Classification: Physical Sciences, Mathematics


The heat kernel as the pagerank of a graph

by
Fan Chung

Department of Mathematics
University of California at San Diego
La Jolla, CA 92093-0112




Corresponding author:
 Fan Chung

Department of Mathematics
University of California at San Diego
La Jolla, CA 92093-0112
858-534-2848 (Office)
858-459-2609 (Fax)
fan@ucsd.edu

# The heat kernel as the pagerank of a graph

Fan Chung*

**Abstract**

The concept of pagerank was first started as a way for determining the ranking of Webpages by Web search engines. Based on relations in interconnected networks, pagerank has become a major tool for addressing fundamental problems arising in general graphs, especially for large information networks with hundreds of thousands of nodes. A notable notion of pagerank, introduced by Brin and Page and denoted by PageRank, is based on random walks as a geometric sum. In this paper we consider a new notion of pagerank which is based on the (discrete) heat kernel and can be expressed as an exponential sum of random walks. The heat kernel satisfies the heat equation and can be used to analyze many useful properties of random walks in a graph. A local Cheeger inequality is established which implies that by focusing on cuts determined by linear orderings of vertices using the heat kernel pageranks, the resulting partition is within a quadratic factor of the optimum. This is true, even if we restrict the volume of the small part separated by the cut to be close to some specified target value. This leads to a graph partitioning algorithm for which the running time is proportional to the size of the targeted volume (instead of the size of the whole graph).

**Introduction**

In the development of quantitative ranking for Webpages, many mathematical methods have come into play. The Hub-and-Authority algorithm by Kleinberg [11] uses eigenvectors. The PageRank introduced by Brin and Page [3] basically uses random walks. These pagerank algorithms mainly rely on the network structure of the Web. The viewpoint is to regard the Web as a graph, with vertices to be Webpages and edges as links between pairs of Webpages. Various notions of pagerank are computed using the Webgraph which are then used for numerous applications, such as identifying communities or finding hot spots in various information networks. Another example is to use PageRank to derive a local graph partitioning algorithm [1], which can be computed very efficiently in the sense that the cost of computing is proportional to the size of the small part of the partition, in contrast with the generic partitioning algorithm having cost depending on the size of the whole graph.

In this paper, we introduce a new notion of pagerank by using the heat kernel of a graph. Similar to PageRank, the heat kernel pagerank is based on random walks but having the extra benefit of satisfying the heat equation. Originally rooted in spectral geometry [15], the heat equation for graphs involves a parameter $t$, the heat, which allows additional control of the rate of diffusion (see detailed definitions later). Using the heat equation, the heat kernel pagerank is amenable to various mathematical analyses of the graph. A key isoperimetric invariant of a graph is the Cheeger constant which provides an evaluation of how good a cut can be found. The classical Cheeger inequality concerns the relationship between the Cheeger constant and eigenvalues of the (normalized) Laplacian of a graph. (A graph can be viewed as a discrete version of a manifold where the original Cheeger inequality applies [4].) Here we will prove several variations of the Cheeger inequality, establishing relationships between the Cheeger constant and the heat kernel pagerank. One of the consequences of the local Cheeger inequality is that, for a given value $s$, the minimum Cheeger ratio of subsets of

---

volume at most $s$ can be approximated up to a quadratic factor by focusing on subsets obtained by using heat kernel pageranks.

A byproduct of the classical Cheeger inequality is a fast partition algorithm using eigenvectors. Here we will show that the heat kernel pagerank leads to efficient local partitioning algorithms with several advantages over previous algorithms. Instead of partitioning graphs into almost equal parts for divide-and-conquer approaches, in a large graph, a local partition algorithm seeks a set with volume bounded by some target size near some specified seed. Recently, there has been progress on developing local partitioning algorithms. Spielman and Teng [14] gave a local partitioning algorithm based on a result of Lovász and Simonovits on rapidly mixing random walks [12, 13]. By using PageRank, an improved local partitioning algorithm was given in [1, 2]. All of these partitioning algorithms are so-called "one-sweep" algorithms that focus on the subsets consisting of the highest $j$ vertices, for some $j$, according to some linear ordering. In this paper, we will give a local partitioning algorithm, which, for a given target size $s$, is a one-sweep algorithm using a truncated version of heat kernel pagerank with support at most $s$, and thus further improves previous work. The algorithm is based on a local version of the Cheeger inequality.

**Preliminaries**

The starting point of the heat kernel pagerank is a typical random walk. In a graph $G$, the transition probability matrix $W$ of a typical random walk on a graph $G = (V, E)$ is a matrix with columns and rows indexed by $V$ and is defined by :

$$W(u,v) \;\; = \;\; \begin{cases} \frac{1}{d_u} & \text{if } \{u,v\} \in E, \\ 0 & \text{otherwise.} \end{cases}$$

Before we proceed to define heat kernel pagerank, we will first describe PageRank, as defined by Brin and Page [3]. The PageRank involves a *preference vector* $f$ (which can be viewed as the probabilistic distribution of the seed(s)) and a *jumping constant* $\alpha$. For example, if we have one starting seed denoted by vertex $u$, then $f$ can be written as the $(0,1)$-indicator function $\chi_u$ of $u$. Another example is to take $f$ to be the constant function with value $1/n$ at every vertex as in the original definition in Brin and Page [3]. The version of PageRank we discuss here is often called the *personalized PageRank*. Throughout this paper, a real-valued function $f : V \to \mathbb{R}$ is taken to be a row vector so that $W$ can act on $f$ from the right by matrix multiplication. The PageRank $\text{pr}(\alpha, f)$, with the scale paramenter $\alpha$ and the preference vector $f$, is defined by:

$$\text{pr}_{\alpha,f} = \alpha \sum_{k=0}^{\infty} (1-\alpha)^k f W^k. \tag{1}$$

The heat kernel pagerank also has two parameters: $t$, a non-negative value (the temperature), and $f$ (a preference vector), defined as follows:

$$\rho_{t,f} = e^{-t} \sum_{k=0}^{\infty} \frac{t^k}{k!} f W^k. \tag{2}$$

If we compare (1) with (2), we can see that $\rho(\cdot)$ is just an exponential sum instead of $\text{pr}(\cdot)$, which is a geometrical sum. It is often the case (e.g., in dealing with generating functions), the exponential sum converges more rapidly.

We note that an equivalent definition of the PageRank is given by the following recurrence:

$$\text{pr}_{\alpha,f} = \alpha f + (1-\alpha)\text{pr}_{\alpha,f} W. \tag{3}$$

Instead, the heat kernel pagerank as defined in (2) satisfies the following heat equation:

$$\frac{\partial}{\partial t}\rho_{t,f} = -\rho_{t,f}(I - W).\tag{4}$$

Let us define $L = I - W$. Then the definition of the heat kernel pagerank in (2) can be rewritten as follows:

$$\begin{aligned}\rho_{t,f} &= e^{-t}\sum_{k=0}^{\infty}\frac{t^k}{k!}fW^k\\ &= fe^{-t(I-W)}\\ &= fe^{-tL}\\ &= \sum_{k=0}^{\infty}\frac{(-t)^k}{k!}fL^k.\end{aligned}$$

For a vertex $v$ in $G$, the *degree* of $v$, denoted by $d_v$, is the number of vertices that $v$ is adjacent to. Let $A$ denote the *adjacency matrix* of $G$ and let $D$ represent the *diagonal degree matrix*. We can write $W = D^{-1}A$ and $H_t = e^{-t(I-W)}$. The discrete heat kernel frist introduced in [8] is a symmetric version of $H_t$.

It is known that a random walk has a stationary distribution if it is irreducible and non-periodic. In graph-theoretical terms, a random walk on a graph $G$ has a stationary distribution $\pi$ if $G$ is connected and non-bipartite with $\pi$ satisfying $\pi(u) = d_u/\sum_v d_v$.

From the above definition, we have the following immediate facts for $\rho_{t,f}$. Namely, $\rho_{0,f} = f$, $\rho_{t,\pi} = \pi$ and $\rho_{t,f}\mathbf{1}^* = f\mathbf{1}^* = 1$ if $f$ satisfies $\sum_v f(v) = 1$. Here $\mathbf{1}$ denotes the all 1's function and $x^*$ denotes the transpose of $x$. Also, we have $DH_t = H_t^*D$ and $H_t = H_{t/2}H_{t/2} = H_{t/2}D^{-1}H_{t/2}^*D$.

To approximate the heat kernel pagerank, one might choose an additive approximation by taking a finite sum (cf. (2)). If one prefers a multiplicative approximation, there is a formula, given by Euler [9], as a sum of two infinite products:

$$\rho_{t,f} = \frac{e^{-t}f}{2}\Big(\prod_{k=0}^{\infty}(I + \frac{4t^2W^2}{(2k+1)^2\pi^2}) + tW\prod_{k=1}^{\infty}(I + \frac{t^2W^2}{k^2\pi^2})\Big).$$

**Isoperimetric properties of the heat kernel**

For a subset $S$ of vertices in $G$, the *volume* of $S$, denoted by $\mathrm{vol}(S)$, is $\sum_{u\in S}d_u$. Also the volume of a graph $G$, denoted by $\mathrm{vol}(G)$ is $\mathrm{vol}(G) = \sum_u d_u$. The *edge boundary* of $S$, denoted by $\partial S$ is defined by

$$\partial S = \{\{u,v\} \in E : u \in S \text{ and } v \notin S\}.$$

Let $\bar{S} = V \setminus S$ denote the complement of $S$. Clearly, $\partial S = \partial\bar{S}$. The *Cheeger ratio* of $S$, denoted by $h_S$, is defined by

$$h_S = \frac{|\partial S|}{\min\{\mathrm{vol}(S), \mathrm{vol}(\bar{S})\}}$$

and the *Cheeger constant* of a graph $G$ is $h_G = \min_{S\subseteq V}h_S$.

For a given set $S$, we consider the distribution $f_S$ with $f_S(u) = d_u/\mathrm{vol}(S)$ if $u \in S$, and 0 otherwise. Note that $f_S$ can be written as $\frac{1}{\mathrm{vol}(S)}\chi_S D$ where $\chi_S$ is the indicator function for $S$. For any function $g : V \to \mathbb{R}$, we define $g(S) = \sum_{v\in S}g(v)$.

We will use the heat equation to derive the following isoperimetric inequality for the heat kernel pagerank:

**Lemma 1** *For a subset $S$ with $\text{vol}(S) \leq \text{vol}(G)/2$, we have*

$$\frac{\partial}{\partial t}\rho_{t,f_S}(S) = -\text{vol}(S)\sum_{u \sim v}\left(\frac{\rho_{t/2,f_S}(u)}{d_u} - \frac{\rho_{t/2,f_S}(v)}{d_v}\right)^2 \leq 0$$

*where the sum is over all unordered pairs of vertices $\{u,v\}$ in $E$ and*

$$|\frac{\partial}{\partial t}\rho_{t,f_S}(S)| \leq h_S.$$

To prove this we see that

$$
\begin{aligned}
\frac{\partial}{\partial t}\rho_{t,f_S}(S) &= f_S\frac{\partial}{\partial t}H_t\chi_S^* \\
&= -f_S(I-W)H_t\chi_S^* \\
&= -f_S H_{t/2}(I-W)D^{-1}H_{t/2}^*D\chi_S^* \\
&= -f_S H_{t/2}D^{-1}(D-A)D^{-1}H_{t/2}^*f_S^*\text{vol}(S) \\
&= -\text{vol}(S)\sum_{u \sim v}\left(\frac{\rho_{t/2,f_S}(u)}{d_u} - \frac{\rho_{t/2,f_S}(v)}{d_v}\right)^2 \\
&\leq 0.
\end{aligned}
$$

Here we use the fact (see [5]) that for any $f,g : V \to \mathbb{R}$, we have

$$f(D-A)g^* = \sum_{u \sim v}(f(u)-f(v))(g(u)-g(v)).$$

In a similar way it can be easily checked that $\frac{\partial^2}{\partial t^2}\rho_{t,f_S}(S) \geq 0$. Therefore

$$
\begin{aligned}
\frac{\partial}{\partial t}\rho_{t,f_S}(S) &\geq \left.\frac{\partial}{\partial t}\rho_{t,f_S}(S)\right|_{t=0} \\
&= -\frac{|\partial S|}{\text{vol}(S)} = -h_S
\end{aligned}
$$

as desired.

**A mixing inequality for the heat kernel pagerank**

For a function $f : V \to \mathbb{R}$, we order the vertices of $G$ so that

$$\frac{f(v_1)}{d_{v_1}} \geq \frac{f(v_2)}{d_{v_2}} \geq \ldots \geq \frac{f(v_n)}{d_{v_n}}. \tag{5}$$

Let $S_i$ denote the set consisting of $v_1,\ldots,v_i$. Let $h_f$ denote the least Cheeger ratio $h_{S_i}$ over all $S_i$. We say that $h_f$ is the Cheeger ratio determined by a sweep of $f$. Our goal is to establish a rapid mixing estimate for the heat kernel page rank in terms of the associated Cheeger ratios. For a vertex $u$, we consider $\rho_{t\chi_u}$ which will also be written as $\rho_{t,u}$.

**Theorem 1** *In a graph $G$, for $t \geq 0$, the heat kernel pagerank satisfies*

$$|\rho_{t,u}(v) - \pi(v)| \leq e^{-t\kappa_t^2/2}\sqrt{\frac{d_v}{d_u}}$$

*where $\kappa_t$ is the minimum Cheeger ratio of all Cheeger ratios obtained by sweeps of $\rho_{t,w}$ over all vertices $w$ in $G$.*

The theorem follows from the following slightly stronger statement:

**Theorem 2**

$$|\rho_{t,u}(v) - \pi(v)| \leq e^{-t(\kappa_{t,u}^2 + \kappa_{t,v}^2)/4}\sqrt{\frac{d_v}{d_u}}$$

where $\kappa_{t,u}$ is the minimum Cheeger ratio determined by a sweep of $\rho_{t,u}$.

**Proof:** We note that

$$
\begin{aligned}
\frac{d_u}{d_v}|\rho_{t,u}(v) - \pi(v)|^2 &= |\chi_u D^{1/2}(H_t - \mathbf{1}^*\pi)D^{-1/2}\chi_v^*|^2 \\
&= |\chi_u D^{1/2}(H_{t/2} - \mathbf{1}^*\pi)D^{-1}(H_{t/2}^* - \pi^*\mathbf{1})D^{1/2}\chi_v|^2 \\
&= \left(\sum_w \left(\chi_u D^{1/2}(H_{t/2} - \mathbf{1}^*\pi)D^{-1/2}\chi_w\right)\left(\chi_v D^{1/2}(H_{t/2} - \mathbf{1}^*\pi)D^{-1/2}\chi_w\right)\right)^2 \\
&\leq \sum_w (\chi_u D^{1/2}(H_{t/2} - \mathbf{1}^*\pi)D^{-1/2}\chi_w)^2 \sum_w (\chi_v D^{1/2}(H_{t/2} - \mathbf{1}^*\pi)D^{-1/2}\chi_w)^2 \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{by Cauchy-Schwarz} \\
&= |\rho_{t,u}(u) - \pi(u)| \; |\rho_{t,v}(v) - \pi(v)|
\end{aligned}
$$

since $\rho_{t,u}(u) - \pi(u) = d_u \sum_w (\rho_{t/2,u}(w) - \pi(w))^2 d_w^{-1}$.

It is enough to show that

$$\rho_{t,u}(u) - \pi(u) \quad \leq \quad e^{-t\kappa_{t,u}^2/2}. \tag{6}$$

We use Lemma 1 and consider the following:

$$
\begin{aligned}
\frac{\frac{\partial}{\partial t}\rho_{t,u}(u)}{\rho_{t,u}(u) - \pi(u)} &= -\frac{d_u \sum_{x \sim y}(\rho_{t/2,u}(x)/d_x - \rho_{t/2,u}(y)/d_y)^2}{\rho_{t,u}(u) - \pi(u)} \\
&= -\frac{\sum_{x \sim y}(\rho_{t/2,u}(x)/d_x - \rho_{t/2,u}(y)/d_y)^2}{\sum_w (\rho_{t/2,u}(w)/d_w - \frac{1}{\text{vol}(G)})^2 d_w}. \tag{7}
\end{aligned}
$$

Now we relabel all the vertices so that $\rho_{t/2,u}(v_1)/d_{v_1} \geq \rho_{t/2,u}(v_2)/d_{v_2} \geq \ldots \geq \rho_{t/2,u}(v_n)/d_{v_n}$. Let $r$ be the largest integer such that $\text{vol}(S_r) \leq \text{vol}(G)/2$. We note that

$$
\begin{aligned}
\sum_w \left(\frac{\rho_{t/2,u}(w)}{d_w} - \frac{1}{\text{vol}(G)}\right)^2 d_w &= \min_c \sum_w \left(\frac{\rho_{t/2,u}(w)}{d_w} - c\right)^2 d_w \\
&\leq \sum_w \left(\frac{\rho_{t/2,u}(w)}{d_w} - \frac{\rho_{t/2,u}(r)}{d_r}\right)^2 d_w.
\end{aligned}
$$

Therefore, we have

$$-\frac{\frac{\partial}{\partial t}\rho_{t,u}(u)}{\rho_{t,u}(u) - \pi(u)} \geq \frac{\sum_{x \sim y}(\rho_{t/2,u}(x)/d_x - \rho_{t/2,u}(y)/d_y)^2}{\sum_w (\rho_{t/2,u}(w)/d_w - \rho_{t/2,u}(r)/d_r)^2 d_w} = X.$$

We consider two functions $f_+$ and $f_-$, defined by:

$$f_+(v) = \begin{cases} \frac{\rho_{t/2,u}(v)}{d_v} - \frac{\rho_{t/2,u}(r)}{d_r} & \text{if } \frac{\rho_{t/2,u}(v)}{d_v} - \frac{\rho_{t/2,u}(r)}{d_r} \geq 0, \\ 0 & \text{otherwise.} \end{cases}$$

6

We also define

$$f_-(v) = \begin{cases} -\frac{\rho_{t/2,u}(v)}{d_v} + \frac{\rho_{t/2,u}(r)}{d_r} & \text{if } \frac{\rho_{t/2,u}(v)}{d_v} - \frac{\rho_{t/2,u}(r)}{d_r} \leq 0, \\ 0 & \text{otherwise.} \end{cases}$$

Then we have

$$\begin{aligned} X &\geq \frac{\sum_{x\sim y}\left((f_+(x)-f_+(y))^2 + (f_-(x)-f_-(y))^2\right)}{\sum_w (f_+(w)^2 + f_-(w)^2)d_w} \\ &\geq \min\{\frac{\sum_{x\sim y}(f_+(x)-f_+(y))^2}{\sum_w f_+(w)^2 d_w}, \frac{\sum_{x\sim y}(f_-(x)-f_-(y))^2}{\sum_w f_-(w)^2 d_w}\}. \end{aligned}$$

Note that each of the above two sums, involving $f_+$ and $f_-$, respectively, are nontrivial. Without loss of generality, we may assume that the minimum is achieved by the sum involving $f_+$. Therefore, we have

$$\begin{aligned} X &\geq \frac{\sum_{x\sim y}(f_+(x)-f_+(y))^2}{\sum_w f_+(w)^2 d_w} \\ &= \frac{\left(\sum_{x\sim y}(f_+(x)-f_+(y))^2\right)\left(\sum_{x\sim y}(f_+(x)+f_+(y))^2\right)}{\sum_w f_+(w)^2 d_w \left(\sum_{x\sim y}(f_+(x)+f_+(y))^2\right)} \\ &\geq \frac{\left(\sum_{x\sim y}|f_+(x)^2-f_+(y)^2|\right)^2}{2(\sum_w f_+(w)^2 d_w)^2}. \end{aligned}$$

Let $\tilde{\text{vol}}(S_i)$ denote the minimum of $\text{vol}(S_i)$ and $\text{vol}(\bar{S}_i)$. Then we have

$$\begin{aligned} X &\geq \frac{\left(\sum_{i=1}^{n-1}(f_+(x_i)^2-f_+(x_{i+1})^2)\,|\partial(S_i)|\right)^2}{2(\sum_w f_+(w)^2 d_w)^2} \\ &\geq \frac{\left(\sum_{i=1}^{n-1}(f_+(x_i)^2-f_+(x_{i+1})^2)\,h_f \tilde{\text{vol}}(S_i)\right)^2}{2(\sum_w f_+(w)^2 d_w)^2} \\ &= \frac{h_f^2\left(\sum_{i=1}^{n} f_+(x_i)^2|\tilde{\text{vol}}(S_i)-\tilde{\text{vol}}(S_{i-1})|\right)^2}{2(\sum_w f_+(w)^2 d_w)^2} \\ &= \frac{h_f^2\left(\sum_{i=1}^{n} f_+(x_i)^2 d_{x_i}\right)^2}{2(\sum_w f_+(w)^2 d_w)^2} \\ &= \frac{h_f^2}{2} \end{aligned}$$

by using the convention that $S_0 = \emptyset$. This implies that

$$-\frac{\partial}{\partial t}\log(\rho_{t,u}(u)-\pi(u)) \geq \frac{h_f^2}{2}.$$

By solving the above equation, we have

$$\rho_{t,u}(u)-\pi(u) \leq c_1 + c_2 e^{-th_f^2/2}.$$

Since $\rho(t,u)(u) \leq 1$ and $\lim_{t\to\infty}\rho(t,u)(u)-\pi(u) = 0$, we can choose $c_1 = 0$ and $c_2 = 1$. We have completed the proof for Theorems 1 and 2.

## A Cheeger inequality using heat kernel pagerank

The classical Cheeger inequality states that

$$2h_G \geq \lambda_1 \geq \frac{\alpha_G^2}{2} \geq \frac{h_G^2}{2}$$

where $\alpha_G$ denotes the minimum Cheeger ratios using a sweep over an eigenvector associated with the spectral gap $\lambda_1$ of the (normalized) Laplacian [5]. Here we will give a local versions of the Cheeger inequality, which relates the Cheeger ratio of a subset to the heat kernel pagerank with seeds as the vertices in the subset.

**Theorem 3** *In a graph $G$, for a subset $S$ of vertices in $G$ with $\mathrm{vol}(S) \leq \mathrm{vol}(G)/2$ and a real value $t \geq 0$, the Cheeger ratio of $S$ satisfies the following:*

$$\frac{h_S}{1-\pi(S)} \geq \frac{\kappa_{t,S}^2}{2} - \frac{1+\log|S|}{t}$$

*where $\kappa_{t,S}$ denotes the minimum Cheeger ratio over all sweeps of $\rho_{t/2,u}$ for all $u \in S$ and $\log$ is the natural logarithm.*

**Proof:** From Theorem 2, we have

$$
\begin{aligned}
\rho_{t,f_S} - \pi(S) &= \frac{1}{\mathrm{vol}(S)}\chi_S D(H_t - \mathbf{1}^*\pi)\chi_S \\
&= \sum_{u,v\in S} \frac{\sqrt{d_u d_v}}{\mathrm{vol}(S)}\chi_u D^{1/2}(H_t - \mathbf{1}^*\pi)D^{-1/2}\chi_v \\
&\leq \frac{(\sum_u \sqrt{d_u})^2}{\mathrm{vol}(S)}e^{-t\kappa_{t,S}^2/2} \\
&\leq e^{-t\kappa_{t,S}^2/2}|S|.
\end{aligned}
\tag{8}
$$

Next, we wish to establish a lower bound for $\rho_{t,f_S}(S) - \pi(S)$. We want to show that

$$\frac{\partial^2}{\partial t^2}\Big(-\log(\rho_{t,f_S}(S) - \pi(S))\Big) \leq 0. \tag{9}$$

This implies that the first derivative of $-\log(\rho_{t,f_S}(S) - \pi(S))$ is decreasing. If this is true, we can use Lemma 1 to get:

$$
\begin{aligned}
F(t,S) &= \frac{\partial}{\partial t}\Big(-\log(\rho_{t,f_S}(S) - \pi(S))\Big) \\
&= \frac{\sum_{x\sim y}(\rho_{t/2,f_S}(x)/d_x - \rho_{t/2,f_S}(y)/d_y)^2}{\sum_w(\rho_{t/2,f_S}(w)/d_w - \frac{1}{\mathrm{vol}(G)})^2 d_w} \\
&\leq F(0,S) \\
&= \frac{h_S}{1-\pi(S)}.
\end{aligned}
$$

This implies

$$\rho_{t,f_S}(S) - \pi(S) \geq (1-\pi(S))e^{-h_S t/(1-\pi(S))}. \tag{10}$$

Combining this with the lower bound in (8), we have

$$2h_S \geq \frac{h_S}{1-\pi(S)} \geq \frac{\kappa_{t,S}^2}{2} - \frac{\log|S|}{t} - \frac{1}{t}$$

as claimed.

It remains to prove (9). We consider

$$-\frac{\partial^2}{\partial t^2}\Big(-\log(\rho_{t,f_S}(S)-\pi(S))\Big) = \frac{\rho_{t,f_S}(I-W)^2(S)(\rho_{t,f_S}(S)-\pi(S))-\big(\rho_{t,f_S}(I-W)(S)\big)^2}{(\rho_{t,f_S}(S)-\pi(S))^2}.$$

It suffices to show that

$$\rho_{t,f_S}(I-W)^2(S)(\rho_{t,f_S}(S)-\pi(S))-\big(\rho_{t,f_S}(I-W)(S)\big)^2 \geq 0.$$

This can be proved by using the Cauchy-Schwarz inequality as follows:

$$(\mathrm{vol}(S))^2 \Big(\rho_{t,f_S}(I-W)^2(S)(\rho_{t,f_S}(S)-\pi(S))-\big(\rho_{t,f_S}(I-W)(S)\big)^2)\Big)$$
$$= \|\rho_{t/2,f_S}(I-W)D^{-1/2}\|^2 \ \|(\rho_{t/2,f_S}-\pi)D^{-1/2}\|^2 - \big(\rho_{t/2,f_S}(I-W)D^{-1}\rho_{t/2,f_S}^*\big)^2$$
$$\geq 0$$

The proof is complete.

### A local Cheeger inequality

In a large graph, given a seed and a target volume $s$ of a set, the goal is to find a good cut separating a subset of volume at most $s$ near the seed. It is desirable to have a local algorithm which has a running time proportional to the target size $s$ instead of generically in terms of the total number of vertices in the graph. In order to do so, we can not afford to consider the minimum Cheeger ratio of a full sweep of a function defined on all vertices of $G$. Instead, we define a $s$-local Cheeger ratio of a sweep $f$, denoted by $h_{f,s}$ to be the minimum Cheeger ratio of the segment $S_i$ with $0 \leq \mathrm{vol}(S_i) \leq 2s$. If no such segment exists, then we set $h_{f,s}$ to be $0$. We note that in order to compute the local $s$-Cheeger ratio, we can ignore most of the entries of $f$ except for those with largest values of $f(u)/d_u$ with total volume not exceeding $2s$. We will prove the following local Cheeger inequality which is weaker than the previous Cheeger inequality by a small constant factor.

**Theorem 4** *In a graph $G$ with a subset $S$ with volume $s$, with $s \leq \mathrm{vol}(G)/4$, for any vertex $u$ in $G$, we have*
$$\rho_{t,u}(S)-\pi(S) \leq \sqrt{\frac{s}{d_u}}e^{-t\kappa_{t,u,s}^2/4}$$
*where $\kappa_{t,u,s}$ denote the minimum $s$-local Cheeger ratio of cuts over a sweep of $\rho_{t,u}$ that separate sets of volume between $0$ and $2s$.*

**Proof:** For a function $f : V \to \mathbb{R}$, we define $f(u,v) = f(u)/d_v$ if $u$ is adjacent to $v$ and $0$ otherwise. For an integer $x$, $0 \leq x \leq \mathrm{vol}(G)/2$, we define

$$f(x) = \max_{\substack{T \subseteq V \times V \\ |T|=x}} \sum_{(u,v)\in T} f(u,v).$$

We can extend $f$ to all real $x = k+r$, with $0 \le r < 1$ by defining $f(x) = (1-r)f(k)+rf(k+1)$. If $x = \mathrm{vol}(S_i)$ where $S_i$ consists of vertices with the $i$ highest values of $f(u)/d_u$, then it follows from the definition that $f(x) = \sum_{u \in S_i} f(u)$. Also $f(x)$ is concave in $x$

We consider the lazy walk $\mathbf{W} = (I+W)/2$. Then

$$
\begin{aligned}
f\mathbf{W}(S) &= \frac{1}{2}\big(f(S)+\sum_{u \sim v \in S} f(u,v)\big) \\
&= \frac{1}{2}\big(\sum_{u \text{ or } v \in S} f(u,v)+\sum_{u \text{ and } v \in S} f(u,v)\big) \\
&\le \frac{1}{2}\big(f(\mathrm{vol}(S)+|\partial S|)+f(\mathrm{vol}(S)-|\partial S|)\big) \\
&= \frac{1}{2}\big(f(\mathrm{vol}(S)(1+h_S))+f(\mathrm{vol}(S)(1-h_S))\big).
\end{aligned}
$$

This can be straightforwardly extended to real $x$ with $0 \le x \le \mathrm{vol}(G)/2$. In particular, we focus on $x$ satisfying $0 \le x \le 2s \le \mathrm{vol}(G)/2$ and we choose $f_t = \rho_{t,u} - \pi$. Then

$$
f_t\mathbf{W}(x) \le \frac{1}{2}\big(f_t(x(1+\kappa_{t,u,s}))+f_t(x(1-\kappa_{t,u,s}))\big).
$$

We now consider for $x \in [0, 2s]$,

$$
\begin{aligned}
\frac{\partial}{\partial t}f_t(x) &= -\rho_{t,u}(I-W)(x) \\
&= -2\rho_{t,u}(I-\mathbf{W})(x) \\
&= -2f_t(x)+2f_t\mathbf{W}(x) \\
&\le -2f_t(x)+f_t(x(1+\kappa_{t,u,s}))+f_t(x(1-\kappa_{t,u,s})) \\
&\le 0
\end{aligned}
\tag{11}
$$

by the concavity of $f_t$. Suppose $g_t(x)$ is a solution of the equation in (11) satisfying $f_0(x) \le g_0(x)$, $f_t(0) = g_t(0)$ and $\frac{\partial}{\partial t}f_t(x)|_{t=0} \le \frac{\partial}{\partial t}g_t(x)|_{t=0}$. Then, we have $f_t(x) \ge g_t(x)$. It is easy to check that $g_t(x) \le e^{-t\kappa_{t,u,s}^2/4}\sqrt{\frac{x}{d_u}}$ using $-2+\sqrt{1+x}+\sqrt{1-x} \le -x^2/4$. Thus,

$$
\begin{aligned}
\rho_{t,u}(S)-\pi(S) &\le \rho_{t,u}(s)-\pi(s) \\
&\le \sqrt{\frac{s}{d_u}}e^{-t\kappa_{t,u,s}^2/4},
\end{aligned}
$$

as desired.

Let $h_s$ denote the minimum Cheeger ratio $h_S$ with $0 \le \mathrm{vol}(S) \le 2s$. Also let $\kappa_{t,2s}$ denotes the minimum of $\kappa_{t,u,2s}$ over all $u$. Combining Theorem 4 and (10), we have

$$
\frac{1}{2}e^{-2t\,h_s} \le \rho_{t,f_S}(s)-\pi(s) \le \sqrt{s}e^{-t\kappa_{t,u,2s}^2/2}.
$$

As an immediate consequence, we have

**Theorem 5** *For $s \le \mathrm{vol}(G)/4$, we have*

$$
h_s \ge \frac{\kappa_{t,2s}^2}{4} - \frac{\log s+1}{2t}.
$$

### A local partition algorithm

The Cheeger inequalities are closely associated with graph partition algorithms which have applications in a wide range of areas, in particular for the divide-and-conquer approaches [10] (also see [7]). The spectral partition algorithm using eigenvectors has a long history and is widely used. However it has several disadvantages. For example, the spectral partition algorithm exercises no control over the size of the small part of the partition (although it can be used recursively to achieve a partition of a desired proportion). In a large graph with hundreds of thousand of nodes, it is prohibitively costly to compute eigenvectors. For very large graphs, it is imperative to develop local partition algorithms which can reduce the cost to be proportional to the volume of the separated part of the cut.

A local partition algorithm has inputs including a vertex as the seed, the volume $s$ of the target set and a target value $\phi$ for the Cheeger ratio of the target set. The local Cheeger inequality in Theorem 4 suggests the following local partition algorithm. In order to find the set achieving the minimum $s$-local Cheeger ratio, one can simply consider a sweep of heat kernel pageranks with further restrictions to the cuts with smaller parts of volume between 0 and $2s$.

How fast is the above local partition algorithm? The running time is basically dominated by the running time of computing the heat kernel pagerank with a seed. Indeed, it is enough to find an approximation of the pagerank with a finite support (no more than $2s$).

How good is the above local partition algorithm? The following theorem shows that there are many seeds (with total volume at least half of $S$) so that the heat kernel pagerank with such a seed will find a partition with Cheeger ratio at most of order $\sqrt{h_S \log s}$. We omit the proof here.

**Theorem 6** *In a graph $G$, for a set $S$ with volume $s \leq \mathrm{vol}(G)/4$, and Cheeger ratio $h_S \leq \phi^2$, there is a subset $S' \subseteq S$ with $\mathrm{vol}(S') \geq \mathrm{vol}(S)/2$ such that for any $u \in S'$, the sweep by using the heat kernel pagerank $\rho_{t,u}$, with $t = \lceil \phi^{-2}/4 \rceil$, will find a set $T$ with $s$-local Cheeger ratio at most $\phi\sqrt{\log s}$.*

We remark that another version of a local algorithm involves restriction to a specified subset and its boundary, which is usually called Dirichlet boundary problem. A variation of a local Cheeger inequality involving Dirichlet eigenvalues is examined in [6]. In this paper, we considered heat kernel pagerank without any specified boundary condition.

### Summary

We introduced the heat kernel pagerank for a graph and established a local Cheeger inequality. This local Cheeger inequality establishes the relations between the Cheeger ratio of a set and the local Cheeger ratios over the sweeps of heat kernel pageranks. Consequently, it leads to a local partition algorithm using heat kernel pagerank with cost proportional to the volume of the separated part. If there is a subset of vertices with volume $s$ and having Cheeger ratio $h_S$, our algorithm using heat kernel pagerank generates a set with volume between 0 and $2s$ and having Cheeger ratio at most $\sqrt{h_S \log s}$. This local partition algorithm can also be used as a subroutine for declustering algorithms or for finding balanced partitions.

# References

[1] Andersen R, Chung F, Lang K (2006) Local graph partitioning using pagerank vectors, *Proceedings of the 47th Annual IEEE Symposium on Founation of Computer Science (FOCS'2006)*, 475–486.

[2] Andersen R, Chung F, Lang K (2007) Detecting sharp drops in PageRank and a simplified local partitioning algorithm, *Theory and Applications of Models of Computation, Proceedings of TAMC 2007*, 1–12.

[3] Brin S, Page L (1998) The anatomy of a large-scale hypertextual Web search engine, *Computer Networks and ISDN Systems*, **30 (1-7)** : 107–117.

[4] Cheeger J (1970) A lower bound for the smallest eigenvalue of the Laplacian, Problems in Analysis, eds Gunning RC (Princeton Univ. Press) 195–199.

[5] Chung F (1997) *Spectral Graph Theory* (AMS Publications).

[6] Chung F (2007) Random walks and local cuts in graphs, *LAA* 423: 22–32.

[7] Chung F (2007) Four proofs for Cheeger's inequality and graph partition algorithms, *Proceedings of ICCM*.

[8] Chung F, Yau S-T (1999) Coverings, heat kernels and spanning trees, *Electronic Journal of Combinatorics* 6: #R12.

[9] Euler L (1748) *Introductio in Analysin Infinitorum*, Lausanne. Also in *Table of Integrals, Series, and Products*, by Gradshteyn IS and Ryzhik IM (2000) eds Jeffrey A, Zwillinger D (Academic Press, San Diego) xlvii+1163 pp.

[10] Kannan R, Vempala S, Vetta A (2004) On clusterings: Good, bad and spectral, *JACM* 51: 497–515.

[11] Kleinberg J (1999) Authoritative sources in a hyperlinked environment, *JACM* 46: 604–632.

[12] Lovász L, Simonovits M (1990) The mixing rate of Markov chains, an isoperimetric inequality, and computing the volume, *31st IEEE Annual Symposium on Foundations of Computer Science*, 346–354.

[13] Lovász L, Simonovits M (1993) Random walks in a convex body and an improved volume algorithm, *Random Structures and Algorithms* 4: 359–412.

[14] Spielman D, Teng S-H (2004) Nearly-linear time algorithms for graph partitioning, graph sparsification, and solving linear systems, *Proceedings of the 36th Annual ACM Symposium on Theory of Computing*, 81–90.

[15] Schoen RM, Yau S-T (1994) *Differential Geometry* (International Press, Cambridge, Massachusetts).