Spring 5-1-2015

Efficient Coupling for Random Walk with Redistribution

Elizabeth Tripp
elizabethtripp92@gmail.com

Efficient Coupling for Random Walk with Redistribution

Elizabeth Anne Tripp, B.S.

University of Connecticut, 2015

ABSTRACT

What can be said on the convergence to stationarity of a finite state Markov chain that behaves 'locally' like a nearest-neighbor random walk on $\mathbb{Z}$? In this work, we looked to obtain sharp bounds for the rate of convergence to stationarity for a particular non-symmetric Markov chain. Our Markov chain is a variant of the simple symmetric random walk on the state space $\{0, \ldots, N\}$ obtained by allowing transitions from 0 to J_0 and from N to J_N. We first looked at the case where J_0 and J_N are fixed, deterministic sites; we then also considered the case where J_0 and J_N are repeatedly sampled from some given probability distribution. For each of these two cases, we constructed an efficient coupling for the model, giving an intuitive and probabilistic explanation for the rates of convergence as well as providing sharp, computable, and non-asymptotic bounds.

Efficient Coupling for Random Walk with Redistribution

Elizabeth Anne Tripp

B.S. Mathematics

B.S. Statistics

A Undergraduate Honors Thesis

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Bachelor of Science

at the

University of Connecticut

2015

APPROVAL PAGE

Bachelor of Science Honors Thesis

Efficient Coupling for Random Walk with Redistribution

Presented by

Elizabeth Anne Tripp

Thesis Advisor _____
Dr. Iddo Ben-Ari

Honors Advisor _____
Dr. Luke Rogers

University of Connecticut

2015

ACKNOWLEDGMENTS

First and foremost, I would like to thank my thesis advisor and research mentor, Dr. Iddo Ben-Ari, for all of the help and guidance he has provided over the past three years. He has always been extremely generous with his time and willing to make every effort to aid me in any way possible. His encouragement throughout our time together often convinced me to pursue opportunities that I would never have considered otherwise, opportunities that have had a profound impact on not only my undergraduate experience but also my future goals. I would not be the student I am today without his mentorship, and I am so grateful to have had this opportunity to work with him so closely.

I would also like to acknowledge the entire Mathematics and Statistics Departments here at the University of Connecticut. The support I have received from numerous professors in each department throughout the past four years has significantly shaped my experiences, for which I am incredibly grateful. In particular, I would like to thank Drs. Luke Rogers and Nalini Ravishanker, my academic advisors from the Mathematics and Statistics Departments (respectively), for all of their guidance and support throughout the past four years. I would also like to thank Dr. William Abikoff, whose mentorship has had a significant impact on my undergraduate career, and Hugo Panzo, a graduate student in the Mathematics Department who aided me greatly this past summer while conducting this project.

Additionally, I must acknowledge the University of Connecticut's Office of Under-

graduate Research, which offered so much support throughout the process of conducting this project. In particular, I am extremely grateful to have received the Trimble Family Award, a Summer Undergraduate Reserach Fund grant, which made it possible for me to conduct the research for this project last summer. Their support made this project possible, for which I am extremely grateful.

Furthermore, I would also like to thank my friends and classmates, who have supported me immeasurably throughout the past four years. Their friendship and kindness has made my time at UConn so special. In particular, I would especially like to thank my suitemates, Kimberly Rebello and Selena Roy, who have stayed by my side in times of both stress and laughter.

Finally, I must thank my parents, Norine and Michael Tripp, and my sister, Hannah Tripp. Their confidence in me and constant guidance have inspired me throughout my life. Their boundless love and support make everything seem possible.

Contents

Ch. 1. Introduction 2
1 Markov Chains . . . 2
2 Coupling . . . 7

Ch. 2. Examples 10
1 Random Walk on an Interval . . . 10
2 Random Walk on a Cycle . . . 14

Ch. 3. Our work 17
1 Background . . . 17
2 Our Model . . . 19
3 Our Results . . . 20
3.1 Assumptions and Preparation . . . 20
3.2 Deterministic Redistribution . . . 22
3.3 Random Redistribution . . . 25
4 Proofs . . . 26
4.1 Coupling Regimes . . . 26
4.2 Deterministic Redistribution . . . 26
4.3 Random Redistribution . . . 34

Bibliography 38

Preface

The focus of this thesis is to provide sharp bounds on the rate of convergence to stationarity of a particular discrete and irreversible Markov chain via probabilistic methods, where the lack of symmetry is created by the model's redistributive behavior at the end-points. In the first chapter, we will provide a theoretical introduction to the the main work. Specifically, we will define an efficient Markovian coupling of Markov chains and present the main theorems employed to obtain our results. Chapter 2 will contain simple examples of Markov chains with both efficient and non-efficient couplings, and our results will be presented in Chapter 3.

Chapter 1

Introduction

1 Markov Chains

First, some necessary definitions and theorems [11]:

Definition 1.1. Let S be a set. A *discrete-time stochastic process* X on *state space* S is a sequence of random variables indexed by the nonnegative integers $X = (X_t : t \in \mathbb{Z}_+ = \{0, 1, 2, \dots\})$, each taking values in S.

In this work, we will only consider discrete-time stochastic processes on finite state spaces.

Definition 1.2. A stochastic process X on state space S is called a *Markov chain* if there exists a function $p : S \times S \to [0, 1]$ such that, for all $t \in \mathbb{Z}_+$ and $x_0, ...x_{t-1}, x, y \in S$, the following holds:

$$P(X_{t+1} = y | X_t = x, X_{t-1} = x_{t-1}, \dots, X_0 = x_0) = p(x, y) \tag{1.1}$$

or, equivalently:

$$P(X_{t+1} = y | X_t = x, X_{t-1} = x_{t-1}, \ldots, X_0 = x_0) = P(X_{t+1} = y | X_t = x). \tag{1.2}$$

The function p is called the *transition function* and can be viewed as an $|S| \times |S|$ matrix.

In simple terms, a Markov chain is a process with "short-term" memory, meaning that the only information from the history of the process relevant when determining its distribution in the future is its current state. It is important to note that p determines the distribution of X **conditioned** on X_0 (in other words, it only determines the probabilities of transitions, not the initial configuration). The distribution of X_0 is called the *initial distribution* of X. However, once X_0 is also specified, the distribution of X is uniquely determined.

We write P_x for the distribution of X conditioned on $X_0 = x$. More generally, if μ is a probability distribution on S, that is $\mu(x) \geq 0$ for all $x \in S$ and $\sum_{x \in S} \mu(x) = 1$, then P_μ denotes the distribution of X when X_0 has law μ. We note that

$$P_\mu = \sum_{x \in S} \mu(x) P_x.$$

Definition 1.3. If a probability distribution μ satisfies $P_\mu(X_1 \in \cdot) = \mu$, then we call μ a *stationary distribution for* X. This is equivalent to the condition

$$P_\mu(X_t \in \cdot) = \mu \ , \forall t \in \mathbb{Z}_+. \tag{1.3}$$

We write p^t to denote the t-th power of p viewed as a matrix. That is:

$$p^0(x,y) = I,\ p^{t+1}(x,y) = \sum_{k\in S} p^t(x,k)p(k,y) \tag{1.4}$$

where I denotes the identity matrix.

It is easy to show that for all $s,t \in \mathbb{Z}_+$ and $x,y \in S$

$$P(X_{t+s} = y | X_s = x) = p^t(x,y) \tag{1.5}$$

Definition 1.4. A Markov chain X is *irreducible* if for any $x,y \in S$, there exists $t \in \mathbb{Z}_+$ such that $p^t(x,y) > 0$.

An irreducible chain is thus one that can (eventually) transition from any state to any state.

A corollary to this is the following:

Proposition 1.5. *Suppose that X is an irreducible Markov chain on a finite state space S. Then it possesses a unique stationary distribution.*

For a proof, see [5], p. 22.

Definition 1.6. A Markov chain is called *aperiodic* if the greatest common divisor of $\{t \geq 1 : p^t(x,x) > 0\}$ is 1 for all states $x \in S$.

This means that no state can appear only on a lattice (for instance, only at even time intervals). The Markov chain X with state space $\mathbb{Z}$ and transition function $p(x,x+1) = p(x,x-1) = \frac{1}{2}$ is an example of a periodic Markov chain, as it can return to its initial position only at even times.

Here is a sufficient condition for aperiodicity:

Proposition 1.7. *Suppose that X is irreducible and that, for some x, $p(x,x) > 0$. Then X is aperiodic.*

Proof. Fix some state y. Because of the irreducibility of X, there exists times t_0 and t_1 with $p^{t_0}(y,x), p^{t_1}(y,x) > 0$. Therefore $p^{t_0+t_1}(y,y) \geq p^{t_0}(y,x)p^{t_1}(x,y) > 0$. However, since $p(x,x) > 0$, it also follows that

$$p^{t_0+1+t_1}(y,y) \geq p^{t_0}(y,x)p(y,y)p^{t_1}(x,y) > 0.$$

In fact, $p^t(y,y) > 0$ for all $t \geq t_0 + t_1$. The result follows.

□

If a Markov chain X on state space S with transition function p is not aperiodic, one can consider the *lazy* version of X, which has transition function $\frac{1}{2}(p+I)$, where I is the identity matrix. More concretely, the transition function of the lazy chain can be described as follows: flip a fair coin, and if it lands 'Heads', then the chain doesn't move; if it lands 'Tails', sample according to p. This allows the chain to 'stay put' with probability at least $\frac{1}{2}$, thereby making the chain aperiodic. Examples of both irreducible and aperiodic Markov chains can be found in Chapter 2.

Definition 1.8. The *total variation distance* between two probability distributions μ and ν on S is defined by

$$\|\mu - \nu\|_{TV} = \max_{A \subset S} |\mu(A) - \nu(A)| \tag{1.6}$$

Note that $\|\mu-\nu\|_{TV} = \max_{f \in A} \int f d\mu - \int f d\nu = \frac{1}{2}\sum_{x \in S} |\mu(x)-\nu(x)|$, where $A = \{f : S \to [0,1]\}$, and thus the total variation distance can be viewed as the (normalized)

ℓ^1-norm of of the difference between the functions μ and ν. Total variation is a metric on probability measures on S.

We define

$$d_t(x,y) = \|P_x(X_t \in \cdot) - P_y(X_t \in \cdot)\|_{TV}, \tag{1.7}$$

where $P_x(X_t \in \cdot)$ denotes the probability distribution of X_t under P_x. Note that $d_t(\cdot,\cdot)$ is a metric on S. We let $d_t = \sup_{x,y} d_t(x,y)$.

Suppose now that π is a stationary distribution for X. Then observe that

$$\begin{aligned}
\|P_x(X_t \in \cdot) - \pi\|_{TV} &= \frac{1}{2}\sum_{y\in S} |P_x(X_t = y) - P_\pi(X_t = y)| \\
&= \frac{1}{2}\sum_{y\in S} |\sum_{z\in S} \pi(z)(P_x(X_t = y) - P_z(X_t = y))| \\
&\leq \frac{1}{2}\sum_{y\in S}\sum_{z\in S} \pi(z)|P_x(X_t = y) - P_z(X_t = y)| \\
&\leq \frac{1}{2}\sum_{z\in S} \pi(z)\sum_{y\in S} |P_x(X_t = y) - P_z(X_t = y)| \\
&\leq d_t
\end{aligned} \tag{1.8}$$

Theorem 1.9 (Ergodic Theorem for Markov Chains)**.** *If a finite-state Markov chain X is aperiodic and irreducible, then there exists $c > 0$ and $\rho \in (0,1)$ such that $d_t \leq c\rho^t$.*

For proof, see [5], p. 264.

In other words, in light of Proposition 1.5 and Equation (1.8), an irreducible and aperiodic Markov chain converges to its unique stationary distribution exponentially fast in total variation.

2 Coupling

Definition 1.10. A *coupling of Markov chains* with transition matrix p is a process $(X, Y) = ((X_t, Y_t) : t \in \mathbb{Z}_+)$ on state space $S \times S$ with the following properties:

1. Each of the *marginal processes* $X = (X_t : t \in \mathbb{Z}_+)$ and $Y = (Y_t : t \in \mathbb{Z}_+)$ is a Markov chain with transition matrix p.

2. If $X_s = Y_s$, then $X_t = Y_t$ for all $t \geq s$.

Definition 1.11. A coupling (X, Y) is *Markovian* if it is a Markov chain on state space $S \times S$, and, in addition, it satisfies:

$$P(X_{t+1} = x' | (X_t, Y_t) = (x, y), \ldots, (X_0, Y_0) = (x_0, y_0)) = P(X_{t+1} = x' | X_t = x), \quad (1.9)$$

and

$$P(Y_{t+1} = y' | (X_t, Y_t) = (x, y), \ldots, (X_0, Y_0) = (x_0, y_0)) = P(Y_{t+1} = y' | Y_t = y). \quad (1.10)$$

The conditions (1.9) and (1.10) imply not only that the marginals are Markovian with respest to their own history, but also that they are Markovian with respect to $\mathbf{q}$, where q denotes the transition function of (X, Y). Then (1.9) and (1.10) are equivalent to

$$\sum_{y'} q((x, y), (x', y')) = p(x, x'), \text{ and } \sum_{x'} q((x, y), (x', y')) = p(y, y'),$$

respectively.

A standard example of a Markovian coupling is when X and Y are indepedent (until they meet). In fact, this coupling is used in the standard proof of Theorem 1.9.

Definition 1.12. The *coupling time* of a coupling of Markov chains is the first time the chains meet:

$$\tau = \inf\{t \geq 0 : X_t = Y_t\} \tag{1.11}$$

Theorem 1.13 (Coupling bound)**.** *Let (X, Y) be a coupling with $(X_0, Y_0) = (x, y)$. Then*

$$d_t(x, y) \leq P_{x,y}(\tau > t) \tag{1.12}$$

Proof. From equation 1.7, we have

$$d_t(x, y) = \max_{A \subset S} \left(P_x(X_t \in A) - P_y(X_t \in A)\right). \tag{1.13}$$

Fix $A \subset S$ and let (X, Y) be any coupling. Then clearly

$$P_x(X_t \in A) - P_y(X_t \in A) = E_{x,y}[\mathbf{1}_A(X_t) - \mathbf{1}_A(Y_t); \tau > t] \leq P_{x,y}(\tau > t).$$

□

The challenge is to construct a coupling that gives not only an upper bound, but also a comparable lower bound, on d_t. This motivates the following definition:

Definition 1.14. A coupling is *efficient* if for some $x, y \in S$,

$$\frac{1}{t} \ln P_{x,y}(\tau > t) \sim \frac{1}{t} \ln d_t \tag{1.14}$$

as $t \to \infty$.

We comment that this definition is slightly weaker than the definition given in [4], which can be stated as follows:

A coupling is *efficient* if there exists x, y such that $cd_t \leq P_{x,y}(\tau > t) \leq c'd_t$ for all $t > 0$.

Our definition is clearly weaker, as we require asymptotic equivalence at the logarithmic scale. This is because, in some cases, the coupling we constructed has a polynomial correction to the exponential decay, an effect which vanishes at the logarithmic scale. [4] gives an example of a Markov chain for which an efficient, Markovian coupling does not exist.

Here is a simple and sufficient condition for the efficiency of a coupling:

Proposition 1.15. *Suppose that $S = \{0, \ldots, N\}$, and let (X, Y) be a coupling satisfying $X_t \leq Y_t$ for all t. Then the coupling is efficient.*

Proof. Let $f : S \to [0, 1]$ be the function $f(k) = k/N$. Then clearly,

$$d_t(X_0, Y_0) \geq E_{X_0,Y_0}[f(Y_t) - f(X_t)] = E_{X_0,Y_0}[f(Y_t) - f(X_t); \tau > t] \geq \frac{1}{N} P_{X_0,Y_0}(\tau > t).$$

□

Note: the processes that we will study are such that the condition for Proposition 1.15 (above) fails.

Chapter 2

Examples

In this chapter, we will present several examples of Markov chains relevant to our work. In all cases, our state space will be $S = S_N = \{0, \ldots, N\}$ for some $N \geq 2$. We begin with two classical models, the random walk on an interval and the random walk on the cycle, and construct efficient couplings for each. The coupling schemes introduced for each model will be the building blocks in the construction of the efficient coupling for the random walk with redistribution in Chapter 3. There, we will use both schemes, switching from one to the other according to the state of the system.

1 Random Walk on an Interval

Consider a Markov chain (X_t) on S_N with transition function q defined so that the chain moves one space to the right or left at each time interval, each with probability $\frac{1}{2}$, where, if the chain tries to move outside the interval (i.e. to -1 or $N+1$), it merely

stays at the end point. This process is called a *random walk* on an interval. The *lazy* random walk (see Chapter 1, Section 1) on S_N is the Markov chain with transition function p given as follows:

if $x \in S_N \setminus \{0, N\}$,

$$p(x,y) = \begin{cases} \frac{1}{4} & \text{if } y = x+1 \\ \frac{1}{2} & \text{if } y = x \\ \frac{1}{4} & \text{if } y = x-1 \end{cases}$$

and if $x \in \{0, N\}$,

$$p(x,y) = \begin{cases} \frac{3}{4} & \text{if } y = x \\ \frac{1}{4} & \text{if } y = x \pm 1, \end{cases}$$

where the $\pm$ is '+' if $x = 0$ and '$-$' otherwise. This process is clearly irreducible and aperiodic.

We will now construct a coupling for this model. Let X_t and Y_t be two Markov chains with state spaces and transition functions as described above, where $X_0 = x$ and $Y_0 = y$ for some $x, y \in S_N$. Flip a coin at each time step to determine whether X_t stays put or moves according to p; if the chain moves, flip another coin to determine whether the chain moves to the right or left (where, if X_t is 0 or N, moving to -1 or $N+1$, respectively, implies staying put). Have Y_t follow the same rule. Thus, at each transition, X_t and Y_t will both be either staying put, moving in the same direction, or, if either is at $\{0, N\}$, moving one step closer together. Thus, the chains are guaranteed to couple at either 0 or N.

We observe that the coupled process is also a Markov chain. Furthermore, it is clear from the construction that, conditioned on $(X_0, Y_0) = (x, y)$, the probability

that $X_1 = x'$ is equal to $p(x, x')$. A similar statement holds for Y. Therefore, this is a Markovian coupling.

We know from Proposition 1.15 that *any* coupling that guarantees the preservation of the order of the two Markov chains up to the coupling time will be efficient. Since this coupling satisfies the condition, it is efficient.

We will refer to this kind of coupling (both copies moving in the same direction when possible) as *rigid coupling*. To discuss the coupling time for this and subsequent processes, we define the following notation:

Definition 2.1. Suppose that $Z = (Z_n : n \in \mathbb{Z}_+)$ is the lazy random walk on $\mathbb{Z}$: that is, Z jumps to a neighboring site with probability $\frac{1}{4}$ and stays put with probability $\frac{1}{2}$. Let $T(L)$ denote the exit time of Z from the set $\{1, \ldots, L\}$:

$$T(L) = \inf\{t \geq 0 : Z_t = 0 \text{ or } Z_t = L + 1. \tag{2.1}$$

Next, we recall a well-known classical result that will serve for estimating the coupling time. Write Q_z for the distribution of Z starting from $Z_0 = z$. Let $T(L)$ be as defined in 2.1, and let

$$\lambda(L) = \frac{1}{2}(\cos(\frac{\pi}{L+1}) + 1). \tag{2.2}$$

Then we have the following well-known lemma:

Lemma 2.2.

1. *There exists a coupling (Z, Z') such that Z' is lazy random walk starting from $Z_0 = \lfloor (L+1)/2 \rfloor$ and $T(L) \leq T'(L)$, where T' is the exit time of Z' from $\{1, \ldots, L\}$.*

2. For any $z \in \{1, \ldots, L\}$,

$$Q_z(T(L) > t) = \frac{2}{L+1} \cot(\frac{\pi}{2(L+1)}) \sin(\frac{\pi}{L+1} z) \lambda(L)^t + \lambda_2(L)^t O(1), \quad (2.3)$$

where $|\lambda_2(L)| < \lambda(L)$.

In light of part (1) of the Lemma, in what follows (with the exception of the proof for Lemma 2.2), we will abuse notation and write $T(L)$ for the distribution of the exit time of Z from $\{1, \ldots, L\}$ starting from $\lfloor (L+1)/2 \rfloor$.

We now return to the coupling presented above. Assuming that $y = x + 1$, it is easy to see that the coupling time is equal to $T(N)$.

Proof.

1. Suppose L is odd so that Z' starts from the unique center $\frac{L+1}{2}$ and Z starts from $z \in \{1, ..., L\}$. Assume, without loss of generality, that $z < \frac{L+1}{2}$. If $\frac{L+1}{2} - z$ is even, we run a reflection coupling (see description in the next section) until Z exits or Z and Z' meet at $z + \frac{1}{2}(\frac{L+1}{2} - z)$ and continue together until exiting. In either case, $T(L) \leq T'(L)$. If $\frac{L+1}{2} - z$ is odd, the method outlined in the description of rigid coupling can be employed. Namely, we toss two fair coins, A and B, and have A determine whether Z or Z' moves while B determines which way it moves. After this procedure, either Z will have exited or $Z' - Z \geq 0$ is even, in which case we can continue with the reflection coupling as before. Now, suppose that L is even. In this case, there are two centers, $\frac{L+1}{2} \pm \frac{1}{2}$. Assume, without loss of generality, that $z < \frac{L+1}{2} - \frac{1}{2}$. If $\frac{L+1}{2} - \frac{1}{2} - z$ is odd, then have Z' start at $\frac{L+1}{2} + \frac{1}{2}$; otherwise, have Z' start at $\frac{L+1}{2} - \frac{1}{2}$. This ensures that $Z' - Z$ is even, so we can run the reflection coupling as before. For a given z, this only establishes that $T(L) \leq T'(L_0)$ for Z' starting from

a particular center. Hence, the exit time starting from z is stochastically dominated by the exit time starting from that particular center. However, since the exit times starting from either center have the same distribution, this is of no consequence, and the stochastic dominance holds for both centers.

2. Observe that the sub-Markovian transition function for Z killed outside $\{1, \ldots, L\}$ is reversible and irreducible on $\{1, \ldots, L\}$. From the Perron-Frobenius Theorem, the largest eigenvalue is simple. The corresponding eigenvector, the Perron eigenvalue, is (without loss of generality) strictly positive on $\{1, \ldots, L\}$. By symmetry, there exists an orthonormal basis with respect to the counting measure consisting of eigenvectors, including the normalized Perron eigenvector as an element. It follows that the Perron root is the unique element in the basis which does not change signs. Finally, note that the function $\sin(\frac{\pi}{L+1}x)$ on $\{1, \ldots, L\}$ and zero elsewhere is an eigenfunction for this transition function, strictly positive on $\{1, \ldots, L\}$. Thus, this a Perron eigenvector. In addition, as is easy to see, the eigenvalue is $\lambda(L) = \frac{1}{2}(\cos\frac{\pi}{L+1} + 1)$. A straightforward computation shows that the ℓ^2 normalized Perron eigenfunction is $\sqrt{\frac{2}{L+1}}\sin(\frac{\pi}{L+1}x)$, and that $\sum_{x=1}^{x=L}\sin(\frac{\pi}{L+1}x) = \cot(\frac{\pi}{2(L+1)})$. The result now follows from the spectral theorem. □

2 Random Walk on a Cycle

Here, we consider essentially the same example as before, but we alter the behavior at the endpoints, making them "neighbors": that is, N is identified as the left neighbor of 0, and 0 is identified as the right neighbor of N. This process is known as a random

walk on an n-cycle, with $n = N + 1$. If n is even, the process will be periodic with period 2, as can be easily seen. Thus, to avoid this issue, we again consider the *lazy random walk*, which has the following transition matrix:

$$p(x, y) = \begin{cases} \frac{1}{2} & \text{if } y = x \bmod N \\ \frac{1}{4} & \text{if } y = x + 1 \bmod N \\ \frac{1}{4} & \text{if } y = x - 1 \bmod N \end{cases}$$

If we try to adapt the rigid coupling scheme from the last section (namely, having the two copies move in the same direction simultaneously), we end up with a coupling in which the distance between the two chains will remain constant and the processes will never meet.

We will introduce a different coupling, the *reflection (or mirror) coupling*. We may assume without loss of generality that $X_0 = 0$. In order to make the discussion simpler, we will also assume that N is odd and $Y_0 = y \in 2, \ldots, N - 1$ is even. The coupling is given as follows: at each time unit, we flip a fair coin, according to which we decide whether the two copies stay put, or move. In the latter case, we flip another fair coin, and if it lands 'Heads', then we increase X by 1, modulo N, and decrease Y by 1, modulo N; if the coin lands 'Tails', we increase Y by 1 and decrease X by 1 (both, modulo N). As before, this coupling is Markovian.

We observe that the distance between the two copies, $Y_t - X_t$, which is initially equal to y, stays put (i.e. does not change) with probability $\frac{1}{2}$, increases by 2 with probability $\frac{1}{4}$, and decreases by 2 with probability $\frac{1}{4}$, all independently of the past, until either the distance is 0 or $N + 1$ (note that $2 \leq Y_0 - X_0 = y < N + 1$ and both y and $N + 1$ are even). Thus, the two copies will meet at the same time that

a lazy random walk would exit the interval $\{1, \ldots, (N-1)/2\}$. Letting $T(L)$ with $L = \frac{N-1}{2}$ denote the exit time from this interval, as shown in Theorem 1.9, there exists a constant $c > 0$, depending on y and N, such that

$$P(T(L) > t) \sim c\lambda^t \tag{2.4}$$

where $\lambda = \frac{1}{2}(\cos(2\pi/(N+1)) + 1)$.

In particular, by the coupling inequality 1.13, there exists some $c' > 0$ such that $d_t \leq c'\lambda^t$. To obtain a matching lower bound, let $f(k) = \cos(2\pi k/(N+1))$. A straightforward computation shows that for any x, $E_k f(X_1) = \lambda f(k)$, and as a result, $E_k f(X_t) = \lambda^t f(k)$. In particular, $d_t \geq \max_{k',k}(f(k) - f(k'))\lambda^t = c''\lambda^t$. This proves that the coupling is efficient.

Chapter 3

Our work

1 Background

The goal of this chapter is to construct an efficient coupling for a discrete version of one-dimensional diffusion with redistribution on an interval, a special case of a model which was studied independently by several groups of authors ([6][3][8][9][10][12][1]). Our coupling gives a probabilistic explanation to the rates of convergence to stationarity for the model and partially answers an open problem posed in [8]. The fact that the coupling does capture the rate of convergence is nontrivial.

Grigorescu and Kang [6] first considered a model they called Brownian Motion on the Figure Eight. The model considered was Brownian motion on an interval $(0, 1)$, which upon hitting the boundary, starts afresh at a point $a \in (0, 1)$. It was shown that the model converges exponentially fast to stationarity with the convergence rate coinciding with the second eigenvalue of the Dirichlet Laplacian $-\frac{1}{2}\frac{d^2}{dx^2}$ on $(0, 1)$. The model was generalized to higher dimension [7], more general diffusion and boundary

behavior in [3][2][12]. However, the one-dimensional model with BM as the underlying diffusion (possibly with constant drift) but more general boundary behavior than BM on the Figure Eight has attracted some attention because, despite its apparent simplicity (and obvious regeneration structure), it exhibits interesting and nontrivial behavior [10][8][9][1]. By more general boundary behavior, we mean that upon hitting the boundary, the process starts afresh in the domain, but with an initial distribution ν_- if exiting from the left and ν_+ if exiting from the right. Of course, this mechanism is repeated indefinitely. It was shown in [10] that if $\nu_- = \nu_+$, then the rate of convergence is the second eigenvalue of the Dirichlet Laplacian, and the first Dirichlet eigenvalue is an (unattainable) infimum of rates of convergence over all choices of ν_- and ν_+. In an unpublished work due to the tragic death of Wenbo Li, it was shown that the third Dirchlet eigenvalue is the maximal rate of convergence. All three results were obtained by Fourier analysis and did not provide any insight on the probabilistic mechanism that governs the rate of convergence. Kolb and Wubker [8] obtained an efficient coupling for the case $\nu_- = \nu_+$, giving a beautiful and intuitive explanation to the rate of convergence, utilizing the fact that $\nu_- = \nu_+$ allows us to guarantee coupling once both copies are redistributed at the same time. This principle does not hold when $\nu_- \neq \nu_+$, a problem left open in [8], and is the main motivation for the present work. Although our main interest is in this latter case, we also provide our version of the coupling in [8] for the discrete setting, as this leads to more questions and completes the picture for the discrete setting.

2 Our Model

We now describe our model. Our process builds from a lazy random walk on the state space $\mathcal{S}_N = \{0, \ldots, N\}$ for some $N > 2$. Let ν_0 be a probability distribution on $\{2, \ldots, N\}$ and ν_N be a probability distribution on $\{0, \ldots, N-2\}$. Slightly abusing notation, we consider ν_0, ν_N also as the probability mass functions with domain $\mathbb{Z}$, through the identification $\nu_x(z) = \nu_x(\{z\})$. Consider the transition function p on the state space given by:

$$p(x,y) = \begin{cases} \frac{1}{2} & x = y \\ \frac{1}{4} & |x-y| = 1 \\ \frac{1}{4}\nu_x(y) & x \in \{0, N\}, |y-x| > 1. \end{cases} \tag{3.1}$$

As is easy to see, the model is never reversible. In what follows, we will always make the following additional assumptions, which will simplify our arguments and allow us to focus more on the main ideas and less on parity-related technicalities (which are still unavoidable, but more manageable):

$$N \in 4\mathbb{N}, \text{ and } \nu_0, \nu_N \text{ are both supported on } \{3, 5, \ldots, N-3\}. \tag{3.2}$$

As we wish to construct a coupling and consequent bounds which are uniform under scaling as $N \to \infty$, these assumptions pose no restriction. Any probability distribution on $(0, 1)$ is a weak limit of scalings of ν_0 and ν_N as above.

3 Our Results

3.1 Assumptions and Preparation

In order to make the main argument simpler and more visible, we will make a reduction to a smaller set of initial distributions. To this end, we need some definitions. Let

$$\rho = 2\lfloor \frac{\min\{x, N-x : \nu_0(x) + \nu_N(x) > 0\}}{2} \rfloor. \tag{3.3}$$

That is, ρ is the largest even number less than or equal to the distance of the union of the support of ν_0 and ν_N to $\{0, N\}$. Observe that, by assumption (3.2), $\rho \geq 2$. Let

$$\tilde{d}_t = \sup_{y-x \in 2\mathbb{N}, y-x \leq \rho} d_t(x, y). \tag{3.4}$$

Then we have the following simple proposition:

Proposition 3.1.

$$\tilde{d}_t \leq d_t \leq \lfloor 1 + N/\rho \rfloor (\tilde{d}_t + \tilde{d}_{t-1}). \tag{3.5}$$

Proof. The first inequality is trivial. We turn to the second.

From the triangle inequality, for any x, y we have

$$d_t(x, y) \leq d_t(x_0, x_1) + \cdots + d_t(x_{n-1}, x_n), \tag{3.6}$$

whenever $x = x_0 < x_1 < \cdots < x_n = y$. Note that $y - x = m\rho + b$ for unique pair (m, b) with $m \in \mathbb{Z}_+$ and $0 \leq b < \rho$. We set the first (possibly empty set of) differences $x_{j+1} - x_j,\ j < m$ each to ρ. If $y - x$ is even, we let $n = m$ or $n = m + 1$ according to whether $b = 0$ or not. In the latter case, we let $x_{m+1} - x_m = b$. When $y - x$ is odd,

we do as follows. If $b = 1$ then set $n = m + 1$ and let $x_n - x_{n-1} = 1$. Otherwise, let $n = m + 2$ and set $x_{m+1} - x_m = b - 1$ and $x_{m+2} - x_m = 1$.

If $y - x$ is even, then we obviously have

$$d_t(x, y) \leq \lfloor 1 + N/\rho \rfloor \tilde{d}_t \tag{3.7}$$

and when $y - x$ is odd, we have

$$d_t(x, y) \leq \lfloor 1 + N/\rho \rfloor \tilde{d}_t + d_t(x_n - 1, x_n). \tag{3.8}$$

It remains to find an upper bound for $d_t(x_n - 1, x_n)$. Choose A such that

$$d_t(x_n - 1, x_n) = P_{x_n - 1}(X_t \in A) - P_{x_n}(X_t \in A). \tag{3.9}$$

We construct a coupling (X, Y) starting from $(x_n - 1, x_n)$ as follows. Let L and R be ν_0 and ν_N distributed random variables, respectively. We will assume that L and R are independent. For the first step, we toss two independent fair coins, independent of L and R. If the first lands H, then X moves and Y stays put. Otherwise, Y moves and X stays put. If the second lands H, then we move the copy we chose one step to the right, meaning redistribution to R if it's Y and Y is at N. If it lands T then we move one step to the left, meaning redistribution to L if it is X and X is at 0. After this first step, both copies continue to evolve independently. Note that in any case, exactly one copy moves. If the copy moved is not redistributed, then $Y_1 - X_1 \in \{0, 2\}$. If X is redistributed from 0 to L, then $x_n = 1$ and $X_1 - Y_1 = L - 1$, so that the distance after one step is even. Similarly, if Y is redistributed form N, then $x_n = N$, so that $X_1 - Y_1 = N - 1 - R$, which is again even.

By the Markov property,

$$d_t(x_n-1,x_n) = E_{(x_n-1,x_n)}\left(\mathbf{1}_A(X_t) - \mathbf{1}_A(Y_t)\right) = E_{(x_n-1,x_n)}E_{(X_1,Y_1)}(\mathbf{1}_A(X_{t-1})-\mathbf{1}_A(Y_{t-1})). \tag{3.10}$$

However, by the argument above and the triangle inequality, it follows from (3.7) that

$$E_{(X_1,Y_1)}(\mathbf{1}_A(X_{t-1}) - \mathbf{1}_A(Y_{t-1})) \leq \lfloor 1 + N/\rho \rfloor \tilde{d}_{t-1}. \tag{3.11}$$

so that $d_t(x_n - 1, x_n) \leq \lfloor 1 + N/\rho \rfloor \tilde{d}_{t-1}$. Plugging this into (3.8) completes the proof.

□

3.2 Deterministic Redistribution

In this section, we will assume, in addition to (3.2), that ν_0 and ν_N are deterministic. Specificially

$$\nu_0 = \delta_{J_0} \text{ and } \nu_N = \delta_{J_N}, \text{ where } J_0, J_N \in \{3, 5, \ldots, N-3\}. \tag{3.12}$$

Let

$$L_0 = \frac{1}{2}\max\{J_0 - 1, N - 1 - J_N, N + J_N - J_0\}. \tag{3.13}$$

Observe that L is a positive integer, and it will serve as the "effective length" that will determine the exponential tail of the coupling time. Roughly speaking, L_0 is the longest interval that one copy of the process needs to exit before the two copies meet, a sort of "bottleneck" for the coupling. Since the coupling is efficient, this actually describes the worst case scenario for convergence to stationarity. The geometric

meaning of L_0 is given as follows: if we think of our state space as consisting of two "loops" (which is incorrect), one from 0 to J_0 (where we identify J_0 with -1), and the other from J_N to N (where we identify J_N with $N+1$), then the first two listed elements of the set on the righthand side represent the lengths of the respective loops. The third, divided by 2, can be viewed as distance between the centers of the loops. Observe that the examples in Chapter 2 serve as "extreme" versions of our model: the random walk on an interval can be thought of as the case where J_0 and J_N are 0 and N, respectively, while the random walk on an n-cycle can be thought of as the case where J_0 and J_N are N and 0, respectively. Note that, given our assumptions on J_0 and J_N, these cases are not covered by our work.

Observe that the largest distance between the centers increases as J_0 decreases and J_N increases and attains a maximum of $N-2$ for $J_0 = 3$, $J_N = N-3$. The minimal distance is 0, attained when $J_0 = N-3$ and $J_N = 3$. We also observe the following additional bounds for L_0:

1. $L_0 = \frac{N}{2}$ whenever $J_0 = J_N$
2. $L_0 \leq N-3$ (attained when $J_0 = 3, J_N = N-3$)
3. $L_0 \geq \frac{2(N-1)}{3}$ (attained when $J_N = \frac{N-1}{3}$).

The main result of this section is the following:

Theorem 3.2. *Suppose that* $0 \leq x < y \leq N$ *and* $y - x \in \{2, 4, \ldots, \rho\}$. *Then there exists a Markovian coupling with* $(X_0, Y_0) = (x, y)$ *such that the coupling time* τ *is dominated by* $\lfloor 6 + N/(\rho+1) \rfloor$ *independent copies of* $T(L_0)$.

In fact, the bound in the statement is weaker than the actual result proved, as the coupling time is dominated by a sum of independent random variables all dominated

by $T(L_0)$. The number of the random variables as well as their distributions depend on x, y. In particular, with the exception of x, y as in Stage 2c in our proof of the coupling, the coupling time is dominated by the independent sum of $T(\max(\frac{J_0-1}{2}, \frac{N+1-J_N}{2})) + T(\frac{N+J_N-J_0}{2}) + T(\frac{J_0-3}{2})$. We will not pursue this further because our main goal is obtaining the exponential rate, and to do so, the statement of Theorem 3.2 would become messy; all possibilities are obtained easily from the proof and Figure 3.1. We comment, however, that for x, y in Stage 2c when $L_0 = \frac{N+J_N-J_0}{2}$, the bound obtained in our construction does contain a sum of at least two independent copies of $T(L_0)$, which implies that the coupling time decays exponentially with a polynomial tail. We do not know whether this is an artifact of our construction or a limitation of Markovian couplings.

Since by Lemma 2.2, $T(L_0)$ has an exponential tail, and a finite sum of independent and identically distributed random variables with an exponential tail also has an exponential tail with the same exponent, it follows from the theorem, Proposition 3.1, and Lemma 2.2-(2) that

Corollary 3.3.

$$\limsup_{t\to\infty} \frac{1}{t} \ln d_t \leq \ln \lambda(L_0). \tag{3.14}$$

To show that the coupling constructed in Theorem 3.2 is efficient, we need a matching lower bound:

Proposition 3.4.

$$d_t \geq \frac{1}{2}(1 - \frac{2\pi}{N})\lambda(L_0)^t. \tag{3.15}$$

3.3 Random Redistribution

Here, we will consider more general redistribution measures under the additional assumption that $\nu_0 = \nu_N$.

As seen in the last section, if $\nu_0 = \nu_N = \delta_{J_0}$, then $\frac{1}{t} \ln d_t \sim \ln \lambda(\frac{N}{2})$ as $t \to \infty$, independently of the choice of J_0. In this section, we show that this remains the same under the present, more general, assumptions. The analogous results for Brownian motion instead of lazy random walk were first obtained by Li and his coauthors [10] through Fourier analysis, and a probabilistic proof using coupling was given in [8]. The coupling we present here is an adaptation of the coupling idea from the latter work; we present it here to distinguish it from the case of the previous section.

Theorem 3.5. *Suppose that* $0 \leq x < y \leq N$ *and* $d = y - x \in \{2, \ldots, \rho\}$. *Then there exists a coupling with* $(X_0, Y_0) = (x, y)$ *such that the coupling time* τ *is dominated by at most* 5 *independent copies of* $T(N/2)$.

The matching lower bound is given by

Proposition 3.6. $d_t \geq P(T(N/2) > t)$.

We highlight the following with regard to Theorem 3.5:

1. As shown in our proof of Theorem 3.5, the coupling is not Markovian, unless ν_0 is a point mass distribution. This raises the question of whether there does exist a Markovian coupling at all, and if so, what is the best bound such a Markovian coupling can give. The same questions are even more interesting for the case $\nu_0 \neq \nu_N$ with none being a point-mass distribution. We leave these for future research.

2. Unlike the coupling of Theorem 3.2, the coupling in Theorem 3.5 ends after at most 5 stages, independently of the parameters.

4 Proofs

4.1 Coupling Regimes

To prove our theorems, we begin by introducing the couplings we will apply. The main idea is to switch between the two coupling regimes described in Chapter 2 (*rigid* and *reflection* coupling) according to the state of the system. It is convenient and simpler to describe the coupling using simple reversible lazy random walk on $\{-1, 0, \ldots, N, N+1\}$, with transition of the random walk from 0 to -1 identified with redistribution to J_0 and transition from N to $N+1$ identified with redistribution to J_N. Switching between the two regimes occurs at hitting times of the joint process.

4.2 Deterministic Redistribution

Proof of Theorem 3.2 .

Without loss of generality we assume that $J_0 \leq N - J_N$. In order to simplify the description of the coupling, we let $\overline{X}_t = \min(X_t, Y_t)$ and $\overline{Y}_t = \max(X_t, Y_t)$.

Suppose that $\overline{X}_0 = x$ and $\overline{Y}_t = y$, and let $D = y - x$. Then D is even and $D < J_0$. Define the "symmetric" points:

$$\ell_0(D) = \frac{J_0 - 1 - D}{2} \text{ and } \ell_N(D) = \frac{N + 1 + J_N - D}{2}. \tag{3.16}$$

The coupling is done in four stages. We begin from the stage that corresponds to

the initial condition of the system. In the description below, the first item describes the initial configuration.

Stage 1.

1. $\overline{X}_0 \in \{\ell_0(D), \ell_N(D)\}$. Apply reflection coupling.
2. Stop when coupling occurs.
3. Time to complete: $T(\frac{J_0-1}{2})$ if $\overline{X}_t = \ell_0(D)$ and $T(\frac{N-J_N-1}{2})$ if $\overline{X}_t = \ell_N(D)$.

When $\overline{X}$ does not begin from either of the symmetric points, we will drive it to one of them. This will be done through rigid coupling. If $\overline{X}$ is between the symmetric points, we apply rigid coupling (Stage 2a). When $\overline{X}_t < \ell_0(D)$ or $\overline{X}_t > \ell_N(D)$, there may be redistribution before $\overline{X}$ reaches either point (Stages 2b and 2c).

Stage 2a.

1. $\overline{X}_0 \in \{\ell_0(D)+1, \ldots, \ell_N(D)-1\}$ and $D < J_0$. Apply rigid coupling.
2. Stop when $\overline{X}$ hits $\{\ell_0(D), \ell_N(D)\}$.
3. Time to complete: $T(\ell_N(D) - \ell_0(D) - 1) = T(\frac{N+J_N-J_0}{2})$.

When this stage ends, we continue to stage 1.

Stage 2b.

1. $\overline{X}_0 < \ell_0(D)$ and $D < J_0$. Apply rigid coupling.
2. Stop when either
   (a) $\overline{X}$ hits $\ell_0(D)$, or

(b) $\overline{X}$ is redistributed from 0.

3. Time to complete: $T(\ell_0(D)) = T(\frac{J_0-1-D}{2})$.

If the first alternative holds, we continue to Stage 1. Otherwise, at the end of the stage, $\overline{Y}$ is at J_0 while $\overline{X}$ is at $D-1$. Thus, the new distance is $D' = J_0 + 1 - D$, which is even, since J_0+1 and D are even, and $D' < J_0$, because D is an even positive integer. Observe then that

$$\ell_0(D') = \frac{J_0 - 1 - D'}{2} = \frac{J_0 - 1 - (J_0 + 1 - D)}{2} = \frac{D}{2} - 1 < D - 1.$$

Therefore, the symmetric point is below the position of $\overline{X}$, and we continue to stage 2a with the new distance D'.

Stage 2c.

1. $\overline{X}_0 > \ell_N(D)$ and $D < J_0$. Apply rigid coupling.

2. Stop when either

   (a) $\overline{X}$ hits $\ell_N(D)$, or

   (b) $\overline{Y}$ is redistributed from N.

3. Time to complete: $T(N - \ell_N(D) - D) = T(\frac{N-J_N-1-D}{2})$.

If the first alternative holds, we continue to Stage 1. Otherwise, at the end of the stage, $\overline{X}$ is at J_N while $\overline{Y}$ is at $N+1-D$. Thus, the new distance is $D' = N+1-D-J_N$. If $D' \geq J_0$, we continue to Stage 3. Otherwise, it is clear that $\overline{X} \leq \ell_N(D')$, so we continue to Stage 2a, with the new distance D'.

Stage 3.

1. $\overline{X}_0 \in \{J_0, J_N\}$, $\overline{X}_0 \leq \ell_N(D)$, and $\overline{Y}_0 = \overline{X}_t + D$ with $D \geq 2$ such that D is even and $D \geq J_0$. Apply rigid coupling.

2. Stop when either

   (a) $\overline{X}$ hits $\ell_N(D)$, or

   (b) $\overline{X}$ is redistributed from 0.

3. Time to complete: $T(\ell_N(D))$. Since $D \geq J_0$, J_0 is odd, and D is even, it follows from (3.16) that the time is dominated by $T(\frac{N+J_N-J_0}{2})$.

If the first alternative holds, we continue to Stage 1. In this case, we adapt Stage 1 slightly, as we now have $D' > J_0$. However, since $\overline{X}_t = \ell_N(D')$ and $\overline{X}_t < \overline{Y}_t$ by definition, $D' < N + 1 - J_N$, so Stage 1 works as before.

Otherwise, by assumption, after redistribution we have that $\overline{X} = J_0$ and $\overline{Y} = D - 1 \geq J_0$, so the new distance is $D' = D - (1 + J_0)$. Let us consider three alternatives:

1. $\overline{X} = J_0 > \ell_N(D')$. This can only occur if 2c started from J_N and $J_0 > J_N$. In this case, we continue to Stage 4.

2. $\overline{X} = J_0 = \ell_N(D')$. Then we continue to Stage 1, as from alternative 2a.

3. $\overline{X} = J_0 < \ell_N(D')$. If $D' < J_0$, we continue to stage 2a. Otherwise, we iterate stage 3. Since in each iteration the distance decreases by $1 + J_0$, the number of iterations does not exceed $\lfloor D/(1 + J_0) \rfloor$, and we eventually continue to Stage 1 or to Stage 2a. Note that, since the distance decreases after each iteration, ℓ_N increases after each iteration.

Stage 4

1. $\overline{X}_0 = J_0 > \ell_N(D)$ and $\overline{Y}_0 = J_0 + D$. Apply reflection coupling.

2. Stop when either

   (a) Coupling occurs, or

   (b) $\overline{Y}$ is redistributed from N.

3. Time to complete: $T(N - J_0 - D/2)$ but since $J_0 > \ell_N(D)$, it follows from (3.16) that $N - J_0 - D/2 < \frac{N-1-J_N}{2}$, so the time is dominated by $T(\frac{N-1-J_N}{2})$.

If the second option occurs, $\overline{X} = J_N$ and $\overline{Y} = J_0 - (N + 1 - (J_0 + D))$. Since $J_N < \overline{Y} < J_0$, $D' = \overline{Y} - \overline{X} < J_0$. Thus, depending on the relation of $\overline{X} = J_N$ to $\ell_0(D')$, we continue to either Stage 1, 2a, or 2b. Furthermore, since, initially, $\overline{X} > \ell_N(D)$, we have that $\overline{X} - J_N > N + 1 - \overline{Y}$ and in particular, $\overline{X} > N + 1 - \overline{Y}$. Thus, we disregard the possibility of $\overline{X}$ redistributing from 0 as $\overline{Y}$ would always be redistributed first.

Let us review the coupling. Figure 3.1 displays all possible implementations of the coupling. Stages 1, 2a, 2b, and 2c are the initial steps, in the sense that the coupling must begin from one of these stages. In the coupling, each of these steps is repeated at most once. Stage 2c is unique in the sense that it takes care of the case that the redistribution may lead to a distance bigger or equal to J_0. Stage 3 is invoked when this happens.

Stage 3 may be repeated a number of times, where the number of iterations is bounded above by $\lfloor 1 + N/(J_0 + 1) \rfloor$. From Stage 3, the coupling proceeds to either stage 1, 2a, or 4. From Stage 4, the coupling can either end (meeting) or continue to one of the Stages 1, 2a, or 2b.

From the point of view of the duration of the coupling, the meeting time is bounded

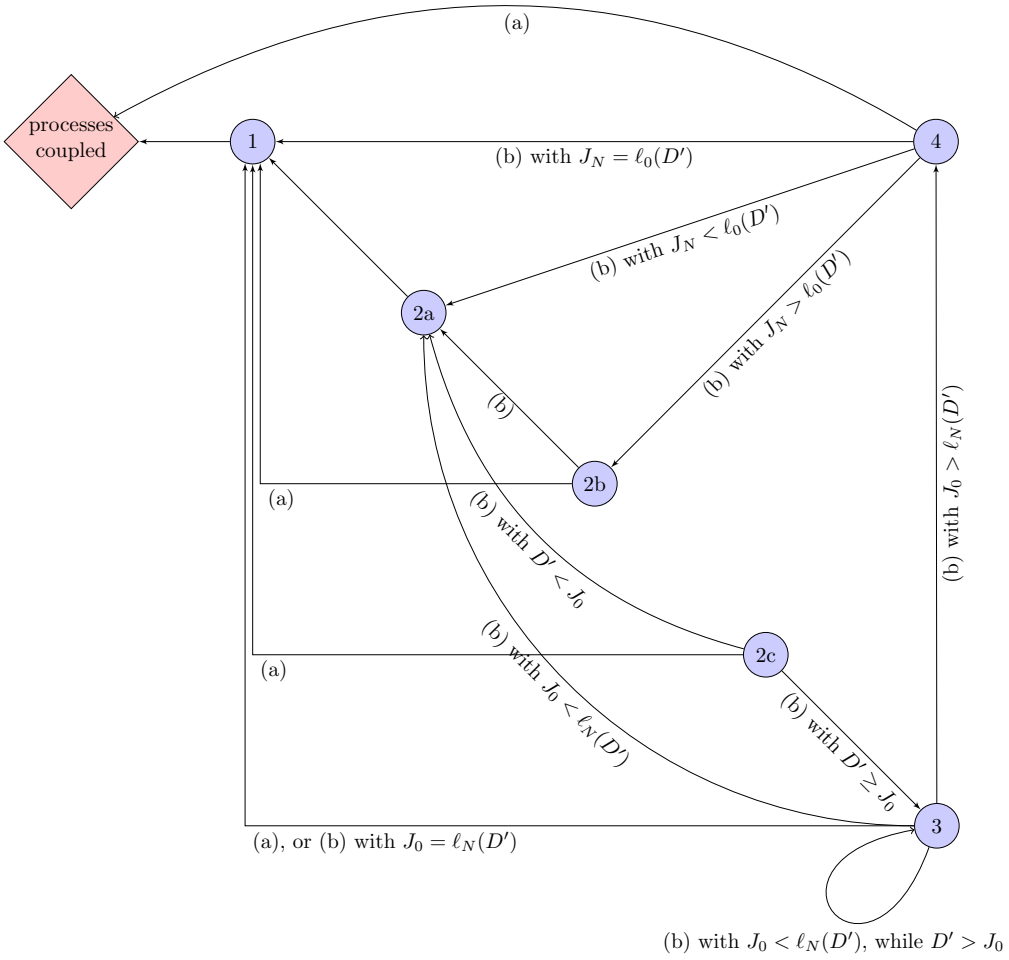

FIGURE 3.1: Summary of the coupling from Section 4.1

above by the independent sum of the times for all Stages, with possible repetitions for Stage 3. Listing the times for each of the stages in order of appearance, we have $T((J_0 - 1)/2)$ or $T((N - J_N - 1)/2)$ (Stage 1), $T(\frac{N+J_N-J_0}{2})$, $T(\frac{J_0-1-D}{2})$ and $T(\frac{N-J_N-1-D}{2})$ (Stage 2), $T(\frac{N+J_N-J_0}{2})$ (Stage 3) and $T(\frac{N-J_N-1}{2})$ (Stage 4). The maximal length among all intervals mentioned above is therefore $\frac{1}{2}\max\{J_0 - 1, N - J_N - 1, N + J_N - J_0\}$, which is L_0. Finally, Figure 3.1 shows that the maximal number of steps (omitting Step 3) is 5, and the result now follows. □

Proof of Proposition 3.4. We will find an eigenfunction f for p which is not constant and has a real eigenvalue λ. Without loss of generality, we may assume that $\sup|f| \leq 1$. For any x, y, $d_t(x,y) \geq \frac{1}{2}\left(E_x f(X_t) - E_y f(X_t)\right)$. Since we must have $\sum_x f(x)\pi(x) = 0$, where π is the stationary distribution of X, it follows that f attains both strictly positive and strictly negative values. In particular, we can choose x and y such that $f(x) > 0 > f(y)$, and it immediately follows that

$$d_t(x,y) \geq \frac{1}{2}(f(x) - f(y))\lambda^t. \tag{3.17}$$

In order to find f, we will choose $f(x) = \sin(\rho x + \omega)$, and we will find choices for the parameters ρ and ω that match the upper bounds from Theorem 3.2. Suppose now that $x \in \{1, \ldots, N-1\}$. Then

$$pf(x) = \frac{1}{4}\left(\sin(\rho(x+1)+\omega) + \sin(\rho(x-1)+\omega)\right) + \frac{1}{2}\sin(\rho x + \omega).$$

Thus, $pf(x) = \frac{1}{2}(\cos\rho + 1)f(x)$. Now, if we extend f to $\{-1, \ldots, N+1\}$ and additionally impose the constraints

$$\begin{cases} \sin(-\rho + \omega) = \sin(\rho J_0 + \omega), \text{ and} \\ \sin(\rho(N+1) + \omega) = \sin(\rho J_N + \omega), \end{cases} \tag{3.18}$$

then it immediately follows that f (restricted to the state space) is indeed an eigenfunction for p with eigenvalue $\lambda = \frac{1}{2}(\cos\rho + 1)$. In order to proceed, we need to find choices for ρ and ω that will satisfy the constraints.

The first constraint is met if $-\rho + \omega + 2\pi = \rho J_0 + \omega$. That is, $\rho(J_0 + 1) = 2\pi$, or $\rho = \frac{2\pi}{J_0+1}$. With this choice of ρ, the first constraint is met for all choices of ω,

which allows to freely chose ω to meet the second constraint. The actual value of ω is irrelevant for the eigenvalue calculation.

As is easy to see, we can repeat the argument by considering the second constraint first. This will give us $\rho = \frac{2\pi}{N+1-J_N}$.

Another way to satisfy the first constraint is to have the arguments of the sin function in the equation symmetric with respect to a maximum or a minimum of the sin function, that is $\pi/2 + \pi k$ for some integer k. In taking $k = 0$, the first constraint is satisfied when

$$\frac{\pi}{2} - (-\rho + \omega) = \rho J_0 + \omega - \frac{\pi}{2},$$

and the second constraint will be met if

$$\frac{3\pi}{2} - (\rho J_N + \omega) = \rho(N+1) + \omega - \frac{3\pi}{2}.$$

Subtract the first equation from the second to obtain $\pi - \rho(J_N + 1) = -\pi + \rho(N + 1 - J_0)$, that is

$$\rho = \frac{2\pi}{N + 1 + J_N - (J_0 - 1)}.$$

Using the first equation, $\omega = \frac{\pi}{2} - \frac{\rho}{2}(J_0 - 1)$, and the second equation is thus satisfied as well.

In light of the above, we see that (3.17) is satisfied when λ is chosen to be λ_1 in the statement of the theorem, and f is the corresponding eigenfunction of the form $\sin(\frac{\pi}{L_0+1}x + \omega_1)$. Observe that $\rho_1 \geq \frac{\pi}{N}$. In particular, the set $\{\omega_1, \rho_1 + \omega, 2\rho_1 + \omega, \ldots, N\rho_1 + \omega\}$ contains at least one element whose distance from $\pi/2 + \pi k_1$ for some $k_1 \in \mathbb{Z}$, is at most $\frac{\pi}{N}$, as well as an element whose distance from πk_2 for some $k_2 \in \mathbb{Z}$ is at most $\frac{\pi}{2}$. Call the first x and the second y. Without loss of generality, we may

assume $f(\pi/2 + k\pi) = 1$ (otherwise change to $-f$). It follows that $f(x) \geq 1 - \frac{\pi}{N}$. Similarly, $f(y) \leq \frac{\pi}{N}$. Therefore, $f(x) - f(y) \geq 1 - \frac{2\pi}{N}$, and the result follows. □

4.3 Random Redistribution

Proof of Theorem 3.5. We first prove the theorem for the case that $x, y \leq N/2$. By symmetry, this also covers the case where $x, y \geq N/2$. After we construct the coupling for this stage, we extend it to the remaining case where $x \leq N/2$ and $y > N/2$.

Assume then that $0 \leq x < y \leq N/2$ and $d = y - x \leq \rho$. Let K, K' be two independent random variables distributed according to ν_0.

Stage 1a.

1. $X_0 = x, Y_0 = y,\ 0 \leq x < y \leq N/2,\ d \leq \rho$. Apply Rigid coupling.
2. Stop when either
   (a) Y hits $N/2$, then continue to 1b, or
   (b) X is redistributed from 0 to K. Continue to Stage 2a.
3. Time is bounded above by $T(N/2)$.

Stage 1b.

1. $Y_0 = N/2$ and $X_0 = N/2 - d$. Apply reflection coupling.
2. Stop when either
   (a) Copies meet, or
   (b) $Y_t - X_t = \rho$. Continue to 1c.

3. Time is bounded above by $T(\rho/2 - 1)$.

Note that the second alternative will occur before a redistribution, because when X hits 0, the distance will be $N/2 - d + d + N/2 - d = N - d$, and since $d \leq \rho$, this quantity is greater or equal to ρ.

Stage 1c.

1. $X_0 = N/2 - d - (\rho - d)/2,\ Y_0 = N/2 + (\rho - d)/2$. Apply Rigid coupling.

2. Stop when either

   (a) $Y_t = (N + \rho)/2$, then continue to Stage 3, or

   (b) X is redistributed from the origin to K, then continue to Stage 3.

3. Time is bounded above by $T(\frac{N-\rho}{2} + 1)$.

If the second alternative holds and $K \leq N/2$, the processes meet. Otherwise, after this stage ends, the copies are at a and $N - a$ for some a.

Stage 2a. $X_0 = K$ and $Y_0 = d - 1$ (Y never jumped). Apply reflection coupling.

1. Stop when either

   (a) Copies meet, or

   (b) When distance is $K + 1$ and continue to 2b.

2. Time bounded above by $T(d/2)$.

Stage 2b.

1. $X_0 = K + d/2, Y_0 = d/2 - 1$ (Y never jumped). Apply Rigid coupling.

2. Stop when either:

   (a) Y is redistributed from 0 to K and copies meet, or

   (b) $Y_t = N/2 - (K+1)/2$; then continue to stage 3.

3. Time is bounded above by $T(N/2 - (K+1)/2)$.

Stage 3.

1. $Y_0 = N - X_0$. Apply reflection coupling.

2. Stop when either

   (a) The copies meet, or

   (b) The copies are redistributed from 0 and N simultaneously to K'.

3. Time is dominated by $T(N/2)$.

As is easy to see, the coupling ends after no more than four Stages, the longest chain being 1a→1b→1c→3, and the times for the stages are all dominated by identical and independently distributed copies of $T(N/2)$. Furthermore, this coupling is not Markovian, since, in Stage 2b, if Y is redistributed from 0, the transition is from $(0, K+1) \to (K, K)$, whereas, if the coupling were Markovian and $\nu_0(y) < 1$, then for any $y < N$, there exists $x \neq y$ and y' such that the transition $(0, y+1) \to (x, y')$ occurs with positive probability.

We have therefore completed the proof for the case where x, y are both $\leq N/2$. Suppose now that $x \leq N/2$ and $y > N/2$. Let $\tilde{y} = N - 2$. Then $x, \tilde{y} \leq N/2$, and furthermore, since $0 < y - x \leq \rho$, $y - x = |y - N/2| + |N/2 - x| = |\tilde{y} - N/2| + |N/2 - x|$.

But $|\tilde{y} - x| \leq \max\{N/2 - \tilde{y}, N/2 - x\} \leq y - x \leq \rho$. Furthermore, since $\tilde{y} - y$ is even, $\tilde{y} - x$ is even too. Thus, we can construct a coupling $(X, \tilde{Y})$ starting from $(x, \tilde{y})$ whose coupling time will be dominated by at most four independent copies of $T(N/2)$. However, letting $Y = N - \tilde{Y}$, at the coupling time τ for $(X, \tilde{Y})$, we have that $Y_\tau = N - X_\tau$. Therefore, applying stage 3 again (with an independent copy of K') guarantees that X and Y will be coupled after no more than 5 independent copies of $T(N/2)$.

□

Proof of Proposition 3.6. Let $f(x) = 1$ if $x \leq N/2$ and $f(x) = 0$ otherwise. Now:

$$d_t \geq E_{N/4} f(X_t) - E_{3N/4} f(Y_t) = E_{(N/4,3N/4)}(f(X_t) - f(Y_t)),$$

for every coupling (X, Y). If we choose reflection coupling until the copies meet and then move them together, since $3N/4 - N/4$ is even, it follows that the meeting time τ occurs exactly when X exists in the set $\{0, \ldots, N/2 - 1\}$. In particular,

$$d_t \geq E_{(N/4,3N/4)} \left[(f(X_t) - f(Y_t)) \, ; \tau > t \right], \tag{3.19}$$

and the righthand side is equal to $P(T(N/2) > t)$. □