Moden Spectral Graph Theory **Winter 2022**

Lecture 15: Log Space Computation

Lecturer: Shayan Oveis Gharan

Disclaimer: *These notes have not been subjected to the usual scrutiny reserved for formal publications.*

A randomized space $S(n)$ touring machine is machine that uses space $S(n)$ on inputs of length n, has two non-deterministic choices at any step of computation depending on an unbiased coin-flip. In words, the machine has a read-only input tape, one work tape of size $S(n)$ and a one way output tape. The machine is said to be halting if for any input and any sequence of coin-flips the computation terminates. We write BPL for the class of languages L for which there is halting randomized $\log(n)$ space machines such that for each input $x \in L$ the machine accepts with probability at least $2/3$ and for $x \notin L$ the machine rejects with probability at least $2/3$.

Given such a machine, we may visualize computation on an input x as a state transition graph on a directed graph G as follows: The vertex set of the graph is the set of configurations in the computation, namely the instantaneous content of the work-tape, the head's position on the input and the state of the finite state machine. Each vertex has two outgoing edges labelled by either 0 or 1. There is a directed edge from a vertex u to v with label b iff starting at configuration u the computation goes to v if the coin flip of this step is b. The graph also has two states ACCEPT and REJECT. Since the configuration is guaranteed to halt, the graph is acyclic (except for the halting configurations). So, any computation must halt in at most the number of vertices of this graph. So, to prove that BPL is equal to L it is enough to show that, assuming such a graph exists, we can deterministically compute the probability of start to ACCEPT within 1/4-additive error in $O(\log n)$ space.

Say graph G has $m = poly(n)$ vertices. The simplest such algorithm simply computes $P^m(start, ACCEPT)$ using $\log m$ matrix multiplication and uses $O(\log^2 m) = O(\log^2 n)$ space. This was improved by Saks and Zhao to $O(\log^{3/2} n)$ using Nisan's pseudorandom generator in 1995. It remains a big open problem, and perhaps the easiest open problem in complexity theory to improve the $\log^{3/2} n$ to $\log n$.

A landmark result in this area is by Reingold [Rei08] who showed that if the graph G is undirected then one can deterministically test whether there is a path from Start to Accept in space $O(\log n)$. Building on [Rei08], [RTV06] designed a deterministic log-space algorithm that given an *Eulerian* directed graph, i.e., a graph that each vertex has an outdegree equal to its indegree, and two vertices s and t, finds a path between s, t if one exists.

The point is Eulerian directed graphs inherit several properties of undirected. For example, it was show in [BPS16] that, similar to undirected graphs, for Eulerian digraphs, the mixing time is $O(mn)$, where m is the number of edges and n is the number of vertices. Such a fact fail miserably for non-Eulerian directed graphs as the mixing time can be exponentially large in the number of vertices.

Reingold, Trevisan and Vadhan [RTV06] proved the following problems is also complete for the class BPL. Given a graph G with N vertices, an integer k given in unary and s, t with a promise that one of the following is true:

Yes Instances: The (simple) random walk on G has a stationary distribution π with $\pi(s), \pi(t) \geq 1/k$ and the walk on G mixes in $O(k \log N)$.

No Instances: There is no path from s to t in G.

In these lectures we will present a recent work of Ahmadinejad, Kelner, Murtagh, Peebles, Sidford and Vadhan [AKMPSV20] who proved the following statement:

Theorem 15.1. *There is a deterministic, $O(\log(kN)\log\log(kN/\epsilon))$-space algorithm that given an* Eulerian digraph G*, two vertices s, t, and a positive integer k, outputs the probability that a k-step random walk in G started at s ends at t, to within additive error of ϵ, where N is the length of the input.*

The proof builds on the recent work of Cohen, Kelner, Kyng, Peebles, Peng, Rao and Sidford [CKKPPRS18] who extended Laplacian solvers machinery of Spielman and Teng [ST14] to Eulerian directed graphs. Namely, they showed that many questions related to random walks such as escape time, hitting time, etc can be solved in near linear time if the underlying directed graph is Eulerian or if it has a polynomial mixing. The work [AKMPSV20] managed to combine ideas from [CKKPPRS18] with classical works of [Rei08] to get the above theorem. However, a big open problem is that if the above theorem can be extended to the setting of directed graphs with polynomial mixing. The tantalizing questions is whether one can extend ideas from [CKKPPRS18] to the setting of small space algorithms.

15.1 Spectral Approximation of Directed Matrices

Adjacency Matrix of Digraphs The adjacency matrix of a directed graph G on n vertices is the matrix $A \in \mathbb{R}^{n\times n}$ where $A_{u,v}$ is the number of edges from vertex u to vertex v in G. Let the degree matrix D be the diagonal matrix containing the out-degrees of the vertices in G. The simple random walk matrix on G is defined as $P = D^{-1}A$, namely, for any function $f :\in \mathbb{R}^V$,

$$Pf(u) = \mathbb{E}_{(u,v)|u} f(v).$$

We write $I - P$ for the corresponding (normalized) Laplacian matrix.

Proof Plan: We first sketch the proof strategy of [AKMPSV20]: First we will find a rough approximation of P^k in small space. We use that to get a rough approximation of $(I-P)^\dagger$ (pseudo-inverse of Laplacian). Then, we use a technique called Richardson iteration to get a very good approximation of $(I-P)^\dagger$. Finally we carry over that to approximate entries of P^k. As we will see the majority of new ideas is just for the first step.

Since we are looking for approximating high powers of random walk matrix, P, it is fundamental to look for notions of matrix approximations which are robust under taking powers: More specifically, say we (spectrally) approximate the Laplacian matrix $I-P$ with $I-\widetilde{P}$. Under what circumstances can we say that $I-\widetilde{P}^k$ is also a good approximation $I-P^k$ for (all or some) large integer k with the *same quality* of approximation? We address several notions of spectral approximation and we see that only the last one, called circular spectral approximation satisfies this property.

Definition 15.2. *For symmetric matrices $W, \widetilde{W} \in \mathbb{R}^{V\times V}$ we say $\widetilde{W}$ is an ϵ-approximation of W, $\widetilde{W} \approx_e \epsilon W$ for short, if for all $x \in \mathbb{R}^V$,*

$$(1-\epsilon)x^\top(I-W)x \le x^\top(I-\widetilde{W})x \le (1+\epsilon)x^T(I-W)x$$

or equivalently,

$$|x^\top(W-\widetilde{W})x| \le \epsilon x^\top(I-W)x = \epsilon\cdot(\|x\|^2 - x^\top Wx)$$

Unless otherwise specified, we assume $\epsilon < 1$. Furthermore, for the above definition to have a solution we need $x^\top Wx \le \|x\|$ for all $x \in \mathbb{R}^n$, i.e., that $\lambda_{\max}(W) \le 1$. This notion is perhaps the most well-known notion of spectral approximation. It was originally proposed by Spielman and Teng in their construction of spectral sparsifiers [ST11].

The above definition naturally extends to directed matrices as follows:

Definition 15.3 (Directed Spectral Approximation). *For (possibly asymmetric) matrices $W, \widetilde{W} \in \mathbb{R}^{n\times n}$ we say $\widetilde{W} \approx_\epsilon W$ if $\forall x, y \in \mathbb{R}^n$:*

$$\begin{aligned} |x^\top (W - \widetilde{W})y| &\le \frac{\epsilon}{2} \cdot (x^\top (I-W)x + y^\top (I-W)y| \\ &= \frac{\epsilon}{2} \cdot (\|x\|^2 + \|y\|^2 - x^\top W x - y^\top W y) \\ &= \frac{\epsilon}{2} \cdot (\|x\|^2 + \|y\|^2 - x^\top U_W x - y^\top U_W y). \end{aligned}$$

where $U_W = \frac{W+W^*}{2}$

Note that it is crucial to use two different vectors x, y for asymmetric matrices. In particular if $x = y$, then $x^\top W x = x^\top W^\top x$, so without loss of generality we can assume W is symmetric.

The reason we use the same notation for (directed) spectral approximation when the underlying matrix is symmetric is that the above two definitions are equivalent for symmetric matrices.

Lemma 15.4. *For (possibly asymmetric) matrices $W, \widetilde{W} \in \mathbb{R}^{n\times n}$, $\widetilde{W} \approx_\epsilon W$ iff*

- $\ker(U_W) \subseteq \ker(\widetilde{W} - W) \cap \ker((\widetilde{W} - W)^\top)$.
- $\left\| U_W^{\dagger/2} (\widetilde{W} - W) U_W^{\dagger/2} \right\| \le \epsilon$.

where for a matrix $M \in \mathbb{R}^{n\times n}$ we write $\|M\|$ to denote the spectral norm *of M, i.e., its largest singular value, namely $\|M\| = \max_{x\in\mathbb{R}^n} \frac{\|Mx\|}{\|x\|}$.*

Before we move forward let us explain a weakness of the above definitions: Suppose $W = \begin{pmatrix} a \end{pmatrix}$ and say $\widetilde{W} = \begin{pmatrix} b \end{pmatrix}$. Then, above definitions are equivalent to $|a - b| \le \epsilon(1-a)$. Notice when $a = 1$ then we must have $b = 1$; however if $a = -1$ then it is enough to have $b = -1 + \epsilon$. It is not hard to see that in such a case indeed $1 - b^2 = 1 - (-1+\epsilon)^2 = 2\epsilon - \epsilon^2$ is only a 2ϵ approximation of $1 - a^2 = 0$. The situation becomes worse as we look at $1 - b^k$ which only gives a $k\epsilon$ approximation of $1 - a^k$.

To put it in context, the spectral approximation defined above acts very poorly in bipartite graph. If W correspond to a random walk matrix of a disconnected graph $\widetilde{W}$ must have exactly the same connected components. However, if W is bipartite $\widetilde{W}$ does not have to be bipartite. This error becomes significant as we look at high (even) powers of the RW in which case W^{2k} is periodic, namely it never jumps from one side to the other whereas $\widetilde{W}^{2k}$ is not necessarily periodic. So although we may have $\widetilde{W} \approx_\epsilon W$ we can only hope $\widetilde{W}^k \approx_{k\epsilon} W$.

Motivated by this, one can suggest that $\widetilde{W}$ is a good approximation of W if, in addition, for every $x \in \mathbb{R}^n$, we have

$$x^\top (I+W)x \le x^\top (I+\widetilde{W})x \le (1+\epsilon)x^\top (I+W)x.$$

It turns out that such an extension is enough for symmetric matrices. But when the matrix is not symmetric we need more conditions:

Definition 15.5 (Unit-circle Spectral Approximation). *Let $W, \widetilde{W} \in \mathbb{C}^{n\times n}$ be possible asymmetric matrices. We say $\widetilde{W}$ is a unit-circle ϵ-approximation of W ($\widetilde{W} \overset{\circ}{\approx}_\epsilon W$) if for all $x, y \in \mathbb{C}^n$,*

$$|x^*(W - \widetilde{W})y| \le \frac{\epsilon}{2} \cdot (\|x\|^2 + \|y\|^2 - |x^* W x + y^* W y|).$$

To understand what we gain from this, suppose $x = y$ is an eigenvector of W with eigenvalue λ where $|\lambda| = 1$. Then, the RHS of the inequality is 0 so the LHS must also be 0, i.e., we must have $x^*Wx = x^*\widetilde{W}x$.

Note that for the above definition to make sense we need that $x^*Wx \leq \|x\|$ for all $x \in \mathbb{C}^n$. Such a condition holds for example when W is a stochastic matrix, $\|W\|_1 \leq 1$.

We leave the proof of the following lemma as an exercise.

Lemma 15.6. *Let $W, \widetilde{W} \in \mathbb{C}^{n\times n}$ be (possibly asymmetric) matrices. Then, the following are equivalent:*

1. $\widetilde{W} \overset{\circ}{\approx}_\epsilon W$,
2. *For all $z \in \mathbb{C}$ with $|z| = 1$, $z \cdot \widetilde{W} \approx_\epsilon z \cdot W$.*
3. *For all $z \in \mathbb{C}$ with $|z| = 1$,*
   - $\ker(U_{I-zW}) \subseteq \ker(\widetilde{W} - W) \cap \ker((\widetilde{W} - W)^\top)$ *and*
   - $\left\| U_{I-zW}^{\dagger/2} (\widetilde{W} - W) U_{I-zW}^{\dagger/2} \right\| \leq \epsilon$.

We note that equivalent of 1 and 2 is kind of immediate from the above discussions. Basically the main difficulty in unit-circle approximation is that the approximation should get tight as we get close to the unit circle in the complex plain. Such a statement is obviously invariant under rotation. Another consequence of the above lemma is that if $\widetilde{W} \overset{\circ}{\approx}_\epsilon W$ then for any $z \in \mathbb{C}$ with $|z| = 1$ we have $z \cdot \widetilde{W} \overset{\circ}{\approx}_\epsilon z \cdot W$.

To see equivalence to (3) one can use the following lemma with $M = N = U_{I-zW} = I - (zW + zW*)/2$:

Lemma 15.7. *For all matrices $A \in \mathbb{C}^{n\times n}$ and Hermitian PSD $M, N \in \mathbb{C}^{n\times n}$ the following are equivalent:*

1. *For all $x, y \in \mathbb{C}^n$,*
$$|x^*Ay| \leq \frac{\epsilon}{2}(x^*Mx + y^*Ny)$$
2. $\left\|M^{\dagger/2} A N^{\dagger/2}\right\| \leq \epsilon$ *and* $\ker(M) \subseteq \ker(A^*)$ *and* $\ker(N) \subseteq \ker(A)$.

15.2 Approximating Cycle-Lifted Graphs and Powers

We start with the following simple example:

Lemma 15.8. *For all rational $0 < \epsilon < 1$ there exists regular digraphs with transition matrices $\widetilde{W}, W$ such that $\widetilde{W} \approx_\epsilon W$ and $-\widetilde{W} \approx_\epsilon -W$ but $\widetilde{W}^4 \not\approx_c W^4$ for any finite c.*

Proof. Let C^4 be the transition matrix of a directed 4-cycle. Fix ϵ and define $\widetilde{C} = (1 - \epsilon/2)C_4 + (\epsilon/2)C_4^\top$, namely $\widetilde{C}$ is the directed 4-cycle with an $\epsilon/2$ probability of traversing backwards. We claim that $\widetilde{C} \approx_\epsilon C_4$ and $-\widetilde{C} \approx_\epsilon -C_4$ but $\widetilde{C}^4$ does not approximate C_4^4.

We just argue the latter and leave the formers as exercises: The point is C_4^4 has 4 connected components (all of the vertices become isolated with self-loops) while $\widetilde{C}^4$ has only two connected components. Therefore eigenvalue of 1 have different multiplicites in C_4^4 and $\widetilde{C}^4$ and that implies that $\widetilde{C}^4$ is not a unit-circle approximation of C_4. ∎

Definition 15.9 (Cycle-Shifted Graphs)**.** *Let C_k denote the transition probability matrix of the k-vertex directed cycle. Given a graph $G = (V, E)$ on n vertices with transition matrix P, the cycle lifted graph of length k, $C_k(G)$ is a layered graph with k layeres of n vertices each, where for every $i \in [k]$ there is an edge from vertex u in layer i to vertex v in layer $i + 1$ mod k iff (u, v) exits in G. To put it differenlty, $C_k(G) = C_k \otimes P$.*

Theorem 15.10. *Fix $W, \widetilde{W} \in C^{n\times n}$. Then, $C_k \otimes \widetilde{W} \approx_\epsilon C_k \otimes W$ iff for all $z \in \mathbb{C}$ such that $z^k = 1$, we have $z \cdot \widetilde{W} \approx_\epsilon z \cdot W$.*

We do not prove the above theorem; instead, let us give a little bit of intuition: First recall that eigenvalues of C_k are the k-th roots of unity: In particular, let $z \in \mathbb{C}$ such that $z^k = 1$ and define $\chi_z = [1, z, z^2, \ldots, z^{k-1}]$. Then, observe

$$C_k \chi_z = z \cdot \chi_z.$$

In addition, for any two (square) matrices A, B it is well known that eigenvalues of $A \otimes B$ are exactly the product of all pairs of eigenvalues of A, B, namely for an pair of eigenvalues λ_A, λ_B of A, B respectively, $\lambda_A \cdot \lambda_B$ is an eigenvalue of $A \otimes B$.

Corollary 15.11. *If $\widetilde{W} \overset{\circ}{\approx}_\epsilon W$ then for* all *integers k, $C_k \otimes \widetilde{W} \overset{\circ}{\approx}_\epsilon C_k \otimes W$.*

Proof. By Lemma 15.6 for any $\omega \in C$ with $|\omega| = 1$ and $z \in \mathbb{C} : z^k = 1$ we have $z \cdot \omega \cdot \widetilde{W} \overset{\circ}{\approx}_\epsilon z \cdot \omega \cdot W$. Therefore, by above theorem, $C_k \otimes \omega \cdot \widetilde{W} \approx_\epsilon C_k \otimes \omega \cdot W$. Finally, the statement follows by the fact that $C_k \otimes \omega \cdot \widetilde{W} = \omega \cdot (C_k \otimes \widetilde{W})$. □

Having that, let us justify that for $C_k \otimes \widetilde{W} \approx_\epsilon C_k \otimes W$ a necessary condition is that $z \cdot \widetilde{W} \approx_\epsilon z \cdot W$ for all k-th roots of unity z. This is because, as alluded to above, $\widetilde{W} \approx_\epsilon W$ only enforces tight condition around eigenvalues of W which are very close to 1. In particular, if W has an eigenvalues z for some $z \in \mathbb{C}$ that is a k-th root of unity, $\widetilde{W}$ may not have z as an eigenvalue. It suffices that $\widetilde{W}$ has an eigenvalue z' with $|z' - z| \le \epsilon$. But then, $z^{-1}\widetilde{W} \not\approx_\epsilon z^{-1}W$. This is W has an eigenvalue of 1 but $\widetilde{W}$ does not. Finally, since z^{-1} is an eigenvalue of C_k, we get that $C_k \otimes \widetilde{W} \not\approx_\epsilon C_k \otimes W$.

The following is the main theorem of this section:

Theorem 15.12. *Let $W, \widetilde{W}$ be transition probability matrices of graphs $G, \widetilde{G}$ such that $\widetilde{W} \overset{\circ}{\approx}_\epsilon W$ then for all integers $k \ge 1$, $\widetilde{W}^k \overset{\circ}{\approx}_{\epsilon/(1-3\epsilon/2)} W^k$.*

To prove the above theorem we need to express the k-th power of a matrix in terms of the Schur complement of its cycle-lifted graph of length k. First, we define the Schur complement.

Definition 15.13 (Schur Complement)**.** *For a matrix $A \in \mathbb{C}^{n\times n}$ and sets $S, T \subseteq [n]$, let $A_{S,T}$ denote the sub-matrix corresponding to the rows in S and columns in T. If S, T partition $[n]$ and $A_{T,T}$ is invertible, then we denote the Schur complement of A onto the set S by*

$$Sc(A, S) = A_{S,S} - A_{S,T} A_{T,T}^{-1} A_{T,S}.$$

For our application consider the case that $A = I - P$ is a (normalized) Laplacian matrix. Then,

$$I_S - Sc(I - P, S) = P_{S,S} - P_{S,\overline{S}}(I - P_{\overline{S},\overline{S}})^{-1} P_{\overline{S},S} = P_{S,S} - P_{S,\overline{S}} \left(\sum_{i=0}^{\infty} P^i_{\overline{S},\overline{S}} \right) P_{\overline{S},S}$$

Note that if $I - P_{\overline{S},\overline{S}}$ is invertible, then $\left\|P_{\overline{S},\overline{S}}\right\| < 1$ so the sum in RHS is convergent. The RHS can be seen as a shortcut graph: If when we jump out of S we shortcut the walk and jump back what would the corresponding transition probability matrix only restricted to S.

The following lemma shows that if $\widetilde{W} \approx_\epsilon W$ then so does any restriction of $\widetilde{W}$ to a subset S.

Lemma 15.14. *Let $W, \widetilde{W} \in \mathbb{C}^{n\times n}$ such that $\widetilde{W} \approx_\epsilon W$ for some $\epsilon < 2/3$. Let $T \subseteq [n]$ such that $I_T - W_{T,T}$ is invertible. Let $S = [n] - T$. Then,*

$$I_S - Sc(I - \widetilde{W}, S) \approx_{\epsilon/(1-3\epsilon/2)} I_S - Sc(I - W, S).$$

Proof of Theorem 15.12. Fix an arbitrary $z \in \mathbb{C}$ with $|z| = 1$. We need to show that $z\cdot\widetilde{W}^k \approx_{\epsilon/(1-3\epsilon/2)} z\cdot W^k$. Let $\omega \in \mathbb{C}$ such that $\omega^k = z$. Since $\widetilde{W} \overset{\circ}{\approx}_\epsilon W$, we have $\omega\cdot\widetilde{W} \overset{\circ}{\approx}_\epsilon \omega\cdot W$. Let $\widetilde{M} := C_k \otimes \omega\cdot\widetilde{W}$ and $M = C_k\otimes\omega\cdot W$. Therefore, by Corollary 15.11, we have that $\widetilde{M} \overset{\circ}{\approx}_\epsilon M$. Note that M can be thought of as the transition probability matrix of the cycle-lifted graph. Define S to be the set of vertices in the first layer of these cycle-lifted graphs. Let $\overline{S}$ be the rest of the vertices. Then,

$$I_S - Sc(I - M, S) = M_{S,S} + M_{S,\overline{S}}(I - M_{\overline{S},\overline{S}})^{-1}M_{\overline{S},S} = \sum_{i=0}^{\infty} M_{S,\overline{S}}M^i_{\overline{S},\overline{S}}M_{\overline{S},S} = (\omega W)^k_{S,S} = (z\cdot W^k)_{S,S}.$$

This is simply because $M_{S,S} = 0$ and terms inside the sum are non-zero only when $i = k-1$. Similarly, we have $I_S - S_c(I - M, S) = (z\cdot\widetilde{W}^k)_{S,S}$. Therefore, by Lemma 15.14 $z\cdot\widetilde{W}^k \approx_{\epsilon/(1-3\epsilon/2)} z\cdot W^k$. Since z was arbitrary we get that $\widetilde{W}^k \overset{\circ}{\approx}_\epsilon W^k$ as desired. □

15.3 Derandomized Squaring with respect to Circular Approximaiton

Definition 15.15 (Two-way Labelling and Rotation Maps)**.** *Let G be a d-regular directed multigraph on n vertices, namely every vertex has in-degree and out-degree d.*

A two-way labeling of G is a labeling of the edges in G such that

1. *Every edge (u, v) has two labels in $[d]$, one as an out-edge of u and one as an in-comming edge of v,*
2. *All in-comming edges of every vertex v has distinct labels and so are the out-going edges.*

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 leaving vertex v leads to vertex w and this edge has label j as an in-comming edge of w.*

The main non-trivial part of the de-randomized construction is the following definition of a derandomized square of a graph. Namely, given a random walk P we want to construct a matrix $\widetilde{P}$ that approximates P^2.

Definition 15.16 (Derandomized Square)**.** *Let G be a d-regular directed graph on n vertices with a two-way labeling. Let H be a c-regular (undirected) graph with d vertices. The* derandomized square *GⓈH is a $c\cdot d$-regular graph on n vertices with rotation map* $\mathrm{Rot}_{G Ⓢ H}$ *defined as follows: For a vertex $v_0 \in [n], i_0 \in [d]$, and $j_0 \in [c]$, we compute* $\mathrm{Rot}_{G Ⓢ H}(v_0, (i_0, j_0))$ *as*

1) Let $(v_1, i_1) = \mathrm{Rot}_G(v_0, i_0)$

2) Let $(i_2, j_1) = \text{Rot}_H(i_1, j_0)$

3) Let $(v_2, i_3) = \text{Rot}_G(v_1, i_2)$

4) Output $(v_2, (i_3, j_1))$

In the application we think of c as a fixed constant whereas d can be (polynomially) large in n.

In the square of a directed graph, for each vertex v, there exists a complete, uni-directional bipartite graph from the in-coming edges of v, to it's out-going edges. This corresponds to a directed edge for every two-step walk that has v in the middle of it. A useful way to view the derandomized square is that it replaces each of these complete bipartite graphs with a uni-directional bipartite expander. The following is the main theorem that we prove in this section.

Theorem 15.17. *Let G be a d-regular directed multigraph with random walk matrix P. Let H be a c-regular expander with $\lambda(H) \leq \epsilon$ and let $\widetilde{P}$ be the random walk matrix of $G \circledS H$. Then*

$$\widetilde{P} \overset{\circ}{\approx}_{2\epsilon} P^2.$$

Lemma 15.18. *Let $P \in \mathbb{R}^{n \times n}$ be the transition matrix of a* regular *(undirected) graph H with $\lambda^*(H) \leq \epsilon$, where $\lambda^* = \max\{\lambda_2, |\lambda_n|\}$ is the second largest eigenvalue in absolute value. Let J be the $n \times n$ matrix with $1/n$ in every entry. Then we have*

$$\begin{bmatrix} 0 & 0 \\ P & 0 \end{bmatrix} \overset{\circ}{\approx}_{\epsilon} \begin{bmatrix} 0 & 0 \\ J & 0 \end{bmatrix}$$

Proof. Let $x, y \in \mathbb{C}^{2n}$. Indeed we prove a stronger claim:

$$\left| x^* \begin{bmatrix} 0 & 0 \\ P - J & 0 \end{bmatrix} y \right| \leq \frac{\epsilon}{2} \left(\|x\|^2 + \|y\|^2 - 2 \left| x^* \begin{bmatrix} 0 & 0 \\ J & 0 \end{bmatrix} x + y^* \begin{bmatrix} 0 & 0 \\ J & 0 \end{bmatrix} y \right| \right)$$

For a vector $v \in \mathbb{C}^{2n}$, write v_1 for the first n coordinates of v and v_2 for the last n coordinates. Then, it is enough to show

$$|x_2^*(P - J)y_1| \leq \frac{\epsilon}{2}(\|x\|^2 + \|y\|^2 - \frac{2}{n}|\langle x_1, \mathbf{1}\rangle\langle \mathbf{1}, x_2\rangle| - \frac{2}{n}|\langle y_1, \mathbf{1}\rangle\langle \mathbf{1}, y_2\rangle|)$$

Let $\alpha_x = |\langle x_1, \mathbf{1}/\sqrt{n}\rangle|, \beta_x = |\langle x_2, \mathbf{1}\rangle|$ and let $x^\perp$ be the vector x projected onto the space orthogonal to vectors span($\mathbf{1}$) on the first n coordinates and span($\mathbf{1}$) on the second n coordinates. It follows that $\|x\| = \alpha_x^2 + \beta_x^2 + \|x^\perp\|^2$. Since $\alpha_x^2 + \beta_x^2 \geq 2\alpha_x\beta_x$ (and similarly for the y vector) the RHS is minimized when $\alpha_x = \alpha_y = \beta_x = \beta_y = 0$. So without loss of generality we assume $\langle x_1, \mathbf{1}\rangle = 0$ and $\langle \mathbf{1}, x_2\rangle = 0$ (and so for y). Then, it is enough to show

$$|x_2^*(P - J)y_1| = |x_2^* P y_1| \underset{\text{Cauchy}}{\leq} \|x_2\| \cdot \|W y_1\| \underset{\lambda^*(H) \leq \epsilon}{\leq} \epsilon \|x_2\| \|y_1\| \underset{\text{AM-GM}}{\leq} \frac{\epsilon}{2}(\|x_2\|^2 + \|y_1\|^2) \leq \frac{\epsilon}{2}(\|x\|^2 + \|y\|^2)$$

as desired. ∎

Fact 15.19. *Let G be a d-regular directed multigraph on n vertices with a two-way labeling and transition matrix P. Let H be a c-regular undirected graph on d vertices with a two-way labeling and transition matrix P_H. Let J be the $d \times d$ matrix with $1/d$ in every entry and let $\widetilde{P}$ be the transition matrix of $G \circledS H$. Define the $2d \times 2d$ matrices*

$$M = \begin{bmatrix} 0 & 0 \\ J & 0 \end{bmatrix}, \widetilde{M} = \begin{bmatrix} 0 & 0 \\ P_H & 0 \end{bmatrix}$$

Furthermore, for each vertex $u \in [n]$ define, $2d \times n$ matrix $T^{(u)}$ as follow

$$T^{(u)}_{i,v} = \begin{cases} 1 & \text{if } i\text{-th edge leaving } u \text{ is to } v \\ 0 & o.w. \end{cases} \qquad \forall i \in [d], v \in [n],$$

$$T^{(u)}_{d+i,v} = \begin{cases} 1 & \text{if } i\text{-th edge entering } u \text{ is from } v \\ 0 & o.w. \end{cases} \qquad \forall i \in [d], v \in [n],$$

Then,

$$P^2 = \frac{1}{d}\sum_u P^{(u)}, \qquad \widetilde{P} = \frac{1}{d}\sum_u \widetilde{P}^{(u)}$$

where $P^{(u)} := (T^{(u)})^\top M T^{(u)}$ and $\widetilde{P}^{(u)} := (T^{(u)})^\top \widetilde{M} T^{(u)}$.

We don't go over the details but the main observation is that for any vertex u, $P^{(u)}$ is exactly the transition matrix of the uni-directional bipartite complete graph from the in-neighbors of vertex u to its out-neighbors in P^2 (with all vertices that are not neighbors of u isolated). Likewise $\widetilde{P}^{(u)}$ is the transition matrix of the uni-directional bipartite expander from the in-neighbors of u to its out-neighbors in $G \circledS H$ (with non-neighbors of u isolated). The factor $\frac{1}{d}$ comes from the fact that every vertex has out-degree d.

Proof of Theorem 15.17. First, by Lemma 15.18, we can write, $\widetilde{M} \overset{\circ}{\approx}_\epsilon M$. For arbitrary $x, y \in \mathbb{C}^n$ write $\tilde{x} = (T^{(u)})^\top x$ and $\tilde{y} = (T^{(u)})^\top y$

$$\begin{aligned} |x^*(\widetilde{P}^{(u)} - P^{(u)})y| &= |\tilde{x}(\widetilde{M} - M)\tilde{y}| \\ &\underset{Lemma\ 15.18}{\leq} (\|\tilde{x}\|^2 + \|\tilde{y}\|^2 - 2|\tilde{x}^* M \tilde{x} + \tilde{y}^* M \tilde{y}|) \\ &= \left(x^*(T^{(u)})^\top T^{(u)} x + y^*(T^{(u)}) T^{(u)} y - 2|x^* P^{(u)} x + y^* P^{(u)} y\right) \end{aligned}$$

Note that the above inequality follows from the proof of Lemma 15.18 (not the specific lemma's statement). Summing up both sides over all vertex u we have

$$\sum_{u \in V} |x^*(\widetilde{P}^{(u)} - P^{(u)})y| \leq \frac{\epsilon}{2} \sum_{u \in V} \left(x^*(T^{(u)})^\top T^{(u)} x + y^*(T^{(u)}) T^{(u)} y - 2|x^* P^{(u)} x + y^* P^{(u)} y|\right)$$

Now, by the triangle inequality, the LHS is at least $d \cdot |x^*(\widetilde{P} - P^2)y|$ and the RHS is at most

$$\frac{\epsilon}{2}\left(x^* \sum_u (T^{(u)})^\top T^{(u)} x + y^* \sum_u (T^{(u)})^\top T^{(u)} y - 2d \cdot |x^* P^2 x + y^* P^2 y|\right)$$

Lastly, observe that

$$\sum_u (T^{(u)})^\top T^{(u)} = 2d \cdot I.$$

The reason is that $(T^{(u)})^\top$ has only one non-zero entry in every column, therefore $(T^{(u)})^\top T^{(u)}$ has non-zero entries only on its diagonal. In particular, it has the sum of the in-degree and out-degree of vertex u on the diagonal. Since G is d-regular, the sum is $2d \cdot I$. Putting these together we get

$$d \cdot |x^*(\widetilde{P} - P^2)y| \leq 2d \cdot \frac{\epsilon}{2}\left(\|x\|^2 + \|y\|^2 - |x^* P^2 x + y^* P^2 y|\right)$$

as desired. □

15.4 Approximating Inverse Laplacian

The following fact is immediate:

Fact 15.20. *Let G_0 be a regular directed graph on n vertices with a two-way labeling and let $H_1, \ldots, H_k$ be c-regular undirected graphs with two-way labelings where for each $i \in [k]$, H_i has $d \cdot c^{i-1}$ vertices. For each $i \in [k]$ let $G_i = G_{i-1} \circledS H_i$.*

Then, given $v \in [n], i_0 \in [d \cdot c^{i-1}], j_0 \in [c]$, $\mathrm{Rot}_{G_i}(v, (i_0, j_0))$ can be computed in space $O(\log(nd) + k \cdot \log c)$ with oracle queries to $\mathrm{Rot}_{H_1}, \ldots, \mathrm{Rot}_{H_k}$.

Running this procedure, we will get a matrix $\widetilde{P}$ which approximates P^{2^k}. So, in particular, we can compute any entry (s, t) of this matrix by counting the fraction $\mathrm{Rot}_{G_k}(s, (i_0, j_)) = t$.

So, combining with statements in previous section, given transition probability matrix P of G, we can find P_1 such that $P_1 \overset{\circ}{\approx}_\epsilon P^2$ and consequntly $P_1^k \overset{\circ}{\approx}_\epsilon P^{2k}$ for all k. Then, we find $P_2 \overset{\circ}{\approx}_\epsilon P_1^2$. Note that the unit circle approximation is not transitive; so we do not immediately get, $P_2 \overset{\circ}{\approx}_\epsilon P^4$. The method given below is a detour to improve the quality of approximation.

Let P_k be the transition probability matrix of a path of length k. First, they observe that for $L = I - P_k \otimes P$ we have

$$L^{-1} = \begin{bmatrix} I & P & P^2 & P^3 \\ 0 & I & P & P^2 \\ 0 & 0 & I & P \\ 0 & 0 & 0 & I. \end{bmatrix}$$

So, in particular, if we can compute the inverse if $I - P_k \otimes P$ exactly, then we can read off P^{k-1} from entries of that matrix. Unfortunately, even if P is transition probability matrix of an Eulerian graph, $P_k \otimes P$ is not Eulerian. Because of that they compute the inverse of the Laplacian of a cycle lifted graph, $C_k \otimes P$ and use that to estimate entries of P^{k-1} by approximating *escape probabilities* of the random walks.

For a random walk matrix P, rather than directly approximation inverse of $I - P$, they approximate the inverse of the Laplacian of the cycle lifted graph $I_{2^k n} - C_{2^k} \otimes P$ for k chosen larger than the mixing time of P. The point is that pseudo-inverse of $I - P$ can then be well-approximated by projection the pseudo-inverse of $I_{2^k n} - C_{2^k} \otimes P$. The approximate inverse of $I_{2^k n} - C_{2^k} \otimes P$, they exploit techniques from recent works on fast Laplacian solvers [CKKPPRS18]. The point is they compute a coarse approximation of the together with a Cholesky decomposition. Say we have a matrix $M \in \mathbb{R}^{n \times n}$ and let S, T be a partition of $[n]$. We can write a LU factorization of M as follows:

$$M = \begin{bmatrix} M_{T,T} & M_{T,S} \\ M_{S,T} & M_{S,S} \end{bmatrix} = \begin{bmatrix} I & 0 \\ M_{S,T} M_{T,T}^{-1} & I \end{bmatrix} \begin{bmatrix} M_{T,T} & 0 \\ 0 & M_{S,S} - M_{S,T} M_{T,T}^{-1} M_{T,S} \end{bmatrix} \begin{bmatrix} I & M_{T,T}^{-1} M_{T,S} \\ 0 & I \end{bmatrix}$$

Recall that $M_{S,S} - M_{S,T} M_{T,T}^{-1} M_{T,S} = Sc(M, S)$ is the Schur complement of M with respect to S. Having that it is straightforward to write the inverse of M:

$$M^{-1} = \begin{bmatrix} I & -M_{T,T}^{-1} M_{T,S} \\ 0 & I \end{bmatrix} \begin{bmatrix} M_{T,T}^{-1} & 0 \\ 0 & Sc(M,S)^{-1} \end{bmatrix} \begin{bmatrix} I & 0 \\ -M_{S,T} M_{T,T}^{-1} & I \end{bmatrix}$$

In other words, we just need to (recursively) compute the inverse of $M_{T,T}$ and $Sc(M, S)$.

Now, let us tailor this technique to our Laplacian matrix of a cycle shifted graph. Consider a cycle of length 2^k, C_{2^k} and arrange its rows and columns inductively such that

$$C_{2^k} = \begin{bmatrix} 0 & C_{2^{k-1}} \\ I_{2^{k-1}} & 0 \end{bmatrix}$$

Let $S = \{2^{k-1}n+1, \dots, 2^k n\}$ be the second half of the coordinates and let T be the (interval) coordinates. Observe that by above definition C_{2^k} alternates between S and T. It follows that

$$Sc(I_{2^k n} - C_{2^k} \otimes P, S) = I_{2^{k-1}n} - C_{2^{k-1}} \otimes P^2$$

So, we can write,

$$L = \begin{bmatrix} I_{2^{k-1}n} & 0 \\ -I_{2^{k-1}} \otimes P & I_{2^{k-1}n} \end{bmatrix} \begin{bmatrix} I_{2^{k-1}n} & 0 \\ 0 & I_{2^{k-1}n} - C_{2^{k-1}} \otimes P^2 \end{bmatrix} \begin{bmatrix} I_{2^{k-1}n} & -C_{2^{k-1}} \otimes P \\ 0 & I_{2^{k-1}n} \end{bmatrix}$$

Now, to invert the above matrix it is enough to compute the inverse of the $I_{2^{k-1}n} - C_{2^{k-1}} \otimes P^2$. This is exactly the Laplacian of the cycle-lifted graph with a cycle of length 2^{k-1} and transition probability matrix P^2. So, we can approximate P^2 using the derandomized squaring. Furthermore, since unit circle approximation is preserved under cycle lifts, we can well approximate $C_{2^{k-1}} \otimes P^2$. Recursively, we compute the inverse of the latter matrix. We just need to choose k big enough such that the random walk mixes by time 2^k. In such a case we case $P^{2^k} \approx J$ so we can easily approximate $(I - P^{2^k})^{-1}$.

Having a course approximation, finally, to boost the quality of approximation the use the following lemma:

Lemma 15.21. *Given matrices $A, B \in \mathbb{R}^{n \times n}$ such that $\|I - BA\| \le \alpha$ for some $\alpha > 0$. Let $P_m = \sum_{i=0}^{m} (I - BA)^i B$. Then,*

$$\|I - P_m A\| \le \alpha^{m+1}.$$

Proof. The proof simply follows from the fact that $I - PMA = (I - BA)^{m+1}$. □

To invoke the above lemma we let A be $I_{2^k n} - C_{2^k} \otimes P$ and B be our coarse approximation of it.

References

[AKMPSV20] A. Ahmadinejad, J. Kelner, J. Murtagh, J. Peebles, A. Sidford, and S. Vadhan. "High-precision Estimation of Random Walks in Small Space". In: *FOCS*. IEEE, 2020, pp. 1295–1306 (cit. on p. 15-2).

[BPS16] L. Boczkowski, Y. Peres, and P. Sousi. "Sensitivity of Mixing Times in Eulerian Digraphs". In: *SIAM Journal on Discrete Mathematics* 32.1 (Mar. 2016) (cit. on p. 15-1).

[CKKPPRS18] M. B. Cohen, J. A. Kelner, R. Kyng, J. Peebles, R. Peng, A. B. Rao, and A. Sidford. "Solving Directed Laplacian Systems in Nearly-Linear Time through Sparse LU Factorizations". In: *FOCS*. IEEE Computer Society, 2018, pp. 898–909 (cit. on pp. 15-2, 15-9).

[Rei08] O. Reingold. "Undirected connectivity in log-space". In: *Journal of the ACM* 55.4 (Sept. 2008), pp. 1–24 (cit. on pp. 15-1, 15-2).

[RTV06] O. Reingold, L. Trevisan, and S. P. Vadhan. "Pseudorandom walks on regular digraphs and the RL vs. L Problem". In: *STOC*. ACM, 2006, pp. 457–466 (cit. on p. 15-1).

[ST11] D. A. Spielman and S. Teng. "Spectral Sparsification of Graphs". In: *SIAM J. Comput.* 40.4 (2011), pp. 981–1025 (cit. on p. 15-3).

[ST14] D. A. Spielman and S. Teng. "Nearly Linear Time Algorithms for Preconditioning and Solving Symmetric, Diagonally Dominant Linear Systems". In: *SIAM J. Matrix Anal. Appl.* 35.3 (2014), pp. 835–885 (cit. on p. 15-2).