

Analysis of Relaxation Time in Random Walk with Jumps

Konstantin Avrachenkov, Ilya Bogdanov

► **To cite this version:**

Konstantin Avrachenkov, Ilya Bogdanov. Analysis of Relaxation Time in Random Walk with Jumps. Algorithms and Models for the Web Graph. WAW 2018., Anthony Bonato; Pawel Pralat; Andrei Raigorodskii, May 2018, Moscow, Russia. pp.70-82. hal-01936065

HAL Id: hal-01936065

<https://hal.inria.fr/hal-01936065>

Submitted on 27 Nov 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Analysis of Relaxation Time in Random Walk with Jumps

Konstantin Avrachenkov¹ and Ilya Bogdanov²

¹ Inria Sophia Antipolis, France
k.avrachenkov@inria.fr

² Higher School of Economics, Russia
ilya160897@gmail.com

Abstract. We study the relaxation time in the random walk with jumps. The random walk with jumps combines random walk based sampling with uniform node sampling and improves the performance of network analysis and learning tasks. We derive various conditions under which the relaxation time decreases with the introduction of jumps.

Keywords: Random walk on graph, random walk with jumps, relaxation time, spectral gap, network sampling, network analysis, learning on graphs.

1 Introduction

In the present work we study the relaxation time or equivalently the spectral gap of the random walk with jumps on a general weighted undirected graph. The random walk with jumps can be viewed as a random walk on a combination of the original graph and the complete graph weighted by a scaled parameter. This parameter determines the rate of jumps from the current node to an arbitrary node. The random walk with jumps has similarities with PageRank [19]. In fact, it coincides with PageRank on the regular graphs but differs on the irregular graphs. In the case of the random walk with jumps, the jump probability depends on the node degree. The higher the degree of the current node is, the less likely the random walk will jump out of the node to an arbitrary node. The random walk with jumps can also be viewed as a particular case of the generalisation of PageRank with node-dependent restart [6].

The random walk with jumps has been introduced in [5] to improve the random walk based network sampling or respondent driven sampling [22] by combining the standard random walk based sampling with uniform node sampling. A big advantage of the random walk with jumps in comparison with PageRank is that, as opposite to PageRank, the random walk with jumps on an undirected graph is a time-reversible process and its stationary distribution is available in a simple, explicit form. In particular, this allows us to unbiasedly sample the random walk, which is node degree biased on irregular graphs. This comes with a price. The price is the difficulty to control the relaxation time. In the case of

PageRank the relaxation time is bounded by the reciprocal of the restart probability [10, 16]. In the case of the random walk with jumps, there is no simple connection with the jump parameter. In [5] the authors have shown that under a natural condition on the clustering structure of the graph, the relaxation time decreases with the increase of the jump parameter.

Let us mention a few more applications of the random walk with jumps beyond network sampling [5, 18, 20]. The random walk with jumps has been used in the context of graph-based semi-supervised learning [2, 15]. The random walk with jumps has also been used as a main building block in the quick algorithm for finding largest degree nodes in a graph [4]. A continuous-time version of the random walk with jumps has an application in epidemiology [13].

All this motivates us to take another look at the relaxation time of the random walk with jumps. In particular, we are now able to give a necessary and sufficient condition for the improvement of the relaxation time on weighted graphs. We give an example showing that there are weighted graphs where introducing jumps increases the relaxation time. The necessary and sufficient conditions are not easy to interpret. Therefore we derive a series of simpler sufficient conditions. One new sufficient condition, similar in spirit to the condition in [5], indicates that on graphs with clusters, the relaxation time improves with the introduction of jumps. The other new sufficient conditions require the spectral gap of the original graph to be smaller than the reciprocal of the squared coefficient of variation of the nodes' degrees, thus establishing a connection between the measure of graph irregularity and the relaxation time. We expect that the derived conditions are satisfied in most complex networks (either due to clustering structure or due to small spectral gap). Thus, the present study confirms that it is safe and in most cases beneficial to use the random walk with jumps for complex network analysis.

The structure of the paper is as follows: in the next section we define the random walk with jumps and provide necessary background material. In Section 3 we discuss the application of Dobrushin coefficient for large jump rates. Then, in Section 4 we discuss necessary and sufficient conditions for the reduction of the relaxation time when the jump rate is small. In Section 5 we provide a series of sufficient conditions, which are easier to interpret and to verify. In particular, we provide a sufficient condition in terms of the coefficient of variation of nodes' degrees. In Section 6 we give interesting numerical illustrations. We conclude the paper with a conjecture in Section 7.

2 Definitions and preliminaries

Most of the analysis in the present article is for the case of a general weighted undirected graph G with vertex set $V(G)$, $|V(G)| = n$, and edge set $E(G)$, $|E| = m$, defined by the *weighted adjacency matrix* $A = (a_{ij})$ with elements

$$a_{ij} = \begin{cases} \text{weight of edge } (i, j), & \text{if } i \sim j, \\ 0, & \text{otherwise.} \end{cases}$$

Unless stated otherwise, we assume that G is connected.

Denote by $\underline{1}$ the column vector of ones of appropriate dimension. Then, $d = A\underline{1}$ is the vector of weighted degrees of vertices and $D = \text{Diag}(d)$ is the diagonal matrix with vector d on the main diagonal.

The Standard Random Walk (SRW) is a discrete-time Markov chain $\{X_t, t = 0, 1, \dots\}$ on the vertex set $V(G)$ with the transition probability matrix

$$P = D^{-1}A, \quad (1)$$

whose elements are

$$p_{ij} = P[X_{t+1} = j | X_t = i] = \begin{cases} w_{ij}/d_i, & \text{if } i \sim j, \\ 0, & \text{otherwise.} \end{cases}$$

SRW is a time-reversible process with the stationary distribution

$$\pi_i = \frac{d_i}{2m}, \quad i = 1, \dots, n. \quad (2)$$

Since SRW is time-reversible, its transition matrix is similar to a symmetric matrix and hence the eigenvalues of the transition matrix are real, semi-simple and can be indexed as follows:

$$1 = \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \geq -1.$$

Denote λ_* the maximum modulus eigenvalue of P different from $+1$ and -1 and call $\gamma(P) = 1 - |\lambda_*|$ the spectral gap.

The relaxation time t_{rel} is then defined by

$$t_{rel} = \frac{1}{\gamma(P)}. \quad (3)$$

One interpretation of the relaxation time is as follows [17]: if $t \geq t_{rel}$, then the standard deviation of $P^t f$ is bounded by $1/e$ times the standard deviation of f . Also, the following inequality (see, e.g., [1, 9, 17]) indicates a strong relation between the relaxation time and mixing time

$$(\log(1/\varepsilon) + \log(1/2))(t_{rel} - 1) \leq t_{mix}(\varepsilon) \leq (\log(1/\varepsilon) + \log(1/\pi_{min}))t_{rel}.$$

In particular, the above inequality suggests that for a finite Markov chain and for small enough ε , the ε -mixing time is very close to $\log(1/\varepsilon)t_{rel}$.

Now let us define the random walk with jumps (RWJ). It is a random walk on a combination of the original graph and the complete graph weighted by a scaled parameter α/n [5]. Specifically, let us modify the adjacency matrix in the following way

$$A(\alpha) = A + \frac{\alpha}{n}\underline{1}\underline{1}^T. \quad (4)$$

Note that the new degree matrix is given by $D(\alpha) = D + \alpha I$, where I is the identity matrix. Then, the random walk on the modified graph is described by the following transition probability matrix

$$P(\alpha) = D^{-1}(\alpha)A(\alpha), \quad (5)$$

with elements

$$p_{ij}(\alpha) = \begin{cases} \frac{w_{ij} + \alpha/n}{d_i + \alpha}, & \text{if } i \sim j, \\ \frac{\alpha/n}{d_i + \alpha}, & \text{otherwise,} \end{cases}$$

Since RWJ is again a random walk on a weighted undirected graph, it is time-reversible Markov chain with semi-simple eigenvalues. The stationary distribution of RWJ also has a simple form

$$\pi_i(\alpha) = \frac{d_i + \alpha}{2m + \alpha n}, \quad i = 1, \dots, n. \quad (6)$$

The modified transition matrix $P(\alpha)$ can be rewritten as follows:

$$P(\alpha) = (D + \alpha I)^{-1} D P + (D + \alpha I)^{-1} \alpha I \underline{1} \left(\frac{1}{n} \underline{1}^T \right). \quad (7)$$

We note that if the graph is regular, i.e., $D = dI$, the above expression reduces to

$$P(\alpha) = \frac{d}{d + \alpha} P + \frac{\alpha}{d + \alpha} \underline{1} \left(\frac{1}{n} \underline{1}^T \right),$$

which is the transition matrix for PageRank (PR) [19] with the damping factor $d/(d + \alpha)$. Thus, in the case of a regular graph RWJ coincides with PR. However, when the graph has inhomogeneous degrees, these two concepts are different. From the equivalence of RWJ to PR on the regular graph, we can immediately conclude that the relaxation time is monotonously decreasing with α when the graph is regular. When the graph is irregular, the situation becomes much more complex.

We also would like to note that RWJ can be viewed as a node-dependent PageRank [6], where the restart probability at each node is given by $\alpha/(d_i + \alpha)$. Thus, in contrast to PR, RWJ restarts with smaller probabilities from higher degree nodes.

3 Application of Dobrushin coefficient

The Dobrushin ergodic coefficient (see, e.g., [8, 12, 21]) can be used to obtain a lower bound on the spectral gap $\gamma(P)$ of a Markov chain. The Dobrushin ergodic coefficient is given by

$$\delta(P) = \frac{1}{2} \max_{i,j \in V} \sum_{k \in V} |p_{ik} - p_{jk}|, \quad (8)$$

or, equivalently,

$$\delta(P) = 1 - \min_{i,j \in V} \sum_{k \in V} p_{ik} \wedge p_{jk}. \quad (9)$$

In the case of RWJ, we can obtain a simple upper bound on $\delta(P)$ by taking the smallest element in the transition probability matrix. Namely, we obtain

$$\delta(P) \leq 1 - \min_{i,j \in V} \sum_{k \in V} \frac{\alpha/n}{d_{max} + \alpha} = 1 - \frac{\alpha}{d_{max} + \alpha}, \quad (10)$$

where d_{max} is the maximal degree in the graph. Since $\gamma(P) \geq 1 - \delta(P)$, we also obtain a lower bound for the spectral gap

$$\gamma(P(\alpha)) \geq \frac{\alpha}{d_{max} + \alpha}. \quad (11)$$

And since $\alpha/(d_{max} + \alpha) \rightarrow 1$ as $\alpha \rightarrow \infty$, we have

Proposition 1 *For any undirected graph G , there always exists $\bar{\alpha} = \bar{\alpha}(G)$ such that for all $\alpha > \bar{\alpha}$, we have $\gamma(P(\alpha)) > \gamma(P)$.*

For the regular graphs with degree d , the bound $\alpha/(d + \alpha)$ is quite tight. In fact, for the regular graphs, as was noted at the end of the previous section $\frac{\alpha}{d+\alpha}$ corresponds to the restart probability and the exact value of the spectral gap is $\frac{d}{d+\alpha}\lambda_*(P)$ [10, 16]. However, we have observed that for irregular graphs the bound (11) can be very loose.

4 Conditions in the case of small jump rate

Let us analyse in this section the effect of small jump rate on the relaxation time.

Denote by $v_*(\alpha)$ the eigenvector corresponding to $\lambda_*(\alpha)$, that is

$$P(\alpha)v_*(\alpha) = \lambda_*(\alpha)v_*(\alpha). \quad (12)$$

For brevity, we shall write $v_* = v_*(0)$ and $\lambda_* = \lambda_*(0)$. We also need the following preliminary result

Lemma 1 *The eigenelements $v_*(\alpha)$ and $\lambda_*(\alpha)$ are analytic functions with respect to α .*

Proof: It is known from [14, Chapter 2] (see also [3, 7]) that if the eigenvalues of the perturbed matrix are semi-simple, then the eigenvalues as well as eigenvectors can be expanded as power series with positive integer powers. Since in our case, RWJ is time-reversible, its eigenvalues are semi-simple and the statement of the lemma follows. \square

Next, we are in a position to provide necessary and sufficient conditions for the improvement of the relaxation time for small jump rates.

Theorem 1. *For sufficiently small α , in the case $\lambda_* < 0$, the spectral gap increases, and equivalently, the relaxation time decreases with respect to α .*

If $\lambda_ > 0$, for sufficiently small α the necessary and sufficient condition for the decrease of the relaxation time is*

$$\frac{1}{n}(\mathbf{1}^T v_*)^2 < \lambda_* v_*^T v_*. \quad (13)$$

In addition, we have the following asymptotics:

$$\lambda_*(\alpha) = \lambda_*(0) + \alpha \frac{\frac{1}{n}(v_*^{(0)})^T \mathbf{1} \mathbf{1}^T v_*^{(0)} - \lambda_*^{(0)}(v_*^{(0)})^T v_*^{(0)}}{(v_*^{(0)})^T D v_*^{(0)}} + O(\alpha^2).$$

Proof: According to Lemma 1, the eigenlements $\lambda_*(\alpha)$ and $v_*(\alpha)$ can be expanded as power series

$$\lambda_*(\alpha) = \lambda_*^{(0)} + \alpha\lambda_*^{(1)} + O(\alpha^2), \quad (14)$$

$$v_*(\alpha) = v_*^{(0)} + \alpha v_*^{(1)} + O(\alpha^2), \quad (15)$$

for sufficiently small α . If we set $\alpha = 0$, we obtain $\lambda_*^{(0)} = \lambda_*(0) = \lambda_*$ and $v_*^{(0)} = v_*(0) = v_*$ such that

$$D^{-1}Av_* = \lambda_*v_*,$$

or equivalently,

$$Av_* = \lambda_*Dv_*. \quad (16)$$

It is also convenient to rewrite (12) as a generalized eigenvalue problem

$$(A + \frac{\alpha}{n}\mathbf{1}\mathbf{1}^T)v_*(\alpha) = \lambda_*(\alpha)(D + \alpha I)v_*(\alpha). \quad (17)$$

Substituting the power series (14) and (15) into equation (17) and equating coefficients in α -terms, i.e.,

$$\begin{aligned} (A + \frac{\alpha}{n}\mathbf{1}\mathbf{1}^T)(v_*^{(0)} + \alpha v_*^{(1)} + O(\alpha^2)) = \\ (\lambda_*^{(0)} + \alpha\lambda_*^{(1)} + O(\alpha^2))(D + \alpha I)(v_*^{(0)} + \alpha v_*^{(1)} + O(\alpha^2)), \end{aligned}$$

yields

$$\frac{1}{n}\mathbf{1}\mathbf{1}^T v_*^{(0)} + Av_*^{(1)} = \lambda_*^{(0)}Iv_*^{(0)} + \lambda_*^{(1)}Dv_*^{(0)} + \lambda_*^{(0)}Dv_*^{(1)}. \quad (18)$$

Now let us multiply the above equation by v_*^T from the left to obtain

$$\begin{aligned} \frac{1}{n}(v_*^{(0)})^T\mathbf{1}\mathbf{1}^T v_*^{(0)} + (v_*^{(0)})^T Av_*^{(1)} = \\ (v_*^{(0)})^T\lambda_*^{(0)}v_*^{(0)} + (v_*^{(0)})^T\lambda_*^{(1)}Dv_*^{(0)} + (v_*^{(0)})^T\lambda_*^{(0)}Dv_*^{(1)}. \end{aligned} \quad (19)$$

Due to symmetry, the equation (16) can be rewritten as

$$v_*^T A = \lambda_*v_*^T D,$$

and hence

$$(v_*^{(0)})^T Av_*^{(1)} = (v_*^{(0)})^T\lambda_*^{(0)}Dv_*^{(1)},$$

which simplifies (19) to

$$\frac{(v_*^{(0)})^T\mathbf{1}\mathbf{1}^T v_*^{(0)}}{n} = (v_*^{(0)})^T\lambda_*^{(0)}v_*^{(0)} + (v_*^{(0)})^T\lambda_*^{(1)}Dv_*^{(0)}.$$

Thus, $\lambda_*^{(1)}$ can be expressed as

$$\lambda_*^{(1)} = \frac{\frac{1}{n}(v_*^{(0)})^T\mathbf{1}\mathbf{1}^T v_*^{(0)} - \lambda_*^{(0)}(v_*^{(0)})^T v_*^{(0)}}{(v_*^{(0)})^T Dv_*^{(0)}}. \quad (20)$$

Clearly, the denominator in (20) is always positive. Now, consider two cases:

Case 1. $\lambda_* < 0$

In this case the numerator in (20) is always positive. Then $\lambda_*^{(1)}$ is also positive. By expansion (14), when α is sufficiently small, the absolute value of λ_* is decreasing with respect to α .

Case 2. $\lambda_* > 0$

Again, by expansion (14), for sufficiently small α the value of $\lambda_*(\alpha)$ decreases in α if and only if $\lambda_*^{(1)}$ is negative, i.e., when the numerator is negative. This is precisely what was stated in the theorem's condition. \square

5 Sufficient conditions with easier interpretation

Even though the condition in Theorem 1 is necessary and sufficient, it is not easy to use and does not have an easy intuitive interpretation. Next, we shall derive a series of sufficient conditions with easier interpretation and verification. Towards this goal, let us transform the condition in Theorem 1 to an equivalent form using the combinatorial Laplacian. Specifically, we shall use the combinatorial Laplacian of the complete graph:

$$L_K = nI - \underline{1}\underline{1}^T.$$

Lemma 2 *The condition (13) is equivalent to*

$$1 - \lambda_* < \frac{v_*^T L_K v_*}{n v_*^T v_*},$$

and if $\lambda_ > 0$, the condition (13) is equivalent to*

$$\gamma(P) < \frac{\sum_{i,j} (v_{*i} - v_{*j})^2}{n \sum_i v_{*i}^2}. \quad (21)$$

Proof: Using the definition of the Laplacian, we can write

$$\frac{1}{n} (\underline{1}^T v_*)^2 = \frac{1}{n} v_*^T \underline{1}\underline{1}^T v_* = \frac{1}{n} v_*^T (nI - L_K) v_* = v_*^T v_* - \frac{1}{n} v_*^T L_K v_*.$$

Thus, we can rewrite condition (13) in a new form:

$$1 - \lambda_* < \frac{v_*^T L_K v_*}{n v_*^T v_*}.$$

The other equivalent form follows immediately from the definitions of the spectral gap and L_K . \square

Next we provide a couple of sufficient conditions with easy interpretation in the context of complex networks.

Corollary 1 *If*

$$\gamma(P) < \frac{1}{n}, \quad (22)$$

then, for sufficiently small $\alpha > 0$, the spectral gap of $P(\alpha)$ is larger than the spectral gap of P .

Proof: Clearly, $\sum_{i,j} (v_{*i}(\alpha) - v_{*j}(\alpha))^2 > \sum_i v_{*i}(\alpha)^2$. Because $v_*(\alpha) \perp \pi(\alpha)$ and so there are both positive and negative numbers among $v_{*i}(\alpha)$. Each number $v_{*i}(\alpha)$ has at least one number of opposite sign $v_{*j}(\alpha)$ and such that $(v_{*i}(\alpha) - v_{*j}(\alpha))^2 > v_{*i}(\alpha)^2$. \square

The above corollary has the following simple and useful interpretation: If P has a sufficiently small gap, the addition of jumps with small rate always improves the relaxation time. It is known that many complex networks have very small spectral gap and thus this corollary gives an explanation why the relaxation time typically improves with the addition of jumps in complex networks. The next corollary refines the above argument.

Corollary 2 *Denote the proportion of negative and positive $v_{*i}(\alpha)$ as $\mu : 1 - \mu$. Then, for sufficiently small α , if*

$$\gamma(P) < \min(\mu, 1 - \mu), \quad (23)$$

the relaxation time decreases with respect to α .

Proof: For each positive number $v_{*i}(\alpha)$ has μn numbers of opposite sign $v_{*j}(\alpha)$ such that $(v_{*i}(\alpha) - v_{*j}(\alpha))^2 > v_{*i}(\alpha)^2$. Analogously, each negative number $v_{*i}(\alpha)$ has $(1 - \mu)n$ numbers of opposite sign $v_{*j}(\alpha)$ and for them $(v_{*i}(\alpha) - v_{*j}(\alpha))^2 > v_{*i}(\alpha)^2$. So, $\sum_{i,j} (v_{*i}(\alpha) - v_{*j}(\alpha))^2 \geq n \min(\mu, 1 - \mu) \sum_i v_{*i}(\alpha)^2$. \square

As a result of Corollary 2, the closer μ is to $1/2$, the better. Often complex networks have clustering structure. The eigenvector v_* can be interpreted as a variant of the Fiedler vector. Thus, if a complex network can be divided into two clusters of similar sizes, the value of μ will be close to $1/2$ or at least far from zero. In such a case, the spectral gap of the original transition matrix P does not need to be small for the addition of jumps to improve the relaxation time. The above statement is similar in spirit to the condition given in [5].

Let us now provide a sufficient condition in terms of nodes' degrees.

Theorem 2. *Let the following condition hold:*

$$\gamma(P) < 4 \frac{(d_1 + d_2 + \dots + d_n)^2}{n(d_1^2 + d_2^2 + \dots + d_n^2)}. \quad (24)$$

Then, if $\alpha > 0$ is sufficiently small, $\gamma(P(\alpha)) > \gamma(P)$.

Proof: We replace the condition (21) with the following more stringent condition:

$$\gamma(P) < \min_{\mathbf{f} \perp \pi} \frac{\sum_{i \sim j} (f_i - f_j)^2}{n \sum f_i^2}. \quad (25)$$

Let us find the minimum of the RHS of the above expression over all vectors \mathbf{f} , orthogonal to $\pi(\alpha)$, i.e., for vectors satisfying $\sum f_i d_i = 0$. So it will be enough for $\gamma(P)$ to be less than the RHS of the condition (25).

Note that the RHS of (25) is homogeneous in \mathbf{f} , so without loss of generality we can set $\sum f_i^2 = 1$.

We shall use the method of Lagrange multipliers for the objective function $\sum_{i,j} (f_i - f_j)^2$ and the following constraints:

$$\sum_{i=1}^n f_i^2 = 1, \quad (26)$$

$$\sum_{i=1}^n f_i d_i = 0. \quad (27)$$

The expression $\sum_{i,j} (f_i - f_j)^2$ taking into account the latter constraint can be rewritten as $2n - 2(f_1 + f_2 + \dots + f_n)^2$. Therefore we are looking for the maximum of $(f_1 + f_2 + \dots + f_n)^2$, which will give us the minimum of $\sum_{i,j} (f_i - f_j)^2$. Let us find the maximal absolute values of the function $(f_1 + f_2 + \dots + f_n)$ with respect to the above constraints. Obviously, these extreme values exist (sums cannot be more than n and less than $-n$). Furthermore, their absolute values coincide, because domains of these two sums is symmetric with the center in the origin. So the set of values of $(f_1 + f_2 + \dots + f_n)$ is centrally symmetric and the maximal absolute values of $(f_1 + f_2 + \dots + f_n)$ and $-(f_1 + f_2 + \dots + f_n)$ are the same.

Looking at the geometry of our maximization problem, we can see that the constraints give us a sphere cut by the hyperplane $\sum f_i d_i = 0$. We are looking for the touching points of this set and some hyperplane $(f_1 + f_2 + \dots + f_n) = c$. Clearly, there will be exactly two centrally-symmetric points of touching, giving us the maxima of the absolute values (the domain of this sum is symmetric with the center of the symmetry in the origin, so the set of values of $(f_1 + f_2 + \dots + f_n)$ is centrally symmetric).

Now, consider the Lagrange function:

$$L = (f_1 + f_2 + \dots + f_n) - \lambda_1 \left(\sum_i f_i^2 - 1 \right) - \lambda_2 \sum f_i d_i. \quad (28)$$

Taking partial derivatives with respect to f_i for $i = 1, \dots, n$, we obtain

$$0 = 1 - 2\lambda_1 f_i - \lambda_2 d_i. \quad (29)$$

Firstly, multiply all these equations with d_i , respectively, and sum all of them (here we also use equation (27)). The result is

$$0 = 2m - \lambda_2 \sum_i d_i^2. \quad (30)$$

Thus, $\lambda_2 = \frac{2m}{\sum_i d_i^2}$, where m is the total number of edges.

Secondly, multiply all equations with f_i , respectively, and sum all of them (here we additionally use equations (26) and (27)). The result is

$$0 = S - 2\lambda_1, \quad (31)$$

where $S = (f_1 + f_2 + \dots + f_n)$. Thus, $\lambda_1 = S/2$.

And finally, just sum all equations (here we also use equations (26) and (27)) to get

$$0 = n - 2\lambda_1 S - \lambda_2(2m). \quad (32)$$

If we know λ_1 and λ_2 , we also know $S^2 = n - \frac{8m^2}{\sum_i d_i^2}$. Substituting this into our initial expression that we want to minimize, we obtain

$$\min_{f \perp \pi} \frac{\frac{1}{n} \sum_{i,j} (f_i - f_j)^2}{\sum f_i^2} = \frac{1}{n} \left(2n - 2 \left(n - \frac{8m^2}{\sum_i d_i^2} \right) \right) = 4 \frac{(d_1 + d_2 + \dots + d_n)^2}{n(d_1^2 + d_2^2 + \dots + d_n^2)}. \quad (33)$$

Thus, if inequality (25) holds, the condition of Theorem 1 also holds, and hence $\gamma(P(\alpha)) > \gamma(P)$ for sufficiently small α . \square

We can also rewrite the condition of the above theorem as follows:

Corollary 3 *Let the following condition hold:*

$$\gamma(P) < 4 \frac{(\bar{d})^2}{\bar{d}^2}, \quad (34)$$

where

$$\bar{d} = \frac{d_1 + d_2 + \dots + d_n}{n},$$

and

$$\bar{d}^2 = \frac{d_1^2 + d_2^2 + \dots + d_n^2}{n}.$$

Then, if $\alpha > 0$ is sufficiently small, $\gamma(P(\alpha)) > \gamma(P)$.

The quantity $(\bar{d})^2/\bar{d}^2$ is the reciprocal of the efficiency or of the squared coefficient of variation. It is also sometimes referred to as the signal-to-noise ratio. From Corollary 3 we can see that the more ‘‘irregular’’ the degree sequence is, the more stringent is the condition on the spectral gap.

The following simple sufficient condition using only \bar{d} and d_{max} also holds.

Corollary 4 *For sufficiently small α , $\gamma(P(\alpha)) > \gamma(P)$ is true for all graphs such that*

$$\gamma(P) < 4\bar{d}/d_{max}. \quad (35)$$

Proof: The RHS of the inequality (34) is greater than \bar{d}/d_{max} . \square

6 Examples

Let us first demonstrate in this section that there exist weighted graphs for which the introduction of jumps actually increases the relaxation time. Towards this end, we consider a weighted graph of size 2 with the following adjacency matrix:

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix}.$$

The characteristic equation for the generalized eigenvalue problem $Av = \lambda Dv$ takes the form

$$\det(A - \lambda D) = \det \begin{bmatrix} a_{11} - \lambda(a_{11} + a_{12}) & a_{12} \\ a_{12} & a_{22} - \lambda(a_{22} + a_{12}) \end{bmatrix} = 0.$$

It has two solutions

$$\lambda_1 = 1 \quad \text{and} \quad \lambda_* = \frac{\det(A)}{(a_{11} + a_{12})(a_{22} + a_{12})}.$$

The eigenvector corresponding to λ_* is given by

$$v_* = \begin{bmatrix} 1 \\ -\frac{a_{11} + a_{12}}{a_{22} + a_{12}} \end{bmatrix} C.$$

Let us calculate the numerator in the expression (20)

$$\begin{aligned} & \frac{1}{2}(v_*^T \mathbf{1})^2 - \lambda_* v_*^T v_* = \\ & \frac{(a_{22} - a_{11})^2(a_{11} + a_{12})(a_{22} + a_{12}) - 2 \det(A)[(a_{11} + a_{12})^2 + (a_{22} + a_{12})^2]}{2(a_{11} + a_{12})(a_{22} + a_{12})^3}. \end{aligned}$$

Now if we choose the weights such that $\det(A)$ is close to zero or just zero and $a_{11} \neq a_{22}$, the spectral gap will decrease, and consequently the relaxation time will increase when α increases from zero. Thus, in general, the conditions of Theorem 1 need to be checked in the case of weighted graphs if one wants to be sure that the relaxation time decreases with the increase of the jump rate.

We also tried to construct an example of unweighted graph when the relaxation time increases with the increase in the jump rate. Somewhat surprisingly, it appears to be hard to construct such an example. In fact, we have checked all non-isomorphic graphs of sizes up to $n \leq 9$ from the Brendan McKay's collection [23] and could not find a single example when the relaxation time increases with the introduction of small jump rate. The code of this verification can be found at [24].

We have also checked various random graph models like Erdős-Rényi graph or non-homogeneous Stochastic Block Model and could not find an example of an unweighted graph for which the relaxation time increases with the introduction of jumps.

7 Conclusion and future research

We have analysed the spectral gap, or equivalently the relaxation time, in the random walk with jumps. We have obtained a necessary and sufficient condition for the decrease of the relaxation time when the jump rate increases from zero. These conditions unfortunately do not have easy interpretation. Therefore, we have proceeded with the derivation of several sufficient conditions with easy interpretation. Some of these sufficient conditions can also be easily verified. The derived sufficient conditions show that in most complex networks the relaxation time should decrease with the introduction of jumps. We have also demonstrated that there exist weighted graphs for which the relaxation time increases with the introduction of jumps. On the other hand, we could not find such an example in the case of unweighted graphs. At the moment, we tend to conjecture that the introduction of jumps always improves the relaxation time in the case of unweighted graphs.

Acknowledgements

The work was partially supported by the joint laboratory Nokia Bell Labs - Inria. This is an author version of the paper accepted to WAW 2018.

References

1. Aldous, D. and Fill, J., 2002. *Reversible Markov chains and random walks on graphs*. Monograph in preparation.
Available at <http://www.stat.berkeley.edu/~aldous/RWG/book.html>
2. Avrachenkov, K., Chebotarev, P. and Mishenin, A., 2017. Semi-supervised learning with regularized Laplacian. *Optimization Methods and Software*, 32(2), pp. 222-236.
3. Avrachenkov, K., Filar J.A. and Howlett P.G., 2013. *Analytic perturbation theory and its application*, SIAM.
4. Avrachenkov, K., Litvak, N., Sokol, M. and Towsley, D., 2014. Quick detection of nodes with large degrees. *Internet Mathematics*, 10(1-2), pp. 1-19.
5. Avrachenkov, K., Ribeiro, B. and Towsley, D., 2010. Improving random walk estimation accuracy with uniform restarts. In *Proceedings of International Workshop on Algorithms and Models for the Web-Graph* (pp. 98-109).
6. Avrachenkov, K., van der Hofstad, R. and Sokol, M., 2014. Personalized PageRank with node-dependent restart. In *Proceedings of International Workshop on Algorithms and Models for the Web-Graph* (pp. 23-33).
7. Baumgärtel, *Analytic perturbation theory for matrices and operators*, 1985. Birkhäuser.
8. Brémaud, P., 1999. *Markov chains: Gibbs fields, Monte Carlo simulation, and queues*. Springer.
9. Chung, F., 1997. *Spectral graph theory*. American Math. Soc.
10. Haveliwala, T. and Kamvar, S., 2003. The second eigenvalue of the Google matrix. Stanford Technical Report.

11. Horn, R.A. and Johnson, C.R., 2013. *Matrix analysis*. 2nd edition. Cambridge University Press.
12. Ipsen, I.C.F. and Selee, T.M., 2011. Ergodicity coefficients defined by vector norms. *SIAM Journal on Matrix Analysis and Applications*, 32(1), pp.153-200.
13. Jacobsen, K.A. and Tien, J.H., 2018. A generalized inverse for graphs with absorption. *Linear Algebra and its Applications*, 537, pp.118-147.
14. Kato, T., 1995. *Perturbation theory for linear operators*. Springer.
15. Kondor, R.I. and Lafferty, J., 2002. Diffusion kernels on graphs and other discrete input spaces. In *Proceedings of ICML* (pp. 315-322).
16. Langville, A.N. and Meyer, C.D., 2006. *Google's PageRank and beyond: The science of search engine rankings*. Princeton University Press.
17. Levin, D.A., Peres, Y. and Wilmer E.L., 2008. *Markov chains and mixing times*. American Math. Soc.
18. Murai, F., Ribeiro, B., Towsley, D. and Wang P., 2017. Characterizing directed and undirected networks via multidimensional walks with jumps. ArXiv preprint arXiv:1703.08252.
19. Page, L., Brin, S., Motwani, R. and Winograd, T., 1999. The PageRank citation ranking: Bringing order to the web. Stanford InfoLab Research Report.
20. Ribeiro, B. and Towsley, D., 2010. Estimating and sampling graphs with multidimensional random walks. In *Proceedings of the 10th ACM SIGCOMM conference on Internet measurement*, (pp. 390-403).
21. Seneta, E., 2006. *Non-negative matrices and Markov chains*. Revised printing edition. Springer.
22. Volz, E. and Heckathorn, D.D., 2008. Probability based estimation theory for respondent driven sampling. *Journal of Official Statistics*, 24(1), p.79.
23. Brendan McKay's graph collection.
Available at <http://users.cecs.anu.edu.au/~bdm/data/graphs.html>
24. GitHub code repository for the numerical experiments of the article.
Available at <https://github.com/ilya160897/Random-walk-with-jumps>