Analyses of Deterministic Processes Analogous to Finite Markov Chains

白髪，丈晴

https://doi.org/10.15017/1807055

出版情報：九州大学, 2016, 博士（工学）, 課程博士
バージョン：

Analyses of Deterministic Processes Analogous to Finite Markov Chains

Takeharu Shiraga

December 2016

Abstract

A *Markov chain*, or a *random walk*, is a simple and important stochastic process, which often appears as mathematical models and analyses in diverse fields including theoretical computer science. For instance, a random walk is intensively studied as a useful approach for network exploration because of its simplicity, locality and robustness in changing networks, and Markov chain Monte Carlo (MCMC) is established as a general scheme for randomized approximate counting algorithms. The performance of these randomized algorithms based on measures of the Markov chain such as *mixing time* and the *cover time*. Many analytic techniques of these values have been developed, and form the basis of the theory of randomized computation.

Recently, a *deterministic random walk*, which is a *deterministic* process analogous to a Markov chain, has been studied as an alternative of a random walk in some contexts such as network exploration and simulations of physical phenomena. In particular, the mixing property and the cover time of a deterministic random walk corresponding to a *simple* random walk (called *rotor-router model*) on specific graphs such as integer lattice and hypercube has been analyzed. However, the theory of the deterministic random walk is still developing compared with the rich theory of general Markov chains in a long history. For example, few studies are concerned with deterministic random walks corresponding to *general* transition probabilities beyond simple random walks. Little is known about the gap between a randomized computation and a deterministic computation, which is one of the important topic on theoretical computer science.

Motivated by understanding the gap between a randomized computation and a deterministic computation, this study constructs a general framework of analyses of deterministic random walks. In particular, we analyze the mixing property and the cover time of deterministic random walks corresponding to general finite Markov chains, and this thesis mainly makes progress on the following three topics.

First, we analyze the vertex-wise discrepancy of token-configurations between a *reversible* Markov chain and its corresponding deterministic random walk. We present an upper bound in terms of the mixing time and the stationary distribution of the Markov chain, and the *local discrepancy* of the deterministic random walk. This result implies polynomial upper bounds for rapidly mixing chains which are developed in the context of MCMC.

Second, we give an analysis of the *total variation discrepancy*, which is an important measure as designing randomized algorithms based on MCMC. This is the first results on an upper bound of the total variation discrepancy between a general finite Markov chain and its corresponding deterministic random walk. The upper bound depends only on the edge size of the graph, the mixing time of the Markov chain, and the local discrepancy of the deterministic random walk, but is independent of the stationary distribution of the Markov chain as well as the total number of tokens. On the other hand, we give an instance for a lower bound that the total variation discrepancy gets as large as the order of magnitude of the number of states, which implies that an extra argument is required to derandomize MCMC in general.

Third, we analyze the discrepancy of visit frequencies between a reversible Markov chain and its corresponding deterministic random walk. We give an upper bound of the discrepancy described in terms of the mixing time and the stationary distribution of the Markov chain, and the local discrepancy of the deterministic random walk. Based on the analyses, we give an upper bound of the cover time of deterministic random walks corresponding to general transition probabilities. This result improves the existing result on the speed up ratio of the rotor-router model as increasing the number of tokens in general graphs.

Acknowledgment

First of all, I am deeply grateful to Associate Professor Shuji Kijima. He is my supervisor, and always gave me perfect supports and advices. I would like to thank Professor Masafumi Yamashita, who was my supervisor when I was a master course student. He gave me many constructive comments. Discussions with Professor Jun'ich Takeuchi, Professor Eiji Takimoto, and Professor Tomoyuki Shirai have been insightful for this work. Assistant Professor Yukiko Yamauchi also gave me valuable comments.

The results in this thesis were published in Proceedings of COCOON 2014 [78], ANALCO 2016 [79], and AofA 2016 [76]. I would like to thank to all anonymous referees.

Last of all, I would like to express the deepest appreciation to my family, my mentor, and all the people who support me.

Table of Contents

Chapter 1

Introduction

A Markov chain is an important stochastic process often appearing in diverse fields, and many analytic techniques have been developed in the literature. Recently, a deterministic process, called rotor-router model, is intensively studied in some contexts as an analogy of a simple random walk. However, the theory for deterministic processes corresponding to general transition probabilities has not been established yet. Motivated by understanding the gap between a randomized computation and a deterministic computation from the viewpoint of theoretical computer science, this thesis studies the *deterministic random walk*, which is a deterministic analogy of a general Markov chain. Then, we construct a general framework of analyses of the mixing property and the cover time of deterministic random walks.

1.1 Markov chains

Markov chains appear in diverse fields as mathematical models and techniques of analyses. For example, a Markov chain is often used as a description of physical phenomena. Brownian motion is a typical one, which is the irregular motion of particles in a solvent like a fluid. This phenomenon is mathematically modeled by Wiener process, which is related to a limit of the random walk [34]. Other diffusion processes occurring in nature are also featured to the diffusion property of Markov chains, e.g., heat diffusion, internal diffusion limited aggregation (IDLA) [61], etc. Markov chains are also used in the simulation of spin systems [80, 72], analyses of the phase transition phenomena of the percolation [44], etc. There are many other appearance of Markov chains in various contexts, e.g., queueing systems in operations research [59], reliability of systems, game theory, the card shuffling in mathematics, a typical model of memoryless information sources in information theory, Pagerank [14], etc.

Markov chain is also an important mathematical tool for algorithm design and analysis in theoretical computer science, e.g., the analysis of randomized algorithm for 2-SAT, (randomized) distributed algorithms such as load balancing [56, 75], voting [18], population protocol [24], etc. In particular, this thesis refers to two successful topics of Markov chains in theoretical computer science. One is Markov chain Monte Carlo (MCMC), which is a general scheme of randomized approximate counting, and the other is the cover time of a random walk, which is concerned in the contexts of network exploration, log space randomized algorithm for the connectivity, etc.

1.1.1 Counting and random sampling

Counting is a fundamental topic in Combinatorics. Valiant [84] proposed the class #P, which is a computational class of the counting version of NP. Several counting problems are known to be #P-complete, such as 0-1 knapsack solutions [55, 85], linear extensions of a partially ordered set (poset) [13], matchings in a graph [85], etc. #P is an important class of polynomial-time complexity theory, which is suggested by the celebrated Toda's theorem [83], $\mathrm{PH} \subseteq \mathrm{P}^{\#\mathrm{P}}$, for instance.

Counting is highly related to sampling, which is a fundamental topic in Probability Theory. Jerrum et al. [53] showed the equivalence in the sense of the polynomial time computation between *almost* uniform generation and randomized approximate counting for self-reducible problems. Markov chain Monte Carlo (MCMC) is a powerful technique of random sampling from a complicated space. The idea of MCMC is simple; design an *ergodic* Markov chain with a desired limit distribution, and sample from the limit distribution simulating the chain (see Section 2.2). A number of fully polynomial-time randomized approximation schemes (FPRAS) based on MCMC have been developed for #P-hard problems, such as the volume of a convex body [35, 66, 23], integral of a log-concave function [66], partition function of the Ising model [50], and counting bipartite matchings [51]. When designing an FPRAS based on the technique, it is important that the *total variation distance* of the approximate distribution from the target distribution is sufficiently small, and hence analyses of the *mixing times* of Markov chains are central issues in a series of works on MCMC for FPRAS to guarantee a small total variation distance. Many techniques are established to estimate the mixing time of a Markov chain, e.g., coupling, conductance, eigenvalue analysis, etc [65], and above FPRAS are constructed by proving polynomial-size bounds of mixing times of Markov chains on target combinatorial objects.

1.1.2 Network exploration

A *random walk* is often used for network exploration, because of its simplicity, locality and robustness in changing networks. The expected cover time (this paper simply says *cover time*) of a random walk on a finite graph is the expected time until every vertex has been visited by a token. The cover time of a random walk has been well investigated since it is a key measure of the network exploration by a random walk. The network exploration by a random walk is related to some basic topics of theoretical computer science such as connectivity, universal traversal sequences, etc (see [5] for detail).

Aleliunas et al. [5] showed that the cover time of a *simple random walk*, in which a neighboring vertex is chosen uniformly at random, is upper bounded by $2m(n-1)$ for any connected graph, where m denotes the number of edges and n denotes the number of vertices. Feige [37, 38] showed that the cover time is lower bounded by $\bigl(1-o(1)\bigr)n\log n$ and upper bounded by $\bigl(1+o(1)\bigr)(4/27)n^3$ for any graph.

Motivated by a faster cover time, the cover time by more than one token has also been investigated. Broder et al. [15] gave an upper bound of the cover time of k independent parallel simple random walks (k-simple random walks) when tokens start from stationary distribution. For an arbitrary initial configuration of tokens, Alon et al. [6] showed that the cover time of k-simple random walks is upper bounded by $\bigl((\mathrm{e}+o(1))/k\bigr)t_{\mathrm{hit}}\log n$ for any graph if $k \le \log n$, where e is Napier's constant and t_{hit} denotes the (maximum) *hitting time*. Elsasser and Sauerwald [36] gave a better upper bound for large k of $\mathrm{O}\bigl(t_{\mathrm{mix}} + (t_{\mathrm{hit}}\log n)/k\bigr)$ for any graph if $k \le n$, where t_{mix} is the *mixing time*.

Ikeda et al. [48] took another approach for speeding up, using *general* transition probabilities (beyond simple random walks). They devised *β-random walk*, consisting of irrational transition probabilities in general, and showed that the cover time is $\mathrm{O}(n^2\log n)$. Nonaka et al. [70] showed that the cover time of a *Metropolis-walk*, which is based on the *Metropolis-Hastings algorithm*, is $\mathrm{O}(n^2\log n)$ for any graph. Abdullah et. al. [1] proposed the *minimum degree weighting scheme* and they showed that the cover time of which is $\mathrm{O}(n^2\log n)$. This upper bound is recently refined by David and Feige [25]: the minimum degree weighting scheme has $\mathrm{O}(n^2)$ cover time.

Little is known about the cover time by multiple tokens with general transition probabilities. Elsasser and Sauerwald [36] gave a general lower bound of $\Omega\big((n \log n)/k\big)$ for any transition probabilities and for any $n^{\varepsilon} \le k \le n$, where $0 < \varepsilon < 1$ is a constant.

1.2 Rotor-router model—a deterministic process

Recently, the *rotor-router model*, also known as the *Propp machine* [71, 22, 57], is studied as a deterministic analogy of a simple random walk on a graph. For example, this model is used for the model of internal diffusion limited aggregation (IDLA) instead of random walk, and diffusion properties are studied in [62, 63, 64, 49]. The rotor-router model is also concerned with the research on information spreading [9, 27, 29, 30, 31, 32, 33, 47].

This section reviews existing results on the discrepancy of token-configurations between a rotor-router model and a simple random walk, considering the ability of a rotor-router model to imitate a stationary distribution of random walks, and results on the cover time of the rotor-router model for a deterministic network exploration algorithm.

1.2.1 Token-configuration of rotor-router models

The idea of the rotor-router model is as follows: tokens distributed over vertices are deterministically served to neighboring vertices in the round-robin fashion, instead of at random (see Section 2.3.2 for the detail of the rotor-router model). Then, it is a natural question whether tokens are entirely distributed over vertices by the rotor-router model, similar to random walks.

Cooper and Spencer [22] investigated the rotor-router model on $\mathbb{Z}^n$, and gave an analysis on the discrepancy of a single vertex: they showed a bound that $|\chi_v^{(t)} - \mu_v^{(t)}| \le c_d$, where $\chi_v^{(t)}$ (resp. $\mu_v^{(t)}$) denotes the number (resp. the expected number) of tokens on vertex $v \in \mathbb{Z}^d$ in a rotor-router model (resp. in the corresponding random walk) at time t on the condition that initial configuration of tokens on "even" vertices and for any configuration of rotor-routers, and c_d is a constant depending only on the dimension d, but independent of the total number of tokens in the system. Later, it is shown that $c_1 \simeq 2.29$ [21] and c_2 is about 7.29 or 7.83 depending on the routers [28]. On the other hand, Cooper et al. [20] gave an example of a rotor-router on the infinite k-regular tree, such that its vertex-wise discrepancy gets $\Omega(\sqrt{kt})$ for an arbitrarily fixed t. Friedrich et al. [40] showed an example of unbounded lower bound of discrepancy on a lattice without the "even" condition on the initial token configuration.

For some specific finite graphs, such as hypercubes and tori, some upper bounds on the discrepancy of polylogarithmic to the size of the transition diagram are known. For d-dimensional hypercube, Kijima et al. [57] gave a bound $\mathrm{O}(d^3)$. Akbari and Berenbrink [2] gave a bound $\mathrm{O}(n^{1.5})$ for lazy random walks on d-dimensional hypercubes, using results by Friedrich et al. [39]. Akbari and Berenbrink [2] also gave a bound $\mathrm{O}(1)$ for constant dimensional tori.

Similar, or essentially the same concepts have been independently developed in load-balancing. Rabani et al. [73] are concerned with a deterministic algorithm similar to the rotor-router model corresponding to a simple random walk on a d-regular graph, and showed for the model that the vertex-wise discrepancy is $\mathrm{O}\left(d\log(n)/(1-\lambda^*)\right)$ where λ^* is the second largest eigenvalue of corresponding transition matrix. Friedrich et al. [39] proposed the "BED" algorithm for load balancing, which uses some extra information in the previous time, and they gave $\mathrm{O}(d^{1.5})$ bound for hypercube and $\mathrm{O}(1)$ bound for constant-dimensional tori. Akbari and Berenbrink [2] discussed the relation between the BED algorithm and the rotor-router model, and gave the same bounds for a rotor-router model. Berenbrink et al. [12] investigated about cumulatively fair balancers algorithms, which includes the rotor-router model, and gave an upper bound $\mathrm{O}(d\min(\sqrt{\log(n)/(1-\lambda^*)}, \sqrt{n}))$ for a lazy version of simple random walks on d-regular graphs. We remark that those results are for *simple* random walks on *regular* graphs.

1.2.2 Cover time of rotor-router model

From the view point of a *deterministic* graph exploration, the rotor-router model is intensively studied recently. Yanovski et al. [88] studied the asymptotic behavior of the rotor-router model, and proved that any rotor-router model always stabilizes in a traversal of an Eulerian cycle after $2mD$ steps at most, where D denotes the diameter of the graph. Bampas et al. [10] gave examples of which the stabilization time gets $\Omega(mD)$. Their results imply that the cover time of a single token version of a rotor-router model is $\Theta(mD)$ in general. Another approach to examine the cover time of the rotor-router model is connecting qualities of a random walk and the *visit frequency* $X_v^{(T)}$ of the rotor-router model, where $X_v^{(T)}$ denotes the total number of times that tokens visited vertex v by time T. Holroyd and Propp [45] showed that $|\pi_v - X_v^{(T)}/T| \leq K\pi_v/T$, where K is an constant independent of T, and π is the stationary distribution of the corresponding random walk. This theorem means that $X_v^{(T)}/T$ converges to π_v as T increasing. Using this fact, Friedrich and Sauerwald [41] gave upper bounds of the cover time for many classes of graphs.

To speed up the cover time, the rotor-router model with $k > 1$ tokens is studied by Dereniowski et al. [26]. They gave an upper bound $\mathrm{O}(mD/\log k)$ for any graph when $k = \mathrm{O}\big(\mathrm{poly}(n)\big)$ or $2^{\mathrm{O}(D)}$, and also gave an example of $\Omega(mD/k)$ as a lower bound. Kosowski and Pajak [60] gave a modified upper bound of the cover time for many graph classes by connecting $X_v^{(T)}$ and the corresponding simple random walk. They showed that the upper bound is $\mathrm{O}\big(t_{\mathrm{mix}} + (\Delta/\delta)(mt_{\mathrm{mix}}/k)\big)$ for general graphs, where Δ, δ are respectively the maximum, minimum degrees. Recently, Chalopin et al. [17] gave upper and lower bounds of the stabilization time for the rotor-router model with many tokens. Cooper et al. [19] used this result and discussed the coalescing of tokens of the rotor-router model.

1.3 Current issues

Compared with the rich theory of Markov chain in a long history, the study on the rotor-router model is still developing and few techniques of general analysis are known. First of all, the rotor-router model imitates only a simple random walk, and the targets of existing analyses are limited. Furthermore, existing works concerning the discrepancy analysis only deal with the vertex-wise discrepancy, but not the total variation discrepancy, which is an important measure in MCMC. Concerning the analysis of the cover time, very few is known about deterministic processes corresponding to general transition probabilities beyond simple random walks. This section discusses these current issues in detail.

1.3.1 Vertex-wise discrepancy of general transition matrices

While there are examples of Markov chains containing irrational transition probabilities, such as Gibbs samplers for the Ising model (cf. [80, 72]), reversible Markov chains for queueing networks (cf. [58]), etc, few studies are concerned with deterministic processes corresponding to *general* transition probabilities beyond simple random walks. Motivated by general rational transition matrices, Kijima et al. [57] investigated a rotor-router model on finite multidigraphs, and gave an upper bound $\mathrm{O}(n|\mathcal{A}|)$ of the vertex-wise discrepancy when P is rational, ergodic and reversible, where $n = |V|$ and $\mathcal{A}$ denotes the set of multiple edges. For an arbitrary rational transition matrix P, Kajino et al. [54] gave an upper bound using the second largest eigenvalue λ^* of P and some other parameters of P. Rabani et al. [73] mentioned that possibility of generalization of their deterministic algorithm for a simple random walk to a reversible transition matrix, but there are no clear analyses of token-configuration discrepancies concerned with general transition probabilities.

Another point is that the existing analyses highly depend on the structures of the specific graphs, e.g., hypercubes, tori [2]. These analyses seems difficult to be extended to general structures appearing in MCMC algorithms. Kijima et al. [57] gave rise to a question if there is a deterministic random walk for #P-hard problems, such as 0-1 knapsack solutions, bipartite matchings, etc., such that $|\chi_v^{(t)} - \mu_v^{(t)}|$ is bounded by a polynomial in the input size.

1.3.2 Total variation discrepancy

As we mentioned, the total variation distance of the Markov chain from the stationary distribution plays a key role for designing an FPRAS based on MCMC. While there are several works on deterministic random walks concerning the vertex-wise discrepancy $\|\chi^{(t)} - \mu^{(t)}\|_\infty$ such as [73, 57, 54, 77, 12], little is known about the total variation discrepancy $\|\chi^{(t)} - \mu^{(t)}\|_1$.

This theme is concerned with designing *deterministic* approximation algorithms for #P-hard problems. Comparing with the progress on randomized approximation algorithms based on MCMC, not many results are known about deterministic approximation algorithms for #P-hard problems. A remarkable progress is the correlation decay technique, independently devised by Weitz [87] and Bandyopadhyay and Gamarnik [11], and there are several recent developments on the technique. For counting 0-1 knapsack solutions, Gopalan et al. [42], and Stefankovic et al. [81] gave deterministic approximation algorithms (see also [43]). Ando and Kijima [7] gave an FPTAS based on approximate convolutions for computing the volume of a 0-1 knapsack polytope. A direct derandomization of MCMC algorithms is not known yet, but it holds a potential for a general scheme of designing deterministic approximation algorithms for #P-hard problems. Deterministic processes may be used as a substitute for Markov chains, for this purpose.

1.3.3 Cover time of general deterministic processes

For the cover time of a random walk, it is known that general transition probabilities beyond simple random walks give a faster cover time, such as the β-random walk using irrational transition probabilities or the minimum degree weighting scheme archiving $\mathrm{O}(n^2)$ cover time for any graphs. For the cover time of a deterministic process, Holroyd and Propp [45] provides the *stack walk* based on the low-discrepancy sequence [82, 8], and showed a connection between the visit frequency and hitting probabilities. However, nothing is known about the cover time of deterministic random walks for general transition probabilities, as far as we know.

Even for the rotor-router model, there are some remaining problems on the speed up by multiple tokens on general graphs. The speed-up parameter $S(k)$ is defined by $T^{(1)}_{\mathrm{cov}}/T^{(k)}_{\mathrm{cov}}$, where $T^{(i)}_{\mathrm{cov}}$ is the cover time of i-tokens version of the rotor-router model. Kosowski and Pajak [60] showed that $S(k) = \Theta(k)$ for the rotor-router model on a *regular* graph with $t_{\mathrm{mix}} = \mathrm{O}(D)$. Since this analysis depends on the regularity of a graph, it is unclear that $S(k) = \Theta(k)$ or not for *irregular* graphs.

1.4 Contributions

This thesis investigates a deterministic random walk for general transition probabilities possibly containing irrational numbers, and presents analyses of the mixing property and the cover time of the deterministic random walk in general. In a deterministic random walk, each vertex v deterministically serves tokens on v to a neighboring vertex u with a ratio about $P_{v,u}$, where $P_{v,u}$ denotes the transition probability from v to u of a corresponding random walk i.e., $z^{(t)}_{v,u}$, which is the number of tokens moving from $v \in V$ to $u \in V$ at time t, is almost $\chi^{(t)}_v P_{v,u}$ (see Chapter 3 for detail). For this model, this study mainly make progress on the following three topics: the vertex-wise discrepancy for general transition matrices, the total variation discrepancy, and the cover time. Those results are derived from our technical key lemma (Lemma 3.1), which implies the discrepancy between a random walk and a deterministic random walk is expressed by a geometric series of the transition matrix of the random walk where the coefficients are the *local discrepancy* of the deterministic random walk.

1.4.1 Analyses of the vertex-wise discrepancy

First, we estimate vertex-wise discrepancies between a Markov chain and its corresponding deterministic random walk. We show $\|\chi^{(t)} - \mu^{(t)}\|_\infty = \mathrm{O}\big(\phi(\pi_{\max}/\pi_{\min})\Delta t_{\mathrm{mix}}\big)$ (Theorem 4.1) for any ergodic and reversible P and for any deterministic random walk, where $\pi_{\max}$ and $\pi_{\min}$ are respectively the maximum, and the minimum values of the stationary distribution π of P, t_{mix} is the *mixing time* of P, and Δ is the maximum degree of the transition diagram. *Local discrepancy* $\phi \stackrel{\mathrm{def}}{=} \max_{v,u,t} |Z_{v,u}^{(t)} - \chi_v^{(t)} P_{v,u}|$, depending on the routing model, plays a key role in the analysis. We investigate the *SRT-router model*, in which ϕ is bounded by a constant (see Section 3.2.4).

Next, we show an improved upper bound $\|\chi^{(t)} - \mu^{(t)}\|_\infty = \mathrm{O}\big(\Phi(\pi_{\max}/\pi_{\min})\Delta\sqrt{t_{\mathrm{mix}}}\big)$ (Theorem 4.2) for ergodic, reversible and *lazy* P, where $\Phi \stackrel{\mathrm{def}}{=} \max_{v,u,T} |\sum_{t=0}^{T-1}(z_{v,u}^{(t)} - \chi_v^{(t)} P_{v,u})|$ is the *cumulative local discrepancy*, which is also constant for its corresponding SRT-router model.

These upper bounds imply that rapidly mixing chain has small vertex-wise discrepancy. Section 4.3 shows some examples of rapidly mixing chains and polynomial upper bounds for them. In Section 4.4, we show examples of lower bounds of the vertex-wise discrepancy, which our upper bounds are tight for some specific examples.

1.4.2 Analyses of the total variation discrepancy

Next, motivated by a derandomization of MCMC, we give analyses about total variation discrepancy. We give two upper bounds: $\|\chi^{(t)} - \mu^{(t)}\|_1 = \mathrm{O}(\phi m t_{\mathrm{mix}})$ for any ergodic P (Theorem 5.1), and $\|\chi^{(t)} - \mu^{(t)}\|_1 = \mathrm{O}(\Phi m \sqrt{t_{\mathrm{mix}}})$ for any ergodic and *lazy* P (Theorem 5.1), where m is the number of edges of the transition diagram of P. Notice that these upper bounds do not require reversibility of P, and free from $\pi_{\max}/\pi_{\min}$, meaning that it is independent of weights of transition probabilities, instead depending on m.

For a lower bound, we give an example of the oblivious model such that $\|\chi^{(t)} - \mu^{(t)}\|_1 = \Omega(n t_{\mathrm{mix}})$ (Proposition 5.14), which implies that the mixing time t_{mix} is non negligible in the L_1 discrepancy for the oblivious model, in general. For a rotor-router model, we show an example satisfying $\|\chi^{(t)} - \mu^{(t)}\|_1 = \Omega(m)$ (Proposition 5.13). We also show $\|\chi^{(t)} - \mu^{(t)}\|_1 = \Omega(n)$ (Proposition 5.12) for an appropriate number of tokens in general. This general lower bound implies that an extra argument is required to derandomized MCMC in general.

1.4.3 Cover time of general deterministic processes

This thesis is also concerned with the cover time of the deterministic random walk according to general transition probabilities with k tokens, while previous results studied the rotor-router model (corresponding to simple transition probabilities). We give an upper bound of the cover time for any deterministic random walk imitating any ergodic and reversible transition matrix possibly containing irrational numbers (Theorem 6.6). Precisely, the upper bound is $\mathrm{O}\big(t_{\mathrm{mix}} + m' t_{\mathrm{mix}}/k\big)$ for any number of tokens $k \geq 1$, where $m' = \max_{u \in V}(\delta(u)/\pi_u)$. This is the first result on an upper bound of the cover time for deterministic random walks imitating general transition probabilities, as far as we know. Theorem 6.6 implies that the upper bound of the cover time of the rotor-router model is $\mathrm{O}\big(t_{\mathrm{mix}} + m t_{\mathrm{mix}}/k\big)$ for any graph (Corollary 6.9). For $k = 1$, this bound matches to the existing bound $\mathrm{O}(mD)$ by [88] when $t_{\mathrm{mix}} = \mathrm{O}(D)$. This bound is better than $\mathrm{O}(mD/\log k)$ by [26] when t_{mix} is small or k is large. Our bound also improves the bound $\mathrm{O}\big(t_{\mathrm{mix}} + (\Delta/\delta)(m t_{\mathrm{mix}}/k)\big)$ by [60] in Δ/δ factor for irregular graphs, and guarantees speed-up parameter $S(k) = \Theta(k)$ for any graphs with $t_{\mathrm{mix}} = \mathrm{O}(D)$.

In our proof, we investigate the connection between the visit frequency $X_v^{(T)}$ of the deterministic random walk and the corresponding multiple random walks with general transition probabilities by extending the arguments of token-configuration analyses. This approach is an extension of [45, 41, 60]. In precise, we show that $|\pi_v - (X_v^{(T)}/kT)| < K\pi_v/T$ holds for any reversible and ergodic transition matrix, where π_v is the stationary distribution of the corresponding transition matrix and K is constant independent of T. This upper bound extends the result of [45] to $k > 1$ tokens and general transition probabilities.

1.5 Organization

This thesis is organized as follows. Chapter 2 introduces the notations, basic facts of Markov chains and the rotor-route model. Chapter 3 describes the model of deterministic random walks, and presents a foundation of discrepancy analyses. Chapter 4 shows upper and lower bounds of the vertex-wise discrepancy. Chapter 5 shows upper and lower bounds of the total variation discrepancy. Chapter 6 deals with the visit frequency and the cover time of deterministic random walks. In Chapter 7, we conclude this thesis and discuss future works.

Chapter 2

Preliminaries

This chapter introduces the notations and terminology of the thesis, and explains some known facts on Markov chains or deterministic processes.

2.1 Notations

Let $\mathbb{R}$ (resp. $\mathbb{R}_{\geq 0}, \mathbb{R}_{>0}$) denote the set of real numbers (resp. nonnegative, positive reals), $\mathbb{Q}$ ($\mathbb{Q}_{\geq 0}, \mathbb{Q}_{>0}$) denote the set of rational numbers (nonnegative, positive rationals), and $\mathbb{Z}$ ($\mathbb{Z}_{\geq 0}, \mathbb{Z}_{>0}$) denote the set of integers (nonnegative, positive integers). Let $V = \{1, 2, \ldots, n\}$ be a finite set and $\xi \in \mathbb{R}^n$ be a vector on V. Then, ξ_v (or $(\xi)_v$) denotes the v-th element of ξ for any $v \in V$, and ξ_A (or $(\xi)_A$) denotes $\sum_{v \in A} \xi_v$ for any subset $A \subseteq V$. Let $P \in \mathbb{R}^{n \times n}$ be a $n \times n$ matrix on V. Then, $(P)_{u,v} = P_{u,v}$ denotes $(u, v) \in V \times V$ entry of P. Let P^t be the t-th power of P, then $P^t_{u,v}$ denotes the (u, v)-entry of P^t, i.e., $(P^t)_{u,v}$. Let $P_{u,\cdot}$ denotes the u-th row vector of P, i.e., $P_{u,\cdot} = e_u P \in \mathbb{R}^n$ and $(P_{u,\cdot})_v = P_{u,v}$, where e_u is the u-th unit vector. For any $\xi \in \mathbb{R}^n$, we define

$$\|\xi\|_\infty = \max_{u \in V} |\xi_u| \quad (\mathbf{L_\infty} \textbf{ norm of } \xi) \tag{2.1}$$

$$\|\xi\|_1 = \sum_{u \in V} |\xi_u| \quad (\mathbf{L_1} \textbf{ norm of } \xi) \quad . \tag{2.2}$$

2.2 Markov chain

Suppose we wish to sample from $V = \{1, 2, \ldots, n\}$ with a probability proportional to a given positive vector $f = (f_1, \ldots, f_n) \in \mathbb{R}^n_{\geq 0}$; for example, we will be in Section 4.3.1

concerned with *uniform* sampling of 0-1 knapsack solutions, where V denotes the set of 0-1 knapsack solutions and $f_v = 1$ for each $v \in V$. The idea of Markov chain Monte Carlo (MCMC) is to sample from a limit distribution of a Markov chain which is equal to the target distribution $f/\|f\|_1$.

Let $P \in \mathbb{R}_{\geq 0}^{n \times n}$ be a transition matrix on V, where $P_{u,v}$ denotes the transition probability from u to v ($u, v \in V$), i.e., $\sum_{v \in V} P_{u,v} = 1$ for any $u \in V$. A transition matrix P is *irreducible* if for any u and v in V there exists $t \geq 0$ such that $P^t_{u,v} > 0$. A transition matrix P is *aperiodic* if $\mathrm{GCD}\{t \in \mathbb{Z}_{>0} \mid P^t_{x,x} > 0\} = 1$ holds for any $x \in V$. An irreducible and aperiodic transition matrix is called *ergodic*. It is well-known that an ergodic P has a *unique* stationary distribution $\pi \in \mathbb{R}^n_{\geq 0}$, which is a probability distribution satisfying

$$\pi P = \pi. \tag{2.3}$$

Theorem 2.1 ([65](Theorem 4.9))**.** *If P is ergodic, then,*

$$\lim_{t \to \infty} \xi P^t = \pi$$

holds for any probability distribution ξ.

An ergodic Markov chain with a transition matrix $P \in \mathbb{R}_{\geq 0}^{n \times n}$ is *reversible* if the *detailed balance equation*

$$f_u P_{u,v} = f_v P_{v,u} \tag{2.4}$$

holds for any $u, v \in V$.

Proposition 2.2 ([65](Proposition 1.19))**.** *If a probability distribution $\xi \in \mathbb{R}^n$ satisfies the detailed balance equation* (2.4)*, then ξ is a stationary distribution of P.*

Proof. By the hypothesis that ξ and P satisfies (2.4),

$$(\xi P)_v \;=\; \sum_{u \in V} \xi_u P_{u,v} \;=\; \sum_{u \in V} \xi_v P_{v,u} \;=\; \xi_v$$

holds for any $v \in V$. By the definition of the stationary distribution (2.3), we obtain the claim. □

Let ξ and ζ be probability distributions on V, then the *total variation distance* $\mathcal{D}_{\mathrm{tv}}$ between ξ and ζ is defined by

$$\mathcal{D}_{\mathrm{tv}}(\xi, \zeta) \stackrel{\mathrm{def}}{=} \max_{A \subset V} \left| \sum_{v \in A} (\xi_v - \zeta_v) \right| = \frac{1}{2} \|\xi - \zeta\|_1 . \tag{2.5}$$

The second equality holds by the following observation.

Observation 2.3. *If* $\xi \in \mathbb{R}^n$ *satisfies* $\sum_{v\in V} \xi_v = 0$, *then,*

$$\max_{A\subseteq V} |\xi_A| = \frac{1}{2}\|\xi\|_1 = \sum_{x\in V:\xi_x>0} \xi_x$$

holds.

Proof. For convenience, let $\xi^+, \xi^- \in \mathbb{R}^n$ be defined by $\xi_i^+ \stackrel{\text{def}}{=} \max\{\xi_i, 0\}$ and $\xi_i^- \stackrel{\text{def}}{=} \max\{-\xi_i, 0\}$. Then, it is easy to check that

$$\xi_i^+ - \xi_i^- = \max\{\xi_i, 0\} - \max\{-\xi_i, 0\} = \xi_i, \tag{2.6}$$

$$\xi_i^+ + \xi_i^- = \max\{\xi_i, 0\} + \max\{-\xi_i, 0\} = |\xi_i| \tag{2.7}$$

hold. (2.6) and (2.7) imply that

$$\sum_{i\in V} \xi_i^+ - \sum_{i\in V} \xi_i^- = \sum_{i\in V} \xi_i = 0, \tag{2.8}$$

$$\sum_{i\in V} \xi_i^+ + \sum_{i\in V} \xi_i^- = \sum_{i\in V} |\xi_i| = \|\xi\|_1, \tag{2.9}$$

and hence

$$\sum_{i\in V} \xi_i^+ = \sum_{i\in V} \xi_i^- = \frac{1}{2}\|\xi\|_1 \tag{2.10}$$

holds. This derives the second equality of (2.5) since $\sum_{i\in V} \xi_i^+ = \sum_{x\in V:\xi_x>0} \xi_x$. The first equality of (2.5) follows from

$$\max_{A\subseteq V} |\xi_A| = \max\left\{\sum_{x\in V:\xi_x>0} \xi_x, \sum_{x\in V:\xi_x<0} (-\xi_x)\right\} = \max\left\{\sum_{i\in V} \xi_i^+, \sum_{i\in V} \xi_i^-\right\} = \frac{1}{2}\|\xi\|_1.$$

□

Note that

$$0 \leq \mathcal{D}_{\text{tv}}(\xi, \zeta) \leq 1 \tag{2.11}$$

holds for any ξ and ζ, since $\|\xi\|_1$ and $\|\zeta\|_1$ are respectively equal to one. Let

$$d(t) \stackrel{\text{def}}{=} \max_{v\in V} \mathcal{D}_{\text{tv}}(P^t_{v,\cdot}, \pi). \tag{2.12}$$

The *mixing time* of a Markov chain is defined by

$$\tau(\varepsilon) \stackrel{\text{def}}{=} \min\{t \in \mathbb{Z}_{\geq 0} \mid d(t) \leq \varepsilon\} \tag{2.13}$$

for any $\varepsilon > 0$. In other words, the distribution $P^t_{v,\cdot}$ of the Markov chain after $\tau(\varepsilon)$ transitions satisfies $\mathcal{D}_{\mathrm{tv}}(P^t_{v,\cdot}, \pi) \leq \varepsilon$, and we obtain an approximate sample from the target distribution. For convenience, let

$$t_{\mathrm{mix}} \stackrel{\text{def}}{=} \tau(1/4), \tag{2.14}$$

which is often used as an important characterization of the mixing time of P (cf. (4.36) of [65]).

2.3 Rotor-router model

This section introduces the *rotor-router model* [57, 22], which is a simple deterministic process analogous to a simple random walk.

2.3.1 Multiple random walks

As a preliminary step, Section 2.3.1 introduces the notations for multiple random walks according to P. Let $\mu^{(0)} = (\mu^{(0)}_1, \ldots, \mu^{(0)}_n) \in \mathbb{Z}^n_{\geq 0}$ denote an initial configuration of k tokens over V. Then, at each time step $t \in \mathbb{Z}_{\geq 0}$, each token move randomly according to P, i.e., each token on $v \in V$ moves independently to $u \in V$ with probability $P_{v,u}$. Let $\mu^{(t)} = (\mu^{(t)}_1, \ldots, \mu^{(t)}_n) \in \mathbb{R}^n_{\geq 0}$ denote the *expected* configuration of tokens at time $t \in \mathbb{Z}_{\geq 0}$. Then

$$(\mu^{(t)})_v = \sum_{u \in V} \mu^{(t-1)}_u P_{u,v} = (\mu^{(t-1)} P)_v = \cdots = (\mu^{(0)} P^t)_v \tag{2.15}$$

holds for any v. Note that

$$\left|\mu^{(T)}_S - k\pi_S\right| \;=\; \left|\sum_{u \in V} \mu^{(0)}_u (P^T_{u,S} - \pi_S)\right| \;\leq\; \sum_{u \in V} \mu^{(0)}_u |P^T_{u,S} - \pi_S| \;\leq\; kd(T) \tag{2.16}$$

holds, thus $\mathcal{D}_{\mathrm{tv}}(\mu^{(t)}/k, \pi) \leq \varepsilon$ holds after $t \geq \tau(\varepsilon)$.

Let $\mathcal{G} = (V, \mathcal{E})$ denote the *transition diagram* of P where $\mathcal{E} = \{(u, v) \in V \times V \mid P_{u,v} > 0\}$. Let $\mathcal{N}^+(v)$ and $\mathcal{N}^-(v)$ respectively denote the out-neighborhood and in-neighborhood of $v \in V$ ($\mathcal{N}^+(v) = \{u \in V \mid P_{v,u} > 0\}$ and $\mathcal{N}^-(v) = \{u \in V \mid P_{u,v} > 0\}$).

Let $\delta^+(v) = |\mathcal{N}^+(v)|$ and $\delta^-(v) = |\mathcal{N}^-(v)|$. Since $\mathcal{N}^+(v) = \mathcal{N}^-(v)$ holds for any $v \in V$ for reversible P, we use $\mathcal{N}(v)(= \mathcal{N}^+(v) = \mathcal{N}^-(v))$ and $\delta(v)(= \delta^+(v) = \delta^-(v))$.

Let $G = (V, E)$ be an undirected graph. Then, in a *simple random walk* on G, each token on v moving to a randomly picked neighbor at each t, i.e., the transition matrix of a simple random walk on G is defined by

$$P_{u,v} = \begin{cases} 1/\delta(u) & (\text{if } \{u,v\} \in E) \\ 0 & (\text{otherwise}) \end{cases} . \tag{2.17}$$

2.3.2 Rotor-router model: model description

The rotor-router model is an analogy of a simple random walk on G. First, we define an arbitrary ordering $\rho_v(0), \rho_v(1), \ldots, \rho_v(\delta(v) - 1)$ over $\mathcal{N}(v)$ for each $v \in V$, and let $\rho_v(j) \stackrel{\text{def}}{=} \rho_v(j \bmod \delta(v))$. Let $\chi^{(0)} \in \mathbb{Z}^n_{\geq 0}$ be an initial configuration of tokens, and let $\chi^{(t)} \in \mathbb{Z}^n_{\geq 0}$ denote the configuration of tokens at time $t \in \mathbb{Z}_{\geq 0}$ in the rotor-router model. A configuration $\chi^{(t)}$ is updated by *rotor-routers* on vertices, as follows. At first time step ($t = 0$), there are $\chi^{(0)}_v$ tokens on vertex v, and each v serves tokens to neighbors according to $\rho_v(0), \rho_v(1), \ldots, \rho_v(\chi^{(0)}_v - 1)$. In other words, $|\{j \in [0, \chi^{(0)}_v) \mid \rho_v(j) = u\}|$ tokens move from v to u, and $\chi^{(1)}_u = \sum_{v \in V} |\{j \in [0, \chi^{(0)}_v) \mid \rho_v(j) = u\}|$. Next time step ($t = 1$), there are $\chi^{(1)}_v$ tokens on vertex v, and each v serves tokens to neighbors according to $\rho_v(\chi^{(0)}_v), \rho_v(\chi^{(0)}_v + 1), \ldots, \rho_v(\chi^{(0)}_v + \chi^{(1)}_v - 1)$, and $\chi^{(2)}$ is defined in a similar way.

In general, $z^{(t)}_{v,u}$, which denotes the number of tokens moving from v to u at time t is defined as

$$z^{(t)}_{v,u} = \left| \left\{ j \in [0, \chi^{(t)}_v) \mid \sigma_v(X^{(t)}_v + j) = u \right\} \right| \tag{2.18}$$

where $X^{(T)}_v \stackrel{\text{def}}{=} \sum_{t=0}^{T-1} \chi^{(t)}_v$ ($X^{(0)}_v = 0$). Then, $\chi^{(t+1)}$ is defined by

$$\chi^{(t+1)}_u = \sum_{v \in V} z^{(t)}_{v,u}. \tag{2.19}$$

It is not difficult to see that

$$\frac{|\{j \in \{0, \ldots, z-1\} \mid \rho_v(j) = u\}|}{z} \xrightarrow{z \to \infty} \frac{1}{\delta(v)} \tag{2.20}$$

holds, which implies that the "outflow ratio" of tokens at v to u approaches asymptotically to $1/\delta(v)$ as T increasing, e.g.,

$$\frac{Z^{(T)}_{v,u}}{X^{(T)}_v} \xrightarrow{T \to \infty} \frac{1}{\delta(v)} \tag{2.21}$$

holds where $Z_{v,u}^{(T)} = \sum_{t=0}^{T-1} z_{v,u}^{(t)}$. Thus, the rotor-router is expected to approximate the distribution of tokens by a simple random walk.

2.4 Geometric convergence of Markov chains

This section gives some analyses of $d(t)$ given by (2.12), related measures, based on the standard argument of the analysis of the mixing time (see [4, 67, 65] for details), which we will use in Chapters 4, 5, 6. Section 2.4.1 gives upper bounds of $d(t)$ and $\bar{d}(t)$ defined by

$$\bar{d}(t) \stackrel{\text{def}}{=} \max_{u,v \in V} \mathcal{D}_{\text{tv}}(P^t_{u,\cdot}, P^t_{v,\cdot}). \tag{2.22}$$

Section 2.4.2 gives the upper bound of $\tilde{d}(t)$ defined by

$$\tilde{d}(t) \stackrel{\text{def}}{=} \max_{v \in V} \mathcal{D}_{\text{tv}}(P^t_{v,\cdot}, P^{t+1}_{v,\cdot}). \tag{2.23}$$

Section 2.4.3 analyses the *separation distance* introduced in [3].

2.4.1 Distances from stationary

Section 2.4.1 establishes the following proposition.

Proposition 2.4.

$$\bar{d}(\ell t_{\text{mix}} + k) \le \frac{1}{2^\ell} \ (\forall \ell \ge 0,\ \forall k \ge 0), \tag{2.24}$$

$$d(\ell t_{\text{mix}} + k) \le \frac{1}{2^{\ell+1}} \ (\forall \ell > 0,\ \forall k \ge 0). \tag{2.25}$$

We will use Proposition 2.4 in Chapters 4, 5, 6. To prove Proposition 2.4, we use the following lemmas.

Lemma 2.5 ([65]). *Let $\xi, \zeta \in \mathbb{R}^n$ be arbitrary probability distributions. Then,*

$$\mathcal{D}_{\text{tv}}(\xi P^t, \zeta P^t) \le \bar{d}(t)$$

holds for any $t \ge 0$.

Proof. From the definition (2.5) of $\mathcal{D}_{\rm tv}$, we have

$$\begin{aligned}2\mathcal{D}_{\rm tv}(\xi P^t,\zeta P^t) &= \left\|\xi P^t-\zeta P^t\right\|_1 = \sum_{w\in V}\left|(\xi P^t)_w-(\zeta P^t)_w\right| \\ &= \sum_{w\in V}\left|\sum_{x\in V}\xi_x P^t_{x,w}-\sum_{y\in V}\zeta_y P^t_{y,w}\right| \\ &= \sum_{w\in V}\left|\sum_{x\in V}\sum_{y\in V}\xi_x\zeta_y P^t_{x,w}-\sum_{x\in V}\sum_{y\in V}\xi_x\zeta_y P^t_{y,w}\right| \\ &= \sum_{w\in V}\left|\sum_{x\in V}\sum_{y\in V}\xi_x\zeta_y (P^t_{x,w}-P^t_{y,w})\right|. \end{aligned} \tag{2.26}$$

Note that we use $\sum_{x\in V}\xi_x=\sum_{y\in V}\zeta_y=1$ in the second equality. We also obtain

$$\begin{aligned}&\sum_{w\in V}\left|\sum_{x\in V}\sum_{y\in V}\xi_x\zeta_y (P^t_{x,w}-P^t_{y,w})\right| \leq \sum_{w\in V}\sum_{x\in V}\sum_{y\in V}\xi_x\zeta_y\left|P^t_{x,w}-P^t_{y,w}\right| \\ &= \sum_{x\in V}\sum_{y\in V}\xi_x\zeta_y\left\|P^t_{x,\cdot}-P^t_{y,\cdot}\right\|_1 \leq \max_{x,y\in V}\left\|P^t_{x,\cdot}-P^t_{y,\cdot}\right\|_1\sum_{x\in V}\sum_{y\in V}\xi_x\zeta_y = 2\bar{d}(t)\end{aligned} \tag{2.27}$$

from the definition (2.22) of $\bar{d}(t)$. Combining (2.26) and (2.27), we obtain the claim. □

Lemma 2.6 ([65](Lemma 4.11))**.**

$$d(t)\leq\bar{d}(t)\leq 2d(t)$$

holds for any $t\geq 0$*.*

Proof. Suppose $d(t)=\mathcal{D}_{\rm tv}(P^t_{v^*,\cdot},\pi)$ holds for $v^*\in V$ (see (2.12)). Then,

$$d(t) = \mathcal{D}_{\rm tv}(P^t_{v,\cdot},\pi) = \mathcal{D}_{\rm tv}(e_{v^*}P^t,\pi P^t) \leq \bar{d}(t)$$

holds by Lemma 2.5, thus we obtain $d(t)\leq\bar{d}(t)$. Similarly, suppose $\bar{d}(t)=\mathcal{D}_{\rm tv}(P^t_{v^*,\cdot},P^t_{w^*,\cdot})$ holds for $v^*,w^*\in V$ (see (2.22)). Then,

$$\begin{aligned}\bar{d}(t) &= \mathcal{D}_{\rm tv}(P^t_{v^*,\cdot},P^t_{w^*,\cdot}) = \frac{1}{2}\sum_{u\in V}\left|P^t_{v^*,u}-P^t_{w^*,u}\right| \\ &= \frac{1}{2}\sum_{u\in V}\left|P^t_{v^*,u}-\pi_u+\pi_u-P^t_{w^*,u}\right| \leq \frac{1}{2}\sum_{u\in V}\left|P^t_{v^*,u}-\pi_u\right|+\frac{1}{2}\sum_{u\in V}\left|\pi_u-P^t_{w^*,u}\right| \leq 2d(t)\end{aligned}$$

holds, and we obtain the claim. □

Lemma 2.7 ([67](Lemma 5.7)). *Suppose that $\xi \in \mathbb{R}^n$ satisfies $\sum_{v \in V} \xi_v = 0$. Then,*

$$\|\xi P^t\|_1 \leq \|\xi\|_1 \bar{d}(t)$$

holds for any $t \geq 0$.

Proof. Let $\xi^+, \xi^- \in \mathbb{R}^n$ be defined by $\xi_i^+ \stackrel{\text{def}}{=} \max\{\xi_i, 0\}$ and $\xi_i^- \stackrel{\text{def}}{=} \max\{-\xi_i, 0\}$. Then,

$$\xi = \xi^+ - \xi^- \tag{2.28}$$

from (2.6), and $\frac{\xi^+}{\frac{1}{2}\|\xi\|_1}$, $\frac{\xi^-}{\frac{1}{2}\|\xi\|_1}$ are probabilistic distributions, respectively from (2.10). Thus, by Lemma 2.5,

$$\left\|\xi P^t\right\|_1 = \left\|\xi^+ P^t - \xi^- P^t\right\|_1 = \frac{1}{2}\|\xi\|_1 \cdot \left\|\frac{\xi^+}{\frac{1}{2}\|\xi\|_1} P^t - \frac{\xi^-}{\frac{1}{2}\|\xi\|_1} P^t\right\|_1 \leq \|\xi\|_1 \bar{d}(t)$$

holds, and we obtain the claim. □

Lemma 2.8 ([4](Lemma 2.20 and (2.17))).

(*a*) $\bar{d}(s+t) \leq \bar{d}(s)\bar{d}(t)$, *for any* $s, t \geq 0$ (**the submultiplicativity property**).

(*b*) $d(s+t) \leq d(s)\bar{d}(t)$, *for any* $s, t \geq 0$.

Proof. Suppose $\bar{d}(s+t) = (1/2)\|e_{v^*}P^{s+t} - e_{w^*}P^{s+t}\|_1$ holds for $v^*, w^* \in V$ (see (2.22)). Then,

$$\begin{aligned} 2\bar{d}(s+t) &= \left\|e_{v^*}P^{s+t} - e_{w^*}P^{s+t}\right\|_1 = \left\|\left(e_{v^*}P^s - e_{w^*}P^s\right) P^t\right\|_1 \\ &\leq \left\|e_{v^*}P^s - e_{w^*}P^s\right\|_1 \bar{d}(t) \leq 2\bar{d}(s)\bar{d}(t) \end{aligned}$$

holds by Lemma 2.7, and we get $\bar{d}(s+t) \leq \bar{d}(s)\bar{d}(t)$.

Similarly, suppose $d(s+t) = (1/2)\|e_{v^*}P^{s+t} - \pi\|_1$ holds for $v^* \in V$ (see (2.12)). Then,

$$2d(s+t) = \left\|e_{v^*}P^{s+t} - \pi\right\|_1 = \left\|\left(e_{v^*}P^s - \pi\right) P^t\right\|_1 \leq \left\|e_{v^*}P^s - \pi\right\|_1 \bar{d}(t) \leq 2d(s)\bar{d}(t)$$

holds by Lemma 2.7, and we get $d(s+t) \leq d(s)\bar{d}(t)$. □

Proof of Proposition 2.4. Lemma 2.8 implies that $d(t)$ and $\bar{d}(t)$ are monotone non increasing, i.e.,

$$\bar{d}(t) \geq \bar{d}(t+1), \tag{2.29}$$

$$d(t) \geq d(t+1) \tag{2.30}$$

hold since $0 \leq \bar{d}(t) \leq 1$ and $0 \leq d(t) \leq 1$ hold by the definitions of $d(t)$, $\bar{d}(t)$ and (2.11). Then,

$$\bar{d}(\ell t_{\mathrm{mix}} + k) \;\leq\; \bar{d}(\ell t_{\mathrm{mix}}) \;\leq\; \bar{d}(t_{\mathrm{mix}})^{\ell} \;\leq\; \big(2d(t_{\mathrm{mix}})\big)^{\ell} \;\leq\; \left(2 \cdot \frac{1}{4}\right)^{\ell} \;=\; \frac{1}{2^{\ell}} \tag{2.31}$$

and

$$d(\ell t_{\mathrm{mix}} + k) \;\leq\; d(\ell t_{\mathrm{mix}}) \;\leq\; d(t_{\mathrm{mix}})\bar{d}\big((\ell-1)t_{\mathrm{mix}}\big) \;\leq\; \frac{1}{4} \cdot \frac{1}{2^{\ell-1}} \;=\; \frac{1}{2^{\ell+1}}$$

hold. □

2.4.2 Distance between $P^t_{u,\cdot}$ and $P^{t+1}_{u,\cdot}$ for lazy chains

A transition matrix P is *lazy* if $P_{x,x} \geq 1/2$ holds for any $x \in V$. This section gives the following proposition.

Proposition 2.9. *For any ergodic and lazy Markov chain,*

$$\tilde{d}(\ell t_{\mathrm{mix}} + k) \leq \frac{12}{2^{\ell}\sqrt{k}}$$

holds for any $\ell \geq 0$ and for any $k > 0$.

Notice that Proposition 2.9 provides a better bound for lazy chains than Proposition 2.4 which implies $\tilde{d}(\ell t_{\mathrm{mix}} + k) = \mathrm{O}(2^{-\ell})$. We use this proposition in the analyses of Chapters 4 and 5. To obtain this proposition, we show the following lemma.

Lemma 2.10.

$$\tilde{d}(s+t) \leq \tilde{d}(s)\bar{d}(t)$$

holds for any $s, t \geq 0$.

Proof. Suppose $\tilde{d}(s+t) = (1/2)\|e_{u^*}P^{s+t} - e_{u^*}P^{s+t+1}\|_1$ holds for $u^* \in V$ (see (2.23)). Then,

$$\begin{aligned} 2\tilde{d}(s+t) \;&=\; \|e_{u^*}P^{s+t} - e_{u^*}P^{s+t+1}\|_1 \;=\; \|(e_{u^*}P^{s} - e_{u^*}P^{s+1})P^t\|_1 \\ &\leq\; \|e_{u^*}P^{s} - e_{u^*}P^{s+1}\|_1\bar{d}(t) \;\leq\; 2\tilde{d}(s)\bar{d}(t) \end{aligned}$$

holds from the definition and Lemma 2.7, thus we obtain the claim. □

For any lazy chain, we have the following lemma.

Lemma 2.11 ([65](Proposition 5.6)). *For any ergodic and lazy Markov chain,*

$$\tilde{d}(t) \leq \frac{12}{\sqrt{t}}$$

holds for any $t > 0$.

Proof of Proposition 2.9. Combining Lemma 2.10, Lemma 2.11 and Proposition 2.4, we obtain

$$\tilde{d}(\ell t_{\mathrm{mix}} + k) \leq \tilde{d}(k)\bar{d}(\ell t_{\mathrm{mix}}) \leq \frac{12}{2^{\ell}\sqrt{k}}. \tag{2.32}$$

□

2.4.3 Separation distance

The *separation distance* introduced in [3], is defined by

$$s(t) \stackrel{\text{def}}{=} \max_{u,v \in V} \left(1 - \frac{P^t_{u,v}}{\pi_v}\right). \tag{2.33}$$

We will use the discussions of this section in Chapter 6. First, we remark that

$$0 \leq s(t) \leq 1 \tag{2.34}$$

holds. The latter inequality is trivial by the definition. For the former inequality, there exists a pair $u^*, v^* \in V$ satisfying $1 - (P^t_{u^*,v^*}/\pi_{v^*}) \geq 0$ $(\Leftrightarrow)$ $\pi_{v^*} - P^t_{u^*,v^*} \geq 0$ since $\sum_{v \in V}(\pi_v - P^t_{u,v}) = 0$ holds for any $u \in V$. This means that $0 \leq s(t)$.

Lemma 2.12 ([3](Lemma 3.7)). *For any ergodic Markov chain,*

$$s(t+u) \leq s(t)s(u) \text{ [\textbf{the submultiplicativity property}]}$$

holds for any $t, u \geq 1$.

Proof. Let

$$Q^{(t)}_{x,y} \stackrel{\text{def}}{=} \frac{1}{s(t)} \left(P^t_{x,y} - \big(1 - s(t)\big)\pi_y\right) \tag{2.35}$$

for any $x, y \in V$. Then, $\sum_{y \in V} Q^{(t)}_{x,y} = 1$ and $Q^{(t)}_{x,y} \geq 0$ holds for any $x, y \in V$ and $t \geq 1$ since $s(t) \geq 1 - (P^t_{x,y}/\pi_y)$ holds for any $x, y \in V$ and $t \geq 1$. Thus $Q^{(t)} \in \mathbb{R}^{n \times n}$ is a transition matrix for any $t \geq 1$, and

$$\begin{aligned}\sum_{x \in V} \pi_x Q^{(t)}_{x,y} &= \frac{1}{s(t)} \left(\sum_{x \in V} \pi_x P^t_{x,y} - \sum_{x \in V} \pi_x \big(1 - s(t)\big) \pi_y \right) \\ &= \frac{1}{s(t)} \left(\pi_y - \big(1 - s(t)\big)\pi_y \right) = \pi_y \end{aligned} \tag{2.36}$$

holds. Since $P^t_{x,y} = s(t) Q^{(t)}_{x,y} + \big(1 - s(t)\big)\pi_y$,

$$\begin{aligned} P^{t+u}_{x,y} &= \sum_{z \in V} P^t_{x,z} P^u_{z,y} \\ &= \sum_{z \in V} \left(s(t) Q^{(t)}_{x,z} + \big(1 - s(t)\big)\pi_z \right) \left(s(u) Q^{(u)}_{z,y} + \big(1 - s(u)\big)\pi_y \right) \\ &= s(t)s(u) \sum_{z \in V} Q^{(t)}_{x,z} Q^{(u)}_{z,y} + s(t) \sum_{z \in V} Q^{(t)}_{x,z} \big(1 - s(u)\big)\pi_y \\ &\quad + \big(1 - s(t)\big) \sum_{z \in V} \pi_z s(u) Q^{(u)}_{z,y} + \big(1 - s(t)\big) \sum_{z \in V} \pi_z \big(1 - s(u)\big)\pi_y \\ &= s(t)s(u) \sum_{z \in V} Q^{(t)}_{x,z} Q^{(u)}_{z,y} + s(t)\big(1 - s(u)\big)\pi_y \\ &\quad + \big(1 - s(t)\big) s(u) \pi_y + \big(1 - s(t)\big)\big(1 - s(u)\big)\pi_y \\ &= \big(1 - s(t)s(u)\big)\pi_y + s(t)s(u) \sum_{z \in V} Q^{(t)}_{x,z} Q^{(u)}_{z,y} \end{aligned} \tag{2.37}$$

holds for any $x, y \in V$ and $t, u \geq 1$, where we used (2.36) in the fourth equality. Suppose that $x^*, y^* \in V$ is a pair satisfying $s(t+u) = 1 - (P^{t+u}_{x^*,y^*}/\pi_{y^*})$. Notice that $s(t+u) \geq 0$ by (2.34). Then, from (2.37),

$$\begin{aligned} s(t)s(u) &= \frac{\pi_{y^*} - P^{t+u}_{x^*,y^*}}{\pi_{y^*} - \sum_{z \in V} Q^{(t)}_{x^*,z} Q^{(u)}_{z,y^*}} = \frac{\pi_{y^*} - P^{t+u}_{x^*,y^*}}{\pi_{y^*}} \cdot \frac{\pi_{y^*}}{\pi_{y^*} - \sum_{z \in V} Q^{(t)}_{x^*,z} Q^{(u)}_{z,y^*}} \\ &= s(t+u) \cdot \frac{\pi_{y^*}}{\pi_{y^*} - \sum_{z \in V} Q^{(t)}_{x^*,z} Q^{(u)}_{z,y^*}} \geq s(t+u) \end{aligned}$$

holds, and we obtain the claim. □

In particular, the following holds for reversible chains.

Lemma 2.13 ([65] (Lemma 19.3))**.** *If P is ergodic and reversible, then,*

$$s(2t) \leq 1 - \big(1 - \bar{d}(t)\big)^2$$

holds for any $t \geq 0$.

Proof. Since P is reversible, we have

$$\begin{aligned}
\frac{P^{2t}_{x,y}}{\pi_y} &= \frac{\sum_{z\in V} P^t_{x,z}P^t_{z,y}}{\pi_y} = \sum_{z\in V} P^t_{x,z}\frac{P^t_{z,y}}{\pi_y} = \sum_{z\in V} P^t_{x,z}\frac{P^t_{y,z}}{\pi_z} = \sum_{z\in V}\frac{P^t_{x,z}P^t_{y,z}}{\pi_z}\sum_{z\in V}\pi_z \\
&\geq \left(\sum_{z\in V}\sqrt{\frac{P^t_{x,z}P^t_{y,z}}{\pi_z}}\sqrt{\pi_z}\right)^2 = \left(\sum_{z\in V}\sqrt{P^t_{x,z}P^t_{y,z}}\right)^2 \\
&\geq \left(\sum_{z\in V}\min\{P^t_{x,z}, P^t_{y,z}\}\right)^2, \qquad (2.38)
\end{aligned}$$

where we used Cauchy-Schwarz inequality for the first inequality. Then, by using Observation 2.3,

$$\begin{aligned}
\sum_{z\in V}\min\{P^t_{x,z}, P^t_{y,z}\} &= \sum_{\substack{z\in V\\ :P^t_{x,z}\leq P^t_{y,z}}} P^t_{x,z} + \sum_{\substack{z\in V\\ :P^t_{x,z}> P^t_{y,z}}} P^t_{y,z} \\
&= \sum_{\substack{z\in V\\ :P^t_{x,z}\leq P^t_{y,z}}} P^t_{x,z} + \left(\sum_{\substack{z\in V\\ :P^t_{x,z}> P^t_{y,z}}} P^t_{x,z} - \sum_{\substack{z\in V\\ :P^t_{x,z}> P^t_{y,z}}} P^t_{x,z}\right) + \sum_{\substack{z\in V\\ :P^t_{x,z}> P^t_{y,z}}} P^t_{y,z} \\
&= \sum_{z\in V} P^t_{x,z} - \sum_{\substack{z\in V\\ :P^t_{x,z}> P^t_{y,z}}} (P^t_{x,z} - P^t_{y,z}) \\
&= 1 - \mathcal{D}_{\mathrm{tv}}(P^t_{x,\cdot}, P^t_{y,\cdot}) \qquad (2.39)
\end{aligned}$$

holds, and we obtain the claim.

□

Chapter 3

Deterministic Random Walks

This chapter introduces *deterministic random walk*, which is a deterministic process analogous to a Markovian process $\mu^{(t)}$ (see (2.15)) with a transition matrix $P \in \mathbb{R}_{\geq 0}^{n \times n}$. Section 3.1 presents the concept of the model, and Sections 3.2 and 3.3 give some specific models of deterministic random walks.

3.1 Framework of deterministic random walks

3.1.1 Model description

Let $V = \{1, 2, \ldots, n\}$ be a finite set and $P \in \mathbb{R}_{\geq 0}^{n \times n}$ be a transition matrix on V. Let $\chi^{(0)} = (\chi_1^{(0)}, \chi_2^{(0)}, \ldots, \chi_n^{(0)}) \in \mathbb{Z}_{\geq 0}^n$ denote a initial configuration of k tokens over V in a deterministic random walk. In particular, we assume

$$\chi^{(0)} = \mu^{(0)} \tag{3.1}$$

holds. Let $\chi^{(t)} \in \mathbb{Z}_{\geq 0}^n$ denote the configuration of tokens at time $t \in \mathbb{Z}_{\geq 0}$ in a deterministic random walk. An update in a deterministic random walk is described by the number of tokens moving from v to u at time t denoted $z_{v,u}^{(t)} \in \mathbb{Z}_{\geq 0}$. Here, $z_{v,u}^{(t)}$ must satisfy the condition that

$$\sum_{u \in \mathcal{N}^+(v)} z_{v,u}^{(t)} = \chi_v^{(t)} \tag{3.2}$$

for any $v \in V$. Then, $\chi^{(t+1)}$ is defined by

$$\chi_u^{(t+1)} \stackrel{\text{def}}{=} \sum_{v \in \mathcal{N}^-(u)} z_{v,u}^{(t)} \tag{3.3}$$

for any $u \in V$. We will explain some specific models of the deterministic random walk in Sections 3.2 and 3.3 by giving precise definitions of $z_{v,u}^{(t)}$.

3.1.2 Foundation of discrepancy analyses

The goal of this study is to analyze the discrepancy between $\chi^{(t)}$ and $\mu^{(t)}$. This section establishes Lemma 3.1, which is the framework of discrepancy analyses. To state Lemma 3.1, we introduce the *local discrepancy* $\phi_{v,u}^{(t)}$ for a deterministic random walk.

Definition 1 (Local discrepancy)**.**

$$\phi_{v,u}^{(t)} \stackrel{\text{def}}{=} z_{v,u}^{(t)} - \chi_v^{(t)} P_{v,u}. \tag{3.4}$$

Let $\phi_{v,u} \stackrel{\text{def}}{=} \max_{t \geq 0} |\phi_{v,u}^{(t)}|$*, and* $\phi \stackrel{\text{def}}{=} \max_{v,u \in V} \phi_{v,u}$*, for convenience.*

Recall that there are $\chi_v^{(t)}$ tokens at vertex $v \in V$ at time t. The local discrepancy $\phi_{v,u}^{(t)}$ is the discrepancy between $\chi_v^{(t)} P_{v,u}$, which is the *expected* number of tokens move from v to $u \in \mathcal{N}^+(v)$, and $z_{v,u}^{(t)}$, which is the *actual* number of tokens move from v to $u \in \mathcal{N}^+(v)$ in a deterministic random walk. We will give upper bounds of the vertex-wise discrepancy (Theorem 4.1) and total variation discrepancy (Theorem 5.1) using ϕ. We also give the upper bound of ϕ for some specific model. Especially, ϕ is bounded by a constant for the *SRT-router model* (Section 3.2.4) and the *oblivious model* (Section 3.3). Now, we state Lemma 3.1.

Lemma 3.1.

$$\chi_w^{(T)} - \mu_w^{(T)} = \sum_{u \in V} \sum_{v \in \mathcal{N}^-(u)} \sum_{t=0}^{T-1} \phi_{v,u}^{(T-t-1)} P_{u,w}^t$$

holds for any $w \in V$ *and for any* $T \geq 0$*.*

Lemma 3.1 implies that the discrepancy between $\chi^{(t)}$ and $\mu^{(t)}$ is expressed by a geometric series of its transition matrix where the coefficients are the local discrepancy at each time. We use this lemma in Chapter 4 for the vertex-wise discrepancy, in Chapter 5 for the total variation discrepancy, and in Chapter 6 for the discrepancy of the visit frequencies concerned with the cover time.

Proof. For convenience, we assume that P^0 is the identity matrix ($P^0_{u,w} = 1$ if $u = w$ and otherwise $P^0_{u,w} = 0$). First, we see that

$$\sum_{t=0}^{T-1} \left(\chi^{(T-t)} P^t - \chi^{(T-t-1)} P^{t+1}\right) = \chi^{(T)} P^0 - \chi^{(0)} P^T$$

holds. Hence we obtain

$$\chi^{(T)} - \mu^{(T)} = \sum_{t=0}^{T-1} \left(\chi^{(T-t)} - \chi^{(T-t-1)} P\right) P^t \tag{3.5}$$

since $\chi^{(0)} P^T = \mu^{(0)} P^T = \mu^{(T)}$ holds by (3.1) and (2.15). From (3.5),

$$\begin{aligned} \chi^{(T)}_w - \mu^{(T)}_w &= \sum_{t=0}^{T-1} \left((\chi^{(T-t)} - \chi^{(T-t-1)} P) P^t\right)_w \\ &= \sum_{t=0}^{T-1} \sum_{u \in V} \left(\chi^{(T-t)} - \chi^{(T-t-1)} P\right)_u P^t_{u,w} \end{aligned} \tag{3.6}$$

and

$$\begin{aligned} \left(\chi^{(t+1)} - \chi^{(t)} P\right)_u &= \chi^{(t+1)}_u - (\chi^{(t)} P)_u = \sum_{v \in V} z^{(t)}_{v,u} - \sum_{v \in V} \chi^{(t)}_v P_{v,u} \\ &= \sum_{v \in \mathcal{N}^-(u)} (z^{(t)}_{v,u} - \chi^{(t)}_v P_{v,u}) = \sum_{v \in \mathcal{N}^-(u)} \phi^{(t)}_{v,u} \end{aligned} \tag{3.7}$$

holds for any $t \geq 0$ and $u \in V$ by (3.3) and (3.4). Thus we obtain the claim combining (3.6) and (3.7). □

3.2 Functional-router model

Recall that on the rotor router model, $Z^{(T)}_{v,u} / X^{(T)}_v$ approaches asymptotically to $1/\delta(v)$ as T increases (2.21), where

$$X^{(T)}_v \stackrel{\text{def}}{=} \sum_{t=0}^{T-1} \chi^{(t)}_v, \tag{3.8}$$

which is the total number of tokens visited vertex v by time T, and

$$Z^{(T)}_{v,u} \stackrel{\text{def}}{=} \sum_{t=0}^{T-1} z^{(t)}_{v,u}, \tag{3.9}$$

which is the total number of tokens moving from v to u by time T (we assume that $X_v^{(0)} = 0$ and $Z_{v,u}^{(0)} = 0$). Now, we consider the model satisfying the generalization of (2.21), e.g.,

$$\frac{Z_{v,u}^{(T)}}{X_v^{(T)}} \xrightarrow{T\to\infty} P_{v,u}. \tag{3.10}$$

In order to get a model satisfying (3.10), we generalize the rotor-router using the following idea: each vertex has a *functional-router* $\sigma_v : \mathbb{Z}_{\geq 0} \to \mathcal{N}^+(v)$ satisfying

$$\frac{|\{j \in \{0,\dots,z-1\} \mid \sigma_v(j) = u\}|}{z} \xrightarrow{z\to\infty} P_{v,u}, \tag{3.11}$$

meaning that local ratio of token moving imitates local transition probability of corresponding random walks. Precisely, for given $\chi^{(0)}$ and σ_v, $z_{v,u}^{(t)}$ is defined by

$$\begin{aligned} z_{v,u}^{(t)} &\stackrel{\text{def}}{=} \left|\left\{j \in [0, \chi_v^{(t)}) \mid \sigma_v\left(X_v^{(t)} + j\right) = u\right\}\right| \\ &= \left|\left\{j \in [X_v^{(t)}, X_v^{(t)} + \chi_v^{(t)}) \mid \sigma_v\left(j\right) = u\right\}\right| \end{aligned} \tag{3.12}$$

for any $v, u \in V$. We in Section 3.2 give some specific functional-routers satisfying (3.10), i.e., satisfying small *cumulative local discrepancy* Φ defined by the following.

Definition 2 (Cumulative local discrepancy)**.**

$$\Phi_{v,u}^{(T)} \stackrel{\text{def}}{=} \sum_{t=0}^{T-1} \phi_{v,u}^{(t)} = Z_{v,u}^{(T)} - X_v^{(T)} P_{v,u}. \tag{3.13}$$

Let $\Phi_{v,u} \stackrel{\text{def}}{=} \max_{T\geq 0} |\Phi_{v,u}^{(T)}|$*, and* $\Phi \stackrel{\text{def}}{=} \max_{v,u\in V} \Phi_{v,u}$*, for convenience.*

The cumulative local discrepancy satisfies $\Phi \leq 1$ for the *SRT-router model* described in Section 3.2.4. We will give upper bounds of the vertex-wise discrepancy (Theorem 4.2), the total variation discrepancy (Theorem 5.2), the discrepancy of visit frequencies (Theorem 6.1) and the cover time (Theorem 6.4) using Φ.

Now, we explain some specific functional-router models, namely the *weighted rotor-router* in Section 3.2.1, the *quasi-random router* in Section 3.2.2, the *billiard router* in Section 3.2.3, and the *SRT router* in Section 3.2.4.

3.2.1 Weighted rotor-router

The rotor-router model described in Section 2.3.2 can be generally considered on digraphs with parallel edges (i.e., multidigraphs). Kijima et al. [57] and Kajino et al. [54] are concerned with the rotor-router model on finite multidigraphs. Suppose that P is a transition matrix with *rational* entries. For each $v \in V$, let $\bar{\delta}(v) \in \mathbb{Z}_{\geq 0}$ be a common denominator (or the least common denominator) of $P_{v,u}$ for all $u \in \mathcal{N}^+(v)$, meaning that $\bar{\delta}(v) \cdot P_{v,u}$ is an integer for each $u \in \mathcal{N}^+(v)$. We define a rotor-router $\sigma_v(0), \sigma_v(1), \dots, \sigma_v(\bar{\delta}(v) - 1)$ arbitrarily satisfying that

$$\left|\{j \in [0, \dots, \bar{\delta}(v)) \mid \sigma_v(j) = u\}\right| = \bar{\delta}(v) \cdot P_{v,u}$$

for any $v \in V$ and $u \in \mathcal{N}^+(v)$. Then, $\sigma_v(i)$ is defined by

$$\sigma_v(i) = \sigma_v(i \bmod \bar{\delta}(v)) \left(\equiv \sigma_v \left(i - \bar{\delta}(v) \cdot \left\lfloor \frac{i}{\bar{\delta}(v)} \right\rfloor\right)\right). \tag{3.14}$$

For the weighted rotor router, it is not difficult to observe the following.

Proposition 3.2. *If $P \in \mathbb{Q}_{\geq 0}^{n \times n}$, then,*

$$\left|\left|\{j \in [z, z') \mid \sigma_v(j) = u\}\right| - (z' - z)P_{v,u}\right| \leq \bar{\delta}(v)P_{v,u}$$

holds for any $v, u \in V$ and for any $z, z' \in \mathbb{Z}_{\geq 0}$ satisfying $z' > z$.

Proof. Considering (3.14),

$$\bar{\delta}(v)P_{v,u} \cdot \left\lfloor \frac{z' - z}{\bar{\delta}(v)} \right\rfloor \leq \left|\{j \in [z, z') \mid \sigma_v(j) = u\}\right| \leq \bar{\delta}(v)P_{v,u} \cdot \left\lceil \frac{z' - z}{\bar{\delta}(v)} \right\rceil$$

holds, and we obtain the claim. □

3.2.2 Quasi-random router

Motivated by deterministic random walks for irrational transition probabilities, this section gives a router σ based on the *van der Corput sequence* [86, 69], which is a well-known low-discrepancy sequence.

The van der Corput sequence $\psi \colon \mathbb{Z}_{\geq 0} \to [0, 1)$ is defined as follows. Suppose $i \in \mathbb{Z}_{>0}$ is represented in binary as $i = \sum_{j=0}^{\lfloor \lg i \rfloor} \beta_j(i) \cdot 2^j$ where $\beta_j(i) \in \{0, 1\}$ ($j \in \{0, 1, \dots, \lfloor \lg i \rfloor\}$) denotes the j-th bit coefficient of i. Then, we define

$$\psi(i) \stackrel{\text{def}}{=} \sum_{j=0}^{\lfloor \lg i \rfloor} \beta_j(i) \cdot 2^{-(j+1)} \tag{3.15}$$

and $\psi(0) \stackrel{\text{def}}{=} 0$. For example, $\psi(1) = 1 \times 1/2 = 1/2$, $\psi(2) = 0 \times 1/2 + 1 \times 1/4 = 1/4$, $\psi(3) = 1 \times 1/2 + 1 \times 1/4 = 3/4$, $\psi(4) = 0 \times 1/2 + 0 \times 1/4 + 1 \times 1/8 = 1/8$, $\psi(5) = 1 \times 1/2 + 0 \times 1/4 + 1 \times 1/8 = 5/8$, $\psi(6) = 0 \times 1/2 + 1 \times 1/4 + 1 \times 1/8 = 3/8$, and so on. Clearly, $\psi(i) \in [0, 1)$ holds for any (finite) $i \in \mathbb{Z}_{\geq 0}$.

Now, given $i \in \mathbb{Z}_{>0}$, we define $\sigma_v(i)$ as follows. Without loss of generality, we may assume that an ordering $u_1, \ldots, u_{\delta(v)}$ is defined on $\mathcal{N}^+(v)$ for each $v \in V$. Then, we define the functional-router $\sigma_v \colon \mathbb{Z}_{\geq 0} \to \mathcal{N}^+(v)$ on $v \in V$ such that $\sigma_v(i) = u_k \in \mathcal{N}^+(v)$ satisfies that

$$\textstyle\sum_{j=1}^{k-1} P_{v,u_j} \leq \psi(i) < \sum_{j=1}^{k} P_{v,u_j}$$

for $k \in \{1, \ldots, \delta^+(v)\}$, where $\sum_{j=1}^{0} P_{v,u_j} = 0$, for convenience.

The following proposition is due to van der Corput [86].

Proposition 3.3 ([86])**.** *For any transition matrix P,*

$$\Big|\big|\{j \in [0, z) \mid \sigma_v(j) = u\}\big| - z \cdot P_{v,u}\Big| \leq \lg(z + 1)$$

holds for any $v, u \in V$ and any $z \in \mathbb{Z}_{>0}$.

More sophisticated bounds are found in [69]. Carefully examining Proposition 3.3, the following proposition is shown.

Proposition 3.4 ([86])**.** *For any transition matrix P,*

$$\Big|\big|\{j \in [z, z') \mid \sigma_v(j) = u\}\big| - (z' - z)P_{v,u}\Big| \leq 2\lg(z' - z + 1)$$

holds for any $v, u \in V$, and for any $z, z' \in \mathbb{Z}_{\geq 0}$ satisfying $z' > z$.

3.2.3 Billiard-router

Billiard sequence is known to be a balanced sequence (cf. [74]). This section presents a functional router based on the billiard sequence. Let $\sigma_v(i)$ be $u^* \in \mathcal{N}^+(v)$ minimizing the value

$$\frac{\big|\{j \in [0, i) \mid \sigma_v(j) = u\}\big| + 1}{P_{v,u}}$$

in all $u \in \mathcal{N}^+(v)$, and if there are two or more such $u \in \mathcal{N}^+(v)$, then let u^* be arbitrary one of them. Then, the following theorem for the billiard sequence is known.

Proposition 3.5 ([74]). *For any transition matrix P,*

$$\left| \left| \left\{ j \in [z, z') \mid \sigma_v(j) = u \right\} \right| - (z' - z) P_{v,u} \right| \leq 1 + \delta^+(v) P_{v,u}$$

holds for any $v, u \in V$, and for any $z, z' \in \mathbb{Z}_{\geq 0}$ satisfying $z' > z$.

3.2.4 SRT-router

This section introduces *SRT router*, which is originally given by Holroyd and Propp [45] and Angel et al. [8] by the name of stack-walk. The SRT-router is similar to the billiard-router, but more sophisticated. The SRT router $\sigma_v(i)$ ($i \in \mathbb{Z}_{\geq 0}$) on $v \in V$ is defined, as follows. Let

$$T_i(v) = \{ u \in \mathcal{N}^+(v) \mid \left| \left\{ j \in [0, i) \mid \sigma_v(j) = u \right\} \right| - (i+1) P_{v,u} < 0 \}. \tag{3.16}$$

Then, let $\sigma_v(i)$ be a vertex $u^* \in T_i(v)$ minimizing the value

$$\frac{\left| \left\{ j \in [0, i) \mid \sigma_v(j) = u \right\} \right| + 1}{P_{v,u}} \tag{3.17}$$

in all $u \in T_i(v)$. If there are two or more such $u \in T_v(i)$, then let u^* be arbitrary one of them.

Considering $\sigma_v(i) \in T_i(v)$, we can see that $\left| \left\{ j \in [0, i+1) \mid \sigma_v(j) = u \right\} \right| - (i+1) P_{v,u} < 1$ holds for any u, v and i, by an induction on $i \in \mathbb{Z}_{\geq 0}$. The following theorem is due to Angel et al. [8] and Tijdeman [82].

Theorem 3.6. *[82, 8] For any transition matrix P,*

$$\left| \left| \left\{ j \in [0, z) \mid \sigma_v(j) = u \right\} \right| - z \cdot P_{v,u} \right| \leq 1$$

holds for any $v, u \in V$ and any $z \in \mathbb{Z}_{>0}$.

Theorem 3.6 was firstly given by Tijdeman [82], where he gave a slightly better bound $\left| \left| \left\{ j \in [0, z) \mid \sigma_v(j) = u \right\} \right| - z \cdot P_{v,u} \right| \leq 1 - (2(\delta^+(v) - 1))^{-1}$, in fact. Angel et al. [8] rediscovered Theorem 3.6 in the context of deterministic random walk (see also [45]), where they also proved a similar statement even when the corresponding probability is time-inhomogeneous.

3.3 Oblivious model

Memoryless is an important property of a Markovian process. This section presents an oblivious deterministic random walk. In the oblivious model, $z_{v,u}^{(t)}$ is defined by $\lceil \chi_v^{(t)} P_{v,u} \rceil$ or $\lfloor \chi_v^{(t)} P_{v,u} \rfloor$ to satisfy (3.2). Precisely, suppose that an arbitrary ordering $u_1, \ldots, u_{\delta^+(v)}$ on $\mathcal{N}^+(v)$ is prescribed for each $v \in V$. Then, let

$$z_{v,u_i}^{(t)} = \begin{cases} \left\lfloor \chi_v^{(t)} P_{v,u_i} \right\rfloor + 1 & (i \leq i^*) \\ \left\lfloor \chi_v^{(t)} P_{v,u_i} \right\rfloor & \text{(otherwise)} \end{cases} \tag{3.18}$$

where $i^* \stackrel{\text{def}}{=} \chi_v^{(t)} - \sum_{i=1}^{\delta^+(v)} \lfloor \chi_v^{(t)} P_{v,u_i} \rfloor$ denotes the number of "surplus" tokens. It is easy to check that the condition (3.2) holds for any $v \in V$ and $t \in \mathbb{Z}_{\geq 0}$. Then, the configuration $\chi^{(t+1)}$ is updated according to (3.3), recursively.

The following is obvious from the definition (3.18) for any oblivious model.

Proposition 3.7. *For any $P \in \mathbb{R}_{\geq 0}^{n \times n}$ and for any its corresponding oblivious models,*

$$\phi_{v,u}^{(t)} \leq 1$$

holds for any $v, u \in V$ and for any $t \geq 0$.

Note that oblivious models described in Section 3.1 do *not* satisfy (3.10) in general.

Chapter 4

Vertex-wise Discrepancy

This chapter investigates the vertex-wise discrepancy (L_∞ discrepancy) between the token configuration $\chi^{(t)}$ of a deterministic random walk (see (3.3)) and the expected token configuration $\mu^{(t)}$ of a random walk (see (2.15)). First, we present general upper bounds (Theorem 4.1 and Theorem 4.2) in Section 4.1. In Section 4.2, we show upper bounds for specific models defined in Section 3.2. Section 4.3 discusses the complexity of a simulation of deterministic random walk for some rapidly mixing chains. Section 4.4 discusses some lower bounds.

4.1 General upper bounds of the vertex-wise discrepancy

This section establishes two theorems for general upper bounds of the vertex-wise discrepancy between a deterministic random walk and a reversible Markov chain.

Theorem 4.1. *If $P \in \mathbb{R}_{\geq 0}^{n \times n}$ is ergodic and reversible, then*

$$\left|\chi_w^{(T)} - \mu_w^{(T)}\right| \leq 3\pi_w \max_{u \in V} \frac{\sum_{v \in \mathcal{N}(u)} \phi_{v,u}}{\pi_u} t_{\mathrm{mix}} = \mathrm{O}\left(\phi \frac{\pi_{\max}}{\pi_{\min}} \Delta t_{\mathrm{mix}}\right)$$

holds for any $w \in V$ and for any $T \geq 0$.

For lazy chains, we show the following better upper bound.

Theorem 4.2. *If $P \in \mathbb{R}_{\geq 0}^{n \times n}$ is ergodic, reversible and lazy, then*

$$\left|\chi_w^{(T)} - \mu_w^{(T)}\right| \leq 96\pi_w \max_{u \in V} \frac{\sum_{v \in \mathcal{N}(u)} \Phi_{v,u}}{\pi_u} \sqrt{t_{\mathrm{mix}}} = \mathrm{O}\left(\Phi \frac{\pi_{\max}}{\pi_{\min}} \Delta \sqrt{t_{\mathrm{mix}}}\right)$$

holds for any $w \in V$ and for any $T \geq 0$.

4.1.1 Proof of Theorem 4.1

To prove Theorem 4.1, we introduce the following lemma derived from Lemma 3.1.

Lemma 4.3.

$$\chi_w^{(T)} - \mu_w^{(T)} = \sum_{u \in V} \sum_{v \in \mathcal{N}^-(u)} \sum_{t=0}^{T-1} \phi_{v,u}^{(T-t-1)} (P_{u,w}^t - \pi_w)$$

holds for any $w \in V$ and for any $T \geq 0$.

Proof. First, we see that

$$\begin{aligned} \sum_{u \in V} \left(\chi^{(t+1)} - \chi^{(t)} P\right)_u &= \sum_{u \in V} \chi_u^{(t+1)} - \sum_{u \in V} \sum_{v \in V} \chi_v^{(t)} P_{v,u} \\ &= \sum_{u \in V} \chi_u^{(t+1)} - \sum_{v \in V} \chi_v^{(t)} = 0 \end{aligned} \tag{4.1}$$

holds for any $t \geq 0$. Combining (3.6) and (4.1), we obtain

$$\begin{aligned} \chi_w^{(T)} - \mu_w^{(T)} &= \sum_{t=0}^{T-1} \sum_{u \in V} \left(\chi^{(T-t)} - \chi^{(T-t-1)} P\right)_u P_{u,w}^t - 0 \\ &= \sum_{t=0}^{T-1} \sum_{u \in V} \left(\chi^{(T-t)} - \chi^{(T-t-1)} P\right)_u P_{u,w}^t - \sum_{t=0}^{T-1} \sum_{u \in V} \left(\chi^{(T-t)} - \chi^{(T-t-1)} P\right)_u \pi_w \\ &= \sum_{t=0}^{T-1} \sum_{u \in V} \left(\chi^{(T-t)} - \chi^{(T-t-1)} P\right)_u (P_{u,w}^t - \pi_w). \end{aligned} \tag{4.2}$$

Thus, we obtain the claim by (4.2) and (3.7). □

The following lemma is corresponding to the property of $d(t)$.

Lemma 4.4.

$$\sum_{t=0}^{T} d(t) \leq \frac{3}{2} t_{\mathrm{mix}}$$

holds for any $T \geq 0$.

Proof. Uisng Proposition 2.4, we obtain

$$\begin{aligned} \sum_{t=0}^{T} d(t) &\leq \sum_{\ell=0}^{\infty} \sum_{k=0}^{t_{\mathrm{mix}}-1} d(\ell t_{\mathrm{mix}} + k) = \sum_{k=0}^{t_{\mathrm{mix}}-1} d(k) + \sum_{\ell=1}^{\infty} \sum_{k=0}^{t_{\mathrm{mix}}-1} d(\ell t_{\mathrm{mix}} + k) \\ &\leq \sum_{k=0}^{t_{\mathrm{mix}}-1} 1 + \sum_{k=0}^{t_{\mathrm{mix}}-1} \sum_{\ell=1}^{\infty} \frac{1}{2^{\ell+1}} \leq \sum_{k=0}^{t_{\mathrm{mix}}-1} 1 + \sum_{k=0}^{t_{\mathrm{mix}}-1} \frac{1}{2} = \frac{3}{2} t_{\mathrm{mix}}. \end{aligned}$$

□

Proof of Theorem 4.1. Note that $\mathcal{N}^-(u) = \mathcal{N}(u)$ holds for any $u \in V$ and $\pi_u P^t_{u,w} = \pi_w P^t_{w,u}$ hold from the assumption of reversibility of P. Using Lemma 4.3, we have

$$\begin{aligned}\chi^{(T)}_w - \mu^{(T)}_w &= \sum_{u\in V}\sum_{v\in\mathcal{N}(u)}\sum_{t=0}^{T-1}\phi^{(T-t-1)}_{v,u}\left(\frac{\pi_w}{\pi_u}P^t_{w,u} - \pi_w\right)\\ &= \pi_w\sum_{u\in V}\frac{1}{\pi_u}\sum_{v\in\mathcal{N}(u)}\sum_{t=0}^{T-1}\phi^{(T-t-1)}_{v,u}(P^t_{w,u} - \pi_u). \qquad (4.3)\end{aligned}$$

Using the definitions (3.4) of $\phi_{v,u}$ and (2.12) of $d(t)$,

$$\begin{aligned}\left|\chi^{(T)}_w - \mu^{(T)}_w\right| &\le \pi_w\sum_{u\in V}\frac{1}{\pi_u}\sum_{v\in\mathcal{N}(u)}\sum_{t=0}^{T-1}\left|\phi^{(T-t-1)}_{v,u}\right|\left|P^t_{w,u} - \pi_u\right|\\ &\le \pi_w\sum_{u\in V}\frac{1}{\pi_u}\sum_{v\in\mathcal{N}(u)}\sum_{t=0}^{T-1}\phi_{v,u}|P^t_{w,u} - \pi_u|\\ &\le \pi_w\max_{u\in V}\frac{\sum_{v\in\mathcal{N}(u)}\phi_{v,u}}{\pi_u}\sum_{t=0}^{T-1}\sum_{u\in V}|P^t_{w,u} - \pi_u|\\ &\le \pi_w\max_{u\in V}\frac{\sum_{v\in\mathcal{N}(u)}\phi_{v,u}}{\pi_u}\sum_{t=0}^{T-1}2d(t) \qquad (4.4)\end{aligned}$$

holds. Thus, we obtain the claim combining (4.4) and Lemma 4.4. □

4.1.2 Proof of Theorem 4.2

For convenience, we assume that P^{-1} is the zero matrix ($P^{-1}_{u,w} = 0$ for any $u, w \in V$). To prove Theorem 4.2, we introduce the following lemma derived from Lemma 3.1.

Lemma 4.5.

$$\chi^{(T)}_w - \mu^{(T)}_w = \sum_{u\in V}\sum_{v\in\mathcal{N}^-(u)}\sum_{t=0}^{T-1}\Phi^{(T-t)}_{v,u}\left(P^t_{u,w} - P^{t-1}_{u,w}\right)$$

holds for any $w \in V$ and for any $T \ge 0$.

Proof. Note that $\Phi^{(t+1)}_{v,u} - \Phi^{(t)}_{v,u} = \phi^{(t)}_{v,u}$ holds for any $t \ge 0$ from the definition (3.13).

Hence

$$\begin{aligned}
\sum_{t=0}^{T-1} \phi_{v,u}^{(T-t-1)} P_{u,w}^t &= \sum_{t=0}^{T-1} \left(\Phi_{v,u}^{(T-t)} - \Phi_{v,u}^{(T-t-1)} \right) P_{u,w}^t \\
&= \sum_{t=0}^{T-1} \Phi_{v,u}^{(T-t)} P_{u,w}^t - \sum_{t=0}^{T-1} \Phi_{v,u}^{(T-t-1)} P_{u,w}^t \\
&= \sum_{t=0}^{T-1} \Phi_{v,u}^{(T-t)} P_{u,w}^t - \sum_{t=-1}^{T-2} \Phi_{v,u}^{(T-t-1)} P_{u,w}^t \\
&= \sum_{t=0}^{T-1} \Phi_{v,u}^{(T-t)} P_{u,w}^t - \sum_{t=0}^{T-1} \Phi_{v,u}^{(T-t)} P_{u,w}^{t-1} \\
&= \sum_{t=0}^{T-1} \Phi_{v,u}^{(T-t)} \left(P_{u,w}^t - P_{u,w}^{t-1} \right) \qquad (4.5)
\end{aligned}$$

where we used the assumptions $P_{v,u}^{-1} = 0$ and $\Phi_{v,u}^{(0)} = 0$ for any $v, u \in V$ in the third equality. We obtain the claim combining Lemma 3.1 and (4.5). □

The following lemma is corresponding to the property of $\tilde{d}(t)$.

Lemma 4.6. *If P is ergodic and lazy, then,*

$$\sum_{t=0}^{T} \tilde{d}(t) \leq 48\sqrt{t_{\mathrm{mix}}} - 23$$

holds for any $T \geq 0$.

Proof. Using Proposition 2.9, we obtain

$$\begin{aligned}
\sum_{t=0}^{T} \tilde{d}(t) &\leq \tilde{d}(0) + \sum_{\ell=0}^{\infty} \sum_{k=1}^{t_{\mathrm{mix}}} \tilde{d}(\ell t_{\mathrm{mix}} + k) \\
&\leq 1 + \sum_{\ell=0}^{\infty} \sum_{k=1}^{t_{\mathrm{mix}}} \frac{12}{2^\ell \sqrt{k}} = 1 + 12 \sum_{\ell=0}^{\infty} \frac{1}{2^\ell} \sum_{k=1}^{t_{\mathrm{mix}}} \frac{1}{\sqrt{k}} \\
&\leq 1 + 12(2 \cdot (2\sqrt{t_{\mathrm{mix}}} - 1)) = 48\sqrt{t_{\mathrm{mix}}} - 23.
\end{aligned}$$

where we used the fact that $\sum_{k=1}^{T} \frac{1}{\sqrt{k}} \leq 2\sqrt{T} - 1$. □

Proof of Theorem 4.2. Note that $\mathcal{N}^-(u) = \mathcal{N}(u)$ holds for any $u \in V$ and $\pi_u P_{u,w}^t = \pi_w P_{w,u}^t$ holds for any $u, w \in V$ and $t \geq 0$ from the assumption of reversibility of P. Then

Lemma 4.5 implies that

$$\chi_w^{(T)} - \mu_w^{(T)} = \pi_w \sum_{u\in V} \frac{1}{\pi_u} \sum_{v\in\mathcal{N}(u)} \sum_{t=0}^{T-1} \Phi_{v,u}^{(T-t)} \left(P_{w,u}^t - P_{w,u}^{t-1}\right).$$

Thus, by definition (3.13),

$$\begin{aligned}
\left|\chi_w^{(T)} - \mu_w^{(T)}\right| &\le \pi_w \sum_{u\in V} \frac{1}{\pi_u} \sum_{v\in\mathcal{N}(u)} \sum_{t=0}^{T-1} \left|\Phi_{v,u}^{(T-t)}\right| \left|P_{w,u}^t - P_{w,u}^{t-1}\right| \\
&\le \pi_w \sum_{u\in V} \frac{1}{\pi_u} \sum_{v\in\mathcal{N}(u)} \sum_{t=0}^{T-1} \Phi_{v,u} \left|P_{w,u}^t - P_{w,u}^{t-1}\right| \\
&\le \pi_w \max_{u\in V} \frac{\sum_{v\in\mathcal{N}(u)} \Phi_{v,u}}{\pi_u} \sum_{t=0}^{T-1} \sum_{u\in V} \left|P_{w,u}^t - P_{w,u}^{t-1}\right| \qquad (4.6)
\end{aligned}$$

holds. Using Lemma 4.6, we have

$$\begin{aligned}
\sum_{t=0}^{T-1} \sum_{u\in V} \left|P_{w,u}^t - P_{w,u}^{t-1}\right| &= \sum_{t=-1}^{T-2} \sum_{u\in V} \left|P_{w,u}^{t+1} - P_{w,u}^t\right| = \sum_{u\in V} \left|P_{w,u}^0\right| + \sum_{t=0}^{T-2} \sum_{u\in V} \left|P_{w,u}^t - P_{w,u}^{t+1}\right| \\
&\le 1 + \sum_{t=0}^{T-2} 2\tilde{d}(t) \le 96\sqrt{t_{\mathrm{mix}}}. \qquad (4.7)
\end{aligned}$$

Thus we obtain the claim combining (4.6) and (4.7). □

4.2 Upper bounds for specific models

Now, we give upper bounds of the vertex-wise discrepancy for specific models. As a preliminary step, we show the following properties for any functional-router model.

Proposition 4.7.

$$\begin{aligned}
\phi_{v,u} &\le \max_{z'>z} \left| \left|\left\{j \in [z, z') \mid \sigma_v(j) = u\right\}\right| - (z'-z)P_{v,u} \right| \\
&\le 2\max_{z} \left| \left|\left\{j \in [0, z) \mid \sigma_v(j) = u\right\}\right| - z\cdot P_{v,u} \right|
\end{aligned}$$

and

$$\begin{aligned}
\Phi_{v,u} &\le \max_{z} \left| \left|\left\{j \in [0, z) \mid \sigma_v(j) = u\right\}\right| - z\cdot P_{v,u} \right| \\
&\le \max_{z'>z} \left| \left|\left\{j \in [z, z') \mid \sigma_v(j) = u\right\}\right| - (z'-z)P_{v,u} \right|
\end{aligned}$$

holds for any functional-router models.

Proof.

$$\begin{aligned}
\phi_{v,u} &= \max_{t\geq 0} |Z^{(t)}_{v,u} - \chi^{(t)}_v P_{v,u}| \\
&= \max_{t\geq 0} \left| \left|\left\{ j \in [X^{(t)}_v, X^{(t)}_v + \chi^{(t)}_v) \mid \sigma_v(j) = u \right\}\right| - \chi^{(t)}_v P_{v,u} \right| \\
&\leq \max_{z'>z} \left| \left|\left\{ j \in [z, z') \mid \sigma_v(j) = u \right\}\right| - (z'-z) P_{v,u} \right|.
\end{aligned}$$

$$\begin{aligned}
\Phi_{v,u} &= \max_{T>0} \left| \sum_{t=0}^{T-1} (Z^{(t)}_{v,u} - \chi^{(t)}_v P_{v,u}) \right| = \max_{T>0} \left| \sum_{t=0}^{T-1} Z^{(t)}_{v,u} - X^{(T)}_v P_{v,u} \right| \\
&= \max_{T>0} \left| \sum_{t=0}^{T-1} \left|\left\{ j \in [X^{(t)}_v, X^{(t)}_v + \chi^{(t)}_v) \mid \sigma_v(j) = u \right\}\right| - X^{(T)}_v P_{v,u} \right| \\
&= \max_{T>0} \left| \left|\left\{ j \in [0, X^{(T)}_v) \mid \sigma_v(j) = u \right\}\right| - X^{(T)}_v P_{v,u} \right| \\
&\leq \max_{z} \left| \left|\left\{ j \in [0, z) \mid \sigma_v(j) = u \right\}\right| - z \cdot P_{v,u} \right|.
\end{aligned}$$

□

4.2.1 SRT-router models

First, for SRT-router models, we give the following proposition by Proposition 4.7 and Theorem 3.6.

Proposition 4.8. *For any $P \in \mathbb{R}^{n\times n}_{\geq 0}$ and for any its corresponding SRT-router models,*

$$\begin{aligned}
\phi_{v,u} &\leq 2, \\
\Phi_{v,u} &\leq 1
\end{aligned}$$

holds for any $v, u \in V$.

Thus, it is easy to check the following upper bounds by Proposition 4.8.

Corollary 4.9. *Suppose $P \in \mathbb{R}^{n\times n}_{\geq 0}$ is ergodic and reversible. Let $\chi^{(T)} \in \mathbb{Z}^n_{\geq 0}$ denote the distribution of tokens of any corresponding SRT-router model. Then,*

$$\left|\chi^{(T)}_w - \mu^{(T)}_w\right| \leq 6\pi_w \max_{u\in V} \frac{\delta(u)}{\pi_u} t_{\mathrm{mix}} = \mathrm{O}\left(\frac{\pi_{\max}}{\pi_{\min}} \Delta t_{\mathrm{mix}}\right)$$

holds for any $w \in V$ and for any $T \geq 0$.

Corollary 4.10. *Suppose $P \in \mathbb{R}_{\geq 0}^{n\times n}$ is ergodic, reversible and lazy. Let $\chi^{(T)} \in \mathbb{Z}_{\geq 0}^{n}$ denote the distribution of tokens of any corresponding SRT-router model. Then,*

$$\left|\chi_w^{(T)} - \mu_w^{(T)}\right| \leq 96\pi_w \max_{u\in V} \frac{\delta(u)}{\pi_u}\sqrt{t_{\mathrm{mix}}} = \mathrm{O}\left(\frac{\pi_{\max}}{\pi_{\min}}\Delta\sqrt{t_{\mathrm{mix}}}\right)$$

holds for any $w \in V$ and for any $T \geq 0$.

4.2.2 Billiard-router models

For billiard-router models, we give the following upper bounds, which are almost same as the bound for SRT-router models. We use the following proposition obtained by Proposition 4.7 and Proposition 3.5.

Proposition 4.11. *For any $P \in \mathbb{R}_{\geq 0}^{n\times n}$ and for any its corresponding billiard-router models,*

$$\begin{aligned} \phi_{v,u} &\leq 1 + \delta^+(v)P_{v,u}, \\ \Phi_{v,u} &\leq 1 + \delta^+(v)P_{v,u} \end{aligned}$$

holds for any $v, u \in V$.

Corollary 4.12. *Suppose $P \in \mathbb{R}_{\geq 0}^{n\times n}$ is ergodic and reversible. Let $\chi^{(T)} \in \mathbb{Z}_{\geq 0}^{n}$ denote the distribution of tokens of any corresponding billiard-router model. Then,*

$$\left|\chi_w^{(T)} - \mu_w^{(T)}\right| \leq 6\pi_w \max_{u\in V} \frac{\delta(u)}{\pi_u} t_{\mathrm{mix}} = \mathrm{O}\left(\frac{\pi_{\max}}{\pi_{\min}}\Delta t_{\mathrm{mix}}\right)$$

holds for any $w \in V$ and for any $T \geq 0$.

Proof. We have $\delta(u) = \delta^+(u) = \delta^-(u)$ from the assumption. Using Proposition 4.11,

$$\frac{\sum_{v\in\mathcal{N}^-(u)} \phi_{v,u}}{\pi_u} \leq \frac{\sum_{v\in\mathcal{N}^-(u)}(1+\delta^+(v)P_{v,u})}{\pi_u} = \frac{\sum_{v\in\mathcal{N}(u)}(1+\delta(v)P_{v,u})}{\pi_u} \tag{4.8}$$

holds for any $u \in V$. Then, we have

$$\begin{aligned} \frac{\sum_{v\in\mathcal{N}(u)}(1+\delta(v)P_{v,u})}{\pi_u} &= \frac{\sum_{v\in\mathcal{N}(u)} 1}{\pi_u} + \sum_{v\in\mathcal{N}(u)} \frac{\delta(v)P_{v,u}}{\pi_u} = \frac{\delta(u)}{\pi_u} + \sum_{v\in\mathcal{N}(u)} \frac{\delta(v)P_{u,v}}{\pi_v} \\ &\leq \frac{\delta(u)}{\pi_u} + \max_{v\in\mathcal{N}(u)} \frac{\delta(v)}{\pi_v} \sum_{v\in\mathcal{N}(u)} P_{u,v} \leq \frac{\delta(u)}{\pi_u} + \max_{v\in V} \frac{\delta(v)}{\pi_v}, \end{aligned}$$

thus

$$\max_{u\in V}\frac{\sum_{v\in\mathcal{N}(u)}(1+\delta(v)P_{v,u})}{\pi_u} \;\le\; \max_{u\in V}\left(\frac{\delta(u)}{\pi_u}+\max_{v\in V}\frac{\delta(v)}{\pi_v}\right) \;=\; 2\max_{u\in V}\frac{\delta(u)}{\pi_u} \tag{4.9}$$

holds, and we obtain the claim combining Theorem 4.1, (4.8) and (4.9). □

Corollary 4.13. *Suppose $P \in \mathbb{R}_{\ge 0}^{n\times n}$ is ergodic, reversible and lazy. Let $\chi^{(T)} \in \mathbb{Z}_{\ge 0}^{n}$ denote the distribution of tokens of any corresponding billiard-router model. Then,*

$$\left|\chi_w^{(T)}-\mu_w^{(T)}\right| \le 192\pi_w\max_{u\in V}\frac{\delta(u)}{\pi_u}\sqrt{t_{\mathrm{mix}}} = \mathrm{O}\left(\frac{\pi_{\max}}{\pi_{\min}}\Delta\sqrt{t_{\mathrm{mix}}}\right)$$

holds for any $w\in V$ and for any $T\ge 0$.

Proof. Using Proposition 4.11,

$$\frac{\sum_{v\in\mathcal{N}^-(u)}\Phi_{v,u}}{\pi_u} \;\le\; \frac{\sum_{v\in\mathcal{N}^-(u)}(1+\delta^+(v)P_{v,u})}{\pi_u} \;=\; \frac{\sum_{v\in\mathcal{N}(u)}(1+\delta(v)P_{v,u})}{\pi_u} \tag{4.10}$$

holds for any $u\in V$. Thus we obtain the claim combining Theorem 4.2, (4.10) and (4.9). □

4.2.3 Weighted rotor-router models

Combining Proposition 4.7 and Proposition 3.2, we give the following.

Proposition 4.14. *Suppose $P \in \mathbb{Q}_{\ge 0}^{n\times n}$. Then, for any its corresponding weighted-rotor-router models,*

$$\begin{aligned}\phi_{v,u} &\le \bar{\delta}(v)P_{v,u},\\ \Phi_{v,u} &\le \bar{\delta}(v)P_{v,u}\end{aligned}$$

holds for any $v,u\in V$.

We obtain the following upper bounds for weighted-rotor-router models. Note that we assume P is rational.

Corollary 4.15. *Suppose $P \in \mathbb{Q}_{\ge 0}^{n\times n}$ is ergodic and reversible. Let $\chi^{(T)} \in \mathbb{Z}_{\ge 0}^{n}$ denote the distribution of tokens of any corresponding weighted rotor-router model. Then,*

$$\left|\chi_w^{(T)}-\mu_w^{(T)}\right| \le 3\pi_w\max_{u\in V}\frac{\bar{\delta}(u)}{\pi_u}t_{\mathrm{mix}} = \mathrm{O}\left(\frac{\pi_{\max}}{\pi_{\min}}\bar{\Delta}t_{\mathrm{mix}}\right)$$

holds for any $w\in V$ and for any $T\ge 0$.

Proof. We have $\delta(u) = \delta^-(u)$ from the assumption. Using Proposition 4.11,

$$\frac{\sum_{v \in \mathcal{N}^-(u)} \phi_{v,u}}{\pi_u} \leq \frac{\sum_{v \in \mathcal{N}(u)} \bar{\delta}(v) P_{v,u}}{\pi_u} \tag{4.11}$$

holds for any $u \in V$. Then,

$$\begin{aligned} \sum_{v \in \mathcal{N}(u)} \frac{\bar{\delta}(v) P_{v,u}}{\pi_u} &= \sum_{v \in \mathcal{N}(u)} \frac{\bar{\delta}(v) P_{u,v}}{\pi_v} \\ &\leq \max_{v \in \mathcal{N}(u)} \frac{\bar{\delta}(v)}{\pi_v} \sum_{v \in \mathcal{N}(u)} P_{u,v} \leq \max_{v \in V} \frac{\bar{\delta}(v)}{\pi_v} \end{aligned} \tag{4.12}$$

holds, and we obtain the claim combining to Theorem 4.1. □

Corollary 4.16. *Suppose $P \in \mathbb{Q}_{\geq 0}^{n \times n}$ is ergodic, reversible and lazy. Let $\chi^{(T)} \in \mathbb{Z}_{\geq 0}^n$ denote the distribution of tokens of any corresponding weighted rotor-router-router model. Then,*

$$\left|\chi_w^{(T)} - \mu_w^{(T)}\right| \leq 96\pi_w \max_{u \in V} \frac{\bar{\delta}(u)}{\pi_u} \sqrt{t_{\mathrm{mix}}} = \mathrm{O}\left(\frac{\pi_{\max}}{\pi_{\min}} \bar{\Delta} \sqrt{t_{\mathrm{mix}}}\right)$$

holds for any $w \in V$ and for any $T \geq 0$.

Proof. Using Proposition 4.11,

$$\frac{\sum_{v \in \mathcal{N}^-(u)} \Phi_{v,u}}{\pi_u} \leq \frac{\sum_{v \in \mathcal{N}(u)} \bar{\delta}(v) P_{v,u}}{\pi_u}. \tag{4.13}$$

holds for any $u \in V$. Thus, we obtain the claim combining Theorem 4.2, (4.13) and (4.13). □

4.2.4 Quasi random-router models

The following proposition is obtained by Proposition 4.7 and Proposition 3.4.

Proposition 4.17. *For any $P \in \mathbb{R}_{\geq 0}^{n \times n}$ and for any its corresponding quasi-random-router models,*

$$\phi_{v,u} \leq 2 \lg(k+1)$$

holds for any $v, u \in V$ and for any $z \geq 0$.

It is easy to check the following upper bound by Proposition 4.17.

Corollary 4.18. *Suppose $P \in \mathbb{R}_{\geq 0}^{n \times n}$ is ergodic and reversible. Let $\chi^{(T)} \in \mathbb{Z}_{\geq 0}^n$ denote the distribution of tokens of any corresponding quasi random-router model. Then,*

$$\left|\chi_w^{(T)} - \mu_w^{(T)}\right| \leq 6\pi_w \max_{u \in V} \frac{\delta(u)}{\pi_u} t_{\mathrm{mix}} \lg(k+1) = \mathrm{O}\left(\frac{\pi_{\max}}{\pi_{\min}} \Delta t_{\mathrm{mix}} \log k\right)$$

holds for any $w \in V$ and for any $T \geq 0$.

4.2.5 Oblivious models

It is easy to check the following upper bound by Proposition 3.7.

Corollary 4.19. *Suppose $P \in \mathbb{R}_{\geq 0}^{n \times n}$ is ergodic and reversible. Let $\chi^{(T)} \in \mathbb{Z}_{\geq 0}^n$ denote the distribution of tokens of any corresponding oblivious model. Then,*

$$\left|\chi_w^{(T)} - \mu_w^{(T)}\right| \leq 3\pi_w \max_{u \in V} \frac{\delta(u)}{\pi_u} t_{\mathrm{mix}} = \mathrm{O}\left(\frac{\pi_{\max}}{\pi_{\min}} \Delta t_{\mathrm{mix}}\right)$$

holds for any $w \in V$ and for any $T \geq 0$.

4.3 Application to rapidly mixing chains

This section shows some examples of bounds suggested by Corollaries 4.9 and 4.12 for some celebrated Markov chains known to be rapidly mixing, namely ones for 0-1 knapsack solutions (Section 4.3.1), linear extensions (Section 4.3.2), and matchings (Section 4.3.3).

4.3.1 0-1 knapsack solutions

Given $\boldsymbol{a} \in \mathbb{Z}_{>0}^n$ and $b \in \mathbb{Z}_{>0}$, the set of 0-1 knapsack solutions is defined by $\Omega_{\mathrm{Kna}} = \{\boldsymbol{x} \in \{0,1\}^n \mid \sum_{i=1}^n a_i x_i \leq b\}$. We define a transition matrix $P_{\mathrm{Kna}} \in \mathbb{R}^{|\Omega_{\mathrm{Kna}}| \times |\Omega_{\mathrm{Kna}}|}$ by

$$P_{\mathrm{Kna}}(\boldsymbol{x}, \boldsymbol{y}) = \begin{cases} 1/2n & (\text{if } \boldsymbol{y} \in \mathcal{N}_{\mathrm{Kna}}(\boldsymbol{x})) \\ 1 - |\mathcal{N}_{\mathrm{Kna}}(\boldsymbol{x})|/2n & (\text{if } \boldsymbol{y} = \boldsymbol{x}) \\ 0 & (\text{otherwise}) \end{cases} \tag{4.14}$$

for $\boldsymbol{x}, \boldsymbol{y} \in \Omega_{\mathrm{Kna}}$, where $\mathcal{N}_{\mathrm{Kna}}(\boldsymbol{x}) = \{\boldsymbol{y} \in \Omega_{\mathrm{Kna}} \mid \|\boldsymbol{x} - \boldsymbol{y}\|_1 = 1\}$. Note that the stationary distribution of P_{Kna} is uniform distribution since P_{Kna} is symmetric. The following theorem is due to Morris and Sinclair [68].

Theorem 4.20. *[68] The mixing time* $\tau(\gamma)$ *of* $P_{\rm Kna}$ *is* $\mathrm{O}(n^{\frac{9}{2}+\alpha}\log\gamma^{-1})$ *for any* $\alpha > 0$ *and for any* $\gamma > 0$.

Thus, Corollaries 4.9 and 4.12 imply the following.

Theorem 4.21. *For the SRT-router model (as well as the billiard-router model) corresponding to* $P_{\rm Kna}$*, the discrepancy between* $\chi^{(T)}$ *and* $\mu^{(T)}$ *satisfies*

$$\left|\chi_w^{(T)} - \mu_w^{(T)}\right| = \mathrm{O}(n^{\frac{11}{2}+\alpha})$$

for any $w \in V$ *and* $T \geq 0$*, where* $\alpha > 0$ *is an arbitrary constant.*

Let $\widetilde{\mu}^{(t)} = \mu^{(t)}/M$, for simplicity, then clearly $\widetilde{\mu}^{(\infty)} = \pi$ holds, since P is ergodic (see Section 2.2). By the definition of the mixing time, $\mathcal{D}_{\rm tv}(\widetilde{\mu}^{(\tau(\varepsilon))}, \pi) \leq \varepsilon$ holds where $\tau(\varepsilon)$ denotes the mixing time of P, meaning that $\widetilde{\mu}$ approximates the target distribution π well. Thus, we hope for a deterministic random walk that the "distribution" $\widetilde{\chi}^{(T)} \stackrel{\rm def}{=} \chi^{(T)}/M$ approximates the target distribution π well. For convenience, a *pointwise distance* $\mathcal{D}_{\rm pw}(\xi,\zeta)$ between $\xi \in \mathbb{R}^N_{\geq 0}$ and $\zeta \in \mathbb{R}^N_{\geq 0}$ satisfying $\|\xi\|_1 = \|\zeta\|_1 = 1$ is defined by

$$\mathcal{D}_{\rm pw}(\xi,\zeta) \stackrel{\rm def}{=} \max_{v\in V} |\xi_v - \zeta_v| = \|\xi - \zeta\|_\infty. \tag{4.15}$$

Corollary 4.22. *For an arbitrary* ε $(0 < \varepsilon < 1)$*, let the total number of tokens* $M := c_1\, n^{\frac{11}{2}+\alpha}\varepsilon^{-1}$ *with some appropriate constants* c_1 *and* α*. Then, the pointwise distance between* $\widetilde{\chi}^{(T)} \stackrel{\rm def}{=} \chi^{(T)}/M$ *and* π *satisfies*

$$\mathcal{D}_{\rm pw}\left(\widetilde{\chi}^{(T)}, \pi\right) \leq \varepsilon \tag{4.16}$$

for any $T \geq c_2\, n^{\frac{9}{2}+\alpha}\log\varepsilon^{-1}$ *with an appropriate constant* c_2*, where* π *is the uniform distribution over* $\Omega_{\rm Kna}$*.*

4.3.2 Linear extensions of a poset

Let $S = \{1, 2, \ldots, n\}$, and $Q = (S, \preceq)$ be a partial order. A linear extension of Q is a total order $X = (S, \sqsubseteq)$ which respects Q, i.e., for all $i, j \in S$, $i \preceq j$ implies $i \sqsubseteq j$. Let $\Omega_{\rm Lin}$ denote the set of all linear extensions of Q. We define a relationship $X \sim_p X'$

($p \in \{1, \dots, n\}$) for a pair of linear extensions X and $X' \in \Omega_{\mathrm{Lin}}$ satisfying that $x_p = x'_{p+1}$, $x_{p+1} = x'_p$, and $x_i = x'_i$ for all $i \neq p, p+1$, i.e.,

$$\begin{aligned} X &= (x_1, x_2, \dots, x_{p-1}, x_p, x_{p+1}, x_{p+2}, \dots, x_n) \\ X' &= (x_1, x_2, \dots, x_{p-1}, x_{p+1}, x_p, x_{p+2}, \dots, x_n) \end{aligned}$$

holds. Then, we define a transition matrix $P_{\mathrm{Lin}} \in \mathbb{R}^{|\Omega_{\mathrm{Lin}}| \times |\Omega_{\mathrm{Lin}}|}$ by

$$P_{\mathrm{Lin}}(X, X') = \begin{cases} F(p)/2 & (\text{if } X' \sim_p X) \\ 1 - \sum_{I \in \mathcal{N}_{\mathrm{Lin}}(X)} P_{\mathrm{Lin}}(X, I) & (\text{if } X' = X) \\ 0 & (\text{otherwise}) \end{cases} \tag{4.17}$$

for $X, X' \in \Omega_{\mathrm{Lin}}$, where $\mathcal{N}_{\mathrm{Lin}}(X) = \{Y \in \Omega_{\mathrm{Lin}} \mid X \sim_p Y (p \in \{1, \dots, n-1\})\}$ and $F(p) = \frac{p(n-p)}{\frac{1}{6}(n^3-n)}$. Note that P_{Lin} is ergodic and reversible, and its stationary distribution is uniform on Ω_{Lin} [16]. The following theorem is due to Bubley and Dyer [16].

Theorem 4.23. *[16] For P_{Lin},*

$$\tau(\gamma) \leq \left\lceil \frac{1}{6}(n^3 - n) \ln \frac{n^2}{4\gamma} \right\rceil$$

holds for any $\gamma > 0$.

It is not difficult to see that the maximum degree $\Delta = n$ (including a self-loop) of the transition diagram P_{LIN}. Thus, Corollaries 4.9 and 4.12 imply the following.

Theorem 4.24. *For the SRT-router model (as well as the billiard-router model) corresponding to P_{LIN}, the discrepancy between $\chi^{(T)}$ and $\mu^{(T)}$ satisfies*

$$\left|\chi_w^{(T)} - \mu_w^{(T)}\right| \leq n \left\lceil \frac{1}{3}(n^3 - n) \ln n \right\rceil = \mathrm{O}(n^4 \log n)$$

for any $w \in V$ and $T \geq 0$.

4.3.3 Matchings in a graph

Counting all matchings in a graph, related to the *Hosoya index* [46], is known to be #P-complete [85]. Jerrum and Sinclair [51] gave a rapidly mixing chain. This section is concerned with a Markov chain for sampling from all matchings in a graph[1].

[1] Remark that counting all *perfect* matchings in a bipartite graph, related to the *permanent*, is also well-known #P-complete problem, and Jerrum et al. [52] gave a celebrated FPRAS based on an MCMC method using annealing. To apply our bound to a Markov chain for sampling perfect matchings, we need some assumptions on the input graph (see e.g., [80, 51, 52]).

Let $H = (U, F)$ be an undirected graph, where $|U| = n$ and $|F| = m$. A matching in H is a subset $\mathcal{M} \subseteq F$ such that no edges in $\mathcal{M}$ share an endpoint. Let $\Omega_{\rm Mat}$ denote the set of all possible matchings of H. Let $N_C(\mathcal{M}) = \{e = \{u, v\} \mid e \notin \mathcal{M}, \text{both } u \text{ and } v \text{ are matched in } \mathcal{M}\}$ and let $\mathcal{N}_{\rm Mat}(\mathcal{M}) = \{e \mid e \notin N_C(\mathcal{M})\}$. Then, for $e = \{u, v\} \in \mathcal{N}_{\rm Mat}(\mathcal{M})$, we define $\mathcal{M}(e)$ by

$$\mathcal{M}(e) = \begin{cases} \mathcal{M} - e & (\text{if } e \in \mathcal{M}) \\ \mathcal{M} + e & (\text{if } u \text{ and } v \text{ are unmatched in } \mathcal{M}) \\ \mathcal{M} + e - e' & (\text{if exactly one of } u \text{ and } v \text{ is matched in } M, \\ & \quad \text{and } e' \text{ is the matching edge}). \end{cases}$$

The we define the transition matrix $P_{\rm Mat} \in \mathbb{R}^{|\Omega_{\rm Mat}| \times |\Omega_{\rm Mat}|}$ by

$$P_{\rm Mat}(\mathcal{M}, \mathcal{M}') = \begin{cases} 1/2m & (\text{if } \mathcal{M}' = \mathcal{M}(e)) \\ 1 - |\mathcal{N}_{\rm Mat}(\mathcal{M})|/2m & (\text{if } \mathcal{M}' = \mathcal{M}) \\ 0 & (\text{otherwise}) \end{cases}$$

for any $\mathcal{M}, \mathcal{M}' \in \Omega_{\rm Mat}$. Note that $P_{\rm Mat}$ is ergodic and reversible, and its stationary distribution is uniform on $\Omega_{\rm Mat}$ [51]. The following theorem is due to Jerrum and Sinclar [51].

Theorem 4.25. *[51] For $P_{\rm Mat}$,*

$$\tau(\gamma) \leq 4mn(n \ln n + \ln \gamma^{-1})$$

holds for any $\gamma > 0$.

It is not difficult to see that the maximum degree $\Delta = m + 1$ (including a self-loop) of the transition diagram $P_{\rm LIN}$. Thus, Corollaries 4.9 and 4.12 imply the following.

Theorem 4.26. *For the SRT-router model (as well as the billiard-router model) corresponding to $P_{\rm LIN}$, the discrepancy between $\chi^{(T)}$ and $\mu^{(T)}$ satisfies*

$$\left|\chi_w^{(T)} - \mu_w^{(T)}\right| \leq 4(m+1)mn(n \ln n + \ln 4) = \mathrm{O}(m^2 n^2 \log n)$$

for any $w \in V$ and $T \geq 0$.

4.4 Lower bounds of the vertex-wise discrepancy

This section shows some lower bounds on some specific structures and specific models.

Proposition 4.27 ([57](Theorem 4.1))**.** *Suppose $P_{u,v} = 1/n$ for any $u, v \in V$. Then there exists an example of its corresponding rotor-router model such that*

$$\left|\chi_w^{(T)} - \mu_w^{(T)}\right| = \Omega(n)$$

holds for any $w \in V$ and for any odd $T \geq 0$.

Note that for such P, it is easy to check that π is uniform distribution, $t_{\mathrm{mix}} = \mathrm{O}(1)$ and $\bar{\delta}(u) = n$ for any u. Thus

$$\left|\chi_w^{(T)} - \mu_w^{(T)}\right| \leq 3n \cdot \mathrm{O}(1) \;=\; \mathrm{O}(n)$$

holds by Corollary 4.15. This P is an example of $|\chi_w^{(T)} - \mu_w^{(T)}| = \Theta(n)$.

Proposition 4.28 ([12])**.** *Suppose that G is d-regular graph and P is the transition matrix of the simple random walk on G. Then there exists an example of its corresponding oblivious model such that*

$$|\chi_w^{(T)} - \mu_w^{(T)}| = \Omega(dD)$$

holds for any $T \geq t_{\mathrm{mix}}$ and for some $w \in V$.

Note that for such P, it is easy to check that π is uniform distribution, $\delta(u) = d$ for any u. Thus

$$\left|\chi_w^{(T)} - \mu_w^{(T)}\right| \leq 3dt_{\mathrm{mix}}$$

holds by Corollary 4.19. This P is an example of $\left\|\chi^{(T)} - \mu^{(T)}\right\|_\infty = \Theta(dt_{\mathrm{mix}})$ when $t_{\mathrm{mix}} = \mathrm{O}(D)$.

Proposition 4.29. *There exists an example of $P \in \mathbb{R}_{\geq 0}^{n \times n}$ and its corresponding oblivious model such that*

$$\left|\chi_w^{(T)} - \mu_w^{(T)}\right| = \Omega(t_{\mathrm{mix}})$$

holds for any $w \in V$ and for any $T \geq t_{\mathrm{mix}}$.

Proof. Suppose $\gamma \geq 2$ is a natural number. Let $V = \{x, y\}$ and $P_{x,y} = P_{y,x} = 1/\gamma$, $P_{x,x} = P_{y,y} = 1 - 1/\gamma$. Then, $\pi_x = \pi_y = 1/2$ and it is not difficult to see that $t_{\mathrm{mix}} = \mathrm{O}(\gamma)$. Now, let $\chi^{(0)} = (\gamma - 1, 0)$. Then,

$$\begin{aligned} z_{x,x}^{(0)} &= \left\lfloor (\gamma - 1) \cdot \frac{\gamma - 1}{\gamma} \right\rfloor + 1 = \gamma - 1 \\ z_{x,y}^{(0)} &= \left\lfloor (\gamma - 1) \cdot \frac{1}{\gamma} \right\rfloor = 0 \end{aligned}$$

holds from the definition of oblivious model, thus $\chi^{(t)} = (\gamma - 1, 0)$ holds for any $t \geq 0$. Hence for any $T \geq t_{\mathrm{mix}}$,

$$\chi_x^{(t)} - \mu_x^{(t)} \;\geq\; \gamma - 1 - \frac{\gamma - 1}{2} - \frac{1}{4} \;=\; \Omega(\gamma) \;=\; \Omega(t_{\mathrm{mix}})$$

holds and we obtain the claim. □

Note that for this P,

$$\left|\chi_w^{(T)} - \mu_w^{(T)}\right| \leq 6t_{\mathrm{mix}} \;=\; \mathrm{O}(t_{\mathrm{mix}})$$

holds by Corollary 4.19, meaning that this P is an example of $|\chi_w^{(T)} - \mu_w^{(T)}| = \Theta(t_{\mathrm{mix}})$ for oblivious model.

Chapter 5

Total Variation Discrepancy

This chapter investigates the total variation discrepancy (L_1 discrepancy) between the token configuration $\chi^{(t)}$ of a deterministic random walk (see (3.3)) and the expected token configuration $\mu^{(t)}$ of a random walk (see (2.15)). First, we state general upper bounds in Section 5.1. Section 5.2 gives upper bounds for some specific models. Section 5.3 discusses lower bounds.

5.1 General upper bounds of the total variation discrepancy

First, we show the following bound, which requires only ergodicity of P, in contrast to Theorems 4.1 and 4.2 assuming reversibility.

Theorem 5.1. *If $P \in \mathbb{R}_{\geq 0}^{n \times n}$ is ergodic, then*

$$\left|\chi_S^{(T)} - \mu_S^{(T)}\right| \leq \sum_{u \in V} \sum_{v \in \mathcal{N}^-(u)} \phi_{v,u} t_{\mathrm{mix}} = \mathrm{O}(\phi m t_{\mathrm{mix}})$$

holds for any $S \subseteq V$ and for any $T \geq 0$.

Theorem 5.1 says that the total variation discrepancy is upper bounded by the edge size m, which is independent of the weights of its stationary distribution. For lazy chains and for deterministic random walks such that Φ is bounded, we obtain the following better bound.

Theorem 5.2. *If $P \in \mathbb{R}_{\geq 0}^{n \times n}$ is ergodic and lazy, then*

$$\left|\chi_S^{(T)} - \mu_S^{(T)}\right| \leq 48 \sum_{u \in V} \sum_{v \in \mathcal{N}^-(u)} \Phi_{u,v} \sqrt{t_{\mathrm{mix}}} = \mathrm{O}(\Phi m \sqrt{t_{\mathrm{mix}}})$$

for any $S \subseteq V$ and for any $T \geq 0$.

5.1.1 Proof of Theorem 5.1

In order to get the idea of the proof of Theorem 5.1, we prove the following weak version, whose proof is very easy.

Theorem 5.3. *Suppose $P \in \mathbb{R}_{\geq 0}^{n\times n}$ is ergodic. Then,*

$$\left|\chi_S^{(T)} - \mu_S^{(T)}\right| \leq \frac{3}{2}\sum_{u\in V}\sum_{v\in\mathcal{N}^-(u)} \phi_{v,u} t_{\mathrm{mix}} = \mathrm{O}(\phi m t_{\mathrm{mix}})$$

for any $S \subseteq V$ and for any $T \geq 0$.

Proof. Lemma 4.3 says that

$$\begin{aligned}\chi_S^{(T)} - \mu_S^{(T)} &= \sum_{w\in S}\sum_{u\in V}\sum_{v\in\mathcal{N}^-(u)}\sum_{t=0}^{T-1}\phi_{v,u}^{(T-t-1)}(P_{u,w}^t - \pi_w)\\ &= \sum_{u\in V}\sum_{v\in\mathcal{N}^-(u)}\sum_{t=0}^{T-1}\phi_{v,u}^{(T-t-1)}(P_{u,S}^t - \pi_S).\end{aligned}$$

Thus

$$\left|\chi_S^{(T)} - \mu_S^{(T)}\right| \leq \sum_{u\in V}\sum_{v\in\mathcal{N}^-(u)}\sum_{t=0}^{T-1}\left|\phi_{v,u}^{(T-t-1)}\right|\left|P_{u,S}^t - \pi_S\right| \leq \sum_{u\in V}\sum_{v\in\mathcal{N}^-(u)}\phi_{v,u}\sum_{t=0}^{T-1} d(t)$$

holds, and we obtain the claim using Lemma 4.4. □

Proof of Theorem 5.1. From (3.5), we have

$$\begin{aligned}\left|\chi_S^{(T)} - \mu_S^{(T)}\right| &= \left|\sum_{w\in S}\sum_{t=0}^{T-1}\left((\chi^{(T-t)} - \chi^{(T-t-1)}P)P^t\right)_w\right| = \left|\sum_{t=0}^{T-1}\left((\chi^{(T-t)} - \chi^{(T-t-1)}P)P^t\right)_S\right|\\ &\leq \sum_{t=0}^{T-1}\left|\left((\chi^{(T-t)} - \chi^{(T-t-1)}P)P^t\right)_S\right| \leq \sum_{t=0}^{T-1}\frac{1}{2}\|(\chi^{(T-t)} - \chi^{(T-t-1)}P)P^t\|_1 \quad (5.1)\end{aligned}$$

since $\sum_{v\in V}\left((\chi^{(T-t)} - \chi^{(T-t-1)}P)P^t\right)_v = 0$ holds. We can apply Lemma 2.7 to (5.1) since (4.1) holds. Then, we have

$$\begin{aligned}\left|\chi_S^{(T)} - \mu_S^{(T)}\right| &\leq \frac{1}{2}\sum_{t=0}^{T-1}\|\chi^{(T-t)} - \chi^{(T-t-1)}P\|_1\bar{d}(t)\\ &= \frac{1}{2}\sum_{t=0}^{T-1}\sum_{u\in V}\left|(\chi^{(T-t)} - \chi^{(T-t-1)}P)_u\right|\bar{d}(t)\\ &= \frac{1}{2}\sum_{t=0}^{T-1}\sum_{u\in V}\left|\sum_{v\in\mathcal{N}^-(u)}\phi_{v,u}^{(T-t-1)}\right|\bar{d}(t) \leq \frac{1}{2}\sum_{t=0}^{T-1}\sum_{u\in V}\sum_{v\in\mathcal{N}^-(u)}\phi_{v,u}\bar{d}(t) \quad (5.2)\end{aligned}$$

Note that we use (3.7) for the second equality. Thus, we obtain the claim since

$$\sum_{t=0}^{T-1} \bar{d}(t) \leq \sum_{\ell=0}^{\infty} \sum_{k=0}^{t_{\mathrm{mix}}-1} \bar{d}(\ell t_{\mathrm{mix}} + k) \leq \sum_{\ell=0}^{\infty} \sum_{k=0}^{t_{\mathrm{mix}}-1} \frac{1}{2^{\ell}} = t_{\mathrm{mix}} \sum_{\ell=0}^{\infty} \frac{1}{2^{\ell}} \leq 2t_{\mathrm{mix}} \tag{5.3}$$

holds by Proposition 2.4. □

5.1.2 Proof of Theorem 5.2

Proof. Lemma 4.5 implies that

$$\begin{aligned} \chi_S^{(T)} - \mu_S^{(T)} &= \sum_{w \in S} \sum_{u \in V} \sum_{v \in \mathcal{N}^-(u)} \sum_{t=0}^{T-1} \Phi_{v,u}^{(T-t)} \left(P_{u,w}^t - P_{u,w}^{t-1} \right) \\ &= \sum_{u \in V} \sum_{v \in \mathcal{N}^-(u)} \sum_{t=0}^{T-1} \Phi_{v,u}^{(T-t)} \left(P_{u,S}^t - P_{u,S}^{t-1} \right). \end{aligned}$$

Thus, by the definition (3.13),

$$\begin{aligned} \left| \chi_S^{(T)} - \mu_S^{(T)} \right| &\leq \sum_{u \in V} \sum_{v \in \mathcal{N}^-(u)} \sum_{t=0}^{T-1} \left| \Phi_{v,u}^{(T-t)} \right| \left| P_{u,S}^t - P_{u,S}^{t-1} \right| \\ &\leq \sum_{u \in V} \sum_{v \in \mathcal{N}^-(u)} \sum_{t=0}^{T-1} \Phi_{v,u} \left| P_{u,S}^t - P_{u,S}^{t-1} \right| \\ &\leq \sum_{u \in V} \sum_{v \in \mathcal{N}^-(u)} \Phi_{v,u} \left(|P_{u,S}^0| + \sum_{t=1}^{T-1} 2\tilde{d}(t-1) \right) \\ &\leq \sum_{u \in V} \sum_{v \in \mathcal{N}^-(u)} \Phi_{v,u} \left(1 + \sum_{t=0}^{T-2} 2\tilde{d}(t) \right) \end{aligned}$$

holds, and we obtain the claim using Lemma 4.6. □

5.2 Upper bounds for specific models

This section states upper bounds for specific deterministic random walks by Theorem 5.1 and Theorem 5.2.

5.2.1 SRT-router models

It is easy to check the following bounds combining Theorem 5.1, Theorem 5.2 and Proposition 4.8.

Corollary 5.4. *Suppose $P \in \mathbb{R}_{\geq 0}^{n \times n}$ is ergodic. Let $\chi^{(T)} \in \mathbb{Z}_{\geq 0}^{n}$ denote the distribution of tokens of any corresponding SRT-router model. Then,*

$$\left|\chi_S^{(T)} - \mu_S^{(T)}\right| \leq 2mt_{\mathrm{mix}}$$

holds for any $S \subseteq V$ and for any $T \geq 0$.

Corollary 5.5. *Suppose $P \in \mathbb{R}_{\geq 0}^{n \times n}$ is ergodic, reversible and lazy. Let $\chi^{(T)} \in \mathbb{Z}_{\geq 0}^{n}$ denote the distribution of tokens of any corresponding SRT-router model. Then,*

$$\left|\chi_S^{(T)} - \mu_S^{(T)}\right| \leq 48m\sqrt{t_{\mathrm{mix}}}$$

for any $S \subseteq V$ and for any $T \geq 0$.

5.2.2 Billiard-router models

We give almost the same upper bounds as one for the SRT-router model for billiard-router models.

Corollary 5.6. *Suppose $P \in \mathbb{R}_{\geq 0}^{n \times n}$ is ergodic and reversible. Let $\chi^{(T)} \in \mathbb{Z}_{\geq 0}^{n}$ denote the distribution of tokens of any corresponding billiard-router model. Then,*

$$\left|\chi_S^{(T)} - \mu_S^{(T)}\right| \leq 2mt_{\mathrm{mix}}$$

holds for any $S \subseteq V$ and for any $T \geq 0$.

Proof. We have

$$\sum_{u \in V} \sum_{v \in \mathcal{N}^-(u)} \phi_{v,u} \leq \sum_{u \in V} \sum_{v \in \mathcal{N}^-(u)} (1 + \delta^+(v) P_{v,u})$$

from Proposition 4.11. Thus

$$\begin{aligned} \sum_{u \in V} \sum_{v \in \mathcal{N}^-(u)} (1 + \delta^+(v) P_{v,u}) &= \sum_{u \in V} \sum_{v \in \mathcal{N}^-(u)} 1 + \sum_{u \in V} \sum_{v \in \mathcal{N}^-(u)} \delta^+(v) P_{v,u} \\ &= m + \sum_{u \in V} \sum_{v \in V} \delta^+(v) P_{v,u} = m + \sum_{v \in V} \delta^+(v) = 2m \end{aligned} \tag{5.4}$$

holds, and we obtain the claim. □

Corollary 5.7. *Suppose $P \in \mathbb{R}_{\geq 0}^{n \times n}$ is ergodic, reversible and lazy. Let $\chi^{(T)} \in \mathbb{Z}_{\geq 0}^{n}$ denote the distribution of tokens of any corresponding billiard-router model. Then,*

$$\left|\chi_S^{(T)} - \mu_S^{(T)}\right| \leq 96m\sqrt{t_{\mathrm{mix}}}$$

for any $S \subseteq V$ and for any $T \geq 0$.

Proof. We have

$$\sum_{u \in V} \sum_{v \in \mathcal{N}^-(u)} \Phi_{v,u} \ \leq\ \sum_{u \in V} \sum_{v \in \mathcal{N}^-(u)} (1 + \delta^+(v) P_{v,u}) \tag{5.5}$$

Thus we obtain the claim combining Theorem 5.2, (5.5) and (5.4). □

5.2.3 Weighted rotor-router models

Corollary 5.8. *Suppose $P \in \mathbb{R}_{\geq 0}^{n \times n}$ is ergodic and reversible. Let $\chi^{(T)} \in \mathbb{Z}_{\geq 0}^{n}$ denote the distribution of tokens of any corresponding weighted rotor-router model. Then,*

$$\left|\chi_S^{(T)} - \mu_S^{(T)}\right| \leq \bar{m} t_{\mathrm{mix}}$$

holds for any $S \subseteq V$ and for any $T \geq 0$.

Proof. We have

$$\sum_{u \in V} \sum_{v \in \mathcal{N}^-(u)} \phi_{v,u} \ \leq\ \sum_{u \in V} \sum_{v \in \mathcal{N}^-(u)} \bar{\delta}(v) P_{v,u}$$

from Proposition 4.14. Thus

$$\sum_{u \in V} \sum_{v \in \mathcal{N}^-(u)} \bar{\delta}(v) P_{v,u} \ \leq\ \sum_{u \in V} \sum_{v \in V} \bar{\delta}(v) P_{v,u} \ =\ \sum_{v \in V} \bar{\delta}(v) \ =\ \bar{m} \tag{5.6}$$

holds, and we obtain the claim. □

Corollary 5.9. *Suppose $P \in \mathbb{R}_{\geq 0}^{n \times n}$ is ergodic, reversible and lazy. Let $\chi^{(T)} \in \mathbb{Z}_{\geq 0}^{n}$ denote the distribution of tokens of any corresponding weighted rotor-router-router model. Then,*

$$\left|\chi_S^{(T)} - \mu_S^{(T)}\right| \leq 48\bar{m}\sqrt{t_{\mathrm{mix}}}$$

for any $S \subseteq V$ and for any $T \geq 0$.

Proof. We have

$$\sum_{u \in V} \sum_{v \in \mathcal{N}^-(u)} \Phi_{v,u} \ \leq\ \sum_{u \in V} \sum_{v \in \mathcal{N}^-(u)} \bar{\delta}(v) P_{v,u} \tag{5.7}$$

Thus we obtain the claim combining Theorem 5.2, (5.7) and (5.6). □

5.2.4 Quasi random-router models

It is easy to check the following upper bound by Proposition 4.17.

Corollary 5.10. *Suppose $P \in \mathbb{R}_{\geq 0}^{n\times n}$ is ergodic and reversible. Let $\chi^{(T)} \in \mathbb{Z}_{\geq 0}^{n}$ denote the distribution of tokens of any corresponding quasi random-router model. Then,*

$$\left|\chi_S^{(T)} - \mu_S^{(T)}\right| \leq 2mt_{\mathrm{mix}} \lg(k+1)$$

holds for any $S \subseteq V$ and for any $T \geq 0$.

5.2.5 Oblivious models

It is easy to check the following upper bound by Proposition 3.7.

Corollary 5.11. *Suppose $P \in \mathbb{R}_{\geq 0}^{n\times n}$ is ergodic and reversible. Let $\chi^{(T)} \in \mathbb{Z}_{\geq 0}^{n}$ denote the distribution of tokens of any corresponding oblivious model. Then,*

$$\left|\chi_S^{(T)} - \mu_S^{(T)}\right| \leq mt_{\mathrm{mix}}$$

holds for any $S \subseteq V$ and for any $T \geq 0$.

5.3 Lower bounds of the total variation discrepancy

This section gives lower bounds of the total variation discrepancy.

5.3.1 General lower bound

First, we observe the following proposition, which is caused by the integral gap between $\chi^{(T)} \in \mathbb{Z}^n$ and $\mu^{(T)} \in \mathbb{R}^n$.

Proposition 5.12. *Suppose that P is ergodic and its stationary distribution is uniform. Then, for any $\chi^{(T)} \in \mathbb{Z}_{\geq 0}^{n}$ with an appropriate number of tokens k,*

$$\max_{S\subseteq V}\left|\chi_S^{(T)} - \mu_S^{(T)}\right| = \Omega(n)$$

holds for any time T after mixing.

Proof of Proposition 5.12. Let $M = (k-(1/2))n$ be the number of tokens for an arbitrary positive integer k. Note that $\widetilde{\mu}_v^{(t)} = \mu_v^{(t)}/M$ converges to $1/n$ for any $v \in V$ since the stationary distribution is uniform. Precisely, for any $S \subseteq V$ and $T \geq \tau(1/(8k))$,

$$\frac{|S|}{n} - \frac{1}{8k} \leq \sum_{v \in S} \widetilde{\mu}_v^{(T)} \leq \frac{|S|}{n} + \frac{1}{8k} \tag{5.8}$$

holds by the definition (2.13) of the mixing time $\tau(\varepsilon)$.

Let T be an arbitrary time, and let $S = \{v \in V \mid \chi_v^{(T)} \geq k\}$. First, we consider the case that $|S| \geq n/2$. Then, we see that $\sum_{v \in S} \chi_v^{(T)} \geq k|S|$ holds. At the same time

$$\sum_{v \in S} \mu_v^{(T)} = \sum_{v \in S} M \widetilde{\mu}_v^{(T)} \leq \left(k - \frac{1}{2}\right) n \cdot \left(\frac{|S|}{n} + \frac{1}{8k}\right) \leq \left(k - \frac{1}{2}\right)|S| + \frac{n}{8}$$

holds. Thus

$$\sum_{v \in S} \left(\chi_v^{(T)} - \mu_v^{(T)}\right) \geq k|S| - \left(\left(k - \frac{1}{2}\right)|S| + \frac{n}{8}\right) = \frac{1}{2}|S| - \frac{n}{8} \geq \frac{n}{8}$$

where the last inequality follows $|S| \geq n/2$. We obtain the claim in the case. Next, we consider the other case, meaning that $|S| < n/2$. Then we see that $\sum_{v \in \overline{S}} \chi_v^{(T)} \leq (k-1)|\overline{S}|$ since $\chi_v^{(T)} < k$ for any $v \in \overline{S}$. At that time,

$$\sum_{v \in \overline{S}} \mu_v^{(T)} = \sum_{v \in \overline{S}} M \widetilde{\mu}_v^{(T)} \geq \left(k - \frac{1}{2}\right) n \cdot \left(\frac{|\overline{S}|}{n} - \frac{1}{8k}\right) \geq \left(k - \frac{1}{2}\right)|\overline{S}| - \frac{n}{8}$$

holds. Thus

$$\sum_{v \in \overline{S}} \left(\mu_v^{(T)} - \chi_v^{(T)}\right) \geq \left(\left(k - \frac{1}{2}\right)|\overline{S}| + \frac{n}{8}\right) - (k-1)|\overline{S}| = \frac{1}{2}|\overline{S}| - \frac{n}{8} \geq \frac{n}{8}$$

where the last inequality follows $|\overline{S}| \geq n/2$. We obtain the claim. □

5.3.2 Lower bounds for specific models

Proposition 5.13. *Suppose n is even and $P_{u,v} = 1/n$ for any $u, v \in V$. Then there exists an example of rotor-router model such that*

$$\max_{S \subseteq V} \left|\chi_S^{(T)} - \mu_S^{(T)}\right| = \Omega(m)$$

holds for any $T > 0$.

Proof. We consider a random walk on a complete graph $K_{2n'}$, i.e., let $V = \{0, 1, \ldots, 2n' - 1\}$ ($n' \in \mathbb{Z}_{>0}$) and $P_{u,v} = 1/(2n')$ for any $u, v \in V$. Let $A = \{0, 1, \ldots, n' - 1\}$, $B = \{n', n' + 1 \ldots, 2n' - 1\}$ and let

$$\chi_u^{(0)} = \begin{cases} (2k+1)n' & (u \in A) \\ 0 & (u \in B), \end{cases}$$

for an arbitrary $k \in \mathbb{Z}_{\geq 0}$. Note that $M = \|\chi^{(0)}\|_1 = (2k+1)(n')^2$. Since this P mixes in a single step, $\mu_A^{(t)} = \mu_B^{(t)} = (2k+1)(n')^2/2$ holds for any $t > 0$. We define the SRT-router $\sigma_u(i)$ as

$$\sigma_u(i \bmod 2n') = i$$

for any $u \in V$. Then, it is not difficult to check that $\chi_A^{(t)} = (k+1)(n')^2$ and $\chi_B^{(t)} = k(n')^2$ when t is odd, as well as that $\chi_A^{(t)} = k(n')^2$ and $\chi_B^{(t)} = (k+1)(n')^2$ when $t > 0$ is even. Thus,

$$\max_{S \subseteq V} |\chi_S^{(t)} - \mu_S^{(t)}| \geq |\chi_A^{(t)} - \mu_A^{(t)}| = \frac{(n')^2}{2} = \frac{n^2}{8}$$

holds for any $t > 0$. We obtain the claim. □

Note that $t_{\mathrm{mix}} = \mathrm{O}(1)$ holds for the Markov chain considered in Proposition 5.13. Thus

$$\left|\chi_S^{(T)} - \mu_S^{(T)}\right| \leq 2m \cdot \mathrm{O}(1) \ = \ \mathrm{O}(m)$$

holds by Corollary 5.4. This P is an example of $\max_{S \subseteq V} |\chi_S^{(T)} - \mu_S^{(T)}| = \Theta(m)$.

We give the following lower bound for an oblivious model. This proposition implies that the factor t_{mix} in total variation discrepancy for an *oblivious* model is not negligible, in general.

Proposition 5.14. *There exists an example of P and its corresponding oblivious model such that*

$$\max_{S \subseteq V} \left|\chi_S^{(T)} - \mu_S^{(T)}\right| = \Omega(n t_{\mathrm{mix}})$$

holds for any time T after mixing.

Proof. Let $V = \{0, \dots, n-1\}$ and let k be an arbitrary integer greater than one. Let a transition matrix P be defined by $P_{u,u} = (k-1)/k$ for any $u \in V$, and $P_{u,v} = 1/k(n-1)$ for any $u, v \in V$ such that $u \neq v$, i.e., P denotes a simple random walk on K_n with a self-loop probability $(k-1)/k$ for any vertex. For this P, it is not difficult to check that

$$\tau(\varepsilon) \leq k \ln \varepsilon^{-1} \tag{5.9}$$

holds for any $0 < \varepsilon < 1$ (see Proposition 5.15). Now, we give a corresponding oblivious deterministic random walk. Let us assume that the prescribed ordering for each $v \in V$ starts with v itself Let

$$\chi_u^{(0)} = \begin{cases} k & (u \in A) \\ 0 & (u \in B), \end{cases}$$

where $A = \{0, \dots, \lceil n/2 \rceil - 1\}$ and $B = \{\lceil n/2 \rceil, \dots, n-1\}$ $(M = k\lceil n/2 \rceil)$. Then, the initial configuration is stable, i.e., $\chi^{(t)} = \chi^{(0)}$ for any $t \geq 0$, since each $v \in A$ serves $\lfloor k \cdot \frac{k-1}{k} \rfloor + 1 = k$ tokens to itself (notice that the "surplus" token stays at v according to the prescribed ordering). Thus it is easy to see that

$$\max_{S \subseteq V} |\chi_S^{(t)} - \mu_S^{(t)}| \geq |\chi_A^{(t)} - \mu_A^{(t)}| \geq k \left\lceil \frac{n}{2} \right\rceil - \left(\frac{kn}{4} + \varepsilon \right) \geq \frac{kn}{4} - \varepsilon \geq \frac{n\tau(\varepsilon)}{4 \ln \varepsilon^{-1}} - \varepsilon = \Omega(nt_{\mathrm{mix}})$$

holds for any $t \geq \tau(\varepsilon)$. We obtain the claim. □

Proposition 5.15. *Let*

$$P_{u,v} = \begin{cases} \frac{k-1}{k} & (\text{if } v = u) \\ \frac{1}{k(n-1)} & (\text{otherwise}). \end{cases}$$

Then

$$\tau(\varepsilon) \leq k \ln \varepsilon^{-1}.$$

Proof. For this P, it is not difficult to check

$$P_{u,v}^t = \begin{cases} \frac{1}{n} + \frac{n-1}{n} \left(1 - \frac{n}{k(n-1)}\right)^t & (\text{if } v = u) \\ \frac{1}{n} - \frac{1}{n} \left(1 - \frac{n}{k(n-1)}\right)^t & (\text{otherwise}) \end{cases}$$

holds for any $t \geq 0$. Thus,

$$\mathcal{D}_{\mathrm{tv}}(P^t_{u,\cdot}, \pi) = \frac{1}{2}\left(|P^t_{u,u} - \pi_u| + \sum_{v \in V, v \neq u} |P^t_{u,v} - \pi_v|\right) = \frac{n-1}{n}\left(1 - \frac{n}{k(n-1)}\right)^t \quad (5.10)$$

holds for any $t \geq 0$. By (5.10),

$$\tau(\varepsilon) = \left\lceil \frac{\ln \varepsilon^{-1} - \ln \frac{n}{n-1}}{\ln \left(1 - \frac{n}{k(n-1)}\right)^{-1}} \right\rceil \leq \ln \varepsilon^{-1} \cdot \frac{k(n-1)}{n} \leq k \ln \varepsilon^{-1}$$

holds, where we used the fact that $\log(1-x)^{-1} \geq x$ holds for any x $(0 < x < 1)$. We obtain the claim. □

Remark that our example on a complete graph, implies only $\Omega(\sqrt{m} t_{\mathrm{mix}})$ lower bound. The gap between the upper and lower bounds remains as open.

Chapter 6

Visit Frequency and Cover Time

Chapter 6 analyses the visit frequency $X_w^{(T)} = \sum_{t=0}^{T-1} \chi_w^{(t)}$ (see (3.8) at vertex $w \in V$). Let $M_w^{(T)}$ denote the expected visit frequency on random walks, i.e.,

$$M_w^{(T)} \stackrel{\text{def}}{=} \sum_{t=0}^{T-1} \mu_w^{(t)} \tag{6.1}$$

where $\mu^{(t)}$ is the expected number of tokens on $w \in V$ at time t (see (2.15)). Then we give an analysis of the cover time of deterministic random walks.

6.1 Analysis of the visit frequency

Section 6.1 establishes the following upper bound of the discrepancy of visit frequency.

Theorem 6.1. *Suppose $P \in \mathbb{R}_{\geq 0}^{n\times n}$ is ergodic and reversible. Then,*

$$\left|X_w^{(T)} - M_w^{(T)}\right| \leq 3\pi_w \max_{u\in V} \frac{\sum_{v\in\mathcal{N}(u)} \Phi_{v,u}}{\pi_u} t_{\mathrm{mix}} = \mathrm{O}\left(\Phi \frac{\pi_{\max}}{\pi_{\min}} \Delta t_{\mathrm{mix}}\right)$$

for any $w \in V$ and for any $T \geq 0$.

This bound says that the discrepancy of visit frequencies depends on Φ, which is independent of number of tokens k and T for SRT/billiard/weighted rotor-router models.

From Theorem 6.1, we get the following Corollary 6.2, like Theorem 4 of [45].

Corollary 6.2. *Suppose $P \in \mathbb{R}_{\geq 0}^{n\times n}$ is ergodic and reversible. Then,*

$$\left|\frac{X_w^{(T)}}{kT} - \pi_w\right| \leq \frac{3t_{\mathrm{mix}}}{2T} + \frac{3\pi_w \max_{u\in V} \frac{\sum_{v\in\mathcal{N}(u)} \Phi_{v,u}}{\pi_u} t_{\mathrm{mix}}}{kT}$$

holds for any $w \in V$ and for any $T > 0$.

Note that Corollary 6.2 gives an upper bound for the SRT-router model with k tokens, while Theorem 4 of [45] is for rotor-router models with a single token. Corollary 6.2 also means that $\left|\pi_w - \frac{X_w^{(T)}}{kT}\right| \leq \varepsilon$ if $T \geq 3\left(\frac{1}{2} + \frac{\pi_w \Delta}{\pi_{\min} k}\right) t_{\mathrm{mix}} \varepsilon^{-1}$.

To prove Theorem 6.1, we introduce the following lemma.

Lemma 6.3.

$$X_w^{(\mathcal{T})} - M_w^{(\mathcal{T})} = \sum_{u \in V} \sum_{v \in \mathcal{N}^-(u)} \sum_{t=0}^{\mathcal{T}-2} \Phi_{v,u}^{(\mathcal{T}-t)} (P_{u,w}^t - \pi_w)$$

holds for any $w \in V$ and for any $\mathcal{T} > 1$.

Proof. From the definitions and Lemma 4.3, we have

$$\begin{aligned} X_w^{(\mathcal{T})} - M_w^{(\mathcal{T})} &= \sum_{T=0}^{\mathcal{T}-1} \left(\chi_w^{(T)} - \mu_w^{(T)}\right) = \sum_{T=1}^{\mathcal{T}-1} \left(\chi_w^{(T)} - \mu_w^{(T)}\right) \\ &= \sum_{u \in V} \sum_{v \in \mathcal{N}^-(u)} \sum_{T=1}^{\mathcal{T}-1} \sum_{t=0}^{T-1} \phi_{v,u}^{(T-t-1)} (P_{u,w}^t - \pi_w). \end{aligned} \tag{6.2}$$

Then, carefully exchanging the variables of the summation, we obtain

$$\sum_{T=1}^{\mathcal{T}-1} \sum_{t=0}^{T-1} \phi_{v,u}^{(T-t-1)} (P_{u,w}^t - \pi_w) = \sum_{t=0}^{\mathcal{T}-2} \sum_{T=t+1}^{\mathcal{T}-1} \phi_{v,u}^{(T-t-1)} (P_{u,w}^t - \pi_w). \tag{6.3}$$

Thus, we obtain the claim combining (6.2), (6.3) and definition (3.13). □

Proof of Theorem 6.1. Note that $\mathcal{N}^-(u) = \mathcal{N}(u)$ holds for any $u \in V$ and $\pi_u P_{u,w}^t = \pi_w P_{w,u}^t$ hold from the assumption of reversibility of P. Lemma 6.3 implies that

$$X_w^{(\mathcal{T})} - M_w^{(\mathcal{T})} = \pi_w \sum_{u \in V} \frac{1}{\pi_u} \sum_{v \in \mathcal{N}(u)} \sum_{t=0}^{\mathcal{T}-2} \Phi_{v,u}^{(\mathcal{T}-t)} (P_{w,u}^t - \pi_u).$$

Thus, by definition (3.13),

$$
\begin{aligned}
\left|X_w^{(\mathcal{T})} - M_w^{(\mathcal{T})}\right| &\leq \pi_w \sum_{u\in V} \frac{1}{\pi_u} \sum_{v\in\mathcal{N}(u)} \sum_{t=0}^{\mathcal{T}-2} \left|\Phi_{v,u}^{(\mathcal{T}-t)}\right| |P_{w,u}^t - \pi_u| \\
&\leq \pi_w \sum_{u\in V} \frac{1}{\pi_u} \sum_{v\in\mathcal{N}(u)} \sum_{t=0}^{\mathcal{T}-2} \Phi_{v,u} |P_{w,u}^t - \pi_u| \\
&\leq \pi_w \max_{u\in V} \frac{\sum_{v\in\mathcal{N}(u)} \Phi_{v,u}}{\pi_u} \sum_{u\in V} \sum_{t=0}^{\mathcal{T}-2} |P_{w,u}^t - \pi_u| \\
&\leq \pi_w \max_{u\in V} \frac{\sum_{v\in\mathcal{N}(u)} \Phi_{v,u}}{\pi_u} \sum_{t=0}^{\mathcal{T}-2} 2d(t)
\end{aligned}
$$

holds, and we obtain the claim by Lemma 4.4. □

Proof of Corollary 6.2. Notice that

$$
\begin{aligned}
\left|\pi_w - \frac{X_w^{(T)}}{kT}\right| &= \frac{\left|kT\pi_w - X_w^{(T)}\right|}{kT} \leq \frac{\left|kT\pi_w - M_w^{(T)}\right| + \left|M_w^{(T)} - X_w^{(T)}\right|}{kT} \\
&\leq \frac{\left|M_w^{(T)} - kT\pi_w\right|}{kT} + \frac{3\pi_w \max_{u\in V} \frac{\max_{u\in V}\sum_{v\in\mathcal{N}(u)}\Phi_{v,u}}{\pi_u} t_{\mathrm{mix}}}{kT},
\end{aligned}
$$

where the last inequality follows Theorem 6.1. Thus, it is sufficient to prove that $|M_w^{(T)} - kT\pi_w| \leq 3kt_{\mathrm{mix}}/2$. Note that $\sum_{t=0}^{T-1}\sum_{u\in V}\mu^{(0)}\pi_w = kT\pi_w$ holds since $\sum_{v\in V}\mu^{(0)} = k$ from the definition, and also note that $M_w^{(T)} = \sum_{t=0}^{T-1}\mu_w^{(t)} = \sum_{t=0}^{T-1}\sum_{u\in V}\mu_u^{(0)}P_{u,w}^t$ holds by the definitions. Then,

$$
\begin{aligned}
\left|M_w^{(T)} - kT\pi_w\right| &= \left|\sum_{t=0}^{T-1}\sum_{u\in V}\mu_u^{(0)}P_{u,w}^t - \sum_{t=0}^{T-1}\sum_{u\in V}\mu^{(0)}\pi_w\right| = \left|\sum_{t=0}^{T-1}\sum_{u\in V}\mu_u^{(0)}(P_{u,w}^t - \pi_w)\right| \\
&\leq \sum_{u\in V}\mu_u^{(0)}\sum_{t=0}^{T-1}|P_{u,w}^t - \pi_w| \qquad (6.4)
\end{aligned}
$$

holds. By Lemma 4.4 and the definition of total variation distance (2.5),

$$
\sum_{t=0}^{T-1}|P_{u,w}^t - \pi_w| \leq \sum_{t=0}^{T-1}\mathcal{D}_{\mathrm{tv}}(P_{u,\cdot}^t, \pi) \leq \frac{3}{2}t_{\mathrm{mix}}. \qquad (6.5)
$$

Combining (6.4) and (6.5), $|M_w^{(T)} - kT\pi_w| \leq 3kt_{\mathrm{mix}}/2$ holds, and we obtain the claim. □

6.2 Cover time

Combining techniques of the analysis of the visit frequency and reversible Markov chains, we get the cover time of deterministic random walks, where the cover time of a deterministic random walk is given by

$$T_{\rm cov} = \min\left\{T \in \mathbb{Z}_{\geq 0} \mid X_v^{(T)} \geq 1 \textit{ holds for any } v \in V\right\}. \tag{6.6}$$

6.2.1 General upper bound of the cover time

First, we show the following theorem.

Theorem 6.4. *Suppose $P \in \mathbb{R}_{\geq 0}^{n\times n}$ is ergodic and reversible. If*

$$T > 2t_{\rm mix} + \frac{12t_{\rm mix}}{k} \max_{u\in V} \frac{\sum_{v\in\mathcal{N}(u)} \Phi_{v,u}}{\pi_u}$$

then $X_w^{(T)} \geq 1$ holds for any $w \in V$.

To prove Theorem 6.4, we introduce the following proposition giving a lower bound of $M_w^{(T)}$.

Proposition 6.5. *Suppose P is ergodic, reversible and $T \geq 2t_{\rm mix}$. Then,*

$$M_w^{(T)} \geq \frac{k\pi_w(T - 2t_{\rm mix})}{4}$$

holds for any $u, w \in V$.

Proof. Note that $s(t) \leq s(2t_{\rm mix})$ holds from the hypothesis $t \geq 2t_{\rm mix}$. Then, we have

$$1 - \frac{P_{u,w}^t}{\pi_w} \leq s(t) \leq s(2t_{\rm mix}) \leq 1 - (1 - \bar{d}(t_{\rm mix}))^2 \leq 1 - \left(1 - \frac{1}{2}\right)^2 = \frac{3}{4}$$

for any $u, w \in V$ and $t \geq 2t_{\rm mix}$. Thus

$$\begin{aligned} M_w^{(T)} &= \sum_{t=0}^{T-1}\sum_{u\in V} \mu_u^{(0)} P_{u,w}^t \geq \sum_{t=2t_{\rm mix}}^{T-1}\sum_{u\in V} \mu_u^{(0)} P_{u,w}^t \\ &\geq \sum_{t=2t_{\rm mix}}^{T-1}\sum_{u\in V} \mu_u^{(0)} \frac{\pi_w}{4} = \frac{k\pi_w(T - 2t_{\rm mix})}{4} \end{aligned} \tag{6.7}$$

holds, and we obtain the claim. □

Proof of Theorem 6.4. By Theorem 6.1 and Proposition 6.5, we have

$$\begin{aligned} X_w^{(T)} &\geq M_w^{(T)} - 3\pi_w \max_{u\in V} \frac{\sum_{v\in\mathcal{N}(u)} \Phi_{v,u}}{\pi_u} t_{\mathrm{mix}} \\ &\geq \frac{k\pi_w(T-2t_{\mathrm{mix}})}{4} - 3\pi_w \max_{u\in V} \frac{\sum_{v\in\mathcal{N}(u)} \Phi_{v,u}}{\pi_u} t_{\mathrm{mix}} \end{aligned} \tag{6.8}$$

for any $w \in V$ and $T \geq 0$. Then, from (6.8) and the assumption of the theorem, we obtain

$$X_w^{(T)} > 0 \tag{6.9}$$

holds for any $w \in V$ and $T > 2t_{\mathrm{mix}} + \frac{12t_{\mathrm{mix}}}{k} \max_{u\in V} \frac{\sum_{v\in\mathcal{N}(u)} \Phi_{v,u}}{\pi_u}$. Thus we obtain the claim since $X_w^{(T)}$ is an integer. □

6.2.2 Upper bounds for specific models

Theorem 6.4 gives upper bounds of deterministic random walks. First, we show the following bound for the SRT-router model.

Corollary 6.6. *Suppose P is ergodic and reversible. Then, the cover time of its corresponding SRT-router model with k tokens satisfies*

$$T_{\mathrm{cov}} \leq 2t_{\mathrm{mix}} + 1 + \frac{12t_{\mathrm{mix}}}{k} \max_{u\in V} \frac{\delta(u)}{\pi_u} = \mathrm{O}\left(\max\left\{\frac{t_{\mathrm{mix}}\Delta}{\pi_{\min}k}, t_{\mathrm{mix}}\right\}\right).$$

Proof. From Proposition 4.8, $\Phi_{v,u} \leq 1$ holds for any $v, u \in V$. Thus Theorem 6.4 guarantees $X_w^{(T)} \geq 1$ for any $T > 2t_{\mathrm{mix}} \frac{12t_{\mathrm{mix}}}{k} \max_{u\in V} \frac{\delta(u)}{\pi_u}$, and we obtain the claim. □

We obtain the following bounds for billiard/weighted rotor-router model, similarly.

Corollary 6.7. *Suppose P is ergodic and reversible. Then, the cover time of its corresponding billiard-router model with k tokens satisfies*

$$T_{\mathrm{cov}} \leq 2t_{\mathrm{mix}} + 1 + \frac{24t_{\mathrm{mix}}}{k} \max_{u\in V} \frac{\delta(u)}{\pi_u} = \mathrm{O}\left(\max\left\{\frac{t_{\mathrm{mix}}\Delta}{\pi_{\min}k}, t_{\mathrm{mix}}\right\}\right).$$

Corollary 6.8. *Suppose P is ergodic and reversible. Then, the cover time of its corresponding weighted-router model with k tokens satisfies*

$$T_{\mathrm{cov}} \leq 2t_{\mathrm{mix}} + 1 + \frac{12t_{\mathrm{mix}}}{k} \max_{u\in V} \frac{\bar{\delta}(u)}{\pi_u} = \mathrm{O}\left(\max\left\{\frac{t_{\mathrm{mix}}\bar{\Delta}}{\pi_{\min}k}, t_{\mathrm{mix}}\right\}\right).$$

Corollary 6.8 implies the following upper bounds for rotor-router model on G.

Corollary 6.9. *The cover time of any rotor-router model with k tokens on $G = (V, E)$ satisfies*

$$T_{\mathrm{cov}} \leq 2t_{\mathrm{mix}} + 1 + \frac{24mt_{\mathrm{mix}}}{k} = \mathrm{O}\left(\max\left\{\frac{mt_{\mathrm{mix}}}{k}, t_{\mathrm{mix}}\right\}\right),$$

where t_{mix} is the mixing time of a simple random walk on G.

Proof. We see that $\bar{\delta}(u) = \delta(u)$ and $\max_{u\in V} \frac{\delta(u)}{\pi_u} = 2m$, since $\pi_u = \frac{\delta(u)}{2m}$. Thus,

$$T_{\mathrm{cov}} \leq 2t_{\mathrm{mix}} + 1 + \frac{24mt_{\mathrm{mix}}}{k}$$

holds. □

The upper bound by [60] (Theorem 4.1, Proposition 4.2, and Theorem 4.5) is $\mathrm{O}\big(t_{\mathrm{mix}} + (\Delta/\delta)(mt_{\mathrm{mix}}/k)\big)$, where Δ/δ is the maximum/minimum degree of the graph. Hence Corollary 6.9 improves their bound for irregular graphs. Compare to the $\mathrm{O}(mD/\log k)$ bound by [26] (Theorem 3.3 and 3.7), our bound is better when $t_{\mathrm{mix}} = \mathrm{O}\big(D(k/\log k)\big)$ (meaning that t_{mix} is small or k is large).

For the rotor-router model with multiple tokens, we have the following theorems.

Theorem 6.10 ([10](Theorem 2))**.** *Let $T^{(i)}_{\mathrm{cov}}$ be the cover time defined by* (6.6) *on the assumption of $\sum_{v\in V} \chi^{(t)}_v = i$ (the total number of tokens is i). Then,*

$$T^{(1)}_{\mathrm{cov}} \geq \frac{1}{4}mD$$

holds for any rotor-router model with 1 *token on $G = (V, E)$.*

Theorem 6.11 ([88](Theorem 3) and [26] (Theorem 4.1))**.** *Let $T^{(i)}_{\mathrm{cov}}$ be the cover time defined by* (6.6) *on the assumption of $\sum_{v\in V} \chi^{(t)}_v = i$ (the total number of tokens is i). Then,*

$$\frac{T^{(1)}_{\mathrm{cov}}}{T^{(k)}_{\mathrm{cov}}} = \mathrm{O}(k)$$

holds for any rotor-router model with k tokens on $G = (V, E)$

Combining Theorem 6.10, Theorem 6.11 and Corollary 6.9, we get the following theorem.

Theorem 6.12. *Let $T_{\rm cov}^{(i)}$ be the cover time defined by* (6.6) *on the assumption of* $\sum_{v\in V}\chi_v^{(t)} = i$ *(the total number of tokens is i). If the mixing time of a simple random walk on G satisfies $t_{\rm mix} = {\rm O}(D)$, then*

$$\frac{T_{\rm cov}^{(1)}}{T_{\rm cov}^{(k)}} = \Theta(k)$$

holds for any rotor-router model on G with $k \le m$ tokens.

Proof. From the assumptions $t_{\rm mix} = {\rm O}(D)$, $k \le m$ and Corollary 6.9, we have

$$T_{\rm cov}^{(k)} = {\rm O}\left(\frac{mD}{k}\right). \tag{6.10}$$

Combining (6.10) and Theorem 6.10, we have

$$\frac{T_{\rm cov}^{(1)}}{T_{\rm cov}^{(k)}} = \Omega(k), \tag{6.11}$$

thus we obtain the claim from (6.11) and Theorem 6.11. □

Theorem 6.12 implies that the rotor-router model achieves k-times speed up ratio for general expander graphs. The same speed up ratio as Theorem 6.12 is proven for simple *random* walks [36] (Theorem 5.1).

Chapter 7

Conclusion

This paper gave analyses of the deterministic random walk. Chapter 4 gave an upper bound of the vertex-wise discrepancy $|\chi_v^{(t)} - \mu_v^{(t)}|$ when Markov chain is ergodic and reversible. It is a future work if the vertex-wise discrepancy is independent of $\pi_{\max}/\pi_{\min}$ and $t_{\rm mix}$.

Chapter 5 gave an upper bounds of the total variation discrepancy $\|\chi^{(t)} - \mu^{(t)}\|_1 = \mathrm{O}(mt_{\rm mix})$ for any ergodic Markov chains, where note that the bound depends on the number of edges m but is independent of $\pi_{\max}/\pi_{\min}$. We also showed some lower bounds. The gap between upper and lower bounds is a future work. Development of a deterministic approximation algorithm based on deterministic random walks for #P-hard problems is a challenge.

Chapter 6 has developed analytic techniques for the visit frequency $X_v^{(T)}$ of deterministic random walks with multiple tokens, and gave an upper bound of the cover time for any ergodic and reversible Markov chains. Also, the upper bound improves the existing results of the rotor-router model with multiple tokens in general case. A better upper bound of the cover time by derandomizing a specific *fast* random walk (e.g., β-random walk, Metropolis walk, the minimum degree weighting scheme) is a future work.

Reference

[1] M. A. Abdullah, C. Cooper, M. Draif, Speeding up cover time of sparse graphs using local knowledge, Proceedings of the 27th International Workshop on Combinatorial Algorithms (IWOCA 2015), 1–12.

[2] H. Akbari and P. Berenbrink, Parallel rotor walks on finite graphs and applications in discrete load balancing, Proceedings of the 25th ACM Symposium on Parallelism in Algorithms and Architectures (SPAA 2013), 186–195.

[3] D. Aldous and P. Diaconis, Strong uniform times and finite random walks, Advances in Applied Mathematics, **8** (1987), 69–97.

[4] D. Aldous and J. Fill, Reversible markov chains and random walks on graphs, http://stat- www.berkeley.edu/pub/users/aldous/RWG/book.html.

[5] R. Aleliunas, R. Karp, R. Lipton, L. Lovasz and C. Rackoff, Random walks, universal traversal sequences, and the complexity of maze problems, Proceedings of the 20th Annual IEEE Symposium on Foundations of Computer Science (FOCS 1979), 218–223.

[6] N. Alon, C. Avin, M. Koucky, G. Kozma, Z. Lotker, M. R. Tuttle, Many random walks are faster than one, Combinatorics, Probability & Computing, **20** (2011), 481–502.

[7] E. Ando and S. Kijima, An FPTAS for the volume computation of 0-1 knapsack polytopes based on approximate convolution integral, Algorithmica, **76** (2016), 1245–1263.

[8] O. Angel, A.E. Holroyd, J. Martin, and J. Propp, Discrete low discrepancy sequences, arXiv:0910.1077.

[9] S. Angelopoulos, B. Doerr, A. Huber, and K. Panagiotou, Tight bounds for quasirandom rumor spreading, The Electronic Journal of Combinatorics, **16** (2009), #R102.

[10] E. Bampas, L. Gasieniec, N. Hanusse, D. Ilcinkas, R. Klasing, and A. Kosowski, Euler tour lock-in problem in the rotor-router model, Proceedings of the 23rd International Symposium on Distributed Computing (DISC 2009), 423–435.

[11] A. Bandyopadhyay and D. Gamarnik, Counting without sampling: asymptotics of the log-partition function for certain statistical physics models, Random Structures & Algorithms, **33** (2008), 452–479.

[12] P. Berenbrink, R. Klasing, A. Kosowski, F. Mallmann-Trenn, and P. Uznanski, Improved analysis of deterministic load-balancing schemes, Proceedings of the 2015 ACM Symposium on Principles of Distributed Computing (PODC2015), 301–310.

[13] G. Brightwell and P. Winkler, Counting linear extensions, Order, **8** (1991), 225–242.

[14] S. Brin and L. Page, The anatomy of a large-scale hypertextual Web search engine, Computer Networks and ISDN Systems, **33** (1998), 7–17.

[15] A. Broder, A. Karlin, P. Raghavan, E. Upfal, Trading space for time in undirected s–t connectivity, SIAM Journal on Computing **23** (1994) 324–334.

[16] R. Bubley and M. Dyer, Faster random generation of linear extensions, Discrete Mathematics, **201** (1999), 81–88.

[17] J. Chalopin, S. Das, P. Gawrychowski, A. Kosowski, A. Labourel and P. Uznanski, Limit behavior of the multi-agent rotor-router system, Proceedings of the 29th International Symposium on Distributed Computing (DISC 2015), 123–139.

[18] C. Cooper, R. Elsasser, T. Radzik, N. Rivera, T. Shiraga, Fast consensus for voting on general expander graphs, Proceedings of the 29th International Symposium on Distributed Computing (DISC 2015), 248–262.

[19] C. Cooper, T. Radzik, N. Rivera,T. Shiraga, Coalescing walks on rotor-router systems, Proceedings of the 22nd International Colloquium on Structural Information and Communication Complexity (SIROCCO 2015), 444–458.

[20] J. Cooper, B. Doerr, T. Friedrich, and J. Spencer, Deterministic random walks on regular trees, Random Structures & Algorithms, **37** (2010), 353–366.

[21] J. Cooper, B. Doerr, J. Spencer, and G. Tardos, Deterministic random walks on the integers, European Journal of Combinatorics, **28** (2007), 2072–2090.

[22] J. Cooper and J. Spencer, Simulating a random walk with constant error, Combinatorics, Probability and Computing, **15** (2006), 815–822.

[23] B. Cousins, S. Vempala: Bypassing KLS: Gaussian cooling and an $O^*(n^3)$ volume algorithm, Proceedings of the 47th Annual Symposium on the Theory of Computing (STOC 2015), 539–548.

[24] J. Czyzowicz, L. Gasieniec, A. Kosowski, E. Kranakis, P. G. Spirakis, P. Uznanski, On convergence and threshold properties of discrete Lotka-Volterra population protocols, Proceedings of the 42nd International Colloquium on Automata, Languages, and Programming (ICALP 2015), 393–405.

[25] R. David and U. Feige, Random walks with the minimum degree local rule have $O(n^2)$ cover time, Proc. SODA 2017, to appear.

[26] D. Dereniowski, A. Kosowski, D. Pajak, and P. Uznanski, Bounds on the cover time of parallel rotor walks, LIPICS, **25** (STACS 2014), 263–275.

[27] B. Doerr, Introducing quasirandomness to computer science, Efficient Algorithms, volume 5760 of Lecture Notes in Computer Science, Springer Verlag, (2009), 99–111.

[28] B. Doerr and T. Friedrich, Deterministic random walks on the two-dimensional grid, Combinatorics, Probability and Computing, **18** (2009), 123–144.

[29] B. Doerr, T. Friedrich, M. Künnemann, and T. Sauerwald, Quasirandom rumor spreading: An experimental analysis, ACM Journal of Experimental Algorithmics, **16** (2011), 3.3:1–3.3:13.

[30] B. Doerr, T. Friedrich, and T. Sauerwald, Quasirandom rumor spreading, Proceedings of the 19th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA 2008), 773–781.

[31] B. Doerr, T. Friedrich, and T. Sauerwald, Quasirandom rumor spreading: Expanders, push vs. pull, and robustness, Proceedings of the 36th International Colloquium on Automata, Languages and Programming, (ICALP 2009), 366–377.

[32] B. Doerr, T. Friedrich, and T. Sauerwald, Quasirandom rumor spreading on expanders, Electronic Notes in Discrete Mathematics, **34** (2009), 243–247.

[33] B. Doerr, A. Huber, and A. Levavi, Strong robustness of randomized rumor spreading protocols, Proceedings of the 20th International Symposium on Algorithms and Computation (ISAAC 2009), 812–821.

[34] R. Durrett, Probability: Theory and examples, Cambridge series in statistical and probabilistic mathematics, 2010.

[35] M. Dyer, A. Frieze, R. Kannan, A random polynomial-time algorithm for approximating the volume of convex bodies, Journal of the ACM, **38** (1991), 1–17.

[36] R. Elsasser and T. Sauerwald, Tight bounds for the cover time of multiple random walks, Theoretical Computer Science, **412** (2011), 2623–2641.

[37] U. Feige, A tight upper bound for the cover time of random walks on graphs, Random Structures and Algorithms, **6** (1995), 51–54.

[38] U. Feige, A tight lower bound for the cover time of random walks on graphs, Random Structures and Algorithms, **6** (1995), 433–438.

[39] T. Friedrich, M. Gairing, and T. Sauerwald, Quasirandom load balancing, SIAM Journal on Computing, **41** (2012), 747–771.

[40] T. Friedrich, M. Katzmann, A. Krohmer, Unbounded discrepancy of deterministic random walks on grids, Proceedings of the 26th International Symposium Algorithms and Computation (ISAAC 2015), 212–222.

[41] T. Friedrich and T. Sauerwald, The cover time of deterministic random walks, The Electronic Journal of Combinatorics, **17** (2010), R167.

[42] P. Gopalan, A. Klivans, and R. Meka, Polynomial-time approximation schemes for knapsack and related counting problems using branching programs, arXiv:1008.3187v1, 2010.

[43] P. Gopalan, A. Klivans, R. Meka, D. Stefankovic, S. Vempala, E. Vigoda, An FPTAS for #knapsack and related counting problems, Proceedings of the 52nd Annual IEEE Symposium on Foundations of Computer Science (FOCS 2011), 817–826.

[44] G. Grimmett, Percolation, 2nd edn. Springer, Berlin, 1999.

[45] A.E. Holroyd and J. Propp, Rotor walks and Markov chains, M. Lladser, R.S. Maier, M. Mishna, A. Rechnitzer, (eds.), Algorithmic Probability and Combinatorics, The American Mathematical Society, 2010, 105–126.

[46] H. Hosoya, Topological index. A newly proposed quantity characterizing the topological nature of structural isomers of saturated hydrocarbons, Bulletin of the Chemical Society of Japan, **44** (1971), 2332–2339.

[47] A. Huber and N. Fountoulakis, Quasirandom broadcasting on the complete graph is as fast as randomized broadcasting, Electronic Notes in Discrete Mathematics, **34** (2009), 553–559.

[48] S. Ikeda, I. Kubo, and M. Yamashita, The hitting and cover times of random walks on finite graphs using local degree information, Theoretical Computer Science, **410** (2009), 94–100.

[49] D. Jerison, L. Levine, and S. Sheffield, Logarithmic fluctuations for internal DLA, Journal of the American Mathematical Society, **25** (2012), 271–301.

[50] M. R. Jerrum and A. Sinclair, Polynomial-time approximation algorithms for the Ising model, SIAM Journal on Computing, **22** (1993), 1087–1116.

[51] M. Jerrum and A. Sinclair, Approximation algorithms for NP-hard problems, D.S. Hochbaum ed., The Markov chain Monte Carlo method: an approach to approximate counting and integration, PWS Publishing, 1996.

[52] M. Jerrum, A. Sinclair, and E. Vigoda, A polynomial-time approximation algorithm for the permanent of a matrix with nonnegative entries, Journal of the ACM, **51** (2004), 671–697.

[53] M. R. Jerrum, L. G. Valiant, and V. V. Vazirani, Random generation of combinatorial structures from a uniform distribution. Theoretical Computer Science, **32** (1986), 169–188.

[54] H. Kajino, S. Kijima, and K. Makino, Discrepancy analysis of deterministic random walks on finite irreducible digraphs, discussion paper.

[55] R.M. Karp, Reducibility among combinatorial problems, in R.E. Miller, J.W. Thatcher, J.D. Bohlinger (eds.), Complexity of Computer Computations, 1972, 85–103.

[56] A. Karzanov and L. Khachiyan, On the conductance of order Markov chains, Order, **8** (1991), 7–15.

[57] S. Kijima, K. Koga, and K. Makino, Deterministic random walks on finite graphs, Random Structures & Algorithms, **46** (2015), 739–761.

[58] S. Kijima and T. Matsui, Approximation algorithm and perfect sampler for closed Jackson networks with single servers, SIAM Journal on Computing, **38** (2008), 1484–1503.

[59] L. Kleinrock, Queueing systems, Vol. 1 Theory, Wiley, New York, 1975.

[60] A. Kosowski and D. Pajak, Does adding more agents make a difference? A case study of cover time for the rotor-router, Proceedings of the 41st International Colloquium on Automata, Languages and Programming (ICALP 2014), 544–555.

[61] G.F. Lawler, M. Bramson, and D. Griffeath, Internal diffusion limited aggregation, The Annals of Probability, **20** (1992), 2117–2140.

[62] L. Levine and Y. Peres, The rotor-router shape is spherical, The Mathematical Intelligencer, **27** (2005), 9–11.

[63] L. Levine and Y. Peres, Spherical asymptotics for the rotor-router model in $\mathbb{Z}^d$, Indiana University Mathematics Journal, **57** (2008), 431–450.

[64] L. Levine and Y. Peres, Strong spherical asymptotics for rotor-router aggregation and the divisible sandpile, Potential Analysis, **30** (2009), 1–27.

[65] D. A. Levine, Y. Peres, and E. L. Wilmer, Markov Chain and Mixing Times, The American Mathematical Society, 2008.

[66] L. Lovasz and S. Vempala, Fast algorithms for logconcave functions: sampling, rounding, integration and optimization, Proceedings of the 47th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2006), 57–68.

[67] R. Montenegro and P. Tetali, Mathematical Aspects of Mixing Times in Markov Chains, NOW Publishers, 2006.

[68] B. Morris and A. Sinclair, Random walks on truncated cubes and sampling 0-1 knapsack solutions, SIAM Journal on Computing, **34** (2004), 195–226.

[69] H. Niederreiter, Quasi-Monte Calro methods and pseudo-random numbers, Bull. Amer. Math .Soc., **84**(1978), 957–1042

[70] Y. Nonaka, H. Ono, K. Sadakane, M. Yamashita, The hitting and cover times of Metropolis walks, Theoretical Computer Science, **411** (2010), 1889–1894.

[71] V. Priezzhev, D. Dhar, A. Dhar, and S. Krishnamurthy, Eulerian walkers as a model of self-organized criticality, Physical Review Letters, **77** (1996), 5079–5082.

[72] J. Propp and D. Wilson, Exact sampling with coupled Markov chains and applications to statistical mechanics, Random Structures Algorithms, **9** (1996), 223–252.

[73] Y. Rabani, A. Sinclair, and R. Wanka, Local divergence of Markov chains and analysis of iterative load balancing schemes, Proceedings of the 39th Annual Symposium on Foundations of Computer Science (FOCS 1998), 694–705.

[74] S. Sano, N. Miyoshi, R. Kataoka, m-Balanced words: A generalization of balanced words, Theoretical Computer Science, **314** (2004), 97–120.

[75] T. Sauerwald and H. Sun, Tight bounds for randomized load balancing on arbitrary network topologies, Proceedings of the 53rd Annual Symposium on Foundations of Computer Science (FOCS 2012), 341–350.

[76] T. Shiraga, The cover time of deterministic random walks for general transition probabilities, Proceedings of the 27th International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms (AofA'16), 328–340.

[77] T. Shiraga, Y. Yamauchi, S. Kijima, and M. Yamashita, Deterministic random walks for rapidly mixing chains, arXiv:1311.3749, 2013.

[78] T. Shiraga, Y. Yamauchi, S. Kijima, and M. Yamashita, L_∞-discrepancy analysis of polynomial-time deterministic samplers emulating rapidly mixing chains, Proceedings of the 20th International Computing and Combinatorics Conference (COCOON 2014), 25–36.

[79] T. Shiraga, Y. Yamauchi, S. Kijima, and M. Yamashita, Total variation discrepancy of deterministic random walks for ergodic Markov chains, Proceedings of the meeting of Analytic Algorithmics and Combinatorics (ANALCO 2016), 138–148.

[80] A. Sinclair, Algorithms for Random Generation & Counting, A Markov chain approach, Birkhäuser, 1993.

[81] D. Stefankovic, S. Vempala, and E. Vigoda, A deterministic polynomial-time approximation scheme for counting knapsack solutions, SIAM Journal on Computing, **41** (2012), 356–366.

[82] R. Tijdeman, The chairman assignment problem, Discrete Math. **32** (1980), 323–330.

[83] S. Toda, PP is as hard as the polynomial-time hierarchy, SIAM Journal on Computing, **20** (1991), 865–877.

[84] L.G. Valiant, The complexity of computing the permanent, Theoretical Computer Science, **8** (1979), 189–201.

[85] L.G. Valiant, The complexity of enumeration and reliability problems, SIAM Journal on Computing, **8** (1979), 410–421.

[86] J. G. van der Corput, Verteilungsfunktionen, Proc. Akad. Amsterdam, **38** (1935), 813–821.

[87] D. Weitz, Counting independent sets up to the tree threshold, Proceedings of the 38th ACM Symposium on Theory of Computing (STOC 2006), 140–149.

[88] V. Yanovski, I.A. Wagner, and A.M. Bruckstein, A distributed ant algorithm for efficiently patrolling a network, Algorithmica, **37** (2003), 165–186.