Electron. Commun. Probab. **30** (2025), article no. 15, 1–13.
https://doi.org/10.1214/25-ECP663
ISSN: 1083-589X

ELECTRONIC COMMUNICATIONS in PROBABILITY

Equivalence of starting point cutoff and the concentration of hitting times on a general state space

David Ledvinka* Jeffrey S. Rosenthal†

Abstract

In this article we extend a result of Martinez and Ycart from their paper "Decay Rates and Cutoff for Convergence and Hitting Times of Markov Chains with Countably Infinite State Space" [7] to show that for a regular Markov chain on a general state space, the existence of a cutoff phenomenon for total variation distance to stationarity as the starting point tends to infinity is equivalent to the concentration of hitting times for any fixed regular set as the starting points tend to infinity. We apply this result to show that all random walks on the half-line with bounded steps exhibit starting point cutoff.

Keywords: Markov chains; cutoff phenomenon.
MSC2020 subject classifications: 60.
Submitted to ECP on September 10, 2023, final version accepted on February 2, 2025.

1 Introduction

In the study of Markov chains, two quantities which are often of interest are hitting times of certain sets and the time it takes to converge to stationarity. In particular one is often interested in whether a Markov chain exhibits a cutoff phenomenon (see [1]): the existence of a time where the total variation distance to stationarity transitions rapidly from near one to near zero. Theorem 4.1 in [7] shows that for continuous-time Markov chains on a countable state space, when the Markov chain is started arbitrarily far from a finite set S, the existence of a cutoff phenomenon is equivalent to the concentration of the hitting times of S, and in this case the hitting times are equivalent to the cutoff times. The primary goal of this paper is to establish a version of this result for discrete-time Markov chains on a general state space with the appropriate adjustments.

In Section 2 we establish our setting, notation and definitions and conclude by stating our main result (Theorem 2.6). The bulk of the proof of Theorem 2.6 is broken up into two propositions which are stated and proved in Section 4. In Section 3 we state and prove two lemmas which are the tools used to extend the proof ideas from [7] to the general state space setting. In Section 5 we use Theorem 2.6 to establish a sufficient condition for starting point cutoff which is applied in Section 6 to show that all regular random walks on the half-line with bounded steps exhibit starting point cutoff. In Section 7 we construct a Markov chain and a sequence of starting points that diverge to infinity but do not exhibit starting point cutoff, and use Theorem 2.6 to justify that it works.

*Department of Mathematics, University of Toronto, Canada. E-mail: david.ledvinka@mail.utoronto.ca
†Department of Statistical Sciences, University of Toronto, Canada. E-mail: jeff@math.toronto.edu

Finally in Section 8 we state and discuss some open problems inspired by the results of this paper. A web appendix can be found at probability.ca/CutoffWeb where we give a detailed proof of starting point cutoff for a toy example.

2 Definitions and statement of equivalence

Throughout the paper we let P be a φ-irreducible, aperiodic Markov kernel with stationary distribution π on a countably generated state space $(\mathcal{X}, \mathcal{B})$. Given $x \in \mathcal{X}$ we can define a Markov chain $\{X_t\}_{t\geq 0}$ starting at x with Markov kernel P by defining $X_0 = x$ and for $t \geq 1$ recursively setting $X_t \sim P(X_{t-1}, \cdot)$. We then denote the law of $\{X_t\}_{t\geq 0}$ by $\mathbb{P}_x$ and expectation with respect to this measure by $\mathbb{E}_x$. Note that for all $t \geq 0$ and $A \in \mathcal{B}$,

$$\mathbb{P}_x(X_t \in A) = P^t(x, A)$$

where P^t is the convolution of the kernel P with itself t times. For any $S \in \mathcal{B}$ define the hitting time τ_S of S as the random variable

$$\tau_S = \inf\{t \geq 0 : X_t \in S\}$$

Note that τ_S is a stopping time with respect to the filtration $\{\sigma(X_0, \ldots, X_t)\}_{t\geq 0}$. Define $\mathcal{B}^+ := \{A \in \mathcal{B} : \pi(A) > 0\}$. We make the additional assumption that P is regular in the sense that for any $A \in \mathcal{B}^+$ and all $x \in \mathcal{X}$, $\mathbb{E}_x[\tau_A] < \infty$. In fact this is not a strong additional assumption since it can be shown that a φ-irreducible, aperiodic Markov kernel with stationary distribution π can be restricted to a regular kernel by throwing out a set of π-measure zero (see [8]).

In order to generalize Theorem 4.1 of [7] to the general state space setting we will need a candidate replacement for the role of finite sets. It turns out that for regular kernels there exists a natural replacement.

Definition 2.1. *We call $S \in \mathcal{B}^+$ a regular set if for any $A \in \mathcal{B}^+$*

$$\sup_{x\in S} \mathbb{E}_x[\tau_A] < \infty$$

In Chapter 11 of [8] it is shown that for φ-irreducible regular kernels there exists regular sets. Note that in the discrete state space setting, finite sets will be regular sets since the supremum becomes a maximum (though infinite sets may or may not be regular).

Now fix a set $S \in \mathcal{B}^+$.

Definition 2.2. *Let x_n be a sequence of starting points in $\mathcal{X}$ and let t_n be an increasing sequence of positive reals such that $\lim_{n\to\infty} t_n = \infty$. We say that τ_S concentrates at time t_n if*

$$\lim_{n\to\infty} \mathbb{P}_{x_n}\left(\left|\frac{\tau_S}{t_n} - 1\right| \leq \varepsilon\right) = 1, \quad \textit{for all } \varepsilon > 0$$

In other words $\tau_S/t_n \to 1$ in probability.

An equivalent definition is given in the following proposition whose proof we omit.

Proposition 2.3. *Let x_n be a sequence of starting points in $\mathcal{X}$ and let t_n be an increasing sequence of positive reals such that $\lim_{n\to\infty} t_n = \infty$. Then τ_S concentrates at time t_n if and only if the following two conditions hold:*

(i) For any $c < 1$, $\lim_{n\to\infty} \mathbb{P}_{x_n}(\tau_S < ct_n) = 0$.

(ii) For any $c > 1$, $\lim_{n\to\infty} \mathbb{P}_{x_n}(\tau_S < ct_n) = 1$.

While the characterization of concentration of hitting times given in Proposition 2.3 may seem less natural, its usefulness comes from its similarity to the definition of the cutoff phenomenon given in Definition 2.5 below. We now turn our discussion to the convergence of $P^t(x, \cdot)$ to the stationary distribution π. To measure distance between these distributions we use the total variation distance.

Definition 2.4. *For any $x \in \mathcal{X}$ and $t \in \mathbb{R}^+$ define*

$$d_x(t) \coloneqq \left\| P^{\lfloor t \rfloor}(x, \cdot) - \pi \right\|_{\mathrm{TV}} = \sup_{A \in \mathcal{B}} \left| P^{\lfloor t \rfloor}(x, A) - \pi(A) \right|$$

We say that $d_x(t)$ is the (total variation) distance to stationarity starting at x after time t. While this makes more sense when t is a non-negative integer, defining it for any $t \in \mathbb{R}^+$ will be convenient for defining the cutoff phenomenon. Since P is a regular aperiodic Markov kernel, it is also an aperiodic Harris recurrent Markov kernel (see Chapter 11 of [8]) and therefore $\lim_{n \to \infty} d_x(t) = 0$ (see Chapter 13 of [8]). This is sometimes called the Harris Ergodic Theorem. While the Harris Ergodic Theorem tells you that a Markov chain converges to its stationary distribution, it gives no qualitative or quantitative information about this convergence. The cutoff phenomenon (see [1] for a review) is the observation that for many natural Markov chains (though not all) the total variation distance to stationarity stays very close to one until it reaches some time (the cutoff time) where it rapidly decays to zero. There are many variations for the definition of cutoff depending on the setting and problem of interest. Often (as in the case in [1]) a sequence of Markov chains on a sequence of finite state spaces are considered and cutoff is a statement about the worst case starting points. In the general state space setting there may not be a worse case starting point, since for any fixed set $A \in \mathcal{B}^+$ there may exist a sequence of starting points x_n that start arbitrarily far from A (more precisely meaning that $\lim_{n\to\infty} \mathbb{E}_{x_n}[\tau_A] = \infty$). Instead we consider "starting point cutoff", similar to the notion in [7] where we fix a Markov kernel and let the starting point vary.

Definition 2.5. *Let x_n be a sequence of starting points in $\mathcal{X}$ and let t_n be an increasing sequence of positive reals such that $\lim_{n\to\infty} t_n = \infty$. We say that P has (starting point) cutoff at time t_n starting from x_n if*

(i) For any $c < 1$, $\lim_{n\to\infty} d_{x_n}(ct_n) = 1$.

(ii) For any $c > 1$, $\lim_{n\to\infty} d_{x_n}(ct_n) = 0$.

Intuitively, if P has cutoff at time t_n starting from x_n, it means that for sufficiently large n, if we start a Markov chain with Markov kernel P at x_n then its distance to stationarity will be close to one for any time before t_n and close to zero at any time after t_n.

We can now state our main theorem which says that starting point cutoff is equivalent to concentration of hitting times on regular sets.

Theorem 2.6. *Let x_n be a sequence of starting points in $\mathcal{X}$ and let t_n be an increasing sequence of positive reals such that $\lim_{n\to\infty} t_n = \infty$. Let S be a regular set. Then P has cutoff at time t_n starting from x_n if and only if τ_S concentrates at time t_n starting from x_n.*

Proof. Using the characterization of concentration of hitting times given in Proposition 2.3 it is sufficient to prove that for regular sets:

a) (i) from Proposition 2.3 is equivalent to (i) from Definition 2.5.

b) (ii) from Proposition 2.3 is equivalent to (ii) from Definition 2.5.

a) is proved in Proposition 4.1 and b) is proved in Proposition 4.2. □

3 Lemmas for a general state space

In this section we prove two lemmas which allow us to extend the ideas from the proof of Theorem 4.1 in [7] to the general state space setting of Theorem 2.6. In the context of Theorem 2.6 we are considering a fixed regular set S and a sequence of starting points x_n. With this context in mind our first lemma says that although the time it takes the Markov chain to hit S starting from x_n may depend on n, once the Markov chain reaches S it will converge to the stationary distribution (up to ε error) in a constant time T.

Lemma 3.1. *For any regular set S, $\lim_{t\to\infty} d_x(t) = 0$ uniformly on S. More precisely for any $\varepsilon \in (0,1)$ there exists $T \in \mathbb{N}$ such that for all $t \geq T, x \in S$ and $A \in \mathcal{B}$, we have that $|P^t(x,A) - \pi(A)| < \varepsilon$.*

In the discrete state space setting where S is replaced with a finite set, this lemma trivially follows from pointwise convergence. For the general state space setting the informal idea of the proof is that (by Egoroff's theorem) we know that there exists some set $B \in \mathcal{B}^+$ such that $\lim_{t\to\infty} d_x(t) = 0$ uniformly on B and we want to transfer this uniform convergence to S. We do this by noting that since S is regular there exists a constant T_0 such that starting anywhere in S, the Markov chain will hit B by time T_0 (with probability $1-\varepsilon$) and since $\lim_{t\to\infty} d_x(t) = 0$ uniformly on B there will exist some time T_1 such that starting anywhere in B the Markov chain will converge to the stationary distribution (up to ε error) in constant time T_1. Therefore starting from anywhere in S the Markov chain will hit B by time T_0 and then converge to the stationary distribution by time $T_0 + T_1$ (with 2ε error which can be made arbitrarily small).

Proof. Since $\mathcal{B}$ is countably generated, for each fixed $t \in \mathbb{N}$, $d_x(t) : \mathcal{X} \to \mathbb{R}$ is measurable (see the Appendix of [9]). Therefore since for any $x \in \mathcal{X}$, $\lim_{t\to\infty} d_x(t) = 0$ by Egoroff's theorem (Theorem 2.33 in [4]) there exists $B \in \mathcal{B}^+$ such that $d_x(t)$ converges to 0 uniformly on B. Since S is a regular set there exists M such that for all $x \in S$, $\mathbb{E}_x[\tau_B] \leq M$.

Let $\varepsilon \in (0,1)$. If we define $T_0 \coloneqq \lceil \frac{M}{\varepsilon} \rceil$, then for all $x \in S$ and $t \geq T_0$

$$\begin{aligned}
\mathbb{P}_x(\tau_B > T_0) &\leq \frac{\mathbb{E}_x[\tau_B]}{(M/\varepsilon)} \quad \text{by Markov's inequality} \\
&\leq \varepsilon
\end{aligned}$$

Since $d_x(t)$ converges to 0 uniformly on B there exists T_1 sufficiently large such that for all $x \in B$ and $t \geq T_1$, $d_x(t) \leq \varepsilon$.

Define $T \coloneqq T_0 + T_1$. Then for any $x \in S$ and $A \in \mathcal{B}$, if $t \geq T$ then

$$\begin{aligned}
&|P^t(x,A) - \pi(A)| \\
&= \left| \sum_{k=0}^{\infty} [\mathbb{P}_x(X_t \in A | \tau_B = k) - \pi(A)] \mathbb{P}_x(\tau_B = k) \right| \\
&\leq \sum_{k=0}^{\infty} |\mathbb{P}_x(X_t \in A | \tau_B = k) - \pi(A)| \, \mathbb{P}_x(\tau_B = k) \\
&\leq \mathbb{P}_x(\tau_B > T_0) + \sum_{k=0}^{T_0} |\mathbb{P}_x(X_t \in A | \tau_B = k) - \pi(A)| \, \mathbb{P}_x(\tau_B = k) \\
&\leq \varepsilon + \sum_{k=0}^{T_0} \sup_{y \in B} \left| P^{t-k}(y,A) - \pi(A) \right| \mathbb{P}_x(\tau_B = k)
\end{aligned}$$

by the strong Markov property

$$
\begin{aligned}
&\le \varepsilon + \sum_{k=0}^{T_0} \sup_{y\in B} d_y(t-k)\mathbb{P}_x(\tau_B = k) \\
&\le \varepsilon + \sum_{k=0}^{T_0} \varepsilon \mathbb{P}_x(\tau_B = k) \quad \text{since } t-k \ge T_1 \\
&= \varepsilon + \varepsilon \mathbb{P}_x(\tau_B \le T_0) \\
&\le 2\varepsilon
\end{aligned}
$$

Therefore (since A was arbitrary) for any $x \in S$ and $t \ge T$,

$$d_x(t) \le 2\varepsilon$$

Since ε was arbitrary this proves

$$\lim_{t\to\infty} d_x(t) = 0$$

uniformly on S. □

Our second lemma says that for any $S \in \mathcal{B}^+$ (not necessarily regular) there exists arbitrarily large sets B (in π-measure) for which there exists a constant time $T \in \mathbb{N}$ such that starting anywhere in B the Markov chain will hit S by time T (with arbitrarily large probability). Again in the context of Theorem 2.6 the point is that while the time it takes the Markov chain to hit B starting from x_n may depend on n, once you hit B the Markov chain will hit S in constant time T (with large probability).

Lemma 3.2. *For any $S \in \mathcal{B}^+$ and $\varepsilon \in (0,1)$ there exists $B \in \mathcal{B}$ and $T \in \mathbb{N}$ such that $\pi(B) > 1-\varepsilon$ and $\sup_{x\in B} \mathbb{P}_x(\tau_S \ge T) \le \varepsilon$.*

Proof. Since P is Harris recurrent, for all $x \in \mathcal{X}$, $\mathbb{E}_x[\tau_S] < \infty$. Thus if we define for each $d \in \mathbb{N}$, $A_d = \{x \in \mathcal{X} : \mathbb{E}_x[\tau_S] \le d\}$, A_d is an increasing sequence of measurable sets whose union is all of $\mathcal{X}$. Therefore by continuity of measure there exists D sufficiently large such that $\pi(A_D) > 1-\varepsilon$ and setting $T = \lceil D/\varepsilon \rceil$ we have that

$$
\begin{aligned}
\sup_{x\in A_D} \mathbb{P}_x(\tau_S \ge T) &\le \sup_{x\in A_D} \mathbb{P}_x(\tau_S \ge D/\varepsilon) \\
&\le (\varepsilon/D) \sup_{x\in A_D} \mathbb{E}_x[\tau_S] \quad \text{by Markov's inequality} \\
&\le \varepsilon \quad \text{by definition of } A_D
\end{aligned}
$$

□

4 Proof of equivalence

As stated in Section 2, Theorem 2.6 follows from Proposition 4.1 and Proposition 4.2 which are stated and proved in this section. It is worth noting however that these propositions prove something slightly stronger, as one direction of the equivalence in each proposition holds regardless of whether the set S is regular (as long as it is in $\mathcal{B}^+$). The proof of both propositions is based heavily on the proof of Theorem 4.1 in [7], but uses the lemmas proved in Section 3 to extend it to a general state space. For each proof direction we give a general overview of the main ideas and intuitions that will be used in the proof before giving the proof in full technical detail.

Proposition 4.1. *Let x_n be a sequence in $\mathcal{X}$ and let t_n be an increasing sequence of positive reals such that $\lim_{n\to\infty} t_n = \infty$. Let $S \in \mathcal{B}^+$ and consider the following conditions:*

(i) For any $c < 1$, $\lim_{n\to\infty} \mathbb{P}_{x_n}(\tau_S < ct_n) = 0$.

(ii) For any $c < 1$, $\lim_{n\to\infty} d_{x_n}(ct_n) = 1$.

If (i) *holds then* (ii) *holds and if* S *is a regular set then* (i) *and* (ii) *are equivalent.*

Proof idea for (i) implies (ii). It is easy to see that (i) implies the weaker statement that for any $c < 1$, $\liminf_{n\to\infty} d_{x_n}(ct_n) \geq \pi(S)$. This is because the probability that the Markov chain starting at x_n is in S at time $\lfloor ct_n \rfloor$ is at most the probability that the Markov chain starting at x_n has hit S by time ct_n. Therefore if the the latter probability vanishes then so does the former and this provides a lower bound on total variation distance to stationarity. To get the stronger statement in (ii) we want to show that we can apply the same argument above to arbitrarily large sets in π-measure. This is where we can use Lemma 3.2 to assert that there exists arbitrarily large sets B and a constant time $T \in \mathbb{N}$ where once the Markov chain has hit B it will hit S in time T (with arbitrarily large probability). Therefore if the Markov chain hits B by time ct_n it will hit S in time $ct_n + T$ (with large probability) which is asymptotically less than $\frac{c+1}{2} t_n$. Therefore since (i) asserts that the probability the Markov chain hits S starting at x_n by time $\frac{c+1}{2} t_n$ is vanishingly small, then the probability that the Markov chain hits B starting at x_n by time ct_n must also be vanishingly small. □

Proof of (i) *implies* (ii). Suppose (i) holds. Let $c < 1$. Let $\varepsilon \in (0, 1)$. By Lemma 3.2 there exists a set $B \in \mathcal{B}$ and $T \in \mathbb{N}$ such that $\pi(B) > 1 - \varepsilon$ and $\sup_{x\in B} \mathbb{P}_x(\tau_S \geq T) \leq \varepsilon$. Since $\lim_{n\to\infty} t_n = \infty$ and $c < 1$ we can choose N_0 sufficiently large such that for all $n \geq N_0$, $T < \frac{1-c}{2} t_n$. Since $\lim_{n\to\infty} \mathbb{P}_{x_n}(\tau_S < \frac{c+1}{2} t_n) = 0$ we can choose N_1 sufficiently large such that for all $n \geq N_1$, $\mathbb{P}_{x_n}(\tau_S < \frac{c+1}{2} t_n) \leq \varepsilon$. Define $N \coloneqq \max\{N_0, N_1\}$. Then for all $n \geq N$

$$
\begin{aligned}
&P^{\lfloor ct_n \rfloor}(x_n, B) \\
&\leq \mathbb{P}_{x_n}(\tau_B \leq ct_n) \\
&= \mathbb{P}_{x_n}(\tau_S \leq (\tau_S - \tau_B) + ct_n) \\
&\leq \mathbb{P}_{x_n}(\tau_S \leq T + ct_n) + \mathbb{P}_{x_n}(\tau_S - \tau_B \geq T) \quad \text{by a union bound} \\
&\leq \mathbb{P}_{x_n}(\tau_S \leq T + ct_n) + \sup_{x\in B} \mathbb{P}_x(\tau_S \geq T) \quad \text{by the strong Markov property} \\
&\leq \mathbb{P}_{x_n}\left(\tau_S \leq \frac{c+1}{2} t_n\right) + \varepsilon \quad \text{since } n \geq N_0 \text{ and by the definition of } B \\
&\leq \varepsilon + \varepsilon \quad \text{since } n \geq N_1 \\
&= 2\varepsilon
\end{aligned}
$$

therefore for all $n \geq N$

$$d_{x_n}(ct_n) \geq \pi(B) - P^{\lfloor ct_n \rfloor}(x_n, B) \geq 1 - 3\varepsilon$$

Since ε was arbitrary this proves

$$\lim_{n\to\infty} d_{x_n}(ct_n) = 1$$

□

Proof idea for (ii) *implies* (i) *when* S *is a regular set.* By Lemma 3.1 we can choose some T such that for all $t \geq T$, $x \in S$ and $A \in \mathcal{B}$, we have that $|P^t(x, A) - \pi(A)| < 1/4$. This implies that if S is hit by time ct_n (starting from x_n) then for any time $t \geq T$ we have that the probability the Markov chain is in A at time $\lfloor ct_n \rfloor + t$ is within $1/4$ of $\pi(A)$. In particular since $\lfloor \frac{c+1}{2} t_n \rfloor$ is asymptotically greater than $\lfloor ct_n \rfloor + T$ we can use this to assert that for any $A \in \mathcal{B}$ with $\pi(A) > 1/2$, $P^{\lfloor \frac{c+1}{2} t_n \rfloor}(x_n, A) \geq \mathbb{P}_{x_n}(\tau_S < ct_n)(\pi(A) - 1/4) >$

$\mathbb{P}_{x_n}(\tau_S < ct_n)(1/4)$. But condition (ii) implies that there exists a sequence $A_n \in B$ for which $\pi(A_n)$ is arbitrarily close to 1 and $P^{\lfloor \frac{c+1}{2} \rfloor}(x_n, A_n)$ is arbitrarily close to 0. Plugging these A_n into the above inequality implies (i). □

Proof of (ii) implies (i) when S is a regular set. Suppose S is a regular set and suppose (ii) holds. Let $c < 1$. Let $\varepsilon \in (0, 1/2)$. Since $\lim_{n\to\infty} d_{x_n}(\frac{c+1}{2} t_n) = 1$ there exists N_0 sufficiently large such that for each $n \ge N_0$, there exists $A_n \in \mathcal{B}$ such that $\pi(A_n) > 1 - \varepsilon$ and $P^{\lfloor \frac{c+1}{2} t_n \rfloor}(x_n, A_n) < \varepsilon$. By Lemma 3.1 we can choose T sufficiently large such that for any $x \in S$ and $t \ge T$, $d_x(t) < 1/4$. In particular this implies for any $n \ge N_0$ and $t \ge T$ that $\inf_{x\in S} P^t(x, A_n) > \pi(A_n) - 1/4 > 1/4$. Furthermore since $\lim_{n\to\infty} t_n = \infty$ and $c < 1$ we can choose N_1 sufficiently large such that for all $n \ge N_1$ and $t \le \lfloor ct_n \rfloor$ we have $\lfloor \frac{c+1}{2} t_n \rfloor - t \ge T$. Define $N = \max\{N_0, N_1\}$. Then for any $n \ge N$ we have

$$\begin{aligned}
&\mathbb{P}_{x_n}(\tau_S < ct_n) \\
&\le \sum_{t=0}^{\lfloor ct_n \rfloor} \mathbb{P}_{x_n}(\tau_S = t) \\
&\le 4 \left(\sum_{t=0}^{\lfloor ct_n \rfloor} \inf_{x\in S} P^{(\lfloor \frac{c+1}{2} t_n \rfloor - t)}(x, A_n) \mathbb{P}_{x_n}(\tau_S = t) \right) \quad \text{since } n \ge N_0 \text{ and } n \ge N_1 \\
&\le 4 \left(\sum_{t=0}^{\lfloor ct_n \rfloor} \mathbb{P}_{x_n}(X_{\lfloor \frac{c+1}{2} t_n \rfloor} \in A_n \mid \tau_S = t) \mathbb{P}_{x_n}(\tau_S = t) \right) \quad \text{by the strong Markov property} \\
&= 4\mathbb{P}_{x_n}(X_{\lfloor \frac{c+1}{2} t_n \rfloor} \in A_n, \tau_S \le \lfloor ct_n \rfloor) \\
&\le 4P^{\lfloor \frac{c+1}{2} t_n \rfloor}(x_n, A_n) \\
&\le 4\varepsilon
\end{aligned}$$

Since ε was arbitrary, this proves

$$\lim_{n\to\infty} \mathbb{P}_{x_n}(\tau_S < ct_n) = 0$$ □

Proposition 4.2. *Let x_n be a sequence in $\mathcal{X}$ and let t_n be an increasing sequence of positive reals such that $\lim_{n\to\infty} t_n = \infty$. Let $S \in \mathcal{B}^+$ and consider the following conditions:*

(i) For any $c > 1$, $\lim_{n\to\infty} \mathbb{P}_{x_n}(\tau_S < ct_n) = 1$.

(ii) For any $c > 1$, $\lim_{n\to\infty} d_{x_n}(ct_n) = 0$.

If (ii) holds then (i) holds and if S is a regular set then (ii) and (i) are equivalent.

Proof idea for (ii) implies (i). It is easy to see that (ii) implies the weaker statement that for any $c > 1$, $\limsup_{n\to\infty} \mathbb{P}_{x_n}(\tau_S < ct_n) \ge \pi(S)$. This is because the probability that the Markov chain starting at x_n has hit S by time ct_n is at least the probability that the Markov chain starting from x_n is in S at time $\lfloor ct_n \rfloor$. Therefore we get the lower bound by noticing that (ii) implies that the probability that the Markov chain starting from x_n is in S at time $\lfloor ct_n \rfloor$ can be made arbitrarily close to $\pi(S)$. Of course this argument applies to any set $B \in \mathcal{B}$, and for B with large π-measure, it shows that $\limsup_{n\to\infty} \mathbb{P}_{x_n}(\tau_B < ct_n)$ must be close to 1. Therefore we again use Lemma 3.2 to find arbitrarily large sets B with constant times $T \in \mathbb{N}$ such that once the Markov chain hits B it will hit S in time T (with arbitrarily large probability). Hence if the Markov chain hits B by time $\lfloor \frac{c+1}{2} t_n \rfloor$ (with arbitrarily large probability) which is asymptotically less than $ct_n - T$, it will hit S by time ct_n (with arbitrarily large probability). □

Proof of (ii) *implies* (i). Suppose (ii) holds. Let $c > 1$. Let $\varepsilon \in (0, 1/2)$. By Lemma 3.2 there exists a set $B \in \mathcal{B}$ and $T \in \mathbb{N}$ such that $\pi(B) > 1 - \varepsilon$ and $\sup_{x \in B} \mathbb{P}_x(\tau_S \geq T) \leq \varepsilon$. Since $\lim_{n\to\infty} d_{x_n}(\frac{c+1}{2}t_n) = 0$, we can choose N_0 sufficiently large such that for all $n \geq N_0$, $d_{x_n}(\frac{c+1}{2}t_n) < \varepsilon$. In particular this implies that for all $n \geq N_0$, $\mathbb{P}_{x_n}(\tau_B < \frac{c+1}{2}t_n) \geq P^{\lfloor \frac{c+1}{2}t_n \rfloor}(x_n, B) \geq \pi(B) - \varepsilon > 1 - 2\varepsilon$. Since $\lim_{n\to\infty} t_n = \infty$ and $c > 1$, we can choose N_1 sufficiently large such that for all $n \geq N_1$ we have $ct_n - \lfloor \frac{c+1}{2}t_n \rfloor > T$. Define $N = \max\{N_0, N_1\}$. Then for any $n \geq N$,

$$
\begin{aligned}
&\mathbb{P}_{x_n}(\tau_S < ct_n) \\
&\geq \mathbb{P}_{x_n}(\tau_S^B < ct_n) \quad \text{where } \tau_S^B = \inf\{t \geq \tau_B : X_t \in S\} \\
&\geq \mathbb{P}_{x_n}\left(\tau_S^B < ct_n, \tau_B \leq \left\lfloor \frac{c+1}{2}t_n \right\rfloor\right) \\
&= \sum_{t=0}^{\lfloor \frac{c+1}{2}t_n \rfloor} \mathbb{P}_{x_n}(\tau_S^B < ct_n \mid \tau_B = t)\mathbb{P}_{x_n}(\tau_B = t) \\
&\geq \sum_{t=0}^{\lfloor \frac{c+1}{2}t_n \rfloor} \inf_{y \in B} \mathbb{P}_y(\tau_S < ct_n - t)\mathbb{P}_{x_n}(\tau_B = t) \quad \text{by the Strong Markov property} \\
&\geq \sum_{t=0}^{\lfloor \frac{c+1}{2}t_n \rfloor} \left(1 - \sup_{y \in B} \mathbb{P}_y(\tau_S \geq T)\right)\mathbb{P}_{x_n}(\tau_B = t) \quad \text{since } n \geq N_1 \\
&\geq \sum_{t=0}^{\lfloor \frac{c+1}{2}t_n \rfloor} (1-\varepsilon)\mathbb{P}_{x_n}(\tau_B = t) \quad \text{since } n \geq N_1 \quad \text{by definition of B} \\
&= (1-\varepsilon)\mathbb{P}_{x_n}\left(\tau_B < \left\lfloor \frac{c+1}{2}t_n \right\rfloor\right) \\
&\geq (1-\varepsilon)(1-2\varepsilon) \quad \text{since } n \geq N_0
\end{aligned}
$$

Since ε was arbitrary this proves

$$\lim_{n\to\infty} \mathbb{P}_{x_n}(\tau_S < ct_n) = 1$$

□

Proof idea for (i) *implies* (ii) *when* S *is a regular set.* Fixing $c > 1$, to prove (ii) we need to show that for any $\varepsilon \in (0, 1)$ there exists $N \in \mathbb{N}$ such that for all $n \geq N$ and $A \in \mathcal{B}$ we have that $|P^{\lfloor ct_n \rfloor}(x_n, A) - \pi(A)| < \varepsilon$. Notice that this is very similar to what we needed to prove in Lemma 3.1 except we are starting from x_n. In fact the proof of (i) implies (ii) will be very similar to the proof of Lemma 3.1 except here S plays the role of B and the time it takes to reach S starting from x_n will depend on n. Condition (i) implies that for sufficiently large n, the Markov chain starting from x_n will reach S by time $\frac{c+1}{2}t_n$ (with probability $1 - \varepsilon$). Since $\lim_{t\to\infty} d_x(t) = 0$ uniformly on S (by Lemma 3.1) there exists some time T such that starting anywhere in S the Markov chain will converge to the stationary distribution (up to ε error) in constant time T. Therefore starting from x_n (for sufficiently large n) the Markov chain will hit S by time $\frac{c+1}{2}t_n$ which is asymptotically less than $\lfloor ct_n \rfloor - T$ and then converge to the stationary distribution by time $\lfloor ct_n \rfloor$ (with 2ε error which can be made arbitrarily small). □

Proof of (i) *implies* (ii) *when* S *is a regular set.* Suppose S is a regular set and suppose (i) holds. Let $c > 1$. Let $\varepsilon \in (0, 1)$. Since $\lim_{n\to\infty} \mathbb{P}_{x_n}(\tau_S \leq \frac{c+1}{2}t_n) = 1$, we can choose N_0 sufficiently large such that for all $n \geq N_0$, $\mathbb{P}_{x_n}(\tau_S \leq \frac{c+1}{2}t_n) > 1 - \varepsilon$. By Lemma 3.1 we can choose T sufficiently large such that for any $x \in S$ and $t \geq T$, $d_x(t) \leq \varepsilon$. Furthermore since $\lim_{n\to\infty} t_n = \infty$ and $c > 1$ we can choose N_1 sufficiently large such that for all

$n \geq N_1$ and $t \leq \lfloor \frac{c+1}{2} t_n \rfloor$ we have $\lfloor ct_n \rfloor - t \geq T$. Define $N = \max\{N_0, N_1\}$. Let $A \in \mathcal{B}$. Then for all $n \geq N$ we have

$$
\begin{aligned}
&|P^{\lfloor ct_n \rfloor}(x_n, A) - \pi(A)| \\
&= \left| \sum_{t=0}^{\infty} [\mathbb{P}_{x_n}(X_{\lfloor ct_n \rfloor} \in A | \tau_S = t) - \pi(A)] \mathbb{P}_{x_n}(\tau_S = t) \right| \\
&\leq \sum_{t=0}^{\infty} \left| \mathbb{P}_{x_n}(X_{\lfloor ct_n \rfloor} \in A | \tau_S = t) - \pi(A) \right| \mathbb{P}_{x_n}(\tau_S = t) \\
&\leq \mathbb{P}_{x_n}\left(\tau_S > \frac{c+1}{2} t_n \right) + \sum_{t=0}^{\lfloor \frac{c+1}{2} t_n \rfloor} \left| \mathbb{P}_{x_n}(X_{\lfloor ct_n \rfloor} \in A | \tau_S = t) - \pi(A) \right| \mathbb{P}_{x_n}(\tau_S = t) \\
&\leq \varepsilon + \sum_{t=0}^{\lfloor \frac{c+1}{2} t_n \rfloor} \sup_{x \in S} \left| P^{\lfloor ct_n \rfloor - t}(x, A) - \pi(A) \right| \mathbb{P}_{x_n}(\tau_S = t)
\end{aligned}
$$

since $n \geq N_0$ and by the strong Markov property

$$
\begin{aligned}
&\leq \varepsilon + \sum_{t=0}^{\lfloor \frac{c+1}{2} t_n \rfloor} \sup_{x \in S} d_x(\lfloor ct_n \rfloor - t) \mathbb{P}_{x_n}(\tau_S = t) \\
&\leq \varepsilon + \sum_{t=0}^{\lfloor \frac{c+1}{2} t_n \rfloor} \varepsilon \mathbb{P}_{x_n}(\tau_S = t) \quad \text{since } n \geq N_1 \\
&\leq \varepsilon + \varepsilon \mathbb{P}_{x_n}\left(\tau_S \leq \left\lfloor \frac{c+1}{2} t_n \right\rfloor \right) \\
&\leq 2\varepsilon
\end{aligned}
$$

Therefore (since A was arbitrary) for any $n \geq N$

$$d_{x_n}(ct_n) \leq 2\varepsilon$$

Since ε was arbitrary

$$\lim_{n \to \infty} d_{x_n}(ct_n) = 0 \qquad \square$$

5 A sufficient condition for cutoff

Theorem 2.6 reduces the problem of exhibiting a starting point cutoff phenomenon for a general state space Markov chain to the problem of showing a concentration of hitting times to some regular set S. In general this may be just as difficult to show, but in some cases this may be much easier. In this section we use Theorem 2.6 to establish a sufficient condition for starting point cutoff. We will use this result in Section 6 to show that any regular random walk on the half-line with bounded steps exhibits the starting point cutoff phenomenon.

Proposition 5.1. *Suppose there exists a constant C such that for all $x \in \mathcal{X} \setminus S$*

$$|\mathbb{E}_{X_1}[\tau_S] - \mathbb{E}_x[\tau_S]| \leq C \quad \textit{almost surely}$$

Then for any sequence of starting points x_n such that $\lim_{n\to\infty} \mathbb{E}_{x_n}[\tau_S] = \infty$ we have that X_t exhibits starting point cutoff at time $t_n = \mathbb{E}_{x_n}[\tau_S]$ starting from x_n.

In order to prove Proposition 5.1 (a qualitative result) we first prove a quantitative concentration inequality (Lemma 5.2). The proof of this lemma is based on a well-known technique for establishing concentration inequalities for hitting times using the classical

Azuma's inequality and the fact that the shifted expected hitting times form a martingale (for example see [6]).

Lemma 5.2. *Suppose there exists a constant C such that for all $x \in \mathcal{X} \setminus S$*

$$|\mathbb{E}_{X_1}[\tau_S] - \mathbb{E}_x[\tau_S]| \leq C \quad \textit{almost surely}$$

Then for any $x \in \mathcal{X}$ and $\varepsilon > 0$,

$$\mathbb{P}_x(|\tau_S - \mathbb{E}_x[\tau_S]| \geq \varepsilon \mathbb{E}_x[\tau_S]) \leq 2\exp\left(\frac{-2\varepsilon^2 \mathbb{E}_x[\tau_S]}{(1+\varepsilon)C^2}\right)$$

To prove this lemma we state (without proof) the version of the Azuma's inequality we will use (see Theorem 5.1 in [2]).

Theorem 5.3. *Suppose $\{M_t\}_{t=0}^T$ is a martingale with respect to a filtration $\{\mathcal{F}_t\}_{t=0}^T$ and there exists $C \geq 0$ such that for each $t \geq 1$,*

$$|M_t - M_{t-1}| \leq C \quad \textit{almost surely}$$

Then for any $\varepsilon > 0$,

$$\mathbb{P}(|M_T - M_0| \geq \varepsilon) \leq 2\exp\left(\frac{-2\varepsilon^2}{TC^2}\right)$$

Proof of Theorem 5.2. Let $X_0 = x$. Fix $\varepsilon > 0$. Let $T = \lfloor (1+\varepsilon)\mathbb{E}_{x_n}[\tau_S] \rfloor$. For t from 0 to T define

$$M_t = \begin{cases} \mathbb{E}_{X_t}[\tau_S] + t & \text{when } t < \tau_S \\ \tau_S & \text{when } t \geq \tau_S \end{cases}$$

It is easy to check that M_t is a martingale with respect to the filtration $\mathcal{F}_t = \sigma(X_0, \dots, X_t)$. Then for any t from 1 to T, either $X_{t-1} \in S$ in which case $|M_t - M_{t-1}| = 0$, or $X_{t-1} \in \mathcal{X} \setminus S$ in which case $|M_t - M_{t-1}| = |\mathbb{E}_{X_t}[\tau_S] - \mathbb{E}_{X_{t-1}}[\tau_S] + 1| \leq C + 1$ almost surely. Therefore $|M_t - M_{t-1}| \leq C + 1$ almost surely. Thus by Theorem 5.3,

$$\mathbb{P}(|M_T - M_0| \geq \varepsilon \mathbb{E}_x[\tau_S]) \leq 2\exp\left(\frac{-2\varepsilon^2 \mathbb{E}_x[\tau_S]}{(1+\varepsilon)C^2}\right)$$

We will finish the proof by showing that the event $|\tau_S - \mathbb{E}_x[\tau_S]| \geq \varepsilon \mathbb{E}_x[\tau_S]$ implies the event $|M_T - M_0| \geq \varepsilon \mathbb{E}_x[\tau_S]$ and therefore

$$\begin{aligned} \mathbb{P}_x(|\tau_S - \mathbb{E}_x[\tau_S] > \varepsilon \mathbb{E}_x[\tau_S]) &\leq \mathbb{P}(|M_T - M_0| \geq \varepsilon \mathbb{E}_x[\tau_S]) \\ &\leq 2\exp\left(\frac{-2\varepsilon^2 \mathbb{E}_x[\tau_S]}{(1+\varepsilon)C^2}\right) \end{aligned}$$

Suppose first $\tau_S - \mathbb{E}_x[\tau_S] > \varepsilon \mathbb{E}_x[\tau_S]$. Then $\tau_S > (1+\varepsilon)\mathbb{E}_x[\tau_S] \geq T$. In particular this implies $M_T = \mathbb{E}_{X_T}[\tau_S] + \lfloor (1+\varepsilon)\mathbb{E}_x[\tau_S] \rfloor$ and since $M_0 = \mathbb{E}_x[\tau_S]$ we have that $|M_T - M_0| \geq \mathbb{E}_{X_T}[\tau_S] + \lfloor (1+\varepsilon)\mathbb{E}_x[\tau_S] \rfloor - \mathbb{E}_x[\tau_S] \geq (\mathbb{E}_{X_T}[\tau_S] - 1) + \varepsilon \mathbb{E}_x[\tau_S] \geq \varepsilon \mathbb{E}_x[\tau_S]$. In the other case suppose $E_x[\tau_S] - \tau_S \geq \varepsilon \mathbb{E}_x[\tau_S]$. Then $T \geq (1-\varepsilon)\mathbb{E}_x[\tau_S] > \tau_S$. In particular this implies that $M_T = \tau_S$ and since $M_0 = \mathbb{E}_x[\tau_S]$ we have that $|M_T - M_0| = |\tau_S - \mathbb{E}_x[\tau_S]| = \mathbb{E}_x[\tau_S] - \tau_S \geq \varepsilon \mathbb{E}_x[\tau_S]$. ☐

Proof of Proposition 5.1. By Theorem 2.6 it suffices to show that τ_S concentrates at times $t_n = \mathbb{E}_{x_n}[\tau_S]$. Let $\varepsilon > 0$. For each $n \in \mathbb{N}$

$$\begin{aligned} \mathbb{P}_{x_n}\left(\left|\frac{\tau_S}{\mathbb{E}_{x_n}[\tau_S]} - 1\right| \geq \varepsilon\right) &= P_x(|\tau_S - \mathbb{E}_x[\tau_S]| \geq \varepsilon \mathbb{E}_x[\tau_S]) \\ &\leq 2\exp\left(\frac{-2\varepsilon^2 \mathbb{E}_x[\tau_S]}{(1+\varepsilon)C^2}\right) \quad \text{by Lemma 5.2} \end{aligned}$$

Taking the limit as $n \to \infty$ shows that

$$\lim_{n\to\infty} \mathbb{P}_{x_n}\left(\left|\frac{\tau_S}{\mathbb{E}_{x_n}[\tau_S]} - 1\right| \geq \varepsilon\right) = 0$$

and therefore

$$\lim_{n\to\infty} \mathbb{P}_{x_n}\left(\left|\frac{\tau_S}{\mathbb{E}_{x_n}[\tau_S]} - 1\right| \leq \varepsilon\right) = 1 \qquad \square$$

6 Random walks on the half line

In this section we apply Proposition 5.1 to show that all random walks on the half-line with bounded step size exhibit the starting point cutoff phenomenon. Let $\{W_t\}_{t\geq 1}$ be a sequence of i.i.d real-valued random variables and let $x \in [0,\infty)$. We define a random walk X_t on the half-line $[0,\infty)$ recursively by setting:

$$\begin{aligned} &X_0 = x, \\ &X_t = [X_{t-1} + W_t]^+ \quad \text{for } t \geq 1 \end{aligned}$$

where $[x]^+ = \max\{x, 0\}$. It is shown in Proposition 11.4.1 of [8] that this Markov chain is regular if and only if $\mu := \mathbb{E}[W_t] < 0$, and in this case all compact sets are regular sets. We will consider in particular the regular set $S = \{0\}$. We say that X_t is a random walk on the half-line with bounded steps if there exists $C > 0$ such that $|W_t| \leq C$ almost surely.

Theorem 6.1. *Suppose X_t is a random-walk on the half-line with bounded steps such that $\mu := \mathbb{E}[W_t] < 0$. Then for any sequence of starting points $x_n \in [0,\infty)$, if $\lim_{n\to\infty} x_n = \infty$ then X_t exhibits starting point cutoff at time $t_n = \mathbb{E}_{x_n}[\tau_{\{0\}}]$ starting from x_n.*

Proof. Fix $x \in \mathcal{X}$ and let $X_0 = x$. Let $S_t = \sum_{i=1}^{t} W_i$. Note that

$$\tau_{\{0\}} = \inf\{t \geq 0 \mid X_t = 0\} = \inf\{t \geq 0 | S_t \leq -x\}$$

Therefore by Wald's equation (see Theorem 2.6.2 in [3]) we have

$$\mathbb{E}_x[\tau_{\{0\}}] = \frac{\mathbb{E}[S_{\tau_{\{0\}}}]}{\mu}$$

but since

$$-(C + x) \leq W_{\tau_{\{0\}}} - x \leq W_{\tau_{\{0\}}} + S_{(\tau_{\{0\}-1})} = S_{\tau_{\{0\}}} \leq -x \quad \text{almost surely:}$$

we have that

$$-(C + x) \leq \mathbb{E}[S_{\tau_{\{0\}}}] \leq -x$$

and therefore

$$-\frac{(C + x)}{\mu} \leq \mathbb{E}_x[\tau_{\{0\}}] \leq -\frac{x}{\mu}$$

thus

$$|(\mathbb{E}_{X_1}[\tau_{\{0\}}] - \mathbb{E}_x[\tau_{\{0\}}])| \leq \frac{(|X_1 - x| + C)}{\mu} = \frac{(|W_1| + C)}{\mu} \leq \frac{2C}{\mu} \quad \text{almost surely}$$

Therefore by Proposition 5.1, X_t exhibits starting point cutoff at time $t_n = \mathbb{E}_{x_n}[\tau_{\{0\}}]$ starting from x_n. $\square$

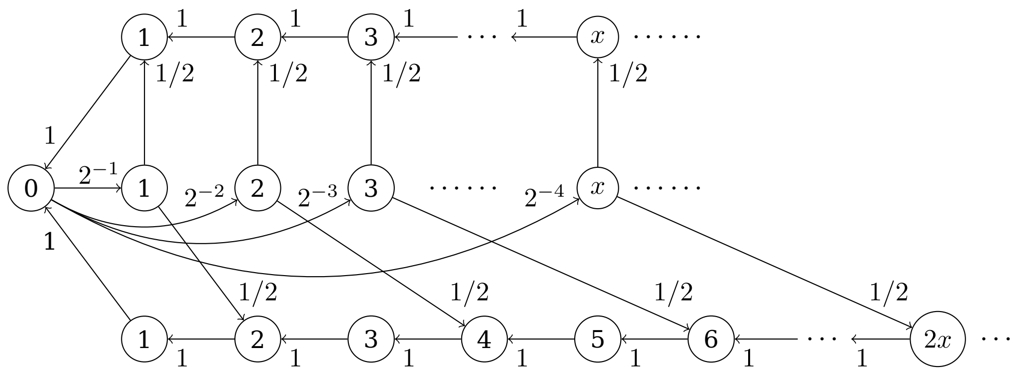

Figure 1: Example Chain.

7 An example without starting point cutoff

In this section we construct an example of a Markov chain with a sequence of starting points x_n that diverges to infinity (in the sense that the expected hitting time of a fixed regular set from those starting points diverges to infinity) where the Markov chain does not exhibit the starting point cutoff phenomenon from those starting points (meaning precisely that there does not exist a sequence of positive reals t_n for which the Markov kernel associated with the chain has starting point cutoff at time t_n starting from x_n).

Consider the Markov chain depicted in Figure 1. Its state space is three disjoint copies of $\mathbb{N}$ glued together at 0. When at 0 the Markov chain jumps to a state in the middle row with a $\mathrm{Geometric}(1/2)$ distribution. When at a non-zero state x in the middle row the Markov chain jumps to x in the top row with probability $1/2$ or jumps to $2x$ in the bottom row with probability $1/2$. Once in the top or bottom row the Markov chain jumps deterministically down in the same row one step at a time until it reaches 0. It is easy to see that this is an aperiodic regular Markov chain and $S = \{0\}$ is a regular set. Now for any increasing sequence of starting points x_n in the middle row we have that $\mathbb{P}_{x_n}(\tau_S = x_n) = \frac{1}{2}$ and $\mathbb{P}_{x_n}(\tau_S = 2x_n) = \frac{1}{2}$ so it is clear that $\lim_{n\to\infty} \mathbb{E}_{x_n}[\tau_S] = \infty$ and τ_S starting from x_n cannot concentrate at any (deterministic) time t_n for any sequence of positive reals t_n. Therefore by Theorem 2.6 for any increasing sequence of starting points x_n in the middle row, the Markov chain does not exhibit the starting point cutoff phenomenon.

8 Open problems

In this section we state and discuss three open problems inspired by the results of this paper. The first proposes a possible extension of the theory, and the latter two propose extensions of the application.

Question 8.1. Can one characterize the starting point cutoff phenomenon for a sequence of (possibly distinct) regular Markov chains by a concentration of hitting times (generalizing both Theorem 2.6 and Theorem 1 of [5])?

Theorem 1 of [5] shows that for a sequence of finite irreducible reversible Markov chains satisfying the product condition, cutoff starting from a sequence of starting distributions can be characterized by concentration of hitting times for a sequence of sets "worst in expectation". While this result only holds for reversible chains and does not apply directly to the general state space setting of Theorem 2.6, one could imagine a generalization that applies to sequences of regular Markov chains (on a general state space). In the case where the sequence of Markov chains is constant this should reduce to Theorem 2.6. Since the main tool of [5], Starr's maximal inequality [10], applies on a general probability space, we believe it is likely that Theorem 1 of [5] can be extended to sequences of reversible regular Markov chains on a general state space in a similar way that this paper extends the results of [7] using ideas from [8].

Question 8.2. Could Proposition 5.1 be used to establish a cutoff phenomenon for other general classes of chains?

One can think of many examples where the hypothesis of Proposition 5.1 likely holds but is hard to verify. This is because there does not seem to be any good tools for computing or bounding $\mathbb{E}_x[\tau_S]$ in general. There may however be other interesting general classes of chains where bounds on this expectation could be derived and the hypothesis of Proposition 5.1 could be shown to hold.

Question 8.3. What are the necessary and sufficient conditions on the (possibly unbounded) distribution of W_t for the random walk on the half-line X_t to exhibit starting point cutoff?

We showed in Theorem 6.1 that bounded steps is a sufficient condition for random walks on the half-line to exhibit the starting point cutoff phenomenon, but it is easy to construct examples which show this is not necessary. It would be interesting to determine precisely under what conditions a random walk on the half-line exhibits starting point cutoff.

References

[1] Persi Diaconis, *The cutoff phenomenon in finite Markov chains.*, Proceedings of the National Academy of Sciences of the United States of America **93 4** (1996), 1659–64. MR1374011

[2] Devdatt Dubhashi and Alessandro Panconesi, *Concentration of Measure for the Analysis of Randomized Algorithms*, 1st ed., Cambridge University Press, USA, 2009. MR2547432

[3] Rick Durrett, *Probability: Theory and examples*, 5 ed., Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press, 2019. MR3930614

[4] G.B. Folland, *Real Analysis: Modern Techniques and Their Applications*, Pure and Applied Mathematics: A Wiley Series of Texts, Monographs and Tracts, Wiley, 1999. MR1681462

[5] Jonathan Hermon, *A technical report on hitting times, mixing and cutoff*, arXiv: Probability (2015). MR3765366

[6] Timo Kötzing, *Concentration of first hitting times under additive drift*, Proceedings of the 2014 Annual Conference on Genetic and Evolutionary Computation (New York, NY, USA), GECCO'14, Association for Computing Machinery, 2014, p. 1391–1398. MR3383680

[7] Servet Martínez and Bernard Ycart, *Decay Rates and Cutoff for Convergence and Hitting Times of Markov Chains with Countably Infinite State Space*, Advances in Applied Probability **33** (2001), no. 1, 188–205. MR1825322

[8] Sean Meyn, Richard L. Tweedie, and Peter W. Glynn, *Markov Chains and Stochastic Stability*, 2 ed., Cambridge Mathematical Library, Cambridge University Press, 2009. MR2509253

[9] Gareth Roberts and Jeffrey Rosenthal, *Geometric Ergodicity and Hybrid Markov Chains*, Electronic Communications in Probability **2** (1997), 13–25. MR1448322

[10] Norton Starr, *Operator limit theorems*, Transactions of the American Mathematical Society **121** (1966), no. 1, 90–115. MR0190757