6. The zig-zag product

We will define a graph product called the zig-zag product, and show that zig-zag product of two expanders is an expander as well.

Again G is an (n, d, α)-graph with normalized adjacency matrix $\hat{A}$. The k-th power G^k is a graph on the same vertex set where we put an edge uv for every walk of length k from u to v. It is easy to check that G^k is an (n, d^k, α^k)-graph.

For an (n, m)-graph G and (m, d)-graph H, we will define their zig-zag product $G ⓩ H$ as an (mn, d^2)-graph. Before defining what it is, we introduce the following theorem stating that the zig-zag product of expanders is also an expander.

Theorem 6.1 (Reingold-Vadhan-Wigderson 2002)**.** *Let G be an (n, m, α)-graph and H be an (m, d, β)-graph. Then $G ⓩ H$ is an $(nm, d^2, \varphi(\alpha, \beta))$-graph where the function φ satisfies the following.*

(1) If $\alpha < 1$ and $\beta < 1$, then $\varphi(\alpha, \beta) < 1$.
(2) $\varphi(\alpha, \beta) < \alpha + \beta$.
(3) $\varphi(\alpha, \beta) \leq 1 - (1 - \beta^2)(1 - \alpha)/2$.

(1) implies that the product is again an expander. When α, β are small (2) is useful. If α, β are big, then (3) is useful.

Using this theorem, we can explicitly construct a family of constant degree expanders. To generate such an infinite family, we need a fixed size expander of certain parameter.

By using exhaustive search, we can find a graph H which is $(d^4, d, 1/4)$-graph for some constant d in constant time. Using this, we define

$$G_1 = H^2, \;\; G_{n+1} = (G_n)^2 ⓩ H. \tag{6.1}$$

Then, using induction, assume that $(G_n)^2$ is a $(d^{4n}, d^4, 1/4)$-graph. Then $(G_n)^2 ⓩ H$, bound (2) implies that G_{n+1} is $(d^{4(n+1)}, d^2, 1/2)$-graph. By repeating this, we obtain a graph with arbitrarily many vertices with degree d^2, which is an expander.

Now we define zig-zag product. Let G be an (n, m, α)-graph and H be an (m, d, β)-graph. For each $v \in V(G)$, we fix some numbering $e_v^1, \ldots, e_v^m$ of the edges incident with v. Let $V(H) = [m]$.

We first define $G ⓡ H$, the replacement product of the two graphs, with the vertex set $V(G) \times V(H)$. We might want to imagine that we replace each vertex v of G into a cloud of m vertices $(v, 1), \ldots, (v, m)$. We add an edge $(v, i)(v, j)$ if $ij \in E(H)$ and $(v, i)(u, j)$ if $e_v^i = e_u^j$ to obtain $G ⓡ H$.

The vertex set of $G ⓩ H$ is again $V(G) \times V(H)$. We add an edge $(v, i)(u, j)$ to $G ⓩ H$ if there exists a walk $(v, i)(v, i')(u, j')(u, j)$ of length three in $G ⓡ H$.

Definition 6.2. *$G ⓩ H = (V(G) \times [m], E')$ where $(v, i)(u, j) \in E'$ iff there exists $k, \ell \in [m]$ such that $ik, \ell j \in E(H)$ and $e_v^k = e_u^\ell$.*

Here, we prove a weaker bound of $\varphi \leq \beta + \max\{\alpha, \beta^2\}$ than (2).

Proof. Let $G' = G ⓩ H$. We analyze the spectral gap of the zig-zag product by considering random walk on G'. Each step in this walk is same as

(i) take a random step on an edge within a cloud
(ii) take a deterministic step on an edge connecting two clouds
(iii) take another random step within a cloud.

We now write down the transition matrix Z of the random walk on G'. Let $B, \hat{B}$ be the adjacency matrix of H and the transition matrix of the corresponding random walk (which is the normalized adjacency matrix of H). The step (i) and (iii) are done by $\tilde{B} = \hat{B} \otimes I_n$.

Let P be the permutation matrix defined by

$$P_{(v,k),(u,\ell)} = \begin{cases} 1 & \text{if } e_v^k = e_u^\ell \\ 0 & \text{otherwise.} \end{cases}$$

Then $Z = \tilde{B}P\tilde{B}$.

To make the equation less clutter, let's write fZf instead of f^TZf, by understanding the left f is a row vector. As the graph G' is an (mn, d^2)-graph, which is regular, so $\mathbf{1}_{mn}$ is an eigenvector. So, we want to claim that for all vectors $f \perp 1_{mn}$, we have

$$\frac{|fZf|}{\|f\|^2} \leq \alpha + \beta + \beta^2.$$

For a given f, we write $f^{\|}$ be the vector we obtain by making it constant on each cloud. In other words, $f^{\|}(x,i) = \frac{1}{m}\sum_{j\in[m]} f(x,j)$. Let $f^{\perp} = f - f^{\|}$. Then $f^{\perp}$ is orthogonal to a constant vector once restricted to a cloud, hence it does not expand more than β when multiplied by $\hat{B}$. We have

$$\begin{aligned} |fZf| &= |f\tilde{B}P\tilde{B}f| \\ &\leq |f^{\|}\tilde{B}P\tilde{B}f^{\|}| + 2|f^{\|}\tilde{B}P\tilde{B}f^{\perp}| + |f^{\perp}\tilde{B}P\tilde{B}f^{\perp}|. \end{aligned}$$

We know that $\hat{B}1_m = 1_m$, so we have $\tilde{B}f^{\|} = f^{\|}$. Also, we know $\|\tilde{B}f^{\perp}\| \leq \beta\|f^{\perp}\|$ as we know $\|\hat{B}u\| \leq \beta\|u\|$ for any u perpendicular to 1_m.

Let $g : V(G) \to \mathbb{R}$ be $g(v) = \frac{1}{\sqrt{m}}\sum_{j\in[m]} f(v,j)$. Then we have $\|g\|^2 = \|f^{\|}\|^2$. By the definition of P, we have $f^{\|}Pf^{\|} = g\hat{A}g$ where $\hat{A}$ is the transition matrix of the random walk of G. (Note that $\hat{A}$ is normalized matrix here, so we divide A by m.) However, $f^{\|} \perp 1_{mn}$ implies that $g \perp 1_n$ and hence $g\hat{A}g \leq \alpha\|g\|^2$. Consequently $|f^{\|}Pf^{\|}| \leq \alpha\|f^{\|}\|^2$. Also, both $\tilde{B}$ and P are doubly stochastic matrices and therefore contractions in ℓ_2. (This can be shown, as any doubly stochastic matrice is a convex combination of permutation matrices. It is easy to see that each permutation matrix is a contraction and a convex combination of them is also a contraction.) Hence, we have

$$|fZf| \leq \alpha\|f^{\|}\|^2 + 2\beta\|f^{\|}\|\|f^{\perp}\| + \beta^2\|f^{\perp}\|^2.$$

However, we have $\|f\|^2 = \|f^{\|}\|^2 + \|f^{\perp}\|^2$ so we have

$$\frac{|fZf|}{\|f\|} \leq \frac{2\beta\|f^{\|}\|\|f^{\perp}\|}{\|f^{\|}\|^2 + \|f^{\perp}\|^2} + \frac{\max\{\alpha,\beta^2\}(\|f^{\|}\|^2 + \|f^{\perp}\|^2)}{\|f^{\|}\|^2 + \|f^{\perp}\|^2} \leq \beta + \max\{\alpha,\beta^2\}.$$

□

Now we consider this as a perspective of entropy. As above, the random walk actually consists of three steps on the replacement product. A random step in one copy of H and a deterministic step to a neighboring cloud, and another random step in the new copy of H. Note that the first and third step are independent random steps on H. If the conditional distribution restricted to one of these clouds is far from uniform, then the entropy grows as H is an expander. The other two steps does not harm as the entropy never drops when we multiply with doubly stochastic matrix.

If the distribution is nearly uniform on the most of clouds, then the Step 1 does not increase the entropy much. In this case, as the distribution is uniform on each cloud, the second step is almost like a real random step on G. Then the entropy of p_G increase. But this middle step is a permutation on $V(G\textcircled{z}H)$, so the entropy of the whole distribution remains unchanged. Hence, the entropy of p_H must have decreased. That means in step 3, the conditional distribution on clouds are not close to uniform, and the entropy increase due to the expansion of H.

This zig-zag product has an application to complexity theory. The following is a definition of Turing machine written on wikipedia. A Turing machine is a mathematical model of computation that defines an abstract machine that manipulates symbols on a strip of tape according to a table of rules. In other words, it is a machine that performs an algorithm following a fixed rule. L (logspace) is a collection of problems that can be solved by a Turing machine using a logarithmic amount of writable memory space. In other words, while remembering a limited amount of information, your Turing machine performs algorithm according to current input and the memory.

However, one specific rule dictating an action might not seem so efficient. For given input (or state) and memory, one might suggests several options. An non-deterministic Turing machine is a theoretical model of computation whose governing rules specify more than one possible action when in some given situations. We consider a collection of problems which can be solved using non-deterministic Turing machine using only logarithmic memory size. We assume one more condition that if transiting from a state A to B is possible then transiting from the state B to A is also possible. We call such collection of problems as SL.

Considering each state as vertex, and add edge between two states if your non-deterministic Turing machine can transit from one state to the other state. This will define an undirected graph. s represent your initial state, and t represent the state you want to reach. Now the question becomes whether there exists a path from s to t, while using logarithmic size (logarithm of the size of the graph) memory. We call such a problem $USTCON$ (undirected s,t-connectivity).

Aleliunas, Karp, Lipton, Lovasz, Rackoff in 1979 showed that this problem can be solved by a probabilistic logspace algorithm, proving $SL \subseteq RL$. To determine if s and t are connected, one simply performs a polynomial length random walk starting at s and checks if the walk ever reaches t.

All we need to remember is the current position and the goal t, so it uses only logarithmic memory (location requires $\log n$ bits of information where n is the number of vertices) We may assume that the graph is regular, as otherwise we replace each vertex v with a cycle of length $d(v)$ to obtain an 3-regular graphs with at most n^2 vertices. This is fine as $\log(n^2) = O(\log n)$.

As every connected graph is (n, d, α)-graph with $\alpha < 1 - \Omega(1/n^2)$. Thus a random walk of length $O(n^3)$ will get exponentially close to the uniform distribution, and if we repeat it n^2 times, resulting in a walk of length n^5, we will not miss a single vertex in this connected component, except with a exponentially small probability.

It is proven to be $S = SL$. In order to prove this, we need to derandomize this approach. How can we do this? If every component of the graph is expander, then we can simply enumerate over all the logarithmically long paths from s and check if one of them arrives at t. So the problem becomes trivial. However, the graph in general is not an expander.

So, we use zig-zag product to turn the graph into an expander.

Consider the D-regular graph G and assume it is connected. Then G is (n, D, α)-graph with $\alpha < 1 - \Omega(1/n^2)$ by connectivity. Assume that $D = d^{16}$, and find an $(d^{16}, d, 1/2)$-graph H. We inductively construct the graphs G_i as follows.

$$G_1 = G, \;\; G_{i+1} = (G_i ⓩ H)^8.$$

We get this to get G_k with $k = O(\log n)$. Then we claim that we have $(nd^{16k}, d^{16}, 3/4)$-expander G_k.

Indeed, by using (3) in the zig-zag theorem, we can show that the spectral gap doubles in each step. If $1-\lambda_i, 1-\mu_i$ are normalized second eigenvalues of G_i and $G_i ⓩ H$ respectively, then (3) implies

$$\mu_i \geq \frac{3}{8}\lambda_i.$$

Hence

$$\lambda_{i+1} = 1 - (1 - \mu_i)^8 \geq 1 - (1 - \frac{3}{8}\lambda_i)^8 \geq \min\{2\lambda_i, 1/2\}.$$

So, for $k = O(\log n)$, we have $\lambda_k \geq 1/2$.

Moreover, the neighborhood queries for G_k can be answered in logspace. (One cannot remember the entire graph G_i in each step using only logarithmic memory space) This is not obvious. This means that the graph G_k is very explicit. Reingold in 2005 proved that this is possible by evaluate the recursion for each query.