Recent Advances in Approximation Algorithms **Spring 2015**

Lecture 4-5: Effective Resistance and Simple Random Walks

Lecturer: Shayan Oveis Gharan *April 8th and 13th*

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

In this lecture we overview the connection of the effective resistance and simple random walks in a graph.

4.1 Electrical Flows

The notion of *electrical flows* arises naturally when we treat our graph as a resistor network. Given a graph $G = (V, E)$ with weights $w(.)$ on the edges, we replace each edge e with a resistance of resistor $1/w(e)$. In other words, think of $w(e)$ as the conductance of the edge e. We can then study how the electricity flows in this network.

Now, we write two underlying properties of electrical flows. The first one is the *flow conservation property*. Say we are sending one unit of flow from s to t. The flow conservation property says that for any vertex $v \neq s, t$, the sum of the flows into v is zero, this sum is $+1$ for s and -1 for t. For any edge (u, v) let $x(e)$ be the flow along edge e, that is $x(e)$ is non-negative if electricity is going from u to v and it is non-positive otherwise. Also, let $\delta^-(v)$ be the neighbors u of v where the edge (u, v) is oriented from u to v, and $\delta^+(v)$ be the rest of the neighbors of v. Then,

$$\sum_{u\in\delta^+(v)} x(e) - \sum_{u\in\delta^-(v)} x(e) = \begin{cases} +1 & \text{if } v = t \\ -1 & \text{if } v = s \\ 0 & \text{otherwise.} \end{cases}$$

We can rewrite the above equality as follows

$$B^\intercal x = b_{s,t}. \tag{4.1}$$

Recall that $b_{s,t}$ is the vector that is 1 in s, -1 in t and 0 otherwise.

The second property is the *Ohm's law*. This property implies that the electrical flows are potential flows. That is, if x is an electrical flow, then we can assign potentials to the vertices $p : V \to \mathbb{R}$ such that for any edge $e = (u, v)$,

$$x(e) = w(e) \cdot (p(u) - p(v)).$$

We use $W \in \mathbb{R}^{E\times E}$ to denote the diagonal matrix where for each edge e, $W(e, e) = w(e)$. We can rewrite the above equality as follows:

$$x = WBp. \tag{4.2}$$

Putting (4.1) and (4.2) together we get

$$B^\intercal WBp = L_G p = b_{s,t}, \tag{4.3}$$

or equivalently,

$$p = L_G^\dagger b_{s,t}. \tag{4.4}$$

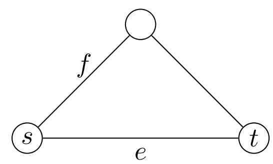

Figure 4.1: In this example if we send one unit of flow form s to t then 2/3 of the flow goes along the edge e and 1/3 goes through f, f'. Therefore, the potential difference between the endpoints of edges f is 1/3.

Note that since $\langle b_{s,t}, \mathbf{1} \rangle = 0$ and G is a connected graph, the linear system in (4.3) is uniquely determined and the solution is described above. The above argument naturally extends to any demand vector b as long as $\langle b, \mathbf{1} \rangle = 0$.

Let

$$Y = W^{1/2} B L_G^{\dagger} B^{\intercal} W^{1/2},$$

be the transfer-current matrix. This is the same matrix that we defined in the last lecture to show that a random spanning tree distribution is a determinantal probability measure. That is for any pair of edges e, f,

$$Y(e,f) = \langle y_e, y_f \rangle = b_e^{\intercal} L_G^{\dagger} b_f \sqrt{w(e)w(f)},$$

where $y_e = L_G^{\dagger/2} b_e \sqrt{w(e)}$. Recall that by (4.4), $L_G^{\dagger} b_f$ is the potential vector when we send one unit of electrical flow from one endpoint of f to the other endpoint. So, $b_e^{\intercal} L_G^{\dagger} b_f$ is the potential of difference of the endpoints of e when we send one unit of electrical flow between the endpoints of f. Also,

$$Y(e,f) \cdot \sqrt{w(f)/w(e)} = w(e) \cdot b_e^{\intercal} L_G^{\dagger} b_f,$$

is the current that flows between the endpoints of e when we send one unit of flow from one endpoint of f to the other endpoint.

Note that

$$Y(e,e) = w(e) b_e^{\intercal} L_G^{\dagger} b_e = \mathbb{P}_{T \sim \mu}[e \in T] \tag{4.5}$$

is the probability that e is chosen in a weighted uniform spanning tree, and by the above argument it is the current that flows along e when we send one unit of electrical flow between the endpoints of e. Let us give a simple example of a graph with 3 vertices.

Example 4.1. *In the last lecture we showed that in an unweighted for any pair of edges e, f,*

$$|\mathbb{P}[e, f \in T] - \mathbb{P}[e \in T] \cdot \mathbb{P}[f \in T]| = (b_e^{\intercal} L_G^{\dagger} b_f)^2.$$

By the above explanations, the RHS is the potential difference between the endpoints of f when we send one unit of flow between the endpoints of e. Now, consider the graph in Figure 4.1. First, observe that this graph has 3 spanning trees, only one of which has both e, f. Therefore, the LHS in the above is $|1/3 - 2/3 \cdot 2/3| = 1/9$. On the other hand, the potential difference between the endpoints of f when we send one unit of flow from s to t is $1/3$, so $b_e^{\intercal} L_G^{\dagger} b_f = 1/3$ and the RHS is also 1/9.

4.1.1 Energy

Say a flow $y : E \to \mathbb{R}$ is feasible if it satisfies the flow conservation property, i.e., (4.1). The *energy* of a (feasible) flow y is defined as follows:

$$\mathcal{E}(y) = \sum_{e \in E} \frac{y(e)^2}{w(e)} = y^{\intercal} W^{-1} y, \tag{4.6}$$

Note that since W is a diagonal matrix, its inverse is simply the inverse of every element in the diagonal.

The energy can be seen as the ℓ_2^2 norm of a flow. Analogously, one can define the ℓ_r norm of a feasible flow y as follows:

$$\left(\sum_{e \in E} \frac{y(e)^r}{w(e)}\right)^{1/r}.$$

It turns out that among all feasible flows that send one unit of flow from s to t, the electrical flow is the one with the smallest energy (or the smallest ℓ_2 norm). To put this into perspective, in the maximum flow problem one is looking for a feasible flow from s to t that with the smallest ℓ_∞ norm.

Lemma 4.2 (Thompson's Law)**.** *For any pair of vertices s, t, among all the flows that send one unit of flow from s to the t, the electrical flow has the smallest energy.*

Proof. Let x be the electrical flow. It follows by (4.2) and (4.4) that

$$x = WBL_G^\dagger b_{s,t}.$$

Therefore, the energy of x is equal to

$$\begin{aligned} \mathcal{E}(x) &= x^\intercal W^{-1} x \\ &= b_{s,t}^\intercal L_G^\dagger B^\intercal W W^{-1} W B L_G^\dagger b_{s,t} \\ &= b_{s,t}^\intercal L_G^\dagger b_{s,t}^\intercal. \end{aligned} \tag{4.7}$$

Now, let y be any feasible flow that sends one unit of flow from s to t, i.e., $B^\intercal y = b_{s,t}$. Therefore,

$$\mathcal{E}(x) = y^\intercal B L_G^\dagger B^\intercal y.$$

Now, all we need to show is that the above quantity is at most $y^\intercal W^{-1} y$. Say, $z = W^{-1/2} y$. We show

$$\mathcal{E}(x) = z^\intercal W^{1/2} B L_G^\dagger B^\intercal W^{1/2} z \leq z^\intercal z = y^\intercal W^{-1} y.$$

To show that above inequality it is enough to show that

$$W^{1/2} B L_G^\dagger B^\intercal W^{1/2} \preceq I.$$

This is proved in Lemma 4.3 below. □

4.1.2 Transfer Current Matrix

In this part we prove spectral properties of the transfer-current matrix. This will complete the proof of Lemma 4.2. In the next lemma, we show that the transfer-current matrix is a projection matrix.

Lemma 4.3 (Spielman and Srivastava [SS11])**.** *The matrix $Y = W^{1/2} B L_G^\dagger B^\intercal W^{1/2}$ is a projection matrix, i.e., all of its eigenvalues are 0 or 1.*

Proof. First, we show that Y is a projection matrix. Note that Y is a symmetric matrix. So, all we need to show that $YY^\intercal = Y$.

$$\begin{aligned} YY^\intercal &= W^{1/2} B L_G^\dagger B^\intercal W^{1/2} W^{1/2} B L_G^\dagger B^\intercal W^{1/2} \\ &= W^{1/2} B L_G^\dagger L_G L_G^\dagger B^\intercal W^{1/2} \\ &= W^{1/2} B L_G^\dagger B^\intercal W^{1/2} = Y. \end{aligned}$$

Therefore all of the eigenvalues of Y are 0 or 1. □

In the next lemma, we try to better understand the transfer-current matrix, in particular its eigenvalues/eigenvectors.

Lemma 4.4 (Spielman and Srivastava [SS11]). *The matrix Y satisfies $\text{im}(Y) = \text{im}(W^{1/2}B)$. Therefore, $n-1$ eigenvalues of Y are 1 and the rest are 0.*

Proof. It is perhaps easier to assume for a moment that $W = I$. Observe that $\text{im}(W^{1/2}B)$ are all of the potential flows in G (up to a normalization). We will see that these flows are also in the image of Y, and indeed Y acts as an identity on these family of flows.

First, note that by definition

$$\text{im}(W^{1/2}BL_G^{\dagger}B^{\intercal}W^{1/2}) \subseteq \text{im}(W^{1/2}B).$$

So, we need to show the converse. Let x be a vector in the image of $W^{1/2}B$, i.e., there is p such that $p \perp \ker(W^{1/2}B)$ and $W^{1/2}Bp = x$. Then,

$$\begin{aligned} Yx &= W^{1/2}BL_G^{\dagger}B^{\intercal}W^{1/2}(W^{1/2}Bp) \\ &= W^{1/2}BL_G^{\dagger}L_Gp \\ &= W^{1/2}Bp = x. \end{aligned}$$

In the second equality we used that $p \perp \ker(W^{1/2}B)$. Observe that the above equation shows that any potential flow x is an eigenvector of Y and in fact these are the only eigenvectors.

Therefore $\text{im}(Y) = \text{im}(W^{1/2}B)$. Now, since $\ker(W^{1/2}B)$ is 1-dimensional (it only has the constant vector), $\text{im}(W^{1/2}B)$ is $n-1$ dimensional, so $\text{im}(Y)$ is $n-1$ dimensional. Therefore, Y has exactly $n-1$ eigenvalues that are 1. □

4.2 Effective Resistance

The effective resistance between a pair of vertices s, t is defined as follows

$$\text{Reff}(s,t) = b_{s,t}^{\intercal} L_G^{\dagger} b_{s,t}. \tag{4.8}$$

By (4.4), $\text{Reff}(s,t)$ is the potential difference between s, t when we send one unit of electrical flow from s to t. Note that $\text{Reff}(s,t)$ is always non-negative because $L_G^{\dagger}$ is a PSD matrix. Equivalently, by (4.7), $\text{Reff}(s,t)$ is the energy of the electrical flow when we send one unit of flow from s to t.

The terminology of *effective resistance* originates from the following observation: If one removes all vertices of G except s, t and replaces the whole network with a resistance of resistor $\text{Reff}(s,t)$ between s, t, then, the energy (and the potential difference) of all electrical flows between s, t remains invariant.

The effective resistance of an edge $e = \{u, v\}$ is usually defined as the effective resistance between its endpoints. Note that if G is an unweighted graph, then the effective resistance of each edge is the probability that the edge is chosen in a uniform spanning tree distribution. In this case the diagonal of the transfer-current matrix Y has the effective resistance of all edges of G.

4.2.1 Properties of Effective Resistance

Lemma 4.5 (Metric Property). *For any triple of vertices s, t, u,*

$$\text{Reff}(s,t) + \text{Reff}(t,u) \geq \text{Reff}(s,u).$$

Proof. By (4.8),

$$\begin{aligned}\text{Reff}(s,u) &= b_{s,u}^{\intercal} L_G^{\dagger} b_{s,u} \\ &= (b_{s,t}+b_{t,u})^{\intercal} L_G^{\dagger}(b_{s,t}+b_{t,u}) \\ &= b_{s,t}^{\intercal} L_G^{\dagger} b_{s,t} + b_{t,u}^{\intercal} L_G^{\dagger} b_{t,u} + 2 b_{s,t}^{\intercal} L_G^{\dagger} b_{t,u} \\ &= \text{Reff}(s,t) + \text{Reff}(t,u) + 2 b_{s,t}^{\intercal} L_G^{\dagger} b_{t,u}.\end{aligned}$$

So, we just need to show that the last term in the RHS is non-positive. The last term is equal to $p(t) - p(u)$ when we send one unit of flow from s to t. But, this means that t has the lowest potential in the network, so $p(u) \geq p(t)$ as required. □

Lemma 4.6 (Monotonicity Property)**.** *For a weight function $w : E \to \mathbb{R}_+$ let $\text{Reff}_w(.,.)$ be the effective resistance function when the conductance of each edge $e \in E$ is $w(e)$. For any w, w' such that $w \leq w'$ and any $s, t \in V$,*

$$\text{Reff}_w(s,t) \geq \text{Reff}_{w'}(s,t).$$

Proof. Let x be the one unit electrical flow from s to t with respect to w. Since $w \leq w'$,

$$\text{Reff}_w(s,t) = \sum_{e\in E} \frac{x(e)^2}{w(e)} \geq \sum_{e\in E} \frac{x(e)^2}{w'(e)}.$$

Since x is a feasible flow that sends one unit of flow from s to t, by Lemma 4.2, the RHS is at least the energy of the electrical flow that sends one unit from s to t w.r.t. w'. □

As an application of the above lemma, we can give a different proof of negative correlation between the edges in a random spanning tree distribution, i.e., we show that

$$\mathbb{P}_{T\sim\mu}\left[e \in T | f \in T\right] \leq \mathbb{P}_{T\sim\mu}\left[e \in T\right].$$

The idea is that conditioning on $[.|f \in T]$ is the same as shortcutting the edge f in the electrical network, or equivalently, making the resistance of f zero, or letting $w(f) = \infty$. By the above lemma this operation decreases the effective resistance between each pair of vertices. But by (4.5), $\mathbb{P}[e \in T] = \text{Reff}(e)w(e)$, so shortcutting f decreases $\mathbb{P}[e \in T]$, i.e., e, f are negatively correlated.

Lemma 4.7 (Convexity)**.** *The effective resistance is convex w.r.t. the conductances and is concave w.r.t. resistances. In particular, for any s, t,*

$$\frac{1}{2}(\text{Reff}_{1/w_1}(s,t) + \text{Reff}_{1/w_2}(s,t)) \quad \leq \quad \text{Reff}_{2/(w_1+w_2)}(s,t). \tag{4.9}$$

$$\frac{1}{2}(\text{Reff}_{w_1}(s,t) + \text{Reff}_{w_2}(s,t)) \quad \geq \quad \text{Reff}_{(w_1+w_2)/2}(s,t), \tag{4.10}$$

Proof. We start by proving (4.9). Let x be the one unit electrical flow from s to t w.r.t. $(w_1 + w_2)/2$.

$$\begin{aligned}\mathcal{E}(x) = \text{Reff}_{2/(w_1+w_2)}(s,t) &= \sum_{e\in E} x(e)^2 \cdot \frac{w_1 + w_2}{2} \\ &= \frac{1}{2}\sum_{e\in E} x(e)^2 w_1(e) + \frac{1}{2}\sum_{e\in E} x(e)^2 w_2(e) \\ &\geq \frac{1}{2}(\text{Reff}_{1/w_1}(s,t) + \text{Reff}_{1/w_2}(s,t)).\end{aligned}$$

Next, we prove (4.10). Let x, y be the one unit electrical flows from s to t w.r.t. $1/w_1$ and $1/w_2$ respectively. Then,

$$\begin{aligned}\frac{1}{2}(\mathrm{Reff}_{w_1}(s,t) + \mathrm{Reff}_{w_2}(s,t)) &= \frac{1}{2}\sum_{e\in E}\frac{x(e)^2}{w_1} + \frac{1}{2}\sum_{e\in E}\frac{y(e)^2}{w_2}.\\ &\geq \sum_{e\in E}\left(\frac{x(e)+y(e)}{2}\right)^2 \frac{2}{w_1+w_2} \geq \mathrm{Reff}_{(w_1+w_2)/2}(s,t).\end{aligned}$$

In the last inequality we used that $(x+y)/2$ is a feasible flow that sends one unit from s to t. The only nontrivial part is the first inequality. To prove that it is sufficient to show for any edge e

$$\frac{x(e)^2}{w_1} + \frac{y(e)^2}{w_2} \geq \frac{(|x(e)| + |y(e)|)^2}{w_1 + w_2}.$$

Multiplying both sides with $2(w_1 + w_2)$ it is enough to show that

$$x(e)^2(1 + w_2/w_1) + y(e)^2(1 + w_1/w_2) \geq x(e)^2 + y(e)^2 + 2|x(e)y(e)|,$$

or equivalently,

$$x(e)^2\frac{w_2}{w_1} + y(e)^2\frac{w_1}{w_2} \geq 2|x(e)y(e)|.$$

Multiplying both sides with $1/|x(e)||y(e)|$ we need to show

$$\frac{|x(e)|w_2}{|y(e)|w_1} + \frac{|y(e)|w_1}{|x(e)|w_2} \geq 2.$$

But for any number $a > 0$, $a + 1/a \geq 2$, so the above equation follows. □

4.2.2 Resistive Embedding

One can define an n dimensional mapping of G, $F : V \to \mathbb{R}^n$, based on effective resistances. For any vertex v, let $F(v) := L_G^{\dagger/2}\mathbf{1}_v$. Then, for any pair of vertices u, v

$$\begin{aligned}\|F(u) - F(v)\|^2 &= \|L_G^{\dagger/2}(\mathbf{1}_u - \mathbf{1}_v)\|^2\\ &= b_{u,v}^{\intercal} L_G^{\dagger} b_{u,v} = \mathrm{Reff}(u,v).\end{aligned}$$

Therefore, by Lemma 4.5, the square of the distances in this mapping is also a metric. In other words, the resistive embedding is a L_2^2 metric.

Since one can optimize over L_2^2 squared metrics using an SDP, they have been one of the fundamental objects in many of the recent works in approximation algorithms, including the $O(\sqrt{\log(n)})$ approximation algorithm of Arora, Rao, Vazirani [ARV09] for the sparsest cut problem. Nonetheless, there have been very few applications of the resistive embedding in the theory literature (see [DLP11] for an application of resistive embedding in designing a deterministic algorithm for estimating the cover time of a graph).

Calculating the effective resistance mapping exactly requires the computation of the inverse of the Laplacian that can take cubic time. Spielman and Srivastava [SS11] used the fast Laplacian solvers [ST04] and the Johnson-Lindenstrauss dimension reduction argument [JL84] to show that the mapping can be approximated in near linear time.

4.2.3 Bounding the Effective Resistance

Lemma 4.8 (Nash Williams Inequality). *Let $S_1, S_2, \ldots, S_k \subseteq V$ such that for all $1 \le i \le k$, $s \in S_i, t \notin S_i$. If for all $1 \le i < j \le k$, $E(S_i, \overline{S_i}) \cap E(S_j, \overline{S_j}) = \emptyset$, then*

$$\text{Reff}(s,t) \ge \sum_{i=1}^{k} \frac{1}{\sum_{e \in E(S_i, \overline{S_i})} w(e)}.$$

Proof. Suppose x sends one unit of flow from s to t. We lower bound $\mathcal{E}(x)$ with the expression in the RHS. Since the cuts corresponding to $S_1, \ldots, S_k$ are disjoint, we can write

$$\mathcal{E}(x) \ge \sum_{i=1}^{k} \sum_{e \in E(S_i, \overline{S_i})} \frac{x^2(e)}{w(e)}.$$

Therefore, it is enough to show that for each $1 \le i \le k$,

$$\sum_{e \in E(S_i, \overline{S_i})} \frac{x(e)^2}{w(e)} \ge \frac{1}{\sum_{e \in E(S_i, \overline{S_i})} w(e)}. \tag{4.11}$$

Since $(S_i, \overline{S_i})$ separates s, t, $\sum_{e \in E(S_i, \overline{S_i})} |x(e)| \ge 1$. Therefore, by Cauchy-Schwarz inequality,

$$\begin{aligned} 1 &\le \left(\sum_{e \in E(S_i, \overline{S_i})} \frac{|x(e)|}{\sqrt{w(e)}} \cdot \sqrt{w(e)} \right)^2 \\ &\le \left(\sum_{e \in E(S_i, \overline{S_i})} \frac{x(e)^2}{w(e)} \right) \cdot \left(\sum_{e \in E(S_i, \overline{S_i})} w(e) \right). \end{aligned}$$

This proves (4.11) and completes the proof of the lemma. □

As a simple application of the above lemma we can show that in a $\sqrt{n} \times \sqrt{n}$ grid there is a pair of vertices s, t such that $\text{Reff}(s,t) \ge \Omega(\log(n))$.

Next, we discuss methods for upper bounding the effective resistance between a pair of vertices s, t. Note that to upper bound the effective resistance it is enough to construct a feasible flow that sends one unit of flow from s to t, then the energy of the flow will give an upper bound on $\text{Reff}(s,t)$.

Suppose there are k edge disjoint paths each of length at most ℓ from s to t. Then we can construct x by sending $1/k$ amount of flow on each path and

$$\mathcal{E}(x) \le \sum_{i=1}^{k} \sum_{e \in P_i} x(e)^2 = \sum_{i=1}^{k} \sum_{e \in P_i} \frac{1}{k^2} \le \frac{k\ell}{k^2}.$$

Unfortunately, we may not be able to find many edge disjoint paths between s, t even though $\text{Reff}(s,t)$ is small. For example, in a k-dimensional hypercube there are at most k edge disjoint paths between each pair of vertices because the degree of each vertex is k. But because the length of each path between $s = 00\ldots0$ and $t = 11\ldots1$ is at least k, the best upper bound that we can get is $O(1)$.

In the next lemma we give an idea to prove significantly better upper bounds on the effective resistance in some families of graphs.

Lemma 4.9 (Anari, Oveis Gharan [AO14]). *Given a d-regular unweighted graph $G = (V, E)$, if for any set $S \subseteq V$ such that $|S| \leq n/2$, $|E(S, \overline{S})| \geq cd|S|^{1/2+\epsilon}$, then for any pair of vertices s, t,*

$$\text{Reff}(s, t) \lesssim \frac{1}{\epsilon \cdot c^2 d}.$$

Proof. Let x be one unit electrical flow from s to t and let p be the potential vector, we assume $p(t) = 0$. So, all we need to do is to upper bound $p(s)$. Let us sort the vertices by their potential, that is let

$$S_r := \{v : p(v) \geq r\}.$$

We will show that for any r where $|S_r| < n/2$ and $f(z) := cdz^{1/2+\epsilon}$.

$$|S_{r-2/f(|S_r|)}| \geq |S_r| + \frac{f(|S_r|)}{2d}. \tag{4.12}$$

Iterating $2d|S_r|/f(|S_r|)$ times

$$|S_{r-4d|S_r|/f(|S_r|)^2}| \geq |S_r| + \frac{f(|S_r|)}{2d} \cdot \frac{2d|S_r|}{f(r)} = 2|S_r|.$$

In other words,

$$|S_{r-4|S_r|^{-2\epsilon}/c^2 d}| \geq 2|S_r|.$$

Running the telescoping sum, we conclude that if

$$r \leq p(s) - \sum_{i=1}^{\log(n)} \frac{4 \cdot 2^{-2i\epsilon}}{c^2 d} \lesssim \frac{1}{\epsilon c^2 d}$$

then $S_r \geq n/2$. By a similar argument at least half of the vertices half potential at most $O(1/\epsilon c^2 d)$. So, $p(s) \leq O(1/\epsilon c^2 d)$.

It remains to prove (4.12). Since $(S_r, \overline{S_r})$ is a cut separating s, t the sum of the flow on the edges in this cut is at least 1. But since every vertex in S_r has a higher potential than those not in S_r, none of the electrical flow is entering S_r; so

$$\sum_{e \in E(S_r, \overline{S_r})} |x(e)| = 1.$$

Therefore,

$$\mathbb{E}_{e \sim E(S_r, \overline{S_r})} [|x(e)|] = \frac{1}{|E(S_r, \overline{S_r})|},$$

where the expectation is over the uniform distribution the edges of $E(S_r, \overline{S_r})$. By the lemma's assumption if $|S_r| < n/2$, then $|E(S_r, \overline{S_r})| \leq f(|S_r|)$. So,

$$\mathbb{E}_{e \sim E(S_r, \overline{S_r})} [|x(e)|] \leq \frac{1}{f(|S_r|)}$$

By Markov's inequality, there is a set $F \subseteq E(S_r, \overline{S_r})$ such that $|F| \geq f(|S_r|)/2$ and for any $e \in F$

$$|x(e)| \leq 2/f(|S_r|).$$

Because the resistance of each edge is 1 the potential difference of the endpoints of every edge in F is at most $2/f(|S_r|)$. Therefore, all of the endpoints of edges of F are in $S_{r-2/f(r)}$. So,

$$|S_{r-2/f(|S_r|)}| \geq |S_r| + \frac{f(|S_r|)}{2d}.$$

This proves (4.12). $\square$

A consequence of the above lemma is that for any $d > 2$, the effective resistance between any pair of vertices in an d-dimensional grid is $O(1/d)$. This can be seen as a surprising fact, since the random walks do not mix rapidly in these families of graphs, i.e., they are not *expanders*. We will talk more about expanders towards the end of the course.

4.3 Random Walks

In this section we overview the connection of the electrical flows and simple random walks in graphs. Part of the materials of this section are based on the lecture notes of Amin Saberi [Sab11].

Let us start by writing a spectral connection between the effective resistance and target probability in a random walk. For simplicity, assume that G is a d-regular graph and note that in a d-regular graph

$$L_G = dI - A = d(I - P),$$

where $P = A/d$ is the transition probability matrix of the simple random walk on G. Let $\lambda_1 \geq \lambda_2 \geq \ldots \lambda_n$ be the eigenvalues of P with the corresponding eigenvectors $x_1, \ldots, x_n$. Observe that by the above equation L_G has the same eigenvectors with corresponding eigenvalues $d(1-\lambda_1), d(1-\lambda_2), \ldots, d(1-\lambda_n)$.

We can write,

$$\begin{aligned} L_G^\dagger &= \sum_{i=2}^{n} \frac{1}{d(1-\lambda_i)} x_i x_i^\intercal \\ &= \sum_{i=2}^{n} \sum_{j=0}^{\infty} \frac{\lambda_i^j}{d} x_i x_i^\intercal \\ &= \frac{1}{d} \sum_{j=0}^{\infty} \sum_{i=2}^{n} \lambda_i^j x_i x_i^\intercal \\ &= \frac{1}{d} \sum_{j=0}^{\infty} (P^j - \mathbf{1}\mathbf{1}^\intercal) \end{aligned}$$

In the last equality we used that for any $j \geq 0$, P^j has the same eigenvectors as P and with corresponding eigenvalues $\lambda_1^j, \ldots, \lambda_n^j$. Since $\mathbf{1} \perp b_{s,t}$, for a pair of vertices $s, t \in V$ we have

$$\text{Reff}(s,t) = \frac{1}{d} \sum_{j=0}^{\infty} b_{s,t}^\intercal P^j b_{s,t} \tag{4.13}$$

$$= \frac{1}{d} \sum_{j=1}^{\infty} (P^j(s,s) + P^j(t,t) - 2P^j(s,t)) \tag{4.14}$$

The above basic identity can be used to directly related the random walk transition probability matrix to the inverse of the Laplacian; we may get back to this later in the course.

Next, we talk about the connection of the potential vector and random walks. Suppose we enforce a potential difference of 1 between s, t, i.e., we send $1/\text{Reff}(s,t)$ electrical flow from s to t. Since the potential differences are invariant under shifting, we assume $p(s) = 1$ and $p(t) = 0$. Let $x = WBp$ be the corresponding electrical flow. Now, for any vertex $u \neq s, t$, the flow conservation property implies

$$\sum_{e \in \delta^+(u)} x(e) - \sum_{e \in \delta^-(u)} x_e = \sum_{v \sim u} p(u) - p(v) = 0,$$

Equivalently,

$$p(u) = \frac{1}{d(u)} \sum_{v \sim u} p(v).$$

It is easy to see that the above system of equations has a unique solution subject to the boundary conditions $p(s) = 1, p(t) = 0$.

Now, consider the following probabilities question. For any vertex $u \in V$, let $q_{s,t}(u)$ be the probability that a simple random walk started at u reaches s before t. By definition, $q(s) = 1, q(t) = 0$. Also, for any vertex u,

$$q_{s,t}(u) = \sum_{v \sim u} P(u, v) \cdot q_{s,t}(v) = \frac{1}{d(u)} \sum_{v \sim u} q_{s,t}(v).$$

Therefore, $q_{s,t}(u) = p(u)$ for all $u \in V$. In other words, the potential vector is the probability that a random walk started at a vertex u reaches the source of the electrical flow sooner than the sink. Note that the connections between electrical flows and random walks essentially follows from the fact that they both satisfy the same system of equations.

Next, we study a probabilistic quantity that is equal to the effective resistance.

Lemma 4.10. *For any edge $\{s,t\} \in E$, the probability that a random walk started at s visits t for the first time using the edge $\{s,t\}$ is equal to* $\text{Reff}(s,t)$.

Proof. First, as a sanity check note that if there is an edge between s,t then $0 \leq \text{Reff}(s,t) \leq 1$. Let z be the solution. Either in the first step we got directly to t, or we go to a neighbor v of s. In the latter case the event specified in the lemma occurs only with probability $q_{s,t}(v)z$, that is if we reach s before t and we recurse. So, we can write

$$\begin{aligned} z &= P(s,t) + \sum_{v \sim s} P(s,v) \cdot q_{s,t}(v) \cdot z \\ &= \frac{1}{d(s)} + \sum_{v \sim s} \frac{z}{d(s)} p(v), \end{aligned}$$

where $p(v)$ is the potential of v when $p(s) = 1$ and $p(t) = 0$. Therefore,

$$z = \frac{1/d(s)}{1 - \sum_{v \sim s} p(v)/d(s)} = \frac{1}{\sum_{v \sim s} 1 - p(v)} = \frac{1}{\sum_{v \sim s} p(s) - p(v)}$$

Let $x = WBp$ be the corresponding electrical flow for the potential vector p. Then, by Ohm's law,

$$z = \frac{1}{\sum_{e \sim s} x(e)} = \frac{1}{1/\text{Reff}(s,t)} = \text{Reff}(s,t).$$

In the second inequality we used that when we send $1/\text{Reff}(s,t)$ electrical flow from s to t the potential difference will be 1 and that $\sum_{e \sim s} x(e)$ is exactly the flow that is being sent to t. □

4.3.1 Cover Time and Effective Resistance

In this part we relate the effective resistance to the cover time of a simple random walk [Cha+96]. The expected *Heating time* $H(u,v)$ of a simple random walk between u,v is the expected number of steps a random walk starting at u takes before it visits vertex v for the first time. The *commute time*, $C(u,v)$ is the expected time to visit v and return back to u,

$$C(u,v) = H(u,v) + H(v,u).$$

The following theorem is proved by Chandra, Ragahavan, Ruzzo, Smolensky and Tiwari [Cha+96].

Theorem 4.11. *For any graph G and any pair of vertices s, t,*

$$C(s,t) = \text{Reff}(s,t) \cdot 2m.$$

It is not hard to prove the above theorem by superposition of two electrical flows, where one sends $d(v)$ units of flow from s to each vertex $v \in V$ and the other sends $d(v)$ units of flow from t to each vertex $v \in V$.

Tetali [Tet91] gave another proof of the above theorem by generalizing Lemma 4.10.

Theorem 4.12 (Tetali [Tet91])**.** *For any pair of vertices s, t and any edge (u, v) the expected number traversals from u to v in a simple random walk that goes from s to t and returns back to s is exactly* $\text{Reff}(s,t)$.

Observe that in Lemma 4.10 we proved the above theorem in the special case where $u = s$ and $v = t$. In other words the above theorem shows that the commute walk that goes from s to t and back to s visits every edge in each direction exactly equal number of times in expectation. That is true no matter how from (u, v) is from nodes s and t. Now, Theorem 4.11 directly follows from the above theorem.

As an application, we can use Theorem 4.11 to give another proof of Lemma 4.5. Say for a triple of vertices s, t, u we want to show

$$\text{Reff}(s,t) + \text{Reff}(t,u) \geq \text{Reff}(s,u)$$

By Theorem 4.11 this is equivalent to

$$C(s,t) + C(t,u) \geq C(s,u).$$

Now, by the definition of expected Hitting time this is equivalent to

$$H(s,t) + H(t,s) + H(t,u) + H(u,t) \geq H(s,u) + H(u,s).$$

Rearranging the terms we need to show

$$(H(s,t) + H(t,u)) + (H(u,t) + H(t,s)) \geq H(s,u) + H(u,s).$$

But, for an triple of vertices s, t, u we have

$$H(s,t) + H(t,u) \geq H(s,u).$$

The expected number of steps to go from s to u is smaller than the expected number of steps to go from s to u while visiting t along the way.

4.3.2 Sampling Random Spanning Trees using Random Walks

Next, we describe an algorithm of Broder and Aldous to sample random spanning trees [Bro89; Ald90] . Note that the running time of the above algorithm is equal to the cover time of the graph. In the worst case

Algorithm 1 Uniform Random Spanning Tree: The Random Walk Algorithm

Run a simple random walk started from an arbitrary vertex u in G until covering all vertices.
For every vertex $v \neq u$, add the edge that we used the first time that we reached v to T.
Return T.

the cover time can be as large as mn so the running time of the above algorithm is $O(mn)$. Wilson designed another algorithm and improved to running to the mean hitting time [Wil96] but still the worst case running time of his algorithm is $O(mn)$. Kelner and Madry [KM09], Propp [Pro10] and Madry, Straszak, Tarnawski [MST15] used Laplacian solvers machinery [ST04; KMP10; KMP11; Kel+13] and improved the running time to $\min\{\tilde{O}(m\sqrt{n}), \tilde{O}(m^{4/3})$. It is a fascinating open problem to design a near linear time algorithm for sampling a random spanning tree in unweighted graphs. Also, note that all of these random walk based algorithm may have an exponential running time in weighted graphs.

In the rest of this section we prove that the above algorithm chooses a uniform spanning tree and that we prove some consequences of this.

As an application of the above algorithm we give another proof of the Kirchoff's theorem, that is the effective resistance of each edge is equal to the probability that the edge is chosen in a uniform spanning tree.

Corollary 4.13. *For any unweighted graph $G = (V, E)$ and any edge $\{s, t\} \in E$,*

$$\mathbb{P}\left[\{s,t\} \in T\right] = \text{Reff}(s,t).$$

Proof. By Lemma 4.10, $\text{Reff}(s,t)$ is the probability that a random walk started at s reaches t for the first time by the edge $\{s,t\}$. Let us choose a random spanning tree by running Algorithm 1 from s. Then, edge $\{s,t\}$ is in the tree if and only if the walk reaches t for the first time by the edge $\{s,t\}$. The lemma follows. □

Next, we show that Algorithm 1 chooses a uniformly random spanning tree. For the sake of brevity we assume that G is unweighted and d-regular. We also assume that the walk is started at a uniformly random vertex of G. Let

$$Y = X_0, X_1, \ldots, X_t, \ldots$$

be a sample path of the random walk in G. Let $Y_t = X_0, \ldots, X_t$ be the first t steps of the above walk. Let us define two directed forests for this walk. In the *forward forest* $\text{FF}(Y_t)$, for any vertex v that is visited in this walk except the first vertex, we add a directed edge from v to the vertex that we entered v from which for the first time, i.e., if $t_v = \min\{\ell : X_\ell = v\}$, then we add $(X_{t_v}, X_{t_v - 1})$ to $\text{FF}(Y_t)$. It is easy to see that $\text{FF}(Y_t)$ does not have any cycles, so it is a forest. In addition, once the walk Y visits all of the vertices at least once, $\text{FF}(Y_t)$ is a directed spanning tree rooted at X_0, i.e., $\text{FF}(Y_t)$ is a rooted tree where every vertex except the root is pointing to its father.

Conversely, we define *backward forest*, $\text{BF}(Y_t)$, as follows: For any vertex v that is visited in the walk except the last vertex, we add a directed arc from v to the vertex that we visit right after v in our last visit to v, i.e., if $t_v = \max\{\ell : X_\ell = v\}$, then we add (X_{t_v}, X_{t_v+1}) to $\text{BF}(Y_t)$. Say $\text{rev}(Y_t)$ is the inverse of the walk Y_t; it is easy to see that

$$\text{BF}(\text{rev}(Y_t)) = \text{FF}(Y_t).$$

The main idea of the proof is as follows: When Y_t covers the whole graph, say time t^*, $\text{FF}(Y_{t^*}), \text{BF}(Y_{t^*})$ will be rooted spanning trees of G. Now, for any $t \geq t^*$, $\text{FF}(Y_t) = \text{FF}(Y_{t^*})$ while $\text{BF}(Y_t)$ is changing. We will show that $\text{BF}(Y_t)$ is a Markov chain on all rooted spanning trees of G with uniform stationary distribution. This means that as $t \to \infty$, $\text{BF}(Y_t)$ will be a uniform rooted spanning tree of G. But because of the reversibility of the random walk and that X_0 is chosen uniformly at random $\mathbb{P}[Y_t] = \mathbb{P}[\text{rev}(Y_t)]$. Therefore, $\text{BF}(\text{rev}(Y_t))$ is identically distributed as $\text{BF}(Y_t)$. But as alluded to in the above, $\text{BF}(\text{rev}(Y_t)) = \text{FF}(Y_t)$. So, $\text{FF}(Y_t)$ is a uniform rooted spanning tree of G. Dropping the directions of edges, $FF(Y_t)$ is a uniform spanning tree of G.

Stationary Distribution and Irreducibility of the Rooted Tree Markov Chain. It remains to show that $\text{BF}(t)$ converges to uniform distribution on all spanning trees of G. We show that for any rooted tree T we go to d rooted trees each with probability $1/d$ and d rooted trees come to T each with probability

$1/d$. This implies that the stationary distribution of the chain is uniform. Say T is a rooted spanning trees with root r. The set of directed trees reachable form T in one step are as follows: First choose a neighbor of r, say v, then add the directed arc (r, v) to T and remove the unique arc that is leaving v. Since G is d-regular we can transit from T to d possible rooted trees each with probability $1/d$. Conversely, to obtain the set of trees that transit to T in one step, first we choose a uniformly random neighbor of r say v. Then we add the directed arc (r, v) to T and we remove the arc point to r in the unique path from v to r, say (u, r).

Finally, it remains to show that the Markov chain is irreducible. i.e., it is possible to reach every rooted spanning tree from any given rooted spanning tree. Say r_1 is the root of T_1 and r_2 is the root of T_2. First we observe that that starting from r_1 any walk that leads to r_2 takes us to a tree, say T_1' with root r_2. Now consider the the depth first search walk starting from r_2 that goes along the tree, visits all of the leaves of T_2 and returns back to r_2. This walk transits T_1' to T_2.

References

[Ald90] D. J. Aldous. “A random walk construction of uniform spanning trees and uniform labelled trees”. In: *SIAM Journal on Discrete Mathematics* 3.4 (1990), pp. 450–465 (cit. on p. 4-11).

[AO14] N. Anari and S. Oveis Gharan. unpublished. 2014 (cit. on p. 4-8).

[ARV09] S. Arora, S. Rao, and U. Vazirani. “Expander flows, Geometric Embeddings and Graph Partitioning”. In: *J. ACM* 56 (2 Apr. 2009), 5:1–5:37 (cit. on p. 4-6).

[Bro89] A. Broder. “Generating random spanning trees”. In: *FOCS*. 1989, pp. 442–447 (cit. on p. 4-11).

[Cha+96] A. Chandra, P. Raghavan, W. Ruzzo, R. Smolensky, and P. Tiwari. “The electrical resistance of a graph captures its commute and cover times”. In: *computational complexity* 6.4 (1996), pp. 312–340 (cit. on p. 4-10).

[DLP11] J. Ding, J. R. Lee, and Y. Peres. “Cover times, blanket times, and majorizing measures”. In: *STOC*. 2011, pp. 61–70 (cit. on p. 4-6).

[JL84] W. B. Johnson and J. Lindenstrauss. “Extensions of Lipschitz mappings into a Hilbert space”. In: *Conference in modern analysis and probability (New Haven, Conn., 1982)*. Providence, RI: Amer. Math. Soc., 1984, pp. 189–206 (cit. on p. 4-6).

[Kel+13] J. A. Kelner, L. Orecchia, A. Sidford, and Z. A. Zhu. “A simple, combinatorial algorithm for solving SDD systems in nearly-linear time”. In: *STOC*. 2013, pp. 911–920 (cit. on p. 4-12).

[KM09] J. Kelner and A. Madry. “Faster generation of random spanning trees”. In: *FOCS*. 2009 (cit. on p. 4-12).

[KMP10] I. Koutis, G. L. Miller, and R. Peng. “Approaching Optimality for Solving SDD Linear Systems”. In: *FOCS*. 2010, pp. 235–244 (cit. on p. 4-12).

[KMP11] I. Koutis, G. L. Miller, and R. Peng. “A Nearly-m log n Time Solver for SDD Linear Systems”. In: *FOCS*. 2011, pp. 590–598 (cit. on p. 4-12).

[MST15] A. Madry, D. Straszak, and J. Tarnawski. “Fast Generation of Random Spanning Trees and the Effective Resistance Metric”. In: *SODA*. 2015, pp. 2019–2036 (cit. on p. 4-12).

[Pro10] J. Propp. Personal communication. 2010 (cit. on p. 4-12).

[Sab11] A. Saberi. *Lecture Notes of Discrete Mathematics and Algorithms*. 2011 (cit. on p. 4-9).

[SS11] D. A. Spielman and N. Srivastava. “Graph Sparsification by Effective Resistances”. In: *SIAM J. Comput.* 40.6 (2011), pp. 1913–1926 (cit. on pp. 4-3, 4-4, 4-6).

[ST04] D. A. Spielman and S.-H. Teng. “Nearly-linear time algorithms for graph partitioning, graph sparsification, and solving linear systems”. In: *STOC*. 2004, pp. 81–90 (cit. on pp. 4-6, 4-12).

[Tet91] P. Tetali. “Random walks and the effective resistance of networks”. In: *Journal of Theoretical Probability* 4.1 (Jan. 1991), pp. 101–109 (cit. on p. 4-11).

[Wil96] D. B. Wilson. “Generating random spanning trees more quickly than the cover time”. In: *Proceedings of the Twenty-eighth Annual ACM Symposium on the Theory of Computing*. ACM, 1996, pp. 296–303 (cit. on p. 4-11).