0368-4283: Space-Bounded Computation 17/4/2018 – Lecture 5

Derandomized squaring

Amnon Ta-Shma and Dean Doron

1 The Derandomized-Squaring Product

There is an obvious way to increase the connectivity of graphs – graph *squaring* (or more generally, graph powering). If G is our original graph, than G^2 is a graph in which every edge corresponds to a path of length 2 in G. Iterating $\log n$ times, in G^n we can in fact check whether s is connected to t in G by checking whether (s,t) is an edge in G^n. As we saw squaring increases both the spectral gap and the degree, if G is undirected degree D than G^t is degree D^t and $\bar{\lambda}(G^t) = \bar{\lambda}(G)^t$. Thus, although G^n is extremely well-connected, the degree of each vertex can be exponentially-large so even enumerating over all neighbours cannot be done in logspace. We will give a solution to that problem, due to Rozenman and Vadhan [1].

1.1 Preliminaries

Even though we will only solve *undirected* st-connectivity, it will be useful to work with regular directed graphs. Consequently, we will have to adapt our spectral definitions.

Definition 1. *We say a directed graph G is D-regular if the in-degree as well as the out-degree of every vertex is D.*

Let M be the transition matrix of a directed D-regular graph over N vertices and let $u = \frac{1}{N}\mathbf{1}$, so by regularity $Mu = u$.

Definition 2. *Let G be a directed regular graph with transition matrix M. We define*

$$\bar{\lambda}(G) \;=\; \max_{x \perp u} \frac{\|Mx\|}{\|x\|}$$

and $\gamma(G) = 1 - \bar{\lambda}(G)$. We say G is a (N, D, λ)-graph if G is a directed D-regular graph over N vertices and $\bar{\lambda}(G) \leq \lambda$.

In the regular directed case as well, $\bar{\lambda}(G)$ measures the rate at which a random walk over G converges to the stationary distribution u. As usual, we call graphs with $\bar{\lambda}(G) \leq 1 - \Omega(1)$ *expanders*. If G is undirected than $\bar{\lambda}(G)$ is the second largest eigenvalue in magnitude of M. In regular directed graphs, it equals the second singular value of M, i.e., the square root of the second largest eigenvalue of $M^\dagger M$ (prove!).

A random step on G can be viewed as going according to the uniform distribution (or, taking a step in a clique) with probability $1 - \bar{\lambda}(G)$ and "not doing too much harm" with probability $\bar{\lambda}(G)$. We already proved it for the undirected case, and the same result holds here as well (prove!).

Claim 3. *Let M be the transition matrix of an (N, D, λ)-graph. Then, $M = (1-\lambda)\frac{1}{N}J + \lambda C$ for some C with $\|C\| \leq 1$.*

1.2 The derandomized-squaring product

Let G be an undirected D_1-regular graph. One way to view the squaring G^2 is that for every vertex v in G we place a clique on its D_1 neighbors. The degree obviously becomes D_1^2. To "derandomize" this approach we use a degree-D_2 expander H on D_1 vertices and place it instead of a clique on the D_1 neighbours of every vertex v. Thus, we pick one edge from the expander, and use its two endpoints as the two instructions for G^2. The resulting graph, $G\textcircled{S}H$, will have degree D_1D_2, which is a significant gain if $D_2 \ll D_1$. The rotation map of $G\textcircled{S}H$ is formally defined as follows.

Definition 4. *Let G be a directed D_1-regular graph on N vertices with a two-way labelling and let H be a directed D_2-regular graph on D_1 vertices with a two-way labelling. The graph $G\textcircled{S}H$ has rotation map $Rot_{G\textcircled{S}H} : [N] \times ([D_1] \times [D_2]) \to [N] \times ([D_1] \times [D_2])$ so that $Rot_{G\textcircled{S}H}(v_0, (a, b))$ is defined as follows:*

1. *Let $(v_1, a') = Rot_G(v, a)$.*
2. *Let $(a'', b') = Rot_H(a', b)$.*
3. *Let $(v_2, a''') = Rot_G(v_1, a'')$.*
4. *Output $(v_2, (a''', b'))$.*

To get a better understanding of derandomized squaring product, note the following phenomena:

- In general, $G\textcircled{S}H$ may not produce an in-regular graph, but it will do so provided G is consistently labelled. If G and H are consistently labeled then $G\textcircled{S}H$ is also consistently labeled.
- Even if G and H are consistently labeled and undirected, $G\textcircled{S}H$ need not be undirected.

We will now prove that although the derandomized squaring graph has smaller degree than D_1^2, it improves connectivity almost just as well as ordinary squaring as long as H is a good expander.

Theorem 5. *Let G be an (N, D_1, λ_1)-graph with a two-way labelling and let H be a (D_1, D_2, λ_2)-graph with a two-way labelling. Then, $G\textcircled{S}H$ is an $(N, D_1D_2, f(\lambda_1, \lambda_2))$-graph, where*

$$f(\lambda_1, \lambda_2) \;=\; 1 - (1 - \lambda_1^2)(1 - \lambda_2) \;\leq\; \lambda_1^2 + \lambda_2.$$

Notice that when λ_2 is very small, $f(\lambda_1, \lambda_2)$ approaches λ_1^2 which is what we would have gotten for G^2.

Proof. Let M be the transition matrix of $G\textcircled{S}H$, and we wish to bound $\|Mv\|$ for every $v \perp u = \frac{1}{N}\mathbf{1}$. In order to translate $G\textcircled{S}H$ to operators notation, we think of a random step on $G\textcircled{S}H$ from a vertex v:

1. We choose a uniformly at random from $[D_1]$ and go to state (v, a).
2. We apply $(v_1, a') = Rot_G(v, a)$.
3. We choose b uniformly at random from $[D_2]$, compute $(a'', b') = Rot_H(a', b)$ and go to state (v_1, a'').

4. We apply $(v_2, b') = Rot_G(v_1, a'')$.

5. Output v_2.

Step (1) corresponds to a linear mapping $L \in \mathbb{R}^{ND_1 \times N}$ that lifts probability distributions on $[N]$ to probability distributions on $[N] \times [D_1]$ by spreading the weights uniformly over each "cloud". Namely, $Lv = v \otimes u$.

Step (2) corresponds to applying the permutation Rot_G, and we let $\dot{A}$ be the corresponding permutation matrix.

Step (3) corresponds to applying the matrix $\tilde{H} = I_N \otimes M_H$ where M_H is the transition matrix of H (i.e., we take a random step over the expander in each cloud).

Step (4) again corresponds to $\dot{A}$.

Step (5) corresponds to a linaer mapping $P \in \mathbb{R}^{N \times ND_1}$ that projects probability distributions on $[N] \times [D_1]$ to probability distributions on $[N]$ by summing the weights of each cloud.

Hence,

$$M \;=\; P\dot{A}\tilde{H}\dot{A}L.$$

By Claim 3 we can write $M_H = (1-\lambda_2)\frac{1}{D_1}J + \lambda_2 C$ for some C with $\|C\| \leq 1$. Thus, also $\tilde{H} = (1-\lambda_2)(I_N \otimes \frac{1}{D_1}J) + \lambda_2(I_N \otimes C) \stackrel{\text{def}}{=} (1-\lambda_2)\tilde{J} + \lambda_2\tilde{C}$. Therefore,

$$M \;=\; (1-\lambda_2)P\dot{A}\tilde{J}\dot{A}L + \lambda_2 P\dot{A}\tilde{C}\dot{A}L.$$

Claim 6. *Let A be the transition matrix of G. Then, $P\dot{A}\tilde{J}\dot{A}L = A^2$.*

Proof. The operator $P\dot{A}\tilde{J}\dot{A}L$ corresponds to using the clique instead of H, which amounts to computing the actual squaring. Formally, note that $LP = \tilde{J}$ so $P\dot{A}\tilde{J}\dot{A}L = P\dot{A}LP\dot{A}L$. The claim follows from the simple observation that $P\dot{A}L = A$. □

Thus, $M = (1-\lambda_2)A^2 + \lambda_2 P\dot{A}\tilde{C}\dot{A}L$. Since $\|L\| = \frac{1}{\sqrt{D_1}}$, $\|P\| = \sqrt{D_1}$, $\left\|\dot{A}\right\| = 1$ and $\|C\| \leq 1$ we get that $\left\|P\dot{A}\tilde{C}\dot{A}L\right\| \leq 1$. Therefore, $M = (1-\lambda_2)A^2 + \lambda_2 D$ for some matrix D with $\|D\| \leq 1$. This implies that $\bar{\lambda}(G \circledS H) \leq (1-\lambda_2)\lambda_1^2 + \lambda_2 = f(\lambda_1, \lambda_2)$, as desired. □

We record the following easy corollary:

Corollary 7. *For $\gamma < \frac{1}{4}$, $1 - f(1-\gamma, \frac{1}{100}) \geq \frac{3}{2}\gamma$.*

The way to think about the above claim is that as long as we use an expander with $\lambda_2 = \frac{1}{100}$ (and not smaller), the spectral gap increases by a constant factor whenever we apply derandomized squaring, as long as it's not too large to begin with.

References

[1] Eyal Rozenman and Salil Vadhan. Derandomized squaring of graphs. In *Approximation, Randomization and Combinatorial Optimization. Algorithms and Techniques*, pages 436–447. Springer, 2005.