Random Walk-Based Algorithms on Networks

Dissertation

zur Erlangung des Grades des
Doktors der Ingenieurwissenschaften
der Naturwissenschaftlich-Technischen Fakultäten
der Universität des Saarlandes

vorgelegt von

Ali Pourmiri

Saarbrücken
2015

Tag des Kolloquiums:
03. Juli. 2015

Dekan der Naturwissenschaftlich-Technischen Fakultät I:
Prof. Dr. Markus Bläser

Prüfungsausschuss:

Prof. Dr. Raimund Seidel (Vorsitzender des Prüfungsausschusses)

Prof. Dr. Dr. h.c. mult. Kurt Mehlhorn (Gutachter)

Dr. Thomas Sauerwald (Gutachter)

Dr. Antonios Antoniadis (Akademischer Beisitzer)

Abstract

The present thesis studies some important random walk-based algorithms, which are randomized rumor spreading and balanced allocation protocols on networks. In the first part of the thesis, we study the Push and the Push-Pull protocols introduced by [DGG+87], which are basic randomized protocols for information dissemination on networks. In Chapter 2, we propose a new model where the number of calls of each node in every round is chosen independently according to a probability distribution R with bounded mean determined at the beginning of the process. In addition to the model being a natural extension of the standard protocols, it also serves as a more realistic model for rumor spreading in a network whose entities are not completely uniform and may have different levels of power. We provide both lower and upper bounds on the rumor spreading time depending on statistical properties of R such as the mean or the variance. While it is well-known that the standard protocols need $\Theta(\log n)$ rounds to spread a rumor on a complete network with n nodes, we show that, if R follows a power law distribution with exponent $\beta \in (2,3)$, then the Push-Pull protocol spreads a rumor in $\Theta(\log\log n)$ rounds. Moreover, when $\beta = 3$, we show a runtime of $\Theta\left(\frac{\log n}{\log\log n}\right)$. In Chapter 3, we analyze the behavior of the standard Push-Pull protocol on a class of random graphs, called random k-trees for every integer $k \geqslant 2$, that are suitable to model poorly connected, small-world and scale free networks. Here, we show that the Push-Pull protocol propagates a rumor from a randomly chosen informed node to almost all nodes of a random k-tree with n nodes in $\mathcal{O}((\log n)^{1+c_k})$ rounds with high probability, where $0 < c_k \leqslant 1$ is a decreasing function in k. We also derive a lower bound of $n^{\Omega(1)}$ for the runtime of the protocol to inform all nodes of the graph. Our technique for proving the upper bound is successfully carried over to a closely related class of random graphs called random k-Apollonian networks.

We devote the rest of the thesis to the study of random walks on graphs, covering both practical and theoretical aspects. In Chapter 4, we show the existence of a *cutoff* phenomenon for simple random walks on Kneser graphs. A cutoff phenomenon for a given sequence of ergodic Markov chains describes a sharp transition in the convergence of the chains to its stationary distribution over a negligible period of time, known as the *cutoff window*. In order to establish the cutoff phenomenon, we combine the spectral information of the transition matrix and a probabilistic technique, known as Wilson's method [Wil04]. And finally in Chapter 5, by using *non-backtracking* random walks introduced by Alon et al. [ABLS07], we propose a new algorithm for sequentially allocating n balls into n bins that are organized as a d-regular graph with n nodes, say G, where $d \geqslant 3$ can be any integer. Let l be a given positive integer. In each round t, $1 \leqslant t \leqslant n$, ball t picks a node of G uniformly at random and performs a non-backtracking random walk of length l from the chosen node. Then it deterministically selects a subset of the visited nodes as the potential choices and allocates itself on one of the choices with minimum load (ties are broken uniformly at random). Provided G has a sufficiently large girth, we establish an upper bound for the maximum number of balls at any bin after allocating n balls by the algorithm. We also show that the upper bound is tight up to a $\mathcal{O}(\log\log n)$ factor. In particular, we show that if we set $l = \lfloor(\log n)^{\frac{1+\epsilon}{2}}\rfloor$, for any constant $\epsilon \in (0,1]$, and G has girth at least $\omega(l)$, then the maximum load is bounded by $\mathcal{O}(1/\epsilon)$ with high probability.

Zusammenfassung

Die vorliegende Arbeit untersucht einige wichtige Zufallspfad-basierte Algorithmen, insbesondere Protokolle zur randomisierte Verbreitung von Gerüchten und Zufallspfade in Netzwerken. Im ersten Teil der Arbeit betrachten wir die von [DGG+87] eingeführten Push und Push-Pull Protokolle, die grundlegende randomisierte Protokolle zur Informationsverbreitung in Netzwerken darstellen. In Kapitel 2 beschreiben wir ein neues Modell, in dem die Anzahl an Aufrufen jedes Knotens in jeder Runde unabhängig von einer Zufallsverteilung R mit beschränktem Erwartungswert gezogen wird, die zu Beginn des Prozesses festgelegt wird. Das Modell ist nicht nur eine natürliche Erweiterung der Standardprotokolle, sondern dient auch als realistischeres Modell der Verbreitung von Gerüchten in Netzwerken deren Entitäten nicht uniform sind und unterschiedlich großen Einfluss haben können. Wir geben untere und obere Schranken für die benötigte Zeit zur Verbreitung der Gerüchte an, in Abhängigkeit von statistischen Eigenschaften von R wie Erwartungswert und Varianz. Während bekannt ist, dass die Standardprotokolle $\Theta(\log n)$ Runden benötigen, um ein Gerücht in einem vollständigen Netzwerk mit n Knoten zu verbreiten, zeigen wir, dass das Push-Pull-Protokoll ein Gerücht in $\Theta(\log\log n)$ Runden verbreitet, wenn R einer Potenzgesetz-Verteilung mit Exponent $\beta \in (2,3)$ folgt. Darüberhinaus zeigen wir, im Falle $\beta = 3$, eine Laufzeit von $\Theta\left(\frac{\log n}{\log\log n}\right)$. In Kapitel 3 analysieren wir das Verhalten des Standard-Push-Pull-Protokolls auf einer Klasse von Zufallsgraphen, den sogenannten Zufalls-k-Bäumen für jede natürliche Zahl $k \geqslant 2$, die sich dafür eignen, schwach zusammenhängende Netzwerke, Small-World-Netzwerke und skalenfreie Netzwerke zu modellieren. Hierbei zeigen wir, dass das Push-Pull-Protokoll ein Gerücht von einem zufällig gewählten informierten Knoten zu fast allen Knoten eines Zufalls-k-Baums mit n Knoten in $O\left((\log n)^{1+c_k}\right)$ Runden mit hoher Wahrscheinlichkeit verbreiten kann, wobei $0 < c_k \leqslant 1$ eine fallende Funktion in k ist. Wir leiten auch eine untere Schranke von $n^{\Omega(1)}$ für die Laufzeit des Protokolls ab, um alle Knoten des Graphen zu informieren. Unsere Technik zum Beweis der oberen Schranke wird erfolgreich auf eine eng verwandte Klasse von Zufallsgraphen, der sogenannten k-Apollonischen Graphen, übertragen.

Den Rest der Dissertation widmen wir der Untersuchung sowohl praktischer als auch theoretischer Aspekte von Zufallspfaden in Graphen. In Kapitel 3 zeigen wir die Existenz eines Cutoff-Phänomens für einfache Zufallspfade in Kneser-Graphen. Ein Cutoff-Phänomen für eine gegebene Sequenz von ergodischen Markovketten beschreibt einen abrupten Übergang bei der Konvergenz der Ketten gegen ihre stationäre Verteilung über einen vernachlässigbaren Zeitraum, bekannt als *Cutoff-Fenster*. Um das Cutoff-Phänomen nachzuweisen kombinieren wir die spektrale Information der Transitionsmatrix und eine probabilistische Technik, bekannt als Wilson's Methode [Wil04]. Und schließlich präsentieren wir in Kapitel 5 unter Einbeziehung von nicht-zurücksetzenden Zufallspfaden, eingeführt von Alon et al. [ABLS07], einen neuen Algorithmus um sequenziell n Bälle n Körben zuzuweisen, die als d-regulärer Graph G mit n Knoten organisiert sind, wobei $d \geqslant 3$ eine beliebige ganze Zahl sein kann. Sei l eine gegebene positive ganze Zahl. In jeder Runde t, $1 \leqslant t \leqslant n$, wählt Ball t einen Knoten von G zufällig mit gleicher Wahrscheinlichkeit und folgt einem nicht-

zurücksetzenden Zufallspfad der Länge l ab diesem gewählten Knoten. Dann wählt der Ball deterministisch eine Teilmenge der besuchten Knoten als potenzielle Kandidaten aus, und weist sich selbst demjenigen Kandidaten mit minimaler Last zu (Gleichstände werden beliebig gelöst). Wenn G hinreichend große Taillenweite hat, können wir eine obere Schranke für die maximale Anzahl an Bällen in jedem Bin nach der Zuweisung von n Bällen durch den Algorithmus angeben. Wir zeigen auch, dass diese obere Schranke bis auf einen $O(\log\log n)$-Faktor scharf ist. Insbesondere zeigen wir, dass die maximale Last mit hoher Wahrscheinlichkeit durch $O(1/\epsilon)$ beschänkt ist, wenn wir $l = \left\lfloor (\log n)^{\frac{1+\epsilon}{2}} \right\rfloor$ setzen, für eine beliebige Konstante $\epsilon \in (0, 1]$, und G Taillenweite mindestens $\omega(l)$ hat. Diese Arbeit ist in englischer Sprache verfasst.

Acknowledgments

I would like to thank my supervisor Thomas Sauerwald for his invaluable guidance and encouragement. I appreciate his patience in correcting several of my drafts and making them into publishable material. I would like to sincerely thank Kurt Mehlhorn for his constant support and reviewing the thesis. I am grateful to thank my co-authors Thomas Sauerwald, Konstantinos Panagiotou and Abbas Mehrabian.

I thank my fellows Megha Khosla, Aruni Choudhary, Sebastian Ott, Pavel Kolev, Michael Dirnberger, Cosmina Croitoru and Bojana Kodric for happy time in MPI.

I would like to deeply thank my father whose memories never fade, my mother and my siblings.

I would like to thank my loving wife Fahimeh and my little son, Mohammad Mahdi, who are source of inspiration and motivation.

Contents

1 Introduction

Randomized algorithms and probabilistic models are playing an increasingly important role in computer science. They are applied in a wide range of problems such as deadlock avoidance, message routing, diffusion networks, distributed consensus, hashing, load balancing, polynomial identity testing, randomized rounding, and primality testing. Although randomized algorithms may not guarantee a correct (and/or optimal) output for an adversarial chosen input, their elegance and efficiency on average input cause them to be essentially in computer science. Furthermore, they often represent the first step towards the design of efficient deterministic algorithms.

Information dissemination on networks is a fundamental and overarching problem in many settings such as broadcasting, load balancing, gossip, sorting, leader election protocols etc. (e.g., see [HKP+05]). In the first part of this thesis we study some random walk-based protocols for information dissemination, known as randomized rumor spreading, introduced by Demers et. al [DGG+87]. These form an important class of protocols that are particularly attractive because of their simplicity, efficiency and robustness in information dissemination on networks [FPRU90]. Not only they exhibit high performance in terms of runtime and communication overhead, they also serve as mathematical models for many phenomena such as spreading computer viruses or diffusion of ideas in real-world and technological networks.

In the second part of the thesis we study another fundamental class of stochastic processes that are random walks on networks, which have been widely applied in the design and analysis of randomized algorithms. One of the first applications of random walks on networks has been to efficiently sample from a stationary distribution, which is the main ingredient of many algorithms such as computing volume of convex bodies approximately [LS93], solving linear programming [KN12], minimizing discrepancy [Ban10] and property testing algorithms [CMOS11]. Furthermore, random walks are applied as a building block in a wide variety of problems ranging from token management [IJ90], small-world routing [Kle00], information propagation and gathering [KKD04], testing expansion [GT12], load balancing and averaging [FGS12].

1.1. Rumor Spreading

The first part of this thesis concerns randomized rumor spreading protocols. One of the most basic randomized rumor spreading protocols is the Push protocol. In this protocol, initially, a node of the network knows of some rumor and then the protocol proceeds in rounds. In each round, every informed node sends (pushes) the rumor to a random neighbor. Similarly, in each round of the Pull protocol, an uninformed node asks (pulls) a random neighbor in order to get the rumor if the neighbor knows it. The Push-Pull protocol is the combination of these two protocols where each node contacts a random neighbor. The Push-Pull protocol is also known as the random phone call model introduced by Demers et al. [DGG$^+$87], which is a well-studied model for information dissemination on networks. One of the most important quantities associated with rumor spreading protocols is the *runtime* of the protocol which is defined as the number of rounds the protocol needs to inform all nodes of a network with probability tending to 1 as n goes to infinity. We refer to this as 'with high probability'.

1.1.1. Multiple-Call Rumor Spreading

Frieze and Grimmett [FG85] showed that the standard Push protocol on a complete graph takes $\log_2 n + \log n \pm o(\log n)$ rounds to inform all n nodes with high probability. This result was later strengthened by Pittel [Pit87]. For the standard Push-Pull protocol on a complete graph, Karp et al. [KSSV00] proved a runtime bound of $\log_3 n + \mathcal{O}(\log\log n)$. In Chapter 2, we study a multiple-call version of the Push and the Push-Pull protocols on complete graphs, where nodes are enabled to make multiple calls in each round. In addition to the model being a natural extension of the standard protocols, it also serves as a more realistic model for rumor spreading in networks, where entities are not completely uniform and may have different levels of power. Specifically, we assume that the power of each node u, denoted by C_u, is determined by a probability distribution R on the positive integers. Note that C_u is the same number for each round. In order to keep the overall communication cost small, we focus on distributions R satisfying $\sum_{u\in V} C_u = \mathcal{O}(n)$ with high probability – in particular, R has bounded mean.

In this work, our aim is to understand the impact of the distribution R on the runtime of the Push and the Push-Pull protocols. In particular, we seek conditions on the distribution R which are necessary (and/or sufficient) for a sub-logarithmic runtime. We provide both lower and upper bounds on the runtime of the protocols, depending on statistical properties of R such as the mean or the variance. For instance, if R follows a power law distribution with exponent $\beta \in (2, 3)$, we show that the Push-Pull protocol spreads a rumor in $\Theta(\log\log n)$ rounds. Moreover, if $\beta = 3$, the Push-Pull protocol spreads a rumor in $\Theta\left(\frac{\log n}{\log\log n}\right)$ rounds. To prove our results, we carefully derive an almost tight growth rate for the size of informed nodes and then establish tight upper (and lower) bounds for the runtime of protocols.

1.1.2. Rumor Spreading on Real-World Networks

Recently, many studies considered the rumor spreading protocols on random graph models that exhibit some fundamental properties of real-world networks such as power-law degree sequence, small diameter and large clustering coefficient. For instance, Doerr, Fouz, and Friedrich [DFF11] studied the Push-Pull protocol on preferential attachment graphs, which is a popular model for real-world networks. They proved an upper bound of $\mathcal{O}(\log n)$ for the runtime of the protocol. Also, Fountoulakis, Panagiotou, and Sauerwald [FPS12] proved the same upper bound $\mathcal{O}(\log n)$ for the runtime of the Push-Pull protocol on the giant component of random graphs with given expected degrees (also known as the Chung-Lu model) with power law degree distribution. They also showed that if degree distribution follows a power law distribution with exponent $\beta \in (2, 3)$, then the Push-Pull protocol takes $\mathcal{O}(\log \log n)$ rounds to inform a constant fraction of nodes with constant probability.

In Chapter 3, we analyze the behavior of the Push-Pull protocol on random k-trees, a class of power law graphs which are scale-free, small-world (i.e., logarithmic diameter in terms of the number of nodes) and have large clustering coefficients. The random k-trees are a class of evolving random graphs built as follows: Initially we have a k-clique. In every step a new node is created, a random k-clique of the current graph is chosen, and the new node is joined to all nodes of the k-clique. For a given random k-tree on n nodes with fixed $k \geqslant 2$, we show that if a random node is initially aware of the rumor, then with probability $1 - o(1)$ after $\mathcal{O}\left((\log n)^{1+\frac{2}{k}} \cdot (\log \log n)^2\right)$ rounds the rumor is propagated to $n - o(n)$ nodes. Since these graphs have polynomially small conductance, vertex expansion $\mathcal{O}(1/n)$, and constant treewidth, these results demonstrate that Push-Pull can be efficient even on poorly connected networks. On the negative side, we prove that with probability $1 - o(1)$ the protocol needs at least $n^{\Omega(1)}$ rounds to inform all nodes. This exponential dichotomy between the time required for informing *almost all* and *all* nodes is striking. Our main contribution is to present, for the first time, a natural class of random graphs in which such a phenomenon can be observed. Our technique for proving the upper bound successfully carries over to a closely related class of graphs, the so-called random k-Apollonian networks, for which we prove an upper bound of $\mathcal{O}\left((\log n)^{1+a_k} \cdot (\log \log n)^2\right)$ rounds for informing $n - o(n)$ nodes with probability $1 - o(1)$, when $k \geqslant 2$ is a constant and $0 < a_k < 1$ is a decreasing function in k.

To prove the upper bounds, a well-known technique is to show the existence of low degree nodes, called *efficient connectors*, which connect different high-degree nodes and speed up the transmission of the rumor among them. By showing the existence of efficient connectors for almost all high degree nodes in random k-trees, we derive our upper bound to inform almost all nodes of a random k-tree.

Source of indication. The results of Chapters 2 and 3 have been published in the Electronic Journal of Combinatorics [PPS15] and in the proceedings of DISC'14 [MP14], respectively. A preliminary version of results of Chapter 2 has been published in the proceedings of ISAAC'13 [PPS13].

1.2. Random Walks on Graphs

The second part of this thesis studies random walks and their applications. Random walks on graphs are key components in many fields of sciences from diffusion processes in statistical physics to enumeration of combinatorial objects through sampling in computer science. Here, we consider two variants of random walks on graphs: *simple* random walks and *non-backtracking* random walks. A simple random walk on a given graph is a stochastic process that proceeds in rounds. Initially, the walker is located on a node of the graph, and then in each round, the walker chooses one of its neighbors uniformly at random and moves to that neighbor. The non-backtracking random walk is similar to a simple random walk with only one difference: the walker never traverses the same edge in two consecutive rounds, that is, the walker never 'backtracks'.

1.2.1. Cutoff Phenomenon for Random walks

A *cutoff* phenomenon for a given sequence of ergodic Markov chains describes a sharp transition in the convergence of the chains to its stationary distribution over a negligible period of time, known as the cutoff window. Due to the elusive behavior of many Markov chains, showing the existence of cutoff phenomenon for a sequence of chains is a challenging question and there is still no necessary and sufficient condition known [Dia96]. In one of the first results in this area, Aldous and Diaconis [AD86] studied a card-shuffling process called *top in at random shuffle* where in each step, a card is picked from the top and inserted in a random position. They proved that for a deck of cards of size n, the chain shows a cutoff at time $n \log n + \mathcal{O}(n)$ over a cutoff window of size $\mathcal{O}(n)$. Recently, Lubetzky and Sly [LS10] considered simple random walks and non-backtracking walks on n-vertex random d-regular graphs, $G(n, d)$, for $d \geqslant 3$ and established a cutoff for these chains. More precisely, they showed that simple random walks on $G(n, d)$ exhibit a cutoff at time $\frac{d}{d-2} \log_{d-1} n$ with a window of size $\mathcal{O}(\sqrt{\log_{d-1} n})$ with high probability. They also derived a cutoff for non-backtracking walks at time $\log_{d-1} n$ with a window of size constant. To prove their result, they elegantly exploited tree-like structures of random regular graphs and estimated the walk distribution.

In Chapter 4, we focus on simple random walks on Kneser graphs and show that they exhibit a cutoff. Given two integers n and k, the Kneser graph $K(2n + k, n)$ is defined as the graph with the vertex set being all subsets of $\{1, \ldots, 2n + k\}$ of size n and two vertices A and B being connected by an edge if $A \cap B = \emptyset$. It is also well-known that the transition matrix of a simple random walk on a Kneser graph $K(2n + k, n)$ has spectral gap $\frac{k}{n+k}$, and its second largest eigenvalue has multiplicity $2n + k$. So by varying $k = \mathcal{O}(n)$, we obtain various families of chains with different spectral gaps. For instance, by setting $k = \Theta(n)$ we obtain a family of transitive expander graphs. In our work, we show that for any $k = O(n)$, the simple random walk on $K(2n+k, n)$ exhibit a cutoff at $\frac{1}{2} \log_{1+k/n} (2n + k)$ with a window of size $O(\frac{n}{k})$. In the case $k = \omega(n)$, every node in $K(2n + k, n)$ is adjacent to almost all nodes, and the simple random walk on $K(2n + k, n)$ converges to its stationary distribution in only one step. For the case $k = \mathcal{O}(n)$, it is necessary to have detailed knowledge of the total variation distance of the walk distribution from its equilibrium at any time step t denoted by $d(t)$. Here, we combine the spectral information of the transition matrix and a probabilistic technique, known as Wilson's method [Wil04] to obtain a precise estimate of $d(t)$.

1.2.2. Balanced Allocation on Graphs

The standard balls-into-bins model is a process that randomly allocates m balls into n bins; each ball picks d bins independently and uniformly at random and the ball is then allocated in a least loaded bin in the set of d choices. In many applications, selecting any random set of choices is costly. For instance, assume the bins are processors that are interconnected as a graph and balls are tasks arriving one by one; the goal is to assign tasks to the processor by minimizing the maximum load in a distributed fashion. Here, having two far away choices is not desirable. Considering this constraint, Kenthapadi and Panigrahy [KP06] proposed a model in which bins are interconnected as a Δ-regular graph where each ball picks a random edge of the graph. It is then placed at one of its endpoints with smaller load. This allocation algorithm results in a maximum load of $\log\log n+\mathcal{O}\left(\frac{\log n}{\log(\Delta/\log^4 n)}\right)+\mathcal{O}(1)$. Following the study of balls-into-bins with correlated choices, Godfrey [God08] generalized the aforementioned result such that each ball picks a random edge of a hypergraph that has $\Omega(\log n)$ bins and satisfies some mild conditions. Recently, Bogdan et al. [BSSS13] studied a model where bins are nodes of a graph and each ball picks a random node and performs a local search from the node to find a node with local minimum load and finally be placed on it. They showed that when the graph is a constant degree expander, the local search guarantees a maximum load of $\Theta(\log\log n)$ with high probability.

In Chapter 5, we propose a new algorithm that uses non-backtracking random walks as a tool for sequentially allocating n balls into n bins that are organized as a d-regular n-vertex graph G, where $d \geqslant 3$ can be any integer. The algorithm takes a positive integer l, G and a sequence of balls as the inputs and proceeds in rounds. In each round t, $1 \leqslant t \leqslant n$, ball t picks a node of G uniformly at random and performs a non-backtracking random walks (NBRW) of length l from the chosen node. Then, it deterministically selects a subset of visited nodes as the potential choices and allocates itself on one of the choices with minimum load (here, ties are broken uniformly at random). Suppose that G has girth at least $\omega(l\log\log n)$. Then we establish an upper bound for the maximum number of balls at any bin after allocating n balls by the algorithm, called *maximum load*, in terms of l with high probability. We also show that the upper bound is at most an $\mathcal{O}(\log\log n)$ factor above the lower bound that is proved for the algorithm. In particular we show that if we set $l = \lfloor(\log n)^{\frac{1+\epsilon}{2}}\rfloor$, for any constant $\epsilon \in (0,1]$ and G has girth at least $\omega(l)$, then the maximum load is bounded by $\mathcal{O}(1/\epsilon)$ with high probability.

To show our result, we apply the *witness tree* technique, which is a well-known method in the balls-into-bins process.

Source of indication. The results of Chapter 4 have been published in the Journal of Discrete Applied Mathematics [PS14] and the results of Chapter 5 are under submission [Pou14].

1.3. Probabilistic Inequalities

In this section, we review some essential tools for the analysis of randomized algorithms, namely deviation bounds. Let us first state some elementary probabilistic inequalities that can be found in [GS01].

Lemma 1.3.1 (Union Bound). *Let $\mathcal{E}_i$, $1 \leqslant i \leqslant n$ be a sequence of events. Then,*

$$\mathbf{Pr}\left[\vee_{i=1}^{n}\mathcal{E}_i\right] \leqslant \sum_{i=1}^{n}\mathbf{Pr}\left[\mathcal{E}_i\right].$$

Lemma 1.3.2 (Markov's Inequality). *Let X be a non-negative random variable and δ be any positive real number. Then, we have*

$$\mathbf{Pr}\left[X \geqslant \delta\right] \leqslant \mathbf{E}\left[X\right]/\delta.$$

Theorem 1.3.3 (Chebychev's Inequality). *Let X be a random variable and δ be any positive real number. Then, we have*

$$\mathbf{Pr}\left[|X - \mathbf{E}\left[X\right]| \geqslant \delta\right] \leqslant \mathbf{Var}\left[X\right]/\delta^2.$$

Now we state two concentration inequalities that we use in the proof of our results. The first one is a Chernoff-type bound. For the proof, see, e.g., [DP09].

Theorem 1.3.4 (Chernoff bounds). *Suppose that $X_1, X_2 \ldots, X_n \in \{0,1\}$ are independent random variables, and let $X := \sum_{i=1}^{n} X_i$. Then, for any $\delta \in (0,1)$, the following inequalities hold:*

$$\mathbf{Pr}\left[X \leqslant (1-\delta)\,\mathbf{E}\left[X\right]\right] \leqslant \mathrm{e}^{-\delta^2\,\mathbf{E}[X]/2},$$
$$\mathbf{Pr}\left[X \leqslant (1+\delta)\,\mathbf{E}\left[X\right]\right] \leqslant \mathrm{e}^{-\delta^2\,\mathbf{E}[X]/3}.$$

In particular,

$$\mathbf{Pr}\left[\left|X - \mathbf{E}\left[X\right]\right| \geqslant \delta\mathbf{E}\left[X\right]\right] \leqslant 2\cdot \mathrm{e}^{-\delta^2\,\mathbf{E}[X]/3}.$$

Panconesi and Srinivasan [PS97] generalized the classic Chernoff bounds for random variables that are negatively correlated. A negative correlation among a set of random variables is defined as follows:

Definition 1.3.5. *The random variables $X_1, \ldots, X_n$, taking values in Ω, are called negatively correlated if for all $I \subseteq [n]$ and every $\omega_i \in \Omega$, $i \in I$, we have*

$$\mathbf{Pr}\left[\wedge_{i\in I} X_i = \omega_i\right] \leqslant \prod_{i\in I}\mathbf{Pr}\left[X_i = \omega_i\right].$$

Theorem 1.3.6 (Chernoff bounds for negatively correlated random variables). *Suppose that $X_1, X_2 \ldots, X_n \in \{0,1\}$ are negatively random variables and let $X := \sum_{i=1}^{n} X_i$. Then, for any $\delta \in (0,1)$, the following inequalities hold:*

$$\mathbf{Pr}\left[X \leqslant (1-\delta)\,\mathbf{E}\left[X\right]\right] \leqslant \mathrm{e}^{-\delta^2\,\mathbf{E}[X]/2},$$
$$\mathbf{Pr}\left[X \geqslant (1+\delta)\,\mathbf{E}\left[X\right]\right] \leqslant \mathrm{e}^{-\delta^2\,\mathbf{E}[X]/3}.$$

In order to deal with moderate independency we have the following lemma whose proof can be found in [AD11, Lemma 1.18].

Lemma 1.3.7 (Deviation bounds for moderate independency)**.** *Let $X_1, \cdots, X_n$ be arbitrary binary random variables. Let $X_1^*, X_2^*, \cdots, X_n^*$ be binary random variables that are mutually independent and such that for all i, X_i, is independent of $X_1, \cdots, X_{i-1}$. Assume that for all i and all $x_1, ..., x_{i-1} \in \{0,1\}$,*

$$\mathbf{Pr}\left[X_i = 1 | X_1 = x_1, \cdots, X_{i-1} = x_{i-1}\right] \geqslant \mathbf{Pr}\left[X_i^* = 1\right].$$

Then for all $k \geqslant 0$, we have

$$\mathbf{Pr}\left[\sum_{i=1}^{n} X_i \leqslant k\right] \leqslant \mathbf{Pr}\left[\sum_{i=1}^{n} X_i^* \leqslant k\right]$$

and the latter term can be bounded by any deviation bound for independent random variables.

The next inequality is known as the Bounded Difference inequality. For the proof see, e.g., [McD98].

Theorem 1.3.8 (Bounded Difference Inequlaity)**.** *Suppose that $X_1, X_2 \ldots, X_n$ are arbitrary independent random variables, and every X_i, $1 \leqslant i \leqslant n$ takes a value from A_i. Let $f : \prod_{1 \leqslant i \leqslant n} A_i \to \mathbb{R}$ be a real-valued function so that there exist $c_1, c_2, \ldots, c_n$ with*

$$\sup_{x_1, x_2, \ldots, x_n, x_i'} |f(x_1, x_2, \ldots, x_i, \ldots, x_n) - f(x_1, x_2, \ldots, x_i', \ldots, x_n)| \leqslant c_i, \text{ for every } 1 \leqslant i \leqslant n.$$

Then, for every $\lambda > 0$,

$$\mathbf{Pr}\left[|f(X_1, X_2, \ldots, X_n) - \mathbf{E}\left[f(X_1, X_2, \ldots, X_n)\right]| \geqslant \lambda\right] \leqslant 2 \cdot \mathrm{e}^{-\frac{\lambda^2}{2\sum_{i=1}^{n} c_i^2}}.$$

Remark. Throughout this thesis, with high probability refers to an event $\mathcal{A}_n$ which holds with probability $1 - o(1)$ as $n \to \infty$. For simplicity, we sometimes abbreviate 'with high probability' to 'whp'. Moreover, $\log n$ denotes the natural logarithm of n.

Part I

Randomized Rumor Spreading

2 Faster Rumor Spreading with Multiple Calls

Randomized rumor spreading is an important primitive for spreading information in networks. The goal is to spread a piece of information, the so-called rumor, from an arbitrary node to all the other nodes. Randomized rumor spreading protocols are based on the simple idea that every node picks a random neighbor and these two nodes are able to exchange information in that round. This paradigm ensures that the protocol is local, scalable and robust against network failures (cf. [FPRU90]). Therefore these protocols have been successfully applied in other contexts such as replicated databases [DGG$^+$87], failure detection [vRMH98], resource discovery [HBLL99], load balancing [BGPS06], data aggregation [KDG03], and analysis of the spread of computer viruses [BBCS05].

The most basic variant of randomized rumor spreading is the Push protocol. At the beginning, there is a single node who knows of some rumor. Then in each of the following rounds every *informed* node calls a random neighbor chosen independently and uniformly at random and informs it of the rumor. The Pull protocol is symmetric, here every *uninformed* node calls a random neighbor chosen independently and uniformly at random, and if that neighbor happens to be informed the node becomes informed. The Push-Pull protocol is simply the combination of both protocols. Most studies in randomized rumor spreading concern the communication overhead produced by these protocols and their runtimes which is the number of rounds required until a rumor initiated by a single node reaches all other nodes of the network (see e.g. [KSSV00]).

In one of the first papers in this area, Frieze and Grimmett [FG85] proved that if the underlying graph is a complete graph with n nodes, then the runtime of the Push protocol is $\log_2 n + \log n \pm o(\log n)$ with high probability. This result was later strengthened by Pittel [Pit87]. For the standard Push-Pull protocol, Karp et al. [KSSV00] proved a runtime bound of $\log_3 n + \mathcal{O}(\log\log n)$. In order to overcome the large number of $\Theta(n \log n)$ calls, Karp et al. also presented an extension of the Push-Pull protocol together with a termination mechanism that spreads a rumor in $\mathcal{O}(\log n)$ rounds using only $\mathcal{O}(n \log\log n)$ messages. More recently Doerr and Fouz [DF11] proposed a new protocol using only Push calls that achieves a runtime of $(1+o(1))\log_2 n$ using only $O(n \cdot f(n))$ calls (and messages), where $f(n)$ is an arbitrarily slow growing function.

Our Results. We study a multiple-calls version of the Push protocol and the Push-Pull protocol on complete networks, where nodes are enabled to make multiple calls in each round. While it is well-known that the classic Push and Push-Pull protocols need $\Theta(\log n)$ rounds to spread a rumor on a complete network with n nodes, we are interested by how much we can speed up the spread of the rumor by enabling nodes to make more than one call in each round. Besides the fact that the model is a natural extension of the standard protocols, it also serves as a more realistic model for rumor spreading in a network whose entities are not completely uniform and may have different levels of power. Specifically, we assume that the power of each node u, denoted by C_u is determined by a probability distribution R on the positive integers which is independent of u and C_u is the same number for each round. In order to keep the overall communication cost small, we focus on distribution R satisfying $\sum_{u\in V} C_u = \mathcal{O}(n)$ with high probability – in particular, R has bounded mean. Our aim is to understand the impact of the distribution R on the runtime of Push and Push-Pull protocols. In particular, we seek for conditions on the distribution R which are necessary (and/or sufficient) for a sub-logarithmic runtime. We estimate the runtime of such multiple-calls protocols on a n-node complete network with a general assumption about mean and variance of R and summarize it as follows: If R has bounded mean and bounded variance, then whp the Push protocol needs $\log_{1+\mathbf{E}[R]} n + \log_{\mathrm{e}^{\mathbf{E}[R]}} n \pm o(\log n)$ rounds to inform all nodes (cf. Theorem 2.4.2). Moreover whp the Push-Pull protocol requires $\Theta(\log n)$ rounds to inform all nodes (Theorem 2.5.2). If R has bounded mean and arbitrary variance, then whp after $\Theta(\log n)$ rounds of the Push protocol, every node gets informed (cf. Theorem 2.4.1).

Note that by putting $R \equiv 1$, we retain the classic result by Frieze and Grimmett for the standard Push protocol. As can be seen, when we assume that R has an unbounded variance, the Push protocol still needs $\Theta(\log n)$ rounds to inform all nodes. Although this result is less precise than Theorem 2.4.2, it demonstrates that it is necessary to consider the Push-Pull protocol with an unbounded variance. An important distribution with bounded mean but unbounded variance is the *power law distribution* with exponent $\beta \leqslant 3$, i.e., there are constants $0 < c_1 \leqslant c_2$ such that $c_1 z^{1-\beta} \leqslant \mathbf{Pr}[C_u \geqslant z] \leqslant c_2 z^{1-\beta}$ for any $z \geqslant 1$, and $\mathbf{Pr}[C_u \geqslant 1] = 1$. We are especially interested in power law distributions, because they are scale invariant and have been observed in a variety of settings in real life. Our main result shows that this natural distribution achieves a sublogarithmic runtime. Notice that if R is a power law distribution with $\beta > 3$, then Theorem 2.5.2 applies because the variance of R is bounded. Hence our results reveal an interesting dichotomy in terms of the exponent β: if $2 < \beta < 3$, then the Push-Pull protocol finishes in $\mathcal{O}(\log\log n)$ rounds, whereas for $\beta > 3$ the Push-Pull protocol finishes in $\Theta(\log n)$ rounds [1]. Moreover if $\beta = 3$ we

[1]We do not consider the case $\beta \leqslant 2$, since then there exists at least one node with degree $\Omega(n)$ and the rumor is spread in constant time. Additionally, $\mathbf{E}[R]$ is no longer bounded.

show that whp the Push-Pull protocol informs all nodes in $\Theta\left(\frac{\log n}{\log\log n}\right)$ rounds.

Finally, we also show that it is necessary that the C_u's are independent of the round t. Instead, suppose we generate a new variable C_u^t according to the distribution R for each round t again, which is the number of calls made by node u in round t. Then we prove that in this model, with high probability, Push-Pull needs $\Omega(logn)$ rounds to inform all nodes.

Techniques. To estimate the runtime of the protocols, we carefully analyze the growth rate of size of informed nodes in several phases depending on the number of informed nodes. For instance, to derive a $\Theta(\log n)$ runtime for a protocol, it is sufficient to show an constant growth rate for the protocol. The technique gets more involved when R is a power low distribution with $2 < \beta \leqslant 3$ and we establish a dichotomy in terms of β. While a very similar dichotomy was shown in [FPS12] for random graphs with a power law degree distribution, our result here concerns the spread of the rumor from one to *all* nodes (and not only to a constant fraction as in [FPS12]). In addition, the distribution of the edges used throughout the execution of the Push-Pull protocol is different from the distribution of the edges in a power law random graph, as the latter is proportional to the product of the weights of the two nodes. Therefore it seems difficult to apply the previous techniques for power law random graphs used for the analysis of the average distance [CL03] and rumor spreading [FPS12].

Outline. In Section 2.1 we formally define the multiple-calls protocols and notations needed to show our results. In Section 2.2 we give some useful facts about power law distributions. In Sections 2.3 and 2.4 we analyze the runtime of the Push protocol on a complete network. In Section 2.5 we continue studying the Push-Pull protocol with a bounded mean and variance distribution R. In Sections 2.6 and 2.7 we analyze the multiple call Push-Pull protocol with power law distribution R where $2 < \beta \leqslant 3$. In the last section we consider a different model where in every round t each node u generates a new C_u^t according to a distribution R.

2.1. Definitions, Notations and Preliminaries

In this section we provide additional definitions and notations. Let us first generalize the classic Push, Pull and Push-Pull to the following statistical model on a complete graph with n nodes. Before the protocol starts, every node u generates a random integer $C_u \geqslant 1$ independent of each other according to a distribution R. Then, a piece of information (rumor, message, ...) is placed on an arbitrary node of the graph. Our generalized Push, Pull and Push-Pull protocol proceed like the classic ones except that every (un)informed node u calls C_u node(s) chosen independently and uniformly at random and sends (request) the rumor.

Let $\mathcal{I}_t$ be the set of all informed nodes in round t (which means after the execution of round t) and $\mathcal{U}_t$ be the complement of $\mathcal{I}_t$, i.e., the set of uninformed nodes. The size of $\mathcal{I}_t$ and $\mathcal{U}_t$ is denoted by I_t and U_t. We indicate the set of newly informed nodes in round $t+1$ by $\mathcal{N}_t$ and its size is denoted by N_t. Let S_t be the number of Push calls in round $t+1$, so $S_t = \sum_{u\in\mathcal{I}_t} C_u \geqslant I_t$. Let us define $\mathcal{N}_t^{\mathsf{Pull}}$ and $\mathcal{N}_t^{\mathsf{Push}}$ to be the set of newly informed nodes by Pull and Push calls in round $t+1$, respectively. The size of $\mathcal{N}_t^{\mathsf{Pull}}$ and $\mathcal{N}_t^{\mathsf{Push}}$ are denoted by N_t^{Pull} and N_t^{Push}. The size of every set divided by n will be denoted by the corresponding small letter, so i_t, n_t and s_t are used to denote I_t/n, N_t/n, and S_t/n, respectively. Further, we define the set

$$\mathcal{L}(z) := \{u \in \mathcal{V} : C_u \geqslant z\}.$$

The size of $\mathcal{L}(z)$ is denoted by $L(z)$. We define Δ to be $\max_{u\in\mathcal{V}} C_u$.

2.2. Some Useful Facts of Power Law Distributions

Let R be a power law probability distribution with exponent β, i.e., there are constants $c_1 > 0$ and $c_2 > 0$ so that for every integer $z \geqslant 1$,

$$c_1 \cdot z^{1-\beta} \leqslant \mathbf{Pr}\left[C_u \geqslant z\right] \leqslant c_2 \cdot z^{1-\beta},$$

and $\mathbf{Pr}\left[C_u \geqslant 1\right] = 1$.

Fact 2.2.1. *If R is a power law distribution with $\beta > 3$, then $\mathbf{Var}\left[R\right] = \mathcal{O}(1)$.*

Proof. Clearly,

$$\mathbf{Var}\left[R\right] \leqslant \mathbf{E}\left[R^2\right] = \sum_{z=1}^{\infty} \mathbf{Pr}\left[R^2 \geqslant z\right] \leqslant 1 + \sum_{z=2}^{\infty} \sqrt{c_2 \cdot z^{1-\beta}} < \infty,$$

since $\beta > 3$. □

Fact 2.2.2. *Let $\{C_u : u \in \mathcal{V}\}$ be a set of n random variables and assume that each C_u is generated according to a power law distribution with exponent $\beta > 2$. Then with probability $1 - o(\frac{1}{\log n})$,*

$$\Delta := \max_{u\in\mathcal{V}} C_u \leqslant n^{\frac{1}{\beta-1}} \cdot \log n.$$

Proof. By definition,

$$\mathbf{Pr}\left[C_u \geqslant n^{\frac{1}{\beta-1}} \log n\right] \leqslant \frac{c_2 \cdot \log^{1-\beta}(n)}{n}.$$

Applying the union bound over the C_u's, $u \in \mathcal{V}$, yields that with probability at least $1 - \frac{c_2}{\log^{\beta-1} n} = 1 - o(\frac{1}{\log n})$,

$$\Delta \leqslant n^{\frac{1}{\beta-1}} \log n.$$

□

Recall that $\mathcal{L}(z) := \{u \in \mathcal{V}, C_u \geqslant z\}$ and $L(z) := |\mathcal{L}(z)|$.

Proposition 2.2.3. *Let $\{C_u \colon u \in \mathcal{V}\}$ be a set of n independent random variables and assume that each C_u is generated according to a power law distribution with exponent $\beta > 2$. Then for every $z = \mathcal{O}(n^{\frac{1}{\beta-1}}/\log n)$, it holds with probability $1 - o(\frac{1}{n})$*

$$\frac{n \cdot c_1 \cdot z^{1-\beta}}{2} \leqslant L(z) \leqslant \frac{3 \cdot n \cdot c_2 \cdot z^{1-\beta}}{2}.$$

Proof. Let us define an indicator random variable I_u for every $u \in \mathcal{V}$ so that

$$I_u := \begin{cases} 1 & \text{if } C_u \geqslant z, \\ 0 & \text{otherwise.} \end{cases}$$

Since the C_u's are independent and identically distributed, so are the I_u's . Further,

$$c_1 \cdot z^{1-\beta} \leqslant \mathbf{Pr}\left[I_u = 1\right] \leqslant c_2 \cdot z^{1-\beta}.$$

We know that $\mathbf{E}\left[L(z)\right] = \mathbf{E}\left[\sum_{u\in\mathcal{V}} I_u\right]$. Hence,

$$n \cdot c_1 \cdot z^{1-\beta} \leqslant \mathbf{E}\left[L(z)\right] \leqslant n \cdot c_2 \cdot z^{1-\beta}.$$

Applying Theorem 1.3.4 to the random variable $X := \sum_{u\in\mathcal{V}} I_u$ yields that

$$\mathbf{Pr}\left[\left|L(z) - \mathbf{E}\left[L(z)\right]\right| > \frac{\mathbf{E}\left[L(z)\right]}{2}\right] < 2 \cdot \mathrm{e}^{-\frac{\mathbf{E}[L(z)]}{10}} \leqslant 2 \cdot \mathrm{e}^{-\frac{n \cdot c_1 \cdot z^{1-\beta}}{10}}$$

Since $z = \mathcal{O}(n^{\frac{1}{\beta-1}} / \log n)$, with probability $1 - o(\frac{1}{n})$

$$\frac{\mathbf{E}\left[L(z)\right]}{2} \leqslant L(z) \leqslant \frac{3 \cdot \mathbf{E}\left[L(z)\right]}{2},$$

and the claim follows. □

2.3. Push Protocol

In this section we will show two general lemmas for the Push protocol with any distribution R. They will be used when analyzing the Push protocol and the Push-Pull protocol.

Lemma 2.3.1. *Consider the* Push *protocol and suppose that $S_t \leqslant \log^c n$, where $c > 0$ is an arbitrary constant. Then with probability at least $1 - \mathcal{O}(\frac{\log^{2c} n}{n})$ we have*

$$I_{t+1} = I_t + S_t.$$

Proof. Recall that S_t is the number of Push calls in round $t+1$. By applying the union bound, the probability that an informed node receives a call in round $t+1$ is bounded by $\frac{S_t I_t}{n}$. So with probability $1 - \frac{S_t I_t}{n}$, none of the calls are sent to a node in $\mathcal{I}_t$. Conditioning on this event, consider all calls one by one in an arbitrary order, the probability that the $i-$th call informs a different node from the previous $i-1$ calls is $1 - \frac{i-1}{U_t}$. Therefore the conditional probability that S_t calls inform S_t different nodes is at least

$$\prod_{i=1}^{S_t - 1}\left(1 - \frac{i}{U_t}\right) > \left(1 - \frac{S_t - 1}{U_t}\right)^{S_t} \geqslant 1 - \frac{S_t^2}{U_t}.$$

So the probability that S_t calls inform S_t different uninformed nodes is at least

$$\left(1 - \frac{S_t I_t}{n}\right) \cdot \left(1 - \frac{S_t^2}{U_t}\right) = 1 - \mathcal{O}\left(\frac{S_t^2}{n}\right),$$

where the above equality holds because $I_t \leqslant S_t \leqslant \log^c n$ and $U_t = n(1 - o(1))$. So with probability at least $1 - \mathcal{O}\left(\frac{\log^{2c} n}{n}\right)$ we have $I_{t+1} = I_t + S_t$ and the claim follows. $\square$

Lemma 2.3.2. *Consider the* Push *protocol. Then with probability at least* $1 - o(\frac{1}{\log n})$ *we have that*

$$s_t - 2s_t^2 - 2\sqrt{\frac{s_t \log\log n}{n}} \leqslant n_t \leqslant s_t.$$

Proof. Since N_t is always bounded by S_t, $n_t \leqslant s_t$. We will prove the lower bound. Let us define Z_v for every $v \in \mathcal{U}_t$ as the indicator random variable with

$$Z_v := \begin{cases} 1 & \text{if } v \in \mathcal{I}_{t+1}, \\ 0 & \text{otherwise.} \end{cases}$$

Then we have $N_t = \sum_{v \in \mathcal{U}_t} Z_v$. Since the Z_v's are identically distributed random variables,

$$\mathbf{E}\left[N_t\right] = U_t \cdot \mathbf{Pr}\left[Z_v = 1\right].$$

Let $X_i \in \mathcal{V}$, $1 \leqslant i \leqslant N = S_t$, denote the target of the i-th call. Define $f(X_1, X_2, ..., X_N) := N_t$ to be the function counting the number of newly informed nodes in round $t+1$. Then $\mathbf{E}\left[f(X_1, X_2, ..., X_N)\right] = \mathbf{E}\left[N_t\right]$. For each change in just one coordinate of f, the following statement holds:

$$\sup_{x_1, x_2, \ldots, x_i, x_i' \in \mathcal{V}} \left|f(x_1, x_2, \ldots, x_i, \ldots, x_N) - f(x_1, x_2, \ldots, x_i', \ldots, x_N)\right| \leqslant 1.$$

Therefore by applying Theorem 1.3.8, we obtain

$$\mathbf{Pr}\left[\left|N_t - \mathbf{E}\left[N_t\right]\right| \geqslant \sqrt{4 \cdot S_t \cdot \log\log n}\right] \leqslant 2 \cdot \mathrm{e}^{\frac{-4S_t \log\log n}{2S_t}} = o\left(\frac{1}{\log n}\right).$$

So with probability $1 - o(\frac{1}{\log n})$ we have

$$N_t > \mathbf{E}\left[N_t\right] - 2\sqrt{S_t \log\log n} = U_t \cdot \mathbf{Pr}\left[Z_v = 1\right] - 2\sqrt{S_t \log\log n}. \tag{2.1}$$

Now we estimate $\mathbf{Pr}\left[Z_v = 1\right]$. We know that

$$\mathbf{Pr}\left[Z_v = 1\right] = 1 - \prod_{u \in \mathcal{I}_t} \left(1 - \frac{1}{n}\right)^{C_u}.$$

Hence using the approximation $1 - x \leqslant \mathrm{e}^{-x} \leqslant 1 - x + x^2$ for any $x \geqslant 0$ results into

$$\mathbf{Pr}\left[Z_v = 1\right] \geqslant 1 - \mathrm{e}^{-\sum_{u \in \mathcal{I}_t} C_u / n} = 1 - \mathrm{e}^{-s_t} \geqslant s_t - s_t^2.$$

We now plug the value obtained by the above formula into (2.1) and normalize it. So we obtain

$$\begin{aligned} n_t &\geqslant (1 - i_t) \cdot \left(s_t - s_t^2\right) - 2\sqrt{\frac{s_t \log\log n}{n}} \\ &= s_t - s_t^2 - i_t \cdot \left(s_t - s_t^2\right) - 2\sqrt{\frac{s_t \log\log n}{n}} \\ &\geqslant s_t - 2s_t^2 - 2\sqrt{\frac{s_t \log\log n}{n}}, \end{aligned}$$

where the last inequality comes from the fact that $i_t \leqslant s_t$.

$\square$

Corollary 2.3.3. *Consider the* Push *protocol. Then with probability at least* $1-o(\frac{1}{\log n})$ *for any round* t *in which* $S_t \leqslant \frac{n}{8}$ *we have*

$$I_{t+1} \geqslant I_t + \frac{S_t}{2}.$$

Proof. If we have $1 \leqslant S_t \leqslant \log n$, then applying Lemma 2.3.1 yields that with probability $1-o(\frac{1}{\log n})$, $N_t = S_t$. If we have $\log n \leqslant S_t \leqslant \frac{n}{8}$, then $2s_t^2 \leqslant \frac{s_t}{4}$ and $2\sqrt{\frac{s_t \log\log n}{n}} \leqslant \frac{s_t}{4}$ which implies that

$$\frac{s_t}{2} \leqslant s_t - 2s_t^2 - 2\sqrt{\frac{s_t \log\log n}{n}}.$$

On the other hand applying Lemma 2.3.2 shows that with probability at least $1 - o(\frac{1}{\log n})$,

$$\frac{s_t}{2} \leqslant s_t - 2s_t^2 - 2\sqrt{\frac{s_t \log\log n}{n}} \leqslant n_t.$$

$\square$

Corollary 2.3.4. *Consider the* Push *protocol. For any round* t *and positive integer* $k = \mathcal{O}(\log n)$ *in which* $S_{t+k} = o(n)$ *with probability* $1 - o(\frac{k}{\log n})$ *we have*

$$I_{t+k} \geqslant I_t \cdot \left(\frac{3}{2}\right)^k.$$

Proof. By assumption we have that for every integer $1 \leqslant i \leqslant k$, $S_{t+i} = o(n)$. Applying Corollary 2.3.3 shows that with probability $1 - o(\frac{1}{\log n})$

$$I_{t+i} \geqslant I_{t+i-1} + \frac{S_{t+i-1}}{2} \geqslant I_{t+i-1} \cdot \frac{3}{2}.$$

Using an inductive argument and the union bound for all k rounds imply that with probability at least $1 - o(\frac{k}{\log n}) = 1 - o(1)$ we have

$$I_{t+k} \geqslant I_t \cdot \left(\frac{3}{2}\right)^k.$$

$\square$

2.4. Push Protocol with a Bounded Mean

In this section we first study the Push protocol for the case where R has bounded mean and arbitrary variance. Afterwards we consider the Push protocol where R has bounded mean and variance. As this is the most basic setting, our runtime bound is even tight up to low-order terms. To this end, let $T_{total} = \min\{t \mid \mathbf{Pr}[I_t = n] \geqslant 1 - o(1)\}$ be the first round in which all nodes are informed whp.

Theorem 2.4.1. *Assume that R is any distribution with $\mathbf{E}[R] = \mathcal{O}(1)$. Then whp, the* Push *protocol needs $\Theta(\log n)$ rounds to inform all nodes.*

Proof. We point out that the results in [FG85, Pit87] for the standard Push protocol imply an upper bound of $\mathcal{O}(\log n)$ rounds. So in what follows we only show that with high probability the protocol needs at least $\Omega(\log n)$ rounds to inform all nodes.In the Push protocol, in round $t+1$, at most S_t randomly chosen uninformed nodes are informed. This implies that $\mathbf{E}[S_{t+1} \mid S_t]$ increases by at most $\mathbf{E}[R] \cdot S_t$. Since the origin of the rumor is chosen without knowing C_u, $\mathbf{E}[S_0] = \mathbf{E}[R]$. Using the law of total expectation yields that

$$\mathbf{E}[S_t] = \mathbf{E}[\ldots \mathbf{E}[\mathbf{E}[S_t|S_{t-1}]\,|S_{t-2}] \ldots |S_0] \leqslant (1+\mathbf{E}[R])^t \cdot \mathbf{E}[R].$$

By applying Markov's inequality, we conclude that

$$\mathbf{Pr}[I_t \geqslant n] \leqslant \mathbf{Pr}[S_t \geqslant n] \leqslant \frac{(1+\mathbf{E}[R])^t \cdot \mathbf{E}[R]}{n}.$$

Hence $\Omega(\log n)$ rounds are necessary to inform all nodes whp. □

Theorem 2.4.2. *Consider the* Push *protocol and assume that R is a distribution with $\mathbf{E}[R] = \mathcal{O}(1)$ and $\mathbf{Var}[R] = \mathcal{O}(1)$. Then $|T_{total} - (\log_{1+\mathbf{E}[R]} n + \log_{e^{\mathbf{E}[R]}} n)| = o(\log n)$.*

Suppose that each random number C_u is generated according to some distribution R with bounded mean and variance. To prove this result, we study the protocol in three consecutive phases. In the following we give a brief overview of the proof.

- **The Preliminary Phase.** This phase starts with just one informed node and ends when $I_t \geqslant \log^5 n$ and $S_t \leqslant \log^{\mathcal{O}(1)} n$. Similar to the Birthday Paradox we show that in each round every Push call informs a different uninformed node and thus the number of informed nodes increases by $S_t \geqslant I_t$. Hence after $\mathcal{O}(\log\log n)$ rounds there are at least $\log^5 n$ informed nodes. Further, since $\mathbf{E}[R] = \mathcal{O}(1)$, after $\mathcal{O}(\log\log n)$ rounds we also have $S_t \leqslant \log^{\mathcal{O}(1)} n$.

- **The Middle Phase.** This phase starts when $\log^5 n \leqslant I_t \leqslant S_t \leqslant \log^{\mathcal{O}(1)} n$ and ends when $I_t \geqslant \frac{n}{\log\log n}$. First we show that the number of Push calls S_t increases by a factor of approximately $1+\mathbf{E}[R]$ as long as the number of informed nodes is $o(n)$. Then we prove that the number of newly informed nodes in round $t+1$ is roughly the same as S_t. Therefore an inductive argument shows that it takes $\log_{1+\mathbf{E}[R]} n \pm o(\log n)$ rounds to reach $\frac{n}{\log\log n}$ informed nodes.

- **The Final Phase.** This phase starts when $I_t \geqslant \frac{n}{\log\log n}$ and ends when all nodes are informed with high probability. In this phase, we first prove that after $o(\log n)$ rounds the number of uninformed nodes decreases to $\frac{n}{\log^5 n}$. Then we show that the probability that an arbitrary uninformed node remains uninformed is $e^{-\mathbf{E}[R] \pm o(\frac{1}{\log n})}$, so U_t decreases by this probability. Finally, an inductive argument establishes that it takes $\log_{e^{\mathbf{E}[R]}} n \pm o(\log n)$ rounds until every node is informed.

In the following we present the detailed proofs of these phases. Before that we show the following proposition.

Proposition 2.4.3. *Let R be a probability distribution with $\mathbf{E}[R] = \mathcal{O}(1)$ and $\mathbf{Var}[R] = \mathcal{O}(1)$. Let t be a round so that there exists a (possibly non-constant) $\delta > 0$ such that $U_t = n^{1-\delta}$. Then with probability $1 - o(\frac{1}{\log n})$,*

$$\sum_{u \in \mathcal{U}_t} C_u = \mathcal{O}(n^{1-\delta/2} \cdot \log^{1+\epsilon} n),$$

where $\epsilon > 0$ is a constant independent of δ.

Proof. Let us define a random variable

$$W_k := \sum_{u \in \mathcal{V}} C_u \cdot \mathbb{1}(C_u \geqslant k),$$

where $\mathbb{1}(C_u \geqslant k)$ is an indicator random variable which takes one if $C_u \geqslant k$ and zero otherwise. By linearity of expectation and the fact that all C_u's are independent and identically distributed random variables we have that

$$\begin{aligned}\mathbf{E}[W_k] &= \sum_{u \in \mathcal{V}} \mathbf{E}[C_u \mathbb{1}(C_u \geqslant k)] = n \cdot \mathbf{E}[C_u \mathbb{1}(C_u \geqslant k)] \\ &= n \cdot \sum_{l \geqslant k} l \cdot \mathbf{Pr}[C_u = l] \leqslant \frac{n}{k} \cdot \sum_{l \geqslant k} l^2 \cdot \mathbf{Pr}[C_u = l].\end{aligned}$$

Since C_u is a random variable with bounded variance,

$$\sum_{l \geqslant k} l^2 \cdot \mathbf{Pr}[C_u = l] = \mathcal{O}(1).$$

Thus, $\mathbf{E}[W_k] = \mathcal{O}(\frac{n}{k})$. Using Markov's inequality implies that with probability $1 - \mathcal{O}(\frac{1}{\log^{1+\epsilon} n}) = 1 - o(\frac{1}{\log n})$, $W_k = \mathcal{O}(\frac{n \cdot \log^{1+\epsilon} n}{k})$. If we set $k = n^{\delta/2}$, then

$$\sum_{u \in \mathcal{U}_t} C_u = \sum_{\{u \in \mathcal{U}_t :\, C_u \geqslant k\}} C_u + \sum_{\{u \in \mathcal{U}_t :\, C_u < k\}} C_u \leqslant W_k + \mathcal{O}(n^{1-\delta} \cdot k) = \mathcal{O}(n^{1-\delta/2} \cdot \log^{1+\epsilon}).$$

□

2.4.1. The Preliminary Phase

This phase starts with one informed node and ends when $I_t \geqslant \log^5 n$ and $S_t \leqslant \log^{\mathcal{O}(1)} n$. Let T_0 be the first round in which the number of informed nodes exceeds $\log^5 n$.

Lemma 2.4.4. *For any round $t = \mathcal{O}(\log \log n)$, with probability at least $1 - \frac{1}{\log^3 n}$ we have $S_t = \log^{\mathcal{O}(1)} n$.*

Proof. We will bound the expected number of calls in each round t as follows:

$$\mathbf{E}\left[S_t \,|\, S_{t-1}\right] = S_{t-1} + \mathbf{E}\left[\sum_{u \in \mathcal{N}_{t-1}} C_u \;\middle|\; S_{t-1}\right] = S_{t-1} + N_{t-1} \cdot \mathbf{E}\left[R\right] \leqslant S_{t-1} \cdot (1 + \mathbf{E}\left[R\right]),$$

where the last inequality comes from the fact that $N_{t-1} \leqslant S_{t-1}$. Since the origin of the rumor is chosen arbitrarily without knowing C_u, $\mathbf{E}\left[S_0\right] = \mathbf{E}\left[R\right]$. Applying the law of total expectation yields

$$\mathbf{E}\left[S_t\right] = \mathbf{E}\left[\dots \mathbf{E}\left[\mathbf{E}\left[S_t \,|\, S_{t-1}\right] \,|\, S_{t-2}\right] \dots \,|\, S_0\right] \leqslant (1 + \mathbf{E}\left[R\right])^t \,\mathbf{E}\left[S_0\right] = (1 + \mathbf{E}\left[R\right])^t \,\mathbf{E}\left[R\right].$$

By using Markov's inequality we have that

$$\mathbf{Pr}\left[S_t \geqslant (1 + \mathbf{E}\left[R\right])^t \cdot \mathbf{E}\left[R\right] \cdot \log^3 n\right] \leqslant \frac{1}{\log^3 n}.$$

So with probability $1 - \frac{1}{\log^3 n}$, for any $t = \mathcal{O}(\log\log n)$,

$$S_t \leqslant (1 + \mathbf{E}\left[R\right])^t \cdot \mathbf{E}\left[R\right] \cdot \log^3 n = \log^{\mathcal{O}(1)} n.$$

$\square$

Corollary 2.4.5. *whp we have $T_0 = \mathcal{O}(\log\log n)$.*

Proof. Using Lemma 2.4.4 gives that with probability at least $1 - \mathcal{O}(\frac{1}{\log^3 n})$, $S_t = \log^{\mathcal{O}(1)} n$ for any round $t = \mathcal{O}(\log\log n)$. Conditioning on this event, we can apply Lemma 2.3.1 and conclude that with probability $1 - \left(\frac{\log^{\mathcal{O}(1)} n}{n}\right)$, for any round $t = \mathcal{O}(\log\log n)$,

$$I_{t+1} = I_t + S_t \geqslant 2I_t,$$

where the inequality comes from the fact that $S_t \geqslant I_t$. Solving the above recursive inequality for any $t = \mathcal{O}(\log\log n)$ shows that $I_t \geqslant 2^t \cdot I_0 = 2^t$. So with probability

$$\left(1 - \frac{1}{\log^3 n}\right)\left(1 - \mathcal{O}(\log\log n) \cdot \frac{\log^{\mathcal{O}(1)} n}{n}\right) = 1 - o(1),$$

there exists a round $T_0 = \mathcal{O}(\log\log n)$ such that $I_{T_0} \geqslant \log^5 n$ and $S_{T_0} \leqslant \log^{\mathcal{O}(1)} n$. $\square$

2.4.2. The Middle Phase

The phase starts when $\log^5 n \leqslant I_t \leqslant S_t \leqslant \log^{\mathcal{O}(1)} n$ and ends when $I_t \geqslant \frac{n}{\log\log n}$. Let T_1 be the first round so that $I_{T_1} \geqslant \frac{n}{\log\log n}$. The main result of this subsection is that $|T_1 - \log_{1+\mathbf{E}[R]} n| = o(\log n)$.

Lemma 2.4.6. *Suppose that for a round t we have $s_t = \Omega\left(\frac{\log^5 n}{n}\right)$ and ,$s_t = o(1)$. Then for any $k = \mathcal{O}(\log n)$ with $(1 + \mathbf{E}\left[R\right])^k s_t = o(1)$, with probability $1 - o\left(\frac{k}{\log n}\right)$,*

$$\text{for all } 1 \leqslant i \leqslant k, \quad (1 + \mathbf{E}\left[R\right])^i \cdot s_t \cdot (1 - o(1)) \leqslant s_{t+i} \leqslant (1 + \mathbf{E}\left[R\right])^i \cdot s_t \cdot (1 + o(1)). \tag{2.2}$$

Proof. Consider the random variable $\sum_{u\in\mathcal{N}_t} C_u$. By linearity of expectation, $\mathbf{E}\left[\sum_{u\in\mathcal{N}_t} C_u\right] = N_t \cdot \mathbf{E}[R]$. Since the C_u's are independent and identically distributed random variables, we have that

$$\mathbf{Var}\left[\sum_{u\in\mathcal{N}_t} C_u\right] = N_t \cdot \mathbf{Var}[R].$$

Chebychev's inequality implies that

$$\mathbf{Pr}\left[\Big|\sum_{u\in\mathcal{N}_t} C_u - N_t\,\mathbf{E}[R]\Big| \geqslant \sqrt{N_t \log^2 n}\right] \leqslant \frac{N_t\,\mathbf{Var}[R]}{N_t \log^2 n} = o\left(\frac{1}{\log n}\right).$$

Since $S_{t+1} = S_t + \sum_{u\in\mathcal{N}_t} C_u$, it follows that with probability $1 - o(\frac{1}{\log n})$,

$$S_t + N_t\cdot\mathbf{E}[R] - \sqrt{N_t\log^2 n} \leqslant S_{t+1} \leqslant S_t + N_t\cdot\mathbf{E}[R] + \sqrt{N_t\log^2 n}. \tag{2.3}$$

Using the above formula and the fact that $N_t \leqslant S_t$ we have

$$S_{t+1} \leqslant S_t + S_t\cdot\mathbf{E}[R] + \sqrt{S_t\log^2 n} \leqslant S_t\cdot\left(1 + \mathbf{E}[R] + \sqrt{\frac{\log^2 n}{S_t}}\right).$$

Since S_t is a non-decreasing function in t and $\log^5 n \leqslant I_t \leqslant S_t$, with probability $1 - o(\frac{1}{\log n})$

$$s_{t+1} \leqslant s_t\cdot(1+\mathbf{E}[R])\left(1 + \sqrt{\frac{\log^2 n}{(1+\mathbf{E}[R])^2\log^5 n}}\right) < s_t\cdot(1+\mathbf{E}[R])\left(1 + \frac{1}{\log^{\frac{3}{2}} n}\right).$$

An inductive argument and the union bound for all k events that violate the above inequality shows that for any $k = \mathcal{O}(\log n)$ with probability $1 - o(\frac{k}{\log n})$,

$$\text{for all } 1\leqslant i\leqslant k,\ \ s_{t+i} \leqslant s_t\cdot(1+\mathbf{E}[R])^i\,(1+o(1)). \tag{2.4}$$

In order to prove the left hand side of (2.2), we use Lemma 2.3.2 which states with probability $1 - o(\frac{1}{\log n})$,

$$n_t \geqslant s_t - 2s_t^2 - 2\sqrt{\frac{s_t\log\log n}{n}}.$$

Using the lower bound in the inequality (2.3) and the above formula implies that with probability $1 - o(\frac{1}{\log n})$,

$$\begin{aligned}
s_{t+1} &\geqslant s_t + n_t\cdot\mathbf{E}[R] - \sqrt{\frac{n_t\log^2 n}{n}}\\
&\geqslant s_t + s_t\cdot\mathbf{E}[R] - 2s_t^2\cdot\mathbf{E}[R] - 2\sqrt{\frac{s_t\log\log n}{n}}\cdot\mathbf{E}[R] - \sqrt{\frac{s_t\cdot\log^2 n}{n}}\\
&\geqslant (1+\mathbf{E}[R])s_t - 2\,\mathbf{E}[R]\,s_t^2 - 2\sqrt{\frac{s_t\log^2 n}{n}}\\
&\geqslant (1+\mathbf{E}[R])s_t - F(s_t),
\end{aligned}$$

where $F(s_t) = 2\,\mathbf{E}[R]\,s_t^2 + 2\sqrt{\frac{s_t \log^2 n}{n}}$. An inductive argument and the union bound for all k events that violate the above inequality show that for any integer k for which $(1+\mathbf{E}[R])^k \cdot s_t = o(1)$ with probability $1 - o(\frac{k}{\log n})$,

$$\text{for all } 1 \leqslant i \leqslant k, \quad s_{t+i} \geqslant (1+\mathbf{E}[R])^i s_t - \sum_{j=0}^{i-1} (1+\mathbf{E}[R])^j F(s_{t+i-j}). \tag{2.5}$$

Inequality (2.4) yields that with probability $1 - o(\frac{k}{\log n})$,

$$\text{for all } 1 \leqslant i \leqslant k = \mathcal{O}(\log n), \quad s_{t+i} \leqslant a \cdot s_t \cdot (1+\mathbf{E}[R])^i,$$

where $a := 1 + o(1)$. $F(s_t)$ is a non-decreasing function in s_t and hence for any $k = \mathcal{O}(\log n)$ and $1 \leqslant j \leqslant k$,

$$\begin{aligned} F(s_{t+i-j}) &\leqslant F(a \cdot (1+\mathbf{E}[R])^{i-j} s_t) \\ &\leqslant 2\,\mathbf{E}[R]\,(1+\mathbf{E}[R])^{2(i-j)} (a \cdot s_t)^2 + 2(1+\mathbf{E}[R])^{\frac{i-j}{2}} \sqrt{\frac{a \cdot s_t \log^2 n}{n}}. \end{aligned}$$

Hence by combining the above inequality and (2.5), we conclude that for any integer k, where $(1+\mathbf{E}[R])^k s_t = o(1)$ and $k = \mathcal{O}(\log n)$ with probability $1 - o(\frac{k}{\log n})$, for all $1 \leqslant i \leqslant k$

$$\begin{aligned} s_{t+i} &\geqslant (1+\mathbf{E}[R])^i s_t - 2\,\mathbf{E}[R] \sum_{j=0}^{i-1} (1+\mathbf{E}[R])^{2i-j} (c \cdot s_t)^2 - 2 \sum_{j=0}^{i-1} (1+\mathbf{E}[R])^{\frac{i+j}{2}} \sqrt{\frac{c \cdot s_t \log^2 n}{n}} \\ &\geqslant (1+\mathbf{E}[R])^i s_t - d_1 \cdot (1+\mathbf{E}[R])^{2i} s_t^2 - d_2 \cdot (1+\mathbf{E}[R])^i \cdot \sqrt{\frac{s_t \log^2 n}{n}} \\ &= (1+\mathbf{E}[R])^i s_t \cdot \left(1 - d_1 \cdot (1+\mathbf{E}[R])^i s_t - d_2 \cdot \sqrt{\frac{\log^2 n}{s_t n}}\right), \end{aligned}$$

where d_1 and d_2 are constants which do not depend on i. Since $(1+\mathbf{E}[R])^k s_t = o(1)$ and $s_t = \Omega(\frac{\log^5 n}{n})$, for any $1 \leqslant i \leqslant k$,

$$s_{t+i} \geqslant (1+\mathbf{E}[R])^i \cdot s_t \cdot (1 - o(1)).$$

□

Lemma 2.4.7. *Suppose that* $\frac{\log^5 n}{n} \leqslant i_t \leqslant s_t \leqslant \frac{\log^{\mathcal{O}(1)} n}{n}$. *Then for any* $k = \mathcal{O}(\log n)$ *with* $(1+\mathbf{E}[R])^k s_t = o(1)$, *whp,*

$$i_t + f_2 \cdot (1+\mathbf{E}[R])^k \cdot s_t \cdot (1 - o(1)) \leqslant i_{t+k} \leqslant i_t + f_1 \cdot (1+\mathbf{E}[R])^k \cdot s_t \cdot (1 + o(1)),$$

where $f_1 > 0$ *and* $f_2 > 0$ *are constants.*

Proof. It is easy to see that

$$i_{t+k} = i_t + \sum_{i=0}^{k-1} n_{t+i} \leqslant i_t + \sum_{i=0}^{k-1} s_{t+i}.$$

Applying Lemma 2.4.6 implies that for any integer k for which $(1+\mathbf{E}[R])^k \cdot s_t = o(1)$, with probability $1 - o(\frac{k}{\log n})$ the following upper bound holds:

$$i_{t+k} \leqslant i_t + \sum_{i=0}^{k-1} s_{t+i} \leqslant i_t + s_t \cdot (1+o(1)) \cdot \sum_{i=0}^{k-1}(1+\mathbf{E}[R])^i = i_t + f_1 \cdot (1+\mathbf{E}[R])^k \cdot s_t \cdot (1+o(1)), \tag{2.6}$$

where $f_1 > 0$ is a constant. On the other hand, Lemma 2.3.2 yields that with probability $1 - o(\frac{1}{\log n})$,

$$n_t \geqslant s_t - 2s_t^2 - 2\sqrt{\frac{s_t \log\log n}{n}}.$$

Another application of Lemma 2.4.6 shows that with probability $1 - o(\frac{k}{\log n})$, for all integers $1 \leqslant i \leqslant k$ in which $(1+\mathbf{E}[R])^k s_t = o(1)$ and $s_t \geqslant \frac{\log^5 n}{n}$,

$$(1+\mathbf{E}[R])^i \cdot s_t \cdot (1-o(1)) \leqslant s_{t+i} \leqslant (1+\mathbf{E}[R])^i \cdot s_t \cdot (1+o(1)).$$

Using these two inequalities, as long as $(1+\mathbf{E}[R])^k s_t = o(1)$, we have with probability $1 - o(\frac{k}{\log n})$,

$$\begin{aligned}
i_{t+k} &= i_t + \sum_{i=0}^{k-1} n_{t+i} \\
&\geqslant i_t + \sum_{i=0}^{k-1} s_{t+i} - \sum_{i=0}^{k-1} \left\{ 2s_{t+i}^2 + 2\sqrt{\frac{s_{t+i}\log\log n}{n}} \right\} \\
&\geqslant i_t + (1-o(1)) \sum_{i=0}^{k-1} (1+\mathbf{E}[R])^i s_t - (2+o(1)) \sum_{i=0}^{k-1} \left\{ (1+\mathbf{E}[R])^{2i} s_t^2 + (1+\mathbf{E}[R])^{i/2} \sqrt{\frac{s_t \log\log n}{n}} \right\} \\
&\geqslant i_t + f_2 \cdot (1+\mathbf{E}[R])^k \cdot s_t - d \cdot \left((1+\mathbf{E}[R])^{2k} s_t^2 + (1+\mathbf{E}[R])^{k/2} \sqrt{\frac{s_t \log\log n}{n}} \right) \\
&\geqslant i_t + f_2 \cdot (1+\mathbf{E}[R])^k \cdot s_t \cdot \left(1 - f \cdot (1+\mathbf{E}[R])^k \cdot s_t - d \cdot (1+\mathbf{E}[R])^{-k/2} \sqrt{\frac{\log\log n}{s_t n}} \right),
\end{aligned} \tag{2.7}$$

where $f_2 > 0$ and $d > 0$ are constants. Since $\frac{\log^5 n}{n} \leqslant i_t \leqslant s_t$, we obtain that

$$i_{t+k} \geqslant i_t + f_2 \cdot (1+\mathbf{E}[R])^k \cdot s_t \cdot (1-o(1)). \tag{2.8}$$

By combining equations (2.8) and (2.7) we infer that with probability $1 - o(\frac{k}{\log n})$,

$$i_t + f_2 \cdot (1+\mathbf{E}[R])^k \cdot s_t \cdot (1-o(1)) \leqslant i_{t+k} \leqslant i_t + f_1 \cdot (1+\mathbf{E}[R])^k \cdot s_t \cdot (1+o(1)).$$

□

Corollary 2.4.8. *whp we have* $|T_1 - \log_{1+\mathbf{E}[R]} n| = o(\log n)$.

Proof. Applying Corollary 2.4.5 shows that whp, $T_0 = \mathcal{O}(\log\log n)$, where T_0 is the first round in which $\frac{\log^5 n}{n} \leqslant i_{T_0} \leqslant s_{T_0} \leqslant \frac{\log^{\mathcal{O}(1)} n}{n}$. Now we can apply Lemma 2.4.7 and set $k = \log_{1+\mathbf{E}[R]} n - o(\log n)$ such that with probability at least $1-o(1)$ we have $\frac{1}{\log\log n} \leqslant i_{T_0+k} \leqslant \frac{A}{\log\log n}$, where $A > 1$ is a constant. Then we conclude that with probability $1-o(1)$, $|T_1 - \log_{1+\mathbf{E}[R]} n| = o(\log n)$. □

2.4.3. The Final Phase

This phase starts with at least $\frac{n}{\log\log n}$ informed nodes and ends when all nodes get informed. Let T_1 be the first round in which $I_{T_1} \geqslant \frac{n}{\log\log n}$ and let T_2 be the first round in which all nodes are informed whp. We will show that whp, $|(T_2 - T_1) - \log_{e^{\mathbf{E}[R]}} n| = o(\log n)$.

Lemma 2.4.9. *whp,*

$$|(T_2 - T_1) - \log_{e^{\mathbf{E}[R]}} n| = o(\log n).$$

Proof. We define the indicator random variable Z_v for every $v \in \mathcal{U}_t$ and any round $t \geqslant T_1$:

$$Z_v = \begin{cases} 1 & \text{if } v \text{ does not get informed in round t+1,} \\ 0 & \text{otherwise.} \end{cases}$$

Thus,

$$\mathbf{E}\left[U_{t+1} \,|\, U_t\right] = \mathbf{E}\left[\sum_{v\in\mathcal{U}_t} Z_v\right] = U_t \cdot \mathbf{Pr}\left[Z_v = 1\right],$$

where for simplicity we omit the conditioning of U_{t+1} on U_t when dealing with the Z_v's. Using the fact that $1 - \frac{1}{n} = e^{-\frac{1}{n} - \mathcal{O}(\frac{1}{n^2})}$, we can approximate the value $\mathbf{Pr}\left[Z_v = 1\right]$ as follows,

$$\mathbf{Pr}\left[Z_v = 1\right] = \prod_{u\in\mathcal{I}_t}\left(1 - \frac{1}{n}\right)^{C_u} = \prod_{u\in\mathcal{I}_t} e^{-\frac{C_u}{n} - \mathcal{O}(\frac{C_u}{n^2})}$$
$$= e^{-\sum_{u\in\mathcal{I}_t}(\frac{C_u}{n} + \mathcal{O}(\frac{C_u}{n^2}))} = e^{-s_t - \mathcal{O}(\frac{s_t}{n})}.$$

Since $\frac{s_t}{n} = \mathcal{O}(\frac{1}{n})$ for any round and $\mathrm{e}^{-\mathcal{O}(\frac{1}{n})} = 1 - \mathcal{O}(\frac{1}{n})$,

$$\mathbf{E}\left[U_{t+1} \,|\, U_t\right] = U_t e^{-s_t} \cdot e^{-\mathcal{O}(\frac{1}{n})} = U_t e^{-s_t} - \mathcal{O}\left(\frac{U_t}{n}\right). \tag{2.9}$$

Since for every $u, v \in \mathcal{U}_t$,

$$\mathbf{Pr}\left[Z_u = 1 \cap Z_v = 1\right] = \mathbf{Pr}\left[Z_u = 1 \mid Z_v = 1\right] \cdot \mathbf{Pr}\left[Z_v = 1\right] \leqslant \mathbf{Pr}\left[Z_v = 1\right] \cdot \mathbf{Pr}\left[Z_u = 1\right],$$

we have that

$$\mathbf{E}\left[Z_u \cdot Z_v\right] \leqslant \mathbf{E}\left[Z_u\right] \cdot \mathbf{E}\left[Z_v\right].$$

Therefore,

$$\begin{aligned}\mathbf{Var}\left[\sum_{v\in\mathcal{U}_t} Z_v\right] &= \sum_{v\in\mathcal{U}_t}\mathbf{E}\left[Z_v^2\right] + \sum_{u\neq v}(\mathbf{E}\left[Z_u\cdot Z_v\right] - \mathbf{E}\left[Z_u\right]\cdot\mathbf{E}\left[Z_v\right])\\ &\leqslant \sum_{v\in\mathcal{U}_t}\mathbf{E}\left[Z_v^2\right] = U_t\cdot\mathbf{Pr}\left[Z_v=1\right] = \mathbf{E}\left[U_{t+1}\mid U_t\right]\leqslant U_t.\end{aligned}$$

Applying Chebychev's inequality implies that with probability $1-o(\frac{1}{\log n})$,

$$\left|U_{t+1}-\mathbf{E}\left[U_{t+1}\,|\,U_t\right]\right| \leqslant \sqrt{U_t\log^2 n}. \tag{2.10}$$

Combining inequalities (2.9) and (2.10) yields that with probability $1-o(\frac{1}{\log n})$,

$$\left|U_{t+1}-U_t e^{-s_t}\right| \leqslant \sqrt{U_t\log^2 n}+\mathcal{O}\left(\frac{U_t}{n}\right)\leqslant 2\sqrt{U_t\log^2 n}. \tag{2.11}$$

According to the value of U_t, we consider two cases.

- Suppose that $U_t \geqslant \frac{n}{\log^5 n}$. Note that $s_t\geqslant i_t\geqslant \frac{1}{\log\log n}$ by the assumption of the lemma. Since s_t is a non-decreasing value in t and $U_t<n$ the recursive formula (2.11) implies that with probability $1-o(\frac{1}{\log n})$,

  $$U_{t+1}\leqslant U_t\cdot e^{\frac{-1}{\log\log n}}+2\sqrt{n\log^2 n}.$$

  Using an inductive argument shows that with probability $1-o(\frac{k}{\log n})$,

  $$U_{t+k}\leqslant U_t\cdot e^{\frac{-k}{\log\log n}}+\sum_{i=0}^{k-1}e^{\frac{-i}{\log\log n}}\cdot\left(2\sqrt{n\log^2 n}\right).$$

  Hence after at most $k_0 = 6\log\log^2 n$ rounds whp the number of uninformed nodes decreases to $\frac{n}{\log^6 n}+\mathcal{O}(\sqrt{n\log^2 n})$, where $c>0$ is a constant.

- Suppose that $U_t\leqslant\frac{n}{\log^5 n}$. If we set $n^\delta=\log^5 n$, then applying Proposition 2.4.3 implies that for any t for which $U_t=\mathcal{O}(\frac{n}{\log^5 n})$ with probability $1-o(\frac{1}{\log n})$,

  $$\sum_{u\in\mathcal{U}_t}C_u = o\left(\frac{n}{\log n}\right). \tag{2.12}$$

  On the other hand, using Chebychev's inequality yields that with probability $1-o(\frac{1}{\log n})$,

  $$\left|\sum_{u\in\mathcal{V}}C_u-n\cdot\mathbf{E}\left[R\right]\right|\leqslant\sqrt{n\cdot\log^2 n}.$$

Combining the above equality and equality (2.12) results into an approximation for s_t which is not best possible but it suffices for our purpose. We know that

$$s_t = \sum_{u\in\mathcal{V}} C_u - \sum_{u\in\mathcal{U}_t} C_u.$$

So,

$$\mathbf{E}\left[R\right] - \sqrt{\frac{\log^2 n}{n}} - o\left(\frac{1}{\log n}\right) \leqslant s_t \leqslant \mathbf{E}\left[R\right] + \sqrt{\frac{\log^2 n}{n}}.$$

Therefore, s_t can be replaced by $\mathbf{E}\left[R\right] \pm o(\frac{1}{\log n})$ with probability $1 - o(\frac{1}{\log n})$. Inequality (2.11) implies that

$$\alpha \cdot U_t - 2\sqrt{U_t \log^2 n} \leqslant U_{t+1} \leqslant \alpha \cdot U_t + 2\sqrt{U_t \log^2 n}, \tag{2.13}$$

where $\alpha = e^{-\mathbf{E}[R] \pm o(1/\log n)}$. So as long as $U_t \geqslant \log^5 n$ with probability $1 - o(\frac{1}{\log n})$,

$$\begin{aligned} U_{t+1} \leqslant \alpha \cdot U_t + 2\sqrt{U_t \log^2 n} &= \alpha \cdot U_t \left(1 + 2\sqrt{\frac{\log^2 n}{\alpha^2 U_t}}\right) \\ &\leqslant \alpha \cdot U_t \left(1 + 2\sqrt{\frac{\log^2 n}{\alpha^2 \log^5 n}}\right) \leqslant \alpha \cdot U_t \left(1 + \frac{2}{\alpha \log^{\frac{3}{2}} n}\right). \end{aligned}$$

Now for any k for which $U_t e^{-k\,\mathbf{E}[R]} \geqslant \log^5 n$, with probability $1 - o(\frac{k}{\log n})$,

$$U_{t+k} \leqslant \alpha^k \cdot U_t \cdot \left(1 + \frac{2}{\alpha \log^{\frac{3}{2}} n}\right)^k = \alpha^k \cdot U_t \cdot (1 + o(1)). \tag{2.14}$$

In order to lower bound U_{t+k} we apply the lower bound (2.13) inductively. So we have that with probability $1 - o(\frac{k}{\log n})$,

$$U_{t+k} \geqslant \alpha^k \cdot U_t - \sum_{i=0}^{k-1} 2 \cdot \alpha^i \cdot \sqrt{U_{t+k-i-1} \log^2 n}.$$

Applying inequality (2.14) yields that with probability $1 - o(\frac{k}{\log n})$,

$$\sqrt{U_{t+k-i} \log^2 n} \leqslant \alpha^{\frac{k-i}{2}} \cdot \sqrt{U_t (1 + o(1)) \log^2 n}.$$

Thus,

$$\begin{aligned} U_{t+k} &\geqslant \alpha^k \cdot U_t - (1 + o(1)) \sum_{i=0}^{k-1} \alpha^{\frac{k-i-1}{2}} \cdot \sqrt{U_t \log^2 n} \\ &\geqslant \alpha^k \cdot U_t - c \cdot \alpha^{\frac{k}{2}} \cdot \sqrt{U_t \log^2 n}, \end{aligned} \tag{2.15}$$

where $c > 0$ is a constant and the last inequality holds because $\sum_{i=0}^{k-1} \alpha^{\frac{k-i-1}{2}} = \mathcal{O}(\alpha^{\frac{k}{2}})$. Combining the inequalities (2.14) and (2.15) yields that for any k satisfying $U_t e^{-k\,\mathbf{E}[R]} \geqslant \log^5 n$ with probability $1 - o(\frac{k}{\log n})$,

$$\alpha^k \cdot U_t(1 - o(1)) \leqslant U_{t+k} \leqslant \alpha^k \cdot U_t(1 + o(1)).$$

Hence by taking $k = \log_{e^{\mathbf{E}[R]}} n - o(\log n)$, whp, the number of uninformed nodes after $T_1 + k_0 + k$ rounds decreases to $\log^5 n$, so we have at most $\log^5 n$ uninformed nodes. Using the fact that for every $x \geqslant 0$, $1 - x \leqslant e^{-x}$, the probability that a node does not get informed after k_1 additional rounds is bounded from above by

$$\prod_{u \in \mathcal{I}_t} \left(1 - \frac{1}{n}\right)^{C_u \cdot k_1} \leqslant e^{-k_1 \sum_{u \in \mathcal{I}_t} C_u}.$$

We already know that $s_t = \mathbf{E}[R] \pm o(\frac{1}{\log n})$ and s_t is an non-decreasing value in t so

$$\sum_{u \in \mathcal{I}_t} C_u = s_t > \frac{\mathbf{E}[R]}{2}.$$

Thus the union bound implies that the probability that every node in $\mathcal{U}_t$ does not get informed is bounded by $\log^5 n \cdot e^{\frac{-k_1 \cdot \mathbf{E}[R]}{2}}$. By choosing $k_1 = \Theta(\log\log n)$ we conclude that with probability $1 - o(1)$ all nodes get informed. So we have with probability at least $1 - o(1)$ that $T_2 \leqslant T_1 + k_0 + k + k_1$, and $k_0 + k + k_1 = \log_{e^{\mathbf{E}[R]}} n + o(\log n)$.

$\square$

2.5. Push-Pull Protocol

In this section we study the Push-Pull protocol where R has bounded mean and variance. Before we present our results about the PUSH-PULL protocol we show the following general lemma.

Lemma 2.5.1. *Consider the* Push-Pull *protocol and let* $\{C_u : u \in \mathcal{V}\}$ *be a sequence of positive integers. Then whp, the* Push-Pull *protocol needs at least* $\Omega\left(\frac{\log n - \log S_0}{\log \sum_u \frac{C_u^2}{n}}\right)$ *rounds to inform all nodes.*

Proof. We know that the probability that an uninformed node u gets informed by Pull in round $t+1$ is bounded by $\frac{I_t \cdot C_u}{n}$. Therefore,

$$\begin{aligned}\sum_{u \in \mathcal{U}_t} \mathbf{E}\left[C_u \mathbb{1}(u \text{ gets informed by Pull}) \mid S_t\right] &= \sum_{u \in \mathcal{U}_t} C_u \cdot \mathbf{Pr}\left[u \text{ gets informed by Pull in round } t+1\right] \\ &\leqslant \sum_{u \in \mathcal{U}_t} C_u \cdot \frac{I_t \cdot C_u}{n} \leqslant I_t \cdot \sum_{u \in \mathcal{V}} \frac{C_u^2}{n}\end{aligned}$$

On the other hand the probability that a node $u \in \mathcal{U}_t$ gets informed by Push in round $t+1$ is at most $\frac{S_t}{n}$. So we get that

$$\begin{aligned}
&\sum_{u \in \mathcal{U}_t} \mathbf{E}\left[C_u \mathbb{1}(u \text{ gets informed by } \mathsf{Push}) \mid S_t\right] \\
&= \sum_{u \in \mathcal{U}_t} C_u \cdot \mathbf{Pr}\left[u \text{ gets informed by } \mathsf{Push} \text{ in round } t+1\right] \\
&\leqslant \sum_{u \in \mathcal{U}_t} C_u \cdot \frac{S_t}{n} \leqslant S_t \cdot \sum_{u \in \mathcal{V}} \frac{C_u^2}{n},
\end{aligned}$$

where the last inequality follows by $C_u \leqslant C_u^2$. Combining the above inequalities implies that

$$\mathbf{E}\left[S_{t+1} \mid S_t\right] \leqslant S_t + (S_t + I_t) \cdot \left(\sum_{u \in \mathcal{V}} \frac{C_u^2}{n}\right) \cdot \leqslant \left(1 + 2 \cdot \sum_{u \in \mathcal{V}} \frac{C_u^2}{n}\right) \cdot S_t,$$

Applying the law of total expectation yields that

$$\mathbf{E}\left[S_t\right] = \mathbf{E}\left[\ldots \mathbf{E}\left[\mathbf{E}\left[S_t | S_{t-1}\right] | S_{t-2}\right] \ldots | S_0\right] \leqslant \left(1 + 2 \cdot \sum_{u \in \mathcal{V}} \frac{C_u^2}{n}\right)^t \cdot S_0.$$

Using Markov's inequality implies that

$$\mathbf{Pr}\left[I_t = n\right] \leqslant \mathbf{Pr}\left[S_t > n/2\right] \leqslant \frac{\mathbf{E}\left[S_t\right]}{n/2} \leqslant \frac{\left(1 + 2 \cdot \sum_{u \in \mathcal{V}} \frac{C_u^2}{n}\right)^t \cdot S_0}{n/2}.$$

Therefore whp, the Push-Pull protocol needs at least $\Omega\left(\frac{\log n - \log S_0}{\log \sum_{u \in \mathcal{V}} \frac{C_u^2}{n}}\right)$ rounds to inform all nodes. □

Theorem 2.5.2. *Assume that R is any distribution with $\mathbf{E}[R] = \mathcal{O}(1)$ and $\mathbf{Var}[R] = \mathcal{O}(1)$. Then for any constant $\epsilon > 0$, with probability $1 - \epsilon$ the* Push-Pull *protocol needs at least $\Theta(\log n)$ rounds to inform all nodes.*

Proof. The Push-Pull protocol with $R \equiv 1$ was studied in [KSSV00] where the authors showed that with high probability the standard Push-Pull informs all nodes in $\Theta(\log n)$ rounds. This result implies that for any distribution R, with high probability, $\mathcal{O}(\log n)$ rounds are sufficient to inform all nodes. Let $\{C_u : u \in \mathcal{V}\}$ be a sequence of positive integers each of which is generated independently according to some distribution R with $\mathbf{E}[R] = \mathcal{O}(1)$ and $\mathbf{Var}[R] = \mathcal{O}(1)$. We call $\{C_u : u \in \mathcal{V}\}$ a *good* sequence if $\sum_{u \in \mathcal{V}} C_u^2 = \mathcal{O}(n)$ and $S_0 = \mathcal{O}(1)$. Since the origin of the rumor is chosen arbitrarily without knowing C_u, $\mathbf{E}[S_0] = \mathbf{E}[R]$. Applying Markov's inequality implies that for any constant $\epsilon > 0$ with probability at least $1 - \epsilon/2$, $S_0 = \mathcal{O}(1)$. Since R is a probability distribution with bounded variance, $\sum_{u \in \mathcal{V}} \mathbf{E}\left[C_u^2\right] = \mathcal{O}(n)$. Another application of Markov's inequality implies that with probability $1 - \epsilon/2$, $\sum_{u \in \mathcal{V}} C_u^2 = \mathcal{O}(n)$. Therefore using the union bound for failure probability of two mentioned events implies that for fixed $\epsilon > 0$ with probability at least $1 - \epsilon$, $\{C_u : u \in \mathcal{V}\}$ is a good sequence. Conditioning on the event that $\{C_u : u \in \mathcal{V}\}$ is a good sequence, using Lemma 2.5.1 implies that with probability at least $1 - o(1)$ the Push-Pull protocol needs $\Omega(\log n)$ rounds to inform n nodes and the result follows. □

2.6. Push-Pull Protocol with Power Law Distribution $2 < \beta < 3$

In this section we analyze the Push-Pull protocol where R is a power law distribution with $2 < \beta < 3$ and show the following theorem.

Theorem 2.6.1. *Assume that R is a power law distribution with $2 < \beta < 3$. Then the* Push-Pull *protocol informs all nodes in $\Theta(\log\log n)$ rounds with probability $1 - o(1)$.*

To prove the upper bound of $\mathcal{O}(\log\log n)$, we study the protocol in three consecutive phases and show that each phase takes only $\mathcal{O}(\log\log n)$ rounds. After that we show the lower bound $\Omega(\log\log n)$.

2.6.1. Proof of the Upper Bound

The Preliminary Phase. This phase starts with just one informed node and ends when $I_t \geqslant n^{\frac{1}{\beta-1}}/(2 \cdot \log n)$. Let T_1 be the first round where $I_{T_1} \geqslant n^{\frac{1}{\beta-1}}/(2\log n)$. We will show that $T_1 = \mathcal{O}(\log\log n)$. First we claim that $\mathcal{O}(\log\log n)$ rounds are sufficient to have $\log^{\mathcal{O}(1)} n$ informed nodes. Then we will show that in round $t+1$ with probability $1 - e^{-\Omega(\log n)}$ there exists a node u with $C_u \geqslant I_t^{1+\gamma}$, $\gamma := \frac{3-\beta}{2(\beta-2)} > 0$, which pulls the rumor and consequently $S_{t+1} \geqslant I_t^{1+\gamma}$. Then considering only Push calls it follows that with probability $1 - o(\frac{1}{\log n})$,

$$I_{t+2} = I_{t+1} + N_{t+1} \geqslant I_{t+1} + S_{t+1}(1 - o(1)) > \frac{1}{2} I_t^{1+\gamma}.$$

So in every two rounds, I_t is increased by a factor of $\frac{1}{2}I_t^{\gamma}$ and hence after $\mathcal{O}(\log\log n)$ rounds the phase ends. For a complete proof see the following lemma.

Lemma 2.6.2. *whp, $T_1 = \mathcal{O}(\log\log n)$.*

Proof. At first we only consider Push calls and apply Lemma 2.3.1 which states that as long as $S_t \leqslant \log^{\frac{2}{3-\beta}} n$, with probability $1 - \mathcal{O}(\frac{\log^{\frac{4}{3-\beta}} n}{n})$,

$$I_{t+1} = I_t + S_t \geqslant 2I_t.$$

Thus as long as $S_t \leqslant \log^{\frac{2}{3-\beta}} n$, in each round the number of informed nodes is at least doubled. So we conclude that whp, $\mathcal{O}(\log\log n)$ rounds are sufficient to inform $\log^{\frac{2}{3-\beta}} n$ nodes. Let T_0 be the first round when $I_{T_0} \geqslant \log^{\frac{2}{3-\beta}} n$. Let us define the constant $\gamma := \frac{3-\beta}{2(\beta-2)} > 0$. Let T be the first round such that

$$I_{T-1}^{(1+\gamma)} \leqslant n^{\frac{1}{\beta-1}}/\log n < I_T^{(1+\gamma)}.$$

Now for any $T_0 \leqslant t \leqslant T$, we can apply Proposition 2.2.3 and conclude that with probability $1 - o(\frac{1}{n})$,

$$\sum_{u \in \mathcal{L}(I_t^{1+\gamma})} C_u \geqslant L(I_t^{1+\gamma}) \cdot I_t^{1+\gamma} \geqslant \frac{n \cdot c_1 \cdot I_t^{(1+\gamma)(2-\beta)}}{2}. \tag{2.16}$$

So,

$$\frac{I_t}{n}\sum_{u\in\mathcal{L}(I_t^{1+\gamma})} C_u \geqslant \frac{c_1\cdot I_t^{1+(1+\gamma)(2-\beta)}}{2} = \frac{c_1\cdot I_t^{3-\beta+\gamma(2-\beta)}}{2}.$$

We will bound the probability that none of $u\in\mathcal{L}(I_t^{1+\gamma})$ gets informed by Pull calls in round $t+1$ as follows,

$$\prod_{u\in\mathcal{L}(I_t^{1+\gamma})}\left(1-\frac{I_t}{n}\right)^{C_u} = \left(1-\frac{I_t}{n}\right)^{\sum_{u\in\mathcal{L}(I_t^{1+\gamma})} C_u} \leqslant e^{-c_1\cdot I_t^{3-\beta+\gamma(2-\beta)}} = e^{-c_1\cdot I_t^{\frac{3-\beta}{2}}}.$$

Since for any $t\geqslant T_0$, $I_t\geqslant \log^{\frac{2}{3-\beta}} n$, we have that with probability at least $1-n^{-c_1}$, at least one node in $\mathcal{L}(I_t^{1+\gamma})$ gets informed by Pull in round $t+1$. Hence we have that

$$S_{t+1}\geqslant I_t^{1+\gamma}.$$

Let us now consider the Push calls in round $t+2$. By applying Lemma 2.3.1 we know that as long as $S_{t+1}=o(n)$ with probability $1-o(\frac{1}{\log n})$,

$$S_{t+1}(1-o(1))\leqslant N_{t+1}.$$

Thus,

$$I_{t+2}\geqslant I_{t+1}+S_{t+1}(1-o(1)) > \frac{I_t^{1+\gamma}}{2}.$$

An inductive argument shows that for any integer $k\geqslant 1$ as long as $I_{T_0+2k-2}^{1+\gamma}\leqslant n^{\frac{1}{\beta-1}}/\log n$, with probability $1-o(\frac{k}{\log n})$

$$I_{T_0+2k} > \left(\frac{1}{2}\right)^{\sum_{i=0}^{k-1}(1+\gamma)^i} I_{T_0}^{(1+\gamma)^k} = \left(\frac{I_{T_0}}{2^\gamma}\right)^{(1+\gamma)^k}\cdot 2^{1/\gamma} > \left(\frac{\log^{\frac{2}{3-\beta}} n}{C'}\right)^{(1+\gamma)^k},$$

where $C'=2^\gamma=\mathcal{O}(1)$. So we conclude that after T_0+2k rounds, where $k=o(\log_{1+\gamma}\log n)$, there are two cases: either $I_{T_0+2k}\geqslant n^{\frac{1}{\beta-1}}/(2\log n)$ which means $T_1\leqslant T_0+2k=\mathcal{O}(\log\log n)$ and we are done, or

$$I_{T_0+2k} < n^{\frac{1}{\beta-1}}/(2\log n) < n^{\frac{1}{\beta-1}}/\log n < I_{T_0+2k}^{1+\gamma}.$$

In the latter case, we change the value γ to γ' which satisfies $I_{T_0+2k}^{1+\gamma'}=n^{\frac{1}{\beta-1}}/\log n$ and a similar argument shows that

$$I_{T_0+2k+2}\geqslant n^{\frac{1}{\beta-1}}/(2\log n).$$

□

The Middle Phase. This phase starts with at least $n^{\frac{1}{\beta-1}}/(2\log n)$ informed nodes and ends when $I_t \geqslant \frac{n}{\log n}$. Let T_2 be the first round in which $\frac{n}{\log n}$ nodes are informed. We will show that $T_2 - T_1 = \mathcal{O}(\log\log n)$. In contrast to the Preliminary Phase where we focus only on an informed node with maximal C_u, we now consider the number of informed nodes u with a C_u above a certain threshold Z_{t+1} which is inversely proportional to I_t.

Lemma 2.6.3. *Suppose that* $I_t \geqslant n^{\frac{1}{\beta-1}}/(2\log n)$ *for some round* t. *Then with probability* $1 - o(\frac{1}{n})$,

$$|\mathcal{L}(Z_{t+1}) \cap \mathcal{I}_{t+1}| \geqslant \frac{1}{4} L(Z_{t+1}),$$

where $Z_{t+1} := \frac{n \log\log n}{I_t}$.

Proof. We consider two cases. If at least $\frac{1}{4}$ of the nodes in $\mathcal{L}(Z_{t+1})$ are already informed (before round $t+1$), then the statement of the lemma is true. Otherwise $|\mathcal{L}(Z_{t+1}) \cap \mathcal{U}_{t+1}| > \frac{3}{4} L(Z_{t+1})$. In the latter case, we define

$$\mathcal{L}'(Z_{t+1}) := \mathcal{L}(Z_{t+1}) \cap \mathcal{U}_{t+1}.$$

Let X_u be an indicator random variable for every $u \in \mathcal{L}'(Z_{t+1})$ so that

$$X_u := \begin{cases} 1 & \text{if } u \text{ gets informed by } \mathsf{Pull} \text{ in round } t+1, \\ 0 & \text{otherwise.} \end{cases}$$

Then we define a random variable X to be $X := \sum_{u \in \mathcal{L}'(Z_{t+1})} X_u$. Since for every $u \in \mathcal{L}'(Z_{t+1})$, $C_u \geqslant Z_{t+1} = \frac{n \log\log n}{I_t}$, it follows that

$$\mathbf{Pr}\left[X_u = 1\right] = 1 - \left(1 - \frac{I_t}{n}\right)^{C_u} \geqslant 1 - \left(1 - \frac{I_t}{n}\right)^{Z_{t+1}} = 1 - e^{-\Omega(\log\log n)} = 1 - o(1).$$

Thus $\mathbf{Pr}\left[X_u = 1\right] > \frac{3}{4}$ and $\mathbf{E}\left[X\right] = \sum_{u \in \mathcal{L}'(Z_{t+1})} \mathbf{Pr}\left[X_u = 1\right] > \frac{3}{4}|\mathcal{L}'(Z_{t+1})|$. Since $|\mathcal{L}'(Z_{t+1})| = |\mathcal{L}(Z_{t+1}) \cap \mathcal{U}_{t+1}| > \frac{3}{4} L(Z_{t+1})$,

$$\mathbf{E}\left[X\right] \geqslant \frac{9}{16} L(Z_{t+1}).$$

We know that $I_t \geqslant n^{\frac{1}{\beta-1}}/(2\log n)$ and also I_t is a non-decreasing function in t, so

$$Z_{t+1} = \frac{n \log\log n}{I_t} \leqslant 2 \cdot n^{\frac{\beta-2}{\beta-1}} \log n \log\log n < n^{\frac{1}{\beta-1}}/\log n,$$

where the last inequality holds because $\beta < 3$. Now we can apply Proposition 2.2.3 (see appendix) to infer that with probability $1 - o(\frac{1}{n})$,

$$L(Z_{t+1}) \geqslant \frac{n \cdot c_1 \cdot Z_{t+1}^{1-\beta}}{2} \geqslant \frac{c_1 \cdot \log^{\beta-1} n}{2}.$$

Therefore,

$$\mathbf{E}\left[X\right] \geqslant \frac{9 \cdot c_1 \cdot \log^{\beta-1} n}{32}.$$

Then applying Theorem 1.3.4 results into

$$\mathbf{Pr}\left[X < \frac{\mathbf{E}[X]}{2}\right] \leqslant \mathbf{Pr}\left[|X - \mathbf{E}[X]| \geqslant \frac{\mathbf{E}[X]}{2}\right] < 2e^{-\frac{\mathbf{E}[X]}{10}} \leqslant 2e^{-\Omega(\log^{\beta-1} n)}. \quad (2.17)$$

So with probability $1 - o(\frac{1}{n})$, we have that

$$|\mathcal{L}(Z_{t+1}) \cap \mathcal{I}_{t+1}| \geqslant X \geqslant \frac{\mathbf{E}[X]}{2} > \frac{3|\mathcal{L}'(Z_{t+1})|}{8} \geqslant \frac{1}{4}L(Z_{t+1}),$$

where the last inequality holds because $|\mathcal{L}'(Z_{t+1})| > \frac{3}{4}L(Z_{t+1})$. □

Lemma 2.6.4. *With probability* $1 - o(1)$, $T_2 - T_1 = \mathcal{O}(\log\log n)$.

Proof. Since $I_t \geqslant n^{\frac{1}{\beta-1}}/(2\log n)$, $Z_{t+1} = \frac{n\log\log n}{I_t} < n^{\frac{1}{\beta-1}}/\log n$, using Proposition 2.6.3 results into a lower bound for $|\mathcal{L}(Z_{t+1}) \cap \mathcal{I}_{t+1}|$. So with probability $1 - o(\frac{1}{n})$,

$$S_{t+1} = \sum_{u \in I_{t+1}} C_u \geqslant |\mathcal{L}(Z_{t+1} \cap \mathcal{I}_{t+1})| \cdot Z_{t+1} \geqslant \frac{1}{4}L(Z_{t+1}) \cdot Z_{t+1}.$$

By applying Proposition 2.2.3, we conclude that with probability $1 - o(\frac{1}{n})$, $L(Z_{t+1}) \geqslant \frac{n \cdot c_1 \cdot Z_{t+1}^{1-\beta}}{2}$. Therefore, with probability $1 - o(\frac{1}{n})$,

$$S_{t+1} \geqslant \frac{n \cdot c_1 \cdot Z_{t+1}^{2-\beta}}{8}.$$

As long as $S_{t+1} = o(n)$, we can apply Lemma 2.3.2 for the Push protocol to round $t+2$ implying that with probability $1 - o(\frac{1}{\log n})$,

$$I_{t+2} = I_{t+1} + N_t \geqslant I_{t+1} + S_{t+1}(1 - o(1)).$$

Thus,

$$I_{t+2} > \frac{S_{t+1}}{2} \geqslant \frac{c_1}{16} n \cdot Z_{t+1}^{2-\beta} = \frac{c_1}{16} \cdot n^{3-\beta} \cdot \log\log^{2-\beta} n \cdot I_t^{\beta-2}.$$

By an inductive argument, we obtain that for any integer $k \geqslant 1$ with $S_{t+k} = o(n)$, it holds with probability $1 - o(\frac{k}{\log n})$,

$$I_{t+2k} > \left(\frac{c}{16} n^{3-\beta} \cdot \log\log^{2-\beta} n\right)^{\sum_{i=0}^{k-1}(\beta-2)^i} I_t^{(\beta-2)^k} = \left(\frac{c}{16} n^{3-\beta} \cdot \log\log^{2-\beta} n\right)^{\frac{1-(\beta-2)^k}{3-\beta}} I_t^{(\beta-2)^k}.$$

Therefore there exists $k = \mathcal{O}(\log_{\frac{1}{\beta-2}} \log n)$ such that

$$\begin{aligned} I_{t+2k} &\geqslant \left(\frac{c}{16} n^{3-\beta} \cdot \log\log^{2-\beta} n\right)^{\frac{1-\mathcal{O}(1/\log n)}{3-\beta}} I_t^{1/\log n} \\ &= \Omega\left(n^{1-\mathcal{O}(1/\log n)} \left(\frac{c}{16} \cdot \log\log^{2-\beta} n\right)^{\frac{1-\mathcal{O}(1/\log n)}{3-\beta}}\right) = \Omega\left(\frac{n}{\log\log^{\delta} n}\right), \end{aligned}$$

where $\delta = \frac{\beta-2}{3-\beta}(1 - \mathcal{O}(1/\log n)) > 0$. Hence $T_2 \leqslant T_1 + 2k = T_1 + \mathcal{O}(\log\log n)$ whp. □

The Final Phase. This phase starts with at least $\frac{n}{\log n}$ informed nodes. Since the runtime of our Push-Pull protocol is stochastically smaller than the runtime of the standard Push-Pull protocol (i.e. $C_u = 1$ for every $u \in V$), we simply use the result by Karp et. al in [KSSV00, Theorem 2.1] for the standard Push-Pull protocol which states that once $I_t \geqslant \frac{n}{\log n}$, additional $\mathcal{O}(\log\log n)$ rounds are whp sufficient to inform all n nodes. $\square$

2.6.2. Proof of the Lower Bound

Since increasing the number of informed nodes can only decrease the runtime of the protocol, we may assume that at the beginning there are $\log^b n$ random informed nodes, where $b := \max\{4, 2 + \frac{3(3-\beta)}{\beta-2}\}$. Applying Markov's inequality to the random variable S_0 implies that with probability $1 - o(\frac{1}{\log n})$, $\log^b n \leqslant S_0 \leqslant \log^{2+b} n$. In the following we lower bound the number of rounds to reach $n^{\frac{1}{\log\log n}}$ informed nodes. We do this by keeping track of the largest value of C_u among all informed nodes and show that this value does not exceed $I_t^{\frac{1}{\beta-2}} \log^{\frac{3}{\beta-2}} n$ with high probability.

By Fact 2.2.2, with probability $1 - o(\frac{1}{\log n})$ we have $\max_{u\in\mathcal{V}} C_u \leqslant n^{\frac{1}{\beta-1}} \log n$. Let i^* be the smallest positive integer so that $2^{i^*} \geqslant n^{\frac{1}{\beta-1}}/\log n$. Then $i^* < \log n$. Let us define the set $\mathcal{M}_i := \{u \in \mathcal{V}\colon 2^{i-1} \leqslant C_u < 2^i\}$ for $1 \leqslant i \leqslant i^* - 1$ and $\mathcal{M}_{i^*} := \{u \in \mathcal{V}\colon 2^{i^*-1} \leqslant C_u \leqslant n^{\frac{1}{\beta-1}} \log n\}$. We denote the size of $\mathcal{M}_i$ with M_i. By definition, for any $1 \leqslant i \leqslant i^*$, $M_i \leqslant L(2^{i-1})$. Applying Proposition 2.2.3 implies that with probability $1 - o(\frac{1}{n})$ for any $1 \leqslant i \leqslant i^*$ we have $M_i \leqslant \frac{3}{2} \cdot c_2 \cdot n \cdot 2^{(i-1)(1-\beta)}$. Let us define the indicator random variable Z_u^i for every $u \in \mathcal{U}_t \cap \mathcal{M}_i$ as follows:

$$Z_u^i := \begin{cases} 1 & \text{if } u \text{ gets informed by Pull in round t+1,} \\ 0 & \text{otherwise.} \end{cases}$$

Hence, $\mathbf{Pr}\left[Z_u^i = 1\right] \leqslant C_u \cdot \frac{I_t}{n} \leqslant \frac{I_t \cdot 2^i}{n}$. Let P_i be the probability that at least one node in $\mathcal{U}_t \cap \mathcal{M}_i$ gets informed by Pull in round $t+1$. Then, for any $1 \leqslant i \leqslant i^* - 1$,

$$P_i \leqslant \sum_{u\in\mathcal{U}_t\cap\mathcal{M}_i} \mathbf{Pr}\left[Z_u^i = 1\right] \leqslant M_i \cdot \frac{I_t}{n} \cdot 2^i \leqslant 3 \cdot c_2 \cdot I_t \cdot 2^{(i-1)(2-\beta)}.$$

Since $2^{i^*} \geqslant n^{\frac{1}{\beta-1}}/\log n$ and $C_u \leqslant n^{\frac{1}{\beta-1}} \log n$ with probability $1 - o(\frac{1}{\log n})$,

$$\begin{aligned} P_{i^*} \leqslant \sum_{u\in\mathcal{U}_t\cap\mathcal{M}_{i^*}} \mathbf{Pr}\left[Z_u^i = 1\right] &\leqslant \frac{3}{2} \cdot c_2 \cdot n \cdot 2^{(i^*-1)(1-\beta)} \cdot \frac{I_t}{n} \cdot n^{\frac{1}{\beta-1}} \log n \\ &\leqslant 6 \cdot c_2 \cdot I_t \cdot n^{\frac{2-\beta}{\beta-1}} \log^{\beta-1} \cdot n. \end{aligned}$$

So as long as $I_t \leqslant n^{\frac{1}{\log\log n}}$, $P_{i^*} = o(\frac{1}{\log^3 n})$. We define $\Delta_t := S_t^{\frac{1}{\beta-2}} \log^{\frac{3}{\beta-2}} n$. Let $1 \leqslant i_t \leqslant i^*$ be the smallest integer so that $2^{i_t} \geqslant \Delta_t$. Then for any $i_t \leqslant i \leqslant i^*$ we have,

$$P_i \leqslant 3 \cdot c_2 \cdot 2^{\beta-2} \cdot I_t \cdot 2^{i(2-\beta)} \leqslant 6 \cdot c_2 \cdot I_t \cdot \Delta_t^{2-\beta} \leqslant 6 \cdot c_2 \cdot \log^{-3} n.$$

Let E_t be the event that no node with $C_u \geqslant \Delta_t$ gets informed by Pull in round $t+1$. Then we have

$$\mathbf{Pr}\left[E_t\right] \geqslant 1 - \sum_{i=i_t}^{i^*} P_i \geqslant 1 - o\left(\frac{1}{\log n}\right). \tag{2.18}$$

Let us define $S_{t+1}^{(1)} := \sum_{u \in \mathcal{N}_t^{\mathsf{Pull}}} C_u$. Conditioning on the event E_t we obtain that

$$\begin{aligned}
\mathbf{E}\left[S_{t+1}^{(1)} \mid S_t\right] &\leqslant \sum_{i=1}^{i_t} \sum_{u \in \mathcal{U}_t \cap \mathcal{M}_i} C_u \cdot \frac{\mathbf{Pr}\left[Z_u^i = 1\right]}{\mathbf{Pr}\left[E_t\right]} \\
&\leqslant (1+o(1)) \cdot \sum_{i=1}^{i_t} 2^i \cdot M_i \cdot \frac{I_t}{n} \cdot 2^i \leqslant (1+o(1)) \cdot \frac{S_t}{n} \cdot \sum_{i=1}^{i_t} 2^i \cdot M_i \cdot 2^i.
\end{aligned}$$

By definition of i_t, we have $2^{i_t} \leqslant 2 \cdot \Delta_t$ and $M_i \leqslant L(2^{i-1}) \leqslant \frac{3}{2} \cdot c_2 \cdot 2^{(i-1)(1-\beta)} \cdot n$. Hence the last sum is bounded by

$$\begin{aligned}
(1+o(1)) \cdot \sum_{i=1}^{i_t} 2^{2i} \cdot I_t \cdot 2^{(i-1)(1-\beta)} &\leqslant 24 \cdot c_2 \cdot I_t \cdot 2^{i_t(3-\beta)} \leqslant 24 \cdot c_2 \cdot I_t \cdot (2 \cdot \Delta_t)^{3-\beta} \\
&\leqslant 24 \cdot c_2 \cdot S_t^{1+\frac{3-\beta}{\beta-2}} \log^{\frac{3(3-\beta)}{\beta-2}} n.
\end{aligned}$$

Conditioning on the event E_t and applying Markov's inequality imply that with probability $1 - o(\frac{1}{\log n})$,

$$S_{t+1}^{(1)} \leqslant \log^2 n \cdot \mathbf{E}\left[S_{t+1}^{(1)} \mid S_t\right] \leqslant 24 \cdot c_2 \cdot S_t^{1+\frac{3-\beta}{\beta-2}} \log^{2+\frac{3(3-\beta)}{\beta-2}} n. \tag{2.19}$$

Let us define the indicator random variable Y_u for every $u \in \mathcal{U}_t$ as follows:

$$Y_u := \begin{cases} 1 & \text{if } u \text{ gets informed by } \mathsf{Push} \text{ in round t+1}, \\ 0 & \text{otherwise.} \end{cases}$$

Then we have $\mathbf{Pr}\left[Y_u = 1\right] \leqslant \frac{S_t}{n}$. Let A denote the event that $\sum_{u \in \mathcal{V}} C_u \leqslant n \cdot \log^2 n$. Since $\mathbf{E}\left[R\right] = \mathcal{O}(1)$, applying Markov's inequality implies that $\mathbf{Pr}\left[A\right] \geqslant 1 - o(\frac{1}{\log n})$. Let us define $S_{t+1}^{(2)} := \sum_{u \in \mathcal{N}_t^{\mathsf{Push}}} C_u$. Conditioning on the event A we have

$$\mathbf{E}\left[S_{t+1}^{(2)} \mid S_t\right] = \sum_{u \in \mathcal{U}_t} C_u \cdot \frac{\mathbf{Pr}\left[Y_u = 1\right]}{\mathbf{Pr}\left[A\right]} \leqslant (1+o(1)) \cdot \sum_{u \in \mathcal{V}} C_u \cdot \frac{S_t}{n} \leqslant (1+o(1)) \cdot S_t \cdot \log^2 n.$$

Conditioning on the event A and applying Markov's inequality implies that with probability $1 - o(\frac{1}{\log n})$,

$$S_{t+1}^{(2)} \leqslant \log^2 n \cdot \mathbf{E}\left[S_{t+1}^{(2)} \mid S_t\right] \leqslant S_t \cdot \log^4 n \tag{2.20}$$

Combining inequalities (2.19) and (2.20) implies that with probability $1 - o(\frac{1}{\log n})$ for every $0 \leqslant t \leqslant \log\log n$

$$S_{t+1} \leqslant S_t + S_{t+1}^{(1)} + S_{t+1}^{(2)} \leqslant S_t + 24 \cdot c_2 \cdot S_t^{1+\frac{3-\beta}{\beta-2}} \log^{2+\frac{3(3-\beta)}{\beta-2}} n + S_t \cdot \log^4 n$$
$$\leqslant S_t + 24 \cdot c_2 \cdot S_t^{b+1} + S_t^2 \leqslant S_t^{b+2},$$

where the last inequality holds because $b = \max\{4, 2 + \frac{3(3-\beta)}{\beta-2}\}$ and $\log^b n \leqslant I_t \leqslant S_t$. We know that with probability $1 - o(\frac{1}{\log n})$ we have $S_0 \leqslant \log^{b+2} n$. An inductive argument shows that for every $1 \leqslant t \leqslant \log\log n$ whp, $S_t \leqslant S_0^{(b+2)^t} \leqslant \log^{(b+2)^{t+1}}$. If we set $T := \frac{1}{2} \cdot \log_{b+2} \log n$, then whp we have $S_T < n^{\frac{1}{\log\log n}}$. Thus $T = \Omega(\log\log n)$ rounds are necessary to inform all nodes whp. This finishes the proof of the lower bound of $\Omega(\log\log n)$. $\square$

2.7. Push-Pull Protocol with Power Law Distribution $\beta = 3$

In this section we analyze the Push-Pull protocol where R is a power law distribution with $\beta = 3$ and show the following theorem.

Theorem 2.7.1. *Assume that R is a power law distribution with $\beta = 3$. Then the* Push-Pull *protocol informs all nodes in $\Theta\left(\frac{\log n}{\log\log n}\right)$ rounds with probability $1 - o(1)$.*

In order to prove this result we first show a lower bound of $\Omega\left(\frac{\log n}{\log\log n}\right)$ and then show an upper bound which is tight up to a constant factor. Throughout this section we assume that the power law distribution with $\beta = 3$ has an additional property that for every positive integer z

$$\mathbf{Pr}\left[R = z\right] \geqslant \frac{c}{z^3}, \tag{2.21}$$

where c is a universal constant. Let us define $\mathcal{L}^*(z) = \{u : C_u = z\}$ and $L^*(z) = |\mathcal{L}_0(z)|$. Furthermore, we define $\mathcal{I}_t(z) = \mathcal{I}_t \cap \mathcal{L}^*(z)$ and $\mathcal{N}_t(z) = \mathcal{N}_t \cap \mathcal{L}^*(z)$, whose sizes are denoted by $I_t(z)$ and $N_t(z)$, respectively. $N_t^{\mathsf{Push}}(z)$ and $N_t^{\mathsf{Pull}}(z)$ are denoting the size of the newly informed nodes with $C_u = z$ by Push and Pull transmissions, respectively. In the following we show a useful fact about $L^*(z)$.

Fact 2.7.2. *Let R be a power law distribution with $\beta = 3$. Then for every $z = \mathcal{O}(n^{1/4})$, with probability $1 - o(\frac{1}{n})$ we have that*

$$\frac{n \cdot \mathbf{Pr}\left[R = z\right]}{2} \leqslant L^*(z) \leqslant \frac{3 \cdot n \cdot \mathbf{Pr}\left[R = z\right]}{2}.$$

Proof. We know that $\mathbf{E}\left[L^*(z)\right] = n \cdot \mathbf{Pr}\left[R = z\right]$. By using the inequality (2.21) we have that for any $z = \mathcal{O}(n^{1/4})$, $\mathbf{Pr}\left[R = z\right] = \Omega(n^{-3/4})$. Then we have that $\mathbf{E}\left[L_0(z)\right] = \Omega(n^{2/5})$ and using a Chernoff bound (see e.g., Theorem 1.3.4) yields that with probability $1 - o(\frac{1}{n})$ the inequality in the statement holds. $\square$

2.7.1. Proof of Lower Bound

Theorem 2.7.3. *whp, the* Push-Pull *needs at least* $\Omega\left(\frac{\log n}{\log\log n}\right)$ *rounds to inform all* n *nodes.*

Proof. Let $\{C_u \colon u \in \mathcal{V}\}$ be a sequence of positive integers where every C_u is generated independently according to a power law distribution with $\beta = 3$. We call a sequence $\{C_u \colon u \in \mathcal{V}\}$ is good if it fulfills three conditions:

1. For every $u \in \mathcal{V}$, $C_u < n$.
2. $S_0 = \mathcal{O}(\log n)$.
3. $\sum_{u\in\mathcal{V}} \frac{C_u^2}{n} = \mathcal{O}(\log^2 n)$.

In the following we show that with probability $1 - o(1)$ every sequence $\{C_u, u \in \mathcal{V}\}$ is good. By definition of the power law distribution for $\beta = 3$ we have that

$$\mathbf{Pr}\left[C_u \leqslant n\right] > 1 - \frac{c_1}{n^2} = 1 - o(1).$$

We know that $\mathbf{E}\left[R\right] = \mathcal{O}(1)$, so Markov's inequality implies that with probability $1 - \mathcal{O}(\frac{1}{\log n})$, $S_0 = \mathcal{O}(\log n)$. Conditioning on the event that for every $u \in \mathcal{V}$, $C_u < n$ we get

$$\mathbf{E}\left[C_u^2 | C_u \leqslant n\right] \leqslant \frac{\sum_{z=1}^{n} \mathbf{Pr}\left[R^2 \geqslant z\right]}{\mathbf{Pr}\left[C_u \leqslant n\right]} \leqslant (1 + o(1)) \cdot c_1 \sum_{z=1}^{n} \frac{1}{z} = (1 + o(1)) \cdot c_1 \cdot \log n.$$

So applying Markov's inequality yields that with probability $1 - \mathcal{O}(\frac{1}{\log n})$,

$$\sum_{u\in\mathcal{V}} \frac{C_u^2}{n} = \mathcal{O}\left(\log^2 n\right).$$

Therefore we have that whp, the sequence $\{C_u \colon u \in \mathcal{V}\}$ is good. Conditioning on this event and then applying Lemma 2.5.1 shows that whp the Push-Pull needs at least

$$\Omega\left(\frac{\log n - \log S_0}{\log \sum_{u\in\mathcal{V}} C_u^2/n}\right) = \Omega\left(\frac{\log n}{\log\log n}\right)$$

rounds to inform n nodes. □

2.7.2. Proof of Upper Bound

Before we present a proof for the upper bound we show the following two lemmas.

Lemma 2.7.4. *Suppose that* $S_t \leqslant \frac{n}{\log^6 n}$ *and* $z \leqslant \min\{\frac{n}{I_t \cdot \log^6 n}, \mathcal{O}(n^{\frac{1}{4}})\}$. *Then with probability* $1 - o(\frac{1}{\log n})$, *for any round* $t = \mathcal{O}(\log n)$ *we have that*

$$|\mathcal{U}_t(z) \cap \mathcal{L}^*(z)| \geqslant \frac{L^*(z)}{2} \geqslant \frac{n \cdot \mathbf{Pr}\left[R = z\right]}{4}.$$

Proof. By considering the Push call we have that the size of newly informed nodes is bounded by S_t. Since they are chosen randomly, we have that

$$\mathbf{E}\left[N_t^{\mathsf{Push}}(z)|S_t\right] \leqslant S_t \cdot \mathbf{Pr}\left[R = z\right]. \tag{2.22}$$

On the other hand we have that

$$\begin{aligned}\mathbf{E}\left[N_t^{\mathsf{Pull}}(z)|I_t\right] &\leqslant \sum_{u \in \mathcal{U}_t \cap \mathcal{L}^*(z)} \mathbf{Pr}\left[u \text{ gets informed by Pull in round } t+1\right] \\ &\leqslant L^*(z) \cdot \mathbf{Pr}\left[u \text{ gets informed by Pull in round } t+1\right] \\ &= L^*(z) \cdot \left(1 - \left(1 - \frac{I_t}{n}\right)^z\right) \leqslant L^*(z) \cdot \frac{2 \cdot I_t \cdot z}{n}\end{aligned}$$

where the second inequality holds because $|\mathcal{U}_t \cap \mathcal{L}^*(z)| \leqslant L^*(z)$ and the last one holds as we assume that $\frac{I_t}{n} \leqslant \frac{S_t}{n} < \frac{1}{2}$ and for any $0 \leqslant a \leqslant \frac{\log 2}{2}$, $\mathrm{e}^{-2a} \leqslant 1 - a \leqslant e^{-a}$. Applying Fact 2.7.2 shows that for any $z = \mathcal{O}(n^{\frac{1}{4}})$ with probability $1 - o(\frac{1}{n})$ we have

$$\frac{n \cdot \mathbf{Pr}\left[R = z\right]}{2} \leqslant L^*(z) \leqslant \frac{3 \cdot n \cdot \mathbf{Pr}\left[R = z\right]}{2}. \tag{2.23}$$

Thus,

$$\mathbf{E}\left[N_t^{\mathsf{Pull}}(z)|S_t\right] \leqslant 3 \cdot I_t \cdot z \cdot \mathbf{Pr}\left[R = z\right]. \tag{2.24}$$

Combining (2.22) and (2.24) implies that

$$\mathbf{E}\left[N_t(z)|S_t, I_t\right] \leqslant S_t \cdot \mathbf{Pr}\left[R = z\right] + 3 \cdot I_t \cdot z \cdot \mathbf{Pr}\left[R = z\right]$$

We know that $I_{t+1}(z) = I_0(z) + \sum_{i=1}^{t} N_i(z)$. Using the linearity of expectation we have that

$$\begin{aligned}\mathbf{E}\left[I_{t+1}(z)|S_i, I_i, 0 \leqslant i \leqslant t\right] &= I_0(z) + \sum_{i=0}^{t} \mathbf{E}\left[N_i(z)|S_i, I_i\right] \\ &\leqslant I_0(z) + \mathbf{Pr}\left[R = z\right] \cdot \sum_{i=0}^{t} (S_i + 3 \cdot I_i \cdot z) \\ &\leqslant 1 + \mathbf{Pr}\left[R = z\right] \cdot (t+1) \cdot (S_t + z \cdot 3 \cdot I_t),\end{aligned}$$

where the last inequality comes from the fact that S_i and I_i are non-decreasing functions in t. By the assumption $z \leqslant \min\{\frac{n}{I_t \cdot \log^6 n}, \mathcal{O}(n^{\frac{1}{4}})\}$ and $S_t \leqslant \frac{n}{\log^6 n}$, for any round $t = \mathcal{O}(\log n)$ we have that

$$\mathbf{E}\left[I_{t+1}(z)|S_i, I_i, 1 \leqslant i \leqslant t\right] \leqslant 2 \cdot (t+1) \cdot (S_t + 3 \cdot I_t \cdot z) \cdot \mathbf{Pr}\left[R = z\right] \leqslant \frac{n \cdot \mathbf{Pr}\left[R = z\right]}{\log^4 n}.$$

Applying Markov's inequality shows that with probability $1 - o(\frac{1}{\log n})$ for any round $t = \mathcal{O}(\log n)$,

$$I_{t+1}(z) \leqslant \log^2 n \cdot \mathbf{E}\left[I_{t+1}(z)|S_i, I_i, 0 \leqslant i \leqslant t\right] \leqslant \frac{n \cdot \mathbf{Pr}\left[R = z\right]}{\log^2 n} \leqslant \frac{L^*(z)}{2},$$

where the last inequality follows from inequality (2.23). Therefore we infer that with probability $1 - o(\frac{1}{\log n})$, $|\mathcal{U}_t(z) \cap \mathcal{L}^*(z)| \geqslant \frac{L^*(z)}{2}$. □

Lemma 2.7.5. *Suppose that* $I_t = \mathrm{e}^{\Omega\left(\frac{\log n}{\log\log n}\right)}$ *and* $S_t \leqslant \frac{n}{\log^6 n}$*. Then whp, the* Push-Pull *protocol needs* $\mathcal{O}\left(\frac{\log n}{\log\log n}\right)$ *rounds to inform at least* $\mathrm{e}^{\log n - \frac{\log n}{\log\log n}}$ *nodes.*

Proof. Let X_u be an indicator random variable for every $u \in \mathcal{U}_t(z) \cap \mathcal{L}^*(z)$ so that

$$X_u := \begin{cases} 1 & \text{if } u \text{ gets informed by Pull in round } t+1, \\ 0 & \text{otherwise.} \end{cases}$$

Then we define the random variable $X_t(z) := \sum_{u \in \mathcal{U}_t(z) \cap \mathcal{L}^*(z)} X_u$. Let us define $z_t = \min\{I_t^{1/4}, \frac{n}{I_t \cdot \log^6 n}\}$. Using the approximation $\mathrm{e}^{-2\cdot a} \leqslant 1 - a \leqslant \mathrm{e}^{-a}$, $0 \leqslant a \leqslant 1/2$, we know that for any $z \leqslant z_t$ we have

$$\mathbf{Pr}\left[X_u = 1\right] = 1 - \left(1 - \frac{I_t}{n}\right)^z \geqslant 1 - \mathrm{e}^{-\frac{I_t \cdot z}{n}} \geqslant \frac{I_t \cdot z}{2 \cdot n},$$

Applying Lemma 2.7.4 shows that with probability $1 - o(\frac{1}{\log n})$ for any $z \leqslant z_t$ and any round $t = \mathcal{O}(\log n)$,

$$\mathbf{E}\left[X_t(z)\right] = \sum_{u \in \mathcal{U}_t(z) \cap \mathcal{L}^*(z)} \mathbf{Pr}\left[X_u = 1\right] > \frac{L^*(z) \cdot I_t \cdot z}{4 \cdot n} \geqslant \frac{I_t \cdot z \cdot \mathbf{Pr}\left[R = z\right]}{8} \geqslant \frac{c \cdot I_t}{I_t^{\frac{3}{4}}}, \tag{2.25}$$

where the last inequality holds because $\mathbf{Pr}\left[R = z\right] \geqslant \frac{c}{z^3}$. Since $I_t = e^{\Omega\left(\frac{\log n}{\log\log n}\right)}$ and X_u's are independent, applying a Chernoff bound (see e.g. Theorem 1.3.4) implies that with probability $1 - o(\frac{1}{n})$,

$$X_t(z) \geqslant \frac{\mathbf{E}\left[X_t(z)\right]}{2}.$$

Using the above inequality and inequality (2.25) shows that with probability $1 - o(\frac{1}{\log n})$ there exists a constant C so that

$$S_{t+1} \geqslant \sum_{z=1}^{z_t} X_t(z) \cdot z \geqslant \frac{I_t}{16} \sum_{z=1}^{z_t} z^2 \cdot \mathbf{Pr}\left[R = z\right] \geqslant \frac{c \cdot I_t}{16} \sum_{z=1}^{z_t} \frac{1}{z} = I_t \cdot C \cdot \log z_t.$$

For any positive integer k such that $I_{t+k} \in [\mathrm{e}^{\Omega\left(\frac{\log n}{\log\log n}\right)}, \mathrm{e}^{\log n - \frac{\log n}{\log\log n}}]$, we have that $\mathrm{e}^{\Omega\left(\frac{\log n}{\log\log n}\right)} \leqslant z_t$. Hence from the above inequality we conclude that here exists a constant C_1 so that

$$S_{t+1} \geqslant C_1 \cdot I_t \cdot \frac{\log n}{\log\log n} \geqslant C_1 \cdot I_t \cdot \sqrt{\log n}.$$

Considering only Push transmission for $S_t = o(n)$ and applying Lemma 2.3.2 implies that with probability $1 - o(\frac{1}{\log n})$

$$I_{t+2} \geqslant \frac{S_{t+1}}{2} \geqslant \frac{C_1 \cdot I_t \cdot \sqrt{\log n}}{2}$$

An inductive argument shows that for any integer k

$$I_{t+2k} \geqslant I_t \cdot \left(\frac{C_1 \cdot \sqrt{\log n}}{2}\right)^k .$$

as long as $S_{t+2k} = \frac{n}{\log^6 n}$ whp. Thus there is a $k = \mathcal{O}\left(\frac{\log n}{\log\log n}\right)$ so that after $t + 2k$ rounds there are at least $\mathrm{e}^{\log n - \frac{\log n}{\log\log n}}$ informed nodes. □

Corollary 2.7.6. *Let R be a power law distribution with $\beta = 3$. Then whp, the* Push-Pull *protocol informs all n nodes in $\mathcal{O}\left(\frac{\log n}{\log\log n}\right)$ rounds.*

Proof. Applying Corollary 2.3.4 shows that as long as $S_t = o(n)$ with probability $1 - o(1)$, for any round $t = \mathcal{O}(\log n)$,

$$I_t \geqslant \left(\frac{3}{2}\right)^t \cdot I_0.$$

So after $\mathcal{O}\left(\frac{\log n}{\log\log n}\right)$ rounds there are at least $\mathrm{e}^{\Omega\left(\frac{\log n}{\log\log n}\right)}$ informed nodes. Now we apply Lemma 2.7.5 and conclude that after $\mathcal{O}\left(\frac{\log n}{\log\log n}\right)$ rounds we have at least $\mathrm{e}^{\log n - \frac{\log n}{\log\log n}}$ informed nodes. Another application of Corollary 2.3.4 implies that after $\mathcal{O}(\frac{\log n}{\log\log n})$ rounds we have at least $\frac{n}{\log\log n}$ informed nodes. Since we have enough informed nodes, the result by Karp et. al in [KSSV00, Theorem 2.1] for the standard Push-Pull protocol can be applied to show that once $I_t \geqslant \frac{n}{\log n}$, whp additional $\mathcal{O}(\log\log n)$ rounds are sufficient to inform all n nodes. □

2.8. Generating a New C_u^t in Each Round

In this section we analyze the Push-Pull protocol for a new model. In this model according to some distribution R, at the beginning of each round t, every node u generates a random natural number $C_u^t \geqslant 1$ independent of all other nodes. Then in round t, the Push-Pull protocol disseminates the information according to $\{C_u^t \colon u \in \mathcal{V}\}$, i.e., node u calls C_u^t random nodes. We prove the following theorem regarding to this model.

Theorem 2.8.1. *Assume that R is any distribution with $\mathbf{E}[R] = \mathcal{O}(1)$. Then whp, the* Push-Pull *protocol needs $\Omega(\log n)$ rounds to inform all nodes.*

Proof. The probability that a node $u \in \mathcal{U}_t$ gets informed by Pull is as follows:

$$\begin{aligned}
&\mathbf{Pr}\left[u \text{ gets informed by } \mathsf{Pull} \text{ in round } t+1\right] \\
&= \sum_{x=1}^{\infty} \mathbf{Pr}\left[u \text{ gets informed by } \mathsf{Pull} \text{ in round } t+1 \mid R_u^{t+1} = x\right] \cdot \mathbf{Pr}\left[R_u^{t+1} = x\right] \\
&= \sum_{x=1}^{\lfloor \frac{n}{2I_t} \rfloor} \left(1 - \left(1 - \frac{I_t}{n}\right)^x\right) \cdot \mathbf{Pr}\left[R_u^{t+1} = x\right] + \sum_{x=\lfloor \frac{n}{2I_t} \rfloor + 1}^{\infty} \left(1 - \left(1 - \frac{I_t}{n}\right)^x\right) \cdot \mathbf{Pr}\left[R_u^{t+1} = x\right] \\
&\leqslant \frac{I_t}{n} \sum_{x=1}^{\lfloor \frac{n}{2I_t} \rfloor} x \cdot \mathbf{Pr}\left[R_u^{t+1} = x\right] + \sum_{x=\lfloor \frac{n}{2I_t} \rfloor + 1}^{\infty} \mathbf{Pr}\left[R_u^{t+1} = x\right] \qquad \left(\text{since } 1 - \left(1 - \frac{I_t}{n}\right)^x \leqslant \frac{I_t \cdot x}{n}\right) \\
&\leqslant \frac{I_t}{n} \cdot \mathbf{E}[R] + \mathbf{Pr}\left[R_u^{t+1} > \left\lfloor \frac{n}{2I_t} \right\rfloor\right] \leqslant \frac{I_t}{n} \cdot \mathbf{E}[R] + \frac{2I_t}{n} \cdot \mathbf{E}[R],
\end{aligned}$$

where the last inequality follows from Markov's inequality. Recall that N_t^{Pull} and N_t^{Push} are the number of newly informed nodes by Pull and Push calls in round $t+1$ respectively. Hence,

$$\begin{aligned}
\mathbf{E}\left[N_t^{\mathsf{Pull}} \mid I_t\right] &= \sum_{u \in \mathcal{U}_t} \mathbf{Pr}\left[u \text{ gets informed by } \mathsf{Pull} \text{ in round } t+1\right] \\
&\leqslant \frac{U_t \cdot I_t \cdot 3 \cdot \mathbf{E}[R]}{n} < 3 \cdot I_t \cdot \mathbf{E}[R].
\end{aligned}$$

Recall that S_t is the number of Push calls by informed nodes in round $t+1$. Therefore, $N_t^{\mathsf{Push}} \leqslant S_t$ and

$$\mathbf{E}\left[N_t^{\mathsf{Push}} \mid I_t\right] \leqslant \mathbf{E}\left[S_t \mid I_t\right] = \sum_{u \in I_t} \mathbf{E}\left[C_u^{t+1}\right] = I_t \cdot \mathbf{E}[R].$$

Hence,

$$\mathbf{E}\left[I_{t+1} \mid I_t\right] \leqslant I_t + \mathbf{E}\left[N_t^{\mathsf{Pull}} \mid I_t\right] + \mathbf{E}\left[N_t^{\mathsf{Push}} \mid I_t\right] \leqslant (1 + 4 \cdot \mathbf{E}[R]) \cdot I_t.$$

By using the law of total expectation, we conclude that $\mathbf{E}[I_t] < (1 + 4 \cdot \mathbf{E}[R])^t$. If we set $T = c \cdot \log n$, where $c > 0$ is a small constant, then

$$\mathbf{Pr}\left[I_T \geqslant \sqrt{n}\right] \leqslant \frac{\mathbf{E}[I_T]}{\sqrt{n}} \leqslant \frac{(1 + 4 \cdot \mathbf{E}[R])^T}{\sqrt{n}} = o(1).$$

So with probability $1 - o(1)$, we need at least $c \cdot \log n$ rounds to inform all nodes. $\square$

3 Randomized Rumor Spreading in Poorly Connected Networks

Randomized rumor spreading is an important primitive for information dissemination in networks and has numerous applications in network science, ranging from spreading information in the WWW and Twitter to spreading viruses and diffusion of ideas in human communities (see [CLP09, DFF11, DFF12a, DFF12b, FPS12]). A well studied rumor spreading protocol is the *Push-Pull protocol*, introduced by Demers et al. [DGG+87]. Suppose that one node in a network is aware of a piece of information, the 'rumor.' The protocol proceeds in rounds. In each round, every *informed* node contacts a random neighbor and sends the rumor to it ('pushes' the rumor), and every *uninformed* nodes contacts a random neighbor and gets the rumor if the neighbor possibly knows it ('pulls' the rumor). Note that this is a synchronous protocol, e.g. a node that receives a rumor in a certain round can only forward it in the next round.

A point to point communication network can be modeled as an undirected graph: the nodes represent the processors and the links represent communication channels between the nodes. Studying rumor spreading has several applications to distributed computing in such networks, of which we mention just two. The first is in broadcasting algorithms: a single processor wants to broadcast a piece of information to all other processors in the network (see [HHL88] for a survey). There are at least three advantages to the Push-Pull protocol: it is simple (each node makes a simple local decision in each round; no knowledge of the global topology is needed; no state is maintained), scalable (the protocol is independent of the size of network: it does not grow more complex as the network grows) and robust (the protocol tolerates random node/link failures without the use of error recovery mechanisms, see [FPRU90]). A second application comes from the maintenance of databases replicated at many sites, e.g., yellow pages, name servers, or server directories. After introducing updates to a few nodes, these should be propagate to all nodes in the network. In each round, a processor communicates with a random neighbor and they share any new updates, so that eventually all copies of the database store to the same contents (see [DGG+87] for details). Other than the aforementioned applications, rumor spreading protocols have successfully been applied in various contexts such as resource discovery [HBLL99], load balancing [BGPS06], data aggregation [KDG03], and the spread of computer viruses [BBCS05].

We only consider simple, undirected and connected graphs. For a graph G, let $\Delta(G)$ and $\mathrm{diam}(G)$ denote the maximum degree and the diameter of G, respectively, and let $\deg(v)$ denote the degree of a vertex v. Most studies in randomized rumor spreading focus on the *runtime* of this protocol, defined as the number of rounds taken until a rumor initiated by one vertex reaches all other vertices. It is clear that $\mathrm{diam}(G)$ is a lower bound for the runtime of the protocol. Feige et al. [FPRU90] showed that for an n-vertex G, the rumor reaches all vertices in $\mathcal{O}(\Delta(G)\cdot(\mathrm{diam}(G)+\log n))$ rounds whp. This protocol has been studied on many graph classes such as complete graphs [KSSV00], Erdös-Réyni random graphs [Els06, FPRU90, FHP10], random regular graphs [BEF08, FP10], and hypercubes [FPRU90]. For most of these classes it turns out that whp the runtime is $\mathcal{O}(\mathrm{diam}(G)+\log n)$, which does not depend on the maximum degree.

Randomized rumor spreading has recently been studied on real-world network models. Doerr, Fouz, and Friedrich [DFF11] proved an upper bound of $\mathcal{O}(\log n)$ whp for the runtime on preferential attachment graphs, and Fountoulakis, Panagiotou, and Sauerwald [FPS12] proved the same upper bound (up to constant factors) for the runtime on the giant component of random graphs with given expected degrees (also known as the Chung-Lu model) with power law degree distribution.

The runtime is closely related to the *expansion profile* of the graph. Let $\Phi(G)$ and $\alpha(G)$ denote the conductance and the vertex expansion of a graph G, respectively. After a series of results by various authors, Giakkoupis [Gia14, Gia11] showed that for any n-vertex graph G, the runtime of the Push-Pull protocol is $\mathcal{O}\left(\min\{\Phi(G)^{-1}\cdot\log n, \alpha(G)^{-1}\cdot\log^2 n\}\right)$. It is known that whp preferential attachment graphs and random graphs with given expected degrees have conductance $\Omega(1)$ (see [CLV03, MPS03]) whp. So it is not surprising that rumors spread fast on these graphs. Censor-Hillel, Haeupler, Kelner, and Maymounkov [CHHKM12] presented a more elaborate rumor spreading protocol that distributes the rumor in $\mathcal{O}(\mathrm{diam}(G)+\mathrm{polylog}(n))$ rounds on any connected n-vertex graph whp. Since this bound does not involve any conductance, it is particularly suitable for poorly connected graphs.

Our Results. We analyze the behavior of the Push-Pull protocol on random k-trees and random k-Apollonian networks, which are small-world networks with power law degrees and have large clustering coefficients. Random k-trees are a class of evolving random networks. Initially our graph is a single k-clique. In every step a new node is born, a random k-clique of the current graph is chosen, and the new node is joined to all nodes of the k-clique. The construction of k-Apollonian networks is slightly different in the sense that whenever a k-clique is chosen it is never chosen again. The definition of random k-trees enjoys a 'the rich get richer' effect, as in the preferential attachment scheme (i.e. the probability that the new vertex attaches to v is proportional to the degree of v). On the other hand, random k-trees have much larger clustering coefficients than preferential attachment graphs, as all neighbors of each new vertex are joined to each other. It is well-known that real-world networks tend to have large clustering coefficients (see, e.g., [WS98, Table 1]).

When $k \geqslant 2$ is fixed, we show that if initially a random node is aware of the rumor, then whp after $\mathcal{O}\left((\log n)^{1+\frac{2}{k}} \cdot \log\log n \cdot f(n)\right)$ rounds the rumor propagates to $n - o(n)$ nodes, where n is the number of nodes and $f(n)$ is any slowly growing function. Since these graphs have polynomially small conductance, vertex expansion $\mathcal{O}(1/n)$ and constant *treewidth* (will be defined in next section), these results demonstrate that Push-Pull can be efficient even on poorly connected networks. On the negative side, we prove that with probability $1 - o(1)$ the protocol needs at least $\Omega\left(n^{(k-1)/(k^2+k-1)}/f^2(n)\right)$ rounds to inform all nodes. This exponential dichotomy between the times required for informing *almost all* and *all* nodes is striking. Our main contribution is to present, for the first time, a natural class of random graphs in which such a phenomenon can be observed. The former implies that if one wishes to inform almost all the vertices, then one only has to wait for a polylogarithmic number of rounds. The latter implies that, however, if one wishes to inform each and every vertex, then one has to wait for polynomially many rounds. The main contribution of this chapter is to present, for the first time, a natural class of random graphs in which this dichotomy can be observed. Using similar techniques we show an upper bound of $\mathcal{O}\left((\log n)^{(k^2-3)/(k-1)^2} \cdot \log\log n \cdot f(n)\right)$ rounds for fixed $k \geqslant 3$, for informing $n - o(n)$ nodes of a random k-Apollonian whp. In fact, in many applications, such as epidemics, viral marketing and voting, it is more appealing to inform 99 percent of the vertices very quickly instead of waiting a long time until everyone gets informed. It is worth mentioning that bounds for the number of rounds to inform *almost all* vertices have already appeared in the literature, see for instance [DFF12a, FPS12]. In particular, for power-law Chung-Lu graphs with exponent $\in (2,3)$, it is shown in [FPS12] that whp after $\mathcal{O}(\log\log n)$ rounds the rumor spreads in $n - o(n)$ vertices, but to inform *all* vertices of the giant component $\Theta(\log n)$ rounds are necessary and sufficient. This result also shows a great difference between the two cases, however in both cases the required time is logarithmically small.

Techniques. To derive an upper bound on the runtime of the protocol on random k-trees, we show that there exist low degree vertices, called *efficient connectors*, facilitating the communication between the high degree vertices. The concept of efficient connectors appeared previously in several papers such as [DFF11] and [FPS12]. To show the existence of efficient connectors, we divide the building process of a random k-tree into two phases, the vertices born before and after round m where $m = o(n)$ is a suitably chosen parameter. We first expose a random k-tree G up to round m and call it $G(m)$. Then we consider the connected components of $G - G(m)$. Most vertices born later than round m have relatively small degree, so most of these components have a small maximum degree (and logarithmic diameter) thus the rumor spreads quickly within each of them. A vertex $v \in V(G(m))$ typically has a large degree, and this means that there is a high chance that v has a neighbor x with small degree, which quickly receives the rumor from v and spreads it (or vice versa). We build an almost-spanning tree T of $G(m)$ with logarithmic height, such that for every edge $\{u, v\}$ of T, one of u and v have a small degree, or u and v have a common neighbor with a small degree. Either of these conditions implies the rumor is exchanged quickly between u and v. This tree T then works as a 'highway system' to spread the rumor among the vertices of $G(m)$ and from them to the components of $G - G(m)$. Our technique for proving the upper bound successfully carries over to a closely related class of graphs, the random k-Apollonian networks.

To prove a polynomial lower bound for the runtime, we define the notion of *barrier* in a graph which is a subset D of edges of size $\mathcal{O}(1)$, whose deletion disconnects the graph and both endpoints of every edge of a barrier D have very large degrees. It is clear that if a graph contains a barrier, then the protocol needs a very large time to pass the rumor through D. Hence it only remains to show a random k-tree has a barrie whp which is done in Section 3.5.

Outline. In Section 3.1 we formally define our random graph models which will be based on generalized urn models. We then apply some useful facts from urn theory and prove some results about the degree sequence of these random graphs. In Sections 3.2 and 3.3 we study basic properties of the networks such as degree sequence and expansion profile of the networks. In Sections 3.4 and 3.5 we show an upper bound on the number of rounds that is required to inform almost all nodes and a lower bound on the number of rounds to inform all nodes by the Push-Pull protocol on random k-trees respectively. In Section 3.6 we present similar results for the Push-Pull protocol on random k-Apollonian networks. Finally, in Section 3.7 we show that the considered networks exhibit several fundamental properties of real-world networks including small-world property, large clustering coefficient and power law degree sequence.

3.1. Definitions, Notations and Preliminaries

In this section we formally define random k-trees, random k-Apollonian networks and the notion of treewidth.

Definition 3.1.1 (Random k-tree process [Gao09]). *Let k be a positive integer. Build a sequence $G_k(0)$, $G_k(1)$, $\ldots$ of random graphs as follows. The graph $G_k(0)$ is just a clique on k vertices. For each $1 \leqslant t \leqslant n$, $G_k(t)$ is obtained from $G_k(t-1)$ as follows: a k-clique of $G_k(t-1)$ is chosen uniformly at random, a new vertex is born and is joined to all vertices of the chosen k-clique. The graph $G_k(n)$ is called a random k-tree on $n+k$ vertices. If k is clear from the text, it is denoted by $G(n)$.*

We remark that this process is different from the random k-tree process defined by Cooper and Uehara [CU10], where in each round a $(k+1)$-clique is chosen uniformly at random and then the new node is connected to k nodes of the chosen clique. This process was further studied in [CF13].

Sometimes it is convenient to view this as a 'random graph evolving in time.' In this interpretation, in every round $1, 2, \ldots,$ a new vertex is born and is added to the evolving graph, and $G(t)$ denotes the graph at the end of round t. Observe that $G(t)$ has $k+t$ many vertices and $kt+1$ many k-cliques.

Gao [Gao09] showed that whp the degree sequence of $G(n)$ asymptotically follows a power law distribution with exponent $2 + \frac{1}{k-1}$. In Subsection 3.7.1 we show that whp the diameter of $G(n)$ is $\mathcal{O}(\log n)$, and its clustering coefficient is at least $1/2$, as opposed to preferential attachment graphs and random graphs with given expected degrees, whose clustering coefficients are $o(1)$ whp. As per these properties, random k-trees serve as more realistic models for real-world networks.

Sometimes it is convenient to view this as a 'random graph evolving in time.' In this interpretation, in every round $1, 2, \ldots$, a new vertex is born and is added to the evolving graph, and $G_k(t)$ denotes the graph at the end of round t. Observe that $G_k(t)$ has $k + t$ many vertices and $kt + 1$ many k-cliques.

A closely related class of graphs is the class of *random k-Apollonian networks*, introduced by Zhang, Comellas, Fertin, and Rong [ZCFR06]. Their construction is very similar to the construction of random k-trees, with just one difference: if a k-clique is chosen in a certain round, it will never be chosen again. Let us define it formally as follows:

Definition 3.1.2 (Random k-Apollonian process [ZCFR06]). *Let k be a positive integer. Build a sequence $A_k(0)$, $A_k(1)$, $\ldots$ of random graphs as follows. The graph $A_k(0)$ is just a clique on k vertices. For each $1 \leqslant t \leqslant n$, $A_k(t)$ is obtained from $A_k(t-1)$ as follows: a k-clique of $A_k(t-1)$ which has not yet picked, called active, is chosen uniformly at random, a new vertex is born and is joined to all vertices of the chosen k-clique. The graph $A_k(n)$ is called a random k-Apollonian on $n + k$ vertices.*

It is known that whp random k-Apollonian networks exhibit a power law degree distribution and large clustering coefficient [ZYW05, Mun11] and have logarithmic diameter [CF13].

Observe that similar to $G_k(n)$, in every round k different k-cliques are added to the network but since in each round one k-clique is chosen and becomes inactive, the number of active k-cliques increases by $k-1$. The embedding of random k-Apollonian networks in $(k-1)$-dimensional space is a distinguishing feature of these family of graphs. In order to have a geometric view of the network, let us think of $A_k(n)$ as a $(k-1)$-dimensional simplex with an additional vertex in its interior which is connected to all other vertices and hence there are k $(k-1)$-dimensional simplices that we call them active, so in the following round we pick one of the active simplices and place the new vertex in its interior and connect the vertex to all vertices of the simplex. Clearly this construction gives a $(k-1)$-dimensional embedding of the network. For $k = 3$, the network is a known as RANs and the above construction gives a triangulated planar graphs (for instance see Figure 3.1).

Definition 3.1.3 (Chapter 2, [Klo94]). *A tree decomposition of a graph $G = (V, E)$ is a tree let say T whose nodes $X_1, X_2, \ldots, X_l$ are subsets of $V(G)$ satisfying following conditions:*

1. $\cup_i^l X_i = V(G)$.

2. *Suppose that $u \in X_i \cap X_j$, then every X_k which belongs to the path from X_i to X_j in T contains u as well.*

3. *For every $\{u, v\} \in E(G)$, there exists at least one $X_i \in V(T)$ so that $\{u, v\} \subset X_i$*

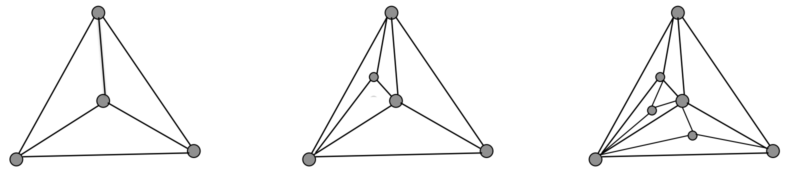

Figure 3.1. RAN process for $n = 1, 2$, and 4 .

The width of a tree decomposition is the size of the largest X_i minus one. The treewidth of a graph is the minimum width among all its tree decompositions.

Fact 3.1.4. *Random k-trees and k-Apollonian networks have treewidth k.*

Proof. We describe a tree decomposition of random k-trees and k-Apollonian networks. We define the set of all $(k+1)$-cliques contained in the network as the vertex set of the tree decomposition and two nodes are connected if and only if they have exactly k common elements. By definition we see that random k-trees and k-Apollonian networks have treewidth at most k. On the other hand the treewidth of every clique is the same as its size minus one. Therefore both networks have treewidth exactly k. □

3.2. Some Results from the Urn Theory

In this subsection we introduce the Polya and generalized Polya urn models and present some known results from urn theory. For more information about urn models and their applications we refer the interested reader to [JK77].

Definition 3.2.1 (Pólya-Eggenberger urn)**.** *Start with W_0 white and B_0 black balls in an urn. In every step a ball is drawn from the urn uniformly at random, the ball is returned to the urn, and s balls of the same color are added to the urn. Let* $\mathrm{Polya}(W_0, B_0, s, n)$ *denote the distribution of the number of white balls right after n draws.*

Proposition 3.2.2. *Let* $X = \mathrm{Polya}(a, b, k, n)$*,* $w = a + b$ *and let* $c \geqslant (a+b)/k$*. Then* $\mathbf{Pr}\left[X = 0\right] \leqslant \left(\frac{c}{c+n}\right)^{a/k}$ *and*

$$\mathbf{E}\left[X^2\right] = \left(a + \frac{a}{w}\,kn\right)^2 + \frac{abk^2 n(kn+w)}{w^2(w+k)} \,.$$

Proof. The expected value and the variance of X are well known (see [Mah03, Corollary 5.1.1] for instance):

$$\begin{aligned}
\mathbf{E}[X] &= a + \frac{a}{w}\,kn\,,\\
\mathbf{Var}[X] &= \frac{abk^2n(kn+w)}{w^2(w+k)}\,.
\end{aligned}$$

For the last inequality, we have

$$\begin{aligned}
\mathbf{Pr}[X=0] &= \frac{b}{a+b}\cdot\frac{b+k}{a+b+k}\cdot\dots\cdot\frac{b+(n-1)k}{a+b+(n-1)k}\\
&= \prod_{i=0}^{n-1}\left(1-\frac{a}{a+b+ik}\right) \leqslant \prod_{i=0}^{n-1}\left(1-\frac{a}{ck+ik}\right)\\
&\leqslant \exp\left(-\sum_{i=0}^{n-1}\frac{a}{ck+ik}\right) = \exp\left(\sum_{i=0}^{n-1}\frac{1}{c+i}\right)^{-a/k}\\
&\leqslant \exp\left(\int_{x=c}^{c+n}\frac{\mathrm{d}x}{x}\right)^{-a/k} = \exp\left(\log((c+n)/c)\right)^{-a/k} = \left(\frac{c}{c+n}\right)^{a/k}\,. \square
\end{aligned}$$

Definition 3.2.3 (Generalized Pólya-Eggenberger urn)**.** *Let $\alpha,\beta,\gamma,\delta$ be nonnegative integers. We start with W_0 white and B_0 black balls in an urn. In every step a ball is drawn from the urn uniformly at random and returned to the urn. Additionally, if the ball is white, then δ white balls and γ black balls are returned to the urn; otherwise, i.e. if the ball is black, then β white balls and α black balls are returned to the urn. Let* $\mathrm{Polya}\left(W_0, B_0, \begin{bmatrix}\alpha & \beta\\ \gamma & \delta\end{bmatrix}, n\right)$ *denote the distribution of the number of white balls right after n draws.*

Note that Pólya-Eggenberger urns correspond to the matrix $\begin{bmatrix}s & 0\\ 0 & s\end{bmatrix}$. The following proposition follows from known results.

Proposition 3.2.4. *Let $X = \mathrm{Polya}\left(W_0, B_0, \begin{bmatrix}\alpha & 0\\ \gamma & \delta\end{bmatrix}, n\right)$ and let r be a positive integer. If $\gamma,\delta > 0$, $\alpha = \gamma+\delta$, and $r\delta \geqslant \alpha$, then we have*

$$\mathbf{E}[X^r] \leqslant \left(\frac{\alpha n}{W_0+B_0}\right)^{r\delta/\alpha}\prod_{i=0}^{r-1}(W_0+i\delta) + \mathcal{O}\left(n^{(r-1)\delta/\alpha}\right)\,.$$

Proof. It is known that

$$\mathbf{E}[X^r] = n^{r\delta/\alpha}\delta^r\frac{\Gamma(W_0/\delta+r)\Gamma((W_0+B_0)/\alpha)}{\Gamma(W_0/\delta)\Gamma((W_0+B_0+r\delta)/\alpha)} + \mathcal{O}\left(n^{(r-1)\delta/\alpha}\right)\,,$$

see [FDP06, Proposition 15] for instance. Note that

$$\frac{\Gamma(W_0/\delta+r)}{\Gamma(W_0/\delta)} = \prod_{i=0}^{r-1}(i+W_0/\delta)\,.$$

Finally, the inequality

$$\frac{\Gamma((W_0+B_0+r\delta)/\alpha)}{\Gamma((W_0+B_0)/\alpha)} \geqslant ((W_0+B_0)/\alpha)^{r\delta/\alpha}$$

follows from $r\delta \geqslant \alpha$ and the following inequality (see, e.g., [Laf84, equation (2.2)])

$$x^{1-s} \leqslant \frac{\Gamma(x+1)}{\Gamma(x+s)} \qquad \forall x > 0, s \in [0,1] .$$

□

3.2.1. Degree Sequence of the Networks

In this subsection we first make a connection between the degrees sequence of the networks and generalized Polya urn models. Then by applying some results from the previous subsection we show a useful lemma and corollary to find a probabilistic bound for the degree of an arbitrary vertex which is born in a certain round. In this subsection we assume that $H(j)$ be from one of the two networks (i.e. random k-trees or random k-Apollonian networks) which is built up to round $j+1$.

Proposition 3.2.5. *Suppose that x be a vertex of $H(j)$ and x has $N > 0$ neighbors, and is contained in B many (active) k-cliques. Conditional on this, if $H(j)$ is a random k-tree process, then the degree of x in $H(n+j)$ is distributed as*

$$N + \left(\text{Polya}\left(B, kj+1-B, \begin{bmatrix} k & 0 \\ 1 & k-1 \end{bmatrix}, n\right) - B\right) \Big/ (k-1) .$$

If $H(j)$ is an k-Apollonian process, then the degree of x in $H(n+j)$ is distributed as

$$N + \left(\text{Polya}\left(B, (k-1)j+1-B, \begin{bmatrix} k-1 & 0 \\ 1 & k-2 \end{bmatrix}, n\right) - B\right) \Big/ (k-1) .$$

Proof. We claim that Polya $\left(B, kj+1-B, \begin{bmatrix} k & 0 \\ 1 & k-1 \end{bmatrix}, n\right)$ is the total number of k-cliques containing x in $G_k(n+j)$. At the end of round j, there are B many k-cliques containing x, and $kj+1-B$ many k-cliques not containing x. In each subsequent round $j+1, \ldots, j+n$, a random k-clique is chosen and k new k-cliques are created. If the chosen k-clique contains x, then $k-1$ new k-cliques containing x are created, and 1 new k-clique not containing x is created. Otherwise, i.e. if the chosen k-clique does not contain x, then no new k-cliques containing x is created, and k new k-cliques not containing x are created. Hence Polya $\left(B, kj+1-B, \begin{bmatrix} k & 0 \\ 1 & k-1 \end{bmatrix}, n\right)$ is exactly the total number of k-cliques containing x in $H(n+j)$. Hence the number of k-cliques that are created in rounds $j+1, \ldots, j+n$ and contain x is Polya $\left(B, kj+1-B, \begin{bmatrix} k & 0 \\ 1 & k-1 \end{bmatrix}, n\right) - B$, and the proof follows by noting that every new neighbor of x creates $k-1$ new k-cliques containing x. When $H(j)$ is a k-Apollonian process, then in each subsequent round $j+1, \ldots, j+n$ the number of active k-cliques and those containing x increases by $k-1$ and $k-2$ respectively and hence the number of active k-cliques that are created in rounds $j+1,\dots,j+n$ and contain x is Polya $\left(B,(k-1)j+1-B,\begin{bmatrix}k-1 & 0\\ 1 & k-2\end{bmatrix},n\right)-B$ and the rest follows similar to the case $H(j)$ is a random k-tree. $\square$

Lemma 3.2.6. *Let $1\leqslant j\leqslant n$ and let q be a positive integer. Let x denote the vertex born in round j. Conditional on any $H(j)$, if $H(j)$ is a random k-tree process then the probability that x has degree greater than $k+q(n/j)^{(k-1)/k}$ in $H(n)$ is $\mathcal{O}\left(q\sqrt{q}\exp(-q)\right)$. If $H(j)$ is a k-Apollonian process, then the probability that x has degree greater than $k+q(n/j)^{(k-2)/(k-1)}$ is $\mathcal{O}\left(q\sqrt{q}\exp(-q)\right)$.*

Proof. Let $X=$ Polya $\left(k,kj-k+1,\begin{bmatrix}k & 0\\ 1 & k-1\end{bmatrix},n-j\right)$. By Proposition 3.2.5, $\deg(x)$ is distributed as $k+(X-k)/(k-1)$. By Proposition 3.2.4,

$$\mathbf{E}\left[X^q\right]\leqslant(1+o(1))\left(\frac{k(n-j)}{kj+1}\right)^{\frac{q(k-1)}{k}}\prod_{i=0}^{q-1}(k+i(k-1))\leqslant\left(\frac{n}{j}\right)^{\frac{q(k-1)}{k}}(k-1)^q(q+1)!\,.$$

Thus,

$$\begin{aligned}\mathbf{Pr}\left[\deg(x)>k+q(n/j)^{(k-1)/k}\right]&=\mathbf{Pr}\left[X-k>q(k-1)(n/j)^{(k-1)/k}\right]\\&\leqslant\frac{\mathbf{E}\left[X^q\right]}{\left(q(k-1)(n/j)^{(k-1)/k}\right)^q}\\&\leqslant(q+1)!q^{-q}=\mathcal{O}\left(q\sqrt{q}\exp(-q)\right)\,.\end{aligned}$$

For the second part, let $Y=$ Polya $\left(k,(k-1)j-k+1,\begin{bmatrix}k & 0\\ 1 & k-1\end{bmatrix},n-j\right)$. From Proposition 3.2.5, $\deg(x)$ is distributed as $k+(Y-k)/(k-1)$ and similar to previous case applying Proposition 3.2.4 implies that

$$\mathbf{Pr}\left[\deg(x)>k+q(n/j)^{(k-2)/(k-1)}\right]=\mathcal{O}\left(q\sqrt{q}\exp(-q)\right)\,.$$ $\square$

Corollary 3.2.7. *With high probability, the maximum degree of $G_k(n)$ and $A_k(n)$ is bounded by $\mathcal{O}(\log n\cdot n^{(k-1)/(k)})$ and $\mathcal{O}(\log n\cdot n^{(k-2)/(k-1)})$ respectively.*

Proof. Let $q=\lfloor 2\log n\rfloor$ and let x be a vertex born in one of the rounds $1,2,\dots,n$ of a random k-tree process. Then applying Lemma 3.2.6 yields that

$$\mathbf{Pr}\left[\deg(x)>k+qn^{1-1/k}\right]=\mathcal{O}(q\sqrt{q}\exp(-q))=o(1/n)\,.$$

A union bound over all vertices shows that whp we have $\Delta(G_k(n))\leqslant k+(2\log n)n^{(k-1)/k}$, as required. Applying similar argument for k-Apollonian process, we conclude that $\Delta(A_k(n))\leqslant k+(2\log n)n^{(k-2)/(k-1)}$. $\square$

3.3. Expansion of the Networks

In this section we prove that random k-trees and k-Apollonian do not expand well confirming our claim in the introduction that the networks are poorly connected graphs and thus existing techniques do not apply.

Definition 3.3.1. *The* vertex expansion *of a graph G (also known as the* vertex isoperimetric number *of G), written $\alpha(G)$, is defined as*

$$\alpha(G) = \min\left\{\frac{|\partial S|}{|S|} : S \subseteq V(G), 0 < |S| \leqslant |V(G)|/2\right\} ,$$

where ∂S denotes the set of vertices in $V(G) \setminus S$ that have a neighbor in S.

Definition 3.3.2. *The* conductance *of a graph G (also known as the* Cheeger constant *of G), written $\Phi(G)$, is defined as*

$$\Phi(G) = \min\left\{\frac{e(S, V(G) \setminus S)}{\mathtt{vol}(S)} : S \subseteq V(G), 0 < \mathtt{vol}(S) \leqslant \mathtt{vol}(V(G))/2\right\} ,$$

where $e(S, V(G) \setminus S)$ denotes the number of edges between S and $V(G) \setminus S$, and $\mathtt{vol}(S) = \sum_{u \in S} \deg(u)$ for every $S \subseteq V(G)$.

Proposition 3.3.3. *Whp $G(n)$ has vertex expansion $\mathcal{O}(k/n)$, and its conductance is $\mathcal{O}\left(\log n \cdot n^{-1/k}\right)$.*

Proof. Let $G = G(n)$. Since G has treewidth k, by [Klo94, Lemma 5.3.1] there exists a partition (A, B, C) of $V(G)$ such that

1. $|C| = k + 1$,
2. $(n-1)/3 \leqslant |A|, |B| \leqslant 2(n-1)/3$, and
3. there is no edge between A and B.

At least one of A and B, say A, has size less than $(n+k)/2$. Then

$$\alpha(G) \leqslant \frac{|\partial A|}{|A|} \leqslant \frac{k+1}{(n-1)/3} = \mathcal{O}(k/n) .$$

At least one of A and B, say B, has volume less than $\mathtt{vol}(G)/2$. Then since all vertices in G have degrees at least k,

$$\Phi(G) \leqslant \frac{e(B, A \cup C)}{\mathtt{vol}(B)} \leqslant \frac{e(B, C)}{k|B|} \leqslant \frac{(k+1)\Delta(G)}{k(n-1)/3} = \mathcal{O}(\Delta(G)/n) .$$

Hence to prove $\Phi(G) = \mathcal{O}\left(\log n \cdot n^{-1/k}\right)$ it is enough to show that whp we have

$$\Delta(G) \leqslant k + (2\log n)n^{1-1/k} .$$

Let $q = \lfloor 2\log n \rfloor$ and let x be a vertex born in one of the rounds 1 to n. By Lemma 3.2.6,

$$\mathbf{Pr}\left[\deg(x) > k + qn^{1-1/k}\right] = \mathcal{O}(q\sqrt{q}\exp(-q)) = o(1/n) .$$

An argument similar to the proof of Lemma 3.2.6 shows that the probability that a vertex in $G(0)$ has degree greater than $k + qn^{1-1/k}$ is $o(1/n)$ as well. A union bound over all vertices shows that whpwe have $\Delta(G) \leqslant k + (2\log n)n^{1-1/k}$, as required. □

3.4. Push-Pull Protocol on Random k-Trees

In this section we analyze the Push-Pull protocol on random k-trees $G(n)$. The main result of this section is the following theorem:

Theorem 3.4.1. *Let $k \geqslant 2$ be constant and let $f(n) = o(\log\log n)$ be an arbitrary function going to infinity with n. If initially a random vertex of an $(n+k)$-vertex random k-tree knows a rumor, then whp after $\mathcal{O}\left((\log n)^{1+\frac{2}{k}} \cdot \log\log n \cdot f(n)\right)$ rounds of the* Push-Pull *protocol, $n - o(n)$ vertices will know the rumor.*

Once we have the following lemma, the proof of the theorem reduces to proving a structural result for random k-trees.

Lemma 3.4.2. *Let G be an n-vertex graph and let $\Sigma \subseteq V(G)$ with $|\Sigma| = n - o(n)$ be such that for every pair of vertices $u, v \in \Sigma$ there exists a (u,v)-path $uu_1u_2\dots u_{l-1}v$ such that $l \leqslant \chi$ and for every $0 \leqslant i \leqslant l-1$ we have $\min\{\deg(u_i), \deg(u_{i+1})\} \leqslant \tau$ (where we define $u_0 = u$ and $u_l = v$). If a random vertex in G knows a rumor, then whp after $6\tau(\chi + \log n)$ rounds of the* Push-Pull *protocol, at least $n - o(n)$ vertices will know the rumor.*

Proof. We show that given any $u, v \in \Sigma$, if u knows the rumor then with probability at least $1 - o\left(n^{-2}\right)$ after $6\tau(\chi + \log n)$ rounds v will know the rumor. The lemma follows by using the union bound and noting that a random vertex lies in Σ with high probability. Consider the (u,v)-path $uu_1u_2\dots u_{l-1}v$ promised by the hypothesis. Using a similar argument as [FPRU90], we bound from below the probability that the rumor is passed through this path. For every $0 \leqslant i \leqslant l-1$, the number of rounds needed for the rumor to pass from u_i to u_{i+1} is a geometric random variable with success probability at least $1/\tau$ (if $\deg(u_i) \leqslant \tau$, this is the number of rounds needed for u_i to push the rumor along the edge, and if $\deg(u_{i+1}) \leqslant \tau$, this is the number of rounds needed for u_{i+1} to pull the rumor along the edge). The random variables corresponding to distinct edges are mutually independent. Hence the probability that the rumor is not passed in $6\tau(\chi + \log n)$ rounds is at most the probability that the number of heads in a sequence of $6\tau(\chi + \log n)$ independent biased coin flips, each having probability $1/\tau$ of being heads, is less than l. Let X denote the number of heads in such a sequence. Then using the Chernoff bound (see, e.g., Theorem 1.3.4) and noting that $\mathbf{E}[X] = 6(\chi + \log n)$ we get

$$\mathbf{Pr}[X < l] \leqslant \mathbf{Pr}[X \leqslant \mathbf{E}[X]/6] \leqslant \exp(-(5/6)^2\,\mathbf{E}[X]/2) \leqslant \exp(-(5/6)^2(6\log n)/2),$$

which is $o\left(n^{-2}\right)$, as required. □

So what remains is to show the following structural result for random k-trees.

Lemma 3.4.3. *Let G be an $(n+k)$-vertex random k-tree. Then whp there exists $\Sigma \subseteq V(G)$ with $|\Sigma| = n - o(n)$ such that for every pair of vertices $u, v \in \Sigma$ there exists a (u,v)-path $uu_1u_2\dots u_{l-1}v$ where $l = \mathcal{O}(\log n + \operatorname{diam}(G))$ and for every $0 \leqslant i \leqslant l-1$ we have $\min\{\deg(u_i), \deg(u_{i+1})\} \leqslant \tau$.*

Let

$$m = \left\lceil \frac{n}{f(n)^{\frac{3}{k-1}} (\log n)^{\frac{2}{k-1}}} \right\rceil .$$

Also let $q = \lceil 4 \log\log n \rceil$ and let

$$\tau = 2k + q(n/m)^{1-1/k} . \tag{3.1}$$

For the rest of this section, G is an $(n+k)$-vertex random k-tree. Recall from Definition 3.1.1 that $G = G(n)$, where $G(0), G(1), \dots,$ is the random k-tree process. Consider the graph $G_1 = G(m)$, which has $k+m$ vertices and $mk+1$ many k-cliques. For an edge e of G_1, let $N(e)$ denote the number of k-cliques of $G(m)$ containing e. We define a spanning forest F of $G(m)$ as follows: initially F is the vertex set of $G(0)$, then for every $1 \leqslant t \leqslant m$, if the vertex x born in round t is joined to the k-clique C, then in F, x is joined to a vertex $u \in V(C)$ such that

$$N(xu) = \max_{v \in V(C)} N(xv) .$$

Note that F has k trees and the k vertices of $G(0)$ lie in distinct trees. Think of these trees as rooted at these vertices. The tree obtained from F by merging these k vertices is the 'highway system' described in the introduction.

Informally speaking, the proof is divided into three parts: first, we show that this tree has a small height (Lemma 3.4.4), second, we show that each edge in this tree quickly exchanges the rumor with a reasonably large probability (Lemma 3.4.6), and finally we show that almost all vertices in $G - G(m)$ have quick access to and from F (Lemma 3.4.7). Let LOG denote the event 'each tree in F has height $\mathcal{O}(\log n)$.

Lemma 3.4.4. *Whp* LOG *happens.*

Proof. We inductively define the notion of *draft* for vertices and k-cliques of $G(m)$. The draft of the vertices of $G(0)$ as well as the k-clique they form equals 0. The draft of every k-clique equals the maximum draft of its vertices. Whenever a new vertex is born and is joined to a k-clique, the draft of the vertex equals the draft of the k-clique plus one. It is easy to see that if $\{x, y\} \in E(G(m))$ and x is born later than y, then $\text{draft}(x) \geqslant \text{draft}(y) + 1$. In particular, if x is a vertex of F with distance h to the root, then $\text{draft}(x) \geqslant h$. Hence we just need to show that whp the draft of each k-clique is $\mathcal{O}(\log n)$. We define an auxiliary tree whose vertices are the k-cliques of $G(m)$. Start with a single vertex corresponding to $G(0)$. Whenever a new vertex x is born and is joined to a k-clique C, k new k-cliques are created. In the auxiliary tree, add these to the set of children of C. The depth of each k-clique in this auxiliary tree equals its draft as defined above. The height of this auxiliary tree is stochastically less than or equal to the height of a random k-ary recursive tree (see [Drm09, Section 1.3.3] for the definition), whose height is $\mathcal{O}(\log n)$ whp, as proved in [Drm09, Theorem 6.47]. □

The following deterministic lemma shows an lower bound for $N(x, y)$, where $\{x, y\} \in E(F)$ which is useful for the next lemma.

Lemma 3.4.5. *Assume that $k > 2$ and $\{x, y\} \in E(F)$ and suppose that x is born later than y. If the degree of x in G_1 is greater than $2k - 2$, then $N(x, y) \geqslant (k^2 - k)/2$.*

Proof. Assume that x is joined to $u_1,\dots,u_k$ when it is born. Also assume that $v_1,v_2,\dots,v_{k-1},\dots$ are the neighbors of x that are born later than x, in the order of birth. Let Ψ denote the number of pairs (u_j,C), where $u_j\in V(G_1)$ and C is a k-clique in G_1 such that $\{x,u_j\}\subseteq V(C)$. Consider the round in which vertex x is born and is joined to $u_1,\dots,u_k$. For every $j\in\{1,\dots,k\}$, the vertex u_j is contained in $k-1$ new k-cliques, so in this round Ψ increases by $k(k-1)$. For each $i\in\{1,\dots,k-1\}$, consider the round in which vertex v_i is born. This vertex is joined to x and $k-1$ neighbors of x. At this round x has neighbor set $\{u_1,\dots,u_k,v_1,\dots,v_{i-1}\}$. Thus at least $k-i$ of the u_j's are joined to v_i in this round. Each vertex u_j that is joined to v_i in this round is contained in $k-2$ new k-cliques, so in this round Ψ increases by at least $(k-i)(k-1)$. Consequently, we have

$$\Psi\geqslant k(k-1)+\sum_{i=1}^{k-1}(k-i)(k-2)=k^2(k-1)k/2. \tag{3.2}$$

By the pigeonhole principle, there exists some $\ell\in\{1,\dots,k\}$ such that the edge xu_ℓ is contained in at least $(k^2-1)/2$ many k-cliques, and this completes the proof. $\square$

A vertex of G is called *young* if it is born later than the end of round m, and is called *old* otherwise. In other words, vertices of G_1 are old and vertices of $G-G_1$ are young. We say edge $uv\in E(G)$ is *fast* if at least one of the following is true: $\deg(u)\leqslant\tau$, or $\deg(v)\leqslant\tau$, or u and v have a common neighbor w with $\deg(w)\leqslant\tau$. For an edge $uv\in E(F)$, let $p_S(uv)$ denote the probability that uv is not fast, and let p_S denote the maximum of p_S over all edges of F.

Lemma 3.4.6. *We have $p_S=o(1/(f(n)\log n))$.*

Proof. Let $\{x,y\}\in E(F)$ be arbitrary. Without loss of generality we may assume that x is born later than y. First, suppose that $k>2$. By Lemma 3.4.5, at least one of the following is true: vertex x has less than $2k-1$ neighbors in G_1, or $N(x,y)\geqslant(k^2-k)/2$. So we may consider two cases.

- Case 1: vertex x has less than $2k-1$ neighbors in G_1. In this case vertex x lies in at most k^2-2k+2 many k-cliques of G_1. Assume that x has A neighbors in G_1 and lies in B many k-cliques in G_1. Let

  $$X=\text{Polya}\left(B,km+1-B,\begin{bmatrix}k&0\\1&k-1\end{bmatrix},n-m\right).$$

  Then by Proposition 3.2.5 the degree of x is distributed as $A+(X-B)/(k-1)$. By Proposition 3.2.4,

  $$\begin{aligned}\mathbf{E}[X^q]&\leqslant(1+o(1))\left(\frac{k(n-m)}{km+1}\right)^{\frac{q(k-1)}{k}}\prod_{i=0}^{q-1}(B+i(k-1))\\&\leqslant(1+o(1))\left(\frac{n}{m}\right)^{\frac{q(k-1)}{k}}(k-1)^q\prod_{i=0}^{q-1}(k+i)\leqslant(k-1)^q(k+q)!\left(\frac{n}{m}\right)^{\frac{q(k-1)}{k}},\end{aligned}$$

where we have used $B \leqslant k(k-1)$ for the second inequality. Therefore,

$$\mathbf{Pr}\left[\deg(x) > 2k + q(n/m)^{\frac{k-1}{k}}\right] \leqslant \mathbf{Pr}\left[X \geqslant (k-1)q(n/m)^{\frac{k-1}{k}}\right]$$

$$\leqslant \frac{\mathbf{E}\left[X^q\right]}{(k-1)^q q^q (n/m)^{\frac{q(k-1)}{k}}} = \mathcal{O}\left(\frac{(k+q)^{k+q}\sqrt{q}}{q^q \exp(k+q)}\right) = o\left(\frac{1}{f(n)\log n}\right).$$

- Case 2: $N(x,y) \geqslant (k^2-k)/2$. In this case we bound from below the probability that there exists a young vertex w that is adjacent to x and y and has degree at most τ. We first bound from above the probability that x and y have no young common neighbors. For this to happen, none of the k-cliques containing x and y must be chosen in rounds $m+1,\ldots,n$. This probability equals $\mathbf{Pr}\left[\text{Polya}(N(x,y), mk+1-N(x,y), k, n-m) = 0\right]$. Since $N(x,y) \geqslant (k^2-1)/2$, by Proposition 3.2.2 we have

$$\mathbf{Pr}\left[\text{Polya}(N(x,y), mk+1-N(x,y), k, n-m) = 0\right] \leqslant \left(\frac{m+1}{n+1}\right)^{\frac{k-1}{2}} = o\left(\frac{1}{f(n)\log n}\right).$$

  Now, assume that x and y have a young common neighbor w. If there are multiple such vertices, choose the one that is born first. Since w appears later than round m, by Lemma 3.2.6,

$$\mathbf{Pr}\left[\deg(w) > k + q(n/m)^{(k-1)/k}\right] = \mathcal{O}\left(q\sqrt{q}\exp(-q)\right) = o\left(\frac{1}{f(n)\log n}\right).$$

The proof for $k=2$ is very similar to the argument for Case 2 above: note that in this case we have $N(x,y) \geqslant 1$ for all edges $\{x,y\} \in E(F)$, and we have

$$\mathbf{Pr}\left[\text{Polya}(1, 2m, 2, n-m) = 0\right] \leqslant \sqrt{\frac{m+1}{n+1}} = \mathcal{O}\left(\sqrt{\frac{m}{n}}\right) = o\left(\frac{1}{f(n)\log n}\right). \qquad \square$$

A old vertex is called *nice* if it is connected to some vertex in $G(0)$ via a path of fast edges. Since F has height $\mathcal{O}(\log n)$ and each edge of F is fast with probability at least $1-p_S$, the probability that a given old vertex is not nice is $\mathcal{O}(p_S \log n)$ by the union bound. A piece H_j is called *nice* if all its young vertices have degrees at most τ, and the vertex r_j is nice. A young vertex is called *nice* if it lies in a nice piece. A vertex/piece is called *bad* if it is not nice.

Lemma 3.4.7. *The expected number of bad vertices is $o(n)$.*

Proof. The total number of old vertices is $k+m=o(n)$ so we may just ignore them in the calculations below. Let $\eta = nf(n)/m = o(\log^3 n)$. Say piece H_j is *sparse* if $|V(H_j)| \leqslant \eta + k$. We first bound the expected number of young vertices in non-sparse pieces. Observe that the number of young vertices in each piece is distributed as $X = (\text{Polya}(1, km, k, n-m)-1)/k$. Using Proposition 3.2.2 we get $\mathbf{E}\left[X^2\right] \leqslant 2kn^2/m^2$. By the second moment method, for every $t>0$ we have

$$\mathbf{Pr}\left[X \geqslant t\right] \leqslant \frac{\mathbf{E}\left[X^2\right]}{t^2} \leqslant \frac{2kn^2}{m^2t^2}.$$

The expected number of young vertices in non-sparse pieces is thus at most

$$\sum_{i=0}^{\infty}(2^{i+1}\eta)(km+1)\,\mathbf{Pr}\left[2^{i}\eta < X \leqslant 2^{i+1}\eta\right] \leqslant \sum_{i=0}^{\infty}(2^{i+1}\eta)(km+1)\frac{2kn^2}{m^2\eta^2 2^{2i}}$$
$$\leqslant \mathcal{O}\left(\frac{n^2}{m\eta}\right)\sum_{i=0}^{\infty}2^{-i} = \mathcal{O}\left(\frac{n^2}{m\eta}\right) = o(n)\,.$$

We now bound the expected number of young vertices in sparse bad pieces. For bounding this value from above we find an upper bound for the expected number of sparse bad pieces, and multiply it by η. A piece H_j can be bad in either two ways:

(1) the representative vertex r_j is bad: the probability of this is $\mathcal{O}\left(p_S \log n\right)$. The expected number of pieces that are bad due to this reason is thus $\mathcal{O}\left(mkp_S \log n\right)$, which is $o(n/\eta)$ by Lemma 3.4.6.

(2) there exists a young vertex in H_j with degree greater than τ: the probability that a given young vertex has degree greater than τ is $\mathcal{O}\left(q\sqrt{q}\exp(-q)\right)$ by Lemma 3.2.6. So the average number of young vertices with degree greater than τ is $\mathcal{O}\left(nq\sqrt{q}\exp(-q)\right)$. Since every young vertex lies in a unique piece, the expected number of pieces that are bad because of this reason is $\mathcal{O}\left(nq\sqrt{q}\exp(-q)\right) = o(n/\log^3 n)$.

So the expected number of bad pieces is $o(n/\eta + n/\log^3 n)$. The expected number of young vertices in sparse bad pieces is thus $o(n + \eta n/\log^3 n) = o(n)$. □

Enumerate the k-cliques of G_1 as $C_1,\dots,C_{mk+1}$. Choose $r_1 \in C_1,\dots,\ r_{mk+1} \in C_{mk+1}$ arbitrarily, and call them the *representative vertices*. Starting from G_1, when young vertices are born in rounds $m+1,\dots,n$ until G is formed, every clique C_i 'grows' to a random k-tree with a random number of vertices, which is a subgraph of G. Enumerate these subgraphs as $H_1,\dots,H_{mk+1}$, and call them the *pieces*. More formally, $H_1,\dots,H_{mk+1}$ are induced subgraphs of G such that a vertex v is in $V(H_j)$ if and only if every path connecting v to a old vertex intersects $V(C_j)$. In particular, $V(C_j) \subseteq V(H_j)$ for all $j \in \{1,\dots,mk+1\}$. Note that the H_j's may intersect as a old vertex may lie in more than one C_j, however every young vertex lies in a unique piece.

We now have all the ingredients to prove Lemma 3.4.3, which concludes the proof of Theorem 3.4.1.

Proof of Lemma 3.4.3. Let Σ denote the set of nice young vertices. By Lemma 3.4.7 and using Markov's inequality, we have $|\Sigma| = n - o(n)$ whp. Let $\{a_1,\dots,a_k\}$ denote the vertex set of $G(0)$. Using an argument similar to the proof of Lemma 3.4.6, it can be proved that given $1 \leqslant i < j \leqslant k$, the probability that edge a_ia_j is not fast is $o(1)$. Since the total number of such edges is a constant, whp all such edges are fast. Let u and v be nice young vertices, and let r_u and r_v be the representative vertices of the pieces containing them, respectively. Since the piece containing u is nice, there exists a (u,r_u)-path whose vertices except possibly r_u all have degrees at most τ. The length of this path is at most $\mathrm{diam}(G)$. Since r_u is nice, for some $1 \leqslant i \leqslant n$ there exists an (r_u,a_i)-path in F consisting of fast edges. By appending these paths we find a (u,a_i)-path with length at most $\mathrm{diam}(G) + \mathcal{O}(\log n)$ such that for every pair of consecutive vertices in this path, one of them has degree at most τ. Similarly, for some $1 \leqslant j \leqslant n$ there exists a (v,a_j)-path of length $\mathcal{O}(\log n + \mathrm{diam}(G))$, such that one of every pair of consecutive vertices in this path has degree at most τ. Since the edge $a_i a_j$ is fast, we can build a (u,v)-path of length $\mathcal{O}(\log n + \mathrm{diam}(G))$ of the type required by the statement of the lemma. $\square$

3.5. A Lower Bound

In this section we show that whpthe Push-Pull protocol needs at least $n^{\Omega(1)}$ many rounds to inform all nodes of a random k-tree $G(n)$. Before we start showing the result of this section let us define the notion of *s-barrier*.

Definition 3.5.1 (s-barrier)**.** *A pair $\{C_1, C_2\}$ of disjoint k-cliques in a connected graph is an s-*barrier *if (i) the set of edges between C_1 and C_2 is a cut-set, i.e. deleting them disconnects the graph, and (ii) the degree of each vertex in $V(C_1) \cup V(C_2)$ is at least s.*

Observe that if G has an s-barrier, then for any starting vertex, whp the Push-Pull protocol needs at least $\Omega(s)$ rounds to inform all vertices.

Lemma 3.5.2. *The graph $G(n)$ has an $\Omega(n^{1-1/k})$-barrier with probability $\Omega(n^{1/k-k})$.*

Proof. Let $u_1, \ldots, u_k$ be the vertices of $G(0)$, and let $v_1, \ldots, v_k$ be the vertices of $G(k) - G(0)$ in the order of appearance. We define two events: Event A is that for every $1 \leqslant i \leqslant k$, when v_i appears, it attaches to $v_1, v_2, \ldots, v_{i-1}, u_i, u_{i+1}, \ldots, u_k$; and for each $1 \leqslant i, j \leqslant k$, u_i and v_j have no common neighbor in $G(n) - G(k)$. Event B is that all vertices of $G(k)$ have degree $\Omega(n^{(k-1)/k})$ in $G(n)$. Note that if A and B both happen, then the pair $\{u_1 u_2 \ldots u_k, v_1 v_2 \ldots v_k\}$ is an $\Omega(n^{(k-1)/k})$-barrier in $G(n)$. Hence to prove the lemma it suffices to show $\mathbf{Pr}[A] = \Omega(n^{1/k-k})$ and $\mathbf{Pr}[B|A] = \Omega(1)$.

For A to happen, first, the vertices $v_1, \ldots, v_k$ must choose the specific k-cliques, which happens with constant probability. Moreover, the vertices appearing after round k must not choose any of the $k^2 - 1$ many k-cliques that contain both u_i's and v_j's. Since $1 - y \geqslant e^{-y-y^2}$ for every $y \in [0, 1/4]$,

$$\begin{aligned}
\mathbf{Pr}[A] &= \mathbf{Pr}\left[\mathrm{Polya}(k^2-1, 2, k, n-k) = 0\right] \\
&= \Omega\left(\prod_{i=0}^{n-k-1}\left(\frac{2+ik}{k^2+1+ik}\right)\right) \\
&\geqslant \Omega\left(\prod_{i=0}^{4k-1}\left(\frac{2+ik}{k^2+1+ik}\right)\prod_{i=4k}^{n-k-1}\left(1-\frac{k^2-1}{ik}\right)\right) \\
&\geqslant \Omega\left(\exp\left(-\sum_{i=4k}^{n-k-1}\left\{\frac{k^2-1}{ik}+\left(\frac{k^2-1}{ik}\right)^2\right\}\right)\right)
\end{aligned}$$

which is $\Omega(n^{1/k-k})$ since $\sum_{i=4k}^{n-k-1} \frac{k^2-1}{ik} \leqslant (k-1/k)\log n + \mathcal{O}(1)$ and $\sum_{i=4k}^{n-k-1}\left(\frac{k^2-1}{ik}\right)^2 = \mathcal{O}(1)$.

Conditional on A and using an argument similar to that in the proof of Proposition 3.2.5, the degree of each of the vertices $u_1, \ldots, u_k, v_1, \ldots, v_k$ in $G(n)$ is at least

$k + (\text{Polya}(1,1,\begin{bmatrix} k & 0 \\ 1 & k-1 \end{bmatrix}, n-k) - 1)/(k-1)$. By [FDP06, Proposition 16], there exists $\delta > 0$ such that

$$\mathbf{Pr}\left[\text{Polya}(1,1,\begin{bmatrix} k & 0 \\ 1 & k-1 \end{bmatrix}, n-k) < \delta n^{(k-1)/k}\right] < 1/(2k+1)\,.$$

By the union bound, the probability that all vertices $u_1, \ldots, u_k, v_1, \ldots, v_k$ have degrees at least $\delta n^{(k-1)/k}/(k-1)$ is at least $1/(2k+1)$, hence $\mathbf{Pr}[B|A] \geqslant 1/(2k+1) = \Omega(1)$. □

Let $f(n) = o(\log \log n)$ be a function going to infinity with n, and let $m = \left\lceil f(n) n^{1-k/(k^2+k-1)} \right\rceil$. (Note that the value of m is different from that in Section 3.4, although its role is somewhat similar.) Consider the random k-tree process up to round m. Enumerate the k-cliques of $G(m)$ as $C_1, \ldots, C_{mk+1}$. Starting from $G(m)$, when new vertices are born in rounds $m+1, \ldots, n$ until $G = G(n)$ is formed, every clique C_i 'grows' to a random k-tree with a random number of vertices, which is a subgraph of G. Enumerate these subgraphs as $H_1, \ldots, H_{mk+1}$, and call them the *pieces*. We say a piece is *moderate* if its number of vertices is between $n/(mf(n))$ and $nf(n)/m$. Note that the number of vertices in a piece has expected value $\Theta(n/m)$. The following lemma is proved by showing that this random variable does not deviate too much from its expected value.

Lemma 3.5.3. *With high probability, there are $o(m)$ non-moderate pieces.*

Proof. We prove that the first piece, H_1, is moderate whp. By symmetry, this would imply that the average number of non-moderate pieces is $o(m)$. By Markov's inequality, this gives that whp there are $o(m)$ non-moderate pieces. Let X denote the number of vertices of H_1. Note that X is distributed as $k + \text{Polya}(1, km, k, n-m)$; so its expected value is $k + \frac{n-m}{1+km} = \Theta(n/m)$. So by Markov's inequality, $\mathbf{Pr}[X > nf(n)/m] = o(1)$. For bounding $\mathbf{Pr}[X < n/(mf(n))]$, we use an alternative way to define the random variable $\text{Polya}(1, km, k, n-m)$ (see [JK77, page 181]): assume Z is a beta random variable with parameters $1/k$ and m. Then $X - k$, which has the same distribution as $\text{Polya}(1, km, k, n-m)$, is distributed as a binomial random variable with parameters $n - m$ and Z. Note that

$$\begin{aligned} \mathbf{Pr}[Z < 3/(mf(n))] &= \frac{\Gamma(m+1/k)}{\Gamma(m)\Gamma(1/k)} \int_0^{3/(mf(n))} x^{1/k-1}(1-x)^{m-1} dx \\ &< \frac{m^{1/k}}{\Gamma(1/k)} \int_0^{3/(mf(n))} x^{1/k-1} dx = 3^{1/k} k/(\Gamma(1/k) f(n)^{1/k}) = o(1)\,, \end{aligned}$$

where we have used the fact $\Gamma(m+1/k) < \Gamma(m) m^{1/k}$ which follows from [Laf84, equation (2.2)]. On the other hand, the Chernoff bound (see, e.g., [MR95, Theorem 4.2]) gives

$$\begin{aligned} \mathbf{Pr}[X < n/(mf(n)) | Z \geqslant 3/(mf(n))] &\leqslant \mathbf{Pr}[\text{Bin}(n-m, 3/(mf(n))) < n/(mf(n))] \\ &< \exp(-3(n-m)/(8mf(n))) = o(1)\,, \end{aligned}$$

thus $\mathbf{Pr}[X < n/(mf(n))] = o(1)$. □

Theorem 3.5.4. *Let $f(n) = o(\log\log n)$ be an arbitrary function going to infinity with n. Suppose that initially one vertex in the random k-tree, $G(n)$, knows a rumor. With high probability, the* Push-Pull *protocol needs at least $n^{(k-1)/(k^2+k-1)} f(n)^{-2}$ rounds to inform all vertices of $G(n)$.*

Proof. Consider an alternative way to generate $G(n)$ from $G(m)$: first, we determine how many vertices each piece has, and then we expose the structure of the pieces. Let Y denote the number of moderate pieces. By Lemma 3.5.3 we have $Y = \Omega(m)$ whp. We prove the theorem conditional on $Y = y$, where $y = \Omega(m)$ is otherwise arbitrary. Note that after the sizes of the pieces are exposed, what happens inside each piece in rounds $m+1, \ldots, n$ is mutually independent from other pieces. Let H be a moderate piece with n_1 vertices. By Lemma 3.5.2, the probability that H has an $\Omega(n_1^{1-1/k})$-barrier is $\Omega(n_1^{1/k-k})$. Since $n/(mf(n)) \leqslant n_1 \leqslant nf(n)/m$, the probability that H has a $\Omega((n/(mf(n)))^{1-1/k})$-barrier is $\Omega((nf(n)/m)^{1/k-k})$. Since there are $y = \Omega(m)$ moderate pieces in total, the probability that no moderate piece has a $\Omega\left((n/(mf(n)))^{1-1/k}\right)$-barrier is at most $(1 - \Omega((nf(n)/m)^{1/k-k}))^y \leqslant \exp(-\Omega(f(n))) = o(1)$, which means whp there exists an $\Omega\left(n^{(k-1)/(k^2+k-1)} f(n)^{-2}\right)$-barrier in $G(n)$, as required. $\square$

3.6. Push-Pull Protocol on Random k-Apollonian Networks

In this section we analyze the Push-Pull protocol on a random k-Apollonian network $A_k(n)$. Since these networks are a sub-family of random k-trees, we reuse the proof techniques in Section 3.4 and find an upper bound for for the number of rounds needed to inform almost all vertices of $A_k(n)$. The main result of this section is the following theorem:

Theorem 3.6.1. *Let $k \geqslant 3$ be constant and let $f(n) = o(\log\log n)$ be an arbitrary function going to infinity with n. If initially a random vertex of a random k-Apollonian knows a rumor, then whp after $\mathcal{O}\left((\log n)^{\frac{k^2-3}{(k-1)^2}} \cdot \log\log n \cdot f(n)\right)$ rounds of the* Push-Pull *protocol, $n - o(n)$ vertices will know the rumor.*

Before proving the above theorem let us define some parameters and show several lemmas. Fix $k > 2$ and let $f(n) = o(\log\log n)$ be an arbitrary function going to infinity with n, and let

$$m = \left\lceil \frac{n}{f(n)^{(2k-2)/(k^2-2k)} (\log n)^{2/k-1}} \right\rceil$$

Finally, let $q = \lceil 4\log\log n \rceil$ and let

$$\tau = 2k + q(n/m)^{(k-2)/(k-1)} \,. \tag{3.3}$$

The proof of Theorem 3.6.1 follows from the following structural result, which we prove in the rest of this section.

Lemma 3.6.2. *Let A be an $(n+k)$-vertex k-RAN. Whp there exists $\Sigma \subseteq V(A)$ satisfying the conditions of Lemma 3.4.3 with τ defined in (3.3) and $\chi = \mathcal{O}(\log n + \mathrm{diam}(A))$.*

The proof of Lemma 3.6.2 is along the lines of that of Lemma 3.4.3. For the rest of this section, $A = A(n)$ is an $(n+k)$-vertex k-RAN. Consider the graph $A(m)$, which has $k+m$ vertices and $m(k-1)+1$ active k-cliques. For any edge e of $A(m)$, let $N^*(e)$ denote the number of active k-cliques of $A(m)$ containing e. Note that, since $k > 2$, for each edge e, the number of active k-cliques containing e does not decrease as the k-RAN evolves. We define a spanning forest F of $A(m)$ as follows: at round 0, F has k isolated vertices, i.e. the vertices of $A(0)$; then for every $1 \leqslant t \leqslant m$, if the vertex x born in round t is joined to the k-clique C, then in F, x is joined to a vertex $u \in V(C)$ such that

$$N^*(xu) = \max_{v \in V(C)} N^*(xv) \, .$$

Note that F has k trees and the k vertices of $A(0)$ lie in distinct trees. Let LOG denote the event 'each tree in F has height $\mathcal{O}(\log n)$.'

Lemma 3.6.3. *Whp* LOG *happens.*

Proof. The proof is very similar to that of Proposition 3.7.1, the only difference being that the built auxiliary tree is indeed a random k-ary recursive tree, whose height is bounded by $\mathcal{O}(\log n)$ whp. □

We prove Lemma 3.6.2 conditional on the event LOG. In fact, we prove it for any $A(m)$ that satisfies LOG. Let A_1 be an arbitrary instance of $A(m)$ that satisfies LOG. All randomness in the following refers to rounds $m+1, \dots, n$. The following deterministic lemma will be used in the proof of Lemma 3.6.5.

Lemma 3.6.4. *Assume that $xy \in E(F)$ and x is born later than y. If the degree of x in A_1 is at least $2k-1$, then $N^*(x,y) \geqslant (k-1)(k-3)/2$.*

Proof. Assume that x is joined to $u_1, \dots, u_k$ when it is born, and that $v_1, v_2, \dots, v_{k-1}, \dots$ are the neighbors of x that are born later than x, in the order of birth. Let Ψ denote the number of pairs (u_j, C), where C is an active k-clique in G_1 such that $\{x, u_j\} \subseteq V(C)$. Consider the round in which vertex x is born and is joined to $u_1, \dots, u_k$. For every $j \in \{1, \dots, k\}$, the vertex u_j is contained in $k-2$ new active k-cliques, and one k-clique containing u_j becomes deactivated. so in this round Ψ increases by $k(k-3)$. For each $i \in \{1, \dots, k-1\}$, consider the round in which vertex v_i is born. At least $k-i$ of the u_j's are joined to v_i in this round. Each vertex u_j that is joined to v_i in this round is contained in $k-2$ new k-cliques, and one k-clique containing u_j becomes deactivated. Hence in this round Ψ increases by at least $(k-i)(k-3)$. Consequently, right after v_{k-1} is born, we have

$$\Psi \geqslant k(k-1) + \sum_{i=1}^{k-1} (k-i)(k-3) = (k-1)^2 k/2 \, .$$

By the pigeonhole principle, there exists some $\ell \in \{1, \dots, k\}$ such that the edge xu_ℓ is contained in at least $(k-1)^2/2$ active k-cliques, and this completes the proof, as the number of active k-cliques containing $\{x, u_\ell\}$ will not decrease later. □

A vertex of A is called *young* if it is born later than the end of round m, and is called *old* otherwise. In other words, vertices of A_1 are old and vertices of $A - A_1$ are young. We say edge $\{u, v\} \in E(A)$ is *fast* if at least one of the following is true: $\deg(u) \leqslant \tau$, or $\deg(v) \leqslant \tau$, or u and v have a common neighbor w with $\deg(w) \leqslant \tau$. For an edge $\{u, v\} \in E(F)$, let $p_S(\{u, v\})$ denote the probability that $\{u, v\}$ is not fast, and let p_S denote the maximum of p_S over all edges of F.

Lemma 3.6.5. *We have $p_S = o(1/(f(n) \log n))$.*

Proof. The proof is similar to that of Lemma 3.4.6. Let $\{x, y\} \in E(F)$ be arbitrary. By symmetry we may assume that x is born later than y. First, assume that $k > 3$. By Lemma 3.6.4, at least one of the following is true: vertex x has less than $2k - 1$ neighbors in A_1, or $N^*(x, y) \geqslant (k - 1)^2/2$. So we may consider two cases.

- Case 1: vertex x has less than $2k - 1$ neighbors in A_1. In this case vertex x lies in at most $k + (k - 2)^2$ many active k-cliques of A_1. Suppose that x has D neighbors in A_1 and lies in B many active k-cliques in A_1. Let

$$X = \text{Polya}\left(B, (k-1)m + 1 - B, \begin{bmatrix} k-1 & 0 \\ 1 & k-2 \end{bmatrix}, n - m\right) .$$

Then by an argument similar to the proof of Proposition 3.2.5, the degree of x is distributed as $D + (X - B)/(k - 2)$. By Proposition 3.2.4,

$$\begin{aligned} \mathbf{E}[X^q] &\leqslant (1 + o(1)) \left(\frac{(k-1)(n-m)}{(k-1)m+1}\right)^{\frac{q(k-2)}{k-1}} \prod_{i=0}^{q-1} (B + i(k-2)) \\ &\leqslant \mathcal{O}\left(\left(\frac{n}{m}\right)^{\frac{q(k-2)}{k-1}} (k-2)^q (k+q)!\right) , \end{aligned}$$

where we have used $B \leqslant k(k - 2)$. Therefore,

$$\begin{aligned} &\mathbf{Pr}\left[\deg(x) > 2k + q(n/m)^{\frac{k-2}{k-1}}\right] \leqslant \mathbf{Pr}\left[X \geqslant (k-2)q(n/m)^{\frac{k-2}{k-1}}\right] \\ &\leqslant \frac{\mathbf{E}[X^q]}{(k-2)^q q^q (n/m)^{\frac{q(k-2)}{k-1}}} = \mathcal{O}\left(\frac{(k+q)!}{q^q}\right) = o\left(\frac{1}{f(n)\log n}\right) . \end{aligned}$$

- Case 2: $N^*(x, y) \geqslant (k-1)^2/2$. In this case we bound from below the probability that there exists a young vertex w that is adjacent to x and y and has degree at most τ. We first bound from above the probability that x and y have no young common neighbors. For this to happen, none of the k-cliques containing x and y must be chosen in rounds $m + 1, \ldots, n$. This probability equals

$$p := \mathbf{Pr}\left[\text{Polya}(N^*(x, y), m(k-1) + 1 - N^*(x, y), k - 1, n - m) = N^*(x, y)\right] .$$

Since $N^*(x, y) \geqslant (k - 1)^2/2$, by Proposition 3.2.2 we have

$$p \leqslant \left(\frac{m+1}{n}\right)^{\frac{k-1}{2}} = o\left(\frac{1}{f(n)\log n}\right) .$$

Now, assume that x and y have a young common neighbor w. If there are multiple such vertices, choose the one that is born first. Since w appears later than round m, by Lemma 3.2.6,

$$\mathbf{Pr}\left[\deg(w) > k + q(n/m)^{(k-2)/(k-1)}\right] = \mathcal{O}\left(q\sqrt{q}\exp(-q)\right) = o\left(\frac{1}{f(n)\log n}\right).$$

The proof for $k = 3$ is very similar to the argument for Case 2 above: note that in this case we have $N^*(x, y) \geqslant 2$ for all edges $\{x, y\} \in E(F)$, and we have

$$\mathbf{Pr}\left[\text{Polya}(2, 2m-1, 2, n-m) = 2\right] \leqslant \frac{m+1}{n} = o\left(\frac{1}{f(n)\log n}\right). \qquad \square$$

Enumerate the k-cliques of A_1 as C_1, C_2, ..., and $C_{m(k-1)+1}$. Then choose $r_1 \in C_1, \ldots, r_{m(k-1)+1} \in C_{m(k-1)+1}$ arbitrarily, and call them the *representative vertices*. Starting from A_1, when young vertices are born in rounds $m+1, \ldots, n$ until A is formed, every clique C_i 'grows' to a k-RAN with a random number of vertices, which is a subgraph of A. Enumerate these subgraphs as $H_1, \ldots, H_{m(k-1)+1}$, and call them the *pieces*. More formally, $H_1, \ldots, H_{m(k-1)+1}$ are induced subgraphs of A such that a vertex v is in $V(H_j)$ if and only if every path connecting v to a old vertex intersects $V(C_j)$.

A old vertex is called *nice* if it is connected to some vertex in $A(0)$ via a path of fast edges. Since F has height $\mathcal{O}(\log n)$ and each edge of F is fast with probability at least $1 - p_S$, the probability that a given old vertex is not nice is $\mathcal{O}(p_S \log n)$ by the union bound. A piece H_j is called *nice* if all its young vertices have degrees at most τ, and the vertex r_j is nice. A young vertex is called *nice* if it lies in a nice piece. A vertex/piece is called *bad* if it is not nice.

Lemma 3.6.6. *The expected number of bad vertices is $o(n)$.*

Proof. The proof is very similar to that of Lemma 3.4.7, except we use Lemmas 3.2.6 and 3.6.5 instead of Lemmas 3.2.6 and 3.4.6, respectively. $\square$

The proof of Lemma 3.6.2 is exactly the same as that of Lemma 3.4.3, except we use Lemmas 3.6.5 and 3.6.6 instead of Lemmas 3.4.6 and 3.4.7, respectively. This concludes the proof of Theorem 3.6.1.

3.7. Real-World Properties of the Networks

The study of real-world networks has exhibited three fundamental properties.

Small-world. This property, popularly known as six degree of separation, says that any two nodes of a real-world network are connected via small number of intermediate nodes or equivalently the network has small diameter, i.e. logarithmic in number of nodes.

Cluster Coefficient. Besides the small-world property, Watts and Strogatz [WS98] discovered that there is another fundamental feature of real-world networks called clustering coefficient, which is the probability of two random nodes with a common neighbor are connected to each other. They found out that in many networks this probability is a relatively large constant. (see, e.g., [WS98, Table 1]).

Scale Freeness. Albert and Barabasi [BA99] observed that the degree distribution of many real-world networks follows a power law distribution with exponent $2 < \beta \leqslant 3$. The authors in [BA99] introduced a model called *preferential attachment* (or briefly PA), to generate a evolving random graph model which is defined as follows. For any given fixed integer k, the model starts with a graph on a fixed number of nodes then in each round a new node is born and attached to k already present nodes with probability proportional to their degrees.

In what follows we show that our models fulfill all of the three properties above.

3.7.1. Diameter of the Networks

Proposition 3.7.1. *whp the diameter of $G_k(n)$ and $A_k(n)$ are $\mathcal{O}(\log n)$.*

Proof. There is a one-to-one correspondence between random k-trees process and a rooted random tree process $R_k(t)$ where in each round a node of the tree is chosen uniformly at random and k leaves append to the node. So if the root represents the initial clique in the random k-tree process, it is easy to see this one-to-one correspondence. Suppose that in round $1 \leqslant t \leqslant n$ in the random k-tree process, a k-clique is chosen, simultaneously corresponding node in $R_k(t-1)$ is chosen and gives birth to k children. Note that since in the Apollonian process every clique is chosen at most once, the corresponding random tree, say $R'_k(t-1)$ is a random k-ary tree. Therefore it is not hard to see that the depth of any node in $R_k(n)$ is equal to the distance of the corresponding clique in the network from the initial k-clique. Hence we have that $\mathrm{diam}(G_k(n))$ and $\mathrm{diam}(A_k(n))$ are bounded by double of the height of $R_k(n)$ and $R'_k(n)$, respectively. It remains to show that whp the height of our auxiliary random tree is bounded by $\mathcal{O}(\log n)$. Although the height of the auxiliary random tree process is an well-studied problem (see [Drm09, Section 1.3.3], for the sake of completeness we give a self-contained proof. Let us label the nodes of the tree according to the round they are born in. Note that nodes are born in the same round have the same label. Suppose that h_k be the height of $R_k(n)$, so there is a sequence of labeled nodes say $t_0, t_1, \ldots t_{h_k}$ where t_{i+1} is connected to t_i in some round $t_i < t \leqslant t_{i+1}$. Since in every round the number of nodes (k-cliques) increases by k, the probability that t_{i+1} is connected to t_i is at most $\sum_{t=t_i+1}^{t_{i+1}} \frac{1}{k \cdot t+1}$. Thus the probability that a given sequence $t_0 = 0, t_1, \ldots, t_{h_k}$ be a path of $R_k(n)$ from the root to some leaf is at most

$$\prod_{i=0}^{h_k-1} \sum_{t=t_i+1}^{t_{i+1}} \frac{1}{k \cdot t + 1}$$

Using the union bound over all possible sequence of length $h_k + 1$, we get

$$\sum_{t_0<t_1<\ldots<t_{h_k}} \prod_{i=0}^{h_k-1} \sum_{t=t_i+1}^{t_{i+1}} \frac{1}{k\cdot t+1} \leqslant \frac{1}{h_k!}\left(\sum_{t=1}^{n} \frac{1}{k\cdot t+1}\right)^{h_k} = \left(\frac{\mathrm{e}\cdot \log n}{h_k}\right)^{h_k}.$$

By setting $h_k = \mathcal{O}(\log n)$ we conclude that whp there is no sequence of length at least $10\log n$ from the root to some leaf. As a consequence the height is bounded by $\mathcal{O}(\log n)$. Note that same argument works for random k-ary tree with only one difference which is the number of nodes giving birth increases by $k-1$ and we have

$$\sum_{t_0<t_1<\ldots<t_{h'_k}} \prod_{i=0}^{h'_k-1} \sum_{t=t_i+1}^{t_{i+1}} \frac{1}{(k-1)\cdot t+1} \leqslant \frac{1}{h'_k!}\left(\sum_{t=1}^{n} \frac{1}{(k-1)\cdot t+1}\right)^{h'_k} = \left(\frac{\mathrm{e}\cdot \log n}{h'_k}\right)^{h'_k},$$

where h'_k is the height of a random k-ary tree. □

3.7.2. Clustering Coefficient of the Networks

Let us first give a formal definition of the clustering coefficient and then show that the clustering coefficient of the considered network models are positive constants. This is in stark cotrast to preferential attachment graphs and random graphs with given expected degrees, whose clustering coefficients are $o(1)$ whp (e.g. see [BR03]).

Definition 3.7.2. *The* clustering coefficient *of a graph G, written $cc(G)$, is defined as*

$$cc(G) = \frac{1}{|V(G)|} \sum_{u\in V(G)} \frac{|\langle N(u)\rangle|}{\binom{\deg(u)}{2}},$$

where $|\langle N(u)\rangle|$ denotes the number of edges $\{x,y\}$ such that both x and y are neighbors of u. Alternatively we can define the clustering coefficient of G as the ratio of the triple number of triangles to the number of pairs of adjacent edges of G.

The authors in [KKV13, Corollary 2. 8], show that the clustering coefficients of random k-Apollonians are strictly positive. In what follows we extend this result to random k-trees.

Proposition 3.7.3. *For every positive integer n, the clustering coefficient of $G_k(n)$ is at least 1/2.*

Proof. Let u be a vertex of $G(n)$. It is not hard to check that $|\langle N(u)\rangle| = (k-1)(\deg(u)-k/2)$, and since $\deg(u) \geqslant k$ we get

$$\frac{|\langle N(u)\rangle|}{\binom{\deg(u)}{2}} \geqslant \frac{k}{\deg(u)}.$$

Using the Cauchy-Schwarz inequality we get

$$cc(G) \geqslant \frac{1}{|V(G)|} \sum_{u\in V(G)} \frac{k}{\deg(u)} \geqslant \frac{k}{n+k}\cdot \frac{(n+k)^2}{2|E(G)|} = \frac{1}{2}.$$ □

3.7.3. Degree Distribution of the Networks

The definition of random k-trees and Apollonian networks enjoys a 'the rich get richer' effect, as in the preferential attachment scheme. Think of the number of k-cliques containing any vertex v as the 'wealth' of v (note that this quantity is linearly related to $\deg(v)$). Then, the probability that the new vertex attaches to v is proportional to the wealth of v, and if this happens, the wealth of v increases by $k-1$. Roughly speaking this phenomenon happens in many power law degree graphs. Gao [Gao09] showed that whp the degree sequence of $G(n)$ asymptotically follows a power law distribution with exponent $2+\frac{1}{k-1}$. Kollosvary et. al [KKV13] showed that the degree distribution of random k-Apollonian follows a power law distribution with exponent $2+\frac{1}{k-2}$ as well.

Part II

Random Walks on Graphs

4 Cutoff Phenomenon for Random Walks on Kneser Graphs

A simple random walk on a finite, non-bipartite graph is a discrete-time ergodic Markov chain, where in each time step the walker, located at some vertex, chooses one of its neighbors uniformly at random and moves to that neighbor. The *cutoff phenomenon* for a sequence of chains describes a sharp transition in the convergence of the chain distribution to its stationary distribution, over a negligible period of time, known as the *cutoff window*. (For a formal definition of cutoff phenomena, see Section 4.1). From a theoretical perspective, establishing a cutoff is often surprisingly challenging, even for simple chains, as it requires very tight bounds on the distribution near the mixing time. For applications such as MCMC, a cutoff is desirable, since running the chain any longer than the mixing time becomes essentially redundant.

Although it is widely believed that many natural families of Markov chains exhibit a cutoff, there are relatively few examples where the cutoff has been shown. In fact, it is quite challenging to prove or disprove the existence of a cutoff even for simple family of chains. The first results exhibiting a cutoff appeared in the studies of card-shuffling processes by Aldous and Diaconis [AD86], and Diaconis and Shahshahani [DS81]. Later, the cutoff phenomenon was also shown for random walks on hypercubes [DGM90], for random walks on distance regular graphs such as Johnson and Hamming graphs [Bel98, DS87], and for randomized riffle shuffles [CSC08b]. For a more general view of Markov chains with and without cutoffs, we refer the reader to [Dia96] or [LPW09, Chapter 18]. A necessary condition, known as product condition, for a family of chains to exhibit a cutoff is that $t^n_{mix}(1/4) \cdot \mathtt{gap}_n$ tends to infinity as n goes to infinity, where $\mathtt{gap}_n$ is the spectral gap of the transition matrix of the n-th chain (see [LPW09, Proposition 18.3].) However, there are some chains where the product condition holds and no cutoff is shown (e.g., see [LPW09, Section 18]), Peres [Per04] conjectured that many natural families of chains satisfying the product condition exhibit cutoffs. For instance, he conjectured that random walks on any family of n-vertex (transitive) expander graphs with $\mathsf{gap}_n = \Theta(1)$ and mixing time $\mathcal{O}(\log n)$ exhibit cutoffs. Chen and Saloff-Coste [CSC08a] verified the conjecture for other distances like the ℓ^p-norm for $p > 1$. Recently, Lubetzky and Sly [LS10] exhibited cutoff phenomena for random walks on random regular graphs. They also showed that there exist families of explicit expanders with and without cutoffs [LS11]. Diaconis [Dia96] pointed out that if the second largest eigenvalues of the transition matrix of a chain has high multiplicity, then this chain is more likely to show a cutoff.

Our Results. In this chapter, we show the cutoff phenomenon for simple random walks on Kneser graphs. Given two integers n and k, the Kneser graph $K(2n+k,n)$ is defined as the graph with the vertex set being all subsets of $\{1,\ldots,2n+k\}$ of size n and two vertices A and B being connected by an edge if $A \cap B = \emptyset$. We show that for any $k = O(n)$, then the random walks on $K(2n+k,n)$ exhibit a cutoff at $\frac{1}{2}\log_{1+k/n}(2n+k)$ with a window of size $O(\frac{n}{k})$. In the case that $k = \omega(n)$, the number of vertices and degree of each vertex are $\binom{2n+k}{n}$ and $\binom{n+k}{n}$, respectively which they have the same magnitude and hence the simple random walks on $K(2n+k,n)$ is mixed in just one step. It is also well-known that the transition matrix of the simple random walk on Kneser graph $K(2n+k,n)$ has spectral gap $\frac{k}{n+k}$ and its second largest eigenvalue has multiplicity $2n+k$ (cf. Corollary 4.3.3). So by varying $k = \mathcal{O}(n)$, we obtain various family of chains with different spectral gaps. For instance by setting $k = \Theta(n)$ we obtain a family of transitive expander graphs. It is worth mentioning that Godsil [God80] shows that for most values of n and k, the graph $K(2n+k,n)$ is not a Cayley graph.

Techniques. For the special case $k = 1$, we obtain the so-called odd graph $K(2n+1,n)$ with large odd cycles of size $2n+1$, which is a subgraph of $K(2n+k,n)$. This proves that $K(2n+k,n)$ is not bipartite for every $k \geqslant 1$. The permutation group on $[2n+k]$ is a subgroup of the automorphism group of $K(2n+k,n)$, and thus the Kneser graph is always transitive. Combining these two observations, we conclude that the simple random walk on $K(2n+k,n)$ is an ergodic and transitive Markov chain. Kneser graphs have been studied frequently in (algebraic) graph theory, in particular due to their connections to chromatic numbers and graph homomorphisms. (See [GR01] for more details and references.) In order to show a cutoff for a simple random walk on Kneser graphs, it is necessary to have a sufficiently tight estimate of its mixing time. Let P be the transition matrix of the simple random walk on Kneser graph $K(2n+k,n)$ with spectrum λ_i, $0 \leqslant i \leqslant \binom{2n+k}{n} - 1$ and $\lambda_0 = 1$. Then, it is shown that [LPW09, Lemma 12.16]

$$d(t) = \max_{x\in\Omega} \|P^t(x,.) - \pi\|_{TV} \leqslant \frac{1}{2}\sqrt{\sum_{i=1}^{|\Omega|-1} \lambda_i^{2t}}, \tag{4.1}$$

where Ω is the vertex set of the graph. It may be surprising that the upper bound obtained by the spectral properties of the transition matrix is sufficiently tight and matches the lower bound, which enables us to show the existence of a cutoff. Besides Kneser graphs, the bound in (4.1) has also been successfully applied in the computing of the mixing time of random walks on Cayley graphs (see [Dia88, DH96]). This may suggest the following question:

Question For which families of transitive ergodic chains is the upper bound in (4.1) tight up to low order terms?

Outline. In Section 4.1 we formally define the cutoff phenomenon for a family of chains. In Section 4.2 we state our main result and give a proof. The proof is based on two propositions shown separately in Sections 4.3 and 4.4.

4.1. Definitions and Notations

In this section, we formally define the cutoff phenomenon for a given sequence of ergodic Markov chains. Let us first recall some basic definitions from the Markov chain theory that can be found in [LPW09].

A finite Markov chain is a stochastic process which moves from one element of a finite state space Ω to another element of Ω, where the movements are governed by a fixed probability distribution or simply by a $|\Omega| \times |\Omega|$ matrix P, which is called the transition matrix. This means that for every $x, y \in \Omega$, at any time the probability of moving from x to y is specified by entry $P(x, y)$ of the matrix. A simple induction argument shows that each entry of the t-th power of P, denoted by $P^t(x, y)$, indicates the probability that a chain started from state x lands on state y. A finite Markov chain is *ergodic* if for every $x, y \in \Omega$, there exists a positive number $t_{x,y}$ so that for every $t \geqslant t_{x,y}$ $P^t(x, y) > 0$. Let $P^t(x, .)$ be the probability distribution of the chain at time $t \in \mathbb{N}$ over Ω with starting state $x \in \Omega$. It is well known that $P^t(x, .)$ converges to a stationary distribution π as t goes to infinity (provided the chain is finite and ergodic). The total variation distance between two probability distributions μ and ν on a probability space Ω is defined by

$$\|\mu - \nu\|_{TV} = \max_{A \subset \Omega} |\mu(A) - \nu(A)| \in [0, 1].$$

Therefore, we can define the worst-case total variation distance to stationarity at time t as

$$d(t) = \max_{x \in \Omega} \|P^t(x, .) - \pi\|_{TV}.$$

For convenience, we define $d(t)$ for non-integer t as $d(t) := d(\lfloor t \rfloor)$. (If the reference is clear from the context, we will also just say total variation distance at time t). The mixing time is defined by

$$t_{mix}(\epsilon) = \min\{t \in \mathbb{N} : d(t) < \epsilon\}.$$

Suppose now that we have a sequence of ergodic finite Markov chains indexed by $n = 1, 2, \ldots$. Let $d_n(t)$ be the total variation distance of the n-th chain at time t and $t_{mix}^{(n)}(\epsilon)$ be its mixing time. Formally, we say that the sequence of chains exhibits a cutoff (in total variation distance), as defined in [LPW09, Section 18.1], if for any fixed $0 < \epsilon < 1$,

$$\lim_{n\to\infty} \frac{t^{(n)}_{mix}(\epsilon)}{t^{(n)}_{mix}(1-\epsilon)} = 1.$$

Equivalently, a sequence of Markov chains has a cutoff at time t_n with a window of size $w_n = o(t_n(1/4))$ if

$$\begin{aligned} \lim_{c\to\infty} \liminf_{n\to\infty} d_n(t_n - cw_n) &= 1, \\ \lim_{c\to\infty} \limsup_{n\to\infty} d_n(t_n + cw_n) &= 0. \end{aligned} \tag{4.2}$$

A transition matrix is said to be *reversible* if it satisfies the following equality

$$\pi(x)P(x,y) = \pi(y)P(y,x).$$

It is claer that simple random walks on a connected non-bipartite graph is finite and ergodic. Moreover, if the graph is regular, then the stationary distribution is the uniform distribution over the vertex set of the graph. Note that if A is an adjacency matrix of a d-regular graph, then $(1/d)A$ is the transition matrix of the simple random walk on the graph. Also, if A is a symmetric matrix, then $(1/d)A$ is a reversible transition matrix. Let P be a reversible transition matrix and λ_2 the second largest eigenvalue of P. Then, $1 - \lambda_2$ is defined as the *spectral gap* of P.

4.2. Main Theorem

In the following we state the main result of the paper.

Theorem 4.2.1. *The simple random walk on $K(2n+k, n)$ exhibits a cutoff at $\frac{1}{2}\log_{1+k/n}(2n+k)$ with a cutoff window of size $O(\frac{n}{k})$ for $k = O(n)$.*

We now give the proof of Theorem 4.2.1 using Proposition 4.3.4 and 4.4.3, whose statements and proofs are deferred to later sections.

Proof. For the proof of the upper bound on the mixing time, we use the spectrum of the transition matrix. Applying Proposition 4.3.4 implies that

$$\lim_{c\to\infty} \liminf_{n\to\infty} d_n\Big(\frac{1}{2}\log_{1+k/n}(2n+k) + c\frac{n}{k}\Big) = 0.$$

We establish the lower bound by considering the vertices visited by a random walk starting from $\{n+1, \ldots, 2n\}$ and their intersection with $[n] = \{1, \ldots, n\}$. For any step, we compute the expected size of the intersection and derive an upper bound on its variance (to stationarity). Then applying Proposition 4.4.3 results into

$$\lim_{c\to\infty} \liminf_{n\to\infty} d_n\left(\frac{1}{2}\log_{1+k/n}(2n+k) - c\frac{n}{k}\right) = 1.$$

Combining these findings establishes a cutoff at $\frac{1}{2}\log_{1+k/n}(2n+k)$ with a cutoff window of size $O(\frac{n}{k})$ for $k = O(n)$. $\square$

4.3. Upper Bound on the Variation Distance

To prove our results, we need two lemmas, the lemma below can be found in [LPW09, Lemma 12.16].

Lemma 4.3.1 ([LPW09, Lemma 12.16]). *Let P be a reversible transition matrix with eigenvalues*

$$1 = \lambda_0 \geqslant \lambda_1 \geqslant \ldots \geqslant \lambda_{|\Omega|-1}.$$

If the Markov chain is transitive, then for every $x \in \Omega$

$$4\|P^t(x,.) - \pi\|_{TV}^2 \leqslant \sum_{i=1}^{|\Omega|-1} \lambda_i^{2t}.$$

To employ Lemma 4.3.1, we need to know all eigenvalues and their multiplicities. The spectrum of the adjacency matrix of Kneser graphs was computed in [GR01, Section 9.4] and [Rei00].

Theorem 4.3.2 ([GR01, Section 9.4] and [Rei00]). *The adjacency matrix of Kneser graphs $K(2n+k,n)$ has the following spectrum*

$$(-1)^i \binom{n+k-i}{n-i} \quad \text{with multiplicity of} \quad \binom{2n+k}{i} - \binom{2n+k}{i-1}, \quad i = 0, \ldots, n,$$

where $\binom{2n+k}{-1} = 0$.

As $K(2n+k,n)$ is a $\binom{n+k}{n}$-regular graph, we immediately obtain the following corollary.

Corollary 4.3.3. *The transition matrix of the simple random walk on $K(2n+k,n)$ has the following spectrum:*

$$(-1)^i \frac{\binom{n+k-i}{n-i}}{\binom{n+k}{n}} \quad \text{with multiplicity of} \quad \binom{2n+k}{i} - \binom{2n+k}{i-1}, \quad i = 0, \ldots, n.$$

Proposition 4.3.4. *We have the following upper bounds on the total variation distance of the simple random walk on $K(2n+k,n)$.*

- *If $k = o(n)$, then for every constant $c \geqslant 1/2$,*

$$d\left(\frac{1}{2}\log_{1+k/n}(2n+k) + c\frac{n}{k}\right) \leqslant e^{-c}.$$

- *If $k = \Omega(n)$, then for every constant c with $(1+\frac{k}{n})^{-c} \leqslant \frac{1}{2}$,*

$$d\left(\frac{1}{2}\log_{1+k/n}(2n+k) + c\right) \leqslant (1+k/n)^{-c}.$$

Proof. By Corollary 4.3.3 we have

$$|\lambda_i| = \left|(-1)^i \frac{n(n-1)(n-2)\cdot\ldots\cdot(n-i+1)}{(n+k)(n+k-1)(n+k-2)\cdot\ldots\cdot(n+k-i+1)}\right| \leqslant \left(\frac{n}{n+k}\right)^i = \left(1-\frac{k}{n+k}\right)^i.$$

Now define

$$g(t) = \left(1-\frac{k}{n+k}\right)^{2t}(2n+k) = \left(1+\frac{k}{n}\right)^{-2t}(2n+k).$$

Applying Lemma 4.3.1 yields,

$$\begin{aligned} 4\|P^t(x,.)-\pi\|_{TV}^2 \leqslant& \sum_{i=1}^{n}\left(1-\frac{k}{n+k}\right)^{i2t}\cdot\left\{\binom{2n+k}{i}-\binom{2n+k}{i-1}\right\} \\ \leqslant& \sum_{i=1}^{n}\frac{\left((1-\frac{k}{n+k})^{2t}(2n+k)\right)^i}{i!} \\ \leqslant& e^{g(t)}-1. \end{aligned}$$

Using the fact that for every x, $0 \leqslant x \leqslant 1/2$, $e^x - 1 \leqslant 2x$, we conclude that for any $0 \leqslant g(t) \leqslant 1/2$,

$$\|P^t(x,.)-\pi\|_{TV} \leqslant \sqrt{g(t)/2} \tag{4.3}$$

We consider two cases:

Case 1. $k = o(n)$. We choose $t = \frac{1}{2}\log_{1+k/n}(2n+k) + c\frac{n}{k}$, where $c \geqslant 1/2$. Hence,

$$g(t) = \left(1+\frac{k}{n}\right)^{-2t}(2n+k) = \left(1+\frac{k}{n}\right)^{-2\frac{cn}{k}} \leqslant e^{-2c} < 1/2,$$

and by inequality (4.3),

$$d\left(\frac{1}{2}\log_{1+k/n}(2n+k) + c\frac{n}{k}\right) \leqslant e^{-c}$$

Case 2. $k = \Omega(n)$. Now we choose $t = \frac{1}{2}\log_{1+k/n}(2n+k) + c$. Then,

$$g(t) = \left(1+\frac{k}{n}\right)^{-2t}(2n+k) = \left(1+\frac{k}{n}\right)^{-2c} \leqslant 1/2,$$

where the last inequality holds due to assumption on c. Hence, inequality (4.3) yields

$$d\left(\frac{1}{2}\log_{1+k/n}(2n+k) + c\right) \leqslant \left(1+\frac{k}{n}\right)^{-c}$$

□

4.4. Lower Bound on the Variation Distance

In order to find a lower bound for variation distance we use the following lemma which was applied in [Wil04]. For further discussion on this method we refer the reader to [SC04]. Let f be a real-valued function on Ω. We use $\mathbf{E}_\mu[f]$ and $\mathbf{Var}_\mu[f]$ to denote the expectation and variance of f under distribution of μ.

Lemma 4.4.1 ([LPW09, Proposition 7.8]). *Let μ and ν be two probability distributions on Ω and $f : \Omega \to \mathbb{R}$ be an arbitrary function. Suppose that* $\max\{\mathbf{Var}_\mu[f], \mathbf{Var}_\nu[f]\} \leqslant \sigma_*^2$. *Then if*

$$|\mathbf{E}_\mu[f] - \mathbf{E}_\nu[f]| \geqslant r\sigma_*,$$

then

$$\|\mu - \nu\|_{TV} \geqslant 1 - \frac{8}{r^2}.$$

Before proceeding, we recall that a random variable $Y \sim H(N, m, n)$ has a hypergeometric distribution if for every $\max\{0, n + m - N\} \leqslant i \leqslant \min\{n, m\}$, $\mathbf{Pr}[Y = i] = \frac{\binom{m}{i}\binom{N-m}{n-i}}{\binom{N}{n}}$. The expected value and variance of Y are $\mathbf{E}[Y] = \frac{nm}{N}$ and $\mathbf{Var}[Y] = \frac{nm(N-m)(N-n)}{N^2(N-1)}$ respectively.

Lemma 4.4.2. *Let X_t be the vertex visited at step t by a simple random walk on $K(2n + k, n)$ which starts at vertex $X_0 = \{n + 1, n + 2, \dots, 2n\}$. Let $f_t = f(X_t) = |X_t \cap [n]|$, so $f_0 = 0$. Moreover, define a random variable $f = |X \cap [n]|$ with X being a vertex chosen uniformly at random from $K(2n + k, n)$. Then for any $t \in \mathbb{N}$,*

$$\mathbf{Var}[f_t] \leqslant C(n, k)\,\mathbf{Var}[f],$$

where $C(n, k) = (1 + o(1))(1 + k/n)$ for $k = O(n)$.

Proof. The random variable f under π has a hypergeometric distribution $H(2n + k, n, n)$. Hence,

$$\mathbf{E}[f] = \frac{n^2}{2n + k}, \tag{4.4}$$

and

$$\mathbf{Var}[f] = \frac{n^2(n + k)^2}{(2n + k)^2(2n + k - 1)}. \tag{4.5}$$

In step $t+1$ of the walk, an n-element subset of the complement of X_t is chosen. If $f_t = s$, $|X_t \cap [n]| = s$, then X_t^c has $n - s$ common elements with $[n]$ and $s + k$ common elements with $[n]^c$. Therefore $f_{t+1} = n - Y$ where Y has hypergeometric distribution $H(n + k, s + k, n)$. Hence,

$$\begin{aligned}\mathbf{E}[f_{t+1} \mid f_t = s] &= \mathbf{E}[n - Y] = n - \frac{(s + k)n}{n + k} = (n - s) \cdot \left(1 - \frac{k}{n + k}\right) \\ &= n\left(1 - \frac{k}{n + k}\right) - \mathbf{E}[f_t]\left(1 - \frac{k}{n + k}\right).\end{aligned}$$

Solving this recursion allows us to compute the expectation of f_t:

$$\mathbf{E}[f_t] = n\sum_{i=1}^{t}\left[(-1)^{i+1}\left(1-\frac{k}{n+k}\right)^i\right] + \underbrace{(-1)^t\,\mathbf{E}[f_0]\,(1-\frac{k}{n+k})^t}_{=0}$$

$$= -n\frac{(\frac{k}{n+k}-1)^{t+1}-(\frac{k}{n+k}-1)}{\frac{k}{n+k}-2} = \frac{n^2}{2n+k} + (-1)^{t+1}\frac{n(n+k)(1-\frac{k}{n+k})^{t+1}}{2n+k}. \tag{4.6}$$

We have already shown that $\mathbf{E}[f_{t+1} \mid f_t] = n(1-\frac{k}{n+k})-f_t(1-\frac{k}{n+k})$, which immediately implies that

$$\mathbf{Var}[\mathbf{E}[f_{t+1} \mid f_t]] = \left(1-\frac{k}{n+k}\right)^2 \mathbf{Var}[f_t].$$

As observed earlier, the random variable f_{t+1} conditioned on f_t has distribution $n-Y$ where $Y \sim H(n+k, f_t+k, n)$ which yields

$$\mathbf{Var}[f_{t+1} \mid f_t] = \mathbf{Var}[n-Y] = \mathbf{Var}[Y] = \frac{(f_t+k)(n-f_t)}{(n+k)^2} \times \frac{nk}{(n+k-1)}.$$

Assume now that A is an upper bound for $\frac{(f_t+k)(n-f_t)}{(n+k)^2}$ for every f_t; A will be specified later. In the following, we use the total law of variance to find a recursive formula for $\mathbf{Var}[f_t]$,

$$\begin{aligned}\mathbf{Var}[f_{t+1}] &= \mathbf{Var}[\mathbf{E}[f_{t+1} \mid f_t]] + \mathbf{E}[\mathbf{Var}[f_{t+1} \mid f_t]]\\ &\leqslant \left(1-\frac{k}{n+k}\right)^2 \mathbf{Var}[f_t] + A\frac{nk}{(n+k-1)}.\end{aligned}$$

Using this recursion, we obtain the following upper bound on $\mathbf{Var}[f_t]$:

$$\mathbf{Var}[f_t] \leqslant A\frac{nk}{n+k-1}\sum_{i=0}^{t-1}\left[\left(1-\frac{k}{n+k}\right)^{2i}\right] + \underbrace{(1-\frac{k}{n+k})^{2t}V(f_0)}_{=0}$$

$$= A\frac{nk}{n+k-1} \times \frac{1-(1-\frac{k}{n+k})^{2t}}{1-(1-\frac{k}{n+k})^2} \leqslant A\frac{n(n+k)^2}{(2n+k)(n+k-1)}.$$

Since always $0 \leqslant f_t \leqslant n$, $\frac{(f_t+k)(n-f_t)}{(n+k)^2} \leqslant 1/4 = A$.

$$\mathbf{Var}[f_t] \leqslant \frac{1}{4}\cdot\frac{n(n+k)^2}{(2n+k)(n+k-1)} = \frac{1}{4}\cdot\frac{n^3(1+k/n)^2}{n^2(2+k/n)(1+k/n-o(1))} = n\frac{(1+k/n)(1+o(1))}{4(2+k/n)}.$$

Moreover,

$$\mathbf{Var}[f] \geqslant \frac{n^4(1+k/n)^2}{n^3(2+k/n)^3}.$$

Using the fact that $1/2 \leqslant \frac{1+x}{2+x}$ for every $x \geqslant 0$,

$$\mathbf{Var}[f]\cdot(1+k/n)\cdot(1+o(1)) \geqslant \frac{n(1+k/n)(1+o(1))}{4(2+k/n)}.$$

By comparing $\mathbf{Var}[f]$ and $\mathbf{Var}[f_t]$, the claim follows. □

We are now ready to apply Lemma 4.4.1 to derive a lower bound on the total variation distance.

Proposition 4.4.3. *For every constant $c > 0$, we have the following lower bounds on the total variation distance for a simple random walk on $K(2n+k, n)$.*

- *If $k = o(n)$,*

$$d\left(\frac{1}{2}\log_{1+k/n}(2n+k) - c\frac{n}{k}\right) \geqslant 1 - 8(1+o(1))(e-o(1))^{-2c}.$$

- *If $k = \Theta(n)$, then*

$$d\left(\frac{1}{2}\log_{1+k/n}(2n+k) - c\right) \geqslant 1 - 8(1+o(1))(1+k/n)^{-2c+4}.$$

Proof. By using Lemma 4.4.2 and (4.5)

$$\sqrt{\max\{\mathbf{Var}\,[f]\,, \mathbf{Var}\,[f_t]\}} \leqslant \sqrt{C(n,k)\,\mathbf{Var}\,[f]} \leqslant C(n,k)\frac{n(n+k)}{(2n+k)\sqrt{2n+k-1}} = \sigma_*.$$

Combining (4.6) and (4.4),

$$|\,\mathbf{E}\,[f_t] - \mathbf{E}\,[f]\,| = \frac{n(n+k)}{2n+k}\left(1 - \frac{k}{n+k}\right)^{t+1} = \frac{1}{C(n,k)}\sigma_*\sqrt{2n+k-1}\left(1+\frac{k}{n}\right)^{-t-1}.$$

Define

$$\tilde{g}(t) = \frac{\sqrt{2n+k-1}}{C(n,k)}\left(1+\frac{k}{n}\right)^{-t-1}.$$

- **Case 1.** $k = o(n)$. By Lemma 4.4.2 we know that $C(n,k) = (1+k/n)(1+o(1)) = (1+o(1))$. We choose $t = \frac{1}{2}\log_{1+k/n}(2n+k) - c\frac{n}{k}$ so that

$$\tilde{g}(t) = \frac{\sqrt{1-o(1)}}{1+o(1)}\left(1+\frac{k}{n}\right)^{c\frac{n}{k}} = (1-o(1))e_n^c,$$

  where $(e_n)_n$ is an increasing sequence tending to e as $n \to \infty$. Applying Lemma 4.4.1 yields,

$$d\left(\frac{1}{2}\log_{1+k/n}(2n+k) - c\frac{n}{k}\right) = \|P^t(X_0,.) - \pi\|_{TV} \geqslant 1 - 8(1+o(1))e_n^{-2c},$$

  where $X_0 = \{n+1, \ldots, 2n\}$ and the equality comes from the fact that the chain is transitive.

- **Case 2.** $k = \Theta(n)$. By Lemma 4.4.2, $C(n,k) = (1+k/n)(1+o(1))$. Take $t = \frac{1}{2}\log_{1+k/n}(2n+k) - c$. Hence,

$$\tilde{g}(t) = \frac{\sqrt{1-o(1)}}{1+o(1)}\,(1+k/n)^{c-2}.$$

  Again, using Lemma 4.4.1 gives

$$d\left(\frac{1}{2}\log_{1+k/n}(2n+k) - c\right) = \|P^t(X_0,.) - \pi\|_{TV} \geqslant 1 - 8(1+o(1)(1+k/n)^{-2c+4}.$$

□

5 Balanced Allocation on Graphs: A Random Walk Approach

The standard balls-into-bins model is a process which randomly allocates m balls into n bins where each ball picks d bins independently and uniformly at random and the ball is then allocated in a least loaded bin in the set of d choices. When $m = n$ and $d = 1$, it is well known that at the end of process the maximum number of balls at any bin, the *maximum load*, is $(1 + o(1))\frac{\log n}{\log\log n}$ whp. Azer et al. [ABKU99] showed that for the d-choice process, $d \geqslant 2$, provided ties are broken randomly, the maximum load is $\frac{\log\log n}{\log d} + \mathcal{O}(1)$. The result implies that the maximum load is constant if and only if $d = \log^{\Omega(1)} n$. For a complete survey on the standard balls-into-bins process we refer the reader to [MRS01]. Many subsequent works consider the settings where the choice of bins are not necessarily independent and uniform. For instance, Vöcking [Völ03] proposed an algorithm called *always-go-left* that uses exponentially smaller number of choices (i.e., $d = \Omega(\log\log n)$) to achieve a constant maximum load. In this algorithm the bins are partitioned into d groups of size n/d and each ball picks one random bin from each group. The ball is then allocated in a least loaded bin among the chosen bins and ties are broken asymmetrically. The algorithm results in a maximum load of $\frac{\log\log n}{d\phi_d} + \mathcal{O}(1)$ whp, where $1 \leqslant \phi_d \leqslant 2$ is constant.

In many applications selecting any random set of choices is costly. For example, in peer-to-peer or cloud-based systems balls (jobs, items,...) and bins (servers, processors,...) are randomly placed in a metric space (e.g., $\mathbb{R}^2$) and the balls have to be allocated on bins that are close to them as it minimizes the access latencies. With regard to such applications, Byer et al. [BCM04] studied a model, where n bins (servers) are uniformly at random placed on a geometric space. Then each ball in turn picks d locations in the space and allocates itself on a nearest neighboring bin with minimum load among other d bins. In this scenario, the probability that a location close to a server is chosen depends on the distribution of other servers in the space and hence there is no a uniform distribution over the potential choices. Here, the authors showed the maximum load is $\frac{\log\log n}{\log d} + \mathcal{O}(1)$ whp. Later on, Kenthapadi and Panigrahy [KP06] proposed a model in which bins are interconnected as a Δ-regular graph and each ball picks a random edge of the graph. It is then placed at one of its endpoints with smaller load. This allocation algorithm results in a maximum load of $\log\log n + \mathcal{O}\left(\frac{\log n}{\log(\Delta/\log^4 n)}\right) + \mathcal{O}(1)$. Following the study of balls-into-bins with correlated choices, Godfrey [God08] generalized the aforementioned result such that each ball picks an random edge of a hypergraph that has $\Omega(\log n)$ bins and satisfies some mild conditions. Then he showed that the maximum load is a constant whp. Recently, Bogdan et al. [BSSS13] studied a model where each ball picks a random node and performs a local search from the node to find a node with local minimum load, where it is finally placed on. They showed that when the graph is a constant degree expander, the local search guarantees a maximum load of $\Theta(\log\log n)$ whp.

Our Results. In this chapter we propose an algorithm for allocating n sequential balls into n bins that are organized as a d-regular n-vertex graph G, where $d \geqslant 3$ can be any integer. Let l be a given positive integer. A non-backtracking random walk (NBRW) W of length l started from a node is a random walk in l steps so that in each step the walker picks a neighbor uniformly at random and moves to that neighbor with an additional property that the walker never traverses an edge twice in a row. Further information about NBRWs can be found in [ABLS07] and [AL09]. Our allocation algorithm, denoted by $\mathcal{A}(G, l)$, is based on a random sampling of bins from the neighborhood of a given node in G by a NBRW from the node. The algorithm proceeds as follows: In each round t, $1 \leqslant t \leqslant n$, ball t picks a node of G uniformly at random and performs a NBRW $W = (u_0, u_1 \ldots, u_l)$. After that a set of potential choices called b-choice, $\boldsymbol{\beta}(W) := \{u_{j\cdot r_G} \mid 0 \leqslant j \leqslant \lfloor l/r_G \rfloor\}$ ($b = \lfloor l/r_G \rfloor + 1$ and $r_G = \lceil 2 \cdot \log_{d-1} \log n \rceil$), is selected and finally the ball is allocated in a least loaded bin of $\boldsymbol{\beta}(W)$ (ties are broken randomly). It is worth to mention if $d \geqslant \log^2 n$, then $r_G = 1$ and $\boldsymbol{\beta}(W)$ is the set of all nodes contained in W. Our result concerns bounding the maximum load attained by $\mathcal{A}(G, l)$, denoted by m^*, in terms of l. Throughout this chapter, we assume that G is a d-regular n-vertex graph with girth at least $\omega(l \log\log n)$, $l \geqslant 20 r_G$ and $l = o(\log_d n)$. However, we will see in the proof of the result, it is sufficient that G has girth at least $\omega(l)$ for some values of l. In order to present the upper bound, we consider two cases:

I. If $l \geqslant 4\gamma_G$, where $\gamma_G = \sqrt{r_G \cdot \log_d n}$, then we show that whp,

$$m^* = \mathcal{O}\left(\frac{\log\log n}{\log(l/\gamma_G)}\right).$$

It is readily checked that for every G, $\gamma_G \leqslant \sqrt{2\log_2 \log n \cdot \log_3 n}$. Therefore, for a given regular graph satisfying the girth condition, if we set $l = \lfloor (\log n)^{\frac{1+\epsilon}{2}} \rfloor$, for any constant $\epsilon \in (0, 1)$ we have $l \geqslant 4\gamma_G$ and then by applying the upper bound we have $m^* = \mathcal{O}(1/\epsilon)$ whp.

II. If $32 \cdot r_G \leqslant l \leqslant 4 \cdot \gamma_G$, then we show that whp,

$$m^* = \mathcal{O}\left(\frac{r_G \cdot \log_d n \cdot \log\log n}{l^2}\right).$$

In addition to the upper bound for m^*, we prove that $m^* = \Omega(r_G \cdot \log_d n/l^2)$ whp (for a proof see Section 5.4). So it is easy to see that the upper bound for m^* is at most $\mathcal{O}(\log\log n)$ factor above the lower bound.

The setting of our work is closely related to [BSSS13]. In this paper in each step a ball picks a node of a graph uniformly at random and performs a local search to find a node with local minimum load and finally allocates itself on it. They showed that with high probability the local search on expander graphs obtains a maximum load of $\Theta(\log\log n)$. In comparison to the mentioned result, our new protocol achieves a further reduction in the maximum load, while still allocating a ball close to its origin. Our result suggests a tradeoff between allocation time and maximum load. In fact we show a constant upper bound for sufficient long walks (i.e., $l = (\log n)^{\frac{1+\epsilon}{2}}$, for any constant $\epsilon \in (0,1)$). Our work can also be related to the one by Kenthapadi and Panigrahy where balls pick a random edge in d-regular graphs with $d = n^{\Omega(1/\log\log n)}$ resulting into a maximum load of $\Theta(\log\log n)$. Godfrey [God08] also studied an allocation algorithm where every ball chooses a random edge e of a hypergraph satisfying some conditions, that is, first the size of each edge is $d = \Omega(\log n)$ and $\mathbf{Pr}\left[u \in e\right] = \Theta(\frac{d}{n})$ for any bin u. The latter one is called *balanced condition*. It is not hard to see that if we have a graph with girth $g = \Omega(\log n)$ and set $l = g/2$, then visited nodes by a ball generates a hyperedge satisfying aforementioned conditions. Berenbrink et al. [BBFN12] simplified Godfrey's proof and slightly weakened the balanced condition but since both analysis apply a Chernoff bound, it seems unlikely that one can extend the analysis for $l = o(\log n)$.

In a different context, Alon and Lubetzky [AL09] showed that if a particle starts a NBRW of length n on n-vertex graph with high-girth then the number of visits to nodes has a poisson distribution. In particular they showed that the maximum visit to a node is at most $(1 + o(1)) \cdot \frac{\log n}{\log\log n}$. Our result can be also seen as an application of the mathematical concept of NBRWs to task allocation in distributed networks.

Techniques. To derive a lower bound for the maximum load we first show that whp there is a path of length l which is traversed by at least $\Omega\left(\log_d n/l\right)$ balls. Also, each path contains l/r_G potential choices and hence, by pigeonhole principle there is a node with load at least $\Omega\left(r_G \log_d n/l^2\right)$, which is a lower bound for m^*. In order to establish the upper bound, we apply the *witness graph* techniques and a key property of the algorithm, which is called (α, n_1)-*uniformity*. We say an allocation algorithm is (α, n_1)-uniform if the probability that ball $1 \leqslant t \leqslant n_1$ is placed on an arbitrary node is bounded by α/n, where $n_1 = \Theta(n)$ and $\alpha = \mathcal{O}(1)$. Note that the intuition behind selecting a subset of visited nodes as a potential choices instead of all of them follows from our technique for showing the (α, n_0)-uniformity of the algorithm. Using this property we conclude that for a given set of nodes of size $\Omega(\log n)$, after allocating n_1 balls, the average load of nodes in the set is some constant whp. Using witness graph method we show that if there is a node with load larger than some threshold then there is a collection of nodes of size $\Omega(\log n)$ where each of them has load larger than some specified constant. Putting these together implies that maximum load is bounded as required whp.

Outline. In Section 5.1, we present notations and some preliminary results that are required for the analysis of the algorithm. In Section 5.2 we show how to construct a witness graph and then in Section 5.3 by applying the results we prove the main theorem. In Section 5.4 we derive a lower bound for the algorithm and finally the last section we show the key property of the algorithm which is then applied to prove the main theorem.

5.1. Notations, Definitions and Preliminaries

In this section we provide notations, definitions and some preliminary results that are needed in this chapter. Throughout this chapter we assume that G is a d-regular n-vertex graph with girth $\omega(l \log\log n)$ and $l \geqslant 20r_G$ is an integer, where $r_G = \lceil 2\log_{d-1}\log n\rceil$. The visited nodes by a non-backtracking walk of length l is called an *l-walk* whose nodes are ordered in terms of their visit. Since G has girth $\omega(l\log\log n)$, any l-walk in G contains $l+1$ nodes and it is a path of length l. For every l-walk $W = (u_0, u_1, \dots, u_l)$ contained in G, we define a b-choice as follows:

$$\boldsymbol{\beta}(W) := \{u_{j\cdot r_G} \mid 0 \leqslant j \leqslant \lfloor l/r_G \rfloor\},$$

where $b = \lfloor l/r_G\rfloor + 1$ denotes the size of $\boldsymbol{\beta}(W)$ and $r_G = \lceil 2\log_{d-1}\log n\rceil$. Also, we define $f(W)$ to be the number of balls in a least-loaded node of $\boldsymbol{\beta}(W)$. The *height* of a ball allocated on a node is the number balls that are placed on the node before the ball.

Definition 5.1.1 (Interference Graph). *For every given pair (G,l), the interference graph $\mathcal{I}(G,l)$ is defined as follows: The vertex set of $\mathcal{I}(G,l)$ is the set of all b-choices that corresponds to the set of all l-walks in G and two vertices $\boldsymbol{\beta}$ and $\boldsymbol{\beta}'$ of $\mathcal{I}(G,l)$ are connected if and only if $\boldsymbol{\beta} \cap \boldsymbol{\beta}' \neq \emptyset$. Note that if the pair (G,l) is clear from the context, then the interference graph is denoted by $\mathcal{I}$.*

Lemma 5.1.2. *Suppose that $V(\mathcal{I})$ and $\Delta(\mathcal{I})$ denote the vertex set and the maximum degree of $\mathcal{I}(G,l)$, respectively. Then we have*

(i) $|V(\mathcal{I})| = nd(d-1)^{l_0-1}/2$,

(ii) $\Delta(\mathcal{I}) \leqslant b^2 d(d-1)^{l_0-1}$,

where $l_0 = \lfloor l/r_G\rfloor \cdot r_G$ and $b = \lfloor l/r_G\rfloor + 1$. Furthermore, the number of rooted λ-vertex trees contained in $\mathcal{I}$ is bounded by $4^\lambda \cdot |V(\mathcal{I})| \cdot \Delta(\mathcal{I})^{\lambda-1}$.

Proof. By definition of b-choice, it is easily seen that the b-choice corresponding to a given l-walk $W = (u_0, u_1, \dots, u_l)$ is exactly the same as the b-choice corresponding to a truncated walk of W, which is $(u_0, u_1, \dots, u_{l_0})$. So we have $\mathcal{I}(G,l) = \mathcal{I}(G,l_0)$. Moreover, for every $\boldsymbol{\beta} \in V(\mathcal{I})$, there are nodes u and u' contained in $\boldsymbol{\beta}$ so that $d(u,u') = (b-1)r_G = l_0$ and since G has girth at least $\omega(l) = \omega(l_0)$, the path of length l_0 connecting u to u' is unique. This implies that there is a one-to-one correspondence between the vertex set of $\mathcal{I}$ and the set of all paths of length l_0 in G. On the other hand, by the girth condition on G we have that the total number of paths of length l_0 in G is $nd(d-1)^{l_0-1}/2$. Therefore we have,

$$|V(\mathcal{I})| = nd(d-1)^{l_0-1}/2.$$

Suppose that $W = (u_0, u_1 \ldots, u_{l_0})$ be an l_0-walk in G corresponding to some $\boldsymbol{\beta} \in V(\mathcal{I})$. By definition of $\boldsymbol{\beta}$, v is an element of $\boldsymbol{\beta}$ if and only if $v = u_{(j-1)r_G}$ for some $1 \leqslant j \leqslant b$. Since the graph locally looks like a d-ary tree, the total number of l_0-walks including v as $(j-1)r_G$-th node (i.e. $u_{(j-1)r_G} = v$) is at most

$$d(d-1)^{(j-1)r_G-1}(d-1)^{l_0-(j-1)r_G} = d(d-1)^{l_0-1}.$$

Index j varies from 1 to b, so v can be an element of at most $bd(d-1)^{l_0-1}$ b-choices. Also, every b-choice contains b elements and hence every b-choice intersects at most $b^2d(d-1)^{l_0-1}$ other b-choices. Thus we get

$$\Delta(\mathcal{I}) \leqslant b^2 \cdot d(d-1)^{l_0-1}.$$

Let us now bound the total number of rooted λ-vertex trees contained in $\mathcal{I}$. It was shown that the total number of different shape rooted trees on λ vertices is 4^λ (For example see [Knu97]); we say two rooted trees have different shapes if they are not isomorphic. For any given shape, there are $|V(\mathcal{I})|$ ways to choose the root. As soon as the root is fixed, each vertex in the first level can be chosen in at most $\Delta(\mathcal{I})$ many ways. By selecting the vertices of the tree level by level we have that each vertex except the root can be chosen in at most $\Delta(\mathcal{I})$ ways. So the total number of rooted λ-vertex trees in $\mathcal{I}$ is bounded by

$$4^\lambda \cdot |V(\mathcal{I})| \cdot \Delta(\mathcal{I})^{\lambda-1}.$$

□

By definition of $\mathcal{I}(G,l)$, the vertex set of $\mathcal{I}(G,l)$ is the set of all potential b-choices that can be made by $\mathcal{A}(G,l)$. Now, let us interpret allocation algorithm $\mathcal{A}(G,l)$ as follows:

For every ball $1 \leqslant t \leqslant n$, the algorithm picks a vertex of $\mathcal{I}(G,l)$, say $\boldsymbol{\beta}_t$, uniformly at random and then allocates ball t on a least-loaded node of $\boldsymbol{\beta}_t$ (ties are broken randomly).

Let $1 \leqslant n_1 \leqslant n$ be a given integer and assume that $\mathcal{A}(G,l)$ has allocated balls until the n_1-th ball. We then define $\mathcal{H}_{n_1}(G,l)$ to be the induced subgraph of $\mathcal{I}(G,l)$ by $\{\boldsymbol{\beta}_t : 1 \leqslant t \leqslant n_1\} \subset V(\mathcal{I})$.

Definition 5.1.3. *Let λ and μ be given positive integers. We say rooted tree $T \subset \mathcal{I}(G,l)$ is a (λ,μ)-tree if T satisfies:*

1) $|V(T)| = \lambda$,

2) $|\cup_{\boldsymbol{\beta}\in V(T)} \boldsymbol{\beta}| \geqslant \mu$.

Note that the latter condition is well-defined because every vertex of T is a b-element subset of $V(G)$. A (λ,μ)-tree T is called c-loaded, if T is contained in $\mathcal{H}_{n_1}(G,l)$, for some $1 \leqslant n_1 \leqslant n$, and every node in $\cup_{\boldsymbol{\beta}\in V(T)}\boldsymbol{\beta}$ has load at least c.

Corollary 5.1.4. *The size of family of (λ, μ)-trees is bounded by $4^\lambda |V(\mathcal{I})| \Delta(I)^{\lambda-1}$.*

Proof. We know that every (λ, μ)-tree T is a rooted λ-vertex subtree of $\mathcal{I}$ with the additional property that $| \cup_{\beta \in V(T)} \boldsymbol{\beta}| \geqslant \mu$. This implies that the size of family of rooted λ-vertex subtrees of $\mathcal{I}$ is an upper bound for the size of family of (λ, μ)-trees and hence by applying Lemma 5.1.2, we reach the upper bound $4^\lambda |V(\mathcal{I})| \Delta(I)^{\lambda-1}$. □

For every node $u, v \in V(G)$ let $d(u, v)$ denote the length of shortest path between u and v in G. Since G has girth at least $\omega(l)$, every path of length at most l is specified by its endpoints, say u and v and we denote it by interval $[u, v]$. Note that for any graph H, $V(H)$ is the vertex set of H.

5.1.1. Appearance Probability of a c-Loaded (λ, μ)-Tree

In this subsection we formally define the notion of (α, n_1)-uniformity for allocation algorithms, and then present our key lemma concerning the uniformity of $\mathcal{A}(G, l)$. By using this lemma we establish an upper bound for the probability that a c-loaded (λ, μ)-tree contained in $\mathcal{H}_{n_1}$ exists.

Definition 5.1.5. *Suppose that $\mathcal{B}$ be an algorithm that allocates n sequential balls into n bins. Then we say $\mathcal{B}$ is (α, n_1)-uniform, if after allocating t balls, for every $1 \leqslant t \leqslant n_1$, then for all $u \in V(G)$,*

$$\mathbf{Pr}\left[\textit{ball } t+1 \textit{ is allocated on } u\ \right] \leqslant \frac{\alpha}{n},$$

where α is some constant and $n_1 = \theta(n)$.

Lemma 5.1.6 (Key Lemma). *$\mathcal{A}(G, l)$ is an (α, n_1)-uniform allocation algorithm, where $n_1 = \lfloor n/(6\mathrm{e}\alpha) \rfloor$.*

Proof Sketch. Let $D_{r_G}(u)$ denote the set of all nodes at distance r_G from node $u \in V(G)$. We first show that if at the end of round $t \leqslant n_1$, for every $u \in V(G)$, a constant fraction of nodes in $D_{r_G}(u)$ are empty, then b-choice $\boldsymbol{\beta}_{t+1}$ contains $\theta(b)$ empty nodes with probability $1 - \mathcal{O}(1/b)$. It is not hard to see that the probability that an arbitrary node, say u, belongs to $\boldsymbol{\beta}_{t+1}$ is b/n. On the other hand, ties are broken randomly so if $u \in \boldsymbol{\beta}_{t+1}$, then ball $t+1$ is placed on u with probability $\mathcal{O}(1/b)$. It may happen that with probability $\mathcal{O}(1/b)$, $\boldsymbol{\beta}_{t+1}$ does not contain $\theta(b)$ empty nodes, which is called a *bad* b-choice. In this case we also show that both events (i.e., $u \in \boldsymbol{\beta}_{t+1}$ and $\boldsymbol{\beta}_{t+1}$ is a bad b-choice) happen with probability $\mathcal{O}(1/n)$. Putting these together implies that ball $t+1$ is being placed on u with probability $\mathcal{O}(1/n)$. Now, it remains to prove that during the allocations of balls up to round n_1, for every $u \in V(G)$, a constant fraction of nodes in $D_{r_G}(u)$ are empty whp. In order to show this, we define potential function $\Phi(t) = \sum_{u \in V(G)} \exp(a_t(u))$, where $a_t(u)$ is the number of nonempty in $D_{r_G}(u)$ after allocation t balls. Then, using a inductive argument we prove that for each $1 \leqslant t \leqslant n_1$, $a_t(u)$ is always bounded from above by a constant fraction of $|D_{r_G}(u)|$ whp as required. For a complete proof see Section 5.5.

In the next lemma, we derive an upper bound for the appearance probability a c-loaded (λ, μ)-tree, whose proof is inspired by [KP06, Lemma 2.1].

Lemma 5.1.7. *Let λ, μ and c be positive integers. Then the probability that there exists a c-loaded (λ, μ)-tree contained in $\mathcal{H}_{n_1}(G, l)$ is at most*

$$n \cdot \exp(4\lambda \log b - c\mu),$$

where $b = \lfloor l/r_G \rfloor + 1$.

Proof. Let us fix an arbitrary (λ, μ)-tree $T \subseteq \mathcal{I}(G, l)$ and p_1 be the probability that using λ balls T is built and contained in $\mathcal{H}_{n_1}$. There are at most $n_1 \leqslant n$ ways to choose one ball per vertex of T and hence at most n^λ ways to choose λ balls that are going to pick the vertices of T. On the other hand, every ball picks a given vertex of T (or a b-choice) with probability $1/|V(\mathcal{I})|$. Thus we get,

$$p_1 \leqslant n^\lambda \cdot (1/V(\mathcal{I}))^\lambda.$$

Now, we have to add c additional balls for very node in $\cup_{\boldsymbol{\beta} \in V(T)} \boldsymbol{\beta}$ and let p_2 denote the probability that such a event happens. Since $\mathcal{A}(G, l)$ is (α, n_1)-uniform with $n_1 = \lfloor n/(6e\alpha) \rfloor$ and $|\cup_{\boldsymbol{\beta} \in V(T)} \boldsymbol{\beta}| = \mu + q$, for some integer $q \geqslant 0$, we get

$$\begin{aligned}
p_2 &\leqslant \sum_{q=0}^{\infty} \binom{n_1}{c \cdot (\mu + q)} \left(\frac{\alpha \cdot (\mu + q)}{n} \right)^{c \cdot (\mu + q)} \\
&\leqslant \sum_{q=0}^{\infty} \left(\frac{\mathrm{e} \cdot n_1}{c \cdot (\mu + q)} \right)^{c \cdot (\mu + q)} \cdot \left(\frac{\alpha \cdot (\mu + q)}{n} \right)^{c \cdot (\mu + q)} \\
&\leqslant \sum_{q=0}^{\infty} \left(\frac{n_1 \cdot \alpha \cdot \mathrm{e}}{n \cdot c} \right)^{c \cdot (\mu + q)} \leqslant 2 \cdot \left(\frac{n_1 \cdot \alpha \cdot \mathrm{e}}{n \cdot c} \right)^{c \cdot \mu},
\end{aligned}$$

where we use the fact that for integers $1 \leqslant a \leqslant b$, $\binom{b}{a} \leqslant (\frac{\mathrm{e}b}{a})^a$ and the last inequality follows from $\left(\frac{n_1 \cdot \alpha \cdot \mathrm{e}}{n \cdot c}\right) < 1/2$. Since balls are mutually independent, $p_1 \cdot p_2$ is an upper bound for the probability that c-loaded (λ, μ)-tree T appears in $\mathcal{H}_{n_1}$. By Corollary 5.1.4 we have an upper bound for the size of family of all (λ, μ)-trees. Hence, taking the union bound over all (λ, μ)-trees gives an upper bound for appearance probability of a c-loaded (λ, μ)-tree in $\mathcal{H}_{n_1}$. Thus we get,

$$4^\lambda |V(\mathcal{I})| \cdot \Delta^{\lambda - 1} \cdot p_1 \cdot p_2 \leqslant 2 \cdot 4^\lambda |V(\mathcal{I})| \cdot \Delta^{\lambda - 1} \left(\frac{n}{V(\mathcal{I})} \right)^\lambda \cdot \left(\frac{n_1 \cdot \alpha \cdot \mathrm{e}}{n \cdot c} \right)^{c \cdot \mu}.$$

By Lemma 5.1.2 we have $V(\mathcal{I}) = nd(d-1)^{l_0}/2$, $\Delta(\mathcal{I}) \leqslant b^2 d(d-1)^{l_0}$, where $l_0 = \lfloor l/r_G \rfloor r_G$. Also we have $b = \lfloor l/r_G \rfloor + 1 > 20$ and $n_1 \leqslant n/6\alpha\mathrm{e}$. So the above bound is simplified as follows,

$$2n \left(8b^2\right)^\lambda \left(\frac{n_1 \cdot \alpha \cdot \mathrm{e}}{n \cdot c} \right)^{c \cdot \mu} \leqslant n b^{4\lambda} (1/6)^{c\mu} \leqslant n \exp(4\lambda \log b - c\mu).$$

$\square$

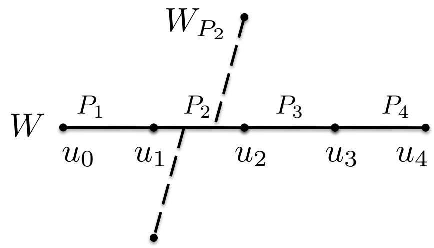

Figure 5.1. The Partition step on W for $k = 4$ and the Branch step for P_2 that gives W_{P_2}, shown by dashed line.

5.2. Witness Graph

In this section, we show that if there is a node whose load is larger than a threshold, then we can construct a c-loaded (λ, μ)-tree contained in $\mathcal{H}_{n_1}(G, l)$. Our construction is based on an iterative application of a 2-step procedure, called Partition-Branch, which we describe as follows:

Partition-Branch. Let $k \geqslant 1$ and $\rho \geqslant 1$ be given integers and W be an l-walk corresponding to a ball at height ρ with $f(W) \geqslant \rho+1$. The Partition-Branch procedure with parameters ρ and k, denoted by $PB(\rho, k)$, is a 2-step procedure that proceeds as follows:

Partition: First, it partitions W into k edge-disjoint subpaths as follows:

$$\mathcal{P}_k(W) = \{[u_i, u_{i+1}] \subset W, 0 \leqslant i \leqslant k-1\},$$

where $d(u_i, u_{i+1}) \in \{\lfloor l/k \rfloor, \lceil l/k \rceil\}$.

Branch: Second, for a given $P_i = [u_i, u_{i+1}] \in \mathcal{P}(W)$, it finds (if exists) another l-walk W_{P_i} corresponding to a ball allocated on $\boldsymbol{\beta}(W)$ that satisfies the following conditions:

(C1) $V(W_{P_i}) \cap V(W) \subseteq V(P_i) \setminus \{u_i, u_{i+1}\}$,

(C2) $f(W_{P_i}) \geqslant f(W) - \rho$.

We say procedure $PB(\rho, k)$ on a given l-walk W is *valid*, if for every $P \in \mathcal{P}_k(W)$, W_P exists. For a graphical view of the Partition-Branch procedure see Figure 5.1.

5.2.1. Construction of Witness Graph

In this subsection conditioning on event $\mathcal{F}_\delta$ which is defined later, we show how to construct a c-loaded (λ, μ)-tree contained in $\mathcal{H}_{n_1}$. Let us define a set of parameters depending on d, n, and l as follows:

$$\begin{aligned} k &:= \max\{4, \lfloor l/\sqrt{r_G \cdot \log_d n} \rfloor\}, \\ \delta &:= \lfloor \lfloor l/k \rfloor /4 \rfloor, \\ \rho &:= \lceil 8 r_G \log_d n / \delta^2 \rceil. \end{aligned}$$

By definition we have $r_G = \lceil 2\log_{d-1}\log n\rceil$, and hence $r_G = o(\sqrt{r_G \cdot \log_d n})$. Also we assume that $l \geqslant 32r_G$. So either $k = 4$ or $k = \lfloor l/\sqrt{r_G \cdot \log_d n}\rfloor$, we have,

$$\delta = \lfloor \lfloor l/k\rfloor/4\rfloor \geqslant 2r_G. \tag{5.1}$$

Now, we define a useful event that if it holds, then the Partition-Branch procedure is valid on an l-walk W with $f(W) \geqslant \rho + 1$.

Definition 5.2.1. *We say that event $\mathcal{F}_\delta$ holds, if after allocating at most n balls by $\mathcal{A}(G, l)$, every path of length δ is contained in less than $6\log_{d-1} n/\delta$ l-walks that are randomly chosen by $\mathcal{A}(G, l)$.*

Lemma 5.2.2. *Suppose that event $\mathcal{F}_\delta$ holds and W be an l-walk with $f(W) \geqslant \rho + 1$. Then the procedure $PB(\rho, k)$ on W is valid.*

Proof. Let us fix an arbitrary subpath $P_i = [u_i, u_{i+1}] \in \mathcal{P}(W)$. By definition, $\delta = \lfloor\lfloor l/k\rfloor/4\rfloor$ and we have $d(u_i, u_{i+1}) \geqslant 4\delta$. Define $P' = [u, v] \subset P_i$ such that

$$d(u_i, u) = d(v, u_{i+1}) = \delta.$$

And we have

$$d(u, v) \geqslant 2\delta.$$

Note that P' is a subpath of $[u_i, u_{i+1}]$ containing neither u_i nor u_{i+1}. By 5.1, $P' \subset W$ has length at least $4r_G$. So if we define $S = \boldsymbol{\beta}(W) \cap V(P')$, then $|S| \geqslant 3$. Let $B(S)$ be the set of all balls allocated on nodes of S with height at least $f(W) - \rho \geqslant 1$. Then $B(S) \neq \emptyset$. Clearly, each ball $t \in B(S)$ represents an l-walk W_t that satisfies $(C2)$, which means $f(W_t) \geqslant f(W) - \rho$ and $\boldsymbol{\beta}(W_t) \cap \boldsymbol{\beta}(W) \neq \emptyset$. So it is sufficient to show there is an l-walk in $\{W_t, t \in B(S)\}$, that satisfies $(C1)$. Recall that $\rho = \lceil 8r_G \log_{d-1} n/\delta^2\rceil$ and by definition of b-choice $\boldsymbol{\beta}(W)$, we have

$$|S| \geqslant 2\delta/r_G - 1.$$

It is easy to see that

$$\begin{aligned}
|B(S)| \geqslant |S|\rho \geqslant (2\delta/r_G - 1)\rho \geqslant (2\delta/r_G - 1)(8r_G \log_{d-1} n/\delta^2)&\\
= 16\log_{d-1} n/\delta - 8r_G \log_{d-1} n/\delta^2&\\
\geqslant 12\log_{d-1} n/\delta,&
\end{aligned}$$

where the last inequality follows from $\delta \geqslant 2r_G$. This means:

$$|\{W_t, t \in B(S)\}| \geqslant 12\log_{d-1} n/\delta.$$

Let x_t, $t \in B(S)$, be an arbitrary node of $V(P') \cap V(W_t)$. Since G has girth $\omega(l)$, we have that if for some $t \in B(S)$, $V(W_t)$ contains u_i (or u_{i+1}), then it also has to contain $[u_i, x_t] \supseteq [u_i, u]$ (or $[x_t, u_{i+1}] \supseteq [v, u_{i+1}]$). Conditioning on $\mathcal{F}_\delta$, $[u_i, u]$ and $[v, u_{i+1}]$ are contained in less than $12\log_{d-1} n/\delta$ l-walks. So there is ball $t_0 \in B(S)$ whose corresponding l-walk W_{t_0}, denoted by W_{P_i}, contains neither u_i nor u_{i+1} and thus it satisfies $(C1)$. Therefore, we conclude that for all $P_i \in \mathcal{P}_k(W)$, W_{P_i} exists and hence $PB(\rho, k)$ on W is valid. □

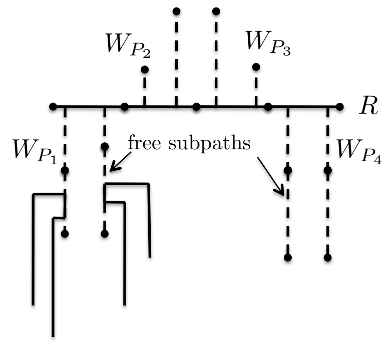

Figure 5.2. The first level $\mathcal{L}_1 = \{W_{P_1}, W_{P_2}, W_{P_3}, W_{P_4}\}$ and the Branch step for free subpaths of W_{P_1}.

Let $U_{n_1,l,h}$ be the event that after allocating at most $n_1 \leqslant n$ balls by $\mathcal{A}(G,l)$ there is a node with load at least $h\rho + c$, where $c = \mathcal{O}(1)$ and $h = \mathcal{O}(\log\log n)$ are positive integers that will be fixed later. Suppose that event $U_{n_1,l,h}$ conditioning on $\mathcal{F}_\delta$ happens. Then there is an l-walk R corresponding to the ball at height $h\rho + c - 1$ with $f(R) \geqslant h\rho + c$. Applying Lemma 5.2.2 shows that $PB(\rho, k)$ on R is valid. So let us define

$$\mathcal{L}_1 := \{W_P, P \in \mathcal{P}_k(R)\},$$

which is called the first level and R is the father of all l-walks in $\mathcal{L}_1$. Condition (C2) in the Partition-Branch procedure ensures that for every $W \in \mathcal{L}_1$,

$$f(W) \geqslant (h-1)\rho + c.$$

Once we have the first level we recursively build the i-th level from the $(i-1)$-th level, for $2 \leqslant i \leqslant h$. Let W be any l-walk in $\mathcal{L}_{i-1}$. We then apply the Partition step on W and get $\mathcal{P}_k(W)$. We say $P \in \mathcal{P}_k(W)$ is a *free* subpath if it does not share any node with W's father. Let W' be W's father. We know that each W except R is created by the Branch step. Thus, by (C1) we have that $\emptyset \neq W \cap W' = [u, v] \subset P$, for some $P \in \mathcal{P}_k(W')$, and hence $d(u, v) \leqslant \lceil l/k \rceil$. Note that since G has girth $\omega(l)$, the intersection of two paths W and W' is a subpath. This implies that $\mathcal{P}_k(W)$ contains at most 2 subpaths that are not free. Let us now choose a set of free subpaths of size $k-2$ denoted by $\mathcal{P}_0(W) \subset \mathcal{P}(W)$. Since for each $W \in \mathcal{L}_{i-1}$, $f(W) \geqslant (h-i+1)\rho + c$, by Lemma 5.2.2 $PB(\rho, k)$ on W is valid. Hence, for each $W \in \mathcal{L}_{i-1}$ we can define set $\mathcal{L}_{i,W} := \{W_P, P \in \mathcal{P}_0(W)\}$, where W is called the father of elements in the set. We now define the i-th level as follows

$$\mathcal{L}_i = \bigcup_{W \in \mathcal{L}_{i-1}} \mathcal{L}_{i,W}.$$

For a graphical view see Figure 5.2. We now present some useful lemmas about the properties of the recursive construction and show how to turn our construction into a subgraph of $\mathcal{H}_{n_1}$. Suppose that $H_j \subset G$, $0 \leqslant j \leqslant h-1$, be the union of all l-walks up to level $j+1$. Then we have the following lemma.

Lemma 5.2.3. *For every $0 \leqslant j \leqslant h-1$, H_j is a tree.*

Proof. When $j=0$, clearly $H_0 = R$, where R is the root. So the diameter of H_0 is l. Assume that for some j_0, $0 \leqslant j_0 < h-1$ the diameter of H_{j_0} is at most $(2j_0+1)l$. We know that every l-walk in the (j_0+1)-th level intersects a path in H_{j_0} so the distance between any two nodes of H_{j_0+1} increases by at most $2l$ and thus the diameter of H_{j_0+1} is at most

$$(2j_0+1)l + 2l = (2(j_0+1)+1)l.$$

So we inductively conclude that H_j, for every $0 \leqslant j \leqslant h-1$, has diameter at most $(2j+1)l$. If for some j, $0 \leqslant j \leqslant h-1$, H_j contains a cycle, then the length of the cycle is at most $2 \cdot \mathtt{diam}(H_i) \leqslant 2(2j+1)l \leqslant 4hl$ which contradicts the fact that $H_j \subset G$ and G has girth at least $\omega(l \log\log n)$. $\square$

Lemma 5.2.4. *For every $1 \leqslant j \leqslant h$, the j-th level contains $k(k-2)^{j-1}$ disjoint l-walks. Moreover every l-walk in the j-the level only intersects one l-walk in the previous levels which is its father.*

Proof. Let us begin with $j=1$. For the sake of a contradiction assume that $W_{P_i}, W_{P_{i'}} \in \mathcal{L}_1$ intersect each other, where $P_i = [u_{i_1}, u_{i+1}], P_{i'} = [u_{i'}, u_{i'+1}] \in \mathcal{P}_k(R)$. l-walks W_{P_i} and $W_{P_{i'}}$ are resulted by the Branch step and hence we can choose two arbitrary nodes $z \in V(P_i) \cap V(W_{P_i})$ and $z' \in V(P_{i'}) \cap V(W_{P_{i'}})$. Also, let $\{u_i, u_{i+1}\}$ and $\{u_{i'}, u_{i'+1}\}$ be the *boundary* of P_i and $P_{i'}$, respectively. Since H_0 is a tree, there is a unique path, say $Q_{z,z'}$, in $H_0 = R$ connecting z to z'. Nodes z and z' have degree 2 in H_0, so $Q_{z,z'}$ contains nodes from boundaries of P_i and $P_{i'}$. By $(C1)$, W_{P_i} and $W_{P_{i'}}$ excludes the boundaries. Thus we get a path from z to z' in $W_{P_i} \cup W_{P_{i'}} \subset H_1$ that excludes the boundaries. This contradicts the fact that there is a unique path in $H_1 \supset H_0$, because H_1 is a tree by Lemma 5.2.3. So we infer that there are k disjoint l-walks in $\mathcal{L}_1$ and they only intersect their father (i.e., R). Also we observe that the nodes contained the free subpaths of each $W \in \mathcal{L}_1$ have degree at most 2 in H_1, which we call the $\mathcal{D}_1$ property. In other word, $\mathcal{D}_1$ property says that any path in H_1 between nodes of two free subpaths in the first level includes nodes from boundaries of the subpaths (see Figure 5.2). Suppose that for some j_0, $1 \leqslant j_0 \leqslant h$, the statement of the lemma and $\mathcal{D}_{j_0}$ hold. Then we show them for the next level as well.

Similar to case $j=1$, toward a contradiction assume that two l-walks $W_P, W_{P'} \in \mathcal{L}_{j_0+1}$ intersect each other. Then, by $(C2)$ we get a path in $W_P \cup W_{P'} \subset H_{j_0+1}$ excluding the boundaries of P and P' that connects one node from P to another node in P'. By $\mathcal{D}_{j_0}$ property, the path in H_{j_0} uses nodes from the boundaries, while we get a path in H_{j_0+1} that exclude boundaries. This is a contradiction because $H_{j_0+1} \supset H_{j_0}$ is a tree by Lemma 5.2.3. So the l-walls in $\mathcal{L}_{j_0+1}$ are disjoint and by the construction we have $|\mathcal{L}_{j_0+1}| = (k-2)|\mathcal{L}_{j_0}|$. It only remains to prove every l-walk only intersect its father in previous levels. Toward a contradiction assume that $W_P \in \mathcal{L}_{j_0+1}$ intersects a path, say W, in previous levels except its father W'. Let $z' \in V(W_P) \cap V(W)$ and $z \in V(W_P) \cap V(P) \subset P$ where $P = [u,v] \in \mathcal{P}_k(W')$. Note that by $(C2)$ z is neither u nor v. We now get a new path from z to z' in H_{j_0+1} excluding u and v that contradicts the fact that there is only one path from z to z' in H_{j_0} including a node from the boundary of P. □

Lemma 5.2.5. *Suppose that G has girth at least $10hl$ and $U_{n_1,l,h}$ conditioning on $\mathcal{F}_\delta$ happens. Then there exists a c-loaded (λ,μ)-tree $T\subset\mathcal{H}_{n_1}$, where $\lambda = 1+k\sum_{j=0}^{h-1}(k-2)^j$ and $\mu = b\cdot k(k-2)^{h-1}$.*

Proof. Since the construction is based on the Partition-Branch procedure, we have that if l-walk W' is the father of W, then $\boldsymbol{\beta}(W)\cap\boldsymbol{\beta}(W')\neq\emptyset$. Let us consider the set of all b-choices that corresponds to l-walks in $\bigcup_{j=0}^{h}\mathcal{L}_j$, where $\mathcal{L}_0=\{R\}$. We connect two b-choices $\boldsymbol{\beta}(W)$ and $\boldsymbol{\beta}(W')$ if W' is the father of W or vice versa. Let $T\subset\mathcal{H}_{n_1}$ denote the resulting graph. By Lemma 5.2.4 for every $1\leqslant j\leqslant h$, the j-th, level contains $k(k-2)^{j-1}$ disjoint l-walks and they intersect either their fathers or their $k-2$ children and consequently we get

$$|V(T)| = \lambda = 1+k\sum_{j=0}^{h-1}(k-2)^j.$$

If we only consider the h-th level, then we get

$$\left|\cup_{\boldsymbol{\beta}\in V(T)}\boldsymbol{\beta}\right| \geqslant \mu = b\cdot k(k-2)^{h-1}.$$

By (C2) in the Partition-Branch procedure we have that $f(W)\geqslant(h-j)\rho+c$, for every $W\in\mathcal{L}_j$, $1\leqslant j\leqslant h$. Hence every node in $\cup_{\boldsymbol{\beta}\in V(T)}\boldsymbol{\beta}$ has load at least c. □

5.3. Main Result

In this section we state our main theorem and its proof. Before that let use recall a set of parameters for given G and l as follows,

$$\begin{aligned}
k &:= \max\{4, \lfloor l/\sqrt{r_G\cdot\log_d n}\rfloor\},\\
\delta &:= \lfloor\lfloor l/k\rfloor/4\rfloor,\\
\rho &:= \lceil 8r_G\log_d n/\delta^2\rceil,
\end{aligned}$$

and $U_{n_1,l,h}$ be the event that at the end of round n_1, there is a nodes with load at least $h\rho+c$, where c is a constant and $h=\left\lceil\frac{\log\log n}{\log(k-2)}\right\rceil$. Note that when $l=(\log n)^{\frac{1+\epsilon}{2}}$ for every constant $\epsilon\in(0,1)$, then k is at least $(\log n)^{\epsilon/3}$ and hence h is a constant. Therefore, in order to apply Lemma 5.2.5 for this case, it is sufficient that G has girth at least $10hl$ or $\omega(l)$. Now we show the following useful lemma.

Lemma 5.3.1. *With probability $1-o(1/n)$, $\mathcal{F}_\delta$ holds.*

Proof. Let us fix an arbitrary path $[u,v]$ of length $\delta=\lfloor\lfloor l/k\rfloor/4\rfloor$. Clearly, if W be an l-walk and $[u,v]\subseteq W=[u_0,u_l]$, then $d(u_0,u)+d(v,u_l)=l-\delta$. Moreover, G is a d-regular graph with girth at least $\omega(l)$, so the total number of different paths of length l containing $[u,v]$ is

$$\sum_{a+b=l-\delta}(d-1)^a(d-1)^b = (l-\delta+1)\cdot(d-1)^{l-\delta}.$$

On the other hand the total number of different paths of length l is $n \cdot d \cdot (d-1)^{l-1}/2$. So the probability that in some round t, $1 \leqslant t \leqslant n$, we get $[u,v] \subseteq W_t$ is at most

$$\frac{2(l-\delta+1)(d-1)^{l-\delta}}{n \cdot d \cdot (d-1)^{l-1}} = \frac{2(l-\delta+1)(d-1)}{n \cdot d \cdot (d-1)^{\delta}} \leqslant \frac{2l}{n(d-1)^{\delta}}.$$

Let $u_\delta = \lceil 6 \log_{d-1} n/\delta \rceil$ and $\{t_1, t_2, \ldots, t_{u_\delta}\} \subset [n]$ be a sequence of distinct rounds of size u_δ. We define indicator random variable $X_{t_1,t_2,\ldots,t_{u_\delta}}([u,v])$, which takes one if $[u,v] \subseteq W_{t_i}$, for every $1 \leqslant i \leqslant u_\delta$, and zero otherwise. Thus we get

$$\begin{aligned}
\mathbf{Pr}\left[X_{t_1,t_2\ldots,t_{u_\delta}}([u,v]) = 1\right] &\leqslant \left(2l/n(d-1)^{\delta}\right)^{u_\delta} \\
&= n^{-u_\delta}\,(d-1)^{(\log_{d-1}(2l)-\delta)u_\delta} \\
&\leqslant n^{-u_\delta}(d-1)^{-u_\delta \cdot \delta/2} = n^{-u_\delta} n^{-3},
\end{aligned}$$

where the last inequality follows from $l = o(\log_d n)$ and hence $\log_{d-1}(2l) \leqslant r_G/2 \leqslant \delta/2$. There are at most n^{u_δ} sequences of rounds of size u_δ and at most $n(d-1)^{\delta-1}$ paths of length δ. Thus, by using the previous upper bound and the union bound over all sequences of rounds and paths of length δ we have

$$\begin{aligned}
&\sum_{\delta\text{-path}} \sum_{t_1,t_2,\ldots,t_{u_\delta}} \mathbf{Pr}\left[X_{t_1,t_2\ldots,t_{u_\delta}}([u,v]) = 1\right] \\
&\leqslant nd(d-1)^{\delta-1} n^{u_\delta}\, \mathbf{Pr}\left[X_{t_1,t_2\ldots,t_{u_\delta}}([u,v]) = 1\right] \\
&\leqslant o(n^2) n^{u_\delta}\, \mathbf{Pr}\left[X_{t_1,t_2\ldots,t_{u_\delta}}([u,v]) = 1\right] = o(1/n),
\end{aligned}$$

where the last inequality follows from $\delta \leqslant l = o(\log_d n)$. This implies that with probability $1 - o(1/n)$ there is no path of length δ contained in at least u_δ l-walks or equivalently $\mathcal{F}_\delta$ holds.

□

Theorem 5.3.2. *With high probability, the maximum load attained by $\mathcal{A}(G,l)$ denoted by m^* is bounded from above as follows:*

I. If $20r_G \leqslant l \leqslant 4\gamma_G$, then we have

$$m^* \leqslant C_1 \cdot \frac{r_G \log_d n \cdot \log\log n}{l^2},$$

where C_1 is a constants.

II. If $4\gamma_G \leqslant l \leqslant \log n$, then we have

$$m^* \leqslant C_2 \cdot \frac{\log\log n}{\log(l/\gamma_G)},$$

where C_2 is a constant.

Proof. By Lemma 5.5.2 we have that $\mathcal{A}(G,l)$ is an (α, n_1)-uniform, where $n_1 = \lfloor n/(6e\alpha)\rfloor$. Let us divide the allocation process into s phases, where s is the smallest integer satisfying $sn_1 \geqslant n$. Let us now focus on the maximum load attained by $\mathcal{A}$ after allocating n_1 balls in the first phase, which is denoted by m_1^*. Let us assume that $U_{n_1,l,h}$ happens. Now, in order to apply Lemma 5.2.5, we only need that G has girth at least $10hl$. By Lemma 5.2.5, if $U_{n_1,l,h}$ conditioning on $\mathcal{F}_\delta$ happens, then there is a c-loaded (λ,μ)-tree T contained in $\mathcal{H}_{n_1}$, where $\lambda = 1 + k\sum_{j=0}^{h-1}(k-2)^j$ and $\mu = b \cdot k(k-2)^{h-1}$. Thus, we get

$$\begin{aligned}\mathbf{Pr}\left[U_{n_1,l,h} \mid \mathcal{F}_\delta\right]\mathbf{Pr}\left[\mathcal{F}_\delta\right] &\leqslant \mathbf{Pr}\left[T \text{ exists} \mid \mathcal{F}_\delta\right]\mathbf{Pr}\left[\mathcal{F}_\delta\right]\\ &= \mathbf{Pr}\left[T \text{ exists and } \mathcal{F}_\delta\right]\\ &\leqslant \mathbf{Pr}\left[T \text{ exists}\right].\end{aligned}$$

Therefore using the above inequality we have

$$\begin{aligned}\mathbf{Pr}\left[U_{n_1,l,h}\right] &= \mathbf{Pr}\left[U_{n_1,l,h} \mid \mathcal{F}_\delta\right]\mathbf{Pr}\left[\mathcal{F}_\delta\right] + \mathbf{Pr}\left[U_{n_1,l,h} \mid \neg\mathcal{F}_\delta\right]\mathbf{Pr}\left[\neg\mathcal{F}_\delta\right]\\ &\leqslant \mathbf{Pr}\left[T \text{ exists}\right] + \mathbf{Pr}\left[\neg\mathcal{F}_\delta\right]\\ &= \mathbf{Pr}\left[T \text{ exists}\right] + o(1/n). \qquad (5.2)\end{aligned}$$

where the last inequality follows from $\mathbf{Pr}\left[\neg\mathcal{F}_\delta\right] = o(1/n)$ by Lemma 5.3.1. By definition of h, we get

$$\lambda = 1 + k(1 + (k-2)^h) \leqslant 2k \log n$$

and

$$\mu = bk(k-2)^{h-1} \geqslant b(k-2)^h \geqslant b \log n.$$

It only remains to bound $\mathbf{Pr}\left[T \text{ exists}\right]$. By applying Lemma 5.1.7 and substituting μ and λ, we conclude that

$$\mathbf{Pr}\left[T \text{ exists}\right] \leqslant n\exp(4\lambda\log b - c\mu) \leqslant n\exp\{-g\log n\},$$

where $g = cb - 8k\log b$. Depending on k we consider two cases. First, $k = 4$. Then it is easy to see there exists a constant c such that $g \geqslant 2$. Second, $k = \lfloor l/\gamma_G\rfloor$. Then we have

$$g \geqslant cl/r_G - 8l\log l^2/\gamma_G = l(c/r_G - 16\log l/\gamma_G),$$

where it follows from $b = l/r_G$ and $b \leqslant l$. Note that we have $1/r_G = 2\log d/\log\log n$ and $l < \log n$. Hence,

$$\frac{\log l}{\gamma_G} = \frac{\log l}{\sqrt{r_G \log_d n}} \leqslant \frac{\log\log n}{\sqrt{2\log\log n\log n/\log d}} = \frac{\log d\sqrt{\log\log n}}{\sqrt{2\log n}} = o(1/r_G).$$

This implies that for some integer $c > 0$, $g = cl/r_G - o(1/r_G) > 2$ and hence in both cases we get $\mathbf{Pr}\left[T \text{ exists}\right] = o(1/n)$. Now, by inequality (5.2) we infer that $m_1^* \leqslant h\rho + c$ with probability $1 - o(1/n)$. In what follows we show the sub-additivity of the algorithm and concludes that in the second phase the maximum load increases by at most m_1^* whp. Assume that we have a copy of G, say G', whose nodes have load exactly m_1^*. Let us consider the allocation process of a pair of balls $(n_1 + t, t)$, for every $0 \leqslant t \leqslant n_1$, by $\mathcal{A}(G,l)$ and $\mathcal{A}(G',l)$. Let $X_u^{n_1+t}$ and Y_u^t, $t \geqslant 0$ denote the load of $u \in V(G) = V(G')$ after allocating balls n_1+t and t by $\mathcal{A}(G,l)$ and $\mathcal{A}(G',l)$, respectively. Now we show that for every integer $0 \leqslant t \leqslant n_1$ and $u \in V(G)$ we have that

$$X_u^{n_1+t} \leqslant Y_u^t. \tag{5.3}$$

When $t=0$, clearly the inequality holds because $Y_u^0 = m_1^*$. We couple the both allocation processes $\mathcal{A}(G,l)$ and $\mathcal{A}(G',l)$ for a given pair of balls (n_1+t,t), $t \geqslant 0$, as follows. We first choose a one-to-one function σ_t from $V(G)$ to $\{1,2,\ldots,n\}$ uniformly at random and let $\sigma_t(u)$, $u \in V(G)$, be the index of u. Note that σ_t is also defined for G' as $V(G) = V(G')$.

Since G' is a copy of G, the coupled process applies $\mathcal{A}(G,l)$ and selects b-choice $\boldsymbol{\beta}_{n_1+t}$ and its copy, say $\boldsymbol{\beta}'_t$, in G'. Then, balls n_1+t and t are allocated on least loaded nodes of $\boldsymbol{\beta}_{n_1+t}$ and $\boldsymbol{\beta}'_t$, respectively, and ties are broken in favor of nodes with minimum index. It is easily checked that the defined process is a coupling. Let us assume that Inequality (5.3) holds for every $t_0 \leqslant t$, then we show it for $t+1$. Let $v \in \boldsymbol{\beta}_{n_1+t+1}$ and $v' \in \boldsymbol{\beta}'_{t+1}$ denote the nodes that are the destinations of pair $(n_1+t+1, t+1)$. Now we consider two cases

1. $X_v^{n_1+t} < Y_v^t$. Then allocating ball n_1+t+1 on v implies that

$$X_v^{n_1+t} + 1 = X_v^{n_1+t+1} \leqslant Y_v^t \leqslant Y_v^{t+1}.$$

So, Inequality (5.3) holds for $t+1$ and every $u \in V(G)$.

2. $X_v^{n_1+t} = Y_v^t$. Since $\boldsymbol{\beta}_{n_1+t+1} = \boldsymbol{\beta}'_{t+1}$, we have that $v \in \boldsymbol{\beta}'_{t+1}$ and $v' \in \boldsymbol{\beta}_{n_1+t+1}$. Also we know that v and v' are nodes with minimum load contained in $\boldsymbol{\beta}_{n+t+1}$ and $\boldsymbol{\beta}_{t+1}$, So we have,

$$X_v^{n_1+t} \leqslant X_{v'}^{n_1+t} \leqslant Y_{v'}^t$$

and also $Y_{v'}^t \leqslant Y_v^t = X_v^{n_1+t}$. Thus,

$$Y_{v'}^t = Y_v^t = X_v^{n_1+t}.$$

If $v \neq v'$ and $\sigma_{t+1}(v') < \sigma_{t+1}(v)$, then it contradicts the fact that ball n_1+t+1 is allocated on v. Similarly, if $\sigma_{t+1}(v') > \sigma_{t+1}(v)$, it contradicts that ball t is allocated on v'. So, we have $v = v'$ and

$$X_v^{n_1+t} + 1 = X_v^{n+t+1} = Y_v^t + 1 = Y_v^{t+1}.$$

So in both cases, Inequality (5.3) holds for every $t \geqslant 0$. If we set $t = n_1$, then the maximum load attained by $A(G',l)$ is at most $2m_1^*$ whp. Therefore, by Inequality (5.3), $2m_1^*$ is an upper bound for the maximum load attained by $\mathcal{A}(G,l)$ in the second phase as well. Similarly, we apply the union bound and conclude that after allocating the balls in s phases, the maximum load m^* is at most sm_1^* with probability $1-o(s/n) = 1-o(1/n)$.

□

5.4. A Lower Bound

In this subsection we derive a lower bound for the maximum load attained by the algorithm based on a second method analysis.

Theorem 5.4.1 (Lower Bound)**.** *Suppose that G be a d-regular n-vertex graph with girth at least $\omega(l)$, where $20r_G \leqslant l \leqslant \mathcal{O}(\gamma_G)$ is an integer, where $\gamma_G = \sqrt{r_G \log_d n}$. Then with probability $1 - n^{-\Omega(1)}$ the maximum load attained by $\mathcal{A}(G, l)$ is at least $\Omega\left(r_G \log_d n / l^2\right)$.*

Proof. We know in each round the algorithm picks a vertex of $V(\mathcal{I}(G, l))$ uniformly at random. Let us define indicator random variable $X_{\boldsymbol{\beta}}$ for every $\boldsymbol{\beta} \in V(\mathcal{I})$ as follows,

$$X_{\boldsymbol{\beta}} := \begin{cases} 1 & \text{if } \boldsymbol{\beta} \text{ is chosen at least } \tau \text{ times by } \mathcal{A}, \\ 0 & \text{otherwise,} \end{cases}$$

where τ will be specified later. By Lemma 5.1.2 we have that

$$|V(\mathcal{I})| = nd(d-1)^{l_0 - 1}/2 \leqslant nd^l/2,$$

where $l_0 = \lfloor l/r_G \rfloor r_G$. Let $\boldsymbol{\beta}$ be an arbitrary vertex of $V(\mathcal{I})$ and $s = |V(\mathcal{I})|$. Thus we get

$$\begin{aligned} \mathbf{Pr}\left[X_{\boldsymbol{\beta}} = 1\right] &= \sum_{i=\tau}^{n} \binom{n}{i} \left(\frac{1}{s}\right)^i \left(1 - \frac{1}{s}\right)^{n-i} \geqslant \left(\frac{n}{s \cdot \tau}\right)^\tau \left(1 - \frac{1}{s}\right)^n \\ &\geqslant \left(\frac{2}{d^l \cdot \tau}\right)^\tau \left(1 - \frac{1}{s}\right)^s \geqslant d^{-(l + \log_d \tau)\tau}/\mathrm{e}, \end{aligned} \tag{5.4}$$

where the second inequality follows from $n \leqslant s \leqslant n \cdot d^l/2$. By setting $\tau = \log_d n/6l$ and using the fact that $\log_d \tau < \log_d \log_d n \leqslant r_G \leqslant l$ we get

$$(l + \log_d \tau)\tau \leqslant \log_d n/6 + \log_d n/6 = \log_d n/3.$$

By substituting the above upper bound in (5.4), we get

$$\mathbf{Pr}\left[X_{\boldsymbol{\beta}} = 1\right] = \Omega(n^{-1/3}).$$

Let us define the random variable $Y = \sum_{\boldsymbol{\beta} \in \mathcal{V}(I)} X_{\boldsymbol{\beta}}$. By linearity of expectation we have

$$\mathbf{E}\left[Y\right] = s \cdot \mathbf{Pr}\left[X_{\boldsymbol{\beta}} = 1\right] = (n \cdot d \cdot (d-1)^{l_0 - 1}/2)\Omega(n^{-1/3}) = \Omega(n^{2/3}). \tag{5.5}$$

It is easily seen that the random variables $X_{\boldsymbol{\beta}}$ and $X_{\boldsymbol{\beta}'}$ are negatively correlated, which means for every $\boldsymbol{\beta}, \boldsymbol{\beta}' \in V(\mathcal{I})$,

$$\mathbf{E}\left[X_{\boldsymbol{\beta}} \cdot X_{\boldsymbol{\beta}'}\right] \leqslant \mathbf{E}\left[X_{\boldsymbol{\beta}}\right] \cdot \mathbf{E}\left[X_{\boldsymbol{\beta}'}\right].$$

This implies that

$$\begin{aligned}\mathbf{Var}\left[Y\right] &= \sum_{\boldsymbol{\beta}\in V(\mathcal{I})} \left(\mathbf{E}\left[X_{\boldsymbol{\beta}}^2\right] - (\mathbf{E}\left[X_{\boldsymbol{\beta}}\right])^2\right) + \underbrace{\sum_{\boldsymbol{\beta}\neq\boldsymbol{\beta}'\in V(\mathcal{I})} \left(\mathbf{E}\left[X_{\boldsymbol{\beta}}X_{\boldsymbol{\beta}'}\right] - \mathbf{E}\left[X_{\boldsymbol{\beta}}\right]\mathbf{E}\left[X_{\boldsymbol{\beta}'}\right]\right)}_{\leqslant 0} \\ &\leqslant \sum_{\boldsymbol{\beta}\in V(\mathcal{I})} \mathbf{E}\left[X_{\boldsymbol{\beta}}^2\right] = \mathbf{E}\left[Y\right].\end{aligned}$$

Applying Chebychev's inequality and above inequality yields that

$$\mathbf{Pr}\left[Y=0\right] \leqslant \mathbf{Pr}\left[|Y-\mathbf{E}\left[Y\right]| \geqslant \mathbf{E}\left[Y\right]\right] = \frac{\mathbf{Var}\left[Y\right]}{(\mathbf{E}\left[Y\right])^2} \leqslant \frac{1}{\mathbf{E}\left[Y\right]}.$$

By equality (5.5) we have that $\mathbf{E}\left[Y\right] = \Omega(n^{2/3})$. Therefore with probability at least $1-\mathcal{O}(n^{-2/3})$ we have $Y \geqslant 1$, which means there exists a vertex $\boldsymbol{\beta}$ that is chosen at least τ times. Since every $\boldsymbol{\beta} \in V(\mathcal{I})$ contains $b = \lfloor l/r_G \rfloor + 1$ elements, by the pigeonhole principle there is a node with load at least

$$\Omega\left(\frac{\tau}{l/r_G}\right) = \Omega\left(\frac{r_G \log_d n}{l^2}\right).$$

□

5.5. Proof of the Key Lemma

In this section we prove our key lemma, which states the algorithm is (α, n_1)-uniform. Before that let us define some notations. For every $S \subseteq V(G)$, $\texttt{Empty}_t(S)$ denotes the number of empty nodes contained in set S after allocating t balls. If it is clear from the context, then we do not mention index t in $\texttt{Empty}_t(S)$. Let $D_r(v)$ denotes the set of all nodes at distance r from node $v \in V(G)$ in graph G. To avoid a lengthy case analysis we do not optimize the constants.

Lemma 5.5.1. *Suppose that with probability* $1 - o(n^{-2})$*, for every* $u \in V(G)$*,* $\texttt{Empty}_t(D_{r_G}(u)) \geqslant |D_{r_G}(u)|/2$*. Then for every* $v \in V(G)$*,*

$$\mathbf{Pr}\left[\textit{ball } t+1 \textit{ is allocated on } v\ \right] \leqslant \frac{\alpha}{n},$$

where α *is a constant.*

Proof. Let $\mathcal{E}_{t+1,v}$ be the event that ball $t+1$ is placed on a given node $v \in V(G)$ and $\mathcal{F}_{t+1}$ be the event that at least $(b-1)/10$ of nodes in $\boldsymbol{\beta}_{t+1}$ are empty. So for every $v \in V(G)$ we have

$$\begin{aligned}\mathbf{Pr}\left[\mathcal{E}_{t+1,v}\right] &= \underbrace{\mathbf{Pr}\left[\mathcal{E}_{t+1,v} | v \notin \boldsymbol{\beta}_{t+1}\right] \cdot \mathbf{Pr}\left[v \notin \boldsymbol{\beta}_{t+1}\right]}_{=0} \\ &\quad + \underbrace{\mathbf{Pr}\left[\mathcal{E}_{t+1,v} | v \in \boldsymbol{\beta}_{t+1} \text{ and } \mathcal{F}_{t+1}\right] \cdot \mathbf{Pr}\left[v \in \boldsymbol{\beta}_{t+1} \text{ and } \mathcal{F}_{t+1}\right]}_{\leqslant (10/b-1)\,\mathbf{Pr}\left[v\in\boldsymbol{\beta}_{t+1}\right]} \\ &\quad + \mathbf{Pr}\left[\mathcal{E}_{t+1,u} | v \in \boldsymbol{\beta}_{t+1} \text{ and } \neg\mathcal{F}_{t+1}\right] \cdot \mathbf{Pr}\left[v \in \boldsymbol{\beta}_{t+1} \text{ and } \neg\mathcal{F}_{t+1}\right] \\ &\leqslant \frac{10}{b-1} \cdot \mathbf{Pr}\left[v \in \boldsymbol{\beta}_{t+1}\right] + \mathbf{Pr}\left[\ \neg\mathcal{F}_{t+1} \mid v \in \boldsymbol{\beta}_{t+1}\right] \mathbf{Pr}\left[v \in \boldsymbol{\beta}_{t+1}\right], \qquad (5.6)\end{aligned}$$

where the first summand follows since if $v \notin \boldsymbol{\beta}_{t+1}$, then ball $t+1$ cannot be placed on v, the second one follows because ties are broken uniformly at random. Let C_i, $1 \leqslant i \leqslant b$, be the event that $v \in \boldsymbol{\beta}_{t+1}$ and $v = u_{(i-1)r_G}$, where $W_{t+1} = (u_0, u_1, \ldots, u_l)$. Conditioning on C_i, without loss of generality, W_{t+1} can be viewed as the union of two edge-disjoint paths W_v^1 and W_v^2 that start with v and their lengths are $(i-1) \cdot r_G$ and $l - (i-1) \cdot r_G$, respectively. Since G has girth at least $\omega(l)$, it locally looks like a d-ary tree and hence the total number of paths of length l with $u_{(i-1)r_G} = v$ is

$$d(d-1)^{(i-1)r_G - 1} \times (d-1)^{l-(i-1)r_G} = d(d-1)^{l-1}.$$

On the other hand in each round, $\mathcal{A}(G, l)$ picks an l-walk randomly from $nd(d-1)^{l-1}$ possible l-walks. Thus we get $\mathbf{Pr}\left[C_i\right] = \frac{d(d-1)^{l-1}}{nd(d-1)^{l-1}} = \frac{1}{n}$, and hence

$$\mathbf{Pr}\left[v \in \boldsymbol{\beta}_{t+1}\right] = \sum_{i=1}^{b} \mathbf{Pr}\left[C_i\right] = \sum_{i=1}^{b} \frac{1}{n} = \frac{b}{n}. \tag{5.7}$$

Now let us compute an upper bound for $\mathbf{Pr}\left[\neg\mathcal{F}_{t+1} | v \in \boldsymbol{\beta}_{t+1}\right]$ that is the second term in (5.6). Conditioning on event $v \in \boldsymbol{\beta}_{t+1}$, we can split W_{t+1} in two subpaths W_v^1 and W_v^2, where both subpaths start with v and $W_{t+1} = W_v^1 \cup W_v^2$. So let us define $\boldsymbol{\beta}_{t+1}^1 = V(W_v^1) \cap \boldsymbol{\beta}_{t+1}$ and $\boldsymbol{\beta}_{t+1}^2 = V(W_v^2) \cap \boldsymbol{\beta}_{t+1}$, where we have $\boldsymbol{\beta}_{t+1} = \boldsymbol{\beta}_{t+1}^1 \cup \boldsymbol{\beta}_{t+1}^2$. Note that by definition of $\boldsymbol{\beta}_{t+1}$, for every $u, u' \in \boldsymbol{\beta}_{t+1}$, $d(u, u') = i \cdot r_G$, where i is an integer. Without loss of generality assume that $s = |\boldsymbol{\beta}_{t+1}^1| \geqslant 2$ and

$$\boldsymbol{\beta}_{t+1}^1 = \{v = u_1, u_2, \ldots, u_s\},$$

where $d(v, u_i) < d(v, u_j)$ for every $1 < i < j \leqslant s$. Then it is clear that every $u_j \in \boldsymbol{\beta}_{t+1}^1$, $2 \leqslant j \leqslant s$, is randomly chosen from a subset of $D_{r_G}(u_{j-1})$, say S_j (because we run a NBRW of length r_G from u_{j-1} to reach u_j). If it happens that the NBRW has already traversed edge $\{z, u_{j-1}\}$, for some z, then the walk cannot take this edge again and hence $|S_j| = (d-1)^{r_G}$. Therefore we have

$$|S_j| \in \{d(d-1)^{r_G - 1}, (d-1)^{r_G}\}.$$

Let $\mathcal{E}_j$ be the event that $\texttt{Empty}_t(D_{r_G}(u_{j-1})) \geqslant d(d-1)^{r_G-1}/2$. If $\mathcal{E}_j$ happens, then the number of nonempty nodes of $D_{r_G}(u_{j-1})$ is at most $d(d-1)^{r_G-1}/2$. We also define an indicator random variable X_u for every $u \in \boldsymbol{\beta}_{t+1} \setminus \{v\}$, which takes one whenever u is empty and zero otherwise. Thus we have

$$\mathbf{Pr}\left[X_{u_j} = 1 \mid \mathcal{E}_j\right] = \frac{\texttt{Empty}_t(S_j)}{|S_j|} \geqslant \begin{cases} \frac{(d-1)^{r_G} - d(d-1)^{r_G-1}/2}{(d-1)^{r_G}} \geqslant 1/4 & \text{if } |S_j| = (d-1)^{r_G}, \\ 1/2 & \text{if } |S_j| = |D_{r_G}(u_{j-1})|, \end{cases}$$

where inequality in the first row follows from $1 - \frac{d}{2(d-1)} \geqslant \frac{1}{4}$ when $d \geqslant 3$. By assumption we have $\mathbf{Pr}\left[\mathcal{E}_j\right] = 1 - o(n^{-2})$ so for every $2 \leqslant j \leqslant s$ we get

$$\begin{aligned} \mathbf{Pr}\left[X_{u_j} = 1\right] &= \mathbf{Pr}\left[X_{u_j} = 1 \mid \mathcal{E}_j\right] \mathbf{Pr}\left[\mathcal{E}_j\right] + \mathbf{Pr}\left[X_{u_j} = 1 \mid \neg\mathcal{E}_j\right] \mathbf{Pr}\left[\neg\mathcal{E}_j\right] \\ &\geqslant 1/4(1 - o(n^{-2})) + o(n^{-2}) \geqslant 1/4 - o(n^{-2}), \end{aligned}$$

Since the above lower bound is independent of any X_{u_i}, $2 \leqslant i \leqslant j$, we have that for every $2 \leqslant j \leqslant s$,

$$\mathbf{Pr}\left[X_{u_j} = 1 \mid X_{u_1} = x_1, \cdots, X_{u_{j-1}} = x_{j-1}\right] \geqslant 1/5.$$

A similar argument also works for $\boldsymbol{\beta}_{t+1}^2$ and we get $\mathbf{Pr}[X_u = 1] \geqslant 1/5$, for every $u \in \boldsymbol{\beta}_{t+1} \setminus \{v\}$. Note that the lower bound for $\mathbf{Pr}[X_u = 1]$, $u \in \boldsymbol{\beta}_{t+1} \setminus \{v\}$ is independent of other X_z's, $z \in \boldsymbol{\beta}_{t+1} \setminus \{u, v\}$. Let $Y = \sum_{u \in \boldsymbol{\beta}_{t+1} \setminus \{v\}} X_u$ be number of empty nodes in $\boldsymbol{\beta}_{t+1} \setminus \{v\}$ then we have that $\mathbf{E}[Y] \geqslant (b-1)/5$. Let Y^* be the summation of $b-1$ independent Bernoulli random variables with success probability $1/5$. Then by applying Lemma 1.3.7 we get,

$$\begin{aligned}\mathbf{Pr}\left[\neg\mathcal{F}_{t+1} | u \in \boldsymbol{\beta}_{t+1}\right] &\leqslant \mathbf{Pr}[Y < (b-1)/10] \\ &\leqslant \mathbf{Pr}[Y^* < \mathbf{E}[Y^*]/2] \leqslant \mathbf{Pr}[|Y^* - \mathbf{E}[Y^*]| \geqslant \mathbf{E}[Y^*]/2].\end{aligned}$$

Now depending on value b we can apply either Chebychev's or a Chernoff inequality to derive an upper bound for the above inequality. We have $\mathbf{Var}[Y^*] \leqslant \mathbf{E}[Y^*]$, so applying Chebychev's bound results into

$$\mathbf{Pr}[|Y^* - \mathbf{E}[Y^*]| \geqslant \mathbf{E}[Y^*]/2] \leqslant \frac{\mathbf{Var}[Y^*]}{\mathbf{E}[Y^*]/4} \leqslant \frac{4}{\mathbf{E}[Y^*]}.$$

Thus we get the following upper bound

$$\mathbf{Pr}\left[\neg\mathcal{F}_{t+1} | u \in \boldsymbol{\beta}_{t+1}\right] \leqslant 4/\mathbf{E}[Y^*] \leqslant 20/(b-1), \tag{5.8}$$

Plugging bounds (5.7) and (5.8) in (5.6) yields that for every $v \in V(G)$,

$$\mathbf{Pr}[\mathcal{E}_{t+1,v}] \leqslant \frac{30b}{n(b-1)}$$

where $30b/(b-1)$ is indeed a constant. □

The proof of the following lemma is similar to [BSSS13, Theorem 1.4].

Lemma 5.5.2 (Key Lemma)**.** *$\mathcal{A}(G, l)$ is an $(\alpha, n/(6\mathrm{e}\alpha))$-uniform allocation algorithm on G, where $1 \leqslant \alpha \leqslant 30b/(b-1)$ is a constant.*

Proof. Let us define potential function $\Phi(t) = \sum_{u \in V(G)} \exp(a_u^t)$, where a_u^t denotes the number of nonempty nodes of $D_{r_G}(u)$ after allocating t balls. It is clear that $\Phi(0) = n$. Let us assume that after allocating t balls we have $\Phi(t) \leqslant n \cdot e^{\Delta/4}$, where $\Delta = d(d-1)^{r_G - 1}$. Then for every $u \in V(G)$,

$$\mathrm{e}^{a_u^t} \leqslant \Phi(t) \leqslant \mathrm{e}^{\log n + \Delta/4}.$$

Recall that $r_G = \lceil 2\log_{d-1}\log n \rceil$. So we get $a_u^t \leqslant \log n + \Delta/4 < \Delta/2$ and consequently $\mathtt{Empty}_t(D_{r_G}(u)) \geqslant \frac{\Delta}{2}$, for every $u \in V(G)$. Let us define indicator random variable $I_{t+1}(u)$ for every $u \in V(G)$ as follows:

$$I_{t+1}(u) := \begin{cases} 1 & \text{if ball } t+1 \text{ is placed on an empty node in } D_{r_G}(u), \\ 0 & \text{otherwise.} \end{cases}$$

Applying Lemma 5.5.1 shows that if $\texttt{Empty}_t(D_{r_G}(u)) \geqslant \frac{\Delta}{2}$, then for every $u \in V(G)$

$$\mathbf{Pr}\left[I_{t+1}(u) = 1\right] \leqslant \frac{\alpha \cdot \texttt{Empty}_t(D_{r_G}(u))}{n} \leqslant \frac{\alpha \cdot \Delta}{n},$$

where α is a constant. So we get

$$\begin{aligned}\mathbf{E}\left[\Phi(t+1) \mid \Phi(t) \leqslant n \cdot e^{\Delta/4}\right] &\leqslant \sum_{u \in V(G)} \left\{\mathbf{Pr}\left[I_{t+1}(u) = 1\right] \cdot e^{a_u^t+1} + \mathbf{Pr}\left[I_{t+1}(u) = 0\right] \cdot e^{a_u^t}\right\} \\ &\leqslant \sum_{u \in V(G)} \left(1 + \frac{\alpha \cdot \mathrm{e} \cdot \Delta}{n}\right) \cdot e^{a_u^t} = \left(1 + \frac{\alpha \cdot \mathrm{e} \cdot \Delta}{n}\right) \Phi(t).\end{aligned}$$

Let us define $\Psi(t) := \min\{\Phi(t), n \cdot e^{\Delta/4}\}$. By using above recursive inequality we have that

$$\mathbf{E}\left[\Psi(t+1)\right] \leqslant \left(1 + \frac{\alpha \cdot \mathrm{e} \cdot \Delta}{n}\right) \Psi(t).$$

Thus, inductively we have that $\mathbf{E}\left[\Psi(t)\right] \leqslant \left(1 + \frac{\alpha \cdot \mathrm{e} \cdot \Delta}{n}\right)^t \Psi(0)$. Let us define $n_1 = n/(6\mathrm{e}\alpha)$. Then applying Markov's inequality implies that

$$\mathbf{Pr}\left[\Psi(n_1) \geqslant n \cdot e^{\Delta/4}\right] \leqslant \frac{\left(1 + \frac{\alpha \cdot \mathrm{e} \cdot \Delta}{n}\right)^{n_1}}{e^{\Delta/4}} \leqslant e^{-\Delta/12}$$

So with probability $1 - n^{-\omega(1)}$, we have $\Phi(n_1) = \Psi(n_1) < n \cdot e^{\Delta/4}$. Since $\Phi(t)$ is an increasing function in t, we have that $\Phi^t \leqslant n \cdot e^{\Delta/4}$, for every $0 \leqslant t \leqslant n_1$, and hence $\texttt{Empty}_t(D_{r_G}(u)) \geqslant \frac{\Delta}{2}$ for every $u \in V(G)$. So applying Lemma 5.5.1 shows that for every $0 \leqslant t \leqslant n_1$ and $u \in V(G)$,

$$\mathbf{Pr}\left[\text{ball } t+1 \text{ is placed on } u \text{ by } \mathcal{A}(G, l)\right] \leqslant \frac{\alpha}{n}.$$

$\square$

Bibliography

[ABKU99] Yossi Azar, Andrei Z. Broder, Anna R. Karlin, and Eli Upfal. Balanced allocations. *SIAM J. Comput.*, 29(1):180–200, 1999.

[ABLS07] Noga Alon, Itai Benjamini, Eyal Lubetzky, and Sasha Sodin. Non-backtracking random walks mix faster. *Communications in Contemporary Mathematics*, 9:585–603, 2007.

[AD86] David Aldous and Persi Diaconis. Shuffling cards and stopping times. *Amer. Math. Monthly*, 93(5):333–348, 1986.

[AD11] Anne Auger and Benjamin Doerr. *Theory of Randomized Search Heuristics: Foundations and Recent Developments.* World Scientific Publishing Co., Inc., River Edge, NJ, USA, 2011.

[AL09] Noga Alon and Eyal Lubetzky. Poisson approximation for non-backtracking random walks. *Israel J. Math.*, 174(1):227–252, 2009.

[BA99] Albert-László Barabási and Réka Albert. Emergence of scaling in random networks. *Science*, 286(5439):509–512, 1999.

[Ban10] Nikhil Bansal. Constructive algorithms for discrepancy minimization. In *Proc. 51st Symp. Foundations of Computer Science (FOCS)*, pages 3–10, 2010.

[BBCS05] Noam Berger, Christian Borgs, Jennifer T. Chayes, and Amin Saberi. On the S pread of V iruses on the I nternet. In *Proc. 16th Symp. Discrete Algorithms (SODA)*, pages 301–310, 2005.

[BBFN12] Petra Berenbrink, André Brinkmann, Tom Friedetzky, and Lars Nagel. Balls into bins with related random choices. *J. Parallel Distrib. Comput.*, 72(2):246–253, 2012.

[BCM04] John W. Byers, Jeffrey Considine, and Michael Mitzenmacher. Geometric generalizations of the power of two choices. In *Proc. 16th Symp. Parallelism in Algorithms and Architectures (SPAA)*, pages 54–63, 2004.

[BEF08] Petra Berenbrink, Robert Elsässer, and Tom Friedetzky. Efficient randomised broadcasting in random regular networks with applications in peer-to-peer systems. In *Proc. 27th Symp. Principles of Distributed Computing (PODC)*, pages 155–164, 2008.

[Bel98] Eric D. Belsley. Rates of convergence of random walk on distance regular graphs. *Probab. Theory Related Fields*, 112(4):493–533, 1998.

[BGPS06] Stephen Boyd, Arpita Ghosh, Balaji Prabhakar, and Devavrat Shah. Randomized gossip algorithms. *IEEE Transactions on Information Theory*, 52(6):2508–2530, 2006.

[BR03] Béla Bollobás and Oliver Riorden. Mathematical results on scale-free random graphs. *Handbook of graphs and networks, Wiley-VCH, Weinheim*, pages 1–34, 2003.

[BSSS13] Paul Bogdan, Thomas Sauerwald, Alexandre Stauffer, and He Sun. Balls into bins via local search. In *Proc. 24th Symp. Discrete Algorithms (SODA)*, pages 16–34, 2013.

[CF13] Colin Cooper and Alan Frieze. The height of random k-trees and related branching processes. *arXiv*, 1309.4342v2 [math.CO], 2013.

[CHHKM12] Keren Censor-Hillel, Bernhard Haeupler, Jonathan A. Kelner, and Petar Maymounkov. Global computation in a poorly connected world: fast rumor spreading with no dependence on conductance. In *Proc. 44th Symp. Theory of Computing (STOC)*, pages 961–970, 2012.

[CL03] Fan R. K. Chung and Linyuan Lu. The average distance in a random graph with given expected degrees. *Internet Mathematics*, 1(1):91–113, 2003.

[CLP09] Flavio Chierichetti, Silvio Lattanzi, and Alessandro Panconesi. Rumor spreading in social networks. In *Proc. 36th Intl. Coll. Automata, Languages and Programming (ICALP)*, pages 375–386, 2009.

[CLV03] Fan R. K. Chung, Linyuan Lu, and Van H. Vu. The spectra of random graphs with given expected degrees. *Internet Mathematics*, 1(3):257–275, 2003.

[CMOS11] Artur Czumaj, Morteza Monemizadeh, Krzysztof Onak, and Christian Sohler. Planar graphs: Random walks and bipartiteness testing. In *Proc. 52nd Symp. Foundations of Computer Science (FOCS)*, pages 423–432, 2011.

[CSC08a] Guan-Yu Chen and Laurent Saloff-Coste. The cutoff phenomenon for ergodic markov processes. *Electron. J. Probab.*, 13(3):26–78, 2008.

[CSC08b] Guan-Yu Chen and Laurent Saloff-Coste. The cutoff phenomenon for randomized riffle shuffles. *Random Structures Algorithms*, 32(3):346–372, 2008.

[CU10] Colin Cooper and Ryuhei Uehara. Scale free properties of random k-trees. *Mathematics in Computer Science*, 3(4):489–496, 2010.

[DF11] Benjamin Doerr and Mahmoud Fouz. Asymptotically optimal randomized rumor spreading. In *Proc. 38th Intl. Coll. Automata, Languages and Programming (ICALP)*, pages 502–513, 2011.

[DFF11] Benjamin Doerr, Mahmoud Fouz, and Tobias Friedrich. Social networks spread rumors in sublogarithmic time. In *Proc. 43th Symp. Theory of Computing (STOC)*, pages 21–30, 2011.

[DFF12a] Benjamin Doerr, Mahmoud Fouz, and Tobias Friedrich. Asynchronous rumor spreading in preferential attachment graphs. In *Proc. 13th Scandinavian Workshop Algorithm Theory (SWAT)*, pages 307–315, 2012.

[DFF12b] Benjamin Doerr, Mahmoud Fouz, and Tobias Friedrich. Why rumors spread so quickly in social networks. *Commun. ACM*, 55(6):70–75, 2012.

[DGG+87] Alan Demers, Mark Gealy, Dan Greene, Carl Hauser, Wes Irish, John Larson, Sue Manning, Scott Shenker, Howard Sturgis, Dan Swinehart, Doug Terry, and Don Woods. Epidemic algorithms for replicated database maintenance. In *Proc. 6th Symp. Principles of Distributed Computing (PODC)*, pages 1–12, 1987.

[DGM90] Persi Diaconis, Ron L. Graham, and John A. Morrison. Asymptotic analysis of a random walk on a hypercube with many dimensions. *Random Structures Algorithms*, 1(1):51–72, 1990.

[DH96] Carl Dou and Martin Hildebrand. Enumeration and random random walks on finite groups. *Ann. Probab.*, 24(2):987–1000, 1996.

[Dia88] Persi Diaconis. *Group Representations in Probability and Statistics.* IMS, 1988.

[Dia96] Persi Diaconis. The cutoff phenomenon in finite markov chains. *Proc. Nat. Acad. Sci. USA*, 93(4):1659–1664, 1996.

[DP09] Devdatt P. Dubhashi and Alessandro Panconesi. *Concentration of Measure for the Analysis of Randomized Algorithms.* Cambridge University Press, 2009.

[Drm09] Michael Drmota. *Random trees: An interplay between combinatorics and probability.* SpringerWienNewYork, Vienna, 2009.

[DS81] Persi Diaconis and Mehrdad Shahshahani. Generating a random permutation with random transposition. *Z. Wahrsch. Verw. Gebiete*, 57(2):159–179, 1981.

[DS87] Persi Diaconis and Mehrdad Shahshahani. Time to reach stationarity in the bernoulli-laplace diffusion model. *SIAM J. Math. Anal.*, 18(1):208–218, 1987.

[Els06] Robert Elsässer. On the communication complexity of randomized broadcasting in random-like graphs. In *Proc. 18th Symp. Parallelism in Algorithms and Architectures (SPAA)*, pages 148–157, New York, NY, USA, 2006.

[FDP06] Philippe Flajolet, Philippe Dumas, and Vincent Puyhaubert. Some exactly solvable models of urn process theory. In *Fourth Colloquium on Mathematics and Computer Science Algorithms, Trees, Combinatorics and Probabilities*, Discrete Math. Theor. Comput. Sci. Proc., AG, pages 59–118. Assoc. Discrete Math. Theor. Comput. Sci., Nancy, 2006.

[FG85] Alain Frieze and Geoffrey Grimmett. The shortest-path problem for graphs with random-arc-lengths. *Discrete Applied Mathematics*, 10:57–77, 1985.

[FGS12] Tobias Friedrich, Martin Gairing, and Thomas Sauerwald. Quasirandom load balancing. *SIAM J. Comput.*, 41(4):747–771, 2012.

[FHP10] Nikolaos Fountoulakis, Anna Huber, and Konstantinos Panagiotou. Reliable broadcasting in random networks and the effect of density. In *Proc. 29th IEEE Conf. Computer Communications (INFOCOM)*, pages 2552–2560, 2010.

[FP10] Nikolaos Fountoulakis and Konstantinos Panagiotou. Rumor spreading on random regular graphs and expanders. In *Proc. 14th Intl. Workshop on Randomization and Comput. (RANDOM)*, pages 560–573, 2010.

[FPRU90] Uriel Feige, David Peleg, Prabhakar Raghavan, and Eli Upfal. Randomized broadcast in networks. *Random Struct. Algorithms*, 1(4):447–460, 1990.

[FPS12] Nikolaos Fountoulakis, Konstantinos Panagiotou, and Thomas Sauerwald. Ultra-fast rumor spreading in social networks. In *Proc. 23th Symp. Discrete Algorithms (SODA)*, pages 1642–1660, 2012.

[Gao09] Yong Gao. The degree distribution of random k-trees. *Theor. Comput. Sci.*, 410(8-10):688–695, 2009.

[Gia11] George Giakkoupis. Tight bounds for rumor spreading in graphs of a given conductance. In *Proc. 28th Symp. Theoretical Aspects of Computer Science (STACS)*, pages 57–68, 2011.

[Gia14] George Giakkoupis. Tight bounds for rumor spreading with vertex expansion. In *Proc. 25th Symp. Discrete Algorithms (SODA)*, pages 801–815, 2014.

[God80] Chris Godsil. More odd graph theory. *Discrete Math.*, 32(2):205–207, 1980.

[God08] Brighten Godfrey. Balls and bins with structure: balanced allocations on hypergraphs. In *Proc. 19th Symp. Discrete Algorithms (SODA)*, pages 511–517, 2008.

[GR01] Chris Godsil and Gordon F. Royle. *Algebraic Graph Theory, Graduate Texts in Mathematics, 207*. Springer-Verlag, 2001.

[GS01] Geoffrey Grimmett and David Stirzaker. *Probability and Random Processes*. Oxford University Press, 3rd edition, 2001.

[GT12] Shayan Oveis Gharan and Luca Trevisan. Approximating the expansion profile and almost optimal local graph clustering. In *Proc. 53rd Symp. Foundations of Computer Science (FOCS)*, pages 187–196, 2012.

[HBLL99] Mor Harchol-Balter, Frank Thomson Leighton, and Daniel Lewin. Resource discovery in distributed networks. In *Proc. 18th Symp. Principles of Distributed Computing (PODC)*, pages 229–237, 1999.

[HHL88] Sandra M. Hedetniemi, Stephen T. Hedetniemi, and Arthur L. Liestman. A survey of gossiping and broadcasting in communication networks. *Networks*, 18(4):319–349, 1988.

[HKP$^+$05] Juraj Hromkovic, Ralf Klasing, Andrzej Pelc, Peter Ruzicka, and Walter Unger. *Dissemination of Information in Communication Networks - Broadcasting, Gossiping, Leader Election, and Fault-Tolerance*. Texts in Theoretical Computer Science. An EATCS Series. Springer, 2005.

[IJ90] Amos Israeli and Marc Jalfon. Token management schemes and random walks yield self-stabilizing mutual exclusion. In *Proc. 9th Symp. Principles of Distributed Computing (PODC)*, pages 119–131, 1990.

[JK77] Norman L. Johnson and Samuel Kotz. *Urn models and their application: An approach to modern discrete probability theory, Wiley Series in Probability and Mathematical Statistics*. John Wiley & Sons, New York-London-Sydney, 1977.

[KDG03] David Kempe, Alin Dobra, and Johannes Gehrke. Gossip-based computation of aggregate information. In *Proc. 44th Symp. Foundations of Computer Science (FOCS)*, pages 482–491, 2003.

[KKD04] David Kempe, Jon M. Kleinberg, and Alan J. Demers. Spatial gossip and resource location protocols. *J. ACM* , 51(6):943–967, 2004.

[KKV13] Istvan Kolossvary, Julia Komjathy, and Lajos Vago. Degress and distane in random evolving apollonian networks. *arXiv:1310.3864 [math.PR]*, 2013.

[Kle00] Jon M. Kleinberg. The small-world phenomenon: an algorithm perspective. In *Proc. 32nd Symp. Theory of Computing (STOC)*, pages 163–170, 2000.

[Klo94] Ton Kloks. *Treewidth: Computations and Approximations*. Springer-Verlag, 1994.

[KN12] Ravindran Kannan and Hariharan Narayanan. Random walks on polytopes and an affine interior point method for linear programming. *Math. Oper. Res.*, 37(1):1–20, 2012.

[Knu97] Donald Knuth. *The Art of Computer Programming, Vol. 1: Fundamental Algorithms*. Adison-Wesley, third edition, 1997.

[KP06] Krishnaram Kenthapadi and Rina Panigrahy. Balanced allocation on graphs. In *Proc. 17th Symp. Discrete Algorithms (SODA)*, pages 434–443, 2006.

[KSSV00] Richard M. Karp, Christian Schindelhauer, Scott Shenker, and Berthold Vöcking. Randomized Rumor Spreading. In *Proc. 41st Symp. Foundations of Computer Science (FOCS)*, pages 565–574, 2000.

[Laf84] Andrea Laforgia. Further inequalities for the gamma function. *Mathematics of Computation*, 42(166):597–600, 1984.

[LPW09] David A. Levin, Yuval Peres, and Elizabeth L. Wilmer. *Markov Chains and Mixing Times*. AMS, 2009.

[LS93] László Lovász and Miklós Simonovits. Random walks in a convex body and an improved volume algorithm. *Random Struct. Algorithms*, 4(4):359–412, 1993.

[LS10] Eyal Lubetzky and Allan Sly. Cutoff phenomenon for random walks on random regular graphs. *Duke Math. J.*, 153(3):475–510, 2010.

[LS11] Eyal Lubetzky and Allan Sly. Explicit expanders with cutoff phenomena. *Electron. J. Probab.*, 16(15):419–435, 2011.

[Mah03] Hosam M. Mahmoud. Pólya urn models and connections to random trees: A review. *Journal of the Iranian Statistical Society*, 2(1):53–114, 2003.

[McD98] Colin McDiarmid. Concentration. In Michel Habib, Colin McDiarmid, Jorge Ramirez-Alfonsin, and Bruce Reed, editors, *Probabilistic Methods for Algorithmic Discrete Mathematics*, pages 195–243. Springer-Verlag, 1998.

[MP14] Abbas Mehrabian and Ali Pourmiri. Randomized rumor spreading in poorly connected small-world networks. In *Proc. 28th Intl. Symp. Distributed Computing (DISC)*, pages 346–360, 2014.

[MPS03] Milena Mihail, Christos H. Papadimitriou, and Amin Saberi. On certain connectivity properties of the internet topology. In *Proc. 44th Symp. Foundations of Computer Science (FOCS)*, pages 28–35, 2003.

[MR95] Rajeev Motwani and Prabhakar Raghavan. *Randomized algorithms*. Cambridge University Press, Cambridge, 1995.

[MRS01] Micheal Mitzenmacher, Andréa W. Richa, and Ramesh Sitaraman. The power of two random choices: A survey of technique and results. *In Handbook of Randomized Computation Volume 1*, pages 255–312, 2001.

[Mun11] Muhittin Mungan. Comment on "apollonian networks: Simultaneously scale-free, small world, euclidean, space filling, and with matching graphs". *Phys. Rev. Lett.*, 106:029802, Jan 2011.

[Per04] Yuval Peres. Sharp thresholds for mixing times. *American Institute of Mathematics (AIM) research workshop (Palo Alto). Summary available at http://www.aimath.org/WWN/mixingtimes*, 2004.

[Pit87] Boris Pittel. On spreading a rumor. *SIAM Journal on Applied Mathematics*, 47(1):213–223, 1987.

[Pou14] Ali Pourmiri. Balanced allocation on graphs: A random walk approach. *CoRR*, abs/1407.2575, 2014.

[PPS13] Konstantinos Panagiotou, Ali Pourmiri, and Thomas Sauerwald. Faster rumor spreading with multiple calls. In *Proc. 24th Intl. Symp. Algorithms and Computation (ISAAC)*, pages 446–456, 2013.

[PPS15] Konstantinos Panagiotou, Ali Pourmiri, and Thomas Sauerwald. Faster rumor spreading with multiple calls. *Electr. J. Comb.*, 22(1):P1.23, 2015.

[PS97] Alessandro Panconesi and Aravind Srinivasan. Randomized distributed edge coloring via an extension of the chernoff-hoeffding bounds. *SIAM J. Comput.*, 26(2):350–368, 1997.

[PS14] Ali Pourmiri and Thomas Sauerwald. Cutoff phenomenon for random walks on kneser graphs. *Discrete Applied Mathematics*, 176:100–106, 2014.

[Rei00] Philipp Reinfeld. Chromatic polynomials and the spectrum of the kneser graph. *CDAM Research Report, LSE-CDAM-2000-02. Available at http://www.cdam.lse.ac.uk/Reports/reports2000.html*, 2000.

[SC04] Laurent Saloff-Coste. Total variation lower bounds for finite markov chains: Wilson's lemma. *Random walks and geometry*, pages 515–532, 2004.

[Vöc03] Berthold Vöcking. How asymmetry helps load balancing. *J. ACM*, 50(4):568–589, 2003.

[vRMH98] Robbert van Renesse, Yaron Minsky, and Mark Hayden. A gossip-style failure detection service. In *Proceedings of the 15th IFIP International Conference on Distributed Systems Platforms (Middleware'98)*, pages 55–70, 1998.

[Wil04] David B. Wilson. Mixing times of lozenge tiling and card shuffling markov chains. *Ann. Appl. Probab.*, 14(1):274–325, 2004.

[WS98] Duncan J. Watts and Steven H. Strogatz. Collective dynamics of 'small-world' networks. *Nature*, 393:440–442, 1998.

[ZCFR06] Zhongzhi Zhang, Francesc Comellas, Guillaume Fertin, and Lili Rong. High-dimensional apollonian networks. *J. Phys. A*, 39(8):1811–1818, 2006.

[ZYW05] Tao Zhou, Gang Yan, and Bing-Hong Wang. Maximal planar networks with large clustering coefficient and power-law degree distribution. *Phys. Rev. E*, 71:046141, Apr 2005.