Lecture 11: ZigZag products and Expanders

October 12, 2004 *Scribe: Peter Pal Zubcsek*

1 Tensor Product

Lemma 1 *For symmetric matrices A,B, the eigenvalues of $A \otimes B$ are exactly $\lambda \cdot \mu$, where λ is the eigenvalue of A, and μ the eigenvalue of B.*

Proof (sketch): Let v, w be the eigenvectors of A and B. Then

$$(A \otimes B)(v \otimes w) = (Av) \otimes (Bw) = (\lambda v) \otimes (\mu w) = \lambda\mu(v \otimes w).$$

Exercise 2 *show that if $v_1, v_2, \ldots, v_n$ are mutually orthogonal and $w_1, w_2, \ldots, w_n$ are mutually orthogonal, then $\{v_i \otimes w_j\}_{i,j}$ are mutually orthogonal as well.*

So, $v \otimes w$ are mutually orthogonal eigenvectors of $A \otimes B$ that form a basis of $\mathbb{R}^{nm}$.

Exercise 3 *show that if $v_1, v_2, \ldots, v_n$ are linearly independent vectors for the $n \times n$ matrix A such that $\forall i A v_i = \lambda_i v_i$, then $\forall \lambda, \forall$ non-zero vector v, if $Av = \lambda v$, then $\lambda = \lambda_i$ for some $1 \le i \le n$.*

Corollary 4 *if G is an (n, d, λ)-expander, then $G \otimes G$ is an (n^2, d^2, λ)-expander.*

Proof: It's easy to see the relation regarding the number of vertices and the degree of the graph. For the second largest eigenvalue, consider that $1 = 1 \cdot 1$ is an eigenvalue of $G \otimes G$. Then the second largest eigenvalue should be $1 \cdot \lambda$.

2 Zig-Zag Expander Construction, 2nd attempt

Let H be an (d^4, d, λ_0)-expander for $\lambda_0 \le \frac{1}{5}$ and let H' be an (d^8, d, λ_0)-expander.

Define

$$G_1 = H^2$$

$$G_{t+1} = (G_t \otimes G_t)^2 \, \textcircled{z} \, H'.$$

Then degree$(G_t) = d^2$. Moreover, if N_t denotes the number of vertices in G_t, then
$N_{t+1} = (N_t)^2 \cdot d^8$. Thus,
$N_t \cong c^{2^t}$, while the time of computing a neighbor of a vertex (here denoted by TimeG_t)
Time$G_t \cong O(4^t) = \text{polylog}(N_t)$.
The problem with this construction is that the N_t's are too far apart.

3 Zig-Zag Expander Construction, 3rd attempt

Use G_1, H' from above, and
Define

$$G_t = (G_{[\frac{t}{2}]} \otimes G_{[\frac{t}{2}]})^2 \textcircled{z} H'.$$

Then $N_t \cong c^t$ and so
Time$(G_t) \cong$ poly(t) = polylog(N_t).

4 The Zig-Zag Product

Recall that the Zig-Zag Product, denoted by $G \textcircled{z} H$ had the following requirements on the graphs G and H:
G is (n_1, d_1, λ_1)-expander, H is (d_1, d_2, λ_2)-expander.

Informal definition of $G \textcircled{z} H$
Take any vertex $u \in G$ with d_1 neighbors $w_1, ..., w_{d_1}$
Make a copy of H around u (see Figure 1)

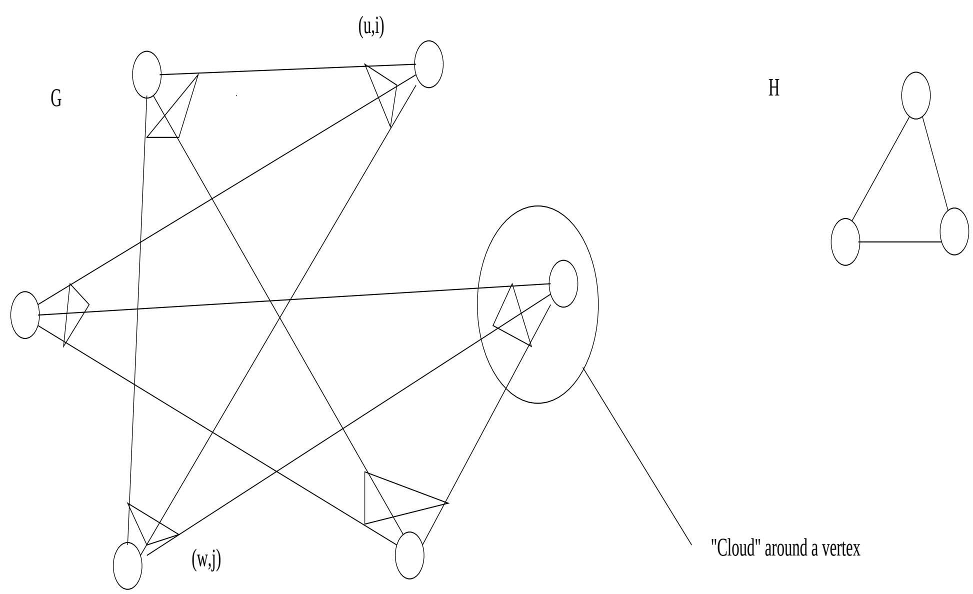

Figure 1: Zig-Zag Product Graph - construction.

A vertex of $G \textcircled{z} H$ is a pair (u, i) for $u \in V(G)$ and $i \in V(H)$.
The first (Zig-) step is a jump to a randomly chosen neighbor in H.
In the second step, we go along the picked edge (our edge may have another number from the other end - see Figure1).
The third (Zag-) step is another random jump, in the cloud of w.

So, from (u, i) we move to (w, j). Wherever we can make the 3-step zig-zag jump, there is an edge in the zig-zag product graph.

A more formal definition of $G ⓩ H$

The vertices of the zig-zag product of the graphs G and H are pairs $(u, i) \in V \times [d_1]$, where $V = V(G)$. Two vertices (u, i) and (v, j) are connected by an edge if $\exists i', j', j \in [d_1]$ such that

- $(i, i') \in E(H)$,
- $(u, v) \in E(G)$ and the edge between them is the (i')-th edge leaving u and the (j')-th edge leaving v,
- $(j, j') \in E(H)$.

Lemma [Zig-Zag Lemma]
For (n_1, d_1, λ_1)-expander G and (d_1, d_2, λ_2)-expander H, $G ⓩ H$ 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: the statements on the number of vertices and the degree of the result are clear.
Intuition: There are two extreme cases of distributions on (u, i).

- I. The distribution is uniform over the u's (the clouds), and very far from uniform within any cloud $H_u = \{(u, i) | i \in [d_i]\}$.
  The Zig step is a random step on an expander H. So, the distribution within each cloud gets closer to uniform. Hence, we get closer to the global uniform distribution.
  The global distribution has the same distance from uniform after the second step.
  The Zag step may or may not get us closer to the global uniform (we might waste some randomness).

- II. The distribution is very far from uniform within the clouds but uniform within each cloud.
  The Zig step has no effect (we waste some randomness).
  On the one hand, the second step is a permutation of vertices and hence does not change the global distance from uniform. On the other hand, it's a random step in the expander G, and hence the distribution over the clouds must get closer to uniform. Since the total distance from uniform stays the same, we must have clouds whose distribution within the cloud is farther from uniform.
  So, the Zag Step makes us closer to the global uniform.

A formal proof: Let A, B, M be the vertices of $G, H, G ⓩ H$.

Then, M can be decomposed as $M = \tilde{B}\hat{A}\tilde{B}$, where $\tilde{B} = I_{n_1} \otimes B$.

$$\hat{A}_{(u,i),(v,j)} = \begin{cases} 1 \text{ if } (u, v) \in E(G) \text{ and the conjunctive edge is the } i\text{-th of } u \text{ and the } j\text{-th of } v \\ 0 \text{ otherwise} \end{cases}$$

By Rayleigh, $\lambda(G ⓩ H) = \max_{x \perp u_{n_1 d_1}, \|x\|=1} (Mx, x)$.
Write $x = x^{\parallel} + x^{\perp}$,
where $x^{\parallel}$ is constant over each cloud (uniform within each cloud) and $x^{\perp}$ is perpendicular to uniform within each cloud. (E.g., $x^{\parallel}$ and $x^{\perp}$ are computed for each cloud separately.)
So, we need to upper bound (Mx, x).

- Case I:
  (a) $\|\tilde{B}x^{\perp}\| \leq \lambda_2 \|x^{\perp}\|$, since B shrinks $x^{\perp}$ within each cloud,

(b) $\forall y \in \mathbb{R}^{n_1 d_1}$, $\|\hat{A}y\| = \|y\|$, since $\hat{A}$ is a permutation.
(c) $\|\tilde{B}y\| \leq \|y\|$, since the largest eigenvalue of B is 1.

- Case II:
  Note that $x^{\|} = w \otimes u_{d_1}$, where $(w, u_{d_1}) = 0$.

  Exercise 5 *Check the above note.*

  (a) $\tilde{B}x^{\|} = x^{\|}$,
  (b) $(\hat{A}x^{\|}, x^{\|}) = \frac{1}{d_1}(Aw, w) \leq \frac{\lambda_1 (w,w)}{d_1} = \lambda_1 (x^{\|}, x^{\|})$.

Finally (since $\tilde{B}^T = \tilde{B}$ as $\tilde{B}$ is symmetric),

$$\begin{aligned}
(Mx, x) &= (\tilde{B}\hat{A}\tilde{B}(x^{\|} + x^{\perp}); x^{\|} + x^{\perp}) \\
&= (\hat{A}\tilde{B}(x^{\|} + x^{\perp}); \tilde{B}(x^{\|} + x^{\perp})) \\
&= (\hat{A}(x^{\|} + \tilde{B}x^{\perp}); x^{\|} + \tilde{B}x^{\perp}).
\end{aligned}$$

So,

$$\begin{aligned}
|(Mx, x)| &\leq |(\hat{A}x^{\|}, x^{\|}) + (\hat{A}\tilde{B}x^{\perp}, x^{\|}) + (\hat{A}x^{\|}, \tilde{B}x^{\perp}) + (\hat{A}\tilde{B}x^{\perp}, \tilde{B}x^{\perp})| && (1) \\
&\leq |(\hat{A}x^{\|}, x^{\|})| + |(\hat{A}\tilde{B}x^{\perp}, x^{\|})| + |(\hat{A}x^{\|}, \tilde{B}x^{\perp})| + |(\hat{A}\tilde{B}x^{\perp}, \tilde{B}x^{\perp})| && (2) \\
&\leq \lambda_1 \|x^{\|}\|^2 + 2\|\tilde{B}x^{\perp}\| \cdot \|x^{\|}\| + \|\tilde{B}x^{\perp}\|^2 && (3) \\
&\leq \lambda_1 \|x^{\|}\|^2 + 2\lambda_2 \|x^{\perp}\| \cdot \|x^{\|}\| + {\lambda_2}^2 \|x^{\perp}\|^2 && (4) \\
&\leq \lambda_1 + \lambda_2 + {\lambda_2}^2. && (5)
\end{aligned}$$

To obtain line (2), we used the triangle inequality. To obtain line (3), we relied on the Cauchy-Schwarz inequality and noted that $\hat{A}$ does not change the norms since it's a permutation. At the last step, we used the following facts:

- $\|x^{\|}\|$ and $\|x^{\perp}\|$ are projections of vectors of norm 1,
- $\sqrt{\|x^{\|}\|^2 \cdot \|x^{\perp}\|^2} \leq \frac{\|x^{\|}\|^2 + \|x^{\perp}\|^2}{2} = \frac{1}{2}$ (inequality between the algebraic and the geometric mean).