CENTERS FOR RANDOM WALKS ON TREES

ANDREW BEVERIDGE*

Abstract. We consider two distinct centers which arise in measuring how quickly a random walk on a tree mixes. Lovász and Winkler [9] point out that stopping rules which "look where they are going" (rather than simply walking a fixed number of steps) can achieve a desired distribution exactly and efficiently. Considering an optimal stopping rule that reflects some aspect of mixing, we can use the expected length of this rule as a mixing measure. On trees, a number of these mixing measures identify particular nodes with central properties. In this context, we study a variety of natural notions of centrality. Each of these criteria identifies the barycenter of the tree as the "average" center and the newly defined focus as the "extremal" center.

Key words. Markov chain, random walk, stopping rule, tree, barycenter

AMS subject classifications. 60510, 60G40, 05C05

1. Introduction. What node is most central with respect to a random walk on a tree $G = (V, E)$? We define $P = \{p_{ij}\}$ to be the matrix of transition probabilities, so $p_{ij} = 1/d(i)$ if $ij \in E$ and $p_{ij} = 0$ otherwise. Let the hitting time $H(i, j)$ be the expected time for a random walk starting at node i to get to node j. A natural definition for centrality is to require that the target node j minimize this hitting time for an appropriately chosen starting node i. We consider two natural choices for this starting node. First, we identify the "average" center c by drawing i from the stationary distribution π:

$$\sum_{i \in V} \pi_i H(i, c) = \min_{j \in V} \sum_{i \in V} \pi_i H(i, j). \tag{1.1}$$

Next we choose the worst possible starting node for each target j. Let j' be a j-*pessimal* node satisfying $H(j', j) = \max_{i \in V} H(i, j)$. A target node a achieving

$$H(a', a) = \min_{j \in V} H(j', j) = \min_{j \in V} \max_{i \in V} H(i, j) \tag{1.2}$$

is the "extremal" center of the tree.

There are two classical centers for trees. A node achieving $\min_{i \in V} \max_{j \in V} d(i, j)$, where $d(i, j)$ is the length of the unique path between i and j, is the *center* of the tree G (or bicenter if there are two adjacent nodes achieving this minimum). In other words, the distance to the furthest node from the center is minimal among all nodes of the tree G. This node does not appear to have any central properties with respect to random walks.

The *barycenter* is the node (or two adjacent nodes) achieving $\min_{i \in V} \sum_{j \in V} d(i, j)$. A barycenter minimizes the total distance to all other nodes. The propostion below reveals that the barycenter is the "average" center of the tree with respect to random walks. If u, v are adjacent nodes, let $V_{u:v}$ be the set of nodes in the subtree rooted at u after the removal of the edge uv. This notation is meant to emphasize that nodes in $V_{u:v}$ are closer to u than to v. For any $S \subset V$ and any distribution τ, let $\tau(S) = \sum_{k \in S} \tau_k$. For an undirected graph, Coppersmith, Tetali and Winkler [5]

*Department of Mathematics and Computer Science, Macalester College, Saint Paul, MN 55105 (abeverid@macalester.edu). Supported in part by NSA Young Investigator Grant H98230-08-1-0064.

define the *central node* to be a node c for which $H(i,c) \leq H(c,i)$ for every node i. The following equivalence has been known to the authors of [5].

PROPOSITION 1.1. *Let $G = (V, E)$ be a tree. The following statements for a node c are equivalent.*

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

(b) The node c satisfies $H(i,c) \leq H(c,i)$ for every node i.

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

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

We introduce the term *focus* to denote the "extremal" center of the tree. There are two types: primary foci and secondary foci.

DEFINITION. Let G be a tree. 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. When all of the a-pessimal nodes are contained in a single subtree $G' \subset G \backslash \{a\}$, the unique a-neighbor b in G' is also a focus of G. If $H(b', b) = H(a', a)$ then b is a primary focus. If $H(b', b) > H(a', a)$ then b is a *secondary focus* of G.

PROPOSITION 1.2. *Every tree G either has one focus or has two adjacent foci.*

When G has a single focus, we say that G is *focal*. When G has two adjacent foci, we say that G is *bifocal*. A bifocal tree may have two primary foci or it may have one primary focus and one secondary focus. For a bifocal tree G with adjacent foci a, b we will see that a has central properties for nodes in $V_{a:b}$ and that b has central properties for nodes in $V_{b:a}$.

We consider some examples of foci for trees. Let P_n denote the path of length n on vertices $v_0, v_1, \ldots, v_n$. It is well known that

$$H(v_i, v_j) = \begin{cases} j^2 - i^2 & i \leq j \\ (n-j)^2 - (n-i)^2 & i > j. \end{cases} \tag{1.3}$$

A v_i-pessimal node must be the furthest leaf from v_i. The node that achieves $\min_{v_i \in P_n} H(v_i', v_i)$ is the center of the path, so the unique focus of P_{2k} is v_k and the foci of P_{2k+1} are v_k and v_{k+1}. In the latter case, both nodes are primary foci since $H(v_k', v_k) = H(v_{2k+1}, v_k) = (k+1)^2 = H(v_0, v_{k+1}) = H(v_{k+1}', v_{k+1})$.

Let $B_{r,s}$ denote the broom graph consisting of a star with r leaves $u_1, u_2, \ldots, u_r$ centered at node c, along with a path of length s on nodes $c = v_0, v_1, \ldots, v_s$. Simple calculations (using formula (2.6) in section 2) for $B_{4k,4k}$ show that node v_k is the primary focus with $H(v_k', v_k) = H(u_1, v_{k-1}) = 9k^2 + 1$ and that v_{k-1} is the secondary focus with $H(v_{k-1}', v_{k-1}) = H(v_{4k}, v_{k-1}) = (3k+1)^2$. Moreover, this broom graph shows that center, barycenter and foci of the tree are distinct notions. Indeed, additional calculations show that the nodes v_{2k}, v_{2k+1} are centers of $B_{4k,4k}$ and the barycenter is $c = v_0$. These three types of centers are pairwise separated by distance $\Theta(k)$.

Having defined the average and extremal centers, we consider a variety of criteria for centrality with respect to random walks on trees. The barycenter satisfies each "average" criterion and one of the foci of the tree (or both) satisfies each "extremal" criterion. Many of these criteria concern exact mixing measures defined via lengths of stopping rules.

Given an initial node i and a target distribution τ, we can follow an optimal stopping rule (see section 2 for a precise definition) to halt a random walk starting at i so that the distribution of the final node is exactly τ. Denote the expected length of

this optimal rule by $H(i,\tau)$. A number of parameterless mixing measures defined via stopping rules have been introduced and studied in [1], [2], [3], [9], [10]. Among the most important measures are the mixing time $T_{\text{mix}} = \max_{i\in V} H(i,\pi)$ and the reset time $T_{\text{reset}} = \sum_{i\in V} \pi_i H(i,\pi)$. We interpret T_{mix} as the pessimal mixing time and T_{reset} as the average mixing time.

Since the barycenter is so closely related to average mixing, a natural question is how $H(c,\pi)$ compares with T_{reset}.

PROPOSITION 1.3. *$H(c,\pi) \le 2T_{\text{reset}}$ where c is a barycenter of the tree. This bound is tight for a star $K_{1,k}$ with $k \ge 2$.*

$H(c,\pi)$ may be considerably smaller than T_{reset}. Indeed, consider a rooted m-ary tree of depth r. By symmetry its root c is the unique barycenter and the unique focus.

THEOREM 1.4. *If G is an m-ary tree of depth r rooted at node c then*

$$H(c,\pi) = \frac{(m+1)(m^r+1)}{(m-1)(m^r-1)}\, r - \frac{m^2+6m+1}{2(m-1)^2} \tag{1.4}$$

and

$$T_{\text{reset}} = \frac{2m^{r+1} - rm^2 - 2m + r}{(m-1)^2}. \tag{1.5}$$

Holding m fixed and letting $r \to \infty$, mixing from the root takes $\Theta(r)$ steps while average mixing takes $\Theta(m^{r-1})$ steps. The exact mixing result of equation (1.4) is complementary to the following approximate mixing result for m-ary trees due to Diaconis and Fill (Example 4.60 in [6]). If σ_t is the distribution achieved by walking t steps from the root c then for fixed m, as $r \to \infty$, the total variational distance $\|\sigma_t - \pi\|$ becomes small after $\frac{m+1}{m-1} r + \alpha r^{1/2}$ steps for a large constant α.

We now turn our attention to the extremal center. Our analysis of the foci relies heavily on stopping rules. We provide a definition for the foci of a distribution τ (which is actually a generalization of the definition of the foci of a tree; see proposition 1.6 below). For any target distribution τ we associate one node or two adjacent nodes that are central with respect to stopping rules from singleton distributions to τ.

DEFINITION. A node i is a *focus* of the distribution τ when $H(i,\tau) < 1 + \sum_{j\in V} p_{ij} H(j,\tau)$.

The left hand side of this equation is the the expected length of an optimal rule. The right hand side of the equation is the expected length of the rule "take one step from i (according to the transition probabilities $P = \{p_{ij}\}$) and then follow an optimal rule starting from this neighbor node to τ." The τ-foci are the nodes for which this composite rule is not optimal.

For example, consider the path P_2 on nodes v_0, v_1, v_2 and take our target to be $\pi = (1/4, 1/2, 1/4)$. We first consider an optimal stopping rule from the center to π. Let $\Gamma_1(v_1,\pi)$ be the rule "with probability $1/2$ take one step to a random neighbor, otherwise stay put." The expected length $E(\Gamma_1(v_1,\pi)) = 1/2$ and equation (2.8) in the next section shows that this rule is an optimal rule from v_1 to π. Now consider the rule $\Gamma_2(v_1,\pi)$; "take one step and then follow an optimal rule to π." This rule is not optimal since $E(\Gamma_2(v_1,\pi)) > 1 > 1/2$.

Finally consider the rule $\Gamma(v_0,\pi)$ from v_0 to π: "take one step and then follow the optimal rule $\Gamma_1(v_1,\pi)$." This turns out to be an optimal rule with $E(\Gamma(v_0,\pi)) = 3/2$. The analogous rule from v_2 to π is also optimal. Therefore according to the definition above, the center v_1 is the unique focus for P_2. An analogous argument shows that

the two internal nodes of P_4 are the foci for π. A similar phenomenon holds for an arbitrary tree.

THEOREM 1.5. *Every distribution τ on a tree has either one focus or two adjacent foci. If τ has a unique focus u then $H(i,\tau) = H(i,u) + H(u,\tau)$ for all i. If τ has two foci u, v then for $i \in V_{u:v}$, $H(i,\tau) = H(i,u) + H(u,\tau)$ and for $i \in V_{v:u}$, $H(i,\tau) = H(i,v) + H(v,\tau)$.*

The key observation of this theorem is that for any node i, the rule "walk from i to the nearest τ-focus and then follow an optimal rule from that focus to τ" is an optimal rule from i to τ. In other words, the foci of τ are central with respect to all walks from nodes to τ.

Naturally, the foci of the tree coincide with the foci of π.

PROPOSITION 1.6. *The foci of the tree G are the foci of the distribution π.*

This proposition shows that the focus of a distribution is indeed a generalization of the focus of the tree.

Another important mixing measure is the forget time $T_{\text{forget}} = \min_\tau \max_{i \in V} H(i,\tau)$ where the minimum is taken over all target distributions. We interpret this quantity as the minimum expected time to "forget" the node we started from by following an optimal rule to some distribution. In spite of its rather unorthodox definition, the forget time is intimately connected to the mixing time and the reset time. For an undirected graph, we have the nontrivial equality $T_{\text{reset}} = T_{\text{forget}}$ (see [10]) and they are within a factor of 4 of T_{mix} (see [2]).

For any graph, Lovász and Winkler [10] show that there is a unique distribution μ achieving the forget time. This distribution μ is central in an extremal sense: μ minimizes the expected length of a rule starting from the worst possible node. For a tree, μ is concentrated on the foci of G:

PROPOSITION 1.7. *If the node a is the unique focus of $G = (V,E)$ then μ is the singleton distribution on the focus a. If the adjacent nodes a and b are the foci of G then*

$$\mu_i = \begin{cases} \frac{1}{2|E|}(H(b',b) - H(a',b)), & i = a \\ \frac{1}{2|E|}(H(a',a) - H(b',a)), & i = b \\ 0, & \textit{otherwise.} \end{cases}$$

where $H(i',i) = \max_{j \in V} H(j,i)$.

Another mixing measure with central properties is $T_{\text{bestmix}} = \min_{i \in V} H(i,\pi)$. The node achieving T_{bestmix} is the best possible starting node for achieving the stationary distribution. This formulation is dual in some sense to that of equation (1.1). As expected, the foci of the tree are central for this extremal problem.

THEOREM 1.8. *The quantity $T_{\text{bestmix}} = \min_{i \in V} H(i,\pi)$ is achieved by a focus of the tree. Specifically, 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} and if $H(a',b) = H(b',a)$ then T_{bestmix} is achieved by both a and b.*

The broom graphs $B_{2,2}$, $B_{2,3}$ and $B_{4,3}$ in figure 1 show that all three possibilities do occur.

Consider another mixing measure similar to the forget time. The start-independent time of a distribution σ is $T_{\text{si}}(\sigma) = \min_\tau \sum_{i \in V} \sigma_i H(i,\tau)$ where the minimum is taken over all target distributions. For a walk started from σ, $T_{\text{si}}(\sigma)$ is the minimum expected time to obtain a sample (from some distribution) that is independent of the initial node of the walk (which was drawn from σ). We may interpret $T_{\text{si}}(\sigma)$ as the fastest way to "forget" that we started our walk from a node drawn from σ. A natural

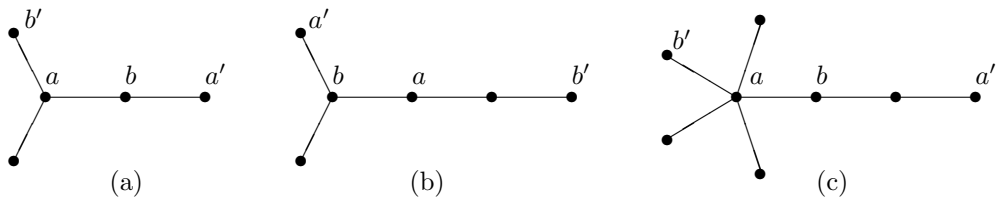

FIG. 1.1. *Trees with different foci achieving* T_{bestmix}. *(a)* $H(a',b) = H(b',a)$ *so* $T_{\text{bestmix}} = H(a,\pi) = H(b,\pi)$. *(b)* $H(a',b) < H(b',a)$ *so* $T_{\text{bestmix}} = H(a,\pi)$. *(c)* $H(a',b) > H(b',a)$ *so* $T_{\text{bestmix}} = H(b,\pi)$.

choice for our initial distribution is the stationary distribution. For a tree, the target distribution achieving $T_{\text{si}}(\pi)$ is central in an average sense and indeed the optimal target is concentrated on the barycenter.

PROPOSITION 1.9. *If c is a barycenter of the tree G then*

$$T_{\text{si}}(\pi) = \min_{\tau} \sum_{k \in V} \pi_k H(k,\tau) = \sum_{k \in V} \pi_k H(k,c) = H(\pi,c).$$

We define the start-independent time of a graph to be $T_{\text{si}} = \max_\sigma T_{\text{si}}(\sigma)$ where the maximum is taken over all initial distributions. For a tree, the target distribution achieving T_{si} is central in an extremal sense, and indeed T_{si} can be achieved by taking either focus as a target. Furthermore, $T_{\text{si}} = T_{\text{forget}}$.

THEOREM 1.10. *For a tree G, we have $T_{\text{si}} = T_{\text{forget}}$. Moreover, if G has a unique focus a, then there exists a distribution ϕ such that $T_{\text{si}} = H(\phi,a)$. If G has two foci a and b then there exists a distribution ϕ such that $T_{\text{si}} = H(\phi,a) = H(\phi,b)$.*

It is an open question how $T_{\text{si}}(\pi)$ and T_{si} compare to the other mixing measures for general graphs.

2. Preliminaries.

2.1. Random walks. Given an undirected, connected graph $G = (V,E)$ a random walk on G is a sequence of nodes $(w_0, w_1, \ldots, w_t, \ldots)$ such that the node w_t at time t is chosen uniformly from the neighbors of w_{t-1}. For non-bipartite G, as t tends to infinity the distribution of the node w_t tends to the so called stationary distribution π where $\pi_i = d(i)/2|E|$ and $d(i)$ is the degree of i. For bipartite G, we have convergence if we consider a "lazy walk" in which at each step we stay at the current node with probability 1/2. For simplicity of exposition we will consider non-lazy walks (laziness simply doubles the expected length of our walks).

For two nodes i,j, the *hitting time* $H(i,j)$ is the expected length of a walk from i to j. The expected number of steps before a walk started at i returns to i is

$$\text{Ret}(i) = \frac{1}{\pi_i}. \tag{2.1}$$

For any undirected graph G, Coppersmith, Tetali and Winkler [5] define a *central node* c of G to be a node which satisfies $H(i,c) \le H(c,i)$ for all i. The existence of such a node for an undirected graph follows from their cycle reversing identity

$$H(i,j) + H(j,k) + H(k,i) = H(i,k) + H(k,j) + H(j,i). \tag{2.2}$$

The "random target identity" (c.f. [8], equation (3.3)) states that $\sum_{j\in V}\pi_j H(i,j)$ is independent of the initial node i. Multiplying equation (2.2) by π_k, summing over all k and using this identity gives

$$\sum_{k\in V}\pi_k H(k,i)+H(i,j)=\sum_{k\in V}\pi_k H(k,j)+H(j,i) \tag{2.3}$$

for any nodes i,j.

2.2. Hitting times for trees. We give a hitting time formula for trees that is equivalent to those found in [4], [7] and [11]. For adjacent nodes i and j,

$$H(i,j)=\sum_{k\in V_{i:j}}d(k)=2|E|\sum_{k\in V_{i:j}}\pi_k=2|E|\,\pi(V_{i:j}). \tag{2.4}$$

Indeed, let G' be the induced subtree on the nodes $V_{i:j}\cup\{j\}$. Let $d'(k)$ be the G'-degree of k and $H_{G'}(i,j)$ be the hitting time from i to j for this graph. By equation (2.1), $H(i,j)=H_{G'}(i,j)=\mathrm{Ret}_{G'}(j)-1=\sum_{k\in G'}d'(k)-1=\sum_{k\in V_{i:j}}d(k)$.

If i and j are neighbors then equation (2.4) immediately gives

$$H(i,j)+H(j,i)=2|E|. \tag{2.5}$$

Furthermore, we can determine a hitting time formula for the general case. Define $\ell(i,k;j)=\frac{1}{2}(d(i,j)+d(k,j)-d(i,k))$, the length of the intersection of the (i,j)-path and the (k,j)-path. This function is symmetric in i and k and is zero if and only if $i=j$, $k=j$, or the nodes i and k are in different connected components of $G\setminus\{j\}$. Assume $d(i,j)=r$ and the (i,j)-path is given by $(i=i_0,i_1,i_2,\ldots,i_r=j)$. Using equation (2.4) and $\ell(i,k;j)$ yields

$$H(i,j)=\sum_{t=0}^{r-1}H(i_t,i_{t+1})=\sum_{k\in V}\ell(i,k;j)d(k). \tag{2.6}$$

Indeed, we can use formula (2.6) to recover formula (1.3) for hitting times on the path.

2.3. Stopping rules. We briefly summarize some results of Lovász and Winkler [9]. Let V^* be the space of finite walks on V, i.e. the set of finite strings $w=(w_0,w_1,w_2,\ldots,w_t)$, $w_i\in V$ 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

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

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 [13], is the "conservation equation"

$$\sum_{i \in V} p_{ij} x_i(\Gamma) - x_j(\Gamma) = \tau_j - \sigma_j. \tag{2.7}$$

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,

$$\Gamma \text{ is an optimal stopping rule } \iff \exists k \in V,\ x_k(\Gamma) = 0. \tag{2.8}$$

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 (2.7). (See [9] 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"

$$H(\sigma, \rho) \leq H(\sigma, \tau) + H(\tau, \rho). \tag{2.9}$$

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 (2.9) 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

$$H(\sigma, \tau) = \max_{j \in V}(H(\sigma, j) - H(\tau, j)). \tag{2.10}$$

In the special case $\sigma = i$ and $\tau = \pi$ we have a particularly nice characterization due to the combination of equations (2.3) and (2.10):

$$j \text{ is an } (i, \pi)\text{-halting state} \iff H(j, i) = \max_{k \in V} H(k, i). \tag{2.11}$$

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

$$H(i, \pi) = H(i', i) - H(\pi, i). \tag{2.12}$$

2.3.1. 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 (2.8), a rule is optimal when it has a halting state. Equation (2.11) 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 $H(v_0, \pi) = |\Gamma(v_0, \pi)| = \frac{1}{6}H(v_0, v_0) + \frac{1}{3}H(v_0, v_1) + \frac{1}{3}H(v_0, v_2) + \frac{1}{6}H(v_0, v_3) = 19/6$ by equation (1.3).

We now consider starting at the node v_1. Equation (2.8) 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. Mixing measures. Stopping rules provide a number of parameterless mixing measures. We define the *mixing time* T_{mix} to be the expected length of an optimal mixing rule starting from the worst initial node: $T_{\text{mix}} = \max_{i \in V} H(i, \pi)$. A node achieving this maximum is called *mixing pessimal*. The *forget time* T_{forget} is the smallest t such that there exists a distribution μ such that for every starting node, the expected time to attain μ via an optimal rule is at most t: $T_{\text{forget}} = \min_\tau \max_{i \in V} H(i, \tau)$. Theorem 4 (and the subsequent remark) in [10] establishes that the forget time is attained by a unique distribution given by

$$\mu_i = \pi_i \left(1 + \sum_{j \in V} p_{ij} H(j, \pi) - H(i, \pi) \right). \tag{2.13}$$

Furthermore, if a node is mixing pessimal then it is also pessimal for μ and every mixing pessimal node is a halting state for an optimal rule from μ to π.

The *reset time* $T_{\text{reset}} = \sum_{i \in V} \pi_i H(i, \pi)$ can be viewed as an average mixing time. Theorem 1 in [10] establishes the remarkable equality

$$T_{\text{forget}} = T_{\text{reset}} \tag{2.14}$$

for a random walk on an undirected graph. Moreover, for an undirected graph we have $T_{\text{reset}} \leq T_{\text{mix}} \leq 4T_{\text{reset}}$ (see [2] corollary 8 and its subsequent remarks).

2.5. Start-independence. The following independence condition arises in applications of random walks. Let Γ be a stopping rule from σ to τ and let $w_0, w_1, \ldots, w_T$ be a walk halted by Γ at time T. The *support* of σ, denoted S_σ, is the the set of nodes i such that $\sigma_i > 0$. We associate a conditional distribution $\tau^{(i)}$ to each $i \in S_\sigma$ given by $\tau^{(i)}_k = \Pr\{w_T = k | w_0 = i\}$. In other words, $\tau^{(i)}_k$ is the probability that Γ stops the walk at k given that the walk started at i (which was drawn from σ). Clearly $\sum_{i \in S_\sigma} \sigma_i \tau^{(i)} = \tau$ and we call the set $\{\tau^{(i)}\}_{i \in S_\sigma}$ the Γ-*decomposition* of τ.

The rule Γ is *start-independent* if $\tau^{(i)} = \tau$ for all $i \in S_\sigma$. The node at which a start-independent rule halts is independent of the initial node. Start-independent rules always exist: the rule "draw w_0 from σ and walk optimally from w_0 to τ" is a start-independent rule of expected length $\sum_{i \in V} \sigma_i H(i, \tau)$.

While start-independent rules are rarely optimal (for example take $\sigma = \tau$), they arise naturally in applications requiring multiple independent samples from the stationary distribution of some state space. We obtain these samples by following an optimal mixing rule, accepting the current state, and then starting a new optimal mixing walk from this state. In this setting, T_{reset} is the expected length of a minimal start-independent rule from π to π. (See [3] for an extremal result concerning start-independent rules whose initial and target distributions are identical.)

We define the start-independent time of a distribution σ to be the minimum expected length of a start-independent rule with initial distribution σ:

$$T_{\text{si}}(\sigma) = \min_\tau \sum_{i \in V} \sigma_i H(i, \tau)$$

A quantity of natural interest is $T_{\text{si}}(\pi)$, the start-independent time for the stationary distribution. We would also like to determine the extremal behavior of $T_{\text{si}}(\sigma)$. The start-independent time of any singleton distribution is zero, so only the maximum case is nontrivial. We define the start-independent time of the graph to be

$$T_{\text{si}} = \max_\sigma T_{\text{si}}(\sigma) = \max_\sigma \min_\tau \sum_{i \in V} \sigma_i H(i, \tau).$$

3. The average and extremal centers. We begin with our characterization of the barycenter of the tree.

Proof of proposition 1.1. The equivalence of (b) and (c) follows from equation (2.3): $H(i,c) \le H(c,i)$ for all i if and only if $\sum_{k \in V} \pi_k H(k,c) \le \sum_{k \in V} \pi_k H(k,i)$ for all i.

We show that (c) and (d) are equivalent. Assume $\pi(V_{i:c}) \le 1/2$ for every node i adjacent to c. For $j \in V_{i:c}$, $H(j,c) \le d(j,c) \sum_{k \in V_{i:c}} d(k) \le d(c,j) \sum_{k \in V_{c:j}} d(k) \le H(c,j)$ by equation (2.6), so c is the central node. Now assume that c is the central node and that $\pi(V_{i:c}) > 1/2$. Then $H(i,c) = \sum_{k \in V_{i:c}} d(k) > \sum_{k \in V_{c:i}} d(k) = H(c,i)$, a contradiction. If $\pi(V_{i:c}) = 1/2$ for some neighbor i of c then i is also a central node.

Finally we prove the equivalence of (a) and (d). For any adjacent nodes i and j, we have

$$\sum_{k \in V} d(k,i) = \sum_{k \in V} d(k,j) - |V_{i:j}| + |V_{j:i}| \tag{3.1}$$

and

$$\pi(V_{i:j}) = \sum_{k \in V_{i:j}} \pi_k = \frac{2|V_{i:j}| - 1}{2|E|}. \tag{3.2}$$

The node j has a neighbor i with $\pi(V_{i:j}) > 1/2 > \pi(V_{j:i})$ if and only if $|V_{i:j}| > |V_{j:i}|$ (by equation (3.2)) if and only if $\sum_{k\in V} d(k,i) < \sum_{k\in V} d(k,j)$ (by equation (3.1)) if and only if j is not the barycenter. □

Proof of proposition 1.2. We divide the proof into two cases, depending on whether G has a focus a such that multiple subtrees of $G\backslash\{a\}$ contain an a-pessimal node.

Case 1. Let a' and a'' be a-pessimal nodes contained in different subtrees of $G\backslash\{a\}$. Any node $u \neq a$ is separated by a from at least one of a', a''. Without loss of generality, assume that a is on the unique (a', u)-path. Then $H(u', u) \geq H(a', u) = H(a', a) + H(a, u) > H(a', a)$ and therefore u is not a focus of the tree. Hence a is the unique focus of G.

Case 2. Suppose that a is a focus of G with all a-pessimal nodes in a single component of $G\backslash\{a\}$. Let b be the unique neighbor of a in this component. By definition, b is a focus of G and $H(b', b) \geq H(a', a)$. For any node $u \in V_{a:b}$, $H(u', u) \geq H(a', u) = H(a', a) + H(a, u) > H(a', a)$, so $u \in V_{a:b}$ is not a focus of G. The b-pessimal node b' must lie in $V_{a:b}$. Indeed, for any node $w \in V_{b:a}$, $H(w, b) < H(w, a) \leq H(a', a) \leq H(b', b)$. Similarly $a' \in V_{b:a}$. Now considering $v \in V_{b:a}\backslash\{b\}$ we have $H(v', v) \geq H(b', v) = H(b', b) + H(b, v) > H(b', b) \geq H(a', a)$, so v is not a focus of G. □

The following corollary is immediate from the proof.

COROLLARY 3.1. *If G is focal with unique focus a then there are multiple subtrees of $G\backslash\{a\}$ containing a-pessimal nodes. If G is bifocal with foci a, b then each a-pessimal node is contained in $V_{b:a}$ and each b-pessimal node is contained in $V_{a:b}$.*

The barycenter is the average center for random walks on trees, so it is natural to compare $H(c, \pi)$ and $T_{\text{reset}} = T_{\text{forget}}$. Mixing from the barycenter never takes more than twice as long as the average mixing time.

Proof of proposition 1.3. Let u be the unique neighbor of c on the path from c to a c-pessimal node c'. By proposition 1.1, $\pi(V_{c:u}) \geq 1/2$. For any node $i \in V_{c:u}$ we have $H(i, \pi) = H(i, c) + H(c, \pi)$. Indeed, $c' \in V_{u:c}$ will also be i-pessimal and the composite rule corresponding to the right hand side preserves c' as a halting state. So equation (2.11) guarantees that this rule is optimal. Therefore

$$T_{\text{reset}} \geq \sum_{i\in V_{c:u}} \pi_i H(i,\pi) = \sum_{i\in V_{c:u}} \pi_i \left(H(i,c) + H(c,\pi)\right) \geq \pi(V_{c:u}) H(c,\pi) \geq \frac{1}{2} H(c,\pi).$$

For a star, we have $H(c, \pi) = 1/2$ and $T_{\text{reset}} = T_{\text{forget}} = 1$, so this bound is tight. □

There are trees for which $H(c, \pi) > T_{\text{reset}}$. Consider a broom graph $B_{4k,4k}$ with path nodes $(c = v_0, v_1, \ldots v_{4k})$. Some simple calculations using equation (2.6) show that v_k and v_{k-1} are the foci of $B_{4k,4k}$. The forget time (and hence the reset time by equation (2.14)) is $T_{\text{forget}} = H(v_{4k}, \mu) < H(v_{4k}, v_{k-1}) = (3k + 1)^2$. Using equations (2.6) and (2.12), the expected time to mix from the barycenter is $H(c, \pi) = H(c, c') - H(\pi, c') = (64k^2 - 1)/6$ which is strictly greater than $(3k + 1)^2$ for $k \geq 4$. Of course, $(64k^2 - 1)/6 < 2(3k)^2 = 2H(v_{4k}, v_k) < 2T_{\text{forget}}$ as stipulated by proposition 1.3.

On the other hand, $H(c, \pi)$ may be markedly smaller than the forget time (and hence the reset time) of the tree. Consider an m-ary tree of depth r with root c. Of course c is the center, the barycenter and the focus of this tree. We adopt the following notation: $S_k = \{i | d(i, c) = k\}$ is the set of all nodes at level k. Let $c = i_0, i_1, \ldots, i_r$ be a path from c to a leaf i_r. The expected behavior of the walk at a node only depends on the level of the node, so we may use i_k as a representative for all nodes in S_k. A node is halting for this mixing walk if and only if it lies in S_r. We explicitly calculate

$H(c,\pi) = H(i_r, c) - H(\pi, c)$ as per equation (2.12). Counting the degrees level-wise, the total number of edges in an m-ary tree of depth r is

$$\frac{1}{2}\left(m + (m+1)\sum_{k=1}^{r-1} m^k + m^r\right) = \frac{m(m^r-1)}{m-1}.$$

Proof of theorem 1.4. We start by showing that

$$H(i_r, i_{r-s}) = \frac{2m^{s+1} - sm^2 - 2m + s}{(m-1)^2}. \tag{3.3}$$

Let G' be the connected component of $G\backslash\{i_{r-s}\}$ containing i_r. We partition $V(G') \cup \{i_{r-s}\}$ into sets $T_k = \{j | \ell(i_r, j; i_{r-s}) = k\}$ so that

$$H(i_r, i_{r-s}) = \sum_{k=1}^{s} k \sum_{j \in T_k} d(j)$$

by equation (2.6). We have $T_s = \{i_r\}$ and for $1 \le k \le s-1$, T_k consists of the node i_{r-s+k} connected to $m-1$ copies of m-ary trees of depth $s-k-1$. Hence,

$$\sum_{j \in T_k} d(j) = (m+1) + (m-1)\left(1 + \frac{2m(m^{s-k-1}-1)}{m-1}\right) = 2m^{s-k}$$

for $1 \le k \le s-1$ so that

$$\begin{aligned} H(i_r, i_{r-s}) &= s + \sum_{k=1}^{s-1} k(2m^{s-k}) = s + 2\sum_{j=1}^{s-1}(s-j)m^j \\ &= s + 2s\sum_{j=1}^{s-1} m^j - 2\sum_{j=1}^{s-1} jm^j \;=\; \frac{2m^{s+1} - sm^2 - 2m + s}{(m-1)^2}. \end{aligned}$$

By equation (2.12),

$$\begin{aligned} H(c,\pi) &= H(i_r, c) - H(\pi, c) = H(i_r, c) - \sum_{j \in V} \pi_j H(j, c) = H(i_r, c) - \sum_{k=0}^{r} \pi(S_k) H(i_k, c) \\ &= \sum_{k=0}^{r} \pi(S_k)\left((H(i_r, i_k) + H(i_k, c)) - H(i_k, c)\right) \;=\; \sum_{k=0}^{r} \pi(S_k) H(i_r, i_k) \\ &= \frac{m-1}{2m(m^r-1)}\left(mH(i_r, i_0) + \sum_{k=1}^{r-1} m^k(m+1)H(i_r, i_{r-(r-k)}) + m^r H(i_r, i_r)\right). \end{aligned}$$

Using equation (3.3) and simplifying (we omit the details) yields the mixing result of equation (1.4).

We can now quickly determine the reset time of an m-ary tree. As per equation (2.14), $T_{\text{reset}} = T_{\text{forget}}$. By symmetry, the root is the unique focus of the m-ary tree and $T_{\text{forget}} = H(i_r, i_0)$. Formula (3.3) for $s = r$ gives equation (1.5). For completeness, we note that $T_{\text{mix}} = H(i_r, \pi) = H(i_r, i_0) + H(i_0, \pi) = T_{\text{forget}} + H(c, \pi) = T_{\text{reset}} + H(c, \pi) = \Theta(m^{r-1})$. ∎

4. The foci of a distribution. Recall that the node k is (i,τ)-halting when $x_k(i,\tau) = 0$. Two nodes i, j have a common halting state for τ when there exists a node k such that $x_k(i,\tau) = 0$ and $x_k(j,\tau) = 0$. A focus of a distribution τ on the tree G is a node u for which the rule "take one step from u and then follow an optimal rule from this random neighbor of u to τ" is not optimal, i.e. $H(u,\tau) < 1 + \sum_{i\in V} p_{ui}H(i,\tau)$. This is equivalent to saying that there is no node that is simultaneously τ-halting for u and all of its neighbors.

For example, the focus for the singleton distribution $\tau = u$ is the node u. Considering mixing walks, equation (2.11) states that k is a π-halting state for i if and only if $H(k,i) = \max_{j\in V} H(j,i) = H(i',i)$. Hence for a path of even length the unique center is the only π-focus, and for a path of odd length, the two central nodes are the π-foci. Also, the center of a star graph is the only π-focus.

We now prove Theorem 1.5, which states that for any tree G, every τ has one focus or two adjacent foci. Fixing τ, let i^* denote a halting state for an optimal stopping rule from i to τ.

LEMMA 4.1. *When i^* is an (i,τ)-halting state then i^* is a (j,τ)-halting state whenever j and i^* are in different subtrees of $G\backslash\{i\}$.*

Proof. We are guaranteed that i is on the unique (j,i^*)-path. Consider the composite rule "walk from j until you reach i and then follow an optimal rule from i to τ." The kth exit frequency of this composite rule is $x_k(j,i)+x_k(i,\tau)$. In particular $x_{i*}(j,i) + x_{i^*}(i,\tau) = 0$ so i^* is a halting state. By equation (2.8) this composite rule is therefore optimal: $H(j,\tau) = H(j,i) + H(i,\tau)$. □

LEMMA 4.2. *If $(j_1,\ldots,i_1,i_2,\ldots,j_2)$ is a path in the tree G then the nodes j_1, j_2 cannot each be τ-halting states for both of the nodes i_1, i_2.*

Proof. Assume the conclusion is false. Equation (2.10) yields $H(i_k,\tau) = H(i_k,j_1) - H(\tau,j_1) = H(i_k,j_2) - H(\tau,j_2)$ for $k = 1,2$ and therefore $-H(i_2,i_1) = H(i_1,j_1) - H(i_2,j_1) = H(i_1,\tau) - H(i_2,\tau) = H(i_1,j_2) - H(i_2,j_2) = H(i_1,i_2)$, a contradiction. □

Proof of theorem 1.5. The case when τ is a singleton is trivial, so assume τ is not a singleton.

Case 1: There exists an edge uv such that u and v do not share a halting state for τ. Note that $u^* \in V_{v:u}$ and $v^* \in V_{u:v}$. Consider a set of optimal rules from the singletons to τ. Lemma 4.1 ensures that each node in $V_{u:v}$ has exactly the same τ-halting states and the nodes in $V_{v:u}$ all have the same τ-halting states. Therefore no node in $V\backslash\{u,v\}$ can be a focus for τ.

We claim that u is a focus of τ. Indeed, we show that the composite rule Γ from u to τ given by "take one step and then follow an optimal rule from that node to τ" is not optimal by proving that $x_k(\Gamma) > 0$ for all $k \in V$. Clearly u is not a halting state for Γ. If u is a leaf, then after our first step, we must be at v. Since every (v,τ)-halting state is contained in $V_{u:v} = \{u\}$, we must have $x_k(\Gamma) > 0$ for all $k \in V$. When u is not a leaf, let $i \in V_{u:v}$ be a neighbor of u. Then $x_k(\Gamma) \geq \frac{1}{d(u)}\left(x_k(v,\tau) + x_k(i,\tau)\right) > 0$. Indeed, after our first step from u, we are at each of v, i with probability $1/d(u)$. Lemma 4.2 ensures that no node is simultaneously halting for both v and u. Lemma 4.1 states that i has the same halting states as u. Combining these observations yields $x_k(v,\tau) + x_k(i,\tau) > 0$ for every node k. Γ is not optimal by equation (2.8) and therefore u is a focus of τ. By a similar proof, v is also a focus of τ.

Case 2: Every neighboring pair of nodes share a τ-halting state. Since τ is not a singleton, there exists a path of the form $(u^*,\ldots,i,u,\ldots,i^*)$ where u^* is a halting state for u but not for i, and u separates i from all of its τ-halting states. If u is not a focus, then the neighbors of u have a common halting state j^*. Let $j \neq i$ be

the neighbor of u on the (u, j^*)-path. The path $(u^*, \ldots, u, j, \ldots, j^*)$ is of the form forbidden by claim 3, a contradiction. So u must be a focus and u is unique. Indeed, u shares a τ-halting state with each of its neighbors, so u must have a halting state in at least two components of $G\backslash\{u\}$. If there were another focus v, this would again imply the existence of a path forbidden by lemma 4.2. □

Proof of proposition 1.6. Case 1: G is bifocal with foci a, b. By corollary 3.1, $a' \in V_{b:a}$ and $b' \in V_{a:b}$ and by equation (2.11), $x_{a'}(a, \pi) = 0$ and $x_{b'}(b, \pi) = 0$. The argument is now identical to case 1 in the proof of theorem 1.5 with $\tau = \pi, u = a, u^* = a', v = b, v^* = b'$.

Case 2: G is focal with focus a. Corollary 3.1 states that at least two subtrees of $G\backslash\{a\}$ contain a-pessimal nodes. By equation (2.11), these nodes are (a, π)-halting states. Lemma 4.1 now ensures every neighboring pair of nodes share a π-halting state, so a is the only potential π-focus. From this point, the argument is identical to the end of case 2 in the proof of theorem 1.5 with $\tau = \pi$ and $u = a$. □

Proof of proposition 1.7. By equation (2.13), when i is not a focus of π we have $\mu_i = 0$, If G is focal then μ is the singleton distribution on a. For G bifocal, rewrite equation (2.13) as

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

When $i \in V_{a:b}$ is a neighbor of a, theorem 1.5 shows $H(i, \pi) - H(a, \pi) = H(i, a)$. Equation (2.4) gives

$$\sum_{i \in V_{a:b}} p_{ai}(H(i, \pi) - H(a, \pi)) = \sum_{i \in V_{a:b}} p_{ai} H(i, a) = \frac{1}{d(a)}(H(a, b) - d(a)).$$

Considering the final neighbor b, equations (2.12) and (2.3) give

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

Thus our formula for μ_a becomes

$$\begin{aligned} \mu_a &= \frac{d(a)}{2|E|}\left[1 + \frac{1}{d(a)}\left(H(a, b) - d(a) + H(b', a) - H(a', b)\right)\right] \\ &= \frac{1}{2|E|}\left(H(a, b) + H(b', a) - H(a', b)\right) \;=\; \frac{1}{2|E|}\left(H(b', b) - H(a', b)\right). \end{aligned}$$

We can calculate μ_b directly as above, or use $\mu_b = 1 - \mu_a$ and equation (2.5). □

COROLLARY 4.3. *For a focal tree,* $T_{\text{forget}} = H(a', a)$. *For a bifocal tree,*

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

Proof. Since a' and b' are mixing pessimal, $T_{\text{forget}} = H(a', \mu) = H(b', \mu)$ and the first statement is obvious. If G is bifocal, the following stopping rule is optimal

from a' to μ: walk until you hit b, then stop with probability μ_b and walk to a with probability μ_a. Hence, $H(a',\mu) = H(a',b) + \mu_a H(b,a)$ and

$$T_{\text{forget}} = H(a',b) + \frac{H(b,a)}{2|E|}(H(b',b) - H(a',b)).$$

Equation (2.5) completes the proof. □

Proof of theorem 1.8. We quickly narrow our search down to the foci of the tree. Recall that a stopping rule is optimal if and only if it has a halting state. Lovász and Winkler [10] show that every mixing pessimal node is a halting state for an optimal rule from μ to π. Hence on a tree, both a' and b' are halting states for an optimal rule from μ to π. Therefore, for any node i, the rule "follow an optimal rule from i to the forget distribution μ and then follow an optimal rule from μ to π" has either a' or b' as a halting state. This rule is optimal and $H(i,\pi) = H(i,\mu) + H(\mu,\pi)$. We may minimize $H(i,\mu)$ rather than $H(i,\pi)$, which is clearly minimized by a focus of the tree.

If G has a unique focus, there is nothing to prove. Assume that G is bifocal with primary focus a and secondary focus b. Then

$$H(a,\mu) - H(b,\mu) = \mu_b H(a,b) - \mu_a H(b,a) = \frac{1}{2|E|}(H(a',b) - H(b',a))(H(a,b) + H(b,a)).$$

Thus $H(a,\mu) \geq H(b,\mu)$ if and only if $H(a',b) \geq H(b',a)$. □

5. Start-independent times. Start-independent stopping rules also identify central nodes: we now prove proposition 1.9 and theorem 1.10, which show that the target distributions achieving $T_{\text{si}}(\pi)$ and T_{si} are concentrated on a barycenter and the foci of the tree, respectively. The following lemma restricts our attention to singleton targets.

LEMMA 5.1. *Let σ and τ be distributions on the tree G. If τ has only one focus, denote this node by u. Otherwise let the foci u, v of τ satisfy $\sigma(V_{u:v})\pi(V_{u:v}) \geq \sigma(V_{v:u})\pi(V_{v:u})$. Then*

$$\sum_{k\in V} \sigma_k H(k,\tau) \geq \sum_{k\in V} \sigma_k H(k,u) = H(\sigma,u).$$

Proof. If u is the only focus for τ then $\sum_{k\in V} \sigma_k H(k,\tau) = \sum_{k\in V} \sigma_k (H(k,u) + H(u,\tau)) \geq \sum_{k\in V} \sigma_k H(k,u)$. If τ has two foci with $\sigma(V_{u:v})\pi(V_{u:v}) \geq \sigma(V_{v:u})\pi(V_{v:u})$, equation (2.4) implies that $\sigma(V_{u:v})H(u,v) \geq \sigma(V_{v:u})H(v,u)$. By theorem 1.5,

$$\begin{aligned}
\sum_{k\in V} \sigma_k H(k,\tau) &= \sum_{k\in V_{u:v}} \sigma_k (H(k,u) + H(u,\tau)) + \sum_{k\in V_{v:u}} \sigma_k (H(k,v) + H(v,\tau)) \\
&= \sum_{k\in V_{u:v}} \sigma_k H(k,u) + \sum_{k\in V_{v:u}} \sigma_k H(k,v) + \sigma(V_{u:v})H(u,\tau) + \sigma(V_{v:u})H(v,\tau) \\
&= \sum_{k\in V} \sigma_k H(k,u) - \sigma(V_{v:u})H(v,u) + \sigma(V_{u:v})H(u,\tau) + \sigma(V_{v:u})H(v,\tau).
\end{aligned}$$

For any rule from u to τ, we must step from u to $V_{v:u}$ with probability $\tau(V_{v:u})$ before halting, hence $H(u,\tau) \geq \tau(V_{v:u})H(u,v)$. Likewise, $H(v,\tau) \geq \tau(V_{u:v})H(v,u)$. Therefore

$$\begin{aligned}
\sigma(V_{u:v})H(u,\tau) + \sigma(V_{v:u})H(v,\tau) &\geq \sigma(V_{u:v})\tau(V_{v:u})H(u,v) + \sigma(V_{v:u})\tau(V_{u:v})H(v,u) \\
&\geq \sigma(V_{v:u})H(v,u)
\end{aligned}$$

so $\sum_{k\in V}\sigma_k H(k,\tau) \geq \sum_{k\in V}\sigma_k H(k,u)$. □

Proof of proposition 1.9. Taking $\sigma = \pi$ in lemma 5.1, a singleton target achieves $T_{\mathrm{si}}(\pi)$. Proposition 1.1 shows that $T_{\mathrm{si}}(\pi) = \min_{i\in V}\sum_{k\in V}\pi_k H(k,i) = \min_{i\in V} H(\pi,i) = H(\pi,c)$. □

Obviously $T_{\mathrm{si}}(\pi) \leq T_{\mathrm{reset}}$ and in fact $T_{\mathrm{si}}(\pi)$ can be arbitrarily small in comparison. Consider the tree consisting of a path of length $2k$ with k^4 leaves connected to the central node c. The focus, center and barycenter of G are all located at c, so the forget distribution is concentrated on this central node. $T_{\mathrm{forget}} = k^2$ while $H(\pi,c)$ becomes arbitrarily close to $1/2$ for large k.

On the other hand, theorem 1.10 states that the forget time and the start-independent time of a tree are identical. The theorem is clearly true for the path on two nodes, so we restrict our proof to trees on three or more nodes. We make the important observation that T_{si} need not be achieved by a unique pair of distributions. For example, consider a star graph with n leaves. Clearly $T_{\mathrm{si}} \leq 1$ since we may always choose the central node c as our target. For any distribution σ concentrated on the leaf set such that $\sigma_i \leq (2n-1)/2n$ for every node i, we have $T_{\mathrm{si}} = \min_{j\in V} H(\sigma,j) = H(\sigma,c) = 1$.

We prove theorem 1.10 by constructing a particular initial distribution ϕ concentrated on two leaves such that $\min_{j\in V} H(\phi,j) = T_{\mathrm{si}}$. Once we have identified such a ϕ, we show that we may choose the target node to be a focus of G.

LEMMA 5.2. *Given a distribution σ, the node u satisfies $H(\sigma,u) = \min_{j\in V} H(\sigma,j)$ if and only if for each neighbor v of u,*

$$\sigma(V_{v:u})H(v,u) \leq \sigma(V_{u:v})H(u,v) \tag{5.1}$$

or equivalently

$$\sigma(V_{v:u}) \leq \frac{H(u,v)}{2|E|} \text{ and } \sigma(V_{v:u}) \geq \frac{H(v,u)}{2|E|}. \tag{5.2}$$

We have equality if and only if $H(\sigma,u) = H(\sigma,v)$ so that v is also a best target for σ. Furthermore, at most one neighbor of u can satisfy equation (5.1) with equality.

Proof. For any neighbor v of u,

$$H(\sigma,v) = \sum_{k\in V_{u:v}}\sigma_k H(k,u) + \sigma(V_{u:v})H(u,v) + \sum_{k\in V_{v:u}}\sigma_k H(k,v).$$

We have $\sigma(V_{v:u})H(v,u) \leq \sigma(V_{u:v})H(u,v)$ if and only if

$$H(\sigma,v) \geq \sum_{k\in V_{v:u}}\sigma_k H(k,u) + \sigma(V_{v:u})H(v,u) + \sum_{k\in V_{v:u}}\sigma_k H(k,v) = H(\sigma,u).$$

Furthermore, equality holds in the first if and only if equality holds in the second. We find the equivalence of equations (5.1) and (5.2) by rewriting $\sigma(V_{v:u})H(v,u) \leq \sigma(V_{u:v})H(u,v) = (1-\sigma(V_{v:u}))H(u,v)$, solving for $\sigma(V_{v:u})$ and then using equation (2.5).

Now suppose that we have equality in equation (5.2) for two distinct u-neighbors v, w. Then

$$1 \geq \sigma(V_{v:u}) + \sigma(V_{w:u}) = \frac{1}{2|E|}\left(H(u,v)+H(u,w)\right) > \frac{1}{2|E|}\left(H(u,v)+H(v,u)\right) = 1,$$

a contradiction (where the inequality follows from equation (2.4)). □

We employ the following terminology for the remainder of the section. Let ϕ be a distribution, let $S_\phi = \{v|\phi_v > 0\} \subset V$ and let u be a node such that $H(\phi, u) = \min_{j\in V} H(\phi, j) = T_{\text{si}}$. Let $v_1, v_2, \ldots, v_{d(u)}$ be the neighbors of u and let $w_i \in V_{v_i:u}$ be a leaf such that $H(w_i, u) = \max_{j\in V_{v_i:u}} H(j, u)$ for $1 \le i \le d(u)$.

Proof of theorem 1.10. We first prove the result for stars $K_{1,k}$, $k \ge 0$. When $k = 0$ the result is trivial. Suppose that $G = K_{1,1}$ is the path on vertices u, v. Then $\max_\sigma \min_j H(\sigma, j) = \max_\sigma \min\{\sigma_u, \sigma_v\} = 1/2$. This value is achieved uniquely by taking $\sigma = (1/2, 1/2)$ and taking either node as the target.

Suppose that G is the star $K_{1,k}$, $k \ge 2$ with center c. When ϕ is divided evenly between two leaves u, v then $\min_j H(\phi, j) = H(\phi, c) = 1$. We now show that $\min_j H(\sigma, j) \ge 1$ for any distribution σ. Note that $\min_j H(\sigma, j) \le H(\sigma, c) = 1 - \sigma_c \le 1$. So assume that a leaf w is the best target for σ. By lemma 5.2, $\sigma(V_{c:w}) \le H(w, c)/2|E| = 1/2k$. In order to maximize $H(\sigma, w)$ we must set $\sigma_w = (2k - 1)/2k$ and $\sigma_v = 1/2k$ for some leaf $v \ne w$. In this case $H(\sigma, w) = \frac{1}{2k} H(v, w) = 1$. Therefore $\max_\sigma \min_j H(\sigma, j) = 1$ for every star.

Now assume that G is a tree on 4 or more vertices and that G is not a star. Let $\sum_{i\in V} \phi_i H(i, u) = T_{\text{si}}$.

Claim 1: The node u is not a leaf.

Assume that u is a leaf and let v be its unique neighbor. Using equation (5.2), our best choice for the initial distribution is $\phi_u = 1/2|E|$ and $\phi_{u'} = 1 - \phi_u$. In this case, $H(\phi, u) = H(\phi, v)$ by lemma 5.2, so v is also a minimizing target. We note that since G is not a star, $H(u', v) > 1$. Lemma 5.2 also guarantees that the remaining v-neighbors have a strict equality in equation (5.1). Therefore we can shift some weight from u to u' while still keeping v as the optimal target. Specifically, there exists some $\epsilon > 0$ such that the distribution ϕ' defined by $\phi'_u = \phi_u - \epsilon$, $\phi'_{u'} = \phi_{u'} + \epsilon$ and $\phi'_i = 0$ otherwise satisfies $\phi'(V_{w:v})H(w, v) < \phi'(V_{v:w})H(v, w)$ for every v-neighbor w. This ensures that v is the unique optimal target, while $H(\phi', v) > H(\phi, v) = H(\phi, u)$, a contradiction. Here the strict inequality follows from the fact that $H(u', v) > 1$.

Claim 2: S_ϕ intersects more than one component of $G\backslash\{u\}$.

Assume that S_ϕ intersects exactly one of $V_{v_1:u}, V_{v_2:u}, \ldots, V_{v_{d(u)}:u}$.

Case 1: $\phi_u = 0$. We may assume $S_\phi \subset V_{v_1:u}$. But $H(\phi, v_1) < H(\phi, u)$, contradicting $\min_{j\in V} H(\phi, j) = H(\phi, u)$.

Case 2: $\phi_u \ne 0$. We may assume $S_\phi \subset V_{v_1:u} \cup \{u\}$. By lemma 5.2 we have $\phi(V_{v_1:u})H(v_1, u) \le \phi(V_{u:v_1})H(u, v_1) = \phi_u H(i, v_1)$. If we have equality here, we may take v_1 as our target, proving the claim. If we have strict inequality, there exists $\epsilon > 0$ such that the distribution ϕ' defined by $\phi'_u = \phi_u - \epsilon$, $\phi'_{v_1} = \phi_{v_1} + \epsilon$ and $\phi'_i = \phi_i$ otherwise satisfies $\phi'(V_{v_1:u})H(v_1, u) < \phi'(V_{u:v_1})H(u, v_1)$. Lemma 5.2 shows that $\min_{i\in V} H(\phi', i) = H(\phi', u)$, while $H(\phi', u) > H(\phi, u) = T_{\text{si}}$, a contradiction.

Claim 3: ϕ may be chosen so that $S_\phi \subset \{w_1, w_2, \ldots, w_{d(u)}\}$.

Assume instead that $S_\phi \not\subset \{w_1, w_2, \ldots, w_{d(u)}\}$.

Case 1: $\phi_u = 0$. Let ϕ' be the distribution given by $\phi'_{w_i} = \phi(V_{v_i:u})$ for $1 \le i \le d(u)$ and zero elsewhere. Lemma 5.2 and $H(\phi, u) = \min_{i\in V} H(\phi, i)$ imply that $H(\phi', u) = \min_{i\in V} H(\phi', i)$ as well. Clearly, $H(\phi', u) \ge H(\phi, u)$, so we may use ϕ' in place of ϕ.

Case 2: $\phi_u \ne 0$. By an argument analogous to case 1, we may choose ϕ so that $S_\phi \subset \{u, w_1, w_2, \ldots, w_{d(u)}\}$. Suppose that $\phi(V_{u:v_i})H(u, v_i) = \phi(V_{v_i:u})H(v_i, u)$ for some i. We may take v_i as our target node in lieu of u by lemma 5.2. Since v_i is an

optimal target, claim 1 ensures that v_i is not a leaf, so $\phi_{v_i} = 0$ and we have reduced ourselves to case 1.

If $\phi(V_{u:v_i})H(u,v_i) > \phi(V_{v_i:u})H(v_i,u)$ for $1 \le i \le d(u)$ then there exists $\epsilon > 0$ such that the distribution ϕ' given by $\phi'_u = \phi_u - \epsilon$, $\phi'_{w_1} = \phi_{w_1} + \epsilon$ and $\phi'_i = \phi_i$ otherwise satisfies $\phi'(V_{v_i:u})H(v_i,u) < \phi'(V_{u:v_i})H(u,v_i)$ for $1 \le i \le d(u)$. By lemma 5.2, $\min_{i\in V} H(\phi',i) = H(\phi',u) > H(\phi,u)$, a contradiction.

Claim 4: ϕ may be chosen to be concentrated on two leaves in $\{w_1, w_2, \dots w_{d(u)}\}$.

Case 1: $\phi(V_{u:v_i})H(u,v_i) = \phi(V_{v_i:u})H(v_i,u)$ for some i. By lemma 5.2, $H(\phi,u) = H(\phi,v_i)$. Notice that ϕ is supported in two components of $G\backslash\{v_i\}$, so using the proof of claim 3 we may define a new distribution ϕ' concentrated on two leaves such that $\min_{k\in V} H(\phi',k) = H(\phi',v_i) \ge H(\phi,v_i) = H(\phi,u) = T_{\text{si}}$.

Case 2: $\phi(V_{u:v_i})H(u,v_i) > \phi(V_{v_i:u})H(v_i,u)$ for all i. We show by induction that there exists a distribution ϕ' supported on two leaves such that $\min_{i\in V} H(\phi',i) \ge H(\phi,u)$. The base case $|S_\phi| = 2$ is trivial. Assume that if $|S_\phi| = k-1$ then there exists a ϕ' concentrated on two leaves satisfying $\min_{j\in V}\sum_{i\in V}\phi'_i H(i,j) = \sum_{i\in V}\phi'_i H(i,u) = \sum_{i\in V}\phi_i H(i,u) = T_{\text{si}}$.

Considering $|S_\phi| = k \le d(u)$, order $S_\phi = \{w_1, w_2, \dots, w_k\}$ so that $H(w_1,u) \ge H(w_2,u) \ge \cdots \ge H(w_k,u)$. There exists $\epsilon > 0$ such that the distribution ϕ^* defined by $\phi^*_{w_1} = \phi_{w_1} + \epsilon$, $\phi^*_{w_k} = \phi_{w_k} - \epsilon$ and $\phi^*_i = \phi_i$ otherwise satisfies $\phi^*(V_{v_i:u})H(v_i,u) < \phi^*(V_{u:v_i})H(u,v_i)$ for all i. If $H(w_1,u) > H(w_k,u)$ then by lemma 5.2, $\min_{i\in V} H(\phi^*,i) = H(\phi^*,u) > H(\phi,u) = T_{\text{si}}$, a contradiction.

Otherwise, we have $H(w_i,u) = H(w_j,u)$ for $1 \le i,j \le k$. If there exists $0 < \epsilon < \phi_{w_k}$ such that $\phi^*(V_{u:v_k})H(u,v_k) = \phi^*(V_{v_k:u})H(v_k,u)$ we have $H(\phi^*,v_k) = H(\phi^*,u) = H(\phi,u) = T_{\text{si}}$. Hence we may take ϕ^* as our starting distribution and v_k as our target node. The support of ϕ^* is contained in two connected components of $V\backslash\{v_k\}$, and so the proof of claim 3 shows that there exists a distribution ϕ' supported on two leaves such that $T_{\text{si}} = \min_{i\in V} H(\phi',i)$. Finally, if $\phi^*(V_{u:v_k})H(u,v_k) > \phi^*(V_{v_k:u})H(v_k,u)$ for all $0 \le \epsilon \le \phi_{w_k}$, taking $\epsilon = \phi_{w_k}$ we have a distribution supported on $k-1$ leaves such that $\min_{i\in V} H(\phi^*,i) = T_{\text{si}}$ and we are done by induction.

Claim 5: ϕ may be chosen so that S_ϕ concentrated on two leaves w_1 and w_2 such that $H(w_1,u) \ge H(w_2,u)$ and the target node u is a focus of G. If the tree G is focal, then ϕ is concentrated on two u-pessimal leaves. If the tree G is bifocal, then w_1 is u-pessimal, v_1 is the other focus of G and w_2 is v_1-pessimal. In this case, ϕ is given by $\phi_{w_1} = \pi(V_{u:v_1})$, $\phi_{w_2} = \pi(V_{v_1:u})$ and $H(\phi,v_1) = H(\phi,u) = T_{\text{si}}$.

By claim 4, we may assume that ϕ is concentrated on leaves $w_1 \in V_{v_1:u}$ and $w_2 \in V_{v_2:u}$ where $H(w_1,u) \ge H(w_2,u)$. In order to maximize the access time, the distribution ϕ must weight w_1 as much as possible while still keeping u as the best target node. By lemma 5.2, we must have $\phi_{w_1}H(v_1,u) \le \phi_{w_2}H(u,v_1)$ and $\phi_{w_2}H(v_2,u) \le \phi_{w_1}H(u,v_2)$. By equation (2.4) this is equivalent to

$$\begin{cases} \phi_{w_1}\pi(V_{v_1:u}) \le \phi_{w_2}\pi(V_{u:v_1}) \\ \phi_{w_2}\pi(V_{v_2:u}) \le \phi_{w_1}\pi(V_{u:v_2}) \end{cases} \tag{5.3}$$

and the optimal choice is $\phi_{w_1} = \pi(V_{u:v_1})$ and $\phi_{w_2} = \pi(V_{v_1:u})$. Note that this choice results in

$$H(\phi,v_1) = H(\phi,u) \tag{5.4}$$

by lemma 5.2.

The node w_1 is u-pessimal. Indeed by claim 3, $H(w_1,u) = \max_{i\in V_{v_1:u}} H(i,u)$, so if $H(w_1,u) < H(u',u)$ then a u-pessimal node u' must lie in one of $V_{v_i:u}$, $2 \le i \le d(u)$.

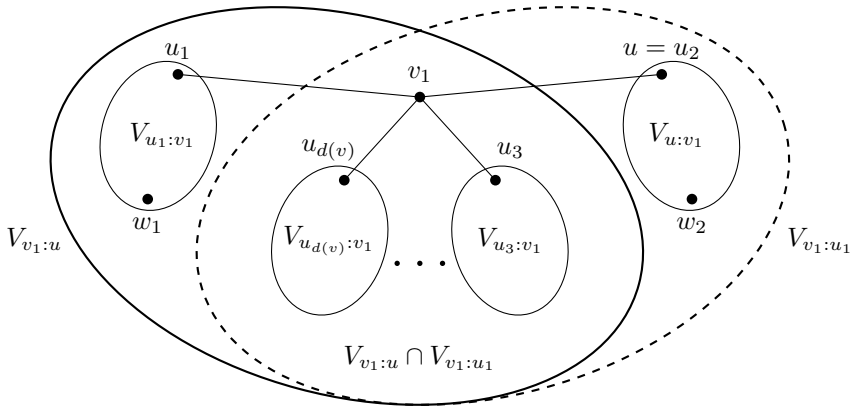

FIG. 5.1. *The tree decomposition in claim 5 of the proof of theorem 1.10.*

Since $H(w_2, u) \leq H(w_1, u)$, w_2 cannot be u-pessimal, so consider the distribution ϕ' given by $\phi'_{w_1} = \phi_{w_1}$, $\phi'_{u'} = \phi_{w_2}$ and $\phi'_i = 0$ otherwise. This distribution ϕ' satisfies the inequalities analogous to equation (5.3) so $\min_{i \in V} H(\phi', i) = H(\phi', u) > H(\phi, u) = T_{\text{si}}$, a contradiction. By a similar argument, w_2 must satisfy $H(w_2, u) = \max_{i \in V_{u:v_1}} H(i, u)$.

If u is the unique focus of G then w_2 must also be u-pessimal. If u is a focus of a bifocal G then v_1 must be the other focus of G and w_2 is v_1-pessimal. If u is not a focus, then the foci of G must be on the unique path between u and the u-pessimal node w_1. Hence w_1 is v_1-pessimal and $H(w_1, v_1) \geq H(w_2, v_1)$. By equation (5.4) we may take v_1 as our target node instead of u. If $H(w_1, v_1) = H(w_2, v_1)$ then v_1 must be the unique focus of G since w_1 is a v_1-pessimal node, and $T_{\text{si}} = H(\phi, v_1)$ as required.

Assume for the sake of contradiction that u is not a focus of G and $H(w_1, v_1) > H(w_2, v_1)$. Let $u_1, u_2, \ldots, u_{d(v_1)}$ be the neighbors of v_1, with $w_1 \in V_{u_1:v_1}$ and $w_2 \in V_{u_2:v_1}$. This ordering ensures that $u = u_2$. Consider the distribution ϕ' given by $\phi'_{w_1} = \pi(V_{v_1:u_1})$, $\phi'_{w_2} = \pi(V_{u_1:v_1})$ and $\phi'_i = 0$ otherwise. By lemma 5.2, $H(\phi', v_1) = \min_{i \in V} H(\phi', i)$ and

$$\begin{aligned}
H(\phi', v_1) &= \pi(V_{v_1,u_1}) H(w_1, v_1) + \pi(V_{u_1:v_1}) H(w_2, v_1) \\
&= \left(\pi(V_{u:v_1}) + \pi(V_{v_1:u} \cap V_{v_1:u_1})\right) H(w_1, v_1) + \pi(V_{u_1:v_1}) H(w_2, v_1) \\
&> \pi(V_{u:v_1}) H(w_1, v_1) + \left(\pi(V_{v_1:u} \cap V_{v_1:u_1}) + \pi(V_{u_1:v_1})\right) H(w_2, v_1) \\
&= \pi(V_{u:v_1}) H(w_1, v_1) + \pi(V_{v_1:u}) H(w_2, v_1) \\
&= H(\phi, v_1) = H(\phi, u) = T_{\text{si}}
\end{aligned}$$

where the second and fourth equalities follow from the decomposition of the tree (as seen in figure 2) and the inequality is due to $H(w_1, v_1) > H(w_2, v_2)$. The resulting inequality $H(\phi', v_1) > T_{\text{si}}$ is a contradiction, so u must be a focus of G.

COMPLETION OF PROOF. If G has a single focus a, then by claim 5, we may take ϕ to be concentrated on two a-pessimal leaves in different components of $G \backslash \{a\}$ and $T_{\text{si}} = H(\phi, a) = H(a', a) = T_{\text{forget}}$ by corollary 4.3. If G is bifocal with foci a and b, then by claim 5, we may take ϕ to be concentrated on an a-pessimal node a' and a

b-pessimal node b'. Also, $T_{\text{si}} = H(\phi, a) = H(\phi, b)$ by equation (5.4). Finally,

$$\begin{aligned} T_{\text{si}} = (\mu_a + \mu_b)T_{\text{si}} &= \mu_a H(\phi, a) + \mu_b H(\phi, b) \\ &= \phi_{a'}(\mu_a H(a', a) + \mu_b H(a', b)) + \phi_{b'}(\mu_a H(b', a) + \mu_b H(b', b)) \\ &= \phi_{a'} H(a', \mu) + \phi_{b'} H(b', \mu) = (\phi_{a'} + \phi_{b'})T_{\text{forget}} = T_{\text{forget}} \end{aligned}$$

by proposition 1.7. □

Acknowledgments. The author would like to thank László Lovász for many insightful conversations, Peter Winkler for his helpful comments, and the anonymous referees for their suggested improvements to the exposition.

REFERENCES

[1] D. Aldous, *Some inequalities for reversible Markov chains*, J. London Math. Soc., 25 (1982), pp. 564–576.
[2] D. Aldous, L. Lovász and P. Winkler, *Mixing times for uniformly ergodic Markov chains*, Stochastic Processes and their Applications, 71 (1997), pp. 165–185.
[3] A. Beveridge and L. Lovász, *Random walks and the regeneration time*, J. Graph Theory, 29 (1998), pp. 57-62.
[4] G. Brightwell and P. Winkler, *Extremal Cover Times for Random Walks on Trees*, J. Graph Theory,14 (1990), pp. 547–554.
[5] D. Coppersmith, P. Tetali, and P. Winkler, *Collisions among Random Walks on a Graph*, SIAM J. on Discrete Mathematics, 6 No. 3 (1993), pp. 363–374.
[6] P. Diaconis and J. Fill, *Strong Stationary Times Via a New Form of Duality*, The Annals of Probability, 18 No. 4 (1990), pp. 1483–1522.
[7] I. Dumitriu, P. Tetali and P. Winkler, *On playing golf with two balls*, SIAM J. on Discrete Mathematics, 16 No. 4 (2003), pp. 604–615.
[8] L. Lovász, *Random Walks on Graphs: a Survey*, Combinatorics, Paul Erdős is Eighty, Vol. II, D. Miklós, V. T. Sós and T. Szőnyi, ed., J. Bolyai Math. Soc. (1996), pp. 353–397.
[9] L. Lovász and P. Winkler, *Efficient stopping rules for Markov chains*, Proceedings of the 27th ACM Symposium on the Theory of Computing (1995), pp. 76–82.
[10] L. Lovász and P. Winkler, *Reversal of Markov Chains and the Forget Time*, Cominatorics, Probability and Computing, 7 (1998), pp. 189–204.
[11] J. W. Moon, *Random Walks on Random Trees*, J. Austral. Math. Soc., 15 (1973), pp. 42–53.
[12] L. H. Pearce, *Random Walks on Trees*, Discrete Math., 30 (1980), pp. 269–276.
[13] J. W. Pitman, *Occupation measures for Markov chains*, Adv. Appl. Prob., 9 (1977), pp. 69-86.
[14] P. Tetali, *Random Walks and Effective Resistance of Networks*, J. Theoretical Prob. No. 1 (1991), pp. 101–109.
[15] P. Tetali and P. Winkler, *Simultaneous Reversible Markov Chains*, Combinatorics, Paul Erdős is Eighty, Vol. I, D. Miklós, V. T. Sós and T. Szőnyi, ed., J. Bolyai Math. Soc. (1993), pp. 433–451.