Merge Times and Hitting Times of Time-inhomogeneous Markov Chains

by

Jiarou Shen

Department of Mathematics
Duke University

Date: ______________________
Approved:

____
David J. Sivakoff, Supervisor

____
[Committee Member: David J. Sivakoff]

____
[Committee Member: Kevin McGoff]

Thesis submitted in fulfillment of the requirements for Graduation with Distinction
in the degree of Bachelor of Science in the Department of Mathematics
in the Undergraduate School of Duke University
2013

Abstract

The purpose of this thesis is to study the long term behavior of time-inhomogeneous Markov chains. We analyze under what conditions they converge, in what sense they converge and what the rate of convergence should be.

A Markov chain is a random process with the memoryless property: the next state only depends on the current state, and not on the sequence of events that preceded it. Time-inhomogeneous Markov chains refer to chains with different transition probability matrices at each step. What makes them interesting is that they don't necessarily have stationary distributions. Instead of comparing the chain to the stationary distribution, we look at the distance between two distributions started at different initial states. We refer to the time until the two distributions get sufficiently close as merge time. As a foundation for our simulations and proofs, we first show the convergence theorem for time-inhomogeneous Markov chains with a sufficient assumption. We then study the various ways of perturbing the random walk on the n-cycle, by forcing the random walker to move in a certain direction with higher probability if it is at a particular site. Changing perturbations at every step results in different one-step transition kernels throughout the chain, therefore making the chain time-inhomogeneous. We then compare the merge times, as a function of the number of states n, to those of the unperturbed time-homogeneous simple random walk on the n-cycle.

One of the perturbations represents the case in which the random walker has a varying probability of staying put at a particular site at every step. Simulations show that the merge times for this perturbed chain is almost identical to those of the unperturbed chains. We are able to show this result by proving an explicit bound on the hitting times.

Contents

Abstract **iii**

List of Tables **vii**

List of Figures **viii**

Acknowledgements **ix**

1 Introduction **1**

1.1 Markov Chains 1

1.2 Coupling 2

1.3 Simple Random Walk on the n-cycle 3

1.4 Mixing times, Merge Times and Hitting Times 4

1.5 Time-inhomogeneous Markov Chains 7

1.6 Perturbations on Simple Random Walk on the n-cycle 8

1.7 Question of Study 8

1.8 Outline 10

2 Theorems and Lemmas **11**

2.1 Convergence of Time-homogeneous Markov Chains 11

2.2 Convergence of Time-inhomogeneous Markov Chains with a Sufficient Assumption 12

2.3 Gambler's Ruin 14

2.4 Conditional Expectation on Reaching Site 0 Before Site k 15

2.5 Hitting Times in Question of Study 17

2.6 Relations Between Hitting Times of the Perturbed and Unperturbed Chains 20

3 Proof of Theorem 2.6.1 **22**

4 Simulations **26**

4.1 Several Kinds of Perturbations 27

4.2 Perturbed Site Alternating Between Two States 29

4.3 Different Angles on a Clock 31

4.4 Varying Stickiness at Site 0 32

4.5 Conclusions for Simulations 33

5 Future Work **36**

Bibliography **39**

List of Tables

4.1 The values of $\alpha' s$ when fitting merge times to $t_{merge}(n)=\alpha n^2$. 35

List of Figures

1.1 Simple Random Walk on the n-cycle 4

1.2 Mixing times and merge times of unperturbed simple random walk on the n-cycle 6

1.3 Perturbed chain with varying stickiness at site 0 9

2.1 Hitting times 18

2.2 Asymmetric paths 19

4.1 Several kinds of perturbations 29

4.2 Perturbed state alternating between 3 and n 30

4.3 Perturbed state alternating between n/4 and n-n/4 31

4.4 Perturbed state alternating between two states 32

4.5 Perturbed state alternating between two states with different angles . 33

4.6 Varying stickiness at site 0 34

Acknowledgements

I would like to thank Prof. Sivakoff for his guidance, inspiration and patience throughout my work on this project, Prof. Kraines for inviting me to events and giving me opportunities to share our results, Prof. Mapes and Prof. Mela for recommending me to the PRUV Program that provides a perfect start for this thesis, the Duke Math Department for financial support through PRUV Fellowship, all PRUV Fellows for useful discussions on my progress, and my family and friends for always supporting and believing in me.

1

Introduction

This thesis addresses a proof for convergence of time-inhomogeneous Markov chains with a sufficient assumption, simulations for the merge times of some time-inhomogeneous Markov chains, and bounds for a perturbed random walk on the n-cycle with varying stickiness at one site. We prove that the hitting times for that specific model converge to the hitting times of the original unperturbed chain.

1.1 Markov Chains

As introduced in the Abstract, a Markov chain is a sequence of stochastic events where the next state of a variable or system is independent of all the past states, except for the current state.

Let X_t denote the state at step t. A Markov chain is said to be time-homogeneous if the transition probabilities are the same at each step. We define the transition probability matrix (kernel), P, of time-homogeneous Markov chain as:

$$P(x, y) = \mathbb{P}\left(X_t = y | X_{t-1} = x\right),$$

$$P^t(x, y) = \mathbb{P}\left(X_t = y | X_0 = x\right).$$

A Markov chain, with transition matrix P and state space S, is said to be **irreducible** if for any $x, y \in S$, $\exists$ positive integer t, such that

$$P^t(x, y) > 0.$$

The **period** of state x is the gcd of $\tau(x) := \{t \geqslant 1 : P^t(x, x) > 0\}$, and all states have the same period if P is irreducible. A Markov chain, with transition matrix P and state space S, is **aperiodic** if the period of all states is 1.

We define a **stationary distribution** of a Markov chain to be the probability distribution π on S that satisfies $\pi = \pi P$ and $\pi(x) > 0 \; \forall \; x \in S$.

Proposition 1.1.1 ([1], Proposition 1.14)**.** *An irreducible Markov chain, with transition matrix P and finite state space S, has a unique stationary distribution.*

If a Markov chain is both irreducible and aperiodic, the chain converges to its stationary distribution. We will formally introduce the convergence theorem for irreducible and aperiodic Markov chains in Section 2.1.

1.2 Coupling

A **coupling** of two probability distributions μ and ν is a construction of a pair of two variables (X, Y) on the same probability space so that each retains its marginal distribution, $X \sim \mu$ and $Y \sim \nu$.

Definition 1.2.1. *The* **total variation distance** *between two distributions* μ *and* ν *on* $[n]$ *is*

$$d_{TV}(\mu,\nu) = ||\mu - \nu||_{TV} = \max_{A \subset [n]} (\mu(A) - \nu(A)). \tag{1.1}$$

Lemma 1.2.2. *[[1], Proposition 4.7] If* μ *and* ν *are two distributions on* $[n]$ *and* $X \sim \mu$*,* $Y \sim \nu$ *then*

$$d_{TV}(\mu,\nu) = \inf \mathbb{P}\left(X \neq Y\right), \tag{1.2}$$

where the infimum is taken over all couplings of X *and* Y*.*

This lemma will be useful for the convergence proof for time-inhomogeneous Markov chains, since it allows us to bound the distance between two distributions if we can find a coupling such that the coupled random variables tend to be equal.

1.3 Simple Random Walk on the n-cycle

A simple random walk on the n-cycle, $[n] = \{1, 2, \ldots, n(or\ 0)\}$, is the Markov chain with transition kernel

$$K(i,j) = \begin{cases} 1/2 & \text{if } |i-j| = 1 \text{ or } n-1 \\ 0 & \text{otherwise.} \end{cases} \tag{1.3}$$

See Figure 1.1.

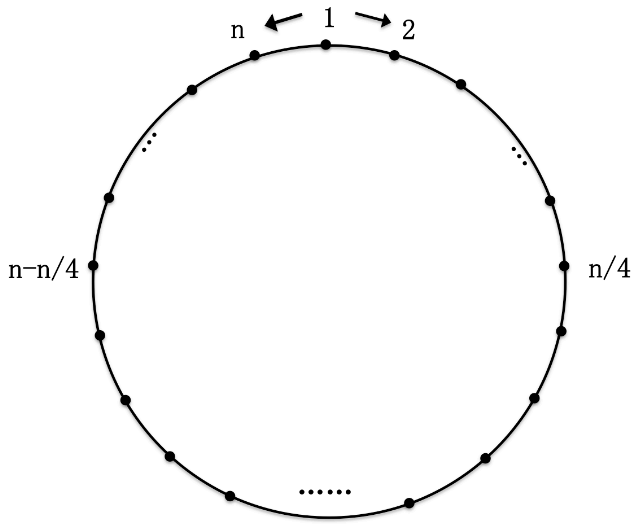


FIGURE 1.1: Simple Random Walk on the n-cycle

Given an initial distribution μ(a row vector), μK^t is the distribution of X_t, the Markov chain after t steps. We see that it has the stationary distribution $\pi(x) = 1/n$ for any $x \in [n]$, because $\pi K = \pi$. We can estimate how long it takes for the distribution of X_t get close to the π in the total variation distance.

1.4 Mixing times, Merge Times and Hitting Times

Let $P^t(x, \cdot)$ be the distribution of X_t when $X_0 = x$, and suppose π is the stationary distribution for this Markov chain. Define

$$d(t) := \max_{x \in [n]} \left|\left|P^t(x, \cdot) - \pi\right|\right|_{TV}, \tag{1.4}$$

$$\bar{d}(t) := \max_{x,y \in [n]} \left|\left|P^t(x, \cdot) - P^t(y, \cdot)\right|\right|_{TV}. \tag{1.5}$$

Mixing time is the time for the distribution of an irreducible Markov Chain to get sufficiently close to its stationary distribution.

Definition 1.4.1. *Suppose X_t is a Markov chain on $[n]$ with stationary distribution π. The* **mixing time** *of X_t is defined to be*

$$t_{mix}(\varepsilon) := \inf\{t : d(t) \leqslant \varepsilon\},$$

$$t_{mix} := t_{mix}(1/4).$$

The concept of mixing time is useful in describing the required time for the chain to be close to its stationary distribution as a function of the size of the state space, which is controlled by the parameter n in the case of the random walk on the n-cycle.

We can also compare the chain to itself instead of to the stationary distribution. Merge time is the time it takes for distributions of the Markov chain to get sufficiently close to one another started from different initial states as a function of the size of the state space.

Definition 1.4.2. *Suppose X_t is any Markov chain on $[n]$. The* **merge time** *is defined to be*

$$t_{merge}(\varepsilon) := \inf\{t : \bar{d}(t) \leqslant \varepsilon\}, \tag{1.6}$$

$$t_{merge} := t_{merge}(1/4). \tag{1.7}$$

If we simulate the simple random walk on the n-cycle, we see that t_{merge} for different number of states are slightly smaller than twice the corresponding t_{mix}. See Figure 1.2.

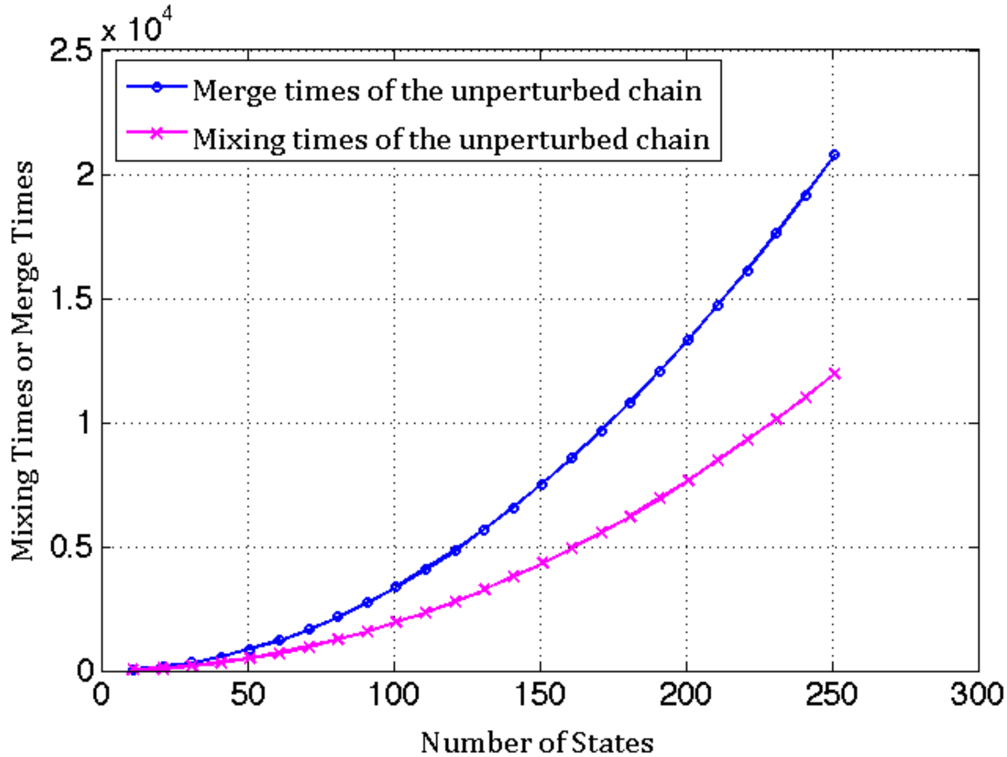


FIGURE 1.2: Mixing times and merge times of unperturbed simple random walk on the n-cycle

Hitting time is the maximum expected time for the Markov chain to travel between any two states.

Definition 1.4.3. *Let X_t be a Markov Chain on S, let $V_y := \min\{t \geqslant 0 : X_t = y\}$, and let E_x denotes expectation with respect to $P(\cdot|X_0 = x)$. The* **hitting time** *corresponding the the chain X_t is*

$$t_{hit} := \max_{x,y \in S} E_x(V_y). \tag{1.8}$$

Hitting times are closely related to merge times as hitting times measure how efficiently the chain explores the state space. In order for a chain to mix or merge, it needs to be able to access most sites on the state space.

1.5 Time-inhomogeneous Markov Chains

A Markov chain with a different transition probability matrix at each step is said to be time-inhomogeneous. Let K_t be the transition matrix of the chain at step t. We have

$$K_t(x, y) = \mathbb{P}\left(X_t = y | X_{t-1} = x\right),$$

$$K_{t,t+r}(x, y) = K_{t+1} K_{t+2} \cdots K_{t+r}(x, y), \forall t \geqslant 0, r \geqslant 1.$$

Time-inhomogeneous Markov chains do not necessarily have stationary distributions. In fact, for a chain with transition matrices alternating between Q_1 and Q_2, where Q_1, Q_2, $Q_1 Q_2$ are all irreducible, if the stationary distributions of Q_1 and Q_2 are different, then $K_{0,t} = Q_1 Q_2 Q_1 Q_2 \ldots$ doesn't converge to a stationary distribution, or even have one. Thus, we compare the chains to themselves by using the concept of merge times. Time-inhomogeneous Markov chains only converge under certain assumptions. In [2, p.759], Griffeath proved a form of convergence for time-inhomogeneous Markov chains which satisfy certain properties, known as the Weak Ergodicity Theorem. In Chapter 2 of this thesis, we will prove the Weak Ergodicity Theorem under slightly stronger assumptions using the ideas of coupling. The proof provides a foundation for our various simulations of inhomogeneous perturbations of the random walk on the n-cycle.

1.6 Perturbations on Simple Random Walk on the n-cycle

In [3], L. Saloff-Coste and J. Zúñiga gave us the idea to look at perturbations on the random walker on the n-cycle. We can make a simple random walk on the n-cycle time-inhomogeneous by applying small inhomogeneous perturbations to the chain and study how the merge times will be affected by these perturbations. Vertices on the n-cycle are numbered clockwise as in Figure 1.1. At every step, we choose a different state to perturb, so the chain will be time-inhomogeneous. The sequence of kernels is independent of the actual paths of the random walkers, and can therefore be chosen in advance. In these time-inhomogeneous cases, we let $K_1, K_2, \ldots, K_t$ be a sequence of one-step transition kernels, and let $P^t = K_1 K_2 \cdots K_t$ be the t step transition matrix.

By comparing the merge times for the perturbed and unperturbed random walks on the n-cycle, as well as using probabilistic arguments to bound their mixing behavior, our goal is to understand the effects of different perturbations.

1.7 Question of Study

Among all perturbed chains we will discuss in this thesis, a particularly interesting strategy is to add some "stickiness" to the state where the random walker starts initially, which is site 0 in our case. At every step, the stickiness can be different. At step i, Instead of only moving right or left, the walker can now stay put for a certain probability δ_j. We take $\delta_j \in (0, 1 - \varepsilon)$, where $\varepsilon > 0$, so that the walker does not get stuck completely. See Figure 1.3.

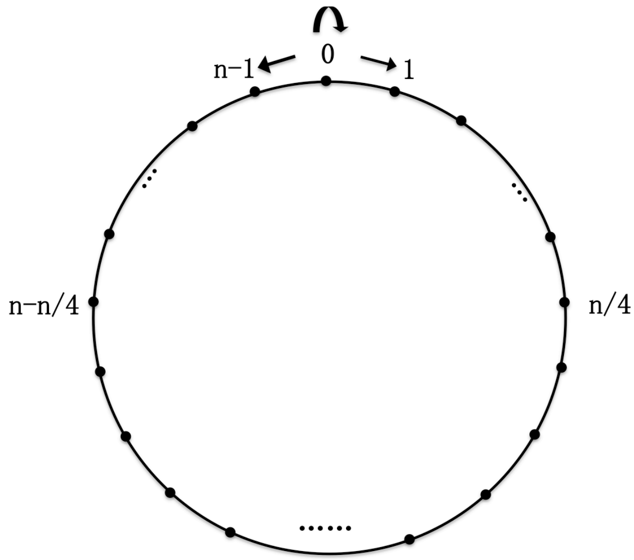


FIGURE 1.3: Perturbed chain with varying stickiness at site 0

Denoted by K^δ, the perturbation at state 0 with intensity of stickiness δ is

$$K^\delta(x,y) := \begin{cases} \frac{1}{2} & \text{if } |x-y| = 1 \text{ and } x \neq 0 \\ \delta & \text{if } x = y = 0 \\ \frac{1}{2} - \frac{\delta}{2} & \text{if } x = 0,\ y = 0 \pm 1 \\ 0 & \text{otherwise,} \end{cases}$$

where addition and subtraction are on the n-cycle, so $n = 1 - 1$ and $1 = n + 1$.

Simulation results indicate that the merge times for the perturbed chain with varying stickiness at site 0 are very similar to those of the unperturbed chain. We want to show that no matter what the sequence of δ_j's is, the merge times of the perturbed and unperturbed chains are asymptotically equal as n $\to \infty$. However, we can only show that the hitting times of the perturbed chain converge to those of the unperturbed chain.

1.8 Outline

An outline of the remainder of this thesis is as follows. Chapter 2 discusses in detail some important theorems and lemma that are essential to understanding the problems or used for computations or proofs in later chapters. Section 2.1 and Section 2.2 lay out the sufficient assumption for a time-inhomogeneous Markov chain to converge and provides a proof of the Weak Ergodicity Theorem using a coupling approach. Section 2.3 studies a famous Markov Chain, Gambler's Ruin. The important results of Gambler's Ruin hitting times are used in the proof in Chapter 3. Section 2.4 and Section 2.5 probe the hitting times in the context of the perturbed random walk with varying stickiness at site 0. Lemmas from these two sections are keys to build the proof in Chapter 3.

Chapter 3 provides a proof for the convergence of the hitting times of the perturbed random walk on the n-cycle with varying stickiness at site 0 to the hitting times of the unperturbed chain. The idea of the proof is to find explicit bounds on the hitting times, and show that the upper and lower bounds converge to the same value.

Chapter 4 introduces four categories of time-inhomogeneous perturbations of simple random walk on the n-cycle with fifteen individual cases, including the one studied in Chapter 3. This chapter presents the simulation results and conclusions for the merge times of the perturbed chain in each category, and compares them to the merge times of the original unperturbed chain.

Finally, Chapter 5 contains discussions of future work on the topic.

2

Theorems and Lemmas

2.1 Convergence of Time-homogeneous Markov Chains

The classical convergence theorem for time-homogeneous Markov chains is the following:

Theorem 2.1.1 ([1], Theorem 4.9)**.** *Suppose that P is irreducible and aperiodic, with stationary distribution π, then $\exists\ \alpha \in (0,1)$ and $C > 0$, s.t.*

$$d(t) \leqslant C\alpha^t. \tag{2.1}$$

The theorem says that a Markov chain on a fixed state space converges exponentially quickly to its stationary distribution.

Proposition 2.1.2 ([1], Exercise 4.3)**.** *For a Markov chain with transition matrix P and any two distributions μ and ν on the state space S,*

$$||\mu P - \nu P||_{TV} \leqslant ||\mu - \nu||_{TV}\,.$$

Proof of Proposition 2.1.2.

$$\begin{aligned}
||\mu P - \nu P||_{TV} &= \frac{1}{2}\sum_{x\in S} |\mu P(x) - \nu P(x)| \\
&= \frac{1}{2}\sum_{x\in S} |\sum_{y\in S} \mu(y)P(y,x) - \sum_{y\in S} \nu(y)P(y,x)| \\
&\leqslant \frac{1}{2}\sum_{x\in S}\sum_{y\in S} P(y,x)|\mu(y) - \nu(y)| \\
&= \frac{1}{2}\sum_{y\in S} |\mu(y) - \nu(y)| \sum_{x\in S} P(y,x) \\
&= \frac{1}{2}\sum_{y\in S} |\mu(y) - \nu(y)| \\
&= ||\mu - \nu||_{TV}\,.
\end{aligned}$$

□

This shows that advancing the chain does not increase the total variation distance between μ and ν.

2.2 Convergence of Time-inhomogeneous Markov Chains with a Sufficient Assumption

In [2], Griffeath proved the convergence of time-inhomogeneous Markov chains under a slightly weaker assumption than the assumption we are using in the following Theorem 2.2.2. The reason we formulate the conditions in Assumption 2.2.1 are motivated by the fact that all time-inhomogeneous Markov chains that we will consider later satisfy this assumption. We hereby provide a proof for the convergence of time-inhomogeneous Markov chains using a coupling argument.

Assumption 2.2.1. *$\exists$ r>0 s.t. $\exists$ $\delta\in$(0,1), so that for every $t \geqslant 0$ and x, y in S,*

$$K_{t,t+r}(x,y) \geqslant \delta. \tag{2.2}$$

Theorem 2.2.2 (Weak Ergodicity)**.**

Under Assumption 2.2.1, $\max_{x,y\in S} ||K_{0,t}(x,\cdot) - K_{0,t}(y,\cdot)||_{TV} \to 0$ *exponentially as* $t \to \infty$.

Proof of Theorem 2.2.2.

Let $X_t \sim K_{0,t}(x,\cdot)$ and $Y_t \sim K_{0,t}(y,\cdot)$. Let (X_t, Y_t) be coupled so that they move independently before they meet, but once $X_t = Y_t$, they stay together.

By Lemma 1.2 and Proposition 2.1.2, we have for any $k \geqslant 0$, $0 \leqslant s \leqslant r-1$,

$$\begin{aligned} ||\delta_x K_{0,kr+s} - \delta_y K_{0,kr+s}||_{TV} &= ||\delta_x K_{0,kr} K_{kr,kr+s} - \delta_y K_{0,kr} K_{kr,kr+s}||_{TV} \\ &\leqslant ||\delta_x K_{0,kr} - \delta_y K_{0,kr}||_{TV} \\ &\leqslant \mathbb{P}\left(X_{kr} \neq Y_{kr}\right), \end{aligned} \tag{2.3}$$

where $\mathbb{P}$ is the probability with respect to the coupled processes.

If we can show

$$\mathbb{P}\left(X_{kr} = Y_{kr} | X_{(k-1)r} \neq Y_{(k-1)r}\right) \geqslant \varepsilon > 0, \tag{2.4}$$

then

$$\begin{aligned} \mathbb{P}\left(X_{kr} = Y_{kr}\right) = \mathbb{P}&\left(X_{(k-1)r} = Y_{(k-1)r}\right) \\ &+ \mathbb{P}\left(X_{(k-1)r} \neq Y_{(k-1)r}\right)\mathbb{P}\left(X_{kr} = Y_{kr} | X_{(k-1)r} \neq Y_{(k-1)r}\right) \\ &\geqslant 1 - \mathbb{P}\left(X_{(k-1)r} \neq Y_{(k-1)r}\right) + \mathbb{P}\left(X_{(k-1)r} \neq Y_{(k-1)r}\right)\varepsilon \\ &= 1 - (1-\varepsilon)\mathbb{P}\left(X_{(k-1)r} \neq Y_{(k-1)r}\right), \end{aligned} \tag{2.5}$$

so that

$$\begin{aligned}\mathbb{P}\left(X_{kr} \neq Y_{kr}\right) &\leqslant (1-\varepsilon)\mathbb{P}\left(X_{(k-1)r} \neq Y_{(k-1)r}\right) \\ &\leqslant (1-\varepsilon)^k \mathbb{P}\left(X_0 \neq Y_0\right) \\ &\leqslant (1-\varepsilon)^k.\end{aligned} \tag{2.6}$$

To show (2.4), we know that for any $z \in S$, by Assumption 2.2.1,

$$K_{t,t+r}(x,z) \geqslant \delta,\ K_{t,t+r}(y,z) \geqslant \delta,$$

Therefore,

$$\begin{aligned}\mathbb{P}\left(X_{kr} = Y_{kr} | X_{(k-1)r} \neq Y_{(k-1)r}\right) &= \sum_{z\in S} \mathbb{P}\left(X_{kr} = z, Y_{kr} = z | X_{(k-1)r} \neq Y_{(k-1)r}\right) \\ &> n\delta^2 \\ &:= \varepsilon > 0.\end{aligned} \tag{2.7}$$

□

2.3 Gambler's Ruin

Gambler's Ruin is a classical Markov chain modeled by a random walk on a path with vertices $0, 1, 2, \cdots, n$. Once the random walker reaches 0 or n, he stays forever. At all interior vertices, it has equal probability to go up or down by 1.

Proposition 2.3.1 ([1], Chapter 2.1)**.** *Let X_t be the random walker's position at time t. Let $V_k = \min\{t \geqslant 0 : X_t = k\}$ and so $V_0 \wedge V_n$ is the time required to be absorbed at one of 0 or n. Assume that $X_0 = k$, where $0 \leqslant k \leqslant n$. Then*

$$\mathbb{P}_k(X_{V_0 \wedge V_n} = n) = k/n, \tag{2.8}$$

$$\mathbf{E}_k(V_0 \wedge V_n) = k(n-k). \tag{2.9}$$

2.4 Conditional Expectation on Reaching Site 0 Before Site k

In this section, we compute the expected time for the walker to reach site 0 before site k starting at site x with respect to the Gambler's Ruin chain, given that the walker hits 0 before k.

By Gambler's Ruin, $E_x(V_0 \wedge V_k) = x(k-x)$. We can split this expected time into two cases: the expected time to hit 0 conditional on the walker reaching 0 before k, and the expected time to hit site k conditional on the walker reaching k before 0.

Define $g^k(x) = E_x(V_0|V_0 < V_k) = E_{k-x}(V_k|V_k < V_0)$. For $0 < x < k$,

$$\begin{aligned}\mathbb{P}\left(X_{n+1} = x+1 | X_n = x, V_0 < V_k\right) &= \frac{\mathbb{P}_x(X_1 = x+1, V_0 < V_k)}{\mathbb{P}_x(V_0 < V_k)} \\ &= \frac{\mathbb{P}_x(X_1 = x+1)\mathbb{P}_{x+1}(V_0 < V_k)}{(k-x)/x} \\ &= \frac{1}{2}\frac{k-x-1}{k-x},\end{aligned} \tag{2.10}$$

and

$$
\begin{aligned}
\mathbb{P}\left(X_{n+1}=x-1|X_n=x,V_0<V_k\right) &= \frac{\mathbb{P}_x(X_1=x-1,V_0<V_k)}{\mathbb{P}_x(V_0<V_k)} \\
&= \frac{\mathbb{P}_x(X_1=x-1)\mathbb{P}_{x-1}(V_0<V_k)}{(k-x)/x} \qquad (2.11) \\
&= \frac{1}{2}\frac{k-x+1}{k-x}.
\end{aligned}
$$

Let

$$
p^k(x)=\mathbb{P}\left(X_{n+1}=x+1|X_n=x,V_0<V_k\right),
$$

$$
q^k(x)=\mathbb{P}\left(X_{n+1}=x-1|X_n=x,V_0<V_k\right).
$$

When $k=2$, $g^2(1)=1$. When $k=3$, we can easily compute $g^3(1)=\frac{5}{3}$ from $g^3(1)=1+p^3(1)(1+q^3(2)g^3(1))$. By solving recursive relations, we get the expected number of steps to reach k conditional on the walker hitting k before 0 to be $g^k(1)=\frac{2k-1}{3}$.

By induction, we generalize that $g^k(x)=g^{k-x+1}(1)+g^{k-x+2}(1)+\cdots+g^k(1)$. Observe that $g^k(k-x)=E_x(V_k|V_k<V_0)=E_{k-x}(V_0|V_0<V_k)$, which represents the expected time to reach k starting at x, conditioning on the random walker reaching k before 0. We have

$$
\begin{aligned}
g^k(k-x) &= \sum_{i=x+1}^{k} g^i(1) \\
&= \frac{2\sum_{i=x+1}^{k} i-(k-x)}{3} \\
&= \frac{2\frac{(x+1+k)(k-x)}{2}-(k-x)}{3} \qquad (2.12) \\
&= \frac{k^2-x^2}{3}.
\end{aligned}
$$

As a check, calculating $g^k(1)$ according to this formula agrees with our previous result.

$$g^k(1) = \frac{k^2 - (k-1)^2}{3} = \frac{2k-1}{3}. \tag{2.13}$$

2.5 Hitting Times in Question of Study

In our question of study, maximizing $E_x V_y$ over all x, y on the n-cycle gives us the hitting time. Note that in our definition of the n-cycle, site 0 is the same as site n and so the state space contains n states. On the n-cycle, let x, y be the clockwise distance between 0 and the point x and y. Suppose $x > y$ and the walker starts at x.

Let T'_y be the time to hit site y in the Markov chain with $p(0,0) = 1-\varepsilon$, let T_y be the time to hit site y in the Markov chain with $p(0,0) = 0$ (so $\varepsilon = 1$), and let T''_y be the time to hit site y in the time inhomogeneous Markov chain where $p_t(0,0) \in [0, 1-\varepsilon]$ (the probability on the t^{th} step of the chain) for some $\varepsilon > 0$. To avoid confusion, we use V_y only referring to the walk on the path in Gambler's Ruin, and use T_y, T'_y, T''_y only referring to the random walk on the n-cycle.

As shown in Figure 2.1, in order to reach y, the walker can take on Path 1 or 2. In other words, the walker either reaches y first without going through 0, or reaches 0 before it reaches y.

If the walker takes on Path 1, we calculate $E_x(T'_y | T'_y < T'_0)$, which equals $E_x(V_y | V_y <$

V_0) in Gambler's Ruin. If we substitute 0 by n, then $E_x(V_y|V_y < V_n) = E_{x-y}(V_0|V_0 < V_{n-y}) = g^{n-y}(x-y)$. If the walker takes on Path 2, we first calculate $E_x(T'_0|T'_0 < T'_y)$, which equals $E_x(V_0|V_0 < V_y)$ in Gambler's Ruin. Now the walker is at 0, and it can stay put, or move either clockwise or counterclockwise to reach y. If it gets back to 0, it is a renewal. After the walker reaches 0, we cannot apply our previous formula to get the expected time to reach y, since the two paths from 0 to y are asymmetric. See Figure 2.2.

By definition, $T'_y := min\{V_y, V_{y-n}\}$. We then have

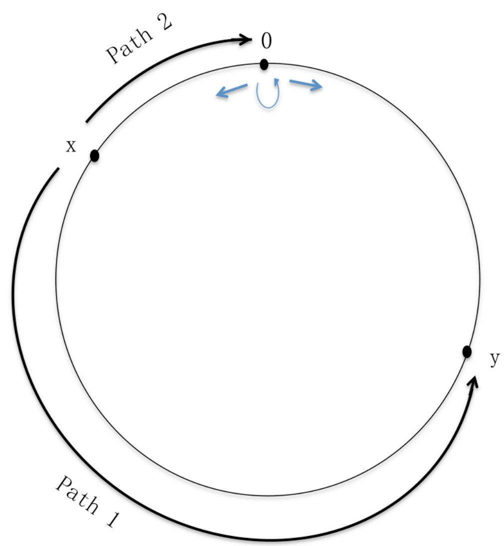


FIGURE 2.1: Hitting times

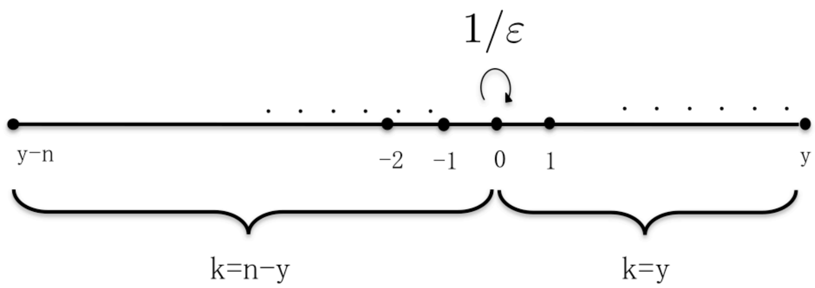


FIGURE 2.2: Asymmetric paths

$$
\begin{aligned}
E_0T'_y = &\frac{1}{2}\{E_0(V_1|V_1 < V_{-1}) + \mathbb{P}_1(V_0 < V_y)[E_1(V_0|V_0 < V_y) + E_0T'_y] \\
&+ \mathbb{P}_1(V_0 > V_y)E_1(V_y|V_0 > V_y)\} \\
&+ \frac{1}{2}\{E_0(V_{-1}|V_1 > V_{-1}) + \mathbb{P}_1(V_0 < V_{n-y})[E_1(V_0|V_0 < V_{n-y}) + E_0T'_y] \\
&+ \mathbb{P}_1(V_0 > V_{n-y})E_1(V_{n-y}|V_0 > V_{n-y})\} \\
= &\frac{1}{2}\left\{\frac{1}{\varepsilon} + \frac{y-1}{y}\left[\frac{y^2-(y-1)^2}{3} + E_0T'_y\right] + \frac{1}{y}\frac{y^2-1}{3}\right\} \\
&+ \frac{1}{2}\left\{\frac{1}{\varepsilon} + \frac{n-y-1}{n-y}\left[\frac{(n-y)^2-(n-y-1)^2}{3} + E_0T'_y\right] + \frac{1}{n-y}\frac{(n-y)^2-1}{3}\right\}.
\end{aligned}
$$

Some simplification yields

$$
E_0T'_y = \frac{y(n-y)((n-2)\varepsilon+2)}{n\varepsilon}. \tag{2.14}
$$

Then, by adding the two paths together, we get

$$
\begin{aligned}
E_x T'_y &= \mathbb{P}(V_0 < V_{n-y}) \left[E_{n-x}(V_0 | V_0 < V_{n-y}) + E_0 T_y \right] + \mathbb{P}(V_0 > V_{n-y}) E_{n-x}(V_{n-y} | V_{n-y} < V_0) \\
&= \frac{x-y}{n-y} \left[\frac{(n-y)^2 - (x-y)^2}{3} + E_0 T_y \right] + \frac{n-x}{n-y} \left[\frac{(n-y)^2 - (n-x)^2}{3} \right] \\
&= \frac{(x-y) \left\{ n(n-x)\varepsilon + y \left[2 + (n-2)\varepsilon \right] \right\}}{n\varepsilon}.
\end{aligned} \tag{2.15}
$$

Note that if we rearrange the equation into the sum of a constant term and a term in some form of ε, the term containing ε should reflect the effect of the varying stickiness at site 0. If we plug in $\varepsilon = 1$, that term should go away and we are left with the hitting time for the unperturbed chain, $E_x T_y$. In Chapter 3, we will be able to determine the explicit form of ε in the term that is only contributed by the perturbations.

2.6 Relations Between Hitting Times of the Perturbed and Unperturbed Chains

From discussion in Section 1.4 in this thesis and Chapter 10 of [1], we see that comparing hitting times give us a good sense of comparing merge times. Therefore, to answer the question of study, we look at hitting times instead of merge times as they are related to one another.

Theorem 2.6.1. *Suppose T_y is the time to hit site y in the simple random walk on the n-cycle with $p(0,0) = 0$, and T''_y is the time to hit site y in the time-inhomogeneous Markov chain with varying stickiness at site 0 added to the simple random walk on* *the n-cycle, where $p_t(0,0) \in [0, 1-\epsilon]$. Then we obtain the following relations between the hitting times of the two chains:*

$$1 \leqslant \frac{\max_{x,y} E_x T''_y}{\max_{x,y} E_x T_y} \leqslant 1 + \frac{2(1-\varepsilon)}{\varepsilon n}. \tag{2.16}$$

We will give a detailed proof for this theorem in Chapter 3.

3

Proof of Theorem 2.6.1

In this chapter, we give a proof for Theorem 2.6.1 using a coupling argument.

For any initial state x (or initial distribution μ), we can construct a coupling between the three random walks so that $T_y \leqslant T''_y \leqslant T'_y$, where T_y, T''_y and T'_y are defined in the beginning of Section 2.5. Let walker A follow the unperturbed chain with hitting time T_y, and let walker B follow the time-inhomogeneous Markov chain with hitting time T''_y, where the stickiness at 0 is different at every step t. Let walker C be the chain with hitting time T'_y, where the perturbation is the same at every step.

Denote X_t^A, X_t^B, X_t^C as the location of walker A, B and C at time t. At time 0, all three walkers start at x, so $X_0^A = X_0^B = X_0^C = x$. We first observe the path of A from x to y. Suppose A visits site 0 for k times. Denote τ_1 as the first time that A hits 0. Let τ_i be the number of steps from the $(i-1)^{st}$ to the i^{th} return of A to 0, for $i \leqslant k$. Let τ_{k+1} be the time between the last time A visits 0 and the time when A eventually hits y.

For walker B, couple B and A so that B follows the path of A until time τ_1. At this time, B might stay at 0 for some time instead of leaving 0 instantly like A. Define $X_{[\tau_1,\tau_1+w]} := (X_{\tau_1}, X_{\tau_1+1}, \cdots, X_{\tau_1+w})$, representing the sequence of states for a walker from time τ_1 to $\tau_1 + w$. Let w_1 be the random number of steps that B takes to leave 0 after τ_1. Conditional on τ_1, the probability of B staying at 0 for j steps after time τ_1 is $\mathbb{P}(w_1 = j|\tau_1) = p_{\tau_1}(0,0)p_{\tau_1+1}(0,0)\cdots p_{\tau_1+w_1-1}(0,0)(1-p_{\tau_1+w_1}(0,0))$. So $X^B[\tau_1, \tau_1+w_1] = (0,0,\cdots,0)$. From step τ_1+w_1 to $\tau_1+w_1-1+\tau_2$, B follows A's pattern in $[\tau_1, \tau_1+\tau_2]$, i.e. we let $X^B[\tau_1+w_1, \tau_1+w_1-1+\tau_2] := X^A[\tau_1+1, \tau_1+\tau_2]$. We continue in this fashion and let $X^B[\sum_{i=1}^{j}\tau_i + w_i, \sum_{i=1}^{j}\tau_i + w_i - 1 + \tau_{i+1}] := X^A[\sum_{i=1}^{j}\tau_i + 1, \sum_{i=1}^{j}\tau_{i+1}]$, where w_i is the random number of steps B takes to leave 0 after hitting 0 for the ith time. Intuitively, right after A's every visit to 0, we simply insert the number of times of B staying at 0 inside A's path to obtain the path for B. For this coupling, because the number of steps in B's path is strictly greater than that of A's, we have $T_y \leqslant T''_y$.

We can couple C and A using the similar method, but now we want to couple C and B so that the time of staying put at 0 on any visit for B is no greater than that for C. Let ξ_t^B, ξ_t^C be independent Bernoulli random variables that indicate whether B and C stay at 0 at time t if they are at 0. If $E(\xi_t^B) = p < q = E(\xi_t^C)$, then there exists a coupling such that $\xi_t^B \leqslant \xi_t^C$. Let $U \sim Uniform(0,1)$, $\xi_t^B = \mathbb{1}_{\{U\leqslant p\}}$, $\xi_t^C = \mathbb{1}_{\{U\leqslant q\}}$. Then whenever $\xi_t^B = 1$, we have $\xi_t^C = 1$, and whenever $\xi_t^C = 0$, we have $\xi_t^B = 0$. Intuitively, we insert a longer time of C staying at 0 than B on any visit to obtain a longer path for C. We can achieve this coupling because the probability of staying at 0 for C is bigger than B's probability of staying at 0 at any step.

Therefore, we have

$$E_xT_y \leqslant E_xT''_y \leqslant E_xT'_y \tag{3.1}$$

for all x and y.

To obtain a simpler expression than in Section 2.5, observe that we assumed $0 < y < x$ and $E_xT^*_y = E_{n-x}T^*_{n-y}$ by reflecting the paths of the random walk over the line through 0 and $n/2$ ($*$ represents either $'$, $''$ or nothing). Therefore

$$\begin{aligned} E_xT'_y &= E_x(V'_0 \wedge V'_y) + P_x(V'_0 < V'_y) \cdot E_0T'_y \\ &= (n-x)(x-y) + \frac{x-y}{n-y} \cdot \frac{y(n-y)\left[n\varepsilon + 2(1-\varepsilon)\right]}{n\varepsilon} \\ &= (x-y)(n-x+y) + \frac{2(x-y)y(1-\varepsilon)}{n\varepsilon} \\ &= E_xT_y + \frac{2(x-y)y(1-\varepsilon)}{n\varepsilon}. \end{aligned}$$

For the first line, we observe that we need $V'_0 \wedge V'_y$ steps no matter what to escape the interval $[0, y]$. If we hit y first, then no more steps are needed, and if we hit 0 first then we need $E_0T'_y$ more steps. In the second line, the first term is just the Gambler's Ruin hitting time, and the first factor in the second term is the Gambler's Ruin hitting probability of 0. The expression for $E_0T'_y$ is from Equation 2.14. In the last line, we use the Gambler's Ruin hitting time again for E_xT_y. As a check, we observe that if $= 1$, the second term goes away and the chain is equivalent to the unperturbed chain. Notice that if we group the terms containing $\frac{1-\varepsilon}{\varepsilon}$ in Equation 2.15, we obtain the same expression for $E_xT'_y$. Now we have

$$
\begin{aligned}
0 \leqslant E_xT''_y - E_xT_y \leqslant E_xT'_y - E_xT_y& \\
\leqslant \frac{2(x-y)y(1-\varepsilon)}{n\varepsilon}& \\
\leqslant \frac{2(n-y)y(1-\varepsilon)}{n\varepsilon} &\quad \leqslant \frac{1-\varepsilon}{2\varepsilon}n.
\end{aligned} \tag{3.2}
$$

In the last line we used the fact that $y(n-y) \leqslant n^2/4$ (maximized when $y = n/2$). Now we need a basic fact about maxima of sets. Let $\{a_k\}_{k=1}^N$ and $\{b_k\}_{k=1}^N$ be finite sets of real numbers, then

$$
\max_k(a_k + b_k) \leqslant \max_k a_k + \max_k b_k.
$$

This can be easily proved by using the definition of maximum. Applying this to our problem gives

$$
\begin{aligned}
1 &\leqslant \frac{\max_{x,y} E_xT''_y}{\max_{x,y} E_xT_y} \leqslant \frac{\max_{x,y}\left[E_xT_y + (1-\varepsilon)n/2\varepsilon\right]}{\max_{x,y} E_xT_y} \\
&\leqslant \frac{\max_{x,y}(E_xT_y) + \max_{x,y}(1-\varepsilon)n/2\varepsilon}{\max_{x,y} E_xT_y} \leqslant 1 + \frac{(1-\varepsilon)n/2\varepsilon}{\max_{x,y} E_xT_y} \\
&\leqslant 1 + \frac{(1-\varepsilon)n/2}{n^2/4} = 1 + \frac{2(1-\varepsilon)}{\varepsilon n}.
\end{aligned} \tag{3.3}
$$

This completes what we set out to prove since the right hand side tends to 1 as $n \to \infty$ – in fact, we proved an explicit bound on the difference.

4

Simulations

As discussed in Chapter 1, simulations on some perturbed random walks on the n-cycle give us some interesting observations on the merge times of some time-inhomogeneous Markov chains.

If we think about two walkers A and B that start at different positions on the n-cycle, we can simulate some perturbed chains with different kinds of perturbations. In most cases, we force A to start at site 1 for easier simulations, but B can start anywhere. With the horizontal axis being the number of states, and the vertical axis being the merge times, we are able to draw histograms for each kind of perturbation and compare it to the original histogram, the merge times for the simple random walk on the n-cycle.

4.1 Several Kinds of Perturbations

In this section, we will talk about four kinds of perturbations. In the first one, the perturbed site moves in the order of 1 to n. At step i, the walker will go clockwise with higher probability if it happens to be at state i mod n. $K_i^{\delta,r}$, the perturbed transition matrix at state i in the clockwise direction with intensity δ is

$$K_i^{\delta,r}(x,y) := \begin{cases} \frac{1}{2} & \text{if } |x-y| = 1 \text{ and } x \neq i \\ \frac{1}{2}+\delta & \text{if } x = i,\ y = i+1 \\ \frac{1}{2}-\delta & \text{if } x = i,\ y = i-1 \\ 0 & \text{otherwise,} \end{cases}$$

where addition and subtraction are on the cycle, so $n = 1-1$ and $1 = n+1$. Likewise, we define $K_i^{\delta,l}$ to be the transition probability with intensity δ of going counterclockwise at state i.

In this first case, we use $K_i^{\delta,r}$ as our transition matrix at state i. Our simulation for this kind of perturbation gives slightly faster merge times than those of the original unperturbed random walk on the n-cycle.

In the second kind of perturbation, let Z_t be the position of the perturbed site at time t. We simulate a case in which Z_0 starts at 1 at moves clockwise with higher probability in the first step. If $Z_t = Z_{t-1}+1$, then we choose $K_{t+1} = K_{Z_t}^{\delta,l}$, where K_{t+1} is the transition kernel at step $t+1$, so that the probability of moving counterclockwise is higher at state Z_t. Conversely, if $Z_t = Z_{t-1}-1$, we choose $K_{t+1} = K_{Z_t}^{\delta,r}$, so that the probability of moving clockwise is higher at state Z_t. Again, note that addition and subtraction are on the n-cycle, so $n = 1-1$ and $1 = n+1$. Essentially, we "guess" the position of walker A through simulation, and make our predicted state the perturbed site at every step. However, the actual paths are still independent of the sequence of the kernels chosen. We conjecture that this perturbation gives us a higher chance to push walker A back to where it starts at, so the walker might be trapped to some extent.

The third kind of perturbation is similar to the second one. In the previous case, the perturbed state is essentially doing a random walk with unequal probabilities. In this case, the perturbed state just does the simple random walk with equal probabilities to move clockwise or counterclockwise. We still "guess" the position of A at every step. If it went clockwise in the previous step according to our "guesses", we force it in counterclockwise direction with higher probability at the perturbed site, and vice versa.

Our simulations show that the merge times for the second and third kinds of perturbed chains are hardly different from the unperturbed chain. See Figure 4.1 for a comparison of the cases in this section.

Which starting position of walker B gives us the maximum $\bar{d}(t)$ that determines t_{merge} in the above cases? Intuition tells us that B has to be farthest from A on the n-cycle, which is not opposite to A's starting state 1, but rather right next to state 1. Even though the particles sit right next to each other, they cannot hit the same spot right away since they will jump over each other. The only way for the particles to meet is to travel around the n-cycle. The results from a small computer program agree with our intuition.

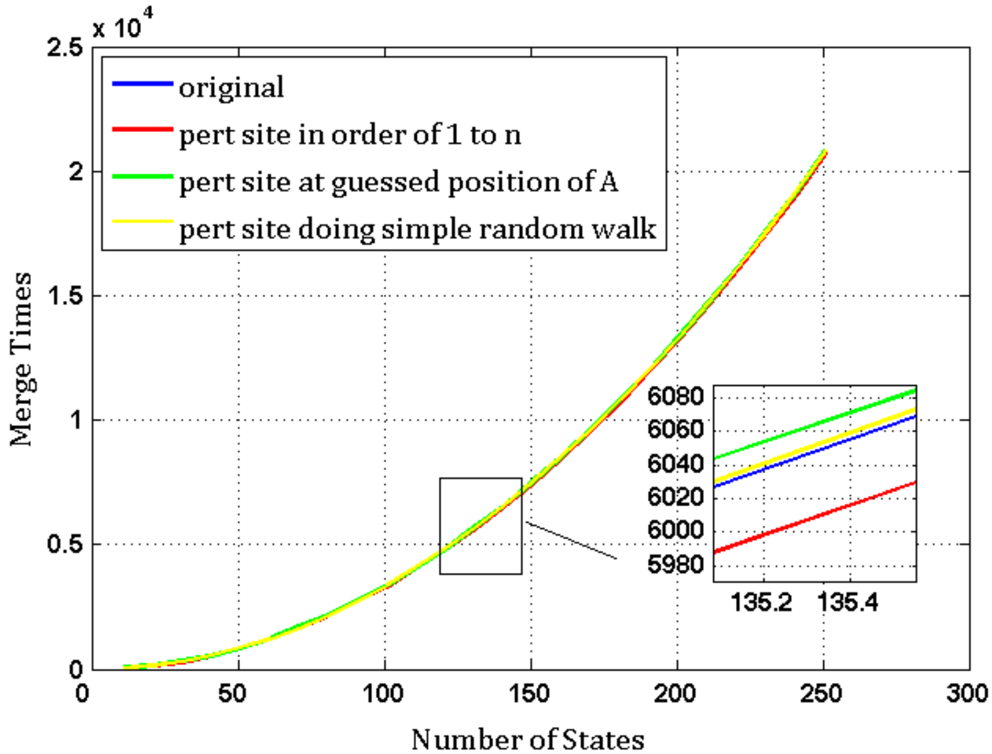


FIGURE 4.1: Several kinds of perturbations

4.2 Perturbed Site Alternating Between Two States

If walker A starts at state 1, it can reach state 3 on an even step or reach state n on an odd step at some point. Therefore, we perturb state 3 at odd steps(force A in counterclockwise direction with higher probability) and perturb state n at even steps (force A in clockwise direction with higher probability), so that A might get trapped for a while. The chain is time-homogeneous if we look at pairs of steps, with transition matrices at time $2t$ is $P^{2t} = \left[K_3^{\delta,\ell} K_n^{\delta,r}\right]^t$, but still time-inhomogeneous if we look at single steps. See Figure 4.2.

We can also make the perturbed state alternate between $n/4$ and $n - n/4$. Again, at odd steps, we perturb state $n/4$ (force in counterclockwise direction with higher

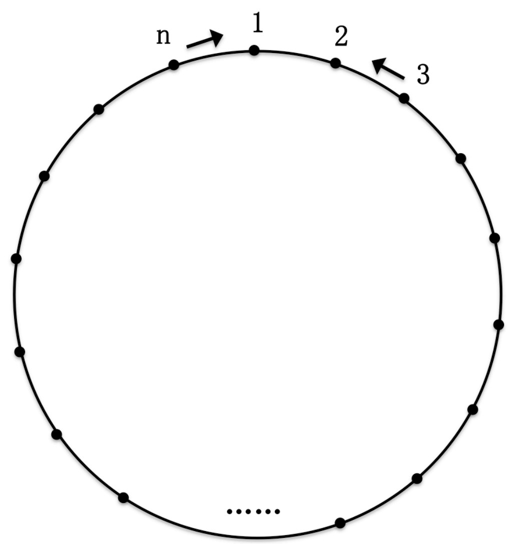


FIGURE 4.2: Perturbed state alternating between 3 and n

probability), and at even steps, we perturb state $n-n/4$ (force in clockwise direction with higher probability). See Figure 4.3. Interestingly, this type of perturbations gives much longer merge times than if we alternate the perturbed state between 3 and n.

In order to see if the starting position of A affects the merge times, we simulate the case in which there are no restrictions on both A and B's starting positions, referred to as "alternating between $n-n/4$ and $n/4$ with random A" in Figure 4.4. Our simulation shows that A starting at anywhere yields very similar results as if A starts at state 1.

There are two ways of combining the above two kinds of perturbations. "Combo1" in Figure 4.4 refers to alternating the perturbed state between 3 and n for $(1/4)n^2$ steps (when A is less likely to reach $n/4$ or $n-n/4$ by that time), and then switch to alternating the perturbed state between $n/4$ and $n-n/4$. "Combo2" in Figure

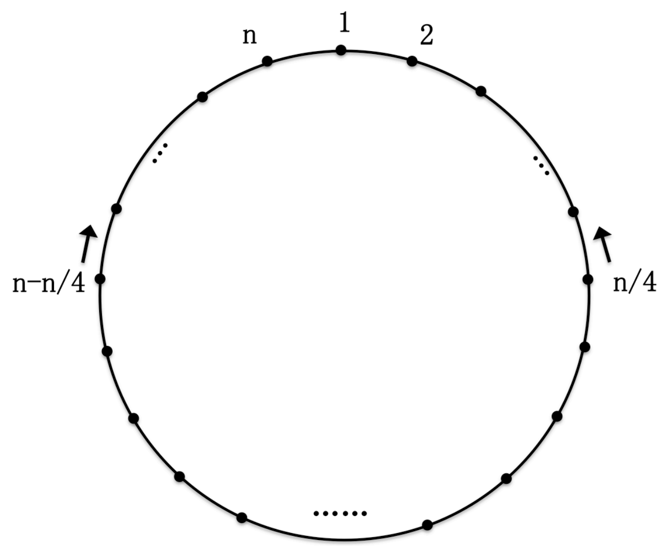


FIGURE 4.3: Perturbed state alternating between n/4 and n-n/4

4.4 simply alternates the perturbed state between these four locations. Both have longer merge times than "alternating between 3 and n". "Combo 1" is the slowest overall and the merge times in "Combo 2" are between "alternating between 3 and n" and "alternating between $n - n/4$ and $n/4$, which is not surprising. See Figure 4.4 for a comparison of merge times for cases in this section.

4.3 Different Angles on a Clock

If we think of the n-cycle as a clock, we can analogize alternating the perturbed state between $n/4$ and $n - n/4$ as alternating between 3 o'clock and 9 o'clock. Interested in whether the results will be different if we change the angles between the two states, we simulate the cases of alternating between 2 and 10 o'clock, 4 and 8 o'clock, 1 and 11 o'clock, as well as 5 and 7 o'clock. The results for each one turn out to be similar.

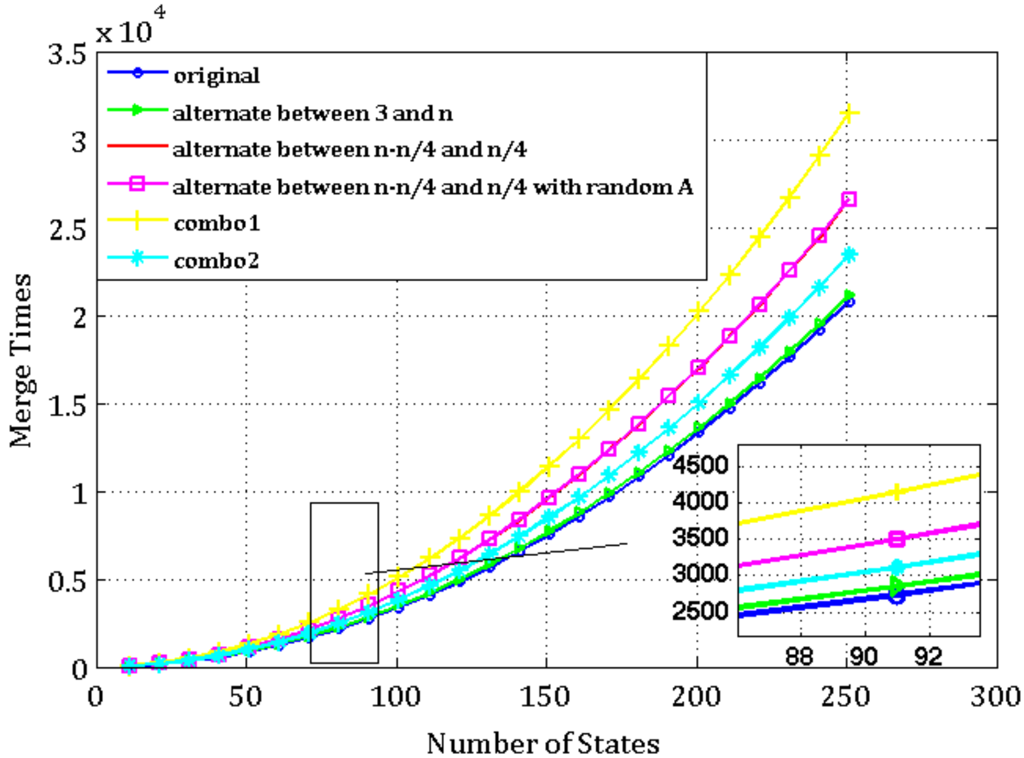


FIGURE 4.4: Perturbed state alternating between two states

See Figure 4.5.

4.4 Varying Stickiness at Site 0

We simulate the merge times of the perturbed chain with varying stickiness at site 0 and compared to those of the unperturbed chain. The simulation exhibit almost identical results for both chains. See Figure 4.6.

This provides a motivation and experimental evidence for proving the hitting times of the perturbed chain in this case asymptotically converge to those of the unper-

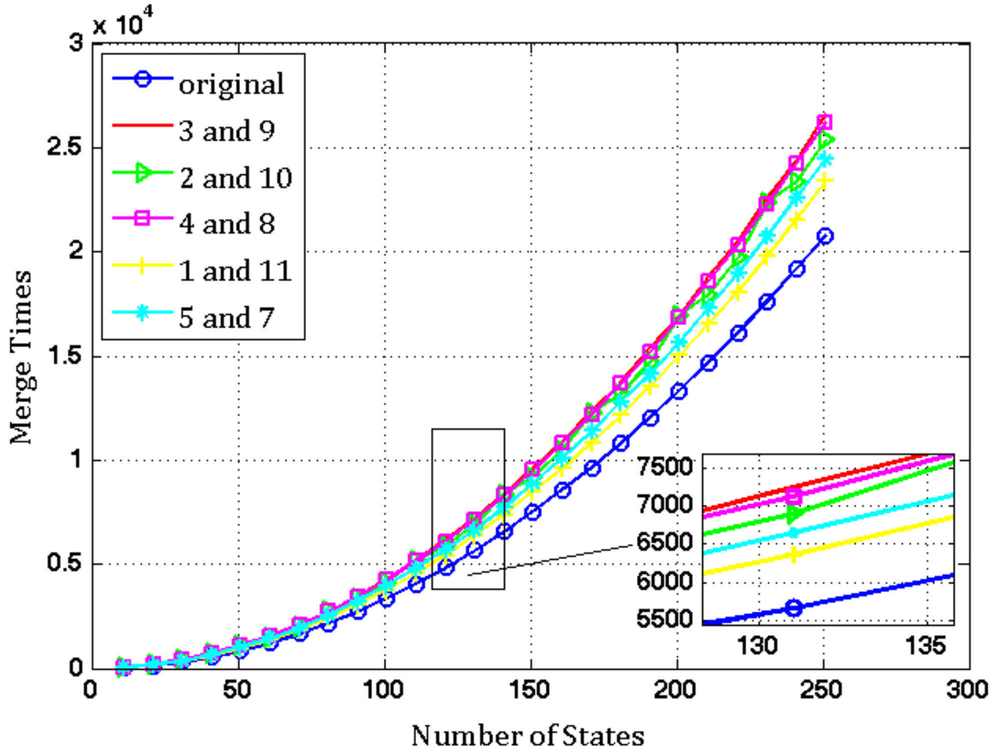


FIGURE 4.5: Perturbed state alternating between two states with different angles

turbed chain in Chapter 3.

4.5 Conclusions for Simulations

As the upper and lower bounds for the merge times of the unperturbed chains are both on the order of n^2 and the shapes of the simulated curves for the perturbed chains appear parabolic, we conjecture that the merge times for these perturbed chains are on the same order. We therefore fit the merge times in every case to $t_{merge}(n) = \alpha n^2$. By comparing the magnitude of the α's, we summarize the speeds of merging for each case in an increasing order in Table 4.1.

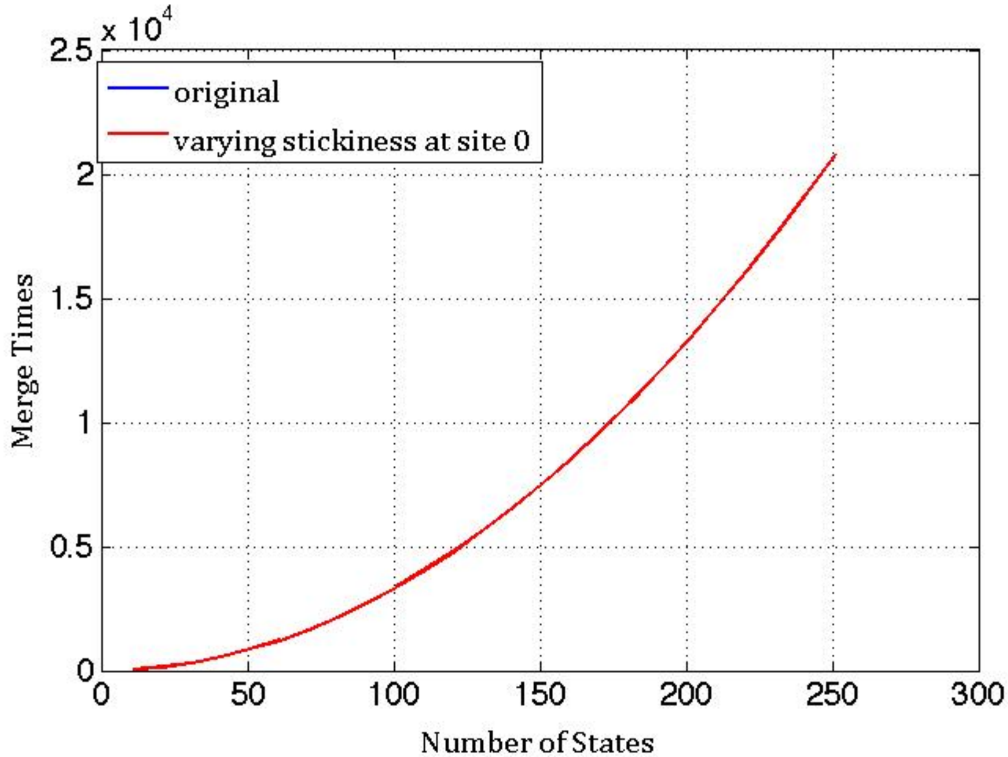


FIGURE 4.6: Varying stickiness at site 0

chain	α fit
unperturbed(comparing to stationary distribution)	0.1897
perturbation site at guessed position of A	0.3298
perturbation site in the order of 1 to n	0.3284
original(unperturbed_comparing to the chain itself)	**0.3299**
varying stickiness at site 0	0.3299
perturbation site doing simple random walk	0.3302
perturbation site alternating between 3 and n	0.3362
"1 and 11 o' clock"	0.3707
combo2(perturbation site alternating between 3, n, $n - n/4$, n/4)	0.3720
"5 and 7 o'clock"	0.3884
"2 and 10 o'clock"	0.4068
"4 and 8 o'clock"	0.4166
perturbation site alternating between $n - n/4$ and $n/4$ or "3 and 9 o'clock"	0.4209
perturbation site alternating between $n - n/4$ and $n/4$ with random A	0.4219
combo1(first $(1/4)n^2$ steps 3 and n, afterwards $n - n/4$ and $n/4$)	0.5000

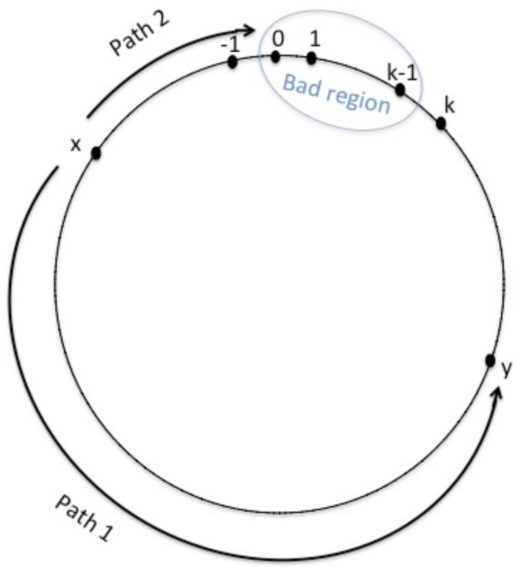
Table 4.1: The values of $\alpha' s$ when fitting merge times to $t_{merge}(n)$=αn^2.

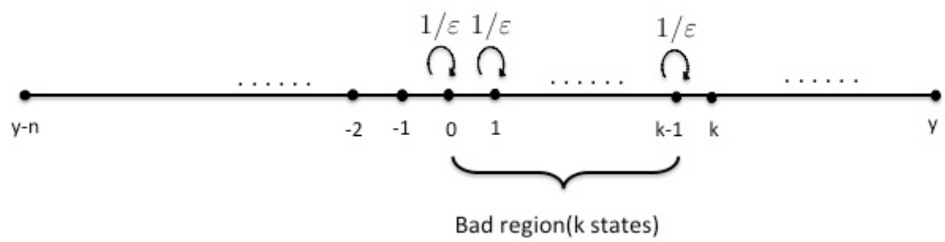
5

Future Work

As an extension of the previous question, suppose we have k perturbed states glued together: $0, 1, \cdots, k-1$. See Figure 5. Instead of making only one site "sticky", we want to make the entire region a trouble for the walker. Note that in this case, for simplicity reasons, 0 and n are separate states and no longer represent the same state, so we have $n+1$ states in total.

If the walker takes on path 1, then it enters the bad region from the side 0. We can compute the probability of escaping the bad region from the side 0 or side $k-1$ using Gambler's Ruins formulas. If the walker escapes through side 0 and reaches n, or -1, it will either go straightly to y from -1, or enter the bad region through side 0 again, and take E_0T_y to eventually get to y. If the walker escapes through side $k-1$ and reaches k, it could either come back to the region or go straightly to y. If it enters back into the region through side $k-1$, the probability of exiting the region from 0 or $k-1$ are now different from what we previously have, since the starting position is $k-1$ instead of 0. We have





$$E_0T_y = \frac{1}{k}\{E_0(V_k|V_k < V_{-1}) + \mathbb{P}_k(V_{k-1} < V_y)[E_k(V_{k-1}|V_{k-1} < V_y) + E_{k-1}V_y]$$
$$+ \mathbb{P}_k(V_y > V_{k-1})E_k(V_y|V_y > V_{k-1})\}$$
$$+ \frac{k-1}{k}\{E_0(V_{-1}|V_k > V_{-1}) + \mathbb{P}_{-1}(V_0 < V_y)[E_{-1}(V_0|V_0 < V_y) + E_0T_y]$$
$$+ \mathbb{P}_{-1}(V_0 > V_y)E_{-1}(V_y|V_0 > V_y)\},$$

where

$$E_{k-1}V_y = P_{k-1}(V_k < V_0)[E_{k-1}(V_k|V_k < V_0) + P_k(V_{k-1} < V_y)E(V_{k-1}|V_{k-1} < V_y) + E_{k-1}V_y] + P_{k-1}(V_0 < V_k)[E_{k-1}(V_0|V_0 < V_k) + E_0T_y].$$

Once we get E_0T_y, previous approach of coupling should still apply. However, we are seeking other alternatives because even the simplified expression for E_0T_y are too cumbersome.

Bibliography

[1] D. A. Levin, Y. Peres, and E. L. Wilmer, *Markov Chains and Mixing Times*. American Mathematical Society, 2009.

[2] D. Griffeath, *Uniform coupling of non-homogeneous Markov chains*, Journal of Applied Probability **12** (1975) no. 4, 753–762.

[3] L. Saloff-Coste and J. Zúñiga, *Convergence of some time inhomogeneous Markov chains via spectral techniques*, Stochastic Processes and their Applications **117** (2007) 961–979.