CMSE 890-001: Spectral Graph Theory and Related Topics, MSU, Spring 2021

Lecture 25: Graph Sparsification via Random Sampling, Part II

April 15, 2021

Lecturer: Matthew Hirn

30.4 The key transformation

From Section 30.2 we know that $\mathbb{E}[\boldsymbol{L}_H] = \boldsymbol{L}_G$. We could try to write $\boldsymbol{L}_H$ as the sum of random matrices and apply Theorem 61 with $\boldsymbol{X} = \boldsymbol{L}_H$ and $\mathbb{E}[\boldsymbol{X}] = \mathbb{E}[\boldsymbol{L}_H] = \boldsymbol{L}_G$, but to make our life easier we will first make an important transformation into an equivalent problem for which $\mu_{\min} = \mu_{\max} = 1$.

For positive definite matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ we have

$$\boldsymbol{A} \preccurlyeq c\boldsymbol{B} \quad \Longleftrightarrow \quad \boldsymbol{B}^{-1/2}\boldsymbol{A}\boldsymbol{B}^{-1/2} \preccurlyeq c\boldsymbol{I}\,.$$

In the above, recall that

$$\boldsymbol{B} = \boldsymbol{\Phi}\boldsymbol{\Lambda}\boldsymbol{\Phi}^T\,,$$

where $\boldsymbol{\Lambda}$ is the diagonal matrix with the eigenvalues $\lambda_1, \ldots, \lambda_n$ of $\boldsymbol{B}$ on its diagonal. Since $\boldsymbol{B}$ is positive definite, $\lambda_i > 0$ for all $1 \leq i \leq n$, and we define

$$\boldsymbol{B}^{-1/2} := \boldsymbol{\Phi}\boldsymbol{\Lambda}^{-1/2}\boldsymbol{\Phi}^T\,, \quad \boldsymbol{\Lambda}^{-1/2}(i,i) := 1/\sqrt{\lambda_i}\,.$$

The same property holds for positive semi-definite matrices so long as they have the same null space, and if we replace the inverse of $\boldsymbol{B}$ with the pseudo-inverse of $\boldsymbol{B}$. More precisely, let $\boldsymbol{B}$ be a positive semi-definite matrix. Then the pseudo-inverse of $\boldsymbol{B}$ is defined as:

$$\boldsymbol{B}^{+} := \boldsymbol{\Phi}\boldsymbol{\Lambda}^{+}\boldsymbol{\Phi}^T\,, \quad \boldsymbol{\Lambda}^{+}(i,i) := \begin{cases} 0 & \lambda_i = 0 \\ \lambda_i^{-1} & \lambda_i > 0 \end{cases}\,.$$

Similarly, we define $\boldsymbol{B}^{+/2}$ as:

$$\boldsymbol{B}^{+/2} := \boldsymbol{\Phi}\boldsymbol{\Lambda}^{+/2}\boldsymbol{\Phi}^T\,, \quad \boldsymbol{\Lambda}^{+/2}(i,i) := \begin{cases} 0 & \lambda_i = 0 \\ \lambda_i^{-1/2} & \lambda_i > 0 \end{cases}\,.$$

Now, if $\boldsymbol{A}$ and $\boldsymbol{B}$ are positive semi-definite matrices with the same null space, then

$$\boldsymbol{A} \preccurlyeq c\boldsymbol{B} \quad \Longleftrightarrow \quad \boldsymbol{B}^{+/2}\boldsymbol{A}\boldsymbol{B}^{+/2} \preccurlyeq c\boldsymbol{B}^{+/2}\boldsymbol{B}\boldsymbol{B}^{+/2}\,.$$

In particular, if H and G are both connected graphs, then both $\boldsymbol{L}_H$ and $\boldsymbol{L}_G$ have the same null space given by $\{\alpha\mathbf{1} : \alpha \in \mathbb{R}\}$. As such:

$$\boldsymbol{L}_H \preccurlyeq (1+\epsilon)\boldsymbol{L}_G \quad \Longleftrightarrow \quad \boldsymbol{L}_G^{+/2}\boldsymbol{L}_H\boldsymbol{L}_G^{+/2} \preccurlyeq (1+\epsilon)\boldsymbol{L}_G^{+/2}\boldsymbol{L}_G\boldsymbol{L}_G^{+/2}\,.$$

Set

$$\boldsymbol{\Pi} := \boldsymbol{L}_G^{+/2} \boldsymbol{L}_G \boldsymbol{L}_G^{+/2} .$$

We conclude that

$$(1-\epsilon)\boldsymbol{L}_G \preccurlyeq \boldsymbol{L}_H \preccurlyeq (1+\epsilon)\boldsymbol{L}_G \quad \Longleftrightarrow \quad (1-\epsilon)\boldsymbol{\Pi} \preccurlyeq \boldsymbol{L}_G^{+/2} \boldsymbol{L}_H \boldsymbol{L}_G^{+/2} \preccurlyeq (1+\epsilon)\boldsymbol{\Pi} .$$

In other words, $\boldsymbol{L}_H$ is an ϵ-approximation of $\boldsymbol{L}_G$ if and only if $\boldsymbol{L}_G^{+/2} \boldsymbol{L}_H \boldsymbol{L}_G^{+/2}$ is an ϵ-approximation of $\boldsymbol{\Pi}$. So rather than prove that $\boldsymbol{L}_H$ is an ϵ-approximation of G with high probability, we will instead prove that $\boldsymbol{L}_G^{+/2} \boldsymbol{L}_H \boldsymbol{L}_G^{+/2}$ is an ϵ-approximation of $\boldsymbol{\Pi}$ with high probability.

At this point, it is instructive to get a feel for what the operator $\boldsymbol{\Pi}$ does. In words, it projects a vector $\boldsymbol{x}$ onto the range $\boldsymbol{L}_G$. Since G is connected, the range of $\boldsymbol{L}_G$ is given by $\text{span}\{\boldsymbol{\psi}_2, \ldots, \boldsymbol{\psi}_n\}$. To see that $\boldsymbol{\Pi}$ projects $\boldsymbol{x}$ onto $\text{span}\{\boldsymbol{\psi}_2, \ldots, \boldsymbol{\psi}_n\}$, we compute

$$\boldsymbol{\Pi} = \boldsymbol{\Psi}\boldsymbol{\Lambda}^{+/2}\boldsymbol{\Psi}^T\boldsymbol{\Psi}\boldsymbol{\Lambda}\boldsymbol{\Psi}^T\boldsymbol{\Psi}\boldsymbol{\Lambda}^{+/2}\boldsymbol{\Psi}^T = \boldsymbol{\Psi}\boldsymbol{\Lambda}^{+/2}\boldsymbol{\Lambda}\boldsymbol{\Lambda}^{+/2}\boldsymbol{\Psi}^T = \boldsymbol{\Psi} \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix} \boldsymbol{\Psi}^T .$$

As such,

$$\boldsymbol{\Pi}\boldsymbol{x} = \sum_{i=2}^{n} \langle \boldsymbol{x}, \boldsymbol{\psi}_i \rangle \boldsymbol{\psi}_i ,$$

and we note that $\lambda_1(\boldsymbol{\Pi}) = 0$ and $\lambda_2(\boldsymbol{\Pi}) = \cdots = \lambda_n(\boldsymbol{\Pi}) = 1$.

30.5 The edge probabilities and graph approximation

In this section we prove that $\boldsymbol{L}_G^{+/2} \boldsymbol{L}_H \boldsymbol{L}_G^{+/2}$ is an ϵ-approximation of $\boldsymbol{\Pi}$ with high probability. In other words, we want to show:

$$\mathbb{P}\left(\boldsymbol{L}_G^{+/2} \boldsymbol{L}_H \boldsymbol{L}_G^{+/2} \preccurlyeq (1-\epsilon)\boldsymbol{\Pi}\right) \ll 1 , \tag{75}$$

$$\mathbb{P}\left(\boldsymbol{L}_G^{+/2} \boldsymbol{L}_H \boldsymbol{L}_G^{+/2} \succcurlyeq (1+\epsilon)\boldsymbol{\Pi}\right) \ll 1 . \tag{76}$$

Note, if $\boldsymbol{L}_G^{+/2} \boldsymbol{L}_H \boldsymbol{L}_G^{+/2} \succcurlyeq (1+\epsilon)\boldsymbol{\Pi}$, then by Theorem 15 we have

$$\lambda_n(\boldsymbol{L}_G^{+/2} \boldsymbol{L}_H \boldsymbol{L}_G^{+/2}) \geq (1+\epsilon)\lambda_n(\boldsymbol{\Pi}) = 1+\epsilon .$$

Thus,

$$\mathbb{P}\left(\boldsymbol{L}_G^{+/2} \boldsymbol{L}_H \boldsymbol{L}_G^{+/2} \succcurlyeq (1+\epsilon)\boldsymbol{\Pi}\right) \leq \mathbb{P}\left(\lambda_n(\boldsymbol{L}_G^{+/2} \boldsymbol{L}_H \boldsymbol{L}_G^{+/2}) \geq 1+\epsilon\right) .$$

As such, to show (76) it is sufficient to show

$$\mathbb{P}\left(\lambda_n(\boldsymbol{L}_G^{+/2} \boldsymbol{L}_H \boldsymbol{L}_G^{+/2}) \geq 1+\epsilon\right) \ll 1 .$$

Similarly, to show (75) it is sufficient to show

$$\mathbb{P}\left(\lambda_2(\boldsymbol{L}_G^{+/2}\boldsymbol{L}_H\boldsymbol{L}_G^{+/2}) \leq 1-\epsilon\right) \ll 1\,.$$

Note that we use λ_2 instead of λ_1 since we are guaranteed to have $\lambda_1(\boldsymbol{L}_G^{+/2}\boldsymbol{L}_H\boldsymbol{L}_G^{+/2}) = \lambda_1(\boldsymbol{\Pi}) = 0$.

At this point you should be thinking that we will want to use Theorem 61. To do so we need to do two things. First, we need to compute $\mathbb{E}[\boldsymbol{L}_G^{+/2}\boldsymbol{L}_H\boldsymbol{L}_G^{+/2}]$. We have:

$$\mathbb{E}[\boldsymbol{L}_G^{+/2}\boldsymbol{L}_H\boldsymbol{L}_G^{+/2}] = \boldsymbol{L}_G^{+/2}\mathbb{E}[\boldsymbol{L}_H]\boldsymbol{L}_G^{+/2} = \boldsymbol{L}_G^{+/2}\boldsymbol{L}_G\boldsymbol{L}_G^{+/2} = \boldsymbol{\Pi}\,.$$

That is great because it means that on average $\boldsymbol{L}_G^{+/2}\boldsymbol{L}_H\boldsymbol{L}_G^{+/2}$ is equal to $\boldsymbol{\Pi}$, and we will be able to use Theorem 61 to prove that any individual realization of $\boldsymbol{L}_G^{+/2}\boldsymbol{L}_H\boldsymbol{L}_G^{+/2}$ should not deviate from $\boldsymbol{\Pi}$ by too much, with high probability. We note that,

$$1+\epsilon = (1+\epsilon)\lambda_n(\boldsymbol{\Pi}) = (1+\epsilon)\lambda_n\left(\mathbb{E}[\boldsymbol{L}_G^{+/2}\boldsymbol{L}_H\boldsymbol{L}_G^{+/2}]\right)\,,$$

and similarly,

$$1-\epsilon = (1-\epsilon)\lambda_2(\boldsymbol{\Pi}) = (1-\epsilon)\lambda_2\left(\mathbb{E}[\boldsymbol{L}_G^{+/2}\boldsymbol{L}_H\boldsymbol{L}_G^{+/2}]\right)\,.$$

The second thing we need to do is write $\boldsymbol{L}_G^{+/2}\boldsymbol{L}_H\boldsymbol{L}_G^{+/2}$ as the sum of random matrices. To that end, for each $(a,b) \in E$ define $\boldsymbol{X}_{a,b}$ as the random matrix such that

$$\mathbb{P}\left(\boldsymbol{X}_{a,b} = \frac{w(a,b)}{p_{a,b}}\boldsymbol{L}_G^{+/2}\boldsymbol{L}_{a,b}\boldsymbol{L}_G^{+/2}\right) = p_{a,b} \quad \text{and} \quad \mathbb{P}(\boldsymbol{X}_{a,b} = \boldsymbol{0}) = 1 - p_{a,b}\,.$$

Recalling how we defined H in (73) and (74), we see that

$$\boldsymbol{L}_G^{+/2}\boldsymbol{L}_H\boldsymbol{L}_G^{+/2} = \sum_{(a,b)\in E} \boldsymbol{X}_{a,b}\,.$$

Now we need to specify the probabilities $p_{a,b}$

$$p_{a,b} := \frac{1}{R}w(a,b)\|\boldsymbol{L}_G^{+/2}\boldsymbol{L}_{a,b}\boldsymbol{L}_G^{+/2}\|\,,$$

where R is a parameter that we will specify shortly. Since

$$\boldsymbol{X}_{a,b} = \frac{w(a,b)}{p_{a,b}}\boldsymbol{L}_G^{+/2}\boldsymbol{L}_G\boldsymbol{L}_G^{+/2}\,,$$

when $(a,b) \in E$ is chosen, we see that

$$\|\boldsymbol{X}_{a,b}\| \leq R\,.$$

Now, there is a chance that $p_{a,b} > 1$. We will address that scenario at the end of our discussion. For now let us assume that $p_{a,b} \leq 1$.

Set

$$R := \frac{\epsilon^2}{\tau \log n}, \quad \tau > 3.$$

Applying Theorem 61 we have:

$$\begin{aligned}
\mathbb{P}\left(\boldsymbol{L}_G^{+/2}\boldsymbol{L}_H\boldsymbol{L}_G^{+/2} \not\preccurlyeq (1+\epsilon)\boldsymbol{\Pi}\right) &\leq \mathbb{P}\left(\lambda_n(\boldsymbol{L}_G^{+/2}\boldsymbol{L}_H\boldsymbol{L}_G^{+/2}) \geq 1+\epsilon\right) \\
&= \mathbb{P}\left(\lambda_n(\boldsymbol{L}_G^{+/2}\boldsymbol{L}_H\boldsymbol{L}_G^{+/2}) \geq (1+\epsilon)\lambda_n(\mathbb{E}[\boldsymbol{L}_G^{+/2}\boldsymbol{L}_H\boldsymbol{L}_G^{+/2}])\right) \\
&\leq n\exp\left(-\frac{\epsilon^2}{3R}\right) \\
&= n\exp\left(-\frac{\tau\log n}{3}\right) \\
&= n \cdot n^{-\tau/3} \\
&= n^{-(\tau-3)/3}.
\end{aligned}$$

Since $\lambda_1(\boldsymbol{X}_{a,b}) = 0$ for all $(a,b) \in E$, we can apply Theorem 61 to λ_2 instead of λ_1 to obtain:

$$\begin{aligned}
\mathbb{P}\left(\boldsymbol{L}_G^{+/2}\boldsymbol{L}_H\boldsymbol{L}_G^{+/2} \not\succcurlyeq (1-\epsilon)\boldsymbol{\Pi}\right) &\leq \mathbb{P}\left(\lambda_2(\boldsymbol{L}_G^{+/2}\boldsymbol{L}_H\boldsymbol{L}_G^{+/2}) \leq 1-\epsilon\right) \\
&= \mathbb{P}\left(\lambda_2(\boldsymbol{L}_G^{+/2}\boldsymbol{L}_H\boldsymbol{L}_G^{+/2}) \leq (1-\epsilon)\lambda_2(\mathbb{E}[\boldsymbol{L}_G^{+/2}\boldsymbol{L}_H\boldsymbol{L}_G^{+/2}])\right) \\
&\leq n\exp\left(-\frac{\epsilon^2}{2R}\right) \\
&= n\exp\left(-\frac{\tau\log n}{2}\right) \\
&= n \cdot n^{-\tau/2} \\
&= n^{-(\tau-2)/2}.
\end{aligned}$$

Remark 33. Based on the above inequalities, one might be tempted to set τ to be very large. However, we will see in Section 30.6 that $\mathbb{E}[|E_H|]$ increases with τ.

30.6 The number of edges

Now we know that with high probability H will be an ϵ-approximation of G, but we still need to compute the expected number of edges in H. To that end define the random variables $X_{a,b}$ as

$$\mathbb{P}(X_{a,b} = 1) = p_{a,b} \quad \text{and} \quad \mathbb{P}(X_{a,b} = 0) = 1 - p_{a,b}.$$

It follows that $\mathbb{E}[X_{a,b}] = p_{a,b}$ and

$$|E_H| = X := \sum_{(a,b)\in E} X_{a,b}.$$

Thus,

$$\mathbb{E}[|E_H|] = \mathbb{E}[X] = \sum_{(a,b)\in E} \mathbb{E}[X_{a,b}] = \sum_{(a,b)\in E} p_{a,b} \,.$$

So we need to compute the sum of the probabilities $p_{a,b}$. Let us first compute $\|\boldsymbol{L}_G^{+/2}\boldsymbol{L}_{a,b}\boldsymbol{L}_G^{+/2}\|$. Since $\boldsymbol{L}_{a,b}$ has rank one, the matrix $\boldsymbol{L}_G^{+/2}\boldsymbol{L}_{a,b}\boldsymbol{L}_G^{+/2}$ also has rank one. Thus it has one non-zero eigenvalue $\lambda_{a,b} > 0$. Therefore:

$$\begin{aligned}
\|\boldsymbol{L}_G^{+/2}\boldsymbol{L}_{a,b}\boldsymbol{L}_G^{+/2}\| &= \lambda_{a,b} \\
&= \mathrm{Tr}(\boldsymbol{L}_G^{+/2}\boldsymbol{L}_{a,b}\boldsymbol{L}_G^{+/2}) \\
&= \mathrm{Tr}(\boldsymbol{L}_G^{+}\boldsymbol{L}_{a,b}) \,.
\end{aligned}$$

As such:

$$\begin{aligned}
\sum_{(a,b)\in E} p_{a,b} &= \frac{1}{R}\sum_{(a,b)\in E} w(a,b)\|\boldsymbol{L}_G^{+/2}\boldsymbol{L}_{a,b}\boldsymbol{L}_G^{+/2}\| \\
&= \frac{1}{R}\sum_{(a,b)\in E} w(a,b)\mathrm{Tr}(\boldsymbol{L}_G^{+}\boldsymbol{L}_{a,b}) \\
&= \frac{1}{R}\mathrm{Tr}\left(\sum_{(a,b)\in E} w(a,b)\boldsymbol{L}_G^{+}\boldsymbol{L}_{a,b}\right) \\
&= \frac{1}{R}\mathrm{Tr}\left(\boldsymbol{L}_G^{+}\sum_{(a,b)\in E} w(a,b)\boldsymbol{L}_{a,b}\right) \\
&= \frac{1}{R}\mathrm{Tr}(\boldsymbol{L}_G^{+}\boldsymbol{L}_G) \\
&= \frac{1}{R}\mathrm{Tr}(\boldsymbol{\Pi}) \\
&= \frac{n-1}{R} \,.
\end{aligned}$$

We conclude that

$$\mathbb{E}[|E_H|] = \frac{n-1}{R} \,.$$

Recall,

$$R = \frac{\epsilon^2}{\tau \log n} \,, \quad \tau > 3 \,,$$

so that

$$\mathbb{E}[|E_H|] = \tau\epsilon^{-2}(n-1)\log n \,.$$

Thus the expected number of edges is $O(\epsilon^{-2} n \log n)$, which is good. But what about any single particular realization of H? Let $0 \le \delta \le 1$ and apply Theorem 60:

$$\begin{aligned}
\mathbb{P}\left(|E_H| \ge (1+\delta)\tau\epsilon^{-2}(n-1)\log n\right) &= \mathbb{P}(|E_H| \ge (1+\delta)\mathbb{E}[|E_H|]) \\
&\le \exp\left(-\frac{\delta^2 \mathbb{E}[|E_H|]}{3}\right) \\
&= \exp\left(-\frac{\delta^2 (n-1)}{3R}\right) \\
&= \exp\left(-\frac{\tau\delta^2 (n-1)\log n}{3\epsilon^2}\right) \\
&\le \exp\left(-\frac{\delta^2 (n-1)\log n}{\epsilon^2}\right) \\
&\le \exp\left(-\frac{\delta^2 (n-1)}{\epsilon^2}\right) .
\end{aligned}$$

We can simplify even further by setting $\delta = \epsilon$, in which case we get:

$$\mathbb{P}\left(|E_H| \ge (1+\epsilon)\tau\epsilon^{-2}(n-1)\log n\right) \le e^{-(n-1)} \le 3e^{-n} .$$

In other words, it is exceedingly unlikely that $|E_H|$ will have more than $O(\epsilon^{-2} n \log n)$ edges, even for small n.

30.7 Collecting everything

In this section we collect all the main results that we proved in the previous sections. Here is our algorithm:

- **Input:** A connected graph $G = (V, E, w)$, a tolerance $0 < \epsilon < 1$, and the parameter $\tau > 3$.

1. For each edge $(a, b) \in E$ compute the probabilities

$$p_{a,b} := \min\left(\frac{1}{R} w(a,b) \|\boldsymbol{L}_G^{+/2} \boldsymbol{L}_{a,b} \boldsymbol{L}_G^{+/2}\| \,,\, 1\right) , \quad \text{with} \quad R := \frac{\epsilon^2}{\tau \log n} .$$

2. Randomly sample the edges from G for inclusion in $H = (V, E_H, w_H)$ according to the probabilities:

$$\mathbb{P}((a,b) \in E_H) = p_{a,b} \quad \text{and} \quad \mathbb{P}((a,b) \notin E_H) = 1 - p_{a,b} .$$

3. For each edge $(a, b) \in E_H$, assign it the weight

$$w_H(a,b) = \frac{w(a,b)}{p_{a,b}} .$$

- **Output:** The graph $H = (V, E_H, w_H)$.

The following theorem collects what we can say about the graph H.

Theorem 62. *In the algorithm above suppose that*

$$p_{a,b} = \frac{1}{R} w(a,b) \|\boldsymbol{L}_G^{+/2} \boldsymbol{L}_{a,b} \boldsymbol{L}_G^{+/2}\| \leq 1 \,, \quad \forall\, (a,b) \in E \,.$$

Then the expected number of edges in H is

$$\mathbb{E}[|E_H|] = \tau \epsilon^{-2} (n-1) \log n \,,$$

and furthermore

$$|E_H| < 2\mathbb{E}[|E_H|] = 2\tau\epsilon^{-2}(n-1)\log n \quad \textit{with probability } 1 - 3e^{-n} \,.$$

Additionally, H is an ϵ-approximation of G with probability $1 - 2n^{-(\tau-3)/3}$.

Remark 34. As we mentioned earlier, what if $R^{-1} w(a,b) \|\boldsymbol{L}_G^{+/2} \boldsymbol{L}_{a,b} \boldsymbol{L}_G^{+/2}\| > 1$? We have two options. One is use the above algorithm, which accounts for this possibility by setting $p_{a,b} = 1$ in this case. Theorem 62 does not apply, but one can adapt the Chernoff bounds to get a similar result.

The other option is to adapt the algorithm. For any $p_{a,b} = R^{-1} w(a,b) \|\boldsymbol{L}_G^{+/2} \boldsymbol{L}_{a,b} \boldsymbol{L}_G^{+/2}\| > 1$, set $k = \lfloor p_{a,b} \rfloor \geq 1$. Create k copies of the edge $(a,b) \in E$ such that each of these copies is put into E_H with probability 1; weight these edges as $w_H(a,b) = w(a,b)/p_{a,b}$. Associated to these edges, create k random matrices $\boldsymbol{X}_{a,b}^{(j)}$, $1 \leq j \leq k$, such that

$$\mathbb{P}\left(\boldsymbol{X}_{a,b}^{(j)} = \frac{w(a,b)}{p_{a,b}} \boldsymbol{L}_G^{+/2} \boldsymbol{L}_{a,b} \boldsymbol{L}_G^{+/2} \right) = 1 \,, \quad 1 \leq j \leq k \,.$$

Create one additional copy of the edge (a,b) such that this copy is placed in E_H with probability $p_{a,b} - k$ and give it the same edge weight $w_H(a,b) = w(a,b)/p_{a,b}$. Associate to this edge the random matrix $\boldsymbol{X}_{a,b}^{(k+1)}$ such that

$$\mathbb{P}\left(\boldsymbol{X}_{a,b}^{(k+1)} = \frac{w(a,b)}{p_{a,b}} \boldsymbol{L}_G^{+/2} \boldsymbol{L}_{a,b} \boldsymbol{L}_G^{+/2} \right) = p_{a,b} - k \quad \text{and} \quad \mathbb{P}\left(\boldsymbol{X}_{a,b}^{(k+1)} = \boldsymbol{0} \right) = 1 - (p_{a,b} - k) \,.$$

Algorithmically, combine all the copies of (a,b) that make it into E_H into a single edge, and assign this single edge the weight equal to the sum of the edge weights of the copies. The proof of Theorem 62 will go through with minor modifications; in particular, one will be able to use the same Chernoff bounds.

References

[1] Daniel A. Spielman. Spectral and algebraic graph theory. Book draft, available at: http://cs-www.cs.yale.edu/homes/spielman/sagt/, 2019.

[2] Michael Perlmutter, Feng Gao, Guy Wolf, and Matthew Hirn. Geometric scattering networks on compact Riemannian manifolds. In *Proceedings of The First Mathematical and Scientific Machine Learning Conference, Proceedings of Machine Learning Research*, volume 107, pages 570–604, 2020.

[3] David I. Shuman, Sunil K. Narang, Pascal Frossard, Antonio Ortega, and Pierre Vandergheynst. The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains. *IEEE Signal Processing Magazine*, 30(3):83–98, 2013.

[4] Stéphane Mallat. *A Wavelet Tour of Signal Processing, Third Edition: The Sparse Way.* Academic Press, 3rd edition, 2008.

[5] Ronald R. Coifman and Stéphane Lafon. Diffusion maps. *Applied and Computational Harmonic Analysis*, 21:5–30, 2006.