CS369E: Expanders May 9 & 16, 2005

Lecture 7: Expander Constructions (Zig-Zag Expanders)

Lecturer: Cynthia Dwork *Scribe: Geir Helleloid*

7.1 Lecture Outline

In this lecture we will see three explicit constructions of expanders. By an "explicit construction", we mean a construction with the following three properties:

1. We can build the entire N-vertex graph in poly(N) time.
2. From a vertex v, we can find the i-th neighbor in poly($\log N, \log D$) time where D is the degree of the graph.
3. Given vertices u and v, we can determine if they are adjacent in poly($\log N$) time.

The first two constructions will be presented without proof, but we will see the proof in the case of the zig-zag construction.

1. The first construction is due to Margulis and Gaber-Galil.
2. The second construction is due to Lubotsky, Phillips, and Sarnak, and achieves optimal spectral expansion $\lambda \approx 2/\sqrt{d}$.
3. The third construction is due to Reingold, Vadhan, and Wigderson. These so-called zig-zag expanders are built via repeated applications of two basic operations that jointly increase the number of nodes but keep the degree and expansion λ small. These operations are graph squaring and the zig-zag product. The proof that these graphs are expanders will use the tensor product of two vectors.

7.2 The First Two Constructions

Construction 7.1 (Margulis [Mar]) Fix a positive integer M and let $[M] = \{1, 2, \ldots, M\}$. Define the bipartite graph $G = (V, E)$ as follows. Let $V = [M]^2 \cup [M]^2$, where vertices in the first partite set are denoted $(x, y)_1$ and vertices in the second partite set are denoted $(x, y)_2$. From each vertex $(x, y)_1$, put in edges to $(x, y)_2$, $(x, x + y)_2$, $(x, x + y + 1)_2$, $(x + y, y)_2$, and $(x + y + 1, y)_2$, where all arithmetic is done modulo M. Then G is an expander. The proof uses Fourier analysis.

Construction 7.2 (Lubotsky-Phillips-Sarnak [LPS]) Fix primes q and p such that $q \equiv 1 \pmod 4$ and $p \equiv 1 \pmod q$. Let i be an integer such that $i^2 \equiv -1 \pmod q$. Define the graph $G = (V, E)$ as follows. Let $V = GF(q) \cup \{\infty\}$. Put an edge between (z, z') if

$$z' = \frac{(a_0 + ia_1)z + (a_2 + ia_3)}{(-a_2 + ia_3)z + (a_0 - ia_1)}$$

for some $a_0, a_1, a_2, a_3 \in \mathbb{N}$ such that $a_0^2 + a_1^2 + a_2^2 + a_3^2 = p$. It can be shown that the number of integral solutions to $a_0^2 + a_1^2 + a_2^2 + a_3^2 = p$ is $p + 1$. Hence, G has degree $d = p + 1$. It can further be shown that the spectral expansion of G is at most $\lambda(G) \leq 2\sqrt{d-1}/d$, which is optimal. Families of graphs with such optimal spectral expansion are called Ramanujan graphs.

7.3 The Zig-Zag Product

In this section, we define the zig-zag product of two graphs. We will use this product in Sections 7.4 and 7.5 to construct expander graphs and prove their spectral properties. This construction is due to Reingold, Vadhan and Wigderson [RVW]. For convenience, we say that G is an (N, d, λ)-expander if G has N vertices, degree d, and spectral expansion λ.

To construct the zig-zag product, we begin with an (N_1, d_1, λ_1)-expander G and a (d_1, d_2, λ_2)-expander H. Assume that $V(H) = [d_1] = \{1, 2, \ldots, d_1\}$. Each vertex in G has d_1 neighbors, and we can label them as the 1st, 2nd, …, and d_1-th neighbors of v. Define a matching Rot_G on $V(G) \times V(H)$ by $\mathrm{Rot}_G(u, i) = (v, j)$ where v is the i-th neighbor of u and u is the i-th neighbor of v. This is the *rotation map* associated to G.

One intuitive way to approach the zig-zag product is to suppose that we want to construct a random walk on $V(G) \times V(H)$. Starting at (u, i), what can we do? We can choose a random neighbor i' of i in H, and use that to pick a random neighbor v of u in G. Since this isn't quite reversible, we need to end by choosing another random neighbor of our current vertex in H. This attempts to motivate the following definition.

The *zig-zag product* of G and H is denoted $G Ⓩ H$. The vertex set of $G Ⓩ H$ is $V(G) \times V(H)$, so the vertices of $G Ⓩ H$ are pairs (v, i) with $v \in V(G)$ and $i \in V(H)$. Put an edge between (u, i) and (v, j) if and only if there exist $i', j' \in V(H)$ such that (i, i') and (j, j') are edges of H and $\mathrm{Rot}_G(u, i') = (v, j')$. (See Figure 1)

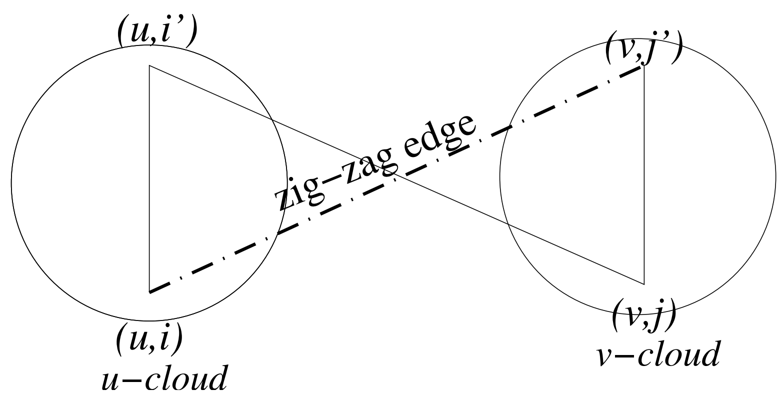

Figure 1: Zig-Zag Product

More formally,

Definition 7.1. *The zig-zag product between rotation map representations of two graphs G, a (N, D_1, λ_1)-graph and H, a (D_1, D_2, λ_2)-graph, is a rotation map representation of a graph, denoted by $G Ⓩ H$. The graph $G Ⓩ H$ and its rotation map are defined as below.*

1. *$G ⓩ H$ has ND_1 vertices.*

2. *$G ⓩ H$ is a D_2^2-regular graph.*

3. $\mathrm{Rot}_{G ⓩ H}((u, i), (a_1, a_2)) = ((v, j), (b_1, b_2))$ *if the following is satisfied: There exist $i', j' \in [D_2]$ such that*

   - $\mathrm{Rot}_H(i, a_1) = (i', b_2)$
   - $\mathrm{Rot}_G(u, i') = (v, j')$
   - $\mathrm{Rot}_H(j', a_2) = (j, b_1)$

Less formally, let's explore what this really looks like. First, form an intermediate graph K by replacing each vertex v of G by a copy H_v of H. For each neighbor w of v in G, choose a vertex in H_w. Then construct a matching between these d_1 vertices and the d_1 vertices of H_v. Of course, the matchings constructed for the vertices in G must be compatible in the obvious way. We refer to H_v as the *cloud* corresponding to v.

Now, if there is an edge between (u, i') and (v, j') in K (with $u \neq v$), then in $G ⓩ H$, all the neighbors of (u, i') in H_u are connected to all the neighbors of (v, j') in H_v. So $G ⓩ H$ is the edge union of many complete bipartite graphs K_{d_2,d_2}. We can easily calculate the degree of $G ⓩ H$: from (u, i), there are d_2 choices for (u, i'), then one choice for $(v, j') = \mathrm{Rot}_G(u, i')$, and finally d_2 choices for (v, j). Thus the degree of $G ⓩ H$ is d^2.

Intuitively, why should $G ⓩ H$ have good expansion when G and H do? If we are given a distribution that is mixed on the G component, then the rapid mixing on H suggests that the distribution will rapidly mix on $G ⓩ H$. Similarly, given a distribution that is mixed on the H component, then the rapid mixing on G suggests that the distribution will rapidly mix on $G ⓩ H$. So we might hope that every distribution mixes rapidly on $G ⓩ H$.

7.4 The Zig-Zag Expander Construction

Using both the zig-zag product and graph squaring, we can demonstrate the zig-zag construction of expanders. Recall that the square of G is denoted G^2; it has the same vertex set as G, and $(x, y) \in E(G^2)$ if and only if there exists a path of length two from x to y in G. If G is an (N, d, λ)-expander, then G^2 is an (N, d^2, λ^2)-expander. In the next section we will show that $\lambda(G ⓩ H) \leq \lambda(G) + \lambda(H) + \lambda(H^2)$. Assuming this result, we can state and prove the zig-zag construction of expanders.

Theorem 7.2. *Let H be a (d^4, d, λ_0)-expander for some $\lambda_0 \leq 1/5$[1]. Define $G_1 = H^2$ and $G_{t+1} = G_t^2 ⓩ H$ for $t \geq 1$. Then for all t, G_t is a (d^{4t}, d^2, λ)-expander with $\lambda \leq 2/5$.*

Proof. The proof is by induction on t. When $t = 1$, based on what we know about the square of a graph, we see that G_1 is a (d^4, d^2, λ_0^2)-expander where $\lambda_0^2 \leq 1/25$.

Now assume that G_{t-1} is a $(d^{4(t-1)}, d^2, \lambda)$-expander with $\lambda \leq 2/5$. It is clear that G_t has d^{4t} nodes since the number of nodes in the zig-zag product of two graphs is the product of the number of nodes in each of the two graphs. Also G_t has degree d^2, since the degree of a zig-zag product is the degree of the second factor.

[1]Since H is a fixed-size graph, such an expander graph can be found by brute-force search.

Finally,

$$\begin{aligned}\lambda(G_t) &\leq \lambda(G_{t-1}^2)+\lambda(H)+\lambda(H^2)\\ &\leq \left(\frac{2}{5}\right)^2+\frac{1}{5}+\frac{1}{25}\\ &= \frac{2}{5}.\end{aligned}$$

□

7.5 Spectral Property of the Zig-Zag Product

It remains to give an upper bound for the spectral expansion of a zig-zag product.

Theorem 7.3. *Suppose G is an (N_1, d_1, λ_1)-expander and H is a (d_1, d_2, λ_2)-expander. Then $G \textcircled{z} H$ is an $(N_1 d_1, d_2^2, f(\lambda_1, \lambda_2))$-expander, where $f(\lambda_1, \lambda_2) \leq \lambda_1 + \lambda_2 + \lambda_2^2$.*

Proof. Recall that given vectors $x \in \mathbb{R}^{N_1}$ and $y \in \mathbb{R}^{N_2}$, their tensor product is given by $x \otimes y = (x_i \cdot y_j) \in \mathbb{R}^{N_1 N_2}$. Let M be the normalized adjacency matrix of $G \textcircled{z} H$. Then

$$\lambda(G \textcircled{z} H) \;=\; \max_{\alpha \perp 1_{N_1 d_1}} \frac{|\langle M\alpha, \alpha\rangle|}{|\langle \alpha, \alpha\rangle|}.$$

So we need to show that for all $\alpha \in \mathbb{R}^{N_1 d_1}$, if $\alpha \perp 1_{N_1 d_1}$, then $|\langle M\alpha, \alpha\rangle| \leq f(\lambda_1, \lambda_2)|\langle \alpha, \alpha\rangle|$.

Let $\alpha \in \mathbb{R}^{N_1 d_1}$ such that $\alpha \perp 1_{N_1 d_1}$. For all $v \in [N_1]$, define $\alpha_v \in \mathbb{R}^{d_1}$ by $(\alpha_v)_k = \alpha_{vk}$. Also define a linear map $C : \mathbb{R}^{N_1 d_1} \to \mathbb{R}^{N_1}$ by $(C\alpha)_v = \sum_{k=1}^{d_1} \alpha_{vk}$. Then $\alpha = \sum_v (e_v \otimes \alpha_v)$. Furthermore, we can decompose α_v into $\alpha_v = \alpha_v^{\perp} + \alpha_v^{\|}$, where $\alpha_v^{\perp} \perp 1_{d_1}$. Thus we can decompose α into $\alpha^{\|}$ and $\alpha^{\perp}$ as follows:

$$\begin{aligned}\alpha &= \sum_v (e_v \otimes \alpha_v^{\|}) + \sum_v (e_v \otimes \alpha_v^{\perp})\\ &=: \alpha^{\|} + \alpha^{\perp}.\end{aligned}$$

Note that $\alpha^{\|}$, if viewed as a distribution on $G \textcircled{z} H$, is uniform within any given cloud. In fact,

$$\alpha^{\|} = \frac{C\alpha \otimes 1_{d_1}}{d_1} = \left(\frac{\text{total on } v_1}{d_1}, \ldots, \frac{\text{total on } v_i}{d_1}, \ldots, \frac{\text{total on } v_N}{d_1}\right).$$

Evidently, since the sum of the entries in $\alpha^{\|}$ equals the sum of the entries in α, namely 0, we have $\alpha^{\|} \perp 1_{N_1 d_1}$ and $C\alpha^{\|} \perp 1_{N_1}$.

Define $\widetilde{B} = I_{N_1} \otimes B$, where B is the normalized adjacency matrix for H. Thus $\widetilde{B}$ is a block diagonal square matrix of size $N_1 d_1$ with blocks B. Furthermore, let $\widetilde{A}$ be the permutation matrix corresponding to the Rot_G mapping. Then $M = \widetilde{B}\widetilde{A}\widetilde{B}$.

Note that

$$\langle M\alpha, \alpha\rangle = \left\langle \widetilde{B}\widetilde{A}\widetilde{B}\alpha, \alpha \right\rangle = \left\langle \widetilde{A}\widetilde{B}\alpha, \widetilde{B}\alpha \right\rangle,$$

since $\widetilde{B}$ is a real symmetric matrix and hence self-adjoint. Also $\widetilde{B}\alpha^{\parallel} = \alpha^{\parallel}$, since the uniform distribution on H is invariant under B. Then

$$\widetilde{B}\alpha = \widetilde{B}(\alpha^{\perp} + \alpha^{\parallel}) = \alpha^{\parallel} + \widetilde{B}\alpha^{\perp}.$$

Computing, we find

$$\begin{aligned}
\langle M\alpha, \alpha\rangle &= \left\langle \widetilde{A}(\alpha^{\parallel} + \widetilde{B}\alpha^{\perp}), (\alpha^{\parallel} + \widetilde{B}\alpha^{\perp})\right\rangle \\
&= \left\langle \widetilde{A}\alpha^{\parallel}, \alpha^{\parallel}\right\rangle + \left\langle \widetilde{A}\alpha^{\parallel}, \widetilde{B}\alpha^{\perp}\right\rangle + \left\langle \widetilde{A}\widetilde{B}\alpha^{\perp}, \alpha^{\parallel}\right\rangle + \left\langle \widetilde{A}\widetilde{B}\alpha^{\perp}, \widetilde{B}\alpha^{\perp}\right\rangle \\
|\langle M\alpha, \alpha\rangle| &\leq \left|\left\langle \widetilde{A}\alpha^{\parallel}, \alpha^{\parallel}\right\rangle\right| + \|\widetilde{A}\alpha^{\parallel}\| \cdot \|\widetilde{B}\alpha^{\perp}\| + \|\widetilde{A}\widetilde{B}\alpha^{\perp}\| \cdot \|\alpha^{\parallel}\| + \|\widetilde{A}\widetilde{B}\alpha^{\perp}\| \cdot \|\widetilde{B}\alpha^{\perp}\| \\
&= \left|\left\langle \widetilde{A}\alpha^{\parallel}, \alpha^{\parallel}\right\rangle\right| + 2\|\alpha^{\parallel}\| \cdot \|\widetilde{B}\alpha^{\perp}\| + \|\widetilde{B}\alpha^{\perp}\|^2,
\end{aligned}$$

where the last line uses the fact that $\widetilde{A}$ is a permutation and hence $\|\widetilde{A}x\| = \|x\|$, for all $x \in \mathbb{R}^{N_1 d_1}$.

To simplify this expression, we first see that

$$\begin{aligned}
\|\widetilde{B}\alpha^{\perp}\|^2 &= \|\widetilde{B}(\sum_v e_v \otimes \alpha_v^{\perp})\|^2 \\
&= \|\sum_v e_v \otimes B\alpha_v^{\perp}\|^2 \\
&= \sum_v \|B\alpha_v^{\perp}\|^2 \\
&\leq \sum_v \lambda_2^2 \|\alpha_v^{\perp}\|^2 \\
&\leq \lambda_2^2 \|\alpha^{\perp}\|^2.
\end{aligned}$$

Secondly, we need to bound $\left|\left\langle \widetilde{A}\alpha^{\parallel}, \alpha^{\parallel}\right\rangle\right|$. Let A be the normalized adjacency matrix for G; we want to relate $\widetilde{A}$ and A. Fix $e_v \in \mathbb{R}^{N_1}$. Then Ae_v gives a uniform distribution on the neighbors of v in G. This means that

$$Ae_v = C\widetilde{A} \cdot \frac{e_v \otimes 1_{d_1}}{d_1}.$$

The tensor product gives a uniform distribution on the cloud corresponding to v, multiplying by $\widetilde{A}$ moves the distribution to the neighbors of v, and multiplying by C adds up the distribution in each cloud.

By linearity, it follows that for all $\beta \in \mathbb{R}^{N_1}$,

$$A\beta = C\widetilde{A} \cdot \frac{\beta \otimes 1_{d_1}}{d_1}.$$

Take $\beta = C\alpha$. Then $\alpha^{\|} = (\beta \otimes 1_{d_1})/d_1$, so we get that $C\widetilde{A}\alpha^{\|} = AC\alpha$. Thus

$$\begin{aligned}
\left\langle \widetilde{A}\alpha^{\|}, \alpha^{\|} \right\rangle &= \left\langle \widetilde{A}\alpha^{\|}, C\alpha \otimes 1_{d_1} \right\rangle /d_1 \\
&= \left\langle C\widetilde{A}\alpha^{\|}, C\alpha \right\rangle /d_1 \\
&= \langle AC\alpha, C\alpha \rangle /d_1 \\
\left| \left\langle \widetilde{A}\alpha^{\|}, \alpha^{\|} \right\rangle \right| &\leq \lambda_1 \langle C\alpha, C\alpha \rangle /d_1 \\
&= \lambda_1 \langle C\alpha \otimes 1_{d_1}, C\alpha \otimes 1_{d_1} \rangle /d_1^2 \\
&= \lambda_1 \left\langle \alpha^{\|}, \alpha^{\|} \right\rangle .
\end{aligned}$$

Combining the two inequalities, we find

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

Take $p = \|\alpha^{\|}\|/\|\alpha\|$ and $q = \|\alpha^{\perp}\|/\|\alpha\|$, so that $p^2 + q^2 = 1$. Then

$$\begin{aligned}
\frac{|\langle M\alpha, \alpha \rangle|}{|\langle \alpha, \alpha \rangle|} &\leq \lambda_1 p^2 + 2\lambda_2 pq + \lambda_2^2 q^2 \\
&= \lambda_1 + \lambda_2 + \lambda_2^2.
\end{aligned}$$

□

References

[GG] Ofer Gabber and Zvi Galil, “Explicit construction of linear size superconcentrators”, JCSS 22(1981), 407–420.

[LPS] Alexander Lubotzky, R. Phillips, P. Sarnak: “Ramanujan graphs”. Combinatorica 8(3): 261–277 (1988)

[Mar] Gregori A. Margulis. “Explicit constructions of expanders”. Problemy Peredaci Informacii, 9(4):71–80, 1973.

[RVW] Omer Reingold, Salil Vadhan, and Avi Wigderson: “Entropy Waves, The Zig-Zag Graph Product, and New Constant-Degree Expanders and Extractors”, Annals of Math 2002.