Mixing of Lazy Random Walk in Monotone Censored Hypercube

B. Tech Project Report Submitted
in Partial Fulfillment of the Requirements
for the Degree of

Bachelor of Technology

by

Prakhar Pandey & Pratham Pekamwar
(200101081 & 200101087)

under the guidance of

Dr. Benny George K

to the

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI
GUWAHATI - 781039, ASSAM

CERTIFICATE

*This is to certify that the work contained in this thesis entitled "**Mixing of Lazy Random Walk in Monotone Censored Hypercube**" is a bonafide work of **Prakhar Pandey & Pratham Pekamwar** (**Roll No. 200101081 & 200101087**), carried out in the Department of Computer Science and Engineering, Indian Institute of Technology Guwahati under my supervision and that it has not been submitted elsewhere for a degree.*

Supervisor: **Dr. Benny George K**

Associate Professor,

April, 2024 Dept. of Computer Science and Engineering,

Guwahati. Indian Institute of Technology Guwahati, Assam.

Acknowledgements

I am deeply grateful to Dr. Benny George K for his unwavering support and invaluable guidance throughout this thesis. His expertise, encouragement and perspective have been instrumental in shaping this work. Further, we would also like to express gratitude to Professor Yuval Peres for his book, literature, and online lectures which served as a cornerstone for our research.

Abstract

The report is a study of mixing times in lazy random walk on monotone subsets of hypercubes, providing a literature survey and simulation study. The aim is to establish a foundation for future research towards proving a better upper bound on the mixing time of lazy random walks within these structures. The report also draws the following conclusions:

1. *The stationary distribution is a uniform distribution over the different states, irrespective of the laziness factor or structure of the monotone subset.*

2. *Mixing time is* ***not*** *directly proportional to laziness.*

3. *Dependence of mixing time on the structure of the montone subset is complex in nature and thus a major hindrance in finding the upper bound.*

Contents

Chapter 1

Introduction

1.1 Problem Statement and Motivation

The problem at hand is to improve the existing upper bound on the mixing time of lazy simple random walks within monotone subsets of hypercubes. Specifically, we focus on a monotone subset V of the hypercube $\{0,1\}^n$, where the volume of V constitutes at least half of the hypercube's volume. Our objective is to refine the upper bound for the mixing time of a lazy simple random walk on V, building upon [DM14] which established an upper bound of order n^3. The ultimate goal is to reach the conjectured potential bound of $O(n \log n)$.

The motivation for studying this problem stems from its relevance to various fields, including:

Percolation theory: Understanding the connectivity and flow properties of random structures.

Monte Carlo simulations: Efficiently sampling from complex probability distributions.

Randomized algorithms: Designing efficient algorithms with probabilistic components.

Statistical mechanics: Analyzing the behavior of systems with many interacting particles.

Monotone subsets of hypercubes offer several advantages for studying mixing times:

Ergodicity: They guarantee the existence of a unique stationary distribution, simplifying the analysis.

Faster mixing times: Their inherent structure is conjectured to lead to faster convergence compared to general graphs.

1.2 Conceptual Framework

This section provides definitions and explanations of key concepts essential for understanding the problem and its analysis:

Markov Chain: A Markov chain is a model for a sequence of random events, where the probability of each event is determined solely by the state achieved in the previous event. This dependency is encapsulated in the *Markov property*, formally expressed as:

Let $X_1, X_2, \ldots, X_n, \ldots$ be a stochastic process with state space S and $n \in \mathbb{N}$.

X_n represents the state of the Markov chain at time step n.

This stochastic process satisfies the Markov property:

$P(X_{n+1} = x | X_n = x_n, X_{n-1} = x_{n-1}, \ldots, X_1 = x_1) = P(X_{n+1} = x | X_n = x_n)$

That is, the conditional distribution of X_{n+1} depends only on the previous state X_n.

Stationary Distribution: A stationary distribution π of a Markov chain is a probability distribution that remains unchanged in the Markov process, satisfying $\pi P = \pi$. Intuitively, it represents a state of equilibrium where, upon repeated application of the transition matrix P, the probability distribution over the states does not change.

Monotone Subset of Hypercube: An n-dimensional hypercube is represented by a graph with 2^n vertices, where each vertex corresponds to an n-digit binary number. The edges are between vertices that differ in only one coordinate. A subset V is monotone if for any binary numbers x and y, with $x \in V$ and $y \geq x$ coordinate-wise, y is also in V. Here the monotone subgraph contains the vertices as in the subset V and the edges from the original hypercube graph for which the corresponding end point vertices are in subset V.

Lazy Random Walk: A lazy random walk on a subset V of the hypercube involves remaining at the current vertex with a probability of 1/2 or moving to any of it's adjacent vertex randomly. On a complete hypercube, this corresponds to selecting a coordinate of the vertex's binary representation at random and flipping the corresponding bit with a probability of 1/2. It is worth noting here that a random walk on hypercube may result in periodicity, therefore refraining us from making a claim for an unique stationary distribution, however, this lazy modification results in an ergodic markov chain in the problem statement.

Mixing Time: Mixing time is a measure of how quickly a Markov chain converges to its stationary distribution. It is defined as the minimum time $t_{\text{mix}}(\varepsilon)$ for which the total variation distance (a metric of distance between two probability distributions) from the stationary distribution is less than a specified threshold ε:

$$t_{\text{mix}}(\varepsilon) = \min\{t : \|P^t(x, \cdot) - \pi\|_{TV} < \varepsilon \text{ for all } x \in X\}$$

Note, here TV subscript indicates total variation distance and $P^t(x, \cdot)$ refers to the transition matrix row corresponding to the initial state as x. π denotes the stationary distribution.

Chapter 2

Review of Prior Works

We begin by reviewing the important concepts from the book [LPW17] and online lectures by Simons institute [Sim16]. Especially those, that lay the foundation for understanding and advancing on the problem statement.

2.1 Basics of Markov Chains

2.1.1 Irreducibility

A Markov chain is irreducible if it is possible to go from any state to any other state. Mathematically, a chain is irreducible if for all states i, j, there exists a $t > 0$ such that $P^t(i, j) > 0$.

2.1.2 Aperiodicity

A state in a Markov chain is aperiodic if there is no fixed periodic pattern in its transitions. Mathematically, state i is aperiodic if $\gcd\{t : P^t(i, i) > 0\} = 1$.

2.1.3 Reversibility

A Markov chain is reversible if it satisfies the detailed balance equation $\pi(x)P(x, y) = \pi(y)P(y, x)$ for all states x, y.

Proposition: If P is aperiodic and irreducible, then there exists an integer r_0 such that $P^r(x, y) > 0$ for all $x, y \in X$ and $r \geq r_0$.

2.2 Intuitive Explanation of Ergodicity Leading to a Unique Stationary State

Convergence Theorem: For an irreducible, aperiodic chain with stationary distribution π, there exist $\alpha \in (0, 1)$ and $C > 0$ such that $\max_{x \in X} \|P^t(x, \cdot) - \pi\|_{TV} \leq C\alpha^t$.

2.2.1 Non-Zero Probability for All States After Time t

Due to irreducibility and aperiodicity, after some time t (which depends on the starting state), every state in the Markov chain will have a non-zero probability of being visited. This time t is the maximum time required to ensure a non-zero probability for all states across all starting distributions.

2.2.2 Decomposition into Stationary Distribution and Residual

Imagine decomposing the probability distribution at time t into two parts: one part that aligns with the stationary distribution π and a residual part. The stationary part remains constant under the action of the transition matrix, while the residual part diminishes over time.

2.2.3 Exponential Decay of the Residual

The key is to observe how the residual distribution behaves over iterations of this process of division into multiple of stationary distribution and residual distribution. The sum of probabilities in this residual distribution is less than 1 after the first step and decreases exponentially with each iterative step. Specifically, it decreases by a factor of $\alpha < 1$ at each step.

2.2.4 Stability of the Stationary Part

The component of the distribution that is a multiple of the stationary distribution remains unaffected by further applications of the transition matrix. Since a stationary distribution, by definition, does not change upon applying the transition matrix, this part of the distribution remains the same.

2.2.5 Convergence to Stationary Distribution

As the residual distribution decays to zero, the overall distribution of the Markov chain converges to the stationary distribution. The unique stationary distribution is reached as the impact of the initial state fades away, and the probability distribution aligns entirely with π.

2.3 Equivalence Classes

Lemma: Let X be the state space of a Markov chain (X_t) with transition matrix P. If $\sim$ is an equivalence relation on X with classes $\{[x]\}$ and $P(x,[y]) = P(x',[y])$ whenever $x \sim x'$, then $([X_t])$ is a Markov chain with state space $\{[x]\}$ and transition matrix $P'([x],[y]) = P(x,[y])$

2.4 Understanding Total Variation Distance

Total variation distance is a metric used to quantify the disparity between two different probability distributions. It is computed by identifying the maximum difference in the probability of any event occurring according to these two distributions.

Total Variation Distance: $\|\mu - \nu\|_{TV} = \frac{1}{2}\sum_{x\in X} |\mu(x) - \nu(x)|$.

Derivation Intuition: The total variation distance is a measure of distance between two probability distributions. It is defined as the maximum difference in the probability of any event happening as per two different probability distributions. The formula for total

variation distance can be derived by finding the event with the maximum difference in probability. This event can be obtained by including either all states which have μ - ν corresponding to those states negative, or all states that have this difference positive. The magnitude of the probability of these events is going to be the sum of the magnitude of the differences of the different states involved. This is because for the events that we are considering, the differences (μ - ν) of the states involved are all of the same sign. Further, the sum of μ - ν for all states is zero. Thus, the two events we are considering must be having equal probability. This is because one event contains all the positive differences, while the other all the negative differences, and the sum of the two is zero. Hence, half of the sum over all states of the magnitude of the difference μ - ν gives us the formula for total variation distance.

2.5 Total Variation Distance and Coupling

Proposition: $\|\mu - \nu\|_{TV} = \inf\{P\{X \neq Y\} : (X, Y)$ is a coupling of μ and $\nu\}$.

Intuition For $P\{X \neq Y\}$ to be minimum, we need $P\{X = Y\}$ to be maximum. This is possible if for all $P(X, X) = min(\mu(X), \nu(Y))$. So, if we consider any such coupling following this particular division

$P\{X \neq Y\} = 1 - \sum \min(\mu(X), \nu(Y))$

$P\{X \neq Y\} = \frac{1}{2} \sum_{x \in X} |\mu(x) - \nu(x)|$.

Chapter 3

Methodology and Implementation

3.1 Methodology

3.1.1 State Space

States represent specific position of the walker within the subset. The size of the state space $\leq 2^d$ where 2^d is the size of the entire d-dimensioned hypercube. This is because each element in the subset can be represented by a d-bit binary string.

3.1.2 Monotone Subset

We can define monotone subset in terms of their *base points*. A *base point* is a point in the monotone subset which is not less than any other points from the monotone subset that are comparable to it. Thus, a monotone subset can be represented entirely by its set of base points.

3.1.3 Transition Probabilities

The transition probabilities define the likelihood of moving from one state to another in a single step. The transition matrix, denoted by P, is a $2^d \times 2^d$ matrix where each entry $P(i, j)$ represents the probability of transitioning from state i to state j. The lazy random

walk can be thought in terms of choosing a bit at random from the d bits, and flipping it with probability $1 - laziness$ and switching if the new state exists. The below pseudo-code outlines the same:

Algorithm 1 Lazy Random Walk on Monotone Subset

Require: Current state $x \in \{0,1\}^n$, laziness parameter $l \in [0,1]$, monotone subset $V \subseteq \{0,1\}^n$

Ensure: New state x'

1: Choose $i \in \{1, 2, \ldots, n\}$ uniformly at random
2: $x' \leftarrow x$ {Initialize new state as current state}
3: Flip x'_i with probability $1 - l$
4: **if** $x' \in V$ **then**
5: $\quad x \leftarrow x'$ {Update state if new state is in the monotone subset}
6: **end if**

3.2 Simulation Design

3.2.1 Types of Monotone Subsets

The simulations considered monotone subsets generated by different configurations of base points. The different configurations are generated using three values (different for different value of dimensions), appending them one by one to the set of base points under use.

The function 'generateMonotoneSubset' takes a set of base points and the dimension as input and returns the set of states that form the monotone subset.

Algorithm 2 generateMonotoneSubset(base_points, dimension)

```
monotone_subset ← []
for i in range(2^dimension) do
    for base_point in base_points do
        if is_greater_than_or_equal(i, base_point, dimension) then
            monotone_subset.append(i)
            break
        end if
    end for
end for
return monotone_subset
```

3.2.2 Transition Matrix

The transition matrix generation is implemented in the function 'generateLazyTransitionMatrix'. This function takes the dimension, the set of states forming the monotone subset, and the laziness parameter as input and returns the transition probability matrix.

Algorithm 3 generateLazyTransitionMatrix(dimension, monotone_subset, laziness)

```
matrix ← zeros(2^dimension, 2^dimension)
for i in range(2^dimension) do
    if i ∉ monotone_subset then
        matrix[i][i] ← 1.0
    else
        matrix[i][i] ← laziness
        for j in range(2^dimension) do
            if is_neighbor(i, j) then
                if  j ∈ monotone_subset then
                    matrix[i][j] ← (1 - laziness) / dimension
                else
                    matrix[i][i] = matrix[i][i] + (1 - laziness) / dimension
                end if
            end if
        end for
    end if
end for
return  matrix
```

3.2.3 Range of Parameters

The simulations explored mixing of the markov chains for different values of parameters to understand the impact of those parameters on the mixing time. The parameters are

- **Laziness:** The laziness parameter, ranging from 0.002 to 0.998, controls the fixed probability of staying at the current state. Varying laziness allows us to observe how it affects the convergence speed.
- **Dimension:** The dimension of the monotone subset was varied from 3 to 13. In-

creasing the dimension leads to a larger state space and potentially longer mixing times.

- **Base Points:** Varying the base point set allows the study of mixing times and structure of monotone subsets.

he primary measure used to estimate mixing time is the total variation distance between the current probability distribution and the stationary distribution. The simulations calculated this distance at each step of the random walk and considered the walk "mixed" when the distance remains below a certain threshold (0.0001 in this case) for 10 consecutive steps. 13

3.2.4 Lazy Random Walk

The lazy random walk is performed using repeated matrix multiplication of the initial distribution with the transition matrix. The final distribution thus obtained would then be used to compare with stationary distribution (uniform distribution) for analyzing the mixing time. he primary measure used to estimate mixing time is the total variation

Algorithm 4 Calculate Mixing Time

```
distribution ← initialDistribution
steps ← 0
streak ← 0
while streak < 10 do
    steps ← steps + 1
    distribution ← MatrixVectorMultiply{transitionMatrix, distribution}
    if CalculateTotalVariationDistance{distribution, stationaryDistribution} < threshold then
        streak ← streak + 1
    else
        streak ← 0
    end if
end while
return steps − 10
```

distance between the current probability distribution and the stationary distribution. The

simulations cal- culated this distance at each step of the random walk and considered the walk "mixed" when the distance remains below a certain threshold (0.0001 in this case) for 10 consecutive steps. 13

3.2.5 Measure used to estimate mixing time

The primary measure used to estimate mixing time is the total variation distance between the current probability distribution and the stationary distribution. The simulations calculated this distance at each step of the random walk and considered the walk "mixed" when the distance remains below a certain threshold (0.0001 in this case) for 10 consecutive steps.

The function 'calculateTotalVariationDistance' is used to measure the distance between two probability distributions.

Algorithm 5 calculateTotalVariationDistance(distribution1, distribution2)

```
distance ← 0.0
for i in range(size(distribution1)) do
    distance += abs(distribution1[i] - distribution2[i])
end for
return distance / 2
```

Chapter 4

Results and Analysis

4.1 Mixing Time Behavior

This section presents and analyzes the results of the simulations, focusing on the relationship between mixing time and the laziness parameter and the configurations of montone subsets generated from different base points.

4.1.1 Impact of Laziness

The simulations explored the effect of varying the laziness parameter on the mixing time of the lazy random walk.

Lemma 1: The stationary distribution of an ergodic Markov chain remains unchanged by the introduction of laziness.

Proof: Let P be the transition matrix of the original Markov chain, and $P_L = LI + (1 - L)P$ the transition matrix of the lazy chain with laziness parameter L, where I is the identity matrix. If π is the stationary distribution of P, then $\pi P = \pi$. We show that $\pi P_L = \pi$:

$$\pi P_L = \pi(LI + (1 - L)P) = L\pi I + (1 - L)\pi P = L\pi + (1 - L)\pi = \pi.$$

Thus, π is also the stationary distribution of the lazy chain P_L.

Lemma 2: If two lazy versions of a non-ergodic Markov chain are ergodic, they share the same stationary distribution.

Proof: Consider $P_{L_1} = L_1 I + (1 - L_1)P$ and $P_{L_2} = L_2 I + (1 - L_2)P$, with $L_1, L_2 > 0$ and $L_1 \neq L_2$. Assume both P_{L_1} and P_{L_2} are ergodic with unique stationary distributions π_{L_1} and π_{L_2}. Then:

$$\pi_{L_1} P_{L_1} = \pi_{L_1}, \quad \pi_{L_2} P_{L_2} = \pi_{L_2}.$$

Rewrite the stationary equations for π_{L_1} and π_{L_2}:

$$\pi_{L_1}(L_1 I + (1 - L_1)P) = \pi_{L_1}, \quad \pi_{L_2}(L_2 I + (1 - L_2)P) = \pi_{L_2}.$$

Expanding both equations, we get:

$$L_1 \pi_{L_1} + (1 - L_1)\pi_{L_1} P = \pi_{L_1}, \quad L_2 \pi_{L_2} + (1 - L_2)\pi_{L_2} P = \pi_{L_2}.$$

By the uniqueness of the stationary distribution for each ergodic chain and the fact that any convex combination (like that formed by the laziness parameters) of π_{L_1} and π_{L_2} that satisfies the stationary condition must equal the unique stationary distribution, it follows that π_{L_1} and π_{L_2} must be equal if they both satisfy the same linear combination equation with P.

Mixing Time: While it might seem that mixing time would be directly proportional to laziness as laziness would just slow down the mixing because of staying in the same state, the simulation results revealed a more complex relationship.

Figure 4.1 illustrates the observed relationship between mixing time and laziness for different scenarios. It can be seen that the relationship is not simply linear.

Intuition: Consider a simple case of a one-dimensional hypercube (two directly con-

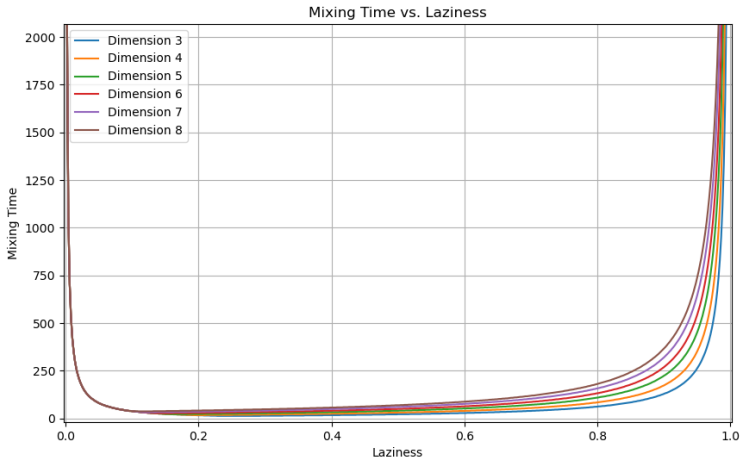

Fig. 4.1 Mixing time vs. Laziness for complete hypercubes for different dimensions

nected states). The stationary distribution is uniform in this case, with each state having a probability of 0.5. The difference in probabilities between the two states during the mixing process can be described as a decreasing geometric progression with a reduction factor inversely proportional to $(1 - laziness) \cdot (laziness)$.

This observation can be extended to higher dimensions by grouping states based on the parity (even or odd) of the number of 1s in their binary representation. Within each group, the behavior is similar to the one-dimensional case, leading to the complex relationship between laziness and overall mixing time. Fundamentally, there is a tug of war between mixing two far away points (between groups of points) and mixing inside the groups. Laziness speeds up the second process but slows down the first process.

Conclusion: As discussed, introducing laziness does not affect the stationary distribution of the random walk. However, it does influence the time to reach the stationary state (mixing time).

4.1.2 Impact of Base Points

The number and location of base points defining the monotone subset were also investigated for their impact on mixing time.

Initial Hypothesis: It was initially thought that adding more number of base points would lead to an increase in mixing time due to the more constrained structure of the monotone subset.

Observations: Contrary to the initial hypothesis, the simulations did not reveal a clear and consistent relationship between the number of base points and the mixing time. The mixing time appears to be more sensitive to the specific location and configuration of the base points rather than their quantity. Overall, the base points define the boundaries of the monotone subset and restrict the possible transitions of the walker. In certain configurations, the base points can slow down the mixing, where the walk gets trapped in specific regions of the state space for longer periods, thus reducing the mixing process. In other configurations, the presence of some base points introduces other points which increase the connectedness of the state space, thus increasing the mixing process. Further analysis is needed to understand the exact nature of this dependence.

4.2 Uniform Stationary Distribution

Empirical evidence from the simulations suggests that the lazy random walk on the considered monotone subsets converges to a uniform stationary distribution. Theoretical proof for the same is as follows:

Lemma: A lazy random walk on a monotone subset V of the hypercube $\{0,1\}^n$ has a uniform stationary distribution.

Proof:

1. The laziness parameter and the connected nature of monotone subsets in the hypercube ensure the ergodicity of the Markov chain, guaranteeing a unique stationary distribution.

2. Each vertex x in V remains with a probability $l + k\frac{1-l}{n}$, where k is the number of adjacent vertices not in V. The transition probability to any neighboring vertex y in

V is $\frac{1-l}{n}$.

3. The detailed balance condition (stronger than the earlier condition, also implying that the Markov chain is reversible), $\pi(x)P(x,y) = \pi(y)P(y,x)$, holds because all transition probabilities between connected vertices x and y in V are equal, ensuring $\pi(x) = \pi(y)$ for all x, y in V.

4. Given $\pi(x) = \pi(y)$ for all x, y in V and $\sum_{x \in V} \pi(x) = 1$, it follows that $\pi(x) = \frac{1}{|V|}$ for each x in V.

   Therefore, the uniform distribution across V is the unique stationary distribution due to the chain's ergodicity, consistent transition probabilities, and the satisfaction of the detailed balance.

Chapter 5

Future Works and Suggestions

This chapter outlines potential future research directions specifically aimed at understanding and finding upper bounds on the mixing time of lazy random walks on monotone subsets of a hypercube, particularly those with a volume greater than half of the hypercube.

5.1 Strong Stationary time Analysis

Strong stationary times provide a precise moment at which the chain can be considered to have reached equilibrium, independent of its initial state. A *strong stationary time* T for a Markov chain $(X_t)_{t\geq 0}$ with stationary distribution π is a random time for which:

- $\Pr(T = t \mid X_0 = x) = \Pr(T = t)$ for all $t \geq 0$ and all states x, indicating that T is independent of the initial state X_0.
- $\Pr(X_T = y \mid T < \infty) = \pi(y)$ for all states y, meaning that the distribution of the Markov chain at time T is exactly the stationary distribution π, regardless of the initial state.

Approach: For all aperiodic irreducible markov chains, there always is a stationary distribution time. If the markov chain is reversible, there always is a strong stationary time whose order is of mixing time. A Markov chain is said to be *reversible* if there exists a

probability distribution π on its state space such that for any two states i and j, the detailed balance condition is satisfied: $\pi(i)P(i,j) = \pi(j)P(j,i)$, where $P(i,j)$ is the probability of transitioning from state i to state j, and $\pi(i)$ is the stationary probability of state i. This condition implies that the chain, when observed in a steady state, appears the same whether it is moving forward or backward in time. We can see our chain is reversible, hence we have a strong stationary time of the order of mixing time. We can use strong stationary times to bound mixing times.

The relevant theory is as follows:

Since, strong stationary time T for a Markov chain provides a clear, probabilistic point at which the chain's distribution matches the stationary distribution, regardless of the initial state. This means that at time T, the total variation distance between the distribution of the Markov chain and the stationary distribution is zero.

If T is a strong stationary time, then for any $\epsilon > 0$, $t_{\text{mix}}(\epsilon)$ is bounded above by $E[T]$, the expected value of T, provided that the probability of T exceeding $E[T]$ by a significant amount is small. Essentially, $t_{\text{mix}}(\epsilon)$ can often be approximated or bounded by $E[T]$, especially when the tail distribution of T decays rapidly.

When T has a light tail (e.g., exponential decay), it implies that the probability of the chain staying away from its stationary distribution for long times after T is low. This can be used to give concrete bounds on $t_{\text{mix}}(\epsilon)$ for small ϵ, typically leading to bounds like:

$$t_{\text{mix}}(\epsilon) \leq E[T] + C \log\left(\frac{1}{\epsilon}\right),$$

where C is a constant depending on the details of the chain and the decay rate of the tail of T.

5.2 Analyzing Structure in Monotone Subsets

It may be explored how different types of structures of monotone subset depending on the set of base points affect the mixing time and other relevant properties. A key area of investigation will be whether the time complexity for a subset defined by a set of base points $S1$ differs significantly from that defined by another set $S2$, and how the time complexity of the mixing time of the structure with base point set as union of these sets, $S1 \cup S2$, is related to the mixing time complexity of structures with base point set $S1$ and $S2$.

5.3 Exploring Coupling for Different Start States

The concept of coupling different start states to derive a measure of τ_{couple} can also be explored, as it is pivotal in understanding the rate at which a Markov chain 'forgets' its initial state. This exploration is fundamental to determining the mixing time of the chain and will involve developing novel coupling strategies to efficiently estimate τ_{couple} for lazy random walks on monotone subsets.

5.4 Ehrenfest Urn Model and Spectral Gap Analysis

Our initial investigations suggest that defining an Ehrenfest Urn Model for monotone subsets is challenging. However, if properties related to base points are established, this model could be useful. In subsets defined by a single base point, the inherent symmetry simplifies the classification to equivalence states, enabling a more feasible approach for projecting Markov chains. While the spectral gap has been used as an upper bound on mixing time, the exponential size of the transition matrix in relation to the hypercube's dimension has made this approach impractical. The projection of states and formulation of equivalence classes might render the spectral gap analysis a viable method for such complex systems.

References

[DM14] Jian Ding and Elchanan Mossel. Mixing under monotone censoring. *Electronic Communications in Probability*, 19:1–6, 2014.

[LPW17] David A. Levin, Yuval Peres, and Elizabeth L. Wilmer. *Markov Chains and Mixing Times*. American Mathematical Society, Providence, Rhode Island, second edition, 2017. With contributions by Elizabeth L. Wilmer.

[Sim16] Simons Institute. Markov chain mixing times and applications i, 2016. Available online: https://youtu.be/svi1j799i0A. Accessed: April 2024.