Diameter of random spanning trees in a given graph

Fan Chung * Paul Horn † Linyuan Lu ‡

June 30, 2008

Abstract

We show that a random spanning tree formed in a general graph G (such as a power law graph) has diameter much larger than the diameter of G. We show, with high probability the diameter of a random spanning tree of G is shown to be between $c\sqrt{n}$ and $c'\sqrt{n}\log n$, where c and c' depend on the spectral gap of G and the ratio of the moments of the degree sequence. For the special case of regular graphs, this result improves the previous lower bound by Aldous by a factor of $\log n$.

1 Introduction

Many information networks or social networks have very small diameters, as dictated by the so-called "small world phenomenon". However, in a recent paper by Liben-Nowell and Kleinberg [8], it was observed that in many social networks, typical spanning trees often have relatively large diameter. Examples of such spanning trees include those resulting from passing information to a small selected number of neighbors in various scenarios such as spam or gossip. A sparse subgraph naturally has very different behavior from its host graph. It is of interest to understand the connections between a graph and its subgraph. What invariants of the host graph can or cannot be translated to its subgraph? Under what conditions, can we predict the behavior of subgraphs? In particular, why does a random spanning tree have a large diameter while the opposite is true for the host graph? In this paper, we would like to address of this paradox by evaluating the diameter of random spanning sub-tree in G.

A spanning tree T of a connected graph G is a subgraph on $V(G)$, which is isomorphic to a tree. The number of spanning tree is determined by the celebrated matrix-tree theorem of Kirchoff [7]. Letting A denote the adjacency matrix and D denote the diagonal matrix of degrees, the matrix-tree theorem

*University of California, San Diego
†University of California, San Diego
‡University of South Carolina

states that the number of spanning tree is equal to the absolute value of the determinant of any $n-1 \times n-1$ sub-matrix of $D - A$.

The diameter of a subgraph is always larger than or equal to the diameter of G. However, the diameter of a spanning tree could be much larger than the diameter of the graph. The case that the host graph G is the complete graph K_n is well-studied in the literature. The number of spanning trees of K_n is n^{n-2} by Cayley's theorem. Rényi and Szekeres [11] showed that the diameter of a random spanning tree is of order $\sqrt{n}$, which contrasts with the fact that the diameter of K_n is 1.

Motivated by these examples, we ask what is true story of the diameter of random spanning trees for a general graph. Previously Aldous [1] proved that in a regular graph G with spectral bound σ (which will be defined later), the expected diameter of a spanning tree T of G, denoted by $diam(T)$ has expected value satisfying

$$\frac{c(1-\sigma)\sqrt{n}}{\log n} \leq \mathbb{E}(diam(T)) \leq \frac{c\sqrt{n}\log n}{\sqrt{1-\sigma}}$$

for some absolute constant c, where here (and throughout this paper) log refers to the natural logarithm.

We partially improve Aldous' result as follows:

Theorem 1. *For a d-regular graph G on n vertices with spectral gap σ, a spanning tree T of G has expected value satisfying*

$$c\sqrt{n} \leq \mathbb{E}(diam(T)) \leq c' \frac{\sqrt{n}\log n}{\sqrt{\log(1/\sigma)}}$$

for some absolute constants c and c' provided that $d \geq \frac{\log^2 n}{\log^2 \sigma}$.

Theorem 1 is an immediate consequence of the following result for general graphs.

Theorem 2. *In a connected graph G on n vertices, we assume that its average degree d, minimum degree δ, and second-order average degree $\tilde{d} = \sum_v d_v^2 / \sum_u d_u$ satisfy, for some given $\epsilon > 0$,*

$$d \gg \frac{\log^2 n}{\log^2 \sigma}.$$

Then with probability $1-\epsilon$, the diameter $diam(T)$ of a random spanning trees T in G satisfies

$$diam(T) \geq (1-\epsilon)\sqrt{\frac{\epsilon n d}{\tilde{d}}}. \tag{1}$$

and

$$diam(T) \leq \frac{c}{\epsilon}\sqrt{\frac{nd}{\delta \log(1/\sigma)}}\log n. \tag{2}$$

for some constant $c \leq 10$.

While the conditions look technical, they are derived from the proofs in Sections 4 and 5. We note that the average degree requirement is satisfied for any graph so long as, for instance, the average degree is $\Omega(\log^2 n)$ (a constant multiple of $\log^2 n$ for some constant), and $\sigma = o(1)$.

For random d-regular graphs, it is known that σ is about $\frac{2}{\sqrt{d}}$. We consider the random graph model $G(\mathbf{w})$ for a given expected degree sequence $\mathbf{w} = (w_1, w_2, \ldots, w_n)$, as introduced in [4]. The probability p_{ij} that there is an edge between v_i and v_j is proportional to the product $w_i w_j$ (as well as the loop at v_i with probability proportional to w_i^2). Namely,

$$p_{ij} = \frac{w_i w_j}{\sum_k w_k} = \frac{w_i w_j}{\mathrm{vol}(G)}. \tag{3}$$

It has been shown in [6] that $G(\mathbf{w})$ has $\sigma = (1 + o(1))\frac{4}{\sqrt{d}}$ provided that the minimum of weights is $\Omega(\log n)$. Theorem 2 implies the diameter of random spanning tree is $\Omega(\sqrt{\frac{d}{\tilde{d}} n})$ if the average degree is $d = \Omega((\frac{\log n}{\log \log n})^2)$. The upper bound is within a multiplicative factor of $\sqrt{\frac{\tilde{d}}{\delta}} \log n$.

It has been observed that many real-world information networks satisfy the so-called power law. We say a graph satisfies power law with exponent β if the degree sequence of the graph satisfies the property that the number of vertices having degree k is asymptotically proportional to $k^{-\beta}$. There are many models being used to capture the behavior of such power law graphs [5], especially for the exponent β in the range between 2 and 3. If we use the random graph model $G(\mathbf{w})$ with $\mathbf{w}$ satisfying the power law. In random graph model $G((w))$, the maximum degree can be as large as $\sqrt{n}$. (In other words, if the maximum degree exceeds $\sqrt{n}$, then $G(\mathbf{w})$ can only be used to model the subgraph with degree no larger than $\sqrt{n}$.) Also in $G(\mathbf{w})$ the second average degree is of order $d^{\beta-1} m^{3-\beta}$. Using Theorem 2, the diameter of a random spanning tree in such random power law graph is at least $cn^{(\beta-2)/4} (\log n)^{(2-\beta)/2}$ and at most $c' \sqrt{n} (\log n)^{3/2}$ for some constant c and c'.

The paper is organized as follows. In section 2, we will give definitions and prove some useful facts on the spectrum of the Laplacian, random walks, and spanning trees. In Section 3, we describe a method of using random walks to generate a uniform spanning tree. In Section 4, we will prove the lower bound for the diameter of a random spanning tree and give an upper bound in section 5.

2 Preliminaries

Suppose G is a connected (non-bipartite) graph on vertex set $[n] = \{1, 2, \ldots, n\}$. Let $A = (a_{ij})$ be adjacency matrix of G defined by

$$a_{ij} = \begin{cases} 1 & \text{if } ij \text{ is an edge;} \\ 0 & \text{otherwise.} \end{cases}$$

For $1 \leq i \leq n$ let $d_i = \sum_j a_{ij}$ be the degree of vertex i. Let $\Delta = \max(d_1, \ldots, d_n)$ be the maximum degree and $\delta = \min(d_1, \ldots, d_n)$ be the minimum degree. For each k, we define the k-th volume, closely related to the k-th moment of the degree sequence, of G to be

$$\mathrm{vol}_k(G) = \sum_{i=1}^{n} d_i^k.$$

The volume $\mathrm{vol}(G)$ is simply the sum of all degrees, i.e. $\mathrm{vol}(G) = \mathrm{vol}_1(G)$.

We define the average degree $d = \frac{1}{n}\mathrm{vol}(G) = \frac{\mathrm{vol}_1(G)}{\mathrm{vol}_0(G)}$ and the second order average degree $\tilde{d} = \frac{\mathrm{vol}_2(G)}{\mathrm{vol}_1(G)}$.

Let $D = diag(d_1, d_2, \ldots, d_n)$ denote the diagonal degree matrix. The Laplacian matrix is defined as

$$\mathcal{L} = I - D^{-\frac{1}{2}} A D^{-\frac{1}{2}}.$$

The spectrum of the Laplacian is the eigenvalues of $\mathcal{L}$ sorted in increasing order.

$$0 = \lambda_0 \leq \lambda_1 \leq \cdots \leq \lambda_{n-1}.$$

The first eigenvalue λ_0 is always equal to 0. $\lambda_1 > 0$ if G is connected and $\lambda_{n-1} \leq 2$ with equality holding only if G is bipartite graph.

Let $\sigma = \max\{1 - \lambda_1, \lambda_{n-1} - 1\}$. Thus $\sigma < 1$ if G is connected and non-bipartite. Note that σ is closely related to the mixing rate of random walks on G.

Let $\alpha_0, \alpha_1, \ldots, \alpha_{n-1}$ be orthonormal eigenvectors of the Laplacian $\mathcal{L}$, $U = (\alpha_0, \alpha_1, \ldots, \alpha_{n-1})$, where α_i is viewed as a column vector. Also we define $\Lambda = diag(\lambda_0, \ldots, \lambda_{n-1})$. We can write

$$\mathcal{L} = U \Lambda U^T.$$

For $0 \leq i \leq n-1$, we define $\phi_i = \alpha_i^T D \alpha_i$. Then we have

Lemma 1. *The degree spectrum $(\phi_0, \phi_1, \ldots, \phi_{n-1})$ satisfies the following properties.*

1. $\phi_1 = \tilde{d}$.
2. *For* $0 \leq i \leq n-1$, $\delta \leq \phi_i \leq \Delta$.
3. $\sum_{i=0}^{n-1} \phi_i = \mathrm{vol}(G)$.

Proof. Note $\alpha_0 = (\frac{\sqrt{d_1}}{\sqrt{\mathrm{vol}(G)}}, \ldots, \frac{\sqrt{d_n}}{\sqrt{\mathrm{vol}(G)}})^T$ since $\mathcal{L}\alpha_0 = 0$. We have

$$\begin{aligned}
\phi_0 &= \alpha_0^T D \alpha_0 \\
&= \sum_{i=1}^{n} \frac{\sqrt{d_i}}{\sqrt{\mathrm{vol}(G)}} d_i \frac{\sqrt{d_1}}{\sqrt{\mathrm{vol}(G)}} \\
&= \frac{\sum_{i=1}^{n} d_i^2}{\mathrm{vol}(G)} \\
&= \tilde{d}.
\end{aligned}$$

We have

$$
\begin{aligned}
|\phi_i - \frac{\delta+\Delta}{2}| &= |\alpha_i^T D \alpha_i - \frac{\delta+\Delta}{2}| \\
&= |\alpha_i^T (D - \frac{\delta+\Delta}{2} I)\alpha_i| \\
&\leq \|D - \frac{\delta+\Delta}{2} I\| \\
&= \frac{\Delta-\delta}{2}.
\end{aligned}
$$

Thus, we have

$$\delta \leq \phi_i \leq \Delta.$$

We also have

$$
\begin{aligned}
\sum_i \phi_i &= \mathrm{Tr}(U^T D U) \\
&= \mathrm{Tr}(D) \\
&= \mathrm{vol}(G).
\end{aligned}
$$

□

Lemma 2. *For any integer $j \geq 1$,*

$$\mathrm{Tr}(A(D^{-1}A)^{j-1}) \leq \tilde{d} + \sigma^j(\mathrm{vol}(G) - \tilde{d}).$$

Proof. We have

$$
\begin{aligned}
\mathrm{Tr}(A(D^{-1}A)^{j-1}) &= \mathrm{Tr}(D(D^{-1}A)^j) \\
&= \mathrm{Tr}(D(D^{-\frac{1}{2}}AD^{-\frac{1}{2}})^j \\
&= \mathrm{Tr}(D(I-\mathcal{L})^j) \\
&= \mathrm{Tr}(DU(I-\Lambda)^j U^T) \\
&= \mathrm{Tr}(U^T DU(I-\Lambda)^j) \\
&= \sum_{i=0}^{n-1} \phi_i (1-\lambda_i)^j \\
&= \tilde{d} + \sum_{i>0} \phi_i (1-\lambda_i)^j \\
&\leq \tilde{d} + \sum_{i>0} \phi_i \sigma^j \\
&= \tilde{d} + (\mathrm{vol}(G) - \tilde{d})\sigma^j.
\end{aligned}
$$

□

A simple random walk on G is a sequence of vertices $v_0, v_1, \ldots, v_k, \ldots$ with

$$\mathbb{P}(v_k = j \mid v_{k-1} = i) = p_{ij} = \begin{cases} \frac{1}{d_i} & \text{if } ij \in E(G) \\ 0 & \text{otherwise} \end{cases}$$

for all $k \geq 1$.

The transition matrix P is a $n \times n$ matrices with entries p_{ij} for $1 \leq i, j \leq n$. We can write $P = D^{-1}A$.

A probability distribution over the set of vertices is a row vector β ($\beta' \in \mathbb{R}^n$) satisfying

1. The entries of β are non-negative.
2. The L_1-norm $\|\beta\|_1$ ($= \beta\mathbf{1}$) equal to 1 where $\mathbf{1}$ denotes a column vector with all entries 1.

If β is a probability distribution, so is βP. The stationary distribution (if exists) is denoted by π satisfying $\pi = \pi P$ and

$$\pi = \frac{1}{\mathrm{vol}(G)}(d_1, d_2, \ldots, d_n).$$

The eigenvalues of P are $1, 1-\lambda_1, \ldots, 1-\lambda_{n-1}$, since $P = D^{-\frac{1}{2}}(I-\mathcal{L})D^{\frac{1}{2}}$. In general, P is not symmetric unless G is regular. The following lemma concerns the mixing rate of the random walks.

Lemma 3. *For any integer $t > 0$, any $\alpha \in \mathbb{R}^n$, and any two probability distributions β and γ, we have*

$$\langle (\beta-\gamma)P^t, \alpha D^{-1}\rangle \leq \sigma^t \|(\beta-\gamma)D^{-1/2})\| \|\alpha D^{-1/2}\|. \tag{4}$$

In particular,

$$\|(\beta-\gamma)P^t D^{-1/2}\| \leq \sigma^t \|(\beta-\gamma)D^{-1/2}\|. \tag{5}$$

Proof. Let $\varphi_0 = \frac{1}{\sqrt{\mathrm{vol}(G)}}(\sqrt{d_1}, \ldots, \sqrt{d_n}) = \mathrm{vol}(G)^{-\frac{1}{2}} D^{\frac{1}{2}}\mathbf{1}$ denote the (row) eigenvector of $I-\mathcal{L}$ for the eigenvalue 1. The matrix $(I-\mathcal{L})^t - \varphi^T\varphi$, which stands for the projection of $(I-\mathcal{L})^t$ to the hyperspace $\varphi^\perp$, has L_2-norm σ^t. Note that

$$(\beta-\gamma)D^{-\frac{1}{2}}\varphi = \frac{1}{\mathrm{vol}(G)}(\beta-\gamma)\mathbf{1} = 0.$$

We have

$$\begin{aligned}\langle (\beta-\gamma)P^t, D^{-1}\alpha\rangle &= (\beta-\gamma)D^{-\frac{1}{2}}[(I-\mathcal{L})^t - \xi\xi']D^{-\frac{1}{2}}\alpha \\ &\leq \|(\beta-\gamma)D^{-\frac{1}{2}}\|_2 \sigma^t \|D^{-\frac{1}{2}}\alpha\|_2.\end{aligned}$$

Now we choose $\alpha = [(\beta-\gamma)P^t]^T$, obtain (5) as desired. □

The mixing rate of the random walks on G measures how fast βP^t converges to the stationary distribution π from an initial distribution β. We can use the above lemma to show that the distribution βP^t converges to π rapidly if σ is strictly less than 1.

3 Random spanning trees generated by random walks

The following so-called groundskeeper algorithm gives a method of generating spanning trees: Start a random walk at a vertex, v. The first time a vertex is visited, we observe the edge it was visited on and add that edge to our spanning tree. Once the graph is covered, the resulting set of edges form a spanning tree. This gives a map Φ from random walks to random spanning trees. Aldous [1] and Broder [2] independently show that the groundskeeper algorithm generates a uniform spanning tree:

Theorem 3 (Groundskeeper Algorithm)**.** *The image of Φ is uniformly distributed over all spanning trees. It is independent of the choice of initial vertex v.*

We pick up an random initial vertex with stationary distribution π. Then at any step t, the distribution remains the same $p_t = \pi$.

For an integer $g \geq 3$, consider the following g-truncated random walks. We construct a random spanning tree by collecting edges $v_{t-1}v_t$ if v_t is first visited. We allow the backtrack step $v_{t+1} = v_{t-i}$ for some $i \leq g-2$. However, if $v_{t+1} = v_{t-i}$ for some $i > g-2$, the random walk stops.

Lemma 4. *The probability that a g-truncated random walk stops before or at time t is at most*

$$\frac{(t-g+3)(t-g+2)\tilde{d}}{2nd} + (t-k)\frac{\sigma^g}{1-\sigma}.$$

Proof. When the truncated random walk stops, there exists a closed walk $C = v_i, v_{i+1}, \ldots, v_t, v_{i+k} = v_i$ of length $k \geq g$ for some $0 \leq i \leq t-k+1$. For a fixed i and k, the probability $f(i,k)$ for such a closed walk is at most

$$\begin{aligned}
f(i,k) &\leq \sum_{\text{closed walk: } v_i,\ldots,v_{i+k}=v_i} \frac{d_i}{\mathrm{vol}(G)} \prod_{j=1}^{k} \frac{1}{d_{v_{i+j-1}}} \\
&= \frac{1}{\mathrm{vol}(G)} \mathrm{Tr}(A(D^{-1}A)^{k-1}) \\
&\leq \frac{\tilde{d}}{\mathrm{vol}(G)} + \sigma^k \left(1 - \frac{\tilde{d}}{\mathrm{vol}(G)}\right) \\
&< \frac{\tilde{d}}{\mathrm{vol}(G)} + \sigma^k.
\end{aligned}$$

By summing up for $i \geq 0$, $k \geq g$, and $i + k \leq t+1$, we have

$$\sum_{i=0}^{t-g+1} \sum_{k=g}^{t-i+1} f(i,k) = \sum_{i=0}^{t-g+1} \sum_{k=g}^{t-i+1} \frac{\tilde{d}}{\mathrm{vol}(G)} + \sigma^k$$

$$\leq \frac{(t-g+3)(t-g+2)}{2} \frac{\tilde{d}}{\mathrm{vol}(G)} + \sum_{i=0}^{t-g+1} \sum_{k=g}^{\infty} \sigma^k$$

$$\leq \frac{(t-g+3)(t-g+2))}{2} \frac{\tilde{d}}{\mathrm{vol}(G)} + (t-g+2)\frac{\sigma^g}{1-\sigma}.$$

□

4 Proving a diameter Lower Bound for random spanning trees

In this section we will prove a diameter lower bound for spanning trees of G as stated in inequality (1) of Theorem 2.

Proof of (1): Let $t = (1-\epsilon)\sqrt{\epsilon \frac{d}{\tilde{d}} n}$ and $g = \lceil \frac{\log\left(\frac{\epsilon(1-\sigma)\sqrt{\delta}}{4t\sqrt{\tilde{d}}}\right)}{\log(\sigma)} \rceil$. Note that g is chosen so that

$$\frac{\sigma^g}{1-\sigma} \leq \frac{\epsilon}{4t}.$$

Apply the g-truncated random walk. By Lemma 1, the g-truncated random walk will survive up to time t with probability at least

$$\begin{aligned} 1 - \frac{(t-g+3)(t-g+2)\tilde{d}}{2nd} - (t-g+2)\frac{\sigma^g}{1-\sigma} &> 1 - \frac{t^2\tilde{d}}{2nd} - t\frac{\sigma^g}{1-\sigma} \\ &> 1 - \frac{\epsilon}{2} - \frac{\epsilon}{4} \\ &\geq 1 - \frac{3\epsilon}{4}. \end{aligned}$$

For $i = 1, \ldots, t$, we say $v_{i-1}v_i$ is a forward step if $v_i \neq v_j$ for some $j < i$; we say $v_{i-1}v_i$ is a k-backward step if $v_i = v_{i-k}$ for some $k \leq g-2$.

Let $X_i = -k$ if $v_{i-1}v_i$ is a k-backward step and $X_i = 1$ otherwise. For all i, we have

$$-(g-2) \leq X_i \leq 1.$$

Let Y be the distance of v_0v_t in the random spanning tree and $X = \sum_{i=1}^t X_i$ Conditioning on that the truncated random walk survives up to time t, we have $Y \geq X$. Or equivalently,

$$\mathbb{P}(Y < X) < \frac{3\epsilon}{4}.$$

Let $\mathcal{F}_i$ be the σ-algebra that $v_0, \ldots, v_i$ is revealed. For $i = 0, \ldots, t$, $\mathbb{E}(X \mid \mathcal{F}_i)$ forms a martingale. We would like to establish a Lipschitz condition for this martingale. For $1 \leq i, j \leq t$, it is enough to bound $|\mathbb{E}(X_j \mid \mathcal{F}_i) - \mathbb{E}(X_j \mid \mathcal{F}_{i-1})|$. For $j < i$, X_j is completely determined by the information on $v_0, v_1, \ldots, v_i$. In this case we have

$$\mathbb{E}(X_j \mid \mathcal{F}_i) = \mathbb{E}(X_j \mid \mathcal{F}_{i-1}).$$

For $j \geq i$, $\mathbb{E}(X_j \mid \mathcal{F}_i)$ and $\mathbb{E}(X_j \mid \mathcal{F}_{i-1})$ are different because v_i is exposed. For $i \leq j \leq i+2g-3$, we apply the trivial bound

$$|\mathbb{E}(X_j \mid \mathcal{F}_i) - \mathbb{E}(X_j \mid \mathcal{F}_{i-1})| \leq g-1.$$

For $j \geq i+2g-2$, X_j only depends on $v_{j-g+2}, v_{j-g+3}, \ldots, v_{j+1}$. Note that the random walk at step i only depends on the current position v_i and is independent of history position $v_0, \ldots, v_{i-1}$. Thus $\mathbb{E}(X_j \mid v_{j-g+2})$ is independent of v_i because of $i < j-g+2$. We use the mixing of our random walk to show that information gained from knowing v_i is quickly lost. Let p be the distribution of v_i giving v_{i-1} and q be the distribution of v_i given v_i (q is a singleton distribution). Let p' be the distribution of v_{j-g+2} giving v_{i-1} and q' be the distribution of v_{j-g+2} giving v_{i-1}. We have

$$\|(p'-q')D^{-1/2}\| \leq \|(p-q)D^{-1/2}\|\sigma^{j-g+2-i} \leq \frac{2}{\sqrt{\delta}}\sigma^{j-g+2-i}.$$

Therefore,

$$\begin{aligned}
|\mathbb{E}(X_j \mid \mathcal{F}_i) - \mathbb{E}(X_j \mid \mathcal{F}_{i-1})| &= |\sum_{u=1}^{n}(p'_u - q'_u)\mathbb{E}(X_j \mid v_{j-g+2} = u)| \\
&\leq \|p'-q'\|_1(g-2) \\
&\leq \sqrt{\mathrm{vol}(G)}\|(p'-q')D^{-1/2}\|(g-2) \\
&\leq 2(g-2)\frac{\sqrt{\mathrm{vol}(G)}}{\sqrt{\delta}}\sigma^{j-g+2-i}.
\end{aligned}$$

We have

$$\begin{aligned}
|\mathbb{E}(X \mid \mathcal{F}_i) - \mathbb{E}(X \mid \mathcal{F}_{i-1})| &\leq \sum_{j=1}^{t}|\mathbb{E}(X_j \mid \mathcal{F}_i) - \mathbb{E}(X_j \mid \mathcal{F}_{i-1})| \\
&\leq 2(g-1)^2 + \sum_{j=i+2g-2}^{t} 2(g-2)\frac{\sqrt{\mathrm{vol}(G)}}{\sqrt{\delta}}\sigma^{j-g+2-i} \\
&\leq 2(g-1)^2 + 2(g-2)\frac{\sqrt{\mathrm{vol}(G)}\sigma^g}{\sqrt{\delta}(1-\sigma)} \\
&\leq 3g^2
\end{aligned}$$

noting that g has been chosen so that

$$\frac{\sigma^g}{1-\sigma} = \frac{\sqrt{\epsilon}\sqrt{\delta}}{4(1-\epsilon)\sqrt{\mathrm{vol}(G)}}$$

is sufficiently small to make the last inequality hold.

Thus we have established that $\mathbb{E}(X|\mathcal{F}_i)$. By applying Azuma's inequality [5], we have

$$\mathbb{P}(X - \mathbb{E}(X) < -\alpha) < e^{-\frac{\alpha^2}{18g^4 t}}$$

Note that

$$\begin{aligned}
\mathbb{E}(X) &= \sum_{i=1}^{t} \mathbb{E}(X_i) \\
&= \sum_{i=1}^{t}\sum_{j=1}^{n} \mathbb{E}(X_i \mid v_{i-1} = j)\mathbb{P}(V_{i-1} = j) \\
&\geq \sum_{i=1}^{t}\sum_{j=1}^{n} \left((1 - \frac{g-1}{d_j}) + \sum_{k=1}^{g-2} \frac{-k}{d_j} \right) \frac{d_j}{\mathrm{vol}(G)} \\
&= \sum_{i=1}^{t}\sum_{j=1}^{n} (1 - \frac{g(g-1)}{2d_j}) \frac{d_j}{\mathrm{vol}(G)} \\
&= \sum_{i=1}^{t} \left(1 - \frac{g(g-1)n}{2\mathrm{vol}(G)} \right) \\
&= (1 - \frac{g(g-1)}{2d})t.
\end{aligned}$$

By choosing $\alpha = \sqrt{18g^4 t \log \frac{4}{\epsilon}}$, we have

$$\mathbb{P}\left(X < (1 - \frac{g(g-1)}{2d})t - \sqrt{18g^4 t \log \frac{4}{\epsilon}} \right) < \frac{\epsilon}{4}.$$

Putting all together, we have

$$\begin{aligned}
\mathbb{P}\left(Y < (1 - \frac{g(g-1)}{2d})t - \sqrt{18g^4 t \log \frac{4}{\epsilon}} \right) &\leq \mathbb{P}(Y < X) + \mathbb{P}\left(X < (1 - \frac{2}{d})t - \sqrt{18g^4 t \log \frac{4}{\epsilon}} \right) \\
&< \frac{3\epsilon}{4} + \frac{\epsilon}{4} \\
&= \epsilon.
\end{aligned}$$

To complete the proof, it suffices to check that our degree conditions imply that

$$(1 - \frac{g(g-1)}{2d})t - \sqrt{18g^4 t \log \frac{4}{\epsilon}} = (1 - \epsilon - o(1))\sqrt{\epsilon \frac{nd}{\tilde{d}}}.$$

In particular it suffices to check that

$$\frac{g}{\sqrt{d}} = o(1),$$

as g^4 is clearly $o(t)$.

Since

$$g \leq \frac{\log\left(\frac{4}{\epsilon(1-\sigma)}\right) + \log\left(t\frac{\tilde{d}}{\delta}\right)}{\log(1/\sigma)} + 1$$

and

$$\log\left(t\frac{\tilde{d}}{\delta}\right) = \log((1-\epsilon)\sqrt{\epsilon}) + \frac{1}{2}\log\left(\frac{d}{\delta}n\right)$$

we have $g/\sqrt{d} = o(1)$, since

$$d \gg \frac{\log^2(n)}{\log^2(1/\sigma)}$$

as hypothesized.

$\square$

5 Proof of Upper Bound

For the upper bound, we follow the general strategy of Aldous in [1]. In particular we provide a (relatively straightforward) generalization of theorem 15 of Aldous' paper to give an upper bound in the general degree case.

Here, we let X_t denote the position of a random walk at time t. We denote by T_B the *hitting time* of a set B; that is

$$T_B = \min\{t : X_t \in B\}.$$

We denote the *return time* of a set B to be

$$T_B^+ = \min\{t \geq 1 : X_t \in B\}.$$

(Note that if the random walker does not start in B, $T_B = T_B^+$.)

When considering the probability that our random walk has some property under some number of steps we use the notation $\mathbb{P}_\rho$ to denote that we condition on our random walker having initial distribution ρ. Likewise, $\mathbb{E}_\rho$ denotes expectation conditioning on the initial distribution. If no distribution is given, it is assumed to be starting from the stationary distribution. As a convenient abuse of notation, for a vertex v, $\mathbb{P}_{\mathrm{v}}$ denotes starting with the distribution that places weight 1 on v.

The first tool is the following, rather standard, mixing lemma

Lemma 5. *For all initial distributions ρ and all $B \subseteq G$, there exists an (absolute) constant K*

$$\mathbb{P}_\rho\left(\mathrm{T_B} > 3\frac{\log \mathrm{n}}{\log(1/\sigma)}\frac{\mathrm{vol(G)}}{\mathrm{vol(B)}}\right) \leq \frac{1}{2}$$

Proof. We begin by bounding $|P^s(\rho, B) - \pi(B)|$; where ρ is an (arbitrary) initial distribution and $\pi(B) = \frac{\mathrm{vol}(B)}{\mathrm{vol}(G)}$. Write

$$\rho' D^{-1/2} = \sum_i \alpha_i \varphi_i'$$

where the φ_i' are left eigenvectors of $(I-\mathcal{L})$ corresponding to eigenvalues $(1-\lambda_i)$. Then

$$\alpha_0 = \langle \rho' D^{-1/2}, \frac{D^{1/2}\mathbf{1}}{\sqrt{\mathrm{vol} G}} \rangle = \frac{1}{\sqrt{\mathrm{vol}(G)}}.$$

Thus

$$\rho' D^{-1/2} = \frac{\mathbf{1}' D^{1/2}}{\mathrm{vol}(G)} + \sum_{i \geq 1} \alpha_i \varphi_i.$$

Then

$$\begin{aligned}
|P^s(\rho, B) - \pi(B)| &= \left| \rho' P^s \chi_B - \frac{\mathbf{1}' D}{\mathrm{vol}(G)} \chi_B \right| \\
&= \left| \left(\rho' D^{-1/2} (I - \mathcal{L})^s - \frac{\mathbf{1}' D^{1/2}}{\mathrm{vol}(G)} \right) D^{1/2} \chi_B \right| \\
&= \left| \left(\frac{\mathbf{1}' D^{1/2}}{\mathrm{vol}(G)} + \sum_{i \geq 1} (1 - \lambda_i)^s \alpha_i \varphi_i' - \frac{\mathbf{1}' D^{1/2}}{\mathrm{vol}(G)} \right) D^{1/2} \chi_B \right| \\
&\leq \sum_{i \geq 1} \sigma^s |\alpha_i| |\varphi_i' D^{1/2} \chi_B| \\
&\leq \sigma^s \sqrt{n} \mathrm{vol}_{1/2}(B) \\
&\leq \sigma^s \sqrt{n |B| \mathrm{vol}(B)|}
\end{aligned}$$

where the last step follows from an application of Cauchy-Schwarz inequality.

Let

$$s = \log \left(\frac{\sqrt{\mathrm{vol}(B)}}{2\sqrt{n|B|} \mathrm{vol}(G)} \right) / \log(\sigma)$$

so

$$\sigma^s \sqrt{n|B| \mathrm{vol}(B)} = \frac{\mathrm{vol}(B)}{2\mathrm{vol}(G)}.$$

Fix $t_i = is$, then

$$\begin{aligned}
\mathbb{P}(\mathrm{T_B > x}) &\leq \mathbb{P}(\mathrm{X_{t_1} \notin B, X_{t_2} \notin B, \ldots, X_{t_{x/s}} \notin B}) \\
&= \mathbb{P}(\mathrm{X_{t_1} \notin B}) \mathbb{P}(\mathrm{X_{t_2} \notin B | X_{t_1} \notin B}) \cdots \mathbb{P}(\mathrm{X_{t_{x/s}} \notin B | X_{t_j} \notin B, \forall j < i}) \\
&\leq \left(1 - \frac{\mathrm{vol}(B)}{\mathrm{vol}(G)} + \sqrt{n|B| \mathrm{vol}(B)} \sigma^s \right)^{x/s} \\
&\leq \left(1 - \frac{\mathrm{vol}(B)}{2\mathrm{vol}(G)} \right)^{x/s}.
\end{aligned}$$

Fix $x = 2\log(2) s \frac{\mathrm{vol}(G)}{\mathrm{vol}(B)}$ and it is easy to check that

$$\mathbb{P}(\mathrm{T_B > x}) \leq \frac{1}{2}.$$

In all, we have

$$x = 2\log(2)\frac{\log\left(\frac{2\sqrt{n|B|\mathrm{vol}(G)}}{\sqrt{\mathrm{vol}(B)}}\right)\mathrm{vol}(G)}{\log(1/\sigma)\mathrm{vol}(B)} \leq 3\frac{\log n}{\log(1/\sigma)}\frac{\mathrm{vol}(G)}{\mathrm{vol}(B)}$$

□

The following result (and it's proof) are due to Aldous [1]. Let $B = \{v_0, \dots, v_c\}$ denote a set of vertices and let $\mathcal{P}_B$ denote the event that the path from v_0 to the root (starting location of our random walker for generating a UST, chosen by the uniform distribution) in a uniform spanning tree starts $v_0, v_1, \dots, v_c$. Then

Lemma 6.

$$\mathbb{P}(\mathrm{T_{v_c}} = \ell|\mathcal{P}_\mathrm{B}) = \frac{\mathbb{P}_{\mathrm{v_c}}(\mathrm{T_B^+} > \ell)}{\mathbb{E}_{\mathrm{v_c}}(\mathrm{T_B^+})}$$

Proof. For $i < c$, we denote the event $\mathcal{D}_i$ to be

$$\mathcal{D}_i = \{T_{\{v_0,\dots,v_i\}} = T_{v_i},\ X_{T_{v_i}-1} = v_{i+1}\}.$$

In words, $\mathcal{D}_i$ is the event that v_i is hit before v_j for $j < i$, and indeed v_i is first hit from v_{i-1}, so $\bigcap_{i<c} \mathcal{D}_i = \mathcal{P}_B$. Then:

$$\{T_{v_c} = \ell\} \cap \mathcal{P}_B = \{T_{v_c} = \ell = T_B\}.$$

Note that, from the Markov property, it is clear that $P(\bigcap_{i<c} \mathcal{D}_i | T_{v_c} = T_B = \ell)$ does not depend on ℓ (this is the motivation for writing $\mathcal{P}_B$ in an obtuse way), thus:

$$\mathbb{P}(\mathrm{T_{v_c}} = \ell|\mathcal{P}_\mathrm{B}) = \alpha\mathbb{P}(\mathrm{T_{v_c}} = \mathrm{T_B} = \ell)$$

for $\ell = 0, 1, \dots$ and for some α which (critically) does not depend on ℓ. We have that:

$$\begin{aligned}
\mathbb{P}_{\mathrm{v_c}}(\mathrm{T_B^+} > \ell) &= \sum_w \mathbb{P}_{\mathrm{v_c}}(\mathrm{X_0 = v_c,\ X_\ell = w,\ X_i \notin B\ for\ } 1 \leq i \leq \ell) \\
&= \frac{1}{\pi(v_c)}\sum_w \mathbb{P}_\pi(\mathrm{X_0 = v_c,\ X_\ell = w,\ X_i \notin B\ for\ } 1 \leq i \leq \ell) \\
&= \frac{1}{\pi(v_c)}\sum_w \mathbb{P}_\pi(\mathrm{X_0 = w, X_\ell = v_c,\ X_i \notin B\ for\ } 1 \leq i \leq \ell) \\
&= \frac{1}{\pi(v_c)}\mathbb{P}_\pi(\mathrm{X_\ell = v_c,\ X_i \notin B\ for\ } 1 \leq i \leq \ell) \\
&= \frac{1}{\pi(v_c)}\mathbb{P}_\pi(\mathrm{T_{v_c} = T_B} = \ell)
\end{aligned}$$

with the third to last equality following from time reversal for the stationary Markov chain. This implies:

$$\mathbb{P}_\pi(\mathrm{T_{v_c}} = \ell|\mathcal{P}_\mathrm{B}) = \alpha\pi(\mathrm{v_c})\mathbb{P}_{\mathrm{v_c}}(\mathrm{T_B^+} > \ell).$$

Note finally, then that

$$1 = \sum_{\ell=0}^{\infty} \mathrm{\mathbb{P}_{\pi}(T_{v_c} = \ell | \mathcal{P}_B)} = \sum_{\ell=0}^{\infty} \mathrm{\alpha\pi(v_c)\mathbb{P}_{v_c}(T_B^+ > \ell)} = \mathrm{\alpha\pi(v_c)\mathbb{E}_{v_c}(T_B^+)}.$$

so $\alpha\pi(v_c) = \mathbb{E}_{v_c}(T_B^+)$, implying the result. □

One can observe that, actually, that while the normalizing constant is easy to compute, the exact value is unnecessary for the proof of the upper bound itself.

We now prove the upper bound, establishing (2) in Theorem 2; whose proof mimics that of Aldous.

Proof of (2): Let us start our random walk from the stationary distribution (unless explicitly noted, all probabilities related with the random walk which generates the spanning tree are taken to start with π).

We begin by fixing a path $v_0, v_1, \ldots, v_c$ in our graph; and B be the set $\{v_0, \ldots, v_c\}$. As above, $\mathcal{P}_B$ will denote the event that the path from v_0 to the root (that is, the starting location of our random walk, X_0) in our uniform spanning tree starts out along the path $v_0, \ldots, v_c$. If we let

$$s = \left\lceil \frac{3}{\log(1/\sigma)} \frac{\mathrm{vol}(G)}{(c+1)\delta} \right\rceil \geq \frac{3}{\log(1/\sigma)} \frac{\mathrm{vol}(G)}{\mathrm{vol}(B)} \log n$$

then, by iterating lemma 5 we have that

$$\mathrm{\mathbb{P}_{v_c}(T_B^+ > js) \leq \frac{1}{2^{j-1}} \mathbb{P}_{v_c}(T_B^+ > s)}.$$

We are now in the position to apply lemma 6 to both sides; note that the normalizing constant will cancel and we are left with:

$$\mathrm{\mathbb{P}(T_{v_c} = js | \mathcal{P}_B) \leq (1/2)^{j-1} \mathbb{P}(T_{v_c} = s | \mathcal{P}_B) \leq (1/2)^{j-1} s^{-1}}.$$

where the last inequality follows from the fact that the right hand side of (6) in lemma 6 is decreasing with l and hence the left hand side must as well. This monotonicity property, and summing gives:

$$\mathrm{\mathbb{P}(js \leq T_{v_c} \leq (j+1)s | \mathcal{P}_B) \leq (1/2)^{j-1}}.$$

Further summing gives

$$\mathrm{\mathbb{P}(js \leq T_{v_c} | \mathcal{P}_B) \leq (1/2)^{j-2}}.$$

If $\mathcal{P}_B$ occurs; then naturally we have that the distance from v_0 to the root, X_0, satisfies

$$d(X_0, v_0) \leq d(X_0, v_c) + c \leq T_{v_c} + c.$$

We also clearly have that if $d(X_0, v_0) > c$, then $\mathcal{P}_B$ occurs for some unique path $v_0, \ldots, v_c$. Thus:

$$\mathrm{\mathbb{P}(d(X_0, v_0) > c + js) \leq (1/2)^{j-2}}.$$

Clearly $diam(T)/2 \leq \max_v d(X_0, v)$; so

$$\mathbb{P}(\mathrm{diam(T)/2 > c + js}) \leq \mathrm{n(1/2)^{j-2}}.$$

This gives us

$$\mathbb{E}(\mathrm{diam(T)}) \leq \mathrm{2c + 3s\log n} \leq \mathrm{2c} + \frac{\mathrm{3vol(G)}}{\mathrm{c}\log(1/\sigma)\delta}\log^2(\mathrm{n}),$$

with the second inequality coming from the definition of S. These terms are the same order of magnitude when setting $c = \sqrt{\frac{\mathrm{vol}G}{\delta\log(1/\sigma)}}\log n$; giving the desired bound. To establish the bound in the form stated in (2), simply apply Markov's inequality. $\square$

Note that by minimizing

$$2c + \frac{3\mathrm{vol}(G)}{c\log(1/\sigma)\delta}\log^2(n)$$

we actually get that

$$\mathbb{E}(\mathrm{diam(T)}) \leq 2\sqrt{6\frac{\mathrm{vol(G)}}{\delta\log(1/\sigma)}}\log \mathrm{n}.$$

References

[1] D. Aldous, The Random Walk construction of Uniform Spanning Trees, *SIAM J. Discrete Math.* **3** (1990), 450-465.

[2] A. Broder, Generating random spanning trees, *Symp. foundations of Computer Sci.,* Institute for Electrical and Electronic Engineers, New York 442-447.

[3] F. Chung, *Spectral Graph Theory*, AMS Publications, 1997

[4] F. Chung and L. Lu, Connected components in a random graph with given degree sequences, *Annals of Combinatorics* **6**, (2002), 125–145.

[5] F. Chung and L. Lu, *Complex Networks and Graphs*, *CBMS* Lecture Series, No. 107, AMS Publications, 2006.

[6] F. Chung, L. Lu and V. Vu, The spectra of random graphs with given expected degrees, *Proceedings of National Academy of Sciences*, **100**, no. 11, (2003), 63 13-6318.

[7] F. Kirchhoff, Über die Auflösung der Gleichungen, auf welche man bei der Undersuchung der linearen Verteilung galvanisher Str"ome gef"uhrt wird, *Ann. Phys. chem.* **72** (1847) 497-508.

[8] David Liben-Nowell and Jon Kleinberg, Tracing Information Flow on a Global Scale Using Internet.

[9] R. Lyons, Asymptotic enumeration of spanning trees. *Combin. Probab. Comput.* **14** (2005), no. 4, 491–522.

[10] R. Pemantle, Uniform random spanning trees, *Topics in contemporary probability and its applications*, J. L. Snell, editor, 1 - 54. CRC Press: Boca Raton (1994).

[11] A. Rényi and G. Szekeres, On the height of trees, *J. Austral. Math. Soc.*, **7** (1967), pp. 497–507.