Entropy Waves, The Zig-Zag Graph Product, and New Constant-Degree Expanders

Omer Reingold* Salil Vadhan† Avi Wigderson‡

August 1, 2001

Abstract

The main contribution of this work is a new type of graph product, which we call the **zig-zag product**. Taking a product of a large graph with a small graph, the resulting graph inherits (roughly) its size from the large one, its degree from the small one, and its expansion properties from both! Iteration yields simple explicit constructions of constant-degree expanders of arbitrary size, starting from one constant-size expander.

Crucial to our intuition (and simple analysis) of the properties of this graph product is the view of expanders as functions which act as "entropy wave" propagators — they transform probability distributions in which entropy is concentrated in one area to distributions where that concentration is dissipated. In these terms, the graph product affords the constructive interference of two such waves.

Subsequent work [ALW01, MW01] relates the zig-zag product of graphs to the standard semidirect product of groups, leading to new results and constructions on expanding Cayley graphs.

Keywords: expander graphs, graph products, entropy

*AT&T Labs - Research. Room A243, 180 Park Avenue, Bldg. 103, Florham Park, NJ, 07932, USA. E-mail: omer@research.att.com. Part of this research was performed while visiting the Institute for Advanced Study, Princeton, NJ.

†Division of Engineering & Applied Sciences, Harvard University, 33 Oxford Street, Cambridge, MA 02138. E-mail: salil@eecs.harvard.edu. URL: http://eecs.harvard.edu/~salil. Work done while at MIT, supported by an NSF Mathematical Sciences Postdoctoral Research Fellowship.

‡Institute for Advanced Study, Princeton and the Hebrew University, Jerusalem. Partially supported by NSF grants CCR-9987007 and CCR-9987845

Contents

1 Introduction

1.1 Expander Graphs

Expanders are graphs which are sparse but nevertheless highly connected. A precise definition will be given in the next section, but here we informally list some properties of such graphs (which are equivalent when formally stated and can serve as alternate definitions)

- The graph satisfies "strong" isoperimetric inequalities.
- Every set of vertices has "many" neighbors.
- Every cut has "many" edges crossing it.
- A random walk on the graph converges quickly to the stationary distribution.

Expander graphs have been used to address many fundamental problems in computer science, on topics including network design (e.g. [Pip87, PY82, AKS83]), complexity theory ([Val77, Sip88, Urq87]), derandomization ([NN93, INW94, IW97]), coding theory ([SS96, Spi96]), and cryptography ([GIL+90]). Expander graphs have also found some applications in various areas of pure mathematics [KR83, Lub94, Gro00, LP01].

Standard probabilistic arguments ([Pin73]) show that almost every constant-degree (≥ 3) graph is an expander. However, explicit and efficient construction of such graphs (which is required by most of the computer science applications above) seems to be much harder. This problem lead to an exciting and extensive body of research, developed mainly by mathematicians intrigued by this computer science challenge.

Most of this work was guided by the algebraic characterization of expanders, developed in [Tan84, AM85, Alo86a]. They showed the intimate relation of (appropriate quantitative versions of) all the properties above to the spectral gap in the adjacency matrix (or, almost equivalently, the Laplacian) of the graph. Using it, expanders can be defined as follows: An infinite family G_n of D-regular graphs is an **expander family** if for all n the second largest (in absolute value) eigenvalue of the adjacency matrix of G_n is bounded *uniformly* from above by the same $\lambda < D$. (Note that the degree D is independent of n; this is what we mean by "constant degree.") [1]

This algebraic definition naturally led researchers to consider algebraic constructions, where this eigenvalue can be estimated. The celebrated sequence of papers [Mar73, GG81, AM85, AGM87, JM87, LPS88, Mar88, Mor94] provided such constant-degree expanders. All these graphs are very simple to describe: given the name of a vertex (in binary), its neighbors can be computed in polynomial time (or even logarithmic space). This level of explicitness is essential for many of the applications. However, the analysis bounding the eigenvalue is quite sophisticated (and often based on deep mathematical results). Thus, it is hard to intuitively understand why these graphs are expanders.

A deviation from this path was taken in [Ajt94], where a combinatorial construction of cubic expanders was proposed. It starts with an arbitrary cubic N-vertex graph and applies a sequence of polynomially many local operations which gradually increase the girth and turn it into an expander. However, the resulting graphs do not have any simply described form, and they lack the explicitness level (and hence applicability) of the algebraic constructions mentioned above.

[1] On an intuitive level, the connection between the spectral gap and the combinatorial and probabilistic properties of expanders listed above should not be surprising. For example, it is well known that the standard random walk on the graph converges exponentially with base λ/D to the stationary uniform distribution. Moreover, equal partitions of the vertices of a graph, thought of as ± 1-vectors, are orthogonal to the uniform distribution, and so the bilinear form representing the number of edges in the cut can be bounded in terms of the gap between D and λ.

In this work, we give a simple, combinatorial construction of constant-degree expander graphs. Moreover, the analysis proving expansion (via the second eigenvalue) is as simple and follows a clear intuition. The construction is iterative, and needs as a basic building block a *single, almost arbitrary* expander of constant size. The parameters required from it can be easily obtained explicitly, but exhaustive search is an equally good solution since it requires only constant time. Simple operations applied to this graph generate another whose size is increased but whose degree and expansion remain unchanged. This process continues, yielding arbitrarily large expanders.

The heart of the iteration is our new "zig-zag" graph product. Informally, taking a product of a large graph with a small graph, the resulting graph inherits (roughly) its size from the large one, its degree from the small one, and its expansion properties from both! (That is, the composed graph has good expansion properties as long as the two original graphs have good expansion properties.)

In the next subsections we give high level descriptions of the iterative construction, the new graph product, the intuition behind it, various extensions. We then mention subsequent work on the relation of the zig-zag product in graphs to the semidirect product in groups and its applications to expanding Cayley graphs.

1.2 Overview of Expander Construction

In this section, we describe a simplified, but less efficient, version of our expander construction and omit formal proofs. Our full construction is described in detail in Section 3. Throughout this section, all graphs are regular, undirected, and may have loops and parallel edges. The **adjacency matrix** of an N-vertex graph G is the matrix M whose (u, v)'th entry is the number of edges between vertices u and v. If the graph is D-regular, the **normalized adjacency matrix** is simply M/D. Note that this stochastic matrix is the transition probability matrix of the natural random walk on G, every step of which moves a "token" from a current vertex along a uniformly chosen edge to a neighboring vertex. It is easy to see that this matrix has an eigenvalue of 1, corresponding to the constant eigenvector, and it turns out that all other eigenvalues have absolute value less than 1. Our primary interest will be the second largest (in absolute value) eigenvalue (which is known to govern the convergence rate of the random walk, and as mentioned above is the essence of expansion).

Thus, three essential parameters play a role in an expander — size, degree and expansion. We classify graphs accordingly.

Definition 1.1 *An* (N, D, λ)**-graph** *is any D-regular graph on N vertices, whose normalized adjacency matrix has 2nd largest (in absolute value) eigenvalue at most λ.*

The Basic Operations. We use two operations on (the adjacency matrices of) graphs — the standard matrix squaring, and our new zig-zag graph product. Here is their effect on these three parameters.

SQUARING: Let G^2 denote the square of G. That is, the edges in G^2 are paths of length 2 in G. Then

Fact 1.2 $(N, D, \lambda)^2 \to (N, D^2, \lambda^2)$

THE ZIG-ZAG PRODUCT: Let $G Ⓩ G_2$ denote the zig-zag product of G_1 and G_2. Then,

Theorem 1.3 $(N_1, D_1, \lambda_1) Ⓩ (D_1, D_2, \lambda_2) \to (N_1 \cdot D_1, D_2^2, \lambda_1 + \lambda_2 + \lambda_2^2)$

(The eigenvalue bound of $\lambda_1 + \lambda_2 + \lambda_2^2$ is improved somewhat in Sections 3 and 4.2.)

The Iterations. Let H be any $(D^4, D, 1/5)$-graph, which will serve as the building block for our construction. We define a sequence of graphs G_i as follows.

- $G_1 = H^2$
- $G_{i+1} = G_i^2 ⓩ H$

From Fact 1.2 and Theorem 1.3 above, it is easy to conclude that this sequence is indeed an infinite family of expanders:

Theorem 1.4 *For every i, G_i is an $(N_i, D^2, 2/5)$-graph with $N_i = D^{4i}$*

This construction is not as efficient as we would like — computing neighborhoods in G_i takes time polynomial in N_i rather than polynomial in $\log N_i$. As we show in Section 3, this is easily overcome by augmenting the iterations with another standard graph operation.

1.3 The Zig-Zag Graph Product

The new product mentioned above takes a large graph and a small one, and produces a graph that (roughly speaking) inherits the size of the large one but the degree of the small one. This was the key to creating arbitrarily large graphs with bounded degrees. Naturally, we are concerned with maintaining the expansion properties of the two graphs. First, we describe the product.

For simplicity, we assume that the edges in our D-regular graphs are D-colored; that is, they are partitioned to D perfect matchings. (This assumption loses generality, and we will remove it in the formal construction in Section 2.) For a color $i \in [D]$ and a vertex v let $v[i]$ be the neighbor of v along the edge colored i. With this simple notation, we can formally define the zig-zag productⓩ (and then explain it).

Definition 1.5 *Let G_1 be an D_1-regular graph on $[N_1]$ and G_2 a D_2-regular graph on $[D_1]$. Then $G_1 ⓩ G_2$ is a D_2^2-regular graph on $[N_1] \times [D_1]$ defined as follows: For all $v \in [N_1], k \in [D_1], i, j \in [D_2]$, the edge (i, j) connects the vertex (v, k) to the vertex $(v[k[i]], k[i][j])$.*

What is going on? Note that the size of the small graph G_2 is the degree of the large graph G_1. Thus a vertex name in $G_1 ⓩ G_2$ has a first component which is a vertex of the large graph, and a second which is viewed both as a vertex of the small graph *and* an edge color of the large one. The edge label in $G_1 ⓩ G_2$ is just a pair of edge labels in the small graph. One step in the new product graph from a vertex (v, k) along the edge (i, j) can be broken into three substeps.

1. $(v, k) \to (v, k[i])$ — A step ("zig") in the small graph moving k to $k[i]$. This affects only the second component, according to the first edge label.
2. $(v, k[i]) \to (v[k[i]], k[i])$ — A step in the large graph, changing the first component according to the second, viewed as an edge color.
3. $(v[k[i]], k[i]) \to (v[k[i]], k[i][j])$ – A step ("zag") in the small graph moving $k[i]$ to $k[i][j]$. This affects only the second component, according to the second edge label.

1.4 Intuition

Why does it work? More precisely, why does Theorem 1.3 hold? What this theorem says intuitively, is that $G_1 ⓩ G_2$ is a good expander as long as both G_1 and G_2 are good expanders. Consider the above three steps as a random walk on $G_1 ⓩ G_2$. Then Steps 1 and 3 are independent random steps on the small graph. If at least one of them "works" as well as it does in the small graph, this would guarantee that the new graph is as good expander as the small one. So let's argue (very intuitively) that indeed one of them "works".

A random step in an expander increases the (H_2-) entropy of a distribution on the vertices, *provided that it is not already too close to uniform.* Let us consider a distribution on the vertices of the new graph (v, k). Roughly speaking, there are two cases.

- If the distribution of the second component k (conditioned on v) is not too uniform, then Step 1 "works". Since Step 2 is just a permutation and Step 3 is a random step on a regular graph, these steps cannot make the distribution less uniform and undo the progress made in Step 1.

- If k (conditioned on v) is very close to uniform, then Step 1 is a "waste". However, Step 2 is then like a real random step in the large expander G_1! This means that the entropy of the first component v increases. Note that Step 2 is a permutation on the vertices of $G_1 ⓩ G_2$, so if entropy increases in the first component, it decreases in the second. That means that in Step 3 we are in the good case (the conditional distribution on the second component is far from uniform), and the entropy of the second component will increase by the expansion of the small graph.

The key to this product is that Step 2 is simultaneously a permutation (so that any progress made in Step 1 is preserved) and an operation whose "projection" to the first component is simply a random step on the large graph (when the second component is random). All previous discussions of expanders focused on the increase of entropy to the vertex distribution by a step along a random edge. We insist on keeping track of that edge name, and consider the joint distribution! In a good expander, if the edge is indeed random, the entropy propagates from it to the vertex. This reduces the (conditional) entropy in the edge. Thus the "entropy wave" in Step 2, in which no fresh randomness enters the distribution on vertices of $G_1 ⓩ G_2$, is what facilitates entropy increase in Steps 1 or 3. Either the "zig" step does it, if there is room for more entropy in k, or if not (which may be viewed as destructive interference of the large and small waves in Step 1), Step 2 guarantees constructive interference in Step 3. Moreover, Step 1 is not redundant as, if there is no or little initial entropy in k, the wave of Step 2 (being a permutation) may flood k with entropy, destroying the effect of Step 3.

The formal proof of Theorem 1.3 follows this intuition quite closely, and separately analyzes these two extreme cases. Indeed, since it becomes linear algebra, these two cases are very natural to define, and the only ones to worry about — all intermediate cases follow by linearity! Moreover, the variational definition of the second eigenvalue better captures the symmetry of the zig and zag steps (and gives a better bound than what can be obtained from this asymmetric intuition).

1.5 Expanders and Extractors

Here we attempt an intuitive explanation of how we stumbled on the definition of the zig-zag product, and the intuition that it does what it should. While this subsection may not be self contained, it will at least lead the interested reader to discover more of the fascinating world of extractors.

The current paper is part of research described in our conference paper [RVW00] which deals with constructions of both expanders and extractors. Extractors are combinatorial objects, defined by [NZ96], which, roughly speaking, "purify" arbitrary nonuniform probability distributions into uniform ones. These objects are as fascinating and as applicable as expanders (see, e.g., the survey papers [Nis96, NT99]). Like expanders, their applications demand explicit construction. Like with expanders, the quest for such constructions has been extremely fruitful and illuminating for complexity theory. Unlike expanders, the construction of optimal extractors is still a challenge, although the best existing ones are quite close to optimal (see the current state of the art, as well as a survey of previous constructions, in [RSW00, TUZ01]).

Expander graphs were ingredients in some previous extractor constructions (as extractors may be viewed as graphs as well). Here the situation is reversed. The expander construction of this paper *followed* our discovery of nearly optimal *high min-entropy* extractors, which handle the "purification" of distributions which are already not too far from being uniform. A key idea in approaching optimality (following [RR99]) was preserving the unused entropy in a random step on an extractor. This lead to a (more complex) type of zig-zag product, and from it, iterative constructions of such extractors. Translating this idea to the expander world turned out to be cleaner and more natural than in the extractor world. It lead to our understanding of the role of the edge-name as a keeper of the unused entropy in a step of a standard random walk, and to the zig-zag product defined above.

1.6 Extensions to the Expander Construction

The list below details the extensions and refinements we obtain to the basic expander construction outlined above. All these will be part of the formal sections which follow.

More Explicit Graphs. As mentioned above, this construction is not as efficient as we would like — computing neighborhoods in G_i takes time polynomial in N_i rather than in $\log N_i$. rather than $\mathrm{polylog}(N_i)$. As we show in Section 3, this is easily overcome by augmenting the iterations with another standard graph operation, namely taking tensor powers of the adjacency matrix.

Describing Graphs by "Rotation Maps". Another explicitness problem in the simple construction above is the assumption that the our D-regular graphs are given together with a proper D-coloring of the edges. This property is not preserved by the zig-zag product. To avoid it, we describe graphs more generally by their "rotation maps," and show how this description is explicitly preserved by all graph operations in our construction.

Smaller Degree. A naive and direct implementation of our graph product yields expanders whose degree is reasonable, but not that small (something under 1000). In Section 3.2, we show how to combine this construction, together with *one, constant-size* cycle, to obtain an infinite family of explicit degree 4 expanders. Again, this combination uses the zig-zag product. In fact, using the replacement product described below, we obtain explicit degree 3 expanders (which is the smallest possible).

Choice of the Base Graph. Our expander construction requires an initial "constant size" base graph H as a building block. While exhaustive search can be used to find such an H (since it is constant size), for completeness we include two elementary explicit constructions (from [Alo86b, AR94]) which can be used instead.

Better Degree vs. Eigenvalue Relation. The best relationship between degree and 2nd largest eigenvalue is obtained by **Ramanujan** graphs, in which the 2nd eigenvalue is $2\sqrt{D-1}/D$. This equals the first eigenvalue of the D-regular infinite tree, and it is known that no finite D-regular graph can have a smaller 2nd largest eigenvalue (cf., [Alo86a, LPS88, Nil91]). Remarkable graphs achieving this optimal bound were first constructed independently by [LPS88] (who coined the term Ramanujan graphs) and by [Mar88].

Our constructions do not achieve this tight relationship. The zig-zag product, applied recursively to one fixed Ramanujan graph, will yield D-regular expanders of 2nd largest eigenvalue $O(1/D^{1/4})$. A "partially derandomized" variant of our zig-zag product, given in Section 6, improves this relation and achieves second eigenvalue $O(1/D^{1/3})$.

A Simpler Product. Perhaps the most natural way to combine G_1 with G_2 when the size of G_2 is the degree of G_1 is simply replace every vertex of G_1 with a copy of G_2 in the natural way, keeping the edges of both graphs. This **replacement product**, which was often used for degree-reduction purposes (e.g., when G_2 is a cycle the resulting graph has degree 3) turns out to enjoy similar properties of the zig-zag product: if both G_1 and G_2 are expanders, so is their replacement product. Moreover, the proof is by a reduction — the zig-zag product is a subgraph of the cube (3rd power) of the replacement product, immediately giving an eigenvalue bound.

1.7 Subsequent Work: Connections with Semidirect Product in Groups

Subsequent to this work, it was shown in [ALW01] that the zig-zag (and replacement) products can be viewed as a generalization of the standard semidirect product of groups. This was used in [ALW01] to construct a family of groups which is expanding with one (constant size) set of generators, but is not expanding with another such set. The connection was further developed in [MW01] to produce new families of expanding Cayley graphs, via bounds on the the number of irreducible representations of different dimensions in terms of the expansion.

1.8 Organization of the Paper

In Section 2, we give preliminary definitions and basic facts. In Section 3, we define the zig-zag graph product, describe the construction of expanders, and state their properties. In particular, it deals with the first four "extensions" listed in the previous subsection. In Section 4, we analyze the expansion of the zig-zag product. In Section 5, we discuss some ways to obtain the base graph used in our expander construction. In Section 6, we give two extensions to the basic zig-zag product. The first is a "derandomized" variant of our basic zig-zag product, which enjoys a better relationship between the degree and the expansion. The second is the simple, natural *replacement* product.

2 Preliminaries

2.1 Graphs and Rotations

All graphs we discuss may have self loops and parallel edges. They are best described by their (nonnegative, integral) adjacency matrix. Such a graph is **undirected** iff the adjacency matrix is symmetric. It is **D-regular** if the sum of entries in each row (and column) is D (so exactly D edges are incident to every vertex).

Let G be a D-regular undirected graph on N vertices. Suppose that the edges leaving each vertex of G are labeled from 1 to D in some arbitrary, but fixed, way. Then for $v, w \in [N]$ and $i \in [D]$, it makes sense (and is standard) to say "the i'th neighbor of vertex v is w". In this work, we make a point to always keep track of the edge traversed to get from v to w. This is formalized as follows:

Definition 2.1 *For a D-regular undirected graph G, the* **rotation map** $\mathrm{Rot}_G : [N] \times [D] \to [N] \times [D]$ *is defined as follows:* $\mathrm{Rot}_G(v, i) = (w, j)$ *if the i'th edge incident to v leads to w, and this edge is the j'th edge incident to w.*

This definition enables us to remove the simplifying assumption made in the introduction, which was that the label of an edge is the same from the perspective of both endpoints, i.e. $\mathrm{Rot}_G(v,i) = (w,j) \Rightarrow i = j$. From Definition 2.1, it is clear that Rot_G is a permutation, and moreover $\mathrm{Rot}_G \circ \mathrm{Rot}_G$ is the identity map.

We will always view graphs as being specified by their rotation maps. Hence we call a family $\mathcal{G}$ of graphs **explicit** if for every $G \in \mathcal{G}$, Rot_G is computable in time $\mathrm{poly}(\log N)$, where N is the number of vertices of G. That is, graphs in $\mathcal{G}$ are indexed by some parameters (such as the number of vertices and the degree, which may be required to satisfy some additional relations) and there should be a single algorithm which efficiently computes Rot_G for any $G \in \mathcal{G}$ when given these parameters as an additional input. The notation $\mathrm{poly}()$ stands for a fixed (but unspecified) polynomial function in the given variables. We will often informally refer to an individual graph as explicit, as shorthand for saying that the graph comes from an explicit family.

Our constructions will be iterative (or recursive), and will be based on a sequence of composition operations, constructing new graphs from given ones. The definition of these compositions (or products) will show how the rotation map of the new graph can be computed using "oracle access" to the rotation maps of the given graphs. (By giving an algorithm "oracle access" to a function f, we mean that the algorithm is given power to evaluate f on inputs of its choice at the cost of 1 time step per evaluation.) Given the time complexity of such a computation *and* the number of oracle calls made, it will be easy to compute the total time required by a recursive construction.

2.2 Eigenvalues and Expansion

The **normalized adjacency matrix** M of G is the adjacency matrix of G divided by D. In terms of the rotation map, we have:

$$M_{u,v} = \frac{1}{D} \cdot \left| \{(i,j) \in [D]^2 : \mathrm{Rot}_G(u,i) = (v,j)\} \right| .$$

M is simply the transition matrix of a random walk on G. By the D-regularity of G, the all-1's vector $1_N = (1,1,\ldots,1) \in \mathbb{R}^N$ is an eigenvector of M of eigenvalue 1. It is turns out that all the other eigenvalues of M have absolute value at most 1, and it is well-known that the second largest eigenvalue of G is a good measure of G's expansion properties [Tan84, AM85, Alo86a]. We will use the following variational characterization of the second largest eigenvalue.

Definition 2.2 $\lambda(G)$ *denotes the* **second largest eigenvalue** *(in absolute value) of G's normalized adjacency matrix. Equivalently,*

$$\lambda(G) = \max_{\alpha \perp 1_N} \frac{|\langle \alpha, M\alpha \rangle|}{\langle \alpha, \alpha \rangle} = \max_{\alpha \perp 1_N} \frac{\|M\alpha\|}{\|\alpha\|}.$$

Above, $\langle \cdot, \cdot \rangle$ refers to the standard inner product in $\mathbb{R}^N$ and $\|\alpha\| = \sqrt{\langle \alpha, \alpha \rangle}$.

The meaning of $\lambda(G)$ can be understood as follows: Suppose $\pi \in [0,1]^N$ is a probability distribution on the vertices of G. By linear algebra, π can be decomposed as $\pi = u_N + \pi^\perp$, where $u_N = 1_N/N$ is the uniform distribution and $\pi^\perp \perp u_N$. Then $M\pi = u_N + M\pi^\perp$ is the probability distribution on vertices obtained by selecting a vertex v according to π and then moving to a uniformly selected neighbor of v. By Definition 2.2, $\|M\pi^\perp\| \le \lambda(G) \cdot \|\pi^\perp\|$. Thus $\lambda(G)$ is a measure of how quickly the random walk on G converges to the uniform distribution. Intuitively, the smaller $\lambda(G)$ is, the better the expansion properties of G. Accordingly, an (infinite) family $\mathcal{G}$ of graphs is called a family of **expanders** if these eigenvalues are bounded away from 1, i.e. there is a constant $\lambda < 1$ such that $\lambda(G) \le \lambda$ for all $G \in \mathcal{G}$. It was shown by Tanner [Tan84] and Alon and Milman [AM85] that this implies (and is in fact equivalent to [Alo86a]) the standard notion of **vertex expansion**: there is a constant $\varepsilon > 0$ such that for every $G \in \mathcal{G}$ and for any set S of at most half the vertices in G, at least $(1+\varepsilon) \cdot |S|$ vertices of G are connected to some vertex in S.

As mentioned in the introduction, we refer to a D-regular undirected graph G on N vertices such that $\lambda(G) \leq \lambda$ as an (N, D, λ)**-graph**. Clearly, achieving expansion is easier as the degree gets larger. The main goal in constructing expanders is to minimize the degree, and, more generally, obtain the best degree-expansion tradeoff. Using the Probabilistic Method, Pinsker [Pin73] showed that most 3-regular graphs are expanders (in the sense of vertex expansion), and this result was extended to eigenvalue bounds in [Alo86a, BS87, FKS89, Fri91]. The best known bound on the eigenvalues of random graphs is due to Friedman [Fri91], who showed that most D-regular graphs have second largest eigenvalue at most $2/\sqrt{D} + O((\log D)/D)$ (for even D). In fact, the bound of $2\sqrt{D-1}/D$ is the best possible for an infinite family of graphs, as shown by Alon and Boppana (cf., [Alo86a, LPS88, Nil91]). Graphs whose second largest eigenvalue meets this optimal bound are called **Ramanujan graphs**. It is easy to verify that this value is the *largest* eigenvalue of the random walk on the *infinite* D-regular tree.

While these probabilistic arguments provide strong existential results, applications of expanders in computer science often require *explicit* families of constant-degree expanders. The first such construction was given by Margulis [Mar73], with improvements and simplifications by Gabber and Galil [GG81], Jimbo and Maruoka [JM87], Alon and Milman [AM85], and Alon, Galil, and Milman [AGM87]. Explicit families of Ramanujan graphs were first constructed by Lubotzky, Phillips, and Sarnak [LPS88] and Margulis [Mar88], with more recent constructions given by Morgenstern [Mor94].The best eigenvalues we know how to achieve using our approach are $O(1/D^{1/3})$.

2.3 Squaring and Tensoring

In addition to the new zig-zag product, our expander construction makes use of two standard operations on graphs — squaring and tensoring. Here we describe these operations in terms of rotation maps and state their effects on the eigenvalues.

Let G be a D-regular multigraph on $[N]$ given by rotation map Rot_G. The t**'th power** of G is the D^t-regular graph G^t whose rotation map is given by $\mathrm{Rot}_{G^t}(v_0, (k_1, k_2, \ldots, k_t)) = (v_t, (\ell_t, \ell_{t-1}, \ldots, \ell_1))$, where these values are computed via the rule $(v_i, \ell_i) = \mathrm{Rot}_G(v_{i-1}, k_i)$.

Proposition 2.3 *If G is an (N, D, λ)-graph, then G^t is an (N, D^t, λ^t)-graph. Moreover,* Rot_{G^t} *is computable in time* $\mathrm{poly}(\log N, \log D, t)$ *with t oracle queries to* Rot_G.

Proof: The normalized adjacency matrix of G^t is the t'th power of the normalized adjacency matrix of G, so all the eigenvalues also get raised to the t'th power. ■

Let G_1 be a D_1-regular multigraph on $[N_1]$ and let G_2 be a D_2-regular multigraph on $[N_2]$. Define the **tensor product** $G_1 \otimes G_2$ to be the $D_1 \cdot D_2$-regular multigraph on $[N_1] \times [N_2]$ given by $\mathrm{Rot}_{G_1 \otimes G_2}((v, w), (i, j)) = ((v', w'), (i', j'))$, where $(v', i') = \mathrm{Rot}_{G_1}(v, i)$ and $(w', j') = \mathrm{Rot}_{G_2}(w, j)$. In order to analyze this construction (and our new graph product), we need some concepts from linear algebra. For vectors $\alpha \in \mathbb{R}^{N_1}$ and $\beta \in \mathbb{R}^{N_2}$, their **tensor product** is the vector $\alpha \otimes \beta \in \mathbb{R}^{N_1 \cdot N_2}$ whose (i, j)'th entry is $\alpha_i \cdot \beta_j$. If A is an $N_1 \times N_1$ matrix and B is an $N_2 \times N_2$ matrix, there is a unique $N_1 N_2 \times N_1 N_2$ matrix $A \otimes B$ (again called the **tensor product**) such that $(A \otimes B)(\alpha \otimes \beta) = (A\alpha) \otimes (B\beta)$ for all α, β.

Proposition 2.4 *If G_1 is an (N_1, D_1, λ_1)-graph and G_2 is an (N_2, D_2, λ_2)-graph, then $G_1 \otimes G_2$ is an $(N_1 \cdot N_2, D_1 \cdot D_2, \max(\lambda_1, \lambda_2))$-graph. Moreover,* $\mathrm{Rot}_{G_1 \otimes G_2}$ *is computable in time* $\mathrm{poly}(\log N_1 N_2, \log D_1 D_2)$ *with one oracle query to* Rot_{G_1} *and one oracle query to* Rot_{G_2}.

Proof: The normalized adjacency matrix of $G_1 \otimes G_2$ is the tensor product of the normalized adjacency matrices of G_1 and G_2. Hence its eigenvalues are the pairwise products of eigenvalues of G_1 and G_2. The largest eigenvalue is $1 \cdot 1$, and the second largest eigenvalue is either $1 \cdot \lambda_2$ or $\lambda_1 \cdot 1$. ■

3 The Zig-Zag Product and the Expander Construction

In the introduction, we described how to obtain a family of expanders by iterating two operations on graphs — squaring and the new "zig-zag" product. That description used a simplifying assumption about the edge labeling. In terms of rotation maps, the assumption was that $\mathrm{Rot}(v, i) = (w, j) \Rightarrow i = j$. In this section, we describe the construction in terms of arbitrary rotation maps and prove its properties. The expander construction given here will also use tensoring to improve the efficiency to polylogarithmic in the number of vertices. This deals with the first two items in the "extensions" subsection of the introduction, which are summarized in Theorem 3.2. The third item — obtaining expanders of degree 4 will follow in Corollary 3.4. The analysis of the zig-zag product is deferred to the following section.

3.1 The Zig-Zag Graph Product

We begin by describing the new graph product in terms of rotation maps. Let G_1 be a D_1-regular multigraph on $[N_1]$ and G_2 a D_2-regular multigraph on $[D_1]$. Their **zig-zag product** is a D_2^2-regular multigraph $G_1 ⓩ G_2$ on $[N_1] \times [D_1]$. We view every vertex v of G_1 as being blown up to a "cloud" of D_1 vertices $(v, 1), \ldots, (v, D_1)$, one for each edge of G_1 leaving v. Thus for every edge $e = (v, w)$ of G_1, there are two associated vertices of $G_1 ⓩ G_2$ — (v, k) and (w, ℓ), where e is the k'th edge leaving v and the ℓ'th edge leaving w. Note that these pairs satisfy the relation $(w, \ell) = \mathrm{Rot}_{G_1}(v, k)$. Since G_2 is a graph on $[D_1]$, we can also imagine connecting the vertices of each such cloud using the edges of G_2. Now, the edges of $G_1 ⓩ G_2$ are defined (informally) as follows: we connect two vertices (v, k) and (w, ℓ) if it is possible to get from (v, k) to (w, ℓ) by a sequence of moves of the following form:

1. Move to a neighboring vertex (v, k') within the initial cloud (using an edge of G_2).
2. Jump across clouds (using edge k' of G_1) to get to (w, ℓ').
3. Move to a neighboring vertex (w, ℓ) within the new cloud (using an edge of G_2).

To make this precise, we describe how to compute the $\mathrm{Rot}_{G_1 ⓩ G_2}$ given Rot_{G_1} and Rot_{G_2}.

Definition 3.1 *If G_1 is a D_1-regular graph on $[N_1]$ with rotation map Rot_{G_1} and G_2 is a D_2-regular graph on $[D_1]$ with rotation map Rot_{G_2}, then their* **zig-zag product** *$G_1 ⓩ G_2$ is defined to be the D_2^2-regular graph on $[N_1] \times [D_1]$ whose rotation map $\mathrm{Rot}_{G_1 ⓩ G_2}$ is as follows:*

$\mathrm{Rot}_{G_1 ⓩ G_2}((v, k), (i, j))$**:**

1. *Let* $(k', i') = \mathrm{Rot}_{G_2}(k, i)$.
2. *Let* $(w, \ell') = \mathrm{Rot}_{G_1}(v, k')$.
3. *Let* $(\ell, j') = \mathrm{Rot}_{G_2}(\ell', j)$.
4. *Output* $((w, \ell), (j', i'))$.

The important feature of this graph product is that $G_1 ⓩ G_2$ is a good expander if both G_1 and G_2 are, as shown by the following theorem.

Theorem 3.2 *If G_1 is an (N_1, D_1, λ_1)-graph and G_2 is a (D_1, D_2, λ_2)-graph, then $G_1 ⓩ G_2$ is a $(N_1 \cdot D_1, D_2^2, f(\lambda_1, \lambda_2))$-graph, where $f(\lambda_1, \lambda_2) \leq \lambda_1 + \lambda_2 + \lambda_2^2$ and $f(\lambda_1, \lambda_2) < 1$ when $\lambda_1, \lambda_2 < 1$. Moreover, $\mathrm{Rot}_{G_1 ⓩ G_2}$ can be computed in time $\mathrm{poly}(\log N, \log D_1, \log D_2)$ with one oracle query to Rot_{G_1} and two oracle queries to Rot_{G_2}.*

Stronger bounds on the function $f(\lambda_1, \lambda_2)$ are given in Section 4.2. Before proving Theorem 3.2, we show how it can be used to construct an infinite family of constant-degree expanders starting from a constant-size expander.

3.2 The Recursion

The construction is like the construction in the introduction, except that we use tensoring to reduce the depth of the recursion and thereby make the construction run in polylogarithmic time (in the size of the graph).

Let H be a (D^8, D, λ)-graph for some D and λ. (Various method for obtaining such an H are described in Section 5.) For every $t \geq 1$, we will define a (D^{8t}, D^2, λ_t)-graph G_t. G_1 is H^2 and G_2 is $H \otimes H$. For $t > 2$, G_t is recursively defined by

$$G_t = \left(G_{\lceil \frac{t-1}{2} \rceil} \otimes G_{\lfloor \frac{t-1}{2} \rfloor} \right)^2 ⓩ H.$$

Theorem 3.3 *For every $t \geq 0$, G_t is an (D^{8t}, D^2, λ_t)-graph with $\lambda_t = \lambda + O(\lambda^2)$. Moreover, Rot_{G_t} can be computed in time $\mathrm{poly}(t, \log D)$ with $\mathrm{poly}(t)$ oracle queries to Rot_H.*

Proof: A straightforward induction establishes that the number of vertices in G_t is D^{8t} and that its degree is D^2. To analyze the eigenvalues, define $\mu_t = \max\{\lambda_1, \ldots, \lambda_t\}$. Then we have $\mu_t \leq \max\{\mu_{t-1}, \mu_{t-1}^2 + \lambda + \lambda^2\}$ for all $t \geq 2$. Solving this recurrence gives $\mu_t \leq \lambda + O(\lambda^2)$ for all t. For the efficiency, note that the depth of the recursion is at most $\log_2 t$ and evaluating the rotation maps for G_t requires 4 evaluations of rotation maps for smaller graphs, so the total number of recursive calls is at most $4^{\log_2 t} = t^2$. ■

In order for Theorem 3.3 to guarantee that graphs $\{G_t\}$ are expanders, the second largest eigenvalue λ of the building block H must be sufficiently small (say, $\lambda \leq 1/5$). This forces the degree of H and hence the degree of the expander family to be rather large, though still constant. However, by zig-zagging the family $\{G_t\}$ with a cycle, we can obtain a family of degree 4 expanders. More generally, we can use this method convert any family of odd-degree expanders into a family of degree 4 expanders:

Corollary 3.4 *For every $\lambda < 1$ and every odd D, there exists a $\lambda' < 1$ such that if G is an (N, D, λ)-graph and C is the cycle on D vertices, then $G ⓩ C$ is a $(ND, 4, \lambda')$-graph.*

Proof: As with any connected and nonbipartite graph, $\lambda(C)$ is strictly less than 1 for an odd cycle C (though $\lambda(C) \to 1$ as $D \to \infty$). Thus, the corollary follows from Theorem 3.2. ■

4 Analysis of the Zig-Zag Product

This section has two subsections. In the first, we give the basic (suboptimal) bound of Theorem 3.2. This bound uses only the intuitive ideas of the introduction, and suffices for the construction of the previous section. In the next, we state and prove a tighter eigenvalue bound. It uses extra information about the zig-zag product (which is less intuitive). It also gives more information about the worst interplay between the two extreme cases studied in the basic analysis, and may hopefully shed a bit of light on the structure of the eigenvectors of the zig-zag product.

4.1 The Basic Eigenvalue Bound

Now we prove Theorem 3.2. Recall the intuition behind the zig-zag product. We aim to show that for any (non-uniform) initial probability distribution π on the vertices of $G_1 Ⓩ G_2$, taking a random step on $G_1 Ⓩ G_2$ results in a distribution that is more uniform. We argued this intuitively in the introduction, by considering two extreme cases, based on the conditional distributions induced by π on the N_1 "clouds" of D_1 vertices each: one in which these conditional distributions are far from uniform, and the second in which they are uniform. The actual linear algebra proof below will restrict itself to these two cases by decomposing any other vector into a linear combination of the two. Also, the argument in the introduction was not symmetric in the first and second steps on the small graph. Using the variational definition of the second largest eigenvalue, we get a cleaner analysis than by following that intuition directly.

Let M be the normalized adjacency matrix of $G_1 Ⓩ G_2$. According to Definition 2.2, we must show that, for every vector $\alpha \in \mathbb{R}^{N_1 \cdot D_1}$ such that $\alpha \perp 1_{N_1 D_1}$, $|\langle M\alpha, \alpha\rangle|$ is smaller than $\langle \alpha, \alpha \rangle$ by a factor $f(\lambda_1, \lambda_2)$. For intuition, α should be thought of as the nonuniform component of the probability distribution π referred to above, i.e. $\pi = u_{N_1 D_1} + \alpha$, where $u_{N_1 D_1} = 1_{N_1 D_1}/N_1 D_1$ is the uniform distribution on $[N_1 D_1]$. Thus, we are showing that π becomes more uniform after a random step on $G_1 Ⓩ G_2$.

For every $v \in [N_1]$, define $\alpha_v \in \mathbb{R}^{D_1}$ by $(\alpha_v)_k = \alpha_{vk}$. Also define a (linear) map $C : \mathbb{R}^{N_1 \cdot D_1} \to \mathbb{R}^{N_1}$ by $(C\alpha)_v = \sum_{k=1}^{D_1} \alpha_{vk}$. Thus, for a probability distribution π on the vertices of $G_1 Ⓩ G_2$, π_v is a multiple of the conditional distribution on "cloud v" and $C\pi$ gives the marginal distribution on set of clouds. By definition, $\alpha = \sum_v e_v \otimes \alpha_v$, where e_v denotes the v'th standard basis vector in $\mathbb{R}^{N_1}$. By basic linear algebra, every α_v can be decomposed (uniquely) into $\alpha_v = \alpha_v^{||} + \alpha_v^{\perp}$ where $\alpha_v^{||}$ is parallel to 1_{D_1} (i.e., all of its entries are the same) and $\alpha_v^{\perp}$ is orthogonal to 1_{D_1} (i.e., the sum of its entries are 0). Thus, we obtain a decomposition of α:

$$
\begin{aligned}
\alpha &= \sum_v e_v \otimes \alpha_v \\
&= \sum_v e_v \otimes \alpha_v^{||} + \sum_v e_v \otimes \alpha_v^{\perp} \\
&\stackrel{\text{def}}{=} \alpha^{||} + \alpha^{\perp}
\end{aligned}
$$

This decomposition corresponds to to the two cases in our intuition: $\alpha^{||}$ corresponds to a probability distribution on the vertices of $G_1 Ⓩ G_2$ such that the conditional distributions on the clouds are all uniform. $\alpha^{\perp}$ corresponds to a distribution such that the conditional distributions on the clouds are all far from uniform. Another way of matching $\alpha^{||}$ with the intuition is to note that $\alpha^{||} = C\alpha \otimes 1_{D_1}/D_1$. Since α and $\alpha^{\perp}$ are both orthogonal to $1_{N_1 D_1}$, so is $\alpha^{||}$ and hence also $C\alpha$ is orthogonal to 1_{N_1}.

To analyze how M acts on these two vectors, we relate M to the normalized adjacency matrices of G_1 and G_2, which we denote by A and B, respectively. First, we decompose M into the product of three matrices, corresponding to the three steps in the definition of $G_1 Ⓩ G_2$'s edges. Let $\tilde{B}$ be the (normalized) adjacency matrix of the graph on $[N_1] \times [D_1]$ where we connect the vertices within each cloud according to the edges of G_2. $\tilde{B}$ is related to B by the relation $\tilde{B} = I_{N_1} \otimes B$, where I_{N_1} is the $N_1 \times N_1$ identity matrix. Let $\tilde{A}$ be the permutation matrix corresponding to Rot_{G_1}. The relationship between $\tilde{A}$ and A is somewhat subtle, so we postpone describing it until later. By the definition of $G_1 Ⓩ G_2$, we have $M = \tilde{B}\tilde{A}\tilde{B}$. Note that both $\tilde{B}$ and $\tilde{A}$ are symmetric matrices, due to the undirectedness of G_1 and G_2.

Recall that we want to bound $|\langle M\alpha, \alpha\rangle| / \langle \alpha, \alpha\rangle$. By the symmetry of $\tilde{B}$, we have

$$\langle M\alpha, \alpha\rangle = \langle \tilde{B}\tilde{A}\tilde{B}\alpha, \alpha\rangle = \langle \tilde{A}\tilde{B}\alpha, \tilde{B}\alpha\rangle. \tag{1}$$

Now note that $\tilde{B}\alpha^{||} = \alpha^{||}$, because $\alpha^{||} = C\alpha \otimes 1_{D_1}/D_1$, $\tilde{B} = I_{N_1} \otimes B$, and $B 1_{D_1} = 1_{D_1}$. This corresponds to the fact that if the conditional distribution within each cloud is uniform, then taking a random G_2-step does nothing. Hence, $\tilde{B}\alpha = \tilde{B}(\alpha^{||} + \alpha^{\perp}) = \alpha^{||} + \tilde{B}\alpha^{\perp}$. Substituting this into (1), we have

$$\langle M\alpha, \alpha\rangle = \langle \tilde{A}(\alpha^{||} + \tilde{B}\alpha^{\perp}), \alpha^{||} + \tilde{B}\alpha^{\perp}\rangle. \tag{2}$$

Expanding and using the fact that $\tilde{A}$ is length-preserving (because it is a permutation matrix), we have

$$|\langle M\alpha, \alpha\rangle| \quad \leq \quad |\langle \tilde{A}\alpha^{||}, \alpha^{||}\rangle| + 2\|\alpha^{||}\| \cdot \|\tilde{B}\alpha^{\perp}\| + \|\tilde{B}\alpha^{\perp}\|^2. \tag{3}$$

Now we apply the expansion properties of G_1 and G_2 to bound each of these terms. First, we bound $\|\tilde{B}\alpha^{\perp}\|$, which corresponds to the intuition that when the conditional distributions within the clouds are far from uniform, they become more uniform when we take a random G_2-step.

Claim 4.1 $\|\tilde{B}\alpha^{\perp}\| \leq \lambda_2 \cdot \|\alpha^{\perp}\|$.

Proof of claim:

$$\begin{aligned}\tilde{B}\alpha^{\perp} &= \tilde{B}\left(\sum_v e_v \otimes \alpha_v^{\perp}\right) \\ &= \sum_v e_v \otimes B\alpha_v^{\perp}.\end{aligned}$$

By the expansion of G_2, $\|B\alpha_v^{\perp}\| \leq \lambda_2 \cdot \|\alpha_v^{\perp}\|$ for all v. Hence, $\|\tilde{B}\alpha^{\perp}\| \leq \lambda_2 \cdot \|\alpha^{\perp}\|$. □

Next, we bound $|\langle \tilde{A}\alpha^{||}, \alpha^{||}\rangle|$, which corresponds to the intuition that when the conditional distribution within each cloud is uniform, the jump between the clouds makes the marginal distribution on clouds themselves more uniform.

Claim 4.2 $|\langle \tilde{A}\alpha^{||}, \alpha^{||}\rangle| \leq \lambda_1 \cdot \langle \alpha^{||}, \alpha^{||}\rangle$.

Proof of claim: To prove this, we must first relate $\tilde{A}$ to A. Recall that, when k is uniformly distributed, $\mathrm{Rot}_{G_1}(v, k)$ gives a pair (w, ℓ) where w is a uniformly selected neighbor of v. Similarly, if $e_v \in \mathbb{R}^{N_1}$ is the v'th standard basis vector, then Ae_v gives the uniform distribution over the neighbors of v. This similarity is captured by the formula $C\tilde{A}(e_v \otimes 1_{D_1}/D_1) = Ae_v$ for all v. (Tensoring e_v with $1_{D_1}/D_1$ corresponds to taking the uniform distribution over k and applying C corresponds to discarding ℓ and looking just at w.) Because the e_v's form a basis, this formula extends to all vectors $\beta \in \mathbb{R}^{N_1}$: $C\tilde{A}(\beta \otimes 1_{D_1}/D_1) = A\beta$. Applying this formula to $\alpha^{||} = C\alpha \otimes 1_{D_1}/D_1$, we have $C\tilde{A}(\alpha^{||}) = AC\alpha$. Thus,

$$\begin{aligned}\langle \tilde{A}\alpha^{||}, \alpha^{||}\rangle &= \langle \tilde{A}\alpha^{||}, C\alpha \otimes 1_{D_1}\rangle / D_1 \\ &= \langle C\tilde{A}\alpha^{||}, C\alpha\rangle / D_1 \\ &= \langle AC\alpha, C\alpha\rangle / D_1.\end{aligned}$$

Recalling that $C\alpha$ is orthogonal to 1_{N_1}, we may apply the expansion of G_1 to obtain:

$$\begin{aligned}|\langle \tilde{A}\alpha^{||}, \alpha^{||}\rangle| &\leq \lambda_1 \cdot \langle C\alpha, C\alpha\rangle / D_1 \\ &= \lambda_1 \cdot \langle C\alpha \otimes 1_{D_1}, C\alpha \otimes 1_{D_1}\rangle / D_1^2 \\ &= \lambda_1 \cdot \langle \alpha^{||}, \alpha^{||}\rangle,\end{aligned}$$

□

Substituting the bounds of Claim 4.1 and 4.2 into (3), we have:

$$|\langle M\alpha, \alpha\rangle| \leq \lambda_1 \cdot \|\alpha^{||}\|^2 + 2\lambda_2 \cdot \|\alpha^{||}\| \cdot \|\alpha^{\perp}\| + \lambda_2^2 \cdot \|\alpha^{\perp}\|^2 \tag{4}$$

If we let $p = \|\alpha^{||}\|/\|\alpha\|$ and $q = \|\alpha^{\perp}\|/\|\alpha\|$, then $p^2 + q^2 = 1$, and the above expression can be rewritten as:

$$\frac{|\langle M\alpha, \alpha\rangle|}{\langle \alpha, \alpha\rangle} \leq \lambda_1 \cdot p^2 + 2\lambda_2 \cdot pq + \lambda_2^2 \cdot q^2 \leq \lambda_1 + \lambda_2 + \lambda_2^2.$$

This shows that we can take $f(\lambda_1, \lambda_2) \leq \lambda_1 + \lambda_2 + \lambda_2^2$. It remains to show that we can set $f(\lambda_1, \lambda_2) < 1$ as long as $\lambda_1, \lambda_2 < 1$. We consider two cases, depending on the length of $\|\alpha^{\perp}\|$. First, suppose that $\|\alpha^{\perp}\| \leq \frac{1-\lambda_1}{3\lambda_2} \cdot \|\alpha\|$. Then, from (4), we have

$$|\langle M\alpha, \alpha\rangle| \leq \lambda_1 \cdot \|\alpha\|^2 + 2\lambda_2 \cdot \left(\frac{1-\lambda_1}{3\lambda_2}\right)\|\alpha\|^2 + \lambda_2^2 \cdot \left(\frac{1-\lambda_1}{3\lambda_2}\right)^2 \|\alpha\|^2 < \left(1 - \frac{1-\lambda_1}{9}\right) \cdot \|\alpha\|^2.$$

Now suppose that $\|\alpha^{\perp}\| > \frac{1-\lambda_1}{3\lambda_2} \cdot \|\alpha\|$. Notice that $\tilde{B}\alpha^{\perp}$ is orthogonal to $\alpha^{||}$: $\langle \tilde{B}\alpha^{\perp}, \alpha^{||}\rangle = \langle \alpha^{\perp}, \tilde{B}\alpha^{||}\rangle = \langle \alpha^{\perp}, \alpha^{||}\rangle = 0$. Using this, we can bound (2) as follows:

$$\begin{aligned} |\langle M\alpha, \alpha\rangle| &= |\langle \tilde{A}(\alpha^{||} + \tilde{B}\alpha^{\perp}), \alpha^{||} + \tilde{B}\alpha^{\perp}\rangle| \leq \|\alpha^{||} + \tilde{B}\alpha^{\perp}\|^2 = \|\alpha^{||}\|^2 + \|\tilde{B}\alpha^{\perp}\|^2 \\ &\leq \|\alpha\|^2 - \|\alpha^{\perp}\|^2 + \lambda_2^2 \cdot \|\alpha^{\perp}\|^2 \leq \|\alpha\|^2 - (1-\lambda_2^2) \cdot \left(\frac{1-\lambda_1}{3\lambda_2}\right)^2 \cdot \|\alpha\|^2. \end{aligned}$$

Thus, we can take

$$f(\lambda_1, \lambda_2) \leq 1 - \min\left\{\frac{1-\lambda_1}{9}, \frac{(1-\lambda_1)^2 \cdot (1-\lambda_2^2)}{9\lambda_2^2}\right\} < 1.$$

4.2 Improved Analysis of the Eigenvalue

In this subsection we state and prove an improved upper bound on the second largest eigenvalue produced by the zig-zag product.

Theorem 4.3 (Thm. 3.2, improved) *If G_1 is an (N_1, D_1, λ_1)-graph and G_2 is a (D_1, D_2, λ_2)-graph, then $G_1 Ⓩ G_2$ is a $(N_1 \cdot D_1, D_2^2, f(\lambda_1, \lambda_2))$-graph, where*

$$f(\lambda_1, \lambda_2) = \frac{1}{2}(1-\lambda_2^2)\lambda_1 + \frac{1}{2}\sqrt{(1-\lambda_2^2)^2\lambda_1^2 + 4\lambda_2^2}.$$

Although the function $f(\lambda_1, \lambda_2)$ looks ugly, it can be verified that it has the following nice properties:

1. $f(\lambda, 0) = f(0, \lambda) = \lambda$ and $f(\lambda, 1) = f(1, \lambda) = 1$ for all $\lambda \in [0, 1]$.

2. $f(\lambda_1, \lambda_2)$ is a strictly increasing function of both λ_1 and λ_2 (except when one of them is 1).

3. If $\lambda_1 < 1$ and $\lambda_2 < 1$, then $f(\lambda_1, \lambda_2) < 1$.

4. $f(\lambda_1, \lambda_2) \leq \lambda_1 + \lambda_2$ for all $\lambda_1, \lambda_2 \in [0, 1]$.

Proof: The proof proceeds along the same lines as the proof of Theorem 3.2, except that we will use a geometric argument to directly bound (2) rather than first passing to (3). That is, we must bound (using the same notation as in that proof)

$$\frac{\langle M\alpha, \alpha\rangle}{\langle \alpha, \alpha\rangle} = \frac{\langle \tilde{A}(\alpha^{||} + \tilde{B}\alpha^{\perp}), \alpha^{||} + \tilde{B}\alpha^{\perp}\rangle}{\|\alpha^{||} + \alpha^{\perp}\|^2}.$$

The key observation is:

Claim 4.4 *$\tilde{A}$ is a reflection through a linear subspace S of $\mathbb{R}^{N_1 D_1}$. Hence, for any any vector v, $\langle \tilde{A}v, v\rangle = (\cos 2\theta)\cdot \|v\|^2$, where θ is the angle between v and S.*

> **Proof of claim:** By the symmetry of $\tilde{A}$, we can decompose $\mathbb{R}^{N_1 D_1}$ into the sum of orthogonal eigenspaces of $\tilde{A}$. Since $\tilde{A}^2 = I_{N_1 D_1}$, the only eigenvalues of $\tilde{A}$ are ± 1. Take S to be the 1-eigenspace of $\tilde{A}$. □

Thus, the expression we want to bound is

$$\frac{|\langle M\alpha, \alpha\rangle|}{\langle \alpha, \alpha\rangle} = |\cos 2\theta| \cdot \frac{\|\alpha^{\|} + \tilde{B}\alpha^{\perp}\|^2}{\|\alpha^{\|} + \alpha^{\perp}\|^2} = |\cos 2\theta| \cdot \frac{\cos^2\phi}{\cos^2\phi'},$$

where θ is the angle between $\alpha^{\|} + \tilde{B}\alpha^{\perp}$ and S, $\phi \in [0, \pi/2]$ is the angle between $\alpha^{\|}$ and $\alpha^{\|} + \alpha^{\perp}$, and $\phi' \in [0, \pi/2]$ is the angle between $\alpha^{\|}$ and $\alpha^{\|} + \tilde{B}\alpha^{\perp}$. If we also let ψ be the angle between $\alpha^{\|}$ and S, then we clearly have $\theta \in [\psi - \phi', \psi + \phi']$.

Now we translate Claims 4.1 and 4.2 into this geometric language. Claim 4.1 constrains the relationship between ϕ' and ϕ by

$$\frac{\tan\phi'}{\tan\phi} = \frac{\|\tilde{B}\alpha^{\perp}\|}{\|\alpha^{\perp}\|} \le \lambda_2.$$

Claim 4.2 says $|\cos 2\psi| \le \lambda_1$. For notational convenience, we will denote the exact values of $(\tan\phi')/(\tan\phi)$ and $|\cos 2\psi|$ by μ_2 and μ_1, respectively. We will work with these values until the end of the proof, at which point we will upper bound them by λ_2 and λ_1.

To summarize, we want to maximize

$$|\cos 2\theta| \cdot \frac{\cos^2\phi}{\cos^2\phi'}. \tag{5}$$

over the variables θ, ϕ, ϕ', and ψ, subject to the following constraints:

1. $\phi, \phi', \psi \in [0, \pi/2]$.
2. $\theta \in [\psi - \phi', \psi + \phi']$.[2]
3. $\tan\phi'/\tan\phi = \mu_2$.
4. $|\cos 2\psi| = \mu_1$.

There are two cases, depending on whether $|\cos 2x|$ ever achieves the value 1 in the interval $[\psi - \phi, \psi + \phi']$.

Case I: $\phi' \le \min\{\psi, \pi/2 - \psi\}$. Then

$$\begin{aligned} |\cos 2\theta| &= \max\{|\cos 2(\psi + \phi')|, |\cos 2(\psi - \phi')|\} \\ &= |\cos 2\psi \cdot \cos 2\phi'| + |\sin 2\psi \cdot \sin 2\phi'|. \end{aligned}$$

After some trigonometric manipulations, we have

$$|\cos 2\theta| \cdot \frac{\cos^2\phi}{\cos^2\phi'} = \frac{1}{2}\left|(1-\mu_2^2)\cos 2\psi + (1+\mu_2^2)\cos 2\psi \cos 2\phi\right| + \frac{1}{2}|2\mu_2 \sin 2\psi \sin 2\phi|$$

[2] We do not require $\theta \in [0, \pi/2]$ so that we do not have to worry about "wraparound" in the interval $[\psi - \phi', \psi + \phi']$. Adding a multiple of $\pi/2$ to θ does not change the value of (5).

The choice of ϕ which maximizes this is to have $(\cos 2\phi, \sin 2\phi)$ be a unit vector in the direction of $(\pm(1+\mu_2^2)\cos 2\psi, 2\mu_2 \sin 2\psi)$, so

$$\begin{aligned} |\cos 2\theta| \cdot \frac{\cos^2\phi}{\cos^2\phi'} &\leq \frac{1}{2}(1-\mu_2^2)|\cos 2\psi| + \frac{1}{2}\sqrt{(1+\mu_2^2)^2\cos^2 2\psi + 4\mu_2^2\sin^2 2\psi} \\ &= \frac{1}{2}(1-\mu_2^2)\mu_1 + \frac{1}{2}\sqrt{(1+\mu_2^2)^2\mu_1^2 + 4\mu_2^2(1-\mu_1^2)}. \end{aligned}$$

Case II: $\phi' > \min\{\psi, \pi/2-\psi\}$. In this case, we cannot obtain any nontrivial bound on $|\cos 2\theta|$, so, after some trigonometric manipulations, the problem is reduced to bounding:

$$|\cos 2\theta| \cdot \frac{\cos^2\phi}{\cos^2\phi'} \leq \frac{\cos^2\phi}{\cos^2\phi'} = \mu_2^2 + (1-\mu_2^2)\cos^2\phi. \tag{6}$$

The condition $\phi' > \min\{\psi, \pi/2-\psi\}$ implies that $\cos 2\phi' < |\cos 2\psi| = \mu_1$. After some trigonometric manipulations, we have

$$\cos 2\phi' = \frac{(1+\mu_2^2)\cos^2\phi - \mu_2^2}{(1-\mu_2^2)\cos^2\phi + \mu_2^2},$$

and the condition $\cos 2\phi' < \mu_1$ is equivalent to

$$\cos^2\phi < \frac{\mu_2^2(1+\mu_1)}{(1-\mu_1)+\mu_2^2(1+\mu_1)}.$$

Substituting this into (6) and simplifying, we conclude that

$$|\cos 2\theta| \cdot \frac{\cos^2\phi}{\cos^2\phi'} < \frac{2\mu_2^2}{1-\mu_1+\mu_2^2(1+\mu_1)}.$$

It can be verified that the bound obtained in Case I is an increasing function of μ_1 and μ_2 and is always greater than or equal to the bound in Case II. Therefore, replacing μ_1 and μ_2 by λ_1 and λ_2 in the Case I bound proves the theorem.

■

5 The Base Graph

Our construction of an infinite family of expanders in Section 3.2 requires starting with a (D^8, D, λ)-graph H (for a sufficiently small λ, say $\leq 1/5$). Since D is a "constant," such a graph can be found by exhaustive search (given that one exists, which can be proven by (nontrivial) probabilistic arguments [Alo86a, BS87, FKS89, Fri91]). However, for these parameters, there are simple explicit constructions known. We describe two of them below. The first is simpler and more intuitive, but the second yields better parameters.

5.1 The Affine Plane

The first construction is based on the "projective plane" construction of Alon [Alo86b], but we instead use the affine plane in order to make N exactly D^2 and then use the zig-zag product to obtain a graph with $N = D^8$. For a prime power $q = p^t$, let $\mathbb{F}_q$ be the finite field of size q; an explicit representation of such a field can be found deterministically in time $\mathrm{poly}(p, t)$ [Sho90]. We define a graph AP_q with vertex set $\mathbb{F}_q^2$, and edge set $\{((a, b), (c, d)) : ac = b + d\}$. That is, we connect the vertex (a, b) to all points on the line $L_{a,b} = \{(x, y) : y = ax - b\}$. (Note that we have chosen the sign of b to make the graph undirected.)

Lemma 5.1 AP_q *is an* $(q^2, q, 1/\sqrt{q})$*-graph. Moreover, a rotation map for* AP_q *can be computed in time* $\mathrm{poly}(\log q)$ *given a representation of the field* $\mathbb{F}_q$.

Proof: The expansion of AP_q will follow from the fact the square of AP_q is almost the complete graph, which in turn is based on the fact that almost all pairs of lines in the plane $\mathbb{F}_q^2$ intersect. Let M be the $q^2 \times q^2$ normalized adjacency matrix of AP_q; we will now calculate the entries of M^2. The entry of M^2 in row (a, b) and column (a', b') is exactly the number of common neighbors of (a, b) and (a', b') in AP_q divided by q^2, *i.e.,* $|L_{a,b} \cap L_{a',b'}|/q^2$. If $a \neq a'$, then $L_{a,b}$ and $L_{a',b'}$ intersect in exactly one point. If $a = a'$ and $b \neq b'$, then their intersection is empty, and if $a = a'$ and $b = b'$, then their intersection is of size q. Thus, if we let I_q denote the $q \times q$ identity matrix and J_q the $q \times q$ all-one's matrix, we have

$$M^2 = \frac{1}{q^2}\begin{pmatrix} qI_q & J_q & \cdots & J_q \\ J_q & qI_q & & J_q \\ \vdots & & \ddots & J_q \\ J_q & J_q & \cdots & qI_q \end{pmatrix} = \frac{I_q \otimes qI_q + (J_q - I_q) \otimes J_q}{q^2}.$$

Now we can calculate the eigenvalues explicitly. J_q has eigenvalues q (multiplicity 1) and 0 (multiplicity $q-1$). So $(J_q - I_q) \otimes J_q$ has eigenvalues $(q-1)\cdot q$, $-1 \cdot q$, and 0. Adding $I_q \otimes qI_q$ increases all these eigenvalues by q, and then we divide by q^2. Hence the eigenvalues of M^2 are 1 (multiplicity 1), 0 (multiplicity $q-1$), and $1/q$ (multiplicity $(q-1)\cdot q$). Therefore, the second largest eigenvalue of M has absolute value $1/\sqrt{q}$.

A rotation map for AP_q is given by

$$\mathrm{Rot}_q((a, b), t) = \begin{cases} ((t/a, t-b), t) & \text{if } a \neq 0 \text{ and } t \neq 0, \\ ((t, -b), a) & \text{if } a = 0 \text{ or } t = 0, \end{cases}$$

where $a, b, t \in \mathbb{F}_q$. ■

Now, define the following graphs inductively:

$$\begin{aligned} \mathrm{AP}_q^1 &= \mathrm{AP}_q \otimes \mathrm{AP}_q \\ \mathrm{AP}_q^{i+1} &= \mathrm{AP}_q^i \text{ⓩ} \mathrm{AP}_q \end{aligned}$$

From Proposition 2.4 and Theorem 3.2, we immediately deduce:

Proposition 5.2 AP_q^i *is a* $(q^{2(i+1)}, q^2, O(i/\sqrt{q}))$*-graph.*[3] *Moreover, a rotation map for* AP_q^i *can be computed in time* $\mathrm{poly}(i, \log q)$ *given a representation of* $\mathbb{F}_q$.

Taking $i = 7$ and a sufficiently large q gives a graph suitable for the expander construction in Section 3.2.

5.2 Low-Degree Polynomials

The graphs we describe here are derived from constructions of Alon and Roichman [AR94], which are Cayley graphs derived from the generator matrix of an error-correcting code. In order to give a self-contained presentation, we specialize the construction to a Reed-Solomon code concatenated with a Hadamard code (as used in, e.g. [AGHP92]).

For a prime power q and $d \in \mathbb{N}$, we define a graph $\mathrm{LD}_{q,d}$ on vertex set $\mathbb{F}_q^{d+1}$ with degree q^2. For a vertex $a \in \mathbb{F}_q^{d+1}$ and $x, y \in \mathbb{F}_q$, the the (x, y)'th neighbor of a is $a + (y, yx, yx^2, \ldots, yx^d)$.

[3] The hidden constant in $O(i/\sqrt{q})$ can be reduced to 1 using the improved analysis of the zig-zag product in Theorem 4.3.

Proposition 5.3 $\mathrm{LD}_{q,d}$ *is a* $(q^{d+1}, q^2, d/q)$*-graph. Moreover, a rotation map for* $\mathrm{LD}_{q,d}$ *can be computed in time* $\mathrm{poly}(\log q, d)$ *given a representation of* $\mathbb{F}_q$.

As above, taking $d = 7$ and sufficiently large q gives a graph suitable for our expander construction. These graphs are better than those of Proposition 5.2 because the the eigenvalue-degree relationship is the optimal $\lambda = O(1/\sqrt{D})$ (as q grows), which implies an eigenvalue of $O(1/D^{1/4})$ for the family constructed in Theorem 3.3.

Proof: To simplify notation, let $\mathbb{F} = \mathbb{F}_q$. Let M be the $q^{d+1} \times q^{d+1}$ normalized adjacency matrix of $\mathrm{LD}_{q,d}$. We view vectors in $\mathbb{C}^{q^{d+1}}$ as functions $f : \mathbb{F}^{d+1} \to \mathbb{C}$. We will now explicitly describe the eigenvectors of M. Let p be the characteristic of $\mathbb{F}$, let $\zeta = e^{2\pi i/p}$ be a primitive p'th root of unity, and let $L : \mathbb{F} \to \mathbb{F}_p$ be any surjective $\mathbb{F}_p$-linear map. (For simplicity, one can think of the special case that $p = q$ and L is the identity map.)

For every sequence $a = (a_0, \ldots, a_d) \in \mathbb{F}^{d+1}$, define the function $\chi_a : \mathbb{F}^{d+1} \to \mathbb{C}$ by $\chi_a(b) = \zeta^{L(\sum a_i b_i)}$. Clearly, $\chi_a(b + c) = \chi_a(b)\chi_a(c)$ for any $b, c \in \mathbb{F}^{d+1}$. Moreover, it can be verified that the $\{\chi_a\}$ are orthogonal under the standard inner product $\langle f, g\rangle = \sum_b f(b)g(b)^*$, and thus form a basis for $\mathbb{C}^{q^{d+1}}$. Hence, if we show that each χ_a is an eigenvector of M, then they are all the eigenvectors of M. This can be done by direct calculation:

$$
\begin{aligned}
(M\chi_a)(b) &= \frac{1}{q^2} \sum_{c \in \mathbb{F}^{d+1}} M_{bc} \cdot \chi_a(c) \\
&= \frac{1}{q^2} \sum_{x,y \in \mathbb{F}} \chi_a(b + (y, yx, \ldots, yx^d)) \\
&= \left(\frac{\sum_{x,y\in\mathbb{F}} \chi_a(y, yx, \ldots, yx^d)}{q^2} \right) \cdot \chi_a(b) \\
&\stackrel{\mathrm{def}}{=} \lambda_a \cdot \chi_a(b).
\end{aligned}
$$

Thus, χ_a is an eigenvector of M with eigenvalue λ_a and all eigenvectors of M are of this form. So we simply need to show that $|\lambda_a| \le d/q$ for all but one $a \in \mathbb{F}^{d+1}$. To do this, note that

$$
\lambda_a = \frac{1}{q^2} \sum_{x,y \in \mathbb{F}} \chi_a((y, yx, \ldots, yx^d)) = \frac{1}{q^2} \sum_{x,y\in\mathbb{F}} \zeta^{L(y \cdot p_a(x))},
$$

where $p_a(x)$ is the polynomial $a_0 + a_1 x + \cdots + a_d x^d$. When x is a root of p_a, then $\zeta^{L(yp_a(x))} = 1$ for all y, and hence x contributes $q/q^2 = 1/q$ to λ_a. When x is not a root of $p_a(x)$, $yp_a(x)$ takes on all values in $\mathbb{F}$ as y varies, and hence $\zeta^{L(yp_a(x))}$ varies uniformly over all p'th roots of unity. Since the sum of all p'th roots of unity is 0, these x's contribute nothing to λ_a. When $a \neq 0$, p_a has at most d roots, so $|\lambda_a| \le d/q$. ■

6 Variants on the Zig-Zag Theme

The two subsections of this section contain two variants of the basic zig-zag product. The first is aimed at improving the relation between the degree and the eigenvalue bound. The second is aimed at simplifying the product, at the cost of deteriorating this relationship.

6.1 A "Derandomized" Zig-Zag Product

In this section we provide a variant of our original zig-zag product, which achieves a better relationship between the degree and the expansion of the resulting graph. The term "derandomized" will become clearer when we define it.

Recall that the optimal second-largest eigenvalue for an infinite family of D-regular graphs is $\Theta(1/D^{1/2})$, and families of graphs meeting this bound (with the right constant) are referred to as Ramanujan. A basic question is how close can we come to this optimal bound using our techniques. Starting with a constant-size Ramanujan graph (or the graphs of Section 5.2), our basic construction of Theorem 3.3 achieves a second-largest eigenvalue of $O(1/D^{1/4})$ for the family of expanders generated..

Here, we define a variant of the zig-zag product, which makes more efficient use of the expansion of the small graph. Using the new product in our iterative construction (of Section 3.2) with an initial constant-size Ramanujan graph or even the graphs of Proposition 5.3, we obtain a second-largest eigenvalue of $O(1/D^{1/3})$ for the family of expanders generated. It is an interesting open problem to construct families of graphs achieving the optimal eigenvalue $O(1/D^{1/2})$ using a similar graph product.

We now turn to the formal definition of the new zig-zag product. It will have two "zig" moves and two "zag" moves, but they will not be independent. The second "zig" and the first "zag" will use the same random bits!

Definition 6.1 *Let G_1 be a D_1-regular graph on $[N_1]$ with rotation map Rot_{G_1} and let G_2 be a D_2-regular graph on $[D_1]$ with rotation map Rot_{G_2}. Suppose that for every $i \in [D_2]$, $\mathrm{Rot}_{G_2}(\cdot, i)$ induces a permutation on $[D_1]$.*[4] *Then the* **modified zig-zag product** *of G_1 and G_2 is defined to be the D_2^3-regular graph $G_1 Ⓩ' G_2$ on $[N_1] \times [D_1]$ whose rotation map $\mathrm{Rot}_{G_1 Ⓩ' G_2}$ is as follows:*

$\mathrm{Rot}_{G_1 Ⓩ' G_2}((v,k),(h,i,j))$**:**

1. *Let $(k', h') = \mathrm{Rot}_{G_2}(k, h)$.*
2. *Let $(k'', i') = \mathrm{Rot}_{G_2}(k', i)$.*
3. *Let $(w, \ell'') = \mathrm{Rot}_{G_1}(v, k'')$.*
4. *Find the unique $\ell' \in [D_1]$ such that $(\ell'', i'') = \mathrm{Rot}_{G_2}(\ell', i)$ for some i''. (ℓ' exists by the assumption on Rot_{G_2}.)*
5. *Let $(\ell, j') = \mathrm{Rot}_{G_2}(\ell', j)$.*
6. *Output $((w, \ell), (j', i, h'))$.*

Again, in this graph product we do *two* random steps on the small graph in both the zig and the zag parts. However, to save random bits (*i.e.,* decrease the degree) we use *the same* random bits for the second move of the zig part and the first move of the zag part. Thus the degree of the new graph is D_2^3. However, we will show that the bound on the eigenvalue will be as if these moves were independent. This proof will follow the lines of the basic analysis of the original zig-zag product.

Theorem 6.2 *If G_1 is an (N_1, D_1, λ_1)-graph and G_2 is a (D_1, D_2, λ_2)-graph, then $G_1 Ⓩ' G_2$ is a $(N_1 \cdot D_1, D_2^3, \lambda_1 + 2\lambda_2^2)$-graph. Moreover, $\mathrm{Rot}_{G_1 Ⓩ' G_2}$ can be computed in time $\mathrm{poly}(\log N, \log D_1, D_2)$ with one oracle query to Rot_{G_1} and $D_2 + 2$ oracle queries to Rot_{G_2}.*

[4] By this we mean that the function $f_i(x) =$ "the first component of $\mathrm{Rot}_{G_2}(x, i)$" $=$ "the i'th neighbor of x" is a permutation for every i.

Proof: We use the same notation as in the proof of Theorem 3.2. Like there, we need to bound $|\langle M\alpha, \alpha\rangle|/\langle\alpha, \alpha\rangle$, where M is the normalized adjacency matrix of $G_1 \textcircled{z}' G_2$ and $\alpha \perp 1_{N_1 D_1}$. Let B_i be the $D_1 \times D_1$ permutation matrix induced by $\mathrm{Rot}_{G_2}(\cdot, i)$, and let $\tilde{B}_i = I_{N_1} \otimes B_i$. Then

$$\tilde{B} = \frac{1}{D_1} \sum_{i=1}^{D_1} \tilde{B}_i.$$

Note that the normalized adjacency matrix corresponding to Steps 2–4 in the definition of $G_1 \textcircled{z}' G_2$ is given by

$$M' = \frac{1}{D_1} \sum_i \tilde{B}_i \tilde{A} \tilde{B}_i^T,$$

where $\tilde{B}_i^T$ is the transpose (equivalently, inverse) of $\tilde{B}_i$. Thus, $M = \tilde{B} M' \tilde{B}$. The main observation is that not only does $\tilde{B}\alpha^{||} = \alpha^{||}$ (as we used in the original analysis), but also $\tilde{B}_i^T \alpha^{||} = \alpha^{||}$ for every i (because B_i is a permutation matrix). Hence,

$$M'\alpha^{||} = \frac{1}{D_1} \sum_i \tilde{B}_i \tilde{A} \tilde{B}_i^T \alpha^{||} = \frac{1}{D_1} \sum_i \tilde{B}_i \tilde{A} \alpha^{||} = \tilde{B} \tilde{A} \alpha^{||}.$$

Applying this (and the symmetry of $\tilde{B}$ and M'), we get

$$\begin{aligned} \langle M\alpha, \alpha\rangle &= \langle M\alpha^{||}, \alpha^{||}\rangle + 2\langle M\alpha^{||}, \alpha^{\perp}\rangle + \langle M\alpha^{\perp}, \alpha^{\perp}\rangle \\ &= \langle \tilde{A}\alpha^{||}, \alpha^{||}\rangle + 2\langle \tilde{A}\alpha^{||}, \tilde{B}^2\alpha^{\perp}\rangle + \langle M'\tilde{B}\alpha^{\perp}, \tilde{B}\alpha^{\perp}\rangle. \end{aligned}$$

Being the normalized adjacency matrix of an undirected, regular graph, M' has no eigenvalues larger than 1 and hence does not increase the length of any vector. Using this together with Claims 4.1 and 4.2, we have

$$\begin{aligned} |\langle M\alpha, \alpha\rangle| &\leq |\langle \tilde{A}\alpha^{||}, \alpha^{||}\rangle| + 2\|\alpha^{||}\| \cdot \|\tilde{B}^2\alpha^{\perp}\| + \|\tilde{B}\alpha^{\perp}\|^2 \\ &\leq \lambda_1 \cdot \|\alpha^{||}\|^2 + 2\lambda_2^2 \cdot \|\alpha^{||}\| \cdot \|\alpha^{\perp}\| + \lambda_2^2 \cdot \|\alpha^{\perp}\|^2. \end{aligned}$$

As in the the proof of Theorem 3.2, using the fact that $\|\alpha^{||}\|^2 + \|\alpha^{\perp}\|^2 = \|\alpha\|^2$ yields the desired bound. ■

6.2 The Replacement Product

In this section, we describe an extremely simple and intuitive graph product, which shares similar properties to the zig-zag product. Namely, when taking the product of two expanders, we get a larger expander whose degree depends only on that of the smaller graph. Here simplicity is the important feature, and the expansion quality is not as good as above. This product is so natural that it was used in various contexts before. Indeed, Gromov [Gro83] even estimates the 2nd eigenvalue of an iterated replacement product of the graph of the Boolean hypercube with smaller copies of itself. (Of course, in this very special case the outcome is not expanding, since the cube is not.) Our proof of its expansion will be a simple reduction to the expansion properties of the zig-zag product. However, one can also prove it directly in a manner similar to the proof of Theorem 3.2 (and thereby obtain a stronger bound).

Assume (as in the basic zig-zag product) that G_1 is a D_1 regular graph on $[N_1]$ and G_2 is a D_2-regular graph on $[D_1]$. A natural idea is to place a "copy" (or "cloud") of G_2 around each vertex of G_1, maintaining the edges of both. More precisely, every vertex will be connected to all its original neighbors in its cloud, as well as to one vertex in the neighboring cloud it defines. For example, if G_1 is the n-dimensional Boolean cube graph, and G_2 is the cycle on n vertices, then the resulting graph is the so-called *cube connected cycle*, which used to be a popular architecture for parallel computers. Note that in this example the small graph had degree 2, and the product graph had degree 3. In general, the resulting graph would have degree $D_2 + 1$. In terms of rotation maps, this product is defined as follows.

Definition 6.3 *If G_1 is a D_1-regular graph on $[N_1]$ with rotation map Rot_{G_1} and G_2 is a D_2-regular graph on $[D_1]$ with rotation map Rot_{G_2}, then their* **replacement product** *$G_1 \circledR G_2$ is defined to be the (D_2+1)-regular graph on $[N_1] \times [D_1]$ whose rotation map $\mathrm{Rot}_{G_1 \circledR G_2}$ is as follows:*

$\mathrm{Rot}_{G_1 \circledR G_2}((v,k),i)$**:**

1. *If $i \le D_2$, let $(m,j) = \mathrm{Rot}_{G_2}(k,i)$ and output $((v,m),j)$.*
2. *If $i = D_2+1$, output $(\mathrm{Rot}_{G_1}(v,k),i)$.*

The expansion properties of the replacement product are given in the next theorem, relating it to those of the zig-zag product.

Theorem 6.4 *If G_1 is an (N_1, D_1, λ_1)-graph and G_2 is a (D_1, D_2, λ_2)-graph, then $G_1 \circledR G_2$ is a $(N_1 \cdot D_1, D_2+1, g(\lambda_1,\lambda_2,D_2))$-graph, where (using the function f from Thm. 3.2 or 4.3)*

$$g(\lambda_1,\lambda_2,D_2) \le (p + (1-p) f(\lambda_1,\lambda_2))^{1/3},$$

and $p = D_2^2/(D_2+1)^3$. In particular, $g(\lambda_1,\lambda_2,D_2) < 1$ when $\lambda_1,\lambda_2 < 1$. Moreover, $\mathrm{Rot}_{G_1 \circledR G_2}$ can be computed in time $\mathrm{poly}(\log N, \log D_1, \log D_2)$ with one oracle query to Rot_{G_1} or Rot_{G_2}.

Proof: The idea of the proof is that the graph of the zig-zag product is a regular subgraph of *the cube* of the graph of the replacement product. Let M denote the normalized adjacency matrix of $G_1 \circledR G_2$. As in the proof of Theorem 3.2, we let A, B respectively denote the normalized adjacency matrices of G_1, G_2, and define their "liftings" $\tilde{A}, \tilde{B}$ in the same way. By inspection, we have $M = (\tilde{A} + D_2\tilde{B})/(D_2+1)$. The key observation is that

$$M^3 = \frac{(\tilde{A} + D_2\tilde{B})^3}{(D_2+1)^3} = p\tilde{B}\tilde{A}\tilde{B} + (1-p)C,$$

where $\tilde{B}\tilde{A}\tilde{B}$ is the normalized adjacency matrix of $G_1 \circledz G_2$, C is the normalized adjacency matrix of an undirected, regular graph (and in particular does not increase the length of any vector), and $p = D_2^2/(D_2+1)^3$. As eigenvalues of powers of matrices are the respective powers of the original eigenvalues (see Proposition 2.3), we have

$$g(\lambda_1,\lambda_2) \le (p + (1-p) f(\lambda_1,\lambda_2))^{1/3}.$$

■

Thus, for "constant" degrees D_2 the replacement product indeed transforms two expanders into a larger one. As in Corollary 3.4, we can use this to get degree 3 expanders.

Corollary 6.5 *For every $\lambda < 1$ and every odd D, there exists a $\lambda' < 1$ such that if G is an (N, D, λ)-graph and C is the cycle on D vertices, then $G \circledR C$ is a $(ND, 3, \lambda')$-graph.*

To make the expansion properties in Theorem 6.4 independent of how large D_2 is, we now slightly modify the replacement product to have D_2 copies of each edge which goes between clouds. This makes the degree of every vertex $2D_2$, of which D_2 stay within the same cloud, and the other D_2 all connect to the same vertex in a neighbor cloud. This "balancing" make the random walk give the same weight to edges defined by G_1 and G_2.

Definition 6.6 *If G_1 is a D_1-regular graph on $[N_1]$ with rotation map Rot_{G_1} and G_2 is a D_2-regular graph on $[D_1]$ with rotation map Rot_{G_2}, then their* **balanced replacement product** *$G_1 \circledb G_2$ is defined to be the $2D_2$-regular graph on $[N_1] \times [D_1]$ whose rotation map $\mathrm{Rot}_{G_1 \circledb G_2}$ is as follows:*

$\mathrm{Rot}_{G_1 \textcircled{b} G_2}((v,k),i)$:

1. *If* $i \le D_2$, *let* $(m,j) = \mathrm{Rot}_{G_2}(k,i)$ *and output* $((v,m),j)$.
2. *If* $i > D_2$, *output* $(\mathrm{Rot}_{G_1}(v,k),i)$.

Theorem 6.7 *If* G_1 *is an* (N_1, D_1, λ_1)*-graph and* G_2 *is a* (D_1, D_2, λ_2)*-graph, then* $G_1 \textcircled{b} G_2$ *is a* $(N_1 \cdot D_1, 2D_2, h(\lambda_1, \lambda_2))$*-graph, where (using the function* f *from Thm. 3.2 or 4.3)*

$$h(\lambda_1, \lambda_2) \le \left(\frac{7}{8} + \frac{1}{8} \cdot f(\lambda_1, \lambda_2)\right)^{1/3}$$

In particular, $h(\lambda_1, \lambda_2) < 1$ *when* $\lambda_1, \lambda_2 < 1$. *Moreover,* $\mathrm{Rot}_{G_1 \textcircled{b} G_2}$ *can be computed in time* $\mathrm{poly}(\log N, \log D_1, \log D_2)$ *with one oracle query to* Rot_{G_1} *and one oracle query to* Rot_{G_2}.

Proof: The proof is the same as that of Theorem 6.4, noting instead that $M = (\tilde{A} + \tilde{B})/2$. ■

As a final note, we observe the weakness of the replacement products relative to the zig-zag product. Informally, in zig-zag the expansion quality of the product improves with those of its component, while in the replacement it does not. More formally, while the function $f(\lambda_1, \lambda_2)$ tends to zero when λ_1 and λ_2 do, the functions $g(\lambda_1, \lambda_2, D_2)$ and $h(\lambda_1, \lambda_2)$ do not.

Acknowledgments

We are grateful to David Zuckerman for illuminating discussions and a number of useful suggestions early in the stages of this work. We thank the organizers of the DIMACS Workshop on Pseudorandomness and Explicit Combinatorial Constructions in October 1999, where we began this research. We are grateful to Peter Winkler for suggesting the name "zig-zag product." We also thank Noga Alon, Oded Goldreich, Peter Sarnak, Ronen Shaltiel, Dan Spielman, and the anonymous referee for helpful comments and pointers.

References

[Ajt94] Miklós Ajtai. Recursive construction for 3-regular expanders. *Combinatorica*, 14(4):379–416, 1994.

[AKS83] Miklós Ajtai, János Komlós, and Endre Szemerédi. Sorting in $c \log n$ parallel steps. *Combinatorica*, 3(1):1–19, 1983.

[Alo86a] Noga Alon. Eigenvalues and expanders. *Combinatorica*, 6(2):83–96, 1986.

[Alo86b] Noga Alon. Eigenvalues, geometric expanders, sorting in rounds, and Ramsey theory. *Combinatorica*, 6(3):207–219, 1986.

[AGM87] Noga Alon, Zvi Galil, and Vitali D. Milman. Better expanders and superconcentrators. *Journal of Algorithms*, 8(3):337–347, 1987.

[AGHP92] Noga Alon, Oded Goldreich, Johan Håstad, and René Peralta. Simple constructions of almost k-wise independent random variables. *Random Structures & Algorithms*, 3(3):289–304, 1992.

[ALW01] Noga Alon, Alex Lubotzky, and Avi Wigderson. Semi-direct product in groups and zig-zag product in graphs: Connections and applications. In *42nd Annual Symposium on Foundations of Computer Science*, Las Vegas, Nevada, 14-17 October 2001. IEEE. To appear.

[AM85] Noga Alon and Vitali D. Milman. λ_1, isoperimetric inequalities for graphs, and superconcentrators. *Journal of Combinatorial Theory. Series B*, 38(1):73–88, 1985.

[AR94] Noga Alon and Yuval Roichman. Random Cayley graphs and expanders. *Random Structures & Algorithms*, 5(2):271–284, 1994.

[BS87] Andrei Broder and Eli Shamir. On the second eigenvalue of random regular graphs (preliminary version). In *28th Annual Symposium on Foundations of Computer Science*, pages 286–294, Los Angeles, California, 12–14 October 1987. IEEE.

[Fri91] Joel Friedman. On the second eigenvalue and random walks in random d-regular graphs. *Combinatorica*, 11(4):331–362, 1991.

[FKS89] Joel Friedman, Jeff Kahn, and Endre Szemerédi. On the second eigenvalue in random regular graphs. In *Proceedings of the Twenty First Annual ACM Symposium on Theory of Computing*, pages 587–598, Seattle, Washington, 15–17 May 1989.

[GG81] Ofer Gabber and Zvi Galil. Explicit constructions of linear-sized superconcentrators. *Journal of Computer and System Sciences*, 22(3):407–420, June 1981.

[GIL+90] Oded Goldreich, Russell Impagliazzo, Leonid Levin, Ramarathnam Venkatesan, and David Zuckerman. Security preserving amplification of hardness. In *31st Annual Symposium on Foundations of Computer Science*, volume I, pages 318–326, St. Louis, Missouri, 22–24 October 1990. IEEE.

[Gro83] Mikhael Gromov. Filling Riemannian manifolds. *Journal of Differential Geometry*, 18(1):1–147, 1983.

[Gro00] Misha Gromov. Spaces and questions. *Geometric and Functional Analysis*, pages 118–161, 2000. Part I of Special Volume on GAFA 2000 (Tel Aviv, 1999).

[INW94] Russell Impagliazzo, Noam Nisan, and Avi Wigderson. Pseudorandomness for network algorithms. In *Proceedings of the Twenty-Sixth Annual ACM Symposium on the Theory of Computing*, pages 356–364, Montréal, Québec, Canada, 23–25 May 1994.

[IW97] Russell Impagliazzo and Avi Wigderson. $P = BPP$ if E requires exponential circuits: Derandomizing the XOR lemma. In *Proceedings of the Twenty-Ninth Annual ACM Symposium on Theory of Computing*, pages 220–229, El Paso, Texas, 4–6 May 1997.

[JM87] Shuji Jimbo and Akira Maruoka. Expanders obtained from affine transformations. *Combinatorica*, 7(4):343–355, 1987.

[KR83] Nigel J. Kalton and James W. Roberts. Uniformly exhaustive submeasures and nearly additive set functions. *Transactions of the American Mathematical Society*, 278(2):803–816, 1983.

[LPS88] Alex Lubotzky, Ralph Phillips, and Peter Sarnak. Ramanujan graphs. *Combinatorica*, 8(3):261–277, 1988.

[Lub94] Alexander Lubotzky. *Discrete groups, expanding graphs and invariant measures*. Birkhäuser Verlag, Basel, 1994.

[LP01] Alexander Lubotzky and Igor Pak. The product replacement algorithm and Kazhdan's property (T). *Journal of the American Mathematical Society*, 14(2):347–363 (electronic), 2001.

[Mar73] Gregory A. Margulis. Explicit constructions of expanders. *Problemy Peredači Informacii*, 9(4):71–80, 1973.

[Mar88] Gregory 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.

[MW01] Roy Meshulam and Avi Wigderson, 2001. In preparation.

[Mor94] Moshe 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.

[NN93] Joseph Naor and Moni Naor. Small-bias probability spaces: Efficient constructions and applications. *SIAM Journal on Computing*, 22(4):838–856, August 1993.

[Nil91] A. Nilli. On the second eigenvalue of a graph. *Discrete Mathematics*, 91(2):207–210, 1991.

[Nis96] Noam Nisan. Extracting randomness: How and why: A survey. In *Proceedings, Eleventh Annual IEEE Conference on Computational Complexity*, pages 44–58, Philadelphia, Pennsylvania, 24–27 May 1996. IEEE Computer Society Press.

[NT99] Noam Nisan and Amnon Ta-Shma. Extracting randomness: A survey and new constructions. *Journal of Computer and System Sciences*, 58(1):148–173, 1999.

[NZ96] Noam Nisan and David Zuckerman. Randomness is linear in space. *Journal of Computer and System Sciences*, 52(1):43–52, February 1996.

[Pin73] Mark S. Pinsker. On the complexity of a concentrator. In *7th Annual Teletraffic Conference*, pages 318/1–318/4, Stockholm, 1973.

[Pip87] Nicholas Pippenger. Sorting and selecting in rounds. *SIAM Journal on Computing*, 16(6):1032–1038, December 1987.

[PY82] Nicholas Pippenger and Andrew C. Yao. Rearrangeable networks with limited depth. *SIAM Journal on Algebraic and Discrete Methods*, 3:411–417, 1982.

[RR99] Ran Raz and Omer Reingold. On recycling the randomness of the states in space bounded computation. In *Proceedings of the Thirty-First Annual ACM Symposium on the Theory of Computing*, Atlanta, GA, May 1999.

[RSW00] Omer Reingold, Ronen Shaltiel, and Avi Wigderson. Extracting randomness via repeated condensing. In *41st Annual Symposium on Foundations of Computer Science*, Redondo Beach, California, 12–14 November 2000. IEEE.

[RVW00] Omer Reingold, Salil Vadhan, and Avi Wigderson. Entropy waves, the zig-zag graph product, and new constant-degree expanders and extractors (extended abstract). In *41st Annual Symposium on Foundations of Computer Science*, Redondo Beach, California, 12–14 November

2000. IEEE. See a more complete version in ECCC TR01-018, `http://www.eccc.uni-trier.de/eccc`.

[Sho90] Victor Shoup. New algorithms for finding irreducible polynomials over finite fields. *Mathematics of Computation*, 54(189):435–447, 1990.

[Sip88] Michael Sipser. Expanders, randomness, or time versus space. *Journal of Computer and System Sciences*, 36(3):379–383, June 1988.

[SS96] Michael Sipser and Daniel A. Spielman. Expander codes. *IEEE Transactions on Information Theory*, 42(6, part 1):1710–1722, 1996.

[Spi96] Daniel A. Spielman. Linear-time encodable and decodable error-correcting codes. *IEEE Transactions on Information Theory*, 42(6, part 1):1723–1731, 1996.

[TUZ01] Amnon Ta-Shma, Christopher Umans, and David Zuckerman. Loss-less condensers, unbalanced expanders, and extractors. In *Proceedings of the Thirty-third Annual ACM Symposium on Theory of Computing*, pages 143–162, Crete, Greece, 6–8 July 2001.

[Tan84] Michael R. Tanner. Explicit concentrators from generalized n-gons. *SIAM Journal on Algebraic Discrete Methods*, 5(3):287–293, 1984.

[Urq87] Alasdair Urquhart. Hard examples for resolution. *Journal of the Association for Computing Machinery*, 34(1):209–219, 1987.

[Val77] Leslie G. Valiant. Graph-theoretic arguments in low-level complexity. In *Mathematical foundations of computer science (Proc. Sixth Sympos., Tatranská Lomnica, 1977)*, pages 162–176. Lecture Notes in Comput. Sci., Vol. 53. Springer, Berlin, 1977.