May 2009

Extremal Random Walks on Trees

Meng Wang
Macalester College, mwang@macalester.edu

Recommended Citation

Wang, Meng, "Extremal Random Walks on Trees" (2009). *Mathematics, Statistics, and Computer Science Honors Projects*. 12.
https://digitalcommons.macalester.edu/mathcs_honors/12

Extremal Random Walks on Trees

Meng Wang

Andrew Beveridge, Advisor

Tom Halverson, Reader

Robert Dobrow (Carleton College), External Reader

May, 2009

Macalester College

Department of Mathematics

Abstract

We study random walks on trees, where we iteratively move from one vertex to a randomly chosen adjacent vertex. We study two quantities arising in random walks: the hitting time and the mixing time. The hitting time is the expected number of steps to walk between a chosen pair of vertices. The mixing time is the expected number of steps before the distribution of the current state is proportional to its degree. For a fixed tree size, we prove that the star is the unique minimizing structure and the path is the unique maximizing structure for both quantities.

Contents

Chapter 1

Introduction

Of all trees of a fixed size, which ones maximize or minimize the extremal hitting times and mixing times? We define a random walk on a tree to be the sequence of nodes visited if one is to start at one node of the tree, and move to a random neighbor of the current node after each period.

- Stationary distribution: Given enough time, there exists a stationary distribution for any random walk under which the probability that one is at any node no longer changes
- Hitting time: The hitting time between two nodes on the tree is the expected length of a random walk from one to another
- Commute time: The commute time is the sum of the hitting time from one to the other and the hitting time from going the opposite direction
- Mixing time: The mixing time is the expected length of an optimal stopping rule starting from the worst possible node
- Best mixing time: The best mixing time is the expected length of an optimal stopping rule starting from the best possible node
- Forget time: The forget time is the minimum length such that there exists a distribution where the expected length of an optimal rule to attain that distribution from the worst possible node is within that length
- Reset time: The reset time is an average mixing time weighted by each node's probability under the stationary distribution, which is equal to the forget time

In general, we found that the path is the maximizing structure for the extremal hitting times, commute times, $H(\pi, i)$, the mixing time, and the best mixing time, while the star is the minimizing structure for the above properties. One exception is that the maximizing structure for the best mixing time is the path when the number of vertices of the tree is even, while it is a close variant of the path when the number of vertices is odd.

Chapter 2

Preliminaries

2.1 Random Walks on Graphs

2.1.1 Random walks

Let G be an undirected graph, $v(G) = n$, $e(G) = m$. Given an initial node X_0, we define a *random walk* recursively by setting X_i to be a random neighbor of X_{i-1} for $i \geq 1$. More rigorously a Markov chain is a random process $(X_0, X_1, \ldots)$ with finite state space $S = \{s_1, \ldots, s_n\}$ and a $n \times n$ transition matrix P such that for all k, all $i, j \in \{1, \ldots, n\}$ and all $i_0, \ldots, i_{k-1} \in \{1, \ldots, n\}$

$$\begin{aligned} & P(X_{k+1} = s_j | X_0 = s_{i_0}, X_1 = s_{i_1}, \ldots, X_{k-1} = s_{i_{k-1}}, X_k = s_i) \\ = \ & P(X_{k+1} = s_j | X_k = s_i) \\ = \ & P_{ij}. \end{aligned}$$

For an undirected graph, the state space is $V(G)$. Elements in the transition matrix P are transition probabilities. P_{ij} is the probability of moving from s_j to s_i in the next period. The transition matrix satisfies that for all $i, j \in n$

$$P_{ij} = \begin{cases} \frac{1}{\deg(i)} & \text{if } i \text{ is adjacent to } j \\ 0 & \text{otherwise.} \end{cases}$$

It follows that

$$P_{ij} \geq 0$$

and

$$\sum_{j=1}^{n} P_{ij} = 1.$$

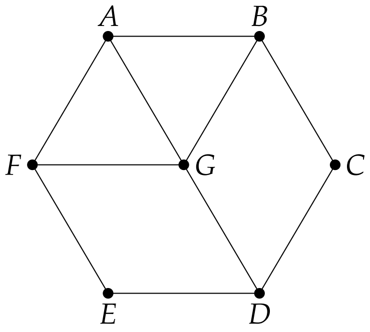

Figure 2.1: A graph on 7 vertices

Example 1. Let's consider a random walk on the graph in Figure 2.1. There are 7 stores in a small town, and adjacent stores are connected with an edge on the graph. If a shopper is at one of the stores at time t, and she moves to one of the neighboring stores with equal probability at time $t+1$, then the sequence of stores visited is a random walk. Assume that at time 0, the shopper is at store A. She decides to move to one of the neighboring stores B, G, or F with equal probability of $\frac{1}{3}$ in the next period. Let X_n denote the store that the shopper is in at time n. The above statement can be expressed as conditional probabilities

$$P(X_1 = B|X_0 = A) = P(X_1 = G|X_0 = A) = P(X_1 = F|X_0 = A) = \frac{1}{3}.$$

Suppose the shopper is at store C at time $t = 3$, then she would move to store B or D with a probability of $\frac{1}{2}$ each at time $t = 4$

$$P(X_4 = B|X_3 = C) = P(X_4 = D|X_3 = C) = \frac{1}{2}$$

and

$$\begin{aligned} & P(X_4 = B|X_3 = C, X_2 = i_2, X_1 = i_1, X_0 = i_0) = P(X_4 = B|X_3 = C) \\ = \; & P(X_4 = D|X_3 = C, X_2 = i_2, X_1 = i_1, X_0 = i_0) = P(X_4 = D|X_3 = C) \\ = \; & \frac{1}{2}. \end{aligned}$$

$\diamondsuit$

As shown above, the probability of the shopper being at a specific store at time n is uniquely determined by her location at time $n-1$, and is independent from all information from time 0 to $n-2$. This is referred to as the *memoryless property*, or Markov property.

The transition matrix for the above random walk is

$$P = \begin{bmatrix} 0 & \frac{1}{3} & 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & 0 & \frac{1}{3} & 0 & 0 & 0 & \frac{1}{3} \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{3} & 0 & \frac{1}{3} & 0 & \frac{1}{3} \\ 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ \frac{1}{3} & 0 & 0 & 0 & \frac{1}{3} & 0 & \frac{1}{3} \\ \frac{1}{4} & \frac{1}{4} & 0 & \frac{1}{4} & 0 & \frac{1}{4} & 0 \end{bmatrix}$$

Two important conditions on Markov chains are *irreducibility* and *aperiodicity*. A Markov chain $(X_0, X_2, \ldots)$ with state space $S = \{s_1, \ldots, s_n\}$ and transition matrix P is irreducible if for all $s_i, s_j \in S$, there exists an n such that $(P^n)_{ij} > 0$. In other words, there is a way to walk between any two given nodes in a finite number of steps. This implies that the underlying graph is connected.

The period $d(s_i)$ of a state $s_i \in S$ is defined as $d(s_i) = gcd\{n \geq 1 : (P^n)_{ii} \geq 0\}$, which is the greatest common divisor of the set of the number of steps s_i can return to the start state. The Markov chain is aperiodic if each state has period 1. Note that all chains on bipartite graphs including trees are periodic. However, they can be turned into aperiodic ones by performing a "lazy walk", where during each period we stay at the current node with probability $\frac{1}{2}$. A lazy walk would double the expected length of the random walk.

Example 2. Our random walk in Figure 1.1 is both irreducible and aperiodic. The Markov chain in Figure 2.2 is aperiodic but reducible, since it is impossible to walk between nodes G and E. The Markov chain in Figure 2.3 is irreducible but periodic, since it takes an even number of steps for any node to take a walk and then return to the starting node. ◇

2.1.2 Stationary distribution

Next we study the long-term behavior of Markov chains. Let $(X_0, X_1, \ldots)$ be a Markov chain with state space $\{s_1, \ldots, s_k\}$ and transition matrix P. As n approaches infinity, we are interested in knowing whether the distribution of X_n approaches a stable state. We define a row vector $\pi = (\pi_1, \ldots, \pi_k)$ to be a *stationary distribution* for the Markov chain if

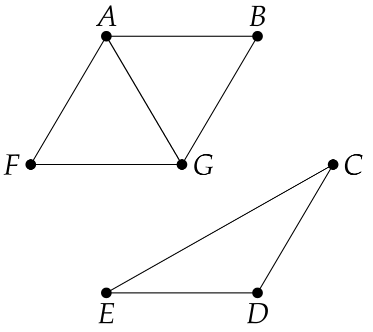

Figure 2.2: A reducible and aperiodic chain

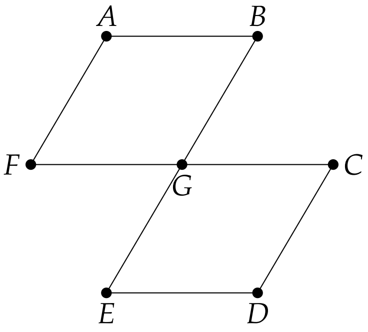

Figure 2.3: An irreducible and periodic chain

(1) $\pi_i \geq 0 \, for \, i = 1, \ldots, k$

(2) $\sum_{i=1}^{k} \pi_i = 1$

(3) $\sum_{i=1}^{k} \pi_i P_{ij} = \pi_j \, for \, j = 1, \ldots, k.$

Conditions (1) and (2) guarantees π to be a probability distribution on S. Condition (3) states that once a distribution equals π, all subsequent distributions would also equal π. We define the *total variation distance* between two probability distributions $v^{(1)} = (v_1^{(1)}, v_2^{(1)}, \ldots, v_k^{(1)})$ and $v^{(2)} = (v_1^{(2)}, v_2^{(2)}, \ldots, v_k^{(2)})$ on state space $S = \{s_1, s_2, \ldots, s_k\}$ to be

$$d_{TV}(v^{(1)}, v^{(2)}) = \frac{1}{2} \sum_{i=1}^{k} |v_i^{(1)} - v_i^{(2)}|.$$

If $v^{(1)}, v^{(1)}, \ldots$, and v are probability distributions on S, then $v^{(n)}$ *converges to* v *in total variation* as $n \longrightarrow \infty$, denoted as $v^{(n)} \xrightarrow{TV} v$, if

$$\lim_{n \to \infty} d_{TV}(v^{(n)}, v) = 0.$$

[(5)] shows the existence and uniqueness of the stationary distribution. Furthermore, it guarantees convergence to stationarity starting from any random distribution.

Theorem 1 *For any given irreducible and aperiodic Markov chain, there exists at least one stationary distribution.*

Theorem 2 *Let* $(X_0, X_1, \ldots)$ *be an irreducible and aperiodic Markov chain with state space* $S = \{s_1, s_2, \ldots, s_k\}$*, transition matrix* P*, and an arbitrary distribution* $\pi^{(0)}$*. Then for any distribution* π *which is stationary for the transition matrix* P*, we have*

$$\pi^{(n)} \xrightarrow{TV} \pi.$$

Theorem 3 *Any irreducible and aperiodic Markov chain has exactly one stationary distribution.*

Thus for any irreducible and aperiodic Markov chain, there is one and only one stationary distribution, and further that if a Markov chain continues long enough, it would converge in total variation to π, implying that the distribution would approach the stationary distribution, called *equilibrium*, regardless of the initial distribution.

A Markov chain is *reversible* if there exists a probability distribution π such that for all $i, j \in S$, $\pi_i P_{i,j} = \pi_j P_{j,i}$. If such a probability distribution exists for the Markov chain, then it must be the stationary distribution for the chain. Under a reversible distribution, during the transition from one time period to the next, the inflow and outflow between two neighboring nodes must be equal to each other, which requires the probability on each node to be proportional to its degree. It follows that $\pi_i = \deg(i)/2|E|$. When the graph is regular, the stationary distribution is uniform.

Example 3. In Figure 2.1 the stationary distribution is

$$\pi = (\frac{3}{20}, \frac{3}{20}, \frac{1}{10}, \frac{3}{20}, \frac{1}{10}, \frac{3}{20}, \frac{1}{5}).$$

During period t to $t+1$, the probability of moving from state B to state C is

$$\pi_B P_{BC} = \frac{3}{20} \times \frac{1}{3} = \frac{1}{20}$$

while the probability of C moving to B is

$$\pi_C P_{CB} = \frac{1}{10} \times \frac{1}{2} = \frac{1}{20} = \pi_B P_{BC}.$$

Therefore under the stationary distribution, the probability of moving from B to C is the same as the probability of moving from C to B. ◇

2.2 Random Walks on Trees

2.2.1 Hitting times for trees

We are interested in random walks on *trees*, which are connected graphs without cycles. There is exactly one path between any two given nodes on a tree.

The *hitting time*, $H(i,j)$, between two nodes i and j on a tree, is the expected length of a walk from i to j. The concept of hitting time provides a sense of average distance between nodes. The *commute time*, $\kappa(i,j)$, between two nodes i and j is the sum of the hitting time between i and j, and the hitting time between j and i:

$$\kappa(i,j) = H(i,j) + H(j,i).$$

The *return time*, $\mathrm{Ret}(i)$, is the expected time for a random walk starting at i to first return to i:

$$\mathrm{Ret}(i) = \frac{1}{\pi_i}.$$

In this paper, we will often be comparing hitting times for distinct trees. When necessary we will use a subscript to identify the tree, for example: $H_T(i,j), \kappa_T(i,j), \mathrm{Ret}_T(i,j)$. Next we define the worst possible start node to hit a given node j. Let j' be a *j-pessimal* node which satisfies $H(j',j) = \max_{i \in V} H(i,j)$.

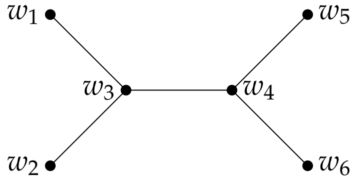

Figure 2.4: A tree on 6 vertices

Example 4. In Figure 2.4, clearly $H(w_1, w_3) = 1$. $H(w_3, w_4)$ is slightly more difficult to calculate, since it is possible that w_3 would move to either w_1 or w_2 first, then come back to w_3, and the loop could repeat itself infinitely. Once the first step is taken, there are three possible nodes, and we have to calculate their separate hitting times to w_4

$$\begin{aligned} H(w_1, w_4) &= H(w_2, w_4) = 1 + H(w_3, w_4) \\ H(w_3, w_4) &= 1 + \left(\frac{1}{3}H(w_4, w_4) + \frac{1}{3}H(w_1, w_4) + \frac{1}{3}H(w_2, w_4)\right) \\ H(w_5, w_4) &= H(w_6, w_4) = 1. \end{aligned}$$

Solve the above system of equations to get $H(w_3, w_4) = 5$ and

$$H(w_1, w_4) = H(w_2, w_4) = 6.$$

Thus w_1 and w_2 are both w_4-pessimal nodes. By symmetry

$$H(w_3, w_4) = H(w_4, w_3).$$

Then we can calculate the commute time between w_3 and w_4:

$$\kappa(w_3, w_4) = H(w_3, w_4) + H(w_4, w_3) = 5 + 5 = 10.$$

We can also calculate the return time for w_4:

$$\begin{aligned} \mathrm{Ret}(w_3) &= 1 + \left(\frac{1}{3}H(w_5, w_4) + \frac{1}{3}H(w_5, w_4) + \frac{1}{3}H(w_3, w_4)\right) \\ &= 1 + \frac{1}{3}(1 + 1 + 5) \\ &= \frac{10}{3}. \end{aligned}$$

♢

2.2.2 Known results for trees

We denote $V_{u:v}$ as the set of nodes in the subtree rooted at u after the removal of edge uv, as shown in Figure 2.5.

From [(2)], for any two adjacent nodes i and j on tree $G = (V, E)$,

$$H(i, j) = \sum_{k \in V_{i:j}} d(k) = 2|E| \sum_{k \in V_{i:j}} \pi_k = 2|E|\pi(V_{I:j}).$$

Note that the commute time between i and j is constant when the two nodes are neighbors

$$H(i, j) + H(j, i) = 2|E|. \tag{2.1}$$

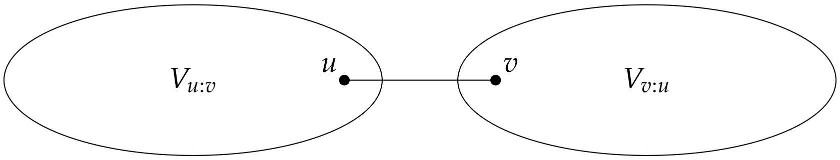

Figure 2.5: The vertex partition into $V_{u:v}$ and $V_{v:u}$.

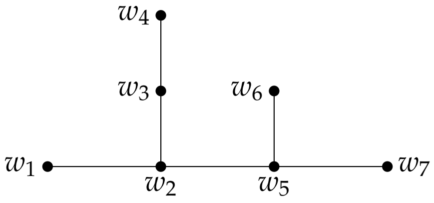

Figure 2.6: A tree on 7 vertices

Define

$$\ell(i,k;j) = \frac{1}{2}\left(d(i,j) + d(k,j) - d(i,k)\right)$$

where i, j, and k are nodes on G. $\ell(i,k;j)$ measures the length of shared portion between the (i,j)-path and the (k,j)-path.

Example 5. On the tree in Figure 2.6,

$$\begin{aligned}\ell(w_1,w_4;w_7) &= \frac{1}{2}\left(d(w_1,w_7) + d(w_4,w_7) - d(w_1,w_4)\right)\\ &= \frac{1}{2} \times (3+4-3) = 2.\end{aligned}$$

Similarly

$$\ell(w_1,w_4;w_7) = \ell(w_1,w_3;w_7) = \ell(w_1,w_2;w_7) = d(w_2,w_7) = 2$$

$$\ell(w_1,w_6;w_7) = \ell(w_1,w_5;w_7) = d(w_5,w_7) = 1.$$

◇

In [(2)], it was shown that the hitting time between any two nodes i and j is

$$H(i,j) = \sum_{k\in V} \ell(i,k;j)d(k). \tag{2.2}$$

It follows that on a path of length n, P_n, where the vertices are labeled $v_1, \ldots, v_n$, the hitting time between two nodes is

$$H(v_i, v_j) = \begin{cases} (j-1)^2 - (i-1)^2 & i \leq j \\ (n-j+1)^2 - (n-i+1)^2 & i < j. \end{cases}$$

2.2.3 Centers for random walks on trees

We summarize some results from [(2)]. There are two important centers for random walks on trees. One is defined as the "extremal" center.

- On a given tree $G = (V, E)$, if $a \in V$ satisfies $H(a', a) = \min_{j \in V} \max_{i \in V} H(i, j)$, then a is a *primary focus* of G.
- If all a-pessimal nodes are in the same component of $G - a$, then the unique a neighbor $b \in G'$ is also a focus.
- If $H(b', b) = H(a', a)$ then b is also a primary focus, otherwise it is a *secondary focus*.
- A tree must either has a single focus or has two adjacent foci. Trees with one focus is referred to as focal, while those with two foci are bifocal.

The second center is defined as an "average" center under the stationary distribution. The *barycenter* is the node or two adjacent nodes that achieves $\min_{i \in V} \sum_{j \in V} \mathrm{d}(i, j)$. In other words, the barycenter minimizes the total distance to all other nodes. The following is proven in [(2)]

Theorem 4 *The following statements for a node c are equivalent*

(I) *The node c is a barycenter of the tree.*

(II) *The node c satisfies $H(i, c) \leq H(c, i)$ for all node i.*

(III) $\sum_{k \in V} \pi_k H(k, c) = \min_{i \in V} \sum_{k \in V} \pi_k H(k, i)$.

(IV) *For every node i adjacent to c, $\pi(V_{i:c}) = \sum_{k \in V_{i:c}} \pi_k \leq \frac{1}{2}$.*

Example 6. In Figure 2.7, we have

$$\min_{j \in V} \max_{i \in V} H(i, j) = H(w_1, w_3) = 4$$

$$H(w_1, w_3) = 4 < 6 = H(w_4, w_2).$$

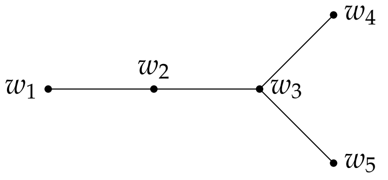

Figure 2.7: A tree on 5 vertices

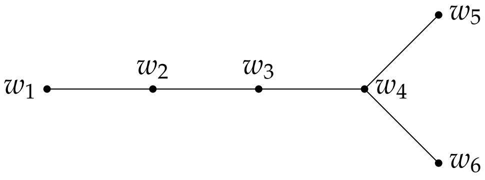

Figure 2.8: A tree on 6 vertices

Therefore node w_3 is a primary focus, and w_2 is a secondary focus. Since node w_3 satisfies condition (III) of the barycenter definition, it is the unique barycenter.

In Figure 2.8, we have

$$\min_{j\in V}\max_{i\in V} H(i,j) = H(w_5, w_3) = H(w_6, w_3) = 6$$

$$H(w_4, w_2) = 6 > 4 = H(w_1, w_3).$$

Therefore node w_3 is a primary focus, and w_4 is a secondary focus. Since node w_3 and w_4 both satisfy condition (III) of the barycenter definition, they are both barycenters of the tree. Note that on an asymmetric tree, foci are not necessarily the same as barycenters. $\diamondsuit$

2.3 Stopping Rules

We briefly summarize some results of Lovász and Winkler [(7)]. Let V^* be the space of finite walks on V, i.e. the set of finite strings $w = (w_0, w_1, w_2, \dots, w_t)$, $w_i \in V$ and w_i adjacent to w_{i-1}. For a given initial distribution σ, the probability of w being the walk after t steps is

$$\Pr(w) = \sigma_{w_0}\prod_{i=0}^{t-1} p_{w_i w_{i+1}}.$$

A *stopping rule* Γ is a map from V^* to $[0,1]$ such that $\Gamma(w)$ is the probability of continuing given that w is the walk so far observed. We assume that with probability 1 the rule stops the walk in a finite number of steps.

Given another distribution τ on V, the *access time* $H(\sigma,\tau)$ is the minimum expected length of a stopping rule Γ that produces τ when started at σ. We say Γ is *optimal* if it achieves this minimum. For example, in the case that $\sigma = \tau$ are both singleton distributions on the node i, the rule "take no steps" is an optimal stopping rule with expected length 0, while the rule "walk until you return to i" is a non-optimal stopping rule with expected length $\mathrm{Ret}(i)$.

Optimal stopping rules exist for any pair σ, τ of distributions and the access time $H(\sigma, \tau)$ has many useful algebraic properties. When σ and τ are concentrated on nodes i and j respectively (we write $\sigma = i, \tau = j$), the access time $H(i,j)$ is the hitting time from i to j. Clearly, $H(\sigma, j) = \sum_{i \in V} \sigma_i H(i,j)$ and $H(\sigma, \tau) \leq \sum_{i \in V} \sigma_i H(i, \tau)$. The latter inequality is usually strict for non-singleton distributions. For example, $0 = H(\pi, \pi) < \sum_{k \in V} \pi_k H(k, \pi) = T_{\text{reset}}$.

Given a stopping rule Γ from σ to τ, the *exit frequency* $x_i(\Gamma)$ is the expected number of times the walk leaves node i before halting. Exit frequencies partition the expected length of the walk: $E(\Gamma) = \sum_{k \in V} x_k(\Gamma)$. Exit frequencies are fundamental to virtually all access time results. A key observation, due to Pitman [(10)], is the "conservation equation"

$$\sum_{i \in V} p_{ij} x_i(\Gamma) - x_j(\Gamma) = \tau_j - \sigma_j. \tag{2.3}$$

It follows that the exit frequencies for two rules from σ to τ differ by $K\pi_i$ where K is the difference between the expected lengths of these rules. Hence the distributions σ and τ uniquely determine the exit frequencies for an optimal stopping rule between them and we denote these optimal exit frequencies by $x_i(\sigma, \tau)$. Moreover,

$$\Gamma \text{ is an optimal stopping rule } \iff \exists k \in V,\ x_k(\Gamma) = 0. \tag{2.4}$$

Otherwise a rule with exit frequencies $x_k(\Gamma) - \pi_k \min_{i \in V}(x_i(\Gamma)/\pi_i)$ will have strictly smaller expected length while also satisfying equation (2.3). (See [(7)] for multiple ways to construct stopping rules from a given set of desired exit frequencies.) When $x_k(\Gamma) = 0$, we call the node k a (σ, τ)-*halting state*, or simply a *halting state* when the initial and target distributions are clear. The presence of a halting state is the single most useful criterion for determining whether a given rule is optimal. Note that an optimal rule may have multiple halting states, but we need only identify one such state to ensure that a rule is optimal.

Any three distributions σ, τ and ρ satisfy the "triangle inequality"

$$H(\sigma, \rho) \leq H(\sigma, \tau) + H(\tau, \rho). \tag{2.5}$$

The right hand side of this equation is the expected length of the composite rule that first follows an optimal stopping rule from σ to τ and then follows an optimal stopping rule from τ to ρ. The exit frequency for node k of this composite rule is $x_k(\sigma,\tau) + x_k(\tau,\rho)$. We have equality in equation (2.5) if and only if this composite rule is optimal. In particular, there must be some node k such that $x_k(\sigma,\tau) = 0$ and $x_k(\tau,\rho) = 0$. Considering the case where ρ is a singleton distribution, $H(\sigma, j) \leq H(\sigma,\tau) + H(\tau, j)$ for any node j and equality holds if and only if j is a halting state for an optimal rule from σ to τ. Hence

$$H(\sigma,\tau) = \max_{j \in V}(H(\sigma, j) - H(\tau, j)). \tag{2.6}$$

In the special case $\sigma = i$ and $\tau = \pi$ we have a particularly nice characterization due to the combination of equations (**??**) and (2.6):

$$j \text{ is an } (i,\pi)\text{-halting state} \iff H(j,i) = \max_{k \in V} H(k,i). \tag{2.7}$$

Let $j = i'$ denote such an *i-pessimal* node. We can reformulate this observation as

$$H(i,\pi) = H(i',i) - H(\pi,i). \tag{2.8}$$

Example: mixing walks on P_3

We describe some optimal stopping rules from singleton distributions on $P_3 = (v_0, v_1, v_2, v_3)$ to $\pi = (1/6, 1/3, 1/3, 1/6)$. First we construct an optimal mixing rule $\Gamma(v_0,\pi)$. By equation (2.4), a rule is optimal when it has a halting state. Equation (2.7) identifies v_3 as the unique halting state. Let $\Gamma(v_0,\pi)$ be the rule "choose a target node according to π and walk to that node." Since v_3 is never exited by this rule, $\Gamma(v_0,\pi)$ is optimal with expected length $H(v_0,\pi) = |\Gamma(v_0,\pi)| = \frac{1}{6}H(v_0,v_0) + \frac{1}{3}H(v_0,v_1) + \frac{1}{3}H(v_0,v_2) + \frac{1}{6}H(v_0,v_3) = 19/6$.

We now consider starting at the node v_1. Equation (2.4) again identifies v_3 as the unique halting state. For this starting node, choosing our target ahead of time does not result in an optimal rule: there is a nonzero chance of reaching v_3 before reaching v_0 (so v_3 would not be a halting state). Instead our heuristic is to try to stop as quickly as possible. The rule $\Gamma(v_1,\pi)$ is: "at $t = 0$, take a step with probability 2/3 (and otherwise halt the walk for good). If the walk is still active at $t = 1$ then we are at either v_0 or v_2. If we are at v_2 then halt the walk. If we are at v_0 then stop with probability 1/2 and otherwise keep walking until you reach v_3." Let us describe the behavior of this rule. At time $t = 0$, our distribution is $(0, 1, 0, 0)$. At time $t = 1$, our distribution is $(1/3, 1/3, 1/3, 0)$. Note that at time $t = 1$ our walk continues to be active only when we are at v_0. In this case we halt (with probability $1/2$) or continue walking (with probability $1/2$) until we reach v_3. When the rule finally terminates, our distribution is $(1/6, 1/3, 1/3, 1/6)$ and v_3 is a halting state. The expected length of this optimal rule is $H(v_1, \pi) = |\Gamma(v_1, \pi)| = \frac{2}{3} + \frac{1}{6}H(v_0, v_3) = 13/6$.

Finally, we consider another optimal (v_0, π)-rule. Let $\Gamma'(v_0, \pi)$ be the rule "take one step and then follow $\Gamma(v_1, \pi)$." Clearly $|\Gamma'(v_0, \pi)| = 1 + |\Gamma(v_1, \pi)| = 19/6 = H(v_0, \pi)$ and indeed v_3 is a halting state for this composite rule. Interestingly, both the rules $\Gamma(v_0, \pi)$ and $\Gamma(v_1, \pi)$ are optimal but they are clearly distinct: $\Gamma'(v_0, \pi)$ always exits v_0 at $t = 0$ while $\Gamma(v_0, \pi)$ halts at $t = 0$ with probabilty $1/6$.

2.4 Optimal Stopping Rules

In this section we introduce some optimal stopping rules and their characteristics.

2.4.1 The filling rule

We define the filling rule recursively: let p_i^k be the probability of being at node i after k steps, and q_i^k be the probability that the walk was stopped at i before k steps have been taken. If we are at node i after k steps, we stop at i with probability $\min(1, (\tau_i - q_i^k)/p_i^k)$. This is a finite stopping rule that achieves τ from σ and there exists a halting state that is never exited by the filling rule.

The filling rule aims to fill each node without overshooting its target probability. We stops at a node with 100% probability if we arrive before a given time, after which the probability of stopping decreases until eventually we would never stop at the given node. Once we exited from a node, we would never stop there in the future.

2.4.2 The threshold rule

In the threshold rule, there is a threshold vector $h = (h_1, \ldots, h_n)$, $h_i \in [0, \infty]$ such that

$$\Gamma_{w_0,\ldots,w_k} = \begin{cases} 0 & \text{if } k \geq h_{w_k} \\ 1 & \text{if } k \leq h_{w_k} - 1 \\ k - h_{w_k} & \text{otherwise.} \end{cases}$$

In effect each node has a critical time, after which times we stop with 100% probability at that node. If we are within time 1 of the critical time, we stop with a certain probability, otherwise if we arrive at the node before the critical time less 1, we would keep going.

The threshold vector may not be uniquely determined by a threshold rule Γ, so we will always use the one whose coordinates are minimal. Γ is optimal if some coordinate of $h(\Gamma)$ is 0. From [(7)]

Theorem 5 *For every target distribution, there exists an optimal filling rule and an optimal threshold rule.*

2.5 Mixing Measures

We define the *mixing time* T_{mix} to be the expected length of an optimal mixing rule starting from the worst possible node: $T_{\text{mix}} = \max_{i \in V} H(i, \pi)$. A node that achieves this maximum is called *mixing pessimal*. The *best mixing time* T_{bestmix} is the expected length of an optimal mixing rule starting from the best possible node: $T_{\text{bestmix}} = \min_{i \in V} H(i, \pi)$. The *forget time* T_{forget} is the minimum length such that there exists a distribution *mu* such that for any start node, the expected length of an optimal rule to attain *mu* is within that length: $T_{\text{forget}} = \min_{\tau} \max_{i \in V} H(i, \tau)$. Theorem 10 in [(9)] shows that the forget time is attained by a unique distribution given by

$$\mu_i = \pi_i \left(1 + \sum_{j \in V} p_{ij} H(j, \pi) - H(i, \pi) \right). \tag{2.9}$$

The *reset time* $T_{\text{reset}} = \sum_{i \in V} \pi_i H(i, \pi)$ is an average mixing time weighted by the node's probability in the stationary distribution. Theorem 1 in [(9)] establishes the remarkable equality

$$T_{\text{forget}} = T_{\text{reset}} \tag{2.10}$$

for a random walk on an undirected graph.

From [(2)], $T_{\text{bestmix}} = \min_{i \in V} H(i, \pi)$ is achieved by a focus of the tree G. Specifically, if G is bifocal, and suppose the two foci are node a and b. If $H(a', b) < H(b', a)$ then node a uniquely achieves T_{bestmix}, if $H(a', b) > H(b', a)$ then node b uniquely achieves T_{bestmix}, if $H(a', b) = H(b', a)$ then both nodes achieve T_{bestmix}.

Again from [(2)], we have the following two theorems

Theorem 6 *If a distribution τ has a unique focus u then for all i*

$$H(i, \tau) = H(i, u) + H(u, \tau).$$

If τ has two foci u and v, then for $i \in V_{u:v}$,

$$H(i,\tau) = H(i,u) + H(u,\tau)$$

and for $i \in V_{v:u}$,

$$H(i,\tau) = H(i,v) + H(v,\tau).$$

Theorem 7 *For a focal tree, $T_{\text{forget}} = H(a',a)$. For a bifocal tree,*

$$\begin{aligned} T_{\text{forget}} &= H(a',\mu) = H(b',\mu) \\ &= \frac{1}{2|E|}\left(H(a,b)H(b,a) + H(a,b)H(a',b) + H(b,a)H(b',a)\right). \end{aligned}$$

Example 7. On the tree in Figure 2.4, we can calculate the hitting time between any two nodes using the formula above. For example

$$\begin{aligned} H(w_1,w_4) &= \sum_{k\in V} \ell(w_1,k;w_4)d(k) \\ &= \ell(w_1,w_1;w_4)d(k) + \ell(w_1,w_3;w_4)d(k) + \ell(w_1,w_2;w_4)d(k) \\ &= 2\cdot 1 + 1\cdot 3 + 1\cdot 1 \\ &= 6. \end{aligned}$$

As a symmetric graph, it is clear that nodes w_3 and w_4 are the two foci as well as the two barycenters of the tree. Thus

$$\begin{aligned} T_{\text{bestmix}} &= H(w_3,\pi) = H(w_5,w_3) - H(\pi,w_3) \\ &= H(w_5,w_3) - \left(\frac{1}{5}H(w_1,w_3) + \frac{3}{10}H(w_4,w_3) + \frac{1}{5}H(w_5,w_3)\right) \\ &= 6 - \left(\frac{1}{5}\times 1 + \frac{3}{10}\times 5 + \frac{1}{5}\times 6\right) \\ &= \frac{31}{10} \\ T_{\text{mix}} &= H(w_1,\pi) = H(w_3,\pi) + H(w_1,w_3) \\ &= \frac{31}{10} + 1 = \frac{41}{10}. \end{aligned}$$

◇

2.6 Special Trees

2.6.1 Stars

A star S_n is a complete bipartite graph $K_{1,n-1}$, on which all the nodes are leaves except for one, which is connected to all other nodes. Because of its simple structure, random walks on a star are often the easiest to study. The central node on the star is the unique focus and barycenter.

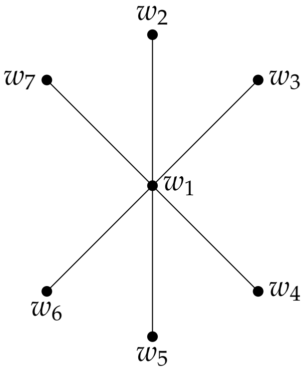

Figure 2.9: A star on 7 vertices

Example 8. On the star S_7 in Figure 2.9, w_1 is the central node. For $i, j = 2, 3, \ldots, 7, i \neq j$

$$H(w_i, w_1) = 1$$

$$H(w_1, w_i) = \sum_{w_k \in V} \ell(w_1, w_k; w_i) d(w_k) = 5 \times 1 \times 1 + 1 \times 6 = 11$$

$$H(w_i, w_j) = H(w_i, w_1) + H(w_1, w_j) = 1 + 11 = 12.$$

Thus every leaf is a w_1-pessimal node. For $2 \geq i \geq 7$, the w_i-pessimal nodes are all other leaves. The stationary distribution is

$$\pi = (\frac{1}{2}, \frac{1}{12}, \frac{1}{12}, \frac{1}{12}, \frac{1}{12}, \frac{1}{12}, \frac{1}{12}).$$

To determine the mixing times, we first calculate

$$H(\pi, w_1) = \sum_{w_k \in V} \pi_k H(w_k, w_1) = 6 \times \frac{1}{12} \times 1 = \frac{1}{2}$$

$$H(\pi, w_i) = \sum_{w_k \in V} \pi_k H(w_k, w_i) = 5 \times \frac{1}{12} \times 12 + \frac{1}{2} \times 11 = \frac{21}{2}$$

and the mixing times are

$$H(w_1, \pi) = H(w_1', w_1) - H(\pi, w_1) = 1 - \frac{1}{2} = \frac{1}{2}$$

$$H(w_i, \pi) = H(w_i', w_i) - H(\pi, w_i) = 12 - \frac{21}{2} = \frac{3}{2}.$$

So

$$T_{\text{bestmix}} = H(w_1, \pi) = \frac{1}{2}$$

$$T_{\text{mix}} = H(w_i, \pi) = \frac{3}{2}.$$

◇

2.6.2 Paths

A path P_n is a sequence of connected nodes such that each node is adjacent to the next node in the sequence. There are only two leaves on a path, and each node has one or two neighbors. The centers for a path depends on whether n is even or odd.

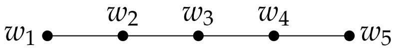

Figure 2.10: A path on 5 vertices

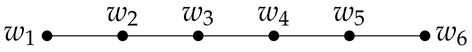

Figure 2.11: A path on 6 vertices

Example 9. On the path P_5 in Figure 2.10, clearly its unique focus and barycenter is w_3. The hitting times are

$H(w_i, w_j)$	w_1	w_2	w_3	w_4	w_5
w_1	0	1	4	9	16
w_2	7	0	3	8	15
w_3	12	5	0	5	12
w_4	15	8	3	0	7
w_5	16	9	4	1	0

Thus w_1 is the pessimal node for w_3, w_4, w_5, while w_5 is the pessimal node for w_1, w_2, w_3. The stationary distribution is

$$\pi = (\frac{1}{8}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{8}).$$

We have

$$H(\pi, w_1) = H(\pi, w_5) = \sum_{w_k \in V} \pi_k H(w_k, w_5) = \frac{1}{8} \times 16 + \frac{1}{4} \times (15 + 12 + 7) = \frac{21}{2}$$

$$H(\pi, w_2) = H(\pi, w_4) = \sum_{w_k \in V} \pi_k H(w_k w_4) = \frac{1}{8} \times (1 + 9) + \frac{1}{4} \times (8 + 5) = \frac{9}{2}$$

$$H(\pi, w_3) = \sum_{w_k \in V} \pi_k H(w_k, w_3) = 2 \times \frac{1}{8} \times 4 + 2 \times \frac{1}{4} \times 3 = \frac{5}{2}.$$

Therefore the mixing times are

$$\begin{aligned} H(w_1, \pi) &= H(w_5, \pi) = H(w_5', w_5) - H(\pi, w_5) \\ &= H(w_1, w_5) - H(\pi, w_5) = 16 - \frac{21}{2} = \frac{11}{2} \\ H(w_2, \pi) &= H(w_4, \pi) = H(w_4', w_4) - H(\pi, w_4) \\ &= H(w_1, w_4) - H(\pi, w_4) = 9 - \frac{9}{2} = \frac{9}{2} \\ H(w_3, \pi) &= H(w_3', w_3) - H(\pi, w_3) \\ &= H(w_1, w_3) - H(\pi, w_3) = 4 - \frac{5}{2} = \frac{3}{2}. \end{aligned}$$

So

$$T_{\text{bestmix}} = H(w_3, \pi) = \frac{3}{2}$$

$$T_{\text{mix}} = H(w_1, \pi) = H(w_5, \pi) = \frac{11}{2}.$$

On the path P_6 in Figure 2.11, there are two foci and two barycenters, w_3 and w_4. Each node has its unique pessimal node, and there are two nodes that achieve the best mix

$$T_{\text{bestmix}} = H(w_3, \pi) = H(w_4, \pi) = \frac{9}{2}$$

$$T_{\text{mix}} = H(w_1, \pi) = H(w_6, \pi) = \frac{17}{2}.$$

◇

2.6.3 Results on stars and paths

Here is a brief summary of important properties of random walks on stars and paths

Tree	Star S_n	Path P_n (n is even)	Path P_n (n is odd)
$\max_{i\in V} H(i',i)$	$2(n-1)$	$(n-1)^2$	$(n-1)^2$
$\min_{i\in V} H(i',i)$	1	$\frac{1}{4}n^2$	$\frac{1}{4}(n^2-2n+1)$
$\max_{i,j\in V} \kappa(i,j)$	$4(n-1)$	$2(n-1)^2$	$2(n-1)^2$
$\min_{i,j\in V} \kappa(i,j)$	$2(n-1)$	$2(n-1)$	$2(n-1)$
$\max_{i\in V} H(\pi,i)$	$2n-\frac{7}{2}$	$\frac{1}{6}(4n^2-8n+3)$	$\frac{1}{6}(4n^2-8n+3)$
$\min_{i\in V} H(\pi,i)$	$\frac{1}{2}$	$\frac{1}{6}(n^2-2n+3)$	$\frac{1}{6}(n^2-2n)$
T_{mix}	$\frac{3}{2}$	$\frac{1}{6}(2n^2-4n+3)$	$\frac{1}{6}(2n^2-4n+3)$
T_{bestmix}	$\frac{1}{2}$	$\frac{1}{12}(n^2+4n-6)$	$\frac{1}{12}(n^2-2n+3)$
$T_{\text{forget}}=T_{\text{reset}}$	1	$\frac{1}{4}(n^2-2n+2)$	$\frac{1}{4}(n^2-2n+1)$

Chapter 3

Extremal Hitting Times

First we find the maximizing and minimizing structures for maximal hitting time.

3.1 Results on hitting times

On a tree T, suppose the two leaves that achieve $\max_{i,j\in T} H_T(i,j)$ are a and b, if all vertices not on the path from a to b are leaves, then T is a *caterpillar*.

Lemma 8 *Given any tree T on n vertices $(n \geq 4)$, there exists a caterpillar T_c on n vertices such that*

$$\max_{i,j\in T} H_T(i,j) = \max_{i,j\in T_c} H_{T_c}(i,j).$$

Proof. Clearly the result is trivially true when T is a caterpillar itself. Let $H_T(a,b) = \max_{i,j\in T} H_T(i,j)$. Suppose T is not a caterpillar, there must exist some leaf x at least distance 2 away from the path from a to b. Let y be the vertex adjacent to x, and let c be the first vertex in common for the (a,b) path and the (x,b) path.

On T (Figure 3.1), we remove x, and add a new leaf z adjacent to c. Let the resulting tree be T_1 (Figure 3.2). Note that T_1 is also on n vertices. Since the path from a to b stays unchanged, and the total number of edges hanging off c remain the same, from Equation 2.2, $H_{T_1}(a,b) = H_T(a,b)$. Furthermore, since $H_{i,j\in T}(i,j) \geq H_{i,j\in T_1}(i,j)$, $\max_{i,j\in T_1} H_{T_1}(i,j) = H_{T_1}(a,b)$.

Repeat the above process until T_k is a caterpillar. We have

$$H_T(a,b) = H_{T_1}(a,b) = \cdots = H_{T_k}(a,b)$$

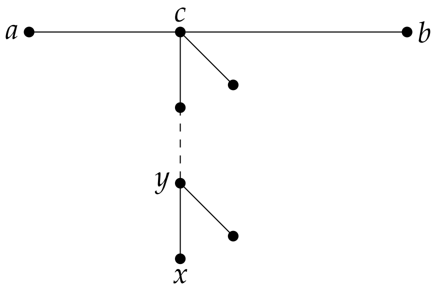

Figure 3.1: T

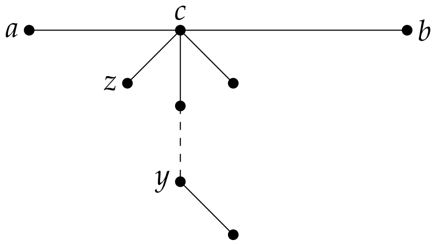

Figure 3.2: T_1

so

$$\max_{i,j\in T} H_T(i,j) = \max_{i,j\in T_k} H_{T_k}(i,j).$$

▮

The above lemma allows us to transform any given tree into a relatively simple structure and preserve its maximal hitting time. The following tells us about the vertices that achieve the maximal hitting time.

Lemma 9 *On any given tree T, let*

$$H_T(b,a) = \max_{i,j\in T} H_T(i,j)$$

Then a and b must both be leaves on T.

Proof. Assume for the sake of contradiction that either a or b is not a leaf.

Case 1. If a is a vertex that is not a leaf, then there exists a vertex c adjacent to a such that c is not on the path between a and b. Since T is a tree, the path between c and b must include a. Then $H(c,b) = H(c,a) + H(a,b)$, so $H(c,b) > H(a,b)$, which is a contradiction.

Case 2. If b is a vertex that is not a leaf, then there exists a vertex c adjacent to b such that c is not on the path between a and b. Since T is a tree, the path between c and a must include b. Then $H(a,c) = H(a,b) + H(b,c)$, so $H(a,c) > H(a,b)$, which is a contradiction.

Therefore a and b must both be leaves. ▮

With the two lemmas above, we can now prove that the maximizing structure for maximal hitting time is the path by a series of transformation of any given tree.

Theorem 10 *Of all trees T on n vertices*

$$\max_T \max_{i,j \in T} H_T(i,j) = (n-1)^2.$$

This value is achieved uniquely by the path P_n.

Proof.

First note that on P_n, i and j would be the two leaves, and $\max_{i \in P_n} H_{P_n}(i,j) = (n-1)^2$.

Given any tree T on n vertices, from Lemma 8, there is a caterpillar T_{cat} (Figure 3.3) on n vertices such that $\max_{i,j \in T} H_T(i,j) = \max_{i,j \in T_{cat}} H_{T_{cat}}(i,j)$. On T_{cat}, suppose $H_{T_{cat}}(a,b) = \max_{i,j \in T_{cat}} H_{T_{cat}}(i,j)$. By Lemma 9, a and b are both leaves. Let c be the adjacent vertex to a.

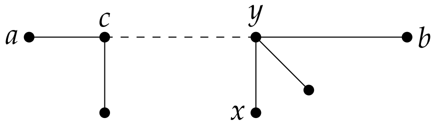

Figure 3.3: T_{cat}

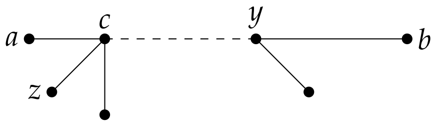

Figure 3.4: T_1

If there exists a leaf x on T_{cat} that is not adjacent to c, except for b, we remove x and then add a new leaf z adjacent to c. Let the resulting tree be T_1 (Figure 3.4), which is also on n vertices. Suppose x was adjacent to y on T_{cat}. All vertices have the same degrees on the two trees except for the following

$$\deg_{T_1}(c) - \deg_{T_{cat}}(c) = 1$$
$$\deg_{T_{cat}}(y) - \deg_{T_1}(y) = 1$$
$$\deg_{T_{cat}}(x) = 1$$
$$\deg_{T_1}(z) = 1.$$

So

$$\begin{aligned} H_{T_{cat}}(a,b) &= \sum_{k \in V(T_{cat})} \ell(a,k;b)d(k) \\ &= \left(\sum_{k \in V(T_1), k \neq c,y,x} \ell(a,k;b)d(k)) \right) \\ &\quad + \ell(a,x;b) + \deg_{T_{cat}}(c)\ell(a,c;b) + \deg_{T_{cat}}(y)\ell(a,y;b) \end{aligned}$$

and

$$\begin{aligned} H_{T_1}(a,b) &= \sum_{k \in V(T_1)} \ell(a,k;b)d(k) \\ &= \left(\sum_{k \in V(T_1), k \neq c,y,z} \ell(a,k;b)d(k)) \right) \\ &\quad + \ell(a,z;b) + \deg_{T_1}(c)\ell(a,c;b) + \deg_{T_1}(y)\ell(a,y;b). \end{aligned}$$

Therefore

$$\begin{aligned} H_{T_1}(a,b) - H_{T_{cat}}(a,b) &= (\ell(a,c;b) - \ell(a,y;b)) + (\ell(a,z;b) - \ell(a,x;b)) \\ &= 2\,(\ell(a,c;b) - \ell(a,y;b)) \\ &= 2\mathrm{d}(c,y) \\ &> 0. \end{aligned}$$

Repeat the above process until on T_k (Figure 3.5), all leaves in $V(T) - \{a,b\}$ are adjacent to c. We have

$$H_{T_{cat}}(a,b) < H_{T_1}(a,b) < \cdots < H_{T_k}(a,b)$$

Figure 3.5: T_k

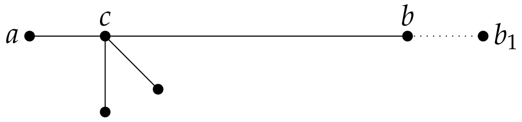

Figure 3.6: T_k^1

On T_k, if there is any leaf x other than a and b, we remove it and attach a new leaf b_1 adjacent to b. Let the resulting tree be T_k^1 (Figure 3.6), which is also on n vertices. The new hitting time maximizing vertices are a and b_1 Similar to the analysis above, we can show that

$$H_{T_k^1}(a,b_1) - H_{T_k}(a,b) \geq \deg_{T_k^1}(c)\ell(a,c;b_1) + 1 > 0.$$

Repeat the above process until the only leaves on the tree are a and b_i. The resulting tree is the unique path P_n. We have

$$H_{T_k}(a,b) < H_{T_k^1}(a,b_1) < \cdots < H_{P_n}(a,b_i).$$

After various transformations from T_0 to P_n, we have

$$\begin{aligned}\max_{i,j\in T_0} H_{T_0}(i,j) &= \max_{i,j\in T_{cat}} H_{T_{cat}}(i,j)\\ &< \max_{i,j\in T_k} H_{T_k}(i,j) < \max_{i,j\in P_n} H_{P_n}(i,j).\end{aligned}$$

Therefore the path P_n is the unique tree on n vertices such that

$$\max_{i\in P_n} H_{P_n}(i',i) = \max_{T}\max_{i\in T} H_T(i',i).$$

∎

After proving the result for the maximizing structure for maximal hitting time, we now consider the easier result for the minimizing structure.

Theorem 11 *Of all trees T on n vertices*

$$\min_{T}\max_{i\in T} H_T(i',i) = 2(n-1).$$

This value is achieved uniquely by the star S_n.

Proof.

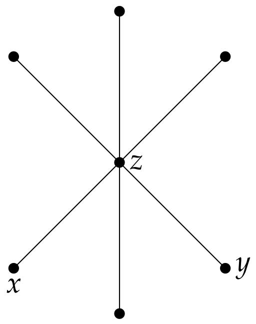

Figure 3.7: S_n

First we show that $\max_{i\in S_n} H_{S_n}(i',i) = 2(n-1)$. Let two arbitrary leaves on S_n (Figure 3.7) be x and y, and the center vertex be z. All vertices in S_n have degree of 1 except for z, whose degree is $n-1$. Note $\ell(x,k;y)=1$ for all $k \neq x,y$, since the only edge they share with the path between x and y is yz. So,

$$\begin{aligned}H(x,y) &= \ell(x,x;y)d(x) + \ell(x,z;y)d(z) + \sum_{k\neq x,y,z} \ell(x,k;y)d(k)\\ &= 2\times 1 + 1\times(n-1) + (n-3)\times 1 = 2(n-1).\end{aligned}$$

Assume for the sake of contradiction that there exists some tree $T_0 \neq S_n$ such that $\max_{i \in T_0} H_{T_0}(i', i) = \min_T \max_{i \in T} H_T(i', i)$. Then $\max_{i \in T_0} H_{T_0}(i', i) \leq 2(n-1)$.

Since T_0 is not a star, there must exist two leaves a and b such that $\mathrm{d}(a, b) > 2$. Then there must be some vertex x that is adjacent to a such that x is not adjacent to b. We have

$$H_{T_0}(b, a) = H_{T_0}(b, x) + H_{T_0}(x, a).$$

Since a and x are adjacent, $H_{T_0}(x, a) = 2(n-1) - 1 = 2n - 3$. In order to make $H_{T_0}(b, a) \leq max_{i \in S_n} H_{S_n}(i', i) = 2(n-1)$, we must have $H_{T_0}(b, x) \leq 1$,which is impossible since b and x are not adjacent (the length of the path between a and b is greater than 2).

Therefore the only tree on n vertices that achieves $\min_T \max_{i \in T} H_T(i', i)$ is the star S_n. ∎

Theorem 12 *Of all focal trees T on n vertices, where n is odd*

$$\max_T \min_{i \in T} H_T(i', i) = (n-1)^2/4.$$

This value is achieved uniquely by the odd path P_n.

Proof. On the odd path P_n, $H_{P_n}(a', a) = \min_{i \in P_n} H_{P_n}(i', i) = (n-1)^2/4$, where a is the unique focus. Note that $H_{P_n}(a', a) = H_{P_n}(a'', a)$.

Assume for the sake of contradiction that there exists a focal tree $T_0 \neq P_n$ on n vertices such that

$$H_{T_0}(a', a) = \min_{i \in T_0} H_{T_0}(i', i) \geq (n-1)^2/4.$$

Since T_0 is focal, we again have $H_{T_0}(a', a) = H_{T_0}(a'', a)$. Then $H_{T_0}(a', a) = H_{T_0}(a'', a) \geq H_{P_n}(a', a) = H_{P_n}(a'', a)$. Let the vertex adjacent to a on T_0 that is on the (a', a) path be x. Since $H_{T_0}(a'', a) \geq H_{P_n}(a'', a)$, $H_{T_0}(a', a) \geq H_{P_n}(a', a)$, and T_0 is not a path, by Theorem 10, we must have $v(V_{x:a}) \geq (n-1)/2$, and $v(V_{a:x}) \geq (n+1)/2$, so $v(V_{x:a}) = (n-1)/2$, and $v(V_{a:x}) = (n+1)/2$.

However, then by Theorem 10, we have $H_{T_0}(a'', a) < H_{P_n}(a'', a)$ or $H_{T_0}(a', a) < H_{P_n}(a', a)$, a contradiction. Therefore of all focal trees T on odd number of vertices, the odd path P_n uniquely achieves $\max_T \min_{i \in T} H_T(i', i)$. ∎

After we have the results for extremal hitting times, we look at the extremal commute times, which are closely related to the hitting times.

3.2 Results on commute times

First we introduce several lemmas on the commute times that follow easily from the previous section.

Lemma 13 *On any tree T, if c is on the path between a and b, we have*

$$\kappa(a,b) = \kappa(a,c) + \kappa(b,c)$$

Proof.

$$\begin{aligned} \kappa(a,b) &= H(a,b) + H(b,a) = (H(a,c) + H(c,b)) + (H(b,c) + H(c,a)) \\ &= (H(a,c) + H(c,a)) + (H(b,c) + H(c,b)) \\ &= \kappa(a,c) + \kappa(b,c) \end{aligned}$$

which also implies $\kappa(a,c) < \kappa(a,b)$ and $\kappa(b,c) < \kappa(a,b)$. ▮

Lemma 14 *Of any given tree T on n vertices*

$$\min_{i,j \in T} \kappa(i,j) = 2(n-1).$$

This value is achieved by two arbitrary adjacent vertices.

Proof. From Equation 2.1, for any two adjacent vertices i, j on any given tree T, we have

$$\kappa(i,j) = H(i,j) + H(j,i) = 2(n-1).$$

Assume for the sake of contradiction that $\kappa(a,b) = \min_T min_{i,j \in T} \kappa(i,j)$, where a and b are not adjacent, then there must exist some vertex c on the path between a and b. From Lemma 13, we have

$$\kappa(a,c) < \kappa(a,b)$$

a contradiction. Therefore $\min_{i,j \in T} \kappa(i,j)$ is achieved by any two adjacent vertices on T. ▮

Lemma 15 *Of any given tree T on n vertices* $\max_{i,j \in T} \kappa(i,j)$ *is achieved by two leaves.*

Proof. Let $\kappa(a,b) = \max_{i,j\in T}\kappa(i,j)$. Assume for the sake of contradiction that at least one of a and b is not a leaf. Since $\kappa(a,b) = \kappa(b,a)$, without loss of generality, assume a is not a leaf. Then there exist some vertex c adjacent to a but not adjacent to b. From Lemma 13, we have

$$\kappa(c,b) > \kappa(a,b)$$

a contradiction. Therefore $\max_{i,j\in T}\kappa(i,j)$ is achieved by two leaves. ▮

From Lemma 14, it directly follows that

Theorem 16 *Of all trees T on n vertices*

$$\min_{T}\min_{i,j\in T}\kappa(i,j) = \max_{T}\min_{i,j\in T}\kappa(i,j) = 2(n-1).$$

This value is achieved by two adjacent vertices on any T.

The following two theorems look at the maximizing and minimizing structures for maximal commute times.

Theorem 17 *Of all trees T on n vertices*

$$\min_{T}\max_{i,j\in T}\kappa(i,j) = 4(n-1).$$

This value is achieved uniquely by the star S_n.

Proof. On S_n, any two leaves achieves $\max_{i,j\in S_n}\kappa(i,j) = 2\times 2(n-1) = 4(n-1)$.

Assume for the sake of contradiction that
$\max_{i,j\in T_0}\kappa(i,j) = \min_T \max_{i,j\in T}\kappa(i,j)$ and that $T_0 \neq S_n$.

Since T_0 is not a star, there must exist two leaves a and b such that $\mathrm{d}(a,b) > 2$. Let c be adjacent to a and on the path between a and b. From Lemma 13 and Lemma 14, we have

$$\kappa(a,b) = \kappa(a,c) + \kappa(b,c) > 2(n-1) + 2(n-1) = 4(n-1)$$

which implies

$$\max_{i,j\in T_0}\kappa(i,j) > \kappa(a,b) > 4(n-1) = \max_{i,j\in S_n}\kappa(i,j)$$

a contradiction. Therefore $\min_T \max_{i,j\in T}\kappa(i,j)$ is achieved uniquely by S_n. ▮

Theorem 18 *Of all trees T on n vertices*

$$\max_{T} \max_{i,j \in T} \kappa(i,j) = 2(n-1)^2$$

This value is achieved uniquely by the path P_n.

Proof. Let the two leaves on P_n be a and b. Since $H_{T_0}(a,b) = H_{T_0}(b,a) = \max_T \max_{i,j \in T} H(i,j)$, it follows that

$$\max_{T} \max_{i,j \in T} \kappa(i,j) = \kappa_{T_0}(a,b) = H_{T_0}(a,b) + H_{T_0}(b,a) = 2 \times (n-1)^2 = 2(n-1)^2$$

Therefore $\min_T \max_{i,j \in T} \kappa(i,j)$ is achieved uniquely by P_n. ▮

Chapter 4

Extremal Results on $H(\pi, i)$

In order to study maximal mixing times, we need to first look at

$$H(\pi, i) = \sum_{k \in V} \pi_k H(k, i),$$

as it is part of the equation 2.8,

$$H(i, \pi) = H(i', i) - H(\pi, i).$$

Recall from section 2.2.2 that for any adjacent vertices $u, v \in V$, let $V_{u:v}$ denote the set of vertices in the subtree rooted at u after the removal of the edge uv, as shown in Figure 2.5.

First we show a lemma for the vertex that achieves the maximal $H(\pi, i)$.

Lemma 19 *On any tree T, the vertex i that achieves $\max_{i \in T} H_T(\pi, i)$ must be a leaf.*

Proof. Assume for the sake of contradiction that $H(\pi, a) = \max_{i \in T} H_T(\pi, i)$ and a is not a leaf. Let $G_1, G_2, ..., G_d$ be the components of $T - a$, where $\pi(G_1) \leq \pi(G_2) \leq \cdots \leq \pi(G_d)$, and $d = \deg(a) \geq 2$.

Let the vertex in G_1 that is adjacent to a be b. Note that $\pi(V_{b:a}) < \pi(V_{a:b})$. Now,

$$\begin{aligned} H_T(\pi, a) &= \sum_{x \in V_{a:b}} \pi_x H_T(x, a) + \sum_{y \in V_{b:a}} \pi_y H_T(y, a) \\ &= \sum_{x \in V_{a:b}} \pi_x H_T(x, a) + \sum_{y \in V_{b:a}} \pi_y \left((H_T(y, b) + H_T(b, a) \right). \end{aligned}$$

Similarly,

$$H_T(\pi, b) \;=\; \sum_{y \in V_{b:a}} \pi_y H_T(y, b) + \sum_{x \in V_{a:b}} \pi_x \left((H_T(x, a) + H_T(a, b)\right).$$

Then,

$$\begin{aligned}
& H_T(\pi, b) - H_T(\pi, a) \\
= \;& \sum_{x \in V_{a:b}} \pi_x H_T(a, b) - \sum_{y \in V_{b:a}} \pi_y H_T(b, a) \\
= \;& \pi(V_{a:b}) H_T(a, b) - \pi(V_{b:a}) H_T(b, a) \\
= \;& \frac{\pi(V_{a:b})}{2|E(T)|} \left(\sum_{x \in V_{a:b}} \deg(x) \right) - \frac{\pi(V_{b:a})}{2|E(T)|} \left(\sum_{y \in V_{b:a}} \deg(y) \right) \\
= \;& \frac{1}{2|E(T)|} \left(\pi(V_{a:b})^2 - \pi(V_{b:a})^2 \right) \\
> \;& 0.
\end{aligned}$$

This contradicts our assumption that $H_T(\pi, a) = \max_{i \in T} H_T(\pi, i)$. Therefore the vertex i that achieves $\max_{i \in T} H_T(\pi, i)$ must be a leaf. ▮

We show the following two results for the minimizing structure of both maximal and minimal $H(\pi, i)$.

Theorem 20 *Of all trees T on n vertices*

$$\min_T \min_{i \in T} H_T(\pi, i) = \frac{1}{2}.$$

This value is achieved uniquely by the star S_n.

Proof.
First we show that $\min_{i \in S_n} H_{S_n}(\pi, i) = \frac{1}{2}$. In the stationary distribution π, the center vertex, denoted as c here, has $\pi_c = \frac{1}{2}$, while all other vertices have $\pi_i = \frac{1}{2(n-1)}$. We know that for all leaves i on S_n, $H(i, c) = 1$. Clearly the vertex that achieves $\min_{i \in S_n} H_{S_n}(\pi, i)$ is c. We have

$$H_{S_n}(\pi, c) = \sum_{i \in V(S_n)} \pi_i H_{S_n}(i, c) = \sum_{i \in V(S_n), i \neq c} \pi_i = \frac{(n-1)}{2(n-1)} = \frac{1}{2}.$$

Assume for the sake of contradiction that there exists some tree $T_0 \neq S_n$ such that $H_{T_0}(\pi, x) = \min_{i \in T_0} H_{T_0}(\pi, i) \leq \frac{1}{2}$. Since T_0 is not a star, $\deg(x) <$

$(n-1)$, so $\pi_x < \frac{1}{2}$. Therefore

$$H_{T_0}(\pi, x) = \sum_{i \neq x} \pi_i H_{T_0}(i, x) \geq \pi(V - \{x\}) \times 1 > \frac{1}{2}$$

Which is a contradiction. Therefore S_n is the unique tree on n vertices that achieves $\min_T \min_{i \in T} H_T(\pi, i)$. ∎

Theorem 21 *Of all trees T on n vertices*

$$\min_T \max_{i \in T} H_T(\pi, i) = 2n - \frac{7}{2}.$$

This value is achieved uniquely by the star S_n.

Proof.

Assume for the sake of contradiction that there exists some tree $T_0 \neq S_n$ such that $\max_{i \in T_0} H_{T_0}(\pi, i) \leq \max_{i \in S_n} H_{S_n}(\pi, i)$. Suppose the vertex that achieves this on T_0 is a. From Lemma 19 we know that a must be a leaf. Let the unique neighbor of a be b. Since T_0 is not a star, $\deg_{T_0}(b) < n - 1 = \deg_{S_n}(b)$. Any vertex $c \in V - \{a, b\}$ must have

$$H_{T_0}(c, a) = H_{T_0}(c, b) + H_{T_0}(b, a) > 1 + 2(n-1) - 1 = H_{S_n}(c, a).$$

Since b is the unique neighbor of leaf a in both T_0 and S_n

$$H_{T_0}(b, a) = H_{S_n}(b, a) < H_{S_n}(c, a).$$

Then

$$\begin{aligned}
& H_{T_0}(\pi, a) - H_{S_n}(\pi, a) \\
= \; & \sum_{i \in V(T_0), i \neq a} \pi_i H_{T_0}(i, a) - \sum_{i \in V(S_n), i \neq a} \pi_i H_{S_n}(i, a) \\
= \; & \left(\pi_{b \in T_0} H_{T_0}(b, a) + \sum_{i \in V(T_0), i \neq b} \pi_i H_{T_0}(i, a) \right) \\
- \; & \left(\pi_{b \in S_n} H_{S_n}(b, a) + \sum_{i \in V(S_n), i \neq b} \pi_i H_{S_n}(i, a) \right) \\
> \; & \left(\pi_{b \in T_0} - \pi_{b \in S_n} \right) H_{S_n}(b, a) + \left(\sum_{i \in V(T_0), i \neq b} \pi_i - \sum_{i \in V(S_n), i \neq b} \pi_i \right) H_{S_n}(c, a) \\
= \; & \left(\pi_{b \in S_n} - \pi_{b \in T_0} \right) \left(H_{S_n}(c, a) - H_{S_n}(b, a) \right) \\
= \; & \left(\pi_{b \in S_n} - \pi_{b \in T_0} \right) H_{S_n}(c, b) \\
> \; & 0
\end{aligned}$$

Which is a contradiction. Therefore S_n is the unique tree on n vertices that achieves $\min_T \max_{i \in T} H_T(\pi, i)$. ▮

We show that the maximizing structure for maximal $H(\pi, i)$ is the path with proof by contradiction and carefully arguing the result of moving a single vertex on a tree.

Theorem 22 *Of all trees T on n vertices*

$$\max_T \max_{i \in T} H_T(\pi, i) = \frac{4n^2 - 8n + 3}{6}.$$

This value is achieved uniquely by the path P_n.

Proof.

Assume for the sake of contradiction that there exists some tree $T_0 \neq P_n$ such that $H_{T_0}(\pi, a) = \max_{i \in T_0} H_{T_0}(\pi, i) = \max_T \max_{i \in T} H_T(\pi, i)$. Let the a-pessimal node be a'.

From Lemma 19 we know that a must be a leaf. Since $T_0 \neq P_n$, there must exist at least one leaf other than a and a'. Let z be the vertex on T_0 with $\deg(z) > 2$ such that all internal vertices on the path between a' and z have degree 2. Let x be a leaf on T_0 other than a' and a, where the path between a' and x contains z. Let the vertex adjacent to x be y.

Now, on T_0 (Figure 4.1) remove vertex x, and add a new leaf a^* adjacent to a'. The resulting tree T_1 (Figure 4.2) still has n vertices.

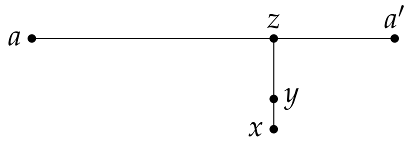

Figure 4.1: T_0

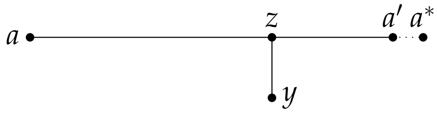

Figure 4.2: T_1

We claim that

$$\max_{i \in T_1} H_{T_1}(\pi, i) \geq H_{T_1}(\pi, a) \geq H_{T_0}(\pi, a) = \max_{i \in T_0} H_{T_0}(\pi, i).$$

Consider the components of $V - \{z\}$. Let the one containing y be V_y, the one containing a' be $V_{a'}$, and the rest of the components collectively $(V_y \cup V_{a'})^c$. Define $\Delta(i, j) = H_{T_1}(i, j) - H_{T_0}(i, j)$. Clearly for all vertices c in $(V_y \cup V_{a'})^c$ in T_1, we have $\Delta(c, z) = 0$, $\deg_{T_1}(c) = \deg_{T_0}(c)$. For all vertices $c \in V_y$, we have $\Delta(c,z) = -(\ell(c,y,z) + \ell(c,x,z)) = -2\ell(c,y,z)$. For all vertices $c \in V_{a'}$, $c \neq a^*$, we have $\Delta(c,z) = \ell(c,a',z) + \ell(c,a^*,z) = 2\ell(c,a',z)$. For i=0,1

$$
\begin{aligned}
H_{T_i}(\pi,a) &= \sum_{k\in V(T_i)} \frac{\deg(k)}{2|E|} H_{T_i}(k,a) \\
&= \left(\sum_{k\in \overline{V(T_a)}} \frac{\deg(k)}{2|E|} H_{T_i}(k,z) + H_{T_i}(z,a) \sum_{k\in \overline{V(T_a)}} \frac{\deg(k)}{2|E|} \right) \\
&+ \sum_{k\in V(T_a)} \frac{\deg(k)}{2|E|} H_{T_i}(k,a)
\end{aligned}
$$

Then,

$$
\begin{aligned}
& H_{T_1}(\pi,a) - H_{T_0}(\pi,a) \\
= & \left(\sum_{k\in V_y\cup V_{a'}} \frac{\deg(k)}{2|E|} H_{T_1}(k,z) + H_{T_1}(z,a) \sum_{k\in V_y\cup V_{a'}} \frac{\deg(k)}{2|E|} \right) \\
& - \left(\sum_{k\in V_y\cup V_{a'}} \frac{\deg(k)}{2|E|} H_{T_0}(k,z) + H_{T_0}(z,a) \sum_{k\in V_y\cup V_{a'}} \frac{\deg(k)}{2|E|} \right) \\
= & \sum_{k\in V_y\cup V_{a'}} \frac{\deg(k)}{2|E|} H_{T_1}(k,z) - \sum_{k\in V_y\cup V_{a'}} \frac{\deg(k)}{2|E|} H_{T_0}(k,z) \\
= & \frac{\deg(a^*)}{2|E|} H_{T_1}(a^*,z) - \frac{\deg(x)}{2|E|} H_{T_0}(x,z) \\
& + \sum_{k\in V_y} \frac{\deg(k)}{2|E|} \Delta(k,z) + \sum_{k\in V_{a'}, k\neq a^*} \frac{\deg(k)}{2|E|} \Delta(k,z) \\
= & \frac{1}{2|E|} (H_{T_1}(a^*,z) - H_{T_0}(x,z) \\
& + 2 \sum_{k\in V_{a'}, k\neq a^*} \deg(k)\ell(k,a',z) - 2 \sum_{k\in V_y} \deg(k)\ell(k,y,z)) \\
= & \frac{1}{2|E|} (H_{T_1}(a^*,z) - H_{T_0}(x,z) + 2H_{T_1}(a',z) - 2H_{T_0}(y,z)).
\end{aligned}
$$

We have,

$$H_{T_1}(a^*,z) > H_{T_1}(a',z) > H_{T_0}(a',z) \geq H_{T_0}(x,z) > H_{T_0}(y,z).$$

So,

$$\max_{i \in T_1} H_{T_1}(\pi, i) \geq H_{T_1}(\pi, a) > H_{T_0}(\pi, a) = \max_{i \in T_0} H_{T_0}(\pi, i)$$

which is a contradiction to our assumption. Therefore the only tree on n vertices that achieves $\max_T \max_{i \in T} H_T(\pi, i)$ is $T = P_n$. ∎

The following theorem shows that the maximizing structure for minimal $H(\pi, i)$ is the path using similar method in the previous theorem. It can take on two values depending on whether the path is even or odd, since an even path is bifocal while an odd path is focal.

Theorem 23 *Of all trees T on n vertices*

$$\max_T \min_{i \in T} H_T(\pi, i) = \begin{cases} \frac{n^2-2n+3}{6} & \text{if } n \text{ is even} \\ \frac{n^2-2n}{6} & \text{if } n \text{ is odd} \end{cases}$$

This value is achieved uniquely by the path P_n.

Proof. Assume for the sake of contradiction that there exists some tree $T_0 \neq P_n$ such that $H_{T_0}(\pi, a) = \min_{i \in T_0} H_{T_0}(\pi, i) = \max_T \min_{i \in T} H_T(\pi, i)$. Let $G_1, G_2, ..., G_d$ be the components of $T_0 - a$, where $\pi(G_1) \geq \pi(G_2) \geq \cdots \geq \pi(G_d)$, and $d = \deg(a) \geq 2$. Suppose the vertex adjacent to a in each component G_i is b_i.

$$\begin{aligned} \max_T \min_{i \in T} H(\pi, i) = H(\pi, a) &= \sum_{k \in V(T_0)} \frac{\deg(k)}{2|E|} H(k, a) \\ = \sum_{k \in V(G_1)} \frac{\deg(k)}{2|E|} H(k, a) &+ \cdots + \sum_{k \in V(G_d)} \frac{\deg(k)}{2|E|} H(k, a). \end{aligned}$$

Within each component G_i

$$\begin{aligned} \sum_{k \in G_i} \frac{\deg(k)}{2|E|} H(k, a) &= \sum_{k \in G_i} \frac{\deg(k)}{2|E|} \left(H(k, b_i) + H(b_i, a)\right) \\ &= \sum_{k \in G_i} \frac{\deg(k)}{2|E|} H(k, b_i) + \frac{H(b_i, a)}{2|E|} \sum_{k \in G_i} \deg(k) \\ &= H_{G_i}(\pi, b_i) + \frac{H_{G_i}(b_i, a)}{2|E|} \sum_{k \in G_i} \deg(k) \end{aligned}$$

Since the right hand side of the expression is constant for each G_i, $H_{G_i}(\pi, b_i)$ must be maximized. From Theorem 22, G_i must each be a path. If $d = 2$, T_0 would be a path, so $d \geq 3$. Since $\pi(G_1) \geq \pi(G_2) \geq \pi(G_3)$, and a is the barycenter, $\pi(G_1) \leq \frac{1}{2}$, so $\pi(G_1) \leq \frac{1}{2} - \frac{1}{n-1}$. Suppose the single leaf in G_3 is x, and x is adjacent to y. Suppose the single leaf in G_2 is a^*.

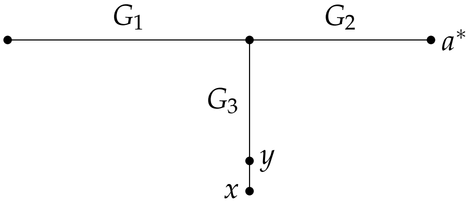

Figure 4.3: T_0

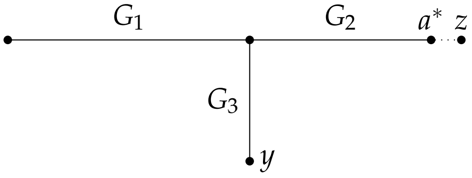

Figure 4.4: T_1

Now on (Figure 4.3), remove x in G_3, and add a new leaf z adjacent to a^* in G_2. Let the resulting tree be T_1 (Figure 4.4), which is still on n vertices. Since the swap of vertices have no effect on components other than G_2 or G_3

$$H_{T_1}(\pi, a) - H_{T_0}(\pi, a) = \sum_{k \in V(T_1)} \frac{\deg(k)}{2|E|} H_{T_1}(k, a) - \sum_{k \in V(T_0)} \frac{\deg(k)}{2|E|} H_{T_0}(k, a)$$

$$= \left(\sum_{k \in G_2 \subset V(T_1)} \frac{\deg_{T_1}(k)}{2|E|} H_{T_1}(k, a) + \sum_{k \in G_3 \subset V(T_1)} \frac{\deg_{T_1}(k)}{2|E|} H_{T_1}(k, a) \right)$$

$$- \left(\sum_{k \in G_2 \subset V(T_0)} \frac{\deg_{T_0}(k)}{2|E|} H_{T_0}(k, a) + \sum_{k \in G_3 \subset V(T_0)} \frac{\deg_{T_0}(k)}{2|E|} H_{T_0}(k, a) \right)$$

Using the same argument as in Theorem 22, we can show that

$$H_{T_1}(\pi, a) > H_{T_0}(\pi, a)$$

Which is a contradiction to our assumption, therefore the only tree on n vertices that achieves $\max_T \min_{i \in T} H_T(\pi, i)$ is $T = P_n$. ∎

Chapter 5

Extremal Mixing Times

The last chapter studies the extremal structures for mixing times. We can calculate mixing time with the equation 2.8,

$$H(i,\pi) = H(i',i) - H(\pi,i).$$

5.1 Results on T_{mix}

Recall from section 2.5,

$$T_{\text{mix}} = \max_{i\in V} H(i,\pi)$$

Lemma 24 *On any tree T, the vertex i that achieves $T_{\text{mix}} = \max_{i\in T} H_T(i,\pi)$ must be a leaf.*

Proof.

Assume for the sake of contradiction that $H(a,\pi) = \max_{i\in T} H_T(i,\pi)$ and a is not a leaf. Suppose the nearest focus to a is u. Let b be a leaf such that $b \in V_{a:u}$. Then, by Theorem 6,

$$\begin{aligned} H(b,\pi) &= H(b,a) + H(a,\pi) \\ &= H(b,a) + H(a,u) + H(u,\pi) \\ &= H(b,a) + H(a,\pi) \\ &> H(a,\pi) \end{aligned}$$

a contradiction. Therefore the vertex i that achieves $\max_{i\in T} H_T(i,\pi)$ must be a leaf. ▮

First we show that the minimizing structure for the mixing time is the star, and that the value is constant regardless of n.

Theorem 25 *Of all trees T on n vertices*

$$\min_T T_{\text{mix}} = \min_T \max_{i\in T} H_T(i,\pi) = \frac{3}{2}$$

This is achieved uniquely by the star S_n.

Proof. By Lemma 24, for any tree T, the vertex that maximizes $H_T(i,\pi)$ must be a leaf a. Since a is a leaf, a is not a focus. Suppose the nearest focus to a is u. By Theorem 6,

$$H_T(a,\pi) = H_T(a,u) + H_T(u,\pi).$$

Since $a \neq u$,

$$H_T(a,u) \geq 1.$$

By Theorem 27,

$$H_T(a,\pi) = H_T(a,u) + H_T(u,\pi) \geq 1 + H_T(u,\pi) \geq 1 + \frac{1}{2} = \frac{3}{2}.$$

where the equality holds if and only if $T = S_n$.

Therefore the only tree on n vertices that achieves $\min_T \max_{i\in T} H_T(i,\pi)$ is $T = S_n$. ∎

To show that the maximizing structure for the mixing time is the path, we use a similar method to that used in the proof of Theorem 22 and combined results from both chapter 3 and 4.

Theorem 26 *Of all trees T on n vertices*

$$\max_T T_{\text{mix}} = \max_T \max_{i\in T} H_T(i,\pi) = \frac{2n^2-4n+3}{6}$$

This value is achieved uniquely by the path P_n.

Proof. Assume for the sake of contradiction that there exists some tree $T_0 \neq P_n$ such that $\max_{i\in T_0} H_{T_0}(i,\pi) = \max_T \max_{i\in T} H_T(i,\pi)$. Let $H_{T_0}(a,\pi) = \max_{i\in T_0} H_{T_0}(i,\pi)$ with a-pessimal node a'. By Lemma 24 we know that a must be a leaf. Since $T_0 \neq P_n$, there must exist at least one leaf other than a and a'. Let z be the vertex on T_0 that satisfies $\deg(z) > 2$ and all internal vertices on the path between a' and z have degree 2. Let x be a leaf on T_0 other than a' and a, where the path between a' and x contains z. Let the vertex adjacent to x be y.

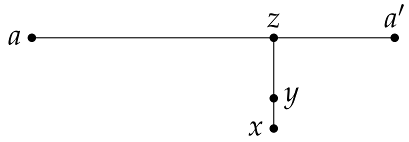

Figure 5.1: T_0

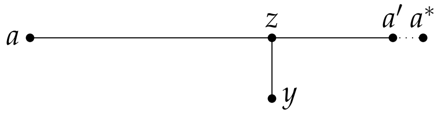

Figure 5.2: T_1

Now, on T_0 (Figure 5.1), remove vertex x, and add a new leaf a^* adjacent to a'. The resulting tree T_1 (Figure 5.2) still has n vertices.

By Equation 2.8, we have $H(i, \pi) = H(i', i) - H(\pi, i)$. Suppose the path from a' to z contains k vertices. We have,

$$H_{T_1}(a, \pi^{(1)}) = H_{T_1}(a^*, a) - H_{T_1}(\pi^{(1)}, a)$$

$$H_{T_0}(a, \pi^{(0)}) = H_{T_0}(a', a) - H_{T_0}(\pi^{(0)}, a)$$

where $\pi^{(i)}$ is the stationary distribution for T_i. So,

$$\begin{aligned} & H_{T_1}(a, \pi^{(1)}) - H_{T_0}(a, \pi^{(0)}) \\ = \ & \Big(H_{T_1}(a^*, a) - H_{T_1}(\pi^{(1)}, a)\Big) - \Big(H_{T_0}(a', a) - H_{T_0}(\pi^{(0)}, a)\Big) \\ = \ & \left(H_{T_1}(a^*, a) - H_{T_0}(a', a)\right) - \Big(H_{T_1}(\pi^{(1)}, a) - H_{T_0}(\pi^{(0)}, a)\Big) \end{aligned}$$

First let us calculate $H_{T_1}(a^*, a) - H_{T_0}(a', a)$. Since $H(i, j) = \sum_{k \in V} \ell(i, k; j) \deg(k)$,

$$\begin{aligned} H_{T_0}(a', a) &= \sum_{k \in V(T_0)} \ell_{T_0}(a', k; a) \deg_{T_0}(k) \\ &= \left(\sum_{k \in V(T_0), k \neq a', x, y} \ell_{T_0}(a', k; a) \deg_{T_0}(k) \right) \\ &\quad + \ell_{T_0}(a', a'; a) + \ell_{T_0}(a', x; a) + \deg_{T_0}(y) \ell_{T_0}(a', y; a) \end{aligned}$$

$$\begin{aligned} H_{T_1}(a^*, a) &= \sum_{k \in V(T_1)} \ell_{T_1}(a^*, k; a) \deg_{T_1}(k) \\ &= \left(\sum_{k \in V(T_1), k \neq a^*, a', y} \ell_{T_1}(a', k; a) \deg_{T_1}(k) \right) \\ &\quad + \ell_{T_1}(a^*, a^*; a) + 2\ell_{T_1}(a^*, a'; a) + \Big(\deg_{T_0}(y) - 1\Big) \ell_{T_1}(a^*, y; a) \end{aligned}$$

$$\begin{aligned} & H_{T_1}(a^*,a) - H_{T_0}(a',a) \\ = \; & \big(\ell_{T_1}(a^*,a^*;a) + 2\ell_{T_1}(a^*,a';a) + \ell_{T_1}(a^*,y;a)\big) \\ & - \big(\ell_{T_0}(a',a';a) + \ell_{T_0}(a',x;a) + 2\ell_{T_0}(a',y;a)\big) \\ = \; & \ell_{T_1}(a^*,a^*;a) + \ell_{T_1}(a^*,a';a) - \ell_{T_0}(a',y;a) - \ell_{T_0}(a',x;a) \end{aligned}$$

Now, to calculate $H_{T_1}(\pi^{(1)},a) - H_{T_0}(\pi^{(0)},a)$, for i=0,1, consider the components of $T_i - z$. Let the one containing y be $V_y^{(i)}$, the one containing a' be $V_{a'}^{(i)}$, and the rest of the components collectively $(V_y^{(i)} \cup V_{a'}^{(i)})^c$. We have

$$\begin{aligned} H_{T_i}(\pi^{(i)},a) &= \sum_{k \in V(T_i)} \pi_k^{(i)} H_{T_i}(k,a) \\ &= \left(\sum_{k \in (V_y^{(i)} \cup V_{a'}^{(i)})^c} \pi_k^{(i)} H_{T_i}(k,z) + H_{T_i}(z,a) \sum_{k \in (V_y^{(i)} \cup V_{a'}^{(i)})^c} \pi_k^{(i)} \right) \\ &+ \sum_{k \in V_y^{(i)} \cup V_{a'}^{(i)}} \pi_k^{(i)} H_{T_i}(k,a) \end{aligned}$$

Define $\Delta(i,j) = H_{T_1}(i,j) - H_{T_0}(i,j)$. Clearly for all vertices c in $(V_y^{(i)} \cup V_{a'}^{(i)})^c$, $\Delta(c,z) = 0$, $\deg_{T_1}(c) = \deg_{T_0}(c)$. For all vertices $c \in V_y^{(0)} \cap V_y^{(1)}$, $\Delta(c,z) = -(\ell(c,y,z) + \ell(c,x,z)) = -2\ell(c,y,z)$. For all vertices $c \in V_{a'}^{(0)} \cap V_{a'}^{(1)}$, $c \neq a^*$, $\Delta(c,z) = \ell(c,a',z) + \ell(c,a^*,z) = 2\ell(c,a',z)$. Then,

$$
\begin{aligned}
& H_{T_1}(\pi^{(1)}, a) - H_{T_0}(\pi^{(0)}, a) \\
= & \left(\sum_{k \in (V_y^{(1)} \cup V_{a'}^{(1)})^c} \pi_k^{(1)} H_{T_1}(k,z) + H_{T_1}(z,a) \sum_{k \in (V_y^{(1)} \cup V_{a'}^{(1)})^c} \pi_k^{(1)} \right) \\
& - \left(\sum_{k \in (V_y^{(0)} \cup V_{a'}^{(0)})^c} \pi_k^{(0)} H_{T_0}(k,z) + H_{T_0}(z,a) \sum_{k \in (V_y^{(0)} \cup V_{a'}^{(0)})^c} \pi_k^{(0)} \right) \\
= & \sum_{k \in (V_y^{(1)} \cup V_{a'}^{(1)})^c} \pi_k^{(1)} H_{T_1}(k,z) + \sum_{k \in (V_y^{(0)} \cup V_{a'}^{(0)})^c} \pi_k^{(0)} H_{T_0}(k,z) \\
= & \frac{\deg(a^*)}{2|E|} H_{T_1}(a^*, z) - \frac{\deg(x)}{2|E|} H_{T_0}(x,z) \\
& + \sum_{k \in V_y^{(1)}} \frac{\deg(k)}{2|E|} \Delta(k,z) + \sum_{k \in V_{a'}^{(1)}, k \neq a^*} \frac{\deg(k)}{2|E|} \Delta(k,z) \\
= & \frac{1}{2|E|} \left(H_{T_1}(a^*, z) - H_{T_0}(x,z) \right) \\
& + \frac{1}{2|E|} \left(2 \sum_{k \in V_{a'}^{(1)}, k \neq a^*} \deg(k) \ell(k, a', z) - 2 \sum_{k \in V_y^{(1)}} \deg(k) \ell(k, y, z) \right) \\
= & \frac{1}{2|E|} \left(H_{T_1}(a^*, z) - H_{T_0}(x,z) + 2H_{T_1}(a', z) - 2H_{T_0}(y,z) \right)
\end{aligned}
$$

Combining the above results,

$$
\begin{aligned}
& H_{T_1}(a, \pi^{(1)}) - H_{T_0}(a, \pi^{(0)}) \\
= & \left(\ell_{T_1}(a^*, a^*; a) + \ell_{T_1}(a^*, a'; a) - \ell_{T_0}(a', y; a) - \ell_{T_0}(a', x; a) \right) \\
& - \left(\frac{1}{2|E|} (H_{T_1}(a^*, z) - H_{T_0}(x,z) + 2H_{T_1}(a', z) - 2H_{T_0}(y,z)) \right) \\
= & (k + (k-1)) \\
& - \left(\frac{1}{2|E|} \left(k^2 + 2(k-1)^2 - H_{T_0}(x,z) - 2H_{T_0}(y,z) \right) \right) \\
> & (2k-1) - \frac{1}{2|E|} (3k^2 - 4k + 2)
\end{aligned}
$$

Then we have $H_{T_1}(a, \pi) - H_{T_0}(a, \pi) > 0$ if and only if $|E| > \frac{3}{4}k - \frac{5}{8} + \frac{3}{16k-8}$, which holds since $|E| > k+2$ and $k \geq 2$. Therefore $H_{T_1}(a, \pi^{(1)}) >$

$H_{T_0}(a, \pi^{(0)})$, which implies $\max_{i \in T_1} H_{T_1}(\pi^{(1)}, i) > \max_{i \in T_0} H_{T_0}(\pi^{(0)}, i)$, a contradiction.

Therefore the only tree on n vertices that achieves $\max_T \max_{i \in T} H_T(i, \pi)$ is $T = P_n$. ▮

5.2 Results on T_{bestmix}

Recall from section 2.5,

$$T_{\text{bestmix}} = \min_{i \in V} H(i, \pi)$$

We can easily show that the minimizing structure for the best mixing time is the star.

Theorem 27 *Of all trees T on n vertices*

$$\min_T T_{\text{bestmix}} = \min_T \min_{i \in T} H_T(i, \pi) = \frac{1}{2}.$$

This value is achieved uniquely by the star S_n.

Proof. First note that for S_n, the vertex that minimizes $H_{S_n}(i, \pi)$ is the center vertex c. Then,

$$H_{S_n}(c, \pi) = H_{S_n}(c', c) - H_{S_n}(\pi, c) = 1 - \frac{1}{2} = \frac{1}{2}.$$

Consider any given tree $T_0 \neq S_n$. Suppose $H_{T_0}(a, \pi) = \min_{i \in T_0} H_{T_0}(i, \pi)$. Since $T_0 \neq S_n$, $\pi_a < 1/2$. Using the threshold rule (see Section 2.4.2), starting at vertex a on T_0, we stay at a with probability π_a, and walk with probability $1 - \pi_a$. Since the threshold rule is optimal, $H_{T_0}(a, \pi) \geq 1 - \pi_a > \frac{1}{2}$.

Therefore the only tree on n vertices that achieves $\min_T \min_{i \in T} H_T(i, \pi)$ is $T = S_n$. ▮

We have a partial result for $\max_T T_{\text{bestmix}} = \max_T \min_{i \in T} H_T(i, \pi)$.

Lemma 28 *If T achieves* $\max_T \min_{i \in T} H_T(i, \pi)$, *then T must be a caterpillar.*

Proof. Assume for the sake of contradiction that there exists some tree T_0, where T_0 is not a caterpillar, such that
$\min_{i \in T_0} H_{T_0}(i, \pi) = \max_T \min_{i \in T} H_T(i, \pi)$. Let a primary focus of T_0 be a. Then the two vertices on T_0 that achieve $\max_{i \in T_0} H_{T_0}(i, j)$ are a-pessimal

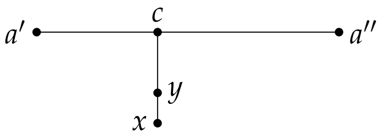

Figure 5.3: T_0

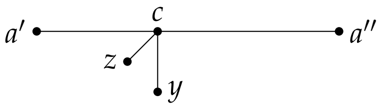

Figure 5.4: T_1

node a' and a'-pessimal node a''. Since T_0 is not a caterpillar, there must exist some leaf x at least distance 2 away from the (a', a'') path. Let y be adjacent to x. Let the first vertex that the path from a' to x and the path from a'' to x share be c.

On T_0 (Figure 5.3), we remove x, and add a new leaf z adjacent to c. Let the resulting tree be T_1 (Figure 5.4). Since the path from a' to a stays unchanged, a is still a primary focus, and we have $H_{T_0}(a', a) = H_{T_1}(a', a)$. We also know $\deg_{T_1}(c) = \deg_{T_0}(c) + 1$, $\deg_{T_1}(y) = \deg_{T_0}(y) - 1$, $\deg_{T_1}(z) = \deg_{T_0}(x) = 0$. Consider the components of $T_1 - c$. Let the one containing y be V_y, and the rest of the components except for z collectively V_c. Define $\Delta(i, j) = H_{T_1}(i, j) - H_{T_0}(i, j)$. Clearly for all vertices w in $(V_y \cup z)^c \in T_1$, $\Delta(w, a) = 0$, $\deg_{T_1}(w) = \deg_{T_0}(w)$. For all vertices $w \in V_y$, since x was removed, $\Delta(w, a) < 0$. Then,

$$\begin{aligned}
& H_{T_0}(x,a) - H_{T_1}(z,a) \\
= \;& (H_{T_0}(x,y) + H_{T_0}(y,c) + H_{T_0}(c,a)) - (H_{T_1}(z,c) + H_{T_1}(c,a)) \\
= \;& (H_{T_0}(c,a) - H_{T_1}(c,a)) + (H_{T_0}(y,c) - H_{T_1}(z,c)) + H_{T_0}(x,y) \\
> \;& 0
\end{aligned}$$

$$
\begin{aligned}
& H_{T_1}(\pi^{(1)}, a) - H_{T_0}(\pi^{(0)}, a) \\
= \; & \left(\sum_{k \in {V_c}^c} H_{T_1}(k,a) \pi_k^{T_1} + H_{T_1}(z,a) \pi_z^{T_1} + H_{T_1}(y,a) \pi_y^{T_1} + H_{T_1}(c,a) \pi_c^{T_1} \right) \\
& - \left(\sum_{k \in {V_c}^c} H_{T_0}(k,a) \pi_k^{T_0} + H_{T_0}(x,a) \pi_x^{T_0} + H_{T_0}(y,a) \pi_y^{T_0} + H_{T_0}(c,a) \pi_c^{T_0} \right) \\
= \; & \frac{1}{2|E|} \left(\sum_{k \in {V_c}^c} \deg_{T_0}(k) H_{T_1}(k,a) - \sum_{k \in {V_c}^c} \deg_{T_0}(k) H_{T_0}(k,a) \right) \\
& + \frac{1}{2|E|} \left(H_{T_1}(z,a) - H_{T_0}(x,a) \right) + \frac{1}{2|E|} \Big(H_{T_1}(y,a) \deg_{T_1}(y) - H_{T_0}(y,a) \deg_{T_0}(y) \Big) \\
& + \frac{1}{2|E|} \Big(H_{T_1}(c,a) \deg_{T_1}(y) - H_{T_0}(c,a) \deg_{T_0}(y) \Big) \\
= \; & \frac{1}{2|E|} \left(H_{T_1}(z,a) - H_{T_0}(x,a) \right) + \frac{\deg_{T_1}(y)}{2|E|} \left(H_{T_1}(y,a) - H_{T_0}(y,a) \right) \\
& - \frac{1}{2|E|} \left(H_{T_1}(c,a) - H_{T_0}(y,a) \right) + \frac{\deg_{T_0}(c)}{2|E|} \left(H_{T_1}(c,a) - H_{T_0}(c,a) \right)
\end{aligned}
$$

We have,

$$
\begin{aligned}
H_{T_1}(z,a) - H_{T_0}(x,a) &< 0 \\
H_{T_1}(c,a) - H_{T_0}(c,a) &= 0 \\
H_{T_1}(y,a) - H_{T_0}(y,a) &< 0 \\
H_{T_1}(c,a) - H_{T_0}(y,a) &< 0
\end{aligned}
$$

Therefore,

$$
H_{T_1}(\pi, a) - H_{T_0}(\pi, a) < 0
$$

$$
\begin{aligned}
& H_{T_1}(a, \pi) - H_{T_0}(a, \pi) \\
= \; & \big(H_{T_1}(a', a) - H_{T_1}(\pi, a) \big) - \big(H_{T_0}(a', a) - H_{T_0}(\pi, a) \big) \\
= \; & \big(H_{T_1}(a', a) - H_{T_0}(a', a) \big) - \left(H_{T_1}(\pi, a) - H_{T_0}(\pi, a) \right) \\
> \; & 0
\end{aligned}
$$

Which is a contradiction to our assumption. Therefore of all trees T on n vertices, the one that achieves $\max_T \min_{i \in T} H_T(i, \pi)$ must be a caterpillar. ∎

We conjecture that of all trees on n vertices, the unique tree achieving $\max_T \min i \in TH(i, \pi)$ must be

1. the tree P_n (Figure 5.5), if n is even;
2. the resulting graph (Figure 5.6) of attaching a leaf to a focus of P_{n-1}, if n is odd.

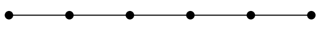

Figure 5.5: n is even

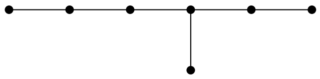

Figure 5.6: n is odd

5.3 Results on T_{forget}

Lastly we look at the maximizing and minimizing structures for forget times. Note that since $T_{\text{forget}} = T_{\text{reset}}$, the same results apply to reset times.

Theorem 29 *Of all trees T on n vertices*

$$\min_T T_{\text{forget}} = 1$$

This value is achieved uniquely by the star S_n.

Proof. From Theorem 7, on any given focal tree T with focus a, we have,

$$T_{\text{forget}} = H(a', a) \geq 1$$

where the equality holds if and only if $T = S_n$.

If T is bifocal with foci a and b,

$$T_{\text{forget}} = H(a', \mu) = \pi_b H(a', b) + \pi_a H(a', a) > 1.$$

Therefore $\min_T T_{\text{forget}} = 1$ is achieved uniquely by the star S_n. ∎

Theorem 30 *Of all trees T on n vertices*

$$\max_T T_{\text{forget}} = \begin{cases} \frac{n^2-2n+2}{4} & \text{if } n \text{ is even} \\ \frac{n^2-2n+1}{4} & \text{if } n \text{ is odd} \end{cases}$$

This value is achieved uniquely by the path P_n.

Proof. From Theorem 7, given a tree T with a single focus a,

$$T_{\text{forget}} = H(a', a)$$

From Theorem 12, of all focal trees, this value is maximized uniquely by the odd path P_n

$$H_{P_n}(a', a) = \frac{n^2 - 2n + 1}{4}.$$

If a tree has two foci a and b,

$$\begin{aligned} T_{\text{forget}} &= H(a', \mu) = H(b', \mu) \\ &= \frac{1}{2|E|}\left(H(a,b)H(b,a) + H(a,b)H(a',b) + H(b,a)H(b',a)\right). \end{aligned}$$

We claim that of all bifocal trees, T_{forget} is maximized uniquely by the even path P_n.

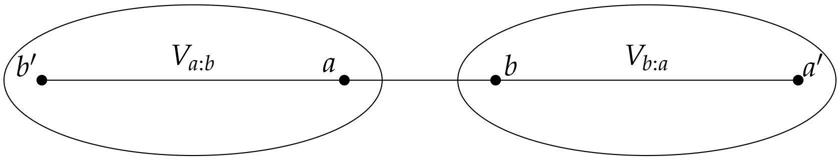

Figure 5.7: T

On any given bifocal tree T in Figure 5.7, let a and b be the two foci, and let $v(V_{a:b}) = k$, $v(V_{b:a}) = l$, where $k + l = n$. We have

$$\begin{aligned} H_T(a,b) &= 2k - 1 \\ H_T(b,a) &= 2l - 1. \end{aligned}$$

If we consider the subgraphs induced by $V_{a:b}$ and $V_{b:a}$, respectively, from Theorem 10, we have

$$\begin{aligned} H_T(b', a) &\leq (k-1)^2 \\ H_T(a', b) &\leq (l-1)^2 \end{aligned}$$

and the equality holds if and only if $V_{a:b}$ and $V_{b:a}$ are paths.

Then,

$$\begin{aligned} T_{\text{forget}} &= \frac{1}{2|E|}\big(H(a,b)H(b,a)+H(a,b)H(a',b)+H(b,a)H(b',a)\big) \\ &\leq (2k-1)(2l-1)+(2k-1)(l-1)^2+(2l-1)(k-1)^2 \\ &= 2k^2l+2l^2k-4kl-k^2-l^2+2k+2l-1 \end{aligned}$$

If we substitute $l = n-k$, and then take the derivate of the above against k, the above is maximized if and only if

$$k = \frac{n}{2}$$

Therefore the even path P_n maximizes T_{forget} of all bifocal trees.

$$T_{\text{forget}} = H_{P_n}(a',\mu) = \frac{n^2-2n+2}{4}$$

Since $\frac{n^2-2n+2}{4} > \frac{n^2-2n+1}{4}$, when n is even, T_{forget} will be maximized uniquely by the even path P_n of all focal and bifocal trees. When n is odd, the odd path P_n would uniquely maximize T_{forget} of all focal trees. We claim that it would uniquely maximize T_{forget} of all focal and bifocal trees.

Assume for the sake of contradiction that there exists a bifocal tree T_0 on n vertices, where n is odd, such that its T_{forget} is higher than that of the odd path P_n. Let the two foci of T_0 be a and b. Then T_0 must satisfy

$$T_{\text{forget}} \geq \frac{n^2-2n+1}{4}$$

Let the number of vertices on the (b',a) path be k, and that on the (a',b) path be l. Since $T_0 \neq P_n$, and T_0 is bifocal, at least one of k and l must be less or equal to $\frac{(n-1)}{2}$. Without loss of generality, let $l \leq \frac{(n-1)}{2}$. Then we have

$$\begin{aligned} T_{\text{forget}} &= H_{T_0}(a',\mu) = \pi_b H(a',b) + \pi_a H(a',a) \\ &< H(a',a) \leq \frac{(n-1)^2}{2} = \frac{n^2-2n+1}{4} \end{aligned}$$

which is a contradiction. Therefore the odd path would uniquely maximize T_{forget} of all focal and bifocal trees when n is odd. Thus the path would uniquely maximize T_{forget} for all trees on any n number of vertices. ▮

Bibliography

[1] J.R. Baxter and R.V. Chacon. Stopping Times for Recurrent Markov processes. *Illinois J. Math* 20 (1976), 467-475.

[2] A. Beveridge. Centers for Random Walks on Trees. *SIAM J. Discrete Math* 23 (2009), 1: 300-318.

[3] A. Beveridge and L. Lovász. Random Walks and the Regeneration Time. submitted to *Journal of Graph Theory*.

[4] P.G. Doyle and J.L. Snell. *Random Walks and Electric Networks*. The Mathematical Association of America (1984).

[5] O. Häggström. Finite Markov Chains and Algorithmic Applications. *Cambridge University Press*, 2002.

[6] L. Lovász. Random Walks on Graphs: a Survey. *Combinatorics, Paul Erdős is Eighty*. Vol II (eds. D. Miklós, V.T. Sós and T. Szőnyi) J. Bolyai Math. Soc. (1996), 355-397.

[7] L. Lovász and P. Winkler. Efficient Stopping Rules for Markov Chains. *Proceedings of the 27th ACM Symposim on the Theory of Computing*, (1995) 76-82.

[8] L. Lovász and P. Winkler. Mixing of Random Walks and Other Diffusions on a Graph. *Surveys in Combinatorics*, (ed. P. Rowlinson), London Math. Soc. Lecture Note Series 218, Cambridge University Press, (1995), 119-154.

[9] L. Lovász and P. Winkler. Reversal of Markov Chains and the Forget Time. *Combin., Probab. Comput.* 7 (1998), 189-204.

[10] J.W. Pitman. Occupation Measures for Markov Chains. *Adv. Appl. Prob.* 9 (1977), 69-86.