博士論文

Algorithmic Study on Randomness in Graphs

(グラフのランダムネスに関するアルゴリズムの研究)

清水 伸高

Algorithmic Study on Randomness in Graphs

Nobutaka Shimizu

Abstract

The general topic of this thesis is the property that a random item drawn from some probability distribution likely to have. We investigate four topics concerning the structure of random graphs and the behavior of randomized algorithms.

The first topic is the average-case complexity, that is the computational complexity of a problem under the assumption that the input is drawn from a probability distribution. The perspective of considering a random input relaxes the pessimism of worst-case complexity that regards the hardest instance of a problem. For graph problems, several researchers have shown that some NP-hard problems such as finding a Hamiltonian cycle admit a polynomial-time algorithm if the input is a random graph. With this in mind, we consider the following natural research question: Is there a problem that is hard to solve even for random graphs? We handle this issue by presenting a problem that is hard in this sense. Specifically, we consider the problem of counting biclique subgraphs and show a "sharp threshold" result: As a positive result, we present an algorithm that solves the problem for any input. As a negative result, we prove that any slightly faster algorithm fails to solve the problem on most of the random graph under a widely-investigated conjecture concerning worst-case hardness. At the heart of this result, we present a general framework of fine-grained hardness amplification, which is inspired by the classical technique from average-case complexity.

Our next topic is a voting process that is a certain type of randomized distributed algorithms on a graph. Voting processes are known as simple models of consensus dynamics and have an application to the consensus problem in the area of distributed computing. Roughly speaking, voting processes on a graph consists of a single dense component such as the Erdős–Rényi graph are known to exhibit simple dynamics and thus converge to consensus quickly. However, voting processes on graphs with more complex structures are much less understood. In response to it, we consider the stochastic block model, which is a random graph consists of two distinct Erdős–Rényi graphs joined by random edges. We obtain a phase transition result concerning the edge density between the components regarding the dynamics: Above the threshold, the dynamics are simple and converge to consensus quickly. Below the threshold, the dynamics expose a "meta-stable" equilibrium and thus require exponentially long rounds to reach consensus. Another contribution of this thesis regarding voting processes is to introduce a new notion of quasi-majority functional voting that is a wide class of voting processes containing several previously-known voting processes. We then prove that the dynamics of any quasi-majority functional voting on graphs consists of a single dense component (i.e., expander graphs) are simple and converge to consensus quickly.

The third topic is random walks on growing networks. A random walk on a graph is a fundamental stochastic process: A walker on a vertex repeats moving to a randomly selected neighbor. Random walks have a wide variety of applications in network analysis. In particular, a random walk on dynamic networks has gathered special attention since real-world networks change their structure over time. However, most of the previous works were concerned with random walks on dynamic graphs with static vertex set (i.e., they considered graphs in which the edge set changes over time). We propose a new model of random walks on a growing graph and then study their performance.

Finally, we explore the distance properties of random graphs. Specifically, we are interested in the average distance and diameter of dense random regular graphs. The average distance and diameter of regular graphs attract special attention in the literature of high-performance computing since a regular graph with a low average distance and diameter yields an efficient network topology for parallel computers. We prove that the diameter of a random regular graph is likely to be asymptotically optimal under a certain mild condition of the degree.

Acknowledgments

First and foremost, the completion of my thesis would not have been possible without the support of my supervisor, Satoru Iwata. During my five years of master and doctoral course, he kindly gave me valuable advice and indicated my research direction. I always enjoyed talking with him. I could not spend such a blessed life at the University of Tokyo without his support and patience.

I would like to extend my gratitude to my collaborators: Takeharu Shiraga, Shuichi Hirahara, and Shuji Kijima. Takeharu Shiraga has been discussing with me every week for more than three years. His fruitful idea and broad knowledge of stochastic process always inspired me. Shuichi Hirahara gave me a deep insight into computational complexity. I could not enrich my understanding of average-case complexity without discussing it with him. It was my great opportunity to have a joint work with Shuji Kijima. Not only his exciting idea and extensive knowledge about random walk but also his research attitude were impressive for me.

I am sincerely grateful to all members of Mathematical Informatics 7th Laboratory at the University of Tokyo. In particular, I am deeply indebted to Shin-ichi Tanigawa and Tasuku Soma for being frank and friendly with me. They also extended a great amount of assistance during my period. I must also thank my classmate, Taihei Oki. He played a decisive role in my research life. He has been friendly and I always enjoyed talking with him. Sometimes, I was surprised at his deep research perspective. My thanks go to Shinsuke Sakaue and Kaito Fujii. They provided me a way of online communication under the circumstance of COVID19. Without their effort, I might have been depressed during the last year of my Ph.D. course.

I am deeply grateful to Professors Kunihiko Sadakane, Tsuyoshi Takagi, Shin-ichi Tanigawa, and Hiroshi Imai for their many constructive comments on this thesis.

I also appreciate the financial support of JST ERATO, JST CREST, and RIKEN AIP. Finally, I would like to offer my special thanks to my family for their support and nurturing.

Contents

Chapter 1

Introduction

Suppose we are given a random item chosen from a specific probability distribution. Then, what property does the item likely to have? This question gather a great deal of attention in a wide range of fields including combinatorics and theoretical computer science. In this thesis, we investigate properties that asymptotically almost all objects satisfy. More precisely, for a set S (e.g., the set of all graphs) and a positive integer $n \in \mathbb{N}$, let $S_n \subseteq S$ be the set of elements in S of size n (e.g., the set of graphs of n vertices). Consider a sequence $(D_n)_{n\in\mathbb{N}}$ where each D_n is a probability distribution over S_n. We are interested in a property $\mathcal{P}$ such that the probability $\mathbf{Pr}_{x\sim D_n}[x \text{ satisfies } \mathcal{P}]$ tends to one as $n \to \infty$. Such a property can be seen as the property that a typical object in S chosen from D_n satisfies for large n. In this thesis, as a random object, we focus on two sources of randomness concerning graphs: random graphs and randomized algorithms that run on graphs.

The structure of a typical graph under specific distribution has been well investigated in random graph theory. A random graph refers to a graph sampled according to some probability distribution over a set of graphs. The distribution is called *random graph model*. Usually, a random graph model is specified by either defining the probability measure or giving a generating algorithm. For example, the *Erdős–Rényi graph* is an n-vertex graph $G(n,p)$ where each vertex pair holds an edge with probability p independently to any other vertex pairs. The distribution of $G(n,p)$, denoted by $\mathcal{G}(n,p)$, is called the Erdős–Rényi model.

Random graphs were initially introduced by Erdős [Erd59] to prove the existence of a graph satisfying a certain property. Soon later, Erdős and Rényi [ER59] studied the connectivity of random graphs. At the same time, independently to Erdős and Rényi [ER59], Gilbert [Gil59] studied the connectivity of a different random graph model. These are known as the first systematic studies of random graphs. Since then, several random graph models (e.g., random geometric graph [Gil61], random regular graph (configuration model) [Bol80], and random planar graph [DVW96]) were introduced and various structural properties of random graphs (e.g., the maximum clique [BE76], Hamiltonicity [Pós76], chromatic number [GM75], and perfect matching [ER66]) have been investigated.

It is widely recognized that randomness provides algorithms with a surprising computational power. Using randomness, one can design simple, low memory, and fast algorithms for various kinds of computational tasks such as the minimum cut [Kar93], the polynomial identity testing [Sax09], random walks [Lov93], and approximate counting [JVV86, SJ89]. The typical behavior of a randomized algorithm is an important issue since the performance of a randomized algorithm relies on its typical behavior.

This thesis consists of four topics regarding random graphs and randomized algorithms: time complexity of a specific subgraph counting problem on random graphs, consensus dynamics of voting processes on graphs, random walks on growing graphs, and distance properties of random regular graphs.

In Chapters 3 and 4, we study the time complexity of graph problems on random graphs. It is known that some NP-hard problems such as finding a Hamiltonian cycle [AV79] and a 2-approximation of the chromatic number [GM75] (for any constant $\epsilon > 0$, it is NP-hard to approximate the chromatic number within an $n^{1-\epsilon}$-factor [FK98, Zuc07]) admit simple polynomial-time algorithms if the input is a random graph. On the other hand, some problems including finding a maximum clique [Kar73] and counting the number of clique subgraphs of a fixed size [GR18a, BABB19] are believed to be hard even when the input is a random graph. We study a subgraph counting problem that is hard to solve even for random graphs in terms of *fine-grained complexity*. More precisely, the problem can be solved in $n^{c+o(1)}$-time for *any* graph but solving it in time $n^{c-\epsilon}$ for an $n^{-\Omega(\epsilon)}$-fraction of random graphs is impossible for any constant

$\epsilon > 0$ unless a widely-investigated conjecture fails. In other words, at least $(1-n^{-\Omega(\epsilon)})$-fraction of random graphs are hard to solve for any $n^{c-\epsilon}$-time algorithms, whereas there is an $n^{c+o(1)}$-time algorithm that solves the problem for all graphs.

In Chapters 5 to 7, we focus on a (synchronous) *voting process* that is a certain type of randomized distributed algorithms. In a voting process, we consider an undirected and connected graph where each vertex holds an opinion from a finite set. In each discrete time step, vertices communicate with their neighbors and simultaneously update their opinion according to a predefined protocol. The aim of the protocol is to reach *consensus* in which all vertices hold the same opinion. In Chapter 5, we introduce a *functional voting process* that is a wide class of voting processes. In Chapter 6, we consider specific voting processes on a *stochastic block model*, which is a random graph model playing key role in the context of community detection of networks. We obtain a phase transition result regarding the behavior of the voting processes. In Chapter 7, we introduce a *quasi-majority functional voting process* that is a subclass of functional voting processes and study the consensus time (i.e., the number of steps to reach consensus) of the process on *expander graphs*.

In Chapter 8, we focus on a random walk on a *growing network*. Although dynamic graphs gather great deal of attention in network analysis since the shapes of real-world networks change over time, most previous works concerning random walks on dynamic graphs consider a graph with a static vertex set (only edges change over time). In view of this, we present the notion of *rondam walk on a growing graph* and study the performance of it for several growing graphs.

In Chapter 9, we study the average distance and diameter of the random regular graph $G_{n,d}$, which is a graph selected uniformly at random from the set of n-vertex d-regular graphs. In contrast to $G(n,p)$, it is nontrivial to sample $G_{n,d}$. Indeed, there is no known efficient algorithm that generates $G_{n,d}$ for $d \gg \sqrt{n}$. This makes the analysis of $G_{n,d}$ difficult for large d. We present asymptotic results concerning the average distance and diameter of $G_{n,d}$.

Finally, in Chapter 10, we conclude this thesis.

In what follows, we present overview of our contributions with their backgrounds.

1.1 Average-Case Complexity (Chapters 3 and 4)

One of the goals of computational complexity theory is to understand the tractability of computational tasks. A standard framework for the tractability of a computational task is the *worst-case complexity*. In this framework, we consider a computational problem Π and seek an algorithm that outputs the correct answer for *all* inputs of Π. However, the worst-case complexity can be too pessimistic since it regards only the hardest instance. To be more optimistic, one may seek an algorithm that perform well on almost all inputs. One common way to formalize this perspective is the framework of *average-case complexity* [Lev86]. In this framework, we consider a *distributional problem* $(\Pi, \mathcal{D})$ that is a pair of a problem Π and sequence $\mathcal{D} = (\mathcal{D}_n)_{n\in\mathbb{N}}$ where each $\mathcal{D}_n$ is a distribution over inputs of size n (e.g., $\mathcal{G}(n,p)$). An algorithm is given $n \in \mathbb{N}$ and $x \sim \mathcal{D}_n$ (i.e., x is sampled according to the distribution $\mathcal{D}_n$). Then we are interested in the *success probability*, that is, the probability that the algorithm outputs the correct answer for the random input. In average-case complexity, we seek a fast algorithm that solves a distributional problem with high success probability.

1.1.1 Biclique counting on random bipartite graphs

It is known that several graph problems such as Hamiltonian Cycle and Graph Isomorphism, which are believed not to be in P, admit a polynomial-time algorithm with a high success probability if the input is sampled from $\mathcal{G}(n,p)$ [FM97] for suitable p. Even for polynomial-time solvable problems, similar gaps between average- and worst-case complexity have been observed. For example, the current fastest algorithm finds a maximum matching in an unweighted m-edge n-vertex graph admits in time $O(m\sqrt{n})$, while we can find it in time $O(m \operatorname{polylog} n)$ on $G(n,p)$ with high probability [Mot94].

On the other hand, recently, an average-case hardness of some subgraph counting problems has been established under the assumption of a worst-case hardness [GR18a, BABB19, DLW20]. For example, Boix-Adserà, Brennan, and Bresler [BABB19] proved that we cannot count the number k-clique subgraphs in $G(n,p)$ in time $n^{o(k)}$ unless the exponential time hypothesis (ETH) of Impagliazzo and Paturi [IP01] fails.

In Chapter 3, we consider the problem of counting the number of biclique (a.k.a. complete bipartite graphs) subgraphs of fixed size in a given graph. Formally, a biclique on partite sets each of size a and b is the graph $K_{a,b} = (A \cup B, E)$ for two disjoint sets A and B such that $|A| = a$ and $|B| = b$, and the

edge set E is defined as $E := \{\{i, j\} : i \in A, j \in B\}$. We focus on the problem in which we are asked to count the number of $K_{a,b}$-subgraphs (i.e., subgraphs that are isomorphic to $K_{a,b}$) in a given graph.

Finding or counting bicliques has been investigated from both practical and theoretical motivations. On the practical side, this study has applications in data mining [AS94, MT17] and bioinformatics [DAB+04]. See [MT17, AVJ98] and the references therein for details and lists of further applications. On the theoretical side, the problem of finding or counting biclique subgraphs has been studied in computational complexity theory [Lin18, GJ79] and several exact exponential algorithms have been proposed [CK12, GKL12, BRFGL10, Kut12].

In this thesis, we obtain a negative result and a positive result. As the negative result, we prove that, under the strong exponential time hypothesis (SETH) of Impagliazzo, Paturi and Zane [IPZ01], for any constants $a \geq 3$ and $\epsilon > 0$, there is a constant $b = b(a, \epsilon)$ satisfying the following: Any $n^{a-\epsilon}$-time algorithm for the $K_{a,b}$-subgraph counting problem has success probability at most $1/\operatorname{polylog}(n)$, where the input is a random bipartite graph. As the positive result, we prove that there is an $n^{a+o(1)}$-time algorithm that counts the number of $K_{a,b}$-subgraphs in *any* graph if $a \geq 8$. As a consequence, we obatin the nearly-tight average-case complexity of the $K_{a,b}$-subgraph counting problem, that is, the $K_{a,b}$-subgraph counting problem admits a worst-case $n^{a+o(1)}$-time algorithm, while it does not admit any $n^{a-\epsilon}$-time algorithm with success probability more than $1 - 1/\operatorname{polylog}(n)$ for any constant $\epsilon > 0$ unless SETH is false.

1.1.2 Fine-grained hardness amplification

The hardness result in Section 1.1.1 was that, any $n^{a-\epsilon}$-time algorithm fails to solve the $K_{a,b}$-subgraph counting problem on a random birpartite graph with probability at least $1/\operatorname{polylog}(n)$. In other words, $1/\operatorname{polylog}(n)$-fraction of random bipartite graphs are hard to count $K_{a,b}$-subgraphs for any $n^{a-\epsilon}$-time algorithms. This result does not represent the hardness of random graphs since the fraction $1/\operatorname{polylog}(n)$ of hard instances is small: For example, there might exist, say, an $O(n^2)$-time algorithm that solves the counting problem on, say, 90% of instances. In Chapter 4, we handle with this issue.

In computational complexity theory, the existence of an average-case hard problem (i.e., a distributional problem with high fraction of hard instances) has gathered special attention. The main reason for this is that the average-case hardness of a problem can serve as the first step towards building secure cryptographic primitives such as pseudorandom generators and one-way functions [HILL99, NW94]. To obtain average-case hard problems, a considerable amount of effort has been devoted to *constructing* a "strongly" average-case hard problem from a "weakly" average-case hard problem. The technique of such construction is known as *hardness amplification*. See [GNW11] for detailed background.

In Chapter 4, we explore hardness amplification in the *fine-grained complexity* setting. Specifically, we consider the following research question.

Question 1.1.1. *Suppose that a function $f(x)$ is hard to compute on more than γ-fraction of inputs x for any $n^{c-\epsilon}$-time algorithm. Then, is there a function $g(y)$ such that computing $g(y)$ on more than $\gamma' \ll \gamma$ fraction of inputs y is impossible for any $n^{c-\epsilon}$-time algorithm?*

As a main result, we prove that, for a certain variant f of subgraph counting problem, we construct an another problem g that holds the property of Question 1.1.1. Consequently, we obtain a problem such that a $(1 - n^{o(1)})$-ratio of random bipartite graphs are hard for any $n^{a-\epsilon}$-time algorithms unless SETH fails. Moreover, the problem is closely related to the $K_{a,b}$ subgraph counting. At the heart of this result, we establish a general framework of *fine-grained hardness amplification* based on the *direct product theorem* [IW97, Tre03, GNW11, IJK09, IJKW10] and *Yao's XOR lemma* [Yao82, GNW11, Tre03, IW97].

1.2 Voting Process (Chapters 6 and 7)

Consider an undirected graph $G = (V, E)$ where each vertex $v \in V$ initially holds an opinion $\sigma_v \in \Sigma$ for a finite set Σ. A voting process is specified with a local updating rule: In each discrete time step, all vertices communicate with their neighboring vertices and simultaneously update their opinion according to the rule. The process aims to reach consensus, that is, a configuration where all vertices have the same opinion. Voting processes appear as simple mathematical models in a wide range of fields, e.g., social behavior, physical phenomena, and biological systems [MNT14, Lig85, AAB+11]. In distributed computing, voting processes are known to be a simple approach for the consensus problem [FLM86, GK10].

In this thesis, we focus on the setting of binary opinion (i.e., $\Sigma = \{0, 1\}$) and consider stochastic updating rule (i.e., each vertex can flip its private coins). Hence, a voting process can be seen as a

Markov chain on 2^V. An element of 2^V is called a *configuration*. Configurations where all vertices have the same opinion are called *consensus*. The main quantity of interest is the *consensus time*, which is the number of steps required to reach consensus.

1.2.1 Examples of voting processes

Pull voting. In *pull voting*, each vertex v picks up a neighbor uniformly at random. Then, the vertex v adopts the opinion of the selected neighbor.

Best-of-two. In *best-of-two* (a.k.a. 2-Choices), each vertex v picks up two neighbors u_1, u_2 uniformly at random (with replacement). If u_1 and u_2 have the same opinion, the vertex v adopt the opinion. Otherwise, v keeps its own opinion.

Best-of-three. In *best-of-three* (a.k.a. 3-Majority), each vertex v picks up three neighbors u_1, u_2, u_3 uniformly at random (with replacement). Then, the vertex v adopts the majority opinion among the three vertices u_1, u_2, u_3. Note that the tie does not occur since we consider the binary opinion setting.

We refer Section 5.2 to previous works of voting processes.

1.2.2 Voting processes on stochastic block model

In Chapter 6, we focus on best-of-two and best-of-three on the *stochastic block model*, a well-known random graph that forms multiple communities. This model has been well-explored in a wide range of fields, including biology [CY06, MPN+99], network analysis [BDLBH17, GZFA10], and machine learning [AS15, Abb18], where it serves as a benchmark for community detection algorithms. The study of the voting processes on the stochastic block model has a potential application in distributed community detection algorithms [BCM+18, BCN+17b, CNS19].

Definition 1.2.1 (Stochastic block model)**.** *Let $n \in \mathbb{N}$ and $p, q \in [0,1]$ with $q \leq p$ be parameters. The stochastic block model $\mathcal{G}(2n,p,q)$ is a random graph defined as follows:*

- *The vertex set is $V_1 \cup V_2$, where $|V_1| = |V_2| = n$ and $V_1 \cap V_2 = \emptyset$.*
- *Each pair $\{u,v\}$ of distinct vertices $u \in V_i$ and $v \in V_j$ forms an edge with probability θ_{ij}, independent of any other edges, where*

$$\theta_{ij} = \begin{cases} p & \text{if } i = j, \\ q & \text{otherwise.} \end{cases}$$

We denote by $G(2n,p,q)$ a graph sampled according to the distribution of $\mathcal{G}(2n,p,q)$.

The behavior of a voting process on $G(2n,p,q)$ depends on the parameters p and q. For example, if $p = q = 1$, then $G(2n,1,1)$ is the complete graph (i.e., the graph where all vertex pairs are connected by an edge). It is known that, on the n-vertex complete graph, best-of-two and best-of-three reach consensus within $\mathrm{O}(\log n)$ steps [DGM+11]. On the other hand, if $p = 1$ and $q = 0$, then $G(2n,1,0)$ consists of two disjoint complete graphs each of size n, meaning that, if one complete graph is in consensus with opinion 0 and the other does with opinion 1, then the voting process keeps the configuration and thus does not reach consensus. In Chapter 6, we prove that there is a threshold r^* depending on the voting process (best-of-two or best-of-three) such that, if $r > r^*$, then the corresponding voting process reaches consensus quickly, while the consensus time can be exponential if $r < r^*$.

1.2.3 General voting processes on expander graphs

In the context of the voting process, it is widely known that the behaviors of best-of-two and best-of-three are similar. For example, both processes on the n-vertex complete graph reach consensus within $\mathrm{O}(\log n)$ rounds with high probability [DGM+11, BCN+16]. Another example is the phase transition result on the stochastic block model presented in the previous part. However, the proofs of these results for best-of-two and best-of-three are obtained independently.

In Chapter 7, we introduce a *quasi-majority functional voting process* as a generalization of the best-of-two, best-of-three, and many other voting processes and consider the consensus time of it on *expander*

graphs. Intuitively speaking, expander graphs are sparse graphs that have strong connectivities (see Chapter 7 for the definition). Expander graphs gather special attention in the context of Markov chains on graphs, yielding a wide range of theoretical applications.

There is a line of works that studied best-of-two and best-of-three on expander graphs [CEOR13, CER14, CER+15]. Roughly speaking, best-of-two and best-of-three reach consensus within $O(\log n)$ rounds with high probability if the initial configuration has a sufficiently large bias (i.e., $||V_0| - |V_1||$ is large, where V_i is the set of vertices that holds opinion $i \in \{0, 1\}$ initially). As a main result, we prove that the quasi-majority functional voting on dense expander graphs reaches consensus within $O(\log n)$ rounds with high probability *without any assumption on the initial configuration*. This extends the previous work [DGM+11] that studies best-of-two on complete graphs. Moreover, we prove that, on a sparse expander graph, the consensus time of a quasi-majority functional voting is $O(\log n)$ if the initial configuration has a bias. This result generalizes previous works of best-of-two and best-of-three on expander graphs. Our result can be applied to obtain the consensus time of quasi-majority functional voting on random graphs such as the Erdős–Rényi graph. Moreover, the result provides an easy criterion for the practicality of voting processes: If someone may come up with a new voting process, then he can ensure the practicality of it on expander graphs by checking that whether the process is quasi-majority functional. Indeed, quasi-majority functional votings are so general that they contains many natural voting processes (see Section 7.1.5).

1.3 Random Walk on Growing Networks (Chapter 8)

Real-world networks change their shapes over time. Nevertheless, what is known about the analyses of algorithms on dynamic networks is quite limited, comparing with a wealth of knowledge on computations in static networks. In response to it, theoretical analyses of models and algorithms on dynamic networks recently attract high attentions, particularly in the context of network science and engineering, concerning such as information spreading [CST15], agreement [KO11], population protocol [MS18], random walks [Coo11] and other stochastic processes [Mic16, JAR16].

Random walk on a graph is a fundamental stochastic process: A walker on a vertex moves to a randomly picked neighbor at each discrete time step. A random walk is a simple and powerful tool in the wide range of computer science [Coo11, SMP15, AKL18, SZ19]. The cover time of a random walk is the time it takes for a walker to visit all vertices of the graph. The cover time is one of the fundamental quantities of a random walk, see e.g., [AKL+79, Ald83, Mat88, Fei95b, Fei95a, DS84, AF, LP17], and it is important with applications such as randomized search. Analyses of *random walks on dynamic graphs* have been actively developed in the context, where the cover time is a central issue [Coo11, CF03, AKL08, AKL18, DR14, YM16, LMS18, SZ19] (see Section 8.3 for more detail).

Those existing works, except for Cooper and Frieze [CF03], about random walks on dynamic networks are concerned only with networks over a static vertex set. However, the real networks change their vertex sets over time. Motivated by a new analysis technique, in Chapter 8, we investigate random walks on graphs with increasing the number of vertices. A dynamic vertex set causes some technical troubles: it is questionable if the "cover time," that is a natural quantity for a static vertex set, is also appropriate for a dynamic vertex set, and also it is hopeless, as Cooper and Frieze [CF03] revealed, to cover vertices beyond a constant ratio when the number of vertices constantly increases.

In view of this, we introduce in Chapter 8 a simple model of *growing graphs*, and presents an analysis of the number of vertices remaining *unvisited* by a random walk as a counterpart to the cover time of a random walk on a static vertex set.

1.4 Average Distance and Diameter (Chapter 9)

The study of the diameter of regular graphs is well-motivated in graph theory [BI73, HS60, Mv05, EFH80, Del85] and gathers special attention in high-performance computing (HPC) [EFH80, HS60, Mv05]. A central question is how to construct an n-vertex d-regular graph with the minimum possible diameter. In the literature of HPC, the performance of a parallel computer depends on the topology of the interconnection network, which is a graph where each vertex corresponds to a calculation node (e.g., CPU) and each edge does a link. If the interconnection network has low average distance and low diameter, data transmission on the network has a small number of hops. On the other hand, the degree of each node is limited due to physical constraints. Therefore, designing an interconnection network with low average distance and diameter under the degree constraint is important issue in HPC [EFH80,

HS60, Mv05]. Indeed, several researchers in the HPC area suggested using random graphs as network topologies (e.g., [SHPG12, KMA+12, KFI+16]).

In Chapter 9, we prove that asymptotically almost all dense regular graphs have the asymptotically optimal diameter. More precisely, we study the average distance and diameter of random regular graphs. A random regular graph $G_{n,d}$ is the graph sampled according to the uniform distribution $\mathcal{G}_{n,d}$ over the set of all n-vertex d-regular graphs. Although several researchers have studied the diameter of random graphs [Bol81, BdlV82, CL01, FR07, RW10, KL81], the diameter of dense random regular graphs is much less understood due to the lack of generation algorithm; The current known efficient algorithm can sample $G_{n,d}$ for $d = \mathrm{o}(\sqrt{n})$ [GW15]. We prove that the diameter of $G_{n,d}$ for $d = d(n) = (\beta + \mathrm{o}(1))n^{\alpha}$ with two arbitrary constants $\alpha \in (0, 1)$ and $\beta > 0$ is equal to $\lfloor \alpha^{-1} \rfloor + 1$ with probability asymptotically one as n tends to infinity. Since any n-vertex d-regular graph for $d(n) = (\beta + o(1))n^{\alpha}$ has diameter at least $\lceil \alpha^{-1} \rceil$ (we will see this fact in Chapter 9), our result implies that $G_{n,d}$ has the minimum diameter among all n-vertex d-regular graphs if $d = (\beta + o(1))n^{\alpha}$ with $\alpha^{-1} \notin \mathbb{N}$. Therefore, our result provides a theoretical guarantee for works suggesting network topologies based on random graphs (e.g., [SHPG12, KMA+12, KFI+16]).

1.5 Publications

The results of this thesis are based on the following publications.

- N. Shimizu, The average distance and the diameter of dense random regular graphs, *The Electronic Journal of Combinatorics*, 27(3), pp. 62:1–62:20, 2020. Preliminary version is in Proceedings of the 29th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA), pp. 1934–1944, 2018.
- N. Shimizu and T. Shiraga, Phase Transitions of Best-of-Two and Best-of-Three on Stochastic Block Models, *Random Structures and Algorithms*, 2020, to appear. Preliminary version is in *Proceedings of the 33rd International Symposium on Distributed Computing (DISC)*, pp. 32:1-32:17, 2019.
- N. Shimizu and T. Shiraga, Quasi-Majority Functional Voting on Expander Graphs, *In Proceedings of the 47th International Colloquium on Automata, Languages, and Programming (ICALP)*, pp. 97:1-97:19, 2020.
- S. Hirahara and N. Shimizu, Nearly Optimal Average-Case Complexity of Counting Bicliques Under SETH, *In Proceedings of the 32nd Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 2021, to appear.
- S. Kijima, N. Shimizu, and T. Shiraga, How Many Vertices Does a Random Walk Miss in a Network with Moderately Increasing the Number of Vertices?, *In Proceedings of the 32nd Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 2021, to appear.

Chapter 2

Preliminaries

2.1 General Notation

We denote by $\mathbb{N}$, $\mathbb{Z}$, and $\mathbb{R}$ the set of positive integers, integers, and real numbers, respectively. We use $\mathbb{Z}_{\geq 0}$ and $\mathbb{R}_{\geq 0}$ to denote the set of nonnegative integers and real numbers, respectively. For $k \in \mathbb{N}$, let $[k]$ be the set $\{1, \ldots, k\}$ of integers. For a finite set S and a positive integer $k \leq |S|$, let $\binom{S}{k} := \{\{x_1, \ldots, x_k\} \subseteq S : |\{x_1, \ldots, x_k\}| = k\}$ be the set of all subsets of size k and

$$(S)_k := \left\{ (x_1, \ldots, x_k) \in \underbrace{S \times \cdots \times S}_{k} : \{x_1, \ldots, x_k\} \in \binom{X}{k} \right\}$$

be the set of all ordered k-tuples.

Unless otherwise noted, $\log(x)$ stands for the logarithm of x to base e. We use $\mathbb{F}_q$ for the finite field of order q.

The probability that an event $\mathcal{E}$ holds is denoted by $\mathbf{Pr}[\mathcal{E}]$. For a random variable X, we denote by $\mathbf{E}[X]$ and $\mathbf{Var}[X]$ the expectation and variance of X, respectively. By $x \sim \mathcal{R}$ we mean that x is sampled according to the distribution $\mathcal{R}$. The binomial distribution of n trials with success probability p is denoted by $\mathrm{Bin}(n, p)$.

For a $p \in \mathbb{R}_{\geq 0} \cup \{\infty\}$ and a vector $x \in \mathbb{R}^n$, the ℓ^p-norm $\|x\|_p$ of x is defined as $\|x\|_p = \left(\sum_{i \in [n]} |x|^p\right)^{1/p}$.

Let $f, g \colon \mathbb{N} \to \mathbb{N}$ be functions. We follow the standard convention of the O-notation. We write $f(n) = \mathrm{O}(g(n))$ if there are constants $C > 0$ and $n_0 \in \mathbb{N}$ such that $f(n) \leq Cg(n)$ holds fol all $n \geq n_0$. We write $f(n) = \mathrm{o}(g(n))$ if, for any small constant $C > 0$, there is $n_0 \in \mathbb{N}$ such that $f(n) < Cg(n)$ holds for all $n \geq n_0$. We write $f(n) = \Omega(g(n))$ if $g(n) = \mathrm{O}(f(n))$, and write $f(n) = \omega(g(n))$ if $g(n) = \mathrm{o}(f(n))$. We write $f(n) = \Theta(g(n))$ if $f(n) = \mathrm{O}(g(n))$ and $f(n) = \Omega(g(n))$ hold. We sometimes use $\mathrm{polylog}(n)$ to denote $(\log n)^{\mathrm{O}(1)}$. We write $f(n) = \widetilde{O}(g(n))$ if $f(n) = \mathrm{O}(g(n)\,\mathrm{polylog}(n))$.

2.2 Graph

In this thesis, by graph we mean a simple undirected graph, that is, a pair (V, E) of a finite set V and set $E \subseteq \binom{V}{2}$ of unordered pairs of elements of V. Each element of V and E are called a *vertex* and *edge*, respectively. For a graph G, we denote by $V(G)$ and $E(G)$, respectively, the vertex set and the edge set of G. For a vertex v, let $N(v) = \{w \in V : \{v, w\} \in E(G)$ be the set of vertices adjacent to v. Note that $N(v)$ does not contain v. The *degree* $\deg(v)$ of a vertex $v \in V(G)$ is defined as $\deg(v) = |N(v)|$. For a vertex $v \in V$ and subset $S \subseteq V$, let $\deg_S(v) = |N(v) \cap S|$. A graph G is *d-regular* if the degrees of all vertices are equal to d. A graph H is a *subgraph* of G, denoted by $H \subseteq G$, if $V(H) \subseteq V(G)$ and $E(H) \subseteq E(G)$ hold. For two graph G and H, we say *G is isomorphic to H* if there is a bijection $\phi : V(G) \to V(H)$ such that $\{u, v\} \in E(G)$ if and only if $\{\phi(u), \phi(v)\} \in E(H)$ holds. For two graphs H and G, an *H-subgraph* of G is a subgraph S of G that is isormorphic to H.

A *graph property* $\mathcal{P}$ is the set of graphs that is invariant under the isomorphism, that is, if $G \in \mathcal{P}$ and H is isomorphic to G, then $H \in \mathcal{P}$. We say that *a graph G satisfies the graph property $\mathcal{P}$* if $G \in \mathcal{P}$.

2.2.1 Random graph

A *random graph model* is a distribution over *finite* graphs. A *random graph* is a graph drawn from a random graph model.

Definition 2.2.1 (Erdős–Rényi graph)**.** *The* Erdős–Rényi model*, denoted by $\mathcal{G}(n,p)$, is the distribution of n-vertex graphs satisfying*

$$\Pr_{G(n,p)\sim\mathcal{G}(n,p)}[G(n,p)=H]=p^{|E(H)|}(1-p)^{\binom{n}{2}-|E(H)|}$$

for any fixed n-vertex graph H. The graph $G(n,p)\sim\mathcal{G}(n,p)$ is called Erdős–Rényi graph.

We often consider the probability that a random graph satisfies a certain graph property $\mathcal{P}$. To state it more formally, consider a sequence $(\Omega_n,P_n,\mathcal{F}_n)_{n\in\mathbb{N}}$ of probability spaces and a sequence $(\mathcal{E}_n)_{n\in\mathbb{N}}$ of events (i.e., $\mathcal{E}_n\in\mathcal{F}_n$ for all $n\in\mathbb{N}$). We say that the event $\mathcal{E}_n$ holds *asymptotically almost surely* (a.a.s.) if $\mathbf{Pr}[\mathcal{E}_n]=1-\mathrm{o}(1)$. We say that the event $\mathcal{E}_n$ holds *with high probability* (w.h.p.) if $\mathbf{Pr}[\mathcal{E}_n \text{ holds}]\geq 1-n^{-\Omega(1)}$. Note that if $\mathcal{E}_n$ holds w.h.p., then $\mathcal{E}_n$ holds a.a.s. For example, it is easy to see that $G(n,p)$ does not contain any edge if $p=\mathrm{o}(n^{-2})$ a.a.s. See Section 2.5 for more details.

2.3 Computational Complexity

We regard a *problem* Π as a function from an input to the solution. The solution of Π for input x is denoted by $\Pi(x)$. A *decision problem* is a problem such that $\Pi(x)\in\{0,1\}$ for any input x. An instance x with $\Pi(x)=1$ (respectively, $\Pi(x)=0$) is called an *YES-instance* (respectively, *NO-instance*) Π. The *size* n of an input x is specified by the problem we consider (for example, if the input is a graph, n stands for the number of vertices).

A randomized algorithm A is said to *solve a problem* Π *in time* $T(n)$ if A runs in time $T(n)$ for any input x of size n and $\mathbf{Pr}_A[A(x)=\Pi(x)]\geq\frac{3}{4}$. Here, by $\mathbf{Pr}_A[\cdot]$ we mean that the probability is taken over the randomness of the algorithm A. Similarly, for an event $\mathcal{E}$ on $x\sim\mathcal{R}$, we sometimes denote by $\mathbf{Pr}_{x\sim\mathcal{R}}[\mathcal{E}]$ the probability of the event $\mathcal{E}$ holds, where the probability is over the choice of $x\sim\mathcal{R}$. We use these notations in order to clarify the randomness in the probability.

2.3.1 Average-case complexity

This thesis follows the common notion of average-case complexity (e.g., [BT06]). A *distributional problem* is a pair $(\Pi,\mathcal{D})$ of a problem Π and a family of distributions $\mathcal{D}=(\mathcal{D}_1,\mathcal{D}_2,\ldots)$, where each $\mathcal{D}_n$ denotes a distribution over inputs of size n. To simplify notations, we shall refer to (Π,D_n) rather than $(\Pi,(D_n)_{n\in\mathbb{N}})$. We say that *a (deterministic) algorithm A solves a distributional problem $(\Pi,\mathcal{D})$ with success probability δ* if, for every $n\in\mathbb{N}$, it holds that $\mathbf{Pr}[A(x)=\Pi(x)]\geq\delta$ where the probability is over the random choice of $x\sim\mathcal{D}_n$. The definition can be extended to a randomized algorithm:

Definition 2.3.1. *Let $(\Pi,\mathcal{D})$ be a distributional problem and $\delta\colon\mathbb{N}\to[0,1]$ be a function. We say that a randomized algorithm A* solves *$(\Pi,\mathcal{D})$ with success probability p if, for every $n\in\mathbb{N}$, $\mathbf{Pr}_{x\sim\mathcal{D}_n}[\mathbf{Pr}_A[A(x)=\Pi(x)]\geq\frac{3}{4}]\geq p$.*

2.4 Finite Markov Chain

We here briefly introduce other terminology for time-homogeneous Markov chains (cf. [LP17]).

Let V be a finite set. A *transition matrix* P over V is a matrix $P\in[0,1]^{V\times V}$ satisfying $\sum_{v\in V}P_{u,v}=1$ for any $u\in V$. A transition matrix P is *irreducible* if for any $u,v\in V$, there exists $t\in\mathbb{N}$ such that $(P^t)_{u,v}>0$, and is *apperiodic* if for any $v\in V$, $\mathrm{GCD}(\{t>0:(P^t)_{v,v}>0\})=1$ holds, where, for a set $S\subseteq\mathbb{N}$ of positive integers, $\mathrm{GCD}(S)$ denotes the greatest commond divisor of S. An irreducible and apperiodic P is said to be *ergodic*.

Let $\pi\in[0,1]^V$ denote the *stationary distribution* of P, that is, a probability distribution over V satisfying $\pi P=\pi$. It is well known that an ergodic P has a unique stationary distribution [LP17]. A transition matrix P is *reversible* if $\pi_uP_{u,v}=\pi_vP_{v,u}$ for any $u,v\in V$. Note that a transition matrix over V defines a Markov chain on V.

For ease of notation, we sometimes identify a matrix $P\in[0,1]^{V\times V}$ as a function $P\colon V\times V\to[0,1]$. Here, we denote $P(u,v)=P_{u,v}$, $P(u,S):=\sum_{v\in S}P(u,v)$, and $\pi(S):=\sum_{v\in S}\pi(v)$.

Example 2.4.1 (Simple random walk)**.** Let $G = (V, E)$ be a graph. Define a transition matrix $P \in [0,1]^{V \times V}$ as

$$P(u,v) = \begin{cases} \frac{1}{\deg(u)} & \text{if } \{u,v\} \in E(G), \\ 0 & \text{otherwise.} \end{cases} \tag{2.1}$$

The matrix P of (2.1) is known as the transition matrix of the *simple random walk* on G. Note that the simple random walk on G is not apperiodic if G is bipartite.

Example 2.4.2 (Simple lazy random walk)**.** A *simple lazy* random walk on an undirected graph G is given by

$$P(u,v) = \begin{cases} \frac{1}{2\deg(u)} & \text{if } \{u,v\} \in E(G), \\ \frac{1}{2} & \text{if } u = v \in V(G), \\ 0 & \text{otherwise.} \end{cases}$$

Note that the simple lazy random walk on G is apperiodic even if G is bipartite.

2.5 Basic Tools

Proposition 2.5.1 (Union Bound)**.** *For any countable set of events $\{\mathcal{E}_i\}_{i\in\Lambda}$,*

$$\mathbf{Pr}\left[\bigcup_{i\in\Lambda} \mathcal{E}_i\right] \leq \sum_{i\in\Lambda} \mathbf{Pr}[\mathcal{E}_i]$$

Proposition 2.5.2 (The Markov Inequality)**.** *Let X be a random variable that takes positive real numbers. Then, for any $a > 0$,*

$$\mathbf{Pr}[X \geq a] \leq \frac{\mathbf{E}[X]}{a}.$$

Example 2.5.3. Consider the Erdős–Rényi graph $G(n,p)$ for $p = o(n^{-2})$. Let $\mathcal{P}_{\text{empty}}$ denote the graph property of being an empty graph (i.e., $\mathcal{P}_{\text{empty}}$ is the set of graphs that does not contain any edge). Then, we can easily prove that $G(n,p)$ satisfies $\mathcal{P}_{\text{empty}}$ a.a.s. using the Markov inequality. To see this, let X be the random variable denoting the number of edges of $G(n,p)$. Then, $\mathbf{E}[X] = \binom{n}{2}p = o(1)$. By applying Proposition 2.5.2 with $a = 1$, we have $\mathbf{Pr}[G(n,p) \notin \mathcal{P}_{\text{empty}}] = \mathbf{Pr}[X \geq 1] = o(1)$. In other words, $G(n,p)$ a.a.s. satisfies $\mathcal{P}_{\text{empty}}$.

Proposition 2.5.4 (The Chebyshev inequality)**.** *Let X be a random variable such that $\mathbf{E}[X] < \infty$ and $\mathbf{Var}[X] < \infty$. Then, for any $t > 0$,*

$$\mathbf{Pr}[|X - \mathbf{E}[X]| \geq t] \leq \frac{\mathbf{Var}[X]}{t^2}.$$

Proposition 2.5.5 (The Chernoff bound; Theorem 1.10.1 and Theorem 1.10.5 of [DN20])**.** *Let $(X_i)_{i\in\mathbb{N}}$ be independent random variables taking values in $[0,1]$. Let $X = \sum_{i\in[n]} X_i$. Then the following hold:*

(i) for any $\delta \geq 0$,

$$\mathbf{Pr}[X \geq (1+\delta)\,\mathbf{E}[X]] \leq \exp\left(-\frac{\min\{\delta,\delta^2\}\,\mathbf{E}[X]}{3}\right).$$

(ii) for any $\delta \in [0,1]$,

$$\mathbf{Pr}[X \leq (1-\delta)\,\mathbf{E}[X]] \leq \exp\left(-\frac{\delta^2\,\mathbf{E}[X]}{2}\right).$$

Example 2.5.6. Let X be the degree of a fixed vertex v of $G(n,p)$. For $i \in V(G(n,p)) \setminus \{v\}$, let X_i be the binary random variable defined as $X_i = 1$ if $\{v,i\} \in E(G(n,p))$ and $X_i = 0$ otherwise. Then, the

random variables $(X_i)_{i\in V\setminus\{v\}}$ are independent and thus satisfy the condition of Proposition 2.5.5. If we set $\delta = 2\sqrt{\frac{\log n}{(n-1)p}}$, we obtain

$$\begin{aligned}\mathbf{Pr}[X \geq (1+\delta)(n-1)p] &\leq \exp(-4\log n/3) = n^{-4/3},\\ \mathbf{Pr}[X \leq (1-\delta)(n-1)p] &\leq \exp(-2\log n) = n^{-2}.\end{aligned}$$

By taking the union bound over all vertices, we have that

$$\mathbf{Pr}\left[\exists v \in V(G(n,p)) : |\deg(v) - (n-1)p| > \delta(n-1)p\right] \leq n\cdot(n^{-4/3}+n^{-2}) \leq 2n^{-1/3}.$$

In other words, if $d_{\min}$ and $d_{\max}$ are the minimum and maximum degree of $G(n,p)$, respectively, then it holds w.h.p. that

$$(n-1)p\left(1-2\sqrt{\frac{\log n}{(n-1)p}}\right) \leq d_{\min} \leq d_{\max} \leq (n-1)p\left(1+2\sqrt{\frac{\log n}{(n-1)p}}\right).$$

Note that this inequality is meaningful when $p = \omega\left(\frac{\log n}{n}\right)$

Proposition 2.5.7 (The inclusion-exclusion principle)**.** *Let $S_1,\ldots,S_k \subseteq E$ be subsets of a finite set E. Then,*

$$\left|\bigcup_{i\in[k]} S_i\right| = \sum_{I\in\subseteq[k]:I\neq\emptyset} (-1)^{|I|-1}\left|\bigcap_{i\in I} S_i\right|.$$

Chapter 3

Biclique Counting on Random Bipartite Graphs

3.1 Results

In this chapter, we consider the problem of counting the number of $K_{a,b}$-subgraphs on a natural distribution. To state it more formally, for a fixed graph H, let $\#\textsc{Emb}^{(H)}$ denote the problem that asks the number of embeddings of H in G for a given graph G. (An *embedding* of H in G is an injective homomorphism from H to G; see Section 3.3 for the formal definition.) The problem $\#\textsc{Emb}^{(H)}$ is equivalent to the H-subgraph counting problem: The number of H-subgraphs in G is equal to the number of embeddings of H in G divided by the number of automorphisms of H. Our main interest is the average-case complexity of $\#\textsc{Emb}^{(K_{a,b})}$ on random bipartite graphs. Specifically, we consider the following distribution.

Definition 3.1.1 (Random bipartite graph $\mathcal{K}_{a,b,n}$)**.** *For given parameters $a, b, n \in \mathbb{N}$, choose α, β uniformly at random from $[a]$ and $[b]$, respectively. Let $\mathcal{K}_{a,b,n}$ be the distribution of a random bipartite graph with $n\alpha$ left vertices and $n\beta$ right vertices, where each possible edge is included independently with probability $1/2$.*

The reader is referred to Section 2.3 for the notions of average-case complexity. The result of this chapter determines a threshold between worst- and average-case complexity of the distributional problem $(\#\textsc{Emb}^{(K_{a,b})}, \mathcal{K}_{a,b,n})$ under *rSETH*, a randomized variant of the strong exponential time hypothesis (SETH) of Impagliazzo, Paturi and Zane [IPZ01] (see Definition 3.3.5).

Theorem 3.1.2 (Worst- and Average-Case Complexity of Counting $K_{a,b}$)**.** *The following hold.*

- *For any constants $a \geq 8$ and b, there is an $n^{a+o(1)}$-time algorithm that solves $\#\textsc{Emb}^{(K_{a,b})}$ for any inputs.*
- *Under rSETH, for any constants $\epsilon > 0$ and $a \geq 3$, there is a constant $b = b(a, \epsilon)$ such that no $n^{a-\epsilon}$-time algorithm solves $(\#\textsc{Emb}^{(K_{a,b})}, \mathcal{K}_{a,b,n})$ with success probability greater than $1 - (1/\log n)^C$, where $C = C(a, b, \epsilon)$ is a constant depending only on a, b and ϵ.*

Theorem 3.1.2 is the first result that determines the nearly optimal average-case complexity of distributional graph problem under a widely investigated hypothesis. This result provides insight towards understanding the hardest instance of subgraph counting problems.

However, there is still an issue that the subgraph $K_{a,b}$ is not fixed in Theorem 3.1.2: The parameter b depends on ϵ. The problem of counting fixed $K_{a,b}$ might be more natural in the context of complexity theory. To cope with this issue, we consider the special case of $b = a$ and obtain the average-case hardness under *randomized Exponential Time Hypothesis (rETH)* (see Definition 3.3.4).

Theorem 3.1.3. *Under rETH, any $n^{o(a)}$-time algorithm solves $(\#\textsc{Emb}^{(K_a)}, \mathcal{K}_{a,a,n})$ with success probability no more than $(1 - (1/\log n)^C)$, where $C = C(a)$ is a constant that depends on a.*

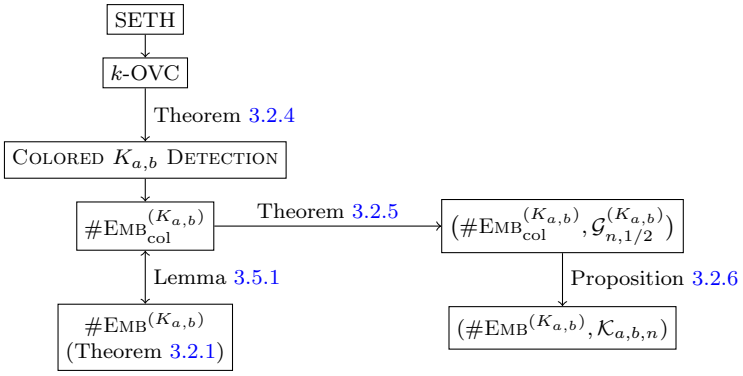


Figure 3.1: The organization of the proof of average-cane hardness of $\#\text{Emb}^{(K_{a,b})}$.

3.2 Overview of Our Techniques

In what follows, we briefly present the ingredients for our results while reviewing the literature. The overall outline of the proof of Theorem 3.1.2 is illustrated in Figure 3.1. In Section 3.2.1, we describe the worst-case complexity of $\#\text{Emb}^{(K_{a,b})}$ under SETH and ETH (see Definitions 3.3.2 and 3.3.3 for the definition). In Section 3.2.2, we present our idea for the worst-case-to-average-case reduction for subgraph counting problems. We review related results in Section 3.2.3.

3.2.1 Worst-case complexity of $K_{a,b}$-subgraph counting

Nearly Optimal Complexity of $\#\text{Emb}^{(K_{a,b})}$. Our first step is to determine the nearly optimal *worst-case* complexity for $\#\text{Emb}^{(K_{a,b})}$ under SETH by proving the following results.

Theorem 3.2.1. *For any constants $\epsilon > 0$ and $a \geq 3$, there exists a constant b such that one cannot solve $\#\text{Emb}^{(K_{a,b})}$ in time $\mathrm{O}(n^{a-\epsilon})$ unless SETH fails.*

Theorem 3.2.2. *If there exists an $n^{\mathrm{o}(a)}$-time algorithm that solves $\#\text{Emb}^{(K_{a,a})}$, then ETH fails.*

Theorem 3.2.3. *If $a \geq 8$, for any $\epsilon > 0$ and $b \in \mathbb{N}$, there is an algorithm that solves $\#\text{Emb}^{(K_{a,b})}$ in time $\mathrm{O}(bn^{a+\epsilon})$.*

We are not aware of previous results that determine the nearly optimal complexity of subgraph counting problems, while the fine-grained complexity of many natural problems, including the All-Pairs Shortest Paths, 3SUM, Orthogonal Vectors, and related problems, has been extensively explored in the research area of hardness in P [Wil15, LPW17].

SETH-hardness of $\#\text{Emb}^{(K_{a,b})}$. The notions in this part are defined in Section 3.3. Our key idea for showing Theorem 3.2.1 is to consider Colored $K_{a,b}$ Detection, which is defined as follows. Let K_n be the n-vertex complete graph. For a graph H, let $K_n \times H$ denote the tensor product. In general, the problem Colored H Detection is defined as follows. For a graph $G \subseteq K_n \times H$, each vertex $v = (u, i) \in V(G)$ is associated with a color $c(v) := i \in V(H)$. We say that an H-subgraph F of $G \subseteq K_n \times H$ is *colored* if F contains every colors from H. The problem Colored H Detection asks, given a pair (n, G) of $n \in \mathbb{N}$ and a graph $G \subseteq K_n \times H$, to decide whether G contains a colored H-subgraph.

Exploiting the fact that Colored $K_{a,b}$ Detection is more "structured" than $\#\text{Emb}^{(K_{a,b})}$, we first present a reduction from k-Orthogonal Vectors (k-OV) to Colored $K_{a,b}$ Detection for $k := a$. Since k-OV is known to be SETH-hard for any $k \geq 2$ [Wil15, Wil05, LPW17], this establishes SETH-hardness of Colored $K_{a,b}$ Detection:

Theorem 3.2.4. *For any constants $a \geq 2$ and $\epsilon > 0$, there exists a constant $b = b(a, \epsilon) \geq a$ such that* Colored $K_{a,b}$ Detection *cannot be solved in time $\mathrm{O}(m^{a-\epsilon})$ unless SETH fails, where m is the number of edges of the input graph.*

To complete the proof of Theorem 3.2.1, we reduce Colored $K_{a,b}$ Detection to $\#\text{Emb}^{(K_{a,b})}$ by using the inclusion-exclusion principle. This technique is well known in the literature of fixed-parameter complexity (see, e.g. [CM14, Cur18]). We will present the detail in Section 3.5.2.

ETH-hardness of #Emb$^{(K_{a,a})}$. The problem of finding a complete bipartite graph $K_{a,a}$ in a given graph has gathered special attention in parameterized complexity. Lin [Lin15, Lin18] proved that the problem is W[1]-hard when a is a parameter. His proof implies that the problem of finding $K_{a,a}$ does not admit any $n^{o(\sqrt{a})}$-time algorithm unless ETH fails. In particular, under ETH, any $n^{o(\sqrt{a})}$-time algorithm fails to solve #Emb$^{(K_{a,a})}$.

Theorem 3.2.2 improves this lower bound by ruling out an $n^{o(a)}$-time algorithm under ETH. A key idea behind this improvement is to take advantage of the structure of *counting* the number of embeddings: We first reduce the problem of finding a clique K_a of size a (which is known to be ETH-hard [CHKX06]) to Colored $K_{a,b}$ Detection, and then reduce it to #Emb$^{(K_{a,a})}$ by using the inclusion-exclusion principle. The latter reduction exploits the structure of counting.

3.2.2 Average-case complexity of subgraph counting problems

Compared to the worst-case hardness, the average-case hardness of subgraph counting problems has not been well understood until very recently [GR18a, BABB19, DLW20]. A breakthrough result of Boix-Adserà, Brennan, and Bresler [BABB19] shows that the worst-case and average-case complexities of counting k-cliques in an n-vertex Erdős–Rényi graph are equivalent up to a polylog(n)-factor. They left as an open question the extension of their results to other subgraph counting problems.

In this chapter, we investigate their open question under a different setting, which is one of our key insights. Specifically, let $G \times H$ be the *tensor product* of two graph G and H (see Section 3.3 for definition). For a fixed graph H, consider the problem #Emb$^{(H)}_{\mathrm{col}}$ of counting *color-preserving* embeddings of H to G that preserves colors (see Figure 3.2 for an illustration). Here, we say that an embedding ϕ *preserves* colors if $u = c(\phi(u))$ holds for any $u \in V(H)$, where $c : V(G) \to V(H)$ is the coloring of G. Note that #Emb$^{(H)}_{\mathrm{col}}(G)$ is the solution of #Emb$^{(H)}_{\mathrm{col}}$ for input G, that is, #Emb$^{(H)}_{\mathrm{col}}(G)$ is equal to the number of embeddings of H in G that preserves colors.

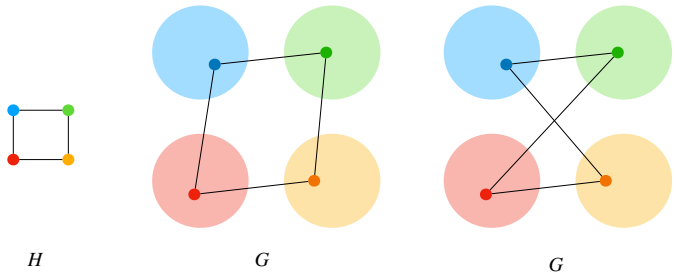


Figure 3.2: An example of color-preserving embedding of H to G for $G \subseteq K_n \times H$. Note that we do not consider the subgraph in right-hand side since $G \subseteq K_n \times H$ contains neither blue-orange nor green-red edges.

Let $\mathcal{G}^{(H)}_{n,1/2}$ be the distribution of graph where a graph $G \subseteq K_n \times H$ is generated by adding edges in $E(K_n \times H)$ independently with probability $1/2$. Based on the techniques of Boix-Adserà, Brennan, and Bresler [BABB19], we prove that #Emb$^{(H)}_{\mathrm{col}}$ is reducible to the distributional problem $(\#\mathrm{Emb}^{(H)}_{\mathrm{col}}, \mathcal{G}^{(H)}_{n,1/2})$.

Theorem 3.2.5 (Worst-case-to-average-case reduction for #Emb$^{(H)}_{\mathrm{col}}$)**.** *Let H be a fixed graph. Suppose that the distributional problem $(\#\mathrm{Emb}^{(H)}_{\mathrm{col}}, \mathcal{G}^{(H)}_{n,1/2})$ can be solved by a $T(n)$-time randomized algorithm A with success probability $1-\delta$, where $\delta = (\log n)^{-C}$ and $C = C_H$ is a sufficiently large constant depending on H. Then, there is a $T(n)\cdot\mathrm{polylog}(n)$-time randomized algorithm B that solves #Emb$^{(H)}_{\mathrm{col}}$ for any input with success probability $2/3$. Moreover, the number of oracle calls of A by B is at most $(\log n)^{\mathrm{O}(|E(H)|)}$.*

It should be noted that, Dalirrooyfard, Lincoln, and Williams [DLW20] proved the same result as Theorem 3.2.5 (their work is independent to us). See Section 3.2.3 for the detail. We then reduce $(\#\mathrm{Emb}^{(K_{a,b})}_{\mathrm{col}}, \mathcal{G}^{(K_{a,b})}_{n,1/2})$ to $(\#\mathrm{Emb}^{(K_{a,b})}, \mathcal{K}_{a,b,n})$ using the inclusion-exclusion principle, which is one of our technical contributions in this chapter.

Proposition 3.2.6. *Suppose that there is a $T(n)$-time randomized algorithm that solves the distributional problem $(\#\mathrm{Emb}^{(K_{a,b})}, \mathcal{K}_{a,b,n})$ with success probability $1-\delta$. Then, there is an $\mathrm{O}(ab2^{a+b}\cdot T(n))$-time randomized algorithm that solves $(\#\mathrm{Emb}^{(K_{a,b})}_{\mathrm{col}}, \mathcal{G}^{(K_{a,b})}_{n,1/2})$ with success probability $1-\mathrm{O}(ab2^{a+b}\delta)$.*

Combining Theorem 3.2.5 and Proposition 3.2.6, we obtain a worst-case-to-average-case reduction from $\#\textsc{Emb}^{(K_{a,b})}$ to $(\#\textsc{Emb}^{(K_{a,b})}, \mathcal{K}_{a,b,n})$.

Theorem 3.2.7. *Let $2 \leq a \leq b$ be arbitrary constants. Suppose that there is a $T(n)$-time randomized algorithm that solves $(\#\textsc{Emb}^{(K_{a,b})}, \mathcal{K}_{a,b,n})$ with success probability $1 - \delta$, where $\delta = (\log n)^{-C}$ and $C = C(a,b)$ is a sufficiently large constant. Then, there is a $T(n) \cdot \mathrm{polylog}(n)$-time randomized algorithm that solves $\#\textsc{Emb}^{(K_{a,b})}$ for any input with success probability $2/3$.*

Theorem 3.2.7 is of interest in its own right; we emphasize that a and b can be chosen arbitrarily unlike Theorem 3.2.1 (i.e., the SETH-hardness of $\#\textsc{Emb}^{(K_{a,b})}$). In the context of subgraph counting, counting $K_{2,2}$ (i.e., 4-cycle) subgraphs in a graph on n vertices with m edges attracts particular interest: The current fastest counting algorithm runs in time $O(n^\omega)$ or $O(m^{1.48})$ [AYZ97], whereas finding a $K_{2,2}$ can be done in time $O(n^2)$ [YZ97] or $O(m^{1.41})$ [AYZ97]. A central question in this context is whether we can beat the $O(n^\omega)$-time algorithm for the $K_{2,2}$-counting problem. The worst-case-to-average-case reduction given in Theorem 3.2.7 indicates that a random bipartite graph is essentially the hardest distribution for the $K_{2,2}$-counting problem.

3.2.3 Related work

Worst-case complexity of subgraph counting

The problem $\#\textsc{Emb}^{(H)}$ is a fundamental task in the context of graph algorithms. For a general subgraph H, we can solve $\#\textsc{Emb}^{(H)}$ in time $f(k) \cdot n^{(0.174+o(1))\ell}$ for some function $f(\cdot)$, where k and ℓ are the number of vertices and edges of H, respectively [CDM17]. If H has some nice structural property (e.g., small treewidth), several faster algorithms are known (see [Cur18] and the references therein). However, to the best of our knowledge, there is no previous result that precisely determines the complexity of counting subgraphs. Chen, Huang, Kanj, and Xia [CHKX06] proved that one cannot find a k-clique in a given graph in time $f(k) \cdot n^{o(k)}$ for any function $f(\cdot)$ unless ETH fails. The current fastest algorithm was given by Nešetřil and Poljak [NP85], who presented an $O(n^{\omega\lceil k/3 \rceil})$-time algorithm that counts the number of k-cliques in a given n-vertex graph. Here, $\omega < 2.373$ is the square matrix multiplication exponent [Gal14, Wil12, AW21]. Lincoln, Williams, and Williams [LWW18] imposed the assumption that detecting a k-clique in an n-vertex graphs requires time $n^{\omega k/3-o(1)}$ and then derived a super-linear lower bound for the shortest cycle problem. However, the precise value of ω is currently not known, and, as a consequence, the precise time complexity of counting k-cliques is not well understood.

Worst-case complexity of biclique counting

We mention in passing some algorithmic results concerned with finding or counting bicliques. The results below consider the case where a and b are given as input. Binkele-Raible, Fernau, Gaspers, and Liedloff [BRFGL10] proved that, for given a, b and a graph G, one can find a $K_{a,b}$ subgraph in G in time $O(1.6914^n)$. Couturier and Kratsch [CK12] gave an $O(1.6107^n)$-time algorithm for $\#\textsc{Emb}^{(K_{a,b})}$. They also provided an $O(1.2691^n)$-time counting algorithm that works on bipartite graphs. It is known that the number of distinct maximal induced biclique subgraphs in any n-vertex graph is $O(3^{n/3}) = O(1.442^n)$ [GKL12]. If a given graph is bipartite, one can solve $\#\textsc{Emb}^{(K_{a,b})}$ by enumerating all maximal $K_{a,b}$ subgraphs using a polynomial delay algorithm [MU04]. Kutzkov [Kut12] presented an $O(1.2491^n)$-time counting algorithm, which is currently the fastest one. If $a \leq b$ are small, we can solve $\#\textsc{Emb}^{(K_{a,b})}$ in time $O(n^{a+1})$ by enumerating all size-a vertex subsets. If $a = 2$, we can solve $\#\textsc{Emb}^{(K_{2,b})}$ in time $O(n^\omega)$ by computing A^2, where $A \in \{0,1\}^{n \times n}$ is the adjacency matrix of a given graph.

Finding a $K_{a,a}$-subgraph in a given graph is NP-complete if a is given as input [GJ79]. The parameterized complexity of finding a $K_{a,a}$-subgraph (parameterized by a) has gathered special attention. Lin [Lin18] proved the W[1]-hardness of the $K_{a,a}$-finding problem parameterized by a. Moreover, his proof implies that, assuming ETH, one cannot find a $K_{a,a}$-subgraph in time $n^{o(\sqrt{a})}$. However, it still remains open whether ETH rules out an $n^{o(a)}$-time algorithm for the $K_{a,a}$-subgraph finding problem. Theorem 3.2.2 rules out an $n^{o(a)}$-time algorithm for the counting variant $\#\textsc{Emb}^{(K_{a,a})}$ under ETH.

Fine-grained complexity

A standard criterion of the tractability of a problem is whether the problem can be solved in polynomial-time. This viewpoint of tractability is called *coarse-grained complexity*. The theory of NP-hardness has

established superpolynomial lower bounds of time complexity for several computational tasks based on the conjecture that some NP problem does not admit a polynomial-time algorithm.

In *fine-grained complexity*, we consider problems that can be solved in time $T(n)$ but no current-known $T(n)^{1-\epsilon}$-time algorithm solves it for any constant $\epsilon > 0$. Under well-known conjectures (e.g., SETH, APSP conjecture, 3-SUM conjecture), lower bounds for several problems (even for polynomial-time solvable problems) have been investigated. See [Wil15] for further details.

Average-case complexity in P

The distributional problem $(\text{Colored } H \text{ Detection}, \mathcal{G}^{(H)}_{n,1/2})$ has been studied in the literature of average-case circuit complexity of the subgraph isomorphism problem (cf. Rossman [Ros18]).

In a pioneering work of Ball, Rosen, Sabin, and Vasudevan [BRSV17], they initiated the study of average-case complexity in the context of fine-grained complexity. Ball et al. [BRSV17] and their subsequent work [BRSV18] constructed average-case hard problems by encoding worst-case problems by a low-degree polynomial over a large finite field. Based on techniques of random self-reducibility (e.g. [CPS99]), they explored the average-case hardness of the evaluation of this polynomial under the worst-case assumptions including the Orthogonal Vector Conjecture, APSP Conjecture, and 3SUM Conjecture, recent hot conjectures in the study of hardness in P [LPW17, Wil15]. Their work is motivated by the construction of PoW systems. Due to the construction, their average-case problems are artificial.

Goldreich and Rothblum [GR18a] studied the average-case complexity of $\#\text{Emb}^{(K_k)}$ for a constant k. They presented a simple distribution over $\widetilde{O}(n)$-vertex graphs on which it is hard to count the number of k-cliques with a success probability better than $3/4$. The distribution is constructed by a gadget reduction, and it is somewhat artificial. The key idea of their reduction is to consider counting weighted cliques: The input graph has node and edge weights in $\mathbb{F}_q$, and the task is to compute the sum of all weights of clique subgraphs. The weight of a clique is defined as the product of all node weights and edge weights contained in the clique. They represented this counting problem as a low-degree polynomial $P : \mathbb{F}_q^{n\times n} \to \mathbb{F}_q$ and used polynomial interpolation to reduce evaluating P to computing $P(r)$, where $r \sim \text{Unif}(\mathbb{F}_q^{n\times n})$. Combining the Chinese Reminder Theorem, a vertex-blowing-up technique and unifying multiple instances into one instance, they further reduced evaluating $P(r)$ to solving $\#\text{Emb}^{(K_k)}$ in a specific random graph. Their result has an error tolerance of constant probability. However, the blowing-up technique and unifying instances yielded an artificial random graph distribution.

The proof of Theorem 3.2.5 is based on techniques of Boix-Adserà, Brennan, and Bresler [BABB19], who reduced $\#\text{Emb}^{(K_k)}$ to $(\#\text{Emb}^{(K_k)}, G(n,p))$. The reduction runs in time $p^{-1}n^2 \operatorname{polylog} n$. Here, the error probability of the average-case solver is assumed to be at most $(\log n)^{-C}$ for a sufficiently large constant $C = C(k)$. They also presented a parity variant of $\#\text{Emb}^{(K_k)}$ and obtained a worst-case-to-average-case reduction with a better error tolerance.

Very recently, independently to this thesis, Dalirrooyfard, Lincoln, and Williams [DLW20] reduced $\#\text{Emb}^{(H)}_{\text{col}}$ to the distributional problem $(\#\text{Emb}^{(H)}, \mathcal{G}_{n,p})$ for a constant p. They first proved the same result as Theorem 3.2.5 and then reduced $(\#\text{Emb}^{(H)}_{\text{col}}, \mathcal{G}^{(H)}_{n,1/2})$ to $(\#\text{Emb}^{(H)}_{\text{col}}, \mathcal{G}_{n,p})$.

3.2.4 Organization

In Section 3.4, we present the worst-case-to-average-case reduction for $\#\text{Emb}^{(H)}_{\text{col}}$. In Section 3.5, we investigate the worst-case complexity of $\#\text{Emb}^{(K_{a,b})}$.

Here is the organization of the proof of our main results.

Theorem 3.1.2. The first statement follows from Theorems 3.2.1 and 3.2.7. We can obtain Theorem 3.2.7 by combining Theorem 3.2.5 and Proposition 3.2.6. See Section 3.4 and Section 3.5.5 for the proofs of Theorem 3.2.5 and Proposition 3.2.6, respectively. The second statement is equivalent to Theorem 3.2.3, which is shown in Section 3.5.4.

Theorem 3.1.3. This result follows from Theorems 3.2.2 and 3.2.7. Theorem 3.2.2 is shown in Section 3.5.6.

3.3 Preliminaries

Our computational model is the $O(\log n)$-Word RAM model. As a consequence, we assume that any field operation can be done in constant-time if the underlying field is $\mathbb{F}_q$ with $q = n^{O(1)}$ (n is specified by the problem).

For simplicity, we sometimes use uv to abbreviate an edge $\{u, v\}$. We identify a graph G with a vector $x_G \in \{0,1\}^{E(H)}$ by regarding x_G as the edge indicator of G.

3.3.1 Subgraph problem

For two graphs G and H, a mapping $\phi : V(H) \to V(G)$ is *homomorphism* from H to G if $\{\phi(u), \phi(v)\} \in E(G)$ whenever $\{u, v\} \in E(H)$. An *embedding* is an injective homomorphism. Let $\#\textsc{Emb}^{(H)}(G)$ be the number of embeddings from H to G. For a fixed graph H, we consider the problem $\#\textsc{Emb}^{(H)}$ of computing $\#\textsc{Emb}^{(H)}(G)$ for an input graph G.

The *tensor product* $X \times Y$ of two graphs X and Y is a graph defined as $V(X \times Y) = V(X) \times V(Y)$ and $\{(x_1, y_1), (x_2, y_2)\} \in E(X \times Y)$ if and only if $\{x_1, x_2\} \in E(X)$ and $\{y_1, y_2\} \in E(Y)$.

For a fixed graph H, a graph G is *H-colored* if $G \subseteq K_n \times H$ for some n. Let G be an H-colored graph $G \subseteq K_n \times H$. A vertex $v \in V(G)$ is associated with a *color* $c(v) \in V(H)$. Formally, if $v = (a, i) \in V(G) \subseteq V(K_n) \times V(H)$, then $c(v) = i$. An embedding $\phi\colon V(H) \to V(G)$ of H to G *preserves color* if $c(\phi(i)) = i$. For a fixed graph H, we consider the problem $\#\textsc{Emb}^{(H)}_{\text{col}}$ of counting the number of color-preserving embeddings of H to G for a given H-colored graph H. See Figure 3.2 for an illustration.

In Colored H Detection, we are asked, given an H-colored graph G, to decide whether there is a color-preserving embedding of H to G. In other words, Colored H Detection is the decision problem that asks whether $\#\textsc{Emb}^{(H)}_{\text{col}}(G) > 0$ for a given H-colored graph.

For a fixed graph H, let $G^{(H)}_{n,1/2} \subseteq K_n \times H$ be a random subgraph of $K_n \times H$ such that each edge $e \in E(K_n \times H)$ is included independently with probability $1/2$. The distribution of $G^{(H)}_{n,1/2}$ is denoted by $\mathcal{G}^{(H)}_{n,1/2}$. For a graph H and a finite set S, let $\mathcal{U}^{(H)}_n(S)$ be the uniform distribution over $S^{E(K_n \times H)}$. Note that $\mathcal{U}^{(H)}_n(\{0,1\})$ is equivalent to $\mathcal{G}^{(H)}_{n,1/2}$.

3.3.2 ETH and SETH

A *Boolean function* ϕ is a function $\phi : \{0,1\}^n \to \{0,1\}$ for some n. Let $x_1, \ldots, x_n$ denote the variables of ϕ. A Boolean function ϕ is a *k-CNF formula* if ϕ can be written as

$$\phi = \bigwedge_{i \in [m]} C_i$$

for some m, where each C_i is of the form $C_i = \bigvee_{j \in S_i} \ell_j$ for literals $\ell_j \in \{x_j, \overline{x_j}\}$ and some $S_i \subseteq [n]$ satisfying $|S_i| \leq k$. Each C_i is called *clause* of ϕ.

Definition 3.3.1 (k-SAT)**.** *In k-SAT, we are given a k-CNF ϕ and are asked to decide the existence of a satifying assignment of ϕ. That is, if $x_1, \ldots, x_n$ are the variables of ϕ, then our task is to decide whether there is an assignment $(b_1, \ldots, b_n) \in \{0,1\}^n$ such that $\phi(b_1, \ldots, b_n) = 1$.*

The problem k-SAT is a classical NP-complete problem. Despite a long line of works, no polynomial-time algorithm for k-SAT is known. There are many algorithms that solves k-SAT in time of the form $2^{(1-\frac{c}{k})n}$, where the constant c depends on the algorithm. In the current fastest (randomized) algorithm of Paturi, Pudlák, Saks, and Zane [PPSZ05], $c \approx 1.64$. In the special case of $k = 3$, slightly faster algorithms are known. Recently, Hansen, Kaplan, Zamir, and Zwick [HKZZ19] proved that there is a 1.307^n-time randomized algorithm that solves 3-SAT.

Definition 3.3.2 (Exponential time hypothesis (ETH); Impagliazzo and Paturi [IP01])**.** *There is an absolute constant $\delta > 0$ such that any $(1+\delta)^n$-time deterministic algorithm cannot solve* 3-SAT*.*

Definition 3.3.3 (Strong exponential time hypothesis (SETH); Impagliazzo, Paturi, and Zane [IPZ01])**.** *For any $\epsilon > 0$, there is $k \geq 3$ such that any $(2-\epsilon)^n$-time deterministic algorithm cannot solve k-SAT.*

In this chapter, we consider a randomized variant of ETH and SETH that exclude randomized algorithms.

Definition 3.3.4 (Randomized exponential time hypothesis (rETH))**.** *There is an absolute constant $\delta > 0$ such that any $(1+\delta)^n$-time randomized algorithm cannot solve* 3-SAT*.*

Definition 3.3.5 (Randomized strong exponential time hypothesis (rSETH))**.** *For any $\epsilon > 0$, there is $k \geq 3$ such that any $(2-\epsilon)^n$-time randomized algorithm cannot solve k-*SAT*.*

3.4 Average-Case Complexity of $\#\mathrm{Emb}^{(H)}_{\mathrm{col}}$

In this section, we present a proof of Theorem 3.2.5, that is, we reduce $\#\textsc{Emb}^{(H)}_{\mathrm{col}}$ to $(\#\textsc{Emb}^{(H)}_{\mathrm{col}}, \mathcal{G}^{(H)}_{n,1/2})$. For ease of notation, we use $z[i]$ to denote z_i for a vector $z \in Z^I$ and $i \in I$.

For a fixed graph H and a prime $q > n^{|V(H)|}$, let $\mathsf{EMBCOL}_{n,H,q} : \mathbb{F}_q^{E(K_n \times H)} \to \mathbb{F}_q$ be a polynomial defined as

$$\mathsf{EMBCOL}_{n,H,q}(x) = \sum_{\substack{v_1,\dots,v_k \in V(K_n \times H) \\ c(v_k)=k \ (\forall k)}} \prod_{\{i,j\} \in E(H)} x[v_i v_j]. \tag{3.1}$$

Suppose $x \in \{0,1\}^{E(K_n \times H)}$ is the edge indicator of a graph $G \subseteq K_n \times H$. Then $\mathsf{EMBCOL}_{n,H,q}(x) = \#\textsc{Emb}^{(H)}_{\mathrm{col}}(G) \bmod q = \#\textsc{Emb}^{(H)}_{\mathrm{col}}(G)$ as $q > n^{|V(H)|}$. We sometimes identify $\mathbb{F}_q$ with the set $\{0,\dots,q-1\}$. The proof of Theorem 3.2.5 consists of two steps.

3.4.1 Step 1: Random self-reducibility of $\mathsf{EMBCOL}_{n,H,q}(\cdot)$

First, we reduce evaluating $\mathsf{EMBCOL}_{n,H,q}(x)$ for a given x to solving $(\mathsf{EMBCOL}_{n,H,q}(\cdot), \mathcal{U}^{(H)}_n(\mathbb{F}_q))$ for a large prime $q > n^{|V(H)|}$. Note that we can obtain such a prime q as follows. Sample a random integer r from $\{n^{|V(H)|}, n^{|V(H)|}+1, \dots, 2n^{|V(H)|}\}$ and then run the primality test for r (according to the Prime Number Theorem, r is prime with probability $\Omega(1/\log n)$.)

The following is well known in the context result of random self-reducibility. A precise estimation of the running time was given by [BABB19, BRSV17].

Lemma 3.4.1 (Essentially given in Lemma 3.2 of [BRSV17])**.** *Let $P : \mathbb{F}_q^N \to \mathbb{F}_q$ be a multivariate polynomial of degree d for a prime $q > 12d$. Suppose that there is a $T(N,q,d)$-time algorithm A satisfying*

$$\Pr_{x \sim \mathrm{Unif}(\mathbb{F}_q^N)}[A(x) = P(x)] \geq 1 - \delta,$$

where $\delta \in (0, 1/3)$. Then, there is a randomized algorithm B that computes $P(y)$ on input $y \in \mathbb{F}_q^N$ with probability $2/3$ in time $\mathrm{O}(Nd^2(\log q)^2 + d^3 + dT(N,q,d))$.

Proof sketch. Ball et al. [BRSV17] proved this result under the condition that $d > 9$. Boix-Adserà, Brennan, and Bresler [BABB19] obtained the same result for a prime power $q > 12d$ (under the same condition) by the same way. The common idea is to invoke the well-known local decoding of the Reed-Muller code (see, e.g., [Lip91, GS92]). In this chapter, we just modify a parameter appeared in their proof to remove the degree condition. We briefly describe the algorithm and refer to the full version of [BRSV17] for the analysis.

For a given $y \in \mathbb{F}_q^N$, sample two random vectors $z_1, z_2 \sim \mathrm{Unif}(\mathbb{F}_q^N)$ independently, and consider the univariate function $f(t) := y + z_1 t + z_2 t^2$. Note that our task is to compute $f(0)$. Set $m = 100d$ (the authors of [BRSV17] set $m = 12d$). Use the oracle algorithm A and compute $A(f(1)), \dots, A(f(m))$. By the Berlekamp–Welch decoding [BW86], obtain a polynomial $\hat{f}$ and output $\hat{f}(0)$. □

By applying this result to our setting, we obtain the following.

Corollary 3.4.2. *For a fixed graph H and a prime $n^{|V(H)|} < q < 2n^{|V(H)|}$, let $\mathsf{EMBCOL}_{n,H,q}(\cdot)$ be the polynomial given in (3.1). Suppose that there is a $T(n)$-time algorithm A satisfying*

$$\Pr_{x \sim \mathcal{U}^{(H)}_n(\mathbb{F}_q)}[A(x) = \mathsf{EMBCOL}_{n,H,q}(x)] \geq \frac{2}{3}.$$

Then, there is a randomized algorithm B that computes $\mathsf{EMBCOL}_{n,H,q}(y)$ on input $y \in \mathbb{F}_q^{E(K_n \times H)}$ with success probability $2/3$ in time $\mathrm{O}(n^2(\log n)^2 + T(n))$.

3.4.2 Step 2: Reduce $\mathsf{EMBCOL}_{n,H,q}(\mathcal{U}_n^{(H)}(\mathbb{F}_q))$ to $\mathsf{EMBCOL}_{n,H,q}(\mathcal{U}_n^{(H)}(\{0,1\}))$

We reduce the problem of computing $\mathsf{EMBCOL}_{n,H,q}(\cdot)$ over the distribution $\mathcal{U}_n^{(H)}(\mathbb{F}_q)$ to that over $\mathcal{U}_n^{(H)}(\{0,1\})$ based on the binary extension technique of [BABB19]. Observe that the distributional problem $(\mathsf{EMBCOL}_{n,H,q}(\cdot), \mathcal{U}_n^{(H)}(\{0,1\}))$ is equivalent to $(\#\textsc{Emb}_{\text{col}}^{(H)}, \mathcal{G}_{n,1/2}^{(H)})$ if $q > n^{|V(H)|}$.

Lemma 3.4.3. *Let H be a fixed graph and q be a prime satisfying $n^{|V(H)|} < q < 2n^{|V(H)|}$. Suppose there is a $T(n)$-time randomized algorithm A satisfying*

$$\Pr_{x\sim\mathcal{U}_n^{(H)}(\{0,1\})}\left[\Pr_A\left[A(x) = \mathsf{EMBCOL}_{n,H,q}(x)\right] \geq \frac{2}{3}\right] \geq 1-\delta,$$

where $\delta = (\log n)^{-C}$ for a sufficiently large constant $C = C_H > 0$ that depends on H.

Then, there is a $T(n)\cdot \operatorname{polylog} n$-time randomized algorithm B satisfying

$$\Pr_{x\sim\mathcal{U}_n^{(H)}(\mathbb{F}_q)}\left[\Pr_B\left[B(x) = \mathsf{EMBCOL}_{n,H,q}(x)\right] > \frac{2}{3}\right] > \frac{2}{3}.$$

Note that Theorem 3.2.5 follows from Corollary 3.4.2 and Lemma 3.4.3.

Observation. Suppose that, for each $uv \in E(K_n \times H)$, $x[uv] \in \mathbb{F}_q$ can be rewritten as

$$x[uv] = \sum_{l=0}^{t-1} 2^l \cdot z^{(l)}[uv] \bmod q \tag{3.2}$$

for some binary variables $z^{(0)}[uv], \ldots, z^{(t-1)}[uv] \in \{0,1\}$. Here, t is some large integer that will be specified later. Then, we obtain

$$\begin{aligned}
\mathsf{EMBCOL}_{n,H,q}(x) &= \sum_{\substack{v_1,\ldots,v_k\in V(G)\\ c(v_i)=i\ (\forall i)}} \prod_{ij\in E(H)} \sum_{l=0}^{t-1} 2^l \cdot z^{(l)}[v_iv_j] \\
&= \sum_{\substack{v_1,\ldots,v_k\in V(G)\\ c(v_i)=i\ (\forall i)}} \sum_{a\in\{0,\ldots,t-1\}^{E(H)}} \prod_{ij\in E(H)} 2^{a[ij]} \cdot z^{(a[ij])}[v_iv_j] \\
&= \sum_{a\in\{0,\ldots,t-1\}^{E(H)}} 2^{\sum_{e\in E(H)} a[e]} \sum_{\substack{v_1,\ldots,v_k\in V(G)\\ c(v_i)\in i\ (\forall i)}} \prod_{ij\in E(H)} z^{(a[ij])}[v_iv_j]. \\
&= \sum_{a\in\{0,\ldots,t-1\}^{E(H)}} 2^{\sum_{e\in E(H)} a[e]} \cdot \mathsf{EMBCOL}_{n,H,q}(\chi^{(a)}).
\end{aligned} \tag{3.3}$$

Here, we define $\chi^{(a)}[uv] := z^{(a[c(u)c(v)])}[uv] \in \{0,1\}$ for each $uv \in E(K_n \times H)$.

Thus, our goal is to sample z such that the distribution of $z^{(a)}$ is closed to $\mathcal{G}_{n,1/2}^{(H)}$ for each $a \in \{0,\ldots,t-1\}^{E(H)}$. In this chapter, we invoke a special case of Lemma 4.3 of [BABB19] and improve the running time of a sampling procedure.

Lemma 3.4.4. *Let $q > 2$ be a prime and t be some integer. For each $x \in \mathbb{F}_q$, let $M_x := \{m \in \{0,\ldots,2^t-1\} : m \bmod q = x\}$ and $Y_x \sim \mathrm{Unif}(M_x)$ be a random variable. Let $\mathcal{Y}_R$ be the distribution of Y_R for $R \sim \mathrm{Unif}(\mathbb{F}_q)$. Then, the following hold.*

1. *$d_{\mathrm{TV}}(\mathcal{Y}_R, \mathrm{Unif}(\{0,\ldots,2^t-1\})) \leq Cq/2^t$ for some absolute constant C.*
2. *For any given $x \in \mathbb{F}_q$, we can sample Y_x in time $\mathrm{O}(t)$.*

Corollary 3.4.5. *Let t be some integer. Let $Z_0,\ldots,Z_{t-1} \sim \mathrm{Unif}(\{0,1\})$ be i.i.d. random variables. Then, for any given $x \in \mathbb{F}_q$, we can sample t random variables $z_0,\ldots,z_{t-1}$ satisfying the following in time $\mathrm{O}(t)$.*

1. *It holds that $\sum_{i=0}^{t-1} 2^i \cdot z_i \bmod q = x$.*

2. *The distribution of* $(z_0, \dots, z_{t-1})$ *when* x *is sampled from* $\mathrm{Unif}(\mathbb{F}_q)$ *is of total variation distance at most* $\mathrm{O}(q/2^t)$ *from the uniform distribution* $(Z_0, \dots, Z_{t-1})$.

Proof. For a given $x \in \mathbb{F}_q$, let $z_0, \dots, z_{t-1}$ be the binary expansion of Y_x of Lemma 3.4.4. Then, $Y_x = \sum_{i=0}^{t-1} 2^i \cdot z_i = x \pmod q$ by the definition of Y_x. Let $Y := \sum_{i=0}^{t-1} 2^i \cdot Z_i \sim \mathrm{Unif}(\{0, \dots, 2^t - 1\})$. Let $f : \{0, \dots, 2^t - 1\} \to \{0,1\}^t$ denote the function that maps $y \in \{0, \dots, 2^t - 1\}$ to the binary representation of y. Note that f is a bijection and $f(Y_x) = (z_0, \dots, z_{t-1})$ holds. Then, from Lemma 3.4.4, for any $A \subseteq \{0,1\}^t$, we have

$$\begin{aligned} |\mathbf{Pr}[(z_0, \dots, z_{t-1}) \in A] - \mathbf{Pr}[(Z_0, \dots, Z_{t-1}) \in A]| &= |\mathbf{Pr}[Y_x \in f^{-1}(A)] - \mathbf{Pr}[Y \in f^{-1}(A)]| \\ &= \mathrm{O}(q/2^t). \end{aligned}$$

This implies the statement 2 of Corollary 3.4.5. □

Remark 3.4.6. Boix-Adserà, Brennan, and Bresler [BABB19] considered the general case of $Z_i \sim \mathrm{Ber}(c_i)$, where $\mathrm{Ber}(c_i)$ is the Bernouli random variable with success probability c_i. Roughly speaking, for some $t = \Theta(c^{-1}(1-c)^{-1}\log(q/\epsilon^2)\log q)$, they proved (1) $d_{\mathrm{TV}}(\mathcal{L}(Y), \mathcal{L}(Y_R)) \le \epsilon$, and (2) For any given $x \in \mathbb{F}_q$, Y_x can be sampled in time $\mathrm{O}(tq)$. Since $q > n^{V(H)}$, the sampling of Y_x cannot be applied directly due to the running time $\mathrm{O}(tq)$. To avoid the large running time, Boix-Adserà, Brennan, and Bresler [BABB19] used the Chinese Reminder Theorem to reduce computing $\mathsf{EMBCOL}_{n,H,q}(\cdot)$ to the computing $\mathsf{EMBCOL}_{n,H,q_1}(\cdot), \dots, \mathsf{EMBCOL}_{n,H,q_m}(\cdot)$, where $q_1, \dots, q_m$ are small primes. In Lemma 3.4.4, we focus on the special case of $c_i = 1/2$ and improve the running time of sampling Y_x.

We will present the proof of Lemma 3.4.4 later.

Proof of Lemma 3.4.3. We describe the randomized algorithm B that computes $\mathsf{EMBCOL}_{n,H,q}(x)$ for a given $x \sim \mathcal{U}_n^{(H)}(\mathbb{F}_q)$.

Set $t = K \log q$ for a sufficiently large constant $K = K(H)$ that will be chosen later depending only on H. For each $e \in E(K_n \times H)$, do the following: For $x = x[e] \in \mathbb{F}_q$, sample $z[e] = (z_0[e], \dots, z_{t-1}[e])$ of Corollary 3.4.5 in time $\mathrm{O}(t)$. Note that (3.2) holds.

After sampling $(z[e])_{e \in E(K_n \times H)}$, the algorithm B computes $\mathsf{EMBCOL}_{n,H,q}(x)$ using (3.3): For each $a \in \{0, \dots, t-1\}^{E(H)}$, construct $\chi^{(a)}$ using $(z[e])_{e \in E(K_n \times H)}$ and compute $\mathsf{EMBCOL}_{n,H,q}(\chi^{(a)})$ using the $T(n,H)$-time algorithm A that solves $(\mathsf{EMBCOL}_{n,H,q}(\cdot), \mathcal{U}_n^{(H)}(\{0,1\}))$ with success probability $1 - \delta$.

We claim that B has success probability $1 - t^{|E(H)|}\delta - \mathrm{O}(n^2|E(H)|q/2^t)$, which completes the proof of Lemma 3.4.3: Indeed, choosing $t = K \log n$ for a sufficiently large constant $K = K(H)$, the success probability of B is at least $1 - \mathrm{O}(\delta(\log n)^{2|E(H)|}) - \mathrm{o}(1) \ge 2/3$ if $\delta = \mathrm{o}((\log n)^{-2|E(H)|})$.

Success probability of B. Since $x[e] \sim \mathrm{Unif}(\mathbb{F}_q)$, Lemma 3.4.4 implies that the distribution of $z[e] := (z_i[e])_{i \in \{0,\dots,t-1\}}$ is total variation distance at most $\epsilon := \mathrm{O}(q/2^t)$ from that of $Z[e] := (Z_0[e], \dots, Z_{t-1}[e])$, where $Z_0[e], \dots, Z_{t-1}[e] \sim \mathrm{Unif}(\{0,1\})$ are i.i.d. random variables. Therefore, the distribution of $z = (z[e])_{e \in E(K_n \times H)}$ is total variation distance at most $|E(K_n \times H)|\epsilon$ from $Z = (Z[e])_{e \in E(K_n \times H)}$ (here, $z[e]$ are independent as well as $Z[e]$).

Let A be the randomized algorithm described in Lemma 3.4.3. Let $\mathcal{S}$ be the set of graphs that is solved by A. Formally,

$$\mathcal{S} = \left\{ F \subseteq K_n \times H : \mathop{\mathbf{Pr}}_{A}[A(F) = \#\mathrm{EMB}_{\mathrm{col}}^{(H)}(F)] \ge \frac{3}{4} \right\}.$$

Let $z := (z[e])_{e \in E(K_n \times H)}$ and $Z := (Z[e])_{e \in E(K_n \times H)}$ be random variables described above. For each $a \in \{0, \dots, t-1\}^{E(H)}$, we have $\mathbf{Pr}_Z\left[\tilde{\chi}^{(a)} \in \mathcal{S}\right] \ge 1 - \delta$, where $\tilde{\chi}^{(a)} = (\tilde{\chi}^{(a)}[e])_{e \in E(K_n \times H)}$ is defined as $\tilde{\chi}^{(a)}[uv] := Z^{(a[c(u)c(v)])}[uv]$. Here, we identify a graph with a binary vector in $\{0,1\}^{E(K_n \times H)}$. Recall that $c : V(K_n \times H) \to V(H)$ maps a vertex to its color. Note that the distribution of $\tilde{\chi}^{(a)}$ is the same as $\mathcal{G}_{n,1/2}^{(H)}$ for every fixed $a \in \{0, \dots, t-1\}^{E(H)}$. By the union bound, we have

$$\mathop{\mathbf{Pr}}_{Z}\left[\forall a \in \{0, \dots, t-1\}^{E(H)} : \tilde{\chi}^{(a)} \in \mathcal{S}\right] \ge 1 - t^{|E(H)|}\delta.$$

Since z is total variation distance at most $|E(K_n \times H)|\epsilon$ from Z, this implies

$$\mathop{\mathbf{Pr}}_{z}\left[\forall a \in \{0, \dots, t-1\}^{E(H)} : \chi^{(a)} \in \mathcal{S}\right] \ge 1 - t^{|E(H)|}\delta - |E(K_n \times H)|\epsilon.$$

This completes the proof of the claim.

Proof of Lemma 3.4.4. Indeed, the statement 1 is a special case of Lemma 4.3 in [BABB19] and the proof is already given (see p. 23 of [BABB19]). For completeness, we present the proof by focusing on the special case. Consider the size of M_x. Let $N := 2^t/q$. Since $x \in \{0, \dots, q-1\}$, it holds that

$$N - 2 \leq \left\lfloor \frac{2^t}{q} - 1 \right\rfloor \leq |M_x| \leq \left\lfloor \frac{2^t}{q} \right\rfloor \leq N.$$

Let $Y \sim \mathrm{Unif}(\{0, \dots, 2^t - 1\})$ and $Y_R \sim \mathcal{Y}_R$ be random variables, where $R \sim \mathrm{Unif}(\mathbb{F}_q)$. For any $A \subseteq \{0, \dots, 2^t - 1\}$, consider the events that $Y \in A$ and $Y_R \in A$. Observe

$$\mathbf{Pr}[Y_R \in A] = \sum_{x \in \mathbb{F}_q} \mathbf{Pr}[Y_x \in A \cap M_x | R = x] \mathbf{Pr}[R = x] = \frac{1}{q} \sum_{x \in \mathbb{F}_q} \frac{|A \cap M_x|}{|M_x|}$$

and

$$\mathbf{Pr}[Y \in A] = \frac{|A|}{2^t} = \frac{1}{q} \sum_{x \in \mathbb{F}_q} \frac{|A \cap M_x|}{N}.$$

Therefore, it holds for any $A \subseteq \{0, \dots, 2^t - 1\}$ that

$$\begin{aligned} |\mathbf{Pr}[Y_R \in A] - \mathbf{Pr}[Y \in A]| &\leq \frac{1}{q} \sum_{x \in \mathbb{F}_q} |A \cap M_x| \left| |M_x|^{-1} - N^{-1} \right| \\ &\leq \frac{|A|}{q} \left(\frac{1}{N-2} - \frac{1}{N} \right) \\ &= \frac{|A|}{q} \cdot \mathrm{O}(N^{-2}) \leq \mathrm{O}(q/2^t). \end{aligned}$$

This completes the proof of the statement 1.

We show the statement 2. The sampling can be done by the following scheme: For a given $x \in \mathbb{F}_q$, let $M := \lfloor (2^t - x - 1)/q \rfloor = |M_x| - 1$ and sample $K \sim \mathrm{Unif}(\{0, \dots, M\})$. Then, output $L := Kq + x$. For any $k \in \{0, \dots, M\}$,

$$\mathbf{Pr}[L = kq + x] = \mathbf{Pr}[K = k] = \frac{1}{M+1}.$$

In other words, $L \sim \mathrm{Unif}(M_x)$ for any x.

3.5 Complexity of Counting $K_{a,b}$-Subgraphs

This section is devoted to the proofs of Theorems 3.2.1 to 3.2.3 and 3.2.7. In Sections 3.5.1 to 3.5.5, we provide several technical results. Finally, in Section 3.5.6, we combine these results to show Theorems 3.2.1 to 3.2.3 and 3.2.7.

3.5.1 Colored subgraph counting vs. (uncolored) subgraph counting

We first prove that the $K_{a,b}$-subgraph counting (i.e., $\#\textsc{Emb}^{(K_{a,b})}$) and colored $K_{a,b}$-subgraph counting (i.e., $\#\textsc{Emb}_{\mathrm{col}}^{(K_{a,b})}$) are computationally equivalent.

Lemma 3.5.1. *Consider $\#\textsc{Emb}_{\mathrm{col}}^{(K_{a,b})}$ and $\#\textsc{Emb}^{(K_{a,b})}$. Given oracle access to one of them, we can solve the other one in time $2^{\mathrm{O}(a+b)} + \mathrm{O}(n^2)$ (in the worst-case sense).*

One direction is well known: The problem $\#\textsc{Emb}_{\mathrm{col}}^{(H)}$ is reducible to $\#\textsc{Emb}^{(H)}$ by using the inclusion-exclusion principle [CM14, Cur18].

Proposition 3.5.2. *Let H be a graph. If $\#\textsc{Emb}^{(H)}$ for n-vertex graphs can be solved in time $T(n)$, then $\#\textsc{Emb}_{\mathrm{col}}^{(H)}$ can be solved in time $\mathrm{O}(2^{|V(H)|}T(n))$ (in the worst-case sense).*

Proof. We identify a vertex of H with a color. Let S be the set of embeddings of H to G. For each color $i \in V(H)$, let $S_i \subseteq S$ be the set of embeddings ϕ of H to G such that the image $\phi(V(H))$ contains the color i. Then, in $\#\mathrm{Emb}^{(H)}_{\mathrm{col}}$, our task is to compute $\left|\bigcap_{i\in V(H)} S_i\right|$. We denote by $\overline{S_i}$ the complement $S \setminus S_i$. Using the inclusion-exclusion principle (Proposition 2.5.7), we can rewrite $\left|\bigcap_{i\in V(H)} S_i\right|$ as

$$\begin{aligned}\left|\bigcap_{i\in V(H)} S_i\right| &= |S| - \left|\bigcup_{i\in V(H)} \overline{S_i}\right| \\ &= |S| + \sum_{I\subseteq V(H): I\neq\emptyset} (-1)^{|I|} \left|\bigcap_{i\in I} \overline{S_i}\right|.\end{aligned}$$

Note that $|S| = \#\mathrm{Emb}^{(H)}(G)$ and $\left|\bigcap_{i\in I} \overline{S_i}\right| = \#\mathrm{Emb}^{(H)}(G_I)$, where G_I is the graph obtained by removing vertices with color in I from G. We can compute these values using the oracle of $\#\mathrm{Emb}^{(H)}$. □

Now we discuss the converse direction: Can we solve $\#\mathrm{Emb}^{(H)}$ given oracle access to $\#\mathrm{Emb}^{(H)}_{\mathrm{col}}$? We show that $\#\mathrm{Emb}^{(H)}$ is reducible to $\#\mathrm{Emb}^{(H)}_{\mathrm{col}}$ when $H = K_{a,b}$. To this end, we consider the problem $\#\mathrm{Hom}^{(H)}$ that asks the number $\#\mathrm{Hom}^{(H)}(G)$ of homomorphisms from H to a given graph G. Recall that a mapping $\phi : V(H) \to V(G)$ is a homomorphism if $\{\phi(u), \phi(v)\} \in E(G)$ whenever $\{u, v\} \in E(H)$.

We reduce $\#\mathrm{Emb}^{(K_{a,b})}$ to $\#\mathrm{Emb}^{(K_{a,b})}_{\mathrm{col}}$ by the following three steps. First, we show that $\#\mathrm{Hom}^{(H)}(G)$ is equal to $\#\mathrm{Emb}^{(H)}_{\mathrm{col}}(G \times H)$ (Fact 3.5.3). Second, we use Lovász's identity [Lov12] to reduce $\#\mathrm{Emb}^{(H)}$ to $\#\mathrm{Hom}^{(H')}$ for some family of graphs H' (Theorem 3.5.4). Finally, we observe that $\#\mathrm{Hom}^{(H')}$ is reducible to $\#\mathrm{Emb}^{(K_{a,b})}$ when $H = K_{a,b}$ (Proposition 3.5.5).

The following well-known fact asserts that $\#\mathrm{Hom}^{(H)}$ is reducible to $\#\mathrm{Emb}^{(H)}_{\mathrm{col}}$.

Fact 3.5.3. *Let H be a k-vertex graph. For any graph G, it holds that $\#\mathrm{Hom}^{(H)}(G) = \#\mathrm{Emb}^{(H)}_{\mathrm{col}}(G \times H)$. Consequently, if $\#\mathrm{Emb}^{(H)}_{\mathrm{col}}$ can be solved in time $T(kn)$ on kn-vertex graphs, then $\#\mathrm{Hom}^{(H)}$ can be solved in time $\mathrm{O}(T(kn) + kn^2)$ on n-vertex graphs.*

Proof. We can solve $\#\mathrm{Hom}^{(H)}$ on input G as follows. Construct $G \times H$ and then run the algorithm for $\#\mathrm{Emb}^{(H)}_{\mathrm{col}}$ on input $G \times H$. Now we show $\#\mathrm{Hom}^{(H)}(G) = \#\mathrm{Emb}^{(H)}_{\mathrm{col}}(G \times H)$. Let ϕ be a homomorphism from H to G. Then, the mapping $\psi : V(H) \ni v \mapsto (\phi(v), v) \in V(G \times H)$ is also a homomorphism and moreover it is injective. This correspondence between ϕ and ψ is one-to-one. □

In light of Fact 3.5.3, it suffices to reduce $\#\mathrm{Emb}^{(H)}$ to $\#\mathrm{Hom}^{(H)}$. To this end, we invoke the following identity.

Theorem 3.5.4 (Lovász [Lov12]; See (2) of [CDM17])**.** *Let H be a fixed graph. Let $\mathcal{P}(H)$ be the set of partitions of $V(H)$ such that, for every $\pi = \{B_1, \ldots, B_t\} \in \mathcal{P}(H)$, each $B_i \subseteq V(H)$ is an independent set $(i = 1, \ldots, t)$.* For each $\pi \in \mathcal{P}(H)$, define H/π as the graph obtained by contracting each vertex set in π. Then

$$\#\mathrm{Emb}^{(H)}(G) = \sum_{\pi\in\mathcal{P}(H)} (-1)^{|V(H)|-|\pi|} \prod_{B\in\pi} (|B|-1)! \cdot \#\mathrm{Hom}^{(H/\pi)}(G).$$

Here, $|\pi|$ denotes the number of subsets in π.

Combining Theorem 3.5.4 and Fact 3.5.3, we can reduce $\#\mathrm{Emb}^{(H)}$ to solving a family of problems $(\#\mathrm{Emb}^{(H/\pi)}_{\mathrm{col}})_{\pi\in\mathcal{P}(H)}$. If $H = K_{a,b}$, we can enumerate all elements of $\mathcal{P}(H)$ in time $\mathrm{O}(2^{a+b})$, and thus the reduction runs in time $\mathrm{O}(n^2 + 2^{a+b})$. Moreover, we show in Proposition 3.5.5 that $\#\mathrm{Emb}^{(K_{a,b}/\pi)}_{\mathrm{col}}$ is reducible to $\#\mathrm{Emb}^{(K_{a,b})}_{\mathrm{col}}$ for every $\pi \in \mathcal{P}(K_{a,b})$, which enables us to reduce $\#\mathrm{Emb}^{(K_{a,b})}$ to $\#\mathrm{Emb}^{(K_{a,b})}_{\mathrm{col}}$.

Proposition 3.5.5. *Assume that $\#\mathrm{Emb}^{(K_{a,b})}_{\mathrm{col}}$ can be solved in time $T(n)$. Let $\pi \in \mathcal{P}(K_{a,b})$. Then, $\#\mathrm{Emb}^{(K_{a,b}/\pi)}_{\mathrm{col}}$ can be solved in time $\mathrm{O}(n^2 + T(n))$.*

Proof. Observe that, for any $\pi \in \mathcal{P}(K_{a,b})$, we have $K_{a,b}/\pi = K_{c,d}$ for some constants $c \le a$ and $d \le b$; therefore, it suffices to reduce $\#\mathrm{Emb}^{(K_{c,d})}_{\mathrm{col}}$ to $\#\mathrm{Emb}^{(K_{a,b})}_{\mathrm{col}}$.

Let (n, G) be an input of $\#\textsc{Emb}_{\mathrm{col}}^{(K_{c,d})}$, where $G \subseteq K_{c,d} \times K_n$. Regard the vertices in $V(K_{a,b})$ as $V(K_{a,b}) = \{l_1, \ldots, l_a, r_1, \ldots, r_b\}$ so that $E(K_{a,b}) = \{l_i, r_j\}_{i\in[a], j\in[b]}$. Then, each vertex $v \in V(G)$ can be represented as the form (r_i, u) or (l_i, u). We write $V(G) = R \cup L$, where R is the set of vertices of the form (r_i, u), and L is that of the form (l_i, u). Fix a vertex $v \in V(K_n)$ and let $L_{\mathrm{add}} = \{(l_i, v)\}_{i=c+1}^{a}$ and $R_{\mathrm{add}} = \{(r_i, v)\}_{i=d+1}^{b}$ be vertex sets. We construct a graph $\hat{G} \subseteq K_{a,b} \times K_n$ as follows.

$$V(\hat{G}) = V(G) \cup L_{\mathrm{add}} \cup R_{\mathrm{add}},$$
$$E(\hat{G}) = E(G) \cup E(R_{\mathrm{add}}, L \cup L_{\mathrm{add}}) \cup E(L_{\mathrm{add}}, R \cup R_{\mathrm{add}}),$$

where, for two vertex subsets S and T, $E(S, T) = \{s, t\}_{s\in S, t\in T}$. See Figure 3.3 for an illustration.

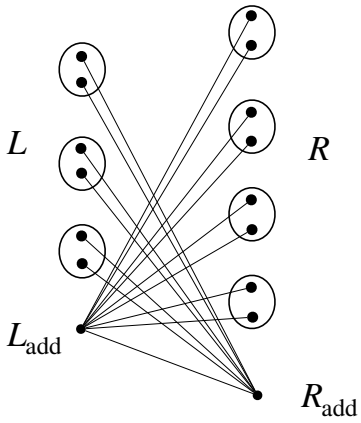


Figure 3.3: The graph $\hat{G}$ of the reduction. In this figure, $\#\textsc{Emb}_{\mathrm{col}}^{(K_{3,4})}(G) = \#\textsc{Emb}_{\mathrm{col}}^{(K_{4,5})}(\hat{G})$ holds.

Note that $\#\textsc{Emb}_{\mathrm{col}}^{(K_{a,b})}(\hat{G}) = \#\textsc{Emb}_{\mathrm{col}}^{(K_{c,d})}(G)$ holds since there is a one-to-one correspondence between copies of $K_{a,b}$ in $\hat{G}$ and that of $K_{c,d}$ in G. □

Lemma 3.5.1 follows from Propositions 3.5.2 and 3.5.5, Fact 3.5.3, and Theorem 3.5.4.

Remark 3.5.6. We comment on the relationship between $\#\textsc{Hom}^{(H)}$ and $\#\textsc{Emb}^{(H)}$. It is easy to see that the problems $\#\textsc{Hom}^{(K_k)}$ and $\#\textsc{Emb}^{(K_k)}$ are equivalent. More generally, such an equivalence holds if H is a *core*; here, a graph H is said to be a core if any homomorphism from H to H is an isomorphism. However, for some H, it is widely believed that there is a gap between $\#\textsc{Hom}^{(H)}$ and $\#\textsc{Emb}^{(H)}$: For example, let M_k be the graph of disjoint k edges. It is known that $\#\textsc{Emb}^{(M_k)}$ (which is the problem of counting the number of matchings of size k) is #W[1]-hard [Cur13], while $\#\textsc{Hom}^{(M_k)}$ can be solved in linear time (observe that $\#\textsc{Emb}^{(M_k)}(G) = (2|E(G)|)^k$).

3.5.2 SETH-hardness of finding a colored $K_{a,b}$-subgraph

Assume $a \leq b$. By enumerating all subsets of size a, we can solve both Colored $K_{a,b}$ Detection in time $\mathrm{O}(n^{a+1})$. If a given graph G is sparse and has m edges, we can solve the problem in time $\mathrm{O}(m^a)$ by enumerating $\binom{N(v)}{a}$ for every vertex v, where $N(v)$ denotes the set of vertices adjacent to v.

Theorem 3.5.7 (Reminder of Theorem 3.2.4)**.** *For any constants $a \geq 2$ and $\epsilon > 0$, there exists a constant $b = b(a, \epsilon) \geq a$ such that* Colored $K_{a,b}$ Detection *cannot be solved in time $\mathrm{O}(m^{a-\epsilon})$ unless SETH fails, where m is the number of edges of the input graph.*

Remark 3.5.8. Theorem 3.2.1 immediately follows from Proposition 3.5.2 and Theorem 3.2.4.

In the proof of Theorem 3.2.4, we consider k-Orthogonal Vectors (k-OV). In k-OV, we are given sets $A_1, \ldots, A_k \subseteq \{0, 1\}^d$ of binary vectors each of cardinarity n and dimension d satisfying $d \leq K \log n$ for a constant K. Our task is to decide whether there exist vectors $\mathbf{a}_1, \ldots, \mathbf{a}_k$ such that $\mathbf{a}_i \in A_i$ for any i and $\sum_{j=1}^{d} \prod_{i=1}^{k} \mathbf{a}_i[j] = 0$. The naïve exhaustive search solves k-OV in time $\mathrm{O}(n^k d) = \mathrm{O}(n^k \log n)$. The current known fastest algorithm solves it in time $\mathrm{O}(n^{k-1/\mathrm{O}(\log(d/\log n))})$ [AWY15]. The *k-Orthogonal Vectors Conjecture* (k-OVC) asserts that k-OV requires time $n^{k-\mathrm{o}(1)}$ for any $d = \omega(\log n)$: More precisely, under k-OVC, for any $k \geq 2$ and $\epsilon > 0$, there exists a constant $K \geq 1$ such that no $\mathrm{O}(n^{k-\epsilon})$-time algorithm solves k-OV of dimension $d \leq K \log n$. It is known that, for every constant $k \geq 2$, SETH implies k-OVC [Wil15, Wil05, LPW17]. Thus, it suffices to reduce k-OV to $\#\textsc{Emb}^{(K_{a,b})}$ for $k = a$.

Reduction (Proof of Theorem 3.2.4)

Fix any constant $a \geq 2$. Assume that there exists a constant $\epsilon > 0$ such that Colored $K_{a,b}$ Detection can be solved in $O(m^{a-\epsilon})$ for every $b \geq a$. We will prove that, under this assumption, there exists a constant $\epsilon' > 0$ such that a-OV of dimension $d = K \log n$ can be solved in time $O(m^{a-\epsilon'})$ for any K. To this end, we present a many-to-one reduction: The reduction maps an instance of a-OV to an equivalent instance of Colored $K_{a,b}$ Detection.

Let $\epsilon > 0$ be a sufficiently small constant that will be specified later. Let $A_1, \ldots, A_k \subseteq \{0,1\}^d$ be an instance of k-OV of dimension $d = K \log n$. We identify a vector $\mathbf{x} \in \{0,1\}^d$ with a subset $x \subseteq [d]$. Thus, each A_i is identified with a family of subsets of $[d]$. Let $\mathcal{P}_1 \cup \cdots \cup \mathcal{P}_C$ be a partition of $[d]$ such that $|\mathcal{P}_i| \leq \epsilon \log n$ holds for every $i \in [C]$, where $C = K/\epsilon$ (we will choose ϵ so that K/ϵ is an integer).

The reduction constructs a graph G and a coloring $c : V(G) \to [a+b]$, resulting in an instance of Colored $K_{a,b}$ Detection of $a := k$ and $b := C$. The vertex set of G is of the form

$$V(G) = V_1 \cup \cdots \cup V_k \cup W_1 \cup \cdots \cup W_C,$$

where each subset $V_1, \ldots, V_k, W_1, \ldots, W_C$ is assigned with a distinct color. For each subset $a \in A_i$, we create a vertex $v_a \in V_i$. For each index $j \in [C]$, enumerate all subsets of $\mathcal{P}_j$. We associate a k-tuple $z = (y_1, \ldots, y_k) \in (2^{\mathcal{P}_j})^k$ of the subsets with a vertex $w_z \in W_j$, if the corresponding vectors $\mathbf{y}_1, \ldots, \mathbf{y}_k$ are orthogonal on $\mathcal{P}_j$. Formally, the vertex set $V(G)$ is

$$V_i := \{v_a : a \in A_i\},$$
$$W_j := \left\{ w_z : z = (y_1, \ldots, y_k) \in (2^{\mathcal{P}_j})^k \text{ satisfies } \sum_{r \in \mathcal{P}_j} \prod_{s \in [k]} \mathbf{y}_s[r] = 0 \right\}.$$

Two vertices $v_a \in V_i$ and $w_z \in W_j$ of $z = (y_1, \ldots, y_k)$ are joined by an edge if $a \cap \mathcal{P}_j = y_i$ holds. The edge set $E(G)$ contains no other edges. Note that $G \subseteq K_n \times K_{a,b}$ in which $V_1, \ldots, V_a, W_1, \ldots, W_b$ obtain distinct colors.

Correctness. Let $A_1, \ldots, A_k \subseteq \{0,1\}^{[d]}$ be the instance of k-OV with $d = K \log n$ and G be the graph constructed by the reduction above. Recall that each A_i is identified with a family of n subsets of $[d]$. Suppose that the given instance is an YES-instance. Then, there is a k-tuple $(a_1, \ldots, a_k) \in A_1 \times \cdots \times A_k$ such that the corresponding vectors $\mathbf{a}_1, \ldots, \mathbf{a}_k$ satisfy $\sum_{r \in [d]} \prod_{s \in [k]} \mathbf{a}_s[r] = 0$. Let

$$U := \bigcup_{i \in [k]} \{v_{a_i}\} \subseteq V(G),$$
$$W := \bigcup_{j \in [C]} \{w_z \in W_j : z = (y_1, \ldots, y_k) \text{ where each } y_i \text{ satisfies } y_i \cap \mathcal{P}_j = a_i \cap \mathcal{P}_j\}.$$

The set $U \cup W$ induces a colored subgraph isomorphic to $K_{a,b}$, where $a = k$ and $b = C$; thus, the pair (G, c) of a graph G and coloring c is an YES-instance of Colored $K_{a,b}$ Detection.

Conversely, suppose that G contains a colored $K_{a,b}$-subgraph. Then we have $|V(H) \cap V_i| = |V(H) \cap W_j| = 1$ for every $i \in [k]$ and $j \in [C]$. Let $v_i \in V(H) \cap V_i$ and $w_j \in V(H) \cap W_j$. As H is isomorphic to $K_{a,b}$, $\{v_i, w_j\} \in E(G)$ for every i, j. Let $\mathbf{a}_i \in \{0,1\}^d$ be the vector associated with the vertex v_i. For every $j \in [C]$, we have $\sum_{r \in \mathcal{P}_j} \prod_{s \in [k]} \mathbf{a}_s[r] = 0$ since each w_j is incident to v_i for all $i \in [k]$. Thus, we have $\sum_{r \in [n]} \prod_{s \in [k]} \mathbf{a}_s[r] = 0$ and hence $(A_1, \ldots, A_k)$ is an YES-instance of k-OV.

Time complexity. The size of the constructed graph G satisfies

$$|V(G)| \leq kn + Cn^{\epsilon k},$$
$$|E(G)| \leq kCn^{1+\epsilon k}.$$

Thus, if Colored $K_{a,b}$ Detection on G can be solved in time $O(m^{a-\epsilon'})$, letting $\epsilon > 0$ be a constant satisfying $(1+\epsilon k)(k-\epsilon') \leq k - \epsilon'/2$ yields an

$$O(m^{a-\epsilon'}) = O(n^{(1+\epsilon k)(k-\epsilon')}) = O(n^{k-\epsilon'/2})$$

time algorithm for k-OV. This falsifies k-OVC as well as SETH.

3.5.3 ETH-hardness of finding a colored $K_{a,a}$-subgraph

Consider the decision problem K_a Detection in which we are asked to decide whether the given graph contains a clique of size a or not. In this section, we reduce K_a Detection to Colored $K_{a,a}$ Detection. Note that K_a Detection does not admit an $f(k) \cdot n^{o(k)}$-time algorithm for any function $f(\cdot)$ unless ETH fails [CHKX06]; thus, the reduction establishes the ETH-hardness of the problem Colored $K_{a,a}$ Detection.

Lemma 3.5.9. *There is an $O(n^2)$-time algorithm that, given a graph G of n vertices, outputs a graph $G' \subseteq K_n \times K_{a,a}$ of $O(an)$ vertices such that G contains an a-clique if and only if G' contains a colored $K_{a,a}$-subgraph.*

Proof. Let G be an instance of K_a Detection. We transform G to the graph G' mentioned in Lemma 3.5.9.

Let $U_1, \ldots, U_a, W_1, \ldots, W_a$ be copies of $V(G)$. We write $V(G) = \{v_1, \ldots, v_n\}$, $U_i = \{u_1^{(i)}, \ldots, u_n^{(i)}\}$, and $W_i = \{w_1^{(i)}, \ldots, w_n^{(i)}\}$. Here, each $u_j^{(i)}$ corresponds to v_j (and so does $w_j^{(i)}$). We set $V(G') = \bigcup_{i \in [a]} (U_i \cup W_i)$; each V_i and W_i is assigned with a distinct color (more formally, each vertex in V_i is assigned with a color i and each vertex in W_i is assigned with a color $a + i$). We construct $E(G')$ such that, for all $i, k \in [a]$ and $j, l \in [n]$, an edge $\{u_j^{(i)}, u_l^{(k)}\}$ is in $E(G')$ if either (1) $i = j$ and $j = l$, or (2) $i \neq j$ and $\{v_j, v_l\} \in E(G)$ holds. The set $E(G')$ does not contain any other edges. This graph can be constructed in time $O(an^2)$.

Now we check the correctness. Suppose that a vertex set $S = \{v_{i_1}, \ldots, v_{i_a}\}$ forms an a-clique in G. Then, the vertex set $\{u_{i_1}^{(1)}, \ldots, u_{i_a}^{(a)}, w_{i_1}^{(1)}, \ldots, w_{i_a}^{(a)}\}$ forms a colored $K_{a,a}$-subgraph in G'. Conversely, if the set $\{u_{i_1}^{(1)}, \ldots, u_{i_a}^{(a)}, w_{j_1}^{(1)}, \ldots, w_{j_a}^{(a)}\}$ forms a colored $K_{a,a}$-subgraph in G', then it holds that $i_1 = j_1, \ldots, i_a = j_a$ and the set $\{v_{i_1}, \ldots, v_{i_a}\}$ forms an a-clique in G. □

3.5.4 An $n^{a+o(1)}$-time algorithm for counting $K_{a,b}$

We now present an algorithm that matches the lower bounds presented so far. Specifically, we design an algorithm that solves $\#\textsc{Emb}_{\text{col}}^{(K_{a,b})}$ in time $O(bn^{a+o(1)})$, thereby proving Theorem 3.2.3. The algorithm of Theorem 3.2.3 is similar to the $O(n^{k+o(1)})$-time algorithm for k-Dominating Set of $k \geq 7$ proposed by Pătraşcu and Williams [PW10]. The algorithm of [PW10] adopts the fast rectangular matrix multiplication [GU18, Gal12, Cop97]. Recently, Le Gall and Urrutia [GU18] proved that we can compute the multiplication of an $n \times n^\gamma$ matrix and an $n^\gamma \times n$ matrix in $n^{2+o(1)}$ arithmetic operations if $\gamma \leq 0.31389$.

Let $G = (V, E)$ be a given instance of $\#\textsc{Emb}^{(K_{a,b})}$. We first consider the case when a is even. We construct an $\binom{n}{a/2} \times n$ matrix B as follows: For each $S \in \binom{V}{a/2}$ and $v \in V$, the corresponding element $B[S][v]$ is defined as

$$B[S][v] := \begin{cases} 1 & \text{if } S \subseteq N(v), \\ 0 & \text{otherwise.} \end{cases}$$

Then, compute the product $BB^\top$ by the fast rectangular matrix multiplication [GU18]. The running time is $O(n^{a+o(1)})$ if $a \geq 8$. Notice that $BB^\top[S_1][S_2]$ is equal to the size of the vertex subset $W(S_1, S_2)$, where $W(S_1, S_2) := \{v \in V \setminus (S_1 \cup S_2) : S_1 \cup S_2 \subseteq N(v)\}$. In other words, the set $W(S_1, S_2)$ contains vertices that is adjacent to all vertices in $S_1 \cup S_2$. For any $S_1, S_2 \in \binom{V}{a/2}$ with $S_1 \cap S_2 = \emptyset$ and $T \in \binom{W(S_1,S_2)}{b}$, the vertex set $S_1 \cup S_2 \cup T$ forms a $K_{a,b}$ subgraph. On the other hand, for a $K_{a,b}$ subgraph, there are $c\binom{a}{a/2}$ ways to take S_1, S_2, T, where $c = 2$ if $a < b$ and $c = 4$ if $a = b$. If $a < b$, the factor c reflects the symmetry of S_1 and S_2; thus $c = 2$. If $a = b$, we further take the symmetry of $S_1 \cup S_2$ and T into account; thus $c = 4$. Then, the number of $K_{a,b}$ subgraphs contained in G is given by

$$c^{-1}\binom{a}{a/2}^{-1} \cdot \sum_{S_1, S_2 \in \binom{V}{a/2} : S_1 \cap S_2 = \emptyset} \binom{BB^\top[S_1][S_2]}{b}.$$

Now consider the case when a is odd. Fix a vertex $u \in V$. Again, we construct an $\binom{n}{(a-1)/2} \times n$ matrix $B^{(u)}$ as follows: For each $S \in \binom{V}{(a-1)/2}$ and $v \in V$,

$$B^{(u)}[S][v] = \begin{cases} 1 & \text{if } \{u, v\} \in E, v \notin S \text{ and } S \subseteq N(v), \\ 0 & \text{otherwise.} \end{cases}$$

Then compute $B^{(u)}(B^{(u)})^\top$ for all $u \in V$. Note that the multiplication can be computed in time $n^{a-1+o(1)}$ for each $u \in V$. Observe that $B^{(u)}(B^{(u)})^\top[S_1][S_2]$ is the number of vertices that is adjacent to all vertices in $S_1 \cup S_2 \cup \{u\}$. Thus, the number of $K_{a,b}$ contained in G is given by

$$c^{-1}\left(a\binom{a-1}{(a-1)/2}\right)^{-1} \cdot \sum_{u\in V} \sum_{S_1,S_2\in\binom{V}{(a-1)/2}:S_1\cap S_2=\emptyset} \binom{B^{(u)}(B^{(u)})^\top[S_1][S_2]}{b},$$

where $c = 2$ if $a < b$ and $c = 4$ if $a = b$.

This yields an $O(bn^{a+o(1)})$ time algorithm (note that $\binom{n}{k}$ can be computed in $O(k \log n)$ time).

3.5.5 Reduce $(\#\mathbf{Emb}^{(K_{a,b})}_{\mathrm{col}}, \mathcal{G}^{(K_{a,b})}_{n,1/2})$ to $(\#\mathbf{Emb}^{(K_{a,b})}, \mathcal{K}_{a,b,n})$

In this section, we present a proof of Proposition 3.2.6, i.e., an average-case-to-average-case reduction from $(\#\textsc{Emb}^{(K_{a,b})}_{\mathrm{col}}, \mathcal{G}^{(K_{a,b})}_{n,1/2})$ to $(\#\textsc{Emb}^{(K_{a,b})}, \mathcal{K}_{a,b,n})$. This will complete a proof of Theorem 3.2.7.

Proof of Theorem 3.2.7. Recall that Theorem 3.2.5 reduces $\#\textsc{Emb}^{(K_{a,b})}_{\mathrm{col}}$ to $(\#\textsc{Emb}^{(K_{a,b})}_{\mathrm{col}}, \mathcal{G}^{(K_{a,b})}_{n,1/2})$. Combined with Proposition 3.5.2, one can reduce $\#\textsc{Emb}^{(K_{a,b})}$ to $(\#\textsc{Emb}^{(K_{a,b})}_{\mathrm{col}}, \mathcal{G}^{(K_{a,b})}_{n,1/2})$. Overall, we obtain a reduction from $\#\textsc{Emb}^{(K_{a,b})}$ to $(\#\textsc{Emb}^{(K_{a,b})}, \mathcal{K}_{a,b,n})$ as stated in Theorem 3.2.7. □

Proof of Proposition 3.2.6. The proof is similar to that of Proposition 3.5.2.

Let $\mathcal{B}_{n,m,1/2}$ be the distribution of a random bipartite graph with left and right vertex sets of size n and m, respectively. Let G be an input of $(\#\textsc{Emb}^{(K_{a,b})}_{\mathrm{col}}, \mathcal{G}^{(K_{a,b})}_{n,1/2})$. Observe that the distribution $\mathcal{G}^{(K_{a,b})}_{n,1/2}$ is identical to $\mathcal{B}_{an,bn,1/2}$. We say that a subgraph $F \subseteq G$ *contains color* i if F contains a vertex of color i. Let S be the set of subgraphs $F \subseteq G$ isomorphic to $K_{a,b}$. Let $S_i \subseteq S$ be the set of subgraphs $F \in S$ that contain color i. Observe that $\#\textsc{Emb}^{(K_{a,b})}_{\mathrm{col}}(G) = \left|\bigcap_{i\in V(K_{a,b})} S_i\right|$. By the inclusion-exclusion principle, we have

$$\begin{aligned}\left|\bigcap_{u\in V(K_{a,b})} S_u\right| &= |S| - \left|\bigcup_{i\in V(K_{a,b})} \overline{S_i}\right| \\ &= |S| - \sum_{\emptyset\neq J\subseteq V(K_{a,b})} (-1)^{|J|-1}\left|\bigcap_{j\in J}\overline{S_j}\right|.\end{aligned}$$

In light of this equality, it suffices to compute $|S|$ and $\left|\bigcap_{j\in J}\overline{S_j}\right|$ for all nonempty $J \subseteq V(K_{a,b})$. Note that the set $\cap_{j\in J}\overline{S_j}$ is equal to the set of $K_{a,b}$ subgraphs in G that does not contain any colors from J. To state it more formally, for a nonempty set $J \subseteq V(K_{a,b})$, let $V_J = \{x \in V(G) : c(x) \in J\}$ and $G_J = G[V_J]$ be the induced subgraph of G by V_J. Then, $\bigcap_{j\in J}\overline{S_j}$ is equal to the set of $K_{a,b}$ subgraphs contained in $G_{\overline{J}}$. Suppose that we have a $T(n)$-time randomized algorithm A that solves $(\#\textsc{Emb}^{(K_{a,b})}, \mathcal{K}_{a,b,n})$ with failure probability δ. Note that, for each $J \subseteq V(K_{a,b})$, the distribution of $G_{\overline{J}}$ for $G \sim \mathcal{B}_{an,bn,1/2}$ is identical to $\mathcal{B}_{cn,dn,1/2}$ for some $c \le a$ and $d \le b$; thus, we can obtain $\left|\bigcap_{j\in J}\overline{S_j}\right|$ with probability at least $1-ab\delta$ since $A(G_{\overline{J}}) = c!d!\left|\bigcap_{j\in J}\overline{S_j}\right|$ (here, $c!d!$ is the number of automorphisms of $K_{c,d}$). Therefore, from the union bound, we can obtain $\left|\bigcap_{j\in J}\overline{S_j}\right|$ for all $\emptyset \neq J \subseteq V(K_{a,b})$ with probability at least $1 - ab2^{a+b}\delta$. Moreover, $A(G) = a!b!|S|$ holds with probability $1 - ab\delta$. Hence, we can solve $(\#\textsc{Emb}^{(K_{a,b})}, \mathcal{G}^{(K_{a,b})}_{n,1/2})$ in time $O(ab2^{a+b} \cdot T(n))$ with probability $1 - O(ab2^{a+b}\delta)$. □

3.5.6 Worst-case complexity of $K_{a,b}$-subgraph counting

We present proofs of Theorems 3.2.1, 3.2.3 and 3.2.7. Theorem 3.2.7 follows from Propositions 3.2.6 and 3.5.2 and Theorem 3.2.5. We can show Theorem 3.2.1 by combining Theorem 3.2.4 and Lemma 3.5.1 (Note that we can solve Colored $K_{a,b}$ Detection using a solver for $\#\textsc{Emb}^{(K_{a,b})}$). Similarly, Theorem 3.2.2 follows from Lemmas 3.5.1 and 3.5.9 and the well-known fact that ETH rules out an $n^{o(a)}$-time algorithm for K_a Detection [CHKX06].

Chapter 4

Fine-Grained Hardness Amplification

4.1 Result

In this chapter, we propose a general framework of fine-grained hardness amplification, that is, the hardness amplification in fine-grained complexity setting. To state it more formally, suppose that there is a distributional problem $(\Pi, \mathcal{D})$ such that any $T(n)$-time algorithm has success probability at most γ. The aim of fine-grained hardness amplification is to construct another distributional problem $(\Pi', \mathcal{D}')$ such that any $T(n)n^{o(1)}$-time algorithm has success probability at most $\gamma' \ll \gamma$.

For a probability distribution $\mathcal{R}$ over a set D, let $\mathsf{supp}(\mathcal{R})$ denote the support of $\mathcal{R}$, that is, if X is the random item sampled from D according to $\mathcal{R}$, then $\mathsf{supp}(\mathcal{R}) = \{x \in D : \mathbf{Pr}[X = x] > 0\}$. For a probability distribution $\mathcal{R}$ and $k \in \mathbb{N}$ let $\mathcal{R}^k$ denote the joint probability distribution of k independent copies each is from $\mathcal{R}$. We define the *direct product* of a distributional problem as follows.

Definition 4.1.1 (Direct product)**.** *Let $k = k(n)$ be any function, and $(\Pi, \mathcal{D})$ be any distributional problem. The* k-wise direct product of $(\Pi, \mathcal{D})$*, denoted by $(\Pi, \mathcal{D})^k$, is defined as the distributional problem $(\Pi^k, \mathcal{D}^k)$ such that*

1. *$(\mathcal{D}^k)_n := \mathcal{D}_n^k$ for each $n \in \mathbb{N}$, and*
2. *$\Pi^k(x_1, \ldots, x_k) := (\Pi(x_1), \ldots, \Pi(x_k))$ for any $(x_1, \ldots, x_k) \in \mathsf{supp}(\mathcal{D}_n^k)$.*

Observe that, if $(\Pi, \mathcal{D})$ has a $T(n)$-time algorithm with success probability γ, then $(\Pi, \mathcal{D})^k$ has a $k \cdot T(n)$-time algorithm with success probability γ^k. Equivalently, if $(\Pi, \mathcal{D})^k$ is hard to solve more than γ^k-fraction of instances, then $(\Pi, \mathcal{D})$ has more than γ-fraction of hard instances (run the algorithm on each of the k inputs). Roughly speaking, the *direct product theorem* claims that the converse direction holds: if $(\Pi, \mathcal{D})$ is weakly hard in average, then $(\Pi, \mathcal{D})$ is strongly hard.

Before going to the detail, we present applications of it to the $K_{a,b}$-subgraph counting problem $\#\textsc{Emb}^{(K_{a,b})}$ on the random bipartite graph drawn from $\mathcal{K}_{a,b,n}$. We consider the k-wise direct product $(\#\textsc{Emb}^{(K_{a,b})}, \mathcal{K}_{a,b,n})^k$.

Theorem 4.1.2 (Average-case complexity of counting $K_{a,b}$-subgraphs for multiple instances)**.** *Under rSETH, for any constants $\epsilon > 0$ and $a \geq 3$, there is a constant $b = b(a, \epsilon)$ such that any $n^{a - \mathrm{O}(\epsilon)}$-time algorithm solve $(\#\textsc{Emb}^{(K_{a,b})}, \mathcal{K}_{a,b,n})$ with success probability at most $n^{-\epsilon}$, where $k = \mathrm{O}(\epsilon \log n)$.*

Theorem 4.1.2 is a "sharp threshold" result: On one side, there is an $n^{a+o(1)}$-time algorithm that solves $\#\textsc{Emb}^{(K_{a,b})}$ for *any* input. On the other side, any $n^{a-\epsilon}$-time algorithm can solve at most $n^{-\Omega(\epsilon)}$-fraction of inputs under rSETH.

Our proof techniques of amplifying average-case hardness can be applied to other subgraph counting problems. Consider the problem $\oplus K_a$-Subgraph of asking the parity of the number of K_a-subgraphs contained in a given input graph, where K_a denotes the complete graph with a vertices. Recall the distribution $\mathcal{G}(n, 1/2)$ of the Erdős–Rényi graph of edge density $1/2$. The *disjoint union* $X \uplus Y$ of two (disjoint) graphs X and Y is the graph defined as $X \uplus Y = (V(X) \cup V(Y), E(X) \cup E(Y))$, where we assume $V(X) \cap V(Y) = \emptyset$. Let $\biguplus^k \mathcal{G}(n, 1/2)$ denote the distribution of the disjoint union of k random graphs $G_1, \ldots, G_k$ each of which is independently drawn from $\mathcal{G}(n, 1/2)$.

We show that the distribution $\biguplus^k \mathcal{G}(n, 1/2)$ is a "hardest" distribution for $\oplus K_a$-Subgraph as follows.

Theorem 4.1.3 (Worst-case to average-case reduction for $\oplus K_a$-Subgraph)**.** *Let $\epsilon > 0$ and $a \in \mathbb{N}$ be arbitrary constants. If there is a $T(n)$-time randomized algorithm that solves the distributional problem $(\oplus K_a\text{-Subgraph}, \biguplus^k \mathcal{G}(n, 1/2))$ with success probability greater than $\frac{1}{2} + n^{-\epsilon}$ for any $k = O(\epsilon \log n)$, then there is a $T(n)n^{O(\epsilon)}$-time randomized algorithm that solves $\oplus K_a$-Subgraph on any input.*

Since any decision problem can be solved with success probability $\frac{1}{2}$ by outputting a uniformly random bit, the success probability of the algorithm in Theorem 4.1.3 is nearly optimal. Therefore, Theorem 4.1.3 shows that the decision problem $\oplus K_a$-Subgraph exhibits *some* sharp threshold between worst- and average-case complexity.[1]

4.2 Overview of Our Framework

In this section, we present a general framework for amplifying average-case hardness in the fine-grained complexity settings, based on the techniques from "coarse-grained" complexity theory. Specifically, we prove fine-grained complexity versions of hardness amplification theorems for any problem Π that admits an efficient *selector* that makes $n^{o(1)}$ queries. In particular, we construct such a selector for $\Pi = \#\text{Emb}^{(K_{a,b})}$ by showing a *doubly-efficient interacive proof system* with low query complexity. We explain the details below.

4.2.1 Direct product theorem

A *direct product theorem* is one of the fundamental hardness amplification results: It states that, if no small circuit can compute a function f on more than a γ-fraction of inputs, then no small circuit can compute the k-wise direct product f^k on a roughly γ^k-fraction of inputs. Here, the k-wise direct product f^k of f is defined as $f^k(x_1, \ldots, x_k) := (f(x_1), \ldots, f(x_k))$. Our plan is to apply a direct product theorem to the function $f := \#\text{Emb}^{(K_{a,b})}$ in order to amplify the average-case hardness of the distributional problem $(\#\text{Emb}^{(K_{a,b})}, \mathcal{K}_{a,b,n})$. We have proved in Theorem 3.1.2 that, under rSETH, no $n^{a-\epsilon}$-time algorithm solves $\#\text{Emb}^{(K_{a,b})}$ on a $(1 - 1/\operatorname{polylog}(n))$-fraction of inputs drawn from $\mathcal{K}_{a,b,n}$. Our strategy is to apply the direct product theorem with $\gamma = 1 - 1/(\log n)^{C+1}$, $k = (\log n)^{C+1}$ and thus $\gamma^k \approx n^{-1}$.

A direct product theorem is a basic way of hardness amplification and there is a long line of works [IW97, Tre03, GNW11, IJK09, IJKW10]. However, there is an obstacle for applying the direct product theorem to our setting. The standard proof of the direct product theorem presents an oracle algorithm A that, given an oracle O that computes f^k for a γ^k-fraction of inputs, produces a list $A_1^O, \ldots, A_m^O$ of oracle algorithms one of which is guaranteed to computes f for a roughly γ-fraction of inputs. Hence, there still remains an issue of identifying a correct algorithm from the list $A_1, \ldots, A_m$. In this chapter, we exploit the direct product theorem of Impagliazzo, Jaiswal, Kabanets, and Wigderson [IJKW10], in which the size m of the list is $m = O(\gamma^{-k})$.

4.2.2 Identifying a correct circuit by a selector

To identify a correct circuit, we use a *selector*, introduced in [Hir15]. For problems Π' and Π, a *selector from Π' to Π* is an efficient algorithm that solves the problem Π' given oracle access to two oracles A_0, A_1 one of which is guaranteed to compute Π. As shown in [Hir15], it is not hard to see that any selector that can identify a correct circuit among *two* circuits can be modified to a selector that can identify a correct circuit among *many* circuits (here, we use circuits as oracle).

In light of this, what is needed for applying the direct product theorem of [IJKW10] is the existence of a selector from $\#\text{Emb}^{(K_{a,b})}$ to the task of solving the distributional problem $(\#\text{Emb}^{(K_{a,b})}, \mathcal{K}_{a,b,n})$ with success probability γ.

In the settings of "coarse-grained" complexity [Hir15], it suffices to consider a polynomial-time selector since polynomial-time algorithms can be composed nicely. However, in the settings of fine-grained complexity, one cannot afford even $n^{\Omega(1)}$ queries for each candidate circuit, because simulating the circuit takes time $n^{a-\epsilon}$. We overcome this difficulty by using the doubly-efficient interactive proof system that makes at most $\operatorname{polylog}(n)$ queries (Theorem 4.2.3) for $\#\text{Emb}^{(K_{a,b})}$. See Section 4.2.3 for details.

[1] The current fastest algorithm [NP85] of counting K_a subgraphs runs in time $O(n^{\omega\lceil a/3\rceil})$ on n-vertex graphs, where ω denotes the matrix multiplication exponent. However, the precise value of ω is not well understood.

Theorem 4.2.1 (Selector for $\#\text{EMB}^{(K_{a,b})}$ using subpolynomial queries)**.** *Let $C_1, \ldots, C_m$ be circuits such that, for some i^*, the circuit C_{i^*} solves $(\#\text{EMB}^{(K_{a,b})}, \mathcal{K}_{a,b,n})$ with success probability $1-(\log n)^{-C}$, where C is a sufficiently large constant that depends only on a, b, and $m = \text{polylog}(n)$. Then, there is a randomized $n^2\,\text{polylog}(n)$-time algorithm that, given the list of circuits $C_1, \ldots, C_m$, solves $\#\text{EMB}_{\text{col}}^{(K_{a,b})}$ (in the worst-case sense) by making $\text{polylog}(n)$ queries for each circuit C_i.*

More generally, we constructed a selector for $\#\text{EMB}_{\text{col}}^{(H)}$ for any fixed graph H.

Theorem 4.2.2 (Selector for $\#\text{EMB}_{\text{col}}^{(H)}$ using subpolynomial queries)**.** *Let $C_1, \ldots, C_m$ be circuits such that, for some i^*, the circuit C_{i^*} solves $(\#\text{EMB}_{\text{col}}^{(H)}, \mathcal{G}_{n,1/2}^{(H)})$ with success probability $1-(\log n)^{-K_H}$, where K_H is a sufficiently large constant that depends only on H, and $m = \text{polylog}(n)$. Then, there is a randomized $n^2\,\text{polylog}(n)$-time algorithm that, given the list of circuits $C_1, \ldots, C_m$, solves $\#\text{EMB}_{\text{col}}^{(H)}$ (in the worst-case sense) by making $\text{polylog}(n)$ queries for each circuit C_i.*

Combining the "almost uniform" direct product theorem of [IJKW10] with the selector of Theorem 4.2.1, we obtain a *completely uniform* and fine-grained version of a direct product theorem for the distributional problem $(\#\text{EMB}^{(K_{a,b})}, \mathcal{K}_{a,b,n})$, which completes a proof of Theorem 4.1.2.

4.2.3 Doubly-efficient interactive proof system

A line of research on interactive proof systems, pioneered by Goldwasser, Micali, and Rackoff [GMR89], revealed the surprising power of interaction (The reader is referred to Section 4.3.1 for the formal definition of an interactive proof system). Early studies of interactive proof systems focused on efficient verification of intractable problems such as PSPACE-complete problems [LFKN92, Sha92]. In contrast, a recent line of research (e.g., [GKR15, RRR16, GR18b, GR18a, BRSV18]) concerns interactive proof systems for tractable problems, which are called *doubly-efficient interactive proof systems*: The goal of a doubly-efficient interactive proof system is to verify a statement in almost linear time by interacting with a polynomial-time prover. We often use n to denote the number of vertices of a given graph; thus, "almost linear time (in the input length)" means $\text{O}(n^2 \operatorname{polylog} n)$ time in our context. It is worth mentioning that a doubly-efficient interactive proof system plays an important role in Proof of Work systems [BRSV17, BRSV18].

At the heart of the proof of Theorems 4.1.2 and 4.1.3, we construct a *doubly-efficient interactive proof system* with *subpolynomial number of queries*.

Theorem 4.2.3 (Interactive proof system for $K_{a,b}$-subgraph counting with subpolynomial queries)**.** *There is an $\text{O}(\log n)$-round interactive proof system for the statement "$\#\text{EMB}^{(K_{a,b})}(G)=C$" such that the verifier runs in time $\text{O}(n^2 \log n)$ and asks the prover to solve $\#\text{EMB}^{(K_{a,b})}$ for $\operatorname{polylog} n$ instances, where n is the number of vertices of the given input graph.*

Roughly speaking, we can construct a selector of Theorem 4.2.1 by simulating the verifier of an interactive proof system of Theorem 4.2.3 by using the candidate circuit as a prover. More precisely, for a given input x and two circuits C_0 and C_1, the selector simulates C_0 and C_1 on input x and obtains the two outputs $C_0(x)$ and $C_1(x)$. Then, the selector runs the interactive proof system to check whether the output is correct. If one of C_0 or C_1 is correct, the verifier accepts the corresponding output and the selector outputs the accepted one.

We emphasize the importance of low query complexity of a doubly-efficient interactive proof system. Suppose that we can simulate the candidate circuits C_0 and C_1 in time $T_C(n)$ and that the verifier runs in time $T_V(n)$, making $Q(n)$ queries in the interactive proof system. The running time of a selector that is constructed from the interactive proof system is roughly $\text{O}(T_V(n) + Q(n)T_C(n))$. In our setting, $T_C(n) = n^{a-\epsilon}$ and thus $Q(n)$ must satisfy $Q(n) = n^{\text{o}(1)}$ to archive our goal.

The salient feature of our interactive proof system is that the amount of communication between a verifier and a prover is at most $\text{polylog}(n)$ bits; equivalently, the number of queries that a verifier makes to a prover is at most $\text{polylog}(n)$. This will be important in the next section—where we prove hardness amplification theorems in a fine-grained setting based on an interactive proof system whose query complexity is subpolynomially small.

Theorem 4.2.4 (interactive proof system for $\#\text{EMB}_{\text{col}}^{(H)}$)**.** *Let H be a graph. There is an $\text{O}(\log n)$-round interactive proof system IP for the statement "$\#\text{EMB}_{\text{col}}^{(H)}(G) = C$" such that, given an input (G, n, C),*

- *The verifier accepts with probability* 1 *for some prover if the statement is true (perfect completeness), while it rejects for any prover with probability at least* 2/3 *otherwise (soundness).*
- *In each round, the verifier runs in time* $n^2(\log n)^{\mathrm{O}(|E(H)|)}$ *and sends* $(\log n)^{\mathrm{O}(|E(H)|)}$ *instances of* $\#\textsc{Emb}_{\mathrm{col}}^{(H)}$ *to a prover.*

Furthermore, for any constant L_0*, there exists a constant* $L_1 = L_1(H, L_0)$ *such that, if the statement is true and the prover has oracle access to a randomized algorithm that solves* $(\#\textsc{Emb}_{\mathrm{col}}^{(H)}, \mathcal{G}_{n,1/2}^{(H)})$ *with success probability* $1 - (\log n)^{-L_1}$*, then the verifier accepts with probability* $1 - (\log n)^{-L_0}$*.*

The "Furthermore" part follows the worst-case-to-average-case reduction of Theorem 3.2.5: We can easily modify an honest prover of IP so that the prover is required to solve polylog(n) instances of the distributional problem $(\#\textsc{Emb}_{\mathrm{col}}^{(H)}, \mathcal{G}_{n,1/2}^{(H)})$.

The interactive proof system of Theorem 4.2.4 can be compared with one given by Goldreich and Rothblum [GR18a]. They presented an O(1)-round $\widetilde{\mathrm{O}}(n)$-query doubly-efficient interactive proof system for $\#\textsc{Emb}^{(K_k)}$. Theorem 4.2.4 significantly improves the query complexity from $\widetilde{\mathrm{O}}(n)$ to polylog(n), at the cost of increasing the round complexity from O(1) to O$(\log n)$. To explain the source of our improvement, we review the ideas of [GR18a]: Their interactive proof system is essentially a variant of the sum-check protocol [LFKN92]. They encoded $\#\textsc{Emb}^{(K_k)}$ as a polynomial over a large finite field and used the following downward self-reducibility of $\#\textsc{Emb}^{(K_k)}(G)$: $\#\textsc{Emb}^{(K_k)}(G) = \sum_{i \in V(G)} \#\textsc{Emb}^{(K_{k-1})}(G - i)$, where $G - i$ denotes the graph obtained by removing the vertex i from G. In each round, the prover sends a polynomial of degree O(n) to the verifier. Each coefficient of the polynomial can be computed by calling a $\#\textsc{Emb}^{(K_k)}$ solver polylog n times. Overall, the number of queries made by the verifier is O$(n \operatorname{polylog} n)$. To summarize, the degree of the polynomial is the main bottleneck for the query complexity.

We improve the query complexity by exploiting a different type of downward self-reducibility. Roughly speaking, at each round, we reduce verifying $\#\textsc{Emb}_{\mathrm{col}}^{(H)}(G)$ for an n-vertex graph G to the verification of $\#\textsc{Emb}_{\mathrm{col}}^{(H)}(G_1), \ldots, \#\textsc{Emb}_{\mathrm{col}}^{(H)}(G_m)$ for $m = \operatorname{polylog}(n)$, where each G_i has $n/2$ vertices. The downward self-reducibility enables us to encode the problem $\#\textsc{Emb}_{\mathrm{col}}^{(H)}$ as a polynomial of degree $|E(H)|(2^{|V(H)|} - 1) = \mathrm{O}(1)$ for a fixed graph H, thereby reducing the query complexity. The details are presented in Section 4.4.

We mention that the existence of a doubly-efficient interactive proof system with communication complexity polylog(n) for $\#\textsc{Emb}_{\mathrm{col}}^{(H)}$ is guaranteed by using the general result of Goldwasser, Kalai, and Rothblum [GKR15]. However, the strategy of an honest prover of their proof system may not be computed efficiently with $\#\textsc{Emb}_{\mathrm{col}}^{(H)}$ oracle. We need an interactive proof system in which an honest prover is simulated with oracle access to $\#\textsc{Emb}_{\mathrm{col}}^{(H)}$, as is guaranteed in Theorem 4.2.4.

4.2.4 Yao's XOR lemma

Let $f : \{0,1\}^n \to \{0,1\}$ be a Boolean function. *Yao's XOR lemma* [Yao82] asserts that, if no small circuit can compute f on more than a γ fraction of inputs, then no small circuit can compute $f^{\oplus k}$ on a roughly $\frac{1}{2} + \gamma^k$ fraction of inputs, where $f^{\oplus k} \colon \{0,1\}^{nk} \to \{0,1\}$ is defined as $f^{\oplus k}(x_1, \ldots, x_k) := f(x_1) \oplus \cdots \oplus f(x_k)$. Yao's XOR lemma is a way to construct a strongly hard-in-average Boolean function from a weakly hard-in-average Boolean function and has been extensively investigated [Yao82, Tre03, GNW11, IJK09] (recall that the direct product theorem does not yield a Boolean function). One can obtain the Yao's XOR lemma by combining the direct product theorem of Impagliazzo, Jaiswal, Kabanets, and Wigderson [IJKW10] with the local list decoding algorithm for the Hadamard code given by Goldreich and Levin [GL89]. Broadly speaking, the algorithm of Goldreich and Levin [GL89] is given an oracle access O computing $f^{\oplus k}$ in $1/2 + \gamma$ and then produces a list $A_1^O, \ldots, A_m^O$ of oracle algorithms one of which is guaranteed to compute f^k with a nonnegligible success probability. The algorithm of [GL89] is simple and efficient and thus we can apply it directly to the fine-grained complexity setting. As a consequence, we can prove a uniform and fine-grained version of Yao's XOR lemma for any problem that admits an efficient selector (Theorem 4.6.3).

We apply the fine-grained version of Yao's XOR lemma to the parity variant of $\#\textsc{Emb}_{\mathrm{col}}^{(H)}$. To state our results formally, let $\oplus\textsc{Emb}_{\mathrm{col}}^{(H)}$ denote the problem of computing the parity $\oplus\textsc{Emb}_{\mathrm{col}}^{(H)}(G) := (\#\textsc{Emb}_{\mathrm{col}}^{(H)}(G) \mod 2)$. Observe that, for k graphs $G_1, \ldots, G_k \subseteq K_n \times H$, computing the k-wise XOR of $\oplus\textsc{Emb}_{\mathrm{col}}^{(H)}(G_1), \ldots, \oplus\textsc{Emb}_{\mathrm{col}}^{(H)}(G_k)$ is equivalent to computing $\oplus\textsc{Emb}_{\mathrm{col}}^{(H)}(G_1 \uplus \cdots \uplus G_k)$ (recall that $F \uplus G$

denotes the disjoint union of two graphs F and G). Let $\biguplus^k \mathcal{G}^{(H)}_{n,1/2}$ denote the distribution of $G_1 \uplus \cdots \uplus G_k$, where each G_i is independently chosen from $\mathcal{G}^{(H)}_{n,1/2}$.

Theorem 4.2.5 (XOR lemma for $\oplus\text{EMB}^{(H)}_{\text{col}}$)**.** *Let H be an arbitrary graph and $c > 0$ be an arbitrary constant. Suppose that there is a $T(n)$-time randomized algorithm that solves $(\oplus\text{EMB}^{(H)}_{\text{col}}, \biguplus^k \mathcal{G}^{(H)}_{n,1/2})$ for any $k = \mathrm{O}(\log n)$ with success probability greater than $\frac{1}{2} + n^{-c}$. Then, there exists a $T(n)n^{\mathrm{O}(c)}$-time randomized algorithm that solves $\oplus\text{EMB}^{(H)}_{\text{col}}$ with probability at least $2/3$ on every input.*

The proof of Theorem 4.2.5 is presented in Section 4.6. The idea is to combine the fine-grained direct product theorem and the local list decoding of [GL89]. Details can be found in Section 4.6.

4.2.5 Related work

Boix-Adserà, Brennan, and Bresler [BABB19] and Goldreich [Gol20] reduced $\oplus K_a$-SUBGRAPH to the distributional problem $(\oplus K_a\text{-SUBGRAPH}, \mathcal{G}(n, 1/2))$. However, their results required an algorithm that solves $(\oplus K_a\text{-SUBGRAPH}, \mathcal{G}(n, 1/2))$ with success probability closed to 1. They left an open question of improving this success probability. Our result Theorem 4.1.3 improves it, albeit for a slightly different distribution.

Goldenberg and Karthik C. S. [GK20] studied hardness amplification of optimization problems, including problems in P. Unlike our settings (in which it is highly non-trivial to construct a selector as in Theorem 4.2.2), it is trivial to construct a selector for any optimization problem; therefore, it is easy to obtain hardness amplification theorems of optimizations problems by using the powerful direct product theorem of Impagliazzo, Jaiswal, Kabanets, and Wigderson [IJKW10].

Goldreich and Rothblum [GR18a] constructed a distribution $\mathcal{D}$ such that $(\#\text{EMB}^{(K_k)}, \mathcal{D})$ is hard to solve efficiently for greater than a n^{-c}-fraction of instances for some constant $c > 0$, based on the worst-case hardness of $\#\text{EMB}^{(K_k)}$. Their distribution $\mathcal{D}$ of graphs is not natural since the distribution is constructed through a number of artificial reductions. Therefore, we cannot exploit their technique since our goal is to obtain a *natural* average-case hard random graph.

4.2.6 Organization of this chapter

In Section 4.4, we present the doubly-efficient proof system of Theorems 3.1.2 and 4.2.4. In Section 4.5, we prove the direct product theorem in the setting of fine-grained complexity. In Section 4.6, we prove our fine-grained XOR Lemma.

The proofs of Theorems 4.1.2 and 4.1.3 are given in Sections 4.5.2 and 4.6.2, respectively.

4.3 Formal Definition

4.3.1 Interactive proof system

We follow the basic notion of interactive proof system (see, e.g., Chapter 8 of [AB08]). For a string $x \in \{0,1\}^*$, let $|x|$ denote the length of x.

Definition 4.3.1 (Interaction of deterministic functions; Definition 8.2 of [AB08])**.** *Let $f, g : \{0,1\}^* \to \{0,1\}^*$ be functions and $k \colon \mathbb{N} \to \mathbb{N}$ be a function. A k-round interaction of f and g on input $x \in \{0,1\}^*$, denoted by $\langle f, g\rangle(x)$, is the sequence of strings $a_1, \ldots, a_k \in \{0,1\}^*$ for $k = k(|x|)$ defined as follows.*

$$\begin{aligned} a_1 &= f(x), \\ a_2 &= g(x, a_1), \\ &\vdots \\ a_{2i+1} &= f(x, a_1, \ldots, a_{2i}) \text{ if } 2i < k, \\ a_{2i+2} &= g(x, a_1, \ldots, a_{2i+1}) \text{ if } 2i + 1 < k. \end{aligned}$$

The output *of f at the end of the interaction, denoted by $\mathsf{out}_f\langle f, g\rangle(x)$, is defined as $\mathsf{out}_f\langle f, g\rangle(x) := f(x, a_1, \ldots, a_k)$.*

Note that an interaction is specified by k and two functions f and g. The string a_{2i+1} can be interpreted as a *message* from f to g, and so does a_{2i+2} vice versa. We can regard f and g as deterministic algorithms: One algorithm sends a new message a_{j+1} to the other given the history $(x, a_1, \ldots, a_j)$ as input. We can extend the notion of interaction of two deterministic algorithms to that of a randomized algorithm f and deterministic algorithm g: We add an additional input $r \in \{0,1\}^*$ to f, where the bit string r stands for the random bit used in the randomized algorithm f.

Definition 4.3.2 (Interaction of functions with private input)**.** *Let $f, g : \{0,1\}^* \to \{0,1\}^*$ be functions and $k \colon \mathbb{N} \to \mathbb{N}$ be a function. A k-round interaction of f and g on input $x \in \{0,1\}^*$ with private input $r \in \{0,1\}^*$, denoted by $\langle f, g\rangle(x; r)$, is the sequence of strings $a_1, \ldots, a_k \in \{0,1\}^*$ for $k = k(|x|)$ defined as follows.*

$$\begin{aligned} a_1 &= f(x, r), \\ a_2 &= g(x, a_1), \\ &\vdots \\ a_{2i+1} &= f(x, r, a_1, \ldots, a_{2i}) \text{ if } 2i < k, \\ a_{2i+2} &= g(x, a_1, \ldots, a_{2i+1}) \text{ if } 2i + 1 < k. \end{aligned}$$

The output *of f at the end of the interaction, denoted by $\mathsf{out}_f\langle f, g\rangle(x; r)$, is defined as $\mathsf{out}_f\langle f, g\rangle(x; r) := f(x, r, a_1, \ldots, a_k)$.*

Note that the function g is not given the string r (thus r is a private input for f). If r is a random string, then the interaction $\langle f, g\rangle(x; r)$ and output $\mathsf{out}_f\langle f, g\rangle(x; r)$ are random variables. Moreover, we can regard f as a randomized algorithm since it is given a randomness.

Definition 4.3.3 (Interactive proof system)**.** *A decision problem Π* has a *k-round interactive proof system if there is an algorithm V (called* verifier*) that performs a k-round interaction satisfying, for any input x and a private random string r,*

1. *(completeness condition) if x is an YES-instance, then there is an algorithm P_{honest} such that*

$$\Pr_r[\mathsf{out}_V\langle V, P_{\text{honest}}\rangle(x; r) = 1] \geq 2/3,$$

2. *(Soundness condition) if x is a NO-instance, then for any algorithm P,*

$$\Pr_r[\mathsf{out}_V\langle V, P\rangle(x; r) = 0] \geq 2/3.$$

An algorithm P that interacts with the verifier is called prover *and the algorithm P_{honest} in the completeness condition is called* honest prover*.*

If the probability in the completeness condition is equal to one, then we say that the verifier has a perfect completeness*.*

Example 4.3.4 (Interactive proof system for NP)**.** Let Hamilton Path be the decision problem that asks whether a given graph G contains a path of length $|V(G)| - 1$. This problem is a well-known NP-complete problem. It is easy to see that Hamilton Path has a one-round interactive proof system with perfect completeness.

Verifier Given a graph G as input, the verifier V sends nothing to the prover P (i.e., the first message a_1 is the empty string).

Prover. Given the input G and an empty message a_1, the prover sends a path P to the verifier as a message.

Verifier. Given a path P, the verifier outputs 1 if $P \subseteq G$ and P is a path of length $|V(G)|-1$. Otherwise, the verifier outputs 0.

If the input G is an YES-instance, an honest prover sends a path of length $|V(G)| - 1$ and then the verifier outputs 1. On the other hand, if G is a NO-instance, no matter what path the prover sends, the verifier outputs 0 since G does not contain such a long path.

In general, any problem in NP has a polynomial-round interactive proof system with a polynomial-time verifier: An honest prover sends the witness and the verifier checks it.

4.3.2 Oracle algorithm

Definition 4.3.5 (Oracle algorithm)**.** *Let $O\colon \{0,1\}^* \to \{0,1\}^*$ be a function. Then, an* oracle algorithm *is an algorithm that uses the function O as a subroutine.*

We denote by A^O if an algorithm A uses O as an oracle. Note that an oracle can be an algorithm (for example, we will consider an oracle that solves a distributional problem with certain success probability in Chapter 4).

Observation 4.3.6. *Consider an oracle algorithm A^O. Suppose that A calls the oracle O at most $Q(n)$ times. If A and O runs in time $T_A(n)$ and $T_O(n)$, respectively, then the total running time of A^O is at most $T_A(n) + T_O(n)Q(n)$.*

We will construct a randomized oracle algorithm A^O that solves $\#\textsc{Emb}^{(K_{a,b})}$ with total running time $n^{a-\Omega(\epsilon)}$, where the oracle O solves the k-wise direct product $(\#\textsc{Emb}^{(K_{a,b})}, \mathcal{K}_{a,b,n})^k$ with success probability at least $n^{-o(\epsilon)}$ in time $n^{a-\epsilon}$ for $k = \mathrm{polylog}(n)$. Here, the number $Q(n)$ of queries has to be at most $n^{o(1)}$.

4.4 Doubly-Efficient Interactive Proof System

This section is devoted to the proof of Theorem 4.2.4. Fix a prime $n^{|V(H)|} < q < 2n^{|V(H)|}$ and consider the polynomial $\mathsf{EMBCOL}_{n,H,q} : \mathbb{F}_q^{E(K_n \times H)} \to \mathbb{F}_q$ defined in (3.1). In our interactive proof system IP, the verifier checks the statement that $\mathsf{EMBCOL}_{n,H,q}(x) = C$ for given $C \in \mathbb{F}_q$ and $x \in \mathbb{F}_q^{E(K_n \times H)}$. Recall that, if x is an edge indicator vector of a graph G, then $\mathsf{EMBCOL}_{n,H,q}(x) = \#\textsc{Emb}^{(H)}(G)$ holds.

4.4.1 Downward reducibility

We may assume without loss of generality that $n = 2^t$ for some $t \in \mathbb{N}$. Since otherwise, we can add isolated vertices to G. For each $i \in V(H)$, let $V_i := \{v \in V(K_n \times H) : c(v) = i\}$.

For each $i \in V(H)$, let $(V_{i,0}, V_{i,1})$ be a partition of V_i such that $|V_{i,0}| = |V_{i,1}| = |V_i|/2$. For $\eta \in \{0,1\}^{V(H)}$, let $E_\eta = \cup_{ij \in E(H)} E(V_{i,\eta(i)}, V_{j,\eta(j)})$, where $E(S,T) = \{e \in E(K_n \times H) : e \cap S \neq \emptyset \text{ and } e \cap T \neq \emptyset\}$ for $S, T \subseteq V(K_n \times H)$ (see Figure 4.1). Since $|V_{i,\eta(i)}| = |V_i|/2$, we can identify E_η with $E(K_{n/2} \times H)$. From the definition (3.1), we have

$$\mathsf{EMBCOL}_{n,H,q}(x) = \sum_{\eta \in \{0,1\}^{V(H)}} \mathsf{EMBCOL}_{n/2,H,q}(x[E_\eta]), \tag{4.1}$$

where $x[E_\eta] \in \mathbb{F}_q^{E_\eta}$ is the restriction of x on E_η.

We identify $\{0,1\}^{V(H)}$ with $\{0,1,\dots,2^{|V(H)|}-1\} \subseteq \mathbb{F}_q$ in the following way. Regard $V(H) = \{0,\dots,k-1\} \subseteq \mathbb{F}_q$ for $k = |V(H)|$ and consider the mapping $\{0,1\}^{V(H)} \ni \eta \mapsto \sum_{i \in V(H)} 2^i \eta(i) \in \{0,\dots,2^{|V(H)|}-1\} \subseteq \mathbb{F}_q$. This mapping is injective and thus we can regard η as an element of $\mathbb{F}_q$. For $\eta \in \{0,\dots,2^{|V(H)|}-1\}$, let $\delta_\eta : \mathbb{F}_q \to \mathbb{F}_q$ be the degree-$(2^{|V(H)|}-1)$ polynomial satisfying

$$\delta_\eta(z) = \begin{cases} 1 & \text{if } z = \eta, \\ 0 & \text{if } z \neq \eta \text{ and } 0 \leq z < 2^{|V(H)|}. \end{cases}$$

Note that these $2^{|V(H)|}-1$ conditions specify δ_η since δ_η has degree $2^{|V(H)|}-1$. Let $m = |E(K_{n/2} \times H)|$. Suppose $E(K_{n/2} \times H) = \{e_1,\dots,e_m\}$ and $E_\eta = \{e_1^\eta,\dots,e_m^\eta\}$ for each $\eta \in \{0,1\}^{V(H)}$. For each $\eta \in \{0,1\}^{V(H)}$ and $v \in \mathbb{F}_q^{E_\eta}$, define $v' \in \mathbb{F}_q^{E(K_{n/2} \times H)}$ by $v'[e_i] = v[e_i^\eta]$. Then, for $x \in \mathbb{F}_q^{E(K_n \times H)}$, let $\tilde{x} : \mathbb{F}_q \to \mathbb{F}_q^{E(K_{n/2} \times H)}$ be $\tilde{x}(\cdot) := \sum_{\eta \in \{0,1\}^{V(H)}} \delta_\eta(\cdot)(x[E_\eta])'$. Note that $\tilde{x}$ satisfies (i) $\tilde{x}(\eta) = x[E_\eta]$ holds for all $\eta \in \{0,1\}^{V(H)}$, and (ii) for each $e \in E(K_{n/2} \times H)$, the function $\mathbb{F}_q \ni z \mapsto \tilde{x}(z)[e] \in \mathbb{F}_q$ is a polynomial of degree $2^{|V(H)|}-1$. In condition (i), we identified E_η with $E(K_{n/2} \times H)$. For each η, the function δ_η can be constructed by $\mathrm{O}(2^{|V(H)|}|V(H)|)$ field operations using the fast univariate polynomial interpolation [Hor72]. Since our computational model is $\mathrm{O}(\log n)$-Word RAM, we can perform any field operation on $\mathbb{F}_q$ in constant time. Thus, the construction of $\tilde{x}$ can be done in time $\mathrm{O}(2^{|V(H)|}|V(H)||E(H)|n^2)$. Using $\tilde{x}(\cdot)$, we can rewrite the recursion formula (4.1) as

$$\mathsf{EMBCOL}_{n,H,q}(x) = \sum_{\eta \in \{0,\dots,2^{|V(H)|}-1\}} \mathsf{EMBCOL}_{n/2,H,q}(\tilde{x}(\eta)). \tag{4.2}$$

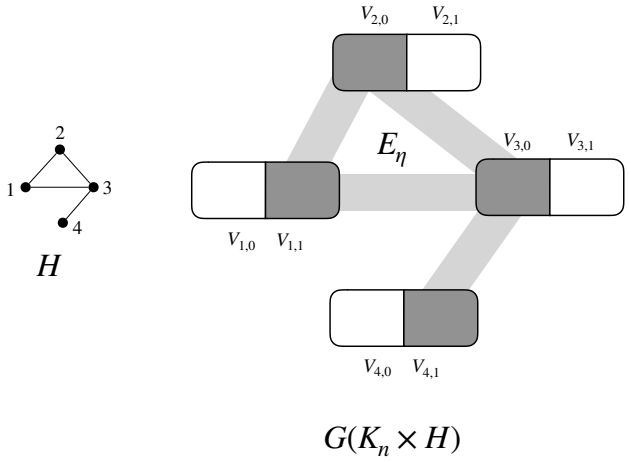


Figure 4.1: An example of E_η. In this example, $\eta = (1,0,0,1) \in \{0,1\}^{V(H)}$. Four grey areas represent $V_{i,\eta(i)}$ for $i = 1,2,3,4$.

Note that $\mathsf{EMBCOL}_{n,H,q}(\tilde{x}(\cdot))$ is a univariate polynomial over $\mathbb{F}_q$ of degree $|E(H)|(2^{|V(H)|}-1)$.

4.4.2 Description and analysis of IP

Now we present IP that verifies the statement "$\mathsf{EMBCOL}_{n,H,q}(x) = C$". Suppose that $n = 2^t$. IP consists of $t+1$ rounds. The verifier is given a vector $x \in \mathbb{F}_q^{E(K_n \times H)}$ (in our case, x is the edge indicator of the input graph G). At each round, the verifier updates the vector x and the constant C. In r-th round, the protocol proceeds as follows.

Verifier. When $r = t+1$, check $\mathsf{EMBCOL}_{n,H,q}(x) = C$ and halt.

Prover. Send a polynomial $G(\cdot)$ of degree at most $|E(H)|(2^{|V(H)|}-1)$ over $\mathbb{F}_q$ to the verifier.

Verifier. Check $C = \sum_{z \in \{0,\dots,2^{|V(H)|}-1\}} G(z)$. If not, reject. Otherwise, construct the polynomial vector $\tilde{x}(\cdot)$ using x. Sample $i \sim \mathrm{Unif}(\mathbb{F}_q)$ and proceed to the next round, where Verifier checks the statement "$\mathsf{EMBCOL}_{n/2,H,q}(\tilde{x}(i)) = G(i)$" recursively.

The task of the prover is to construct the function $\mathsf{EMBCOL}_{n/2,H,q}(\tilde{x}(\cdot))$. Note that we can construct $\mathsf{EMBCOL}_{n/2,H,q}(\tilde{x}(\cdot))$ by evaluating $\mathsf{EMBCOL}_{n/2,H,q}(\tilde{x}(\cdot))$ at $|E(H)|(2^{|V(H)|}-1)+1$ points. The evaluation is reducible to $\#\textsc{Emb}_{\mathrm{col}}^{(H)}$ via (3.3). Thus, one can modify IP above such that the verifier asks the prover to solve $\#\textsc{Emb}_{\mathrm{col}}^{(H)}$ for $(\log n)^{\mathrm{O}(1)}$ instances. In what follows, we analyze this modified protocol.

Running time

Let n be the size of the original input. In the beginning of r-th round, the size of x is $|E(K_n \times H)|/4^{r-1} = |E(H)| \cdot 4^{t-r+1}$. Thus, in the $(t+1)$-th round, the verifier runs in constant time. Any other task of the verifier can be done in time $n^2(\log n)^{\mathrm{O}(1)}$ (the bottleneck is the simulation of the reduction of constructing $\mathsf{EMBCOL}_{n/2,H,q}(\tilde{x}(\cdot))$ to $\#\textsc{Emb}_{\mathrm{col}}^{(H)}$).

Completeness and soundness

The perfect completeness of IP is easy: If the statement is true, an honest prover convinces the verifier with probability 1 by sending the polynomial $\mathsf{EMBCOL}_{n/2,H,q}(\tilde{x}(\cdot))$ at each round (recall that the polynomial $\mathsf{EMBCOL}_{n,H,q}(\tilde{x}(\cdot))$ satisfies the recurence formula (4.2)).

Now we show the soundness.

Proposition 4.4.1. *Let $n = 2^t$. If the statement "$\mathsf{EMBCOL}_{n,H,q}(x) = C$" is false, then for any provers, the verifier rejects with probability $\left(1 - \frac{D}{q}\right)^{t-1}$.*

Proof. The proof is based on the standard argument of sumcheck protocols. We show Proposition 4.4.1 by induction on the number of rounds. Suppose that the statement "$\mathsf{EMBCOL}_{n,H,q}(x) = C$" is false. In the last t-th round, the verifier immediately reject (with probability 1).

Let $D := |E(H)|(2^{|V(H)|} - 1)$ be the degree of $\mathsf{EMBCOL}_{n/2,H,q}(\tilde{x}(\cdot))$. Fix $1 \leq r < t$ and suppose that the verifier rejects with probability $(1 - D/q)^{t-r}$ during $(r+1)$-th to t-th rounds. Consider r-th round with the assumption that the the statement "$\mathsf{EMBCOL}_{n,H,q}(x) = C$" is false. Then, any provers cheat with sending a polynomial $G(\cdot)$ that is not equal to $\mathsf{EMBCOL}_{n/2,H,q}(\tilde{x}(\cdot))$. It holds that

$$\Pr_{i \sim \mathrm{Unif}(\mathbb{F}_q)} \left[G(i) \neq \mathsf{EMBCOL}_{n/2,H,q}(\tilde{x}(i)) \right] \geq 1 - \frac{D}{q}$$

since both G and $\mathsf{EMBCOL}_{n/2,H,q}(\tilde{x}(\cdot))$ are degree at most D. Therefore, with probability $1 - D/q$, the rounds proceeds to the next $(t+1)$-th round with a false statement. By the induction assumption, the verifier rejects with probability at least $(1 - D/q) \cdot (1 - D/q)^{t-r} = (1 - D/q)^{1+t-r}$, which completes the proof of Proposition 4.4.1. □

Ability of an honest prover

In IP, the prover is required to send a polynomial G that is expected to be $\mathsf{EMBCOL}_{n/2,H,q}(\tilde{x}(\cdot))$. As mentioned above, this can be reduced to solving $m = \mathrm{polylog}(n)$ instances of $\#\textsc{Emb}^{(H)}_{\mathrm{col}}$. Moreover, by Theorem 3.2.5, each of the m instances can be reduced to $\mathrm{polylog}(n)$ instances of $(\#\textsc{Emb}^{(H)}_{\mathrm{col}}, \mathcal{G}^{(H)}_{n,1/2})$. In other words, an honest prover can construct the polynomial $\mathsf{EMBCOL}_{n/2,H,q}(\cdot)$ by solving $m\,\mathrm{polylog}(n) = \mathrm{polylog}(n)$ instances of the distributional problem $(\#\textsc{Emb}^{(H)}_{\mathrm{col}}, \mathcal{G}^{(H)}_{n,1/2})$.

Suppose that the prover has oracle access to a randomized algorithm solving $(\#\textsc{Emb}^{(H)}_{\mathrm{col}}, \mathcal{G}^{(H)}_{n,1/2})$ with success probability $1 - (\log n)^{-L_1}$. Then, by the union bound, the probability that the oracle outputs at last one wrong answer is at most $\mathrm{polylog}(n) \cdot (\log n)^{-L_1} \leq (\log n)^{-L_0}$ if $L_1 = L_1(H, L_0)$ is sufficiently large.

This completes the proof of Theorem 4.2.4.

4.4.3 Interactive proof system for counting $K_{a,b}$

By combining Theorems 3.2.7 and 4.2.4 and Lemma 3.5.1, we obtain an interactive proof system for $\#\textsc{Emb}^{(K_{a,b})}$ as follows.

Corollary 4.4.2 (IP for $\#\textsc{Emb}^{(K_{a,b})}$)**.** *Let H be a fixed graph. There is an $\mathrm{O}(\log n)$-round interactive proof system IP for the statement "$\#\textsc{Emb}^{(K_{a,b})}(G) = C$" such that, given an input (G, n, C),*

- *The verifier accepts with probability 1 if the statement is true (perfect completeness), while it rejects for any prover with probability at least $2/3$ otherwise (soundness).*
- *The verifier runs in time $n^2(\log n)^{\mathrm{O}(ab)}$ and sends $(\log n)^{\mathrm{O}(ab)}$ instances of $\#\textsc{Emb}^{(K_{a,b})}$ at each round.*

Furthermore, for any contant L_0, there exists a constant $L_1 = L_1(a, b, L_0)$ such that, if the prover solves $(\#\textsc{Emb}^{(K_{a,b})}, \mathcal{K}_{a,b,n})$ with success probbalitiy $1 - (\log n)^{-L_1}$, then the verifier accepts with probability $1 - (\log n)^{-L_0}$.

Proof. For a given graph G, the verifier applies Lemma 3.5.1 and reduces $\#\textsc{Emb}^{(K_{a,b})}$ to solving $m = \mathrm{O}(1)$ instances of $\#\textsc{Emb}^{(K_{a,b})}_{\mathrm{col}}$.

Then, the verifier solves each of the m instances $G'_1, \ldots, G'_m$ of $\#\textsc{Emb}^{(K_{a,b})}_{\mathrm{col}}$ using the reduction of Theorem 3.2.5 with the help of the prover. Let $C'_1, \ldots, C'_m$ be the values obtained by the reduction. Now, the verifier suffices to check that, for each $i \in [m]$, the answer of the i-th instance G'_i of $\#\textsc{Emb}^{(K_{a,b})}_{\mathrm{col}}$ is C'_i (if all of these m values are the correct one, then the verifier could solve the original instance G of $\#\textsc{Emb}^{(K_{a,b})}$).

To this end, run IP of Theorem 4.2.4 with letting $H = K_{a,b}$. Here, an honest prover suffices to solve $(\#\textsc{Emb}^{(K_{a,b})}, \mathcal{K}_{a,b,n})$ since $\#\textsc{Emb}^{(K_{a,b})}_{\mathrm{col}}$ reduces to solving $\mathrm{polylog}(n)$ instances of $(\#\textsc{Emb}^{(K_{a,b})}, \mathcal{K}_{a,b,n})$ by combining the reductions of Lemma 3.5.1 and Theorem 3.2.7. □

4.5 Fine-Grained Direct Product Theorem

In this section, we provide a sufficient condition for a direct product theorem to hold. We will present a direct product theorem for any distributional problem that admits a selector. The notion of selector that we use in this chapter is defined below.

Definition 4.5.1 ((Oracle) Selector; [Hir15])**.** *A randomized oracle algorithm S is said to be a* selector *from Π to a distributional problem $(\Pi, \mathcal{D})$ with success probability $1-\delta$ if*

1. *given access to two oracles A_0, A_1 one of which solves $(\Pi, \mathcal{D})$ with success probability $1-\delta$, on input x, the oracle algorithm S^{A_0,A_1} computes $\Pi(x)$ with high probability (say, probability $\geq \frac{3}{4}$), and*
2. *for any $n \in \mathbb{N}$ and any input $x \in \mathsf{supp}(\mathcal{D}_n)$, each query q of S to the oracles A_0 and A_1 satisfies that $q \in \mathsf{supp}(\mathcal{D}_n)$.*

In order to obtain a direct product theorem in the settings of fine-grained complexity, it will be crucial to consider a selector with $\mathrm{polylog}(n)$ queries.

4.5.1 Selector for subgraph counting problems

In this subsection, we show the existence of a selector with $\mathrm{polylog}(n)$ queries for $\#\textsc{Emb}_{\mathrm{col}}^{(H)}$. We first recall the notion of instance checker, which is known to imply the existence of selector ([Hir15]).

Definition 4.5.2 (Instance Checker; Blum and Kannan [BK95])**.** *For a problem Π, a randomized oracle algorithm M is said to be an* instance checker *for Π if, for every instance x of Π and any oracle A,*

1. $\mathbf{Pr}_M[M^A(x) = \Pi(x)] = 1$ *if A solves Π correctly on every input, and*
2. $\mathbf{Pr}_M[M^A(x) \notin \{\Pi(x), \mathsf{fail}\}] = \mathrm{o}(1)$.

The existence of an instance checker for a problem Π is implied by an efficient interactive proof system for Π where the computation of an honest prover is efficiently reducible to Π. By using the interactive proof system of Theorem 4.2.4, we obtain the following instance checker with $\mathrm{polylog}(n)$ queries for $\#\textsc{Emb}_{\mathrm{col}}^{(H)}$.

Theorem 4.5.3. *There exists an instance checker* Checker *for $\#\textsc{Emb}_{\mathrm{col}}^{(H)}$ such that, given a graph $G \subseteq K_n \times H$,*

1. Checker *runs in time $\widetilde{O}(n^2)$,*
2. *for any oracle A, $\mathsf{Checker}^A$ calls the oracle A at most $(\log n)^{C_H}$ times, where C_H is a constant that depends only on H, and*
3. *each query G' of* Checker *satisfies $G' \subseteq K_n \times H$.*

Proof. Recall IP of Theorem 4.2.4. For a given oracle A, Checker obtains $C := A(G)$ and then runs IP using A as a prover to verify $\#\textsc{Emb}_{\mathrm{col}}^{(H)}(G) = C$. If the verifier accepts, then Checker outputs C and otherwise it outputs fail.

Suppose the oracle A solves $\#\textsc{Emb}_{\mathrm{col}}^{(H)}$ correctly. Then, Checker output the correct answer with probability 1 by the perfect completeness of IP.

Now we check the second condition of Definition 4.5.2. If $A(G)$ is the correct answer, then the output of Checker is either $A(G)$ or fail. Otherwise, IP proceeds with the false statement that $\#\textsc{Emb}_{\mathrm{col}}^{(H)}(G) = A(G)$. It follows from the soundness of IP (c.f. Proposition 4.4.1) that Verifier rejects with probability $(1 - \mathrm{O}(q^{-1}))^t$. Hence, $\mathbf{Pr}[\mathsf{Checker}(G) = \mathsf{fail}] \geq 1 - \mathrm{O}(t/q) = 1 - \mathrm{o}(1)$. □

Theorem 4.5.4 (Restatement of Theorem 4.2.2)**.** *Let H be a fixed graph. There exists a selector S from $\#\textsc{Emb}_{\mathrm{col}}^{(H)}$ to $(\#\textsc{Emb}_{\mathrm{col}}^{(H)}, \mathcal{G}_{n,1/2}^{(H)})$ with success probability $1 - 1/\mathrm{polylog}(n)$ such that*

1. *S runs in time $\widetilde{O}(n^2)$, and*
2. *S makes at most $\mathrm{polylog}(n)$ queries.*

Proof. We combine the instance checker C of Theorem 4.5.3 with a worst-case to average-case reduction R (Theorem 3.2.5).

Here is the algorithm of a selector S. Given a graph G and oracle access to A_0, A_1, for each $b \in \{0,1\}$, the selector S simulates the instance checker $C(G)$, and answer any query q of the instance checker by running the reduction $R^{A_b}(q)$. If the checker outputs some answer other than fail, the selector S outputs the answer and halts.

The correctness of S can be shown as follows. Let A_b be an oracle that solves $(\#\textsc{Emb}_{\mathrm{col}}^{(H)}, \mathcal{G}_{n,1/2}^{(H)})$ with success probability $1 - 1/\operatorname{polylog}(n)$, where $b \in \{0,1\}$. By the correctness of the reduction R, the algorithm R^{A_b} solves $\#\textsc{Emb}_{\mathrm{col}}^{(H)}$ with high probability. Therefore, if the instance checker C is simulated with oracle access to R^{A_b}, then by the property of an instance checker, C outputs the correct answer with high probability. Moreover, C outputs a wrong answer with probability at most $\mathrm{o}(1)$; thus, the selector outputs the correct answer with high probability. □

Corollary 4.5.5. *There is an $\widetilde{O}(n^2)$-time selector S from $\#\textsc{Emb}^{(K_{a,b})}$ to $(\#\textsc{Emb}^{(K_{a,b})}, \mathcal{K}_{a,b,n})$ with success probability $1 - 1/\operatorname{polylog}(n)$. Moreover, S makes at most $\operatorname{polylog}(n)$ queries.*

Proof. From Theorem 4.5.4, we obtain a selector S from $\#\textsc{Emb}_{\mathrm{col}}^{(K_{a,b})}$ to $(\#\textsc{Emb}_{\mathrm{col}}^{(K_{a,b})}, \mathcal{G}_{n,1/2}^{(K_{a,b})})$. Here, we let $H = K_{a,b}$. Invoke Proposition 3.2.6 that reduces $(\#\textsc{Emb}_{\mathrm{col}}^{(K_{a,b})}, \mathcal{G}_{n,1/2}^{(K_{a,b})})$ to $(\#\textsc{Emb}^{(K_{a,b})}, \mathcal{K}_{a,b,n})$ in $2^{\mathrm{O}(a+b)} = \mathrm{O}(1)$ time. Note that the reduction of Proposition 3.2.6 preserves the success probability within a constant factor. Thus, each oracle query of S can be replaced by the reduction and we obtain a selector from $\#\textsc{Emb}_{\mathrm{col}}^{(K_{a,b})}$ to $(\#\textsc{Emb}^{(K_{a,b})}, \mathcal{K}_{a,b,n})$. By Lemma 3.5.1, $\#\textsc{Emb}^{(K_{a,b})}$ is efficiently reducible to $\#\textsc{Emb}_{\mathrm{col}}^{(K_{a,b})}$, from which the existence of a selector from $\#\textsc{Emb}^{(K_{a,b})}$ to $(\#\textsc{Emb}^{(K_{a,b})}, \mathcal{K}_{a,b,n})$ follows. □

4.5.2 Direct product theorem for any problem with selector

Using the notion of selector, we provide a direct product theorem in the context of fine-grained complexity. A direct product of a distributional problem is formally defined as follows. Recall Definition 4.1.1 of the direct product of a distributional problem.

The following direct product theorem gives an *almost* uniform direct product, in the sense that it requires $\mathrm{O}(\log 1/\epsilon)$ bits of non-uniform advice in order to identify which is a correct algorithm. We observe that the direct product theorem is quite efficient and useful even in the setting of fine-grained complexity.

Theorem 4.5.6 (Impagliazzo, Jaiswal, Kabanets, and Wigderson [IJKW10])**.** *Let $k \in \mathbb{N}$, $\epsilon, \delta > 0$ be parameters that satisfy $\epsilon > \exp(-\Omega(\delta k))$. There exists a randomized oracle algorithm M that, given access to an oracle C that solves $(\Pi, \mathcal{D})^k$ with success probability ϵ, with high probability, produces a list of deterministic oracle algorithms $M_1, \cdots, M_m$ such that M_i^C computes $(\Pi, \mathcal{D})$ with success probability $1 - \delta$ for some $i \in \{1, \ldots, m\}$, where $m = \mathrm{O}(1/\epsilon)$.*

If an oracle C can be computed in $T_C(n)$ time, then each M_i^C runs in time $\mathrm{O}(T_C(n)\epsilon^{-1} \log \delta^{-1}))$. The running time of M is at most $\mathrm{O}(\epsilon^{-1} T_C(n))$.

Remark 4.5.7. The algorithm M of Theorem 4.5.6 works even when Π is not a decision problem (see Theorem 1.6 of [IJKW10]).

Lemma 4.5.8. *Let $(\Pi, \mathcal{D})$ be a distributional problem. Suppose there exists a selector S from Π to $(\Pi, \mathcal{D})$ with success probability $1 - \delta$ that calls an oracle at most $Q(n)$ times. Let $M_1, \ldots, M_m$ be a list of deterministic algorithms such that, (1) for some $i \in \{1, \ldots, m\}$, M_i solves $(\Pi, \mathcal{D})$ with success probability $1 - \delta$, (2) for all $i \in \{1, \ldots, m\}$, M_i runs in time $t_M(n)$, and (3) for all $i, j \in \{1, \ldots, m\}$, the selector S^{M_i, M_j} runs in time $t_S(n)$ (here, $t_S(n)$ does not take the running times of M_i and M_j into account).*

Then, there exists a $t(n)$-time randomized algorithm that solves Π with high probability, where $t(n) = \mathrm{O}(m^2(t_M(n)Q(n) + t_S(n)) \log m / \log \delta^{-1})$.

Proof. Let x be an input. From the assumption, there exists a selector S such that $\mathbf{Pr}_S[S^{A_0, A_1}(x) = \Pi(x)] \geq 1 - 1/16m$ for any oracles A_0, A_1 and any input x. Here, at least one of A_0 and A_1 solves $(\Pi, \mathcal{D})$ with success probability $1 - \delta$. Note that the probability $\mathbf{Pr}_S[S^{A_0, A_1}(x) = \Pi(x)]$ can be assumed to be $1 - 1/16m$ since one can repeat the computation of S for $\mathrm{O}(\log m)$ times and then output the majority. Let us call this selector $\widetilde{S}$.

We present a randomized algorithm B that solves Π. For each $i, j \in \{1, \dots, m\}$ with $i \neq j$, B runs $\widetilde{S}^{M_i, M_j}(x)$ and let c_{ij} be its output. If there exists $i \in \{1, \dots, m\}$ such that $c_{ij} = c$ for all $j \in \{1, \dots, m\}$, B outputs c. If B outputs nothing during the iteration, B outputs anything. Since the overall running time of $\widetilde{S}^{M_i, M_j}$ is at most $(t_S(n) + t_M(n)Q(n)) \cdot \mathrm{O}(\log m)$ for every i, j, the algorithm B runs in time $\mathrm{O}(m^2(t_M(n)Q(n) + t_S(n)) \log m)$.

We claim the correctness of the algorithm B. From the assumption, there exists $i \in \{1, \dots, m\}$ such that M_i solves $(\Pi, \mathcal{D})$ with success probability $1 - \delta$. For this M_i, we have $\mathbf{Pr}_{\widetilde{S}}[\widetilde{S}^{M_i, M_j}(x) = \Pi(x)] \geq 1 - 1/16m$ for every j. By the union bound, with probability at least $15/16$, $c_{i,j} = \Pi(x)$ for every j. Similarly, we also have $c_{j,i} = \Pi(x)$ for every j with probability at least $15/16$. These two properties guarantee that the output of B is equal to $\Pi(x)$. Overall, with probability at least $1 - 3/16$, the algorithm B outputs $\Pi(x)$. □

We usually set $Q(n) = \mathrm{polylog}(n)$. We now present a *completely uniform* direct product theorem for any problem that admits a $\mathrm{polylog}(n)$-query selector.

Theorem 4.5.9 (Direct Product Theorem for Any Problem with Selector)**.** *Let $k \in \mathbb{N}$, $\epsilon, \delta > 0$ be parameters that satisfy $\epsilon > \exp(-\Omega(\delta k))$. Let $(\Pi, \mathcal{D})$ be a distributional problem. Suppose there exists a $t_S(n)$-time selector S from Π to $(\Pi, \mathcal{D})$ with success probability $1 - \delta$ that calls an oracle at most $Q(n)$ times.*

Suppose that there exists a $t(n)$-time algorithm solving $(\Pi, \mathcal{D})^k$ with success probability ϵ. Then, there exists an $\mathrm{O}((t(n)Q(n)\epsilon^{-1} \log \delta^{-1} + t_S(n))\epsilon^{-2} \log \epsilon^{-1})$-time algorithm that solves Π with high probability.

Proof. Let A be a $t(n)$-time algorithm solving $(\Pi, \mathcal{D})^k$ with success probability ϵ. By using the algorithm M of Theorem 4.5.6, M^A produces a list of oracle algorithms $M_1, \cdots, M_m$ such that M_i^A computes $(\Pi, \mathcal{D})$ with success probability $1 - \delta$ for some $i \in \{1, \dots, m\}$, where $m = \mathrm{O}(1/\epsilon)$. Then, we apply Lemma 4.5.8 using $M_1^A, \dots, M_m^A$ as the list of algorithms. Note that, from Theorem 4.5.6, each M_i^A runs in time $\mathrm{O}(t(n)\epsilon^{-1} \log \delta^{-1})$ and M runs is time $t(n)/\epsilon$. Thus, the algorithm solving Π of Lemma 4.5.8 runs in time $\mathrm{O}((t(n)Q(n)\epsilon^{-1} \log \delta^{-1} + t_S(n))\epsilon^{-2} \log \epsilon^{-1})$. □

We obtain the main result (Theorem 3.1.2) by combining a selector (Corollary 4.5.5) with the direct product theorem (Theorem 4.5.9).

Proof of Theorem 4.1.2. We prove the contrapositive. Assume that there exists a $t(n)$-time algorithm that solves $(\#\textsc{Emb}^{(K_{a,b})}, \mathcal{K}_{a,b,n})^k$ with success probability $\epsilon = n^{-\alpha/4}$, where $t(n) = n^{a-\alpha}$.

By Corollary 4.5.5, there is a $\widetilde{O}(n^2)$-time selector using $Q(n) = \mathrm{polylog}(n)$ queries from $\#\textsc{Emb}^{(K_{a,b})}$ to $(\#\textsc{Emb}^{(K_{a,b})}, \mathcal{K}_{a,b,n})$ with success probability $1 - \delta$, where $\delta = (\log n)^{-C_H}$ for a constant $C_H > 0$ that depends only on H. We choose $k = \mathrm{O}(\delta^{-1} \log \epsilon^{-1}) \leq \mathrm{O}(\alpha(\log n)^{C_H+1}) = \mathrm{polylog}(n)$ large enough so that the assumption of Theorem 4.5.9 is satisfied. By Theorem 4.5.9, we obtain a $t'(n)$-time algorithm that solves $\#\textsc{Emb}^{(K_{a,b})}$, where $t'(n) = \widetilde{O}((n^2 + t(n)) \cdot n^{\alpha/2}) \leq \widetilde{O}(n^{a-\alpha/2})$. This contradicts Theorem 3.2.1. □

4.6 Fine-Grained XOR Lemma

In this section, we show a XOR lemma in the context of fine-grained complexity. We focus on the *XOR problem* $\Pi^{\oplus k}$ defined as follows.

Definition 4.6.1. *Let Π be a problem such that $\Pi(x) \in \{0, 1\}$ for any input x. For a parameter $k \in \mathbb{N}$, let $\Pi^{\oplus k}$ be the problem of computing $\sum_{i=1}^{k} \Pi(x_i) \pmod 2$ on input $(x_1, \dots, x_k)$.*

Throughout this section, we consider decision problems unless otherwise noted. For a distributional problem $(\Pi, \mathcal{D})$, let $\mathcal{D}^k$ be the direct product of $\mathcal{D}$ (see Definition 4.1.1). Suppose that there is a selector from Π to $(\Pi, \mathcal{D})$ that makes at most $\mathrm{polylog}(n)$ queries. The aim of this section is to derive the average-case hardness of the distributional problem $(\Pi^{\oplus k}, \mathcal{D}^k)$ from the worst-case hardness assumption of Π (see Theorem 4.6.3). To this end, we combine Direct Product Theorem (Theorem 4.5.9) and the well-known list-decoding technique for the Hadamard code due to Goldreich and Levin [GL89]. Let us restate the Goldreich-Levin theorem as follows.

Theorem 4.6.2 (Goldreich-Levin Theorem [GL89])**.** *Let $(\Pi, \mathcal{D})$ be a distributional problem and let $k = k(n) \in \mathbb{N}, \epsilon = \epsilon(n) > 0$ be parameters. Then, there exists an oracle algorithm M that, given an oracle A solving $(\Pi^{\oplus k}, \mathcal{D}^k)$ with success probability $1/2 + \epsilon$, produces with high probability a list of*

deterministic oracle algorithms $M_1, \dots, M_m$ such that, for some $t \in \{1, \dots, m\}$, the oracle algorithm M_t^A solves $(\Pi, \mathcal{D})^{2k}$ with success probability $\epsilon/(6\sqrt{k})$. Here, $m = O(k/\epsilon^2)$.

If an oracle A can be computed in time $T_A(n)$, then each M_i^A runs in time $O\big(T_A(n)k^{2.5}/\epsilon^2\big)$ for any i, and M runs in time $O\big(m \cdot T_A(n)k^{2.5}/\epsilon^2\big) = O\big(T_A(n)k^{3.5}/\epsilon^4\big)$.

Proof. The proof is essentially given in [GL89]. Let $(\Pi, \mathcal{D})$ be the distributional problem and $k = k(n), \epsilon = \epsilon(n)$ be the parameters mentioned in Theorem 4.6.2. Consider the following problem Π': Given $2k$ instances $x_1, \dots, x_{2k}$ of Π and $r \in \{0,1\}^{2k}$, compute $\sum_{i=1}^{2k} r_i \cdot \Pi(x_i) \bmod 2$. Let $(\Pi', \mathcal{D}')$ be the distributional problem, where, in $\mathcal{D}'$, an input $(x_1, \dots, x_{2k}, r)$ is sampled as $(x_1, \dots, x_{2k}) \sim \mathcal{D}^{2k}$ and $r \sim \mathrm{Unif}(\{0,1\}^{2k})$. Note that, if $r \sim \mathrm{Unif}(\{0,1\}^{2k})$, with probability at least $\binom{2k}{k}/2^{2k} \geq 1/(2\sqrt{k})$, the vector $r \in \{0,1\}^{2k}$ has exactly k ones (note that $\binom{2k}{k} \geq \big(1 - \frac{1}{8k}\big)\frac{4^k}{\sqrt{\pi k}}$). Conditioned on this event, the distributional problem $(\Pi', \mathcal{D}')$ is equivalent to $(\Pi^{\oplus k}, \mathcal{D}^k)$. Let A' be the algorithm that takes $2k$ instances $x_1, \dots, x_{2k}$ and $r \in \{0,1\}^{2k}$ as input and outputs $A(x_{i_1}, \dots, x_{i_k})$ if r contains exactly k ones in the position of $i_1 < \cdots < i_k$; otherwise outputs a random bit. This algorithm A' runs in time $O(t(n))$ and solves $(\Pi', \mathcal{D}')$ with success probability at least $1/2 + \epsilon/(2\sqrt{k})$. In other words,

$$\mathop{\mathbf{Pr}}_{A', x_1, \dots, x_{2k}, r \sim \mathrm{Unif}(\{0,1\}^{2k})} \left[A'(x_1, \dots, x_k, r) = \sum_{i=1}^{2k} r_i \Pi(x_i) \bmod 2 \right] \geq \frac{1}{2} + \frac{\epsilon}{2\sqrt{k}}.$$

Now we present the algorithm M mentioned in Theorem 4.6.2. We say that an input $(x_1, \dots, x_{2k})$ is *good* if

$$\mathop{\mathbf{Pr}}_{A', r \sim \mathrm{Unif}(\{0,1\}^{2k})} \left[A'(x_1, \dots, x_k, r) = \sum_{i=1}^{2k} r_i \Pi(x_i) \bmod 2 \right] \geq \frac{1}{2} + \frac{\epsilon}{4\sqrt{k}}.$$

We claim that at least $\epsilon/(4\sqrt{k})$ fraction of $(x_1, \dots, x_k)$ are good. To see this, let $\mathcal{E}$ be the event that A' success (i.e., $A'(x_1, \dots, x_{2k}, r) = \sum_{i=1}^{2k} r_i \Pi(x_i) \pmod 2$) and let $\mathcal{F}$ be the event that $(x_1, \dots, x_{2k})$ is good. Assume $\mathbf{Pr}[\mathcal{F}] < \epsilon/(4\sqrt{k})$. Then, from the property of A' and the assumption, we have

$$\begin{aligned} \frac{1}{2} + \frac{\epsilon}{2\sqrt{k}} &\leq \mathbf{Pr}[\mathcal{E}] \\ &\leq \mathbf{Pr}[\mathcal{E}|\mathcal{F}]\,\mathbf{Pr}[\mathcal{F}] + \mathbf{Pr}[\mathcal{E}|\text{not } \mathcal{F}]\,\mathbf{Pr}[\text{not } \mathcal{F}] \\ &< \frac{\epsilon}{4\sqrt{k}} + \left(\frac{1}{2} + \frac{\epsilon}{4\sqrt{k}}\right) = \frac{1}{2} + \frac{\epsilon}{2\sqrt{k}}. \end{aligned}$$

Thus we have $\mathbf{Pr}[\mathcal{F}] = \mathbf{Pr}[(x_1, \dots, x_{2k}) \text{ is good}] \geq \epsilon/(4\sqrt{k})$.

Let $m = 24k/\epsilon^2$ and ℓ be the minimum integer satisfying $m \leq 2^\ell$. The algorithm M produces a list $M_1, \dots, M_{2^\ell}$ such that, for some $i \in \{1, \dots, 2^\ell\}$, $M_i^{A'}$ solves $(\Pi, \mathcal{D})^{2k}$ for good inputs (it outputs anything for non-good inputs). Let $s^{(1)}, \dots, s^{(\ell)} \sim \mathrm{Unif}(\{0,1\}^{2k})$ be ℓ i.i.d. random vectors. Construct m distinct nonempty subsets $T_1, \dots, T_m \subseteq [\ell]$ in a canonical way and let $r^{(i)} := \sum_{j \in T_i} s_j^{(i)}$. Note that, for every $i \neq i'$, $r^{(i)}$ and $r^{(i')}$ are pairwise independent random vectors and each $r^{(i)}$ is drawn from $\mathrm{Unif}(\{0,1\}^{2k})$.

Now we present the list of oracle algorithms $M_1, \dots, M_{2^\ell}$. For each $t \in \{1, \dots, 2^\ell\}$, the algorithm $M_t^{A'}$ works as follows. Write $t = \sum_{j=1}^{\ell} 2^{j-1} w_j$ as a binary extension. In other words, $(w_1, \dots, w_\ell)$ can be seen as an ℓ-bits of advice. The bit w_j tells us the value of $\langle \Pi(x), s^{(j)} \rangle := \sum_{i=1}^{2k} \Pi(x_i) s_i^{(j)} \pmod 2$. Note that, for some t, this equality holds for all $j = 1, \dots, \ell$.

Suppose that the input $(x_1, \dots, x_{2k})$ is good. Given $(w_1, \dots, w_\ell)$, for every $i = 1, \dots, m$, $M_t^{A'}$ does the following: First, $M_t^{A'}$ computes $W^{(i)} := \sum_{j \in T_i} w_j$. Note that, for some t, we have $W^{(i)} = \sum_{j \in T_i} \langle \Pi(x), s^{(j)} \rangle = \langle \Pi(x), r^{(i)} \rangle$. Then, for every index $l \in \{1, \dots, 2k\}$, $M_t^{A'}$ calls the oracle and obtain $A'(x_1, \dots, x_{2k}, r_1^{(i)}, \dots, r_{l-1}^{(i)}, \overline{r_l^{(i)}}, r_{l+1}^{(i)}, \dots, r_{2k}^{(i)})$, where $\overline{z} := 1 - z$ for $z \in \{0,1\}$. The output $O^{(i)}$ satisfies $W^{(i)} + O^{(i)} = \Pi(x_l) \pmod 2$ if A' success. This happens with probability $1/2 + \epsilon/(4\sqrt{k})$ since $(x_1, \dots, x_{2k})$ is good. We repeat this for $Q = 96k^{1.5}/\epsilon^2$ times and then we can compute $\Pi(x_l)$ by taking the majority among the Q trials with successes probability at least $1 - \frac{1}{12k}$ for each $l = 1, \dots, 2k$. To see this, let Z_i be a binary indicator random variable such that $Z_i = 1$ if and only if $W^{(i)} + O^{(i)} = \Pi(x_l)$. Let $Z = Z_1 + \cdots + Z_Q$. It suffices to show $\mathbf{Pr}[Z > Q/2] \geq 1 - \frac{1}{3k}$. Note that $\mathbf{E}[Z] \geq \frac{Q}{2} + \frac{\epsilon Q}{4\sqrt{k}}$

and $\mathbf{Var}[Z] = \sum_{i=1}^{Q} \mathbf{Var}[Z_i] \le Q$ since the random variables Z_i are pairwise independent. From the Chebyshev inequality, we obtain

$$\begin{aligned}\mathbf{Pr}\left[Z \le \frac{Q}{2}\right] &\le \mathbf{Pr}\left[|Z - \mathbf{E}[Z]| \ge \frac{\epsilon Q}{4\sqrt{k}}\right] \\ &\le \mathbf{Pr}\left[|Z - \mathbf{E}[Z]| \ge \frac{\epsilon\sqrt{Q}}{4\sqrt{k}}\sqrt{\mathbf{Var}[Z]}\right] \\ &\le \frac{16\sqrt{k}}{\epsilon^2 Q} \le \frac{1}{6k}.\end{aligned}$$

Here, recall that the Chebyshev inequality asserts

$$\mathbf{Pr}\left[|Z - \mathbf{E}[Z]| \ge \xi\sqrt{\mathbf{Var}[Z]}\right] \le \frac{1}{\xi^2}$$

for any $\xi > 0$. Then, from the union bound over $2k$ indices, $M_t^{A'}$ (for the appropriate t) computes $(\Pi(x_1), \ldots, \Pi(x_{2k}))$ with probability at least $2/3$.

Note that $M_i^{A'}$ is deterministic without loss of generality since the coin flips can be given by M. The success probability of $M_i^{A'}$ is at least $(2/3) \cdot (\epsilon/(4\sqrt{k})) \ge \epsilon/(6\sqrt{k})$ since input $(x_1, \ldots, x_{2k})$ is good with probability at least $\epsilon/(4\sqrt{k})$. The running time of $M_i^{A'}$ is $O(Qk) = O(k^{2.5}/\epsilon^2)$ for all $i \in \{1, \ldots, m\}$. Thus, if A' is a $T_A(n)$-time algorithm, then we can construct M_i as a deterministic $O(T_A(n)k^{2.5}/\epsilon^2)$-time algorithm. The total running time of M is at most $m \cdot O(T_A(n)k^{3.5}/\epsilon^4)$ since M constructs $m = O(k/\epsilon^2)$ algorithms each of them runs in time $O(T_A(n)k^{2.5}/\epsilon^2)$. □

Now we prove the main result of this section.

Theorem 4.6.3 (XOR lemma for any problem with selector)**.** *Let $k \in \mathbb{N}, \epsilon, \delta > 0$ be parameters satisfying $\epsilon > \exp(-\Omega(\delta k))$. Let $(\Pi, \mathcal{D})$ be a distributional decision problem.*

Suppose there exists a $t_S(n)$-time selector S from Π to $(\Pi, \mathcal{D})$ with success probability $1 - \delta$ that calls an oracle at most $Q(n)$ times. Suppose that there exists a $t(n)$-time algorithm solving $(\Pi^{\oplus k}, \mathcal{D}^k)$ with success probability $\frac{1}{2} + \epsilon$.

Then, there exists a $t'(n)$-time randomized algorithm that solves Π with high probability. Here $t'(n) = O\big((t_S(n) + Q(n)t(n)) \cdot \log(1/\delta)(k/\epsilon)^8\big)$.

Proof. From the assumption, we have a $t(n)$-time algorithm A solving $(\Pi^{\oplus k}, \mathcal{D}^k)$ with success probability $\frac{1}{2} + \epsilon$. Then, by Theorem 4.6.2 using A as the oracle, we obtain a list of deterministic oracle algorithms $M_1, \ldots, M_m$ such that M_i^A solves $(\Pi, \mathcal{D})^{2k}$ with success probability $\Omega\big(\epsilon/\sqrt{k}\big)$ for some $i \in \{1, \ldots, m\}$, where $m = O(k/\epsilon^2)$. Each of M_i runs in time $O(t(n)k^{2.5}\epsilon^{-2})$ if we take the running time of A into account. This list can be constructed in time $O(t(n)k^{3.5}\epsilon^{-4})$.

Let $\delta > 0$ be the parameter mentioned in Theorem 4.6.3. For each $i \in \{1, \ldots, m\}$, apply Theorem 4.5.6 using M_i as the oracle. This yields a list $M_{i,1}, \ldots, M_{i,m'}$ of deterministic algorithms for each $i \in \{1, \ldots, m\}$, where $m' = O(\epsilon^{-1})$. Moreover, if M_{i^*} solves $(\Pi, \mathcal{D})^{2k}$, then M_{i^*,j^*} solves Π with success probability $1 - \delta$ for some $j^* \in \{1, \ldots, m'\}$. For every i, j, $M_{i,j}$ runs in time $O(t(n)k^{2.5}\epsilon^{-2} \cdot \epsilon^{-1} \log \delta^{-1}) \le O(t(n)k^{2.5}\epsilon^{-3} \log \delta^{-1})$.

Now we have a list $(M_{i,j})$ of $mm' = O(k\epsilon^{-3})$ deterministic algorithms. From Lemma 4.5.8, there exists an algorithm B that solves Π with high probability. The overall running time of B is at most $O\big((mm')^2(t_S(n) + Q(n)t(n)k^{2.5}\epsilon^{-3}) \log(\sqrt{k}/\epsilon) \log(mm')\big) \le O\big((t_S(n) + Q(n)t(n)) \cdot (k/\epsilon)^8\big)$. □

4.6.1 Application 1: $\oplus\mathsf{EMB}_{\mathrm{col}}^{(H)}$

Let H be a fixed graph. Consider the problem $\oplus\textsc{Emb}_{\mathrm{col}}^{(H)}$ of computing the parity of $\#\textsc{Emb}_{\mathrm{col}}^{(G)}(H)$ for a given graph G. For a parameter k, let $\biguplus^k \mathcal{G}_{n,1/2}^{(H)}$ be the distribution of random graphs that is a direct sum of k i.i.d. graphs drawn from $\mathcal{G}_{n,1/2}^{(H)}$. That is, let $G_1, \ldots, G_k \sim \mathcal{G}_{n,1/2}^{(H)}$ be i.i.d. random graphs. Suppose that $G(V_i) \cap G(V_j) = \emptyset$ for any $i \ne j$. Then, the graph G defined by $V(G) = \bigcup_{i=1}^k V(G_i)$ and $E(G) = \bigcup_{i=1}^k E(G_i)$ forms the distribution $\biguplus^k \mathcal{G}_{n,1/2}^{(H)}$. Let $\textsc{Emb}_{\mathrm{col}}^{(H)}$ be the decision problem in which we are asked to decide whether $\#\textsc{Emb}_{\mathrm{col}}^{(H)}(G) > 0$ or not for a given graph G. This subsection is devoted to the proof of the following result.

Theorem 4.6.4. *Suppose that there exists a $t(n)$-time algorithm solving the distributional problem $(\oplus\mathrm{EMB}^{(H)}_{\mathrm{col}}, \biguplus^k \mathcal{G}^{(H)}_{n,1/2})$ with success probability $\frac{1}{2}+\epsilon$ far any $k = \mathrm{polylog}(n)$. Then, there exists a $t(n)\cdot(\log n/\epsilon)^{O(1)}$-time randomized algorithm solving $\mathrm{EMB}^{(H)}_{\mathrm{col}}$ with probability $2/3$.*

The proof of Theorem 4.6.4 consists of the following three steps. First, we present a randomized reduction of $\mathrm{EMB}^{(H)}_{\mathrm{col}}$ to $\oplus\mathrm{EMB}^{(H)}_{\mathrm{col}}$ in the worst-case sense. Then, we check that the parity problem $\oplus\mathrm{EMB}^{(H)}_{\mathrm{col}}$ admits a $\widetilde{O}(n^2)$-time selector with $\mathrm{polylog}(n)$ queries. Finally, we apply Theorem 4.6.3 to boost the error tolerance. The second and third steps imply Theorem 4.2.5. More specifically, we obtain the following.

Theorem 4.6.5 (Refinement of Theorem 4.2.5)**.** *Let H be an arbitrary graph. Suppose that there exists a $T(n)$-time randomized algorithm that solves $(\oplus\mathrm{EMB}^{(H)}_{\mathrm{col}}, \biguplus^k \mathcal{G}^{(H)}_{n,1/2})$ with success probability greater than $\frac{1}{2}+\epsilon$ for any $k = O(\log \epsilon^{-1})$. Then, there exists a $T(n)(\log n/\epsilon)^{O(1)}$-time randomized algorithm that solves $\oplus\mathrm{EMB}^{(H)}_{\mathrm{col}}$ with probability at least $2/3$ for any input.*

Remark 4.6.6. Theorem 4.2.5 immediately follows from Theorem 4.6.5 (substitute $\epsilon = n^{-c}$ to Theorem 4.6.5).

Parity vs. Detection.

Lemma 4.6.7. *Suppose that there exists a $t(n)$-time randomized algorithm solving $\oplus\mathrm{EMB}^{(H)}_{\mathrm{col}}$ for any input with probability at least $2/3$. Then, there exists a $t'(n)$-time randomized algorithm that solves $\mathrm{EMB}^{(H)}_{\mathrm{col}}$ with probability at least $2/3$. Here, $t'(n) = O(2^{|E(H)|}t(n))$.*

Proof. The proof is essentially given in Appendix A of [BABB19]. For completeness, we present the proof. Consider the polynomial $P_G : \mathbb{F}_2^{E(G)} \to \mathbb{F}_2$ defined as

$$P_G(x) := \sum_{\substack{F\subseteq E(G):\\ F \text{ is isomorphic to } H}} \prod_{e\in F} x_e.$$

Then, G does not contain H if and only if $P_G(\cdot) \equiv 0$. The degree of P_G is $|E(H)|$. Moreover, if $P_G(\cdot) \not\equiv 0$, then $P_G(z) = 1$ for at least $2^{-|E(H)|}$ fraction of $z \in \mathbb{F}_2^{|E(G)|}$ (see, e.g., Lemma 2.6 of [NS94]).

We present an algorithm that solves $\mathrm{EMB}^{(H)}_{\mathrm{col}}$ using an oracle for $\oplus\mathrm{EMB}^{(H)}_{\mathrm{col}}$. Let $m = 100\cdot 2^{|E(H)|}$ and sample m i.i.d. random vectors $z_1,\dots,z_m \sim \mathrm{Unif}(\mathbb{F}_2^{E(G)})$. Then, compute $P_G(z_1),\dots,P_G(z_m)$. If $P_G(z_i) = 1$ for some i, output YES. Otherwise, output NO. Note that one can compute $P_G(\cdot)$ by solving $\oplus\mathrm{EMB}^{(H)}_{\mathrm{col}}$ since $P_G(\cdot)$ is a polynomial over $\mathbb{F}_2$.

If G does not contain H, the algorithm outputs NO with probability 1. If G contains H, the probability that the algorithm outputs NO is at most $(1-2^{-|E(H)|})^m \le \mathrm{e}^{-100}$. □

Selector for $\oplus\mathsf{EMB}^{(H)}_{\mathrm{col}}$.

Theorem 4.6.8. *There exists a selector S from $\oplus\mathrm{EMB}^{(H)}_{\mathrm{col}}$ to $(\oplus\mathrm{EMB}^{(H)}_{\mathrm{col}}, \mathcal{G}^{(H)}_{n,1/2})$ with success probability $1-1/\mathrm{polylog}(n)$ such that (1) S runs in time $\widetilde{O}(n^2)$, and (2) The number of oracle accesses is at most $\mathrm{polylog}(n)$.*

Proof. The proof is essentially the same as that of Theorem 4.5.4. First, we encode $\oplus\mathrm{EMB}^{(H)}_{\mathrm{col}}$ to the low-degree polynomial $\mathsf{EMBCOL}_{n,H,\mathbb{F}_{2^t}}$. Note that, since $\mathbb{F}_{2^t}$ has characteristic 2 (i.e., $a+a=0$ for any $a\in\mathbb{F}_{2^t}$), computing the polynomial $\mathsf{EMBCOL}_{n,H,\mathbb{F}_{2^t}}(x)$ for $x\in\{0,1\}^{E(H)\times K_n}$ is equivalent to solving $\oplus\mathrm{EMB}^{(H)}_{\mathrm{col}}$ by regarding the input x as the edge indicator of a graph. Using Corollary 3.4.2, we reduce computing $\mathsf{EMBCOL}_{n,H,\mathbb{F}_{2^t}}$ to solving the distributional problem $(\mathsf{EMBCOL}_{n,H,\mathbb{F}_{2^t}}, \mathcal{U}^{(H)}_n(\mathbb{F}_{2^t}))$. Moreover, we can reduce $(\mathsf{EMBCOL}_{n,H,\mathbb{F}_{2^t}}, \mathcal{U}^{(H)}_n(\mathbb{F}_{2^t}))$ to $(\mathsf{EMBCOL}_{n,H,\mathbb{F}_{2^t}}, \mathcal{U}^{(H)}_n(\mathbb{F}_2))$ with query complexity $(\log n)^{O(|E(H)|)}$ using the technique of [BABB19]. This yields a worst-case-to-average-case reduction for $\oplus\mathrm{EMB}^{(H)}_{\mathrm{col}}$ (c.f., Theorem 3.2.5).

Similarly, a slight modification of the interactive proof system IP of Theorem 4.2.4 works for $\oplus\mathrm{EMB}^{(H)}_{\mathrm{col}}$. To be more specifically, let us consider an interactive proof system IP' for the statement "$\oplus\mathrm{EMB}^{(H)}_{\mathrm{col}}(G) =$

b". The protocol IP' is the same as IP except for using $\mathbb{F}_{2^t}$ instead of $\mathbb{F}_q$. Note that the equation (4.2) holds even for $\mathsf{EMBCOL}_{n,H,\mathbb{F}_{2^t}}$. Moreover, computing the polynomial $\mathsf{EMBCOL}_{n,H,\mathbb{F}_{2^t}}$ can be reduced to computing $\mathsf{EMBCOL}_{n,H,\mathbb{F}_2}$ using the aforementioned technique of [BABB19].

Using the interactive proof system IP' for $\oplus\mathrm{EMB}^{(H)}_{\mathrm{col}}$, we can construct an $\widetilde{O}(n^2)$-time $\mathrm{polylog}(n)$-query instance checker C' for $\oplus\mathrm{EMB}^{(H)}_{\mathrm{col}}$ (see Theorem 4.5.3). Combining the instance checker C' and the worst-case-to-average-case reduction, we can construct the desired selector (see the proof of Theorem 4.5.4). □

XOR lemma for $\oplus\mathsf{EMB}^{(H)}_{\mathrm{col}}$ (proof of Theorem 4.6.5). Assume that there exists a $t(n)$-time algorithm A that solves $(\oplus\mathrm{EMB}^{(H)}_{\mathrm{col}}, \biguplus^k \mathcal{G}^{(H)}_{n,1/2})$ with success probability ϵ. Note that the distributional problem $(\bigoplus(\oplus\mathrm{EMB}^{(H)}_{\mathrm{col}})^k, (\mathcal{G}^{(H)}_{n,1/2})^k)$ is equivalent to the distributional problem $(\oplus\mathrm{EMB}^{(H)}_{\mathrm{col}}, \biguplus^k \mathcal{G}^{(H)}_{n,1/2})$. Hence, the algorithm A also solves $(\bigoplus(\oplus\mathrm{EMB}^{(H)}_{\mathrm{col}})^k, (\mathcal{G}^{(H)}_{n,1/2})^k)$. From Theorem 4.6.8, there exists an $\widetilde{O}(n^2)$-time selector using $\mathrm{polylog}(n)$ oracle accesses from $\oplus\mathrm{EMB}^{(H)}_{\mathrm{col}}$ to $(\oplus\mathrm{EMB}^{(H)}_{\mathrm{col}}, \mathcal{G}^{(H)}_{n,1/2})$ with success probability $1-(\log n)^{-C}$ for some constant $C>0$ that depends only on H. Let $\delta=(\log n)^{-C}$ and $k=k(n)$ be parameters such that the assumptions of Theorem 4.6.3 is satisfied. Note that we can set $k=(\log n)^{O(C)}=\mathrm{polylog}(n)$. Then, by Theorem 4.6.3, we have an $t'(n)$-time ranndomized algorithm that solves $\oplus\mathrm{EMB}^{(H)}_{\mathrm{col}}$ with high probability, where $t'(n)=\widetilde{O}((n^2+t(n))\cdot(k/\epsilon)^8=t(n)\cdot(\log n/\epsilon)^{O(1)}$ (here we assume $t(n)\geq n^2$).

Proof of Theorem 4.6.4. We combine Lemma 4.6.7 and Theorem 4.6.5. Suppose that there exists a $t(n)$-time algorithm solving $(\oplus\mathrm{EMB}^{(H)}_{\mathrm{col}}, \biguplus^k \mathcal{G}^{(H)}_{n,1/2})$ with success probability ϵ. From Theorem 4.6.5, there exists a $t(n)\cdot(\log n/\epsilon)^{O(1)}$-time randomized algorithm for $\oplus\mathrm{EMB}^{(H)}_{\mathrm{col}}$. Then, from Lemma 4.6.7, we obtain a $2^{|E(H)|}\cdot t(n)\cdot(\log n/\epsilon)^{O(1)}$-time randomized algorithm for $\mathrm{EMB}^{(H)}_{\mathrm{col}}$.

4.6.2 Application 2: $\oplus K_a$-Subgraph

Recall that $\oplus K_a$-Subgraph is the problem of computing the parity of the number of K_a subgraphs contained in a given graph. This subsection is devoted to the proof of Theorem 4.1.3. Recall that $\mathcal{G}(n,1/2)$ is the distribution of the Erdős–Rényi graph $G(n,1/2)$, and $\biguplus^k \mathcal{G}(n,1/2)$ is the distribution of the disjoint union of k random graphs $G_1,\dots,G_k$ each of which is independently drawn from $\mathcal{G}(n,1/2)$.

Theorem 4.6.9 (Refinement of Theorem 4.1.3)**.** *Suppose that there exists a $T(n)$-time randomized algorithm that solves $(\oplus K_a\text{-Subgraph}, \biguplus^k \mathcal{G}(n,1/2))$ with success probability $\frac{1}{2}+\varepsilon$ for any $k=\mathrm{polylog}(n)$. Then, there exists a $T(n)(\log n/\varepsilon)^{O(1)}$-time randomized algorithm that solves $\oplus K_a$-Subgraph for any input with probability $2/3$.*

Proof of Theorem 4.1.3. Theorem 4.6.9 directly implies Theorem 4.1.3 (let $\varepsilon=n^{-\epsilon}$). □

The core of the proof of Theorem 4.6.9 is the existence of the following efficient selector.

Lemma 4.6.10. *There exists an $\widetilde{O}(n^2)$-time selector S from $\oplus K_a$-Subgraph to the distributional problem $(\oplus K_a\text{-Subgraph}, \mathcal{G}(n,1/2))$ with success probability $1-1/\mathrm{polylog}(n)$. Moreover, the number of oracle accesses of S is at most $\mathrm{polylog}(n)$.*

Proof. The proof of Lemma 4.6.10 is similar to that of Corollary 4.5.5. From Theorem 4.6.8, we have obtain a selector from $\oplus\mathrm{EMB}^{(K_a)}_{\mathrm{col}}$ to $(\oplus\mathrm{EMB}^{(K_a)}_{\mathrm{col}}, \mathcal{G}^{(K_a)}_{n,1/2})$. Then, we use the reduction by Boix-Adserá, Brennan, and Bresler [BABB19]. They reduced $(\oplus\mathrm{EMB}^{(K_a)}_{\mathrm{col}}, \mathcal{G}^{(K_a)}_{n,1/2})$ to $(\oplus K_a\text{-Subgraph}, \mathcal{G}(n,1/2))$ with preserving the success probability up to a constant factor (Lemma 3.10 of [BABB19]). Using their reduction, each query of the selector S can be replaced by the reduction. This yields a selector from $\oplus\mathrm{EMB}^{(K_a)}_{\mathrm{col}}$ to $(\oplus K_a\text{-Subgraph}, \mathcal{G}(n,1/2))$. We then use the reduction of Lemma 3.3 of Boix-Adserá, Brennan, and Bresler [BABB19]. They reduced $\oplus K_{a,b}$-Subgraph to $\oplus\mathrm{EMB}^{(K_a)}_{\mathrm{col}}$. Specifically, if $\oplus\mathrm{EMB}^{(K_a)}_{\mathrm{col}}$ can be solved in time $t(n)$, then there exists a $t(n)+O(n^2)$-time algorithm for $\oplus K_a$-Subgraph. □

Proof of Theorem 4.6.9. Suppose that there exists a $T(n)$-time randomized algorithm that solves the distributional problem $(\oplus K_a\text{-Subgraph}, \biguplus^k \mathcal{G}(n,1/2))$ with success probability $\frac{1}{2}+\epsilon$ for any $k=\mathrm{polylog}(n)$. Note that $(\oplus K_a\text{-Subgraph}, \biguplus^k \mathcal{G}(n,1/2))$ is equivalent to $((\oplus K_{a,b}\text{-Subgraph})^{\oplus k}, (\mathcal{G}(n,1/2))^k)$. From Theorem 4.6.3 and Lemma 4.6.10, we obtain a $t'(n)$-time randomized algorithm solving $\oplus K_{a,b}$-Subgraph

with probability 2/3, where $t'(n) = (n^2 + t(n)(\log n/\epsilon)^{O(1)} = t(n) \cdot (\log n/\epsilon)^{O(1)}$ (here, we assume $t(n) = \Omega(n^2)$ and let $\delta = 1 - (\log n)^{-C}$) and $k = (\log n)^{O(1)}$.

□

Chapter 5

Functional Voting

5.1 Model

In this chapter, we introduce *functional voting*, which contains the pull voting, best-of-two, and best-of-three as special cases. Then we investigate basic properties of functional voting. The model in this chapter will play a central role in Chapters 6 and 7.

Definition 5.1.1 (Functional voting process)**.** *Let V be a finite set. Let $f\colon \mathbb{R} \to \mathbb{R}$ be a function satisfying $f([0,1]) = f([0,1])$, $f(0) = 0$, and $f(1) = 1$. Let $P \in [0,1]^{V\times V}$ be a transition matrix over V. For a fixed subset $A \subseteq V$, let $B = V \setminus A$ and let $(X_v)_{v\in V}$ be independent binary random variables defined as*

$$\begin{aligned} \mathbf{Pr}[X_v = 1] &= f\big(P(v,A)\big) \quad \text{if } v \in B, \\ \mathbf{Pr}[X_v = 0] &= f\big(P(v,B)\big) \quad \text{if } v \in A. \end{aligned} \tag{5.1}$$

For $A \subseteq V$ and (X_v) above, define $A' = \{v \in V : X_v = 1\}$. A functional voting with respect to f *is the Markov chain $(A_t)_{t\in\mathbb{Z}_{\geq 0}}$ over 2^V given by $A_{t+1} = (A_t)'$. We call the function f a* betrayal function*.*

In particular, a functional voting with respect to f on a graph G *is the functional voting with respect to f where transition matrix P is given as the simple random walk (2.1) on G.*

Note that, if $A_t \in \{\emptyset, V\}$ then $A_{t+1} = A_t$ since $f(0) = 0$ and $f(1) = 1$.

Definition 5.1.2. *Consider a functional voting on V. For $A \subseteq V$, the* consensus time*, denoted by $T_{\mathrm{cons}}(A)$, is defined as*

$$T_{\mathrm{cons}}(A) := \min\{t \geq 0 \,:\, A_t \in \{\emptyset, V\},\, A_0 = A\}\,.$$

Definition 5.1.3 (Best-of-two and best-of-three)**.** *The* best-of-two *is a functional voting with respect to $f_{\mathrm{Bo2}}\colon x \mapsto x^2$. The* best-of-three *is a functional voting with respect to $f_{\mathrm{Bo3}}\colon x \mapsto 3x^2 - 2x^3$.*

Remark 5.1.4. The definitions of best-of-two and best-of-three on a graph given in Definition 5.1.3 coincide with the best-of-two and best-of-three introduced in Section 1.2.1. Note that the pull voting is a functional voting with respect to $f_{\mathrm{pull}}\colon x \to x$.

5.2 Previous Works on Voting Processes

In this part, we review previous results concerning voting processes on graphs.

The pull voting is a simple and well-studied voting process [NIY99, HP01]. The pull voting and its variants have been studied as a model of *interactive particle systems* in the context of statistical physics. In the context of voting process, Hassin and Peleg [HP01] showed that the expected consensus time is $\mathrm{O}(n^3 \log n)$ for any non-bipartite graphs and any initial opinion configuration, where n is the number of vertices. Note that, for bipartite graphs, there is an initial opinion configuration that never reaches consensus. Cooper, Elsässer, Ono, and Radzik [CEOR13] improved the bound of Hassin and Peleg [HP01] by showing that the expected consensus time is $\mathrm{O}(n^3)$. Cooper and Rivera [CR16] proposed the *linear voter model*, that is a generalization of the pull voting, push voting and several other voting processes.

Doerr Goldberg, Minder, Sauerwald, and Scheideler. [DGM+11] introduced best-of-two and showed that, for complete graphs initially involving two possible opinions, the consensus time of best-of-two is $O(\log n)$ with high probability. Since best-of-two reaches consensus much faster than the pull voting, the study of best-of-two has gathered special attention in the area of distributed computing.

Motivated by the application in distributed computing, several researchers have studied voting process on complete graphs initially involving $k \geq 2$ opinions [BCN+16, BCN+17a, BCE+17, GL18]. Becchetti, Clemanti, Natale, Pasquale, and Trevisan [BCN+16] introduced the best-of-three and obtained an upper bound on the consensus time of the process on complete graphs. Interestingly, it is known that best-of-three outperforms best-of-two in the multi-opinion setting: Berenbrink, Clemanti, Elsässer, Kling, Mallmann-Trenn, and Natale [BCE+17] proved that, on the complete graph, best-of-three reaches consensus within $O(n^{3/4}(\log n)^{7/8})$ rounds from any initial configuration (k is arbitrary), while best-of-two requires $\Omega(n/\log n)$ rounds to reach consensus for some initial configuration. Ghaffari and Lengler [GL18] proved that the consensus time of the best-of-three on the complete graph is $\widetilde{O}(n^{2/3})$, where the term "$\widetilde{O}(\cdot)$" hides a $(\log n)^{O(1)}$ factor.

Several researchers considered the best-of-two and best-of-three on general graphs. Cruciani, Natale, Nusser, and Scornavacca [CNNS18] studied best-of-two on the core periphery network. Cruciani, Natale, and Scornavacca [CNS19] studied best-of-two on the (a, b)-*regular stochastic block model*, which is a graph consisting of two a-regular graphs connected by a b-regular bipartite graph. Kang and Rivera [KR19] considered the best-of-three on graphs with minimum degree n^γ for $\gamma = \Omega((\log\log n)^{-1})$. Under the assumption that the initial configuration is randomly sampled from a biased distribution, they proved that the process reaches consensus within $O(\log\log n)$ rounds. There is a line of works that studies these voting processes on *expander graphs* [CEOR13, CER14, CER+15] (see Section 7.1.1 for details).

The *best-of-k* is a natural generalization of pull voting, best-of-two and best-of-three. In each round, every vertex v randomly selects k neighbors (with replacement) and then if at least $\lfloor k/2 \rfloor + 1$ of them have the same opinion, the vertex v adopts it. Note that the best-of-1 is equivalent to pull voting. Abdullah and Draief [AD15] studied a variant of best-of-k ($k \geq 5$ is odd) on a specific class of sparse graphs that contains the random regular graph $G_{n,d}$ of $d = o(\sqrt{\log n})$ with a random initial configuration. To the best of our knowledge, best-of-k has not been studied explicitly so far.

In *Majority* (a.k.a. *local majority*), each vertex v updates its opinion to match the majority opinion among the neighbors. This simple model has been extensively studied in previous works [BCO+16, Ber01, GZ18, Pel98, Pel02, Zeh18]. For example, Majority on certain families of graphs including the Erdős–Rényi random graph [BCO+16, Zeh18], random regular graphs [GZ18] have been investigated. See [Pel02] for further results.

5.3 Basic Property

In this section, we explore basic properties of a functional voting on an n-vertex graph. For fixed $A \subseteq V$, let $\alpha = |A|/n$ and $\alpha' = |A'|/n$. Note that α' is a random variable that can be written as the sum of independent random variables: $\alpha' = \frac{1}{n}\sum_{v \in V} \mathbb{1}_{v \in A'}$. Hence, from the Hoeffding bound (Proposition 5.4.3), we obtain the following.

Proposition 5.3.1 (Concentration of α')**.** *For any $\kappa > 0$,*

$$\mathbf{Pr}\left[|\alpha' - \mathbf{E}[\alpha']| \geq \kappa\right] \leq 2\exp\left(-\frac{2\kappa^2}{n}\right).$$

In particular, it holds w.h.p. that $\alpha' = \mathbf{E}[\alpha'] + \mathrm{O}(\sqrt{\log n/n})$.

Consider a functional voting on the n-vertex complete graph with self-loops (that is, the transition matrix P satisfies $P_{u,v} = 1/n$ for all $u, v \in V$). Then, we can write $\mathbf{E}[\alpha']$ as

$$\begin{aligned}\mathbf{E}[\alpha'] &= \alpha + \frac{1}{n}\sum_{v \in B} \mathbf{Pr}[v \in A'] - \frac{1}{n}\sum_{v \in A} \mathbf{Pr}[v \in B'] \\ &= \alpha + (1-\alpha)f(\alpha) - \alpha f(1-\alpha).\end{aligned}$$

Therefore, we can view the sequence $(\alpha_t)_{t \in \mathbb{Z}_{\geq 0}}$ ($\alpha_t = |A_t|/n$) as an iteration of applying the mapping $x \mapsto x + (1-x)f(x) - xf(1-x)$.

5.4 Tool

We present inequalities that will be used in Chapters 6 and 7.

5.4.1 Concentration inequalities

Proposition 5.4.1 (A variant of the Chernoff bound; Corollary 1.10.4 of [DN20]). *Let $X_1, X_2, \dots, X_n$ be independent random variables taking values in $[0,1]$. Let $X = \sum_{i\in[n]} X_i$. Then, for any $k \geq 2\mathrm{e}\,\mathbf{E}[X]$, we have*

$$\mathbf{Pr}\left[X \geq k\right] \leq 2^{-k}.$$

Proposition 5.4.2 (Additive Chernoff bound; Theorems 10.10 and 10.11 of [DN20]). *Let $X_1, X_2, \dots, X_n$ be independent random variables taking values in $[0,1]$. Let $X = \sum_{i\in[n]} X_i$. Then for any $\delta \geq 0$,*

$$\mathbf{Pr}[X \geq \mathbf{E}[X] + \delta] \leq \exp\left(-\frac{1}{3}\min\left\{\frac{\delta^2}{\mathbf{E}[X]}, \delta\right\}\right),$$
$$\mathbf{Pr}[X \leq \mathbf{E}[X] - \delta] \leq \exp\left(-\frac{\delta^2}{2\,\mathbf{E}[X]}\right).$$

Proposition 5.4.3 (Hoeffding bound; Theorem 10.9 of [DN20]). *Let $X_1, X_2, \dots, X_n$ be independent random variables. Assume that each X_i takes values in a real interval $[a_i, b_i]$ of length $c_i := b_i - a_i$. Let $X = \sum_{i\in[n]} X_i$. Then for any $\delta > 0$,*

$$\mathbf{Pr}\left[X \geq \mathbf{E}[X] + \delta\right] \leq \exp\left(-\frac{2\delta^2}{\sum_{i\in[n]} c_i^2}\right),$$
$$\mathbf{Pr}\left[X \leq \mathbf{E}[X] - \delta\right] \leq \exp\left(-\frac{2\delta^2}{\sum_{i\in[n]} c_i^2}\right).$$

Corollary 5.4.4. *Let $X_1, X_2, \dots, X_n$ be independent random variables. Assume that each X_i takes values in a real interval $[a_i, b_i]$ of length $c_i := b_i - a_i$. Let $X = \sum_{i\in[n]} X_i$. Then, for any $\delta > 0$,*

$$\mathbf{Pr}\left[|X| \geq |\mathbf{E}[X]| + \delta\right] \leq 2\exp\left(-\frac{2\delta^2}{\sum_{i\in[n]} c_i^2}\right),$$
$$\mathbf{Pr}\left[|X| \leq \left|\mathbf{E}[X]\right| - \delta\right] \leq 2\exp\left(-\frac{2\delta^2}{\sum_{i\in[n]} c_i^2}\right).$$

Proof. For the first inequality, it is straightforward to see that

$$\begin{aligned}\mathbf{Pr}\left[|X| \geq |\mathbf{E}[X]| + \delta\right] &= \mathbf{Pr}\left[|X| - |\mathbf{E}[X]| \geq \delta\right] \leq \mathbf{Pr}\left[|X - \mathbf{E}[X]| \geq \delta\right] \\ &\leq 2\exp\left(-\frac{2\delta^2}{\sum_{i\in[n]} c_i^2}\right).\end{aligned}$$

Note that $|x| - |y| \leq |x - y|$ for any $x, y \in \mathbb{R}$. Similarly, it holds that

$$\begin{aligned}\mathbf{Pr}\left[|X| \leq |\mathbf{E}[X]| - \delta\right] &= \mathbf{Pr}\left[|\mathbf{E}[X]| - |X| \geq \delta\right] \leq \mathbf{Pr}\left[|\mathbf{E}[X] - X| \geq \delta\right] \\ &\leq 2\exp\left(-\frac{2\delta^2}{\sum_{i\in[n]} c_i^2}\right),\end{aligned}$$

and we obtain the claim. □

Proposition 5.4.5 (The Janson inequality; Theorem 21.12 of [FK16]). *Let $I_1, \dots, I_M$ be independent binary random variables and let $F_1, F_2, \dots, F_N$ be subsets of $[M]$. Consider a random variable Y defined as*

$$Y := \sum_{i\in[N]} \prod_{e\in F_i} I_e.$$

Then, it holds for any $t \le \mathbf{E}[Y]$ that

$$\mathbf{Pr}\left[Y \le \mathbf{E}[Y] - t\right] \le \exp\left(-\frac{t^2}{2\nabla}\right)$$

where

$$\nabla := \sum_{\substack{i\in N, j\in N:\\ F_i\cap F_j\neq\emptyset}} \mathbf{E}\left[\left(\prod_{e\in F_i} I_e\right)\left(\prod_{e'\in F_j} I_{e'}\right)\right].$$

Proposition 5.4.6 (The Kim-Vu concentration inequality; Main Theorem of [KV00])**.** *For a given set $[M] = \{1, 2, \ldots, M\}$, let $I_1, I_2, \ldots, I_M$ be independent binary random variables. Now, let $\mathcal{E} \subseteq 2^{[M]}$ be a collection of subsets of $[M]$ and let*

$$Y := \sum_{F\in\mathcal{E}} w(F) \prod_{e\in F} I_e,$$

where $w(F)$ are positive coefficients. For a subset $A \subseteq [M]$, define Y_A as

$$Y_A = \sum_{\substack{F\in\mathcal{E}:\\ F\supseteq A}} w(F) \prod_{e\in F\setminus A} I_e.$$

If the polynomial Y has degree at most k (i.e., $\max_{F\in\mathcal{E}} |F| \le k$), then for any positive $\lambda > 1$, it holds that

$$\mathbf{Pr}\left[|Y - \mathbf{E}[Y]| \ge \sqrt{k! \max_{A\subseteq[M]} \mathbf{E}[Y_A] \max_{A\subseteq[M]:A\neq\emptyset} \mathbf{E}[Y_A]}(8\lambda)^k\right] \le 2\exp(2 + (k-1)\log M - \lambda).$$

5.4.2 Other inequalities

Proposition 5.4.7 (The Berry-Esseen theorem; Theorem 1 of [She10])**.** *Let $X_1, X_2, \ldots, X_n$ be independent random variables such that $\mathbf{E}[X_i] = 0$, $\mathbf{E}[X_i^2] > 0$, $\mathbf{E}[|X_i|^3] < \infty$ for all $i \in [n]$ and $\sum_{i\in[n]} \mathbf{E}[X_i^2] = 1$. Let $X = \sum_{i\in[n]} X_i$ and let $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \mathrm{e}^{-y^2/2}\mathrm{d}y$ (the cumulative distribution function of the standard normal distribution). Then*

$$\sup_{x\in\mathbb{R}} \left|\mathbf{Pr}\left[X \le x\right] - \Phi(x)\right| \le 5.6 \sum_{i\in[n]} \mathbf{E}[|X_i|^3].$$

Corollary 5.4.8. *Let $X_1, X_2, \ldots, X_n$ be independent random variables and let $X := \sum_{i\in[n]} X_i$. Suppose that $\mathbf{Var}[X] \neq 0$ and $|X_i - \mathbf{E}[X_i]| \le c_i < \infty$ for all $i \in [n]$. Then for any $x \in \mathbb{R}$,*

$$\left|\mathbf{Pr}\left[\frac{X - \mathbf{E}[X]}{\sqrt{\mathbf{Var}[X]}} \le x\right] - \Phi(x)\right| \le \frac{5.6C}{\sqrt{\mathbf{Var}[X]}}.$$

Proof. For all $i \in [n]$, let

$$Z_i := \frac{X_i - \mathbf{E}[X_i]}{\sqrt{\mathbf{Var}[X]}},$$
$$Z := \sum_{i\in[n]:\mathbf{E}[Z_i^2]>0} Z_i = \sum_{i\in[n]} Z_i.$$

Note that $\mathbf{E}[Z_i^2] = 0 \iff \sum_z z^2 \mathbf{Pr}[Z_i = z] = 0 \iff \mathbf{Pr}[Z_i = 0] = 1$. Then, for all $i \in \{j \in [n] : \mathbf{E}[Z_j^2] > 0\}$, it is easy to check that $\mathbf{E}[Z_i] = 0$, $\mathbf{E}[Z_i^2] > 0$, and $\mathbf{E}[|Z_i|^3] \le \frac{C^3}{\mathbf{Var}[X]^{3/2}} < \infty$. Furthermore,

$$\sum_{i\in[n]:\mathbf{E}[Z_i^2]>0} \mathbf{E}[Z_i^2] = \sum_{i\in[n]} \mathbf{E}[Z_i^2] = \frac{\sum_{i\in[n]} \mathbf{E}[(X_i - \mathbf{E}[X_i])^2]}{\mathbf{Var}[X]} = 1.$$

Thus we can apply Proposition 5.4.7 to Z and it holds that

$$\begin{aligned}\left|\mathbf{Pr}\left[\frac{X-\mathbf{E}[X]}{\sqrt{\mathbf{Var}[X]}}\le x\right]-\Phi(x)\right| &= \left|\mathbf{Pr}\left[\sum_{i\in[n]} Z_i \le x\right]-\Phi(x)\right| \\ &= \left|\mathbf{Pr}\left[Z\le x\right]-\Phi(x)\right| \\ &\le 5.6 \sum_{i\in[n]:\mathbf{E}[Z_i^2]>0} \mathbf{E}[|Z_i|^3] \\ &\le \frac{5.6C}{\sqrt{\mathbf{Var}[X]}}\sum_{i\in[n]}\mathbf{E}[Z_i^2] = \frac{5.6C}{\sqrt{\mathbf{Var}[X]}}.\end{aligned}$$

□

Corollary 5.4.9. *Let $X_1, X_2, \ldots, X_n$ be independent random variables, $c = (c_1, \ldots, c_n) \in \mathbb{R}^n$ be a vector, and $X := \sum_{i\in[n]} X_i$. Suppose that, for all $i \in [n]$, $|X_i - \mathbf{E}[X_i]| \le c_i < \infty$ and $\mathbf{Var}[X] > 0$. Let $\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} \mathrm{e}^{-y^2/2}\mathrm{d}y$. Then, for any positive $x \in \mathbb{R}$,*

$$\mathbf{Pr}\left[|X| \le x\sqrt{\mathbf{Var}[X]}\right] \le \Phi(x) + \frac{5.6\|c\|_3^3}{\mathbf{Var}[X]^{3/2}}.$$

Proof. For each $i \in [n]$, let

$$Z_i := \frac{X_i - \mathbf{E}[X_i]}{\sqrt{\mathbf{Var}[X]}}, \quad Z := \sum_{i\in[n]:\mathbf{E}[Z_i^2]>0} Z_i = \sum_{i\in[n]} Z_i.$$

For any $i \in [n]$ satisfying $\mathbf{E}[Z_i^2] > 0$, it is easy to check that $\mathbf{E}[Z_i] = 0$, $\mathbf{E}[Z_i^2] > 0$, and $\mathbf{E}[|Z_i|^3] \le \frac{c_i^3}{\mathbf{Var}[X]^{3/2}} < \infty$. Furthermore,

$$\sum_{i\in[n]:\mathbf{E}[Z_i^2]>0}\mathbf{E}[Z_i^2] = \sum_{i\in[n]}\mathbf{E}[Z_i^2] = \frac{\sum_{i\in[n]}\mathbf{E}[(X_i-\mathbf{E}[X_i])^2]}{\mathbf{Var}[X]} = 1.$$

Thus, we can apply Proposition 5.4.7 to Z and it holds that

$$\begin{aligned}\left|\mathbf{Pr}\left[\frac{X-\mathbf{E}[X]}{\sqrt{\mathbf{Var}[X]}}\le x\right]-\Phi(x)\right| &= \left|\mathbf{Pr}\left[\sum_{i\in[n]} Z_i \le x\right]-\Phi(x)\right| = \left|\mathbf{Pr}\left[Z\le x\right]-\Phi(x)\right| \\ &\le 5.6 \sum_{i\in[n]:\mathbf{E}[Z_i^2]>0} \mathbf{E}[|Z_i|^3] \\ &\le 5.6 \sum_{i\in[n]} \frac{c_i^3}{\mathbf{Var}[X]^{3/2}} = \frac{5.6\|c\|_3^3}{\mathbf{Var}[X]^{3/2}}. \qquad (5.2)\end{aligned}$$

Next we observe that

$$\mathbf{Pr}\left[|X| \ge x\sqrt{\mathbf{Var}[X]}\right] = \mathbf{Pr}\left[X \ge x\sqrt{\mathbf{Var}[X]}\right] + \mathbf{Pr}\left[Y \le -x\sqrt{\mathbf{Var}[X]}\right] \qquad (5.3)$$

holds. If $\mathbf{E}[X] \ge 0$, we have

$$\begin{aligned}\mathbf{Pr}\left[|X| \ge x\sqrt{\mathbf{Var}[X]}\right] &\ge \mathbf{Pr}\left[X \ge x\sqrt{\mathbf{Var}[X]} + \mathbf{E}[X]\right] \\ &\ge 1 - \mathbf{Pr}\left[X - \mathbf{E}[X] \le x\sqrt{\mathbf{Var}[X]}\right] \ge 1 - \Phi(x) - \frac{5.6\|c\|_3^3}{\mathbf{Var}[X]^{3/2}}\end{aligned}$$

from (5.2). Similarly, if $\mathbf{E}[X] \le 0$, (5.2) yields

$$\begin{aligned}\mathbf{Pr}\left[|X| \ge x\sqrt{\mathbf{Var}[X]}\right] &\ge \mathbf{Pr}\left[X \le -x\sqrt{\mathbf{Var}[X]} + \mathbf{E}[X]\right] \\ &= \mathbf{Pr}\left[X - \mathbf{E}[X] \le -x\sqrt{\mathbf{Var}[X]}\right] \ge \Phi(-x) - \frac{5.6\|c\|_3^3}{\mathbf{Var}[X]^{3/2}}.\end{aligned}$$

Thus, the claim holds for both cases. Note that $\Phi(-x) = 1 - \Phi(x)$ holds. □

A function $f : \{0,1\}^M \to \mathbb{R}$ is *monotone increasing* if $f(\mathbf{x}) \leq f(\mathbf{y})$ whenever $\mathbf{x} = (x_1, \ldots, x_M), \mathbf{y} = (y_1, \ldots, y_M) \in \{0,1\}^M$ satisfies $x_i \leq y_i$ for every $i = 1, \ldots, M$.

Proposition 5.4.10 (The FKG inequality; Theorem 21.5 of [FK16])**.** *Let $I_1, I_2, \ldots, I_M$ be independent binary random variables. Then for any two monotone increase functions $f, g : \{0,1\}^M \to \mathbb{R}$, it holds that*

$$\mathbf{E}[f(\mathbf{I})g(\mathbf{I})] \geq \mathbf{E}[f(\mathbf{I})]\,\mathbf{E}[g(\mathbf{I})]$$

where $\mathbf{I} = (I_1, I_2, \ldots, I_M) \in \{0,1\}^M$.

Chapter 6

Voting Process on Stochastic Block Model

6.1 Our Results

In this chapter, we consider best-of-two and best-of-three on the graph $G(2n, p, q)$ drawn from the stochastic block model $\mathcal{G}(2n, p, q)$ (see Definition 1.2.1 for the definition). Throughout this chapter, we assume $p = \omega(\log n/n)$, in which regime each community is connected w.h.p [FK16]. We denote by V the vertex set of the underlying graph.

Our voting process proceeds as follows. We first generate a graph $G(2n, p, q)$ according to $\mathcal{G}(2n, p, q)$, and then set an initial opinion configuration $A_0 \subseteq V$. Then consider functional voting (see Definition 5.1.1) on the graph $G(2n, p, q)$. We are interested in the consensus time T_{cons}.

Remark 6.1.1 (Two sources of randomness)**.** A functional voting on the graph $G(2n, p, q)$ drawn from $\mathcal{G}(2n, p, q)$ involves two sources of randomness: the generation of $G(2n, p, q)$ and the Markov chain $(A^{(t)})_{t=0,1,\ldots}$. We say $T_{\text{cons}}(A) \leq f(n)$ *for any* $A \subseteq V$ *w.h.p.* if there is a set $\mathcal{P}$ of finite graphs satisfying

- It holds w.h.p. that $G(2n, p, q) \in \mathcal{P}$, and
- For any n-vertex graph $G \in \mathcal{P}$ and any $A \subseteq V(G)$, $T_{\text{cons}}(A) \leq f(n)$ w.h.p.

In the former (respectively, latter) condition, the probability is taken over $\mathcal{G}(2n, p, q)$ (respectively, the process). The term "$T_{\text{cons}}(A) \geq g(n)$ *for some* $A \subseteq V$ *w.h.p.*" is defined in a similar way.

The set $\mathcal{P}$ can be interpreted as a graph property (strictly speaking, $\mathcal{P}$ may not be a graph property since $G(2n, p, q)$ has a vertex set $V_1 \cup V_2$ and $\mathcal{P}$ may not invariant under the isomorphism).

Cooper, Elsässer, and Radzik [CER14] used the framework of Remark 6.1.1 to consider the consensus time of best-of-two on the random regular graph $G_{n,d}$.

Example 6.1.2. Consider the best-of-two on the graph $G(2n, 1, 1)$. Then, $T_{\text{cons}}(A) = \mathrm{O}(\log n)$ for any $A \subseteq V$ w.h.p. To see this, let $\mathcal{P} = \bigcup_{n\in\mathbb{N}}\{K_{2n}\}$, where K_{2n} is the $2n$-vertex complete graph. Obviously, $\mathcal{G}(2n, 1, 1)$ is in $\mathcal{P}$ with probability 1. Moreover, from [DGM+11], it is known that $T_{\text{cons}}(A) = \mathrm{O}(\log n)$ w.h.p. for any $A \subseteq V$.

Example 6.1.3. Consider the best-of-two on the random regular graph $G_{n,d}$ for $d = \omega(\log n)$. Then, $T_{\text{cons}}(A) = \mathrm{O}(\log n)$ w.h.p. for any $A \subseteq V$ satisfying $||A| - |V \setminus A|| = \Omega(n)$. To see this, let $\mathcal{P}$ be the set of all λ-expander graphs for an appropriate $\lambda = \mathrm{O}(1/\sqrt{np})$ (see Chapter 7 for the definition of λ-expander graphs). It is known that $\mathcal{G}_{n,d}$ is $\mathrm{O}(1/\sqrt{d})$-expander w.h.p. [CGJ18, TY19]. Moreover, from [CER14], it is known that, the best-of-two on any $\mathrm{O}(1/\sqrt{d})$-expander graph reach consensus within $\mathrm{O}(\log n)$ rounds w.h.p. if the initial configuration $A \subseteq V$ satisfies $||A| - |V \setminus A|| = \Omega(n)$.

In this chapter we obtain two results. The first result concerns the *phase transition* of voting processes on $\mathcal{G}(2n, p, q)$. In the second result, we consider the *worst-case* consensus time of voting processes on $\mathcal{G}(2n, p, q)$. Here, the term "worst-case" refers to the worst initial configuration, that is, the configuration $A \subseteq V$ that attains the maximum consensus time.

Result I: Phase transition. We obtain a "sharp threshold result" on the consensus time of voting processes on the stochastic block model.

Theorem 6.1.4 (Phase transition of best-of-three on $G(2n,p,q)$)**.** *Let $p=p(n)$ and $q=q(n)$ be functions such that $p,q=\omega(\log n/n)$ and $r=r(n):=q(n)/p(n)$ is a constant. Consider best-of-three on $G(2n,p,q)$. Then, for any constant $\epsilon>0$, the following hold.*

(i) If $r\geq\frac{1}{7}+\epsilon$, then $T_{\mathrm{cons}}(A)=\mathrm{O}(\log\log n+\log n/\log(np))$ for any $A\subseteq V$ satisfying $||A|-|V\setminus A||=\Omega(n)$ w.h.p.

(ii) If $r\leq\frac{1}{7}-\epsilon$, then $T_{\mathrm{cons}}(A)=\exp(\Omega(n))$ for some $A\subseteq V$ w.h.p.

Theorem 6.1.5 (Phase transition of best-of-two on $G(2n,p,q)$)**.** *Let $p=p(n)$ and $q=q(n)$ be functions such that $p,q=\omega(\log n/n)$ and $r=r(n):=q(n)/p(n)$ is a constant. Consider best-of-two on $G(2n,p,q)$. Then, for any constant $\epsilon>0$, the following hold.*

(i) If $r\geq\sqrt{5}-2+\epsilon$, then $T_{\mathrm{cons}}(A)=\mathrm{O}(\log\log n+\log n/\log(np))$ for any $A\subseteq V$ satisfying $||A|-|V\setminus A||=\Omega(n)$ w.h.p.

(ii) If $r\leq\sqrt{5}-2-\epsilon$, then $T_{\mathrm{cons}}(A)=\exp(\Omega(n))$ for some $A\subseteq V$ w.h.p.

The bound $T_{\mathrm{cons}}(A)=\mathrm{O}(\log\log n+\log n/\log(np))$ is tight up to a constant factor if $\log n/\log(np)\geq\log\log n$. To see this, observe that there is a set $A\subseteq V$ such that $T_{\mathrm{cons}}(A)$ is at least half of the diameter. In addition, it is easy to see that the diameter of $G(2n,p,q)$ is $\Theta(\log n/\log(np))$ w.h.p. [FK16].

We also note that the consensus time of the pull voting is $\mathrm{poly}(n)$ w.h.p. for any connected non-bipartite graph [HP01]. Therefore, Theorems 6.1.4 and 6.1.5 imply that best-of-two and best-of-three can be exponentially slower than the pull voting.

Result II: Worst-case consensus time. The most difficult part in the analysis of voter processes is the *symmetry breaking*, i.e. the number of iterations required to cause a small bias starting from the half-and-half state. Here, we are interested in the worst-case consensus time with respect to initial opinion configurations. To the best of our knowledge, all current results on worst-case consensus time of best-of-two and best-of-three deal with complete graphs [DGM+11, BCN+17a, BCE+17, GL18]. All previous work on non-complete graphs has involved some special bias setting (e.g. an initial bias [CER14, CER+15, CRRS17], or a random initial opinion configuration [AD15, CNS19, KR19]). In this chapter, we present the first worst-case result concerning the consensus time of best-of-two and best-of-three on non-complete graphs.

Theorem 6.1.6 (Worst-case analysis of best-of-three on $G(2n,p,q)$)**.** *Let $p\geq q>0$ be constants and consider best-of-three on $G(2n,p,q)$. If $\frac{q}{p}>\frac{1}{7}$, then $T_{\mathrm{cons}}(A)=\mathrm{O}(\log n)$ w.h.p. for any $A\subseteq V$.*

Theorem 6.1.7 (Worst-case analysis of best-of-two on $G(2n,p,q)$)**.** *Consider best-of-two on $G(2n,p,q)$ for positive constants p and q. If $\frac{q}{p}>\sqrt{5}-2$, then $T_{\mathrm{cons}}(A)=\mathrm{O}(\log n)$ w.h.p. for any $A\subseteq V$.*

6.1.1 Our strategy: Structual analysis of $\mathcal{G}(2n,p,q)$

Consider a functional voting on a graph $G=(V,E)$ (see Definition 5.1.1). Then, if $A_t=A$ is fixed, then $|A'|=\sum_{v\in V}\mathbb{1}_{v\in A'}$ is the sum of independent random variables; thus, the size $|A'|$ concentrates on $\mathbf{E}[|A'|]$.

As mentioned in Section 5.3, if the underlying graph is the complete graph (with self loops), the state space of the functional voting becomes $\{0,\dots,n\}$ (each state represents $|A|$). Therefore, the expectation $\mathbf{E}[|A_{t+1}|\mid A_t=A]$ can be written as a function $\mathbf{E}[|A_{t+1}|\mid A_t=A]=F(|A|)$, where $F(\cdot)$ is a function depending on f. For example, in best-of-three, from a straightforward calculation, we have $\mathbf{E}[|A'|]=nf_{\mathrm{Bo3}}\big(3\frac{|A|}{n}\big)$, where $f_{\mathrm{Bo3}}(x):=3x^2-2x^3$. Doerr et al. [DGM+11] exploited this idea for best-of-two and obtained the worst-case consensus time on complete graphs.

The key observation of the best-of-three on complete graphs is that, the property that the expectation $\mathbf{E}[|A_{t+1}|\mid A_t=A]$ can be written as a function $F(|A|)$ makes the analysis of best-of-three on complete graphs tractable. Our strategy is to extend this observation. For simplicity, we state our strategy for the analysis of best-of-three on $G(n,p)$. The first idea is to define a graph property $\mathcal{P}$ as the set of graphs such that $\mathbf{E}[|A_{t+1}|\mid A_t=A]$ can be *approximated* by a function $F(|A|)$. Next, we show that

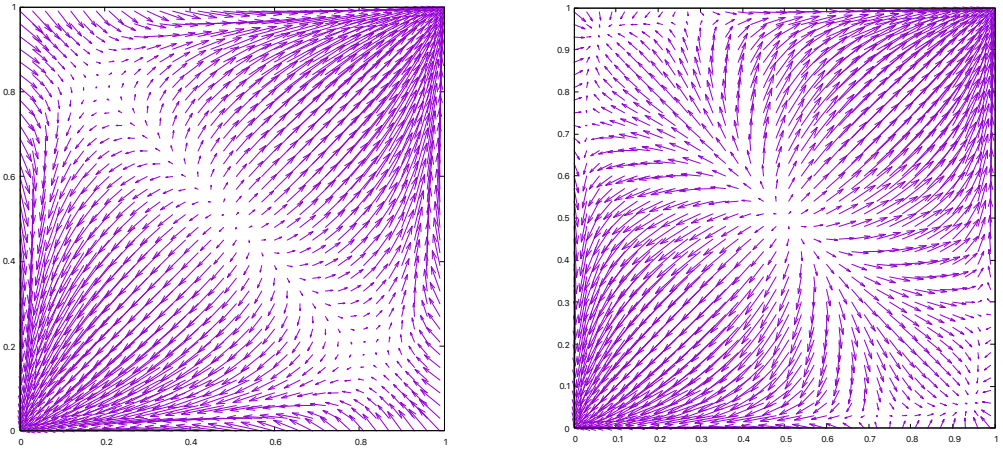
$\mathcal{G}(n,p)$ satisfies the graph property $\mathcal{P}$. Then, the graph property $\mathcal{P}$ makes the analysis of best-of-three on $\mathcal{G}(n,p)$ tractable.

We further extend this strategy to $\mathcal{G}(2n,p,q)$. Recall that, the graph $G(2n,p,q)$ has the vertex set $V = V_1 \cup V_2$. For $A \subseteq V$ and $i \in \{1,2\}$, let $A_i := A \cap V_i$. Since $|A'_i|$ can be written as the sum of random variables, we focus on $\mathbf{E}[|A'_i|]$. We define a set $\mathcal{P}_{\mathrm{approx}}$ of graphs as the set of graphs with vertex set $V_1 \cup V_2$ on which the best-of-three satisfies $\mathbf{E}[|A'_i|] = F_i(|A_1|,|A_2|) \pm \mathrm{O}(\sqrt{n/p})$ for all $A \subseteq V$, where $F_1, F_2 : \mathbb{N}^2 \to \mathbb{N}$ are fixed functions.

The technical contribution of this chapter is to show that $G(2n,p,q) \in \mathcal{P}_{\mathrm{approx}}$ w.h.p. Indeed, in best-of-three, we show that $\mathbf{E}[|A'_i|] = F_i(|A_1|,|A_2|) \pm \mathrm{O}(\sqrt{n/p})$ for all $A \subseteq V$, where $F_i : \mathbb{N}^2 \to \mathbb{N}$ is some function ($i \in \{1,2\}$). We prove that the similar approximation result holds for the best-of-two. Our key tool is the concentration inequalities, specifically the Janson inequality (Proposition 5.4.5) and the Kim-Vu concentration inequality (Proposition 5.4.6).

6.1.2 Proof overview: Voting process on $G(2n,p,q)$

We briefly present our idea by considering the best-of-three on $G(2n,p,q)$. The following arguments on the best-of-three also works for the best-of-two, which implies Theorem 6.1.5. Consider a sequence $(\boldsymbol{\alpha}^{(t)})_{t\in\mathbb{Z}_{\geq 0}} = ((\alpha_1^{(t)}, \alpha_2^{(t)}))_{t\in\mathbb{Z}_{\geq 0}}$ of random variables, where $\alpha_i^{(t)} = |A_t \cap V_i|/n$ for $i \in \{1,2\}$. From our technical result that $G(2n,p,q) \in \mathcal{P}_{\mathrm{approx}}$ w.h.p., we can approximate the sequence $(\boldsymbol{\alpha}^{(t)})_{t\in\mathbb{Z}_{\geq 0}}$ by a sequence $(\mathbf{a}^{(t)})_{t\in\mathbb{Z}_{\geq 0}}$ defined as $\mathbf{a}^{(t+1)} = H(\mathbf{a}^{(t)})$ and $\boldsymbol{\alpha}^{(0)} = \mathbf{a}^{(0)}$ for some function $H : [0,1]^2 \to [0,1]^2$. Specifically, we show that $\|\boldsymbol{\alpha}^{(t)} - \mathbf{a}^{(t)}\|_2 = O(1)^t \cdot (1/\sqrt{np} + \sqrt{\log n/n})$ for all $t = 0,\dots,n^{o(1)}$ (Theorem 6.2.3). The function H defines a two-dimensional dynamical system, which we call the *induced dynamical system* (see Figure 6.1).

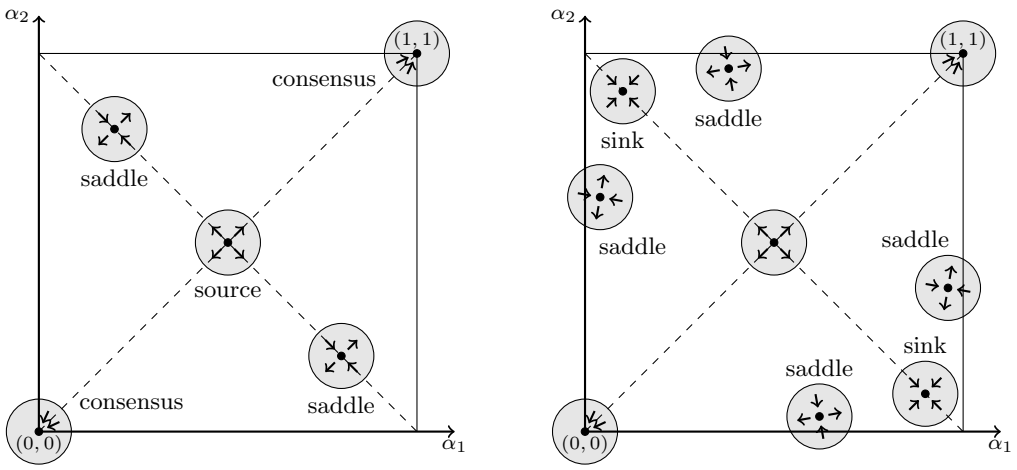


(a) $r > 1/7$ (b) $r < 1/7$

Figure 6.1: The induced dynamical system of the best-of-three on $G(2n,p,q)$ is illustrated, where $r = q/p$. Note that the best-of-three exhibits a phase transition at threshold $r = 1/7$ (Theorem 6.1.4). The existence of a "sink point" affects the behavior of the best-of-three.

In terms of induced dynamical systems, we obtain two results concerning $\boldsymbol{\alpha}^{(t)}$. Let H be the induced dynamical system. First, we show that, for any initial configuration A, $\boldsymbol{\alpha}^{(t)}$ is arbitrary closed to one of the fixed points of H for some $t = \mathrm{O}(1)$. Figure 6.2 illustrates the fixed points of the induced dynamical system H of best-of-three on $G(2n,p,q)$. In general, it is quite difficult to predicate the orbit of a dynamical system since some dynamical systems exhibit chaos property. Moreover, some dynamical system has a loop of period two or more (i.e., there might exist two distinct points $\mathbf{a},\mathbf{b}$ such that $H(\mathbf{a}) = \mathbf{b}$ and $H(\mathbf{b}) = \mathbf{a}$). Therefore, the proof of the convergence of the sequence $(\mathbf{a}^{(t)})$ generated by H is difficult in general. Fortunately, the induced dynamical system of best-of-two and best-of-three on $G(2n,p,q)$ is *competitive*, which is a nice property of dynamical systems [HS05] (see Section 6.2.4 for definition).

Second, we investigate the behavior of $\boldsymbol{\alpha}^{(t)}$ starting from a point closed to a fixed point of H. The fixed points are classified into four types using the Jacobian matrix: consensus, sink, saddle and source points.Around consensus points, we show that the process reaches consensus within $\mathrm{O}(\log\log n +$



(a) $r > 1/7$ (b) $r < 1/7$

Figure 6.2: Four fixed points of the induced dynamical system H are illustrated. Note that, the horizontal and vertical axis correspond to α_1 and α_2, respectively. The points $(\alpha_1, \alpha_2) = (0, 0), (1, 1)$ represents consensus. If $r > 1/7$, sink areas (excluding consensus points) do not exist.

$\log n / \log(np)$) steps. Around sink points, we show that the process remains there for at least $\exp(\Omega(n))$ steps, and also show that sink points do not appear if $r > 1/7$. This yields the lower bound of the consensus time. Around saddle and source points, we show that the process "escapes" from there within $O(\log n)$ steps if p and q are constants. Intuitively speaking, in these two kinds of fixed points, there are drifts towards outside and we can apply the framework of [DGM+11].

6.1.3 Organization of this chapter

As a preliminary, we introduce precise definition of our model and auxiliary results of the stochastic block model and dynamical systems in Section 6.2. In Section 6.4, we prove Theorems 6.1.4 to 6.1.7 using the auxiliary results. In Sections 6.5 and 6.6, we prove the auxiliary results of the stochastic block model and dynamical systems, respectively.

6.2 Auxiliary Results

In this section, we introduce auxiliary results for Theorems 6.1.4 to 6.1.7.

6.2.1 Our model

We say that a functional voting is a *polynomial voting process* if it is a functional voting with respect to a *polynomial*. We consider a polynomial voting process on $G(2n, p, q)$.

6.2.2 Concentration result for the stochastic block model

Consider a polynomial voting process on the graph $G(2n, p, q)$. For a fixed $A \subseteq V$, let $\alpha_i = |A \cap V_i|/n$ for $i \in \{1, 2\}$ and $\alpha'_i = |A' \cap V_i|/n$ (see Definition 5.1.1 for the definition of A'). Since $\alpha'_i = \frac{1}{n} \sum_{v \in V_i} X_v$ is the sum of independent random variables, the Hoeffding bound (Proposition 5.4.3) implies that

$$|\alpha'_i - \mathbf{E}[\alpha'_i]| = O\left(\sqrt{n \log n}\right) \tag{6.1}$$

holds w.h.p. (here, the probability is over the randomness of the process). For any fixed $A \subseteq V$, we have

$$\begin{aligned}\mathbf{E}[\alpha_i'] &= \frac{1}{n}\sum_{v\in V_i}\mathbf{E}[X_v] \\ &= \frac{1}{n}\sum_{v\in A_i}(1-\mathbf{Pr}[X_v=0]) + \frac{1}{n}\sum_{v\in V\setminus A_i}\mathbf{Pr}[X_v=1] \\ &= \frac{1}{n}\sum_{v\in A_i}\left(1-f\left(1-\frac{\deg_A(v)}{\deg(v)}\right)\right) + \frac{1}{n}\sum_{v\in V\setminus A_i} f\left(\frac{\deg_A(v)}{\deg(v)}\right). \end{aligned} \tag{6.2}$$

In general, (6.2) is a random variable since $G(2n,p,q)$ is a random graph. Our key ingredient is the following concentration result for this random variable.

Definition 6.2.1 (f-good graph)**.** *Let $G=(V,E)$ be a graph on $2n$ vertices. Let $f:[0,1]\to[0,1]$ be a function, $V_1,V_2\subseteq V$ be a partition of V such that $|V_1|=|V_2|=n$, and $p,q\in[0,1]$ be parameters satisfying $p\geq q>0$ ($p=p(n)$ and $q=q(n)$ may depend on n). We say a graph $G=(V,E)$ is f-good for a partition V_1,V_2 and parameters p,q if the graph satisfies the following conditions.*

(P1) *It is connected and non-bipartite.*

(P2) *There is a positive constant C_1 such that, for all $A,S\subseteq V$ and $i\in\{1,2\}$,*

$$\left|\sum_{v\in S\cap V_i} f\left(\frac{\deg_A(v)}{\deg(v)}\right) - |S\cap V_i| f\left(\frac{|A_i|p+|A_{3-i}|q}{n(p+q)}\right)\right| \leq C_1\sqrt{\frac{n}{p}},$$

where $A_j=A\cap V_j$ for $j\in\{1,2\}$.

(P3) *There is a positive constant C_2 such that, for all $A\subseteq V$, $S\in\{A,V\setminus A,V\}$ and $i\in\{1,2\}$,*

$$\sum_{v\in S\cap V_i} f\left(\frac{\deg_A(v)}{\deg(v)}\right) \leq |S\cap V_i| f\left(\frac{|A_i|p+|A_{3-i}|q}{n(p+q)}\right) + C_2|A|\sqrt{\frac{\log n}{np}}.$$

Theorem 6.2.2 (Main technical theorem)**.** *Consider a stochastic block model $G(2n,p,q)$ on a vertex set $V_1\cup V_2$. Let $f:[0,1]\to[0,1]$ be a polynomial, and $p=\omega(\log n/n)$ and $q\geq\log n/n^2$ be functions. Then $G(2n,p,q)$ is f-good for a partition (V_1,V_2) and parameters p,q w.h.p.*

The proof of (P1) is not difficult since $p=\omega(\log n/n)$ and $q\geq\log n/n^2$ (see, e.g., [FK16]). However, proving (P2) and (P3) is challenging: we prove them in Section 6.5.

Let $f:[0,1]\to[0,1]$ be a polynomial and define $\overline{f}(x)=1-f(1-x)$. Note that $g:[0,1]\to[0,1]$. From Theorem 6.2.2, $G(2n,p,q)$ is f- and $\overline{f}$-good for a partition (V_1,V_2) and parameters p,q w.h.p. Henceforth, we consider a polynomial voting process on a (fixed) f- and $\overline{f}$-good graph for a partition (V_1,V_2) and parameters p,q. Let $r:=\frac{q}{p}$, $\alpha_i:=\frac{|A_i|}{n}$ and $\alpha=\alpha_1+\alpha_2$. From (6.2), (P2) and (P3), we have

$$\mathbf{E}[\alpha_i'] = \alpha_i\overline{f}\left(\frac{\alpha_i+r\alpha_{3-i}}{1+r}\right) + (1-\alpha_i)f\left(\frac{\alpha_i+r\alpha_{3-i}}{1+r}\right) + \mathrm{O}\left(\min\left\{\frac{1}{\sqrt{np}}, \alpha\sqrt{\frac{\log n}{np}}\right\}\right) \tag{6.3}$$

for all $A\subseteq V$ and $i=1,2$. Here, we note that the additive error $\mathrm{O}\left(\min\left\{\sqrt{\frac{1}{np}}, \alpha\sqrt{\frac{\log n}{np}}\right\}\right)$ depends on α. This property plays a key role in consensus.

6.2.3 Induced dynamical system

Consider a polynomial voting process with respect to a polynomial f on a fixed graph that is f-good for a partition (V_1,V_2) and parameters p,q. Suppose that $r=\frac{q}{p}$ is a constant. Define two functions $H_1,H_2:[0,1]^2\to[0,1]$ as

$$H_i(a_1,a_2) = a_i\overline{f}\left(\frac{a_i+ra_{3-i}}{1+r}\right) + (1-a_i)f\left(\frac{a_i+ra_{3-i}}{1+r}\right) \quad \text{for } i\in\{1,2\}. \tag{6.4}$$

From (6.1) and (6.3), for all $A \subseteq V$ and $i \in \{1,2\}$, it holds w.h.p. that

$$|\alpha_i' - H_i(\alpha_1, \alpha_2)| = \mathrm{O}\left(\sqrt{\frac{1}{np}} + \sqrt{\frac{\log n}{n}}\right). \tag{6.5}$$

Throughout this chapter, we use $\boldsymbol{\alpha} = (\alpha_1, \alpha_2)$ and $\boldsymbol{\alpha}' = (\alpha_1', \alpha_2')$ as vector-valued random variables. Equation (6.5) leads us to the dynamical system H, where we define $H : \mathbb{R}^2 \to \mathbb{R}^2$ as

$$H : \mathbf{a} \mapsto (H_1(\mathbf{a}), H_2(\mathbf{a})). \tag{6.6}$$

By combining (6.5) with the Lipschitz condition (see Section 6.3.3), it is not difficult to show the following result.

Theorem 6.2.3. *Consider a polynomial voting process with respect to f on an f-good graph for a partition (V_1, V_2) and parameters p, q. For the mapping H given by (6.4) and (6.6), let $(\mathbf{a}^{(t)})_{t=0}^{\infty}$ be the sequence defined as*

$$\begin{cases} \mathbf{a}^{(0)} = \boldsymbol{\alpha}^{(0)}, \\ \mathbf{a}^{(t+1)} = H(\mathbf{a}^{(t)}). \end{cases} \tag{6.7}$$

Then, there is a constant $C > 0$ such that

$$\forall 0 \le t \le n^{\mathrm{o}(1)}, \forall A_0 \subseteq V \,:\, \mathbf{Pr}\left[\|\boldsymbol{\alpha}^{(t)} - \mathbf{a}^{(t)}\|_\infty \le C^t \left(\frac{1}{\sqrt{np}} + \sqrt{\frac{\log n}{n}}\right)\right] \ge 1 - n^{-\Omega(1)}.$$

Broadly speaking, Theorem 6.2.3 approximates the behavior of $\boldsymbol{\alpha}^{(t)}$ by the orbit $\mathbf{a}^{(t)}$ of the dynamical system determined by H. We call the mapping H *induced dynamical system*.

Proof of Theorem 6.2.3. From (6.5), for all $A \subseteq V$, there is some positive constant C_1 such that $\|\boldsymbol{\alpha}' - H(\boldsymbol{\alpha})\|_\infty \le C_1 \left(\frac{1}{\sqrt{np}} + \sqrt{\frac{\log n}{n}}\right)$ holds w.h.p. Since f and $\overline{f}$ are polynomials, the function H satisfies the Lipschitz condition. That is, there is a positive constant C_2 such that $\|H(\mathbf{x}) - H(\mathbf{y})\|_\infty \le C_2 \|\mathbf{x} - \mathbf{y}\|_\infty$ holds for any $\mathbf{x}, \mathbf{y} \in [0,1]^2$ (see Section 6.3.3). Then, we have

$$\begin{aligned} \|\boldsymbol{\alpha}^{(t)} - \mathbf{a}^{(t)}\|_\infty &= \|\boldsymbol{\alpha}^{(t)} - H(\boldsymbol{\alpha}^{(t-1)}) + H(\boldsymbol{\alpha}^{(t-1)}) - H(\mathbf{a}^{(t-1)})\|_\infty \\ &\le \|\boldsymbol{\alpha}^{(t)} - H(\boldsymbol{\alpha}^{(t-1)})\|_\infty + C_2 \|\boldsymbol{\alpha}^{(t-1)} - \mathbf{a}^{(t-1)}\|_\infty \\ &\le C_2 \|\boldsymbol{\alpha}^{(t-1)} - \mathbf{a}^{(t-1)}\|_\infty + C_1 \left(\frac{1}{\sqrt{np}} + \sqrt{\frac{\log n}{n}}\right) \\ &\le C^t \left(\frac{1}{\sqrt{np}} + \sqrt{\frac{\log n}{n}}\right), \end{aligned}$$

where C is a sufficiently large constant. □

Now, we change the coordinate of H for convenience. Let $\boldsymbol{\delta}$ and $\boldsymbol{\delta}'$ be

$$\begin{aligned} \boldsymbol{\delta} = (\delta_1, \delta_2) &:= (\alpha_1 - \alpha_2, \alpha_1 + \alpha_2 - 1), \\ \boldsymbol{\delta}' = (\delta_1', \delta_2') &:= (\alpha_1' - \alpha_2', \alpha_1' + \alpha_2' - 1). \end{aligned} \tag{6.8}$$

Axes δ_1 and δ_2 correspond to the dotted lines of Figure 6.2. From (6.3), for any $A \subseteq V$ and any $i = 1, 2$, $\mathbf{E}[\delta_i'] = T_i(\delta_1, \delta_2) + \mathrm{O}\left(\frac{1}{\sqrt{np}}\right)$ holds, where

$$\begin{aligned} T_1(d_1, d_2) &:= H_1\left(\frac{1+d_1+d_2}{2}, \frac{1-d_1+d_2}{2}\right) - H_2\left(\frac{1+d_1+d_2}{2}, \frac{1-d_1+d_2}{2}\right), \\ T_2(d_1, d_2) &:= H_1\left(\frac{1+d_1+d_2}{2}, \frac{1-d_1+d_2}{2}\right) + H_2\left(\frac{1+d_1+d_2}{2}, \frac{1-d_1+d_2}{2}\right) - 1. \end{aligned}$$

This suggests a dynamical system $T(\mathbf{d}) = (T_1(\mathbf{d}), T_2(\mathbf{d}))$. Here, we use $\mathbf{d} = (d_1, d_2)$ as a specific point and $\boldsymbol{\delta} = (\delta_1, \delta_2)$ as a vector-valued random variable. Note that $\boldsymbol{\delta}$ satisfies $|\delta_1| + |\delta_2| \le 1$. In addition,

the dynamical system T is symmetric: Specifically, $T_1(\pm d_1, \mp d_2) = \pm T_1(d_1, d_2)$ and $T_2(\pm d_1, \mp d_2) = \mp T_2(d_1, d_2)$ hold. To see this, observe $H_i(\alpha_2, \alpha_1) = H_{3-i}(\alpha_1, \alpha_2)$ (exchange V_1 and V_2) and $H_i(1 - \alpha_1, 1 - \alpha_2) = 1 - H_i(\alpha_1, \alpha_2)$ (consider $(V \setminus A)'$ instead of A'). Consider $\boldsymbol{\delta}^{(t)} = (\delta_1^{(t)}, \delta_2^{(t)}) = (\alpha_1^{(t)} - \alpha_2^{(t)}, \alpha_1^{(t)} + \alpha_2^{(t)} - 1)$ and $(\mathbf{d}^{(t)})_{t=0}^{\infty}$, where $\mathbf{d}^{(0)} = \boldsymbol{\delta}^{(0)}$ and $\mathbf{d}^{(t+1)} = T(\mathbf{d}^{(t)})$ for each $t \geq 0$. From Theorem 6.2.3, it holds w.h.p. that

$$\|\boldsymbol{\delta}^{(t)} - \mathbf{d}^{(t)}\|_\infty \leq C^t \left(\frac{1}{\sqrt{np}} + \sqrt{\frac{\log n}{n}} \right) \tag{6.9}$$

for sufficiently large constant $C > 0$, any $0 \leq t \leq n^{o(1)}$ and any initial configuration $A_0 \subseteq V$.

Let

$$S := \{(d_1, d_2) \in [0,1]^2 : d_1 + d_2 \leq 1\}. \tag{6.10}$$

We will show in Lemma 6.4.1 that, $\mathbf{d}^{(t+1)} \in S$ holds for any $\mathbf{d}^{(t)} \in S$ in best-of-two and best-of-three. Therefore, we focus on the behavior $(\mathbf{d}^{(t)})_{t=0}^{\infty}$ within S.

6.2.4 Orbit convergence

In this subsection, for a map $T : S \to S$ and an initial point $\mathbf{x} \in S$, we present a sufficient condition for the convergence of the orbit (i.e., the sequence $(T^n(\mathbf{x}))_{n \geq 0}$). We call a point $\mathbf{x}$ a *fixed point* if $T(\mathbf{x}) = \mathbf{x}$ holds.

Theorem 6.2.4. *Let $T : S \to S$ be an injective and C^1 (i.e., differentiable and its derivation is continuous) function where S is defined in (6.10). Let $J = (j_{ij})_{i,j \in [2]}$ be the Jacobian matrix of T at $\mathbf{x} \in S$. Suppose that J satisfies*

(C1) For any $\mathbf{x} \in S$, it hold that $j_{11}, j_{22} \geq 0$ and $j_{12}, j_{21} \leq 0$, and

(C2) For any $\mathbf{x} \in S \setminus \{(0,1)\}$, the determinant satisfies $\det J > 0$.

Then, for any $\mathbf{x} \in S$, there is the limit $\lim_{n \to \infty} T^n(\mathbf{x})$ and the limit is a fixed point of T.

We will show that the dynamical system T of the best-of-two (and best-of-three) satisfies both (C1) and (C2). Roughly speaking, from (6.9) and Theorem 6.2.4, it holds w.h.p. that $\boldsymbol{\delta}^{(t)}$ approaches around a fixed point after constant steps (see Section 6.4 for details).

To show Theorem 6.2.4, we introduce the notion of *competitive dynamical system*.

Definition 6.2.5 (Competitive dynamical system)**.** *For two points $\mathbf{x} = (x_1, x_2)$ and $\mathbf{y} = (y_1, y_2)$, we write $\mathbf{x} \leq_K \mathbf{y}$ if $x_1 \leq y_1$ and $y_2 \leq x_2$ hold. For $S \subseteq \mathbb{R}^2$, a map $T : S \to S$ is* competitive *if $T(\mathbf{x}) \leq_K T(\mathbf{y})$ whenever $\mathbf{x} \leq_K \mathbf{y}$.*

See [HS05] for the background of competitive dynamical systems. For two points $\mathbf{x} = (x_1, x_2)$ and $\mathbf{y} = (y_1, y_2)$, we write $\mathbf{x} \leq \mathbf{y}$ if $x_1 \leq y_1$ and $x_2 \leq y_2$. We write $\mathbf{x} \ll \mathbf{y}$ if $x_1 < y_1$ and $x_2 < y_2$. The following known result provides a sufficient condition for the orbit convergence of a competitive dynamical system.

Theorem 6.2.6 (Theorem 5.28 of [HS05])**.** *Suppose that a competitive map $T : S \to S$ satisfies $\mathbf{x} \leq \mathbf{y}$ for any $\mathbf{x}, \mathbf{y} \in S$ of $T(\mathbf{x}) \ll T(\mathbf{y})$. Then, for any $\mathbf{x} \in S$, the sequence $(T^n(\mathbf{x}))_{n \geq 0}$ converges to some fixed point of T.*

Proof of Theorem 6.2.4. It suffices to check the condition of Theorem 6.2.6 holds. First, we claim that T is competitive. Let $T(\mathbf{x}) = (T_1(\mathbf{x}), T_2(\mathbf{x}))$ for $\mathbf{x} = (x_1, x_2) \in S$. From (C1), the function T_i is nondecreasing on x_i and is nonincreasing on x_{3-i}. Therefore, for any (a, b), $(c, d) \in S$ of $(a, b) \leq_K (c, d)$, we have $T(a, b) \leq_K T(c, d)$. In other words, T is competitive.

Second, we claim that the inverse T^{-1} satisfies $T^{-1}(\mathbf{x}) \leq T^{-1}(\mathbf{y})$ whenever $\mathbf{x} \leq \mathbf{y}$. Let $U := T^{-1}$ and $U(\mathbf{x}) = (U_1(\mathbf{x}), U_2(\mathbf{x}))$ for $\mathbf{x} = (x_1, x_2) \in S$. By the Inverse Function Theorem (Proposition 6.3.5), the Jacobian matrix K of U at $\mathbf{x} \in S \setminus \{(0,1)\}$ is the inverse of that of T. Thus, from (C2), we have $\frac{\partial U_i}{\partial x_j}(\mathbf{x}) \geq 0$ for any $\mathbf{x} \in S \setminus \{(0,1)\}$. Hence the functions $U_1(x_1, x_2)$ and $U_2(x_1, x_2)$ are nondecreasing on both x_1 and x_2. Therefore for any two points $(a, b), (c, d) \in S$ of $(a, b) \leq (c, d)$, we have $U_1(a, b) \leq U_1(c, d)$ and $U_2(a, b) \leq U_2(c, d)$ (note that if $(a, b) = (0, 1)$ then (c, d) must be $(0, 1)$ and we are done).

For any points $\mathbf{x}, \mathbf{y} \in S$ of $T(\mathbf{x}) \ll T(\mathbf{y})$, the second claim implies that $\mathbf{x} = T^{-1}(T(\mathbf{x})) \leq T^{-1}(T(\mathbf{y})) = \mathbf{y}$. Therefore, we can apply Theorem 6.2.6. □

6.2.5 Local dynamics around fixed points

Consider a polynomial voting process with respect to f on an f- and $\overline{f}$-good graph for a partition (V_1, V_2) and parameters p, q (recall that $\overline{f}(x) = 1 - f(1-x)$). Let H be the induced dynamical system.

In this subsection, we focus on the behavior of $(\boldsymbol{\alpha}^{(t)})_{t=0}^{\infty}$ when the initial point $\boldsymbol{\alpha}^{(0)}$ is around a fixed point of H (i.e., a point $\mathbf{x}$ such that $H(\mathbf{x}) = \mathbf{x}$ holds). In this case, Theorem 6.2.3 does not provide enough information about the dynamics. In dynamical system theory, a common approach for the local behavior around fixed points is to consider the Jacobian matrix. In what follows, we will investigate the local dynamics from the viewpoint of the maximum singular value and eigenvalue of the Jacobian matrix. For the readability, we put the proofs of each statements in Section 6.6.

Sink point. We begin with defining the notion of sink points. Recall that the singular value of a matrix M is the positive square root of the eigenvalue of $M^\top M$ (see Section 6.3.2 for formal definition and basic properties). For $\mathbf{x} \in [0,1]^2$ and $r > 0$, let $B(\mathbf{x}, r) := \{\mathbf{y} \in \mathbb{R}^2 : \|\mathbf{x} - \mathbf{y}\|_\infty < r\}$ denote the open ball of radius r with respect to the ℓ^∞-norm.

Definition 6.2.7 (sink point)**.** *Consider a dynamical system H. A fixed point $\mathbf{a}^* \in \mathbb{R}^2$ is a* sink point *if the Jacobian matrix J at $\mathbf{a}^*$ satisfies $\sigma_{\max} < 1$, where $\sigma_{\max}$ is the largest singular value of J.*

Proposition 6.2.8. *Consider a polynomial voting process with respect to f on an f- and $\overline{f}$-good graph for a partition (V_1, V_2) and parameters p, q such that $r = \frac{q}{p}$ is a constant. Let H be the induced dynamical system. Then, for any sink point $\mathbf{a}^*$ and any sufficiently small $\epsilon = \omega(\sqrt{1/np})$,*

$$\mathbf{Pr}\left[\boldsymbol{\alpha}' \notin B(\mathbf{a}^*, \epsilon) \,|\, \boldsymbol{\alpha} \in B(\mathbf{a}^*, \epsilon)\right] \leq \exp(-\Omega(\epsilon^2 n))$$

holds. In particular, let $\tau := \inf\left\{t \in \mathbb{N} : \boldsymbol{\alpha}^{(t)} \notin B(\mathbf{a}^*, \epsilon)\right\}$ be a stopping time. Then, $\tau \geq \exp(\Omega(\epsilon^2 n))$ holds w.h.p. conditioned on $\boldsymbol{\alpha}^{(0)} \in B(\mathbf{a}^*, \epsilon)$ for any ϵ satisfying $\epsilon = \omega(\max\{1/\sqrt{np}, \sqrt{\log n/n}\})$.

Fast consensus. We consider the case in which the initial opinion configuration A_0 is closed to consensus. We first observe that, in the best-of-two and best-of-three, the Jacobian matrix at the consensus point (i.e., $\boldsymbol{\alpha} = (0,0), (1,1)$) is the all-zero matrix.

Proposition 6.2.9. *Consider a polynomial voting process with respect to f on an f- and $\overline{f}$-good graph for a partition (V_1, V_2) and parameters p, q such that $\frac{p}{q}$ is a constant. Suppose that the Jacobian matrix at the point $\boldsymbol{\alpha} = (0,0)$ is the all-zero matrix. Then, there are constants $C_1, C_2, \delta > 0$ such that*

$$\mathbf{Pr}\left[T_{\mathrm{cons}}(A) \leq C_1\left(\log\log n + \frac{\log n}{\log np}\right)\right] \geq 1 - n^{-C_2}$$

hold for any $A \subseteq V$ satisfying $|A| \leq \delta n$.

Escape from a fixed point. Let $\mathbf{a}^* \in \mathbb{R}^2$ be a fixed point of the induced dynamical system H. Let J be the Jacobian matrix of H at $\mathbf{a}^*$ and λ_1, λ_2 be its eigenvalues. Let $\mathbf{u}_i$ be the eigenvector of J corresponding to λ_i. Suppose that $\mathbf{u}_1, \mathbf{u}_2$ are linearly independent. Then, we can rewrite J as

$$J = U^{-1}\Lambda U,$$

where $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2)$ and $U = (\mathbf{u}_1\ \mathbf{u}_2)^{-1}$. For a fixed point $\mathbf{a}^* \in \mathbb{R}^2$, let $\boldsymbol{\beta} = (\beta_1, \beta_2)$ be a vector-valued random variable defined as

$$\boldsymbol{\beta} = U(\boldsymbol{\alpha} - \mathbf{a}^*). \tag{6.11}$$

From the Taylor expansion of H at $\mathbf{a}^*$, we have $\mathbf{E}[\boldsymbol{\beta}'] \approx \Lambda\boldsymbol{\beta}$ if $\|\boldsymbol{\beta}\|_\infty$ is sufficiently small.

Recall that $B(\mathbf{a}, R)$ is the open ball of radius R (with respect to the ℓ^∞-norm) centered at $\mathbf{a}$. If $|\lambda_i| > 1$ for some $i \in [2]$, one may expect that $\boldsymbol{\alpha}^{(\tau)} \notin B(\mathbf{a}^*, \epsilon_0)$ holds for any $A_0 \subseteq V$ and for some constant $\epsilon_0 > 0$. We prove this under some assumptions.

Assumption 6.2.10 (Basic assumptions)**.** *Consider an (f_1, f_2)-polynomial voting process on an f_1- and f_2-good graph for a partition (V_1, V_2) and parameters p, q, where $p \geq q \geq 0$ are constants. Let H be the induced dynamical system. Let $\mathbf{a}^*$ be a fixed point and J be the Jacobian matrix of H at $\mathbf{a}^*$. We assume that J satisfies*

(A1) *The eigenvectors* $\mathbf{u}_1$ *and* $\mathbf{u}_2$ *are linearly independent.*

(A2) *There is a constant* $\epsilon_0 > 0$ *such that* $\mathbf{Var}[\alpha'_i] \geq \Omega(n^{-1})$ *for any* $i \in \{1,2\}$ *and any* $A \subseteq V$ *of* $\boldsymbol{\alpha} \in B(\mathbf{a}^*, \epsilon_0)$.

(A3) *The matrix* J *contains an eigenvalue* λ *satisfying* $|\lambda| > 1$.

Proposition 6.2.11. *Let* $\mathbf{a}^*$ *be a fixed point satisfying Assumption 6.2.10. Suppose that the eigenvalues* λ_1, λ_2 *of the Jacobian matrix* J *at* $\mathbf{a}^*$ *satisfies* $|\lambda_i| \neq 1$ *for all* $i \in [2]$. *Then, for some* $t = \mathrm{O}(\log n)$ *and some constant* $\epsilon' > 0$, *it hold w.h.p. that* $\|\beta^{(t)}\|_\infty > \epsilon'$, *and* $|\beta_j^{(t)}| \leq \epsilon'$ *for any* $j \in [2]$ *of* $|\lambda_j| \leq 1$.

We consider the case of $\lambda_i = 1$ for some i as follows.

Proposition 6.2.12. *Let* $\mathbf{a}^*$ *be a fixed point satisfying Assumption 6.2.10. Suppose that there is a constant* $\epsilon^* > 0$ *satisfying*

(B1) *There are two positive constants* ϵ_1, C *such that* $|\mathbf{E}[\beta'_i]| \geq (1+\epsilon_1)|\beta_i| - \frac{C}{\sqrt{n}}$ *holds for any* $A \subseteq V$ *of* $\|\boldsymbol{\beta}\| \leq \epsilon^*$ *and any* $i \in [2]$ *of* $|\lambda_i| > 1$.

(B2) *For any* $i \in [2]$ *of* $|\lambda_i| \leq 1$ *and any* $A \subseteq V$ *of* $|\beta_i| \leq \epsilon^*$, *it holds that* $\mathbf{Pr}[|\beta'_i| \leq \epsilon^*] \geq 1 - n^{-\Omega(1)}$.

Then, for some $t = \mathrm{O}(\log n)$ *and some constant* $\epsilon' > 0$, *it hold w.h.p. that* $\|\beta^{(t)}\|_\infty > \epsilon'$, *and* $|\beta_j^{(t)}| \leq \epsilon'$ *for any* $j \in [2]$ *of* $|\lambda_j| \leq 1$.

6.3 Tool

6.3.1 Probability

Proposition 6.3.1 (Lemma 4.5 of [CGG+18])**.** *Consider a Markov chain* $(X_t)_{t=1}^\infty$ *with finite state space* Ω *and a function* $f : \Omega \to \{0, \ldots, n\}$. *Let* C_3 *be arbitrary constant and let* $m = C_3\sqrt{n \log n}$. *Suppose that* Ω, f *and* m *satisfies the following conditions:*

(1) *For any positive constant* h, *there is a positive constant* $C_1 < 1$ *such that*

$$\mathbf{Pr}\left[f(X_{t+1}) < h\sqrt{n} \,\middle|\, f(X_t) \leq m\right] < C_1.$$

(2) *There are three positive constants* ϵ, C_2 *and* h *such that, for any* $x \in \Omega$ *satisfying* $h\sqrt{n} \leq f(x) < m$,

$$\mathbf{Pr}\left[f(X_{t+1}) < (1+\epsilon)f(X_t) \,|\, X_t = x\right] < \exp\left(-C_2\frac{f(x)^2}{n}\right).$$

Then $f(X_\tau) \geq m$ *holds for some* $\tau = \mathrm{O}(\log n)$.

Corollary 6.3.2. *Consider a Markov chain* $(X_t)_{t=1}^\infty$ *with finite state space* Ω *and a function* $f : \Omega \to \{0, \ldots, n\}$. *Let* C_3 *be arbitrary constant and* $m = C_3\sqrt{n \log n}$. *Consider a set* $\mathcal{B} \subseteq \Omega$ *such that*

$$\mathcal{B} \subseteq \{x \in \Omega \,:\, f(x) < m\}.$$

Suppose that Ω, f, m *and* $\mathcal{B}$ *satisfy the following conditions:*

(1′) *For any positive constant* h, *there is a positive constant* $C_1 < 1$ *such that*

$$\mathbf{Pr}\left[f(X_{t+1}) < h\sqrt{n} \,\middle|\, f(X_t) \leq m, X_t \in \mathcal{B}\right] < C_1.$$

(2′) *There are three positive constants* ϵ, C_2, h *such that, for any* $x \in \mathcal{B}$ *satisfying* $h\sqrt{n} \leq f(x) < m$,

$$\mathbf{Pr}\left[f(X_{t+1}) < (1+\epsilon)f(X_t) \,|\, X_t = x\right] < \exp\left(-C_2\frac{f(x)^2}{n}\right).$$

(3′) *For some constant* $C_4 > 0$,

$$\mathbf{Pr}\left[X_{t+1} \notin \mathcal{B} \text{ and } f(X_{t+1}) < m \,|\, X_t \in \mathcal{B}\right] \leq \mathrm{O}(n^{-C_2}).$$

Then,

$$\mathbf{Pr}\left[f(X_\tau) \geq m \mid X_0 \in \mathcal{B}\right] \geq 1 - n^{-\Omega(1)}$$

holds for some $\tau = \mathrm{O}(\log n)$.

Proof. Let $\Omega' = \mathcal{B} \cup \{a, b\}$ be the state space with two special states a and b. We consider a Markov chain $(X'_t)_{t+1}^{\infty}$ on Ω' by

$$\mathbf{Pr}[X'_{t+1} = x \mid X'_t = y] = \begin{cases} \mathbf{Pr}[X_{t+1} = x \mid X_t = y] & \text{if } x, y \in \mathcal{B}, \\ \mathbf{Pr}[X_{t+1} \notin \mathcal{B} \wedge f(X_{t+1}) < m \mid X_t = y] & \text{if } x = a \text{ and } y \in \mathcal{B}, \\ \mathbf{Pr}[f(X_{t+1}) \geq m \mid X_t = y] & \text{if } x = b \text{ and } y \in \mathcal{B}, \\ 1 & \text{if } x = y \in \{a, b\}. \end{cases}$$

In other words, the special state a corresponds to the event "$f(x) < m$ and $x \notin \mathcal{B}$ ", and b does "$f(x) \geq m$".

Suppose that $X'_0 \in \mathcal{B}$ and let $\tau' = \min\{t : X'_t \notin \mathcal{B}\} > 0$ be the stopping time. Then, the above definition of X'_t naturally yields a coupling $(X_t, X'_t)_{t<\tau'}$ satisfying $X_t = X'_t$ for $t < \tau'$.

Let $f' : \Omega' \to \{0, \ldots, n\}$ be a function given by

$$f'(x) = \begin{cases} f(x) & \text{if } x \in \mathcal{B}, \\ n & \text{if } x \in \{a, b\}. \end{cases}$$

Then, the Markov chain (X'_t) on Ω' and the function f' satisfies the conditions (1) and (2) of Proposition 6.3.1. Hence, for some $\tau = \mathrm{O}(\log n)$, it holds that $X'_\tau \in \{a, b\}$. We insist that $X'_\tau = b$, that is, $f(X_\tau) \geq m$. Indeed, from the condition (3′), we have $\mathbf{Pr}[X'_\tau = a \mid X'_0 \in \mathcal{B}] \leq \tau \cdot \mathrm{O}(n^{-c_2}) \leq n^{-\Omega(1)}$. □

6.3.2 Linear algebra

Definition 6.3.3 (singular value)**.** *For a real matrix* $A \in \mathbb{R}^{m \times n}$, singular values $\sigma_1, \ldots, \sigma_m$ *of* A *are nonnegative square roots of eigenvalues of* $AA^\top$. *We write* $\sigma_i(A)$ *when we specify* A. *In particular,* the maximum singular value*, denoted by* $\sigma_{\max}$*, is the largest value among all singular values.*

Proposition 6.3.4. *For a real matrix* $A \in \mathbb{R}^{m \times n}$*, it holds that*

$$\sigma_{\max} = \max_{\mathbf{v} \in \mathbb{R}^n : \|\mathbf{v}\|_2 = 1} \|A\mathbf{v}\|_2,$$

where the norm $\|\cdot\|_2$ *is the* ℓ^2 *norm.*

In particular, it holds that

$$\|A\mathbf{v}\|_2 \leq \sigma_{\max} \|\mathbf{v}\|_2$$

for any vector $\mathbf{v} \in \mathbb{R}^n$.

6.3.3 Real analysis

Proposition 6.3.5 (The inverse function theorem; Theorem 12.17 of [Kra16])**.** *Let* f *be a continuously differentiable function from an open set* $U \subseteq \mathbb{R}^k$ *into* $\mathbb{R}^k$. *Suppose that the Jacobian matrix* J *at* $\mathbf{p} \in U$ *is invertible. Then there is a neighborhood* V *of* $\mathbf{p}$ *such that the restriction of* f *to* V *is invertible. Moreover, the Jacobian matrix of* f^{-1} *at* p *is given by* J^{-1}.

Definition 6.3.6. *Consider a function* $H : S \to T$*, where* $S \subseteq \mathbb{R}^m$ *and* $T \subseteq \mathbb{R}^n$ *are closed sets. The function* H *satisfies the* Lipschitz condition *if there is a constant* $C > 0$ *such that*

$$\|H(\mathbf{x}) - H(\mathbf{y})\|_\infty \leq C \|\mathbf{x} - \mathbf{y}\|_\infty$$

holds for any $\mathbf{x}, \mathbf{y} \in S$.

It should be noted that the definition of the Lipschitz condition does not depends on the norm.

Proposition 6.3.7 (Exercise 1D.3 of [DR14]). *Let $O \subseteq \mathbb{R}^k$ be an open set and $S \subseteq O$ be a compact convex subset of O. Suppose that $H : O \to \mathbb{R}^k$ is continuously differentiable on an open set O. Then H is Lipschitz continuous on C and*

$$\|H(\mathbf{x}) - H(\mathbf{y})\|_\infty \leq \max_{\mathbf{p}\in S} \sigma_{\max}(J_{\mathbf{p}})\|\mathbf{x} - \mathbf{y}\|_\infty,$$

where $J_{\mathbf{p}}$ is the Jacobian matrix at $\mathbf{p}$.

Corollary 6.3.8. *Let $H : \mathbb{R}^m \to \mathbb{R}^n$ be a function given by*

$$H(\mathbf{x}) = (H_1(\mathbf{x}), \ldots, H_n(\mathbf{x})),$$

where $H_i(\mathbf{x}) = H_i(x_1, \ldots, x_m)$ is a polynomial on $x_1, \ldots, x_m$ for all $i \in [n]$. Then, H satisfies the Lipschitz condition on $[0, 1]^m$.

6.4 Best-of-Two and Best-of-Three on Stochastic Block Model

This section is devoted proving our main results Theorems 6.1.4 to 6.1.7. For notational convenience, let $f_{\mathrm{Bo3}}(x) := 3x^2 - 2x^3$, $f_{\mathrm{Bo2}}(x) := x^2$, and $\overline{f_{\mathrm{Bo2}}}(x) := 1 - f_{\mathrm{Bo2}}(1-x) = x(2-x)$. Recall that the best-of-two (best-of-three) is a polynomial voting with respect to f_{Bo2} (f_{Bo3}, respectively). Consider the best-of-two on an f_{Bo2}- and $\overline{f_{\mathrm{Bo2}}}$-good graph, or the best-of-three on an f_{Bo3}-good graph (note that f_{Bo3} satisfies $f_{\mathrm{Bo3}}(x) = 1 - f_{\mathrm{Bo3}}(1-x)$ for every $x \in [0, 1]$). We consider the behavior of $\boldsymbol{\delta}$ defined as (6.8).

Let $u := \frac{1-r}{1+r}$. Then we have $\mathbf{E}[\delta_i'] = T_i(\delta_1, \delta_2) + \mathrm{O}\left(\frac{1}{\sqrt{np}}\right)$, where, in the best-of-three,

$$T_1(d_1, d_2) := \frac{ud_1}{2}\left(3 - (ud_1)^2 - 3d_2^2\right),\ T_2(d_1, d_2) := \frac{d_2}{2}\left(3 - 3(ud_1)^2 - d_2^2\right), \tag{6.12}$$

and in the best-of-two,

$$T_1(d_1, d_2) := \frac{d_1}{2}\left((2u+1) - (ud_1)^2 - (2u+1)d_2^2\right),\ T_2(d_1, d_2) := \frac{d_2}{2}\left(3 - u(2+u)d_1^2 - d_2^2\right). \tag{6.13}$$

Note that $T\colon S \to S$, where S is defined as (6.10). For notational convenience, we refer $\mathbf{d}'$ to $T(\mathbf{d})$. The dynamical system H of the best-of-three is illustrated in Figure 6.3.

Lemma 6.4.1. *Consider the best-of-two or best-of-three. For any $\mathbf{d} \in S$, it holds that $\mathbf{d}' \in S$.*

Proof. In this chapter, we prove Lemma 6.4.1 for the best-of-three. The case of the best-of-two can be shown in the same way. If $(d_1, d_2) \in S$, we have $3d_2^2 + (ud_1)^2 \leq \max\{3, u^2\} \leq 3$ and $d_2^2 + 3(ud_1)^2 \leq \max\{1, 3u^2\} \leq 3$. Hence, from (6.12), we have $d_1' \geq 0$ and $d_2' \geq 0$. Let $x = \frac{1+d_1+d_2}{2}$ and $y = \frac{1+d_2-d_1}{2}$. Then, $(d_1, d_2) \in S$ implies $\frac{1}{2} \leq x \leq 1$ and $0 \leq y \leq 1$. In addition, a simple calculation yields $T_1(d_1, d_2) + T_2(d_1, d_2) = 3\left(\frac{x+ry}{1+r}\right)^2 - 2\left(\frac{x+ry}{1+r}\right)^3 \leq 1$, where $r = \frac{1-u}{1+u}$. Note that $0 \leq \frac{x+ry}{1+r} \leq 1$ and the function $f : z \mapsto 3z^2 - 2z^3$ satisfies $f(z) \leq f(1) = 1$ for all $0 \leq z \leq 1$. Therefore, $\mathbf{d}' \in S$. □

From Lemma 6.4.1 and the symmetry of T, it suffices to consider the case of $\boldsymbol{\delta}^{(0)} \in S$.

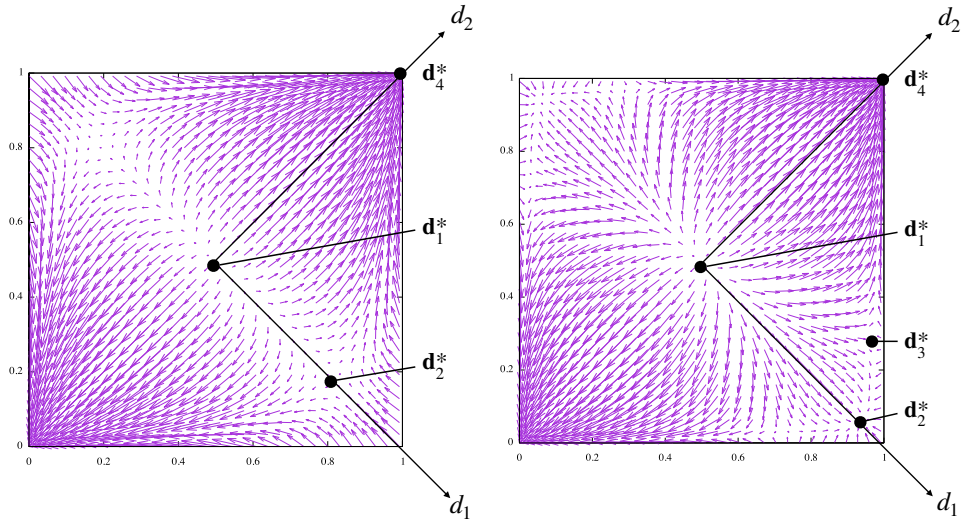
6.4.1 Best-of-three

It is straightforward to check that fixed points of (6.12) in S are $\mathbf{d}_1^*, \mathbf{d}_2^*, \mathbf{d}_3^*, \mathbf{d}_4^*$, where

$$\mathbf{d}_i^* := \begin{cases} (0,0) & \text{if } i = 1, \\ \left(\sqrt{\frac{3u-2}{u^3}}, 0\right) & \text{if } i = 2 \text{ and } u \geq \frac{2}{3}, \\ \left(\sqrt{\frac{1}{4u^3}}, \sqrt{\frac{4u-3}{4u}}\right) & \text{if } i = 3 \text{ and } u \geq \frac{3}{4}, \\ (0,1) & \text{if } i = 4. \end{cases} \tag{6.14}$$

The Jacobian matrix at (d_1, d_2) of the dynamical system (6.12) is

$$J = \frac{3}{2}\begin{pmatrix} u(1 - (ud_1)^2 - d_2^2) & -2ud_1d_2 \\ -2u^2d_1d_2 & 1 - (ud_1)^2 - d_2^2 \end{pmatrix}. \tag{6.15}$$



(a) $r = 1/6$ (b) $r = 1/9$

Figure 6.3: The induced dynamical system H of the best-of-three. The horizontal and vertical axes correspond to α_1 and α_2, respectively. The points $\mathbf{d}_i^*$ are the fixed points of T in S. In (a), the only sink point is $\mathbf{d}_4^*$, which is the consensus point.

Proposition 6.4.2 (Orbit convergence)**.** *Consider a sequence* $(\mathbf{d}^{(t)})_{t=0}^{\infty}$ *such that* $\mathbf{d}^{(0)} \in S$ *and* $d^{(t+1)} = T(d^{(t)})$*. Then* $\lim_{t\to\infty} \mathbf{d}^{(t)} = \mathbf{d}_i^*$ *for some* $i \in \{1,2,3,4\}$*. Additionally, suppose that* $u < \frac{3}{4}$ *and* $\mathbf{d}^{(0)} = (d_1^{(0)}, d_2^{(0)}) \in S$ *satisfies* $d_2^{(0)} > c$ *for some constant* $c > 0$*, then* $\lim_{t\to\infty} \mathbf{d}^{(t)} = \mathbf{d}_4^*$.

Proof of Proposition 6.4.2. If $u = 1$, we have $d_1' + d_2' = f(d_1 + d_2)$ and $d_1' - d_2' = f(d_1 - d_2)$, where $f : z \mapsto \frac{1}{2}z(3-z^2)$. Since $f(z) > z$ for $z \in (0,1)$, we have $\lim_{t\to\infty} \mathbf{d}^{(t)} \in \{(0,0),(0,1),(1,0)\}$. In addition, if $d_2^{(0)} > 0$ then $d_2^{(t)} \to 1$ as $t \to \infty$. Suppose that $0 \le u < 1$. Note that T is C^1 and injective[1] on S. It is straightforward to check that the conditions (C1) and (C2) hold (see (6.15)). Therefore, from Theorem 6.2.4, we obtain the first claim of Proposition 6.4.2.

We show the second claim of Proposition 6.4.2. Suppose that $u < \frac{3}{4}$. From (6.14), there are at most three fixed points $\mathbf{d}_1^*, \mathbf{d}_4^*$, and $\mathbf{d}_4^*$ of T (if $u \ge \frac{2}{3}$). From the first statement of Proposition 6.4.2, we have $\lim_{t\to\infty} \mathbf{d}^{(t)} = \mathbf{d}_i^*$ for some $i \in \{1,2,4\}$. If $i = 4$, we are done. Suppose that $i = 1$. Then, for any $\epsilon > 0$, there is $T \in \mathbb{N}$ such that $\|\mathbf{d}^{(t)} - \mathbf{d}_1^*\|_\infty < \epsilon$ for all $t \ge T$. Recall that, from the assumption, $d_2^{(0)} > 0$. From (6.12), it is easy to check that $d_2' > d_2^2$ if $d_2 > 0$. Fix a sufficiently small constant $\epsilon > 0$ and a point $\mathbf{d}$ such that $\|\mathbf{d} - \mathbf{d}_1^*\|_\infty < \epsilon$ and $d_2 > 0$. Then, from (6.12), we have $d_2' \ge 1.5(d_2 - 3u^2\epsilon^2 - \epsilon^2) > 1.49 d_2$. Therefore, if $d_2^{(0)}$ satisfies $\|\mathbf{d}^{(0)} - \mathbf{d}_1^*\|_\infty < \epsilon$ and $d_2^{(0)} > 0$, then either $d_1^{(t)} > \epsilon$ or $d_2^{(t)} > \epsilon$ holds for some $t = O_{\epsilon\to 0}(\log \epsilon^{-1})$. This contradicts to the assumption that $\lim_{t\to\infty} \mathbf{d}^{(t)} = \mathbf{d}_1^*$. Therefore, we have $\lim_{t\to\infty} \mathbf{d}^{(t)} \ne \mathbf{d}_1^*$. Similarly, we can show that $\lim_{t\to\infty} \mathbf{d}^{(t)} \ne \mathbf{d}_2^*$ (when $\frac{2}{3} \le u < \frac{3}{4}$). □

Now we focus on the behavior of $\boldsymbol{\delta}^{(t)}$ when $\boldsymbol{\delta}^{(0)}$ is around a fixed point. Table 6.1 shows the property of eigenvalues of the Jacobian matrix at $\mathbf{d}_i^*$ for each $i \in \{1,2,3,4\}$.

Table 6.1: Each (c_1, c_2) represents the property of the eigenvalues $\lambda_1 \ge \lambda_2$ of the corresponding Jacobian matrix. Specifically, c_i represents the sign of $\lambda_i - 1$. For example, $(+,0)$ means that $\lambda_1 > \lambda_2 = 1$ and $(+,-)$ means that $\lambda_1 > 1 > \lambda_2$.

points	$0 < u < \frac{2}{3}$	$u = \frac{2}{3}$	$\frac{2}{3} < u < \frac{3}{4}$	$u = \frac{3}{4}$	$\frac{3}{4} < u \le 1$
$\mathbf{d}_1^*$	$(+,-)$	$(+,1)$	$(+,+)$	$(+,+)$	$(+,+)$
$\mathbf{d}_2^*$	undefined	$(+,1)$	$(+,-)$	$(1,-)$	$(-,-)$
$\mathbf{d}_3^*$	undefined	undefined	undefined	$(1,-)$	$(+,-)$
$\mathbf{d}_4^*$	$(-,-)$	$(-,-)$	$(-,-)$	$(-,-)$	$(-,-)$

[1]Note that $\det J > 0$ for any $\mathbf{d} \in S \setminus \{(0,1)\}$. Then, from the Inverse Function Theorem (Proposition 6.3.5), T is injective on S.

Recall that $B(\mathbf{x}, r)$ is the open ball of radius r with respect to the ℓ^∞-norm. For $\mathbf{d} = (d_1, d_2) \in \mathbb{R}^2$, let $\langle \mathbf{d} \rangle_+ := (|d_1|, |d_2|) \in \mathbb{R}^2$.

Proposition 6.4.3. *Consider the best-of-three on an f_{Bo3}-good graph for a partition (V_1, V_2) and parameters p, q such that $r = q/p < 1/7$ is a constant. Then there is a constant $\epsilon = \epsilon(r) > 0$ satisfying*

$$\mathbf{Pr}\left[\langle \boldsymbol{\delta}' \rangle_+ \notin B(\mathbf{d}_2^*, \epsilon) \,\middle|\, \langle \boldsymbol{\delta} \rangle_+ \in B(\mathbf{d}_2^*, \epsilon)\right] \le \exp(-\Omega(n)).$$

In particular, $T_{\mathrm{cons}}(A) = \exp(\Omega(n))$ w.h.p. for any A satisfying $\langle \boldsymbol{\delta} \rangle_+ \in B(\mathbf{d}_2^*, \epsilon)$.

Proof. From Table 6.1, it is straightforward to check that the points $\mathbf{d}_2^*$ and $-\mathbf{d}_2^*$ are sink if $r < \frac{1}{7}$ (or equivalently, $u > \frac{3}{4}$). Therefore, Proposition 6.4.3 immediately follows from Proposition 6.2.8. □

Proposition 6.4.4. *Consider the best-of-three on an f_{Bo3}-good graph for a partition (V_1, V_2) and parameters p, q such that $r = q/p$ is a constant. Then, for some constant $\epsilon = \epsilon(r) > 0$, $T_{\mathrm{cons}}(A) \le \mathrm{O}(\log\log n + \log n/\log(np))$ holds w.h.p. for any $A \subseteq V$ satisfying $\min\{|A|, 2n - |A|\} \le \epsilon n$.*

Proof. Note that the Jacobian matrix at $\mathbf{d}_4^*$ is the all-zero matrix and the same holds for $-\mathbf{d}_4^*$. Let $\epsilon > 0$ be sufficiently small constant. If A satisfies $|A| \le \epsilon n$, apply Proposition 6.2.9. If A satisfies $|A| \ge (2-\epsilon)n$, apply Proposition 6.2.9 for $V \setminus A$. □

Proposition 6.4.5. *Consider the best-of-three on an f_{Bo3}- and $f_{\mathrm{Bo3}}(1 - f_{\mathrm{Bo3}})$-good graph for a partition (V_1, V_2) and parameters p, q such that p and q are constants. If $q/p > 1/7$ and $|\delta_2^{(0)}| = \mathrm{o}(1)$, then it holds w.h.p. that $|\delta_2^{(t)}| > \kappa$ for some $t = \mathrm{O}(\log n)$ and some constant $\kappa > 0$.*

Proof. Suppose that $u < \frac{3}{4}$ (or equivalently, $r > \frac{1}{7}$) and that $p \ge q > 0$ are constants. From Proposition 6.4.2, we may assume

$$\boldsymbol{\delta}^{(0)} \in \bigcup_{i \in \{1,2\}} B(\mathbf{d}_i^*, \epsilon_2) \tag{6.16}$$

for a sufficiently small constant $\epsilon_2 > 0$. We use Propositions 6.2.11 and 6.2.12.

First, we check the condition (A2) of Assumption 6.2.10. Note that variance $\mathbf{Var}[|A_i'| \mid A]$ can be written as $\mathbf{Var}[|A_i'|] = \sum_{v \in V_i} g\left(\frac{\deg_A(v)}{\deg(v)}\right)$ for any $A \subseteq V$, where $g(x) := f_{\mathrm{Bo3}}(x)(1 - f_{\mathrm{Bo3}}(x))$. Therefore, from the property (P2) of g-goodness, there are two constants $C_1 > 0, C_2 > 0$ such that

$$\forall A \subseteq V, \forall i \in \{1,2\} : \left|\mathbf{Var}\left[|A_i'|\right] - n \cdot g\left(\frac{|A_i|p + |A_{3-i}|q}{n(p+q)}\right)\right| \le C_2\sqrt{\frac{n}{p}}. \tag{6.17}$$

Using $z_i := \frac{a_i + r a_{3-i}}{1+r}$, we rewrite $\mathbf{Var}[\alpha_i'] = \mathbf{Var}[|A_i'|]/n^2$ as

$$\mathbf{Var}[\alpha_i'] = \frac{z_i^2(3 - 2z_i)(1 - z_i)^2(2z_i + 1)}{n} \pm \mathrm{O}\left(\frac{1}{\sqrt{n^3 p}}\right).$$

Note that $\mathbf{Var}[\alpha_i'] = \Omega(n^{-1})$ if $\alpha_1 < 1 - \epsilon_3$ or $\alpha_2 < 1 - \epsilon_3$ for some constant $\epsilon_3 > 0$. Therefore, the statement (A2) holds for every $\boldsymbol{\delta}$ satisfying $\boldsymbol{\delta} \in \bigcup_{i \in \{1,2\}} B(\mathbf{d}_i^*, \epsilon_2)$ with sufficiently small constant $\epsilon_2 < 1$ mentioned in (6.16).

We consider two cases: $u \neq \frac{2}{3}$ and $u = \frac{2}{3}$.

The case of $u \neq \frac{2}{3}$. A straightforward calculation of the Jacobian matrix implies that both $\mathbf{d}_1^*$ and $\mathbf{d}_2^*$ satisfies the conditions (A1) and (A3) of Assumption 6.2.10. Moreover, the condition of Proposition 6.2.11 holds (see Table 6.1). Therefore, we can apply Proposition 6.2.11.

Suppose that $u < \frac{2}{3}$. Then the fixed point $\mathbf{d}_2^*$ does not exist and thus we may assume $\boldsymbol{\delta}^{(0)} \in B(\mathbf{d}_1^*, \epsilon_2)$. From Proposition 6.2.11, we have $|\delta_2^{(t)}| > \epsilon_2$ for some $t = \mathrm{O}(\log n)$ (note that, here, $\boldsymbol{\beta} = \boldsymbol{\delta}$ and the eigenvalues satisfy $0 \le \lambda_1 < 1 < \lambda_2$).

Suppose that $u > \frac{2}{3}$. Both eigenvalues of J_1 are strictly larger than 1. Hence, for $\mathbf{d}_1^*$, Proposition 6.2.11 implies that either $|\delta_1^{(t)}| > \epsilon_2$ or $|\delta_2^{(t)}| > \epsilon_2$ holds for some $t = \mathrm{O}(\log n)$ if $\boldsymbol{\delta}^{(0)} \in B(\mathbf{d}_1^*, \epsilon_2)$. If the former holds with $|\delta_2^{(t)}| = \mathrm{o}(1)$, then $\boldsymbol{\delta}^{(t+T)} \in B(\mathbf{d}_2^*, \epsilon_2)$ holds for some constant $T = T(\epsilon_2)$ since $d_1' > d_1$ holds whenever $0 < d_1 < \sqrt{\frac{3u-2}{u^3}}$ and $d_2 = 0$. Note that, at the point $\mathbf{d}_2^*$, the Jacobian matrix J_2 has eigenvalues λ_1, λ_2 satisfying $0 < \lambda_1 < 1 < \lambda_2$. Moreover, in look at (6.11), we have $\boldsymbol{\beta} = \boldsymbol{\delta} - \mathbf{d}_2^*$. Thus, Proposition 6.2.11 yields that $|\delta_2^{(t')}| > \epsilon_2$ holds for some $t' = \mathrm{O}(\log n)$ and for any $\boldsymbol{\delta}^{(0)} \in B(\mathbf{d}_2^*, \epsilon_2)$.

The case of $u = \frac{2}{3}$. In this case, we have $\mathbf{d}_1^* = \mathbf{d}_2^* = (0,0)$. We claim that this point satisfies (B1) and (B2) and then apply Proposition 6.2.12.

Let $\epsilon_2 > 0$ be sufficiently small constant mentioned in (6.16). The Jacobian matrix $J_1 = J_2$ has eigenvalues 1 and $\frac{3}{2}$. Suppose that $\|\boldsymbol{\delta}^{(0)}\|_\infty \le \epsilon_2$ for sufficiently small constant $\epsilon_2 > 0$. Then, we have $|\mathbf{E}[\delta_2']| = \left|\frac{\delta_2}{2}(3 - 3(u\delta_1)^2 - \delta_2^2)\right| \pm \mathrm{O}(n^{-0.5}) \ge 1.49|\delta_2| - \mathrm{O}(n^{-0.5})$. This verifies the assumption (B1). Now we check that the assumption (B2) holds. Note that (B2) is equivalent to

$$\mathbf{Pr}\left[|\delta_1'| \le \epsilon_2 \,|\, |\delta_1| \le \epsilon_2\right] \ge 1 - n^{-\Omega(1)}.$$

For any $\boldsymbol{\delta}$ of $|\delta_1| \le \epsilon_2$, we have

$$|\mathbf{E}[\delta_1']| = \left|\frac{u\delta_1}{2}\right| \left|3 - (u\delta_1)^2 - 3\delta_2^2\right| \pm \mathrm{O}(n^{-0.5}) \le |\delta_1|\left(1 - \frac{4}{27}\delta_1^2\right) + \mathrm{O}(n^{-0.5}).$$

Therefore, from the Hoeffding inequality (Proposition 5.4.3), if $|\delta_1| \le \epsilon_2$, it holds w.h.p. that

$$|\delta_1'| \le |\delta_1| - \frac{4}{27}|\delta_1|^3 + C\sqrt{\frac{\log n}{n}}$$

for sufficiently large constant $C > 0$ and large n. If $|\delta_1|^3 \ge \frac{27C}{4}\sqrt{\frac{\log n}{n}}$, we have $|\delta_1'| \le |\delta_1| \le \epsilon_2$ holds w.h.p. If $|\delta_1|^3 < \frac{27C}{4}\sqrt{\frac{\log n}{n}}$, we have $|\delta_1'| = \mathrm{O}\left(\sqrt{\frac{\log n}{n}}\right) \le \epsilon_2$ holds w.h.p.

Thus, from Proposition 6.2.11, we have $|\delta_2^{(t)}| > \epsilon_2$ w.h.p. for some $t = \mathrm{O}(\log n)$. This completes the proof of Proposition 6.4.5. □

Here, we prove Theorems 6.1.4 and 6.1.6 using Propositions 6.4.2 to 6.4.5.

Proof of Theorem 6.1.4. From Theorem 6.2.2, $G(2n,p,q)$ is f_{Bo3}-good. If $r > \frac{1}{7}$ and $A_0 \subseteq V$ satisfies $\left||A_0| - n\right| = \Omega(n)$, then we have $|d_2^{(0)}| = |\delta_2^{(0)}| > \kappa$ for some constant $\kappa > 0$. Next, for any constant $\epsilon > 0$, Proposition 6.4.2 implies $\langle \mathbf{d}^{(l)} \rangle_+ \in B(\mathbf{d}_4^*, \epsilon)$ for some constant $l = l(\epsilon)$. From (6.9), we have $\langle \boldsymbol{\delta}^{(l)} \rangle_+ \in B(\mathbf{d}_4^*, \epsilon)$ for sufficiently large n. Set ϵ be the constant mentioned in Proposition 6.4.4. Then, from Proposition 6.4.4, it holds w.h.p. that $T_{\mathrm{cons}}(A_0) \le l + T_{\mathrm{cons}}(A_l) \le \mathrm{O}(\log\log n + \log n/\log(np))$.

If $r < \frac{1}{7}$, Proposition 6.4.3 yields $T_{\mathrm{cons}}(A_0) \ge \exp(\Omega(n))$ w.h.p. for any $A_0 \subseteq V$ with $\boldsymbol{\delta}^{(0)} \in B(\mathbf{d}_2^*, \epsilon)$, where $\epsilon > 0$ is the constant from Proposition 6.4.3. This completes the proof of (ii). □

Proof of Theorem 6.1.6. From Theorem 6.2.2, $G(2n,p,q)$ is both f_{Bo3}- and $f_{\mathrm{Bo3}}(1-f_{\mathrm{Bo3}})$-good w.h.p. If $|\delta^{(0)}| = \mathrm{o}(1)$, then Proposition 6.4.5 yields that $|\delta^{(t)}| > \kappa$ for some constant $\kappa > 0$ and some $t = \mathrm{O}(\log n)$. Then, from Theorem 6.1.4, we have $T_{\mathrm{cons}}(A_t) \le \mathrm{O}(\log\log n + \log n/\log(np))$. Thus, $T_{\mathrm{cons}}(A_0) \le t + T_{\mathrm{cons}}(A_t) \le \mathrm{O}(\log n)$. □

6.4.2 Best-of-two

The induced dynamical system (6.13) of the best-of-two has the same form as that of the best-of-three. A straightforward calculation yields that $\mathbf{d}' = \mathbf{d} \in S$ holds if and only if $\mathbf{d} \in \{\mathbf{d}_i^*\}_{i=1}^4$, where

$$\mathbf{d}_i^* := \begin{cases} (0,0) & \text{if } i = 1, \\ \left(\sqrt{\frac{2u-1}{u^2}}, 0\right) & \text{if } i = 2 \text{ and } u \ge \frac{1}{2}, \\ \left(\sqrt{\frac{u^2+u-1}{(u+1)^2}}, \sqrt{\frac{1}{u(u+1)^2}}\right). & \text{if } i = 3 \text{ and } u \ge \frac{\sqrt{5}-1}{2}, \\ (0,1) & \text{if } i = 4. \end{cases} \tag{6.18}$$

The Jacobian matrix J at (d_1, d_2) is

$$J = \frac{1}{2}\begin{pmatrix} 2u + 1 - 3(ud_1)^2 - (2u+1)d_2^2 & -2(2u+1)d_1d_2 \\ -2u(u+2)d_1d_2 & 3 - u(2+u)d_1^2 - 3d_2^2 \end{pmatrix}. \tag{6.19}$$

See Table 6.2 for the eigenvalues of J at each $\mathbf{d}_i^*$.

Proposition 6.4.6. *For any sequence $(\mathbf{d}^{(t)})_{t=0}^\infty$, $\lim_{t\to\infty}\langle \mathbf{d}^{(t)} \rangle_+ = \mathbf{d}_i^*$ for some $i \in \{1,2,3,4\}$. Furthermore, if $u < \frac{\sqrt{5}-1}{2}$ and there is a positive constant $\kappa > 0$ such that the initial point $\mathbf{d}^{(0)} = (d_1^{(0)}, d_2^{(0)}) \in S$ satisfies $|d_2^{(0)}| > \kappa$, then $\lim_{t\to\infty}\langle \mathbf{d}^{(t)} \rangle_+ = \mathbf{d}_4^*$.*

Table 6.2: Each (c_1, c_2) represents the property of the eigenvalues $\lambda_1 \geq \lambda_2$ of the corresponding Jacobian matrix. Specifically, c_i represents the sign of $\lambda_i - 1$.

points	$0 < u < \frac{1}{2}$	$u = \frac{1}{2}$	$\frac{1}{2} < u < \frac{\sqrt{5}-1}{2}$	$u = \frac{\sqrt{5}-1}{2}$	$\frac{\sqrt{5}-1}{2} < u \leq 1$
$\mathbf{d}_1^*$	$(+,-)$	$(+,1)$	$(+,+)$	$(+,+)$	$(+,+)$
$\mathbf{d}_2^*$	undefined	$(+,1)$	$(+,-)$	$(1,-)$	$(-,-)$
$\mathbf{d}_3^*$	undefined	undefined	undefined	$(1,-)$	$(+,-)$
$\mathbf{d}_4^*$	$(-,-)$	$(-,-)$	$(-,-)$	$(-,-)$	$(-,-)$

Proposition 6.4.7. *Consider the best-of-two on an f_1^{Bo2}- and f_2^{Bo2}-good graph for a partition (V_1, V_2) andparameters p, q such that $r = q/p < \sqrt{5} - 2$ is a constant. Then, there is a constant $\epsilon = \epsilon(r) > 0$ satisfying*

$$\mathbf{Pr}\left[\boldsymbol{\delta}' \notin B(\mathbf{d}_2^*, \epsilon) \,\middle|\, \boldsymbol{\delta} \in B(\mathbf{d}_2^*, \epsilon)\right] \leq \exp(-\Omega(n)).$$

In particular, $T_{\mathrm{cons}}(A) = \exp(\Omega(n))$ holds w.h.p. for any $A \subseteq V$ satisfying $\langle \boldsymbol{\delta} \rangle_+ \in B(\mathbf{d}_2^*, \epsilon)$.

Proposition 6.4.8. *Consider the best-of-two on an f_{Bo3}-good graph for a partition (V_1, V_2) and parameters p, q such that $r = q/p$ is a constant. Then, for some constant $\epsilon = \epsilon(r) > 0$, $T_{\mathrm{cons}}(A) \leq \mathrm{O}(\log\log n + \log n/\log(np))$ holds w.h.p. for any $A \subseteq V$ satisfying $\min\{|A|, 2n - |A|\} \leq \epsilon n$.*

Proposition 6.4.9. *Consider the best-of-two on an f_{Bo2}-, $f_{\mathrm{Bo2}}(1-f_{\mathrm{Bo2}})$-, $\overline{f_{\mathrm{Bo2}}}$- and $\overline{f_{\mathrm{Bo2}}}(1-\overline{f_{\mathrm{Bo2}}})$-good graph for a partition (V_1, V_2) and parameters p, q such that p and q are constants. If $q/p > \sqrt{5} - 2$ and $|\delta_2^{(0)}| = \mathrm{o}(1)$, then it holds w.h.p. that $|\delta_2^{(t)}| > \kappa$ for some $t = \mathrm{O}(\log n)$ and some constant $\kappa > 0$.*

We omit proofs of Propositions 6.4.6 to 6.4.9 since they are substantially the same as that of Propositions 6.4.2 to 6.4.5.

Proof of Theorems 6.1.5 and 6.1.7. The proof of Theorem 6.1.5 is the same as that of Theorem 6.1.4 except for the threshold and using Propositions 6.4.6 to 6.4.9 instead of Propositions 6.4.2 to 6.4.5. □

6.5 Proof of f-Goodness

In this section we show Theorem 6.2.2. In Section 6.5.1, we show that the property (P2) is obtained from Lemma 6.5.1.

Lemma 6.5.1. *For a finite set V with $|V| = N$, let $(I_e)_{e \in \binom{V}{2}}$ be $\binom{N}{2}$ independent binary random variables. Let $p := \max_{e \in \binom{V}{2}} \mathbf{E}[I_e]$. Suppose that $Np \geq 1$. For $\ell + 1$ vertex subsets $S_0, S_1, \ldots, S_\ell$, let*

$$W(S_0; S_1, \ldots, S_\ell) := \sum_{s \in S_0} \prod_{i \in [\ell]} \deg_{S_i}(s), \tag{6.20}$$

$$\hat{W}(S_0; S_1, \ldots, S_\ell) := \sum_{s \in S_0} \prod_{i \in [\ell]} \mathbf{E}[\deg_{S_i}(s)] \tag{6.21}$$

where $\deg_S(v) = \sum_{s \in S \setminus \{v\}} I_{\{v,s\}}$ for $S \subseteq V$ and $v \in V$.

Then there are two positive constants C_1, C_2 depending only on ℓ such that the following holds with probability $1 - N^{-C_1}$:

$$\forall S_0, S_1, \ldots, S_\ell : \left|W(S_0; S_1, \ldots, S_\ell) - \hat{W}(S_0; S_1, \ldots, S_\ell)\right| \leq C_2 N (Np)^{\ell - 1/2}.$$

Our proof of Lemma 6.5.1 consists of three parts. First, we give a concentration of W (Lemma 6.5.3). Next, we give an upper bound on the gap between $\mathbf{E}[W]$ and $\hat{W}$ (Lemma 6.5.4). At the end, we show Lemma 6.5.2 which plays a key role in showing Lemmas 6.5.3 and 6.5.4.

6.5.1 Reduction to the concentration of W

Proof of (P2) of Theorem 6.2.2 via Lemma 6.5.1. Let $f(x) = \sum_{j=0}^{\ell} c_j x^j$. For notational convenience, let $x_v = \frac{\deg_A(v)}{\deg(v)}$, $\bar{x}_v = \frac{\mathbf{E}[\deg_A(v)]}{\mathbf{E}[\deg(v)]}$ and $\hat{x} = \frac{|A_i|p+|A_{3-i}|q}{n(p+q)}$. Then from the triangle inequality, it holds that

$$\left|\sum_{v\in S\cap V_i}\big(f(x_v)-f(\hat{x})\big)\right| \le \left|\sum_{v\in S\cap V_i}\big(f(x_v)-f(\bar{x}_v)\big)\right| + \left|\sum_{v\in S\cap V_i}\big(f(\bar{x}_v)-f(\hat{x})\big)\right|. \tag{6.22}$$

For the second term of the right hand of (6.22), there are two positive constants C_1, C_2 such that

$$\left|\sum_{v\in S\cap V_i}\big(f(\bar{x}_v)-f(\hat{x})\big)\right| \le \sum_{v\in S\cap V_i}|f(\bar{x}_v)-f(\hat{x})| \le C_1\sum_{v\in S\cap V_i}|\bar{x}_v-\hat{x}| \le C_2\frac{|S\cap V_i|}{n} \le C_2.$$

The second inequality follows from the Lipschitz condition of f (cf. Section 6.3.3). The third inequality holds since $\mathbf{E}[\deg(v)] = (n-1)p+nq$ and $(|A_i|-1)p+|A_{3-i}|q \le \mathbf{E}[\deg_A(v)] \le |A_i|p+|A_{3-i}|q$ for any $v\in V_i$.

For the first term of the right hand of (6.22), since

$$\left(\frac{\deg_A(v)}{\deg(v)}\right)^j - \left(\frac{\mathbf{E}[\deg_A(v)]}{\mathbf{E}[\deg(v)]}\right)^j = \frac{\left(\mathbf{E}[\deg(v)]^j-\deg(v)^j\right)\left(\frac{\deg_A(v)}{\deg(v)}\right)^j + \left(\deg_A(v)^j-\mathbf{E}[\deg_A(v)]^j\right)}{\mathbf{E}[\deg(v)]^j}$$

for any j and $v\in V$, we have

$$\begin{aligned}\left|\sum_{v\in S\cap V_i}\big(f(x_v)-f(\bar{x}_v)\big)\right| &= \left|\sum_{j=1}^{\ell} c_j\sum_{v\in S\cap V_i}\big((x_v)^j-(\bar{x}_v)^j\big)\right| \\ &\le \sum_{j=1}^{\ell}\frac{|c_j|}{((n-1)p)^j}\left(\left|\sum_{v\in S\cap V_i}\left(\mathbf{E}[\deg(v)]^j-\deg(v)^j\right)(x_v)^j\right| + \left|\sum_{v\in S\cap V_i}\left(\deg_A(v)^j-\mathbf{E}[\deg_A(v)]^j\right)\right|\right).\end{aligned}$$

Note that $\mathbf{E}[\deg(v)] = (n-1)p+nq \ge (n-1)p$ for any $v\in V$. Since

$$\begin{aligned}\left|\sum_{v\in S\cap V_i}\left(\mathbf{E}[\deg(v)]^j-\deg(v)^j\right)(x_v)^j\right| &\le \max_{U\subseteq V}\left|\sum_{u\in U}\left(\mathbf{E}[\deg(u)]^j-\deg(u)^j\right)\right| \\ &= \max_{U\subseteq V}\left|W(U;\overbrace{V,\dots,V}^{j}) - \hat{W}(U;\overbrace{V,\dots,V}^{j})\right|\end{aligned}$$

and $\sum_{v\in S\cap V_i}\left(\deg_A(v)^j-\mathbf{E}[\deg_A(v)]^j\right) = W(S\cap V_i;\overbrace{A,\dots,A}^{j}) - \hat{W}(S\cap V_i;\overbrace{A,\dots,A}^{j})$, we obtain the claim from Lemma 6.5.1. Note that, for any $S\subseteq V$, $a\in\mathbb{R}^V$ and $x\in[0,1]^V$, $|\sum_{s\in S}a_s x_s| \le \max_{U\subseteq V}|\sum_{u\in U}a_u|$ since $\sum_{s\in S:a_s\le 0}a_s \le \sum_{s\in S}a_s x_s \le \sum_{s\in S:a_s\ge 0}a_s$. □

Now we introduce the following lemma, which we will use in Sections 6.5.2 and 6.5.3.

Lemma 6.5.2. *Let V be a finite set of size N and fix $l+1$ subsets $S_0, S_1, \dots, S_l \subseteq V$. For any $\mathbf{s} = (s_0, s_1, \dots, s_l) \in S_0\times S_1\times\cdots\times S_l$, define*

$$U(\mathbf{s}) := \{s_i : i\in\{0\}\cup[l]\}.$$

Consider $\sum_{\mathbf{s}\in\mathcal{S}} p^{|\mathcal{F}(\mathbf{s})|}$, where $p\in[1/N,1]$, $\mathcal{S}\subseteq S_0\times S_1\times\cdots\times S_l$ and

$$\mathcal{F}: S_0\times S_1\times\cdots\times S_l \to 2^{\binom{V}{2}}.$$

Suppose that the following three conditions hold for any $\mathbf{s}\in\mathcal{S}$: (1) $|\mathcal{F}(\mathbf{s})|\le k$, (2) $\mathcal{F}(\mathbf{s})\subseteq\binom{U(\mathbf{s})}{2}$ and (3) the graph $G(\mathbf{s}) = \big(U(\mathbf{s}),\mathcal{F}(\mathbf{s})\big)$ is connected. Let $L\subseteq\{0\}\cup[l]$ be a set of indices such that $S_i\cap S_j = \emptyset$ for any $i,j\in L$ $(i\ne j)$. Then

$$\sum_{\mathbf{s}\in\mathcal{S}} p^{|\mathcal{F}(\mathbf{s})|} \le \mathcal{B}_{l+1}N(Np)^k\frac{\prod_{i\in L}|S_i|}{N^{|L|}}$$

where $\mathcal{B}_l$ denotes the l-th Bell number.

The l-th *Bell number* $\mathcal{B}_l$ is the number of possible partitions of a set with l labeled elements. It is known that $\mathcal{B}_l < \left(\frac{0.792l}{\ln(l+1)}\right)^l$ for all positive integer l [DT10].

6.5.2 Concentration of W

Lemma 6.5.3. *Let W and $\hat{W}$ be the values defined in (6.20) and (6.21), respectively. Then there are two positive constants C_1, C_2 depending only on ℓ such that the following holds with probability $1 - N^{-C_1}$:*

$$\forall S_0, S_1, \dots, S_\ell : \left|W(S_0; S_1, \dots, S_\ell) - \mathbf{E}[W(S_0; S_1, \dots, S_\ell)]\right| \le C_2 N (Np)^{\ell - 1/2}.$$

Proof. For $\ell + 1$ vertex subsets $S_0, S_1, \dots, S_\ell$, let

$$\mathbf{S} := \{(s_0, s_1, \dots, s_\ell) : s_0 \in S_0, s_i \in S_i \setminus \{s_0\} \text{ for every } i \in [\ell]\}$$

and for any $\mathbf{s} = (s_0, s_1, \dots, s_\ell) \in \prod_{i=0}^{\ell} S_i = S_0 \times S_1 \times \cdots \times S_\ell$, let

$$F(\mathbf{s}) := \big\{\{s_0, s_i\} : i \in [\ell]\big\} \setminus \{s_0\}.$$

For example, $F\big((a, b, a, c, d, b, f, a)\big) = \big\{\{a, b\}, \{a, c\}, \{a, d\}, \{a, f\}\big\}$. Then,

$$W(S_0; S_1, \dots, S_\ell) = \sum_{s_0 \in S_0} \prod_{i \in [\ell]} \left(\sum_{s_i \in S_i \setminus \{s_0\}} I_{\{s_0, s_i\}} \right) = \sum_{\mathbf{s} \in \mathbf{S}} \prod_{i \in [\ell]} I_{\{s_0, s_i\}} = \sum_{\mathbf{s} \in \mathbf{S}} \prod_{e \in F(\mathbf{s})} I_e. \tag{6.23}$$

Lower bound on W. First, we claim the following: There are two positive constants C_3, C_4 such that

$$\mathbf{Pr}\left[\forall S_0, S_1 \dots, S_\ell : W(S_0; S_1, \dots, S_\ell) \ge \mathbf{E}[W(S_0; S_1, \dots, S_\ell)] - C_4 N (Np)^{\ell - 1/2}\right] \ge 1 - N^{-C_3}. \tag{6.24}$$

To obtain (6.24), we apply Janson's inequality (Proposition 5.4.5) to (6.23). Then we have

$$\begin{aligned} &\mathbf{Pr}\left[\exists S_0, S_1, \dots, S_\ell : W(S_0; S_1, \dots, S_\ell) \le \mathbf{E}[W(S_0; S_1, \dots, S_\ell)] - t\right] \\ &\le \left(2^N\right)^{\ell+1} \exp\left(-\frac{t^2}{2\nabla(S_0; S_1, \dots, S_\ell)}\right) \le \exp\left((\ell+1)N - \frac{t^2}{2\nabla(S_0; S_1, \dots, S_\ell)}\right) \end{aligned} \tag{6.25}$$

where

$$\nabla(S_0; S_1, \dots, S_\ell) = \sum_{\substack{\mathbf{s} \in \mathbf{S}, \mathbf{s}' \in \mathbf{S}: \\ F(\mathbf{s}) \cap F(\mathbf{s}') \neq \emptyset}} \mathbf{E}\left[\prod_{e \in F(\mathbf{s})} I_e \prod_{e' \in F(\mathbf{s}')} I_{e'}\right].$$

Thus it suffices to show that $\nabla(S_0; S_1, \dots, S_\ell) = \mathrm{O}(N(Np)^{2\ell-1})$. Since $\max_{e \in \binom{V}{2}} \mathbf{E}[I_e] = p$, it holds that

$$\nabla(S_0; S_1, \dots, S_\ell) = \sum_{\substack{\mathbf{s} \in \mathbf{S}, \mathbf{s}' \in \mathbf{S}: \\ F(\mathbf{s}) \cap F(\mathbf{s}') \neq \emptyset}} \mathbf{E}\left[\prod_{e \in F(\mathbf{s})} I_e \prod_{e' \in F(\mathbf{s}')} I_{e'}\right] \le \sum_{\substack{\mathbf{s} \in \mathbf{S}, \mathbf{s}' \in \mathbf{S}: \\ F(\mathbf{s}) \cap F(\mathbf{s}') \neq \emptyset}} p^{|F(\mathbf{s}) \cup F(\mathbf{s}')|}. \tag{6.26}$$

To bound (6.26), we apply Lemma 6.5.2 which we will prove in Section 6.5.4. Consider $2\ell + 2$ vertex subsets $S'_0, S'_1, \dots, S'_{2\ell+1}$ where $S'_i := S_{i \bmod (\ell+1)}$. For any $i \in \{0\} \cup [2\ell + 1]$, let

$$\mathcal{S} := \big\{(s_0, s_1, \dots, s_{2\ell+1}) \in \mathbf{S} \times \mathbf{S} : F\big((s_0, \dots, s_\ell)\big) \cap F\big((s_{\ell+1}, \dots, s_{2\ell+1})\big) \neq \emptyset\big\} \subseteq \prod_{i=0}^{2\ell+1} S'_i,$$

$$\mathcal{F}(\mathbf{s}) := F\big((s_0, \dots, s_\ell)\big) \cup F\big((s_{\ell+1}, \dots, s_{2\ell+1})\big) \text{ for any } \mathbf{s} = (s_0, s_1, \dots, s_{2\ell+1}) \in \prod_{i=0}^{2\ell+1} S'_i.$$

Then for any $\mathbf{s} \in \mathcal{S}$, $G(\mathbf{s}) = \big(U(\mathbf{s}), \mathcal{F}(\mathbf{s})\big)$ is a connected graph and $|\mathcal{F}(\mathbf{s})| \le 2\ell - 1$. Thus, for any $i_* \in \{0\} \cup [\ell]$, Lemma 6.5.2 with letting $l = 2\ell + 1, k = 2\ell - 1$ and $L = \{i_*\}$ yields

$$\sum_{\substack{\mathbf{s} \in \mathbf{S}, \mathbf{s}' \in \mathbf{S}: \\ F(\mathbf{s}) \cap F(\mathbf{s}') \neq \emptyset}} p^{|F(\mathbf{s}) \cup F(\mathbf{s}')|} = \sum_{\mathbf{s} \in \mathcal{S}} p^{|\mathcal{F}(\mathbf{s})|} \le \mathcal{B}_{2(\ell+1)} N (Np)^{2\ell-1} \frac{|S'_{i_*}|}{N} = \mathcal{B}_{2(\ell+1)} |S_{i_*}| (Np)^{2\ell-1}. \tag{6.27}$$

Equations (6.26) and (6.27) imply the following statement: For any $\ell+1$ vertex subsets $S_0, S_1, \dots S_\ell$ and for any $i_* \in \{0\} \cup [\ell]$,

$$\nabla(S_0; S_1, \dots, S_\ell) \le \mathcal{B}_{2(\ell+1)} |S_{i_*}| (Np)^{2\ell-1} \le \mathcal{B}_{2(\ell+1)} N (Np)^{2\ell-1}. \tag{6.28}$$

Thus by substituting $t = C_4 N (Np)^{\ell-1/2}$ with $C_4 = \sqrt{2(\ell+1+C_3)\mathcal{B}_{2(\ell+1)}}$ to (6.25), we obtain the claim (6.24).

Upper bound on W. To complete the proof of Lemma 6.5.3, we combine the claim (6.24) and the following claim: There are two positive constants C_5, C_6 such that

$$\mathbf{Pr}\left[\forall S_0, S_1 \dots, S_\ell : W(S_0; S_1, \dots, S_\ell) \le \mathbf{E}[W(S_0; S_1, \dots, S_\ell)] + C_6 N (Np)^{\ell-1/2}\right] \ge 1 - N^{-C_5}. \tag{6.29}$$

To show the claim, we consider the following expression of W. For any $S_0, S_1, \dots, S_\ell$, let $W_0 := W(S_0; S_1, \dots, S_\ell)$ and let $W_i := W(\overbrace{V; V \dots, V}^{i}, S_i, S_{i+1}, \dots, S_\ell)$ for each $i \in [\ell+1]$. Since $W_{i+1} - W_i = W(\overbrace{V; V, \dots, V}^{i}, V \setminus S_i, S_{i+1}, \dots, S_\ell)$ for any $i \in \{0\} \cup [\ell+1]$ and $\sum_{i=0}^{\ell}(W_{i+1} - W_i) = W_{\ell+1} - W_0$, we have

$$W(S_0; S_1, \dots, S_\ell) = W(\overbrace{V; V, \dots, V}^{\ell+1}) - \sum_{i=0}^{\ell} W(\overbrace{V; V, \dots, V}^{i}, V \setminus S_i, S_{i+1}, \dots, S_\ell). \tag{6.30}$$

We can apply (6.24) for the second term of the right hand of (6.30). Now we try to get an upper bound on $W(\overbrace{V; V, \dots, V}^{\ell+1})$. For the notational convenience, let $Y = W(\overbrace{V; V, \dots, V}^{\ell+1})$. Let $S_i = V$ for every $i \in \{0\} \cup [\ell]$ and let

$$\mathcal{E} := \{F(\mathbf{s}) : \mathbf{s} \in \mathbf{S}\}.$$

From (6.23), we have

$$Y = \sum_{\mathbf{s}\in\mathbf{S}} \prod_{e \in F(\mathbf{s})} I_e = \sum_{F \in \mathcal{E}} |\{\mathbf{s} \in \mathbf{S} : F(\mathbf{s}) = F\}| \prod_{e \in F} I_e.$$

Thus applying Kim-Vu inequality (Proposition 5.4.6) to Y yields

$$\mathbf{Pr}\left[|Y - \mathbf{E}[Y]| \ge \sqrt{\ell! \max_{A \subseteq \binom{V}{2}} \mathbf{E}[Y_A] \max_{A \subseteq \binom{V}{2} : A \neq \emptyset} \mathbf{E}[Y_A]} (8\lambda)^\ell\right] \le 2\exp(2 + 2(\ell-1)\log N - \lambda) \tag{6.31}$$

where

$$Y_A = \sum_{\substack{F \in \mathcal{E}: \\ F \supseteq A}} |\{\mathbf{s} \in \mathbf{S} : F(\mathbf{s}) = F\}| \prod_{e \in F \setminus A} I_e = \sum_{\substack{\mathbf{s}\in\mathbf{S}: \\ F(\mathbf{s}) \supseteq A}} \prod_{e \in F(\mathbf{s}) \setminus A} I_e.$$

Now, we give an upper bound on $\mathbf{E}[Y_A]$. Since $\max_{e \in \binom{V}{2}} \mathbf{E}[J_e] = p$, it holds that

$$\mathbf{E}[Y_A] = \sum_{\substack{\mathbf{s}\in\mathbf{S}: \\ F(\mathbf{s}) \supseteq A}} \mathbf{E}\left[\prod_{e \in F(\mathbf{s}) \setminus A} I_e\right] \le \sum_{\substack{\mathbf{s}\in\mathbf{S}: \\ F(\mathbf{s}) \supseteq A}} p^{|F(\mathbf{s}) \setminus A|} = \sum_{\substack{\mathbf{s}\in\mathbf{S}: \\ F(\mathbf{s}) \supseteq A}} p^{|F(\mathbf{s})| - |A|}.$$

If $A = \emptyset$, a direct application of Lemma 6.5.2 with letting $l = k = \ell$ and $L = \emptyset$ yields

$$\mathbf{E}[Y_A] = \mathbf{E}[Y] \le \sum_{\mathbf{s}\in\mathbf{S}} p^{|F(\mathbf{s})|} \le \mathcal{B}_{\ell+1} N (Np)^\ell. \tag{6.32}$$

Note that $|F(\mathbf{s})| \le \ell$ and $G(\mathbf{s}) = (U(\mathbf{s}), F(\mathbf{s}))$ is a connected graph for any $\mathbf{s} \in \mathbf{S} \subseteq \prod_{i=0}^{\ell} S_i$.

Now we consider the case $|A| = \kappa \ge 1$. Let $V(A)$ be the set of vertices induced by the edge set $A \subseteq \binom{V}{2}$. If $F(\mathbf{s}) \supseteq A$ for some $\mathbf{s} \in \prod_{i=0}^{\ell} V$, the graph $G' = (V(A), A)$ is a star graph and hence

$|V(A)| = |A| + 1 = \kappa + 1$. Let $V(A) = \{a_0, a_1, \dots, a_\kappa\}$. Now consider $(\ell+1)+(\kappa+1)$ vertex subsets $S'_0, S'_1, \dots S'_{\ell+\kappa+1}$ where $S'_i = S_i$ for any $0 \le i \le \ell$ and $S'_i = \{a_{i-(\ell+1)}\}$ for any $\ell+1 \le i \le \ell+\kappa+1$. Let

$$\mathcal{S} := \left\{ (s_0, s_1, \dots, s_{\ell+\kappa+1}) \in \mathbf{S} \times \prod_{i=0}^{\kappa} \{a_i\} : F\big((s_0, \dots, s_\ell)\big) \supseteq A \right\} \subseteq \prod_{i=0}^{\ell+\kappa+1} S'_i,$$

$$\mathcal{F}(\mathbf{s}) := F\big((s_0, \dots, s_\ell)\big) \text{ for any } \mathbf{s} = (s_0, s_1, \dots, s_{\ell+\kappa+1}) \in \prod_{i=0}^{\ell+\kappa+1} S'_i.$$

Note that, for any $\mathbf{s} \in \mathcal{S}$, the graph $G(\mathbf{s}) = \big(U(\mathbf{s}), \mathcal{F}(\mathbf{s})\big)$ is connected and $|\mathcal{F}(\mathbf{s})| \le \ell$. Thus Lemma 6.5.2 with letting $l = \ell+\kappa+1$, $k = \ell$ and $L = \{\ell+1, \ell+2, \dots, \ell+\kappa+1\}$ (note that $a_i \ne a_j$ for any $i \ne j$ and $\prod_{i=\ell+1}^{\ell+\kappa+1} |S'_i| = \prod_{i=\ell+1}^{\ell+\kappa+1} |\{a_{i-(\ell+1)}\}| = 1$) yields

$$\mathbf{E}[Y_A] \le \sum_{\substack{\mathbf{s}\in\mathbf{S}:\\ F(\mathbf{s})\supseteq A}} p^{|F(\mathbf{s})|-|A|} = \frac{1}{p^\kappa} \sum_{\mathbf{s}\in\mathcal{S}} p^{|\mathcal{F}(\mathbf{s})|} \le \frac{1}{p^\kappa} \mathcal{B}_{\ell+\kappa+2} N (Np)^\ell \frac{\prod_{i=\ell+1}^{\ell+\kappa+1} |S'_i|}{N^{\kappa+1}} \le \mathcal{B}_{2(\ell+1)} (Np)^{\ell-\kappa}. \quad (6.33)$$

Combining (6.32) and (6.33), we have

$$\max_{A\subseteq\binom{V}{2}:|A|\ge 1} \mathbf{E}[Y_A] \le \max_{A\subseteq\binom{V}{2}:|A|\ge 1} \mathcal{B}_{2(\ell+1)} (Np)^{\ell-|A|} = \mathcal{B}_{2(\ell+1)} (Np)^{\ell-1},$$

$$\max_{A\subseteq\binom{V}{2}} \mathbf{E}[Y_A] = \max \left\{ \max_{A\subseteq\binom{V}{2}:|A|\ge 1} \mathbf{E}[Y_A], \mathbf{E}[d_\ell(V)] \right\} \le \mathcal{B}_{2(\ell+1)} N (Np)^\ell.$$

Thus from (6.31) with $\lambda = (2(\ell-1) + C_7/2) \log N$ and $C_8 = \sqrt{\ell!}\mathcal{B}_{2(\ell+1)}(16(\ell - 1 + C_7/2))^\ell$, we obtain

$$\mathbf{Pr}\left[|Y - \mathbf{E}[Y]| \ge C_8 \sqrt{N} (\log N)^\ell (Np)^{\ell-1/2} \right] \le 2\mathrm{e}^2 / N^{C_7}. \quad (6.34)$$

Combining (6.30), (6.24) and (6.34), the following holds with probability at least $1 - 2e^2/N^{C_7} - 1/N^{C_3}$:

$$\forall S_0, S_1, \dots, S_\ell :$$
$$W(S_0; S_1, \dots, S_\ell) \le \mathbf{E}[W(S_0; S_1, \dots, S_\ell)] + C_9 \sqrt{N} (\log N)^\ell (Np)^{\ell-1/2} + (\ell+1) C_4 N (Np)^{\ell-1/2}.$$

Thus we obtain the claim (6.29) and combining the claims (6.24) and (6.29) complete the proof of Lemma 6.5.3. □

6.5.3 Expectation evaluation

Lemma 6.5.4. *Suppose the same setting of Lemma 6.5.1. Then for any vertex subsets $S_0, S_1, \dots, S_\ell$ and for any $i_* \in \{0\} \cup [\ell]$, there is a positive constant C such that*

$$\left| \mathbf{E}[W(S_0; S_1, \dots, S_\ell)] - \hat{W}(S_0; S_1, \dots, S_\ell) \right| \le C |S_{i_*}| (Np)^{\ell-1}.$$

Proof of Lemma 6.5.4. We show

$$\sum_{s\in S_0} \prod_{i\in[\ell]} \mathbf{E}[\deg_{S_i}(s)] \le \mathbf{E}[W(S_0; S_1, \dots, S_\ell)] \le \sum_{s\in S_0} \prod_{i\in[\ell]} \mathbf{E}[\deg_{S_i}(s)] + \mathcal{B}_{\ell+1} |S_{i_*}| (Np)^{\ell-1}$$

for any $i_* \in \{0\} \cup [\ell]$. The first inequality follows directly from the FKG inequality (Proposition 5.4.10) since $\deg_{S_i}(s)$ is a monotone increase function on $(I_e)_{e\in\binom{V}{2}}$ for every i. Now we show the second inequality. We write each element $\mathbf{s} \in \mathbf{S}$ as $\mathbf{s} = (s_0, s_1, \dots, s_\ell)$. Then we have

$$\mathbf{E}[W(S_0; S_1, \dots, S_\ell)] = \sum_{\substack{\mathbf{s}\in\mathbf{S}:\\ |F(\mathbf{s})|=\ell}} \mathbf{E}\left[\prod_{i\in[\ell]} I_{\{s_0,s_i\}} \right] + \sum_{\substack{\mathbf{s}\in\mathbf{S}:\\ |F(\mathbf{s})|\le\ell-1}} \mathbf{E}\left[\prod_{i\in[\ell]} I_{\{s_0,s_i\}} \right]$$

since $\mathbf{E}[W(S_0;S_1,\dots,S_\ell)] = \sum_{\mathbf{s}\in\mathbf{S}} \mathbf{E}\left[\prod_{i\in[\ell]} I_{\{s_0,s_i\}}\right]$. For the first term, since $s_i \neq s_j$ for any $i,j\in[\ell]$ $(i\neq j)$ if $|F(\mathbf{s})|=\ell$, we obtain

$$\sum_{\substack{\mathbf{s}\in\mathbf{S}:\\ |F(\mathbf{s})|=\ell}} \mathbf{E}\left[\prod_{i\in[\ell]} I_{\{s_0,s_i\}}\right] = \sum_{\substack{\mathbf{s}\in\mathbf{S}:\\ |F(\mathbf{s})|=\ell}} \prod_{i\in[\ell]} \mathbf{E}\left[I_{\{s_0,s_i\}}\right] \le \sum_{\mathbf{s}\in\mathbf{S}} \prod_{i\in[\ell]} \mathbf{E}\left[I_{\{s_0,s_i\}}\right] = \sum_{s\in S_0} \prod_{i\in[\ell]} \mathbf{E}[\deg_{S_i}(s)].$$

For the second term, from Lemma 6.5.2,

$$\sum_{\substack{\mathbf{s}\in\mathbf{S}:\\ |F(\mathbf{s})|\le\ell-1}} \mathbf{E}\left[\prod_{e\in F(\mathbf{s})} I_e\right] \le \sum_{\substack{\mathbf{s}\in\mathbf{S}:\\ |F(\mathbf{s})|\le\ell-1}} p^{|F(\mathbf{s})|} \le \mathcal{B}_{\ell+1}|S_{i_*}|(Np)^{\ell-1}.$$

Note that $G(\mathbf{s}) = (U(\mathbf{s}), F(\mathbf{s}))$ is a connected graph for any $\mathbf{s}\in\mathbf{S}$. □

6.5.4 Proof of the key result

To complete the proof of Lemma 6.5.1, we show Lemma 6.5.2 in this section.

Proof of Lemma 6.5.2. It is easy to see that

$$\forall \mathbf{s}\in\mathcal{S} : |U(\mathbf{s})| - 1 \le |\mathcal{F}(\mathbf{s})| \le k$$

since $G(\mathbf{s}) = \big(U(\mathbf{s}), \mathcal{F}(\mathbf{s})\big)$ is a connected graph from the assumption. Hence we have

$$\sum_{\mathbf{s}\in\mathcal{S}} p^{|\mathcal{F}(\mathbf{s})|} \le \sum_{\mathbf{s}\in\mathcal{S}} p^{|U(\mathbf{s})|-1} = \sum_{\mathbf{s}\in\mathcal{S}:|U(\mathbf{s})|\le k+1} p^{|U(\mathbf{s})|-1} \le \sum_{\mathbf{s}\in\prod_{i=0}^{l} S_i:|U(\mathbf{s})|\le k+1} p^{|U(\mathbf{s})|-1}. \tag{6.35}$$

To estimate above, we introduce the following notations. For any $(l+1)$-dimensional vector $\mathbf{s} = (s_0,s_1,\dots,s_l)\in S_0\times S_1\times\cdots\times S_l$, let

$$R(\mathbf{s}) := \big\{\{j\in\{0\}\cup[l] : s_j = s_i\} : i\in\{0\}\cup[l]\big\}.$$

For example, $R\big((a,b,a,c,d,b,f,a)\big) = \big\{\{0,2,7\},\{1,5\},\{3\},\{4\},\{6\}\big\}$. Note that $R(\mathbf{s})$ is a partition of $\{0\}\cup[l]$. From the definition, we have $|R(\mathbf{s})| = |U(\mathbf{s})|$. For example, $|U\big((a,b,a,c,d,b,f,a)\big)| = |\{a,b,c,d,f\}| = 5 = |R\big((a,b,a,c,d,b,f,a)\big)|$. Let $\mathcal{R}_l$ be the family of all partitions of $\{0\}\cup[l]$. For example,

$$\mathcal{R}_2 = \Big\{\big\{\{0\},\{1\},\{2\}\big\},\big\{\{0\},\{1,2\}\big\},\big\{\{1\},\{0,2\}\big\},\big\{\{2\},\{0,1\}\big\},\big\{\{0,1,2\}\big\}\Big\}.$$

Note that $|\mathcal{R}_l| = \mathcal{B}_{l+1}$. Then we have

$$\sum_{\substack{\mathbf{s}\in\prod_{i=0}^{l} S_i:\\ |U(\mathbf{s})|\le k+1}} p^{|U(\mathbf{s})|} = \sum_{\substack{R\in\mathcal{R}_l:\\ |R|\le k+1}} \sum_{\substack{\mathbf{s}\in\prod_{i=0}^{l} S_i:\\ R(\mathbf{s})=R}} p^{|U(\mathbf{s})|} = \sum_{\substack{R\in\mathcal{R}_l:\\ |R|\le k+1}} p^{|R|} \left|\left\{\mathbf{s}\in\prod_{i=0}^{l} S_i : R(\mathbf{s}) = R\right\}\right|. \tag{6.36}$$

From the definition of $R(\mathbf{s})$, for any $r\in R(\mathbf{s})$, $s_i = s_j$ for any $i,j\in r$. Thus

$$\left|\left\{\mathbf{s}\in\prod_{i=0}^{l} S_i : R(\mathbf{s}) = R\right\}\right| = \sum_{\substack{\mathbf{s}\in\prod_{i=0}^{l} S_i:\\ R(\mathbf{s})=R}} 1 \le \sum_{s_0\in S_0}\sum_{s_1\in S_1}\cdots\sum_{s_l\in S_l}\prod_{r\in R}\prod_{i,j\in r}\mathbb{1}_{s_i=s_j} \le \prod_{r\in R}\left|\bigcap_{i\in r} S_i\right|. \tag{6.37}$$

For example, consider four vertex subsets S_0,S_1,S_2,S_3, let $R = \big\{\{0,1\},\{2\},\{3\}\big\}\in\mathcal{R}_3$ and let $l = \{i_*\}\subseteq\{0\}\cup[3]$ where $i_*\in\{0\}\cup[3]$. Then (6.37) means that

$$\begin{aligned}
&\left|\left\{\mathbf{s}\in\prod_{i=0}^{3} S_i : R(\mathbf{s}) = R\right\}\right| \\
&= |\{(s_0,s_1,s_2,s_3)\in S_0\times S_1\times S_2\times S_3 : s_0 = s_1, s_0\neq s_2, s_0\neq s_3, s_2\neq s_3\}| \\
&\le \sum_{s_0\in S_0}\sum_{s_1\in S_1}\sum_{s_2\in S_2}\sum_{s_3\in S_3}\mathbb{1}_{s_0=s_1} \le |S_0\cap S_1||S_2||S_3| = \prod_{r\in\{\{0,1\},\{2\},\{3\}\}}\left|\bigcap_{i\in r} S_i\right|.
\end{aligned}$$

For an index $i \in \{0\} \cup [l]$, let r_i be the element of R such that $r_i \ni i$. Now let us consider the set L described in the statement (of Lemma 6.5.2). First we assume that there are $i, j \in L$ with $i \neq j$ such that both i and j in the same $r_* = r_i = r_j \in R$. In this case, since $S_i \cap S_j = \emptyset$ from the definition of L, we have

$$\prod_{r\in R}\left|\bigcap_{i\in r} S_i\right| = \left|\bigcap_{i\in r_*} S_i\right| \prod_{r\in R\setminus r_*}\left|\bigcap_{i\in r} S_i\right| = 0. \tag{6.38}$$

Now we assume that $r_i \neq r_j$ for any $i, j \in L$. Then since $|\{r_i : i \in L\}| = |L|$ and $R = \{r_i : i \in L\} \cup R \setminus \{r_i : i \in L\}$, we have

$$\prod_{r\in R}\left|\bigcap_{i\in r} S_i\right| = \prod_{i\in L}\left|\bigcap_{j\in r_i} S_j\right| \prod_{r\in R\setminus\{r_i : i\in L\}}\left|\bigcap_{j\in r} S_j\right| \leq \left(\prod_{i\in L}|S_i|\right) N^{|R|-|L|}. \tag{6.39}$$

Finally, by combining (6.35) to (6.39), we obtain

$$\begin{aligned}
\sum_{\mathbf{s}\in\mathcal{S}} p^{|\mathcal{F}(\mathbf{s})|} &\leq \frac{1}{p} \sum_{\substack{R\in\mathcal{R}_l:\\ |R|\leq k+1}} p^{|R|} \left|\left\{\mathbf{s} \in \prod_{i=0}^{l} S_i : R(\mathbf{s}) = R\right\}\right| \\
&\leq \frac{1}{p} \sum_{\substack{R\in\mathcal{R}_l:\\ |R|\leq k+1}} p^{|R|} N^{|R|} \frac{\prod_{i\in L}|S_i|}{N^{|L|}} \leq \frac{1}{p}\left(\frac{\prod_{i\in L}|S_i|}{N^{|L|}}\right)(Np)^{k+1} \sum_{\substack{R\in\mathcal{R}_l:\\ |R|\leq k+1}} 1 \\
&\leq |\mathcal{R}_l|\left(\frac{\prod_{i\in l}|S_i|}{N^{|L|}}\right) N(Np)^k = \mathcal{B}_{l+1}\left(\frac{\prod_{i\in L}|S_i|}{N^{|L|}}\right) N(Np)^k.
\end{aligned}$$

Note that the third inequality follows since $Np \geq 1$ from the assumption. □

Proof of Lemma 6.5.1. Combining Lemmas 6.5.3 and 6.5.4, we obtain the proof. □

6.5.5 Concentration of sum of degree powers for small $|A|$

We prove that $G(n, p, q)$ satisfies the property (P3) of Theorem 6.2.2. We begin with showing the following two lemmas.

Lemma 6.5.5. *Suppose that $0 \leq q \leq p = \omega(\log n/n)$. Then there are two positive constants C_1, C_2 such that $G(2n, p, q)$ satisfies the following with probability $1 - \mathrm{O}(n^{-C_1})$:*

$$\forall v \in V : |\deg(v) - n(p+q)| \leq C_2\sqrt{np\log n}.$$

Proof. Applying the Chernoff bound (Lemma 5.4.2),

$$\begin{aligned}
&\mathbf{Pr}\left[\exists v \in V : |\deg(v) - n(p+q)| > t\right] \\
&\leq \sum_{v\in V}\left(\exp\left(-\frac{t^2}{3\,\mathbf{E}[\deg(v)]}\right) + \exp\left(-\frac{t}{3}\right) + \exp\left(-\frac{t^2}{2\,\mathbf{E}[\deg(v)]}\right)\right) \\
&\leq n\left(2\exp\left(-\frac{t^2}{6np}\right) + \exp\left(-\frac{t}{3}\right)\right) \\
&\leq 2\exp\left(\log n - \frac{t^2}{6np}\right) + \exp\left(\log n - \frac{t}{3}\right).
\end{aligned}$$

Note that $\mathbf{E}[\deg(v)] = (n-1)p + nq$ for any $v \in V$ and $\mathbf{E}[\deg(v)] \leq n(p+q) \leq 2np$. Thus we obtain the claim letting $t = C_2\sqrt{np\log n}$ since $t = C_2\sqrt{np\log n} \geq C\log n$ for some constant C.

□

Lemma 6.5.6. *Suppose that $0 \leq q \leq p = \omega(\log n/n)$. Let $\mathcal{S}(A) = \{S \cap U : S \in \{A, V\setminus A, V\}, U \in \{V_1, V_2, V\}\}$ for $A \subseteq V$. For any constant ℓ, there are two positive constants C_1, C_2 such that $G(2n, p, q)$ satisfies the following with probability $1 - \mathrm{O}(n^{-C_1})$:*

$$\begin{aligned}
&\forall A \subseteq V, \forall S_0, \ldots, S_{\ell-1} \in \mathcal{S}(A) : \\
&\left|W(S_0; S_1, \ldots, S_{\ell-1}, A) - \hat{W}(S_0; S_1, \ldots, S_{\ell-1}, A)\right| \leq C_2|A|\sqrt{\log n}(np)^{\ell-1/2}.
\end{aligned}$$

Proof.
Lower bound. First we claim the following: There are two positive constants C_3, C_4 such that the following holds with probability $1 - n^{-C_3}$:

$$\begin{aligned}&\forall A \subseteq V, \forall S_0, \dots, S_{\ell-1} \in \mathcal{S}(A) : \\ &W(S_0; S_1, \dots, S_{\ell-1}, A) \geq \hat{W}(S_0; S_1, \dots, S_{\ell-1}, A) - C_4|A|\sqrt{\log n}(np)^{\ell-1/2}. \end{aligned} \tag{6.40}$$

From Janson's inequality (Proposition 5.4.5) and (6.28) with a constant C_5 and $C_6 = \sqrt{2(C_5+1)\mathcal{B}_{2(\ell+1)}}$, we have

$$\begin{aligned}&\mathbf{Pr}\left[\exists_{S_0,\dots,S_{\ell-1}\in\mathcal{S}(A)}^{\exists A\subseteq V,} : W(S_0; S_1, \dots, S_{\ell-1}, A) \leq \mathbf{E}[W(S_0; S_1, \dots, S_{\ell-1}, A)] - C_6|A|\sqrt{\log N}(Np)^{\ell-1/2}\right] \\ &\leq \binom{N}{|A|}|\mathcal{S}(A)|^\ell \exp\left(-\frac{2(C_5+1)\mathcal{B}_{2(\ell+1)}|A|(\log N)(Np)^{2\ell-1}}{2\mathcal{B}_{2(\ell+1)}|A|(Np)^{2\ell-1}}\right) \\ &\leq 9^\ell \exp\left(|A|\log N - (C_5+1)|A|\log N\right) \leq 9^\ell / N^{C_5}. \end{aligned} \tag{6.41}$$

Thus combining (6.41) and Lemma 6.5.4 yields the claim (6.40).
Upper bound. Now we show the following claim: There are two positive constants C_7, C_8 such that the following holds with probability $1 - n^{-C_7}$:

$$\begin{aligned}&\forall A \subseteq V, \forall S_0, \dots, S_{\ell-1} \in \mathcal{S}(A) : \\ &W(S_0; S_1, \dots, S_{\ell-1}, A) \leq \hat{W}(S_0; S_1, \dots, S_{\ell-1}, A) + C_8|A|\sqrt{\log n}(np)^{\ell-1/2}. \end{aligned} \tag{6.42}$$

From the same discussion of (6.30),

$$W(S_0; S_1, \dots, S_{\ell-1}, A) = W(\overbrace{V; V, \dots, V}^{\ell}, A) - \sum_{i=0}^{\ell-1} W(\overbrace{V; V, \dots, V}^{i}, V \setminus S_i, S_{i+1}, \dots, S_{\ell-1}, A) \tag{6.43}$$

since $W_\ell - W_0 = \sum_{i=0}^{\ell-1}(W_{i+1} - W_i)$. Thus we consider an upper bound on $W(S_0; S_1, \dots, S_{\ell-1}, A)$. Let $d_{\max} := \max_{v\in V} \deg(v)$. Since $\sum_{v\in V} \deg_A(v) = \sum_{a\in A} \deg(v)$, we have

$$W(\underbrace{V; V, \dots, V}_{\ell}, A) = \sum_{v\in V} \deg(v)^{\ell-1} \deg_A(v) \leq d_{\max}^{\ell-1} \sum_{v\in V} \deg_A(v) \leq d_{\max}^{\ell-1} \sum_{a\in A} \deg(a) \leq d_{\max}^\ell |A|.$$

From Lemma 6.5.5, it holds with high probability that

$$d_{\max}^\ell = \left(n(p+q) + \mathrm{O}(\sqrt{np\log n})\right)^\ell = \left(n(p+q)\right)^\ell \left(1 + \mathrm{O}\left(\sqrt{\frac{\log n}{np}}\right)\right).$$

The second equality holds since $(\log n)/(np) = \mathrm{o}(1)$ and ℓ is a constant. Hence we have

$$\begin{aligned}W(\overbrace{V; V, \dots, V}^{\ell}, A) &\leq d_{\max}^\ell |A| = |A|\left(n(p+q)\right)^\ell \left(1 + \mathrm{O}\left(\sqrt{\frac{\log n}{np}}\right)\right) \\ &\leq \hat{W}(\overbrace{V; V, \dots, V}^{\ell}, A) + \mathrm{O}(|A|\sqrt{\log n}(np)^{\ell-1/2}). \end{aligned} \tag{6.44}$$

Note that $\hat{W}(\overbrace{V; V, \dots, V}^{\ell}, A) = |A|\left((n-1)p + nq\right)^\ell$. Thus we obtain the claim (6.42) by applying (6.40) and (6.44) to (6.43). Combining the claims (6.40) and (6.42) complete the proof of Lemma 6.5.6. □

Proof of (P3) of Theorem 6.2.2. Let $d_{\min} := \min_{v\in V} \deg(v)$. Then for any $j \in [\ell]$,

$$\sum_{s\in S\cap V_i} \left(\frac{\deg_A(s)}{\deg(s)}\right)^j \leq d_{\min}^{-j} \sum_{s\in S\cap V_i} \deg_A(s)^j = d_{\min}^{-j} W(S \cap V_i; \overbrace{A, \dots, A}^{j}).$$

From Lemma 6.5.5, it holds with high probability that

$$d_{\min}^{-j} = \left(n(p+q) - \mathrm{O}(\sqrt{np\log n})\right)^{-j} = \left(1 + \mathrm{O}\left(\sqrt{\frac{\log n}{np}}\right)\right)\left(n(p+q)\right)^{-j}.$$

The second equality holds since $(\log n)/(np) = \mathrm{o}(1)$ and $j \in [\ell]$ is a constant. Thus from Lemma 6.5.6, we have

$$\begin{aligned}\sum_{s\in S\cap V_i}\left(\frac{\deg_A(s)}{\deg(s)}\right)^j &= \left(1+\mathrm{O}\left(\sqrt{\frac{\log n}{np}}\right)\right)\left(\frac{\hat{W}(S\cap V_i;\overbrace{A,\dots,A}^{j})}{(n(p+q))^j} + \mathrm{O}\left(|A|\sqrt{\frac{\log n}{np}}\right)\right)\\ &\le |S\cap V_i|\left(\frac{|A_i|p+|A_{3-i}|q}{n(p+q)}\right)^j + \mathrm{O}\left(|A|\sqrt{\frac{\log n}{np}}\right).\end{aligned}$$

Note that $\frac{|S\cap V_i|(|A_i|p+|A_{3-i}|q)^j}{(n(p+q))^j} = \frac{|S\cap V_i|(|A_i|p+|A_{3-i}|q)}{n(p+q)}\left(\frac{|A_i|p+|A_{3-i}|q}{n(p+q)}\right)^{j-1} \le \frac{|A|p}{(p+q)} \le |A|$. Thus we obtain the claim. □

6.6 Proof of local dynamics around fixed points

In this section, we consider a polynomial voting process with respect to f on an f- and $\overline{f}$-good graph G for a partition (V_1, V_2) and parameters p, q (recall that $\overline{f}(x) = 1 - f(1-x)$). Throughout this section, the randomness is the generation of A' at each step. Let $H = (H_1, H_2) : [0,1]^2 \to [0,1]^2$ be the induced dynamical system.

6.6.1 Dynamics around sink points

Let $B_2(\mathbf{x}, r)$ denote the open ball of radius r (with respect to the ℓ^2-norm) centered at $\mathbf{x}$. Let $\mathbf{a}^*$ be the sink point. From the property of singular value (Proposition 6.3.4) and the Taylor expansion, there are constants $r, K > 0$ such that, for any $\mathbf{x} \in B(\mathbf{x}^*, r)$, it holds that

$$\|H(\mathbf{x}) - \mathbf{x}^*\|_2 = \|H(\mathbf{x}) - H(\mathbf{x}^*)\|_2 \le \sigma_{\max}\|\mathbf{x} - \mathbf{x}^*\|_2 + O_{\mathbf{x}\to\mathbf{x}^*}(\|\mathbf{x} - \mathbf{x}^*\|_2^2) < (1-K)r.$$

Let $\epsilon > 0$ be such that $\epsilon < r$ and $\epsilon = \omega(1/\sqrt{np})$. From the Hoeffding inequality (Proposition 5.4.3), for any $A \subseteq V$ of $\boldsymbol{\alpha} \in B(\mathbf{a}^*, \epsilon)$, we have

$$\begin{aligned}\mathbf{Pr}[\|\boldsymbol{\alpha}' - \mathbf{a}^*\|_2 \ge \epsilon] &\le \mathbf{Pr}[\|\boldsymbol{\alpha}' - \mathbf{E}[\boldsymbol{\alpha}']\|_2 + \|\mathbf{E}[\boldsymbol{\alpha}'] - H(\boldsymbol{\alpha})\|_2 + \|H(\boldsymbol{\alpha}) - \mathbf{a}^*\|_2 \ge \epsilon]\\ &\le \mathbf{Pr}[\|\boldsymbol{\alpha}' - \mathbf{E}[\boldsymbol{\alpha}']\|_2 \ge K\epsilon - \mathrm{O}(1/\sqrt{np})]\\ &\le \mathbf{Pr}\left[\|\boldsymbol{\alpha}' - \mathbf{E}[\boldsymbol{\alpha}']\|_\infty \ge \frac{K\epsilon}{\sqrt{2}} - \mathrm{O}\left(\frac{1}{\sqrt{np}}\right)\right]\\ &\le \exp(-\Omega(K\epsilon^2 n)).\end{aligned}$$

Fix an initial set $A_0 \subseteq V$ such that $\boldsymbol{\alpha}^{(0)} \in B(\mathbf{a}^*, \epsilon)$. For any $T \ge 0$, from the union bound over the time $t = 1, \dots, T$, we obtain

$$\mathbf{Pr}\left[\exists t \in [T] : \boldsymbol{\alpha}^{(t)} \notin B(\mathbf{a}^*, \epsilon)\right] \le T\exp(-\Omega(\epsilon^2 n)).$$

Suppose that $\epsilon = \omega(\max\{\sqrt{\log n/n}, \sqrt{1/np}\})$. If we set $T = \exp(D\epsilon^2 n)$ for some constant $D > 0$, the stopping time $\tau = \min\{t : \boldsymbol{\alpha}^{(t)} \notin B(\mathbf{a}^*, \epsilon)\}$ satisfies $\tau \ge \exp(\Omega(\epsilon^2 n))$ w.h.p. Note that $T_{\mathrm{cons}}(A_0) \ge \tau$ and we are done.

6.6.2 Dynamics around consensus points

This subsection is devoted to the proof of Proposition 6.2.9. We begin with the following result which is of independent interest.

Proposition 6.6.1. *Suppose that there are absolute constants $C, \delta > 0$ and a function $\epsilon = \epsilon(n) = \mathrm{o}(1)$ such that*

$$\mathbf{E}[|A'|] \le \frac{C|A|^2}{n} + \epsilon|A|$$

holds for all $A \subseteq V$ satisfying $|A| \le \delta n$. Then, there are positive constants δ', C', C'' such that

$$\mathbf{Pr}\left[T_{\mathrm{cons}}(A) \le C'\left(\log\log n + \frac{\log n}{\log \epsilon^{-1}}\right)\right] \ge 1 - n^{-C''}$$

holds for any $A \subseteq V$ satisfying $|A| \le \delta' n$.

Proof of Proposition 6.6.1. Note that we may assume $\epsilon(n) = \Omega(\sqrt{\log n/n})$: If $\epsilon = \mathrm{o}(\sqrt{\log n/n})$, we have $\log n/\log \epsilon^{-1} = \mathrm{O}(1)$ and we will obtain the claim by applying Proposition 6.6.1 with letting $\epsilon = \sqrt{\log n/n}$.

Take a positive constant δ' such that

$$10\left(\frac{CM^2}{n} + \epsilon M\right) \le M, \tag{6.45}$$

$$\delta' \le \min\left\{\delta, \frac{1}{16C}\right\} \tag{6.46}$$

hold for any $0 \le M \le \delta' n$. We can take such constant $\delta' > 0$ since $\epsilon = \mathrm{o}(1)$ and thus the inequality (6.45) holds if the ratio $\frac{M}{n}$ is sufficiently small.

Consider $(A_t)_{t\in\mathbb{Z}_{\ge 0}}$ given by the polynomial voting process such that $|A_0| \le \delta n$. To exploit the assumption of the expectation, we first claim that $|A_t| \le \delta' n \le \delta n$ holds w.h.p. for all $t = 0, \dots, n^{\mathrm{o}(1)}$. Let $\mathcal{B}^{(t)}$ be the event that $|A_i| \le \delta n$ for all $i = 0, \dots, t$. Note that $\mathcal{B}^{(0)}$ holds. Consider $\mathbf{Pr}[\mathcal{B}^{(t+1)}|\mathcal{B}^{(t)}]$. If $\mathbf{E}[|A'|] \ge \log n$, from the Chernoff bound ((i) of Proposition 2.5.5), for any $A \subseteq V$ such that $|A| \le \delta' n$, we obtain

$$\mathbf{Pr}\left[|A'| \ge 10\left(\frac{C|A|^2}{n} + \epsilon|A|\right)\right] \le \mathbf{Pr}\left[|A'| \ge 10\,\mathbf{E}[|A'|]\right] \le \exp\left(-\frac{10}{3}\log n\right) \le n^{-3}.$$

Then, for $|A| \le \delta' n$, it holds with probability $1 - \mathrm{O}(n^{-3})$ that

$$|A'| \le 10\left(\frac{C|A|^2}{n} + \epsilon|A|\right) \le |A| \le \delta' n. \tag{6.47}$$

Here, we used (6.45) with letting $M = |A|$. If $\mathbf{E}[|A'|] \le \log n$, from the Chernoff bound (Proposition 5.4.1), we have

$$|A'| \le 6\log n = \mathrm{o}(\delta' n) \tag{6.48}$$

with probability at least $1 - \mathrm{O}(n^{-3})$. From (6.47) and (6.48), we obtain $\mathbf{Pr}\left[\mathcal{B}^{(t+1)} \,\middle|\, \mathcal{B}^{(t)}\right] \ge 1 - \mathrm{O}(n^{-3})$ for each t and thus $\mathcal{B}^{(t)}$ holds for $t = n^{0.01}$ with probability $1 - \mathrm{O}(n^{-2.99})$.

Now we look at $|A_t|$. Note that, if $|A| \le \delta' n$, then

$$\mathbf{E}[|A'|] \le \begin{cases} \frac{2C|A|^2}{n} & \text{if } \frac{\epsilon}{C} n \le |A| \le \delta' n, \\ 2\epsilon|A| & \text{if } 0 \le |A| \le \frac{\epsilon}{C}. \end{cases}$$

We consider the following two cases.

Case I: $\frac{\epsilon}{C} n \le |A|^{(t)} \le \delta' n$. From the Chernoff bound (Proposition 5.4.1), we have

$$\mathbf{Pr}\left[|A_{t+1}| \ge \frac{12C|A_t|^2}{n} \,\middle|\, \mathcal{B}^{(t)}\right] \le 2^{-\Omega(\log n)}.$$

In the last inequality, we used $|A_t| \ge \frac{\epsilon}{C} n = \Omega(\sqrt{n\log n})$. Hence, conditioned on $\mathcal{B}^{(t)}$ and $|A|^{(i)} \ge \frac{\epsilon}{C} n$ $(i = 0, \dots, t)$, it holds w.h.p. that

$$|A_t| \le \frac{12C(|A_{t-1}|)^2}{n} \le \frac{n}{12C}\left(\frac{12C|A_0|}{n}\right)^{2^t} \le \frac{0.75^{2^t} n}{12C}.$$

Here, we used (6.46). Therefore, for some $\tau_1 = \mathrm{O}(\log\log n)$, $|A_{\tau_1}| \le \frac{\epsilon}{C}$ holds w.h.p.

Case II: $0 \le |A_t| \le \frac{\epsilon}{C}n$. Conditioned on $|A_0| \le \frac{\epsilon}{C}n$, we claim that $\mathbf{E}[|A_{\tau_2}|] \le n^{-\Omega(1)}$ for some $\tau_2 = \mathrm{O}(\log n/\log \epsilon^{-1})$. Note that this completes the proof of Proposition 6.6.1 since $\mathbf{Pr}[A_{\tau_2} \neq \emptyset] \le \mathbf{E}[|A_{\tau_2}|] = n^{-\Omega(1)}$ from the Markov inequality.

To show the claim, we exploit the property that $\mathbf{E}[|A'|] \le 2\epsilon|A|$ if $|A| \le \frac{\epsilon n}{C}$. Before using this, we show that $|A_t| \le \frac{\epsilon n}{C}$ holds for all $t = 1, \dots, n^{\mathrm{o}(1)}$. Conditioned on $|A| \le \frac{\epsilon n}{C}$, we have $\mathbf{E}[|A'|] \le 2\epsilon|A| \le \mathrm{O}(\epsilon^2 n)$ and thus, from the Chernoff bound (Proposition 5.4.1), for any $A \subseteq V$ such that $|A| \le \frac{\epsilon n}{C}$, we obtain $\mathbf{Pr}\left[|A'| \ge \frac{\epsilon n}{C}\right] \le 2^{-\Omega(\epsilon n)} = n^{-\Omega(1)}$. Therefore, it holds w.h.p. that $|A_t| \le \frac{\epsilon}{C}n$ for all $t = 0, \dots, n^{\mathrm{o}(1)}$. Let $\mathcal{C}^{(t)}$ be the event that $|A_i| \le \frac{\epsilon}{C}n$ holds for all $i = 0, \dots, t$. Then, from the tower property of the conditional expectation, we have

$$\begin{aligned}
\mathbf{E}[|A_{\tau_2}| \mid \mathcal{C}^{(\tau_2)}] &\le \mathbf{E}[\mathbf{E}[|A_{\tau_2}| \mid A_{\tau_2-1}, \mathcal{C}^{(\tau_2)}] \mid \mathcal{C}^{(\tau_2)}] \\
&\le \mathbf{E}[2\epsilon|A_{\tau_2-1}| \mid \mathcal{C}^{(\tau_2)}] \\
&\le (2\epsilon)^{\tau_2} \cdot \frac{\epsilon n}{C} \\
&\le n^{-\Omega(1)}
\end{aligned}$$

for some $\tau_2 = \mathrm{O}(\log n/\log \epsilon^{-1})$. This shows the aforementioned claim, which completes the proof of Proposition 6.6.1. □

Proof of Proposition 6.2.9. It suffices to check the condition of Proposition 6.6.1 for $\epsilon = \Theta\left(\sqrt{\frac{\log n}{np}}\right)$. Using (P3) and the Taylor expansion, there is a constant $C = C(H)$ such that

$$\begin{aligned}
\mathbf{E}[|A'_i|] &= nH_i(\alpha_1, \alpha_2) \pm \mathrm{O}\left(|A|\sqrt{\frac{\log n}{np}}\right) \le nC\left((\alpha_1+\alpha_2)^2 + |A|\sqrt{\frac{\log n}{np}}\right) \\
&= C\frac{|A|^2}{n} + C|A|\sqrt{\frac{\log n}{np}}
\end{aligned}$$

holds if $\|\boldsymbol{\alpha}\| \le \delta$ for sufficiently small constant δ. □

6.6.3 Dynamics around source and saddle points

Let $\mathbf{a}^*$ be a fixed point satisfying Assumption 6.2.10. Recall the random variable $\boldsymbol{\beta}$ defined in (6.11). From the definition (6.11), each element β_i of $\boldsymbol{\beta}$ can be rewritten as

$$\beta_i = \sum_{j=1}^{2} u_{ij}\alpha_j - (U\mathbf{a}^*)_i, \tag{6.49}$$

where we let $U = (u_{ij})$. Each element u_{ij} of the matrix U does not depend on n. Hence, the Hoeffding bound (Proposition 5.4.3) implies

$$\mathbf{Pr}[|\beta'_i - \mathbf{E}[\beta'_i]| \ge t] \le \exp\left(-\Omega(t^2 n)\right). \tag{6.50}$$

From (6.3) and the Taylor expansion, it holds w.h.p. for any $A \subseteq V$ that

$$\begin{aligned}
\mathbf{E}[\boldsymbol{\beta}'] &= U(H(\boldsymbol{\alpha}) - \mathbf{a}^*) + \mathrm{O}\left(\frac{1}{\sqrt{np}}\right) \cdot \mathbf{1} \\
&= UJ(\boldsymbol{\alpha} - \mathbf{a}^*) + \left(O_{\boldsymbol{\alpha}\to\mathbf{a}^*}(\|\boldsymbol{\alpha} - \mathbf{a}^*\|_\infty^2) + \mathrm{O}\left(\frac{1}{\sqrt{np}}\right)\right) \cdot \mathbf{1} \\
&= \Lambda\boldsymbol{\beta} + \left(O_{\|\boldsymbol{\beta}\|\to 0}(\|\boldsymbol{\beta}\|_\infty^2) + \mathrm{O}\left(\frac{1}{\sqrt{np}}\right)\right) \cdot \mathbf{1}.
\end{aligned}$$

Hence, the i-th element β_i of $\boldsymbol{\beta} = (\beta_1\ \beta_2)^\top$ satisfies

$$|\mathbf{E}[\beta'_i]| = |\lambda_i\|\beta_i| + O_{\|\boldsymbol{\beta}\|\to 0}(\|\boldsymbol{\beta}\|_\infty^2) + \mathrm{O}\left(\frac{1}{\sqrt{np}}\right). \tag{6.51}$$

It is convenient to consider the behavior of $\boldsymbol{\beta}$ instead of $\boldsymbol{\alpha}$. Note that $\boldsymbol{\alpha} \to \mathbf{a}^*$ implies $\boldsymbol{\beta} \to \mathbf{0}$ and vice versa since the matrix U is nonsingular. By substituting $t = \Theta\left(\sqrt{\frac{\log n}{n}}\right)$ to (6.50), for sufficiently large constant $C > 0$, it holds w.h.p. that

$$||\beta'_i| - |\lambda_i||\beta_i|| \leq C\|\beta\|_\infty^2 + C\sqrt{\frac{\log n}{n}}. \tag{6.52}$$

Proof of Proposition 6.2.12 Suppose that the fixed point $\mathbf{a}^*$ satisfies the condition of Proposition 6.2.12 and Assumption 6.2.10. Let $I_{>1} := \{i \in [2] : |\lambda_i| > 1\}$ and $I_{\leq 1} := [2] \setminus I_{>1}$. Fix a sufficiently large constant $K > 0$ and let ϵ^* be the constant mentioned in Proposition 6.2.12. Define

$$\mathcal{A}_1 = \left\{ A \subseteq V \,:\, \|\boldsymbol{\beta}\|_\infty \leq \epsilon^* \text{ and } |\beta_j| < K\sqrt{\frac{\log n}{n}} \text{ for all } j \in I_{>1} \right\},$$
$$\mathcal{A}_2 = \left\{ A \subseteq V \,:\, \|\boldsymbol{\beta}\|_\infty \leq \epsilon^* \text{ and } |\beta_j| \geq K\sqrt{\frac{\log n}{n}} \text{ for some } j \in I_{>1} \right\},$$
$$\mathcal{A}_3 = \{A \subseteq V \,:\, \|\boldsymbol{\beta}\|_\infty > \epsilon^* \text{ and } |\beta_j| \leq \epsilon^* \text{ for all } j \in I_{\leq 1}\}.$$

We claim that, for each $i = 1, 2$ and any $A_0 \in \mathcal{A}_i$, it holds w.h.p. that $A_\tau \in \mathcal{A}_{i+1}$ for some $\tau = \mathrm{O}(\log n)$. This completes the proof of Proposition 6.2.12.

Case I: $A_0 \in \mathcal{A}_1$**.** Let $f(A) := \lfloor n \max\{|\beta_i| : i \in I_{>1}\} \rfloor$ and $m = K\sqrt{n \log n}$. We use Corollary 6.3.2 to show $A_\tau \in \mathcal{A}_2$ for some $\tau = \mathrm{O}(\log n)$. Here, we use $\mathcal{A}_1$ as $\mathcal{B}$ of Corollary 6.3.2. Note that $A \in \mathcal{A}_1$ implies $f(A) < m$.

From (6.49) and (A2), we have $\mathbf{Var}[|\beta_i|] = \sum_{j \in [2]} u_{ij}^2 \mathbf{Var}[\alpha_j] = \Omega(n^{-1})$ for any $A \in \mathcal{A}_1$. Here, note that, for every $i \in [2]$, there is $j \in [2]$ such that $u_{ij} \neq 0$, since otherwise, it contradicts to the fact that the matrix U is nonsingular. Thus, from Corollary 5.4.8, it holds that, for any constant $h > 0$, there is a positive constant $C_1 < 1$ such that $\mathbf{Pr}[f(A') < h\sqrt{n}] < C_1$ holds for any $A \subseteq V$ with $f(A) \leq m$ This verifies the condition (1′) of Corollary 6.3.2.

Now we check the condition (2′). Let $z \in [2]$ be the least index satisfying $|\beta_z| = \max\{|\beta_i| : i \in [2]\}$. Suppose that $A \in \mathcal{A}_1$ satisfies $f(A) = \lfloor n|\beta_z| \rfloor \geq h\sqrt{n}$ for sufficiently large constant $h > \frac{100C}{\epsilon_1}$ (recall that the constant ϵ_1 is mentioned in (B1)). Then, from (B1), we have

$$|\mathbf{E}[\beta'_z]| \geq (1 + 0.99\epsilon_1)|\beta_z| + 0.01\epsilon_1|\beta_z| - \frac{C}{\sqrt{n}} \geq (1 + 0.99\epsilon_1)|\beta_z|.$$

Thus, from the Hoeffding inequality (Proposition 5.4.3), we obtain

$$\mathbf{Pr}[f(A') < (1 + 0.98\epsilon_1)f(A)] \leq \mathbf{Pr}\left[f(A') < \frac{1 + 0.98\epsilon_1}{1 + 0.99\epsilon_1}\mathbf{E}[f(A')]\right] \leq \exp\left(-\Omega\left(\frac{f(A)^2}{n}\right)\right)$$

holds for every $A \in \mathcal{A}_1$ satisfying $f(A) \geq h\sqrt{n}$. This verifies the condition (2′).

Finally, we check the condition (3′) of Corollary 6.3.2. From (B2), for any $A \in \mathcal{A}_1$, it holds that

$$\mathbf{Pr}[A' \notin \mathcal{A}_1 \text{ and } f(A') < m] \leq \mathbf{Pr}[\exists j \in I_{\leq 1}, |\beta'_j| > \epsilon^*] \leq n^{-\Omega(1)}.$$

Therefore, from Corollary 6.3.2, we have $f(A_\tau) \geq m = K\sqrt{n \log n}$ (i.e. $A_\tau \in \mathcal{A}_2$) holds w.h.p. for some $\tau = \mathrm{O}(\log n)$.

Case II: $A_0 \in \mathcal{A}_2$**.** Suppose that $A_0 \in \mathcal{A}_2$ and let $j \in I_{>1}$ be the index satisfying $|\beta_j| > K\sqrt{\frac{\log n}{n}}$. We remark that K is sufficiently large. From (B1) and (6.52), we have $|\beta'_j| \geq (1 + 0.99\epsilon_1)|\beta_j|$. Thus, for some $\tau = \mathrm{O}(\log n)$, we have $|\beta_j^{(\tau)}| > \epsilon^*$. Moreover, from (B2), we have $|\beta_i^{(\tau)}| \leq \epsilon^*$ for all $i \in I_{\leq 1}$. Therefore, $A_\tau \in \mathcal{A}_3$ holds w.h.p.

Proof of Proposition 6.2.11. Suppose that the fixed point $\mathbf{a}^*$ satisfies the condition of Proposition 6.2.11. Let

$$I_{<1} := \{i \in [2] : |\lambda_i| < 1\},$$
$$I_{>1} := \{i \in [2] : |\lambda_i| > 1\}.$$

Note that $I_{<1} \cup I_{>1} = [2]$. Moreover, there is some constant $\epsilon > 0$ such that

$$||\lambda_i| - 1| > 3\epsilon \tag{6.53}$$

holds for every $i \in [2]$. For $A \subseteq V$, let $z = z(A) \in [2]$ be the least index satisfying $|\beta_z| = \|\boldsymbol{\beta}\|_\infty$. We use four constants: In (6.52) and (6.53), we defined C and ϵ. Let $K := \frac{C}{\epsilon}$ and $\epsilon' := \frac{\epsilon}{C}$. Consider four events

$$\mathcal{B}_1 = \left\{ A \subseteq V \,:\, K\sqrt{\frac{\log n}{n}} < \|\boldsymbol{\beta}\|_\infty \le \epsilon' \text{ and } z(A) \in I_{<1} \right\},$$
$$\mathcal{B}_2 = \left\{ A \subseteq V \,:\, \|\boldsymbol{\beta}\|_\infty \le K\sqrt{\frac{\log n}{n}} \right\},$$
$$\mathcal{B}_3 = \left\{ A \subseteq V \,:\, K\sqrt{\frac{\log n}{n}} < \|\boldsymbol{\beta}\|_\infty \le \epsilon' \text{ and } z(A) \in I_{>1} \right\},$$
$$\mathcal{B}_4 = \left\{ A \subseteq V \,:\, \|\boldsymbol{\beta}\|_\infty > \epsilon' \text{ and } |\beta_j| \le \epsilon' \text{ for all } j \in I_{<1} \right\}.$$

We claim that, if $A_0 \in \mathcal{B}_i$, then $A_\tau \in \mathcal{B}_j$ holds w.h.p. for some $j > i$ and some $\tau = \mathrm{O}(\log n)$. This completes the proof of Proposition 6.2.11.

Case I: $A_0 \in \mathcal{B}_1$. Suppose $A_0 \in \mathcal{B}_1$. We claim that, if $A_t \in \mathcal{B}_1$, then either $A_{t+1} \in \mathcal{B}_3$ or $\|\boldsymbol{\beta}^{(t+1)}\|_\infty \le (1-\epsilon)\|\boldsymbol{\beta}^{(t)}\|_\infty$ holds w.h.p. For any $j \in I_{<1}$, the bound (6.52) yields that

$$\begin{aligned}
|\beta'_j| &\le (1-3\epsilon)\|\boldsymbol{\beta}\|_\infty + C\|\boldsymbol{\beta}\|_\infty^2 + C\sqrt{\frac{\log n}{n}} \\
&\le (1-\epsilon)\|\boldsymbol{\beta}\|_\infty - 2\epsilon\|\beta\|_\infty + C\epsilon'\|\beta\|_\infty + C\sqrt{\frac{\log n}{n}} \\
&= (1-\epsilon)\|\beta\|_\infty - \epsilon\|\beta\|_\infty + C\sqrt{\frac{\log n}{n}} \\
&\le (1-\epsilon)\|\beta\|_\infty
\end{aligned}$$

holds w.h.p. If $A_{t+1} \notin \mathcal{B}_3$, then $\|\boldsymbol{\beta}'\|_\infty = |\beta'_j|$ for some $j \in I_{<1}$; thus, we have $\|\boldsymbol{\beta}'\|_\infty \le (1-\epsilon)\|\boldsymbol{\beta}\|_\infty$ w.h.p. Therefore, for some $\tau = \mathrm{O}(\log n)$, it holds w.h.p. that $A_\tau \in \mathcal{B}_2 \cup \mathcal{B}_3$.

Case II: $A_0 \in \mathcal{B}_2$. Suppose $A_0 \in \mathcal{B}_2$. Our strategy is to apply Corollary 6.3.2. We will prove the following result in the last part of this subsection.

Lemma 6.6.2. *For any fixed $A \in \mathcal{B}_2$, the following hold w.h.p.:*

(i) For every $i \in I_{<1}$, it holds that $|\beta'_i| \le K\sqrt{\frac{\log n}{n}}$, and

(ii) there is a constant $h > 0$ such that $|\mathbf{E}[\beta'_i]| \ge (1+\epsilon)|\beta_i|$ for every $i \in I_{>1}$ satisfying $|\beta_i| \ge \frac{h}{\sqrt{n}}$.

Let $m = K\sqrt{n \log n}$ and define $f(A) := \lfloor n \cdot \max_{i \in I_{>1}} |\beta_i| \rfloor$. Suppose that $f(A_\tau) \ge K\sqrt{n \log n}$ holds w.h.p. for some $\tau = \mathrm{O}(\log n)$. Then, we have $A_\tau \notin \mathcal{B}_1 \cup \mathcal{B}_2$ w.h.p. since $|\beta_i^{(\tau)}| \le K\sqrt{\frac{\log n}{n}}$ holds w.h.p. for any $i \in I_{<1}$. Here, we used (i) of Lemma 6.6.2. To show $f(A_\tau) \ge K\sqrt{n \log n}$, we check the condition (1′) to (3′) of Corollary 6.3.2 and then apply it.

First, we check the condition (1′) of Corollary 6.3.2. We use the same argument described in the Case I in Section 6.6.3. From (6.49), we have $\mathbf{Var}[\beta'_i] \ge \sum_{j=1}^2 u_{ij}^2 \mathbf{Var}[\alpha'_j]$. Moreover, for every $i \in [2]$ there is $j \in [2]$ such that $u_{ij} \neq 0$, since otherwise, it contradicts to the fact that U is nonsingular. From (A2), we have $\mathbf{Var}[\beta'_i] = \Omega(n^{-1})$; thus, from Corollary 5.4.8, it holds that, for any constant $h > 0$, there is a positive constant $C_1 < 1$ such that $\mathbf{Pr}[f(A') \ge h\sqrt{n}] < C_1$ holds for any $A \subseteq V$ with $f(A) < m$.

We check the condition (2′) of Corollary 6.3.2. For every $i \in I_{>1}$, from Lemma 6.6.2, we obtain

$$|\,\mathbf{E}[\beta_i^{(t+1)} \mid A_t \in \mathcal{B}_2]| \geq (1+\epsilon)|\beta_i^{(t)}|. \tag{6.54}$$

In look at (6.49), from the Hoeffding inequality (Proposition 5.4.3), it holds for any set $A_t \in \mathcal{B}_2$, any index $i \in I_{>1}$ and any constant $\epsilon' > 0$ that

$$\mathbf{Pr}[|\beta_i'| \leq (1-\epsilon')|\,\mathbf{E}[\beta_i']|] \leq \exp\left(-\Omega\left(\epsilon'^2\,\mathbf{E}[\beta_i']^2 n\right)\right) \leq \exp\left(-\Omega\left(\epsilon'^2 \frac{f(A)^2}{n}\right)\right). \tag{6.55}$$

From (6.54) and (6.55), by letting $\epsilon' = \frac{\epsilon}{2(1+\epsilon)}$, we obtain

$$\begin{aligned}\mathbf{Pr}\left[|\beta_i^{(t+1)}| \leq \left(1+\frac{\epsilon}{2}\right)\cdot|\beta_i^{(t)}| \,\middle|\, A_t \in \mathcal{B}_2\right] &\leq \mathbf{Pr}\left[|\beta_i'| \leq (1-\epsilon')\cdot|\,\mathbf{E}[\beta_i' \mid A_t \in \mathcal{B}_2]|\right] \\ &\leq \exp\left(-\Omega\left(\frac{f(A)^2}{n}\right)\right).\end{aligned}$$

In other words, for any $A \in \mathcal{B}_2$ satisfying $f(A) \geq h\sqrt{n}$ for some constant $h > 0$, we have

$$\mathbf{Pr}\left[f(A_{t+1}) < \left(1+\frac{\epsilon}{2}\right) f(A_t) \,\middle|\, A_t = A\right] \leq \exp\left(-\Omega\left(\frac{f(A)^2}{n}\right)\right).$$

Finally, we check the condition (3′) of Corollary 6.3.2. From Lemma 6.6.2, we have

$$\begin{aligned}\mathbf{Pr}[A_{t+1} \notin \mathcal{B}_2 \wedge f(A_{t+1}) \leq m \mid A_t \in \mathcal{B}_2] &\leq \mathbf{Pr}\left[\exists j \in I_{<1} : |\beta_j'| > K\sqrt{\frac{\log n}{n}} \,\middle|\, A \in \mathcal{B}_2\right] \\ &\leq n^{-\Omega(1)}.\end{aligned}$$

Now, from Corollary 6.3.2, there is some $\tau = \mathrm{O}(\log n)$ such that $f(A_\tau) \geq K\sqrt{\frac{\log n}{n}}$ and $|\beta_j^{(\tau)}| \leq K\sqrt{\frac{\log n}{n}}$ hold w.h.p. for every $j \in I_{<1}$. Consequently, $A_\tau \in \mathcal{B}_3 \cup \mathcal{B}_4$ holds w.h.p.

Case III: $A_0 \in \mathcal{B}_3$**.** Suppose that $A_0 \in \mathcal{B}_3$. From (6.52), it holds w.h.p. that

$$\begin{aligned}|\beta_z'| &\geq |\lambda_z||\beta_z| - C\|\boldsymbol{\beta}\|_\infty - C\sqrt{\frac{\log n}{n}} \\ &\geq (1+\epsilon)|\beta_z| + (\epsilon|\beta_z| - C|\beta_z|^2) + \left(\epsilon|\beta_z| - C\sqrt{\frac{\log n}{n}}\right) \\ &\geq (1+\epsilon)|\beta_z|.\end{aligned}$$

Moreover, for any $j \in I_{<1}$, it holds w.h.p. that

$$|\beta_j'| \leq (1-3\epsilon)\|\boldsymbol{\beta}\|_\infty + C\|\boldsymbol{\beta}\|_\infty^2 + C\sqrt{\frac{\log n}{n}} \leq (1-\epsilon)|\beta_z|$$

These imply that $A_{t+1} \notin \mathcal{B}_1 \cup \mathcal{B}_2$ holds w.h.p. whenever $A_t \in \mathcal{B}_3$. Let τ be the stopping time given by $\tau := \min\{t : A_t \notin \mathcal{B}_3\}$. Then, $\|\boldsymbol{\beta}^{(t+1)}\|_\infty \geq (1+\epsilon)\|\boldsymbol{\beta}^{(t)}\|_\infty$ holds w.h.p. for all $t < \tau$. Therefore, we have $A_\tau \in \mathcal{B}_4$ with $\tau = \mathrm{O}(\log n)$, and $|\beta_j^{(\tau)}| \leq \epsilon'$ for all $j \in I_{<1}$.

Proof of Lemma 6.6.2. Suppose $A \in \mathcal{B}_2$ and recall the definition $K = \frac{C}{\epsilon}$. For any $i \in I_{<1}$, the bound (6.52) yields that

$$|\beta_i'| \leq (1-3\varepsilon)K\sqrt{\frac{\log n}{n}} + K^2\frac{\log n}{n} + C\sqrt{\frac{\log n}{n}} \leq K\sqrt{\frac{\log n}{n}}$$

holds w.h.p. This completes the proof of the statement (i).

Now we consider the statement (ii). Suppose that $A \in \mathcal{B}_2$ and $|\beta_i| \geq \frac{C}{\epsilon} \cdot \frac{1}{\sqrt{n}}$ (we expect $h = \frac{C}{\epsilon}$). For $i \in I_{>1}$, the bound (6.51) implies $|\mathbf{E}[\beta_i']| \geq |\lambda_i||\beta_i| - C\|\boldsymbol{\beta}\|_\infty^2 - \frac{C}{\sqrt{n}}$. Since $\|\boldsymbol{\beta}\|_\infty \leq K\sqrt{\frac{\log n}{n}}$ and $|\beta_i| \geq \frac{h}{\sqrt{n}} = \frac{C}{\epsilon} \cdot \frac{1}{\sqrt{n}}$, we have

$$C\|\boldsymbol{\beta}\|_\infty^2 \leq \frac{CK \log n}{n} \leq \epsilon|\beta_i| \quad \text{(for sufficiently large } n\text{)},$$
$$\frac{C}{\sqrt{n}} \leq \epsilon|\beta_i|.$$

This leads to $|\mathbf{E}[\beta_i']| \geq (1+\epsilon)|\beta_i|$, which completes the proof of the statement (ii).

Chapter 7

Quasi-Majority Functional Voting

7.1 Introduction

In this chapter, we focus on voting processes on expander graphs. We say that a graph G is λ-*expander* if $\max\{|\lambda_2|, |\lambda_n|\} \le \lambda$, where $1 = \lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_n \ge -1$ are the eigenvalues of the transition matrix P of the simple random walk on G. For example, the Erdős–Rényi graph $G(n,p)$ of $p \ge (1+\epsilon)\frac{\log n}{n}$ for an arbitrary constant $\epsilon > 0$ is $\mathrm{O}(1/\sqrt{np})$-expander w.h.p. [CO07]. The random d-regular graph $G_{n,d}$ of $3 \le d \le n/2$ is $\mathrm{O}(1/\sqrt{d})$-expander w.h.p. [CGJ18, TY19].

7.1.1 Voting processes on expander graphs

There is a line of works concerning voting processes on expander graphs. Cooper, Elsässer, Ono, and Radzik [CEOR13] showed that the expected consensus time of the pull voting is $\mathrm{O}(n/(1-\lambda))$ on λ-expander regular graphs for any initial configuration. Compared to the pull voting, the study of the best-of-two on general graphs seems much harder. Most of the previous works concerning the best-of-two on expander graphs put some assumptions on the initial configuration. Let A denote the set of vertices of opinion 0 and let $B := V \setminus A$. Cooper, Elsässer, and Radzik [CER14] showed that, for any regular λ-expander graph, the consensus time is $\mathrm{O}(\log n)$ w.h.p. if $\big||A| - |B|\big| = \Omega(\lambda n)$. This result was improved by Cooper, Elsässer, Radzik, Rivera, and Shiraga [CER+15]. Roughly speaking, they proved that, on λ-expander graphs, the consensus time is $\mathrm{O}(\log n)$ if $|d(A) - d(B)| = \Omega(\lambda^2 d(V))$, where $d(S) = \sum_{v \in S} \deg(v)$ denotes the volume of $S \subseteq V$. To the best of our knowledge, any previous works that studies voting process other than pull voting on non-complete graphs put some assumption on the initial configuration (e.g., random initial configuration [KR19, AD15] and initial bias [CEOR13, CER14, CER+15]).

7.1.2 Our model

In this part, we introduce *quasi-majority functional voting*, a subclass of functional voting (see Definition 5.1.1 for the definition of functional voting).

Definition 7.1.1 (update function)**.** *For a functional voting, the function*

$$H_f(x) := x\big(1 - f(1-x)\big) + (1-x)f(x)$$

is called an update function.

The intuition behind the update function H_f is that, on a complete graph K_n (with self-loops), the functional voting with respect to f satisfies $\mathbf{E}[\alpha'] = \frac{|A|}{n}\left(1 - f\left(\frac{|B|}{n}\right)\right) + \frac{|B|}{n} f\left(\frac{|A|}{n}\right) = H_f(\alpha)$, where $\alpha = |A|/n$ and $\alpha' = |A'|/n$ (see Section 5.3).

Example 7.1.2. The pull voting, best-of-two, and best-of-three are functional votings with respect to x, x^2 and $3x^2 - 2x^3$, respectively. The update function of them are x,$3x^2 - 2x^3$, and $3x^2 - 2x^3$, respectively. In general, the best-of-k is a functional voting with respect to

$$f_k(x) = \sum_{i=\lfloor k/2 \rfloor + 1}^{k} \binom{k}{i} x^i (1-x)^{k-i}. \tag{7.1}$$

It is straightforward to check that $H_{f_k}(x) = f_k(x)$ if k is odd and $H_{f_k}(x) = f_{k+1}(x)$ if k is even. Majority is a functional voting with respect to

$$f(x) = \begin{cases} 0 & \text{if } x < \frac{1}{2}, \\ \frac{1}{2} & \text{if } x = \frac{1}{2}, \\ 1 & \text{if } x > \frac{1}{2} \end{cases} \tag{7.2}$$

if a vertex adopts the random opinion when it meets the tie. It is easy to see that $H_f(x) = f(x)$ for Majority.

We focus on a functional voting with respect to f satisfying the following property.

Definition 7.1.3 (Quasi-majority functional voting)**.** *A function f is* quasi-majority *if f satisfies the following conditions.*

(1) f is in class C^2 (i.e., the derivatives f' and f'' of f exist and are continuous),

(2) $0 < f(1/2) < 1$,

(3) $H_f(x) < x$ whenever $x \in (0, 1/2)$.

(4) $H'_f(1/2) > 1$,

(5) $H'_f(0) < 1$.

A voting process is a quasi-majority functional voting *if it is a functional voting with respect to a quasi-majority function f.*

Note that $H_f(x)$ is symmetric (i.e., $H_f(1-x) = 1 - H_f(x)$) and thus the condition (3) implies $H_f(x) > x$ for every $x \in (1/2, 1)$. Intuitively speaking, the conditions (3) to (5) ensure the drift towards consensus. The conditions (1) and (2) are due to a technical reasons.

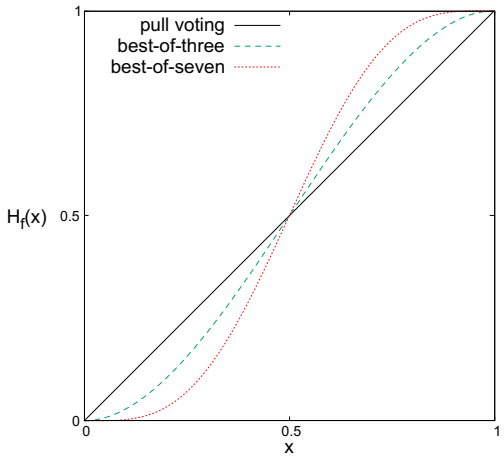


Figure 7.1: The update functions $H_f(x)$ of pull voting (solid line), best-of-three (dashed line) and best-of-seven (dotted line). One can easily observe that best-of-three and best-of-seven are quasi-majority functional voting. Intuitively speaking, quasi-majority functional voting processes have update functions H_f with the property so-called "the rich get richer", which coincides with Definition 7.1.3.

For each constant $k \geq 2$, the best-of-k is quasi-majority functional voting but the pull voting and Majority are not. Indeed, if H_{f_k} is the update function of best-of-k, then $H'_{f_{2\ell}}(x) = H'_{f_{2\ell+1}}(x) = (2\ell+1)\binom{2\ell}{\ell}x^\ell(1-x)^\ell$. It is straightforward to check that this function satisfies the conditions (3) to (5) if $\ell \neq 0$ (this condition excludes the pull-voting). See Figure 7.1 for depiction of update functions of the pull voting, best-of-three and best-of-seven.

7.1.3 Our result

We study the consensus time T_{cons} of a quasi-majority functional voting on expander graphs. Throughout this chapter, we consider sufficiently large $n = |V|$. Let $T_{\mathrm{cons}}(A)$ denote the consensus time starting from the initial configuration $A \subseteq V$. For a graph $G = (V, E)$, let $\pi = (\pi(v))_{v \in V}$ denote the *degree distribution* defined as

$$\pi(v) = \frac{\deg(v)}{2|E|}. \tag{7.3}$$

Note that π is the stationary distribution of the transition matrix (2.1) of the simple random walk on G.

For $A \subseteq V$, let $\pi(A) := \sum_{v \in A} \pi(v)$. Let

$$\delta(A) := \pi(A) - \pi(V \setminus A) = 2\pi(A) - 1$$

denote the *bias* between A and $V \setminus A$.

Theorem 7.1.4 (Main theorem)**.** *Consider a quasi-majority functional voting with respect to f on an n-vertex λ-expander graph with degree distribution π. Then, the following hold:*

(i) Let $C_1 > 0$ be an arbitrary constant and $\varepsilon : \mathbb{N} \to \mathbb{R}$ be an arbitrary function satisfying $\varepsilon(n) \to 0$ as $n \to \infty$. Suppose that $\lambda \le C_1 n^{-1/4}$, $\|\pi\|_2 \le C_1/\sqrt{n}$ and $\|\pi\|_3 \le \varepsilon/\sqrt{n}$. Then, for any $A \subseteq V$, $T_{\mathrm{cons}}(A) = \mathrm{O}(\log n)$ w.h.p.

(ii) Let C_2 be a positive constant depending only on f. Suppose that $\lambda \le C_2$ and $\|\pi\|_2 \le C_2/\sqrt{\log n}$. Then, for any $A \subseteq V$ satisfying $|\delta(A)| \ge C_2 \max\{\lambda^2, \|\pi\|_2\sqrt{\log n}\}$, $T_{\mathrm{cons}}(A) = \mathrm{O}(\log n)$ w.h.p.

The following result which we show in Section 7.5 indicates that the consensus time of Theorem 7.1.4(i) is optimal up to a constant factor.

Theorem 7.1.5 (Lower bound)**.** *Under the same assumption of Theorem 7.1.4(i), $T_{\mathrm{cons}}(A) = \Omega(\log n)$ w.h.p. for some $A \subseteq V$.*

Theorem 7.1.6 (Fast consensus for $H_f'(0) = 0$)**.** *Consider a quasi-majority functional voting with respect to f on an n-vertex λ-expander graph with degree distribution π. Let $C > 0$ be a constant depending only on f. Suppose that $H_f'(0) = 0$, $\lambda \le C$ and $\|\pi\|_2 \le C/\sqrt{\log n}$. Then, for any $A \subseteq V$ satisfying $|\delta(A)| \ge C \max\{\lambda^2, \|\pi\|_2\sqrt{\log n}\}$, it holds w.h.p. that*

$$T_{\mathrm{cons}}(A) = \mathrm{O}\left(\log\log n + \log|\delta(A)|^{-1} + \frac{\log n}{\log \lambda^{-1}} + \frac{\log n}{\log(\|\pi\|_2\sqrt{\log n})^{-1}}\right).$$

For example, for each constant $k \ge 2$, best-of-k is quasi-majority with $H_f'(0) = 0$.

Remark 7.1.7. Roughly speaking, for $p \ge 2$, $\|\pi\|_p$ measures the imbalance of the degrees. For any graphs, $\|\pi\|_p \ge n^{-1+1/p}$ and the equality holds if and only if the graph is regular. For star graphs, we have $\|\pi\|_p \approx 1$.

Results of best-of-k. Our results above do not explore Majority since it is not quasi-majority. A plausible approach is to consider best-of-k for $k = k(n) = \omega(1)$ since each vertex is likely to choose the majority opinion if the number of neighbor sampling increases. Also, note that the betrayal function f_k of best-of-k given in (7.1) converges to that of Majority (i.e., $f_k(x) \to f(x)$ as $k \to \infty$ for each $x \in [0, 1]$, where f is the betrayal function (7.2) of Majority). On the other hand, if $k = \mathrm{O}(1)$, there is a tremendous gap between best-of-k and Majority: For any functional voting on the complete graph K_n, $T_{\mathrm{cons}}(A) = \Omega(\log n)$ for some $A \subseteq V$ from Theorem 7.1.5. Majority on K_n reaches the consensus in a single step if $|A| < |V \setminus A| - 1$. This motivates us to consider best-of-k for $k = k(n) \to \infty$ as $n \to \infty$. For simplicity, we focus on best-of-$(2k+1)$ and prove the following result in Section 7.6.

Theorem 7.1.8. *Let $k = k(n)$ be such that $k = \omega(1)$ and $k = \mathrm{o}(n/\log n)$. Let C be an arbitrary positive constant. Consider the best-of-$(2k+1)$ on an n-vertex λ-expander graph with degree distribution π such that $\lambda \le Ck^{-1/2}n^{-1/4}$, $\|\pi\|_2 \le Cn^{-1/2}$ and $\|\pi\|_3 \le Ck^{-1/6}n^{-1/2}$. Then, $T_{\mathrm{cons}}(A) = \mathrm{O}\left(\frac{\log n}{\log k}\right)$ holds w.h.p. for any $A \subseteq V$. Furthermore, there exists a set $A \subseteq V$ such that $T_{\mathrm{cons}}(A) = \Omega\left(\frac{\log n}{\log k}\right)$ holds w.h.p.*

7.1.4 Application

Here, we apply our main theorem to specific graphs and derive some useful results.

Erdős–Rényi graph

For any $p \geq (1+\epsilon)\frac{\log n}{n}$ for an arbitrary constant $\epsilon > 0$, the Erdős–Rényi graph $G(n,p)$ is connected and $\mathrm{O}(1/\sqrt{np})$-expander w.h.p [CO07, FK16].

Corollary 7.1.9. *Consider the best-of-k on the Erdős–Rényi graph $G(n,p)$ for an arbitrary constant $k \geq 2$. Then, $G(n,p)$ w.h.p. satisfies the following:*

(i) Suppose that $p = \Omega(n^{-1/2})$. Then

(a) for any $A \subseteq V$, $T_{\mathrm{cons}}(A) = \mathrm{O}(\log n)$ w.h.p.

(b) for some $A \subseteq V$, $T_{\mathrm{cons}}(A) = \Omega(\log n)$ w.h.p.

(ii) Suppose that $p \geq (1+\epsilon)\frac{\log n}{n}$ for an arbitrary constant $\epsilon > 0$. Then, for any $A \subseteq V$ satisfying $|\delta(A)| \geq C\max\left\{\frac{1}{np}, \sqrt{\frac{\log n}{n}}\right\}$, $T_{\mathrm{cons}}(A) = \mathrm{O}\left(\log\log n + \log|\delta(A)|^{-1} + \frac{\log n}{\log(np)}\right)$ w.h.p., where $C > 0$ is a constant depending only on f.

If $\frac{\log n}{\log(np)} = \mathrm{O}(\log\log n)$ (or equivalently, $np = n^{\Omega(1/\log\log n)}$), Corollary 7.1.9(ii) implies $T_{\mathrm{cons}}(A) = \mathrm{O}(\log\log n + \log|\delta(A)|^{-1})$ w.h.p.

Corollary 7.1.10. *Let $k = k(n)$ be such that $k = \omega(1)$ and $k = \mathrm{O}(\sqrt{n})$. Consider the best-of-$(2k+1)$ on $G(n,p)$ for $p = \Omega(k/\sqrt{n})$. Then, for any $A \subseteq V$, $T_{\mathrm{cons}}(A) = \mathrm{O}\left(\frac{\log n}{\log k}\right)$ holds w.h.p.*

From Corollary 7.1.10, best-of-n^ϵ on $G(n, n^{-1/2+\epsilon})$ for any constant $\epsilon \in (0, 1/2)$ reaches consensus in $\mathrm{O}(1)$ steps. It is known that Majority on $G(n, Cn^{-1/2})$ satisfies $T_{\mathrm{cons}}(A) \leq 4$ for large constant C and random $A \subseteq V$ with constant probability [BCO+16].

Random regular graph

For $3 \leq d \leq n/2$, it is known that the random d-regular graph $G_{n,d}$ is connected and $\mathrm{O}(1/\sqrt{d})$-expander w.h.p. [CGJ18, TY19].

Corollary 7.1.11. *Consider the best-of-k on the random d-regular graph $G_{n,d}$ for an arbitrary constant $k \geq 2$. Then, $G_{n,d}$ w.h.p. satisfies the following:*

(i) Suppose that $d = \Omega(n^{1/2})$ and $d \leq n/2$. Then,

(a) for any $A \subseteq V$, $T_{\mathrm{cons}}(A) = \mathrm{O}(\log n)$ w.h.p.

(b) for some $A \subseteq V$, $T_{\mathrm{cons}}(A) = \Omega(\log n)$ w.h.p.

(ii) Suppose that $d \geq C$ and $d \leq n/2$ for a constant $C > 0$ depending only on f. Then, for any $A \subseteq V$ satisfying $|\delta(A)| \geq C\max\left\{\frac{1}{d}, \sqrt{\frac{\log n}{n}}\right\}$, $T_{\mathrm{cons}}(A) = \mathrm{O}\left(\log\log n + \log|\delta(A)|^{-1} + \frac{\log n}{\log d}\right)$ w.h.p.

Corollary 7.1.12. *Let $k = k(n)$ be such that $k = \omega(1)$ and $k = \mathrm{O}(\sqrt{n})$. Consider the best-of-$(2k+1)$ on the random d-regular graph $G_{n,d}$ such that $d = \Omega(k\sqrt{n})$ and $d \leq n/2$. Then, for any $A \subseteq V$, $T_{\mathrm{cons}}(A) = \mathrm{O}\left(\frac{\log n}{\log k}\right)$ holds w.h.p.*

Roughly-regular expander graphs

We can apply Theorems 7.1.4 and 7.1.6 if the ratio of the maximum and average degree is constant as follows.

Corollary 7.1.13. *Consider a quasi-majority functional voting with respect to f on an n-vertex λ-expander graph with degree distribution π. Suppose that $d_{\max} \leq C_1 d_{\mathrm{ave}}$ for an arbitrary constant $C_1 > 0$, where $d_{\max}$ and d_{avr} denote the maximum and average degree, respectively. Then, the following hold:*

(i) Suppose that $\lambda \leq C_1 n^{-1/4}$. Then

(a) *for any* $A \subseteq V$, $T_{\mathrm{cons}}(A) = \mathrm{O}(\log n)$ *w.h.p.*
(b) *for some* $A \subseteq V$, $T_{\mathrm{cons}}(A) = \Omega(\log n)$ *w.h.p.*

(ii) *Suppose that* $\lambda \le C_2$ *for some constant* $C_2 > 0$ *depending only on* f. *Then, for any* $A \subseteq V$ *satisfying* $|\delta(A)| \ge C_2 \max\{\lambda^2, \sqrt{\frac{\log n}{n}}\}$, $T_{\mathrm{cons}}(A) = \mathrm{O}(\log n)$ *w.h.p.*

(iii) *In addition to the same assumption as (ii), suppose that* $H_f'(0) = 0$. *Then, it holds w.h.p. that* $T_{\mathrm{cons}}(A) = \mathrm{O}\left(\log\log n + \log|\delta(A)|^{-1} + \frac{\log n}{\log \lambda^{-1}}\right)$.

Corollary 7.1.14. *Let* $k = k(n)$ *be such that* $k = \omega(1)$ *and* $k = \mathrm{o}(n/\log n)$. *Let* C *be an arbitrary constant. Consider the best-of-*$(2k+1)$ *on an* n*-vertex* λ*-expander graph with degree distribution* π *such that* $\lambda \le Ck^{-1/2}n^{-1/4}$, *and* $d_{\max} \le Cd_{\mathrm{avr}}$, *where* $d_{\max}$ *and* d_{avr} *denote the maximum and average degree, respectively. Then,* $T_{\mathrm{cons}}(A) = \mathrm{O}\left(\frac{\log n}{\log k}\right)$ *holds w.h.p. for any* $A \subseteq V$.

Corollaries 7.1.13 and 7.1.14 immediately follow from Theorems 7.1.4 to 7.1.6 and 7.1.8 since $\|\pi\|_2 = \mathrm{O}(n^{-1/2})$. Note that if the ratio of the maximum degree $d_{\max}$ and average degree d_{avr} is constant, $\|\pi\|_p = \Theta(1/n^{1-1/p})$ since $\pi(v) = \mathrm{O}(1/n)$ for all $v \in V$. We obtain Corollaries 7.1.9 to 7.1.12 from Corollaries 7.1.13 and 7.1.14.

7.1.5 Other quasi-majority functional votings

We can consider the *ρ-lazy* variant of a voting process, i.e., every vertex v individually tosses its private coin and operates the voting process with probability ρ, while v does nothing with probability $1-\rho$. Berenbrink, Giakkoupis, Kermarrec, and Mallmann-Trenn [BGKMT16] studies 1/2-lazy pull voting. If the original voting process is a quasi-majority functional voting with respect to f, then the corresponding ρ-lazy variant is quasi-majority functional voting with respect to $\rho f : x \mapsto \rho f(x)$. Indeed, $H_{\rho f}(x) = (1-\rho)x + \rho H_f(x)$.

Corollary 7.1.15. *Consider a* ρ*-lazy quasi-majority functional voting on* $G(n,p)$ *for an arbitrary constant* $\rho \in (0,1]$. *Suppose that* $p = \Omega(1/\sqrt{n})$. *Then, for any* $A \subseteq V$, $T_{\mathrm{cons}}(A) = \mathrm{O}(\log n)$ *w.h.p.*

This implies the following interesting observation. In voting processes, the number of neighbor sampling queries per each vertex at each step affects the performance. In pull voting, each vertex communicates with one neighbor but it has a drawback on the slow consensus time. In the best-of-two, each vertex communicates with two random neighbors and its consensus time is much faster than that of pull voting. In ρ-lazy best-of-two, each vertex queries 2ρ vertices at each round in expectation, that is less queries than pull voting if $\rho < 1/2$. On the other hand, the consensus time is much faster than pull voting.

Additionally, we can deal with *k-careful voting*. In this model, each vertex v selects k random neighbors (with replacement), and if these sampled k opinions are the same one, v adopts it. Note that the one-careful voting and two-careful voting are equivalent to the pull voting and best-of-two, respectively. One can check easily that, for any constant $k \ge 2$, this model is a quasi-majority functional voting with respect to $f(x) = x^k$. Note that $H_f'(0) = 0$ and $H_f'(1/2) = 1 + \frac{k-1}{2^{k-1}}$.

Corollary 7.1.16. *Consider a* k*-careful voting on* $G(n,p)$ *for an arbitrary constant* $k \ge 2$. *Suppose that* $p = \Omega(1/\sqrt{n})$. *Then, for any* $A \subseteq V$, $T_{\mathrm{cons}}(A) = \mathrm{O}(\log n)$ *w.h.p.*

7.2 Preliminary

7.2.1 Technical background

Consider best-of-two on a complete graph K_n (with self loop on each vertex) with a current configuration $A \subseteq V$. Let $\alpha = |A|/n$. We have $P(v,A) = \alpha$ for any $v \in V$ and $A \subseteq V$. Then, for any $A \subseteq V$, $\mathbf{E}[\alpha'] = H_f(\alpha) = 3\alpha^2 - 2\alpha^3$. Thus, in each round, $\alpha' = 3\alpha^2 - 2\alpha^3 \pm \mathrm{O}(\sqrt{\log n/n})$ holds w.h.p. from Proposition 5.3.1. Therefore, we can regard the behavior of α as the iteration of applying H_f.

The most technical part is the symmetry breaking at $\alpha = 1/2$. Note that $H_f(1/2) = 1/2$ and thus, the argument above does not work in the case of $|\alpha - 1/2| = \mathrm{o}(\sqrt{\log n/n})$. To analyze this case, the authors of [DGM+11, CGG+18] proved the following technical lemma asserting that α w.h.p. escapes from the area in $\mathrm{O}(\log n)$ rounds.

Lemma 7.2.1 (Lemma 4.5 of [CGG+18] (informal)). *For any constant C, it holds w.h.p. that $|\alpha-1/2| \geq C\sqrt{\log n/n}$ in $\mathrm{O}(\log n)$ rounds (the hidden constant factor depends on C) if*

(i) For any constant h, there is a constant $C_0 > 0$ such that, if $|\alpha - 1/2| = \mathrm{O}(\sqrt{\log n/n})$ then $\mathbf{Pr}[|\alpha' - 1/2| > h/\sqrt{n}] > C_0$.

(ii) If $|\alpha - 1/2| = \mathrm{O}(\sqrt{\log n/n})$ and $|\alpha - 1/2| = \Omega(1/\sqrt{n})$, $\mathbf{Pr}[|\alpha' - 1/2| \leq (1+\epsilon)|\alpha - 1/2|] \leq \exp(-\Theta((\alpha-1/2)^2 n))$ for some constant $\epsilon > 0$.

Intuitively speaking, the condition (ii) means that the bias $|\alpha' - 1/2|$ is likely to be at least $(1+\epsilon)|\alpha - 1/2|$ for some constant $\epsilon > 0$. The condition (ii) is easy to check using the Hoeffding bound. The condition (i) means that α' has a fluctuation of size $\Omega(1/\sqrt{n})$ with a constant probability. We can check condition (i) using the Central Limit Theorem (or the Berry-Esseen bound, see Proposition 5.4.7). The Central Limit Theorem implies that the normalized random variable $(\alpha' - \mathbf{E}[\alpha'])/\sqrt{\mathbf{Var}[\alpha']}$ converges to the standard normal distribution as $n \to \infty$. In other words, α' has a fluctuation of size $\Theta(\sqrt{\mathbf{Var}[\alpha']})$ with constant probability. Now, to verify the condition (i), we evaluate $\mathbf{Var}[\alpha']$. On K_n, it is easy to show that $\mathbf{Var}[\alpha'] = \Theta(1/n)$, which implies the condition (i).

The authors of [CER+15, CRRS17] considered best-of-two on expander graphs. They focused on the behavior of $\pi(A)$ instead of α. Roughly speaking, they proved that $\mathbf{E}[\pi(A') - 1/2] \geq (1+\epsilon)(\pi(A) - 1/2) - \mathrm{O}(\lambda^2)$. Then, from the Hoeffding bound, we have $\mathbf{E}[\pi(A') - 1/2] \geq (1+\epsilon)(\pi(A) - 1/2) - \mathrm{O}(\lambda^2 + \|\pi\|_2\sqrt{\log n})$. Thus, if the initial bias $|\pi(A) - 1/2|$ is $\Omega(\max\{\lambda^2, \sqrt{\log n/n}\})$, we can show that the consensus time is $\mathrm{O}(\log n)$.

Unfortunately, we can not apply the same technique to evaluate $\mathbf{Var}[\pi(A')]$ on expander graphs, and due to this reason, the worst-case consensus time on expander graphs seems to be a nontrivial task. Actually, any previous works put assumptions on the initial bias due to the same reason. It should be noted that Lemma 7.2.1 is well-known in the literature. For example, Cruciani, Natale, and Scornavacca [CNS19] used Lemma 7.2.1 from random initial configurations (under this assumption, we can trivially evaluate $\mathbf{Var}[\pi(A_0)]$, where the randomness is taken over the initial configuration).

The techniques of evaluating $\mathbf{E}[\pi(A')]$ by Cooper, Elsässer, Radzik, Rivera, and Shiraga [CER+15] and Cooper, Radzik, Rivera, and Shiraga [CRRS17] are specialized in best-of-two. Thus, it is not straightforward to prove the evaluation of $\mathbf{E}[\pi(A')]$ for voting processes other than best-of-two.

7.2.2 Our technical contribution

For a C^2 function $h : \mathbb{R} \to \mathbb{R}$, let

$$K_1(h) := \max_{x\in[0,1]} |h'(x)|, \quad K_2(h) := \max_{x\in[0,1]} |h''(x)| \tag{7.4}$$

be some constants depending only on h. The following technical result enables us to evaluate $\mathbf{E}[\pi(A')]$ and $\mathbf{Var}[\pi(A')]$ of a functional voting with respect to any C^2 function.

Lemma 7.2.2. *Consider a functional voting with respect to a C^2 function f on a λ-expander graph. Then, for all $A \subseteq V$,*

$$\big|\mathbf{E}[\pi(A')] - H_f\big(\pi(A)\big)\big| \leq K_2(f)\lambda\big(|2\pi(A) - 1| + \lambda\big)\pi(A)\big(1 - \pi(A)\big).$$

Lemma 7.2.3. *Consider a functional voting with respect to a C^2 function f on a λ-expander graph. Let $g(x) := f(x)(1 - f(x))$. Then, for all $A \subseteq V$,*

$$\left|\mathbf{Var}[\pi(A')] - \|\pi\|_2^2 g\left(\frac{1}{2}\right)\right| \leq K_1(g)\left(\frac{1}{2}\|\pi\|_2^2\,|2\pi(A) - 1| + 2\|\pi\|_3^{3/2}\lambda\sqrt{\pi(A)\big(1 - \pi(A)\big)}\right).$$

7.2.3 Proof sketch of the main result

We present proof sketch of Theorem 7.1.4(i). From the assumption of Theorem 7.1.4(i) and Lemma 7.2.3, if $|\pi(A) - 1/2| = \mathrm{o}(1)$, we have $\mathbf{Var}[\pi(A')] = \Theta(\|\pi\|_2^2 g(\pi(A))) = \Theta(\|\pi\|_2^2 g(1/2 + \mathrm{o}(1))) = \Theta(1/n)$. Moreover, $\mathbf{E}[\pi(A')] = H_f(\pi(A)) \pm \mathrm{O}(\pi(A)/\sqrt{n})$ holds for any $A \subseteq V$. Hence, from the Hoeffding bound, $\pi(A') = H_f(\pi(A)) + \mathrm{O}(\sqrt{\log n/n})$ holds w.h.p. for any $A \subseteq V$.

- If $|\pi(A)-1/2| = \mathrm{O}(\sqrt{\log n/n})$, we use Lemma 7.2.1 to obtain an $\mathrm{O}(\log n)$ round symmetry breaking. In this phase, since $|\pi(A) - 1/2| = \mathrm{o}(1)$, $\mathbf{Var}[\pi(A') - 1/2] = \Theta(1/n)$. Then, from the Berry-Esseen theorem (Proposition 5.4.7), we can check the condition (i). To check the condition (ii), we invoke the condition $H_f'(1/2) > 1$ of the quasi-majority function. From Taylor's theorem and the assumption of Lemma 7.2.1(ii) ($\pi(A) - 1/2 = \Omega(1/\sqrt{n})$), $\mathbf{E}[\pi(A') - 1/2] = H_f(\pi(A)) - H_f(1/2) - \mathrm{O}(1/\sqrt{n}) \approx (1+\epsilon_1)(\pi(A) - 1/2)$ for some positive constant $\epsilon_1 > 0$. Note that $H_f(1/2) = 1/2$.
- If $C_1\sqrt{\log n/n} \le |\pi(A) - 1/2| \le C_2$ for sufficiently large constant C_1 and some constant $C_2 > 0$, we use the Hoeffding bound and then obtain $\pi(A') - 1/2 \approx (1+\epsilon_1)(\pi(A) - 1/2) - \mathrm{O}(\sqrt{\log n/n}) \ge (1 + (\epsilon_1/2))(\pi(A) - 1/2)$ w.h.p. Hence, $\mathrm{O}(\log n)$ rounds suffice to yield a constant bias. (Note that this argument holds when $|\pi(A) - 1/2| \le C_2$ due to the remainder term of Taylor's theorem.)
- If $C_3 \le \pi(A) < 1/2$, it is straightforward to see that $\pi(A') = H_f(\pi(A)) + \mathrm{O}(\sqrt{\log n/n}) \le \pi(A) - \epsilon_2$ w.h.p. for some constant $\epsilon_2 > 0$. Note that we invoke the property that $H_f(x) < x$ whenever $0 < x < 1/2$.
- If $\pi(A) \le C_3$ for sufficiently small constant C_3, we use the Markov inequality to show $\pi(A_t) = \mathrm{O}(n^{-3})$ w.h.p. for some $t = \mathrm{O}(\log n)$. Since $\pi(A) \ge 1/n^2$ whenever $A \neq \emptyset$, this implies that the consensus time is $\mathrm{O}(\log n)$ w.h.p. Note that, since $H_f'(0) < 1$, we have $\mathbf{E}[\pi(A')] \le H_f(\pi(A)) + \mathrm{O}(\pi(A)/\sqrt{n}) \approx H_f'(0)\pi(A) + \mathrm{O}(\pi(A)/\sqrt{n}) \le (1-\epsilon_3)\pi(A)$ for some constant $\epsilon_3 > 0$.

In the proof of Theorem 7.1.8, we modify Lemma 7.2.1 and apply the same argument.

7.2.4 Tools

Let $\lambda_1 \ge \cdots \ge \lambda_n$ be eigenvalues of a transition matrix P. If P is reversible, it is known that $\lambda_i \in \mathbb{R}$ for all i. Let $\lambda = \max\{|\lambda_2|, |\lambda_n|\}$ be the second largest eigenvalue in absolute value. The quantity

$$Q(S,T) := \sum_{v\in S} \pi(v)P(v,T) \tag{7.5}$$

is well known as *edge measure* [LP17] or *ergodic flow* [AF, MT06] in the context of Markov chain. Note that, for any reversible P and subsets $S, T \subseteq V$, $Q(S,T) = Q(T,S)$ holds.

Proposition 7.2.4 (Expander mixing lemma; p.163 of [LP17])**.** *Suppose P is reversible. Then, for any $S, T \subseteq V$,*

$$|Q(S,T) - \pi(S)\pi(T)| \le \lambda\sqrt{\pi(S)\pi(T)\big(1-\pi(S)\big)\big(1-\pi(T)\big)}.$$

Proposition 7.2.5 (Lemma 3 of [CRRS17])**.** *Suppose that P is reversible. Then, for any $S \subseteq V$,*

$$\left|\sum_{v\in V} \pi(v)P(v,S)^2 - \pi(S)^2\right| \le \lambda^2\pi(S)\big(1-\pi(S)\big).$$

Corollary 7.2.6. *Suppose that P is reversible. Then, for any $S \subseteq V$,*

$$\sum_{v\in V} \pi(v)\big(P(v,S) - \pi(S)\big)^2 \le \lambda^2\pi(S)\big(1-\pi(S)\big).$$

Proof. Since $Q(V,S) = Q(S,V) = \pi(S)$ for any reversible P and $S \subseteq V$, we have

$$\begin{aligned}\sum_{v\in V} \pi(v)\big(P(v,S) - \pi(S)\big)^2 &= \sum_{v\in V} \pi(v)P(v,S)^2 + \pi(S)^2 - 2\pi(S)Q(V,S)\\ &= \sum_{v\in V} \pi(v)P(v,S)^2 - \pi(S)^2 \le \lambda^2\pi(S)\big(1-\pi(S)\big).\end{aligned}$$

Here, we invoked Proposition 7.2.5 in the last inequality. □

7.3 Evaluate the Expectation and Variance of $\pi(A')$

7.3.1 Extension to reversible Markov chains

The main results of this chapter (Theorems 7.1.4 to 7.1.6) hold for any quasi-majority functional voting specified by a reversible Markov chain (see Definition 5.1.1). For example, we can apply our results to a voting process on an edge-weighted graph $G = (V, E, w)$, where $w : E \to \mathbb{R}_{\geq 0}$ denotes an edge weight function by considering the transition matrix P defined as follows: $P(u, v) = w(\{u, v\})/\sum_{x:\{u,x\}\in E} w(\{u, x\})$ for $\{u, v\} \in E$ and $P(u, v) = 0$ for $\{u, v\} \notin E$. A *weighted functional voting with respect to* f is determined by $\mathbf{Pr}[v \in A' \mid v \in B] = f(P(v, B))$ and $\mathbf{Pr}[v \in B' \mid v \in A] = f(P(v, A))$. The weighted variant can be interpreted as follows: For example, in the best-of-three, a vertex selects three neighbors uniformly at random. In the weighted best-of-three, each edge has a positive weight and the probability of the selection is in proportional to the weight of the edge.

For a function $h : \mathbb{R} \to \mathbb{R}$ and subsets $S, T \subseteq V$, we are interested in the quantity $Q_h(S, T)$ defined as

$$Q_h(S, T) := \sum_{v\in S} \pi(v)h\big(P(v, T)\big). \tag{7.6}$$

Note that the edge measure (7.5) is a special case of Q_h where $h(x) = x$. We show the following lemma which gives a useful evaluation of $Q_h(S, T)$.

Lemma 7.3.1. *Suppose P is reversible. Then, for any $S, T \subseteq V$ and any C^2 function $h : \mathbb{R} \to \mathbb{R}$,*

$$\Big|Q_h(S, T) - \pi(S)h\big(\pi(T)\big) - h'\big(\pi(T)\big)\big(Q(S, T) - \pi(S)\pi(T)\big)\Big| \leq \frac{K_2(h)}{2}\lambda^2\pi(T)\big(1 - \pi(T)\big).$$

Proof of Lemma 7.3.1. From Taylor's theorem, it holds for any $x, y \in [0, 1]$ that

$$|h(x) - h(y) - h'(y)(x - y)| \leq \frac{K_2(h)}{2}(x - y)^2.$$

Hence

$$\begin{aligned}
&\Big|Q_h(S, T) - \pi(S)h\big(\pi(T)\big) - h'\big(\pi(T)\big)\big(Q(S, T) - \pi(S)\pi(T)\big)\Big| \\
&= \left|\sum_{v\in S} \pi(v)\Big(h\big(P(v, T)\big) - h\big(\pi(T)\big) - h'\big(\pi(T)\big)\big(P(v, T) - \pi(T)\big)\Big)\right| \\
&\leq \sum_{v\in S} \pi(v)\Big|h\big(P(v, T)\big) - h\big(\pi(T)\big) - h'\big(\pi(T)\big)\big(P(v, T) - \pi(T)\big)\Big| \\
&\leq \sum_{v\in S} \pi(v)\frac{K_2(h)}{2}\big(P(v, T) - \pi(T)\big)^2 \leq \frac{K_2(h)}{2}\sum_{v\in V} \pi(v)\big(P(v, T) - \pi(T)\big)^2 \\
&\leq \frac{K_2(h)}{2}\lambda^2\pi(T)\big(1 - \pi(T)\big).
\end{aligned}$$

Note that the last inequality follows from Corollary 7.2.6. □

Next, consider

$$R_h(S, T) := \sum_{v\in S} \pi(v)^2 h\big(P(v, T)\big) \tag{7.7}$$

for a function $h : \mathbb{R} \to \mathbb{R}$ and $S, T \subseteq V$. For ease of notation, let $\pi_2(S) := \sum_{v\in S} \pi(v)^2$ for $S \subseteq V$. We show the following result that evaluates $R_h(S, T)$.

Lemma 7.3.2. *Suppose that P is reversible. Then, for any $S, T \subseteq V$ and any C^2 function $h : \mathbb{R} \to \mathbb{R}$,*

$$\big|R_h(S, T) - \pi_2(S)h\big(\pi(T)\big)\big| \leq K_1(h)\|\pi\|_3^{3/2}\lambda\sqrt{\pi(T)\big(1 - \pi(T)\big)}.$$

Proof. We first observe that

$$\left|h(x) - h(y)\right| \le K_1(h)|x - y| \tag{7.8}$$

holds for any $x, y \in [0,1]$ from Taylor's theorem. Hence,

$$\begin{aligned}
&\left|R_h(S,T) - \pi_2(S)h\big(\pi(T)\big)\right| \\
&= \left|\sum_{v\in S} \pi(v)^2 \Big(h\big(P(v,T)\big) - h\big(\pi(T)\big)\Big)\right| \le \sum_{v\in S} \pi(v)^2 \left|h\big(P(v,T)\big) - h\big(\pi(T)\big)\right| \\
&\le \sum_{v\in S} \pi(v)^2 K_1(h)\left|P(v,T) - \pi(T)\right| \le K_1(h) \sum_{v\in V} \pi(v)^2 \left|P(v,T) - \pi(T)\right|.
\end{aligned}$$

Then, applying the Cauchy-Schwarz inequality and Corollary 7.2.6,

$$\begin{aligned}
\sum_{v\in V} \pi(v)^2 \left|P(v,T) - \pi(T)\right| &\le \sqrt{\left(\sum_{v\in V} \pi(v)^3\right)\left(\sum_{v\in V} \pi(v)\big(P(v,T) - \pi(T)\big)^2\right)} \\
&\le \|\pi\|_3^{3/2} \lambda \sqrt{\pi(T)\big(1-\pi(T)\big)}
\end{aligned}$$

and we obtain the claim. □

For simplicity, in this chapter, we do not explore the weighted variant and focus on the usual setting where P is the matrix (2.1) and its stationary distribution π is (7.3).

7.3.2 Expectation and Variance

This section is devoted to evaluating $\mathbf{E}[\pi(A')]$ and $\mathbf{Var}[\pi(A')]$ for a quasi-majority functional voting. Recall that we use $B = V \setminus A$ for $A \subseteq V$. Then, it is clear that

$$\mathbf{E}[\pi(A')] = \pi(A) - \sum_{v\in A} \pi(v) f\big(P(v,B)\big) + \sum_{v\in B} \pi(v) f\big(P(v,A)\big), \tag{7.9}$$

$$\mathbf{Var}[\pi(A')] = \sum_{v\in A} \pi(v)^2 g\big(P(v,B)\big) + \sum_{v\in B} \pi(v) g\big(P(v,A)\big). \tag{7.10}$$

Proof of Lemma 7.2.2. From Definition 5.1.1, (7.6) and (7.9), we have

$$\mathbf{E}[\pi(A')] = \pi(A) - Q_f(A,B) + Q_f(B,A), \tag{7.11}$$

$$H_f\big(\pi(A)\big) = \pi(A) - \pi(A) f\big(\pi(B)\big) + \pi(B) f\big(\pi(A)\big). \tag{7.12}$$

For notational convenience, for $S, T \subseteq V$, let

$$\begin{aligned}
\Delta_f(S,T) &:= Q_f(S,T) - \pi(S) f\big(\pi(T)\big) - f'\big(\pi(T)\big)\big(Q(S,T) - \pi(S)\pi(T)\big) \\
&= Q_h(S,T) - \pi(S) f\big(\pi(T)\big) - f'\big(\pi(T)\big)\big(Q(T,S) - \pi(T)\pi(S)\big).
\end{aligned}$$

The equality follows from the reversibility of P (see Section 7.3). From Lemma 7.3.1, we have

$$|\Delta_f(S,T)| \le \frac{K_2(f)}{2} \lambda^2 \pi(T)\big(1-\pi(T)\big).$$

Then, combining (7.11) and (7.12), we have

$$\begin{aligned}
&\left|\mathbf{E}[\pi(A')] - H_f\big(\pi(A)\big)\right| \\
&= \left|Q_f(B,A) - \pi(B) f\big(\pi(A)\big) - Q_f(A,B) + \pi(A) f\big(\pi(B)\big)\right| \\
&= \Big|\Delta_f(B,A) + f'\big(\pi(A)\big)\big(Q(A,B) - \pi(A)\pi(B)\big) \\
&\qquad -\Delta_f(A,B) - f'\big(\pi(B)\big)\big(Q(A,B) - \pi(A)\pi(B)\big)\Big| \\
&\le |\Delta_f(B,A)| + |\Delta_f(A,B)| + \left|f'\big(\pi(A)\big) - f'\big(\pi(B)\big)\right| \left|Q(A,B) - \pi(A)\pi(B)\right| \\
&\le K_2(f)\lambda^2 \pi(A)\pi(B) + K_2(f)\left|\pi(A) - \pi(B)\right| \lambda \pi(A)\pi(B).
\end{aligned}$$

and we obtain the claim. Note that the last inequality follows from Taylor's theorem (7.8) and Proposition 7.2.4. □

Proof of Lemma 7.2.3. From (7.7) and (7.10),

$$\mathbf{Var}[\pi(A')] = R_g(A,B) + R_g(B,A).$$

Thus, applying Lemma 7.3.2 yields

$$\left|\mathbf{Var}\left[\pi(A')\right] - \Big(\pi_2(A)g\big(\pi(B)\big) + \pi_2(B)g\big(\pi(A)\big)\Big)\right| \le 2K_1(g)\|\pi\|_3^{3/2}\lambda\sqrt{\pi(A)\pi(B)}. \tag{7.13}$$

Next, using Taylor's theorem (7.8),

$$\begin{aligned}
&\left|\pi_2(A)g\big(\pi(B)\big) + \pi_2(B)g\big(\pi(A)\big) - \|\pi\|_2^2 g\left(\frac{1}{2}\right)\right| \\
&= \left|\pi_2(A)\left(g\big(\pi(B)\big) - g\left(\frac{1}{2}\right)\right) + \pi_2(B)\left(g\big(\pi(A)\big) - g\left(\frac{1}{2}\right)\right)\right| \\
&\le K_1(g)\pi_2(A)\left|\pi(B) - \frac{1}{2}\right| + K_1(g)\pi_2(B)\left|\pi(A) - \frac{1}{2}\right| = K_1(g)\|\pi\|_2^2\left|\pi(A) - \frac{1}{2}\right|.
\end{aligned} \tag{7.14}$$

The last equality follows since $|\pi(A) - 1/2| = |\pi(B) - 1/2|$. Combining (7.13) and (7.14), we obtain the claim. □

7.3.3 Symmetric functions

In this subsection, we consider a special case that $f(x)$ is symmetry (i.e., $f(1-x) = 1 - f(x)$ holds). We present the following simplified variant of Lemma 7.2.2, which we will use in Section 7.6. Note that, for a symmetric function f, it holds that $H_f(x) = f(x)$ hor all $x \in [0,1]$.

Lemma 7.3.3 (Special case of Lemma 7.2.2)**.** *Consider a functional voting with respect to a symmetric C^2 function f on a λ-expander graph with degree distribution π. Suppose that f satisfies $f(1-x) = 1 - f(x)$ for every $x \in [0,1]$. Then, for all $A \subseteq V$,*

$$\left|\mathbf{E}[\pi(A')] - f(\pi(A))\right| \le \frac{K_2(f)}{2}\lambda^2\pi(A)\big(1-\pi(A)\big).$$

Proof. Note that, for a functional voting with respect to a symmetric f, we have $\mathbf{Pr}[v \in A'] = f(P(v,A))$ for any $v \in V$. Thus we have

$$\mathbf{E}[\pi(A')] = \sum_{v \in V}\pi(v)f\big(P(v,A)\big) = Q_f(V,A).$$

By substituting V to S of Lemma 7.3.1, we obtain

$$\left|Q_f(V,A) - f\big(\pi(A)\big)\right| \le \frac{K_2(f)}{2}\lambda^2\pi(A)\big(1-\pi(A)\big).$$

Note that $Q(V,T) = Q(T,V) = \pi(T)$ from the reversibility of P. □

7.4 Proof of main results

We prove Theorems 7.1.4 and 7.1.6. Consider a quasi-majority functional voting with respect to f on an n-vertex λ-expander graph with degree distribution π. Let $A_0, A_1, \ldots,$ be the sequence given by the functional voting with initial configuration $A_0 \subseteq V$. Theorems 7.1.4 and 7.1.6 follow from the following lemma.

Lemma 7.4.1. *Consider a quasi-majority functional voting with respect to f on an n-vertex λ-expander graph with degree distribution π. For the betrayal function f, let $\epsilon_h(f) := H_f'(1/2) - 1$, $\epsilon_c(f) := 1 - H_f'(0)$, and $K(f) := \max\{K_2(f), K_2(H_f)\}$ be three positive constants depending only on f. Then, the following hold:*

(I) Let $C_1 > 0$ be an arbitrary constant and $\varepsilon : \mathbb{N} \to \mathbb{R}$ be an arbitrary function satisfying $\varepsilon(n) \to 0$ as $n \to \infty$. Suppose that $\lambda \le C_1 n^{-1/4}$, $\|\pi\|_2 \le C_1/\sqrt{n}$ and $\|\pi\|_3 \le \varepsilon/\sqrt{n}$. Then, for any $A_0 \subseteq V$ such that $|\delta(A_0)| \le c_1 \log n/\sqrt{n}$ for an arbitrary constant $c_1 > 0$, $|\delta(A_t)| \ge c_1 \log n/\sqrt{n}$ within $t = \mathrm{O}(\log n)$ steps w.h.p.

(II) *Suppose that* $\lambda \le \frac{\epsilon_h(f)}{2K(f)}$. *Then, for any* $A_0 \subseteq V$ *s.t.* $\frac{2\max\{K(f),8\}}{\epsilon_h(f)} \max\{\lambda^2, \|\pi\|_2\sqrt{\log n}\} \le |\delta(A_0)| \le \frac{\epsilon_h(f)}{K(f)}$, $|\delta(A_t)| \ge \frac{\epsilon_h(f)}{K(f)}$ *within* $t = \mathrm{O}(\log|\delta(A_0)|^{-1})$ *steps w.h.p.*

(III) *Let* c_2, c_3 *be two arbitrary constants satisfying* $0 < c_2 < c_3 < 1/2$ *and* $\epsilon(f) := \min_{x\in[c_2,c_3]}\big(x - H_f(x)\big)$ *be a positive constant depending* f, c_2, c_3. *Suppose that* $\lambda \le \frac{\epsilon(f)}{2K(f)}$ *and* $\|\pi\|_2 \le \frac{\epsilon(f)}{4\sqrt{\log n}}$. *Then, for any* $A_0 \subseteq V$ *satisfying* $c_2 \le \pi(A_0) \le c_3$, $\pi(A_t) \le c_2$ *within constant steps w.h.p.*

(IV) *Suppose that* $\lambda \le \frac{\epsilon_c(f)}{2K(f)}$ *and* $\|\pi\|_2 \le \frac{\epsilon_c(f)^2}{32K(f)\sqrt{\log n}}$. *Then, for any* $A_0 \subseteq V$ *satisfying* $\pi(A_0) \le \frac{\epsilon_c(f)}{8K(f)}$, $\pi(A_t) = 0$ *within* $t = \mathrm{O}(\log n)$ *steps w.h.p.*

(V) *Suppose that* $H_f'(0) = 0$, $\lambda \le \frac{1}{10K(f)}$ *and* $\|\pi\|_2 \le \frac{1}{64K(f)\sqrt{\log n}}$. *Then, for any* $A_0 \subseteq V$ *satisfying* $\pi(A_0) \le \frac{1}{7K(f)}$, *it holds w.h.p. that* $\pi(A_t) = 0$ *within*

$$t = \mathrm{O}\left(\log\log n + \frac{\log n}{\log\lambda^{-1}} + \frac{\log n}{\log(\|\pi\|_2\sqrt{\log n})^{-1}}\right) \text{ steps.}$$

Proof of Theorem 7.1.4(ii). Since $\|\pi\|_2 \ge 1/\sqrt{n}$, we have $|\delta(A_0)| = \Omega(\sqrt{\log n/n})$. This implies that Phase (II) takes at most $\mathrm{O}(\log n)$. Thus, we obtain the claim since we can merge Phases (II) to (IV) by taking appropriate constants c_2, c_3 in Phase (III). □

Proof of Theorem 7.1.4(i). Under the assumption of Theorem 7.1.4(i), for any positive constant C, a positive constant C' exists such that $C(\lambda^2 + \|\pi\|_2\sqrt{\log n}) \le C'\frac{\log n}{\sqrt{n}}$. Thus, we can combine Phase (I) and Theorem 7.1.4(ii), and we obtain the claim. □

Proof of Theorem 7.1.6. Combining Phases (II), (III) and (V), we obtain the claim. □

In the rest of this section, we show Phases (I) to (V) of Lemma 7.4.1. For notational convenience, let

$$\alpha := \pi(A),\ \alpha' := \pi(A'),\ \alpha_t := \pi(A_t),$$
$$\delta := \delta(A) = 2\alpha - 1,\ \delta' := \delta(A'),\ \delta_t := \delta(A_t).$$

7.4.1 Phase (I): $0 \le |\delta| \le c_1 \log n/\sqrt{n}$

We invoke Proposition 6.3.1 to show Lemma 7.4.1(I). We begin with showing the following result concerning the growth rate of $|\delta|$ to prove (I) and (II) of Lemma 7.4.1.

Lemma 7.4.2. *Consider a quasi-majority functional voting with respect to* f *on an* n*-vertex* λ*-expander graph with degree distribution* π. *Let* $\epsilon_h(f) := H_f'(1/2) - 1$ *and* $K(f) := \max\{K_2(f), K_2(H_f)\}$ *be positive constants depending only on* f. *Suppose that* $\lambda \le \frac{\epsilon_h(f)}{2K(f)}$. *Then, for any* $A \subseteq V$ *satisfying* $\frac{2K(f)}{\epsilon_h(f)}\lambda^2 \le |\delta| \le \frac{\epsilon_h(f)}{K(f)}$,

$$\mathbf{Pr}\left[|\delta'| \le \left(1 + \frac{\epsilon_h(f)}{8}\right)|\delta|\right] \le 2\exp\left(-\frac{\epsilon_h(f)^2\delta^2}{128\|\pi\|_2^2}\right).$$

Proof. Combining Lemma 7.2.2 and Taylor's theorem, we have

$$\begin{aligned}
\left|\mathbf{E}[\delta'] - H_f'\left(\frac{1}{2}\right)\delta\right| &= 2\left|\mathbf{E}[\alpha'] - \frac{1}{2} - H_f'\left(\frac{1}{2}\right)\left(\alpha - \frac{1}{2}\right)\right| \\
&= 2\left|\mathbf{E}\left[\alpha'\right] - H_f\left(\alpha\right) + H_f\left(\alpha\right) - H_f\left(\frac{1}{2}\right) - H_f'\left(\frac{1}{2}\right)\left(\alpha - \frac{1}{2}\right)\right| \\
&\le 2K_2(f)\lambda\left(|\delta| + \lambda\right)\alpha(1-\alpha) + K_2(H_f)\left(\alpha - \frac{1}{2}\right)^2 \\
&\le \left(\frac{K(f)}{2}\lambda + \frac{K(f)}{4}|\delta|\right)|\delta| + \frac{K(f)}{2}\lambda^2 \qquad (7.15)
\end{aligned}$$

Note that $H_f(1/2) = 1/2$ from the definition. From assumptions of $\lambda \le \frac{\epsilon_h(f)}{2K(f)}$, $|\delta| \le \frac{\epsilon_h(f)}{K(f)}$ and $\lambda^2 \le \frac{\epsilon_h(f)}{2K(f)}|\delta|$, we have

$$\left|H_f'\left(\frac{1}{2}\right)\delta\right| - |\mathbf{E}[\delta']| \le \left|H_f'\left(\frac{1}{2}\right)\delta - \mathbf{E}[\delta']\right| \le \frac{3}{4}\epsilon_h(f)|\delta|.$$

Hence, it holds that

$$\left|\mathbf{E}[\delta']\right| \ge \left|H_f'\left(\frac{1}{2}\right)\delta\right| - \frac{3}{4}\epsilon_h(f)|\delta| = (1+\epsilon_h(f))|\delta| - \frac{3}{4}\epsilon_h(f)|\delta| = \left(1+\frac{\epsilon_h(f)}{4}\right)|\delta|.$$

We observe that, for any $\kappa > 0$,

$$\mathbf{Pr}\left[|\delta'| \le \left|\mathbf{E}[\delta']\right| - \kappa\right] \le 2\exp\left(-\frac{\kappa^2}{2\|\pi\|_2^2}\right) \tag{7.16}$$

from Corollary 5.4.4. Note that $\delta' = \sum_{v\in V}\pi(v)(2X_v - 1)$ for independent indicator random variables $(X_v)_{v\in V}$ (see Definition 5.1.1 for the definition of X_v). Thus,

$$\begin{aligned}\mathbf{Pr}\left[|\delta'| \le \left(1+\frac{\epsilon_h(f)}{8}\right)|\delta|\right] &= \mathbf{Pr}\left[|\delta'| \le \left(1+\frac{\epsilon_h(f)}{4}\right)|\delta| - \frac{\epsilon_h(f)}{8}|\delta|\right] \\ &\le \mathbf{Pr}\left[|\delta'| \le \left|\mathbf{E}[\delta']\right| - \frac{\epsilon_h(f)}{8}|\delta|\right] \le 2\exp\left(-\frac{\epsilon_h(f)^2\delta^2}{128\|\pi\|_2^2}\right)\end{aligned}$$

and we obtain the claim. □

Proof of Lemma 7.4.1(I). We check the conditions (1) and (2) of Proposition 6.3.1 with letting $\Psi(A) = \lfloor n|\delta(A)|\rfloor$ and $m = c_1\sqrt{n}\log n$.

Condition (1). First, we show the following claim that evaluates $\mathbf{Var}[\delta']$.

Claim 7.4.3. *Under the same assumption as Lemma 7.4.1(I),*

$$\frac{\epsilon_{\mathrm{var}}(f)}{n} \le \mathbf{Var}[\delta'] \le \frac{5C_1^2}{n}$$

where $\epsilon_{\mathrm{var}}(f) := f(1/2)(1-f(1/2))$ is a positive constant depending only on f.

Proof of the claim. From Lemma 7.2.3 and assumptions, we have

$$\begin{aligned}\left|\frac{\mathbf{Var}[\delta']}{4} - \|\pi\|_2^2 g\left(\frac{1}{2}\right)\right| = \left|\mathbf{Var}[\alpha'] - \|\pi\|_2^2 g\left(\frac{1}{2}\right)\right| &\le K_1(g)\left(\|\pi\|_2^2\frac{|\delta|}{2} + \|\pi\|_3^{3/2}\lambda\right) \\ &\le \frac{K_1(g)}{n}\left(C_1^2 c_1\frac{\log n}{\sqrt{n}} + C_1\epsilon^{3/2}\right) = \frac{1}{n}\cdot\mathrm{o}(1).\end{aligned}$$

Note that $\mathbf{Var}[\delta'] = \mathbf{Var}[2\alpha' - 1] = 4\,\mathbf{Var}[\alpha']$. Since $\|\pi\|_2^2 \ge 1/n$, we have

$$\frac{\epsilon_{\mathrm{var}}(f)}{n} \le \frac{4\epsilon_{\mathrm{var}}(f) - \mathrm{o}(1)}{n} \le \mathbf{Var}[\delta'] \le \frac{4C_1^2 + \mathrm{o}(1)}{n} \le \frac{5C_1^2}{n}.$$

□

From Corollary 5.4.9 with letting $Y_v = \pi(v)(2X_v - 1)$, we have

$$\begin{aligned}\mathbf{Pr}\left[|\delta'| \le x\sqrt{\frac{\epsilon_{\mathrm{var}}(f)}{n}}\right] \le \mathbf{Pr}\left[|\delta'| \le x\sqrt{\mathbf{Var}[\delta']}\right] &\le \Phi(x) + \frac{5.6\|\pi\|_3^3}{\mathbf{Var}[\delta']^{3/2}} \\ &\le \Phi(x) + 5.6\frac{\epsilon^3}{n^{3/2}}\cdot\frac{n^{3/2}}{\epsilon_{\mathrm{var}}(f)^{3/2}} = \Phi(x) + \mathrm{o}(1)\end{aligned} \tag{7.17}$$

for any $x \in \mathbb{R}$, where $\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}\mathrm{e}^{-y^2/2}\mathrm{d}y$. Thus, for any constant $h > 0$, there exists some constant $C > 0$ such that

$$\mathbf{Pr}[\Psi(A') < h\sqrt{n} \mid \Psi(A) \le m] < C,$$

which verifies the condition (1).

Condition (2). Set $h = \frac{2K(f)}{\epsilon_h(f)} C_1^2$ and assume $h\sqrt{n} \le \Psi(A) < m$. Then

$$\frac{2K(f)}{\epsilon_h(f)}\lambda^2 n \le \frac{2K(f)}{\epsilon_h(f)} C_1^2 \sqrt{n} = h\sqrt{n} \le \Psi(A) \le |\delta| n = \mathrm{o}(n).$$

Thus, we can apply Lemma 7.4.2 and positive constants γ, C exist such that, for any $h\sqrt{n} \le \Psi(A) \le c_1 \sqrt{n}\log n$,

$$\mathbf{Pr}[\Psi(A') < (1+\gamma)\Psi(A)] < \exp\left(-C\frac{\Psi(A)^2}{n}\right).$$

Note that $\|\pi\|_2^2 = \Theta(1/n)$ from the assumption. This verifies the condition (2).

Thus, we can apply Proposition 6.3.1 and we obtain the claim. □

7.4.2 Phase (II): $\frac{2\max\{K(f),8\}}{\epsilon_h(f)} \max\{\lambda^2, \|\pi\|_2\sqrt{\log n}\} \le |\delta| \le \frac{\epsilon_h(f)}{K(f)}$

Proof of Lemma 7.4.1(II). Since $|\delta| \ge \frac{16}{\epsilon_h(f)}\|\pi\|_2\sqrt{\log n}$ from assumptions, applying Lemma 7.4.2 yields

$$\mathbf{Pr}\left[|\delta'| \le \left(1 + \frac{\epsilon_h(f)}{8}\right)|\delta|\right] \le \frac{2}{n^2}.$$

Thus, it holds with probability larger than $(1-2/n^2)^t$ that $|\delta_t| \ge \left(1+\frac{\epsilon_h(f)}{8}\right)^t |\delta_0|$ and we obtain the claim by substituting $t = \mathrm{O}(\log|\delta_0|^{-1})$. □

7.4.3 Phase (III): $0 < c_2 \le \alpha \le c_3 < 1/2$

Proof of Lemma 7.4.1(III). We first observe that, for any $\kappa > 0$,

$$\mathbf{Pr}\left[|\alpha' - \mathbf{E}[\alpha']| \ge \kappa\|\pi\|_2\sqrt{\log n}\right] \le 2n^{-2\kappa} \tag{7.18}$$

from Proposition 5.4.3. Note that $\alpha' = \sum_{v\in V}\pi(v)X_v$ for independent indicator random variables $(X_v)_{v\in V}$. Hence, applying Lemma 7.2.2 yields

$$|\alpha' - H_f(\alpha)| \le |\alpha' - \mathbf{E}[\alpha']| + |\mathbf{E}[\alpha'] - H_f(\alpha)| \le \|\pi\|_2\sqrt{\log n} + \frac{K_2(f)}{4}(|\delta|+\lambda)\lambda \tag{7.19}$$

with probability larger than $1-2/n^2$. Then, for any $\alpha \in [c_2, c_3]$, it holds with probability larger than $1-2/n^2$ that

$$\alpha' \le H_f(\alpha) + \frac{K(f)}{2}\lambda + \|\pi\|_2\sqrt{\log n} \le \alpha - \epsilon(f) + \frac{\epsilon(f)}{4} + \frac{\epsilon(f)}{4} \le \alpha - \frac{\epsilon(f)}{2}.$$

Thus, for $\alpha_0 \in [c_2, c_3]$, $\alpha_t \le c_2$ within $t = 2(c_3 - c_2)/\epsilon(f) = \mathrm{O}(1)$ steps w.h.p. □

7.4.4 Phase (IV): $0 \le \alpha \le \frac{\epsilon_c(f)}{8K(f)}$

We show the following lemma which is useful for proving (IV) and (V) of Lemma 7.4.1.

Lemma 7.4.4. *Let $\epsilon \in (0,1]$ be an arbitrary constant. Consider functional voting on an n-vertex connected graph with degree distribution π. Suppose that, for some $\alpha_* \in [0,1]$ and $K \in [0, 1-\epsilon]$,*

$$\mathbf{E}[\alpha'] \le K\alpha$$

for any $A \subseteq V$ satisfying $\alpha \le \alpha_*$ and $\|\pi\|_2 \le \frac{\epsilon\alpha_*}{2\sqrt{\log n}}$. Then, for any $A_0 \subseteq V$ satisfying $\alpha_0 \le \alpha_*$, $\alpha_t = 0$ w.h.p. within $\mathrm{O}\left(\frac{\log n}{\log K^{-1}}\right)$ steps.

Proof. For any $\alpha \le \alpha_*$, from (7.18) and assumptions of $\mathbf{E}[\alpha'] \le \alpha$ and $\|\pi\|_2 \le \frac{\epsilon\alpha_*}{2\sqrt{\log n}}$, it holds with probability larger than $1-2/n^4$ that

$$\alpha' \le \mathbf{E}[\alpha'] + 2\|\pi\|_2\sqrt{\log n} \le K\alpha + \epsilon\alpha_* \le (1-\epsilon)\alpha_* + \epsilon\alpha_* = \alpha_*.$$

Thus, for any $\alpha_0 \le \alpha_*$, we have

$$\begin{aligned}\mathbf{E}[\alpha_t] &= \sum_{x \le a_*} \mathbf{E}\left[\alpha_t | \alpha_{t-1} = x\right] \mathbf{Pr}\left[\alpha_{t-1} = x\right] + \sum_{x > a_*} \mathbf{E}\left[\alpha_t | \alpha_{t-1} = x\right] \mathbf{Pr}\left[\alpha_{t-1} = x\right] \\ &\le \sum_{x \le a_*} Kx \, \mathbf{Pr}\left[\alpha_{t-1} = x\right] + \mathbf{Pr}\left[\alpha_{t-1} > a_*\right] \le K \, \mathbf{E}[\alpha_{t-1}] + \frac{2t}{n^4} \\ &\le \cdots \le K^t \alpha_0 + \frac{2t^2}{n^4} \le K^t + \frac{2t^2}{n^4}.\end{aligned}$$

This implies that, $\mathbf{E}[\alpha_t] = \mathrm{O}(n^{-3})$ within $t = \mathrm{O}\left(\frac{\log n}{\log K^{-1}}\right)$ steps. Let $\pi_{\min} := \min_{v \in V} \pi(v) \ge 1/(2|E|) \ge 1/n^2$. Markov inequality yields

$$\mathbf{Pr}[\alpha_t = 0] = 1 - \mathbf{Pr}[\alpha_t \ge \pi_{\min}] \ge 1 - \frac{\mathbf{E}[\alpha_t]}{\pi_{\min}} = 1 - \mathrm{O}(1/n)$$

and we obtain the claim. □

Proof of Lemma 7.4.1 of (IV). Combining Lemma 7.2.2 and Taylor's theorem,

$$\begin{aligned}\left|\mathbf{E}[\alpha'] - H_f'(0)\alpha\right| &= \left|\mathbf{E}[\alpha'] - H_f(\alpha) + H_f(\alpha) - H_f(0) - H_f'(0)(\alpha - 0)\right| \\ &\le K_2(f)\lambda\left(|\delta| + \lambda\right)\alpha(1-\alpha) + \frac{K_2(H_f)}{2}\alpha^2 \\ &\le 2K(f)\lambda\alpha + \frac{K(f)}{2}\alpha^2. \end{aligned} \tag{7.20}$$

Hence, for any $\alpha \le \frac{\epsilon_c(f)}{8K(f)}$, we have

$$\begin{aligned}\mathbf{E}[\alpha'] &\le \left(H_f'(0) + 2K(f)\lambda + \frac{K(f)}{2}\alpha\right)\alpha \\ &\le \left(1 - \epsilon_c(f) + \frac{\epsilon_c(f)}{4} + \frac{\epsilon_c(f)}{4}\right)\alpha = \left(1 - \frac{\epsilon_c(f)}{2}\right)\alpha.\end{aligned}$$

Letting $\epsilon = \epsilon_c(f)/2$, $K = 1 - \epsilon_c(f)/2$ and $\alpha_* = \frac{\epsilon_c(f)}{8K(f)}$, from the assumption, $\|\pi\|_2 \le \frac{\epsilon_c(f)^2}{32K(f)\sqrt{\log n}} = \frac{\epsilon\alpha_*}{2\sqrt{\log n}}$. Thus, we can apply Lemma 7.4.4 and we obtain the claim. □

7.4.5 Phase (V): $H_f'(0) = 0$ and $0 \le \alpha \le \frac{1}{7K(f)}$

Proof of Lemma 7.4.1(V). In this case, from (7.20),

$$\mathbf{E}[\alpha'] \le 2K(f)\lambda\alpha + \frac{K(f)}{2}\alpha^2. \tag{7.21}$$

We consider the following two cases.

Case 1. $\max\left\{\lambda, \sqrt{\frac{\|\pi\|_2\sqrt{\log n}}{K(f)}}\right\} \le \alpha \le \frac{1}{7K(f)}$**:** In this case, combining (7.18) and (7.21), it holds with probability larger than $1 - 2/n^2$ that

$$\alpha' \le \left(\frac{2K(f)\lambda}{\alpha} + \frac{K(f)}{2} + \frac{\|\pi\|_2\sqrt{\log n}}{\alpha^2}\right)\alpha^2 \le \frac{7K(f)}{2}\alpha^2.$$

Applying this inequality iteratively, for any $\alpha_0 \le 7K(f)^{-1}$,

$$\alpha_t \le \frac{7K(f)}{2}\alpha_{t-1}^2 \le \cdots \le \frac{2}{7K(f)}\left(\frac{7K(f)}{2}\alpha_0\right)^{2^t} \le \frac{2}{7K(f)2^{2^t}}.$$

holds with probability larger than $(1 - 2/n^2)^t$. This implies that, within $t = \mathrm{O}(\log\log n)$ steps, $\alpha_t \le \max\left\{\lambda, \sqrt{\frac{\|\pi\|_2\sqrt{\log n}}{K(f)}}\right\}$ w.h.p. Note that $\max\left\{\lambda, \sqrt{\frac{\|\pi\|_2\sqrt{\log n}}{K(f)}}\right\} \ge \sqrt{\frac{\|\pi\|_2\sqrt{\log n}}{K(f)}} \ge \sqrt{\frac{\sqrt{\log n/n}}{K(f)}}$ since $\|\pi\|_2^2 \ge 1/n$.

Case 2. $\alpha \le \max\left\{\lambda, \sqrt{\frac{\|\pi\|_2\sqrt{\log n}}{K(f)}}\right\}$**:** Set $\alpha_* = \max\left\{\lambda, \sqrt{\frac{\|\pi\|_2\sqrt{\log n}}{K(f)}}\right\} \ge \sqrt{\frac{\|\pi\|_2\sqrt{\log n}}{K(f)}}$, $K = \frac{5K(f)}{2}\lambda + \frac{1}{2}\sqrt{K(f)\|\pi\|_2\sqrt{\log n}}$ and $\epsilon = 1/4$. Then, from $\lambda \le \frac{1}{10K(f)}$ and $\|\pi\|_2 \le \frac{1}{64K(f)\sqrt{\log n}}$,

$$\begin{aligned}
\|\pi\|_2 = (\sqrt{\|\pi\|_2})^2 &\le \frac{\sqrt{\|\pi\|_2}}{8\sqrt{K(f)\sqrt{\log n}}} = \sqrt{\frac{\|\pi\|_2\sqrt{\log n}}{K(f)}}\frac{\epsilon}{2\sqrt{\log n}} \le \frac{\epsilon\alpha_*}{2\sqrt{\log n}},\\
K &\le \frac{1}{2} + \frac{1}{16} \le 1 - \epsilon,\\
\mathbf{E}[\alpha'] &\le \left(2K(f)\lambda + \frac{K(f)}{2}\alpha\right)\alpha \le \left(2K(f)\lambda + \frac{K(f)}{2}\lambda + \frac{1}{2}\sqrt{K(f)\|\pi\|_2\sqrt{\log n}}\right)\alpha = K\alpha.
\end{aligned}$$

$K \le 1/4 + 1/2 = 3/4$. Thus, applying Lemma 7.4.4, we obtain the claim. □

7.5 Lower Bound of Consensus Time

This section is devoted to prove Theorem 7.1.5. In particular, we show the following theorem.

Theorem 7.5.1. *Let $C > 0$ be an arbitrary constant. Consider a quasi-majority functional voting with respect to f on an n-vertex λ-expander graph with degree distribution π. Suppose that $\max\{\lambda, \|\pi\|_2\} \le n^{-C}$. Then, for any $A \subseteq V$ satisfying $|\delta(A)| \le n^{-C}$, $T_{\mathrm{cons}}(A) = \Omega(\log n)$ w.h.p.*

Proof of Theorem 7.5.1. From (7.15),

$$\begin{aligned}
|\mathbf{E}[\delta']| &\le H'_f\left(\frac{1}{2}\right)|\delta| + \left(\frac{K(f)}{2}\lambda + \frac{K(f)}{4}|\delta|\right)|\delta| + \frac{K(f)}{2}\lambda^2\\
&\le \left(1 + \epsilon_h(f) + \frac{3K(f)}{4}\right)|\delta| + K(f)\lambda^2.
\end{aligned}$$

Recall that $\delta' = \sum_{v\in V}(2\pi_v - 1)$ for independent indicator random variables $(X_v)_{v\in V}$ Definition 5.1.1. Thus, for any $\kappa > 0$,

$$\mathbf{Pr}\left[|\delta'| \ge |\mathbf{E}[\delta']| + \kappa\right] \le \exp\left(-\frac{\kappa^2}{2\|\pi\|_2^2}\right)$$

from Corollary 5.4.4. Hence, it holds with probability larger than $1 - 2/n^2$ that

$$|\delta'| \le c|\delta| + K(f)\lambda^2 + 2\|\pi\|_2\sqrt{\log n},$$

where we put $c := 1 + \epsilon_h(f) + \frac{3K(f)}{4} > 1$. Then, applying this inequality iteratively with $t = (C/2)\log_c n$ steps,

$$\begin{aligned}
|\delta_t| &\le c|\delta_{t-1}| + K(f)\lambda^2 + 2\|\pi\|_2\sqrt{\log n}\\
&\le \cdots \le c^t|\delta_0| + tc^t\left(K(f)\lambda^2 + 2\|\pi\|_2\sqrt{\log n}\right)\\
&\le \frac{n^{C/2}}{n^C} + n^{C/2}\log_c n^{C/2}\left(\frac{K(f)}{n^{2C}} + \frac{2\sqrt{\log n}}{n^C}\right) = \mathrm{o}(1)
\end{aligned}$$

w.h.p., and we obtain the claim. Note that we use our assumptions of $|\delta_0|, \max\lambda, \|\pi\|_2 \le n^{-C}$ in the last inequality. □

7.6 Best-of-$(2k+1)$ on Expander Graphs

We show Theorem 7.1.8. The proof is almost same as the one given in Section 7.4 but we need some special care. We assume $k = \omega(1)$ and thus k is sufficiently large. Consider the best-of-$(2k+1)$ on an n-vertex λ-expander graph with degree distribution π. Suppose that the graph satisfies the conditions of Theorem 7.1.8. Let $A_0, A_1, \ldots,$ be the sequence given by the best-of-$(2k+1)$ with initial configuration $A_0 \subseteq V$. For notational convenience, let

$$\begin{aligned}
&\alpha := \pi(A),\ \alpha' := \pi(A'),\ \alpha_t := \pi(A_t),\\
&\delta := \delta(A) = 2\alpha - 1,\ \delta' := \delta(A'),\ \delta_t := \delta(A_t).
\end{aligned}$$

The dynamics of best-of-$(2k+1)$ are divided into four phases. More specifically, we prove the following key result that corresponds to Lemma 7.4.1.

Lemma 7.6.1. *Consider the best-of-$(2k+1)$ on an n-vertex λ-expander graph with degree distribution π. Suppose that the graph satisfies the conditions of Theorem 7.1.8. Then, the following hold:*

(I) For any $A_0 \subseteq V$ satisfying $|\delta_0| \le 300C\log n/\sqrt{n}$, $|\delta_t| \ge 300C\log n/\sqrt{n}$ within $t = \mathrm{O}(\log n/\log k)$ steps w.h.p.

(II) For any $A_0 \subseteq V$ satisfying $|\delta_0|$ satisfying $300C\log n/\sqrt{n} \le |\delta_0| \le \frac{1.25}{\sqrt{k}}$, $|\delta_t| > \frac{1.25}{\sqrt{k}}$ within $t = \mathrm{O}(\log n/\log k)$ steps w.h.p. Moreover, there exists $A_0 \subseteq V$ satisfying $300C\log n/\sqrt{n} \le |\delta_0| \le \frac{1.25}{\sqrt{k}}$ such that $|\delta_t| \le \frac{1.25}{\sqrt{k}}$ w.h.p. for any $t = \mathrm{o}(\log n/\log k)$

(III) For any $A_0 \subseteq V$ satisfying $\frac{1.25}{\sqrt{k}} \le |\delta_0| \le 0.9$, $|\delta_1| > 0.9$ w.h.p.

(IV) For any $A_0 \subseteq V$ satisfying $0.9 \le |\delta_0| < 1$, $|\delta_t| = 1$ (or equivalently, the voting process reaches consensus) within $t = \mathrm{O}(\log n/\log k)$ steps w.h.p.

Proof of Theorem 7.1.8 using Lemma 7.6.1. Theorem 7.1.8 is straightforward from Lemma 7.6.1. For any initial configuration $A_0 \subseteq V$, A_0 satisfies one of (I) to (IV). If A_0 satisfies (IV), the consensus time is $\mathrm{O}(\log n/\log k)$. Otherwise, from Lemma 7.6.1, for some $t = \mathrm{O}(\log n/\log k)$, A_t satisfies $|\delta(A_t)| > 0.9$ and then apply Lemma 7.6.1(IV).

Note that, for an initial configuration A_0 satisfying (II), we have $T_{\mathrm{cons}}(A_0) = \Omega(\log n/\log k)$ w.h.p. □

The rest of this section is devoted to prove Lemma 7.6.1. We begin with preparing useful facts concerning with best-of-$(2k+1)$. Let f_{2k+1} be the betrayal function of best-of-$(2k+1)$. Then, we have

$$\begin{aligned}
&1.05\sqrt{k} \le (2k+1)\binom{2k}{k}4^{-k} = \left|f'_{2k+1}\left(\frac{1}{2}\right)\right| \le 2\sqrt{k},\\
&|f'_{2k+1}(x)| \le \left|f'_{2k+1}\left(\frac{1}{2}\right)\right| \le \frac{3}{\sqrt{\pi}}\sqrt{k} \le 2\sqrt{k},\\
&|f''_{2k+1}(x)| \le \left|f''_{2k+1}\left(\frac{1}{2}+\frac{1}{2\sqrt{2k-1}}\right)\right| < 1.6k
\end{aligned}$$

for sufficiently large k. Here, we used $\frac{4^k}{\sqrt{\pi k}}\left(1-\frac{1}{8k}\right) \le \binom{2k}{k} \le \frac{4^k}{\sqrt{\pi k}}$.

From Lemmas 7.2.3 and 7.3.3 (note that $f_{2k+1}(x)$ satisfies $f_{2k+1}(x)+f_{2k+1}(1-x)=1$), it holds for all $A \subseteq V$ that

$$\left|\mathbf{E}[\alpha'] - f_{2k+1}(\alpha)\right| \le 0.8k\lambda^2\alpha(1-\alpha) \tag{7.22}$$

$$\left|\mathbf{Var}[\alpha'] - g_{2k+1}(1/2)\|\pi\|_2^2\right| \le 2\sqrt{k}\left(\frac{\|\pi\|_2^2}{2}|\delta| + \lambda\|\pi\|_3^{3/2}\right), \tag{7.23}$$

where $g_{2k+1}(x) = f(x)(1-f(x))$. Note that $g'_{2k+1}(x) = f'_{2k+1}(x)(1-2f_{2k+1}(x))$ satisfies $|g'_{2k+1}(x)| \le |f'_{2k+1}(x)| \le 2\sqrt{k}$. Thus, from the Hoeffding bound (Proposition 5.4.3), it holds w.h.p. that

$$\begin{aligned}
|\alpha' - f_{2k+1}(\alpha)| &\le 0.8k\lambda^2\alpha(1-\alpha) + \|\pi\|_2\sqrt{\log n}\\
&\le 0.8k\lambda^2\alpha(1-\alpha) + \sqrt{\frac{C\log n}{n}}.
\end{aligned} \tag{7.24}$$

On the other hand, it is routine to check the following facts.

$$\lambda\sqrt{k\|\pi\|_3^3} = \mathrm{o}(n^{-1}), \tag{7.25}$$

$$\frac{1}{n} \le \|\pi\|_2^2 \le \frac{C}{n}, \tag{7.26}$$

$$k\lambda^2 = \mathrm{O}(1/\sqrt{n}). \tag{7.27}$$

We begin with proving the following result.

Lemma 7.6.2. *There exist constants $h, c, C, C' > 0$ such that the following hold for any $A \subseteq V$ satisfying $h/\sqrt{nk} \le |\delta| \le 1.25/\sqrt{k}$:*

$$\mathbf{Pr}[|\delta'| < 0.025\sqrt{k}|\delta|] \le \exp\left(-ckn\delta^2\right),$$

and

$$\mathbf{Pr}[|\delta'| > C\sqrt{k}|\delta|] \le \exp(-C'kn\delta^2).$$

Proof. Let h be a sufficiently large constant and let $A \subseteq V$ be a configuration satisfying $\frac{h}{\sqrt{kn}} \le |\delta| \le \frac{1.25}{\sqrt{k}}$. From (7.22), (7.27) and Taylor's theorem, we have

$$\begin{aligned}
|\mathbf{E}[\delta']| &\ge \left|2f_{2k+1}\left(\frac{1}{2}+\frac{\delta}{2}\right)-1\right| - 0.8k\lambda^2\alpha(1-\alpha) \\
&\ge f'_{2k+1}\left(\frac{1}{2}\right)|\delta| - \max_{0\le z\le 1}|f''_{2k+1}(z)|\frac{\delta^2}{2} - 0.2k\lambda^2 \\
&\ge 0.05\sqrt{k}|\delta| + (\sqrt{k}|\delta| - 0.8k\delta^2) - 0.2k\lambda^2 \\
&\ge 0.05\sqrt{k}|\delta| + \frac{0.01h}{\sqrt{n}} - 0.2k\lambda^2 \\
&\ge 0.05\sqrt{k}|\delta|.
\end{aligned}$$

In the fourth inequality, note that $\sqrt{k}|\delta| \ge 0.8k\delta^2$ holds if $|\delta| \le 1.25/\sqrt{k}$. In the last inequality, we used $\lambda = \mathrm{O}(k^{-0.5}n^{-0.25})$ and thus $k\lambda^2 = \mathrm{O}(1/\sqrt{n}) \le 0.01h/\sqrt{n}$ for sufficiently large constant h. Then, from Corollary 5.4.4, we have

$$\begin{aligned}
\mathbf{Pr}[|\delta'| < 0.025\sqrt{k}|\delta|] &\le \mathbf{Pr}\left[|\delta'| < 0.5|\,\mathbf{E}[\delta']|\right] \\
&\le 2\exp\left(-\frac{0.5|\,\mathbf{E}[\delta']|^2}{\|\pi\|_2^2}\right) \\
&\le \exp\left(-ckn\delta^2\right)
\end{aligned}$$

for some suitable constant $c > 0$. In the last inequality, we used (7.26).

Similarly, we obtain

$$\begin{aligned}
|\mathbf{E}[\delta']| &\le \left|2f_{2k+1}\left(\frac{1}{2}+\frac{\delta}{2}\right)-1\right| + 0.8k\lambda^2\alpha(1-\alpha) \\
&\le f'_{2k+1}\left(\frac{1}{2}\right)|\delta| + \max_{0\le z\le 1}|f''_{2k+1}(z)|\frac{\delta^2}{2} + 0.2k\lambda^2 \\
&\le 2\sqrt{k}|\delta| + 0.8k\delta^2 + 0.2k\lambda^2 \\
&= \mathrm{O}(\sqrt{k}|\delta|)
\end{aligned}$$

and thus, from Corollary 5.4.4, two constants C and C' exist such that

$$\mathbf{Pr}[|\delta'| > C\sqrt{k}|\delta|] \le \exp(-C'kn\delta^2).$$

□

7.6.1 Phase (I): $0 \le |\delta| \le 300C\log n/\sqrt{n}$

In this part, we show Lemma 7.6.1(I). The proof is almost same as that of Lemma 7.4.1(I) that is presented in Section 7.4.1. The difference is that we use the following result, which is a slight modification of Proposition 6.3.1.

Lemma 7.6.3 (Modification of Proposition 6.3.1)**.** *Consider a Markov chain $(X_t)_{t=1}^{\infty}$ with finite state space Ω and a function $\Psi : \Omega \to [0, n]$. Let C_1 be an arbitrary constant and $m = C_1\sqrt{n}\log n$. Let $k = k(n)$ be a function such that $k(n) \to \infty$ as $n \to \infty$. Suppose that Ω, Ψ and m satisfies the following conditions:*

(*i*) *For any positive constant h, there exists a positive constant $C_2 < 1$ such that*

$$\mathbf{Pr}\left[\Psi(X_{t+1}) < h\sqrt{\frac{n}{k}} \,\middle|\, \Psi(X_t) \le m\right] < \frac{C_2}{\sqrt{k}}.$$

(*ii*) *Three positive constants C_3, C_4 and h exist such that, for any $x \in \Omega$ satisfying $h\sqrt{n/k} \le \Psi(x) < m$,*

$$\mathbf{Pr}\left[\Psi(X_{t+1}) < C_3\sqrt{k}\Psi(X_t) \,\middle|\, X_t = x\right] < \exp\left(-C_4\frac{k\Psi(x)^2}{n}\right).$$

Then, $\Psi(X_t) \ge m$ holds w.h.p. for some $t = \mathrm{O}(\log n/\log k)$.

We prove Lemma 7.6.3 in Section 7.6.5. Lemma 7.6.1(I) is immediate from Lemma 7.6.3 with letting $\Psi(A) = n|\delta|$ and $C_1 = 300C$. Hence, it suffices to verify the conditions (*i*) and (*ii*).

Condition (*i*). First we evaluate the variance $\mathbf{Var}[\delta']$.

Claim 7.6.4. *Under the same assumption as that of Lemma 7.6.1(I),*

$$\mathbf{Var}[\delta'] \ge \frac{0.99}{n}.$$

Proof of the claim. Note that $\mathbf{Var}[\delta'] = 4\,\mathbf{Var}[\alpha']$. From (7.23), we can evaluate the variance $\mathbf{Var}[\alpha']$ as follows:

$$\begin{aligned}
\mathbf{Var}[\alpha'] &\ge g_{2k+1}(1/2)\|\pi\|_2^2 - \sqrt{k}|\delta|\|\pi\|_2^2 - 2\sqrt{k}\lambda\|\pi\|_3^{3/2} \\
&\ge \frac{1}{4n} - 3\sqrt{k}\left(300C^2\sqrt{\frac{\log n}{n^3}} + \lambda\|\pi\|_3^{3/2}\right) \quad \text{(from (7.26))} \\
&= \frac{1-\mathrm{o}(1)}{4n} \quad \text{(since } k = \mathrm{o}(\log n/n)\text{) and (7.25))} \\
&\ge \frac{0.99}{4n}.
\end{aligned}$$

□

From Corollary 5.4.9, for any positive real x, we have

$$\mathbf{Pr}\left[|\delta'| \le x\sqrt{\frac{0.99}{n}}\right] \le \Phi(x) + \frac{5.6\|\pi\|_3^3}{\mathbf{Var}[\delta']^{3/2}} = \Phi(x) + \mathrm{o}(1),$$

where $\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} \mathrm{e}^{-y^2/2}\mathrm{d}y$ (see (7.17)). This yields the condition (*i*).

Condition (*ii*). This condition directly follows Lemma 7.6.2 by substituting $|\delta| = \frac{\Psi(A)}{n}$.

7.6.2 Phase (II): $300C\log n/\sqrt{n} \le |\delta| \le 1.25/\sqrt{k}$

Since $|\delta| \ge 300C\log n/\sqrt{n}$, from Lemma 7.6.2, we have

$$\mathbf{Pr}[|\delta'| < 0.025\sqrt{k}|\delta|] \le \exp\left(-\frac{ckn(\log n)^2}{n}\right) \le n^{-2}$$

if k is sufficiently large. Thus, we have $|\delta_t| \ge (0.025\sqrt{k})^t \cdot 300C\log n/\sqrt{n}$ with probability $(1-n^{-2})^t$. Therefore, for some $t = \mathrm{O}(\log n/\log k)$, $|\delta_t| \ge 1.25/\sqrt{k}$ with probability $1 - n^{-1.9}$.

Now we show the lower bound that $t = \Omega(\log n/\log k)$. We assume that $k = n^{\mathrm{o}(1)}$ (otherwise, $\log n/\log k$ is a constant and the lower bound is trivial). From Lemma 7.6.2, we have $|\delta'| = \mathrm{O}(\sqrt{k}|\delta|)$ w.h.p. Suppose that $|\delta| = 300C\log n/\sqrt{n}$. If $t = \mathrm{o}(\log n/\log k)$, we have $|\delta_t| \le n^{\mathrm{o}(1)}\mathrm{O}(\log n/\sqrt{n}) \le \mathrm{O}(n^{-1/3}) \le 1.25/\sqrt{k}$ if $k = n^{\mathrm{o}(1)}$. This implies that it requires $\Omega(\log n/\log k)$ steps to reach a configuration that $|\delta| \ge 1.25/\sqrt{k}$.

7.6.3 Phase (III): $1.25/\sqrt{k} < |\delta| \leq 0.9$

We may assume that $\delta \geq 0$ without loss of generality (otherwise, consider A^c). From (7.24), we have

$$\begin{aligned}\delta' &\geq 2f_{2k+1}\left(\frac{1}{2}+\frac{\delta}{2}\right) - 1 - 0.4k\lambda^2 - 2\sqrt{\frac{K\log n}{n}} \\ &\geq 2f_{2k+1}\left(\frac{1}{2}+\frac{\delta}{2}\right) - 1 - \mathrm{o}(1).\end{aligned}$$

We claim that $2f_{2k+1}\left(\frac{1}{2}+\frac{\delta}{2}\right) - 1 > 0.9$ during this phase (for sufficiently large n and k). Let $\mathrm{Bin}(N,p)$ denote the random variable of the binomial distribution with N trials and probability p. Then, from the definition of f_{2k+1}, it holds that

$$\begin{aligned}f_{2k+1}\left(\frac{1}{2}+\delta\right) &= \mathbf{Pr}\left[\mathrm{Bin}\left(2k+1,\frac{1}{2}+\delta\right) \geq k+1\right] \\ &= 1 - \mathbf{Pr}\left[\mathrm{Bin}\left(2k+1,\frac{1}{2}+\delta\right) \leq k\right]. \end{aligned} \tag{7.28}$$

Let $\mu = (2k+1)(1/2+\delta)$ be the expectation of $\mathrm{Bin}(2k+1, 1/2+\delta)$. Then, since $\mu - k \geq 2k\delta$, we have

$$\begin{aligned}\mathbf{Pr}\left[\mathrm{Bin}\left(2k+1,\frac{1}{2}+\delta\right) \leq k\right] &\leq \mathbf{Pr}\left[\mathrm{Bin}\left(2k+1,\frac{1}{2}+\delta\right) \leq \mu - (\mu-k)\right] \\ &\leq \mathbf{Pr}\left[\mathrm{Bin}\left(2k+1,\frac{1}{2}+\delta\right) \leq \mu - 2k\delta\right] \\ &\leq \exp(-2k\delta^2). \end{aligned} \tag{7.29}$$

In the third inequality, we applied the Hoeffding bound (Proposition 5.4.3). If $\delta \geq \frac{1.25}{\sqrt{k}}$, by combining (7.28) and (7.29), we obtain

$$\begin{aligned}f_{2k+1}\left(\frac{1}{2}+\delta\right) &\geq 1 - \exp(-2k\delta^2) \\ &\geq 1 - \mathrm{e}^{-3.125} \\ &> 0.92.\end{aligned}$$

Thus, from (7.24), $\delta' \geq 0.92 - \mathrm{o}(1) > 0.9$ holds w.h.p.

7.6.4 Phase (IV): $0.9 < |\delta| \leq 1$

We may assume $\pi(A_0) \leq 0.1$ without loss of generality. We claim that $\pi(A_t) < \frac{1}{n^2}$ for some $t = \mathrm{O}(\log n/\log k)$, which implies $A_t = \emptyset$ (since $\pi(S) \geq \frac{1}{2m} \geq \frac{1}{n^2}$ whenever $S \neq \emptyset$).

Observe that

$$\begin{aligned}f_{2k+1}(x) &= \sum_{i=k+1}^{2k+1} \binom{2k+1}{i} x^i (1-x)^{2k+1-i} \\ &\leq (4x)^k \\ &\leq \frac{x}{4k}\end{aligned}$$

whenever $x \leq 0.1 \leq 4^{-1}(16k)^{1/(k-1)}$ with $k \geq 2$. Therefore, from (7.22), we have

$$\mathbf{E}[\alpha'] \leq \left(\frac{1}{4k} + 0.4k\lambda^2\right)\alpha.$$

From (7.24) and the upper bound of $\mathbf{E}[\alpha']$ above, it holds with probability $1 - \mathrm{O}(n^{-3})$ that $\alpha' \leq 0.9$ conditioned on $\alpha \leq 0.1$. Thus,

$$\mathbf{E}[\pi(A_t)] \leq \left(\frac{1}{4k} + 0.4k\lambda^2\right)^t + n^{-3+\mathrm{o}(1)} \leq n^{-3+\mathrm{o}(1)}$$

for some $t = \mathrm{O}(\log n/\log k + \log n/\log \lambda^{-1}) = \mathrm{O}(\log n/\log k)$ (note that $\lambda^{-1} = \Omega(n^{1/4})$ from (7.27)). For this t, we have $\mathbf{Pr}[A_t \neq \emptyset] \leq \mathbf{Pr}[\pi(A_t) \geq n^{-2}] \leq n^2\,\mathbf{E}[\pi(A_t)] = \mathrm{O}(n^{-1})$. This completes the proof of Lemma 7.6.1 as well as Theorem 7.1.8.

7.6.5 Symmetry breaking lemma for best-of-k

The proof is essentially given in [CGG+18]. By inspecting the proof of [CGG+18] with evaluating constant terms carefully, we obtain Lemma 7.6.3. For completeness, let us present it here.

Let $m = C_1\sqrt{n}\log n$. Let $\tau = \inf\{t \in \mathbb{N} : \Psi(X_t) \geq m\}$ and $\{\tau(i)\}_{i\in\mathbb{N}}$ be the hitting times defined as

$$\begin{cases} \tau(0) = 0, \\ \tau(i) = \inf_{t\in\mathbb{N}}\{t : \tau(i-1) < t < \tau, f(X_t) \geq h\sqrt{n/k}\}. \end{cases}$$

Let $R_1, R_2, \ldots$ be the sequence of random variables defined as $R_i = X_{\tau(i)}$. It is shown in [CGG+18] that

- The sequence $(R_i)_{i\in\mathbb{N}}$ is a Markov chain.
- The sequence $(R_i)_{i\in\mathbb{N}}$ satisfies

$$\mathbf{Pr}[\Psi(R_{i+1}) < C_3\sqrt{k}\Psi(R_i)|R_i = x] < \exp\left(-C_4\frac{k\Psi(x)^2}{n}\right)$$

  for any $x \in \Omega$ that $h\sqrt{n/k} \leq \Phi(x) < m$.

We claim that $\Psi(R_i) \geq m$ for some $i = \mathrm{O}(\log n/\log k)$. To prove this, we use the Markov inequality. Fix a state $x \in \Omega$ such that $h\sqrt{n/k} \leq \Psi(x) < m$ for a sufficiently large constant h. Let $Y_i = \exp(-\frac{\Psi(R_i)}{\sqrt{n}})$ for each i. Let $y = \exp(-\frac{\Psi(x)}{\sqrt{n}})$ and $z = z(x) = \frac{\sqrt{k}\Psi(x)}{\sqrt{n}} \geq h$ for $x \in \Omega$. Note that $\mathrm{e}^z = y^{-\sqrt{k}}$. Then, we have

$$\begin{aligned}
&\mathbf{E}[Y_{i+1}|R_i = x] \\
&\leq \mathbf{Pr}[\Psi(R_{i+1}) < C_3\sqrt{k}\Psi(x)] + \mathbf{Pr}[\Psi(R_{i+1}) \geq C_3\sqrt{k}\Psi(x)] \cdot \exp\left(-C_3\sqrt{k}\frac{\Psi(x)}{\sqrt{n}}\right) \\
&\leq \exp\left(-C_4\frac{k\Psi(x)^2}{n}\right) + \exp\left(-C_3\frac{\sqrt{k}\Psi(x)}{\sqrt{n}}\right) \\
&= \exp\left(-C_4z^2\right) + \exp\left(-C_3z\right) \\
&= y^{-\frac{C_3}{2}\sqrt{k}}\left(\exp\left(\frac{C_3}{2}z - C_4z^2\right) + \exp\left(-\frac{C_3}{2}z\right)\right) \\
&\leq \frac{1}{2}y^{\frac{C_3}{2}\sqrt{k}} \qquad \text{(since } z \geq h \text{ is sufficiently large and } C_2 > 1\text{)} \\
&\leq \begin{cases} \frac{1}{2} & \text{if } \frac{1}{2} < y_i \leq 1, \\ \frac{y}{C_3\sqrt{k}} & \text{if } y_i \leq \frac{1}{2}. \end{cases}
\end{aligned}$$

In the second part of the last inequality, we assume that $k \geq 2$; hence, it holds that $r^a \leq \frac{r}{a}$ for $0 \leq r \leq \frac{1}{2}$ if $a \geq 2$. Note that for each $i \geq 1$, the random variable $\Psi(R_i) = \Psi(X_{\tau(i)})$ satisfies $h\sqrt{n/k} \leq \Psi(R_i) < m$. Then, we have

$$\mathbf{E}[Y_i] \leq \frac{1}{2}\left(\frac{1}{C_3\sqrt{k}}\right)^{i-2}$$

and thus, by the Markov inequality,

$$\begin{aligned}
\mathbf{Pr}[\Psi(R_i) < m] = \mathbf{Pr}\left[Y_i > \exp\left(-\frac{m}{\sqrt{n}}\right)\right] &\leq \exp\left(\frac{m}{\sqrt{n}}\right)\frac{1}{2}\left(\frac{1}{C_3\sqrt{k}}\right)^{i-2} \\
&= \frac{n^{C_1}}{2(C_3\sqrt{k})^{i-2}} \\
&\leq n^{-1}
\end{aligned}$$

for $i = \lfloor C_5\log n/\log k\rfloor$ for some constant C_5 that depends on C_1 and C_3.

Finally, we consider $\tau(\lfloor C_5 \log n / \log k \rfloor)$. Let $W_0, W_1, \ldots$ be binary random variables defined as

$$W_t = \begin{cases} 1 & \text{if } \Psi(X_t) \geq h\sqrt{\frac{n}{k}}, \\ 0 & \text{otherwise.} \end{cases}$$

Note that $\mathbf{Pr}[\tau(T_1) \geq T_2] = \mathbf{Pr}[\sum_{t=1}^{T_2} W_t \leq T_1]$. Let $\hat{W}_0, \hat{W}_1, \ldots$ be i.i.d. binary random variables such that $\mathbf{E}[\hat{W}_t] = 1 - \frac{C_1}{\sqrt{k}}$. From the condition (*i*), for every T, the sum $\sum_{t=1}^{T} \hat{W}_t$ has stochastic dominance over $\sum_{t=1}^{T} W_t$.Therefore, setting $T_1 = \lfloor \frac{C_4 \log n}{\log k} \rfloor$ and $T_2 = \lceil \frac{2C_4 \log n}{\log k} \rceil$, we obtain

$$\begin{aligned}
\mathbf{Pr}\left[\tau\left(\left\lfloor \frac{C_5 \log n}{\log k} \right\rfloor\right) \geq T_2\right] &= \mathbf{Pr}\left[\sum_{t=1}^{T_2} W_t \leq \left\lfloor \frac{C_5 \log n}{\log k} \right\rfloor\right] \\
&\leq \mathbf{Pr}\left[\sum_{t=1}^{T_2} W_t \leq \frac{C_5 \log n}{\log k}\right] \\
&\leq \mathbf{Pr}\left[\sum_{t=1}^{T_2} \hat{W}_t \leq \frac{C_5 \log n}{\log k}\right] \\
&\leq \mathbf{Pr}\left[\sum_{t=1}^{T_2} (1 - \hat{W}_t) \geq T_2 - \frac{C_5 \log n}{\log k}\right] \\
&\leq \mathbf{Pr}\left[\sum_{t=1}^{T_2} (1 - \hat{W}_t) \geq \frac{C_5 \log n}{\log k}\right] \\
&\leq 2^{T_2} \left(\frac{C_1}{\sqrt{k}}\right)^{\frac{C_5 \log n}{\log k}} \\
&\leq n^{\mathrm{O}(1/\log k) - \frac{C_5}{2}}.
\end{aligned}$$

In the fifth inequality, we used the union bound over the choice for $\hat{W}_t$. Note that $1 - \hat{W}_t = 1$ with probability $\frac{C_1}{\sqrt{k}}$. □

Chapter 8

Random Walk on Growing Networks

8.1 Model and Quantities

In this chapter we introduce a new notion of growing graph and investigate a random walk on it.

8.1.1 Random walk on a growing graph

A *growing graph* is a sequence $\mathcal{G} = (\mathcal{G}_t)_{t\in\mathbb{Z}_{\geq 0}}$ of graphs where each $\mathcal{G}_t = (\mathcal{V}_t, \mathcal{E}_t)$ is a graph such that $\mathcal{V}_t \subseteq \mathcal{V}_{t+1}$. A *random walk on a growing graph* is a stochastic process $Z = (Z_t)_{t\in\mathbb{Z}_{\geq 0}}$ for $Z_t \in \mathcal{V}_t$, where the transition probability from Z_t to Z_{t+1} is provided as a random walk on $\mathcal{G}_t$. Note that $Z_t \in \mathcal{V}_{t-1}$ holds for $t \in \mathbb{N}$.

We are particularly concerned with a simple model of growing graphs with moderate changes. Roughly speaking, a growing graph $\mathcal{G}$ in this chapter keeps being a graph $G^{(n)}$ unchanged for some duration of steps, then changes its shape to $G^{(n+1)}$ by adding a single vertex and connecting it to $G^{(n)}$. Let $\mathfrak{d}\colon \mathbb{N} \to \mathbb{N}$ be a function, denoting the *duration* of keeping the graph unchanged. For $n \in \mathbb{N}$, let $G^{(n)} = (V^{(n)}, E^{(n)})$ be a graph such that $V^{(n)} = \{v_1, \ldots, v_n\}$, $E^{(1)} = \emptyset$, and $E^{(n)} = E^{(n-1)} \cup \bigcup_{u\in S}\{\{v_n, u\}\}$ for some $S \subseteq V^{(n-1)}$. Then, $\mathcal{G}$ is given as $\mathcal{G}_t = G^{(n(t))}$, where $n(t) \in \mathbb{N}$ is the least positive integer satisfying $\mathfrak{d}(i) \leq t < \sum_{i=1}^{n} \mathfrak{d}(i)$. Notice that $\mathcal{G}_0$ is a graph of a single vertex[1]. In other words, $\mathfrak{d}(n)$ denotes the duration of $|\mathcal{V}_t| = n$, and hence $\mathfrak{d}(n) = \min\{t : |\mathcal{V}_t| = n+1\} - \min\{t : |\mathcal{V}_t| = n\}$ holds. For convenience, let $T_n := \sum_{i=1}^{n-1} \mathfrak{d}(i) = \min\{t \geq 0 : |\mathcal{V}_t| = n\}$ for $n \in \mathbb{N}$. For example, if $\mathfrak{d}(n) = n$, then $\mathcal{G}^{(0)} = G^{(1)}$, $\mathcal{G}^{(1)} = \mathcal{G}^{(2)} = G^{(2)}$, $\mathcal{G}^{(3)} = \mathcal{G}^{(4)} = \mathcal{G}^{(5)} = G^{(3)}$, and $T_1 = 0$, $T_2 = 1$, $T_3 = 3$, $T_4 = 6$.

This chapter is also concerned with a particular model of random walks on growing graphs. For simplicity, we assume that a random walk on a growing graph $\mathcal{G}$ is temporarily time-homogeneous, meaning that a random walk is formally represented by a common $n \times n$ transition matrix $P^{(n)}$ such that $\mathbf{Pr}[Z_{t+1} = v \mid Z_t = u] = (P^{(n)})_{u,v}$ when $\mathcal{G}_t = G^{(n)}$. We simply represent a random walk on a growing graph (*RWoGG*, for short) by a triple $R = (\mathfrak{d}, (G^{(n)})_{n\in\mathbb{N}}, (P^{(n)})_{n\in\mathbb{N}})$. Strictly speaking, an RWoGG is specified by a pair $(\mathfrak{d}, (P^{(n)})_{n\in\mathbb{N}})$ and the sequence $(G^{(n)})_{n\in\mathbb{N}}$ is not essential. However, we define an RWoGG as a triple $(\mathfrak{d}, (G^{(n)})_{n\in\mathbb{N}}, (P^{(n)})_{n\in\mathbb{N}})$ in order to emphasize that the random walk takes place on a growing graph.

We are concerned with the number of vertices unvisited by an RWoGG, formally given by

$$\mathcal{U}_t := |\{v \in \mathcal{V}_{t-1} : v \neq Z_s \text{ for any } s \in \{0, 1, \ldots, t\}\}|,$$

where recall the fact that $Z_t \in \mathcal{V}_{t-1}$. Particularly, let $U(n)$ (or simply U without confusion) denote $\mathcal{U}_{T_{n+1}}$, i.e., $U(n) = n - \left|\bigcup_{t=0}^{T_{n+1}} \{Z_t\}\right|$, and we will be concerned with it.

8.2 Our Results

This chapter investigates the behavior of $\mathbf{E}[U]$ regarding $\mathfrak{d}$ for an RWoGG $R = (\mathfrak{d}, (G^{(i)})_{i\in\mathbb{N}}, (P^{(i)})_{i\in\mathbb{N}})$. As a warm-up, we first study the simple random walk on a growing complete graph. Then, we obtain upper bounds of $\mathbf{E}[U]$ for general growing graph in terms of the hitting and mixing time of $G^{(n)}$. Finally, we obtain a lower bound of $\mathbf{E}[U]$ for a growing path. This verifies that our upper bounds for general case is tight for the growing path.

[1] This is just for convenience of descriptions, but not essential in our later analyses.

8.2.1 Complete graph (Section 8.5)

As an introductory example of our analyses, we firstly concerned with a random walk on a growing complete graph. Let $R_{\mathrm{c}} = (\mathfrak{d}, (G^{(i)})_{i\in\mathbb{N}}, (P^{(i)})_{i\in\mathbb{N}})$ be a the simple random walk on a growing *complete* graph, where $G^{(i)}$ is the i-vertex complete graph (with self-loops), and $(P^{(i)})(u,v) = 1/i$ for any $u, v \in V^{(i)}$.

Theorem 8.2.1. *For $R_{\mathrm{c}} = (\mathfrak{d}, (G^{(i)})_{i\in\mathbb{N}}, (P^{(i)})_{i\in\mathbb{N}})$, the following hold:*

(1) If there is a constant $C > 0$ such that $\mathfrak{d}(i) \geq Ci$ for all $i \in [n]$, then $\mathbf{E}[U] = \mathrm{O}(1)$.

(2) If $\mathfrak{d}(i)/i \to \infty$ as $i \to \infty$, then $\mathbf{E}[U] \to 0$ as $n \to \infty$.

(3) If $\mathfrak{d}$ is unbounded (i.e., $\mathfrak{d}(i) \to \infty$ as $i \to \infty$) and satisfies for all $i \in \mathbb{N}$ that $\frac{\mathfrak{d}(i)}{i} \geq \frac{\mathfrak{d}(i+1)}{i+1}$ and $\mathfrak{d}(i) \leq \mathfrak{d}(i+1)$, then $\mathbf{E}[U] = (1 - \mathrm{o}(1))\frac{n}{\mathfrak{d}(n)+1}$.

(4) If $\mathfrak{d}(i) = c$ for all $i \in \mathbb{N}$, then $\mathbf{E}[U] = (1 - \mathrm{O}(n^{-1}))\frac{n}{c+1}$.

Claim (3) means that, in case of $\mathfrak{d}(i) = \mathrm{o}(i)$ and $\mathfrak{d}(i) = \omega(1)$, we have $\mathbf{E}[U] \approx \frac{n}{\mathfrak{d}(n)}$. Claim (4) is the counterpart of (3) for constant $\mathfrak{d}$. For example, if a new vertex appears every step ($\mathfrak{d}(i) = 1$), a random walk on a growing complete graph misses a half of the number of vertices.

8.2.2 General upper bound (Section 8.6)

In the rest of this chapter, let $t_{\mathrm{hit}}(i)$, $t_{\mathrm{cov}}(i)$ and $t_{\mathrm{mix}}(i)$ denote the hitting, cover and mixing times of $P^{(i)}$, respectively. We obtain several bounds of $\mathbf{E}[U]$ for an RWoGG $(\mathfrak{d}, (G^{(i)})_{i\in\mathbb{N}}, (P^{(i)})_{i\in\mathbb{N}})$ in terms of $t_{\mathrm{hit}}(i), t_{\mathrm{cov}}(i)$ and $t_{\mathrm{mix}}(i)$.

Theorem 8.2.2. *Let $(\mathfrak{d}, (G^{(i)})_{i\in\mathbb{N}}, (P^{(i)})_{i\in\mathbb{N}})$ be an arbitrary RWoGG.*

(1) If there is a constant $C > 1$ such that $\mathfrak{d}(i) \geq Ct_{\mathrm{hit}}(i)$ for all $i \in [n]$, then $\mathbf{E}[U] = \mathrm{O}(1)$.

(2) If $\mathfrak{d}(i)/t_{\mathrm{hit}}(i) \to \infty$ as $i \to \infty$, then $\mathbf{E}[U] \to 0$ as $n \to \infty$.

Theorem 8.2.2 can be seen as an generalization of Theorem 8.2.1 (1) and (2): it is known that $t_{\mathrm{hit}}(i) = \Theta(i)$ if $G^{(i)}$ is the i-vertex complete graph.

In Theorem 8.2.2, we obtain a *general* upper bound of $\mathbf{E}[U]$ in the case of $\mathfrak{d}(i) \geq (1+\epsilon)t_{\mathrm{hit}}(i)$, where $\epsilon > 0$ is a constant. In contrast, the case of $\mathfrak{d}(i) \leq (1+\mathrm{o}(1))t_{\mathrm{hit}}(i)$ seems not easy: it contains an issue of "short random walks". that is a challenging topic in the literature of the cover time of multiple random walks, see e.g., [KMTS19]. Intuitively speaking, if $\mathfrak{d}(i)$ is large enough (say, $\mathfrak{d}(i) \geq t_{\mathrm{mix}}(i)$), then at time $t = T_i$ (recall that $T_i = \min\{t \in \mathbb{Z}_{\geq 0} : \mathcal{G}^{(t)} = G_i\}$), the distribution of the position Z_{T_i} is close to the stationary distribution of $G^{(i-1)}$. This property makes the analysis of $\mathbf{E}[U]$ tractable.

RWoGG on graphs with $t_{\mathrm{mix}} \ll t_{\mathrm{hit}}$. We focus on lazy and reversible random walks. For "rapidly" mixing random walks such that $t_{\mathrm{mix}} \ll t_{\mathrm{hit}}$, we obtain the following upper bound.

Theorem 8.2.3. *Let $(\mathfrak{d}, (G^{(i)})_{i\in\mathbb{N}}, (P^{(i)})_{i\in\mathbb{N}})$ be an RWoGG such that $P^{(i)}$ is lazy and reversible. Let $C > 0$ and $\gamma \in [0,1]$ be arbitrary constants. If $t_{\mathrm{hit}}(i)/t_{\mathrm{mix}}(i) \geq i^{\gamma}/C$ and $\mathfrak{d}(i) \geq \frac{3Ct_{\mathrm{hit}}(i)}{i^{\gamma}}$ for all $1 < i \leq n$, then $\mathbf{E}[U] \leq \frac{8n^{\gamma}}{C} + 32$.*

Observe that Theorem 8.2.3 for $\gamma = 0$ claims that $\mathbf{E}[U] = \mathrm{O}(1)$ if $\mathfrak{d}(i) = \Theta(t_{\mathrm{hit}}(i))$ under the mild condition. A natural question remains unsettled whether $\mathbf{E}[U] = \mathrm{O}(1)$ requires $\mathfrak{d}(i) = \Omega(t_{\mathrm{hit}}(i))$ for any RWoGG $(\mathfrak{d}, (G^{(i)})_{i\in\mathbb{N}}, (P^{(i)})_{i\in\mathbb{N}})$. As a consequence of Theorem 8.2.3, for example, we obtain a bound for degree restricted expander graphs, for which $t_{\mathrm{hit}}(i) = \mathrm{O}(i)$ and $t_{\mathrm{mix}}(i) = \mathrm{O}(\log i)$ hold, that $\mathbf{E}[U] = \mathrm{O}(n^{\gamma})$ if $\mathfrak{d}(i) = \Omega(i^{1-\gamma})$ for $\gamma \in [0,1)$; see Corollary 8.2.6, for detail.

RWoGG on sparse graphs. Though the condition of $t_{\mathrm{mix}} \ll t_{\mathrm{hit}}$ covers interesting examples of rapidly mixing random walks, it misses some representative examples, such as random walks on paths and lollipop graphs, interested in the context of hitting and cover times. To cover those examples, we consider RWoGG on "sparse" graph. Our first concern is a graph sequence $(G^{(i)})_{i\in\mathbb{N}}$ satisfying $\frac{|E^{(i)}|}{|E^{(i-1)}|} = 1 + \mathrm{O}(i^{-1})$. For example,

Theorem 8.2.4. *Let $(\mathfrak{d}, (G^{(i)})_{i\in\mathbb{N}}, (P^{(i)})_{i\in\mathbb{N}})$ be an RWoGG such that $P^{(i)}$ is lazy and simple, and that for all $i \in \{3, \dots, n\}$, $\frac{|E^{(i)}|}{|E^{(i-1)}|} \le 1 + \frac{L}{i}$ hold for some positive constant L. Let $C > 0$ and $\gamma \in [0, 1]$ be arbitrary constants. If $\mathfrak{d}(i) \ge \left(\frac{C}{i^\gamma} + \frac{L+1}{2i}\right) t_{\mathrm{hit}}(i)$ holds for any $1 < i \le n$, then $\mathbf{E}[U] \le \sqrt{L+1}\frac{n^\gamma}{C}$.*

We can obtain an upper bound of $\mathbf{E}[U]$ for a growing path by applying Theorem 8.2.4. Indeed, we will prove that the upper bound is tight in Theorem 8.2.9. We will also demonstrate another applications of Theorem 8.2.4 to a growing *lollipop graph* (see Corollary 8.2.7), where the static lollipop graph is well-known as a tight example for the bounds $t_{\mathrm{hit}} = \mathrm{O}(n^3)$ and $t_{\mathrm{cov}} = \mathrm{O}(n^3)$ for a simple random walk for any graph.

In the following result, we focus on a lazy simple and symmetric random walk with $\mathfrak{d} \ll t_{\mathrm{hit}}$.

Theorem 8.2.5. *Let $(\mathfrak{d}, (G^{(i)})_{i\in\mathbb{N}}, (P^{(i)})_{i\in\mathbb{N}})$ be an RWoGG such that $P^{(i)}$ is lazy and symmetric. Let $C > 0$ and $\gamma \in [0, 1]$ be arbitrary constants. If $\mathfrak{d}(i) \ge \left(\frac{C}{i^\gamma} + \frac{2}{i}\right) t_{\mathrm{hit}}(i)$ for all $1 < i \le n$, then $\mathbf{E}[U] \le \frac{\sqrt{3}n^\gamma}{C}$.*

A typical application of Theorem 8.2.5 is a lazy Metropolis walk with the uniform stationary distribution (see Corollary 8.2.8 for details), which often appears in the context of Markov chain Monte Carlo. Nonaka, Ono, Sadakane, and Yamashita [NOSY10] proved that the Metropolis achieves $t_{\mathrm{hit}}(i) = \mathrm{O}(i^2)$ for any connected graph. The upper bound by Theorem 8.2.5 is also tight for a Metropolis walk on a growing path (Theorem 8.2.9 and Corollary 8.2.10).

Example: Degree restricted expander graph. For a graph $G = (V, E)$, let $d_{\mathrm{ave}}(G)$ and $d_{\min}(G)$ denote the average and the minimum degree of G, respectively. Suppose that P is the transition matrix of the lazy simple random walk on G and let $\lambda_2(P)$ denote the second largest eigenvalue of P. We call a graph G *degree restricted expander graph* if both $\frac{d_{\mathrm{ave}}(G)}{d_{\min}(G)}$ and $\frac{1}{1-\lambda_2(P)}$ are upper bounded by some positive constant. For example, it is easy to see that $G(n, p)$ is a degree restricted expander graph if $p = \omega(\log n/n)$. For any degree restricted expander graph, we have $t_{\mathrm{hit}}(P) = \mathrm{O}(|V|)$ and $t_{\mathrm{mix}}(P) = \mathrm{O}(\log |V|)$ (See Lemma 8.4.6 in Section 8.4 and Theorem 12.4 in [LP17]). Thus Theorem 8.2.3 implies the following.

Corollary 8.2.6. *Suppose that $G^{(i)}$ is a degree restricted expander graph and $P^{(i)}$ is the transition matrix of the lazy simple random walk on $G^{(i)}$ in $R = (\mathfrak{d}, (G^{(i)})_{i\in\mathbb{N}}), (P^{(i)})_{i\in\mathbb{N}})$. Let $\gamma \in [0, 1]$ and $C > 0$ be arbitrary constants. Then two positive constants K_1, K_2 satisfying the following exist: If $\mathfrak{d}(i) \ge CK_1 i^{1-\gamma} + K_2 \log i$ for all $i \in [n]$, then $\mathbf{E}[U] \le 8\frac{n^\gamma}{C} + 32$.*

Proof. Since there exist some positive constants K_1, K_2 satisfying $t_{\mathrm{hit}}(i) \le K_1 i$ and $t_{\mathrm{mix}} \le K_2 \log i$, we obtain the claim from Theorem 8.6.2. □

Example: Lollipop graph. A graph $G = (V, E)$ is (a, b)-*lollipop graph* if G is obtained by connecting the complete graph K_a with the path on b vertices (thus G has $a+b$ vertices and $\binom{a}{2}+b$ edges). Lollipop graphs gather special attention in the literature of random walk since the $(2n/3, n/3)$-lollipop graph attains the maximum possible cover and connected graphs [Fei95b].

In this thesis we consider an RWoGG on which each $G^{(i)}$ is the $(\lceil i/2\rceil, \lfloor i/2\rfloor)$-lollipop graph. Formally, at each round $i \in [n]$, the sets of odd-indexed vertices $V^{(i)}_{\mathrm{odd}} := \{v_{2i-1} : 1 \le i \le \lceil i/2\rceil\}$ and even-indexed vertices $V^{(i)}_{\mathrm{even}} := \{v_{2i} : 1 \le i \le \lfloor i/2\rfloor\}$ form the complete graph and path graph, respectively. Then these two components are connected by an edge $\{v_1, v_2\}$. Let $P^{(i)}$ be the transition matrix of the simple lazy random walk on $G^{(i)}$. For such $P^{(i)}$, it is well known that $t_{\mathrm{hit}}(i) = \mathrm{O}(i^3)$ [Fei95b].

Corollary 8.2.7. *Consider $R = (\mathfrak{d}, (G^{(i)})_{i\in\mathbb{N}}, (P^{(i)})_{i\in\mathbb{N}})$ where $G^{(i)}$ is the lollipop graph defined above and $(P^{(i)})_{i\in[n]}$ is the transition matrix of the lazy simple random walk on $G^{(i)}$. Let $\gamma \in [0, 1]$ be an arbitrary constants. If $\mathfrak{d}(i) \ge C_1 i^{3-\gamma}$ for all i, then $\mathbf{E}[U] \le C_2 n^\gamma$. Here, C_1, C_2 are some positive constants.*

Proof. From definition, $|E^{(2i)}| = 1 + \frac{i(i-1)}{2} + i - 1 = \frac{i(i+1)}{2}$ and $|E^{(2i+1)}| = 1 + \frac{(i+1)i}{2} + i - 1 = \frac{i(i+3)}{2}$. Thus for any i, $\frac{|E^{(i)}|}{|E^{(i-1)}|} \le 1 + \frac{K_1}{i}$ for some constant K_1. Furthermore, $t^{(i)}_{\mathrm{hit}} \le K_2 i^3$ holds for some constant K_2. Applying Theorem 8.2.4, we obtain the claim. □

Figure 8.1: The transition diagram of (8.2).

Example: Metropolis walk. For a given $G = (V, E)$, the transition matrix P of the lazy *Metropolis walk* on G is defined by

$$(P)_{u,v} = \begin{cases} \frac{1}{2\max\{d_u, d_v\}} & (\text{if } \{u,v\} \in E) \\ 1 - \sum_{w:\{u,w\}\in E} (P)_{u,w} & (\text{if } u = v) \\ 0 & (\text{otherwise}). \end{cases} \tag{8.1}$$

Nonaka, Ono, Sadakane and Yamashita [NOSY10] proved that $t_{\mathrm{hit}}(P) = \mathrm{O}(|V|^2)$ for any connected graphs. Since P is symmetric matrix, we can apply Theorem 8.2.5 directly.

Corollary 8.2.8. *Consider $R = (\mathfrak{d}, (G^{(i)})_{i\in\mathbb{N}}, (P^{(i)})_{i\in\mathbb{N}})$, where each $P^{(i)}$ is the lazy Metropolis walk on a connected graph $G^{(i)}$. Let $\gamma \in [0,1]$ and $C > 0$ be arbitrary constants. If $\mathfrak{d}(i) \geq \left(\frac{C}{i^\gamma} + \frac{2}{i}\right) t_{\mathrm{hit}}(i)$ for all $1 < i \leq n$, then $\mathbf{E}[U] \leq \sqrt{3}\frac{n^\gamma}{C}$.*

8.2.3 Lower bound for paths (Section 8.7)

In contrast to upper bounds, an analysis of a lower bound requires more technically complicated arguments. We establish a lower bound of $\mathbf{E}[U]$ for a random walk on a growing path graph, which implies that the upper bound by Theorem 8.2.5 is tight in the case. Let $R_{\mathrm{p}} = (\mathfrak{d}, (G^{(i)})_{i\in\mathbb{N}}, (P^{(i)})_{i\in\mathbb{N}})$ be a random walk on a growing *path* graph, where $G^{(i)} = (V^{(i)}, E^{(i)})$ is given by $V^{(i)} = \{v_1, \ldots, v_i\}$, and $E^{(i)} = \{\{v_1, v_2\}, \ldots, \{v_{i-1}, v_i\}\}$, and $P^{(i)}$ is given by

$$(P^{(i)})_{u,v} = \begin{cases} p & \text{if } u = v = v_1 \text{ or } u = v = v_i, \\ 1-p & \text{if } (u,v) \in \{(v_1, v_2), (v_i, v_{i-1})\}, \\ q & \text{if } \{u,v\} = \{v_j, v_{j+1}\} \text{ for } j \in \{2, 3, \ldots, i-1\}, \\ 1-2q & \text{if } u = v = v_j \text{ for } j = \{2, 3, \ldots, i-1\}, \\ 0 & \text{otherwise} \end{cases} \tag{8.2}$$

for two parameters $p, q \in [0,1]$ satisfying $p \geq q$ and $q \leq 1/2$ (see Figure 8.1). For example, if $(p,q) = (\frac{1}{2}, \frac{1}{4})$, the corresponding walk is the lazy simple random walk. If $(p,q) = (\frac{3}{4}, \frac{1}{4})$ the corresponding one is the lazy Metropolis random walk.

Theorem 8.2.9. *If $\mathfrak{d}(i) \leq Ci^{2-\gamma}$ in R_{p} for some constants $C > 0$ and $\gamma \in [0,1]$ then $\mathbf{E}[U] = \Omega(n^\gamma/C)$.*

Theorems 8.2.4, 8.2.5 and 8.2.9 imply the following tight bounds of $\mathbf{E}[U]$ on a growing path.

Corollary 8.2.10. *For $R_{\mathrm{p}} = (\mathfrak{d}, (G^{(i)})_{i\in\mathbb{N}}, (P^{(i)})_{i\in\mathbb{N}})$, where $P^{(i)}$ is the transition matrix of either the lazy simple random walk or the lazy Metropolis random walk. Then*

(1) If $\mathfrak{d}(i) \geq Ci^{2-\gamma}$ for some constants $C > 0$ and $\gamma \in [0,1]$ then $\mathbf{E}[U] = \mathrm{O}(n^\gamma/C)$.

(2) If $\mathfrak{d}(i) \leq Ci^{2-\gamma}$ for some constants $C > 0$ and $\gamma \in [0,1]$ then $\mathbf{E}[U] = \Omega(n^\gamma/C)$.

8.3 Related Works

The cover time is a fundamental topic of analyses of random walks. Here, we review some representative results about the cover times of random walks on static graphs, and on dynamic graphs.

8.3.1 Cover times of random walks on static graphs

It is known that the cover time of a simple random walk satisfies $t_{\mathrm{cov}} \leq 2m(n-1)$ for any undirected graph, see Aleliunas, Karp, Lipton, Lovász, and Rackoff [AKL+79] and Aldous [Ald83]. Mathews [Mat88] devised a technique of upper and lower bounding t_{cov} by t_{hit}, of which a celebrated implication is $t_{\mathrm{cov}} \leq t_{\mathrm{hit}} \log n$. The lolipop graph is famous for $t_{\mathrm{hit}} = \Omega(n^3)$, and hence $t_{\mathrm{cov}} = \Omega(n^3)$. Fiege gave a tight upper bound of the cover times of simple random walks on any graphs such that $t_{\mathrm{cov}} \leq \frac{4}{27}n^3 + \mathrm{O}(n^{5/2})$ in [Fei95b], while he in [Fei95a] gave a tight lower bound of the cover time of simple random walks on any graphs such that $t_{\mathrm{cov}} \geq n \ln n + \mathrm{o}(n \ln n)$, using a Mathews' argument [Mat88]. The connection between the hitting time and electric circuits is well known (see e.g., [DS84, AF, LP17]).

Motivated by a faster covering by a random walk, Ikeda et al. [IKOY03] (see also [IKY09]) proposed β-random walk, which makes transitions only using local information, and proved that the cover time of a β-random walk is upper bounded by $\mathrm{O}(n^2 \log n)$ for any graph. Nonaka et al. [NOSY10] proved the same bound holds for a Metropolis walk, which is simpler and more popular than β-random walk. Recently, David and Feige [DF17] (see also [DF18]) proved that a biased random walk achieves $\mathrm{O}(n^2)$ cover time for any graph, and affirmatively settled the question posed by Ikeda et al. [IKOY03].

8.3.2 Cover time of random walks on dynamic graphs

An early work [CF03] by Cooper and Frieze investigated random walks on growing "web-graphs". Specifically, they considered a random walk on a growing preferential attachment graph with a constant duration (i.e., the number of vertices increases every constant steps). They proved that $\mathbf{E}[U]/n$ converges to some constant as n tends to infinity. Note that our RWoGG contains their model as a special case.

There are several results about the cover times of random walks on dynamic graphs, sometimes called "evolving graphs," with static vertex sets. Avin et al. [AKL08] (see also [AKL18]) investigated the hitting times, mixing times and cover times of random walks on evolving graphs with static vertex sets. They gave a prescribed sequence of graphs on which the hitting time of a simple random walk gets $2^{\Omega(n)}$, and hence the cover time is as well. On the other hand, they proved that the cover time of a max-degree random walk is $\mathrm{O}(d_{\max} n^3 (\log n)^2)$ where $d_{\max}$ is the maximum degree of the evolving graph. Denysyuk and Rodrigues [DR14] were concerned with *ρ-recurrent family* of evolving graphs, where preferable graphs are assumed to appear frequently in the graph sequence. Then, for max-degree random walks on ρ-recurrent families, they gave upper and lower bounds of the cover time in terms of the hitting time, as well as gave an upper bound of the mixing time. Lamprou, Martin, and Spirakis [LMS18] were concerned with two random walks of "random walk with a delay" (RWD), where at each step, the walker chooses an edge of underlying graph and moves when it appears, and "random walk on what is available" (RWA), where the walker chooses an edge of current graph and moves immediately. Then, they investigated the cover times of RWD and RWA for edge-uniform stochastically evolving graphs. Sauerwald and Zanetti [SZ19] extended the argument by Avin et al. [AKL18] in the case that a sequence of graphs have the same stationary distribution, and presented an upper bound $\mathrm{O}(n^2)$ of the cover time on d-regular dynamic graphs.

8.3.3 Other related works

Saloff-Coste and Zúñiga investigated time-inhomogeneous Markov chains, and provided some Nash and log-Sobolev inequalities [SCZ09, SCZ11]. Recently, Cai, Sauerwald, and Zanetti [CSZ20] investigated the relation between the density of edge-Markovian dynamic graphs and mixing times. They showed for fast-changing dynamic graphs that $t_{\mathrm{mix}} = \infty$ in sparse case while $t_{\mathrm{mix}} = \mathrm{O}(\log n)$ in dense case. They also showed for slowly-changing dynamic graphs that $t_{\mathrm{mix}} = \Omega(n)$ in sparse case while $t_{\mathrm{mix}} = \mathrm{O}(\log n)$ in dense case.

There are many works on other stochastic processes on dynamic graphs, such as exploration, information spreading, rumor spreading, gossiping and voter model, see e.g., [JAR16, CST15, BGKMT16]. Theoretical analyses of algorithms on dynamic graphs attract high attentions in the context of distributed computing, and there are many works concerning the topics, such as exploration, agreement, and population protocol, on dynamic networks, see e.g., [MS18, Mic16, KO11].

8.4 Preliminaries

8.4.1 Random walk

We briefly introduce notions of random walk. Let $P \in [0,1]^{V\times V}$ be a transition matrix over V. A (time-homogeneous) *random walk* (or *Markov chain*) is a sequence $X = (X_t)_{t\in\mathbb{Z}_{\geq 0}}$ of random variables given as $\mathbf{Pr}[X_{t+1} = v \mid X_t = u] = P_{u,v}$ for each $t \in \mathbb{Z}_{\geq 0}$. Note that a random walk is characterized by the transition matrix P.

A random walk is *lazy* if $P_{v,v} \geq 1/2$ for all $v \in V$, is *reversible* if P is reversible, and is *symmetric* if $P_{u,v} = P_{v,u}$ holds for all $u, v \in V$.

For a transition matrix P, The *hitting time* t_{hit} (also denoted by $t_{\mathrm{hit}}(P)$) is the random variable given by $t_{\mathrm{hit}} := \max_{u,v\in V} \mathbf{E}[\min\{t \geq 0 : X_0 = u \text{ and } X_t = v\}]$. The *cover time* t_{cov} (or $t_{\mathrm{cov}}(P)$) is given by $t_{\mathrm{cov}} := \max_{u\in V} \mathbf{E}[\min\{t \geq 0 : [X_0 = u] \text{ and } [\forall v \in V, \exists s \leq t, X_s = v]\}]$. The *mixing time* t_{mix} is given by $t_{\mathrm{mix}} := \min\{t > 0 : (1/2)\max_{u\in V} \sum_{v\in V} |P^t(u,v) - \pi(v)| \leq 1/4\}$.

8.4.2 Notation

For ease of notation, we sometimes use $x(v)$ to denote x_v for a vector $x \in \mathbb{R}^V$ and $v \in V$. For two vectors $x, y \in \mathbb{R}$ and a probability vector $\pi \in (0,1]^V$, let $\langle x, y\rangle_\pi := \sum_{v\in V} \pi(v)x(v)y(v)$. Then, the $\ell_2(\pi)$-norm of x is defined by $\|x\|_{2,\pi} := \sqrt{\langle x, x\rangle_\pi} = \sqrt{\sum_{v\in V} \pi(v)x(v)^2}$. For two vectors $x, y \in \mathbb{R}^V$ where $y(v) \neq 0$ holds for all $v \in V$, define $\frac{x}{y} \in \mathbb{R}^V$ by $\frac{x}{y}(v) = \frac{x(v)}{y(v)}$. Note that, for any probability vector $\xi \in [0,1]^V$, $\left\|\frac{\xi}{\pi} - \mathbb{1}^{(|V|)}\right\|_{2,\pi}^2 = \left\|\frac{\xi}{\pi}\right\|_{2,\pi}^2 - 1$ holds. Here, $\mathbb{1}^{(n)}$ denotes the n-dimensional vector where all elements are equal to one.

8.4.3 Tools

Lemma 8.4.1 (Theorem 4.1 of [OP19])**.** *Let $P \in [0,1]^{V\times V}$ be an irreducible, reversible and lazy transition matrix over V, and let $\pi \in (0,1]^V$ denote its stationary distribution. Let $(X_t)_{t\in\mathbb{Z}_{\geq 0}}$ denote the Markov chain according to P. Let $\tau_v(P) = \min\{t \geq 0 : X_t = v\}$ and let $t_{\mathrm{hit}}(P) = \max_{u,v\in V} \mathbf{E}_u[\tau_v(P)]$. Then for any $t \geq 0$ and any $h_0, h_1, \ldots, h_t \in V$,*

$$\mathop{\mathbf{Pr}}_{\pi}[X_s \neq h_s \text{ for all } s \in \{0,\ldots,t\}] \leq \left(1 - \frac{1}{t_{\mathrm{hit}}(P)}\right)^t.$$

By taking $h_0 = \cdots = h_t = v$ in Lemma 8.4.1, we immediately obtain the following.

Corollary 8.4.2. *Let $P \in [0,1]^{V\times V}$ be an irreducible, reversible and lazy transition matrix over V, and let $\pi \in (0,1]^V$ denote its stationary distribution. Let $(X_t)_{t\in\mathbb{Z}_{\geq 0}}$ denote the Markov chain according to P. Let $\tau_v(P) = \min\{t \in \mathbb{Z}_{\geq 0} : X_t = v\}$ and let $t_{\mathrm{hit}}(P) = \max_{u,v\in V} \mathbf{E}_u[\tau_v(P)]$. Then for any $v \in V$ and $t > 0$,*

$$\mathop{\mathbf{Pr}}_{\pi}[\tau_v(P) > t] \leq \left(1 - \frac{1}{t_{\mathrm{hit}}(P)}\right)^t \leq \exp\left(-\frac{t}{t_{\mathrm{hit}}(P)}\right).$$

Lemma 8.4.3 (See Section 3.6.5 of [AF] or Theorem 4.1 of [OP19])**.** *Let $P \in [0,1]^{V\times V}$ be an irreducible and reversible transition matrix over V, and let $\pi \in (0,1]^V$ denote its stationary distribution. For a subset $S \subseteq V$, define $P_{\overline{S}} \in [0,1]^{V\times V}$ by $(P_{\overline{S}})_{u,v} = P_{u,v}$ if $u, v \in V \setminus S$ and $(P_{\overline{S}})_{u,v} = 0$ otherwise. Let $\lambda(M)$ denote the largest eigenvalue of a matrix M. Then for any $S \notin \{\emptyset, V\}$,*

$$\lambda(P_{\overline{S}}) \leq 1 - \frac{1}{t_{\mathrm{hit}}(P)}.$$

Furthermore, for any $S \notin \{\emptyset, V\}$ and any $f \in \mathbb{R}^V$,

$$\langle f, P_{\overline{S}} f\rangle_\pi \leq \lambda(P_{\overline{S}})\langle f, f\rangle_\pi.$$

Since $\|P_{\overline{S}} f\|_{2,\pi}^2 = \langle P_{\overline{S}} f, P_{\overline{S}} f\rangle_\pi = \left\langle f, P_{\overline{S}}^2 f\right\rangle_\pi$, we have the following corollary.

Corollary 8.4.4. *Let $P \in [0,1]^{V\times V}$ be an irreducible, reversible and lazy transition matrix over V, and let $\pi \in (0,1]^V$ denote its stationary distribution. Suppose that $P_{\overline{S}}$ is a matrix defined in Lemma 8.4.3. Then for any $S \notin \{\emptyset, V\}$ and any $f \in \mathbb{R}^V$,*

$$\|P_{\overline{S}} f\|_{2,\pi}^2 \le \lambda_1(P_{\overline{S}})^2 \|f\|_{2,\pi}^2 \le \left(1 - \frac{1}{t_{\mathrm{hit}}(P)}\right)^2 \|f\|_{2,\pi}^2$$

Here, $\lambda_1(M)$ denotes the largest eigenvalue in absolute value of a matrix M.

Lemma 8.4.5 (See e.g. (12.8) of [LP17])**.** *Let $P \in [0,1]^{V\times V}$ be a reversible transition matrix with respect to $\pi \in (0,1]^V$. Then for any probability vector $f \in [0,1]^V$, $\left\|\frac{f}{\pi} - \mathbb{1}\right\|_{2,\pi}^2 = \left\|\frac{f}{\pi}\right\|_{2,\pi}^2 - 1$ and*

$$\left\|P\frac{f}{\pi} - \mathbb{1}\right\|_{2,\pi}^2 \le \lambda_2(P)^2 \left\|\frac{f}{\pi} - 1\right\|_{2,\pi}^2$$

holds where $\lambda_2(P)$ is the second largest eigenvalue (in absolute value) of P.

Lemma 8.4.6 (Lemmas 4.24 and 4.25 of [AF])**.** *Let P be reversible transition matrix and let π be its stationary distribution. Then*

$$\frac{1}{1-\lambda_2(P)} \le t_{\mathrm{hit}}(P) \le \frac{2}{\pi_{\min}(1-\lambda_2(P))}.$$

8.5 Complete Graph

This section is devoted to the proof of Theorem 8.2.1. Consider a random walk $(Z_t)_{t\in\mathbb{Z}_{\ge 0}}$ on a RWoGG. For convenience, we divide the T_{n+1} step random walk $Z_0, \dots, Z_{T_{n+1}}$ into n random walks each of length $\mathfrak{d}(i)$ (for $i = 1, \dots, n$). We call each period a *round*. For each $i \in [n]$, let $(X_s^{(i)})_{s\in\mathbb{Z}_{\ge 0}}$ denote a random walk in the i-th round (specifically, it is a random walk according to $P^{(i)}$) with the initial state $X_0^{(i)} = Z_{T_i} = X_{\mathfrak{d}(i-1)}^{(i-1)}$. Note that $(X_s^{(i)})_{s\in\mathbb{Z}_{\ge 0}}$ is a random walk on $G^{(i)}$. Table 8.1 illustrates the correspondence between Z_t and $X_s^{(i)}$ in the case of $\mathfrak{d}(i) = i$.

	Z_0	Z_1	Z_2	Z_3	Z_4	Z_5	Z_6	Z_7	Z_8	$\cdots$
$G^{(1)}$	$X_0^{(1)}$	$X_1^{(1)}$	$\cdots$							
$G^{(2)}$		$X_0^{(2)}$	$X_1^{(2)}$	$X_2^{(2)}$	$\cdots$					
$G^{(3)}$				$X_0^{(3)}$	$X_1^{(3)}$	$X_2^{(3)}$	$X_3^{(3)}$	$\cdots$		
$G^{(4)}$							$X_0^{(4)}$	$X_1^{(4)}$	$X_2^{(4)}$	$\cdots$

Table 8.1: Correspondence between Z_t and $X_s^{(i)}$ when $\mathfrak{d}(i) = i$. For each $i \in \mathbb{N}$, $(X_s^{(i)})_{s\in\mathbb{Z}_{\ge 0}}$ is a random walk on $G^{(i)}$. Note that $X_0^{(i)} = X_{\mathfrak{d}(i-1)}^{(i-1)} = Z_{T_i}$ for $i \ge 2$. In this example, $U(3) = 3 - \left|\bigcup_{t=0}^{T_{3+1}} \{Z_t\}\right| = 3 - \left|\bigcup_{i=1}^{3}\bigcup_{s=0}^{i} \{X_s^{(i)}\}\right|$.

For $v \in V^{(n)}$ let $\mathcal{E}(v)$ denote the event that $v \notin \bigcup_{i=1}^{n}\bigcup_{s=0}^{\mathfrak{d}(i)} \{X_s^{(i)}\}$. In other words, $\mathcal{E}(v)$ means that the random walk $Z_0, Z_1, \dots, Z_{T_{n+1}}$ does not visit the vertex v.

Consider the RWoGG of Theorem 8.2.1. For the vertex v_k attached to $\mathcal{G}$ at time T_k, we see that $\mathbf{Pr}[\mathcal{E}(v_k)] = \prod_{i=k}^{n} \left(1 - \frac{1}{i}\right)^{\mathfrak{d}(i)}$ holds, and thus

$$\mathbf{E}[U] = \sum_{k=1}^{n} \mathbf{Pr}[\mathcal{E}(v_k)] = \sum_{k=1}^{n}\prod_{i=k}^{n} \left(1 - \frac{1}{i}\right)^{\mathfrak{d}(i)}$$

holds. Theorem 8.2.1 follows from the next lemma.

Lemma 8.5.1. *For a function $f : \mathbb{N} \to \mathbb{N}$, let $S(n) := \sum_{k=1}^{n} \prod_{i=k}^{n} \left(1 - \frac{1}{i}\right)^{f(i)}$.*

(i) If $f(i) \ge Ci$ for some constant C, then $S(n) = \mathrm{O}(1)$.

(ii) If f satisfies $f(i) \le f(i+1)$ for all $i \in \mathbb{N}$, then $S(n) \ge \frac{n}{f(n)+1}\left(1-\frac{1}{n}\right)^{f(n)}$.

(iii) If f satisfies $\frac{f(i)}{i} \ge \frac{f(i+1)}{i+1}$, then for all $n \in \mathbb{N}$, $S(n) \le \frac{n}{f(n)}$.

(iv) If there is a constant $c \in \mathbb{N}$ such that $f(i) = c$ for all $i \in \mathbb{N}$, then for all $n \in \mathbb{N}$, $S(n) \le \frac{n}{c+1}$.

Proof of (i). Since $1 + x \le \mathrm{e}^x$, we have

$$S(n) \le \sum_{k=1}^{n} \exp\left(-\sum_{i=k}^{n} \frac{f(i)}{i}\right) \le \sum_{k=1}^{n} \exp\left(-(n-k+1)C\right) = \mathrm{O}(1).$$

□

Proof of (ii). Observe that $S(1) = 0$ and for all $n \ge 1$,

$$S(n+1) = \sum_{k=1}^{n+1} \prod_{i=k}^{n+1} \left(1-\frac{1}{i}\right)^{f(i)} = \left(1-\frac{1}{n+1}\right)^{f(n+1)} (S(n)+1). \tag{8.3}$$

We prove (ii) by induction on n. In the base case, $S(1) = 0$ and we are done. If $S(n) \ge \frac{n}{f(n)+1}\left(1-\frac{1}{n}\right)^{f(n)}$, then

$$\begin{aligned} S(n)+1 &\ge \frac{n}{f(n)+1}\left(1-\frac{1}{n}\right)^{f(n)} + 1 \ge \frac{n}{f(n)+1}\left(1-\frac{f(n)}{n}\right) + 1 \\ &= \frac{n-f(n)}{f(n)+1} + 1 = \frac{n+1}{f(n)+1} \ge \frac{n+1}{f(n+1)+1}. \end{aligned} \tag{8.4}$$

Here, we used $(1+x)^r \ge 1 + rx$ in the second inequality and $f(n) \le f(n+1)$ in the last inequality. Combining (8.3) and (8.4), $S(n+1) \ge \left(1-\frac{1}{n+1}\right)^{f(n+1)} \frac{n+1}{f(n+1)+1}$ and we are done. □

Proof of (iii). The proof is obtained by induction on $n \ge 1$. When $n = 1$, $S(1) = 0 \le 1/f(1)$. Assume $S(n) \le n/f(n)$. Then,

$$S(n+1) = \left(1-\frac{1}{n+1}\right)^{f(n+1)} (S(n)+1) \le \frac{\frac{n}{f(n)}+1}{1+\frac{f(n+1)}{n+1}} \le \frac{\frac{n+1}{f(n+1)}+1}{1+\frac{f(n+1)}{n+1}} = \frac{n+1}{f(n+1)}.$$

Note that $(1-x)^y \le 1/(1+xy)$ for all $x \in [0,1]$ and $y \ge 0$. The second inequality follows from $\frac{f(n+1)}{n+1} \le \frac{f(n)}{n}$. □

Proof of (iv). The proof is obtained by induction on n. First $S(1) = 0 \le 1/(f(1)+1)$. Assume $S(n) \le n/(f(n)+1)$. Then, from (8.3) and the induction assumption, we have

$$S(n+1) \le \frac{\frac{n}{f(n)+1}+1}{1+\frac{f(n+1)}{n+1}} = \frac{\frac{n}{f(n)+1}+1}{1+\frac{f(n)}{n+1}} = \frac{\frac{n+1}{f(n)+1}\left(\frac{n}{n+1}+\frac{f(n)+1}{n+1}\right)}{\frac{n}{n+1}+\frac{f(n)+1}{n+1}} = \frac{n+1}{f(n)+1} = \frac{n+1}{f(n+1)+1}.$$

Note that we use $f(n) = f(n+1)$ in the first and the last equality. □

We are ready to prove Theorem 8.2.1.

Proof of Theorem 8.2.1. Recall that $\mathbf{E}[U] = S(n)$. Statement (1) follows from Lemma 8.5.1(i). Statement (3) follows from (ii) and (iii) of Lemma 8.5.1. (4) follows from (ii) and (iv) of Lemma 8.5.1.

Now, we prove Statement (2). More precisely, we prove that, for any $\epsilon > 0$, there is $n_0 \in \mathbb{N}$ such that for all $n \ge n_0$, $S(n) \le \epsilon$ holds. From the assumption that $\mathfrak{d}(i) = \omega(i)$, for any large constant $C > 0$, we can take $i_0 \in \mathbb{N}$ such that for all $i \ge i_0$, $f(i) > Ci$ holds. Fix a constant $C > 0$ and take i_0 in this way. Since $1 + x \le \mathrm{e}^x$ and $f(k)/k > C$ for all $k \ge i_0$, we have

$$\begin{aligned} S(n) &\le \sum_{i=1}^{i_0} \exp\left(-\sum_{k=i_0}^{n} \frac{f(k)}{k}\right) + \sum_{i=i_0+1}^{n} \exp\left(-\sum_{k=i}^{n} \frac{f(k)}{k}\right) \\ &\le i_0 \exp(-(n-i_0+1)C) + \sum_{i=i_0+1}^{n} \exp(-(n-i+1)C) \\ &\le i_0 \exp(-(n-i_0+1)C) + \frac{\mathrm{e}^{-C}}{1-\mathrm{e}^{-C}}. \end{aligned}$$

Let $\epsilon > 0$ be an arbitrary small constant. Then, take $C > 0$ such that $\frac{\mathrm{e}^{-C}}{1-\mathrm{e}^{-C}} < \frac{\epsilon}{2}$ holds. According to this constant C, we can take i_0 such that $f(i) > Ci$ for all $i \geq i_0$ holds. Now C and i_0 are fixed. Hence, for sufficiently large n, we have $i_0 \exp(-(n - i_0 + 1)C) \leq \frac{\epsilon}{2}$. This implies $S(n) \leq \epsilon$ and we are done. □

8.6 General Upper Bound

In this section we prove Theorems 8.2.2 to 8.2.5. Consider an RWoGG $R = (\mathfrak{d}, (G^{(i)})_{i\in\mathbb{N}}, (P^{(i)})_{i\in\mathbb{N}})$. Recall that, at each round i, $(X_t^{(i)})_{t\in\mathbb{Z}_{\geq 0}}$ denotes the random walk according to $P^{(i)}$ where $X_0^{(i)} = X_{\mathfrak{d}(i-1)}^{(i-1)}$ holds (See Table 8.1 for an example). Let $\pi^{(i)}$ denote the stationary distribution of $P^{(i)}$. Let $\tau_v^{(i)} := \min\{t \geq 0 : X_t^{(i)} = v\}$, i.e., $\tau_v^{(i)}$ denotes the time taken for a random walk $(X_t^{(i)})_{t\in\mathbb{Z}_{\geq 0}}$ to reach $v \in V^{(i)}$. Note that $\mathbf{E}[\tau_u^{(i)}] \leq \max_{u,v\in V} \mathbf{E}[\tau_v^{(i)} | X_0^{(i)} = u] = t_{\mathrm{hit}}(i)$. Recall that $V(G^{(i)}) = \{v_1, \ldots, v_i\}$ and thus $X_0^{(1)} = v_1$. For any round $k \leq n$, the probability that the walker does not visit the vertex v_k until the end of the round n is equal to $\mathbf{Pr}\left[\bigwedge_{i=k}^{n} \left\{\tau_{v_k}^{(i)} > \mathfrak{d}(i)\right\}\right]$. Hence we have

$$\begin{aligned}
\mathbf{E}[U] &= \sum_{k=1}^{n} \mathbf{Pr}\left[\bigwedge_{i=k}^{n} \left\{\tau_{v_k}^{(i)} > \mathfrak{d}(i)\right\}\right] \\
&= \sum_{k=2}^{n} \mathbf{Pr}\left[\bigwedge_{i=k}^{n} \left\{\tau_{v_k}^{(i)} > \mathfrak{d}(i)\right\}\right] \\
&= \sum_{k=2}^{n} \sum_{v\in V^{(k-1)}} \mathbf{Pr}\left[X_0^{(k)} = v\right] \mathbf{Pr}\left[\bigwedge_{i=k}^{n} \left\{\tau_{v_k}^{(i)} > \mathfrak{d}(i)\right\} \middle| X_0^{(k)} = v\right] \qquad (8.5) \\
&\leq \sum_{k=2}^{n} \max_{v\in V^{(k-1)}} \mathbf{Pr}\left[\bigwedge_{i=k}^{n} \left\{\tau_{v_k}^{(i)} > \mathfrak{d}(i)\right\} \middle| X_0^{(k)} = v\right]. \qquad (8.6)
\end{aligned}$$

The second equality follows from $\mathbf{Pr}[X_1^{(1)} \neq v_1] = 0$. The rest of this section is devoted giving upper bounds of (8.5) and (8.6).

8.6.1 Upper bound for large $\mathfrak{d}$

We show Theorem 8.2.2 in this section. To begin with, we show the following lemma.

Lemma 8.6.1. *For any $R = (\mathfrak{d}, (G^{(i)})_{i\in\mathbb{N}}, (P^{(i)})_{i\in\mathbb{N}})$, we have*

$$\mathbf{E}[U] \leq \sum_{k=2}^{n} \prod_{i=k}^{n} \max_{v\in V^{(i)}} \mathbf{Pr}\left[\tau_{v_k}^{(i)} > \mathfrak{d}(i) \middle| X_0^{(i)} = v\right].$$

Proof. Consider a fixed vertex v_k with $k > 1$. For a round $i \geq k$ and a vertex $u \in V^{(i)}$, let $\mathcal{E}_u^{(i)} = \mathcal{E}_u^{(i)}(v_k)$ denote the event that the walker is in vertex u at the end of the i-th round without visiting vertex v_k during the round. Formally $\mathcal{E}_u^{(i)}(v_k)$ is defined as the event of $\{\tau_{v_k}^{(i)} > \mathfrak{d}(i)\} \wedge \{X_{\mathfrak{d}(i)}^{(i)} = u\}$. Then for any $u_{k-1} \in V^{(k-1)}$,

$$\mathbf{Pr}\left[\bigwedge_{i=k}^{n} \left\{\tau_{v_k}^{(i)} > \mathfrak{d}(i)\right\} \middle| X_0^{(k)} = u_{k-1}\right] = \sum_{u_k\in V^{(k)}} \cdots \sum_{u_n\in V^{(n)}} \mathbf{Pr}\left[\bigwedge_{i=k}^{n} \mathcal{E}_{u_i}^{(i)} \middle| X_0^{(k)} = u_{k-1}\right]. \qquad (8.7)$$

To bound (8.7), we first observe that, for any vertices, $u_{k-1} \in V^{(k-1)}, u_k \in V^{(k)}, \ldots, u_n \in V^{(n)}$,

$$\mathbf{Pr}\left[\bigwedge_{i=k}^{n} \mathcal{E}_{u_i}^{(i)} \middle| X_0^{(k)} = u_{k-1}\right] = \frac{\mathbf{Pr}\left[X_0^{(k)} = u_{k-1}, \mathcal{E}_{u_k}^{(k)}\right]}{\mathbf{Pr}\left[X_0^{(k)} = u_{k-1}\right]} \prod_{\ell=k+1}^{n} \frac{\mathbf{Pr}\left[X_0^{(k)} = u_{k-1}, \bigwedge_{i=k}^{\ell} \mathcal{E}_{u_i}^{(i)}\right]}{\mathbf{Pr}\left[X_0^{(k)} = u_{k-1}, \bigwedge_{i=k}^{\ell-1} \mathcal{E}_{u_i}^{(i)}\right]} \qquad (8.8)$$

holds. Then, from the definition of the conditional probability, we have

$$\frac{\mathbf{Pr}\left[X_0^{(k)} = u_{k-1}, \mathcal{E}_{u_k}^{(k)}\right]}{\mathbf{Pr}\left[X_0^{(k)} = u_{k-1}\right]} = \mathbf{Pr}[\mathcal{E}_{u_k}^{(k)} | X_0^{(k)} = u_{k-1}]$$

and

$$\frac{\mathbf{Pr}\left[X_0^{(k)} = u_{k-1}, \bigwedge_{i=k}^{\ell} \mathcal{E}_{u_i}^{(i)}\right]}{\mathbf{Pr}\left[X_0^{(k)} = u_{k-1}, \bigwedge_{i=k}^{\ell-1} \mathcal{E}_{u_i}^{(i)}\right]} = \mathbf{Pr}\left[\mathcal{E}_{u_\ell}^{(\ell)} \middle| X_0^{(k)} = u_{k-1}, \bigwedge_{i=k}^{\ell-1} \mathcal{E}_{u_i}^{(i)}\right]$$
$$= \mathbf{Pr}\left[\mathcal{E}_{u_\ell}^{(\ell)} \middle| X_{\mathfrak{d}(\ell-1)}^{(\ell-1)} = u_{\ell-1}\right] = \mathbf{Pr}\left[\mathcal{E}_{u_\ell}^{(\ell)} \middle| X_0^{(\ell)} = u_{\ell-1}\right]. \tag{8.9}$$

We use the Markov property in the second equality. The last equality follows from our assumption of $X_{\mathfrak{d}(\ell-1)}^{(\ell-1)} = X_0^{(\ell)}$. Hence combining (8.7) to (8.9), we have

$$\mathbf{Pr}\left[\bigwedge_{i=k}^{n} \left\{\tau_{v_k}^{(i)} > \mathfrak{d}(i)\right\} \middle| X_0^{(k)} = u_{k-1}\right]$$
$$= \sum_{u_k \in V^{(k)}} \cdots \sum_{u_n \in V^{(n)}} \prod_{\ell=k}^{n} \mathbf{Pr}\left[\tau_{v_k}^{(\ell)} > f(\ell), X_{\mathfrak{d}(\ell)}^{(\ell)} = u_\ell \middle| X_0^{(\ell)} = u_{\ell-1}\right] \tag{8.10}$$
$$= \sum_{u_k \in V^{(k)}} \mathbf{Pr}\left[\mathcal{E}_{u_k}^{(k)} \middle| X_0^{(k)} = u_{k-1}\right] \cdots \sum_{u_n \in V^{(n)}} \mathbf{Pr}\left[\mathcal{E}_{u_n}^{(n)} \middle| X_0^{(n)} = u_{n-1}\right]$$
$$\leq \prod_{\ell=k}^{n} \max_{u \in V^{(\ell)}} \sum_{u_\ell \in V^{(\ell)}} \mathbf{Pr}\left[\mathcal{E}_{u_\ell}^{(\ell)} \middle| X_0^{(\ell)} = u\right] = \prod_{\ell=k}^{n} \max_{u \in V^{(\ell)}} \mathbf{Pr}\left[\tau_{v_k}^{(\ell)} > \mathfrak{d}(\ell) \middle| X_0^{(\ell)} = u\right]. \tag{8.11}$$

We obtain the claim from (8.6) and (8.11). □

Proof of Theorem 8.2.2(1). From the Markov inequality, for any $k \leq i$ and $v \in V^{(i)}$, we have

$$\mathbf{Pr}\left[\tau_{v_k}^{(i)} > \mathfrak{d}(i) \middle| X_0^{(i)} = v\right] \leq \frac{\mathbf{E}\left[\tau_{v_k}^{(i)} \middle| X_0^{(i)} = v\right]}{\mathfrak{d}(i)} \leq \frac{t_{\mathrm{hit}}(i)}{\mathfrak{d}(i)}.$$

Hence from Lemma 8.6.1, we obtain

$$\mathbf{E}[U] \leq \sum_{k=1}^{n} \prod_{i=k}^{n} \frac{t_{\mathrm{hit}}(i)}{\mathfrak{d}(i)} \leq \sum_{k=1}^{n} C^{-(n-k+1)} = \sum_{k=1}^{n} C^{-k} \leq \frac{1}{C-1}.$$

□

Proof of Theorem 8.2.2(2). For an arbitrary (small) $\epsilon > 0$, let $C = C(\epsilon) = \frac{2}{\epsilon} + 1$. From assumption on (2), we can take some $i_0 = i_0(\epsilon)$ such that $\mathfrak{d}(i) \geq C t_{\mathrm{hit}}(i)$ for all $i \geq i_0$. Let $K = \max_{i \in [i_0]} \frac{t_{\mathrm{hit}}(i)}{\mathfrak{d}(i)}$. From Lemma 8.6.1,

$$\mathbf{E}[U] \leq \sum_{i=1}^{i_0} \left(\prod_{k=i}^{i_0} \frac{t_{\mathrm{hit}}(k)}{\mathfrak{d}(k)}\right) \left(\prod_{k=i_0+1}^{n} \frac{t_{\mathrm{hit}}(k)}{\mathfrak{d}(k)}\right) + \sum_{i=i_0+1}^{n} \prod_{k=i}^{n} \frac{t_{\mathrm{hit}}(k)}{\mathfrak{d}(k)}$$
$$\leq C^{-(n-i_0)} \sum_{i=1}^{i_0} K^{i-i_0+1} + \sum_{i=i_0+1}^{n} C^{-(n-i+1)}$$
$$= C^{-(n-i_0)} \sum_{i=1}^{i_0} K^i + \sum_{i=1}^{n-i_0} C^{-i}$$
$$\leq C^{-(n-i_0)} \frac{K(1-K^{i_0})}{1-K} + \frac{1}{C-1}.$$

Then we can take some $n_0 = n_0(\epsilon)$ satisfying $C^{-(n-i_0)} \frac{K(1-K^{i_0})}{1-K} \leq \epsilon/2$. Hence for any $n \geq n_0$, $\mathbf{E}[U] \leq \epsilon$ and we obtain the claim. □

8.6.2 Upper bound for random walks with small mixing times

In this section we show the following generalized version of Theorem 8.2.3.

Theorem 8.6.2. *Suppose that $P^{(i)}$ is reversible and lazy in $R = (\mathfrak{d}, (G^{(i)})_{i\in\mathbb{N}}, (P^{(i)})_{i\in\mathbb{N}})$. Let $N > 0$ be an arbitrary positive number. If $\mathfrak{d}(i) \geq \frac{t_{\mathrm{hit}}(i)}{N} + 2t_{\mathrm{mix}}(i)$ for all $i \in [n]$, then $\mathbf{E}[U] \leq 8N + 32$.*

Proof of Theorem 8.2.3. For all i, it is straight forward to see that

$$\mathfrak{d}(i) \geq \frac{Ct_{\mathrm{hit}}(i)}{i^\gamma} + \frac{2Ct_{\mathrm{hit}}(i)}{i^\gamma} \geq \frac{t_{\mathrm{hit}}(i)}{n^\gamma/C} + 2t_{\mathrm{mix}}(i)$$

from assumptions. Taking $N = n^\gamma/C$ in Theorem 8.6.2, we obtain the claim. □

To show Theorem 8.6.2, we introduce following two lemmas. The first one generalizes Lemma 8.5.1(i). The second one is a variant of Lemma 8.6.1.

Lemma 8.6.3. *For $f, h : \mathbb{N} \to \mathbb{N}$ and $n \in \mathbb{N}$, let*

$$S(n) := \sum_{k=1}^{n} \prod_{i=k}^{n} \left(1 - \frac{1}{h(i)}\right)^{f(i)}.$$

Let $N > 0$ be an arbitrary number. If $f(i) \geq \frac{h(i)}{N}$ for all $i \in [n]$, then $S(n) \leq N$.

Proof. It is easy to check that

$$\begin{aligned} S(n) &\leq \sum_{k=1}^{n} \prod_{i=k}^{n} \exp\left(-\frac{f(i)}{h(i)}\right) = \sum_{k=1}^{n} \exp\left(-\sum_{i=k}^{n} \frac{f(i)}{h(i)}\right) \leq \sum_{k=1}^{n} \exp\left(-\frac{n+k-1}{N}\right) \\ &= \sum_{k=1}^{n} \exp\left(-\frac{k}{N}\right) \leq \frac{\mathrm{e}^{-1/N}}{1-\mathrm{e}^{-1/N}} = \frac{1}{\mathrm{e}^{1/N}-1} \leq N. \end{aligned}$$

Note that we use $1 + x \leq \mathrm{e}^x$ in the first and the last inequalities. □

Lemma 8.6.4. *For any $R = (\mathfrak{d}, (G^{(i)})_{i\in\mathbb{N}}, (P^{(i)})_{i\in\mathbb{N}})$ and any function $s : \mathbb{N} \to \mathbb{N}$ such that $s(i) < \mathfrak{d}(i)$ holds for all i, we have*

$$\mathbf{E}[U] \leq \sum_{k=2}^{n} \prod_{i=k}^{n} \max_{u\in V^{(i)}} \left(\sum_{v\in V^{(i)}} \left((P^{(i)})^{s(i)}\right)_{u,v} \mathbf{Pr}\left[\tau_{v_k}^{(i)} > \mathfrak{d}(i) - s(i) \middle| X_0^{(i)} = v\right] \right).$$

Proof. From Lemma 8.6.1, we evaluate

$$\mathbf{Pr}\left[\tau_{v_k}^{(i)} > \mathfrak{d}(i) \middle| X_0^{(i)} = u\right] = \sum_{v\in V^{(i)}} \mathbf{Pr}\left[\tau_{v_k}^{(i)} > \mathfrak{d}(i), X_{s(i)}^{(i)} = v \middle| X_0^{(i)} = u\right].$$

Fix $k \geq 2$ and i satisfying $k \leq i \leq n$. For any $u, v \in V^{(i)}$, observe that

$$\begin{aligned} &\mathbf{Pr}\left[\tau_{v_k}^{(i)} > \mathfrak{d}(i), X_{s(i)}^{(i)} = v \middle| X_0^{(i)} = u\right] \\ &= \mathbf{Pr}\left[\tau_{v_k}^{(i)} > \mathfrak{d}(i) \middle| X_{s(i)}^{(i)} = v, X_0^{(i)} = u, \tau_{v_k}^{(i)} > s(i)\right] \mathbf{Pr}\left[X_{s(i)}^{(i)} = v, \tau_{v_k}^{(i)} > s(i) \middle| X_0^{(i)} = u\right] \\ &= \mathbf{Pr}\left[\tau_{v_k}^{(i)} > \mathfrak{d}(i) - s(i) \middle| X_0^{(i)} = v\right] \mathbf{Pr}\left[X_{s(i)}^{(i)} = v, \tau_{v_k}^{(i)} > s(i) \middle| X_0^{(i)} = u\right] \end{aligned}$$

holds. In the second inequality, we used the Markov property. Then, since

$$\mathbf{Pr}\left[X_{s(i)}^{(i)} = v, \tau_{v_k}^{(i)} > s(i) \middle| X_0^{(i)} = u\right] \leq \mathbf{Pr}\left[X_{s(i)}^{(i)} = v \middle| X_0^{(i)} = u\right] = ((P^{(i)})^{s(i)})_{u,v},$$

we obtain

$$\begin{aligned} \mathbf{Pr}\left[\tau_{v_k}^{(i)} > \mathfrak{d}(i) \middle| X_0^{(i)} = u\right] &= \sum_{v\in V^{(i)}} \mathbf{Pr}\left[\tau_{v_k}^{(i)} > \mathfrak{d}(i), X_{s(i)}^{(i)} = v \middle| X_0^{(i)} = u\right] \\ &\leq \sum_{v\in V^{(i)}} \left((P^{(i)})^{s(i)}\right)_{u,v} \mathbf{Pr}\left[\tau_{v_k}^{(i)} > \mathfrak{d}(i) - s(i) \middle| X_0^{(i)} = v\right] \end{aligned} \tag{8.12}$$

for any $u \in V^{(i)}$. Combining Lemma 8.6.1 and (8.12), we obtain the claim. □

Proof of Theorem 8.6.2. We use Lemma 8.6.4 with letting $s(i) = 2t_{\mathrm{mix}}(i)$. If $P^{(i)}$ is reversible, for any $i \in [n]$ and $u, v \in V^{(i)}$, there is a transition matrix $\hat{P}^{(i)} \in [0,1]^{V^{(i)} \times V^{(i)}}$ satisfying

$$\left((P^{(i)})^{2t_{\mathrm{mix}}(i)}\right)_{u,v} = \frac{1}{4}\pi^{(i)}(v) + \frac{3}{4}(\hat{P}^{(i)})_{u,v} \tag{8.13}$$

holds (see, e.g., p.338 of [LP17]). Hence it holds for any $u \in V^{(i)}$ that

$$\begin{aligned}
&\sum_{v \in V^{(i)}} \left((P^{(i)})^{2t_{\mathrm{mix}}(i)}\right)_{u,v} \mathbf{Pr}\left[\tau_{v_k}^{(i)} > \mathfrak{d}(i) - 2t_{\mathrm{mix}}(i) \middle| X_0^{(i)} = v\right] \\
&= \frac{1}{4} \sum_{v \in V^{(i)}} \pi^{(i)}(v)\, \mathbf{Pr}\left[\tau_{v_k}^{(i)} > \mathfrak{d}(i) - 2t_{\mathrm{mix}}(i) \middle| X_0^{(i)} = v\right] \\
&\quad + \frac{3}{4} \sum_{v \in V^{(i)}} (\hat{P}^{(i)})_{u,v}\, \mathbf{Pr}\left[\tau_{v_k}^{(i)} > \mathfrak{d}(i) - 2t_{\mathrm{mix}}(i) \middle| X_0^{(i)} = v\right] \\
&\leq \frac{1}{4} \exp\left(-\frac{\mathfrak{d}(i) - 2t_{\mathrm{mix}}(i)}{t_{\mathrm{hit}}(i)}\right) + \frac{3}{4} \leq \frac{1}{4}\exp\left(-\frac{1}{N}\right) + \frac{3}{4}.
\end{aligned} \tag{8.14}$$

Here, we used Corollary 8.4.2 in the inequality above. Now, for a positive integer L, consider a random variable $X \sim \mathrm{Bin}(L, 1/4)$. Here, $\mathrm{Bin}(m,p)$ denotes the binomial distribution of m trials with success probability p. Then, it is straightforward to see that

$$\begin{aligned}
\left(\frac{1}{4}\exp\left(-\frac{1}{N}\right) + \frac{3}{4}\right)^L &= \sum_{i=0}^{L} \binom{L}{i} \left(\frac{1}{4}\exp\left(-\frac{1}{N}\right)\right)^i \left(\frac{3}{4}\right)^{L-i} \\
&= \sum_{i=0}^{L} \exp\left(-\frac{i}{N}\right) \mathbf{Pr}\left[X = i\right] \\
&\leq \sum_{i=0}^{\lfloor L/8 \rfloor} \exp\left(-\frac{i}{N}\right) \mathbf{Pr}\left[X = i\right] + \sum_{i=\lceil L/8 \rceil}^{L} \exp\left(-\frac{i}{N}\right) \mathbf{Pr}\left[X = i\right] \\
&\leq \mathbf{Pr}\left[X \leq \frac{L}{8}\right] + \exp\left(-\frac{L}{8N}\right) \leq \exp\left(-\frac{L}{32}\right) + \exp\left(-\frac{L}{8N}\right).
\end{aligned} \tag{8.15}$$

The last inequality follows since

$$\mathbf{Pr}\left[X \leq \frac{L}{8}\right] = \mathbf{Pr}\left[X \leq \frac{\mathbf{E}[X]}{2}\right] \leq \exp\left(-\frac{\mathbf{E}[X]}{8}\right) = \exp\left(-\frac{L}{32}\right)$$

holds from the Chernoff inequality (Proposition 2.5.5). By combining Lemma 8.6.4 and (8.14) and (8.15), we obtain

$$\begin{aligned}
\mathbf{E}[U] &\leq \sum_{k=1}^{n} \left(\frac{1}{4}\exp\left(-\frac{1}{N}\right) + \frac{3}{4}\right)^{n-k+1} \\
&\leq \sum_{k=1}^{n} \left(\exp\left(-\frac{n-k+1}{32}\right) + \exp\left(-\frac{n-k+1}{8N}\right)\right) \\
&= \sum_{k=1}^{n} \exp\left(-\frac{k}{32}\right) + \sum_{k=1}^{n} \exp\left(-\frac{k}{8N}\right) \leq 32 + 8N.
\end{aligned}$$

□

8.6.3 Upper bounds for simple or symmetric random walks

This section is devoted proving Theorem 8.6.5, which is a generalized version of Theorems 8.2.4 and 8.2.5.

Theorem 8.6.5. *Suppose that $P^{(i)}$ is reversible and lazy in $R = (\mathfrak{d}, (G^{(i)})_{i\in\mathbb{N}}), (P^{(i)})_{i\in\mathbb{N}})$. Let $r_i = \max_{v \in V^{(i-1)}} \frac{\pi^{(i-1)}(v)}{\pi^{(i)}(v)}$ for $1 < i \leq n$. Let N be an arbitrary number. If $\mathfrak{d}(i) \geq \left(\frac{1}{N} + \frac{i(r_i-1)+1}{2i}\right) t_{\mathrm{hit}}^{(i)}$ for all i, then $\mathbf{E}[U] \leq N\sqrt{\max_{1<i\leq n} i(r_i - 1) + 1}$.*

Proof of Theorem 8.2.4. Let $d_v^{(i)}$ denote the degree of a vertex $v \in V^{(i)}$ at round i. Then, for all $v \in V^{(i)}$,

$$\frac{\pi^{(i-1)}(v)}{\pi^{(i)}(v)} = \frac{d_v^{(i-1)}}{2|E^{(i-1)}|}\frac{2|E^{(i)}|}{d_v^{(i)}} \le \frac{|E^{(i)}|}{|E^{(i-1)}|}$$

Note that $d_v^{(i-1)} \le d_v^{(i)}$ holds from our assumption. Combining the assumptions on $\mathfrak{d}(i)$ and $E^{(i)}$, we have $\mathfrak{d}(i) \ge \frac{t_{\mathrm{hit}}(i)}{i^\gamma/C} + \frac{L+1}{2i}t_{\mathrm{hit}}(i) \ge \frac{t_{\mathrm{hit}}(i)}{n^\gamma/C} + \frac{L+1}{2i}t_{\mathrm{hit}}(i)$. Thus we obtain the claim by taking $N = n^\gamma/C$ in Theorem 8.6.5. □

Proof of Theorem 8.2.5. Since $P^{(i)}$ is symmetric, $\pi^{(i)}(v) = \frac{1}{i}$ and thus $r_i = \frac{i}{i-1} \le 1 + \frac{2}{i}$ for all $i > 1$. From the assumption of Theorem 8.2.5, $\mathfrak{d}(i) \ge \frac{t_{\mathrm{hit}}(i)}{i^\gamma/C} + \frac{2t_{\mathrm{hit}}(i)}{i} \ge \frac{t_{\mathrm{hit}}(i)}{n^\gamma/C} + \frac{t_{\mathrm{hit}}(i)(2+1)}{2i}$ for all $i > 1$. Thus we obtain the claim by taking $N = n^\gamma/C$ in Theorem 8.6.5. □

For a matrix $M \in \mathbb{R}^{V\times V}$ let $\lambda_j(M)$ denote the j-th largest (in absolute value) eigenvalue of M. For any round $1 < \ell \le n$ and $0 \le t \le \mathfrak{d}(\ell)$, define a probability vector $\nu_t^{(\ell)} \in [0,1]^{V^{(\ell)}}$ by

$$\nu_t^{(\ell)}(v) = \mathbf{Pr}[X_t^{(\ell)} = v] \tag{8.16}$$

for all $v \in V^{(\ell)}$. For any rounds k, ℓ satisfying $k - 1 \le \ell \le n - 1$, define $\mu_{v_k}^{(\ell)} \in [0,1]^{V^{(\ell)}}$ by

$$\mu_{v_k}^{(\ell)}(v) = \mathbf{Pr}\left[\bigwedge_{i=\ell+1}^{n}\left\{\tau_{v_k}^{(i)} > \mathfrak{d}(i)\right\}\middle| X_{\mathfrak{d}(\ell)}^{(\ell)} = v\right] \tag{8.17}$$

for all $v \in V^{(\ell)}$. For $\ell = n$, we define $\mu_{v_k}^{(n)} := \mathbb{1}^{(n)}$. Here, recall the notation of Section 8.4.2. Observe

$$\mathbf{E}[U] = \sum_{k=2}^{n}\sum_{v\in V^{(k-1)}} \nu_{\mathfrak{d}(k-1)}^{(k-1)}(v)\mu_{v_k}^{(k-1)}(v).$$

Then, combining the Cauchy-Schwarz inequality, (8.5), (8.16) and (8.17), we have

$$\begin{aligned}
\mathbf{E}[U] &= \sum_{k=2}^{n}\sum_{v\in V^{(k-1)}} \frac{\nu_{\mathfrak{d}(k-1)}^{(k-1)}(v)}{\sqrt{\pi^{(k-1)}(v)}}\cdot \mu_{v_k}^{(k-1)}(v)\sqrt{\pi^{(k-1)}(v)} \\
&\le \sum_{k=2}^{n}\sqrt{\sum_{v\in V^{(k-1)}}\frac{\nu_{\mathfrak{d}(k-1)}^{(k-1)}(v)^2}{\pi^{(k-1)}(v)}\sum_{v\in V^{(k-1)}}\pi^{(k-1)}(v)\mu_{v_k}^{(k-1)}(v)^2} \\
&= \sum_{k=2}^{n}\left\|\frac{\nu_{\mathfrak{d}(k-1)}^{(k-1)}}{\pi^{(k-1)}}\right\|_{2,\pi^{(k-1)}}\left\|\mu_{v_k}^{(k-1)}\right\|_{2,\pi^{(k-1)}} \qquad (8.18)\\
&= \sum_{k=2}^{n}\sqrt{1 + \left\|\frac{\nu_{\mathfrak{d}(k-1)}^{(k-1)}}{\pi^{(k-1)}} - \mathbb{1}^{(k-1)}\right\|_{2,\pi^{(k-1)}}^2}\left\|\mu_{v_k}^{(k-1)}\right\|_{2,\pi^{(k-1)}}. \qquad (8.19)
\end{aligned}$$

The rest of this section is devoted to proving the following bounds, which imply Theorem 8.6.5.

Lemma 8.6.6. *Consider an RWoGG $R = (\mathfrak{d}, (G^{(i)})_{i\in\mathbb{N}}, (P^{(i)})_{i\in\mathbb{N}})$ such that each $P^{(i)}$ is reversible and lazy. Let $r_i = \max_{v\in V^{(i-1)}}\frac{\pi^{(i-1)}(v)}{\pi^{(i)}(v)}$ for $1 < i \le n$. If $\mathfrak{d}(i) \ge \frac{i(r_i-1)+1}{2i(1-\lambda_2(P^{(i)}))}$, then $\left\|\frac{\nu_{\mathfrak{d}(k)}^{(k)}}{\pi^{(k)}} - \mathbb{1}^{(k)}\right\|_{2,\pi^{(k)}}^2 < \max_{1<i\le n} i(r_i - 1)$ for all $k \in [n]$.*

Lemma 8.6.7. *Consider an RWoGG $R = (\mathfrak{d}, (G^{(i)})_{i\in\mathbb{N}}, (P^{(i)})_{i\in\mathbb{N}})$ such that each $P^{(i)}$ is reversible and lazy. Let $r_i = \max_{v\in V^{(i-1)}}\frac{\pi^{(i-1)}(v)}{\pi^{(i)}(v)}$ for $1 < i \le n$. Let N be an arbitrary positive number such that $\mathfrak{d}(i) \ge \left(\frac{1}{N} + \frac{r_i-1}{2}\right)t_{\mathrm{hit}}(i)$ for all $1 < i \le n$. Then $\sum_{k=2}^{n}\left\|\mu_{v_k}^{(k-1)}\right\|_{2,\pi^{(k-1)}} \le N$.*

Proof of Theorem 8.6.5. Suppose $\mathfrak{d}(i) \geq \frac{t_{\mathrm{hit}}(i)}{N} + \frac{(i(r_i-1)+1)t_{\mathrm{hit}}(i)}{2i}$ for all $1 < i \leq n$. Then, $\mathfrak{d}(i) \geq \frac{i(r_i-1)+1}{2i(1-\lambda_2(P^{(i)}))}$ from Lemma 8.4.6. Furthermore, $\mathfrak{d}(i) \geq \frac{t_{\mathrm{hit}}(i)}{N} + \frac{r_i-1}{2} t_{\mathrm{hit}}(i)$. Thus applying Lemmas 8.6.6 and 8.6.7 to (8.19),

$$\mathbf{E}[U] \leq \sum_{k=2}^{n} \sqrt{\max_{1<i\leq n} i(r_i-1)+1} \left\| \mu_{v_k}^{(k-1)} \right\|_{2,\pi^{(k-1)}} \leq N \sqrt{\max_{1<i\leq n} i(r_i-1)+1}$$

and we obtain the claim. □

Now it suffices to prove Lemmas 8.6.6 and 8.6.7. To this end, we show the following.

Lemma 8.6.8. *Consider an RWoGG $R = (\mathfrak{d}, (G^{(i)})_{i\in\mathbb{N}}, (P^{(i)})_{i\in\mathbb{N}})$ such that each $P^{(i)}$ is reversible and lazy. Let $r_i = \max_{v\in V^{(i-1)}} \frac{\pi^{(i-1)}(v)}{\pi^{(i)}(v)}$ for $1 < i \leq n$. Then for any round $1 \leq k \leq n$,*

$$\left\| \frac{\nu^{(k)}_{\mathfrak{d}(k)}}{\pi^{(k)}} - \mathbb{1}^{(k)} \right\|^2_{2,\pi^{(k)}} \leq \sum_{i=2}^{k} \left(\prod_{j=i}^{k} r_j \lambda_2(P^{(j)})^{2\mathfrak{d}(j)} \right) \left(1 - \frac{1}{r_i}\right).$$

Proof of Lemma 8.6.8. To obtain the claim, we show the following recurrence inequality:

$$\left\| \frac{\nu^{(\ell)}_{\mathfrak{d}(\ell)}}{\pi^{(\ell)}} - \mathbb{1}^{(\ell)} \right\|^2_{2,\pi^{(\ell)}} \leq r_\ell \lambda_2(P^{(\ell)})^{2\mathfrak{d}(\ell)} \left\| \frac{\nu^{(\ell-1)}_{\mathfrak{d}(\ell-1)}}{\pi^{(\ell-1)}} - \mathbb{1}^{(\ell-1)} \right\|^2_{2,\pi^{(\ell-1)}} + (r_\ell - 1)\lambda_2(P^{(\ell)})^{2\mathfrak{d}(\ell)}. \tag{8.20}$$

Write $x_\ell = \left\| \frac{\nu^{(\ell)}_{\mathfrak{d}(\ell)}}{\pi^{(\ell)}} - \mathbb{1}^{(\ell)} \right\|^2_{2,\pi^{(\ell)}}$, $c_\ell = r_\ell \lambda_2(P^{(\ell)})^{2\mathfrak{d}(\ell)}$ and $d_\ell = (r_\ell - 1)\lambda_2(P^{(\ell)})^{2\mathfrak{d}(\ell)}$ for notational convenience. If (8.20) holds for any $\ell > 1$, applying (8.20) repeatedly yields

$$x_k \leq c_k x_{k-1} + d_k \leq c_k c_{k-1} x_{k-2} + c_k d_{k-1} + d_k \leq \cdots \leq \left(\prod_{i=2}^{k} c_i \right) x_1 + \sum_{i=2}^{k} \left(\prod_{j=i+1}^{k} c_j \right) d_i.$$

Since $x_1 = \left\| \frac{\nu^{(1)}_{\mathfrak{d}(1)}}{\pi^{(1)}} - \mathbb{1}^{(1)} \right\|^2_{2,\pi^{(1)}} = 0$, we obtain the claim.

Now we show (8.20). From the reversibility of $P^{(\ell)}$, it is easy to see that, for all $v \in V^{(\ell)}$,

$$\left(\frac{\nu^{(\ell)}_t}{\pi^{(\ell)}} \right)(v) = \frac{\sum_{u\in V^{(\ell)}} \nu^{(\ell)}_0(u) \left((P^{(\ell)})^t\right)_{u,v}}{\pi^{(\ell)}(v)} = \sum_{u\in V^{(\ell)}} \frac{\nu^{(\ell)}_0(u) \left((P^{(\ell)})^t\right)_{v,u}}{\pi^{(\ell)}(u)} = \left((P^{(\ell)})^t \frac{\nu^{(\ell)}_0}{\pi^{(\ell)}} \right)(v). \tag{8.21}$$

From (8.21) and Lemma 8.4.5, it holds that

$$\left\| \frac{\nu^{(\ell)}_{\mathfrak{d}(\ell)}}{\pi^{(\ell)}} - \mathbb{1}^{(\ell)} \right\|^2_{2,\pi^{(\ell)}} \leq \lambda_2(P^{(\ell)})^{2\mathfrak{d}(\ell)} \left\| \frac{\nu^{(\ell)}_0}{\pi^{(\ell)}} - \mathbb{1}^{(\ell)} \right\|^2_{2,\pi^{(\ell)}} = \lambda_2(P^{(\ell)})^{2\mathfrak{d}(\ell)} \left(\left\| \frac{\nu^{(\ell)}_0}{\pi^{(\ell)}} \right\|^2_{2,\pi^{(\ell)}} - 1 \right). \tag{8.22}$$

Note that, for $v_\ell \in V^{(\ell)} \setminus V^{\ell-1}$, it holds that $\nu^{(\ell)}_0(v_\ell) = \mathbf{Pr}[X^{(\ell)}_0 = v_\ell] = 0$. Therefore, we have

$$\begin{aligned} \left\| \frac{\nu^{(\ell)}_0}{\pi^{(\ell)}} \right\|^2_{2,\pi^{(\ell)}} &= \sum_{v\in V^{(\ell-1)}} \pi^{(\ell)}(v) \frac{\nu^{(\ell)}_0(v)^2}{\pi^{(\ell)}(v)^2} = \sum_{v\in V^{(\ell-1)}} \frac{\pi^{(\ell-1)}(v)}{\pi^{(\ell)}(v)} \pi^{(\ell-1)}(v) \frac{\nu^{(\ell-1)}_{\mathfrak{d}(\ell-1)}(v)^2}{\pi^{(\ell-1)}(v)^2} \\ &\leq r_\ell \sum_{v\in V^{(\ell-1)}} \pi^{(\ell-1)}(v) \frac{\nu^{(\ell-1)}_{\mathfrak{d}(\ell-1)}(v)^2}{\pi^{(\ell-1)}(v)^2} = r_\ell \left\| \frac{\nu^{(\ell-1)}_{\mathfrak{d}(\ell-1)}}{\pi^{(\ell-1)}} \right\|^2_{2,\pi^{(\ell-1)}}. \end{aligned} \tag{8.23}$$

The claim (8.20) follows from (8.22) and (8.23). □

Proof of Lemma 8.6.6. Observe that $\log\left(r_j(\frac{j+1}{j})\right) = \log(1+(r_j-1)) + \log(1+\frac{1}{j}) \le (r_j-1)+\frac{1}{j}$ and thus $2\mathfrak{d}(j) \ge \frac{\log\left(r_j(\frac{j+1}{j})\right)}{1-\lambda_2(P^{(j)})}$. Hence we obtain

$$\lambda_2(P^{(j)})^{2\mathfrak{d}(j)} \le \left(1-\left(1-\lambda_2(P^{(j)})\right)\right)^{\frac{\log\left(r_j(\frac{j+1}{j})\right)}{1-\lambda_2(P^{(j)})}} \le \frac{1}{r_j}\cdot\frac{j}{j+1}.$$

From Lemma 8.6.8, we have

$$\begin{aligned}\sum_{i=2}^{k}\left(\prod_{j=i}^{k} r_j\lambda_2(P^{(j)})^{2\mathfrak{d}(j)}\right)\left(1-\frac{1}{r_i}\right) &\le \sum_{i=2}^{k}\left(\prod_{j=i}^{k}\frac{j}{j+1}\right)\frac{r_i-1}{r_i} \le \sum_{i=2}^{k}\frac{i}{k+1}(r_i-1)\\ &\le \max_{1<i\le n} i(r_i-1)\frac{k-1}{k+1} < \max_{1<i\le n} i(r_i-1).\end{aligned}$$

□

Now we prove Lemma 8.6.7. We begin with showing the following lemma.

Lemma 8.6.9. *Consider an RWoGG $R=(\mathfrak{d},(G^{(i)})_{i\in\mathbb{N}},(P^{(i)})_{i\in\mathbb{N}})$ such that each $P^{(i)}$ is reversible and lazy. Let $r_i = \max_{v\in V^{(i-1)}}\frac{\pi^{(i-1)}(v)}{\pi^{(i)}(v)}$ for $1<i\le n$. Then, for any $1<k\le n$,*

$$\left\|\mu_{v_k}^{(k-1)}\right\|_{2,\pi^{(k-1)}} \le \prod_{i=k}^{n}\sqrt{r_i}\left(1-\frac{1}{t_{\mathrm{hit}}(i)}\right)^{\mathfrak{d}(i)}.$$

Proof. For a transition matrix $P\in[0,1]^{V\times V}$ and a vertex $w\in V$, define $P_{\overline{w}}\in[0,1]^{V\times V}$ by

$$(P_{\overline{w}})_{u,v} = \begin{cases} P_{u,v} & (\text{if } u\ne w \text{ and } v\ne w)\\ 0 & (\text{otherwise})\end{cases}.$$

In other words, $(P_{\overline{w}})_{u,v} = P_{u,v}\mathbb{1}_{u\ne w}\mathbb{1}_{v\ne w}$ for $u,v\in V$. Note that $P_{\overline{w}}$ is a substochastic matrix, i.e., $\sum_{v\in V}(P_{\overline{w}})_{u,v}\le 1$ for any $u\in V$. Observe for any $u,v\in V$ and $T>0$ that

$$\begin{aligned}(P_{\overline{w}}^T)_{u,v} &= \sum_{v_1\in V\setminus\{w\}}\cdots\sum_{v_{T-1}\in V\setminus\{w\}}\mathbb{1}_{u\ne w}P_{u,v_1}P_{v_1,v_2}\cdots P_{v_{T-1},v}\mathbb{1}_{v\ne w}\\ &= \mathbf{Pr}\left[\tau_w>T, X_T=v\middle|X_0=u\right].\end{aligned} \tag{8.24}$$

Here, $(X_t)_{t\in\mathbb{Z}_{\ge 0}}$ denotes a random walk according to P and τ_w denotes the hitting time of $(X_t)_{t\in\mathbb{Z}_{\ge 0}}$ to w. In other words, $(P_{\overline{w}}^T)_{u,v}$ denotes the probability that the random walk of length T ends up at v starting from u without visiting w.

Fix $k>1$. For ease of notation, we write $\mu^{(\ell)}=\mu_{v_k}^{(\ell)}$ and $Q^{(\ell)}=(P_{\overline{v}_k}^{(\ell)})^{\mathfrak{d}(\ell)}$ for $k-1\le\ell\le n-1$ (see (8.17) for the definition of $\mu_{v_k}^{(\ell)}$). We begin with observing the following recurrence equation: For all $k-1\le\ell\le n-1$ and $v\in V^{(\ell)}$, it holds that

$$\mu^{(\ell)}(v) = \left(Q^{(\ell+1)}\mu^{(\ell+1)}\right)(v). \tag{8.25}$$

Indeed, for any $u_\ell\in V^{(\ell)}$, combining (8.10) and (8.24) yields

$$\begin{aligned}\mu^{(\ell)}(u_\ell) &= \sum_{u_{\ell+1}\in V^{(\ell+1)}}\cdots\sum_{u_n\in V^{(n)}}\prod_{i=\ell+1}^{n}\left((P_{\overline{v}_k}^{(i)})^{\mathfrak{d}(i)}\right)_{u_{i-1},u_i}\\ &= \sum_{u_{\ell+1}\in V^{(\ell+1)}}Q^{(\ell+1)}_{u_\ell,u_{\ell+1}}\mu^{(\ell+1)}(u_{\ell+1})\\ &= \left(Q^{(\ell+1)}\mu^{(\ell+1)}\right)(u_\ell).\end{aligned}$$

Using (8.25) and Corollary 8.4.4, we obtain

$$\begin{aligned}\left\|\mu^{(\ell)}\right\|_{2,\pi^{(\ell)}}^2 &= \sum_{v\in V^{(\ell)}} \pi^{(\ell)}(v)\mu^{(\ell)}(v)^2 = \sum_{v\in V^{(\ell)}} \frac{\pi^{(\ell)}(v)}{\pi^{(\ell+1)}(v)}\pi^{(\ell+1)}(v)\left(Q^{(\ell+1)}\mu^{(\ell+1)}\right)(v)^2 \\ &\le r_{\ell+1}\sum_{v\in V^{(\ell+1)}} \pi^{(\ell+1)}(v)\left(Q^{(\ell+1)}\mu^{(\ell+1)}\right)(v)^2 = r_{\ell+1}\left\|Q^{(\ell+1)}\mu^{(\ell+1)}\right\|_{2,\pi^{(\ell+1)}}^2 \\ &\le r_{\ell+1}\lambda_1(Q^{(\ell+1)})^2\left\|\mu^{(\ell+1)}\right\|_{2,\pi^{(\ell+1)}}^2. \end{aligned} \tag{8.26}$$

Hence applying (8.26) repeatedly, it holds that

$$\left\|\mu^{(\ell)}\right\|_{2,\pi^{(\ell)}}^2 \le \prod_{i=\ell+1}^{n} r_i\lambda_1(Q^{(i)})^2. \tag{8.27}$$

From the definition of $Q^{(i)}$ and $P^{(i)}_{\overline{v}_k}$, Lemma 8.4.3 implies

$$\lambda_1(Q^{(i)}) = \lambda_1(P^{(i)}_{\overline{v}_k})^{\mathfrak{d}(i)} \le \left(1-\frac{1}{t_{\mathrm{hit}}(i)}\right)^{\mathfrak{d}(i)}. \tag{8.28}$$

Thus we obtain the claim from (8.27) and (8.28). □

Proof of Lemma 8.6.7. Since $\log\sqrt{r_i} = \frac{1}{2}\log r_i = \frac{1}{2}\log(1+(r_i-1)) \le \frac{r_i-1}{2}$, we have

$$\sqrt{r_i}\left(1-\frac{1}{t_{\mathrm{hit}}(i)}\right)^{\mathfrak{d}(i)} \le \left(1-\frac{1}{t_{\mathrm{hit}}(i)}\right)^{\mathfrak{d}(i)-t_{\mathrm{hit}}(i)\log\sqrt{r_i}} \le \left(1-\frac{1}{t_{\mathrm{hit}}(i)}\right)^{\mathfrak{d}(i)-\frac{r_i-1}{2}t_{\mathrm{hit}}(i)}. \tag{8.29}$$

Thus combining Lemma 8.6.9 and (8.29),

$$\sum_{k=2}^{n}\left\|\mu_{v_k}^{(k-1)}\right\|_{2,\pi^{(k-1)}} \le \sum_{k=2}^{n}\prod_{i=k}^{n}\left(1-\frac{1}{t_{\mathrm{hit}}(i)}\right)^{\mathfrak{d}(i)-\frac{r_i-1}{2}t_{\mathrm{hit}}(i)} \le N.$$

We invoke Lemma 8.6.3 in the last inequality. □

8.7 A Lower Bound for a Growing Path

This section is devoted to the proof of Theorem 8.2.9. We will use the following well-known inequality.

Lemma 8.7.1 (The Kolmogorov inequality; Theorem 2.5.5 of [Dur19])**.** *Let $Z_1,\dots,Z_n$ be i.i.d. random variables such that $\mathbf{E}[Z_i]=0$ and $\mathbf{Var}[Z_i]<\infty$. Let $S_i := \sum_{j=1}^{i} Z_i$. Then,*

$$\mathbf{Pr}\left[\max_{1\le j\le n}|S_j|\ge M\right] \le \frac{\mathbf{Var}[S_n]}{M^2}.$$

Let $L,R\in[n]$ be parameters satisfying $L<R$. For a vertex $v\in V^{(n)}$, let $\mathcal{E}(v)$ be the event that $v\notin\bigcup_{i=1}^{n}\bigcup_{t=0}^{\mathfrak{d}(i)}\{X_t^{(i)}\}$. In other words, $\mathcal{E}(v)$ means that the walker does not visit the vertex v during the walk. For two vertices $v_i,v_j\in V^{(n)}$, we write $v_i\preceq v_j$ if $i\le j$. Note that, for any two vertices $u\preceq v$ and any round $k\in[n]$, it holds that $\mathbf{Pr}[\mathcal{E}(v)|X_0^{(k)}\preceq u]\ge\mathbf{Pr}[\mathcal{E}(v)|X_0^{(k)}=u]$. Then, we have

$$\begin{aligned}\mathbf{E}[U] &= \sum_{k=1}^{n}\mathbf{Pr}[\mathcal{E}(v_k)] \ge \sum_{k=R}^{n}\mathbf{Pr}[\mathcal{E}(v_k)] \ge \sum_{k=R}^{n}\mathbf{Pr}\left[\mathcal{E}(v_k)\wedge X_0^{(k)}\preceq v_L\right] \\ &= \sum_{k=R}^{n}\mathbf{Pr}\left[\mathcal{E}(v_k)\middle|X_0^{(k)}\preceq v_L\right]\mathbf{Pr}[X_0^{(k)}\preceq v_L] \\ &\ge (n-R)\,\mathbf{Pr}\left[\mathcal{E}(v_R)\middle|X_0^{(R)}=v_L\right]\min_{R\le k\le n}\left\{\mathbf{Pr}\left[X_0^{(k)}\preceq v_L\right]\right\}. \end{aligned} \tag{8.30}$$

We will determine the parameters R and L such that $n-R=\Omega(n^\gamma)$, $\mathbf{Pr}\left[\mathcal{E}(v_R)\middle|X_0^{(R)}=v_L\right]=\Omega(1/C)$ and $\mathbf{Pr}[X_0^{(k)}\le L]=\Omega(1)$ for all $R\le k\le n$. This yields the lower bound $\mathbf{E}[U]=\Omega(n^\gamma/C)$. For fixed parameter R, let $T:=\sum_{i=R}^{n}\mathfrak{d}(i)$ denote the number of steps of the walk during the last $n-R+1$ rounds.

Lemma 8.7.2. *Let $L, R \in \mathbb{N}$ be parameters satisfying $L < R$ and let $T := \sum_{i=R}^{n} \mathfrak{d}(i)$. Then, the following hold.*

(i) $\mathbf{Pr}\left[\mathcal{E}(v_R) \middle| X_0^{(R)} = v_L\right] \geq 1 - \frac{T}{4(R-L)^2}$, *and*

(ii) $\mathbf{Pr}[X_0^{(k)} \preceq v_L] \geq 1 - \frac{L}{n}$ *for all* $k \in [n]$.

Proof of (i). Let $(Z_t)_{t\in\mathbb{N}}$ be i.i.d. random variables sampled from the uniform distribution over $\{-1, +1\}$ and $S_c := \sum_{j=0}^{c} Z_j$ denote the sum. For a vertex $v_i \in V^{(n)}$, let $\mathrm{pos}(v_i) = i$ denote the position of v_i. Then the complementary event $\overline{\mathcal{E}(v_R)}$ conditioned on $X_0^{(R)} = v_L$ is equivalent to the event that $\max_{R\leq i\leq n, 0\leq j\leq \mathfrak{d}(i)}\{\mathrm{pos}(X_j^{(i)}) - \mathrm{pos}(X_0^{(R)})\} \geq R - L$. Moreover, $\max_{R\leq i\leq n, 0\leq j\leq \mathfrak{d}(i)} |\mathrm{pos}(X_j^{(i)}) - \mathrm{pos}(X_0^{(R)})|$ is dominated[2] by $\max_{1\leq c\leq T} |S_c|$ (recall $T = \sum_{i=R}^{n} \mathfrak{d}(i)$). This is because the distribution of $\mathrm{pos}(X_j^{(i)}) - \mathrm{pos}(X_{j-1}^{(i)})$ conditioned on $\mathrm{pos}(X_j^{(i)}) - \mathrm{pos}(X_{j-1}^{(i)}) \neq 0$ is uniform on $\{-1, +1\}$. Thus we obtain

$$\begin{aligned}
\mathbf{Pr}\left[\overline{\mathcal{E}(v_R)} \middle| X_0^{(R)} = v_L\right] &\leq \mathbf{Pr}\left[\max_{R\leq i\leq n, 0\leq j\leq \mathfrak{d}(i)} |\mathrm{pos}(X_j^{(i)}) - \mathrm{pos}(X_0^{(R)})| \geq R - L \middle| X_0^{(R)} = L\right] \\
&\leq \mathbf{Pr}[\max_{1\leq c\leq T} |S_c| \geq R - L] \\
&\leq \frac{\mathbf{Var}[S_T]}{(R-L)^2} = \frac{T}{4(R-L)^2}.
\end{aligned}$$

In the last inequality, we used the Kolmogorov inequality (Lemma 8.7.1). □

Proof of (ii). It suffices to show that

$$\mathbf{Pr}[X_0^{(k)} = v_i] \geq \mathbf{Pr}[X_0^{(k)} = v_{i+1}] \tag{8.31}$$

holds for any $1 \leq i \leq k-1$. To see this, assuming (8.31), we obtain

$$\frac{\mathbf{Pr}[X_0^{(k)} \preceq v_L]}{L} \geq \mathbf{Pr}[X_0^{(k)} = v_L] \geq \frac{1 - \mathbf{Pr}[X_0^{(k)} \leq L]}{n - L},$$

which implies the claim (ii). Here, in the second inequality, note that $\mathbf{Pr}[X_0^{(k)} = v_L] \geq \mathbf{Pr}[X_0^{(k)} = v_j]$ for all $j > L$ and thus, the average $\frac{1}{n-L}\sum_{j>L} \mathbf{Pr}[X_0^{(k)} = v_j]$ is at most $\mathbf{Pr}[X_0^{(k)} = v_L]$.

Now we prove the inequality (8.31). Let $x_j^{(i)} \in [0,1]^{V_i}$ denote the distribution of $X_j^{(i)}$. To simplify notations, for a vector $y \in [0,1]^{V^{(i)}}$, we write $y[u]$ for the u-th element of y. We call the distribution $y \in [0,1]^{V^{(i)}}$ *monotone* if $y[v_k] \geq y[v_{k+1}]$ holds for any $1 \leq k \leq i-1$. Our aim here is to prove that $x_0^{(k)}$ is monotone, which is equivalent to (8.31).

Indeed, we will prove a stronger statement: $x_j^{(i)}$ is monotone for any i and j. We prove this statement inductively. First, the vector $x_j^{(1)} = (1)$ is obviously monotone. Secondly, if $x_{\mathfrak{d}(i)}^{(i)}$ is monotone, so does $x_0^{(i+1)}$. To see this, note that $x_0^{(i+1)}$ is obtained by concatenating $x_{\mathfrak{d}(i)}^{(i)}$ with 0. More precisely, $x_0^{(i+1)} \in [0,1]^{i+1}$ satisfies

$$x_0^{(i+1)}[j] = \begin{cases} x_{\mathfrak{d}(i)}^{(i)}[j] & \text{if } 1 \leq j \leq i, \\ 0 & \text{if } j = i+1. \end{cases}$$

Finally, we check that $x_{j+1}^{(i)}$ is monotone if $x_j^{(i)}$ is monotone. From (8.2), we have

$$x_{j+1}^{(i)}[v_k] = \begin{cases} px_j^{(i)}[v_1] + (1-p)x_j^{(i)}[v_2] & \text{if } k = 1, \\ qx_j^{(i)}[v_{k-1}] + (1-2q)x_j^{(i)}[v_k] + qx_j^{(i)}[v_{k+1}] & \text{if } 1 < k < i, \\ (1-p)x_j^{(i)}[v_{i-1}] + px_j^{(i)}[v_i] & \text{if } k = i. \end{cases}$$

[2]For two random variables X and Y, we say X *dominates* Y if, for any $r \in \mathbb{R}$, $\mathbf{Pr}[X \geq r] \geq \mathbf{Pr}[Y \geq r]$ holds.

By the induction assumption, $x_j^{(i)}$ is monotone. Now we check that $x_j^{(i)}$ is monotone. For $k = 1$, since $p \geq q$, we have

$$x_{j+1}^{(i)}[v_1] - x_{j+1}^{(i)}[v_2] = (p-q)(x_j^{(i)}[v_1] - x_j^{(i)}[v_2]) + q(x_j^{(i)}[v_2] - x_j^{(i)}[v_3]) \geq 0.$$

For $1 < k < i-1$, since $q \leq \frac{1}{2}$, we have

$$\begin{aligned} x_{j+1}^{(i)}[v_i] - x_{j+1}^{(i)}[v_{i+1}] &= qx_j^{(i)}[v_{k-1}] + (1-3q)x_j^{(i)}[v_k] - (1-3q)x_j^{(i)}[v_{k+1}] - qx_j^{(i)}[v_{k+2}] \\ &\geq (1-2q)(x_j^{(i)}[v_k] - x_j^{(i)}[v_{k+1}]) \geq 0. \end{aligned}$$

Finally, for $k = i$, since $p \geq q$, we have

$$x_{j+1}^{(i)}[v_{i-1}] - x_{j+1}^{(i)}[v_i] = q(x_j^{(i)}[v_{i-2}] - x_j^{(i)}[v_{i-1}]) + (p-q)(x_j^{(i)}[v_{i-1}] - x_j^{(i)}[v_i]) \geq 0.$$

Therefore $x_{j+1}^{(i)}$ is monotone. □

Now we are ready to prove Corollary 8.2.10. Recall $\mathfrak{d}(i) \leq Ci^{2-\gamma}$. Fix a small positive constant ϵ such that $\epsilon < \min\{1/C, 0.1\}$. Set $R := n - \epsilon n^\gamma$ and $L := R - 0.6n \in [0.3n, 0.4n]$. Then we have $T \leq (n-R)\mathfrak{d}(n) \leq C\epsilon n^2 \leq n^2$ and thus $1 - \frac{T}{4(R-L)^2} \geq 1 - \frac{1}{4\times 0.36} > 0.3$ and $1 - \frac{L}{n} \geq 0.6$. Then, from (8.30) and Lemma 8.7.2, we have

$$\mathbf{E}[U] \geq \epsilon n^\gamma \cdot 0.3 \cdot 0.6 = \Omega\left(\frac{n^\gamma}{C}\right),$$

which completes the proof of Theorem 8.2.9 (here, we take $\epsilon > 0$ such that $\epsilon = \Omega(1/C)$). □

8.8 Note on the Initial Round

For a $n_0 > 0$, we consider the case where n_0 vertices exist at the first round.

Theorem 8.8.1. *Let $G^{(i)} = K_{n_0+i}$, i.e., the complete graph with $n_0 + i$ vertices, and let $(P^{(i)})_{u,v} = 1/(n_0+i)$ for all $u, v \in V^{(i)}$ in $R = (\mathfrak{d}, (G^{(i)})_{i=1}^\infty, (P^{(i)})_{i=1}^\infty)$. Let N be an arbitrary positive number. If $\mathfrak{d}(i) \geq 2i/N$ for all i, then $\mathbf{E}[U(n)] \leq 2n_0 + N$.*

Proof. If $n \leq n_0$, $|V^{(n)}| = n_0 + n \leq 2n_0$ and we are done. Suppose that $n > n_0$. Then it is straightforward to see that

$$\begin{aligned} \mathbf{E}[U(n)] &= n_0 \prod_{i=1}^n \left(1 - \frac{1}{n_0+i}\right)^{\mathfrak{d}(i)} + \sum_{k=1}^n \prod_{i=k}^n \left(1 - \frac{1}{n_0+i}\right)^{\mathfrak{d}(i)} \\ &\leq n_0 + n_0 + \sum_{k=n_0+1}^n \prod_{i=k}^n \left(1 - \frac{1}{n_0+i}\right)^{\mathfrak{d}(i)} \\ &\leq 2n_0 + \sum_{k=n_0+1}^n \prod_{i=k}^n \left(1 - \frac{1}{2i}\right)^{\mathfrak{d}(i)} \\ &\leq 2n_0 + N. \end{aligned}$$

Note that we use Lemma 8.6.3 in the last inequality. □

Chapter 9

Average Distance and Diameter

9.1 Results

In this chapter, we consider the average distance and diameter of the random regular graph $G_{n,d}$. Recall that the random regular graph $G_{n,d}$ is the graph sampled according to the uniform distribution $\mathcal{G}_{n,d}$ of all n-vetex d-regular graphs. Let $\mathrm{AD}(G)$ and $\mathrm{diam}(G)$ be the average distance and diameter of a graph G, respectively (see Section 9.1.3 for the definition). We define $\mathrm{AD}(G) = \mathrm{diam}(G) = \infty$ if G is not connected. The main results of this chapter are the following.

Theorem 9.1.1. *For two constants $\alpha \in (0,1)$ and $\beta > 0$, let $d = (\beta + \mathrm{o}(1))n^{\alpha}$ be an integer. For every constant $\epsilon > 0$, it holds a.a.s. that*

$$|\mathrm{AD}(G_{n,d}) - \mu| < \epsilon,$$

where

$$\mu = \begin{cases} \alpha^{-1} + \exp(-\beta^{1/\alpha}) & \text{if } \alpha^{-1} \in \mathbb{N}, \\ \lceil \alpha^{-1} \rceil & \text{otherwise.} \end{cases} \tag{9.1}$$

Theorem 9.1.2. *For two constants $\alpha \in (0,1)$ and $\beta > 0$, let $d = (\beta + \mathrm{o}(1))n^{\alpha}$ be an integer. It holds a.a.s. that*

$$\mathrm{diam}(G_{n,d}) = \lfloor \alpha^{-1} \rfloor + 1.$$

The diameter of regular graphs has gathered special attention in graph theory [EFH80, HS60, Mv05] and has an application in designing efficient network topologies. Note that for every vertex v, there are at most $d(d-1)^k$ vertices having distance k from v. Thus, for every n-vertex d-regular graph G of diameter D with $d \geq 3$, we have

$$\begin{aligned} D &\geq \min\left\{ D \in \mathbb{N} : n \leq 1 + \sum_{i=1}^{D} d(d-1)^{i-1} \right\} \\ &= \left\lceil \log_{d-1} n + \log_{d-1}\left(1 - \frac{2}{d}\left(1 - \frac{1}{n}\right)\right) \right\rceil \\ &= \frac{\log n}{\log(d-1)} - \mathrm{O}(1). \end{aligned} \tag{9.2}$$

We denote by $D' = D'(n,d)$ the lower bound (9.2), which is known as the *Moore bound* [Mv05].

From our result Theorem 9.1.2 and the Moore bound (9.2), we have that the random d-regular graph $G_{n,d}$ of $d = (\beta + \mathrm{o}(1))n^{\alpha}$ with two arbitrary constants $\alpha \in (0,1)$ and $\beta > 0$ a.a.s. satisfies

$$\lim_{n\to\infty} \mathrm{diam}(G_{n,d}) - D'(n,d)) = \begin{cases} 0 & \text{if either } \alpha^{-1} \notin \mathbb{N} \text{ or } (\alpha^{-1} \in \mathbb{N} \text{ and } \beta < 1), \\ 1 & \text{if } \alpha^{-1} \in \mathbb{N} \text{ and } \beta > 1. \end{cases} \tag{9.3}$$

This means that, for example, $G_{n,d}$ has the minimum possible diameter among all n-vertex d-regular graphs a.a.s. if $\alpha^{-1} \notin \mathbb{N}$.

9.1.1 Background of $\mathcal{G}_{n,d}$

The study of the random regular graph $\mathcal{G}_{n,d}$ originated from the *configuration model* introduced by Bollobás [Bol80]. The configuration model is an algorithm that generates the random d-regular *multigraph* (i.e., the genareted graph may contain multiple edges or self-loops). The probability that the graph generated by the configuration model is simple (i.e., does not contain neither multiple edges nor self-loops) is $(1+\mathrm{o}(1))\exp\left(\frac{1-d^2}{4}\right)$ for fixed d. Therefore, we can study $G_{n,d}$ by considering the graph generated by the configuration model conditioned on being simple. Specifically, if $C_{n,d}$ is the graph generated by the configuration model, then we can show that, for any graph property $\mathcal{P}$,

$$\mathbf{Pr}[G_{n,d} \text{ satisfies } \mathcal{P}] = (1+\mathrm{o}(1))\exp\left(\frac{d^2-1}{4}\right)\mathbf{Pr}[C_{n,d} \text{ satisfies } \mathcal{P}]$$

This enables us to study $G_{n,d}$ for a constant d. The case of $d = d(n) \gg 1$ is much less understood, though there is a well-known successful approach called the *switching method*, introduced by McKay [McK81]. Roughly speaking, the switching method is a framework of algorithms that generates $G_{n,d}$. It starts from $C_{n,d}$ and repeat eliminating multiple edges and self-loops randomly. See [Wor99] for a detailed survey on $G_{n,d}$. However, since the switching method fails to generate $G_{n,d}$ with some probability depending on d, results based on the switching method usually require the condition that $d = \mathrm{o}(n^\gamma)$ for some reasonable constant $\gamma \le 1$. Therefore, $G_{n,d}$ of $d = (\beta+\mathrm{o}(1))n^\alpha$ with arbitrary constant $\alpha \in (0,1)$ seems to be far from these methods.

Another recent remarkable approach for the study of $G_{n,d}$ is to compare $G_{n,d}$ with the Erdős–Rényi graph $G(n,p)$ of $p = \frac{d}{n}$. Recall that the degree is concentrated on np (Example 2.5.6) if $p = \omega\left(\frac{\log n}{n}\right)$; thus we may expect that $G(n,p)$ and $G_{n,d}$ share several structural properties if $d = (1+\mathrm{o}(1))np$. For $\log n \ll d \ll n^{1/3}/(\log n)^2$, Kim and Vu [KV04] presented a coupling of $G_{n,d}$ and $G_{n,d}$ of $p = (1-\mathrm{o}(1))\frac{d}{n}$ such that $G(n,p) \subseteq G_{n,d}$ holds a.a.s. Dudek, Frieze, Ruciński, and Šileikis [DFRv17] improved this result by presenting a coupling having the same property for $\log n \ll d \ll n$. Their result is called the *embedding theorem*. The embedding theorem enables us to bound $\mathrm{diam}(G_{n,d})$ and $\mathrm{AD}(G_{n,d})$ from above by $\mathrm{diam}(G(n,p))$ and $\mathrm{AD}(G(n,p))$, respectively. Very recently, Gao, Isaev, and McKay [GIM20] proved that there is a coupling of $G(n,p)$ and $G_{n,d}$ satisfying $G(n,p) \supseteq G_{n,d}$ if $p \ge \frac{Cd\log n}{n}$ for some constant C, $d = \omega(\log n)$ and $d = \mathrm{o}(n)$. We can immediately obtain Theorem 9.1.2 by combining the coupling of [GIM20] and known results cencerning the diameter of $G(n,p)$. However, due to the $\mathrm{O}(\log n)$ factor in the condition $p \ge \frac{Cd\log n}{n}$, Theorem 9.1.1 does not follow from [GIM20] immediately.

To study $\mathrm{diam}(G_{n,d})$ and $\mathrm{AD}(G_{n,d})$, we shall look at $\mathrm{diam}(G(n,p))$ and $\mathrm{AD}(G(n,p))$ of $p = \frac{d}{n}$. It is well known that $G(n,p)$ of $p = (\beta+\mathrm{o}(1))n^{-1+\alpha}$ has diameter $\lfloor\alpha^{-1}\rfloor+1$ [Bol01, Bol81, FK16]. As for the average distance, we obtain a concentration result of $\mathrm{AD}(G(n,p))$, which is of independent interest.

Theorem 9.1.3. *For two constants $\alpha \in (0,1)$ and $\beta > 0$, let $p = \beta n^{-1+\alpha}$ and*

$$\mu := \begin{cases} \alpha^{-1} + \exp(-\beta^{1/\alpha}) & \text{if } \alpha^{-1} \in \mathbb{N}, \\ \lceil \alpha^{-1} \rceil & \text{otherwise.} \end{cases}$$

Then, there are absolute constants $C_1, C_2 > 0$ such that

$$|\mathrm{AD}(G(n,p)) - \mu| \le C_1 n^{-C_2}$$

holds a.a.s.

9.1.2 Related results and trivial bounds

Diameter of $G(n,p)$

There is a long line of works on the diameter of $G(n,p)$ [KL81, Bol81, CL01, FR07, RW10]. For dense $G(n,p)$, Bollobas [Bol81] proved the following result.

Theorem 9.1.4 (Theorem 6 of [Bol81])**.** *Fix a positive constant c. Let $D = D(n) \ge 2$ be a positive integer and $p = p(n) \in [0,1]$ be a real number satisfying*

$$p^D n^{D-1} = \log(n^2/c).$$

Suppose that $np = \omega(\log n)$. *Then,* $G(n,p)$ *satisfies*

$$\lim_{n\to\infty} \mathbf{Pr}[\mathrm{diam}(G(n,p)) = k] = \begin{cases} \exp(-c/2) & \text{if } k = D, \\ 1-\exp(-c/2) & \text{if } k = D+1, \\ 0 & \text{otherwise.} \end{cases}$$

Corollary 9.1.5. *Suppose that* $p = (\beta + \mathrm{o}(1))n^{-1+\alpha}$, *where* $\alpha \in (0,1)$ *and* $\beta > 0$ *are any constants. Then,* $\mathrm{diam}(G(n,p)) = \lfloor \alpha^{-1} \rfloor + 1$ *holds a.a.s.*

It should be noted that Corollary 9.1.5 also follows from the main result of Klee and Larman [KL81].

The diameter of $G(n,p)$ of small p has gathered special attention [Bol84, RW10, CL01]. In this line of work, there is a convention that the diameter of a disconnected graph is the maximum among all diameters of its connected components. Bollobás [Bol84] proved that $\mathrm{diam}(G(n,p)) \in A$ holds a.a.s. if $np - \log n = \omega(1)$, where $A = A(n) \subseteq \mathbb{N}$ satisfies $|A| \le 4$. Chung and Lu [CL01] studied $\mathrm{diam}(G(n,p))$ with $1 < np \le c \log n$ where c is some constant. For example, they proved that $\mathrm{diam}(G(n,p)) = (1+\mathrm{o}(1))\frac{\log n}{\log np}$ holds a.a.s. if $\omega(1) = np < \log n$. Riordan and Wormald [RW10] strengthened the results of [CL01], providing the tight estimate for $\mathrm{diam}(G(n,p))$ for $1 + \mathrm{o}(1) \le np = \mathrm{O}(1)$. For smaller p, Łuczak [Luc98] investigated $\mathrm{diam}(G(n,p))$ with $np < 1$.

Average distance of $G(n,p)$

The average distance of random graphs with a power law degree sequence has gathered a great deal of attention in network analysis [KNb+15, NKKB16, BGHJ07, vdHHM05, CL04, vdHH08]. Focusing on $G(n,p)$ with $np = \omega(\log n)$, one may observe that $\mathrm{AD}(G(n,p)) \approx \mathrm{diam}(G(n,p))$. More precisely, it is easy to see that $\mathrm{AD}(G(n,p)) \le \mathrm{diam}(G(n,p)) = (1+\mathrm{o}(1))\frac{\log n}{\log np}$ and $\mathrm{AD}(G(n,p)) \ge (1-\mathrm{o}(1))\frac{\log n}{\log np}$ hold by considering the maximum degree of $G(n,p)$).

Katzav, Nitzan, ben-Avraham, Krapisky, Künh, Ross, and Biham [KNb+15] presented analytical results on $\mathrm{AD}(G(n,p))$ for dense $G(n,p)$ that coincide with Theorem 9.1.3. However, to the best of our knowledge, there are no known results with rigorous proofs for $\mathrm{AD}(G(n,p))$ with $np = n^{\Omega(1)}$.

Diameter of $G_{n,d}$

For random regular graphs $\mathcal{G}_{n,d}$, Bollobás and de la Vega [BdlV82] proved that

$$\mathrm{diam}(G_{n,d}) = D'(n,d) \pm \mathrm{O}\left(\frac{\log\log n}{\log(d-1)}\right)$$

holds a.a.s. if the degree $d \ge 3$ is a constant. If $\log n \ll d \le n^{\mathrm{o}(1)}$, the embedding theorem of Dudek et al. [DFRv17] and the lower bound (9.2) together imply that

$$\mathrm{diam}(G_{n,d}) = (1+\mathrm{o}(1))\frac{\log n}{\log d} = (1+\mathrm{o}(1))D'(n,d)$$

holds a.a.s.

Suppose that $d = (\beta + \mathrm{o}(1))n^{\alpha}$, where $\alpha \in (0,1)$ and $\beta > 0$ are constants. From the embedding theorem, we have $\mathrm{diam}(G_{n,d}) \le \lfloor \alpha^{-1} \rfloor + 1$ holds a.a.s., as we will confirm in Section 9.2. On the other hand, by substituting $d = (\beta + \mathrm{o}(1))n^{\alpha}$ to (9.2), we obtain

$$\lim_{n\to\infty} D' = \begin{cases} \lfloor \alpha^{-1} \rfloor + 1 & \text{if } \alpha^{-1} \notin \mathbb{N} \text{ or } (\alpha^{-1} \in \mathbb{N} \wedge \beta < 1), \\ \alpha^{-1} & \text{if } \alpha^{-1} \in \mathbb{N} \wedge \beta > 1, \\ \text{depends on the term } \mathrm{o}(1) & \text{if } \alpha^{-1} \in \mathbb{N} \wedge \beta = 1. \end{cases} \tag{9.4}$$

By combining Theorem 9.1.2 and (9.4), we obtain (9.3). As mentioned earlier, Theorem 9.1.1 immediately follows from the result of Gao, Isaev, and McKay [GIM20]. In this chapter, we prove Theorem 9.1.2 by combining the upper bound from the embedding theorem [DFRv17] and Theorem 9.1.1 (note that $\mathrm{diam}(G) \ge \lceil \mathrm{AD}(G) \rceil$).

Average distance of $G_{n,d}$

Let N_k be the number of vertex pairs of distance k. We use the same argument as for (9.2) to obtain a lower bound of $\mathrm{AD}(G)$ for any d-regular graph with $d \geq 3$. Suppose $\mathrm{diam}(G) = D'$ and thus $N_1 + \cdots + N_{D'} = \binom{n}{2}$. Moreover, for every $k = 1, \ldots, D' - 1$, we have $N_k \leq d(d-1)^{k-1}$. Therefore, we obtain

$$\begin{aligned}
\mathrm{AD}(G) &= \binom{n}{2}^{-1} (N_1 + 2N_2 + \cdots + D'N_{D'}) \\
&= D' - \binom{n}{2}^{-1} ((D'-1)N_1 + (D'-2)N_2 + \cdots + N_{D'-1}) \\
&\geq D' - \binom{n}{2}^{-1} \sum_{k=1}^{D'-1} (D'-k) d(d-1)^{k-1} \\
&= D' - \frac{d(d-1)^{D'}}{(n-1)(d-2)^2} + \frac{dD'}{(n-1)(d-2)} + \frac{d}{(n-1)(d-2)^2} \qquad (9.5) \\
&= \log_{d-1} n - \mathrm{O}(1).
\end{aligned}$$

Let $\mathrm{AD}' = \mathrm{AD}(n,d)$ denote the lower bound (9.5). Then, we have

$$\frac{\log n}{\log(d-1)} - \mathrm{O}(1) \leq \mathrm{AD}(G_{n,d}) \leq \mathrm{diam}(G_{n,d}).$$

This implies that

$$\mathrm{AD}(G_{n,d}) = (1 + \mathrm{o}(1)) \frac{\log n}{\log(d-1)}$$

holds a.a.s. if $d \geq 3$ is constant or $\log n \ll d \leq n^{\mathrm{o}(1)}$.

Suppose that $d = (\beta + \mathrm{o}(1)) n^\alpha$, where $\alpha \in (0,1)$ and $\beta > 0$ are constants. From the lower bound (9.5), we have

$$\lim_{n\to\infty} \mathrm{AD}' = \begin{cases} \lfloor \alpha^{-1} \rfloor + 1 & \text{if } \alpha^{-1} \notin \mathbb{N}, \\ \alpha^{-1} & \text{if } \alpha^{-1} \in \mathbb{N} \text{ and } \beta > 1, \\ \alpha^{-1} - \beta^{1/\alpha} + 1 & \text{if } \alpha^{-1} \in \mathbb{N} \text{ and } \beta < 1, \\ \text{depends on the term } \mathrm{o}(1) & \text{otherwise.} \end{cases} \qquad (9.6)$$

9.1.3 Preliminaries

Throughout this chapter, the number of vertices of a graph is denoted by n, and the vertex set is identified with $[n]$.

For two graphs G and H, let

$$\begin{aligned}
G \cup H &= (V(G) \cup V(H), E(G) \cup E(H)), \\
G \cap H &= (V(G) \cap V(H), E(G) \cap E(H)).
\end{aligned}$$

Note that both G and H are labelled graph.

A *path* is a graph $P = (\{v_0, \ldots, v_\ell\}, \{\{v_0, v_1\}, \ldots, \{v_{\ell-1}, v_\ell\}\})$ for distinct vertices $v_0, \ldots, v_\ell$. The vertices of degree one in a path are called *endpoints*. We call a path of endpoints s and t an *st-path*. The *length* of a path is the number of edges. For a graph G and its two distinct vertices s and t, the *distance* $\mathrm{dist}_G(s,t)$ is the minimum length among all st-paths contained in G. We define $\mathrm{dist}_G(s,t) = \infty$ if G does not contain any st-paths. If the graph G is clear from the context, we use $\mathrm{dist}(s,t)$ for $\mathrm{dist}_G(s,t)$.

For a graph $G = (V, E)$ of n vertices, the *average distance* $\mathrm{AD}(G)$ of G is

$$\mathrm{AD}(G) = \binom{n}{2}^{-1} \sum_{\{s,t\} \in \binom{V}{2}} \mathrm{dist}_G(s,t).$$

The *diameter* $\mathrm{diam}(G)$ of G is

$$\mathrm{diam}(G) = \max_{s \neq t} \mathrm{dist}_G(s,t).$$

Note that $\mathrm{diam}(G) = \mathrm{AD}(G) = \infty$ if G is not connected.

9.1.4 Tools

Lemma 9.1.6 (Multivariate version of Brun's sieve; Lemma 2.8 of [Wor99]). *Let $S_n^{(1)}, \dots, S_n^{(k)}$ be random variables defined on the same space Ω_n such that each $S_n^{(i)}$ can be written as the sum of binary random variables. Suppose that there exist positive constants $\lambda_1, \dots, \lambda_k$ satisfying*

$$\lim_{n\to\infty} \mathbf{E}\left[\prod_{i=1}^{k} (S_n^{(i)})_{r_i}\right] = \prod_{i=1}^{k} \lambda_i^{r_i}$$

for every fixed integers $r_1, \dots, r_k \geq 0$.

Then, for any constants $j_1, \dots, j_k \geq 0$, it holds that

$$\lim_{n\to\infty} \mathbf{Pr}\left[\bigwedge_{i=1}^{k} [S_n^{(i)} = j_i]\right] = \prod_{i=1}^{k} \exp(-\lambda_i)\frac{\lambda^{j_i}}{j_i!}.$$

Lemma 9.1.7 (Lemma 2.1 of [KSV07]). *Suppose that $1 \ll d \ll n$. For any fixed graph H, it holds that*

$$\mathbf{Pr}[H \subseteq G_{n,d}] = (1 + \mathrm{o}(1)) \left(\frac{d}{n}\right)^{|E(H)|}.$$

Let $G[n, m]$ be a graph selected uniformly at random from the set of all graphs of n vertices with exactly m edges.

Lemma 9.1.8 (The embedding theorem; Theorem 10.10 of [FK16]). *There is a constant $C > 0$ that satisfies the following. For any real $\gamma = \gamma(n)$, integer $d = d(n)$ satisfying*

$$C\left(\left(\frac{d}{n} + \frac{\log n}{d}\right)^{1/3}\right) \leq \gamma < 1, \tag{9.7}$$

and $m = \lfloor (1-\gamma)nd/2 \rfloor$, there exists a joint distribution π of $G[n, m]$ and $G_{n,d}$ such that

$$\lim_{n\to\infty} \mathop{\mathbf{Pr}}_{\pi}[G[n, m] \subseteq G_{n,d}] = 1$$

holds.

In other words, for $\log n \ll d \ll n$, we can choose $m = (1 - \mathrm{o}(1))nd/2$ and couple $G[n, m]$ and $G_{n,d}$ such that $G[n, m] \subseteq G_{n,d}$ holds a.a.s.

9.2 Upper bounds of $\mathrm{AD}(G_{n,d})$ and $\mathrm{diam}(G_{n,d})$

In this section we obtain upper bounds of $\mathrm{AD}(G_{n,d})$ and $\mathrm{diam}(G_{n,d})$ using Lemma 9.1.8. As noted in [DFRv17], in Lemma 9.1.8, one can replace $G[n, m]$ by $G(n, p)$ of $p = (1 - 2\gamma)d/(n - 1)$. This yields the following result.

Corollary 9.2.1. *For $d = d(n)$ satisfying $\log n \ll d \ll n$, there exists $p = (1 - \mathrm{o}(1))\frac{d}{n}$ such that $\mathrm{AD}(G_{n,d}) \leq \mathrm{AD}(G(n, p))$ and $\mathrm{diam}(G_{n,d}) \leq \mathrm{diam}(G(n, p))$ hold a.a.s.*

For $d = (\beta + \mathrm{o}(1))n^{\alpha}$, take γ of Lemma 9.1.8 satisfying $\gamma = \mathrm{o}(1)$, and let $p = (1 - 2\gamma)\frac{d}{n-1} = (\beta + \mathrm{o}(1))n^{-1+\alpha}$. Then, from Theorem 9.1.3 and Corollary 9.2.1, it holds a.a.s. that

$$\mathrm{AD}(G_{n,d}) \leq \mathrm{AD}(G(n, p)) \leq \mu + \mathrm{o}(1). \tag{9.8}$$

Similarly, from Corollaries 9.1.5 and 9.2.1, the random regular graph $G_{n,d}$ a.a.s. satisfies

$$\mathrm{diam}(G_{n,d}) \leq \mathrm{diam}(G(n, p)) \leq \lfloor \alpha^{-1} \rfloor + 1. \tag{9.9}$$

9.3 Lower bounds of $\mathrm{AD}(G_{n,d})$ and $\mathrm{diam}(G_{n,d})$

If $\alpha^{-1} \notin \mathbb{N}$, the lower bound (9.6) and the upper bound (9.8) yield that

$$\mathrm{AD}(G_{n,d}) = \lfloor \alpha^{-1} \rfloor + 1 - \mathrm{o}(1)$$

holds a.a.s. Now we focus on the case where $\alpha^{-1} \in \mathbb{N}$. This section is devoted to proving the following.

Lemma 9.3.1. *Let $d = (\beta + \mathrm{o}(1))n^{\alpha}$, where $\alpha \in (0,1)$ and $\beta > 0$ are any constants satisfying $\alpha^{-1} \in \mathbb{N}$. For any constant $\epsilon > 0$,*

$$\lim_{n\to\infty} \mathbf{Pr}[\mathrm{AD}(G_{n,d}) \le \mu - \epsilon] = 0,$$

where $\mu = \alpha^{-1} + \exp(-\beta^{1/\alpha})$.

Remark 9.3.2. By combining (9.8) and Lemma 9.3.1, we complete the proof of Theorem 9.1.1. Moreover, Lemma 9.3.1 implies

$$\mathrm{diam}(G_{n,d}) \ge \lceil \mathrm{AD}(G_{n,d}) \rceil = \lfloor \alpha^{-1} \rfloor + 1$$

holds a.a.s., which completes the proof of Theorem 9.1.2.

Proof of Lemma 9.3.1. Note that

$$\begin{aligned}
\mathrm{AD}(G_{n,d}) &= \binom{n}{2}^{-1} \sum_{\{s,t\} \in \binom{V}{2}} \mathrm{dist}(s,t) \\
&= \sum_{\ell=1}^{\infty} \binom{n}{2}^{-1} \sum_{\{s,t\} \in \binom{V}{2}} \mathbb{1}_{[\mathrm{dist}(s,t) \ge \ell]} \\
&\ge \sum_{\ell=1}^{\alpha^{-1}+1} \binom{n}{2}^{-1} \sum_{\{s,t\} \in \binom{V}{2}} \mathbb{1}_{[\mathrm{dist}(s,t) \ge \ell]}.
\end{aligned}$$

For $\ell \in \{1, \dots, \alpha^{-1}+1\}$, let $p_\ell = p_\ell(G_{n,d}) = \binom{n}{2}^{-1} \sum_{\{s,t\} \in \binom{V}{2}} \mathbb{1}_{[\mathrm{dist}(s,t) \ge \ell]}$. We evaluate p_ℓ using the following result.

Lemma 9.3.3. *Consider $G_{n,d}$ of $d = (\beta + \mathrm{o}(1))n^{\alpha}$. Fix two constants $\alpha \in (0,1)$ and $\beta > 0$ satisfying $\alpha^{-1} \in \mathbb{N}$. For any constant $k \in \mathbb{N}$, fix $2k$ distinct vertices $s_1, \dots, s_k, t_1, \dots, t_k$. For any fixed $\ell_1, \dots, \ell_k \in \{1, \dots, \alpha^{-1}+1\}$, it holds that*

$$\lim_{n\to\infty} \mathbf{Pr}\left[\bigwedge_{i=1}^{k} [\mathrm{dist}(s_i, t_i) \ge \ell_i]\right] = \exp(-M\beta^{1/\alpha})$$

where $M = |\{i \in \{1, \dots, k\} : \ell_i = \alpha^{-1}+1\}|$.

We will prove Lemma 9.3.3 in Section 9.3.1. For $\ell \in \{1, \dots, \alpha^{-1}+1\}$, let

$$\mu_\ell = \begin{cases} 1 & \text{if } 1 \le \ell \le \alpha^{-1}, \\ \exp(-\beta^{1/\alpha}) & \text{if } \ell = \alpha^{-1}+1. \end{cases}$$

From Lemma 9.3.3, we have

$$\begin{aligned}
\mathbf{E}[p_\ell] &= \binom{n}{2}^{-1} \sum_{\{s,t\} \in \binom{V}{2}} \mathbf{Pr}[\mathrm{dist}(s,t) \ge \ell] \\
&= \mathbf{Pr}[\mathrm{dist}(1,2) \ge \ell] = \mu + \mathrm{o}(1)
\end{aligned}$$

and

$$\begin{aligned}\mathbf{E}[p_\ell^2] &= \binom{n}{2}^{-2} \sum_{\{s,t\},\{s',t'\}\in\binom{V}{2}} \mathbf{Pr}[\mathrm{dist}(s,t)\geq\ell \wedge \mathrm{dist}(s',t')\geq\ell] \\ &= \binom{n}{2}^{-2}\left(\mathrm{O}(n^3) + \sum_{\substack{\{s,t\},\{s',t'\}\in\binom{V}{2}:\\ \{s,t\}\cap\{s',t'\}=\emptyset}} \mathbf{Pr}[\mathrm{dist}(s,t)\geq\ell \wedge \mathrm{dist}(s',t')\geq\ell]\right) \\ &= \mathbf{Pr}[\mathrm{dist}(1,2)\geq\ell \wedge \mathrm{dist}(3,4)\geq\ell] + \mathrm{o}(1) = \mu^2 + \mathrm{o}(1).\end{aligned}$$

From the Chebyshev inequality, for every constant $\epsilon > 0$, we have

$$\mathbf{Pr}[|p_\ell - \mathbf{E}[p_\ell]| \geq \epsilon] \leq \frac{\mathbf{Var}[p_\ell]}{\epsilon^2} = \mathrm{o}(1).$$

Thus we obtain

$$\mathbf{Pr}\left[\left|\left(\sum_{\ell=1}^{\alpha^{-1}+1} p_\ell\right) - \mu\right| > \epsilon\right] \leq \sum_{\ell=1}^{\alpha^{-1}+1} \mathbf{Pr}\left[|p_\ell - \mu_\ell| > \epsilon/(\alpha^{-1}+1)\right] = \mathrm{o}(1).$$

Therefore, it holds a.a.s. that

$$\mathrm{AD}(G_{n,d}) \geq \sum_{\ell=1}^{\alpha^{-1}+1} p_\ell \geq \mu - \mathrm{o}(1),$$

which completes the proof of Lemma 9.3.1. □

9.3.1 Distances of fixed vertex pairs of $G_{n,d}$

This part is devoted to proving Lemma 9.3.3. We start with establishing the following result.

Lemma 9.3.4. *Consider $G_{n,d}$ of $d = (\beta + \mathrm{o}(1))n^\alpha$ for constants $\alpha \in (0,1)$ and $\beta > 0$. For two fixed distinct vertices s and t, it holds a.a.s. that* $\mathrm{dist}(s,t) \in \{\lceil\alpha^{-1}\rceil, \lfloor\alpha^{-1}\rfloor + 1\}$.

Proof. For two fixed vertices s, t of $G_{n,d}$ and an integer ℓ, we denote by $\mathcal{P}$ the set of paths of length ℓ connecting s and t in a complete graph. Let $X_\ell = X_\ell(G_{n,d})$ be the number of paths $P \in \mathcal{P}$ contained in $G_{n,d}$, that is,

$$X_\ell = |\{P \in \mathcal{P} : P \subseteq G_{n,d}\}|. \tag{9.10}$$

Fix an integer ℓ satisfying $\ell\alpha < 1$ (or equivalently, $\ell \leq \lceil\alpha^{-1}\rceil - 1$). Then, from Lemma 9.1.7, we have

$$\begin{aligned}\mathrm{E}(X_\ell) &= \sum_{P\in\mathcal{P}} \mathbf{Pr}[P \subseteq G_{n,d}] \\ &= (1+\mathrm{o}(1))n^{\ell-1}\left(\frac{d}{n}\right)^\ell \\ &= \mathrm{o}(1).\end{aligned}$$

From the Markov's inequality, we obtain

$$\begin{aligned}\mathbf{Pr}[\mathrm{dist}(s,t)\leq\ell] &\leq \mathbf{Pr}[X_1 + \cdots + X_\ell > 0] \\ &\leq \sum_{i=1}^{\ell} \mathrm{E}(X_i) \\ &= \mathrm{o}(1).\end{aligned}$$

In other words, $\mathrm{dist}(s,t) \geq \ell + 1 \geq \lceil\alpha^{-1}\rceil$ holds a.a.s.

On the other hand, from (9.9), we have $\mathrm{dist}(s,t) \leq \mathrm{diam}(G_{n,d}) \leq \lfloor\alpha^{-1}\rfloor + 1$. This completes the proof of Lemma 9.3.4. □

Proof of Lemma 9.3.3. Fix an integer $k > 0$ and $2k$ distinct vertices $s_1, \dots, s_k, t_1, \dots, t_k$ of $G_{n,d}$, where $d = (\beta + \mathrm{o}(1))n^\alpha$. From Lemma 9.3.4, it holds a.a.s. that $\mathrm{dist}(s,t) \in \{\alpha^{-1}, \alpha^{-1}+1\}$.

Suppose that $\ell_1 \le \alpha^{-1}$ and thus $\mathrm{dist}(s_1,t_1) \ge \ell_1$ holds a.a.s. Then we have

$$\mathbf{Pr}\left[\bigwedge_{i=2}^{k}[\mathrm{dist}(s_i,t_i) \ge \ell_i]\right] - \mathbf{Pr}[\mathrm{dist}(s_1,t_1) < \ell_1] \le \mathbf{Pr}\left[\bigwedge_{i=1}^{k}[\mathrm{dist}(s_i,t_i) \ge \ell_i]\right] \le \mathbf{Pr}\left[\bigwedge_{i=2}^{k}[\mathrm{dist}(s_i,t_i) \ge \ell_i]\right]$$

and thus

$$\mathbf{Pr}\left[\bigwedge_{i=1}^{k}[\mathrm{dist}(s_i,t_i) \ge \ell_i]\right] = \mathbf{Pr}\left[\bigwedge_{i=2}^{k}[\mathrm{dist}(s_i,t_i) \ge \ell_i]\right] - \mathrm{o}(1).$$

Hence, we may assume that $\ell_i = \alpha^{-1}+1$ for all $i = 1, \dots, k$ (i.e., $M = k$ in Lemma 9.3.3).

Let $\mathcal{P}^{(i)}$ denote the set of s_it_i-paths of length α^{-1} contained in the complete graph K_n. Define $X^{(i)}$ as the number of paths of $\mathcal{P}^{(i)}$ contained in $G_{n,d}$, that is,

$$X^{(i)} = |\{P \in \mathcal{P}^{(i)} : P \subseteq G(n,p)|.$$

Then, we have

$$\begin{aligned}\mathbf{Pr}\left[\bigwedge_{i=1}^{k}[\mathrm{dist}(s_i,t_i) \ge \alpha^{-1}+1]\right] &= \mathbf{Pr}\left[\bigwedge_{i=1}^{k}[\mathrm{dist}(s_i,t_i) \ge \alpha^{-1}] \wedge \bigwedge_{i=1}^{k}[X^{(i)} = 0]\right] \\ &= \mathbf{Pr}\left[\bigwedge_{i=1}^{k}[X^{(i)} = 0]\right] - \mathrm{o}(1). \end{aligned} \tag{9.11}$$

We evaluate (9.11) using the following result, which will be shown in Section 9.3.2.

Lemma 9.3.5. *Consider $G_{n,d}$ of $d = (\beta + \mathrm{o}(1))n^\alpha$, where $\alpha \in (0,1)$ and $\beta > 0$ are any constants satisfying $\alpha^{-1} \in \mathbb{N}$. Fix $2k$ distinct vertices $s_1, \dots, s_k, t_1, \dots, t_k$, where k is any constant. For $i = 1, \dots, k$, let $X^{(i)}$ denote the number of s_it_i-paths of length $\alpha^{-1} \in \mathbb{N}$ contained in $G(n,p)$. Fix arbitrary nonnegative integers $r_1, \dots, r_k$. Then, it holds that*

$$\mathbf{E}\left[\prod_{i=1}^{k}(X^{(i)})_{r_i}\right] = (\beta^{1/\alpha})^R + \mathrm{o}(1),$$

where $R = r_1 + \cdots + r_k$.

From Lemma 9.3.5 and the Poisson approximation theorem (Lemma 9.1.6), we have

$$\mathbf{Pr}\left[\bigwedge_{i=1}^{k}[X^{(i)} = 0]\right] = \exp(-k\beta^{1/\alpha}) + \mathrm{o}(1). \tag{9.12}$$

By combining (9.11) and (9.12), we have

$$\mathbf{Pr}\left[\bigwedge_{i=1}^{k}[\mathrm{dist}(s_i,t_i) \ge \alpha^{-1}+1]\right] = \exp(-k\beta^{1/\alpha}) - \mathrm{o}(1).$$

This completes the proof of Lemma 9.3.3 and thus Lemma 9.3.1.

9.3.2 Proof of Lemma 9.3.5

We first prove the following result and then show Lemma 9.3.5.

Lemma 9.3.6. *Fix an integer $\ell \geq 1$ and consider $G(n,p)$ satisfying $(np)^\ell = \Omega(n)$. Fix $2k$ distinct vertices $s_1, \dots, s_k, t_1, \dots, t_k$, where k is arbitrary constant. For $i = 1, \dots, k$, let $X^{(i)}$ denote the number of $s_i t_i$-paths of length $\ell \in \mathbb{N}$ contained in $G(n,p)$.*

Then, for any fixed nonnegative integers $r_1, \dots, r_k$,

$$\mathbf{E}\left[\prod_{i=1}^{k}\left(X^{(i)}\right)_{r_i}\right] = n^{R(\ell-1)}p^{R\ell}\left(1 \pm \mathrm{O}\left(\frac{1}{np}\right)\right),$$

where $R = r_1 + \cdots + r_k$.

Corollary 9.3.7. *Consider $G(n,p)$ of $p = (\beta + \mathrm{o}(1))n^{-1+\alpha}$, where $\alpha \in (0,1)$ and $\beta > 0$ are any constants satisfying $\alpha^{-1} \in \mathbb{N}$. Fix arbitrary nonnegative integers $r_1, \dots, r_k$. Then, it holds that*

$$\mathbf{E}\left[\prod_{i=1}^{k}(X^{(i)})_{r_i}\right] = (\beta^{1/\alpha})^R + \mathrm{o}(1),$$

where $R = r_1 + \cdots + r_k$.

Proof of Lemma 9.3.6. For a positive constant k, fix $2k$ distinct vertices $s_1, \dots, s_k, t_1, \dots, t_k$. For every $i \in \{1, \dots, k\}$, let $\mathcal{P}^{(i)}$ denote the set of all $s_i t_i$-paths of length ℓ contained in the complete graph. We denote by $X^{(i)}$ the number of paths of $\mathcal{P}^{(i)}$ contained in $G(n,p)$.

Fix nonnegative integers $k, r_1, \dots, r_k$. We may assume that $r_i > 0$ for every $i = 1, \dots, k$. Let $\mathcal{A} = (\mathcal{P}^{(1)})_{r_1} \times \cdots \times (\mathcal{P}^{(k)})_{r_k}$. Each element $A \in \mathcal{A}$ is a tuple

$$A = ((P_1^{(1)}, \dots, P_{r_1}^{(1)}), \dots, (P_1^{(k)}, \dots, P_{r_k}^{(k)})),$$

where each $P_j^{(i)} \in \mathcal{P}_i$ is an $s_i t_i$-path of length ℓ and $P_j^{(i)} \neq P_{j'}^{(i)}$ holds for every i and $j \neq j'$. For notational convention, we write $A = (P_1, \dots, P_R) \in \mathcal{A}$. Since $r_k > 0$, it holds that $P_R \in \mathcal{P}^{(k)}$.

For a tuple $A = (P_1, \dots, P_t)$ of t paths, let $E(A) = \bigcup_{i=1}^{t} E(P_i)$ and $V(A) = \bigcup_{i=1}^{t} V(P_i)$ (we will use induction on R and hence we assume $t \leq R$ here). For $\mathcal{S} \subseteq \mathcal{A}$, we consider

$$\Gamma_{\mathcal{S}} = \sum_{A \in \mathcal{S}} p^{|E(A)|}.$$

Note that $\mathbf{E}\left[\prod_{i=1}^{k}\left(X^{(i)}\right)_{r_i}\right] = \sum_{A \in \mathcal{A}} \mathbf{Pr}[E(A) \subseteq E(G(n,p))] = \Gamma_{\mathcal{A}}$. We claim

$$n^{R(\ell-1)}p^{R\ell}\left(1 - \mathrm{O}\left(\frac{1}{n}\right)\right) \leq \Gamma_{\mathcal{A}} \leq n^{R(\ell-1)}p^{R\ell}\left(1 + \mathrm{O}\left(\frac{1}{np}\right)\right), \tag{9.13}$$

which completes the proof of Lemma 9.3.6.

For any $A \in \mathcal{A}$, it holds that $|E(A)| \leq R\ell$ and the equality holds if and only if any two distinct paths P_i, P_j of A shares no edges (see Figure 9.1). Let

$$\begin{aligned}\mathcal{F} &= \{A \in \mathcal{A} : |E(A)| < R\ell\} \\ &= \{(P_1, \dots, P_R) \in \mathcal{A} : \exists i \neq j,\ E(P_i) \cap E(P_j) \neq \emptyset\}.\end{aligned} \tag{9.14}$$

Figure 9.2 illustrates an example. Then, $\Gamma_{\mathcal{A}}$ can be decomposed into

$$\Gamma_{\mathcal{A}} = \Gamma_{\mathcal{F}} + \Gamma_{\mathcal{A}\setminus\mathcal{F}}. \tag{9.15}$$

The second term $\Gamma_{\mathcal{A}\setminus\mathcal{F}}$ satisfies

$$\begin{aligned}\Gamma_{\mathcal{A}\setminus\mathcal{F}} &= p^{R\ell}\,|\{A \in \mathcal{A} : |E(A)| = R\ell\}| \\ &\geq p^{R\ell}\,|\{A \in \mathcal{A} : |E(A)| = R\ell \text{ and } |V(A)| = R(\ell-1) + 2k\}| \\ &= (n-2k)_{R(\ell-1)}p^{R\ell} \\ &\geq n^{R(\ell-1)}p^{R\ell}\left(1 - \mathrm{O}\left(\frac{1}{n}\right)\right).\end{aligned}$$

This implies the lower bound $\Gamma_{\mathcal{A}} \geq \Gamma_{\mathcal{A}\setminus\mathcal{F}} \geq n^{R(\ell-1)}p^{R\ell}\left(1 - \mathrm{O}\left(\frac{1}{n}\right)\right)$.

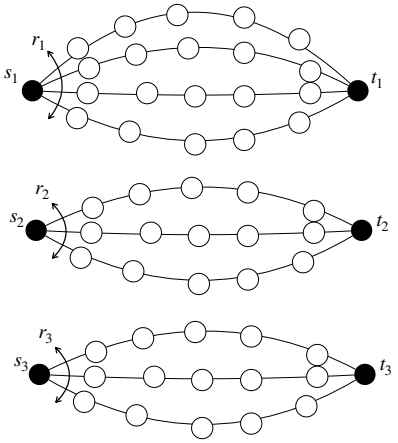


Figure 9.1: A tuple $A \in \mathcal{A} \setminus \mathcal{F}$.

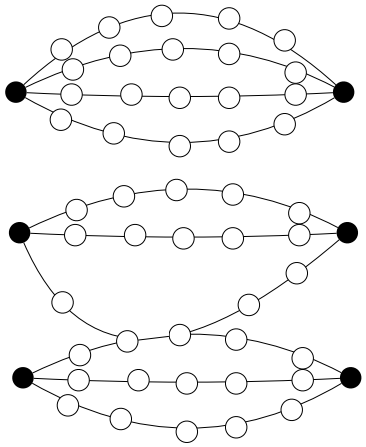


Figure 9.2: A tuple $A \in \mathcal{F}$.

Now it suffices to bound $\Gamma_{\mathcal{A}}$ from above. Observe that $\Gamma_{\mathcal{A}\setminus\mathcal{F}}$ satisfies

$$\Gamma_{\mathcal{A}\setminus\mathcal{F}} = p^{R\ell} \left|\{A \in \mathcal{A} : |E(A)| = R\ell\}\right| \leq n^{R(\ell-1)} p^{R\ell}. \tag{9.16}$$

We show that this term is dominating in $\Gamma_{\mathcal{A}}$. Lemma 9.3.6 immediately follows from (9.15) and (9.16) and the following result:

Lemma 9.3.8. *Suppose that* $(np)^\ell = \Omega(n)$. *Define* $\mathcal{F}$ *as* (9.14). *It holds that*

$$\Gamma_{\mathcal{F}} = \mathrm{O}\left(\frac{n^{R(\ell-1)} p^{R\ell}}{np}\right).$$

Proof. We use induction on R. For the base case of $R = 1$, we have $\mathcal{F} = \emptyset$ and thus

$$\Gamma_{\mathcal{A}} \leq n^{\ell-1} p^\ell,$$
$$\Gamma_{\mathcal{F}} = 0.$$

Suppose that $R \geq 2$ and that Lemma 9.3.8 holds for $R-1$. Note that Lemma 9.3.6 also holds for $R-1$ since Lemma 9.3.8 implies Lemma 9.3.6. Let

$$\mathcal{A}' = (\mathcal{P}^{(1)})_{r_1} \times \cdots \times (\mathcal{P}^{(k-1)})_{r_k - 1}.$$

Then, each element $A = (P_1, \ldots, P_R) \in \mathcal{A}$ can be decomposed into $A' = (P_1, \ldots, P_{R-1}) \in \mathcal{A}'$ and $P_R \in \mathcal{P}^{(k)}$. Note that the edge set $E(A')$ for $A' \in \mathcal{A}'$ are defined in the same way as $E(A)$ and it holds that $|E(A')| \leq (R-1)\ell$. Let

$$\mathcal{F}' = \{A' \in \mathcal{A}' : |E(A')| < (R-1)\ell\}.$$

By the induction assumption on $\mathcal{F}'$ and $\mathcal{A}'$, we have

$$\Gamma_{\mathcal{A}'} \leq n^{(R-1)(\ell-1)} p^{(R-1)\ell} \left(1 + \frac{C_1}{np}\right),$$
$$\Gamma_{\mathcal{F}'} \leq C_2 \left(\frac{n^{(R-1)(\ell-1)} p^{(R-1)\ell}}{np}\right)$$

for some constants $C_1, C_2 > 0$. For $A = (P_1, \ldots, P_R) \in \mathcal{F}$, let $A' = (P_1, \ldots, P_{R-1}) \in \mathcal{A}'$. Since $A \in \mathcal{F}$, either

(i) $E(P_R) \cap E(P_i) \neq \emptyset$ for some $1 \leq i < R$, or

(ii) $E(P_R) \cap E(A') = \emptyset$ and $E(P_i) \cap E(P_j) \neq \emptyset$ for some $1 \leq i < j < R$ (thus $A' \in \mathcal{F}'$)

holds. Therefore, we have

$$\begin{aligned}\Gamma_{\mathcal{F}} &= \sum_{A\in\mathcal{F}} p^{|E(A)|} \\ &\le \sum_{A'\in\mathcal{A}'} \sum_{\substack{P_R\in\mathcal{P}^{(k)}:\\ E(A)\cap E(P_R)\neq\emptyset}} p^{|E(A')\cup E(P_R)|} + \sum_{A'\in\mathcal{F}'} \sum_{\substack{P_R\in\mathcal{P}^{(k)}:\\ E(P_R)\cap E(A')=\emptyset}} p^{|E(A')\cup E(P_R)|}. \end{aligned} \tag{9.17}$$

From the induction assumption, the second term satisfies

$$\begin{aligned}\sum_{A'\in\mathcal{F}'} \sum_{\substack{P_R\in\mathcal{P}^{(k)}:\\ E(P_R)\cap E(A')=\emptyset}} p^{|E(A')\cup E(P_R)|} &= \sum_{A'\in\mathcal{F}'} p^{|E(A')|} \sum_{\substack{P_R\in\mathcal{P}^{(k)}:\\ E(P_R)\cap E(A')=\emptyset}} p^{|E(P_R)|} \\ &\le \Gamma_{\mathcal{F}'}\cdot n^{\ell-1}p^{\ell}. \end{aligned} \tag{9.18}$$

The first term can be rewritten as

$$\sum_{A'\in\mathcal{A}'} \sum_{\substack{P_R\in\mathcal{P}^{(k)}:\\ E(A')\cap E(P_R)\neq\emptyset}} p^{|E(A')\cup E(P_R)|} = \sum_{A'\in\mathcal{A}'} p^{|E(A')|} \sum_{\substack{P_R\in\mathcal{P}^{(k)}:\\ E(A)\cap E(P_R)\neq\emptyset}} p^{|E(P_R)\setminus E(A')|}.$$

Fix $A' = (P_1,\dots,P_{R-1}) \in \mathcal{A}'$. Let $S = \{s_1,\dots,s_k,t_1,\dots,t_k\}$ be the endpoints of the paths and let $V_1 = S\cup V(P_1)\cup\cdots\cup V(P_{R-1})$. To bound the number of P_R satisfying the condition (ii), we consider two cases: $E(P_R)\not\subseteq E(A')$ and $E(P_R)\subseteq E(A')$.

Case I. $E(P_R)\not\subseteq E(A')$**.** The edge set $E(P_R)\cap E(A')$ forms a forest. Since $E(P_R)\not\subseteq E(A')$, this forest is not connected and thus we have $|V(P_R)\cap V_1| - |E(P_R)\cap E(A')| \ge 2$. This yields

$$\begin{aligned}|V(P_R)\setminus V_1| &= |V(P_R)| - |V(P_R)\cap V_1| \\ &\le \ell - |E(P_R)\cap E(A')| - 1.\end{aligned}$$

Let $|E(P_R)\cap E(A')| = t < \ell$. Then, P_R consists of two type of vertices: at most $\ell - t - 1$ from $V\setminus V_1$ and the others from V_1. Therefore, there are at most $n^{\ell-t-1}|V_1|^t \le C^t n^{\ell-t-1}$ candidates for the path P_R satisfying $|E(P_R)\cap E(A')| = t < \ell$, where $C = (R-1)(\ell+1)$ (recall that two endpoints of P_R are fixed and thus they are not taken into account).

Case II. $E(P_R)\subseteq E(A')$**.** We claim $A'\in\mathcal{F}'$. If not, it holds that $E(P_i)\cap E(P_j) = \emptyset$ for any $i<j<R$. Hence, $E(P_R)\subseteq E(A')$ implies $P_R = P_i$ for some $i < R$. This contradicts to the definition of $\mathcal{A}$ ($P_i \neq P_j$ for any $i<j\le R$). Moreover, the number of $P_R\in\mathcal{P}^{(k)}$ satisfying $E(P_R)\subseteq E(A')$ is at most $|V_1|^{\ell-1} \le C^{R(\ell-1)}$. Therefore, we have

$$\begin{aligned}&\sum_{A'\in\mathcal{A}'} \sum_{\substack{P_R\in\mathcal{P}^{(k)}:\\ E(A')\cap E(P_R)\neq\emptyset}} p^{|E(A')\cup E(P_R)|} \\ &\le \sum_{A'\in\mathcal{A}'} p^{|E(A')|} \left(\sum_{t=1}^{\ell-1} \sum_{\substack{P_R\in\mathcal{P}^{(k)}:\\ |E(A)\cap E(P_R)|=t}} p^{|E(P_R)\setminus E(A')|} \right) + \sum_{A'\in\mathcal{F}'} p^{|E(A')|} C^{R(\ell-1)} \\ &\le \sum_{A'\in\mathcal{A}'} p^{|E(A')|}\cdot \sum_{t=1}^{\ell-1} C^t n^{\ell-t-1} p^{\ell-t} + C^{R(\ell-1)}\Gamma_{\mathcal{F}'} \\ &\le \Gamma_{\mathcal{A}'}\cdot \frac{Cn^{\ell-1}p^{\ell}}{np}\left(1+\frac{1.01C}{np}\right) + C^{R(\ell-1)}\Gamma_{\mathcal{F}'}. \end{aligned} \tag{9.19}$$

From (9.17) to (9.19) and the induction assumption, we have

$$\begin{aligned}\Gamma_{\mathcal{F}} &\le \Gamma_{\mathcal{F}'}\cdot n^{\ell-1}p^{\ell} + \Gamma_{\mathcal{A}'}\cdot \frac{Cn^{\ell-1}p^{\ell}}{np}\left(1+\frac{1.01C}{np}\right) + C^{R(\ell-1)}\Gamma_{\mathcal{F}'} \\ &\le \mathrm{O}\left(\frac{n^{R(\ell-1)}p^{R\ell}}{np}\right).\end{aligned}$$

This completes the proof of Lemma 9.3.8 and thus Lemma 9.3.6 (Here, we have used the assumption that $(np)^\ell = \Omega(n)$). □

□

Proof of Lemma 9.3.5. Let $d = (1+\mathrm{o}(1))np = (\beta+\mathrm{o}(1))n^\alpha$. From Lemma 9.1.7, we have $\mathbf{Pr}[H \subseteq G(n,p)] = (1+\mathrm{o}(1))\,\mathbf{Pr}[H \subseteq G_{n,d}]$ for any fixed graph H. Let $R = r_1 + \cdots + r_k$ and $\mathcal{A} = (\mathcal{P}^{(1)})_{r_1} \times \cdots \times (\mathcal{P}^{(k)})_{r_k}$. We write each element $A \in \mathcal{A}$ as a tuple $A = (P_1, \ldots, P_R)$ of R paths. Then, from Corollary 9.3.7, we have

$$\begin{aligned}
\mathbf{E}_{G_{n,d}}\left[\prod_{i=1}^{k}\left(X^{(i)}\right)_{r_i}\right] &= \sum_{(P_1,\ldots,P_R)\in\mathcal{A}} \mathbf{Pr}[E(P_1\cup\cdots P_R)\subseteq G_{n,d}] \\
&= (1+\mathrm{o}(1)) \sum_{(P_1,\ldots,P_R)} \mathbf{Pr}[E(P_1\cup\cdots P_R)\subseteq G(n,p)] \\
&= (1+\mathrm{o}(1))\,\mathbf{E}_{G(n,p)}\left[\prod_{i=1}^{k}\left(X^{(i)}\right)_{r_i}\right] \\
&= (\beta+\mathrm{o}(1))^{1/\alpha}.
\end{aligned}$$

□

9.4 Concentration of $\mathrm{AD}(G(n,p))$

We prove Theorem 9.1.3. We use $\mathrm{AD} = \mathrm{AD}(G(n,p))$ and $\mathrm{diam} = \mathrm{diam}(G(n,p))$ as random variables. Let $D = \lceil \mu \rceil = \lfloor \alpha^{-1} \rfloor + 1$. From Corollary 9.1.5, we have

$$\begin{aligned}
\mathbf{Pr}\left[|\mathrm{AD}-\mu| > \epsilon\right] &\le \mathbf{Pr}\left[|\mathrm{AD}-\mu| > \epsilon \,|\, \mathrm{diam} = D\right]\mathbf{Pr}[\mathrm{diam} = D] + \mathbf{Pr}[\mathrm{diam} \neq D] \\
&\le \mathbf{Pr}\left[|\mathrm{AD}-\mu| > \epsilon \,|\, \mathrm{diam} = D\right] + \mathrm{o}(1)
\end{aligned}$$

for any $\epsilon = \epsilon(n) > 0$. Therefore, we may put the condition that $\mathrm{diam} = D$.

For $i = 1, \ldots, D$, let

$$N_i = \left|\left\{\{s,t\} \in \binom{V}{2} : \mathrm{dist}(s,t) = i\right\}\right|.$$

We will prove the following result in Section 9.4.1:

Lemma 9.4.1. *Let $C > 0$ be a sufficiently large constant and $\epsilon = \epsilon(n) := \sqrt{\frac{\log n}{np}}$. Then, $|N_i - M_i| \le C\epsilon M_i$ holds a.a.s. for all $i = 1, \ldots, D-1$, where*

$$M_i = \begin{cases} \frac{(np)^i}{n}\binom{n}{2} & \text{if } i < \alpha^{-1}, \\ (1-\exp(-\beta^{1/\alpha}))\binom{n}{2} & \text{if } i = \alpha^{-1} \in \mathbb{N}. \end{cases}$$

Upper bound of AD. Conditioned on $\mathrm{diam} = D$, it immediately holds that $\mathrm{AD} \le \mathrm{diam} \le D$. Thus, if $\alpha^{-1} \notin \mathbb{N}$, we have

$$\mathrm{AD} \le D = \mu$$

with probability $1 - \exp(-n^{\Omega(n)})$.

Now we focus on the case where $\alpha^{-1} \in \mathbb{N}$. Let $\epsilon = C\sqrt{\frac{\log n}{np}}$ for sufficiently large constant $C > 0$. Conditioned on $\mathrm{diam} = D$, Lemma 9.4.1 implies

$$\begin{aligned}
N_D &= \binom{n}{2} - N_1 - \ldots - N_{D-1} \\
&\le (1+\mathrm{O}(\epsilon))\exp(-\beta^{1/\alpha})\binom{n}{2}.
\end{aligned}$$

Therefore, conditioned on diam $= D$, we have

$$\begin{aligned}\binom{n}{2} \cdot \mathrm{AD} &= \sum_{i=1}^{D} iN_i \\ &\le DN_D + (D-1)\left(\binom{n}{2} - N_D\right) \\ &= N_D + (D-1)\binom{n}{2} \\ &\le (1+\mathrm{O}(\epsilon))\mu\binom{n}{2}.\end{aligned}$$

In other words, $\mathrm{AD} \le \mu + \mathrm{O}(\epsilon)$ holds a.a.s.

Lower bound of AD. Conditioned on diam $= D$, we have $N_1 + \cdots + N_D = \binom{n}{2}$ and thus

$$\begin{aligned}\binom{n}{2} \cdot \mathrm{AD} &= \sum_{i=1}^{D} iN_i \\ &= N_1 + 2N_2 + \cdots + (D-1)N_{D-1} + D\left(\binom{n}{2} - N_1 - \cdots - N_{D-1}\right) \\ &= D\binom{n}{2} - (D-1)N_1 - (D-2)N_2 - \cdots - N_{D-1} \\ &\ge (1-\mathrm{O}(\epsilon))\mu\binom{n}{2}.\end{aligned}$$

In the last inequality, we used Lemma 9.4.1. This completes the proof of Theorem 9.1.3.

9.4.1 Proof of Lemma 9.4.1

The proof of Lemma 9.4.1 is a slight modification of the proof of Theorem 7.1 of [FK16].

Consider $G(n,p)$ of $p = (\beta + \mathrm{o}(1))n^{-1+\alpha}$. Let $D = \lfloor \alpha^{-1} \rfloor + 1$. We consider the breadth first search process on $G(n,p)$ from a fixed vertex. Fix a vertex v. For $k \ge 0$, let

$$N_k(v) = \{w \in V : \mathrm{dist}(v,w) = k\}.$$

Note that $N_0(v) = \{v\}$. For sufficiently large constant $C > 0$ and $\epsilon := \sqrt{\frac{\log n}{np}}$, let $\mathcal{F}_k$ be the event of $G(n,p)$ that

$$\left| |N_i(v)| - \frac{2M_i}{n} \right| \le \frac{C\epsilon M_i}{n} \text{ for all } i = 1, \dots, k,$$

where M_i is given in Lemma 9.4.1. Note that $\mathcal{F}_0$ must hold. The degree of v is denoted by $\deg(v)$. We denote by $\mathrm{Bin}(m,q)$ the binomial distributed random variable with m trials and success probability q. Note that, if we are given $N_0(v), \dots, N_{k-1}(v)$, the random variable $|N_k(v)|$ is distributed as a binomial random variable, that is,

$$|N_k(v)| \sim \mathrm{Bin}\left(n - \sum_{i=0}^{k-1} |N_i(v)|,\ 1 - (1-p)^{|N_{k-1}(v)|}\right).$$

Consider $\mathbf{E}[|N_k(v)| \,|\, \mathcal{F}_{k-1}]$. For every $k = 1, \dots, D-1$, conditioned on $\mathcal{F}_{k-1}$, we have

$$n \ge n - \sum_{i=0}^{k-1} |N_i(v)| \ge (1-\mathrm{O}(\epsilon))n$$

Here, recall that $(np)^{D-1} = \mathrm{O}(n)$. Using the inequality $\mathrm{e}^{-\frac{x}{1-x}} \le 1 - x \le \mathrm{e}^{-x}$ for every $x \in [0,1)$ (cf. Lemma 21.1 of [FK16]), we obtain

$$1 - (1-p)^{|N_{k-1}(v)|} = \begin{cases} (1 \pm \mathrm{O}(\epsilon))p(np)^{k-1} & \text{if } k = 1, \dots, D-2, \\ (1 \pm \mathrm{O}(\epsilon))\exp(-\beta^{1/\alpha}) & \text{if } k = D-1. \end{cases}$$

Therefore, we have

$$\begin{aligned}\mathbf{E}\left[|N_k(v)| \,|\, \mathcal{F}_{k-1}\right] &= \begin{cases}(1 \pm \mathrm{O}(\epsilon))(np)^k & \text{if } k = 1, \ldots, D-2, \\ (1 \pm \mathrm{O}(\epsilon)) \exp(-\beta^{1/\alpha})n & \text{if } k = D-1\end{cases} \\ &= (1 \pm \mathrm{O}(\epsilon))\frac{2M_k}{n}.\end{aligned}$$

From the Chernoff bound (Proposition 2.5.5), we have

$$\begin{aligned}\mathbf{Pr}[\mathcal{F}_k \,|\, \mathcal{F}_{k-1}] &\geq 1 - \exp\left(-\Theta\left(\epsilon^2 (np)^k\right)\right) \\ &\geq 1 - \mathrm{O}(n^{-2})\end{aligned}$$

if the constant C is sufficiently large (recall that C is the constant in the definition of $\mathcal{F}_k$). Therefore, $\mathcal{F}_{D-1}$ holds with probability $1 - \mathrm{O}(n^{-2})$ for sufficiently large C. Taking the union bound, it holds a.a.s. that $|N_i(v)| = (1 \pm \mathrm{O}(\epsilon))\frac{2M_i}{n}$ for all v. Consequently, we have $N_i = \frac{1}{2}\sum_{v \in V} |N_i(v)| = (1 \pm \mathrm{O}(\epsilon))M_i$, which completes the proof of Lemma 9.4.1.

Chapter 10

Conclusion

In this thesis, we have studied the average-case complexity of graph problems and the behavior of algorithms on random graphs.

In Chapters 3 and 4, we obtained the nearly-tight average-case complexity of counting biclique-subgraphs in random graphs. This reveals a computational hardness aspect of random graphs. In Chapter 3, we proved that no $n^{a-\epsilon}$-time algorithm counts $K_{a,b}$-subgraphs even when the input is a random bipartite graph for any constant $\epsilon > 0$ under a widely investigated conjecture of worst-case complexity. On the other hand, we presented an $n^{a+o(1)}$-time algorithm that solves the counting problem for *any* input. The main issue in this result was the fraction of hard instances: Actually, the hardness result above implies that counting biclique-subgraphs on more than a $(1 - 1/\operatorname{polylog}(n))$-fraction of instances is hard for any $n^{a-\epsilon}$-time algorithms. In other words, it ensures that a $1/\operatorname{polylog}(n)$-fraction of random graphs is hard. We handled this issue in Chapter 4 by presenting a general framework of fine-grained hardness amplification. The core of this framework was a doubly-efficient interactive proof system with low query complexity. We presented such an interactive proof system for the subgraph counting problem, which is of independent interest.

The topic of Chapters 5 to 7 was voting processes. In Chapter 5, we introduced the notion of the functional voting process, which contains several previously-known voting processes. In Chapter 6, we showed phase transition results of the best-of-two and best-of-three on the stochastic block model. Our technical contribution here is to present a framework for studying voting processes on the stochastic block model based on induced dynamical systems. This framework is general and we can apply it to the analysis of voting processes on graphs having more than two community structures. In Chapter 7, we introduced the notion of quasi-majority functional voting, which contains several known voting processes such as the best-of-two and best-of-three. In Chapter 7, we studied quasi-majority functional voting, which is a wide class of voting processes containing best-of-two and best-of-three as special cases. Then we obtained upper bounds of the consensus time of the quasi-majority functional voting on expander graphs. This result generalized and improved several previous works. In particular, we obtained the tight consensus time of the best-of-two and best-of-three on the Erdős–Rényi graph and random regular graph.

In Chapter 8, motivated by studying real-world networks, we introduced the model of random walks on growing graphs (RWoGG). The novelty of the presented model is that the size of the underlying graph increases over time, which reflects the convention that most real world networks are growing all the time. We then evaluate the performance of an RWoGG using the expected number of unvisited vertices during the random walk. We obtained several bounds of this value in terms of hitting and mixing time.

In Chapter 9, we obtained the asymptotic behavior of the average distance and diameter of random regular graphs. In particular, we proved that the diameter of dense random regular graphs are asymptotically optimal if the degree satisfies some mild condition.

This thesis investigated random graphs drawn from simple distributions such as random bipartite graph, stochastic block model, and random regular graph. However, these kinds of random graphs usually do not appear in the real-world. Our future direction is the analysis of randomness on real-world networks such as preferential attachment model.

Bibliography

[AAB+11] Y. Afek, N. Alon, O. Barad, E. Hornstein, N. Barkai, and Z. Bar-Joseph. A biological solution to a fundamental distributed computing problem. *Science*, 331(6014):183–185, 2011.

[AB08] S. Arora and B. Barak. *Computational Complexity: A Modern Approach*. Cambridge University Press, 2008.

[Abb18] E. Abbe. Community detection and stochastic block models: recent developments. *Journal of Machine Learning Research*, 18(177):1–86, 2018.

[AD15] M. A. Abdullah and M. Draief. Global majority consensus by local majority polling on graphs of a given degree sequence. *Discrete Applied Mathematics*, 1(10):1–10, 2015.

[AF] D. Aldous and J. Fill. Reversible Markov chains and random walks on graphs. http://statwww.berkeley.edu/pub/users/aldous/RWG/book.html.

[AKL+79] R. Aleliunas, R. M. Karp, R. J. Lipton, L. Lovász, and C. Rackoff. Random walks, universal traversal sequences, and the complexity of maze problems. *In Proceedings of the 20th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 218–223, 1979.

[AKL08] C. Avin, M. Kouský, and Z. Lotler. How to explore a fast-changing world (cover time of a simple random walk on evolving graphs). *In Proceedings of the 35th International Colloquium on Automata, Languages, and Programming (ICALP)*, pages 121–132, 2008.

[AKL18] C. Avin, M. Kouský, and Z. Lotler. Cover time and mixing time of random walks on dynamic graphs. *Random Structures & Algorithms*, 52(4):576–596, 2018.

[Ald83] D. J. Aldous. On the time taken by random walks on finite groups to visit every state. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 62:361–374, 1983.

[AS94] R. Agrawal and R. Srikant. Fast Algorithms for Mining Association Rules in Large Databases. *In Proceedings of the 20th International Conference on Very Large Data Bases (VLDB)*, pages 487–499, 1994.

[AS15] E. Abbe and C. Sandon. Recovering communities in the general stochastic block model without knowing the parameters. *In Proceedings of the 28th International Conference on Neural Information Processing Systems (NIPS)*, 1:676–684, 2015.

[AV79] D. Angluin and L. G. Valiant. Fast probabilistic algorithms for hamiltonian circuits and matchings. *Journal of Computer and System Sciences*, 18(2):155–193, 1979.

[AVJ98] J. Amilhastre, M. C. Vilarem, and P. Janssen. Complexity of minimum biclique cover and minimum biclique decomposition for bipartite domino-free graphs. *Discrete Applied Mathematics*, 86(2–3):125–144, 1998.

[AW21] J. Alman and V. V. Williams. A Refined Laser Method and Faster Matrix Multiplication. *In Proceedings of the 32nd Annual ACM-SIAM Symposium on Discrete Algorithms (SODA), to appear*, 2021.

[AWY15] A. Abboud, R. Williams, and H. Yu. More applications of the polynomial method to algorithm design. *In Proceedings of the 26th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 218–230, 2015.

[AYZ97] N. Alon, R. Yuster, and U. Zwick. Finding and counting given length cycles. *Algorithmica*, 17(3):209–223, 1997.

[BABB19] E. Boix-Adserà, M. Brennan, and G. Bresler. The average-case complexity of counting cliques in Erdős-Rényi hypergraphs. *In Proceedings of the 60th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 1256–1280, 2019.

[BCE+17] P. Berenbrink, A. Clementi, R. Elsässer, P. Kling, F. Mallmann-Trenn, and E. Natale. Ignore or comply? On breaking symmetry in consensus. *In Proceedings of the ACM Symposium on Principles of Distributed Computing (PODC)*, pages 335–344, 2017.

[BCM+18] L. Becchetti, A. Clementi, P. Manurangsi, E. Natale, F. Pasquale, P. Raghavendra, and L. Trevisan. Average whenever you meet: Opportunistic protocols for community detection. *In Proceedings of the 26th Annual European Symposium on Algorithms (ESA)*, 7:1–13, 2018.

[BCN+16] L. Becchetti, A. Clementi, E. Natale, F. Pasquale, and L. Trevisan. Stabilizing consensus with many opinions. *In Proceedings of the 27th annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 620–635, 2016.

[BCN+17a] L. Becchetti, A. Clementi, E. Natale, F. Pasquale, R. Silvestri, and L. Trevisan. Simple dynamics for plurality consensus. *Distributed Computing*, 30(4):293–306, 2017.

[BCN+17b] L. Becchetti, A. Clementi, E. Natale, F. Pasquale, and L. Trevisan. Find your place: Simple distributed algorithms for community detection. *In Proceedings of the 28th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 940–959, 2017.

[BCO+16] I. Benjamini, S.-O. Chan, R. O'Donnell, O. Tamuzc, and L.-Y. Tand. Convergence, unanimity and disagreement in majority dynamics on unimodular graphs and random graphs. *Stochastic Processes and their Applications*, 126(9):2719–2733, 2016.

[BDLBH17] P. Barbillon, S. Donnet, E. Lazega, and A. Bar-Hen. Stochastic block models for multiplex networks: an application to a multilevel network of researchers. *Journal of the Royal Statistical Society Series A*, 180(1):295–314, 2017.

[BdlV82] B. Bollobás and W. F. de la Vega. The diameter of random regular graphs. *Combinatorica*, 2(2):125–134, 1982.

[BE76] B. Bollobás and P. Erdős. Cliques in random graphs. *Mathematical Proceedings of the Cambridge Philosophical Society*, 80(3):419–427, 1976.

[Ber01] E. Berger. Dynamic monopolies of constant size. *Journal of Combinatorial Theory Series B*, 83(2):191–200, 2001.

[BGHJ07] V. D. Blondel, J.-L. Guillaume, J. M. Hendrickx, and R. M. Jungers. Distance distribution in random graphs and application to network exploration. *Physical Review E*, 76(066101), 2007.

[BGKMT16] P. Berenbrink, G. Giakkoupis, A.-M. Kermarrec, and F. Mallmann-Trenn. Bounds on the voter model in dynamic networks. *In Proceedings of the 43rd International Colloquium on Automata, Languages, and Programming (ICALP)*, pages 146:1–146:15, 2016.

[BI73] E. Bannai and T. Ito. On finite Moore graphs. *Journal of the Faculty of Science, the University of Tokyo.*, 20(2):191–208, 1973.

[BK95] M. Blum and S. Kannan. Designing programs that check their work. *Journal of the ACM*, 42(1):269–291, 1995.

[Bol80] B. Bollobás. A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European Journal of Combinatorics*, 1(4):311–316, 1980.

[Bol81] B. Bollobás. The diameter of random graphs. *Transactions of the American Mathematical Society*, 267(1):41–52, 1981.

[Bol84] B. Bollobás. *The evolution of sparse graphs*, pages 35–57. Graph Theory and Combinatorics, Academic Press, 1984.

[Bol01] B. Bollobás. *Random Graphs*. Cambridge University Press, 2 edition, 2001.

[BRFGL10] D. Binkele-Raible, H. Fernau, S. Gaspers, and M. Liedloff. Exact exponential-time algorithms for finding bicliques. *Information Processing Letters*, 111(2):64–67, 2010.

[BRSV17] M. Ball, A. Rosen, M. Sabin, and P. N. Vasudevan. Average-case fine-grained hardness. *In Proceedings of the 49th Annual ACM Symposium on Theory of Computing (STOC)*, pages 483–496, 2017.

[BRSV18] M. Ball, A. Rosen, M. Sabin, and P. N. Vasudevan. Proofs of Work from worst-case assumptions. *Advances in Cryptology – CRYPTO 2018, In Proceedings of the 38th Annual International Cryptology Conference*, pages 789–819, 2018.

[BT06] A. Bogdanov and L. Trevisan. Average-Case Complexity. *Foundations and Trends in Theoretical Computer Science*, 2(1), 2006.

[BW86] E. R. Berlekamp and L. R. Welch. Error correction of algebraic block codes. *US Patent Number 4633470*, 1986.

[CDM17] R. Curticapean, H. Dell, and D. Marx. Homomorphisms are a good basis for counting small subgraphs. *In Proceedings of the 49th Annual ACM Symposium on Theory of Computing (STOC)*, pages 210–223, 2017.

[CEOR13] C. Cooper, R. Elsässer, H. Ono, and T. Radzik. Coalescing random walks and voting on connected graphs. *SIAM Journal on Discrete Mathematics*, 27(4):1748–1758, 2013.

[CER14] C. Cooper, R. Elsässer, and T. Radzik. The power of two choices in distributed voting. *In Proceedings of the 41st International Colloquium on Automata, Languages, and Programming (ICALP)*, 2:435–446, 2014.

[CER+15] C. Cooper, R. Elsässer, T. Radzik, N. Rivera, and T. Shiraga. Fast consensus for voting on general expander graphs. *In Proceedings of the 29th International Symposium on Distributed Computing (DISC)*, pages 248–262, 2015.

[CF03] C. Cooper and A. Frieze. Crawling on simple models of web graphs. *Internet Mathematics*, 1(1):57–90, 2003.

[CGG+18] A. Clementi, M. Ghaffari, L. Gualà, E. Natale, F. Pasquale, and G. Scornavacca. A tight analysis of the parallel undecided-state dynamics with two colors. *In Proceedings of the 43rd International Symposium on Mathematical Foundations of Computer Science (MFCS)*, 117(28):1–15, 2018.

[CGJ18] N. Cook, L. Goldstein, and T. Johnson. Size biased couplings and the spectral gap for random regular graphs. *The Annals of Probability*, 46(1):72–125, 2018.

[CHKX06] J. Chen, X. Huang, I. A. Kanj, and G. Xia. Strong computational lower bounds via parameterized complexity. *Journal of Computer and System Sciences*, 72(8):1346–1367, 2006.

[CK12] J.-F. Couturier and D. Kratsch. Bicolored independent sets and bicliques. *Information Processing Letters*, 112(8–9):329–334, 2012.

[CL01] F. Chung and L. Lu. The diameter of sparse random graphs. *Advances in Applied Mathematics*, 26(4):257–279, 2001.

[CL04] F. Chung and L. Lu. The average distance in a random graph with given expected degrees. *Internet Mathematics*, 1(1):91–113, 2004.

[CM14] R. Curticapean and D. Marx. Complexity of counting subgraphs: Only the boundedness of the vertex-cover number counts. *In Proceedings of the 55th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 130–139, 2014.

[CNNS18] E. Cruciani, E. Natale, A. Nusser, and G. Scornavacca. Phase transition of the 2-choices dynamics on core-periphery networks. *In Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS)*, pages 777–785, 2018.

[CNS19] E. Cruciani, E. Natale, and G. Scornavacca. Distributed community detection via metastability of the 2-choices dynamics. *In Proceedings of the 33rd AAAI conference on artificial intelligence (AAAI)*, pages 6046–6053, 2019.

[CO07] A. Coja-Oghlan. On the Laplacian eigenvalues of $G_{n,p}$. *Combinatorics, Probability and Computing*, 16(6):923–946, 2007.

[Coo11] C. Cooper. Random walks, interacting particles, dynamic networks: randomness can be helpful. *In Proceeedings of the 18th International Colloquium on Structural Information and Communication Complexity (SIROCCO)*, pages 1–14, 2011.

[Cop97] D. Coppersmith. Rectangular matrix multiplication revisited. *Journal of Complexity*, 13(1):42–49, 1997.

[CPS99] J.-Y. Cai, A. Pavan, and D. Sivakumar. On the hardness of permanent. *In Proceedings of the 16th Annual Symposium on Theoretical Aspects of Computer Science (STACS)*, pages 90–99, 1999.

[CR16] C. Cooper and N. Rivera. The linear voting model. *In Proceedings of the 43rd International Colloquium on Automata, Languages, and Programming (ICALP)*, 55(144):1–12, 2016.

[CRRS17] C. Cooper, T. Radzik, N. Rivera, and T. Shiraga. Fast plurality consensus in regular expanders. *In Proceedings of the 31st International Symposium on Distributed Computing (DISC)*, 91(13):1–16, 2017.

[CST15] A. Clementi, R. Silvestri, and L. Trevisan. Information spreading in dynamic graphs. *Distributed Computing*, 28:55–73, 2015.

[CSZ20] L. Cai, T. Sauerwald, and L. Zanetti. Random walks on randomly evolving graphs. *In Proceedings of the 27th International Colloquium on Structural Information and Communication Complexity (SIROCCO)*, pages 111–128, 2020.

[Cur13] R. Curticapean. Counting matchings of Size k Is #W[1]-hard. *In Proceedings of the 40th International Colloquium on Automata, Languages, and Programming (ICALP)*, pages 352–363, 2013.

[Cur18] R. Curticapean. Counting problems in parameterized complexity. *In Proceedings of the 13th International Symposium on Parameterized and Exact Computation (IPEC)*, pages 1:1–1:18, 2018.

[CY06] J. Chen and B. Yuan. Detecting functional modules in the yeast protein-protein interaction network. *Bioinformatics*, 22(18):2283–2290, 2006.

[DAB+04] A. C. Driskell, C. Ané, J. G. Burleigh, M. M. McMahon, and M. J. Sanderson. Prospects for building the tree of life from large sequence databases. *Science*, 306(5699):1172–1174, 2004.

[Del85] C. Delorme. Grands graphes de degré et diamè et diamètre donnés. *European Journal of Combinatorics*, 6(4):291–302, 1985.

[DF17] R. David and U. Feige. Random walks with the minimum degree local rule have $O(N^2)$ cover time. *In Proceedings of the 28th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1839–1848, 2017.

[DF18] R. David and U. Feige. Random walks with the minimum degree local rule have $O(n^2)$ cover time. *SIAM Journal on Computing*, 47(3):755–768, 2018.

[DFRv17] A. Dudek, A. Frieze, A. Ruciński, and M. Šileikis. Embedding the Erdős-Rényi hypergraph into the random regular hypergraph and Hamiltonicity. *Journal of Combinatorial Theory, Series B*, 122:719–740, 2017.

[DGM+11] B. Doerr, L. A. Goldberg, L. Minder, T. Sauerwald, and C. Scheideler. Stabilizing consensus with the power of two choices. *In Proceedings of the 23rd Annual ACM symposium on Parallelism in algorithms and architectures (SPAA)*, pages 149–158, 2011.

[DLW20] M. Dalirrooyfard, A. Lincoln, and V. V. Williams. New Techniques for Proving Fine-Grained Average-Case Hardness. *In Proceedings of the 61st Annual IEEE Symposium on Foundations of Computer Science (FOCS), to appear*, 2020.

[DN20] B. Doerr and F. Neumann. *Theory of evolutionary computation: Recent developments in discrete optimization*. Springer International Publishing, 2020.

[DR14] O. Denysyuk and L. Rodrigues. Random walks on evolving graphs with recurring topologies. *In Proceedings of the 28th International Symposium on Distributed Computing (DISC)*, pages 333–345, 2014.

[DS84] P. G. Doyle and J. L. Snell. Random Walks and Electric Networks. *Mathematical Association of America*, pages 361–374, 1984.

[DT10] B. Daniel and T. Tassa. Improved bounds on Bell numbers and on moments of sums of random variables. *Probability and Mathematical Statistics*, 30(2):185–205, 2010.

[Dur19] R. Durrett. *Probability: Theory and Examples*. Campridge University Press, 2019.

[DVW96] A. Denise, M. Vasconcellos, and D. J. A. Welsh. The Random Planar Graph. *Congressus Numerantium*, 113:61–79, 1996.

[EFH80] P. Erdős, S. Fajtlowicz, and A. J. Hoffman. Maximum degree in graphs of diameter 2. *Networks*, 10(1):87–90, 1980.

[ER59] P. Erdős and A. Rényi. On random graphs I. *Publicationes Mathematicae (Debrecen),6 (1959)*, 6:290–297, 1959.

[ER66] P. Erdős and A. Rényi. On the existence of a factor of degree one of a connected random graph. *Acta Mathematica Academiae Scientiarum Hungarica*, 17:359–368, 966.

[Erd59] P. Erdős. Graph Theory and Probability. *Canadian Journal of Mathematics*, 11:34–38, 1959.

[Fei95a] U. Feige. A tight lower bound on the cover time for random walks on graphs. *Random Structures & Algorithms*, 6(4):433–438, 1995.

[Fei95b] U. Feige. A tight upper bound on the cover time for random walks on graphs. *Random Structures & Algorithms*, 6(1):51–54, 1995.

[FK98] U. Feige and J. Kilian. Zero Knowledge and the Chromatic Number. *Journal of Computer and System Sciences*, 57(2):187–199, 1998.

[FK16] A. Frieze and M. Karońsky. *Introduction to Random Graphs*. Campridge University Press, 2016.

[FLM86] M. Fischer, N. Lynch, and M. Merritt. Easy impossibility proofs for distributed consensus problems. *Distributed Computing*, 1(1):26–39, 1986.

[FM97] A. Frieze and C. McDiarmid. Algorithmic theory of random graphs. *Random Structures & Algorithms*, 10(1–2):5–42, 1997.

[FR07] D. Fernholz and V. Ramachandran. The diameter of sparse random graphs. *Random Structures & Algorithms*, 31(4):482–516, 2007.

[Gal12] F. L. Gall. Faster algorithms for rectangular matrix multiplication. *In Proceedings of the 53rd Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 514–523, 2012.

[Gal14] F. L. Gall. Powers of tensors and fast matrix multiplication. *In Proceedings of the 39th International Symposium on Algorithms and Computation (ISAAC)*, pages 296–303, 2014.

[Gil59] E. N. Gilbert. Random Graphs. *The Annals of Mathematical Statistics*, 30(4):1141–1144, 1959.

[Gil61] E. N. Gilbert. Random Plane Networks. *Journal of the Society for Industrial and Applied Mathematics*, 9(4):533–543, 1961.

[GIM20] P. Gao, M. Isaev, and B. D. McKay. Sandwiching random regular graphs between binomial random graphs. *In Proceedings of the 31st Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 690–701, 2020.

[GJ79] M. R. Garey and D. S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W. H. Freeman and Company, 1979.

[GK10] S. Gilbert and D. Kowalski. Distributed agreement with optimal communication complexity. *In Proceedings of the 21st Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 965–977, 2010.

[GK20] E. Goldenberg and Karthik C. S. Hardness amplification of optimization problems. *In Proceedings of the 11th Innovations in Theoretical Computer Science Conference (ITCS)*, pages 1:1–1:13, 2020.

[GKL12] S. Gaspers, D. Kratsch, and M. Liedloff. On independent sets and bicliques in graphs. *Algorithmica*, 62:637–658, 2012.

[GKR15] S. Goldwasser, Y. T. Kalai, and G. N. Rothblum. Delegating computation: Interactive proofs for muggles. *Journal of the ACM*, 62(4), 2015.

[GL89] O. Goldreich and L. A. Levin. A hard-core predicate for all one-way functions. *In Proceedings of the 21st Annual ACM Symposium on Theory of Computing (STOC)*, pages 25–32, 1989.

[GL18] M. Ghaffari and J. Lengler. Nearly-tight analysis for 2-choice and 3-majority consensus dynamics. *In Proceedings of the 37th ACM Symposium on Principles of Distributed Computing (PODC)*, pages 305–313, 2018.

[GM75] G. R. Grimmett and C. J. H. McDiarmid. On colouring random graphs. *Mathematical Proceedings of the Cambridge Philosophical Society*, 77(2):313–324, 1975.

[GMR89] S. Goldwasser, S. Micali, and C. Rackoff. The knowledge complexity of interactive proof-systems. *SIAM Journal on Computing*, 18(1):291–304, 1989.

[GNW11] O. Goldreich, N. Nisan, and A. Wigderson. On Yao's XOR-Lemma. *Studies in Complexity and Cryptography. Miscellanea on the Interplay between Randomness and Computation*, pages 273–301, 2011.

[Gol20] O. Goldreich. On counting t-Cliques mod 2. *Electronic Colloquium on Computational Complexity (ECCC)*, 20-104, 2020.

[GR18a] O. Goldreich and G. N. Rothblum. Counting t-cliques: Worst-case to average-case reductions and direct interactive proof systems. *In Proceedings of the 59th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 77–88, 2018.

[GR18b] O. Goldreich and G. N. Rothblum. Simple doubly-efficient interactive proof systems for locally-characterizable sets. *In Proceedings of the 9th Innovations in Theoretical Computer Science (ITCS)*, pages 18:1–18:19, 2018.

[GS92] P. Gemmell and M. Sudan. Highly resilient correctors for polynomials. *Information Processing Letters*, 43(4):169–174, 1992.

[GU18] F. L. Gall and F. Urrutia. Improved rectangular matrix multiplication using powers of the Coppersmith-Winograd tensor. *In Proceedings of the 29th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1029–1046, 2018.

[GW15] P. Gao and N. C. Wormald. Uniform generation of random regular graphs. *In Proceedings of the 56th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 1218–1230, 2015.

[GZ18] B. Gärtner and A. N. Zehmakan. Majority model on random regular graphs. *In Proceedings of the 13th Latin American Symposium on Theoretical Informatics (LATIN)*, pages 572–583, 2018.

[GZFA10] A. Goldenberg, A. X. Zheng, S. E. Fienberg, and E. M. Airoldi. A survey of statistical network models. *Foundations and Trends in Machine Learning*, 2(2):129–233, 2010.

[HILL99] J. Håstad, R. Impagliazzo, L. A. Levin, and M. Luby. A pseudorandom generator from any one-way function. *SIAM Journal on Computing*, 28(4):1364–1396, 1999.

[Hir15] S. Hirahara. Identifying an honest EXPˆNP oracle among many. *In Proceedings of the 30th Conference on Computational Complexity (CCC)*, pages 244–263, 2015.

[HKZZ19] T. D. Hansen, H. Kaplan, O. Zamir, and U. Zwick. Faster k-SAT algorithms using biased-PPSZ. *In Proceedings of the 51st Annual ACM Symposium on Theory of Computing (STOC)*, pages 578–589, 2019.

[Hor72] E. Horowitz. A fast method for interpolation using preconditioning. *Information Processing*, 1(4):157–163, 1972.

[HP01] Y. Hassin and D. Peleg. Distributed probabilistic polling and applications to proportionate agreement. *Information and Computation*, 171(2):248–268, 2001.

[HS60] A. J. Hoffman and R. R. Singleton. On Moore graphs with diameters 2 and 3. *IBM Journal of Research and Development*, 4(5):497–504, 1960.

[HS05] M. Hirsch and H. L. Smith. Monotone dynamical systems. *In Handbook of Differential Equations: Ordinary Differential Equations*, 2(4):239–357, 2005.

[IJK09] R. Impagliazzo, R. Jaiswal, and V. Kabanets. Approximate list-decoding of direct product codes and uniform hardness amplification. *SIAM Journal on Computing*, 39(2):564–605, 2009.

[IJKW10] R. Impagliazzo, R. Jaiswal, V. Kabanets, and A. Wigderson. Uniform direct product theorems: Simplified, optimized, and derandomized. *SIAM Journal on Computing*, 39(4):1637–1665, 2010.

[IKOY03] S. Ikeda, I. Kubo, N. Okumoto, and M. Yamashita. Impact of local topological information on random walks on finite graphs. *In Proceedings of the 30th International Colloquium on Automata, Languages and Programming (ICALP)*, pages 1054–1067, 2003.

[IKY09] 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(1):94–100, 2009.

[IP01] R. Impagliazzo and R. Paturi. On the complexity of k-SAT. *Journal of Computer and System Sciences*, 62(2):367–375, 2001.

[IPZ01] R. Impagliazzo, R. Paturi, and F. Zane. Which problems have strongly exponential complexity? *Journal of Computer and System Sciences*, 63(4):512–530, 2001.

[IW97] R. Impagliazzo and A. Wigderson. P = BPP if E requires exponential circuits: Derandomizing the XOR lemma. *In Proceedings of the 29th Annual ACM Symposium on Theory of Computing (STOC)*, pages 220–229, 1997.

[JAR16] G. P. J. Augustine and P. Robinson. Distributed algorithmic foundations of dynamic networks. *SIGACT News*, 47(1):69–98, 2016.

[JVV86] M. Jerrum, L. G. Valiant, and V. V. Vazirani. Random generation of combinatorial structures from a uniform distribution. *Theoretical Computer Science*, 43:169–188, 1986.

[Kar73] A. V. Karzanov. O nakhozhedenii maksimal'nogo potoka v setykh spetsial'nogo vida i nekotorykh prilozheniyakh. *Matematicheskie Voprosy Upravleniya Proizvodstvom*, 31(1):81–94, 1973.

[Kar93] D. R. Karger. Global min-cuts in RNC, and other ramifications of a simple min-cut algorithm. *In Proceedings of the 4th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 21–30, 1993.

[KFI+16] M. Koibuchi, I. Fujiwara, K. Ishii, S. Namiki, F. Chaix, H. Matsutani, H. Amano, and T. Kudoh. Optical network technologies for HPC: computer-architects point of view. *IEICE Electronics Express*, 13(6):1–14, 2016.

[KL81] V. Klee and D. Larman. Diameters of random graphs. *Canadian Journal of Mathematics*, 33:618–640, 1981.

[KMA+12] M. Koibuchi, H. Matsutani, H. Amano, D. F. Hsu, and H. Cassanova. A case for random shortcut topologies for HPC interconnects. *In Proceedings of the 39th Annual International Symposium on Computer Architecture (ISCA)*, pages 177–188, 2012.

[KMTS19] V. Kanade, F. Mallmann-Trenn, and T. Sauerwald. On coalescence time in graphs: When is coalescing as fast as meeting? *In Proceedings of the 30th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 956–965, 2019.

[KNb+15] E. Katzav, M. Nitzan, D. ben-Avraham, P. L. Krapisky, R. Künh, N. Ross, and O. Biham. Analytical results for the distribution of shortest path lengths in random networks. *A Letters Journal Exploring the Frontiers of Physics*, 111(26006), 2015.

[KO11] F. Kuhn and R. Oshman. Dynamic networks: models and algorithms. *SIGACT News*, 42(1):82–96, 2011.

[KR19] N. Kang and R. Rivera. Best-of-Three voting on dense graphs. *In Proceedings of the 31st ACM Symposium on Parallelism in Algorithms and Architectures (SPAA)*, pages 115–121, 2019.

[Kra16] S. G. Krantz. *Real analysis and foundations.* CRC Press, third edition, 2016.

[KSV07] J. H. Kim, B. Sudakov, and V.H.Vu. Small subgraphs of random regular graphs. *Discrete Mathematics*, 307(15):1961–1967, 2007.

[Kut12] K. Kutzkov. An exact exponential time algorithm for counting bipartite cliques. *Information Processing Letters*, 112(13):535–539, 2012.

[KV00] J. H. Kim and V. H. Vu. Concentration of multivariate polynomials and its applications. *Combinatorica*, 20:417–434, 2000.

[KV04] J. H. Kim and V. H. Vu. Sandwiching random graphs: universality between random graph models. *Advances in Mathematics*, 188(2):444–469, 2004.

[Lev86] L. A. Levin. Average case complete problems. *SIAM Journal on Computing*, 15(1):285–286, 1986.

[LFKN92] C. Lund, L. Fortnow, H. Karloff, and N. Nisan. Algebraic methods for interactive proof systems. *Journal of the ACM*, 39(4), 1992.

[Lig85] T. M. Liggett. *Interacting particle systems.* Springer-Verlag, 1985.

[Lin15] B. Lin. The parameterized complexity of k-biclique. *In Proceedings of the 26th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 605–615, 2015.

[Lin18] B. Lin. The parameterized complexity of k-biclique. *Journal of the ACM*, 65(5), 2018.

[Lip91] R. Lipton. New directions in testing. *In Distributed Computing and Cryptography*, 2, 1991.

[LMS18] I. Lamprou, R. Martin, and P. Spirakis. Cover time in edge-uniform stochastically-evolving graphs. *Algorithms*, 11(10), 2018.

[Lov93] L. Lovász. Random Walks on Graphs: A Survey. *Combinatorics, Paul Erdős is Eighty*, pages 353–398, 1993.

[Lov12] L. Lovász. *Large Networks and Graph Limits*. American Mathematical Society Providence, 2012.

[LP17] D. A. Levin and Y. Peres. *Markov Chain and Mixing Times: Second Edition*. The American Mathematical Society, 2017.

[LPW17] M. Lewenstein, S. Pettie, and V. V. Williams. Structure and Hardness in P (Dagstuhl Seminar 16451). *Dagstuhl Reports*, 6(11):1–34, 2017.

[Łuc98] T. Łuczak. Random trees and random graphs. *Random Structures & Algorithms*, 13(3–4), 1998.

[LWW18] A. Lincoln, V. V. Williams, and R. Williams. Tight hardness for shortest cycles and paths in sparse graphs. *In Proceedings of the 29th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1236–1252, 2018.

[Mat88] P. Matthews. Covering problems for Markov chains. *The Annals of Probability*, 16(3):1215–1228, 1988.

[McK81] B. D. McKay. Subgraphs of random graphs with specified degrees. *Congressus Numeratium*, 33:213–223, 1981.

[Mic16] O. Michail. An introduction to temporal graphs: an algorithmic perspective. *Internet Mathematics*, 12(4):239–280, 2016.

[MNT14] E. Mossel, J. Neeman, and O. Tamuz. Majority dynamics and aggregation of information in social networks. *Autonomous Agents and Multiagent Systems*, 28(3):408–429, 2014.

[Mot94] R. Motwani. Average-case analysis of algorithms for matchings and related problems. *Journal of the ACM*, 41(6):1329–1356, 1994.

[MPN+99] E. M. Marcotte, M. Pellegrini, H.-L. Ng, D. W. Rice, T. O. Yeates, and D. Eisenberg. Detecting protein function and protein-protein interactions from genome sequences. *Science*, 285(5428):751–753, 1999.

[MS18] O. Michail and P. G. Spirakis. Elements of the theory of dynamic networks. *Communications of the ACM*, 61(2):72–81, 2018.

[MT06] R. Montenegro and P. Tetali. *Mathematical aspects of mixing times in Markov chains*. NOW Publishers, 2006.

[MT17] A. P. Mukherjee and S. Tirthapura. Enumerating maximal bicliques from a large graph using MapReduce. *IEEE Transactions on Services Computing*, 10(5):771–784, 2017.

[MU04] K. Makino and T. Uno. New algorithms for enumerating all maximal cliques. *In Proceedings of 9th Scandinavian Workshop on Algorithm Theory (SWAT)*, 3111:260–272, 2004.

[Mv05] M. Miller and J. Širáň. Moore graphs and beyond: A survey of the degree/diameter problem. *The Electronic Journal of Combinatorics, Dynamic Survey*, (14), 2005.

[NIY99] T. Nakata, H. Imahayashi, and M. Yamashita. Probabilistic local majority voting for the agreement problem on finite graph. *In Proceedings of the 5th Annual International Computing and Combinatorics Conference (COCOON)*, pages 330–338, 1999.

[NKKB16] M. Nitzan, E. Katzav, R. Kühn, and O. Biham. Distance distribution in configuration-model networks. *Physical Review E*, 93(062309), 2016.

[NOSY10] Y. Nonaka, H. Ono, K. Sadakane, and M. Yamashita. The hitting and cover times of Metropolis walks. *Theoretical Computer Science*, 411(16–18):1889–1894, 2010.

[NP85] J. Nešetřil and S. Poljak. On the complexity of the subgraph problem. *Commentationes Mathematicae Universitatis Carolinae*, 26(2):415–419, 1985.

[NS94] N. Nisan and M. Szegedy. On the degree of boolean functions as real polynomials. *computational complexity*, 4:301–313, 1994.

[NW94] N. Nisan and A. Wigderson. Hardness vs randomness. *Journal of Computer and System Sciences*, 49(2):149–167, 1994.

[OP19] R. I. Oliveira and Y. Peres. Random walks on graphs: New bounds on hitting, meeting, coalescing and returning. *In Proceedings of the 16th Workshop on Analytic Algorithmics and Combinatorics (ANALCO)*, pages 119–126, 2019.

[Pel98] D. Peleg. Size bounds for dynamic monopolies. *Discrete Applied Mathematics*, 86(2–3):263–273, 1998.

[Pel02] D. Peleg. Local majorities, coalitions and monopolies in graphs: a review. *Theoretical Computer Science*, 282(2):231–257, 2002.

[Pós76] L. Pósa. Hamiltonian circuits in random graphs. *Discrete Mathematics*, 14(4):359–364, 1976.

[PPSZ05] R. Paturi, P. Pudlák, M. E. Saks, and F. Zane. An improved exponential-time algorithm for k-SAT. *Journal of the ACM*, 2005.

[PW10] M. Pătraşcu and R. Williams. On the possibility of faster SAT algorithms. *In Proceedings of the 21st Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1065–1075, 2010.

[Ros18] B. Rossman. Lower bounds for subgraph isomorphism. *In Proceedings of International Congress of Mathematicians (ICM)*, 3:3409–3430, 2018.

[RRR16] O. Reingold, G. N. Rothblum, and R. D. Rothblum. Constant-round interactive proofs for delegating computation. *In Proceedings of the 48th Annual ACM Symposium on Theory of Computing (STOC)*, pages 49–62, 2016.

[RW10] O. Riordan and N. Wormald. The diameter of sparse random graphs. *Combinatorics, Probability and Computing*, 19:835–926, 2010.

[Sax09] N. Saxena. Progress on Polynomial Identity Testing. *Bulletin of the EATCS*, (90):49–79, 2009.

[SCZ09] L. Saloff-Coste and J. Zúñiga. Merging for time inhomogeneous finite Markov chains. I. Singular values and stability. *Electronic Journal of Probability*, 14(49):1456–1494, 2009.

[SCZ11] L. Saloff-Coste and J. Zúñiga. Merging for inhomogeneous finite Markov chains Part II: Nash and log-Sobolev inequalities. *Annals of Probability*, 39(3):1161–1203, 2011.

[Sha92] A. Shamir. IP=PSPACE. *Journal of the ACM*, 39(4), 1992.

[She10] I. G. Shevtsova. An improvement of convergence rate estimates in the Lyapunov theorem. *Doklady Mathematics*, 82(3):862–864, 2010.

[SHPG12] A. Singla, C.-Y. Hong, L. Popa, and P. B. Godfrey. Jellyfish: Networking data centers randomly. *In Proceedings of the 9th USENIX Symposium on Networked Systems Design and Implementation (NSDI)*, pages 225–238, 2012.

[SJ89] A. Sinclair and M. Jerrum. Approximate counting, uniform generation and rapidly mixing Markov chains. *Information and Computation*, 82(1):93–133, 1989.

[SMP15] A. D. Sarma, A. R. Molla, and G. Pandurangan. Distributed computation in dynamic networks via random walks. *Theoretical Computer Science*, 581:45–66, 2015.

[SZ19] T. Sauerwald and L. Zanetti. Random walks on dynamic graphs: Mixing times, hitting times, and return probabilities. *In Proceedings of the 46th International Colloquim on Automata, Languages, and Programming (ICALP)*, pages 93:1–93:15, 2019.

[Tre03] L. Trevisan. List-decoding using the XOR lemma. *In Proceedings of the 44th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 126–135, 2003.

[TY19] K. Tikhomirov and P. Youssef. The spectral gap of dense random regular graphs. *The Annals of Probability*, 47(1):362–419, 2019.

[vdHH08] R. van der Hofstad and G. Hooghiemstra. Universality for distances in power-law random graphs. *Journal of Mathematical Physics*, 49(125209), 2008.

[vdHHM05] R. van der Hofstad, G. Hooghiemstra, and P. V. Mieghem. Distances in random graphs with finite variance degrees. *Random Structures & Algorithms*, 27(1):76–123, 2005.

[Wil05] R. Williams. A new algorithm for optimal 2-constraint satisfaction and its implications. *Theoretical Computer Science*, 348(2):357–365, 2005.

[Wil12] V. V. Williams. Multiplying matrices faster than Coppersmith-Winograd. *In Proceedings of the 44th Annual ACM Symposium on Theory of Computing (STOC)*, pages 887–898, 2012.

[Wil15] V. V. Williams. Hardness of easy problems: Basing hardness on popular conjectures such as the Strong Exponential Time Hypothesis. *In Proceedings of the 10th International Symposium on Parameterized and Exact Computation (IPEC)*, pages 17–29, 2015.

[Wor99] N. C. Wormald. Models of random regular graphs. *Surveys in Combinatorics*, 267:239–298, 1999.

[Yao82] A. C. Yao. Theory and application of trapdoor functions. *In Proceedings of the 23rd Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 80–91, 1982.

[YM16] W. Yu and J. A. McCann. Random walk with restart over dynamic graphs. *In Proceedings of the 16th IEEE International Conference on Data Mining (ICDM)*, pages 589–598, 2016.

[YZ97] R. Yuster and U. Zwick. Finding even cycles even faster. *SIAM Journal on Discrete Mathematics*, 10(2):209–222, 1997.

[Zeh18] A. N. Zehmakan. Opinion forming in Erdős-Rényi random graph and expanders. *In Proceedings of the 29th International Symposium on Algorithms and Computation (ISAAC)*, pages 4:1–4:13, 2018.

[Zuc07] D. Zuckerman. Linear Degree Extractors and the Inapproximability of Max Clique and Chromatic Number. *Theory of Computing*, 3(6):103–128, 2007.