputting $(\pi_1(q_0(v^{(2)})), \ldots, \pi_1(q_{k-1}(v^{(2)})))$ and $U_{[D_1]^k}$ is the uniform distribution over $[D_1]^k$.

We say $\bar{\gamma}$ is ε–pseudorandom with respect to G_1, if G_1 has a local inversion function ϕ_{G_1}, ψ is defined as in Equation (1) and $\bar{\gamma}$ is ε–pseudorandom with respect to ψ.

In the next section (in Lemma 5) we shall show that for every D–regular locally invertible graph, almost every $\bar{\gamma}$ is ε–pseudorandom with respect to it.

We are now ready to define when $\bar{H}$ is good:

DEFINITION 4. *Let $\bar{H} = (H_1, \ldots, H_k)$ be a k-tuple of D_2–regular graphs over V_2. We say $\bar{H}$ is* ε–pseudorandom *with respect to ψ, if we can express each graph H_i as $H_i = \frac{1}{D_2} \sum_{j=1}^{D_2} \mathcal{H}_{i,j}$ such that:*

- *$\mathcal{H}_{i,j}$ is the transition matrix of a permutation $\gamma_{i,j} \in \mathbb{S}_{V_2}$.*
- *For any $1 \le \ell_1 \le \ell_2 \le k$, $j_{\ell_1}, \ldots, j_{\ell_2} \in [D_2]$, the sequence $\gamma_{\ell_1, j_{\ell_1}}, \ldots, \gamma_{\ell_2, j_{\ell_2}}$ is ε–pseudorandom with respect to ψ.*

We say $\bar{H}$ is ε–pseudorandom with respect to G_1, if G_1 has a local inversion function ϕ_{G_1}, ψ is defined as in Equation (1) and $\bar{H}$ is ε–pseudorandom with respect to ψ. If, in addition, for each $i = 1, \ldots, k$ we have $\bar{\lambda}(H_i) \le \lambda_{\text{Ram}}(D_2) + \varepsilon$, we say that $\bar{H}$ is ε–good *with respect to G_1 (or ψ).*

In Section 4 we prove that for every locally invertible graph G_1, almost all $\bar{H}$ are good with respect to G_1. In fact, it turns out that there exists a sequence $\bar{H}$ that is good for all D_1-regular, locally invertible graphs.[3]

[3]The original claim we had only showed that for every G_1 there is a good sequence $\bar{H}$. We thank the anonymous referee for noticing that the bound in Lemma 5 actually proves this stronger claim.

In the following section (in Theorem 7) we shall prove that almost any $\bar{H}$ is ε–good with respect to any D_1–regular locally invertible graph.

Our main result states that, whenever $\bar{H}$ is good with respect to G_1, the k-step zigzag product does not lose much in the spectral gap. Formally,

THEOREM 2. *Let $G_1 = (V_1 = [N_1], E_1)$ be a (D_1, λ_1) locally invertible graph with a local inversion function ϕ_{G_1}. Let $\bar{H} = (H_1, \ldots, H_k)$ be a sequence of $(N_2 = D_1^{4k}, D_2, \lambda_2)$ graphs that is ε–good with respect to G_1, and assume $\lambda_2 \leq \frac{1}{2}$. Then, $G_{\text{new}} = G \textcircled{z} \bar{H}$ is a $(N_1 \cdot N_2, D_2^k, f(\lambda_1, \lambda_2, \varepsilon, k))$ graph for $f(\lambda_1, \lambda_2, \varepsilon, k) = \lambda_2^{k-1} + 2(\varepsilon + \lambda_1) + \lambda_2^k$.*

A word about the parameters is in place. Say our goal is to construct a $D = D_2^k$–regular graph that is as good algebraic expander as possible. By increasing D_1 we can decrease λ_1. In fact, we can make λ_1 any small constant we choose, while still keeping D_1 and $N_2 = D_1^{4k}$ constants. The crucial point is that we can still pick a good sequence $\bar{H}$ on this larger number of vertices, with degree D_2 (as before) and $\bar{\lambda} = \lambda_2$ (as before). Namely, we can decrease λ_1 to any constant we wish, while keeping D_2 and λ_2 as before, and the only (negligible) cost is making N_2 a somewhat larger constant. In particular, the final degree $D = D_2^k$ of the graph G_{new} stays unchanged. The same argument can be applied to decrease ε, and, in fact, ε in Theorem 7 is already much smaller than λ_2^k. We therefore consider λ_1 and ε as *negligible* terms. In this view the graph we construct has $\bar{\lambda} = \lambda_2^{k-1} + \lambda_2^k$ plus some negligible terms. In other words, we do k zig-zag steps and almost all of them ($k-1$ out of k) "work" for us.

4. ALMOST ANY $\bar{H}$ IS GOOD

4.1 A Hyper-Geometric lemma

We shall need the following tail estimate:

THEOREM 3. *([10], Theorem 2.10) Let Ω be a universe and $S_1 \subseteq \Omega$ a fixed subset of size m_1. Let $S_2 \subseteq \Omega$ be a uniformly random subset of size m_2. Set $\mu = \mathbb{E}_{S_2}[|S_1 \cap S_2|] = \frac{m_1 m_2}{|\Omega|}$. Then for every $\varepsilon > 0$,*

$$\Pr_{S_2}[\,|\,|S_1 \cap S_2| - \mu| \geq \varepsilon\mu] \;\leq\; 2e^{-\varepsilon^2/3\mu}.$$

A simple generalization of this gives:

LEMMA 4. *Let Ω be a universe and $S_1 \subseteq \Omega$ a fixed subset of size m. Let $S_2, \ldots, S_k \subseteq \Omega$ be uniformly random subsets of size m. Set $\mu_k = \mathbb{E}_{S_2,\ldots,S_k}[\,|\,S_1 \cap S_2 \ldots \cap S_k\,|\,] = \frac{m^k}{|\Omega|^{k-1}}$. Then for every $0 < \varepsilon \leq \frac{1}{4k}$,*

$$\Pr_{S_2,\ldots,S_k}[|\,|S_1 \cap S_2 \ldots \cap S_k| - \mu_k| \geq 2\varepsilon k \mu_k] \;\leq\; 2ke^{-\frac{\varepsilon^2}{6}\mu_k}.$$

Proof: By induction on k. $k = 2$ is Theorem 3. Assume for k, and let us prove for $k+1$. Let $A = S_1 \cap \ldots S_k \subseteq \Omega$. By the induction hypothesis we know that, except for probability $\delta_k = 2ke^{-\frac{\varepsilon^2}{6}\mu_k}$, the set A has size in the range $[(1-2(k-1)\varepsilon)\mu_k, (1+2(k-1)\varepsilon)\mu_k]$ for $\mu_k = \frac{m^k}{|\Omega|^{k-1}}$. When this happens, by Theorem 3, $|A \cap S_{k+1}|$ is in the range $[(1-\varepsilon)\frac{|A|m}{|\Omega|}, (1+\varepsilon)\frac{|A|m}{|\Omega|}] \subseteq [(1-2k\varepsilon)\mu_k, (1+2k\varepsilon)\mu_k]$ except for probability $2e^{-\frac{\varepsilon^2}{3}\frac{|A|m}{|\Omega|}} \leq 2e^{-\frac{\varepsilon^2}{3}(1-2(k-1)\varepsilon)\mu_{k+1}} \leq 2e^{-\frac{\varepsilon^2}{6}\mu_{k+1}}$. Thus, $|A \cap S_{k+1}|$ is in the required range except for probability $\delta_k + 2e^{-\frac{\varepsilon^2}{6}\mu_{k+1}} \leq 2(k+1)e^{-\frac{\varepsilon^2}{6}\mu_{k+1}}$ and this completes the proof. ■

4.2 Almost any $\bar{\gamma}$ is pseudorandom

The main lemma we prove in this section is:

LEMMA 5. *For every $\varepsilon > 0$, the probability that a sequence of uniformly random and independent permutations $(\gamma_1, \ldots, \gamma_{k-1})$ is not ε–pseudorandom with respect to G_1 is at most $2ke^{-\Omega(\varepsilon\frac{D_1^{3k}}{k^2})}$.*

Proof: Let $q_0, \ldots, q_{k-1} : V_2 \to V_2$ be the permutations induced by $(\bar{\gamma} = (\gamma_1, \ldots, \gamma_{k-1}), \psi)$, where ψ is as defined in Equation (1). Let A denote the distribution $\pi_1(q_1(U)) \circ \ldots \circ \pi_1(q_k(U))$ and B the uniform distribution over $[D_1]^k$. Fix an arbitrary $\bar{r} = (r_1, \ldots, r_k) \in [D_1]^k$. For $1 \leq i \leq k$, denote $S_i = \{x \in V_2 \mid \pi_1(q_i(x)) = r_i\}$. Since q_i is a permutation and π_1 is a regular function, $|S_i| = \frac{|V_2|}{D_1}$. We observe that for each i, q_i is a random permutation distributed uniformly in $\mathbb{S}_{V_2}$. Moreover, these permutations are independent. It follows that the sets $S_2, \ldots, S_k$ are random $\frac{|V_2|}{D_1}$–subsets of V_2, and they are independent as well.

By definition $A(\bar{r}) = \frac{|S_1 \cap S_2 \ldots \cap S_k|}{|V_2|}$. Notice that

$$\mathbb{E}[|S_1 \cap S_2 \ldots \cap S_k|] = \mu = \frac{(|V_2|/D_1)^k}{|V_2|^{k-1}} = \frac{|V_2|}{D_1^k} = D_1^{3k}.$$

By Lemma 4 the probability we deviate from this by a multiplicative factor of $1+\varepsilon$ is at most $2ke^{-\Omega(\frac{\varepsilon}{k^2}\mu)} = 2ke^{-\Omega(\varepsilon\frac{D_1^{3k}}{k^2})}$. It follows that:

$$\Pr_{\gamma_1,\ldots,\gamma_k}[|A(\bar{r}) - B(\bar{r})| \geq \varepsilon D_1^{-k}] \leq 2ke^{-\Omega(\varepsilon\frac{D_1^{3k}}{k^2})}.$$

Therefore, using a simple union bound, the event $\exists \bar{r}\, |A(\bar{r}) - B(\bar{r})| \geq \varepsilon D_1^{-k}$ happens with probability that is at most $D_1^k \cdot 2ke^{-\Omega(\varepsilon\frac{D_1^{3k}}{k^2})}$. However, $|A - B|_1 = \sum_{\bar{r}} |A(\bar{r}) - B(\bar{r})| \leq D_1^k \cdot \max_{\bar{r}}\{|A(\bar{r}) - B(\bar{r})|\}$ and therefore except for the above failure probability we have $|A - B|_1 \leq \varepsilon$ as desired. ■

4.3 The spectrum of random D-regular graphs

Friedman [7] proved the following theorem regarding the spectrum of random regular graphs. The distribution $G_{N,D}$ is described in Section 2.

THEOREM 6. *([7]) For every $\delta > 0$ and for every even D, there exists a constant $c > 0$, independent of N, such that*

$$\Pr_{G \sim G_{N,D}}[\bar{\lambda}(G) > \lambda_{\text{Ram}}(D) + \delta] \;\leq\; c \cdot N^{-\lceil(\sqrt{D-1}+1)/2\rceil - 1}.$$

4.4 Almost any $\bar{H}$ is good

THEOREM 7. *For every even $D_2 \geq 4$, there exists a constant B, such that for every $D_1 \geq B$ and every $k \geq 3$ the following holds. Set $N_2 = D_1^{4k}$ and $\varepsilon = D_2^{-k}$. Pick $\bar{H} = (H_1, \ldots, H_k)$ with each H_i sampled independently and uniformly from G_{N_2,D_2}. Then,*

- *Each H_i is locally invertible.*

- *With probability at least half, $\bar{H}$ is ε–good with respect to any D_1–regular locally invertible graph.*

Proof: We first show that for any fixed D_1–regular locally invertible graph G_1, almost any $\bar{H}$ is good for it. We then use a union bound (over all possible local inversion functions for D_1–regular graphs) to deduce the theorem.

Let us fix a D_1–regular locally invertible graph G_1. We randomly pick $\bar{H} = (H_1, \ldots, H_k)$ as in the lemma. I.e., let $\{\gamma_{i,j}\}_{i\in[k],\ j\in[D_2/2]}$ be a set of random permutations chosen uniformly and independently from $\mathbb{S}_{V_2}$. For $1 \le i \le k$, let H_i be the undirected graph over V_2 formed from the permutations $\{\gamma_{i,j}\}_{j\in[D_2/2]}$ and their inverses. Notice that H_i is locally invertible, simply by labeling the directed edge $(v, \gamma_{i,j}(v))$ with the label j, and $(v, \gamma_{i,j}^{-1}(v))$ with the label $D_2/2 + j$ (recall that each edge needs to be labeled twice, once by each of its vertices).

Notice that the inverse of a uniform random permutation is also a uniform random permutation. Therefore, for every $j_1, \ldots, j_k \in [D_2/2]$ and for every $p_1, \ldots, p_k \in \{1, -1\}$, the k-tuple $\bar{\gamma} = (\gamma_{1,j_1}^{p_1}, \ldots, \gamma_{k,j_k}^{p_k})$ is uniform in $(\mathbb{S}_{|V_2|})^k$. It follows from Lemma 5 that $\bar{H}$ is not ε–pseudorandom with respect to G_1 with probability at most $k^2 \cdot D_2^k \cdot D_1^k \cdot 2ke^{-\Omega(\varepsilon \frac{D_1^{3k}}{k^2})}$.[4] Taking $\varepsilon = D_2^{-k} \ge D_1^{-k}$ we see that the error term is at most $\delta \stackrel{def}{=} D_1^{3k} e^{-\Omega(\frac{D_1^k}{k^2})}$.

To see that a single sequence $\bar{H}$ is, with high probability, good for any D_1-regular locally invertible graph, we use a union bound. Notice that there are only $D_1!$ local inversion functions over D_1 vertices (compare this with the $N_2!$ permutations over V_2). The probability a random $\bar{H}$ is bad for any of them is at most δ, and therefore the probability over $\bar{H}$ that it is bad for any of them is at most $D_1! \cdot \delta$. Taking D_1 large enough this term is at most $\frac{1}{10}$.

Also, by Theorem 6, the probability that there exists a graph H_i in $\bar{H}$ with $\bar{\lambda}(H_i) \ge \lambda_{\mathrm{Ram}}(D_2) + \varepsilon$ is at most $k \cdot c \cdot |V_2|^{-\lceil(\sqrt{D_2-1}+1)/2\rceil - 1} \le k \cdot c \cdot |V_2|^{-1} = \frac{kc}{{D_1}^{4k}}$ for some universal constant c independent of $|V_2|$ and therefore also independent of D_1. Taking D_1 large enough (depending on the unspecified constant c) this term also becomes smaller than $\frac{1}{10}$.

Altogether, $\bar{H}$ is always locally invertible, and with probability at least $\frac{1}{2}$ is ε–good with respect to any D_1–regular locally invertible graph. ■

5. ANALYSIS OF THE PRODUCT

We want to express the k-step walk described in Section 3.1 as a composition of linear operators. We define vector spaces $\mathcal{V}_i$ with $\dim(\mathcal{V}_i) = |V_i| = N_i$, and we identify an element $v^{(i)} \in V_i$ with a basis vector $\overrightarrow{v^{(i)}}$. Notice that $\{\overrightarrow{v^{(1)}} \otimes \overrightarrow{v^{(2)}} \mid v^{(1)} \in V_1, v^{(2)} \in V_2\}$ is a basis for $\mathcal{V}$. On this basis we define the linear operators $\tilde{H}_i(\overrightarrow{v^{(1)}} \otimes \overrightarrow{v^{(2)}}) = \overrightarrow{v^{(1)}} \otimes \overrightarrow{H_i v^{(2)}}$ and $\dot{G}_1(\overrightarrow{v^{(1)}} \otimes \overrightarrow{v^{(2)}}) = \overrightarrow{v^{(1)}[\pi_1(v^{(2)})]} \otimes \overrightarrow{\psi(v^{(2)})}$, where ψ is as defined in Equation 1. Having this terminology, the adjacency matrix of the new graph G_{new} is the linear transformation on $\mathcal{V}$ defined by $\tilde{H}_k \dot{G}_1 \tilde{H}_{k-1} \dot{G}_1 \ldots \tilde{H}_2 \dot{G}_1 \tilde{H}_1$.

Proof of Theorem 2: G_{new} is a regular, directed graph and our goal is to bound $s_2(G_{\mathrm{new}})$. Fix unit vectors $x, y \perp \mathbf{1}$ for which $s_2(G_{\mathrm{new}}) = |\langle G_{\mathrm{new}} x, y\rangle|$. As in the analysis of the zig-zag product, we decompose $\mathcal{V} = \mathcal{V}_1 \otimes \mathcal{V}_2$ to its parallel and perpendicular parts. $\mathcal{V}^{||}$ is defined by $\mathcal{V}^{||} = \mathrm{Span}\{\overrightarrow{v^{(1)}} \otimes \mathbf{1} : v^{(1)} \in V_1\}$ and $\mathcal{V}^{\perp}$ is its orthogonal complement. For any vector $\tau \in \mathcal{V}$ we denote by $\tau^{||}$ and $\tau^{\perp}$ the projections of τ on $\mathcal{V}^{||}$ and $\mathcal{V}^{\perp}$ respectively.

In G_{new} we take $k-1$ steps on $\dot{G}_1$. As a result, in the analysis we need to decompose not only $x_0 = x$ and $y_0 = y$, but also the vectors $x_1, \ldots, x_{k-1}$ and $y_1, \ldots, y_{k-1}$ where $x_i = \dot{G}_1 \tilde{H}_i x_{i-1}^{\perp}$ and $y_i = \dot{G}_1 \tilde{H}_{k-i+1}\ y_{i-1}^{\perp}$. Observe that $\|x_i\| \le \lambda_2^i \|x_0\|$ and $\|y_i\| \le \lambda_2^i \|y_0\|$.

Now look at $y_0^\dagger \tilde{H}_k \dot{G}_1 \ldots \tilde{H}_2 \dot{G}_1 \tilde{H}_1 x_0$ and decompose x_0 to $x_0^{||}$ and $x_0^{\perp}$. Focusing on $x_0^{\perp}$ we see that, by definition, $y_0^\dagger \tilde{H}_k \dot{G}_1 \ldots \tilde{H}_2 \dot{G}_1 \tilde{H}_1 x_0^{\perp} = y_0^\dagger \tilde{H}_k \dot{G}_1 \ldots \tilde{H}_3 \dot{G}_1 \tilde{H}_2 x_1$. We continue by decomposing x_1. This results in: $y_0^\dagger \tilde{H}_k x_{k-1}^{\perp} + \sum_{i=1}^k y_0^\dagger \tilde{H}_k \dot{G}_1 \ldots \tilde{H}_{i+1} \dot{G}_1 \tilde{H}_i x_{i-1}^{||}$.

We can now do the same decomposition on y_0, using the fact that both $\dot{G}_1$ and $\tilde{H}_j$ are Hermitian and so $(y_j^{\perp})^\dagger \tilde{H}_{k-j} \dot{G}_1$ equals $(\dot{G}_1 \tilde{H}_{k-j} y_j^{\perp})^\dagger = y_{j+1}^\dagger$. This gives the expression $y_0^\dagger \tilde{H}_k x_{k-1}^{\perp} + \sum_{i=1}^k (y_{k-i}^{||})^\dagger x_{i-1}^{||} + \sum_{i=1}^k (y_{k-i}^{\perp})^\dagger x_{i-1}^{||} + \sum_{1\le i<j\le k} (y_{k-j}^{||})^\dagger \dot{G}_1 \tilde{H}_{j-1} \ldots \tilde{H}_{i+1} \dot{G}_1 x_{i-1}^{||}$.

Now,

- $y_0^\dagger \tilde{H}_k x_{k-1}^{\perp} \le \tilde{H}_k x_{k-1}^{\perp} \le \lambda_2\ x_{k-1}^{\perp} \le \lambda_2 \|x_k\| \le \lambda_2 \lambda_2^{k-1} \|x_0\| = \lambda_2^k$.
- Since $\mathcal{V}^{\perp} \perp \mathcal{V}^{||}$, we get $\sum_{i=1}^k (y_{k-i}^{\perp})^\dagger x_{i-1}^{||} = 0$.
- The term $\sum_{i=1}^k (y_{k-i}^{||})^\dagger x_{i-1}^{||} \le \sum_{i=1}^k\ y_{k-i}^{||} \cdot x_{i-1}^{||}$ is bounded in Lemma 13 by λ_2^{k-1}.
- Finally, we take advantage of the way we selected $\bar{H}$. As $\bar{H}$ is ε–pseudorandom with respect to G_1, the action of $\dot{G}_1 \tilde{H}_{j-1} \ldots \tilde{H}_{i+1} \dot{G}_1$ on $\mathcal{V}^{||}$ is ε–close to the action of G^{j-i} on it. Formally, we use Lemma 10 to get:

$$\begin{aligned}
&\sum_{1\le i<j\le k} (y_{k-j}^{||})^\dagger \dot{G}_1 \tilde{H}_{j-1} \ldots \tilde{H}_{i+1} \dot{G}_1 x_{i-1}^{||} \\
&\le \sum_{1\le i<j\le k} (\lambda_1^{j-i} + \varepsilon)\ y_{k-j}^{||} \quad x_{i-1}^{||} \\
&= \sum_{t=1}^{k-1} (\lambda_1^t + \varepsilon) \sum_{i=1}^{k-t} y_{k-i-t}^{||} \quad x_{i-1}^{||} \\
&\le (\lambda_1 + \varepsilon) \sum_{t=1}^{k-1} \lambda_2^{k-t-1} \\
&= (\lambda_1 + \varepsilon) \sum_{i=0}^{k-2} \lambda_2^i \le 2(\lambda_1 + \varepsilon),
\end{aligned}$$

where we have used Lemma 13 and the assumption $\lambda_2 \le \frac{1}{2}$.

Altogether, $|y^\dagger G_{\mathrm{new}} x| \le \lambda_2^{k-1} + 2(\varepsilon + \lambda_1) + \lambda_2^k$ as desired. ■

[4] The k^2 term appears because ε–pseudorandomness of $\bar{H}$ requires every subsequence of permutations to have this property; taking a union bound over the choice of the starting and ending indices $1 \le \ell_1 \le \ell_2 \le k$ of the subsequence amounts to k^2

5.1 The action of $\dot{G}_1\tilde{H}_{i+\ell}\dot{G}_1\ldots\tilde{H}_{i+1}\dot{G}_1$ on $\mathcal{V}^{||}$

The heart of this section is the following lemma.

LEMMA 8. *Suppose $\bar{\gamma}=(\gamma_1,\ldots,\gamma_\ell)$ is ε–pseudorandom with respect to G_1 and denote by $\tilde{\Gamma}_1,\ldots,\tilde{\Gamma}_\ell$ the operators corresponding to $\gamma_1,\ldots,\gamma_\ell$. Any $\tau,\xi\in\mathcal{V}^{||}$ can be written as $\tau=\tau^{(1)}\otimes\mathbf{1}$ and $\xi=\xi^{(1)}\otimes\mathbf{1}$. For any such τ,ξ:*

$$\left\langle\dot{G}_1\tilde{\Gamma}_\ell\dot{G}_1\ldots\tilde{\Gamma}_1\dot{G}_1\tau,\xi\right\rangle-\left\langle G^{\ell+1}\tau^{(1)},\xi^{(1)}\right\rangle\ \leq\ \varepsilon\cdot\|\tau\|\cdot\|\xi\|.$$

Proof: G_1 is D_1–regular, hence it can be represented as $G_1=\frac{1}{D_1}\sum_{i=1}^{D_1}\mathcal{G}_i$, where $\mathcal{G}_i$ is the adjacency matrix of some permutation in $\mathbb{S}_{V_1}$. Let ψ be as in Equation (1) and $\bar{q}=(q_0,\ldots,q_{k-1})$ be the permutations induced by $(\bar{\gamma},\psi)$. A simple calculation (that is given in Lemma 11 in Subsection 5.2) shows that there exists some $\sigma\in\mathbb{S}_{V_2}$, such that for any $u^{(1)}\in V_1$ and $u^{(2)}\in V_2$:

$$\begin{aligned}&\dot{G}_1\tilde{\Gamma}_\ell\dot{G}_1\ldots\tilde{\Gamma}_1\dot{G}_1(\overrightarrow{u^{(1)}}\otimes\overrightarrow{u^{(2)}})\\&\quad=\ \mathcal{G}_{\pi_1(q_\ell(u^{(2)}))}\cdots\mathcal{G}_{\pi_1(q_0(u^{(2)}))}(\overrightarrow{u^{(1)}})\otimes\overrightarrow{\sigma(u^{(2)})}.\end{aligned}\tag{2}$$

Now, we analyze the action of $\dot{G}_1\tilde{\Gamma}_\ell\dot{G}_1\ldots\tilde{\Gamma}_1\dot{G}_1$ on vectors $\tau=\tau^{(1)}\otimes\mathbf{1}$ and $\xi=\xi^{(1)}\otimes\mathbf{1}$ in $\mathcal{V}^{||}$. Using Equation (2) we can show that (see Lemma 12 in Subsection 5.2):

$$\begin{aligned}&\left\langle\dot{G}_1\tilde{\Gamma}_\ell\dot{G}_1\ldots\tilde{\Gamma}_1\dot{G}_1\tau,\xi\right\rangle\\&\quad=\ \frac{1}{N_2}\sum_{v^{(2)}\in V_2}\left\langle\mathcal{G}_{\pi_1(q_\ell(v^{(2)}))}\cdots\mathcal{G}_{\pi_1(q_0(v^{(2)}))}\tau^{(1)},\xi^{(1)}\right\rangle.\end{aligned}$$

Restating the above, $\left\langle\dot{G}_1\tilde{\Gamma}_{j+\ell}\dot{G}_1\ldots\tilde{\Gamma}_{j+1}\dot{G}_1\tau,\xi\right\rangle$ equals $\mathbb{E}_{z_1,\ldots,z_\ell\sim\mathcal{Z}}\left[\left\langle\mathcal{G}_{z_\ell}\ldots\mathcal{G}_{z_1}\tau^{(1)},\xi^{(1)}\right\rangle\right]$, where $\mathcal{Z}$ is the distribution on $[D_1]^k$ obtained by picking $v^{(2)}$ uniformly at random in V_2 and outputting $z_1,\ldots,z_\ell$ where $z_i=\pi_1(q_i(v^{(2)}))$. Notice also that $G_1^k=\mathbb{E}_{z\in[D_1]^k}[\mathcal{G}_{z_\ell}\ldots\mathcal{G}_{z_1}]$. As $(\gamma_1,\ldots,\gamma_k)$ is ε–pseudorandom with respect to G_1 we can deduce that $\left\|\mathcal{Z}-U_{[D_1]^k}\right\|_1\leq\varepsilon$. We now use:

CLAIM 9. *Let P,Q be two distributions over Ω and let $\{\mathcal{L}_i\}_{i\in\Omega}$ be a set of linear operators over Λ, each with operator norm bounded by 1. Define $\mathcal{P}=\mathbb{E}_{x\sim P}[\mathcal{L}_x]$ and $\mathcal{Q}=\mathbb{E}_{x\sim Q}[\mathcal{L}_x]$. Then, for any $\tau,\xi\in\Lambda$, $|\langle\mathcal{P}\tau,\xi\rangle-\langle\mathcal{Q}\tau,\xi\rangle|\leq|P-Q|_1\cdot\|\tau\|\cdot\|\xi\|$.*

Proof: First, notice that $\|\mathcal{P}-\mathcal{Q}\|_\infty\leq\sum_x|P(x)-Q(x)|\cdot\|\mathcal{L}_x\|_\infty\leq|P-Q|_1$. Therefore, it follows that $|\langle\mathcal{P}\tau,\xi\rangle-\langle\mathcal{Q}\tau,\xi\rangle|=|\langle(\mathcal{P}-\mathcal{Q})\tau,\xi\rangle|\leq\|\mathcal{P}-\mathcal{Q}\|_\infty\cdot\|\tau\|\cdot\|\xi\|\leq|P-Q|_1\cdot\|\tau\|\cdot\|\xi\|$. ■

Thus, $\left\langle\dot{G}_1\tilde{\Gamma}_\ell\dot{G}_1\ldots\tilde{\Gamma}_1\dot{G}_1\tau,\xi\right\rangle-\left\langle G^{\ell+1}\tau^{(1)},\xi^{(1)}\right\rangle$ is upper bounded by $\varepsilon\cdot\ \tau^{(1)}\ \cdot\ \xi^{(1)}\ =\varepsilon\cdot\|\tau\|\cdot\|\xi\|$ (because $\|\tau\|=\ \tau^{(1)}\otimes\mathbf{1}\ =\ \tau^{(1)}\ \cdot\|\mathbf{1}\|=\ \tau^{(1)}\ $) and this completes the proof of Lemma 8. ■

Having Lemma 8 we can prove:

LEMMA 10. $\left\langle\dot{G}_1\tilde{H}_{i+\ell}\dot{G}_1\ldots\tilde{H}_{i+1}\dot{G}_1\tau,\xi\right\rangle\ \leq\ (\lambda_1^{\ell+1}+\varepsilon)\|\tau\|\|\xi\|$ *for every $\ell\geq1$ and $\tau,\xi\in\mathcal{V}^{||}$, $\tau,\xi\perp\mathbf{1}_V$.*

Proof: Since $\bar{H}$ is ε–good with respect to G_1, we can express each H_i as $H_i=\frac{1}{D_2}\sum_{j=1}^{D_2}\mathcal{H}_{i,j}$ such that $\mathcal{H}_{i,j}$ is the transition matrix of a permutation $\gamma_{i,j}\in\mathbb{S}_{V_2}$ and each of the D_2^k sequences $\gamma_{1,j_1},\ldots,\gamma_{k,j_k}$ is ε–pseudorandom with respect to G_1. Let $\Gamma_{i,j}$ be the operator on $\mathcal{V}_2$ corresponding to the permutation $\gamma_{i,j}$ and $\tilde{\Gamma}_{i,j}=I\otimes\Gamma_{i,j}$ be the corresponding operator on $\mathcal{V}_1\otimes\mathcal{V}_2$.

Now, $\left\langle\dot{G}_1\tilde{H}_{i+\ell}\dot{G}_1\ldots\tilde{H}_{i+1}\dot{G}_1\tau,\xi\right\rangle$ is equal to the expectation $\mathbb{E}_{j_1,\ldots,j_\ell\in[D_2]}\left[\left\langle\dot{G}_1\tilde{\Gamma}_{i+\ell,j_\ell}\dot{G}_1\ldots\tilde{\Gamma}_{i+1,j_1}\dot{G}_1\tau,\xi\right\rangle\right]$. Notice that not only $\bar{H}$ is ε–pseudorandom with respect to G_1, but also every subsequence of $\bar{H}$ is. Thus, by Lemma 8,

$$\begin{aligned}&\left\langle\dot{G}_1\tilde{H}_{i+\ell}\dot{G}_1\ldots\tilde{H}_{i+1}\dot{G}_1\tau,\xi\right\rangle-\left\langle G^{\ell+1}\tau^{(1)},\xi^{(1)}\right\rangle\\&\leq\ \varepsilon\cdot\|\tau\|\cdot\|\xi\|.\end{aligned}$$

Since $\tau,\xi\perp\mathbf{1}$, so does their $\tau^{(1)},\xi^{(1)}$ components. Therefore, $\left\langle G^{\ell+1}\tau^{(1)},\xi^{(1)}\right\rangle\ \leq\ \lambda_1^{\ell+1}\ \tau^{(1)}\ \ \xi^{(1)}\ $. The fact that $\|\tau\|=\ \tau^{(1)}\ $ and $\|\xi\|=\ \xi^{(1)}\ $ completes the proof. ■

5.2 The action of the composition

LEMMA 11. *There exists $\sigma\in\mathbb{S}_{V_2}$, such that for any $u^{(1)}\in V_1$ and $u^{(2)}\in V_2$:*

$$\begin{aligned}&\dot{G}_1\tilde{\Gamma}_\ell\dot{G}_1\ldots\tilde{\Gamma}_1\dot{G}_1(\overrightarrow{u^{(1)}}\otimes\overrightarrow{u^{(2)}})\\&\quad=\ \mathcal{G}_{\pi_1(q_\ell(u^{(2)}))}\cdots\mathcal{G}_{\pi_1(q_0(u^{(2)}))}(\overrightarrow{u^{(1)}})\otimes\overrightarrow{\sigma(u^{(2)})}.\end{aligned}$$

Proof: The action of $\dot{G}_1\tilde{\Gamma}_\ell\dot{G}_1\ldots\tilde{\Gamma}_1\dot{G}_1$ on a basis element $\overrightarrow{u^{(1)}}\otimes\overrightarrow{u^{(2)}}$, where $u^{(1)}\in V_1$ and $u^{(2)}\in V_2$, is as follows.

- We first check which of the $[D_1]$ labels we use at the i'th application of $\dot{G}_1$ (for $i=0,\ldots,\ell$). We see that $q_0(u^{(2)})=u^{(2)}$ and that for $i=1,\ldots,\ell$ we have $q_i(u^{(2)})=\gamma_i(\phi(q_{i-1}(u^{(2)})))$.
- Hence, the action of $\dot{G}_1\tilde{\Gamma}_i\dot{G}_1\ldots\tilde{\Gamma}_1\dot{G}_1$ on the first component (for $i=1,\ldots,\ell$) is given by the linear operator $\mathcal{G}_{\pi_1(q_i(u^{(2)}))}\cdots\mathcal{G}_{\pi_1(q_0(u^{(2)}))}$.
- Next, we notice that the $\mathcal{V}_2$ component evolves independently of $u^{(1)}$. At the beginning it is $u^{(2)}$. After applying one step of $\dot{G}_1$ and one of $\tilde{\Gamma}_1$ it evolves to $\gamma_1(\phi(u^{(2)}))$. Eventually, this component becomes $\phi(\gamma_\ell(\phi(\ldots\gamma_1(\phi(u^{(2)}))\ldots)))$. The crucial thing to notice here is that $\{\gamma_i\}$ and ϕ are all permutations in $\mathbb{S}_{V_2}$. We define σ to be the permutation $\phi\gamma_\ell\phi\ldots\gamma_1\phi$.

This completes the proof of the lemma. ■

LEMMA 12. *For any $\tau=\tau^{(1)}\otimes\mathbf{1}$ and $\xi=\xi^{(1)}\otimes\mathbf{1}$ in $\mathcal{V}^{||}$,*

$$\begin{aligned}&\left\langle\dot{G}_1\tilde{\Gamma}_\ell\dot{G}_1\ldots\tilde{\Gamma}_1\dot{G}_1\tau,\xi\right\rangle\\&\quad=\ \frac{1}{N_2}\sum_{v^{(2)}\in V_2}\left\langle\mathcal{G}_{\pi_1(q_\ell(v^{(2)}))}\cdots\mathcal{G}_{\pi_1(q_0(v^{(2)}))}\tau^{(1)},\xi^{(1)}\right\rangle.\end{aligned}$$

Proof: A simple calculation yields that $\left\langle\dot{G}_1\tilde{\Gamma}_\ell\dot{G}_1\ldots\tilde{\Gamma}_1\dot{G}_1\tau,\xi\right\rangle$ equals $\frac{1}{N_2}\sum_{v^{(2)},u^{(2)}\in V_2}\left\langle\mathcal{G}_{\pi_1(q_\ell(v^{(2)}))}\cdots\mathcal{G}_{\pi_1(q_0(v^{(2)}))}\tau^{(1)},\xi^{(1)}\right\rangle\cdot\left\|\overrightarrow{\sigma(v^{(2)})},\overrightarrow{u^{(2)}}\right\|$. However, as σ is a permutation over V_2, for

every $v^{(2)} \in V_2$ there is exactly one $u^{(2)}$ that does not vanish. Hence,

$$\left\langle \dot{G}_1\tilde{\Gamma}_\ell\dot{G}_1\ldots\tilde{\Gamma}_1\dot{G}_1\tau,\xi\right\rangle = \frac{1}{N_2}\sum_{v^{(2)}\in V_2}\left\langle \mathcal{G}_{\pi_1(q_\ell(v^{(2)}))}\ldots\mathcal{G}_{\pi_1(q_0(v^{(2)}))}\tau^{(1)},\xi^{(1)}\right\rangle.$$

■

5.3 A lemma on partial sums

In the following lemma we have a sum of k terms. Each of magnitude at most λ_2^{k-t-1}. Surprisingly, we can bound the sum by λ_2^{k-t-1}, improving upon the trivial bound of $k\cdot\lambda_2^{k-t-1}$.

LEMMA 13. *Let $t \geq 0$. Then $\sum_{i=1}^{k-t} y^{||}_{k-i-t} \cdot x^{||}_{i-1} \leq \lambda_2^{k-t-1}$.*

Proof:

$$\sum_{i=1}^{k-t} y^{||}_{k-i-t}\cdot x^{||}_{i-1} = \lambda_2^{k-t-1}\sum_{i=1}^{k-t}\frac{y^{||}_{k-i-t}}{\lambda_2^{k-i-t}}\cdot\frac{x^{||}_{i-i}}{\lambda_2^{i-1}}$$
$$\leq \lambda_2^{k-t-1}\cdot\frac{1}{2}\left(\sum_{i=0}^{k-t-1}\left(\frac{y^{||}_i}{\lambda_2^i}\right)^2+\sum_{i=0}^{k-t-1}\left(\frac{x^{||}_i}{\lambda_2^i}\right)^2\right).$$

Now, we bound $\sum_{i=0}^{k-t-1}\left(\frac{x^{||}_i}{\lambda_2^i}\right)^2$ and the bound for the expression $\sum_{i=0}^{k-t-1}\left(\frac{y^{||}_i}{\lambda_2^i}\right)^2$ is similarly obtained. Denote

$$\Delta_\ell = \left(\frac{x^{\perp}_\ell}{\lambda_2^\ell}\right)^2+\sum_{i=0}^{\ell}\left(\frac{x^{||}_i}{\lambda_2^i}\right)^2.$$

Then

$$\Delta_\ell = \left(\frac{x_\ell}{\lambda_2^\ell}\right)^2+\sum_{i=0}^{\ell-1}\left(\frac{x^{||}_i}{\lambda_2^i}\right)^2 \leq \left(\frac{\lambda_2 x^{\perp}_{\ell-1}}{\lambda_2^\ell}\right)^2+\sum_{i=0}^{\ell-1}\left(\frac{x^{||}_i}{\lambda_2^i}\right)^2 = \Delta_{\ell-1}.$$

In particular, $\Delta_{k-t-1}\leq\Delta_0 = \left(x^{||}_0\right)^2$. It follows that

$$\sum_{i=0}^{k-t-1}\left(\frac{x^{||}_i}{\lambda_2^i}\right)^2 \leq \left(x^{||}_0\right)^2-\left(\frac{x^{\perp}_{k-t-1}}{\lambda_2^{k-t-1}}\right)^2 \leq \|x_0\|^2 = 1.$$

■

6. THE ITERATIVE CONSTRUCTION

In [21] an iterative construction of expanders was given, starting with constant-size expanders, and constructing at each step larger constant-degree expanders. Each iteration is a sequence of tensoring (which makes the graph much larger, the degree larger and the spectral gap the same), powering (which keeps the graph size the same, increases the spectral gap and the degree) and a zig-zag product (that reduces the degree back to what it should be without harming the spectral gap much). Here we follow the same strategy, using the same sequence of tensoring, powering and degree reduction, albeit we use k-step zigzag products rather than zig-zag products to reduce the degree. We do it for degrees D of the special form $D = 2D_2^k$. Giving an iterative construction for a general D is a bit more technical and will appear in the full version of this paper.

Let D_2 be an arbitrary even number greater than 2. We are given a degree D of the form $D = 2D_2^k$. Set $\varepsilon = D_2^{-k}$ and $\lambda_2 = \lambda_{\mathrm{Ram}}(D_2)+\varepsilon$. We find a sequence $\bar{H} = (H_1,\ldots,H_k)$ of $(D^{16k}, D_2, \lambda_2)$ graphs, that is ε-good with respect to D^4–regular locally invertible graphs. We find it by brute force; its existence is guaranteed by Theorem 7. Verify that a given $\bar{H}$ can be done in time depending only on D, D_2 and k, independent of N_1.

We start with two constant-size graphs G_1 and G_2. G_1 is a (N_0, D, λ) graph, and G_2 is a (N_0^2, D, λ) graph, for $N_0 = D^{16k}$ and $\lambda = 2\lambda_2^{k-1}$. We find both graphs by a brute force search (the existence of such graphs follows from Theorem 6 given in Subsection 4.3). Now, for $t > 2$:

- Define $G_{\mathrm{temp}} = (G_{\lfloor\frac{t-1}{2}\rfloor}\otimes G_{\lceil\frac{t-1}{2}\rceil})^2$. G_{temp} is over N_0^{t-1} vertices and has degree D^4.
- Let $G_t = \frac{1}{2}[G_{\mathrm{temp}}ⓩ\bar{H} + G_{\mathrm{temp}}ⓩ\bar{H}^{\dagger}]$.

We claim:

THEOREM 14. *The family of undirected graphs $\{G_t\}$ is fully-explicit and each graph G_t is a (N_0^t, D, λ) graph.*

The proof is immediate from the following two lemmas.

LEMMA 15. *For every $t \geq 1$, G_t is a (N_0^t, D, λ) undirected graph.*

Proof: It is easy to verify that G_t is over N_0^t vertices and has degree $D = 2D_2^k$. We turn to prove the bound on its spectral gap. Let α_t denote the second-largest eigenvalue of G_t and let $\beta_t = \max_{i\leq t}\{\alpha_i\}$. We shall prove by induction that $\beta_t \leq \lambda$. For $t = 1, 2$ this follows from the way G_1 and G_2 were chosen. For $t > 2$, using the properties of tensoring, powering and the k-step zig-zag product, we get the recursive relation $\beta_t = \max\left\{\beta_{t-1}, \lambda_2^{k-1}+\lambda_2^k+2(\beta_{t-1}^2+\varepsilon)\right\}$. Bounding β_{t-1} by $2\lambda_2^{k-1}$ and plugging $\varepsilon \leq \lambda_2^{2k}$ we get

$$\beta_t \leq \lambda_2^{k-1}(1+\lambda_2+10\cdot\lambda_2^{k-1}) \leq 2\lambda_2^{k-1} = \lambda,$$

where in the last inequality we used the fact that $\lambda_2 \leq \lambda_{\mathrm{Ram}}(D_2)+\varepsilon \leq 1/4$. ■

LEMMA 16. *$\{G_t\}$ is a fully explicit family of graphs, each having an explicit local inversion function.*

Proof: We prove the lemma by the induction. The cases $t = 1, 2$ are immediate. Assume we have a local inversion function $\phi_i : [D] \to [D]$ for all $\{G_i\}_{i\leq t}$, written as a constant-size table. This defines the local inversion function $\phi : [D^4] \to [D^4]$ for $G_{\mathrm{temp}} = (G_r \otimes G_\ell)^2$, simply by taking $\phi((r_1,\ell_1),(r_2,\ell_2)) = ((\phi_r(r_2),\phi_\ell(\ell_2)),(\phi_r(r_1),\phi_\ell(\ell_1)))$.

We next explain how to write down the inversion function $\phi_{t+1} : [2D_2^k] \to [2D_2^k]$ for G_{t+1}. G_{t+1} has $2D_2^k$ directed edges, and we label the edges coming from $G_{\mathrm{temp}}ⓩ\bar{H}$ with the labels $(0, i_1, \ldots, i_k)$ and the edges coming from $G_{\mathrm{temp}}ⓩ\bar{H}^{\dagger}$ with $(1, i_k, \ldots, i_1)$, where i_j describes the step on H_j. We then set the function to be $\phi_{t+1}(b, i_1, \ldots, i_k) = (1-b, \phi_{H_k}(i_k), \ldots, \phi_{H_1}(i_1))$.

We need to show how to compute $\mathrm{Rot}_{G_{t+1}}(v, w) = (v[w], w')$. We already saw how to compute $w' = \phi_{t+1}(w)$. We now show how to compute $v[w]$. Say $w = (1, i_1, \ldots, i_k) \in \{0,1\}\times[D_2]^k$ and $v = (v_1^{(1)}, v_2^{(1)}, v^{(2)})$ with $v_1^{(1)} \in [N_0^{t_1}], v_2^{(1)} \in$

$[N_0^{t_2}], v^{(2)} \in [N_0 = D^{16k}]$ and $t_1 + t_2 = t$. One can compute $v[w]$ by the following the walk starting at v, each time taking a step on H_j or on $(G_{t_1} \otimes G_{t_2})^2$. This takes time poly-logarithmic in the number of vertices of G_{t+1}. ■

The resulting eigenvalue is $\lambda = 2\lambda_2^{k-1}$ where λ_2 is about the Ramanujan value for D_2, whereas the best we can hope for $\bar{\lambda}_{\mathrm{Ram}}(D) = \frac{2\sqrt{D-1}}{D}$. As explained in the introduction, our losses come from two different sources. First we lose one application of H out of the k different H applications, and this loss amounts to, roughly, $\sqrt{D_2}$ multiplicative factor. We also have a second loss of 2^{k-1} multiplicative factor emanating from the fact that $\lambda_{\mathrm{Ram}}(D_2)^k =\approx 2^{k-1}\lambda_{\mathrm{Ram}}(D_2^k)$. Balancing losses we roughly have $D = D_2^k$ and $D_2 = 2^k$ which is solved by $k = \log(D_2)$ and $D = 2^{\log^2(D_2)}$. I.e., our loss is about $2^k = 2^{\sqrt{\log(D)}}$. Formally,

COROLLARY 17. *Let D_2 be an arbitrary even number that is greater than 2, and let $D = 2D_2^{\log D_2}$. Then, there exists a fully explicit family of $(D, D^{-\frac{1}{2}+O(\frac{1}{\sqrt{\log D}})})$ graphs.*

Proof: Set $k = \log D_2$ in the above construction. Clearly the resulting graphs are D–regular and fully explicit. Also, for every graph G in the family,

$$\bar{\lambda}(G) \;\le\; 2(\lambda_{\mathrm{Ram}}(D_2) + D_2^{-k})^{k-1} \;\le\; D^{-\frac{1}{2}+\frac{2}{\sqrt{\log D}}}.$$

■

Acknowledgements

We thank the anonymous referees for several useful suggestions that improved the presentation of the paper. We thank one of the referees for strengthening Theorem 7 (see footnote 3).

7. REFERENCES

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

[2] 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 FOCS*, pages 630–637, 2001.

[3] N. Alon and V. Milman. λ_1, isoperimetric inequalities for graphs, and superconcentrators. *Journal of Combinatorial Theory. Series B*, 38(1):73–88, 1985.

[4] Y. Bilu and N. Linial. Lifts, discrepancy and nearly optimal spectral gap. *Combinatorica*, 26(5):495–519, 2006.

[5] M. Capalbo, O. Reingold, S. Vadhan, and A. Wigderson. Randomness conductors and constant-degree expansion beyond the degree / 2 barrier. In *Proceedings of the 34th STOC*, pages 659–668, 2002.

[6] J. Dodziuk. Difference equations, isoperimetric inequality and transience of certain random walks. *Trans. Amer. Math. Soc.*, 284(2):787–794, 1984.

[7] J. Friedman. A proof of Alon's second eigenvalue conjecture. *Memoirs of the AMS*, to appear.

[8] O. Gabber and Z. Galil. Explicit Constructions of Linear-Sized Superconcentrators. *Journal of Computer and System Sciences*, 22(3):407–420, 1981.

[9] S. Hoory, N. Linial, and A. Wigderson. Expander graphs and their applications. *Bulletin of the AMS*, 43(4):439–561, 2006.

[10] S. Janson, T. Łuczak, and A. Ruciński. *Random graphs*. John Wiley New York, 2000.

[11] S. Jimbo and A. Maruoka. Expanders obtained from affine transformations. *Combinatorica*, 7(4):343–355, 1987.

[12] N. Kahale. Eigenvalues and expansion of regular graphs. *Journal of the ACM*, 42(5):1091–1106, 1995.

[13] A. Lubotzky, R. Philips, and P. Sarnak. Ramanujan graphs. *Combinatorica*, 8:261–277, 1988.

[14] G. A. Margulis. Explicit constructions of expanders. *Problemy Peredaci Informacii*, 9(4):71–80, 1973.

[15] G. A. Margulis. Explicit group-theoretic constructions of combinatorial schemes and their applications in the construction of expanders and concentrators. *Problemy Peredachi Informatsii*, 24(1):51–60, 1988.

[16] R. Meshulam and A. Wigderson. Expanders in group algebras. *Combinatorica*, 24(4):659–680, 2004.

[17] M. Morgenstern. Existence and explicit constructions of $q + 1$ regular Ramanujan graphs for every prime power q. *Journal of Combinatorial Theory. Series B*, 62(1):44–62, 1994.

[18] A. Nilli. On the second eigenvalue of a graph. *Discrete Mathematics*, 91(2):207–210, 1991.

[19] M. Pinsker. On the complexity of a concentrator. In *7th Internat. Teletraffic Confer.*, pages 318/1–318/4, 1973.

[20] O. Reingold. Undirected st-connectivity in log-space. In *Proceedings of the 37th STOC*, pages 376–385, 2005.

[21] O. Reingold, S. Vadhan, and A. Wigderson. Entropy waves, the zig-zag graph product, and new constant-degree expanders. *Annals of Mathematics*, 155(1):157–187, 2002.

[22] E. Rozenman, A. Shalev, and A. Wigderson. Iterative construction of cayley expander graphs. *Theory of Computing*, 2(5):91–120, 2006.

[23] E. Rozenman and S. Vadhan. Derandomized squaring of graphs. In *Proceedings of the 7th RANDOM*, pages 436–447, 2005.