A Gradient Estimate for PageRank

Paul Horn[1] and Lauren M. Nelsen[2(✉)]

[1] University of Denver, Denver, CO 80208, USA
paul.horn@du.edu
[2] University of Indianapolis, Indianapolis, IN 46227, USA
nelsenl@uindy.edu
https://cs.du.edu/~paulhorn,
https://sites.google.com/view/laurennelsen

Abstract. Personalized PageRank has found many uses in not only the ranking of webpages, but also algorithmic design, due to its ability to capture certain geometric properties of networks. In this paper, we study the diffusion of PageRank: how varying the jumping (or teleportation) constant affects PageRank values. To this end, we prove a gradient estimate for PageRank, akin to the Li-Yau inequality for positive solutions to the heat equation (for manifolds, with later versions adapted to graphs).

Keywords: PageRank · Discrete curvature · Random walks · Gradient estimate

1 Introduction/Background

Personalized PageRank, developed by Brin and Page [3] ranks the importance of webpages 'near' a seed. PageRank can be thought of in a variety of ways, but one of the most important points of view of PageRank is that it is the distribution of a random walk allowed to diffuse for a geometrically distributed number of steps. A key parameter in PageRank, then, is the 'jumping' or 'teleportation' constant which controls the expected length of the involved random walks. As the jumping constant controls the length, it controls locality – that is, how far from the seed the random walk is (likely) willing to stray.

When the jumping constant is small, the involved walks are (on average) short, and the mass of the distribution will remain concentrated near the seed. As the jumping constant increases, then the involved walk will (likely) be much longer. This allows the random walk to mix, and the involved distribution tends towards the stationary distribution of the random walk. As the PageRank of individual vertices (for a fixed jumping constant) can be thought of as a measure of importance to the seed, then as the jumping constant increases this importance diffuses.

In this paper, we are interested in how this importance diffuses as the jumping constant increases. This diffusion is related to the network's geometry; in particular, the importance can get 'caught' by small cuts. This partially accounts for

H. Cherifi et al. (Eds.): COMPLEX NETWORKS 2019, SCI 881, pp. 15–26, 2020.
https://doi.org/10.1007/978-3-030-36687-2_2

PageRank's importance in web search but has other uses as well – for instance Andersen, Chung and Lang use PageRank to implement local graph partitioning algorithms in [1].

This paper seeks to understand the diffusion of influence (as the jumping constant changes) in analogy to the diffusion of heat. The study of solutions to the heat equation $\Delta u = \frac{\partial}{\partial t} u$ on both graphs and manifolds has a long history, motivated by its close ties to geometric properties of graphs. On graphs, the relationship between heat flow and PageRank has been exploited several times. For instance, Chung [4] introduced the notion of heat kernel PageRank and used it to improve the algorithm of Anderson, Chung, Lang for graph partitioning.

A particularly useful way of understanding positive solutions to the heat equation is through curvature lower bounds, which can be used to prove 'gradient estimates', which bound how heat diffuses locally in space and time and which can be integrated to obtain Harnack inequalities. Most classical of these is the Li-Yau inequality [13], which (in it's simplest form) states that if u is a positive solution on a non-negatively curved n-dimensional compact manifold, then u satisfies

$$\frac{|\nabla u|^2}{u^2} - \frac{u_t}{u} \leq \frac{n}{2t}. \tag{1}$$

In the graph setting, Bauer, et al. proved a gradient estimate for the heat kernel on graphs in [2]. In this paper we aim to prove a similar inequality for PageRank. Our gradient estimate, which is formally stated as Theorem 1 below, is proved using the exponential curvature dimension inequality CDE, introduced by Bauer et. al.

We mention that, in some ways, our inequality is more closely related to another inequality of Hamilton [9] which bounds merely $\frac{|\nabla u|^2}{u^2}$, and was established for graphs by Horn in [10]. Other related works establish gradient estimates for eigenfunctions for the Laplace matrix; these include [6].

This paper is organized as follows: In the next section we introduce definitions for both PageRank and the graph curvature notions used. We further establish a useful 'time parameterization' for PageRank, which allows us to think of increasing the jumping constant as increasing a time parameter, and makes our statements and proofs cleaner. In Sect. 3 we prove a gradient estimate for PageRank. In Sect. 4 we use this gradient estimate to prove a Harnack-type inequality that allows us to compare PageRank at two vertices in a graph.

2 Preliminaries

2.1 Spectral Graph Theory and Graph Laplacians

Spectral graph theory involves associating a matrix (or operator) with a graph and investigating how eigenvalues of the associated matrix reflect graph properties. The most familiar such matrix is the adjacency matrix A, whose rows and columns are indexed with vertices and

$$a_{ij} = \begin{cases} 1 & \text{if } v_i \sim v_j \\ 0 & \text{else.} \end{cases}$$

In this work, the principal matrix that we will consider is the normalized Laplace operator

$$\Delta = I - D^{-1}A,$$

where D is the diagonal matrix of vertex degrees and $D^{-1}A$ is the transition probability matrix for a simple random walk.

As a quick observation, note that Δ is non-positive semidefinite. This is contrary to usual sign conventions in graph theory, but is the proper sign convention for the Laplace-Beltrami operator in Riemannian manifolds, the analogy to which we emphasize in this paper. Also note that this matrix is (up to sign) the unsymmetrized version of the normalized Laplacian popularized by Chung (see [7]), $\mathcal{L} = (D^{-1/2}AD^{-1/2}) - I$.

2.2 PageRank

(Personalized) PageRank was introduced as a ranking mechanism [12], to rank the importance of webpages with respect to a seed. To define personalized PageRank, we introduce the following operator which we call the PageRank operator. This operator, $P(\alpha)$, is defined as follows:

$$P(\alpha) = (1-\alpha)\sum_{k=1}^{\infty} \alpha^k W^k,$$

where $W = D^{-1}A$ is the transition probability matrix for a simple random walk. Here the parameter α is known as the jumping or teleportation constant. For a finite n-vertex graph, $P(\alpha)$ is a square matrix; the *personalized PageRank vector* of a vector $u : V \to \mathbb{R}$ is

$$u^T P(\alpha) = (1-\alpha)\sum_{k=1}^{\infty} \alpha^k u^T W^k$$

It has been noticed ([4,5]) that PageRank has many similarities to the heat kernel $e^{t\Delta}$. Chung defined the notion of 'Heat Kernel PageRank' to exploit these similarities. In this work, we take inspiration in the opposite direction: we are interested in understanding the action of the PageRank operator in analogy to solutions of the heat equation. In order to emphasize our point of view, we note that graph theorists view the heat kernel operator in two different ways: For a vector $u : V \to \mathbb{R}$ studying the evolution of $u^T e^{t\Delta}$ as $t \to \infty$ is really studying the evolution of the *continuous time random walk*, while studying the evolution of $\left(e^{t\Delta}\right) u$ as $t \to \infty$ is studying the solutions to the heat equation $\Delta u = u_t$. The differing behavior of these two evolutions comes from the fact that (for irregular graphs) the left and right eigenvectors of $\Delta = I - W$ are different: the left Perron-Frobenius eigenvector of Δ is proportional to the degrees of a graph (as it captures the stationary distribution of the random walk) while the right Perron-Frobenius eigenvector is the constant vector. In particular, as $t \to \infty$ the vector $e^{t\Delta}u$ tends to a constant. Physically, this represents the 'heat' on a graph evening out, and this regularization (and the rate of regularization) is related to a number of geometric features of a graph.

A similar feature holds for PageRank. As $\alpha \to 1$, $u^T P(\alpha)$ tends to a vector proportional to degrees, but $P(\alpha)u$ regularizes. In this paper we study this regularization. Although we do not study the PageRank vector explicitly, we note that the left and right action of the PageRank operator are closely related. For an undirected graph $u^T P(\alpha) = (P(\alpha)^T u)^T = (DP(\alpha)D^{-1}u)^T$, so that the regularization of $D^{-1}u$ can be translated into information on the 'mixing' of the personalized PageRank vector seeded at u.

To complete the analogy between $P(\alpha)u$ and $e^{t\Delta}u$, it is helpful to come up with a time parameterization $t = t(\alpha)$ so we can view the regularization as a function of 'time', in analogy to the heat equation. To do this in the best way, it is useful to think of $\alpha = \alpha(t)$ and compute $\frac{\partial}{\partial t} P_\alpha$.

Proposition 1

$$\frac{\partial}{\partial t} P_\alpha = \frac{\alpha'}{(1-\alpha)^2} \Delta P_\alpha^2,$$

where $\Delta P_\alpha^2 = \Delta P_\alpha(P_\alpha)$.

Proof. Notice that, the chain rule and algebra reveals,

$$\begin{aligned} \frac{\partial}{\partial t} P_\alpha &= \frac{\partial}{\partial t}(1-\alpha)(I-\alpha W)^{-1} \\ &= \alpha'((\alpha W - I)(I-\alpha W)^{-2} + (1-\alpha)W(I-\alpha W)^{-2} = \frac{\alpha'}{(1-\alpha)^2}\Delta P_\alpha^2. \end{aligned}$$

□

This is remarkably close to the heat equation if $\alpha'(t) = (1-\alpha)^2$; solving this separable differential equation yields that $\alpha = \alpha(t) = 1 - \frac{1}{t+C}$. Since we desire a parameterization so that $\alpha(0) = 0$ and $\alpha \to 1$ as $t \to \infty$, this gives us that $C = 1$ from whence we obtain:

$$\boxed{\begin{aligned} \alpha(t) &= 1 - \frac{1}{t+1} \qquad (2) \\ t &= \frac{\alpha}{1-\alpha} \qquad (3) \end{aligned}}$$

Given the time parameterization in Eq. 2, we get the following Corollary to Proposition 1.

Corollary 2

$$\frac{\partial}{\partial t} P_\alpha = \Delta P_\alpha^2,$$

where $\Delta P_\alpha^2 = \Delta P_\alpha(P_\alpha)$.

Proof. From Proposition 1 and our choice of parameterization, we see that

$$\frac{\partial}{\partial t} P_\alpha = \frac{\alpha'}{(1-\alpha)^2} \Delta P_\alpha^2 = \frac{\frac{1}{(t+1)^2}}{\left(\frac{1}{t+1}\right)^2} \Delta P_\alpha^2 = \Delta P_\alpha^2.$$

□

Fix a vector $u : V \to \mathbb{R}$. From now on, we let

$$\boxed{f = P_\alpha u.} \tag{4}$$

Lemma 1. *For $f = P_\alpha u$ and $t = \frac{\alpha}{1-\alpha}$, we have that $\Delta f = \frac{f-u}{t}$.*

Proof. We know that $W = D^{-1}A$ and $\Delta = W - I$, so

$$\begin{aligned} \Delta P_\alpha &= (W-I)(1-\alpha)(I-\alpha W)^{-1} \\ &= -\frac{1}{\alpha}(I-\alpha W)(1-\alpha)(I-\alpha W)^{-1} + \frac{1-\alpha}{\alpha} \cdot (1-\alpha)(I-\alpha W)^{-1} \\ &= \frac{1-\alpha}{\alpha}(P_\alpha - I). \end{aligned}$$

Hence

$$\Delta f = \Delta P_\alpha u = \frac{(1-\alpha)}{\alpha}(P_\alpha - I)u = \frac{f-u}{t}.$$

□

2.3 Graph Curvature

In this paper we study the regularization of $P(\alpha)u$ for an initial seed u as $\alpha \to 1$. On one hand, the information about this regularization is contained in the spectral decomposition of the random walk matrix W. The eigenvalues of $P(\alpha)$ are determined by the eigenvalues of W: indeed, if λ is an eigenvalue of W, then $\frac{1-\alpha}{1-\alpha\lambda}$ is an eigenvalue of P_α. One may observe that, then, as $\alpha \to 1$ all eigenvalues of P_α tend to zero except for the eigenvalue, 1, of W, and this is what causes the regularization. Thus the difference between $P_\alpha u$ and the constant vector can be bounded in terms of (say) the infinity norms of eigenvectors of P_α and α itself.

On the other hand, curvature lower bounds (in graphs and manifolds) have proven to be important ways to understand the local evolution of solutions to the heat equation. As we have already noted important similarities between heat solutions and PageRank, we seek similar understanding in the present case. Curvature, for graphs and manifolds, gives a way of understanding the local geometry of the object. A manifold (or graph) satisfying a curvature lower bound at every point has a locally constrained geometry which allows a *local* understanding of heat flow through where a 'gradient estimate' can be proved. These gradient estimates can then be 'integrated' over space-time to yield Harnack inequalities which compare the 'heat' of different points at different times.

While a direct analogue of the Ricci curvature is not defined in a graph setting, a number of graph theoretical analogues have been developed recently in an attempt to apply geometrical ideas in the graph setting. In the context of proving gradient estimates of heat solutions, a new notion of curvature known as the *exponential curvature dimension inequality* was introduced in [2]. In order to discuss the exponential curvature dimension inequality, we first need to introduce some notation. The Laplace operator, Δ, on a graph G is defined at a vertex x by

$$\Delta f(x) = \sum_{y \sim x} (f(y) - f(x)).$$

Definition 3. *The gradient form* Γ *is defined by*

$$\begin{aligned} 2\Gamma(f,g)(x) &= (\Delta(f \cdot g) - f \cdot \Delta(g) - \Delta(f) \cdot g)(x) \\ &= \sum_{y \sim x} (f(y) - f(x))(g(y) - g(x)), \end{aligned}$$

and we write $\Gamma(f) = \Gamma(f,f)$*.*

In general, there is no "chain rule" that holds for the Laplacian on graphs. However, the following formula does hold for the Laplacian and will be useful to us:

$$\Delta f = 2\sqrt{f}\Delta\sqrt{f} + 2\Gamma(\sqrt{f}). \tag{5}$$

We define an iterated gradient form, Γ_2, that will be of use to us for the notion of graph curvature that we are using.

Definition 4. *The gradient form* Γ_2 *is defined by*

$$2\Gamma_2(f,g) = \Delta\Gamma(f,g) - \Gamma(f,\Delta g) - \Gamma(\Delta f, g),$$

and we write $\Gamma_2(f) = \Gamma_2(f,f)$*.*

At the heart of the exponential curvature dimension inequality is an idea that had been used previously based on the Bochner formula. The Bochner formula reveals a connection between solutions to the heat equation and the curvature of a manifold. Bochner's formula tells us that if M is a Riemannian manifold and f is in $C^\infty(M)$, then

$$\frac{1}{2}\Delta|\nabla f|^2 = \langle \nabla f, \nabla \Delta f \rangle + ||\mathrm{Hess} f||_2^2 + \mathrm{Ric}(\nabla f, \nabla f).$$

The Bochner formula implies that for an n-dimensional manifold with Ricci curvature at least K, we have

$$\frac{1}{2}\Delta|\nabla f|^2 \geq \langle \nabla f, \nabla \Delta f \rangle + \frac{1}{n}(\Delta f)^2 + K|\nabla f|^2. \tag{6}$$

An important insight of Bakry and Emery was that an object satisfying an inequality like (6) could be used as a *definition* of a curvature lower bound even when curvature could not be directly defined. Such an inequality became known as a *curvature dimension inequality*, or the CD inequality. Bauer, et al. introduced a modification of the CD inequality that defines a new notion of curvature on graphs that we will use here [2], the exponential curvature inequality.

Definition 5. *A graph is said to satisfy the* **exponential curvature dimension inequality** $CDE(n, K)$ *if, for all positive* $f : V \to \mathbb{R}$ *and at all vertices* $x \in V(G)$ *satisfying* $(\Delta f)(x) < 0$

$$\Delta\Gamma(f) - 2\Gamma(f, \frac{\Delta f^2}{2f}) \geq \frac{2}{n}(\Delta f)^2 + 2K\Gamma(f), \tag{7}$$

where the inequality in (7) *is taken pointwise.*

While the inequality (7) may seem somewhat unwieldy it, as shown in [2], arises from 'baking in' the chain rule and is actually equivalent to the standard curvature dimension inequality (6) in the setting of diffusive semigroups (where the Laplace operator satisfies the chain rule.) Additionally, in [2], it is shown that some graphs including the Ricci flat graphs of Chung and Yau satisfy $CDE(n, 0)$ (and hence are non-negatively curved for this curvature notion) and some general curvature lower bounds for graphs are given.

An important observation is that this notion of curvature only requires looking at the second neighborhood of a graph, and hence this kind of curvature is truly a local property (and hence a curvature lower bound can be certified by only inspecting second neighborhoods of vertices.)

3 Gradient Estimate for PageRank

Our main result will make use of the following lemma, adapted from a lemma in [2].

Lemma 2. **([2]).** *Let* $G(V, E)$ *be a (finite or infinite) graph, and let* $f, H : V \times \{t^\star\} \to \mathbb{R}$ *be functions. If* $f \geq 0$ *and* H *has a local maximum at* $(x^\star, t^\star) \in V \times \{t^\star\}$*, then*

$$\Delta(fH)(x^\star, t^\star) \leq (\Delta f)H(x^\star, t^\star).$$

Our goal is to show that $\frac{\Gamma(\sqrt{f})}{\sqrt{f \cdot M}} \leq \frac{C(t)}{t}$ for some function $C(t)$. However, $\frac{C(t)}{t}$ is badly behaved as $t \to 0$. The way that we handle this is by showing that $H := t \cdot \frac{\Gamma(\sqrt{f})}{\sqrt{f \cdot M}} \leq C(t)$. If H is a function from $V \times [0, \infty) \to \mathbb{R}$, then instead consider H as a function from $V \times [0, T] \to \mathbb{R}$ for some $T > 0$. Then, by compactness, there is a point $(x^\star, t^\star)$ in $V \times [0, T]$ at which $H(x, t)$ is maximized. At this maximum, we know that $\Delta H \leq 0$ and $\frac{\partial}{\partial t}H \geq 0$. Since $\mathcal{L} = \Delta - \frac{\partial}{\partial t}$, this implies that at the maximum point, $\mathcal{L}H \leq 0$. Using the CDE inequality, along with some other lemmas and an identity, we are able to relate H^2 with itself. This allows us to find an upper bound for H, and thus for $\frac{\Gamma(\sqrt{f})}{\sqrt{f \cdot M}}$. Our situation is a little easier, because we consider a fixed t.

A simple computation shows the following:

Lemma 3. *Let G be a graph, and suppose $0 \leq f(x) \leq M$ for all $x \in V(G)$ and $t \in [0,\infty)$, and let $H = \frac{t\Gamma(\sqrt{f})}{\sqrt{f \cdot M}}$. Then*

$$\Delta\sqrt{f} = \frac{f-u}{2t\sqrt{f}} - \frac{\sqrt{M}H}{t}.$$

This identity plays a similar role to the identity $\Delta \log u = \frac{\Delta u}{u} - \frac{|\nabla u|^2}{u^2}$ that is key in the Li-Yau inequality on manifolds, and the identity $\frac{\Delta\sqrt{u}}{\sqrt{u}} = \frac{\Delta u}{u} - \frac{|\nabla\sqrt{u}|^2}{u}$, which is behind the Li-Yau inequality for graphs. Lemma 3 is similar to these other identities and the CDE inequality allows us to exploit this relationship.

Theorem 1. *Let G be a graph satisfying $CDE(n,0)$. Suppose $0 \leq f \leq M$ for all $x \in V(G)$ and $t \in (0,\infty)$. Then*

$$\frac{\Gamma(\sqrt{f})}{\sqrt{f \cdot M}} \leq \frac{n+4}{n+2} \cdot \frac{1}{t} + 2\sqrt{\frac{n}{n+2}} \cdot \frac{1}{\sqrt{t}}.$$

Note that this theorem actually is more akin to a 'Hamilton-type' gradient estimate, as it is an estimate in space only (and not time). Due to space constraints, the full proof of Theorem 1 is deferred to the full version of the paper. It proceeds by the maximum principle, similarly to the proof of the Li-Yau inequality in [2] but requires additional care in handling some terms since the heat equation is not specified; these give rise to its form. For convenience of the reader, we sketch the main ideas in the proof in an appendix.

4 Harnack-Type Inequality

We can use Theorem 1 to prove a result comparing PageRank at two vertices in a graph depending on the distance between them. This result is similar to a *Harnack inequality.* The classical form of a Harnack inequality is the following.

Proposition 6 *([2]). Suppose G is a graph satisfying $CDE(n,0)$. Let $T_1 < T_2$ be real numbers, let $d(x,y)$ denote the distance between $x, y \in V(G)$, and let $D = \max_{v \in V(G)} \deg(v)$. If u is a positive solution to the heat equation on G, then*

$$u(x,T_1) \leq u(y,T_2)\left(\frac{T_2}{T_1}\right)^n \exp\left(\frac{4Dd(x,y)^2}{T_2 - T_1}\right).$$

This result allows one to compare heat at different points and different times. This can make it possible to deduce geometric information about the graph, such as bottlenecking. Delmotte [8] showed that Harnack inequalities do not only allow us to compare heat at different points in space and time – they also have geometric consequences, such as volume doubling and satisfying the Poincaré inequality. Horn, Lin, Liu, and Yau [11] completed the work of Delmotte by proving that even more geometric information can be obtained from Harnack inequalities.

Using Theorem 1, we are able to relate PageRank at different vertices, but our result is not quite of the right form to be a Harnack inequality. In Theorem 1, an ideal conclusion would be to have an f instead of $\sqrt{f \cdot M}$ in the denominator. Since we do not, this makes proving a "Harnack-type" inequality, directly comparing the two values in terms of themselves and their distance, more difficult. (A somewhat similar technique is used by Horn in [10] on the heat equation, but in the case of [10] the gradient estimate is scaled better, yielding stronger results.)

To prove our Harnack-type inequality, we will use a lemma comparing PageRank at adjacent vertices. From now on, we will consider t fixed and write $f(x)$ instead of $f(x,t)$. If a vertex, w, is adjacent to a vertex, z, then we want to lower bound $\sqrt{f(z)}$ by a function only involving $f(w)$. The trick to this is to rewrite $\sqrt{\frac{f(w)}{f(z)}}$ so that we can use Theorem 1 in order to get rid of the '$\sqrt{f(z)}$' in the denominator.

Lemma 4. *Let $D = \max_{v \in V(G)} \deg(v)$. If $w \sim z$, then*

$$\sqrt{\frac{f(w)}{f(z)}} \leq \frac{2CD\sqrt{M}}{\sqrt{t}} \cdot \frac{1}{\sqrt{f(w)}} + 2.$$

Proof. If $\sqrt{f(z)} \geq \frac{1}{2}\sqrt{f(w)}$, then $\sqrt{\frac{f(w)}{f(z)}} \leq 2 \leq \frac{2CD\sqrt{M}}{\sqrt{t}} \cdot \frac{1}{\sqrt{f(w)}} + 2$.

If $\sqrt{f(z)} < \frac{1}{2}\sqrt{f(w)}$, then

$$\begin{aligned}\sqrt{\frac{f(w)}{f(z)}} &= \frac{\sqrt{f(w)} - \sqrt{f(z)} + \sqrt{f(z)}}{\sqrt{f(z)}} \\ &= \frac{\sqrt{f(w)} - \sqrt{f(z)}}{\sqrt{f(z)}} + 1 \\ &= \frac{D(\sqrt{f(w)} - \sqrt{f(z)})^2}{D\sqrt{f(z)}(\sqrt{f(w)} - \sqrt{f(z)})} + 1. \qquad (8)\end{aligned}$$

Now applying the gradient estimate (Theorem 1) yields,

$$\begin{aligned}(8) &\leq \frac{CD\sqrt{M}}{\sqrt{t}} \cdot \frac{1}{\sqrt{f(w)} - \sqrt{f(z)}} + 1 \\ &\leq \frac{CD\sqrt{M}}{\sqrt{t}} \cdot \frac{2}{\sqrt{f(w)}} + 1 \text{ since } \sqrt{f(z)} < \frac{1}{2}\sqrt{f(w)} \\ &\leq \frac{2CD\sqrt{M}}{\sqrt{t}} \cdot \frac{1}{\sqrt{f(w)}} + 2.\end{aligned}$$

□

We note that this can be carefully iterated to compare PageRank of vertices of a given distance. This proof, however (and even its rather complicated statement) are deferred to the full journal version of the paper due to space considerations.

5 Conclusions, Applications, and Future Work

In this paper we investigated PageRank as a diffusion, using recently developed notions of discrete curvature. These results, while theoretical (and in some cases not as strong as would be desired due to the dependence on the maximum value 'M' in the gradient estimate), show that curvature aspects of graphs can be used to understand relative importance in networks – at least when ranking is based on random walk based diffusions.

Regarding these points, we highlight the following:

– Curvature is a local property – based only on second neighborhood conditions. An upshot of this is that it can be certified quickly. While the work here focuses on situations where the entire graph is non-negatively curved for simplicity, work in [2,10] show that these methods can be used when only parts of the graph satisfy such a radius by using cut-off functions. In principle these yield algorithms that are linear, either in the size of the graph – or even in a considered portion of the graph – verifying curvature conditions and elucidating PageRank's diffusion in bounded degree graphs.
– The influence of the jumping constant on PageRank has been important for certain algorithms (such as in [1]), but was originally picked rather arbitrarily (see, eg. [14]). A more rigorous study of this phenomenon seems important for the analysis of complex networks and this paper should be seen as part of this thrust.
– There are several interesting areas for improvement here: The non-ideal scaling in Theorem 1 leads to a weaker than ideal result in Lemma 4. While Lemma 4 seems a reasonable result, when iterated it quickly loses power (unlike the Harnack inequality from a 'properly scaled' gradient estimate like in Proposition 6). While a 'properly scaled' Theorem 1 may not even be true, we suspect the scaling can be improved. An interesting question is whether a true 'Hamilton type' gradient estimate is true: Is $\Gamma(\sqrt{f})/f \leq C\log(M/f)t^{-1}$? Note that the addition of the logarithmic term damages a Harnack inequality, but the results obtainable from this are far better than we obtain. Also, a version including the time derivative term is also desirable.

Acknowledgments. Horn's work was partially supported by Simons Collaboration grant #525309.

Appendix

The proof of Theorem 1 includes some rather lengthy computations, and is deferred for the full paper. For the benefit of readers, however, we have included a sketch here which highlights the initial part of the proof where one relates the quantity to be bounded with its own square using CDE.

Proof. **(Proof sketch for Theorem 1).** Let $H = \dfrac{t\Gamma(\sqrt{f})}{\sqrt{f \cdot M}}$. Fix $t > 0$. Let $(x^\star, t)$ be a point in $V \times \{t\}$ such that $H(x, t)$ is maximized. We desire to bound $H(x^\star, t)$.

Our goal, then is to apply the CDE inequality to $\Delta(\sqrt{f}H)$. In order to do this, we must ensure that $\Delta\sqrt{f} < 0$, but a computation shows that if $\Delta\sqrt{f} \geq 0$, then $H \leq \frac{1}{2}$ so this is allowable.

Following this, one computes by bounding the arising $\Delta\Gamma(\sqrt{f})$ by CDE. One bounds the ensuing terms; clearly $\frac{2}{t}\Gamma(\sqrt{f}) - \frac{2}{t}\Gamma\left(\sqrt{f}, \frac{u}{\sqrt{f}}\right) \geq -\frac{2}{t}\Gamma\left(\sqrt{f}, \frac{u}{\sqrt{f}}\right)$ and by Lemma 3, $\Delta\sqrt{f} = \left(\frac{f-u}{2t\sqrt{f}} - \frac{\sqrt{M}H}{t}\right)$. Then one bounds:

$$\begin{aligned}
\left(\Delta\sqrt{f}\right) H &\geq \frac{2}{\sqrt{M}nt}\left(\frac{(f-u)^2}{2\sqrt{f}} - \frac{(f-u)\sqrt{M}H}{\sqrt{f}} + MH^2\right) \\
&\qquad + \frac{2\Gamma(\sqrt{f})}{\sqrt{M}} - \frac{2}{\sqrt{M}}\Gamma\left(\sqrt{f}, \frac{u}{\sqrt{f}}\right) \\
&\geq \frac{2\left(MH^2 - \sqrt{f}\sqrt{M}H\right)}{\sqrt{M}nt} \\
&\qquad - \frac{1}{\sqrt{M}}\widetilde{\sum_{y\sim x}}\left(u(x)\left(1 - \sqrt{\frac{f(y)}{f(x)}}\right) + u(y)\left(1 - \sqrt{\frac{f(x)}{f(y)}}\right)\right) \\
&\geq \frac{2\left(MH^2 - \sqrt{f}\sqrt{M}H\right)}{\sqrt{M}nt} - \sqrt{M},
\end{aligned}$$

Now one proceeds carefully, noting that we have related H and its square and thus, in principle at least, have recorded an upper bound for H. Now we continue to compute to recover the result.

Remark: In a typical application of the maximum principle, one maximizes over $[0, T]$ and then uses information from the time derivative. Here, we don't do this. This is important because one obtains an inequality of the form

$$H^2 \leq C_1 \cdot H + C_2 \cdot t$$

Because of the dependence of this inequality on the time where the maximum occurs, if the $t^\star$ maximizing the function over all $[0, T]$ is considered, then the result will depend on $t^\star$, giving a bound like $H \leq \frac{\sqrt{2C_2 t^\star}}{t}$. However, since we are able to do the computation at t, this problem does not arise.

References

1. Andersen, R., Chung, F., Lang, K.: Local partitioning for directed graphs using PageRank. Internet Math. **5**(1–2), 3–22 (2008)
2. Bauer, F., Horn, P., Lin, Y., Lippner, G., Mangoubi, D., Yau, S.-T.: Li-Yau inequality on graphs. J. Differ. Geom. **99**(3), 359–405 (2015)
3. Brin, S., Page, L.: The anatomy of a large-scale hypertextual web search engine. Comput. Netw. ISDN Syst. **30**(1), 107–117 (1998). Proceedings of the Seventh International World Wide Web Conference

4. Chung, F.: The heat kernel as the pagerank of a graph. Proc. Natl. Acad. Sci. **104**(50), 19735–19740 (2007)
5. Chung, F.: PageRank as a discrete Green's function. In: Geometry and Analysis, no. 1, volume 17 of Advanced Lectures in Mathematics (ALM), pp. 285–302. International Press, Somerville (2011)
6. Chung, F., Lin, Y., Yau, S.-T.: Harnack inequalities for graphs with non-negative Ricci curvature. J. Math. Anal. Appl. **415**(1), 25–32 (2014)
7. Chung, F.R.K.: Spectral Graph Theory, volume 92 of CBMS Regional Conference Series in Mathematics. Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence (1997)
8. Delmotte, T.: Parabolic Harnack inequality and estimates of Markov chains on graphs. Rev. Mat. Iberoamericana **15**(1), 181–232 (1999)
9. Hamilton, R.S.: A matrix Harnack estimate for the heat equation. Comm. Anal. Geom. **1**(1), 113–126 (1993)
10. Horn, P.: A spacial gradient estimate for solutions to the heat equation on graphs. SIAM J. Discrete Math. **33**(2), 958–975 (2019)
11. Horn, P., Lin, Y., Liu, S., Yau, S.-T.: Volume doubling, poincaré inequality and Gaussian heat kernel estimate for non-negatively curved graphs. J. Reine Angew. Math. (to appear)
12. Jeh, G., Widom, J.: Scaling personalized web search. In: Proceedings of the 12th World Wide Web Conference (WWW), pp. 271–279 (2003)
13. Li, P., Yau, S.-T.: On the parabolic kernel of the Schrödinger operator. Acta Math. **156**(3–4), 153–201 (1986)
14. Page, L., Brin, S., Motwani, R., Winograd, T.: The pagerank citation ranking: bringing order to the web. In: Proceedings of the 7th International World Wide Web Conference, Brisbane, Australia, pp. 161–172 (1998)