Efficient Stopping Rules for Markov Chains

(EXTENDED ABSTRACT)

László Lovász
Dept. of Computer Science, Yale University, New Haven CT 06510; lovasz@cs.yale.edu
Peter Winkler
AT&T Bell Laboratories 2D-147, Murray Hill NJ 07974; pw@research.att.com

Abstract

Let M be the transition matrix, and σ the initial state distribution, for a discrete-time finite-state irreducible Markov chain. A *stopping rule* for M is an algorithm which observes the progress of the chain and then stops it at some random time Γ; the distribution of the final state is denoted by σ^Γ.

We give a useful characterization for stopping rules which are optimal for given target distribution τ, in the sense that $\sigma^\Gamma = \tau$ and the expected stopping time $\mathrm{E}\Gamma$ is minimal. Four classes of optimal stopping rules are described, including a unique "threshold" rule which also minimizes $\max(\Gamma)$. The minimum value of $\mathrm{E}\Gamma$, which we denote by $H(\sigma,\tau)$, is easily computable from the hitting times of M.

For applications in computing, the most important case is when σ is concentrated on a single starting state s and τ is the stationary distribution π. We describe a simple, practical stopping rule which achieves target distribution close to π in expected time of order $T_{\rm mix} = \max_s H(s,\pi)$.

Finally, we give a stopping rule that runs in time polynomial in the maximum hitting time of M and achieves the stationary distribution exactly, even though the transition probabilities of the chain are unknown.

Some of the work described herein is joint with David Aldous.

1 Introduction

In the past ten years many new applications have been found (see, e.g., [3]) for sampling via Markov chains; these have resulted in polynomial-time randomized approximation schemes for computing the volume of a convex body [12] and counting combinatorial objects such as matchings [13], linear extensions [14], and Eulerian orientations [18]. Typically, in these applications, sampling from an approximately correct distribution is accomplished by walking randomly on a graph for a fixed number of steps, after which the distribution of the occupied vertex is nearly stationary.

There is no particular reason why such a walk must be run for a fixed number of steps; in fact, more general stopping rules which "look where they are going" are capable of achieving the stationary distribution exactly, and just as quickly. Such stopping rules (times) have been studied in the theory of Markov chains (see e.g. [7]), but not from the point of view of algorithmic efficiency.

In this work we study stopping rules that achieve any given distribution, when starting from some other given distribution. It turns out that there is a surprising variety of such rules, several of which are optimal in the sense that they entail the smallest possible expected number of steps; one of them also minimizes the *maximum* number of steps. These rules are related to important parameters of the chain, like hitting times and conductance.

From a theoretical standpoint, our stopping rules result in a generalization of the notion of "hitting time" to state-distributions. This measure of distance between distributions behaves quite nicely.

The most important special case arises when one wishes to generate a state from the stationary distribution, starting from a given state. The minimum time to do this, which we call $T_{\rm mix}$, is easily calculable (especially in the reversible case) from the hitting times. However, in applications the latter are not generally known. Furthermore, the hitting times are typically large (say, around n^2 for a graph on n nodes), while the mixing time, in the cases of interest for sampling algorithms, is polylogarithmic in n: it can expressed as a difference of hitting times, and hence even asymptotic values for the hitting times are insufficient to bound it. In Section 7 below we describe a simple, practical stopping rule which achieves an approximation to the stationary distribution in time $\mathrm{O}(T_{\rm mix})$.

From the point of view of applications to sampling algorithms, our results imply that we don't gain too much in the efficiency of the algorithms by considering sophisticated stopping rules (cf. Theorem 7.2). On the other hand, the theory of stopping rules, in particular the algebraic machinery developed here, offers new ways of proving bounds on the mixing time and rate of Markov chains. We do not elaborate on these applications here.

In the final section we describe a stopping rule which achieves the stationary distribution exactly even though nothing is known about the transition probabilities. The expected number of steps for this rule is nowhere near optimal, but is nonetheless polynomial in the maximum hitting time of the chain.

Work described herein will be found in three papers, [6], [15] and [16]; the first of these is joint with David Aldous, of the Dept. of Statistics, U.C. Berkeley. Research of the first author was supported in part by NSF grant No. CCR-9402916.

2 Notation and Preliminaries

For most of this paper we will assume a fixed irreducible Markov chain with transition matrix $M = \{p_{ij}\}$ and initial distribution σ, over a state space S of finite cardinality n. Most of our results generalize in straightforward manner in the presence of transient states and/or a countably infinite state space, but with messier conditions and arguments which we prefer to avoid. We assume discrete-time transitions but, again, most or our results can be modified to apply to the continuous-time case.

We denote by S^* be the space of finite "walks" on S, that is, the set of finite strings $w = (w_0, w_1, \ldots, w_t)$, $w_i \in S$. Relative to the Markov chain we then have

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

The distribution of w_t will be denoted by σ^t, so that $\sigma^0 = \sigma$ and $\sigma_i^t = \sum_{w_t=i} \Pr((w_0, \ldots, w_t))$.

It will be useful to us to regard a stopping rule Γ both as a random stopping time and as a (partial) function from S^* to $[0,1]$ giving the probability of continuing from a walk w. Formally:

Definition. A *stopping rule* Γ is a partial map from S^* to $[0,1]$, such that $\Gamma(w)$ is defined for $w = (w_0, \ldots, w_t)$ just when $\Pr(w) > 0$ and $\Gamma(w_0, \ldots, w_i)$ is defined and non-zero for each i, $0 \le i < t$. If it happens that $\Gamma(w)$ depends only on w_t and t then we say that Γ is *balanced* and write $\Gamma(t; w_t) := \Gamma(w)$.

We interpret $\Gamma(w)$ as the probability of continuing given that w is the walk so far observed, each such stop-or-go decision being made independently. We can also regard Γ as a random variable with values in $\{0, 1, \ldots\}$, so that we stop at w_Γ. The probability of stopping at state j, given starting distribution σ, is

$$\sigma_j^\Gamma \ := \sum_{w=(w_0,\ldots,w_t=j)} \sigma_{w_0} \left(\prod_{i=0}^{t-1} \Gamma(w_0, \ldots, w_i) p_{w_i, w_{i+1}} \right) (1 - \Gamma(w)) \ .$$

The *mean length* $\mathrm{E}\Gamma$ of the stopping rule Γ is its expected duration; if $\mathrm{E}\Gamma < \infty$ then with probability 1 the walk eventually stops, and thus σ^Γ is a probability distribution. A stopping rule Γ for which $\sigma^\Gamma = \tau$ is also called a stopping rule from σ to τ.

Note that if $\mathbf{X}$ is a random variable whose values are stopping rules, then $\mathbf{X}$ is equivalent to a stopping rule Γ whose probability of continuing at w is a sum or integral of $\mathbf{X}(w)$ conditioned on $\mathbf{X}$ not having stopped the chain so far. Thus our stopping rules are not less general for insisting on independent randomization at each step.

Note that for any Markov chain and any τ, there is at least one finite stopping rule Γ such that $\sigma^\Gamma = \tau$; namely, we select a target state j in accordance with τ and walk until we reach j. We call this the "naive" stopping rule Ω_τ. Obviously, the mean length of Ω_τ is given by

$$\sum_{i,j} \sigma_i \tau_j H(i,j)$$

where $H(i,j)$ is the ordinary mean first passage time, or "hitting time," from i to j. In the case when $\tau = \pi$ is the stationary distribution, this formula can be simplified using what we call the "random target identity," which says that

$$\sum_j \pi_j H(i,j) = C \tag{1}$$

is independent of the choice of i (see e.g. the "right averaging principle" in Aldous [1]). Thus the mean length of the naive rule to reach π is C, independently of the starting distribution.

The *maximum length* $\max(\Gamma)$ of a stopping rule Γ is the maximum length of a walk that has positive probability. Γ is said to be *bounded* (for σ) if this quantity is finite. It is clear that for a fixed M and Γ, $\max(\Gamma)$ only depends on the support of σ. We will see later that bounded stopping rules are available when the the target distribution τ has sufficiently large support, thus in particular when $\tau = \pi$.

We often think of a stopping rule Γ as a means of moving from a starting distribution σ to a given target distribution $\tau = \sigma^\Gamma$. Such a Γ is said to be *mean-optimal* or simply *optimal* (for σ and τ) if $\mathrm{E}\Gamma$ is minimal, *max-optimal* if $\max(\Gamma)$ is minimal.

The mean length of a mean-optimal stopping rule from σ to τ will be called the *access time* from σ to τ and denoted $H(\sigma, \tau)$. As suggested by the notation, we think of $H(\sigma, \tau)$ as a generalized hitting time.

3 Elementary Properties

Trivially, $H(\sigma, \tau) = 0$ if and only if $\sigma = \tau$. It is easy to see that the following *triangle inequality* is satisfied for any three distributions σ, ρ and τ:

$$H(\sigma, \tau) \le H(\sigma, \rho) + H(\rho, \tau). \tag{2}$$

(To generate τ from σ, we can first use an optimal rule to generate ρ from σ and then use the state obtained as a starting state for an optimal rule generating τ from ρ). We should warn the reader, however, that $H(\sigma, \tau) \ne H(\tau, \sigma)$ in general.

Clearly if τ is concentrated at j (for which we write, rather carelessly, "$\tau = j$") then

$$H(\sigma, j) = \sum_i \sigma_i H(i,j) \tag{3}$$

since clearly the only optimal stopping rule in this case is Ω_j, "walk until state j is reached." This is therefore both mean-optimal and max-optimal, although $\max(\Omega_j)$ will generally be infinite.

By considering the naive rule Ω_τ, we get the inequality

$$H(\sigma, \tau) \le \sum_{i,j} \sigma_s \tau_j H(i,j) \ . \tag{4}$$

This may be quite far from equality; for example, $H(\sigma, \sigma) = 0$ for any σ.

The access time $H(\sigma, \tau)$ is convex in both its arguments, i.e.,

$$H(c\sigma + (1-c)\sigma', \tau) \le cH(\sigma, \tau) + (1-c)H(\sigma', \tau) \tag{5}$$

and

$$H(\sigma, c\tau + (1-c)\tau') \leq cH(\sigma,\tau) + (1-c)H(\sigma,\tau') \qquad (6)$$

for $0 \leq c \leq 1$ and any distributions $\sigma, \sigma', \tau, \tau'$ on S.

We set $H_{\max}(\tau) = \max_\sigma H(\sigma,\tau) = \max_i H(i,\tau)$. From the point of view of applications, stopping rules generating states from the stationary distribution are of particular interest. The value $T_{\text{mix}} = H_{\max}(\pi) = \max_i H(i,\pi)$ (the mean time of an optimal rule, starting from the worst point) is a natural and very useful definition of the mixing time.

It turns out that for given target distribution τ there are at least four interesting optimal stopping rules: the *filling rule*, the *local rule*, the *chain rule* and the *threshold rule*. We describe these rules, together with some important non-optimal stopping rules, in upcoming sections.

4 Exit Frequencies

Let us now fix M, σ and a finite stopping rule Γ.

Definition. The *exit frequencies* $x = \{x_i\}_{i\in S}$ for Γ are given by setting x_i equal to the expected number of times the walk leaves state i before stopping.

Exit frequencies are a special case of what Pitman [19] calls "pre-T occupation measures" for stopping times T. Indeed, versions of Lemma 4.2 and Theorem 4.4 below are proved in [19] using Pitman's "occupation measure identity."

We begin with some simple facts about exit frequencies. The first result is from Aldous [1]; several related formulas can be derived using relations to electrical networks, as in [9], [11], or [20].

Lemma 4.1 *The exit frequencies $\tilde{x}$ for the naive stopping rule Ω_j in reaching state j from state i are given by*

$$\tilde{x}_k = \pi_k \left(H(i,j) + H(j,k) - H(i,k) \right) .$$

More generally, the exit frequencies for the naive stopping rule Ω_τ from initial distribution σ are given by

$$\begin{aligned}\tilde{x}_k &= \pi_k \sum_{i,j} \sigma_i \tau_j (H(i,j) + H(j,k) - H(i,k)) \\ &= \pi_k \left(\sum_{i,j} \sigma_i \tau_j H(i,j) + H(\tau,k) - H(\sigma,k) \right) .\end{aligned}$$

Counting how often a state i is entered and exited, we get the following relation between exit frequencies and starting and ending distributions (Pitman [19]):

Lemma 4.2 *The exit frequencies of any stopping rule that reaches distribution τ from distribution σ satisfy the equations*

$$\sum_i p_{i,j} x_i - x_j = \tau_j - \sigma_j .$$

Theorem 4.3 *Fix σ and let Γ and Γ' be two finite stopping rules with exit frequencies x and x' respectively. Let $D = \mathrm{E}\Gamma' - \mathrm{E}\Gamma$ be the difference between their mean lengths. If $\sigma^\Gamma = \sigma^{\Gamma'}$ then $x' - x = D\pi$, where π is the stationary distribution of M; conversely if x' and x differ by a multiple of π then $\sigma^\Gamma = \sigma^{\Gamma'}$.*

Proof. By Lemma 4.2,

$$\sigma_j^\Gamma = \sigma_j + \sum_i p_{i,j} x_i - x_j .$$

In case $\sigma^\Gamma = \sigma^{\Gamma'}$, we can subtract the similar identity involving x' to get

$$x'_j - x_j = \sum_i p_{i,j} (x'_i - x_i)$$

which means that $x' - x$ is a multiple of the stationary distribution, say $x' - x = D\pi$. Since $\sum_j x_j = \mathrm{E}\Gamma$, the difference in mean length is D. The converse statement follows by the same formulas. □

It follows from theorem 4.3 that the exit frequencies of any mean-optimal stopping rule from σ to τ are the same. We denote them by $x_i(\sigma,\tau)$. In particular,

Theorem 4.4 *The exit frequencies of a mean-optimal stopping rule from σ to τ are given by*

$$x_k(\sigma,\tau) = \pi_k \left(H(\tau,k) - H(\sigma,k) + H(\sigma,\tau) \right).$$

5 Halting States and Optimal Rules

Any state j for which $x_j = 0$ is called a *halting state*. By definition we stop immediately if and when any halting state is entered. (Of course we may stop in other states too, just not all the time.) The following theorem gives an extremely useful characterization of optimality.

Theorem 5.1 *A stopping rule Γ is mean-optimal for target distribution τ if and only if it has a halting state.*

Proof. Since x_j cannot be negative, it is a trivial corollary of Theorem 4.3 that if some $x_j = 0$ then Γ is mean-optimal. The converse is less obvious: we must provide, for any two distributions σ and τ, a stopping rule that has a halting state.

The proof makes use of the discrete version of the "filling scheme," introduced by Chacon and Ornstein [8] and shown by Baxter and Chacon [7] to minimize expected number of steps. We call it the *filling rule* Φ_τ (or $\Phi_{\sigma,\tau}$ when dependence on initial distribution must be made explicit) and define it recursively as follows.

Let p_i^k be the probability of being at state i after k steps (and thus not having stopped at a prior step); let q_i^k be the probability of stopping at state i in *fewer* than k steps. Then if we are at state i after step k, we stop with probability $\min(1, (\tau_i - q_i^k)/p_i^k)$.

Thus, Φ_τ stops myopically as soon as it can without overshooting the target probability of its current state. We will see that Φ_τ is a finite stopping rule and thus it does in fact achieve τ when started at σ.

As an example, suppose we take a random walk starting at vertex 3 of the path 1—2—3—4—5, with target distribution $\tau = (1/4, 0, 1/2, 0, 1/4)$. There is a stopping rule here that is both mean-optimal and max-optimal, namely "take two steps and quit." In contrast the filling rule stops dead at the starting point half the time, otherwise it walks until it reaches an endpoint (which takes 4 steps on the average). It thus achieves the same mean length 2, but unbounded maximum length. For both rules the endpoints are halting states. Note that the naive rule Ω_τ is worse, choosing its stopping vertex beforehand to get mean length $H(3,1)/4 + H(3,5)/4 = 6$.

A straightforward induction, omitted here, shows that the filling rule always has a halting state. □

Theorem 5.2 *For all distributions σ and τ,*

$$H(\sigma,\tau) = \max_j \left(H(\sigma,j) - H(\tau,j)\right) .$$

This maximum is achieved precisely when j is a halting state of some optimal rule which stops at τ when started at σ.

Proof. The inequality $H(\sigma,\tau) \geq H(\sigma,j) - H(\tau,j)$ is a special case of the triangle inequality (2). On the other hand, if j is a halting state for (σ,τ) then it costs nothing to pass through τ on the way from σ to j, i.e. $H(\sigma,j) = H(\sigma,\tau) + H(\tau,j)$. □

In the important special case of time-reversible chains, we can use the "cycle-reversing" identity

$$H(i,j) + H(j,k) + H(k,i) = H(i,k) + H(k,j) + H(j,i).$$

of Coppersmith et al. [10] to obtain the following formula for the exit frequencies of the naive rule:

$$\begin{aligned} \tilde{x}_k &= \pi_k \sum_{i,j} \sigma_i \tau_j \left(H(j,i) + H(k,j) - H(k,i)\right) \\ &= \pi_k \left(\sum_{i,j} \sigma_i \tau_j H(j,i) \right. \\ &\quad + \left. \sum_i (\tau_i - \sigma_i) H(k,i) \right) . \end{aligned}$$

It follows that the exit frequencies of a mean-optimal rule from σ to τ are

$$\begin{aligned} x_k(\sigma,\tau) &= \pi_k \left(\sum_i (\tau_i - \sigma_i) H(k,i) \right. \\ &\quad - \left. \min_j \sum_i (\tau_i - \sigma_i) H(j,i) \right) \end{aligned}$$

and

$$\begin{aligned} H(\sigma,\tau) &= \sum_k x_k(\sigma,\tau) \\ &= \sum_i (\tau_i - \sigma_i) H(\pi,i) \\ &\quad - \min_j \sum_i (\tau_i - \sigma_i) H(j,i) . \end{aligned}$$

In the case when $\tau = \pi$ and σ is concentrated at a single state i (which is perhaps the most common in applications of Markov chain techniques to sampling), this formula can be simplified using the random target identity (1) to get that any mean-optimal rule from i to π has exit frequencies

$$x_k = \pi_k (\max_j H(j,i) - H(k,i)) \tag{7}$$

and

$$H(i,\pi) = \max_j H(j,i) - H(\pi,i) . \tag{8}$$

We have thus identified the halting state of a reversible chain, in attaining the stationary distribution from a fixed state i, as the state j *from which* the mean time to hit i is greatest. This seems slightly perverse in that we are interested in getting from i to j, not the other way 'round!

Our results concerning exit frequencies can be used to describe two more mean-optimal stopping rules; then in the following section, we will describe a stopping rule which achieves the least *maximum* number of steps as well as the least expectation.

Let τ be arbitrary and let x_i be the exit frequencies for a mean-optimal stopping rule from σ to τ. Consider the following "local" rule Λ_τ: if we are at node i, we stop with probability $\tau_i/(x_i + \tau_i)$, and move on with probability $x_i/(x_i + \tau_i)$ (if $x_i + \tau_i = 0$ the stopping probability does not need to be defined). Thus the probability of stopping under Λ_τ depends only on the current state, not the time; in effect, it treats "stop" like a new (absorbing) state of the Markov chain.

Theorem 5.3 *The local rule generates τ, i.e., $\sigma^{\Lambda_\tau} = \tau$. Moreover, it is mean-optimal.*

Let ρ be a probability distribution on the *subsets* of the state space S, and observe that it provides a stopping rule: "choose a subset U from ρ, and walk until some state in U is hit." The naive rule is of course a special case, with ρ concentrated on singletons. Another special case, which we call the *chain rule*, is described in the next theorem.

Theorem 5.4 *For every target distribution τ, there exists a unique distribution ρ which is concentrated on a chain of subsets and gives an optimal stopping rule for generating τ.*

It is interesting that this result defines an ordering (with ties—technically, a "preorder") of the states for every σ and τ, starting with the state having the largest surplus (σ_i/τ_i maximum), and ending with the halting state. (This ordering is in general different from the ordering by exit frequencies, the ordering implicit in the filling rule, or the ordering by thresholds defined in the next section.)

The chain rule shares with the filling rule the "now-or-never" property that once a state is exited, it can never be the state at which the rule stops. In fact, once it exits state i_k it can never stop at any i_j for $j \leq k$ in the notation of the proof.

Another rather neat way to think of the chain rule is to assign the real value $r(i) = \sum\{\rho(U) : i \in U\}$ to each state i. The chain rule is then implemented by choosing a random real r uniformly from $[0,1]$ and walking until a state j with $r(j) \leq r$ is reached.

6 Threshold Rules

A stopping rule Γ is said to be of the "threshold type", or for short to be a *threshold rule*, if there is a threshold vector $h = (h_1, \ldots, h_n)$, $h_i \in [0,\infty]$ such that

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

Thus in effect each state is provided with a critical time. If we hit that state past the critical time we stop; otherwise we continue, except that if we're within 1 we use the fractional part to randomize. The threshold vector may not be uniquely determined by a threshold rule Γ (e.g. all possible thresholds h_i smaller than the time before any possible walk reaches i are equivalent), but by convention we always associate with Γ the vector $h(\Gamma)$ each of whose coordinates is minimal. Then $h(\Gamma)$ is finite if and only if Γ is bounded. Moreover, in view of Theorem 5.1, Γ is mean-optimal just when some coordinate of $h(\Gamma)$ is zero.

Theorem 6.1 *For every target distribution τ there is a mean-optimal threshold rule.*

Proof. We construct such a rule, denoted Θ_τ, recursively. In a sense Θ_τ is the antithesis of the filling rule Φ_τ; it keeps moving whenever possible without overshooting the mean-optimal exit frequencies $x(\sigma, \tau)$.

We see from the definition that threshold rules are balanced, hence it suffices to define $\Theta_\tau(t; i)$. Accordingly, we assume $\Theta_\tau(t'; i)$ has been defined for all states i and all times $t' < t$ and that some thresholds h_i have already been set, including $h_j = 0$ for all halting states j. Let $q_i(t) > 0$ be the probability of being in state i at time t (under the influence of Θ_τ), and let $x_i(t)$ be the expected number of times state i has been exited before time t. If $x_i(t) = x_i$ then h_i has already been set to some value $\leq t-1$, and thus we stop at state i at time t and thereafter. If $x_i - x_i(t) > q_i(t)$ we continue from state i at time t, so that $x_i(t+1) = x_t + q_i(t) < x_i$. Otherwise we continue with probability $(x_i - x_i(t))/q_i(t)$, achieving $x_i(t+1) = x_i$, and we set h_i accordingly at $t-1+(x_i - x_i(t))/q_i(t)$.

If it happens that $x_i(t) < x_i$ for all t then we set $h_i = \infty$.

Let us observe that the expected number of times we *stop* at state j can never exceed τ_j, because the expected number of times j is entered by time t is

$$\sigma_j + \sum_i x_i(t-1) \leq \sigma_j + \sum_i x_i$$

which is the corresponding quantity for the greedy stopping rule; and j is exited at every opportunity until x_j is attained. However, Θ_τ is finite (since the sum of its exit times is bounded) and must does achieve an ending distribution. Thus $\sigma^{\Theta_\tau} = \tau$; and Θ_τ is mean-optimal by Theorem 2. $\square$

Note that the proof gives an algorithm to compute the threshold function from the given target distribution. The algorithm is polynomial in the number of states (which is unfortunately not good enough in a typical application of these techniques to sampling, where the size of the state space is exponential.)

It turns out that that Θ_τ is bounded whenever τ has sufficient support.

Theorem 6.2 *Suppose there is no closed walk*

$$w = w_0, w_1, \ldots, w_t, w_0$$

of length $t > 1$ for which $\tau_{w_i} = 0$ for each i. Then Θ_τ is bounded.

The threshold rule is special among all rules in the following sense:

Theorem 6.3 *The threshold rule Θ_τ is max-optimal.*

Let us note that the filling rule Φ can also be used to define thresholds g_i, but for use in the opposite manner: we stop at state i if we hit it *before* its threshold, so that a halting state j gets $g_j = \infty$. For clarity, however, we choose to reserve the term "threshold rule" for for the stopping rule Θ described above.

The continuous-time versions of both Φ and Θ enjoy more elegant threshold descriptions: in the first case we stop whenever we transfer to a state ahead of its threshold, in the second whenever a threshold is reached while we sit in the corresponding state. In both cases the thresholds themselves are uniquely defined.

7 Blind Rules

We now turn to a type of stopping rule which is not generally capable of achieving arbitrary target distributions, and is almost never optimal. We call a stopping rule Γ *blind* if it is balanced and $\Gamma(t; i)$ depends only on t.

The simplest blind stopping rule is the stopping rule used most often: "stop after t steps." Equal practicality and better convergence guarantees can be obtained from *lazy* random walks, considered e.g. in Lovász and Simonovits [17]. The lazy version of a Markov chain is obtained by flipping a coin before each move and staying where we are if we see "heads." Stopping the lazy version of a Markov chain after T steps is equivalent to a blind rule for the original chain in which we stop after t steps with probability

$$\binom{T}{t} \Big/ \left(\binom{T}{t} + \binom{T}{t+1} + \ldots + \binom{T}{T} \right) .$$

It is often more convenient to describe a blind stopping rule by the probabilities a_t that v_t is selected (not conditioning on not having stopped before). Trivially $a_t \geq 0$ and the finite termination of the rule is equivalent to $\sum_t a_t = 1$. The rule is bounded if and only if the sequence a_t has a finite number of non-0 terms. Thus a blind rule can always be thought of as an averaging, using the distribution $\{a_t\}$.

One cannot generate any distribution by a blind stopping rule; for example, starting from the stationary distribution, every blind rule generates the stationary distribution itself. We shall restrict our attention to stopping rules generating the stationary distribution (or at least approximations of it).

Let $\lambda_1, \ldots, \lambda_q$ be the eigenvalues of M, $\lambda_1 = 1$. From the Perron-Frobenius Theorem we know that $|\lambda_i| \leq 1$. Our next theorem gives a characterization for the existence of a blind stopping rule for the stationary distribution.

Theorem 7.1 *(a) If λ_k is positive real for some $k \geq 2$, then there exists a state s from which no blind stopping rule can generate π.*

(b) If every λ_k, $k \geq 2$, is either non-real, negative or zero, then from any starting distribution σ there is a finite blind stopping rule that which generates π.

Interestingly, the condition formulated in the theorem is most restrictive for time-reversible chains; then all the eigenvalues are real, and typically many of them are positive. For example, for random walks on graphs, only complete multipartite graphs give a spectrum with just one positive eigenvalue. More generally, it is easy to argue that for any two states i and j that are not "twins" (i.e., $p_{ik} \neq p_{jk}$ or $p_{ki} \neq p_{kj}$ for some k) we must have $p_{ij} > 0$. (If the chain is symmetric and $p_{ii} = 0$ for all i, then the condition is equivalent to saying that states can be represented by points in some euclidean space so that the distance between i and j is $\sqrt{p_{ij}}$.)

It is virtually never possible to find a blind rule for generating π which is mean-optimal. If we consider only stopping at some fixed time, even approximating π may take far more time than a mean-optimal rule (for example when the chain is almost periodic). However, the following theorem shows that we can get within ε of π (from below) with a fast blind rule.

Let $\varepsilon > 0$, put $k = \lceil \log(1/\varepsilon) \rceil$, and let $t = 8\lceil T_{\text{mix}} \rceil$. Let Z be the sum of k independent random variables $Y_1, \ldots, Y_k$ distributed uniformly in $\{0, \ldots, t-1\}$. Finally, let Γ_ε be the blind rule which stops the chain at time Z.

Theorem 7.2 *For any starting distribution σ, the rule Γ_ε produces a distribution τ satisfying*

$$\tau \geq (1-\varepsilon)\pi,$$

and has mean length $O(T_{\text{mix}} \log(1/\varepsilon))$.

We note that similar convergence speed can be achieved using other "averaging" rules, e.g., the Poisson stopping time of Aldous [2].

8 Exact Mixing in an Unknown Chain

In the previous section we saw that there is at least one practical stopping rule for approximating the stationary distribution when all that is known about the chain is its mixing time (or a good upper bound for it). Even this requirement can be circumvented: Aldous [4] gives a rule that comes within total variation ε of the stationary distribution in expected time $O(h/\varepsilon^2$ where $h = \max_{i,j} H(i,j)$ is the maximum hitting time of the chain, provided the number n of states is known.

Amazingly, one can even achieve the the stationary distribution *exactly* in an unknown chain. A stopping rule for doing so is described in Asmussen et al. [5], but it is complex and requires perfect generation of random variables with certain exponential distributions. The expected number of steps required appears to be super-polynomial in the maximum hitting time, although no bound or estimate is given in the paper.

It turns out, however, that there is a simple stopping rule Γ which can reach the stationary distribution exactly, in any irreducible, n-state Markov chain; the rule finishes in time polynomial in h, requires only coin-flips for its randomization, and can even be made deterministic unless the chain itself is completely deterministic.

We describe Γ as follows. Given a state i and a positive integer t, a *t-exit* of i consists of waiting X steps after being in state i, where X is uniformly chosen from $\{0, 1, \ldots, t-1\}$, then setting i' equal to the resulting state. A *t-pass* entails selecting a state j from the uniform distribution on S and then performing independent t-exits from every state i *except* j; a pass is deemed *successful* if the directed graph consisting of the $n-1$ arcs (i, i') has no loop or cycle (and is thus an in-directed spanning tree rooted at j).

The stopping rule Γ performs $3n$ t-passes, doubles t, and repeats (starting, say, from $t = 2$). As soon as the first successful pass is completed, Γ proceeds to the next occurrence of the state j which was the root of the successful pass, and stops.

Theorem 8.1 *Γ runs in expected number of steps bounded by $O(h^3 \log h)$ and stops at state j with probability exactly π_j.*

Proof. In outline: to show that $\sigma^\Gamma = \pi$, we show that for any t, the sum of the probabilities of the exit-trees rooted at a state j is proportional to π_j. In general the probability that a t-pass will succeed is exponentially small if t is small, but approaches $1/n$ (regardless of π) as t grows past nh, assuring polynomial stopping time.

We can give a better bound of $O(hT_{\text{mix}} n \log n)$ on the expected length of Γ. Here h majorizes the parameters n and T_{mix}, and can in fact be much larger (e.g. if we do random walk on an expander graph). One might ask if we could do better and and get rid of the dependence of h. The following simple observation shows that this is not possible.

Proposition 8.2 *Let the number n of states be fixed. Consider any stopping rule that works for an unknown Markov chain (i.e., it decides when to stop based on the pattern of the states seen before), and assume that for every Markov chain on n states, the distribution of the state where it stops is the stationary distribution. Then it never stops without visiting all states.*

It follows that the cover time, and hence also the maximum hitting time h, is a lower bound on the expected number of steps of such a rule.

References

[1] D.J. Aldous, *Reversible Markov Chains and Random Walks on Graphs* (book), to appear.

[2] D.J. Aldous, On the Markov chain simulation method for uniform combinatorial distributions and simulated annealing, *Probability in the Engineering and Informational Sciences,* **1** (1987), 33–46.

[3] D.J. Aldous, Applications of random walks on graphs, preprint (1989).

[4] D.J. Aldous, On simulating a Markov chain stationary distribution when transition probabilities are unknown, prepint (1993).

[5] S. Asmussen, P. W. Glynn and H. Thorisson, Stationary detection in the initial transient problem, *ACM Transactions on Modeling and Computer Simulation* **2** (1992), 130–157.

[6] D.J. Aldous, L. Lovász and P. Winkler, Fast mixing in a Markov chain, in preparation.

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

[8] R.V. Chacon and D.S. Ornstein, A general ergodic theorem, *Illinois J. Math.* **4** (1960), 153-160.

[9] A.K. Chandra, P. Raghavan, W.L. Ruzzo, R. Smolensky, and P. Tiwari, The Electrical Resistance of a Graph Captures its Commute and Cover Times, *Proceedings of the 21st Annual ACM Symposium on Theory of Computing*, May (1989).

[10] D. Coppersmith, P. Tetali, and P. Winkler, Collisions among Random Walks on a Graph, *SIAM J. on Discrete Mathematics* **6** No. 3 (1993), 363-374.

[11] P.G. Doyle and J.L. Snell, *Random Walks and Electric Networks*, Mathematical Assoc. of America, Washington, DC 1984.

[12] M. Dyer, A. Frieze and R. Kannan, A random polynomial time algorithm for estimating volumes of convex bodies, *Proc. 21st Annual ACM Symposium on the Theory of Computing* (1989), 375-381.

[13] M. Jerrum and A. Sinclair, Conductance and the rapid mixing property for Markov chains: the approximation of the permanent resolved, *Proc. 20nd Annual ACM Symposium on Theory of Computing* (1988), 235-243.

[14] A. Karzanov and L. Khachiyan, On the conductance of order Markov chains, Technical Report DCS TR 268, Rutgers University, June 1990.

[15] L. Lovász and P. Winkler, Exact mixing in an unknown Markov chain, in preparation.

[16] L. Lovász and P. Winkler, The forget time of a Markov chain, in preparation.

[17] L. Lovász and M. Simonovits, Random walks in a convex body and an improved volume algorithm, *Random Structures and Alg.* **4** (1993), 359–412.

[18] M. Mihail and P. Winkler, On the number of Eulerian orientations of a graph, *Proc. 3rd ACM-SIAM Symposium on Discrete Algorithms* (1992), 138-145.

[19] J.W. Pitman, Occupation measures for Markov chains, *Adv. Appl. Prob.* **9** (1977), 69-86.

[20] P. Tetali, Random walks and effective resistance of networks, *J. Theoretical Prob.* No. 1 (1991), 101-109.