A Probabilistic Analysis of Kademlia Networks

Xing Shi Cai

Master of Science

School of Computer Science

McGill University

Montreal, Quebec

2012-08-01

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Master of Science

ACKNOWLEDGEMENTS

I would like to acknowledge the following persons who have made this work possible:

My supervisor, Luc Devroye, who intrigued my interest in probability theory with his excellent course COMP 690, and who has spent countless hours giving me invaluable, insightful, and sometimes hilarious instructions and encouragement.

Carlton Davis of École Polytechnique de Montréal, who initially explained Kademlia to us and sparked our interest in the problem.

Mahshid Yassaei, a co-student in 690, who first introduced Kademlia to us.

Shitai Liu, my friend and fellow computer science student at McGill, who encouraged me to ask Luc to be my supervisor.

Omkar Deshmukh, who is such a nice office mate.

My parents, who have given me so much support.

ABSTRACT

Nowadays Kademlia [17] is one of the most widely used DHTs (Distributed Hash Table) [2] in P2P (peer-to-peer) networks. This work studies one essential question about Kademlia overlay networks from a mathematical perspective: how long does it take to locate a node? To answer it, we introduce a random graph $\mathcal{K}$ to model a Kademlia overlay and study how long it takes to locate a given vertex in $\mathcal{K}$ by using Kademlia's routing algorithm.

ABRÉGÉ

Aujourd'hui Kademlia [17] est l'un des les plus utilisés DHTs (Distributed Hash Tableau) dans les réseaux P2P (peer-to-peer) [2]. Cet article étudie une question essentielle des réseaux "overlay" de Kademlia d'un point de vue mathématique: combien de temps faut-il pour localiser un noeud? Pour y répondre, nous introduisons un graphe aléatoire $\mathcal{K}$ pour modéliser un réseau de Kademlia et étudier la complexité d'un algorithme de routage de Kademlia.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1
Introduction

1.1 Peer-to-peer Networks

A peer-to-peer (P2P) network is defined as a computer network in which each computer can both act as client and server at the same time, which allows computing, storage and bandwidth resources to be shared among all participants of the network [27]. While in traditional client-server architectures, a client only consumes and a server only provides resources, in a P2P network each participating computer, often referred to as a *node*, is both a consumer and a supplier of resources.

P2P networks have many advantages compared to centralized client-sever systems [19]. For example, since all the resources come from participating nodes, the cost to set up a P2P network is very low. When new nodes join a P2P network, although the consumption of services increases, the total capacity also increases. By removing centralized servers, a P2P network increases fault-tolerance as there is no longer a single point of failure. For users who care about privacy issues, P2P networks provide better anonymity because in client-sever systems the central server is usually able to identify a client.

The first P2P file sharing service, Napster [18], which was released in 1999, reportedly had 26.4 million unique users in one month at its peak. The more recent P2P file sharing protocol BitTorrent [3] estimated to be responsible for more than 40% of all internet traffic in 2009. Skype, the P2P voice-over-internet service, announced that they had more than 600 million registered users in 2009. Today, P2P networks have fundamentally changed how we use

the internet and have exerted considerable social and economic impact on our society.

P2P networks have been a popular research domain for years and plenty of literature on this topic has been written. Oram [22] compiled a book which covers most pioneering P2P projects that emerged around the beginning of this century. Androutsellis-Theotokis and Spinellis [1] wrote a comprehensive survey on content distribution P2P services. Lua et al. [16] compared and categorized various structures of P2P overlay networks. Risson and Moors [25]'s survey focused on classifying P2P networks by their searching methods. Liu et al. [15] surveyed recent applications of P2P techniques for video streaming. Steinmetz and Wehrle [31] wrote a comprehensive book which discusses a broad spectrum of topics about P2P techniques.

1.2 Distributed Hash Table

The great success of P2P networks raises many interesting challenges for computer scientists. One of great importance is the *lookup problem*. Given a data item, say a file, how do we find it efficiently? This problem can be handled by having a powerful central database which maps file names to network addresses of nodes storing the file, just as Napster did for music files. This approach greatly raises costs, introduces a single point of failure, and has an efficiency problem. Some systems, like DNS (Domain Name System) [20] [21], bring in hierarchy to alleviate the pressure on central servers. But failure of high level nodes in the hierarchy can still be disastrous. Some P2P networks, like Gnutella [24], try to avoid central databases and hierarchy by flooding the whole system with messages requesting a data item. However, there is a considerable cost of bandwidth and computation.

To overcome these problems, in the first few years of this century, researchers invented a group of P2P algorithms called DHT which all provide a simple and general hash-table style interface, including CAN [23], Chord [32], Pastry [26], Tapestry [34], and Kademlia [17]. A DHT allows data items to be stored in its participating nodes in a load-balanced way and allows them to be retrieved efficiently. We briefly describe how DHT works here as the background, although it is not needed to understand our model.

In a DHT, participating nodes are identified with a numeric ID. Given a data item, a DHT first maps it to an ID, often called a *key*, which does not necessarily belong to any node in the network. Then nodes whose IDs are close to the key are required to store the data item. To be precise about what "close" means, given two IDs x and y, a DHT must define a *distance function* $\delta(x, y)$ to calculate the "closeness" between them.

To store data in and retrieve data from nodes with specific IDs, their network addresses must be found. Thus, an important problem that a DHT must deal with is the *routing problem*: given an ID, how to find the network addresses of nodes that are close to it efficiently? Due to the potential huge size of a P2P network, it is impossible to require each node to have omniscient knowledge of the whole system. Instead, a node can only maintain a data structure called a *routing table*, which contains information, like the IDs and the network addresses, of a few other nodes, which are referred to as its *neighbors*.

The neighbor relationships in a DHT can be represented with a directed graph $G = (\mathcal{V}, \mathcal{E})$. Let each vertex in $\mathcal{V}$ represent a node. Let an arc $(u, v) \in \mathcal{E}$ if and only if v is in the routing table of u. Some call this graph G the *topology* of a DHT. The routing problem can be rephrased as a graph theory problem: for all vertices $u \in \mathcal{V}$, given an ID y, how to efficiently find a path from u to

y', where y' is the vertex whose ID is closest to y among all vertices in $\mathcal{V}$. The routing problem in Kademlia is the main topic of this work.

For further information of DHT, Balakrishnan et al. [2] wrote a clean and concise summary of all the major DHT algorithms.

1.3 Kademlia as a DHT

Kademlia [17] as a DHT chooses $\{0,1\}^d$ as its ID space, where d is usually 128 [29] or 160 [4]. When a node joins Kademlia, it picks an ID uniformly at random from $\{0,1\}^d$ so that the uniqueness of IDs can be guaranteed by the huge cardinality of $\{0,1\}^d$. Therefore, IDs are all d-bit binary vectors in Kademlia. Given two such vectors, Kademlia defines the distance between the two as the bit-wise XOR (exclusive-or) of them. To be precise, given $x = (x_1, \ldots, x_d)$ and $y = (y_1, \ldots, y_d)$, Kademlia defines its distance function by

$$\delta(x, y) = \sum_{i=1}^{d} (x_i \oplus y_i) \times 2^{d-i},$$

where $\oplus$ denotes XOR operation

$$u \oplus v = \begin{cases} 1 & \text{if } u \neq v, \\ 0 & \text{otherwise.} \end{cases}$$

Kademlia stipulates that given a data item, it should be stored in the k nodes whose IDs are closest to the its key, where k is a system-wide constant, which usually equals 8 [4], 10 [29] or 20 [7]. Note that mapping a data item to a key, storing and retrieving it in Kademlia are not discussed in this work.

The routing algorithm of Kademlia is sometimes referred to as "tree-like routing" [2], as each node maintains a tree-like structure as its routing table. More specifically, Kademlia defines its routing table based on the data structure called a *prefix tree*, also known as a *trie*. A *trie* is an ordered tree data structure

invented by Fredkin [9], which is generally used to store associative arrays where keys are strings. In a trie, a node's position determines a unique string with which it is associated. All descendants of this node share the same prefix of the string associated with it. Thus values can be stored in the node associated with the corresponding key and be retrieved efficiently by traversing downwards from the root. Figure 1–1 gives an example of a trie that stores some English words. For more information on tries, see [33].

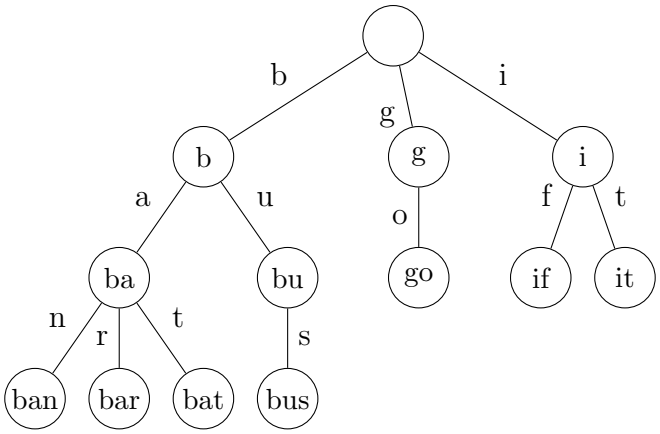

Figure 1–1: An example of a trie for alphabetic strings.

Let $G = (\mathcal{V}, \mathcal{E})$ be Kademlia's *topology*. If we write an ID as a string consisting of zeros and ones, from higher bits to lower bits, then we can completely represent $\mathcal{V}$ as a trie, with each leaf associated with a unique ID, as depicted in Figure 1–2. It is easy to verify that, in this ID trie, the position of the lowest common ancestor of two leaves bounds their XOR distance from below and above. Let x and y be two different IDs. Let $\ell(x, y)$ be the length of the path from the root to x and y's lowest common ancestor, i.e., the length of x and y's common prefix, then we have

$$\ell(x, y) = \max\{i : x_1 = y_1, \ldots, x_i = y_i\},$$

which bounds $\delta(x, y)$, the XOR distance of x and y, on both sides as

$$2^{d-\ell(x,y)-1} \leq \delta(x, y) < 2^{d-\ell(x,y)}.$$

Thus given an ID x, if we partition all other IDs as follows,

$$\mathcal{D}_i(x) = \{y \in \mathcal{V} - \{x\} : \delta(x, y) \in [2^{i-1}, 2^i)\}, \quad i = 1, \ldots, d,$$

then $\mathcal{D}_i(x)$ is equivalent to a subtree in the trie, in which each node shares a common prefix of length i with x. Put differently, we have the equivalent definition

$$\mathcal{D}_i(x) = \{y \in \mathcal{V} - \{x\} : \ell(x, y) = d - i\}, \quad i = 1, \ldots, d.$$

Kademlia requires that node x can only have up to k neighbors in each of these subtrees, where k is the system wide constant mentioned before. For each subtree, x maintains a list which contains the information of its neighbors in that subtree. This list is referred to as the k-bucket in Kademlia literature, since at most k nodes' information can be added to it. Together, all x's k-buckets form its routing table and decide its neighbors. We denote the k-bucket of x that is associated with subtree $\mathcal{D}_i(x)$ as $\mathcal{B}_i(x)$, i.e.,

$$\mathcal{B}_i(x) = \{y \in \mathcal{D}_i(x) : (x, y) \in \mathcal{E}\}.$$

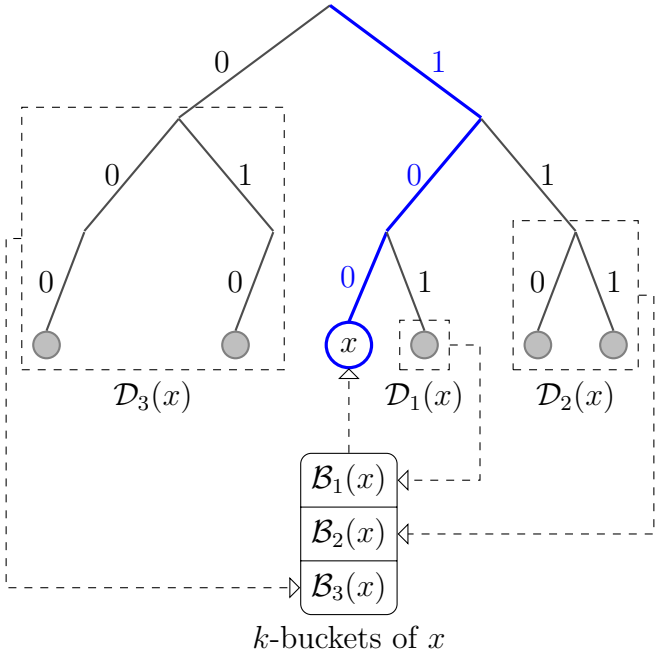

Figure 1–2: An example of Kademlia ID trie and k-buckets. Given an ID $x = (1, 0, 0)$, the trie is partitioned into subtrees $\mathcal{D}_1(x), \mathcal{D}_2(x), \mathcal{D}_3(x)$. x maintains a k-bucket for each of these subtrees containing the information of up to k nodes in the subtree.

To deal with the routing problem, Kademlia defines a message FIND_NODE. When node x needs to find the node that is closest to ID y, it sends FIND_NODE to α adjacent nodes that are closest to y among its neighbors, where α is a system-wide constant, sometimes chosen as 3 [30] or 10 [7]. The recipient of this message must return information of k of its neighbors that are closest to y. From this additional information, x again sends FIND_NODE to the α nodes that are closest to y. This process continues until x can no longer find nodes that are closer to y than those it has already contacted.

We briefly describe how Kademlia updates routing tables here as part of the motivation of our model, but it is not a must for understanding this work. Initially, when x joins Kademlia, its k-buckets are all empty. Later when x receives a message from another node y which is not x's neighbor yet, the

sender's information is added to the appropriate k-bucket, say $\mathcal{B}_i(x)$, unless $\mathcal{B}_i(x)$ is already full. In the latter case, x contacts the least recently seen node in $\mathcal{B}_i(x)$, say z. If z does not respond in time, then z is removed from and y is added to $\mathcal{B}_i(x)$; otherwise x leaves $\mathcal{B}_i(x)$ unchanged. Kademlia usually relies on the traffic between nodes to keep routing tables up to date. But if it happens that x does not communicate with its neighbors in a particular subtree, e.g., $\mathcal{D}_i(x)$, for a long time, Kademlia requires x to generate a random ID within $[2^{i-1}, 2^i)$ and try to locate it, which creates traffic to and from its neighbors in $\mathcal{D}_i(x)$.

Kademlia has been implemented in many popular P2P networks and has aroused great interest among researchers. Crosby and Wallach [4] examined two major Kademlia overlay networks used by BitTorrent clients, each of which has more than a million live nodes. Steiner et al. [29] measured various properties of the Kad network, a DHT based on Kademlia with several million simultaneous users. As Grizzard et al. [11] presented, Kademlia is also used by attackers to establish a more resilient botnet, i.e., a collection of infiltrated computers controlled by attackers. Davis et al. [6] assessed different disinfection strategies against a Kademlia-based botnet.

1.4 A random graph model for Kademlia

To model a Kademlia DHT with n nodes, we define a directed graph $\mathcal{K}$ whose edges are all chosen at random. Let $\mathcal{V}(\mathcal{K})$ be the vertex set of $\mathcal{K}$. We define $\mathcal{V}(\mathcal{K}) = \{X_1, \ldots, X_n\}$, where $X_1, \ldots, X_n$ are selected from $\{0,1\}^d$ without replacement. (Note that we put aside how $X_1, X_2, \ldots, X_n$ are selected, deterministic or at random, for now.) Thus vertices in $\mathcal{V}(\mathcal{K})$ can be seen as the n nodes in Kademlia. Given a vertex $x \in \mathcal{V}(\mathcal{K})$, we randomly decide its k-bucket $\mathcal{B}_i(x)$ as follows. If $|\mathcal{D}_i(x)| \leq k$, then $\mathcal{B}_i(x) = \mathcal{D}_i(x)$; otherwise let $\mathcal{B}_i(x)$

be a sample of size k uniformly chosen from $\mathcal{D}_i(x)$ without replacement. Put differently, we fill up each k-bucket of x uniformly at random.

We introduce two models which differ in how $X_1, X_2, \ldots, X_n$ are determined. In the first one, which we call the *deterministic ID model*, we have $X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n$, where $x_1, x_2, \ldots, x_n$ are n unique and fixed d-bit vectors. We study this model in Chapter 3. In the second one, which we call the *random ID model*, $X_1, X_2, \ldots, X_n$ are selected uniformly at random from $\{0,1\}^d$ without replacement. We study this model in Chapter 4.

To analyze the time complexity of Kademlia's routing algorithm, we define a simple path $\rho_x(y)$ in $\mathcal{K}$, where $x \in \mathcal{V}(\mathcal{K})$ is a vertex and $y \in \{0,1\}^d$ is an ID. This path starts from x, then jumps to the one that is closest to y among all the neighbors of x; from there, it again jumps to the adjacent vertex that is closest to y. In other words, we use a greedy strategy and try to get as close to y as possible at each hop. This approach repeats until we can not get closer to y anymore. Thus $\rho_x(y)$ imitates how node x locates the node that is closest to ID y in Kademlia. Figure 1–3 gives an example of such a path. Write $u \rightsquigarrow v$ for $(u,v) \in \mathcal{E}(\mathcal{K})$, which is the edge set of $\mathcal{K}$. Let $\rho_x(y) = (z_0, z_1, \ldots, z_r)$. We can precisely define the value of $z_0, z_1, \ldots, z_r$ as:

- Let $z_0 = x$. Then repeat the following step.
- Given z_t and $t \geq 0$, let

$$z' = \arg\min_{w : z_t \rightsquigarrow w} \delta(w, y).$$

  If z' exists and $\delta(z', y) < \delta(z_t, y)$, then $z_{t+1} = z'$, otherwise the path terminates.

Let $T_x(y) = r$, i.e., the length of $\rho_x(y)$. Then $T_x(y)$ is a random variable which counts the number of rounds of FIND_NODE that a node x has sent before it finally finds the node that is closest to y. We call $T_x(y)$ the *routing time* from x

to y. It is worthy of note that, in the *random ID model*, by symmetry, we have $T_{X_1}(y) \stackrel{L}{=} T_{X_i}(y)$ for all $y \in \{0,1\}^d$ and for all $i \in \{1,\ldots,n\}$.

Now consider the *deterministic ID model*, in which $X_1 = x_1, \ldots, X_n = x_n$ where $x_1, \ldots, x_n$ are fixed d-bit vectors. Then we can define

$$\begin{aligned} T_{\text{MAX}} &= \sup_{x_1,\ldots,x_n} \sup_{1\le i\le n} \sup_{y\in\{0,1\}^d} \mathbb{E}\left[T_{x_i}(y)\right] \\ &= \sup_{x_1,\ldots,x_n} \sup_{y\in\{0,1\}^d} \mathbb{E}\left[T_{x_1}(y)\right], \end{aligned}$$

i.e., the supremum of $\mathbb{E}\left[T_{x_1}(y)\right]$ over all choices of $x_1, \ldots, x_n$ and y. We call T_{MAX} the *maximal routing time*.

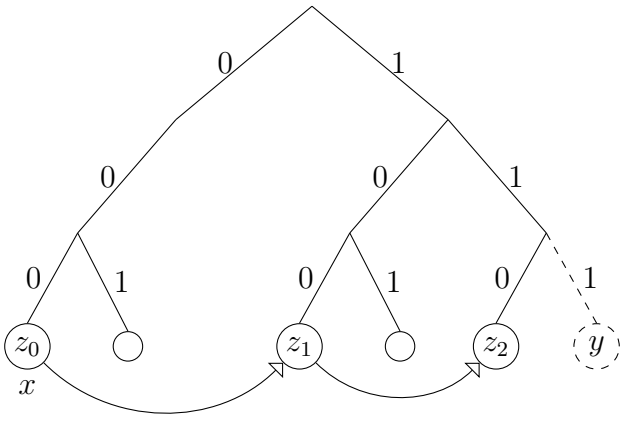

Figure 1–3: An example of $\rho_x(y)$ where $x = (0,0,0)$ and $y = (1,1,1)$. Note that in this example, there is no leaf node having ID y. Thus the path terminates at the node $(1,1,0)$, which is the closest to y.

Note that how we add an edge into $\mathcal{K}$ differs from how nodes in Kademlia add entries into their routing tables. But this is a sound simplification. In Kademlia, since each node uniformly chooses its ID at random from $\{0,1\}^d$ and each node is required to routinely locate randomly selected IDs periodically, no particular node should have much higher chance to be added into k-buckets of others. Therefore it is reasonable to assume that all candidates of a k-bucket have the same probability to be added to it.

Kademlia is a dynamic system in which nodes keep leaving and joining, and routing tables keep changing at every second, whereas our model $\mathcal{K}$ is static graph. But since locating an ID generally takes only a couple of seconds, we can still take our graph $\mathcal{K}$ as a valid approximation of the system during this short period and see how efficient Kademlia's routing algorithm is, and how its efficiency changes when the size of the network grows.

Although our definition of $T_{X_1}(y)$ only measures how long it takes to locate a node when parallel searching is not allowed, i.e., when we have system parameter $\alpha = 1$, it can be seen as an upper bound for situations with $\alpha > 1$. It also provides a framework for future studying of how α affects the efficiency of routing. For now we leave it as an open question.

We do not discuss how data items are stored and retrieved in Kademlia, although it is a large part of Kademlia's protocol, since these two operations fundamentally depend on routing — to store data to and retrieve them from nodes with specified IDs, we need to find their network addresses first.

Let $H_k = \sum_{i=1}^{k} 1/i$, which is also known as the k-th harmonic number. Our first main result is that

$$T_{\text{MAX}} \leq (1 + o(1))\frac{\log(n)}{H_k},$$

which we prove in Theorem 3.0.8. This conclusion matches the $\lceil \log n \rceil$ bound given by Maymounkov and Mazières [17], but improves it by a constant factor. We also show that, there is strong indication that this upper bound is tight. More specifically, Theorem 4.5.10 states that, when all IDs are uniformly chosen at random from $\{0,1\}^d$, the time it takes for a node, say x, to locate the ID having the largest XOR distance to x, dived by $\log n$, converges in probability to $1/g(k)$ where $g(k) = H_k + o(1)$ is a function of k.

This work is an exhibition of how probabilistic methods and theories, like stochastic ordering, concentration inequalities, and coupling of random variables, etc., could surprisingly simplify the analysis of a complex model and provide rigorous proofs which were considered too troublesome to construct.

This work is organized as follows. In Section 2, we briefly discuss the connectivity of $\mathcal{K}$. In Section 3 we give an upper bound on T_{MAX}. In Section 4, we study the case that all IDs are selected uniformly at random from $\{0,1\}^d$ without replacement. Finally, in Section 5, we list some open questions which indicate directions of future research.

CHAPTER 2
Connectivity

In this chapter, we study the connectivity of $\mathcal{K}$, i.e., the random graph defined in the last chapter for modeling a Kademlia network. We first establish a simple lemma.

Lemma 2.0.1 (Existence of Mixing Point)**.** *For all rooted trees with at least two leaves, if we paint some but not all of its leaves white and the remaining leaves black, then there is always at least one internal node having two subtrees, one which has only white leaves and another which has only black leaves. We call this node a Mixing Point.*

Proof. Let h be the height of the tree. We use induction on h. When $h = 1$, the root must be a Mixing Point because the tree has more than two leaves and we can not paint all of them with the same color. Given the lemma holds when $h \leq i$ for some integer $i \geq 1$, we show that it also holds for trees of height $i + 1$. If the root has a single subtree, then this subtree of height h, which, by induction hypothesis, must contain a Mixing Point. On the other hand, if the root has more than one subtree, but none of them contains a Mixing Point, then all of them contain only monochromatic leaves. Since not all leaves are painted the same color, one of these subtrees must contain only white leaves, and another must contain only black leaves. Thus the root is a Mixing Point. □

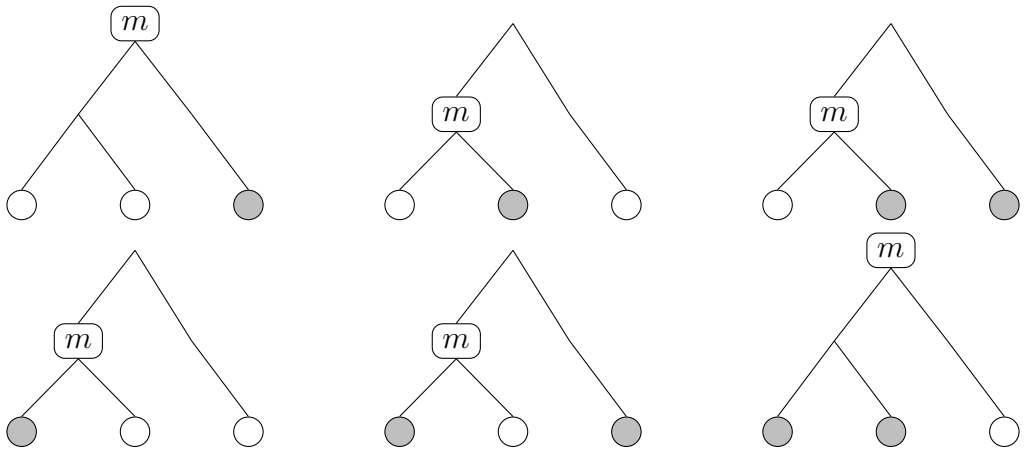

Figure 2–1: An example of Mixing Point. However we paint the three leaves, we are always able to find a Mixing Point (the node m).

We say that a directed graph is *strongly connected* if there is a path from each vertex in the graph to every other vertex.

Theorem 2.0.2 (Connectivity)**.** *$\mathcal{K}$ is strongly connected.*

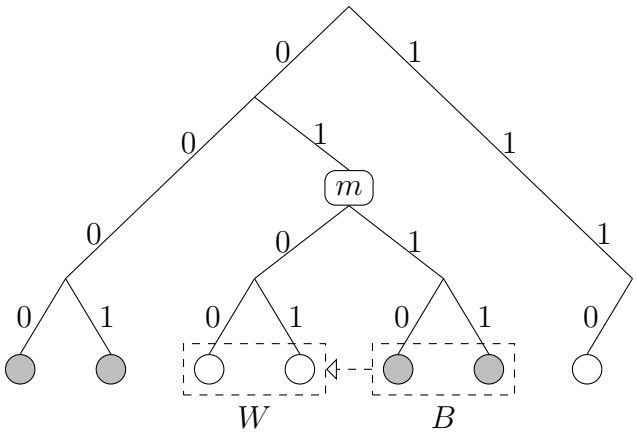

Figure 2–2: An example of painting an ID trie. Painting the ID trie's leaves black and white, we can always find a Mixing Point, such as the node m in the picture. Let the two subtrees of m be B and W. We claim that there is at least one edge in $\mathcal{K}$ that goes from a leaf in B to a leaf in W.

Proof. In the trivial case that $\mathcal{K}$ contains only one vertex, $\mathcal{K}$ is strongly connected. Thus we assume that $\mathcal{K}$ contains more than one vertex, i.e., $n > 1$. Represent all IDs in $\mathcal{V}(\mathcal{K})$ in a trie as shown in Figure 2–2. It follows from

Lemma 2.0.1 that if we paint a proper subset of the leaves white and the remaining ones black, there is always a Mixing Point m in the trie. Let the two subtrees of m be B and W. For an ID $b \in B$, it is easy see that $W = \mathcal{D}_i(b)$ for some integer i. Put differently, b has a k-bucket $\mathcal{B}_i(b)$ that contains only nodes from W. Thus there must exist at least one leaf $w \in W$ such that $b \rightsquigarrow w$. Put differently, however we partition $\mathcal{V}(\mathcal{K})$ into two nonempty sets, there is always an edge that goes from one to another, which implies the strong connectivity of $\mathcal{K}$. □

CHAPTER 3
The deterministic ID model

In this chapter, we assume that $X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n$ where $x_1, \ldots, x_n$ are fixed d-bit vectors, i.e., the IDs are deterministically decided. Recall that in Chapter 1 we define the *maximal routing time* by

$$T_{\text{MAX}} = \sup_{x_1,\ldots,x_n} \sup_{y\in\{0,1\}^d} \mathbb{E}\left[T_{x_1}(y)\right].$$

The main result of this chapter is Theorem 3.0.8, which states that

$$T_{\text{MAX}} \le (1+o(1))\frac{\log n}{H_k},$$

where $H_k = \sum_{i=1}^{k} 1/i$.

Note that for all $y_1, y_2 \in \{0,1\}^d$, we have

$$\sup_{x_1,\ldots,x_n} \mathbb{E}\left[T_{x_1}(y_1)\right] = \sup_{x_1,\ldots,x_n} \mathbb{E}\left[T_{x_1}(y_2)\right].$$

In other words, the supremum of $\mathbb{E}\left[T_{x_1}(y)\right]$ over all choices of $x_1, \ldots, x_n$ is the same for all $y \in \{0,1\}^d$. That is because $\mathbb{E}\left[T_{x_1}(y)\right]$ is only decided by the XOR distances between vertices and the XOR distance between x_1 and y, and has nothing to do with what particular IDs $x_1, \ldots, x_n$ and y are. Therefore, we have

$$T_{\text{MAX}} = \sup_{x_1,\ldots,x_n} \mathbb{E}\left[T_{x_1}(\overline{1})\right].$$

where $\overline{1} = (1, 1, \ldots, 1)$. Thus, if we can get an upper bound on $\mathbb{E}\left[T_{x_1}(\overline{1})\right]$ regardless of the choice of $x_1, \ldots, x_n$, then we also get an upper bound on T_{MAX}. We write $\rho = \rho_{x_1}(\overline{1})$ and $T = T_{x_1}(\overline{1})$.

Let m and b be two positive integers. We define $Z(m, b)$ as a discrete random variable, which is the minimum of a sample of up to size m selected from $\{0, 1, \ldots, b-1\}$ uniformly at random. We define a similar continuous random variable $U(m)$ as the minimum of m continuous i.i.d. uniform random variables on the unit interval.

Figure 3–1 depicts the first hop of ρ in the ID trie. A simple fact about the ID trie is that, if we always arrange branches representing 1 to the right hand side, which we take as a convention from now, then the closer a leaf is to the right, the smaller its XOR distance to $\overline{1}$ is. Thus the rightmost leaf is the node that is closest to $\overline{1}$. We denote this leaf (ID) by y' in this chapter. Also note that, the closer a leaf to the right, the smaller its XOR distance to y'.

Let us look at the first hop more carefully. Let $z_0 = x_1$ be the starting ID. z_0 must choose the one that is closest $\overline{1}$ as the next hop z_1. In other words, z_1 is z_0's rightmost neighbor. Let $i = d - \ell(z_0, y')$, then $\mathcal{B}_i(x)$ is the k-bucket that contains y'. It is easy to see from Figure 3–1 that $\mathcal{D}_i(z_0)$ is the highest subtree on the right hand side of z_0. Let $S_0 = \mathcal{D}_i(z_0)$. We claim that the second hop $z_1 \in S_0$. (Since S_0 is not empty, z_0 must have at least one neighbor, say z', in S_0. If $z_1 \notin S_0$, then it must be on the left hand side of z', which is contradictory to that z_1 is z_0's rightmost neighbor.) Remember that we decide z_0's neighbors in S_0 by selecting a sample of size up to k uniformly at random from S_0. Thus we can think that z_1 is decided by selecting up to k leaves from S_0 uniformly at random without replacement, and let z_1 be the rightmost one. Repeating our argument recursively, we can give $\rho = (z_0, z_1, \ldots, z_T)$ an equivalent definition as follows:

- Let $z_0 = x_1$. Repeat the following step.

- Given z_r and $r \geq 0$, let S_r be highest subtree on the right hand side of z_r. If $S_r = \emptyset$, then the path terminates. Otherwise select up to k leaves from S_r uniformly at random without replacement, and let the rightmost one be z_{r+1}.

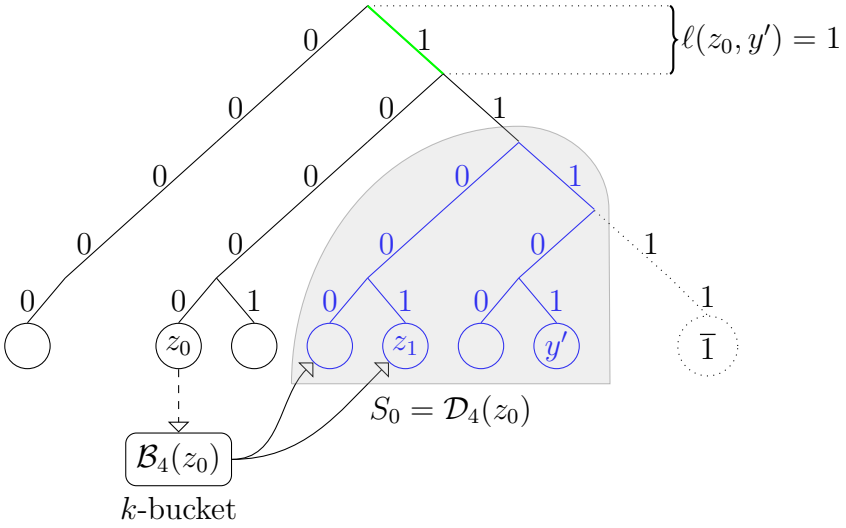

Figure 3–1: An example of the first hop of ρ with parameter $d = 5$ and $k = 2$. Since $d - \ell(z_0, y') = 4$, y' must be in subtree $S_0 = \mathcal{D}_4(z_0)$, which is the highest subtree on the right of z_0. z_0 chooses up to k nodes from S_0 uniformly at random without replacement, and let z_1 be the rightmost one.

As shown in Figure 3–2, it is obvious that $S_0 \supset S_1 \supset \ldots \supset S_T$. Thus $|S_0|, |S_1|, \ldots, |S_T|$ is a sequence of strictly decreasing random variables, which suggests that we might be able to derive an upper bound of $\mathbb{E}[T]$ by studying how fast this sequence decreases.

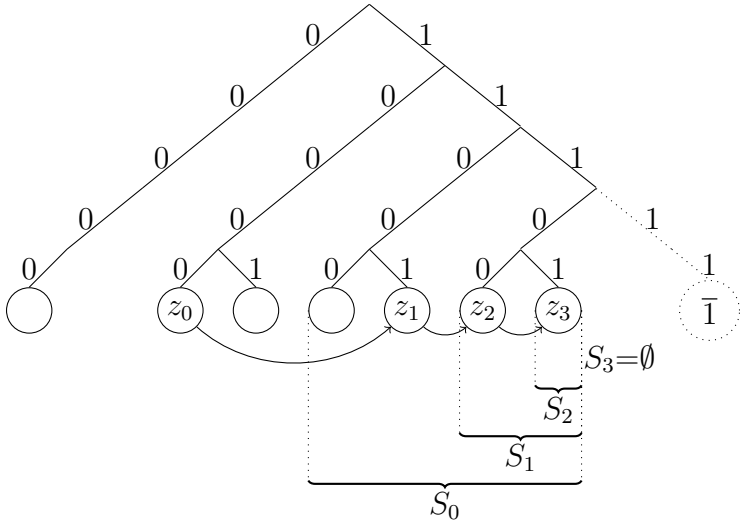

Figure 3–2: An example of ρ whose length $T = 3$. Note that $S_0 \supset S_1 \supset S_2 \supset S_3 = \emptyset$.

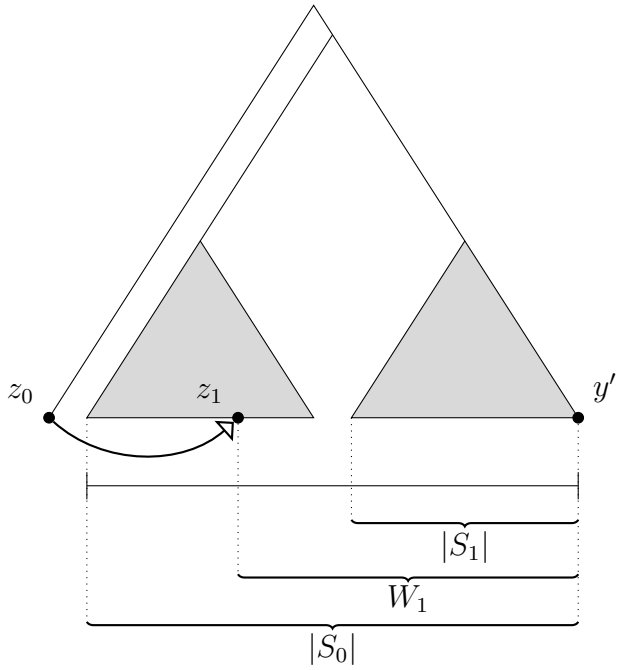

Figure 3–3: The intuition that leads us to our final proof.

It is still not clear how to write the distribution of $|S_0|, |S_1|, \ldots, |S_T|$. But if we define a random variable W_1 to be the number of nodes on the right hand side of z_1, as shown in Figure 3–3, then it is relatively easy to

write W_1's distribution. To be precise, it follows from how z_1 is selected that $W_1 \stackrel{L}{=} Z(k, |S_0|)$, the minimum of a sample of size up to k selected uniformly at random from $\{0, 1, \ldots, |S_0| - 1\}$. And we know for sure that $W_1 \geq |S_1|$, as shown in Figure 3–3. If we define a random variable $W_2 = Z(k, W_1)$, then intuitively we know that $|S_2| < W_2$ is more likely to happen than the opposite. If we define a sequence of random variables $(W_i)_{i \geq 1}$ recursively in this way, and let $W_{T'}$ be the first one that takes value 0, then it is likely that this sequence decreases slower than $|S_0|, |S_1|, \ldots, |S_T|$, and thus the expectation of T' must be bigger than T, which gives us the upper bound. In the remainder of this chapter, we turn this intuition into a real proof.

To be precise about what we mean by $A < B$ *is more likely to happen than the opposite*, we introduce the *stochastically smaller than* notation $\preceq$. Given two random variables A and B, we say $A \preceq B$ if and only if

$$P\{A \geq r\} \leq P\{B \geq r\} \quad \text{for all } r \in (-\infty, \infty).$$

Thus, our intuition about $W_1, \ldots, W_T$ can be precisely expressed as

$$|S_t| \preceq W_t \quad \text{for all } 1 \leq t \leq T.$$

However, the distributions of the $W_1, W_2, \ldots$ are still too troublesome for analysis. Thus we define an infinite sequence of continuous random variables $W_1', W_2' \ldots$, which have similar distributions as $(W_i)_{i \geq 1}$, but are much more analysis-friendly. Let $B_1, B_2, \ldots$ be i.i.d. random variables with distribution $U(k)$. We define

$$W_t' = \begin{cases} B_1 \times |S_0| & \text{if } t = 1, \\ B_t \times W_{t-1}' & \text{if } t > 1. \end{cases}$$

In other words

$$W'_t = |S_0| \times \prod_{i=1}^{t} B_i, \quad t = 1, 2, \ldots.$$

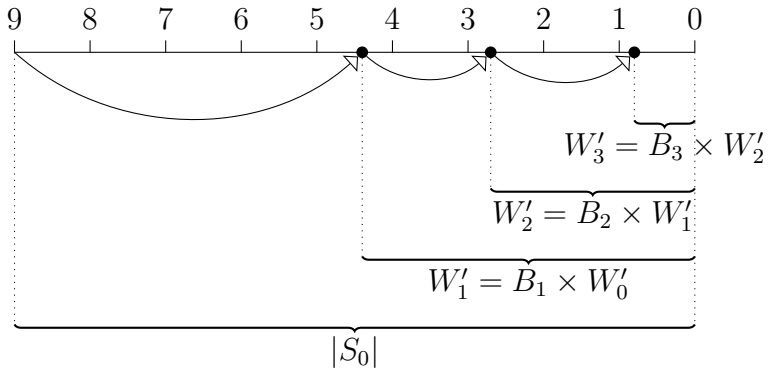

Figure 3–4: An example of $(W'_i)_{i\geq 1}$.

This infinite sequence of random variables $(W'_t)_{t\geq 1}$ can be interpreted as a process of randomly cutting a line of $|S_0|$ unit length as shown in Figure 3–4. We first cut the line to B_1 fraction of the original length; then we cut the remaining to B_2 fraction. This approach can be repeated for ever and $W'_1, W'_2, \ldots$ are the lengths of line after each cutting operation.

$(W'_t)_{t\geq 1}$ are easier to analyze because $U(k)$ is beta distribution. A beta random variable $B(a, b)$ has its probability density function defined by

$$f(x; a, b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1}, \quad 0 \leq x \leq 1,$$

where a and b are two parameters and $\Gamma(z)$ is the gamma function defined by

$$\Gamma(z) = \int_0^\infty e^{-t} t^{z-1} dt,$$

which has the following properties

$$\Gamma(1) = 1, \qquad \Gamma(1+z) = \Gamma(z) \qquad \text{for all } z > 0.$$

In order statistics theory, a basic result [5, chap. 2.3] is that the r-th smallest of a sample of size m from a continuous uniform distribution on the unit interval has beta distribution $B(r, m+1-r)$. Plugging in $r = 1$ and $m = k$, we have

$$B_t \stackrel{L}{=} U(k) \stackrel{L}{=} B(1, k), \quad t = 1, 2, \ldots.$$

For more about beta distribution, see [13, chap. 25].

To establish the stochastic order between $|S_1|, |S_2|, \ldots$ and $W_1', W_2', \ldots$, we introduce two lemmas.

Lemma 3.0.3. *We have*

$$Z(m, b) \preceq b \times U(m)$$

for all positive integer b and m.

Proof. In the trivial case when $b \leq m$, we always have $Z(m, b) = 0$ and the lemma holds because $b \times U(m) > 0$. Therefore we assume $b > m$. By definition of $Z(m, b)$, we can write

$$Z(m, b) = \min\{Z_1, Z_2, \ldots, Z_m\},$$

where $\{Z_1, \ldots, Z_m\}$ is a sample of size m uniformly selected at random from $\{0, \ldots, b-1\}$ without replacement. Since it is a sample selected without replacement, Z takes maximum when $\{Z_1, \ldots, Z_m\} = \{b-m, \ldots, b-1\}$, which implies

$$Z \leq b - m.$$

Thus for all real numbers $a > b - m$, we have

$$P\{Z \geq a\} = 0 \leq P\{b \times U(m) \geq a\}.$$

On the other hand, for a real numbers $a \in (0, b-m]$, we have

$$\begin{aligned}
P\{Z \geq a\} &= P\{\cap_{i=1}^{m}[Z_i \geq a]\} \\
&= P\{Z_1 \geq a\} P\left\{Z_2 \geq a \mid Z_1 \geq a\right\} \; P\left\{Z_3 \geq a \mid Z_1 \geq a \cap Z_2 \geq a\right\} \\
&\quad \ldots P\left\{Z_m \geq a \mid \cap_{j=1}^{m-1} Z_j \geq a\right\} \\
&= \prod_{i=0}^{m-1} \frac{b-a-i}{b-i} \\
&\leq \prod_{i=0}^{m-1} \frac{b-a}{b} \\
&= P\{b \times U(m) \geq a\}.
\end{aligned}$$

Finally, for $a \leq 0$, we have

$$P\{Z \geq a\} = P\{b \times U(m) \geq a\} = 1.$$

Therefore, $Z(m, b) \preceq b \times U(m)$. □

Lemma 3.0.4. *For all $t \geq 1$ we have*

$$|S_t| \preceq B_t \times |S_{t-1}|.$$

Proof. Let Z be the number of leaves to the right of z_t. Then it follows from how z_t is selected that $Z \stackrel{L}{=} Z(k, |S_{t-1}|)$. Thus by Lemma 3.0.3, for all real numbers a and positive integers b we have

$$P\left\{Z \geq a \mid |S_{t-1}| = b\right\} \leq P\left\{b \times U(k) \geq a\right\}.$$

Note that we always have $Z \geq |S_t|$, which can be easily seen from Figure 3–5. It follows from the fact $Z \geq |S_t|$ that

$$\begin{aligned} & P\left\{|S_t| \geq a \mid |S_{t-1}| = b\right\} \\ & \quad \leq P\left\{Z \geq a \mid |S_{t-1}| = b\right\} \\ & \quad \leq P\left\{b \times U(k) > a\right\}. \end{aligned}$$

Note that by definition for all $t \geq 1$ we have $B_t \stackrel{L}{=} U(k)$. Let $\mathbb{N}$ be the set of positive integers. For all real numbers a, we have

$$\begin{aligned} P\{|S_t| \geq a\} &= \sum_{b \in \mathbb{N}} P\{|S_t| \geq a \mid |S_{t-1}| = b\} P\{|S_{t-1}| = b\} \\ &\leq \sum_{b \in \mathbb{N}} P\{bB_t \geq a\} P\{|S_{t-1}| = b\} \\ &= \sum_{b \in \mathbb{N}} P\{bB_t \geq a \cap |S_{t-1}| = b\} \\ &= P\{B_t \times |S_{t-1}| \geq a\}. \end{aligned}$$

Therefore, we have $|S_t| \preceq B_t |S_{t-1}|$. □

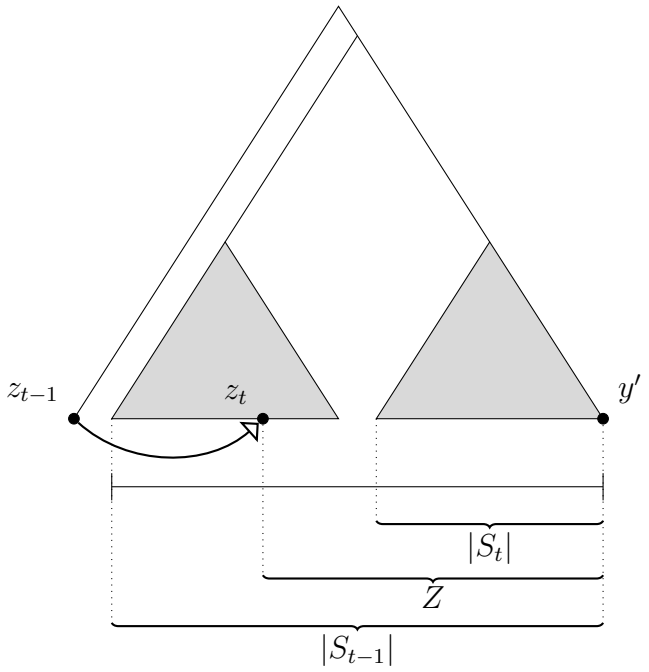

Figure 3–5: Proof of Lemma 3.0.4. It can be easily seen from the picture that $Z \geq |S_t|$.

To finally establish the relationship between $|S_1|, |S_2|, \ldots$ and $W'_1, W'_2, \ldots$, we need the following result from stochastic order theory [28, thm. 1.A.3].

Theorem 3.0.5. *Let $Y_1, Y_2, \ldots, Y_m$ be a set of independent random variables and let $Z_1, Z_2, \ldots, Z_m$ be another set of independent random variables. If $Y_i \preceq Z_i$ for $i = 1, 2, \ldots, m$, then, for any increasing function $\Phi : \mathbb{R}^m \to \mathbb{R}$, one has $\Phi(Y_1, Y_2, \ldots, Y_m) \preceq \Phi(Z_1, Z_2, \ldots, Z_m)$.*

An immediate result of this theorem is that given two non-negative random variables $Y_1 \preceq Y_2$, and another positive random variable Z that is independent of both Y_1 and Y_2, we have

$$Y_1 \times Z \preceq Y_2 \times Z. \tag{3.1}$$

Lemma 3.0.6. *For all $t \geq 1$, we have $|S_t| \preceq W'_t$.*

Proof. Recursively applying Lemma 3.0.3 and equation (3.1) gives us

$$\begin{aligned} |S_t| &\preceq B_t|S_{t-1}| \\ &\preceq B_t B_{t-1}|S_{t-2}| \\ &\preceq B_t B_{t-1} \dots B_1|S_0| \\ &= W'_t. \end{aligned}$$

□

To finish the proof, we use the moment bound to get an upper bound on the expectation of $\prod_{i=1}^{t} B_i$.

Lemma 3.0.7. *Let $B_1, B_2, \dots B_t$ be t i.i.d. beta random variables with distribution $B(1,k)$. For any $r > 0$, we have*

$$\mathbb{E}\left[\left(\prod_{i=1}^{t} B_t\right)^r\right] = \left(\frac{k!}{\prod_{i=1}^{k}(r+i)}\right)^t.$$

Proof. A beta random variable $B(a,b)$ has r-th raw moment [13, chap. 25],

$$\mathbb{E}\left[B(a,b)^r\right] = \frac{\Gamma(a+r)\Gamma(a+b)}{\Gamma(a)\Gamma(a+b+r)}.$$

Since $B_1 \overset{L}{=} B(1,k)$, we have

$$\begin{aligned} \mathbb{E}\left[B_1^r\right] &= \frac{\Gamma(1+r)\Gamma(1+k)}{\Gamma(1)\Gamma(1+r+k)} \\ &= \frac{k!}{\prod_{i=1}^{k}(r+i)}. \end{aligned}$$

As $B_1, B_2, \dots, B_t$ are independent, we have

$$\mathbb{E}\left[\left(\prod_{i=1}^{t} B_t\right)^r\right] = \mathbb{E}\left[B_1^r\right]^t = \left(\frac{k!}{\prod_{i=1}^{k}(r+i)}\right)^t.$$

□

Theorem 3.0.8. *We have*

$$T_{\text{MAX}} = \sup_{x_1,\dots,x_n} \sup_{y \in \{0,1\}^d} \mathbb{E}\left[T_{x_1}(y)\right] \leq (1+o(1))\frac{\log n}{H_k},$$

where $H_k = \sum_{i=1}^{k} 1/i$, also known as the k-th harmonic number.

Proof. Recall that T is the routing time from x_1 to $\overline{1}$, i.e., the length of path $\rho_{x_1}(\overline{1})$. Note that however $x_1, \ldots, x_n$ are chosen, the event $[T \geq t]$ is always equivalent to event $[|S_t| \geq 1]$, which gives us an upper bound of $P\{T \geq t\}$:

$$\begin{aligned}
P\{T \geq t\} &= P\{|S_t| \geq 1\} \\
&\leq P\left\{\prod_{i=1}^{t} B_t \geq \frac{1}{|S_0|}\right\} \quad \text{(Lemma 3.0.6)} \\
&\leq P\left\{\prod_{i=1}^{t} B_t \geq \frac{1}{n}\right\} \quad (n = |\mathcal{V}(\mathcal{K})| > |S_0|) \\
&\leq \mathbb{E}\left[\left(\prod_{i=1}^{t} B_t\right)^r\right] \times n^r \quad (r > 0, \text{ moment bound}) \\
&= \left(\frac{k!}{\prod_{i=1}^{k}(r+i)}\right)^t \times n^r \quad \text{(Lemma 3.0.7)} \\
&= N_t,
\end{aligned}$$

where

$$N_t = \left(\frac{k!}{\prod_{i=1}^{k}(r+i)}\right)^t \times n^r, \quad t = 1, 2, \ldots.$$

Since a probability is at most 1, we have

$$P\{T \geq t\} \leq \min\{1, N_t\}, \quad t = 1, \ldots.$$

As $(N_t)_{t \geq 1}$ decreases exponentially, we can find a

$$t^* = \max\{t : N_t \geq 1\}.$$

Therefore, we have

$$\begin{aligned}\mathbb{E}[T] &= \sum_{t=1}^{\infty} P\{T \geq t\} \\ &\leq \sum_{t=1}^{\infty} \min\{1, N_t\} \\ &= \sum_{t=1}^{t^*} 1 + \sum_{t=t^*+1}^{\infty} N_t \\ &\leq t^* + C,\end{aligned}$$

where

$$C = \frac{1}{1 - k!/\prod_{i=1}^{k}(r+i)}$$

is a constant with respect to n. Take logarithm of both sides of $N_{t^*} \geq 1$, we have

$$t^* \left\{ \sum_{i=1}^{k} \log \frac{i}{r+i} \right\} + r \log n \geq 0,$$

which implies

$$\begin{aligned}t^* &\leq \frac{r \log(n)}{\sum_{i=1}^{k} \log(1 + r/i)} \\ &\leq \frac{r \log(n)}{\sum_{i=1}^{k} f_i(r)},\end{aligned} \tag{3.2}$$

where

$$f_i(r) = \log(1 + \frac{r}{i}), \quad i = 1, \ldots, k.$$

Taking derivatives of f_i, we have

$$\begin{aligned}f_i(0)' &= \frac{1}{i}, \\ f_i(0)^{(2)} &= -\frac{1}{i^2}, \\ f_i(r)^{(3)} &= \frac{2}{(i+r)^3}.\end{aligned}$$

It follows from Taylor Series that

$$\begin{aligned} f_i(r) &= f(0) + f'(0)r + \frac{f''(0)}{2!}r^2 + \frac{f^{(3)}(\zeta)}{3!}r^3 \\ &\geq \frac{r}{i} - \frac{r^2}{2i^2} \\ &\geq \frac{r}{i} - \frac{r^2}{i^2}, \end{aligned}$$

where $0 \leq \zeta \leq r$. Together with inequality (3.2), we have

$$\begin{aligned} t^* &\leq \frac{r \log n}{\sum_{i=1}^{k}(r/i - r^2/i^2)} \\ &\leq \frac{\log n}{\sum_{i=1}^{k}(1/i - r/i^2)} \\ &= \frac{\log n}{H_k - r \times \sum_{i=1}^{k} 1/i^2}. \end{aligned}$$

Thus, for any $\epsilon > 0$, if we choose

$$r = \frac{\epsilon}{\sum_{i=1}^{k} 1/i^2},$$

we have

$$\mathbb{E}[T] \leq \frac{\log n}{H_k - \epsilon} + C,$$

which implies

$$\mathbb{E}[T] \leq (1 + o(1))\frac{\log n}{H_k}.$$ □

Note that this upper bound matches the $\lceil \log n \rceil$ upper bound given by Maymounkov and Mazières [17] in their sketch of proof.

CHAPTER 4
The random ID model

4.1 Plan of the proof

In this chapter, we assume that $\mathcal{V}(\mathcal{K})$ is sample of size n selected uniformly at random without replacement from $\{0,1\}^d$. We study the distribution of $T_{X_1}(X_1^c)$ where X_1^c is the *polar opposite* of X_1, i.e., the ID in $\{0,1\}^d$ that has the largest XOR distance to X_1. Given two IDs $x = (x_1, \ldots, x_d)$ and $y = (y_1, \ldots, y_d)$, we define the XOR operator on these two vectors by

$$x \oplus y = (x_1 \oplus y_1, x_2 \oplus y_2, \ldots, x_d \oplus y_d),$$

which is useful for rotating hypercubes. Recalling that $\overline{1} = (1, 1, \ldots, 1)$, we can define X_1^c (the *polar opposite of* X_1) using this operator by

$$X_1^c = X_1 \oplus \overline{1}.$$

The last equation implies $\delta(X_1, X_1^c) = 2^d - 1$, which is the largest possible XOR distance. We call $T_{X_1}(X_1^c)$ the *polar routing time* and denote it by $\hat{T}$. We also denote the path $\rho_{X_1}(X_1^c)$ by $\hat{\rho}$.

Let $Y_1, Y_2, \ldots$ be random variables on some probability space, we say $Y_n \to Y$ in probability [10, chap. 7], written $Y_n \xrightarrow{p} Y$, if for all $\epsilon > 0$, we have

$$\lim_{n\to\infty} P\{|Y_n - Y| \geq \epsilon\} = 0.$$

With $\xrightarrow{p}$ defined, Theorem 4.5.10, the main result of this chapter, states that

$$\frac{\widehat{T}}{\log n} \xrightarrow{p} \frac{1}{g(k)},$$

as $n \to \infty$, where

$$g(k) = H_k + o(1)$$

is a function of k.

A simple observation that simplifies the analysis of $\widehat{T}$ is that if we relabel all vertices with a group of IDs $X'_1, \ldots, X'_d$ where

$$X'_i = X_i \oplus X_1, \quad i = 1, \ldots, n,$$

then we have $X'_1 = (0, 0, \ldots, 0)$ and $X'^c_1 = \overline{1}$. We also have

$$\delta(X_i, X_j) = \delta(X'_i, X'_j), \quad i, j \in 1, 2, \ldots, n.$$

In other words, this relabeling does not change the distance between vertices and thus does not affect our analysis of $\widehat{T}$, but it allows us to see the routing process as starting from $(0, 0, \ldots, 0)$, which is the leftmost leaf in the ID trie, denoted by $\overline{0}$, and going towards the rightmost leaf. Thus, in this chapter, we always assume that $X_1 = \overline{0}$.

We define a group of discrete random variables $N_1, \ldots, N_d$ as

$$N_i = \sum_{x \in \mathcal{V}(\mathcal{K})} \mathbf{1}_{[\ell(x,\overline{1})=i]}, \quad i = 1, \ldots, d,$$

where we recall that $\ell(x, \overline{1})$ is the length of the common prefix of x and $\overline{1}$. Put differently, N_i is the number of IDs which have the following format,

$$(\underbrace{1, 1, \ldots, 1}_{i}, 0, \ldots).$$

From the perspective of the ID trie, N_1 is the number of leaves under internal node that represents string "10", N_2 is the number of leaves under internal node that represents string "110", etc., as shown in Figure 4–1. It follows from how $\mathcal{V}(\mathcal{K})$ is chosen that N_i is binomial (n, p_i), i.e.,

$$P\{N_i = j\} = \binom{n}{j} (p_i)^j (1 - p_i)^{n-j}, \quad i = 1, \ldots, d,$$

where

$$p_i = \begin{cases} 1/2^{i+1}, & \text{if } 1 \leq i < d, \\ 1/2^d, & \text{if } i = d. \end{cases}$$

Thus we have

$$\mathbb{E}[N_i] = \begin{cases} n/2^{i+1}, & \text{if } 1 \leq i < d, \\ n/2^d, & \text{if } i = d. \end{cases}$$

For more about the binomial distribution, see [14, chap. 3].

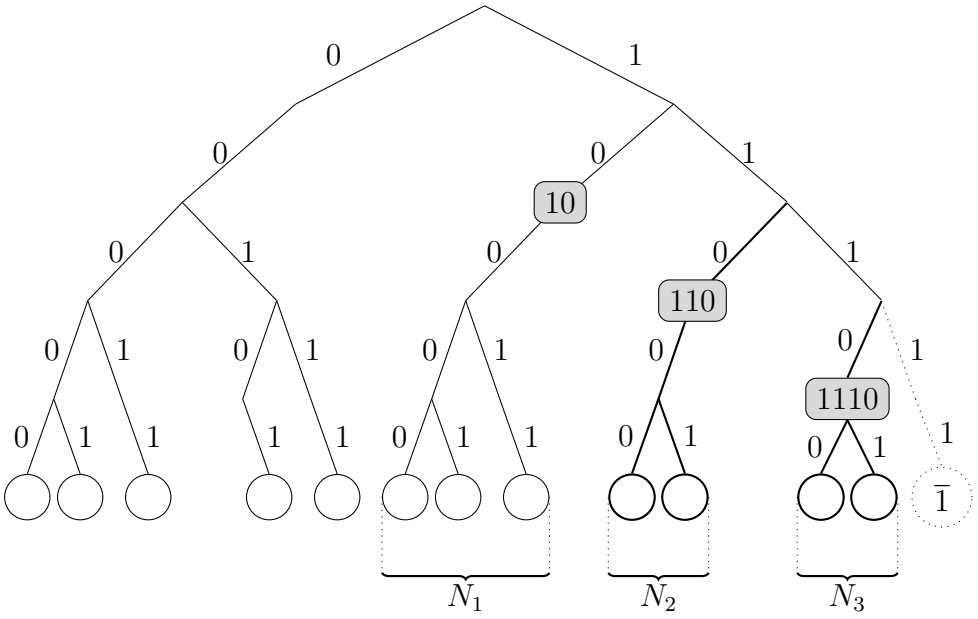

Figure 4–1: An example of $N_1, \ldots, N_d$. Each leaf node in this trie represents a randomly selected ID.

Let $\hat{\rho} = (z_0, z_1, \ldots, z_{\widehat{T}})$, where $z_0 = X_1 = \overline{0}$. To decide next hop z_1, we select up to k leaves uniformly at random from $\mathcal{D}_d(z_0)$, which is the right half the ID trie, and let z_1 be the rightmost one. We define a random variable $R_1 = \ell(z_1, \overline{1})$, which is the length of the common prefix shared by z_1 and $\overline{1}$, as shown in Figure 4–2. R_1 has a random distribution which is decided by the values of $N_1, \ldots, N_d$.

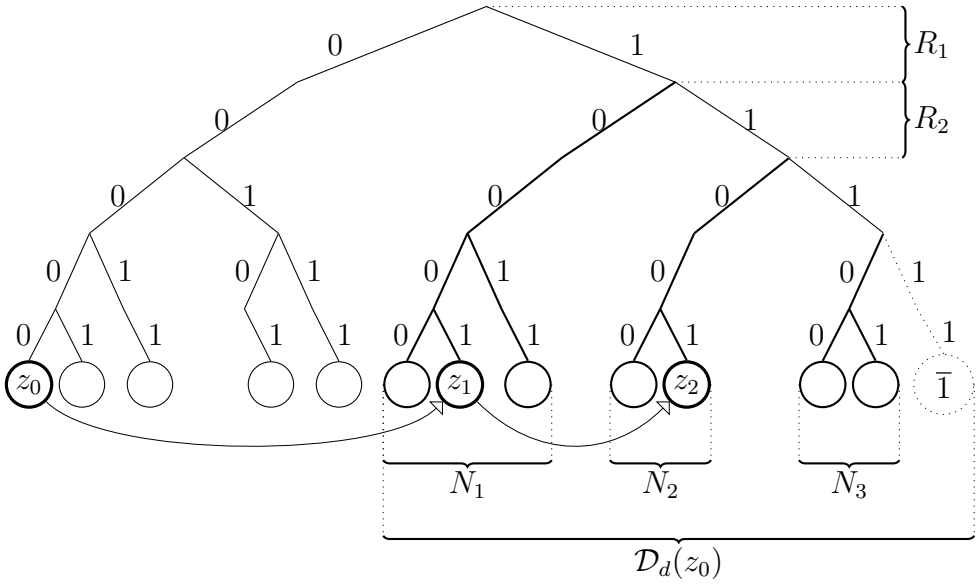

Figure 4–2: An example of $(R_t)_{t\geq 1}$. Each leaf node in this trie represents a randomly selected ID.

In the case $k = 1$, since z_0 can have at most one neighbor in $\mathcal{D}_d(z_0)$, z_1 can be seen as selected uniformly at random from the leaves in $\mathcal{D}_d(z_0)$, which number $\sum_{j=1}^{d} N_j$ in all. Therefore, the probability of event $[R_1 = j]$ is proportional to the number of leaves that share a j-length prefix with $\overline{1}$, which is N_j. Thus we have

$$P\{R_1 = j\} = \frac{N_j}{\sum_{j=1}^{d} N_j}.$$

When n is large, $N_1, N_2, \ldots$ are close to their expectations $n/4, n/8, \ldots$ respectively, and the total number of leaves in the right half of the trie is about $n/2$.

Therefore, roughly speaking, R_1 has a geometric (1/2) distribution when $k = 1$, i.e.,

$$P\{R_1 = j\} = \frac{n/2^{j+1}}{n/2} = 1/2^j.$$

(The inaccuracy is gradually cleaned up later in this chapter.) It is easy to see from Figure 4–2 that after the first hop, the leaves left as candidate for next hop z_2 is approximately $N_{\ell(z_1,\overline{1})}$, i.e.,

$$\frac{n}{2^{\ell(z_1,\overline{1})+1}} = \frac{n}{2^{R_1+1}}.$$

Let

$$R_i = \ell(z_i,\overline{1}) - \ell(z_{i-1},\overline{1}), \quad i = 1, 2, \ldots.$$

Repeating our argument recursively, we can see that $R_2, R_3, \ldots$ also have approximately a geometric (1/2) distribution, and the candidate leaves for z_{i+1} number about

$$\frac{n}{2^{\ell(z_i,\overline{1})+1}} = \frac{n}{2^{1+R_1+\ldots+R_i}}.$$

Since after $\widehat{T}$-th hop, the path $\hat{\rho}$ terminates, we have

$$\frac{n}{2^{1+R_1+R_2+\ldots+R_{\widehat{T}}}} < 1. \tag{4.1}$$

In the case that $k > 1$, although $R_1, \ldots, R_{\widehat{T}}$ do not have geometric-like distribution anymore, they still have distributions which are close to each other, and Equation (4.1) still holds. Therefore, for all $k \geq 1$, if we assume $R_1, \ldots, R_{\widehat{T}}$ are i.i.d. discrete random variables with the same distribution, then we are able to apply Chernoff's bound to $\sum_{i=1}^{\widehat{T}} R_i$ to show that $\widehat{T}$ has a distribution which concentrates its mass at $\log_2(n)/\mathbb{E}[R_1]$. When n is large enough, this conclusion should be very close to the face. In the remainder of this chapter, we turn the previous argument into a mathematical proof.

4.2 Concentration of Mass

The following theorem serves as the cornerstone of this chapter.

Theorem 4.2.1. *Let $G_1, G_2, \ldots$ be i.i.d. positive discrete random variables, such that $\mathbb{E}[\exp(\theta_0 G_1)]$ exists for some $\theta_0 > 0$. Let*

$$T_n = \min\left\{t : \log_2 n \leq \sum_{i=1}^{i} G_i\right\}.$$

For all $\epsilon > 0$, we have

$$(1) \quad \lim_{n\to\infty} P\left\{T_n > \left(\frac{1}{\mathbb{E}[G_1]} + \epsilon\right)\log_2 n\right\} = 0,$$

and

$$(2) \quad \lim_{n\to\infty} P\left\{T_n < \left(\frac{1}{\mathbb{E}[G_1]} - \epsilon\right)\log_2 n\right\} = 0.$$

Or more concisely, we have

$$\frac{T_n}{\log_2 n} \xrightarrow{p} \frac{1}{\mathbb{E}[G_1]}.$$

Proof. If $\mathbb{E}[\exp(\theta_0 G_1)]$ exists, then for all $\theta < \theta_0$, $\mathbb{E}[\exp(\theta G_1)]$ also exists. Thus we have the cumulant-generating function [14, chap. 1.2.8] of G_1 defined on domain $(-\infty, \theta_0]$ as

$$f(\theta) = \log \mathbb{E}\left[e^{\theta G_1}\right] = \sum_{i=1}^{\infty} \kappa_i \frac{\theta^i}{i!},$$

where $\kappa_1, \kappa_2, \ldots$ are known as G_1's cumulants. Note that

$$\kappa_1 = \mathbb{E}[G_1].$$

For part 1 of the theorem, write

$$t' = \lfloor (1/\kappa_1 + \epsilon)\log_2 n \rfloor.$$

Since T_n is the smallest integer such that $\log_2 n \leq \sum_{i=1}^{T_n} G_i$, $T_n > t'$ implies

$$\log_2 n > \sum_{i=1}^{t'} G_i.$$

We can apply Chernoff's bound for the left tail to show that the probability of this event goes to 0 when n goes to infinity. For $\theta < 0$, we have

$$\begin{aligned}
&P\{T_n > (1/\kappa_1 + \epsilon)\log_2 n\} \\
&\leq P\left\{\sum_{i=1}^{t'} G_i < \log_2 n\right\} \\
&\leq \left(\mathbb{E}\left[e^{\theta G_1}\right]\right)^{t'} / e^{\theta \log_2 n} \\
&= e^{f(\theta)t'} \times n^{-\theta/\log 2} \\
&\leq e^{f(\theta)((1/\kappa_1+\epsilon)\log_2(n)-1)} \times n^{-\theta/\log 2} \\
&= e^{-f(\theta)} n^{(f(\theta)(1/\kappa_1+\epsilon)-\theta)/\log 2}.
\end{aligned}$$

To show that this goes to 0, we only need to find a $\theta < 0$, such that

$$f(\theta)(1/\kappa_1 + \epsilon) - \theta < 0.$$

As $f(\theta) < 0$ when $\theta < 0$, this inequality is identical to

$$\frac{\theta}{f(\theta)} < \frac{1}{\kappa_1} + \epsilon.$$

Note that $\theta/f(\theta)$ goes to $1/\kappa_1$ when θ goes to 0 from below, which means that for any $\epsilon > 0$, we can always find a $\theta < 0$ that fulfills this inequality. Thus the first part of the theorem follows.

The proof of part 2 is similar to that of part 1. But instead of Chernoff's bound for the left tail, we use the bound for the right tail. Write

$$t^* = \lceil (1/\kappa_1 - \epsilon)\log_2 n \rceil.$$

Since T_n is the smallest integer such that $\log_2 n \leq \sum_{i=1}^{T_n} G_i$, $T_n < t^*$ implies

$$\log_2 n < \sum_{i=1}^{t^*} G_i,$$

we can apply Chernoff's bound for the right tail to show that the probability of this event goes to 0 when n goes to infinity. For $\theta \in (0, \theta_0)$, we have

$$\begin{aligned}
&P\left\{T_n < (1/\kappa_1 - \epsilon)\log_2 n\right\} \\
&\leq P\left\{\sum_{i=1}^{t^*} G_i > \log_2 n\right\} \\
&\leq \left(\mathbb{E}\left[e^{\theta G_1}\right]\right)^{t^*} / e^{\theta \log_2 n} \\
&= e^{f(\theta)t^*} \times n^{-\theta/\log 2} \\
&< e^{f(\theta)((1/\kappa_1 - \epsilon)\log_2(n)+1)} \times n^{-\theta/\log 2} \\
&= e^{f(\theta)} n^{(f(\theta)(1/\kappa_1 - \epsilon) - \theta)/\log 2}.
\end{aligned}$$

To show that this goes to 0, we need to find a θ, such that $\theta_0 > \theta > 0$ and

$$f(\theta)(1/\kappa_1 - \epsilon) - \theta < 0.$$

As $f(\theta) > 0$ when $\theta > 0$, this inequality is identical to

$$\frac{\theta}{f(\theta)} > \frac{1}{\kappa_1} - \epsilon.$$

Note that $\theta/f(\theta)$ goes to $1/\kappa_1$ when θ goes to 0 from above, which means that for any $\epsilon > 0$, we can always find a $\theta > 0$ that fulfills this inequality. The second part of the theorem follows. □

4.3 Approximate results

Since $k = 1$ implies that $R_1, R_2, \ldots$ all have approximately geometric $(1/2)$ distribution, we can apply Theorem 4.2.1 to get the following corollary which give us a rough idea of the distribution of $\widehat{T}$ when $k = 1$.

Corollary 4.3.1. *Let $G_1, G_2, \ldots$ be i.i.d. geometric $(1/2)$ random variables, i.e.,*

$$P\{G_1 = i\} = 1/2^i, \quad i = 1, 2 \ldots \quad .$$

Let

$$T_n = \min \left\{ t : \log_2 n \leq \sum_{i=1}^{t} G_i \right\}.$$

We have

$$\frac{T_n}{\log_2 n} \xrightarrow{p} \frac{1}{2}.$$

Proof. Note for all $\theta_0 \in (0, \log 2)$, $\mathbb{E}\left[\exp(\theta_0 G_1)\right]$ exists because

$$\mathbb{E}\left[\exp(\theta_0 G_1)\right] = \sum_{i=1}^{\infty} \left(\frac{e^{\theta_0}}{2}\right)^i = \frac{e^{\theta_0}}{2 - e^{\theta_0}}.$$

Therefore G_1 fulfills all the conditions of Theorem 4.2.1. Since $\mathbb{E}\left[G_1\right] = 2$, it follows that

$$\frac{T_n}{\log_2 n} \xrightarrow{p} \frac{1}{2}.$$

□

When $k > 1$, if we ignore that IDs are selected into a bucket from candidates without replacement for the moment, then we can think how R_1 is decided as follows. Draw a line of length $\sum_{j=1}^{d} N_j$. Cut the line from left to right into segments of lengths $N_1, N_2, \ldots$ respectively. Then we choose k points on this line uniformly at random, and let the index of the segment in which the rightmost point is selected be R_1. Therefore, event $[R_1 \leq i]$ happens if and only

if all the k points fall into the first i segments, which implies

$$P\{R_1 \leq i\} = \left(\frac{\sum_{r=1}^{i} N_r}{\sum_{j=1}^{d} N_j}\right)^k, \quad i = 1, \ldots, d.$$

If we assume that $d \to \infty$ and $(N_j)_{j\geq 1}$ take values close to $n/4, n/8, \ldots$, which is close to the truth when n is large, then the last equation simplifies to

$$P\{R_1 \leq i\} = \left(1 - \frac{1}{2^i}\right)^k, \quad i = 1, \ldots \quad .$$

Thus, when $k > 1$, the distribution of R_1 is about

$$P\{R_1 = i\} = \left(1 - \frac{1}{2^i}\right)^k - \left(1 - \frac{1}{2^{i-1}}\right)^k, \quad i = 1, \ldots, \tag{4.2}$$

and the expectation of R_1 is about

$$\mathbb{E}[R_1] = \sum_{i=0}^{\infty} 1 - (1 - 1/2^i)^k. \tag{4.3}$$

Note that these two equations still work for $k = 1$. Thus we get the general form of the approximate distribution and expectation of $R_1, \ldots$ for all $k \geq 1$.

Although Equation (4.3) is a complex polynomial of k, the following lemma gives us tight bounds on it.

Lemma 4.3.2. *Let G_1 be a random variable with distribution*

$$P\{G_1 = i\} = (1 - 1/2^i)^k - (1 - 1/2^{i-1})^k, \quad k \geq 1, i \geq 1.$$

We have its expectation bounded from below and above as

$$\frac{H_k}{\log 2} \leq \mathbb{E}[G_1] \leq \frac{H_k}{\log 2} + 1.$$

where $H_k = \sum_{i=1}^{k} 1/i$, also known as the k-th harmonic number.

Proof. We can bound $\mathbb{E}[G_1]$ by integration. First we show that

$$\begin{aligned}
&\int_0^\infty \left(1-\left(1-\frac{1}{2^x}\right)^k\right)\mathrm{d}x \\
&= \int_0^1 \left(1-\left(1-\frac{1}{2^x}\right)^k\right)\frac{\mathrm{d}\left(1-\frac{1}{2^x}\right)}{\log 2\left(1-\left(1-\frac{1}{2^x}\right)\right)} \\
&= \int_0^1 \frac{1}{\log 2}\frac{1-w^k}{1-w}\mathrm{d}w \\
&= \int_0^1 \frac{1}{\log 2}\left(\sum_{i=1}^k w^{i-1}\right)\mathrm{d}w \\
&= \frac{1}{\log 2}\sum_{i=1}^k \frac{1}{i} \\
&= \frac{H_k}{\log 2}.
\end{aligned}$$

Since $\mathbb{E}[G_1]$ can be represented as

$$\mathbb{E}[G_1] = \sum_{i=1}^\infty P\{G_1 \geq i\} = \sum_{i=0}^\infty [1-P\{G_1 \leq i\}] = \sum_{i=0}^\infty \left[1-\left(1-\frac{1}{2^i}\right)^k\right],$$

and $1-(1-1/2^i)^k$ is a decreasing function of i, we have

$$\mathbb{E}[G_1] \geq \int_0^\infty \left(1-\left(1-\frac{1}{2^x}\right)^k\right)\mathrm{d}x = \frac{H_k}{\log 2}.$$

We also have

$$\begin{aligned}
\mathbb{E}[G_1] &= 1+\sum_{i=1}^\infty \left[1-\left(1-\frac{1}{2^i}\right)^k\right] \\
&\leq 1+\int_0^\infty \left(1-\left(1-\frac{1}{2^x}\right)^k\right)\mathrm{d}x \\
&= 1+\frac{H_k}{\log 2}.
\end{aligned}$$

□

We can apply Theorem 4.2.1 to get the following corollary which gives us a rough idea of the distribution of $\widehat{T}$ for all $k > 1$.

Corollary 4.3.3. *Let k be a positive integer. Let $G_1, G_2, \ldots$ be i.i.d. random variables with distribution*

$$P\{G_1 = i\} = (1 - 1/2^i)^k - (1 - 1/2^{i-1})^k, \quad k \geq 1, i \geq 1.$$

Let

$$T_n = \min\left\{t : \log_2 n \leq \sum_{i=1}^{t} G_i\right\}.$$

We have

$$\frac{T_n}{\log n} \xrightarrow{p} \frac{1}{g(k)},$$

where $g(k) = \log(2) \times \mathbb{E}[G_1] = H_k + o(1)$.

Proof. Note for all $\theta_0 \in (0, \log 2)$, $\mathbb{E}[\exp(\theta_0 G_1)]$ exists because

$$\begin{aligned}
&\mathbb{E}[\exp(\theta_0 G_1)] \\
&= \sum_{i=1}^{\infty} e^{\theta_0 i}\left(\left(1 - \frac{1}{2^i}\right)^k - \left(1 - \frac{1}{2^{i-1}}\right)^k\right) \\
&\leq \sum_{i=1}^{\infty} e^{\theta_0 i}\frac{k}{2^i} \\
&= \frac{k e^{\theta_0 i}}{2 - e^{\theta_0 i}}.
\end{aligned}$$

Therefore G_1 fulfills all the conditions of Theorem 4.2.1 and we have

$$\frac{T_n}{\log_2 n} \xrightarrow{p} \frac{1}{\mathbb{E}[G_1]}.$$

It follows from Lemma 4.3.2 that $\mathbb{E}[G_1] = g(k)/\log 2$ where

$$g(k) = H_k + o(1).$$

Together with last statement, we have

$$\frac{T_n}{\log n} \xrightarrow{p} \frac{1}{g(k)}.$$

□

4.4 The distribution of IDs

We establish a relationship between the distribution of $(G_i)_{i\geq 1}$ and $(R_i)_{i\geq 1}$. We first study the distribution of IDs. Let

$$M_r = \sum_{j=r+1}^{d} N_j, \quad r = 0, 1, \ldots, d-1.$$

In other words, we have

$$M_r = \sum_{i=1}^{n} \mathbf{1}_{[\ell(X_i,\overline{1})>r]}, \quad r = 0, \ldots, d-1.$$

Since $\mathbf{1}_{[\ell(X_i,\overline{1})>r]}$ is Bernoulli $(1/2^{r+1})$, we have

$$\mathbb{E}\left[M_r\right] = \frac{n}{2^{r+1}}, \quad r = 0, 1, \ldots, d-1.$$

From the perspective of the ID trie, M_0 is the total number of leaves under the internal node that is associated with string "1", M_1 is the total number of leaves under the internal node that is associated with string "11", etc., as shown in Figure 4–3. Therefore, if for z_i, the i-th hop of $\hat{\rho}$, we have $\ell(z_i, \overline{1}) = r$, then the number of candidate leaves for the next hop z_{i+1} is M_r. The following lemma shows how "close" $N_1, \ldots, N_d$, and $M_0, \ldots, M_{d-1}$ are to their expectations.

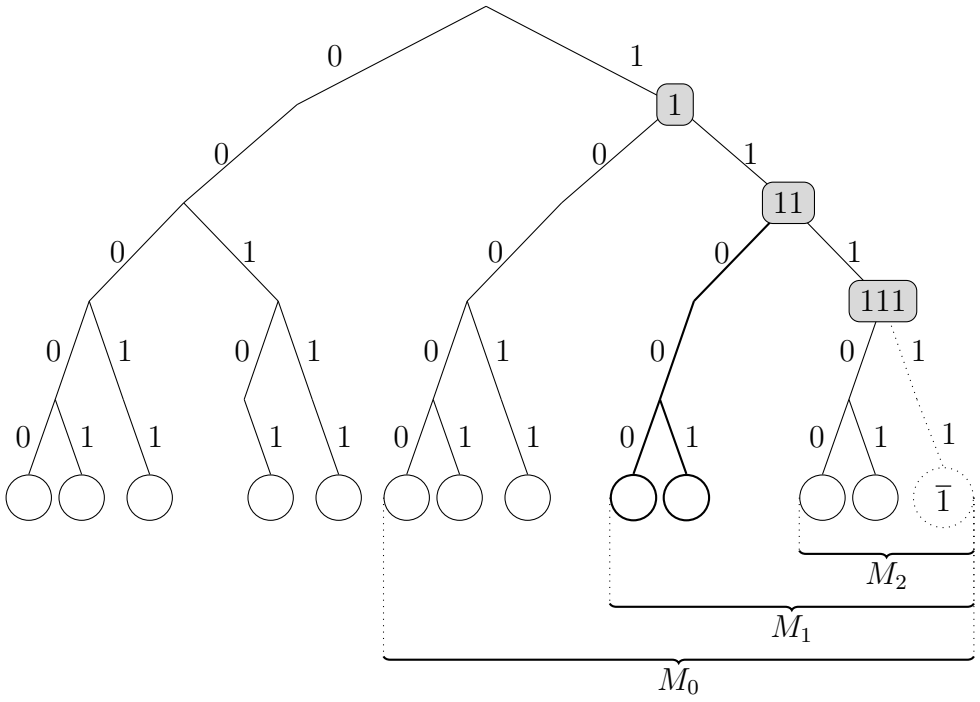

Figure 4–3: An example of $M_0, \ldots, M_{d-1}$. Each leaf node in this trie represents a randomly selected ID.

Lemma 4.4.1. *For $\epsilon > 0$, we have*

$$P\left\{\left|N_r - \frac{n}{2^{r+1}}\right| > \frac{\epsilon n}{2^{r+1}}\right\} \leq 2\exp\left\{-\frac{n}{2^{r+1}} \times \frac{\epsilon^2}{2(1+\epsilon/3)}\right\} \quad r = 1, \ldots, d-1.$$

and

$$P\left\{\left|M_r - \frac{n}{2^{r+1}}\right| > \frac{\epsilon n}{2^{r+1}}\right\} \leq 2\exp\left\{-\frac{n}{2^{r+1}} \times \frac{\epsilon^2}{2(1+\epsilon/3)}\right\} \quad r = 0, \ldots, d-1.$$

To prove this lemma, we need Bernstein's inequality [12, (2.13)].

Theorem 4.4.2. *If $Z_1, \ldots, Z_n$ are independent random variables which have*

$$\mathbb{E}[Z_i] = 0, \quad \mathbf{Var}(Z_i) = \sigma^2, \quad Z_i \leq b, \qquad i = 1, 2 \ldots, n,$$

then for $0 < t < b$, we have

$$P\left\{\frac{\sum_{i=1}^n Z_i}{n} \geq t\right\} \leq \exp\left\{-\tau \frac{\lambda}{2(1+\lambda/3)}\right\},$$

where

$$\lambda = \frac{bt}{\sigma^2}, \quad \tau = \frac{nt}{b}.$$

Proof of Lemma 4.4.1. Given $1 \leq r \leq d-1$, let

$$Y_i = \mathbf{1}_{[\ell(X_i,\overline{1})=r]}, \quad i = 1, \ldots, n.$$

Then we have $N_r = \sum_{i=1}^{n} Y_i$. Note that $Y_1, \ldots, Y_n$ are not independent: they can be seen as samples from a finite population without replacement. Note that Theorem 4.4.2 still applies [12]. To fulfill all conditions of Bernstein's inequality, we define

$$Z_i = Y_i - \frac{1}{2^{r+1}}, \quad i = 1, 2, \ldots, n,$$

which are random variables with zero mean. Since $Y_1, \ldots, Y_n$ are Bernoulli $(1/2^{r+1})$, we have

$$\sigma^2 = \mathbf{Var}\left(Z_1\right) = \mathbf{Var}\left(Y_1\right) = \frac{1}{2^{r+1}}\left(1 - \frac{1}{2^{r+1}}\right),$$

and $Z_1, \ldots, Z_n$ have upper bound

$$b = 1 - \frac{1}{2^{r+1}}.$$

Writing

$$t = \frac{\epsilon}{2^{r+1}},$$

we have

$$\tau = \frac{nt}{b} = \frac{n\epsilon}{2^{r+1}-1}, \quad \lambda = \frac{bt}{\sigma^2} = \epsilon.$$

It follows form Theorem 4.4.2 that

$$\begin{aligned}
&P\left\{N_r - \frac{n}{2^{r+1}} \geq \frac{\epsilon n}{2^{r+1}}\right\} \\
&= P\left\{\frac{\sum_{i=1}^n Z_i}{n} \geq t\right\} \\
&\leq \exp\left\{-\tau \frac{\lambda}{2\left(1+\lambda/3\right)}\right\} \\
&= \exp\left\{-\frac{n}{2^{r+1}-1} \times \frac{\epsilon^2}{2(1+\epsilon/3)}\right\} \\
&\leq \exp\left\{-\frac{n}{2^{r+1}} \times \frac{\epsilon^2}{2(1+\epsilon/3)}\right\}.
\end{aligned}$$

Note that as $r \geq 1$, we have

$$\begin{aligned}
-Z_i &= -Y_i + \frac{1}{2^{r+1}} \\
&\leq \frac{1}{2^{r+1}} \\
&\leq 1 - \frac{1}{2^{r+1}} \\
&= b.
\end{aligned}$$

Since $-Z_1, \ldots, -Z_n$ have the same upper bound b and variance σ^2 as $Z_1, \ldots, Z_n$ do, we also have

$$\begin{aligned}
&P\left\{N_r - \frac{n}{2^{r+1}} \leq -\frac{\epsilon n}{2^{r+1}}\right\} \\
&= P\left\{\frac{\sum_{i=1}^n (-Z_i)}{n} \geq t\right\} \\
&\leq \exp\left\{-\frac{n}{2^{r+1}} \times \frac{\epsilon^2}{2(1+\epsilon/3)}\right\}.
\end{aligned}$$

Together, we have,

$$\begin{aligned}
&P\left\{\left|N_r - \frac{n}{2^{r+1}}\right| > \frac{\epsilon n}{2^{r+1}}\right\} \\
&= P\left\{N_r - \frac{n}{2^{r+1}} \geq \frac{\epsilon n}{2^{r+1}}\right\} + P\left\{N_r - \frac{n}{2^r} \leq -\frac{\epsilon n}{2^{r+1}}\right\} \\
&\leq 2\exp\left\{-\frac{n}{2^{r+1}} \times \frac{\epsilon^2}{2(1+\epsilon/3)}\right\}.
\end{aligned}$$

Since

$$M_r = \sum_{i=1}^{n} \mathbf{1}_{[\ell(X_i,\bar{1})>r]}, \quad r = 0, \ldots, d-1,$$

and $\mathbf{1}_{[\ell(X_i,\bar{1}]\geq r]}$ are Bernoulli $(1/2^{r+1})$, the previous argument about N_r also works for M_r. □

Lemma 4.4.3. *For all $\epsilon > 0$, and all integers $r \in [0, d), j \in [1, d-r), m \in [1, 2^d]$, we have*

$$P\left\{\left|N_{r+j} - \frac{m}{2^j}\right| > \frac{\epsilon m}{2^j} \;\middle|\; M_r = m\right\} < 2\exp\left\{-\frac{m}{2^j} \times \frac{\epsilon^2}{2(1+\epsilon/3)}\right\}.$$

Proof. Given $M_r = m$, we know that m IDs in $X_1, \ldots, X_n$ have the format

$$(\underbrace{1, 1, \ldots, 1}_{r+1}, \ldots).$$

It can be seen as that these m IDs have their first $r+1$ bits fixed. If we only look at the remaining $d-r-1$ bits, these IDs can be seen as chosen uniformly at random from $\{0,1\}^{d-r-1}$ without replacement. Therefore, the same argument for $N_1, N_2, \ldots$ in Lemma 4.4.1 can be applied to $N_{r+1}, \ldots, N_d$ and give us the same result. □

4.5 Exact Results

In this section, we use coupling to establish the precise relationship between $(G_t)_{t\geq 1}$ and $(R_t)_{t\geq 1}$, which allows us to relate T_n and $\widehat{T}$, and to give the expectation of $\widehat{T}$.

We slightly change how each hop of the path $\hat{\rho}$ is chosen in a way that does not change the distributions of $(R_i)_{i\geq 1}$. At the beginning, let $z_0 = X_0 = \overline{1}$. Let $L_1 = \ell(z_0, \overline{1}) = 0$. Let $T'_n = 0$. Then we repeat step (1), (2), (3):

(1) Given $t \geq 1$ and L_t, then as shown in Figure 4–4, the number of candidate leaves for next hop z_t in the ID trie is M_{L_t}, and these candidates can be partitioned in to $d - L_t$ subtrees of size $N_{L_t+1}, \ldots, N_d$ according the length of their common prefix with $\overline{1}$. Instead of directly choosing k samples uniformly at random from the M_{L_t} leaves without replacement and making the rightmost one z_t, we label the candidate leaves from left to right with index set $\{1, \ldots, M_{L_t}\}$. Let $(B_{t,i})_{i\geq 1}$ be a sequence of i.i.d. continuous random variables uniformly distributed on $[0, M_{L_t}]$. Let $\overline{B}_{t,i} = \lceil B_{t,i} \rceil$ for all $i \geq 0$. Then among the first k unique members of $(\overline{B}_{t,i})_{i\geq 1}$, we choose the largest one as B_t, and choose the candidate leaf B_t as next hop z_t. In other words, we are still choosing k leaves uniformly at random without replacement, but we are doing it by choosing with replacement and ignoring duplicates. Let $R_t = \ell(z_t, \overline{1}) - \ell(z_{t-1}, \overline{1})$ as before.

(2) Given $(B_{t,i})_{i\geq 1}$, let B'_t be the largest one among $B_{t,1}, \ldots, B_{t,k}$. Let $\mathbb{N} = \{1, 2, 3, \ldots\}$. Let

$$G_t = \min \left\{ j \in \mathbb{N} : B'_t < M_{L_t} \left(1 - \frac{1}{2^j}\right) \right\}.$$

Since $B_{t,1}, \ldots, B_{t,k}$ are i.i.d. uniformly distributed on $[0, M_{L_t}]$, the probability that $G_t \le j$ for an integer $j \ge 1$ is that given M_{L_t},

$$P\{G_t \le j \mid M_{L_t}\} = P\left\{ \bigcap_{r=1}^{k} \left[B_{t,r} \le M_{L_t} \left(1 - \frac{1}{2^j}\right)\right] \;\middle|\; M_{L_t} \right\}$$
$$= \left(1 - \frac{1}{2^j}\right)^k.$$

Therefore we have

$$P\{G_t = j \mid M_{L_t}\} = \left(1 - \frac{1}{2^j}\right)^k - \left(1 - \frac{1}{2^{j-1}}\right)^k, \quad j \in \mathbb{N}.$$

(3) Let $L_{t+1} = L_t + R_t = \ell(z_t, \overline{1})$. If $L_{t+1} \ge \log_2 n - \sqrt{\log_2 n}$, then let $T'_n = t$. If $L_{t+1} = d$ or $M_{L_{t+1}} = 0$, i.e., we have reached the rightmost leaf of the ID trie, then let $\widehat{T} = t$ and halt.

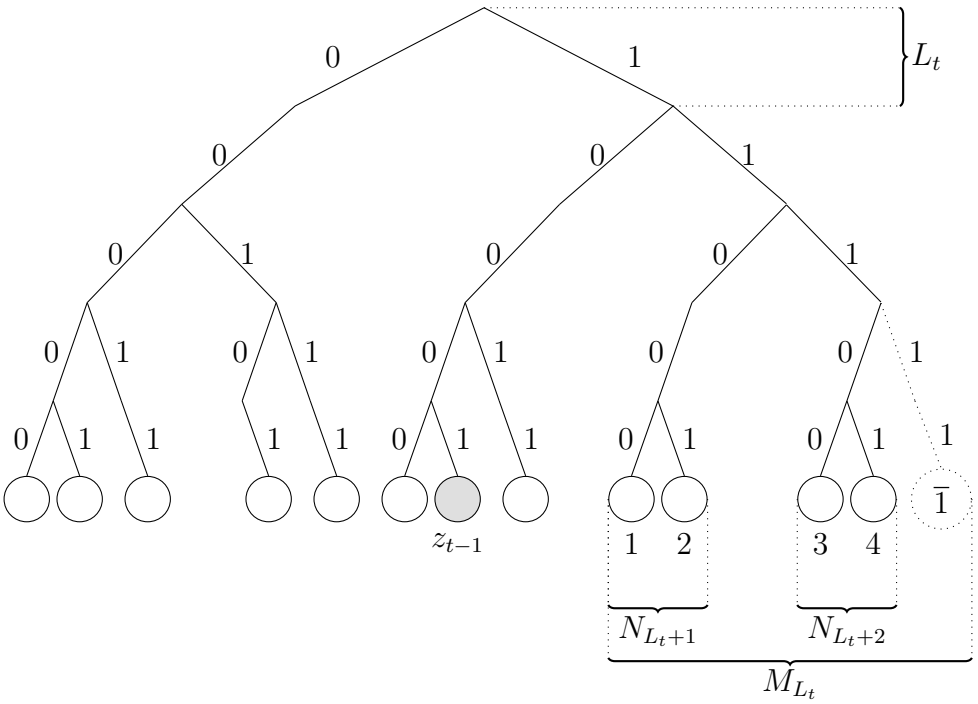

Figure 4–4: An example of L_t and M_{L_t} in which $M_{L_t} = 4$ and $L_t = 1$. The M_{L_t} candidate leaves for z_t are labeled with $1, \ldots, M_{L_t}$ from left to right.

We show that $T'_n > 0$ with high probability. Let $\ell' = \left\lceil \log_2 n - \sqrt{\log_2 n} \right\rceil - 1$. Put differently, letting $\mathcal{Z}$ be the set of all integers, we have

$$\ell' = \max \left\{ i \in \mathcal{Z} : i < \log_2 n - \sqrt{\log_2 n} \right\}.$$

Throughout this chapter, we assume that $n \geq 2^4$ so $\ell' > 0$. Note that for a positive integer t, $t < T'_n$ implies $L_t < \ell'$ because $(L_t)_{t \geq 1}$ is strictly increasing. For the same reason, we have that $T'_n \leq \ell' + 1$.

It might be easier to understand this construction from a geometric perspective, as shown in Figure 4–5. Given L_t and $t \geq 1$, consider line segments ℓ_G and ℓ_R of length M_{L_t} which are parallel to the x-axis in a plane, with all their leftmost endpoints on the y-axis. We partition ℓ_G into segments of length $M_{L_t}/2, M_{L_t}/4, \ldots$ from left to right and label them with index $\{1, 2 \ldots\}$. We partition ℓ_R into $d - L_t$ segments of length $N_{L_t+1}, N_{L_t+2}, \ldots, N_d$ and label them with index $\{1, 2 \ldots\}$. Since $(N_{L_t+i})_{i \geq 1}$ are random variables, the i-th segment on ℓ_R do not necessarily overlap with the i-th segment on ℓ_G. Let Δ_i be the interval between the right endpoints of the i-th segments. Note that the discrepancies between segments can be complex, as shown in Figure 4–6.

Figure 4–7 shows the uniform points $(B_{t,i})_{i \geq 1}$. Let $B'_t = \max(B_{t,i})_{i \leq k}$. Now if the number of unique unit intervals containing these k points is less than k, then keep generating uniform points until k unique intervals are found. Let K_t be the random total number of points that we have generated. Therefore, we have $K_t \geq k$. Let $B_t = \max(B_{t,i})_{i \leq K_t}$.

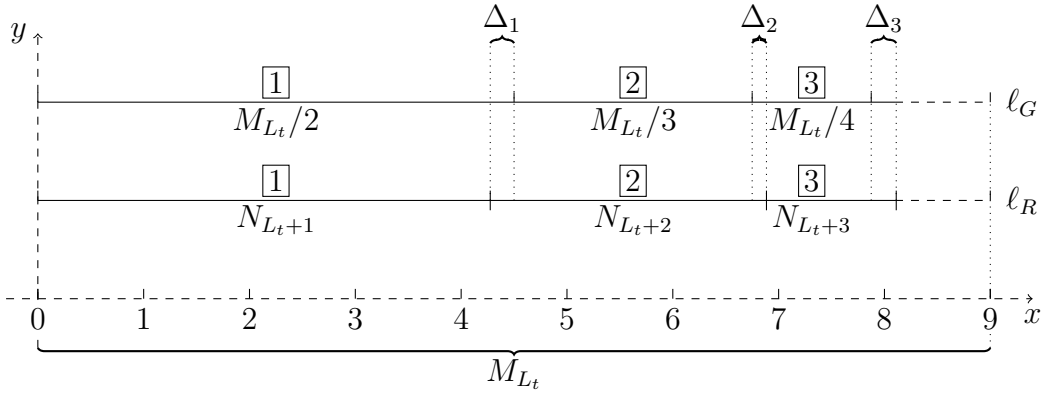

Figure 4–5: An example of the geometric construction in which $M_{L_t} = 9$. ℓ_R is partitioned into segments of length $(N_{L_t+i})_{i\geq 1}$. ℓ_G is partitioned into segments of length of $(M_{L_t}/2^i)_{i\geq 1}$. Δ_i is the interval between the right endpoint of the i-th segment of ℓ_R and that of the i-th segment of ℓ_G. The labels of segments are drawn in boxes.

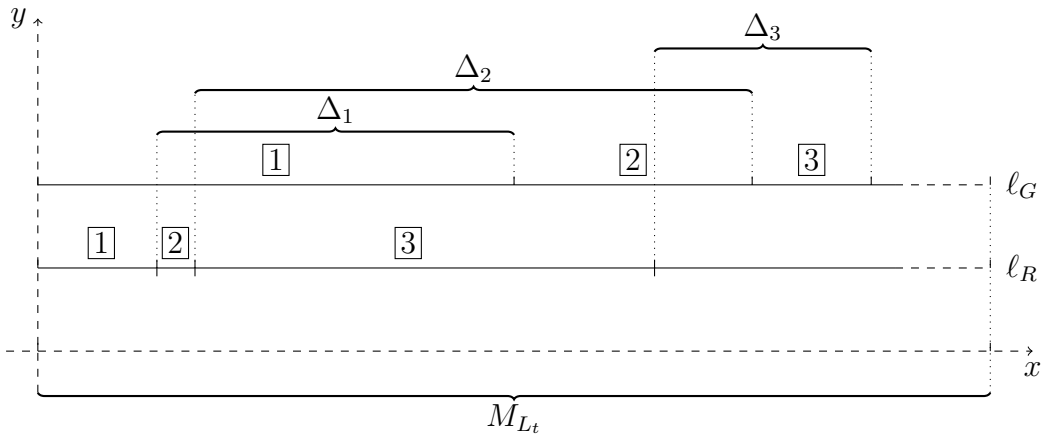

Figure 4–6: An example to illustrate a more complex situation of the geometric construction.

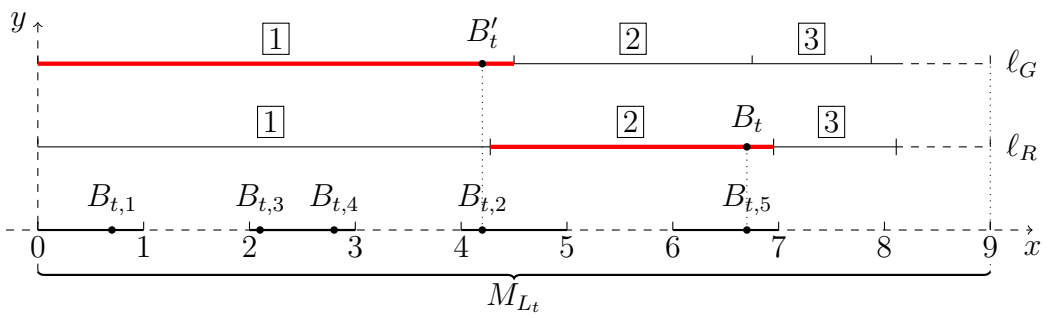

Figure 4–7: An example of $(B_{t,i})_{i\geq 1}$. We uniformly choose points $B_{t,1}, B_{t,2}, \ldots$ until there are k unique unit intervals containing these points. We choose the rightmost one among all of these points as B_t, and the rightmost one among the first k points as B'_t. In this example, $k = 4$, $K_t = 5$, $B_t = B_{t,5}$, $B'_t = B_{t,2}$, $R_t = 2$ and $G_t = 1$.

To define R_t, we draw a line parallel to the y-axis through $(B_t, 0)$. We choose the index of the segment of ℓ_R that intersects with this line as R_t. To define G_t, we draw another line parallel to the y-axis through $(B'_t, 0)$. We choose the index of the segment of ℓ_G that intersects this line to be G_t. It is easy to see that this geometric construction of $(R_t)_{t\geq 1}$ and $(G_t)_{t\geq 1}$ is equivalent to the ID trie construction.

In the remainder, we first show that we have $G_t = R_t$ for all $t < T'_n$ with high probability. Then we use our previous conclusion about $(G_t)_{t\geq 1}$ in chapter 4.3 to prove

$$\frac{T'_n}{\log n} \xrightarrow{p} \frac{1}{g(k)},$$

where $g(k) = H_k + o(1)$ is a function of k. It follows that we also have

$$\frac{\widehat{T}}{\log n} \xrightarrow{p} \frac{1}{g(k)}.$$

We define some notation used in following proofs. Define the events

$$A_r = \left[\bigcap_{j=1}^{\ell'-r} \left| N_{r+j} - \frac{M_r}{2^j} \right| \leq \frac{\epsilon M_r}{2^j}\right], \quad r = 0, \ldots, \ell' - 1, \quad \epsilon \in (0, 1).$$

If we have $L_t = r$, then A_r implies that at step t, ℓ_R and ℓ_G are partitioned in very similar ways. Define the events

$$B_r = \left[\left|M_r - \frac{n}{2^{r+1}}\right| \le \frac{\epsilon n}{2^{r+1}}\right], \quad r = 0, \ldots, \ell' - 1, \quad \epsilon \in (0, 1).$$

Define

$$f(\epsilon) = \frac{\epsilon^2}{2(1 + \epsilon/3)}, \quad \epsilon > 0.$$

Lemma 4.5.1. *We have*

$$P\left\{T'_n = 0\right\} \le \min_{\epsilon > 0} \left[2 \exp\left\{-\frac{n f(\epsilon)}{2^{\ell'+1}}\right\}\right].$$

Proof. Let the rightmost leaf in the ID trie be y', which must be the last hop of $\hat{\rho}$, i.e., $z_{\widehat{T}} = y'$. Thus we have $L_{\widehat{T}+1} = \ell(y', \overline{1})$ which is the length of the common prefix of y' and $\overline{1}$, as shown in Figure 4–8.

By definition of T'_n, $T'_n = 0$ if and only if that for all $t \in [1, \widehat{T}]$, we have

$$L_{t+1} < \log_2 n - \sqrt{\log_2 n}.$$

Since $(L_t)_{t \ge 1}$ is a strictly increasing sequence, we have

$$L_{\widehat{T}+1} < \log_2 n - \sqrt{\log_2 n}.$$

Recalling that $\ell' = \left\lceil \log_2 n - \sqrt{\log_2 n} \right\rceil - 1$, the last inequality implies that $L_{\widehat{T}} < \ell'$.

As y' is the rightmost ID in the trie, there must be no ID on its right hand side, which implies that $M_{L_{\widehat{T}}} = 0$. Since $(M_i)_{i \ge 0}$ is a decreasing sequence and $L_{\widehat{T}} < \ell'$, we have

$$M_{\ell'} \le M_{L_{\widehat{T}}} = 0.$$

It follows from Lemma 4.4.1 that for all $\epsilon > 0$

$$\begin{aligned} P\{T'_n = 0\} &\le P\{M_{\ell'} = 0\} \\ &\le P\left\{\left|M_{\ell'} - \frac{n}{2^{\ell'+1}}\right| > \frac{\epsilon n}{2^{\ell'+1}}\right\} \\ &\le 2\exp\left\{-\frac{nf(\epsilon)}{2^{\ell'+1}}\right\}. \end{aligned}$$

□

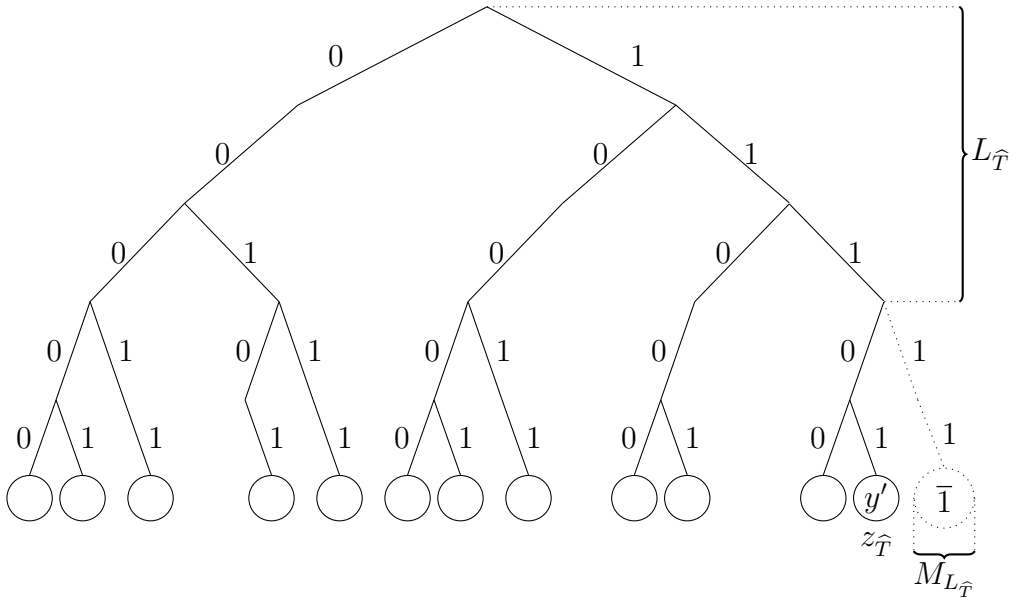

Figure 4–8: An example of $L_{\widehat{T}}$ and $M_{L_{\widehat{T}}}$. The rightmost leaf y' must be $z_{\widehat{T}}$, the last hop of $\hat{\rho}$. It is easy to see that $M_{L_{\widehat{T}}} = 0$.

Lemma 4.5.2. *For all integers $r \in [0, \ell' - 1]$, we have*

$$P\{A_r^c \cap B_r\} \le \min_{\epsilon\in(0,1)} \left[2(\epsilon n + 1)\ell' \exp\left\{-\frac{(1-\epsilon)n}{2^{\ell'+1}} \times f(\epsilon)\right\}\right].$$

Proof. We denote the interval

$$\left[\frac{(1-\epsilon)n}{2^{r+1}}, \frac{(1+\epsilon)n}{2^{r+1}}\right]$$

by $\mathcal{I}$. Let $|\mathcal{I}|$ be the length of $\mathcal{I}$, i.e., $|\mathcal{I}| = \epsilon n/2^r$. In the trivial case that $|\mathcal{I}| < 1$, there is no integer in $\mathcal{I}$, which implies that B_r cannot happen. Thus we

have

$$P\{A_r^c \cap B_r\} \leq P\{B_r\} = 0.$$

In the other case that $|\mathcal{I}| \geq 1$, let $m \in \mathcal{I}$ be an integer. It follows from Lemma 4.4.3 that

$$\begin{aligned} P\{A^c \cap [M_r = m]\} &\leq P\left\{ \bigcup_{j=1}^{\ell'-r} \left[\left| N_{r+j} - \frac{m}{2^j} \right| > \frac{\epsilon m}{2^j} \right] \;\middle|\; M_r = m \right\} \\ &\leq \sum_{j=1}^{\ell'-r} 2\exp\left\{ -\frac{m}{2^j} \times f(\epsilon) \right\} \\ &\leq 2\ell' \exp\left\{ -\frac{m}{2^{\ell'-r}} \times f(\epsilon) \right\} \\ &\leq 2\ell' \exp\left\{ -\frac{(1-\epsilon)n}{2^{\ell'+1}} \times f(\epsilon) \right\}. \end{aligned}$$

The number of integers in $\mathcal{I}$ is at most

$$|\mathcal{I}| + 1 = \frac{\epsilon n}{2^r} + 1 \leq \epsilon n + 1.$$

Summing over all $m \in \mathcal{I}$, we have

$$\begin{aligned} P\{A_r^c \cap B_r\} &= \sum_{m \in \mathcal{I}} P\{A^c \cap [M_r = m]\} \\ &\leq 2(\epsilon n + 1)\ell' \exp\left\{ -\frac{(1-\epsilon)n}{2^{\ell'+1}} \times f(\epsilon) \right\}. \end{aligned}$$

□

Lemma 4.5.3. *For all integers $r \in [0, \ell' - 1]$, $t > 0$ and $k > 0$, we have*

$$P\{R_t \neq G_t \mid E(r,t)\} \leq \min_{\epsilon \in (0,1)} \left[\frac{2^{\ell'+1}k^2}{(1-\epsilon)^2 n} + \frac{2k\epsilon\ell'}{(1-\epsilon)} \right],$$

where $E(r,t)$ is defined by

$$E(r,t) = [t < T_n'] \cap [L_t = r] \cap A_r \cap B_r.$$

Proof. Given that $0 < t < T_n'$, there are two reasons that can cause $R_t \neq G_t$. First, we might have $K_t > k$, which implies that among $(B_{t,i})_{i \leq k}$, more than

one fall into the same unit interval. Therefore we have to choose more than k points to decide R_t, which might lead to $R_t \neq G_t$, as in the example shown in Figure 4–7. Put differently, we have $\overline{B}_{t,r} = \overline{B}_{t,s}$ for some $r, s \in [1, k]$ and $r \neq s$, i.e., $\lceil B_{t,r} \rceil = \lceil B_{t,s} \rceil$.

Let $N' = \sum_{j=1}^{\ell'-r} N_{r+j}$. Given that $t < T'_n$ and $L_t = r$, all the first k points must be selected from interval $[0, N']$, otherwise we would have $t \geq T'_n$. Therefore, the probability that two of these points fall in the same unit interval is $1/N'$. Given that B_r happens, we have

$$M_r \geq (1 - \epsilon) \frac{n}{2^{r+1}}.$$

It follows from A_r, this inequality, and $r < \ell'$ that

$$N_{r+1} \geq (1 - \epsilon)\frac{M_r}{2} \geq (1 - \epsilon)^2 \frac{n}{2^{r+2}} \geq (1 - \epsilon)^2 \frac{n}{2^{\ell'+1}}.$$

Therefore, we have

$$N' = \sum_{j=1}^{\ell'-r} N_{r+j} \geq N_{r+1} \geq (1 - \epsilon)^2 \frac{n}{2^{\ell'+1}}.$$

Thus we can get

$$\begin{aligned} P\{k \neq K_t \mid E(t, r)\} &= P\left\{ \cup_{\substack{r,s \in [1,k] \\ r \neq s}} \left[\overline{B}_{t,r} = \overline{B}_{t,s} \right] \;\middle|\; E(t, r) \right\} \\ &\leq k^2 P\left\{ \left[\overline{B}_{t,1} = \overline{B}_{t,2} \right] \mid E(t, r) \right\} \\ &\leq \frac{k^2}{N'} \\ &\leq \frac{2^{\ell'+1} k^2}{(1 - \epsilon)^2 n}. \end{aligned}$$

Another reason that might cause the discrepancy between G_t and R_t is that the segments of ℓ_G and ℓ_R do not overlap completely. For instance, for an integer $i \in [1, \ell' - r]$, the right endpoint of the i-th segment of ℓ_G has

x-coordinate $\sum_{j=1}^{i} M_r/2^j$, but the right endpoint of the i-th segment of ℓ_R has x-coordinate $\sum_{j=1}^{i} N_{r+j}$. Let the interval between the two coordinates be Δ_i, i.e.,

$$\Delta_i = \left[\min\left\{ \sum_{j=1}^{i} \frac{M_r}{2^j}, \sum_{j=1}^{i} N_{r+j} \right\}, \max\left\{ \sum_{j=1}^{i} \frac{M_r}{2^j}, \sum_{j=1}^{i} N_{r+j} \right\} \right].$$

If we have $B'_t = \max(B_{t,j})_{j \le k} \in \Delta_i$, then we might have that one of R_t and G_t is at most i and the other is at least $i+1$, as shown in Figure 4–9.

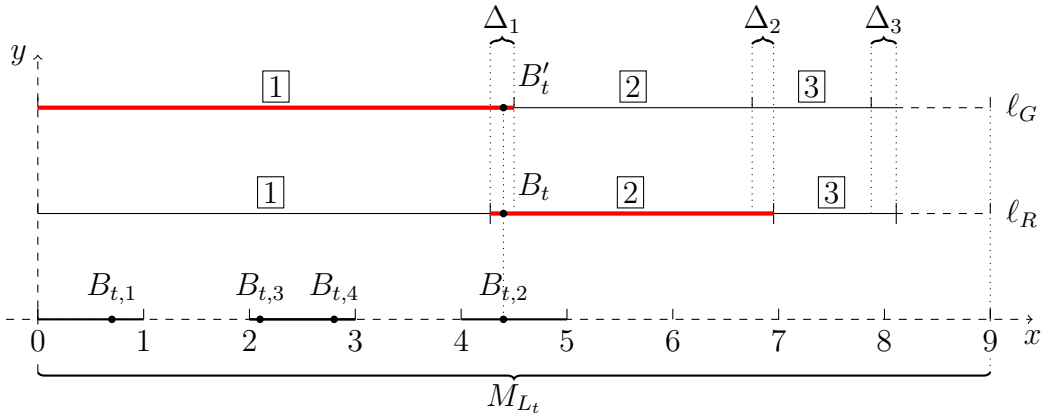

Figure 4–9: An example of the second possible reason for $R_t \neq G_t$. Δ_i is the interval between the right endpoint of the i-th segment of ℓ_R and that of the i-th segment of ℓ_G. In this example, $K_t = k = 4$, $B_t = B'_t \in \Delta_1$, $G_t = 1$ and $R_t = 2$.

Given that event A_r happens, by triangle inequality, we can bound the width of Δ_i, denoted by $|\Delta_i|$, as follows:

$$\begin{aligned} |\Delta_i| &= \left| \sum_{j=1}^{i} \frac{M_r}{2^j} - \sum_{j=1}^{i} N_{r+j} \right| \\ &\le \sum_{j=1}^{i} \left| \frac{M_r}{2^j} - N_{r+j} \right| \\ &\le \sum_{j=1}^{i} \frac{\epsilon M_r}{2^j} \\ &\le \epsilon M_r. \end{aligned}$$

Let the union of all such intervals be $\Delta = \cup_{i=1}^{\ell'-r}\Delta_i$. Then we can bound the width of Δ as

$$|\Delta| \leq \sum_{i=1}^{\ell'-r} |\Delta_i| \leq \epsilon \ell' M_r.$$

Again, given that $t > 0$ and $t < T'_n$, we must have that $(B_{t,j})_{j\leq k}$ all fall in the interval $[0, N']$. Since they are selected uniformly at random, we have

$$P\left\{B_{t,1} \in \Delta \mid E(t,r)\right\} = \frac{|\Delta|}{N'} \leq \frac{|\Delta|}{N_{r+1}} \leq \frac{\epsilon\ell' M_r}{(1-\epsilon)M_r/2} = \frac{2\epsilon\ell'}{(1-\epsilon)}.$$

Therefore, the probability that $B_t \in \Delta$ is bounded by

$$P\left\{B_t \in \Delta \mid E(t,r)\right\} \leq kP\left\{B_{t,1} \in \Delta \mid E(t,r)\right\} \leq \frac{2k\epsilon\ell'}{(1-\epsilon)}.$$

In summary, we have

$$\begin{aligned} P\left\{R_t \neq G_t \mid E(t,r)\right\} &\leq P\left\{[k \neq K_t] \cup [B_t \in \Delta] \mid E(t,r)\right\} \\ &\leq \frac{2^{\ell'+1}k^2}{(1-\epsilon)^2 n} + \frac{2k\epsilon\ell'}{(1-\epsilon)}. \end{aligned}$$

□

Lemma 4.5.4. *For all integers $t > 0$ and $k > 0$, we have*

$$\begin{aligned} P\left\{[R_t \neq G_t] \cap [t < T'_n]\right\} \leq \min_{\epsilon\in(0,1)} \Bigg[&\frac{2^{\ell'+1}k^2\ell'}{(1-\epsilon)^2 n} + \frac{2k\epsilon(\ell')^2}{(1-\epsilon)} \\ &+ 2(\epsilon n + 1)(\ell')^2 \exp\left\{-\frac{(1-\epsilon)n}{2^{\ell'+1}} \times f(\epsilon)\right\} \\ &+ 4\ell' \exp\left\{-\frac{n}{2^{\ell'+1}} \times f(\epsilon)\right\}\Bigg]. \end{aligned}$$

Proof. Let $r < \ell'$ be a positive integer. Write the event

$$E(r,t) = [t < T'_n] \cap [L_t = r] \cap A_r \cap B_r.$$

It is easy to verify that

$$
\begin{aligned}
&P\left\{[G_t \neq R_t] \cap [t < T'_n] \cap [L_t = r]\right\} \\
&\quad \leq P\left\{[G_t \neq R_t] \cap [t < T'_n] \cap [L_t = r] \cap A_r \cap B_r\right\} + P\left\{A_r^c\right\} + P\left\{B_r^c\right\} \\
&\quad = P\left\{[G_t \neq R_t] \cap E(r,t)\right\} + P\left\{A_r^c\right\} + P\left\{B_r^c\right\} \\
&\quad \leq P\left\{[G_t \neq R_t] \mid E(r,t)\right\} + P\left\{A_r^c \cap B_r\right\} + 2P\left\{B_r^c\right\}.
\end{aligned}
$$

Plugging the inequalities given by Lemma 4.4.1, Lemma 4.5.2 and Lemma 4.5.3, we have

$$
\begin{aligned}
&P\left\{[G_t \neq R_t] \cap [t < T'_n] \cap [L_t = r]\right\} \\
&\quad \leq \frac{2^{\ell'+1}k^2}{(1-\epsilon)^2 n} + \frac{2k\epsilon\ell'}{(1-\epsilon)} + 2(\epsilon n + 1)\ell' \exp\left\{-\frac{(1-\epsilon)n}{2^{\ell'+1}} \times f(\epsilon)\right\} \\
&\qquad + 4\exp\left\{-\frac{n}{2^{r+1}} \times f(\epsilon)\right\}.
\end{aligned}
$$

Summing over all possible r, we have

$$
\begin{aligned}
&P\left\{[G_t \neq R_t] \cap [t < T'_n]\right\} \\
&\quad = \sum_{r=1}^{\ell'-1} P\left\{[G_t \neq R_t] \cap [t < T'_n] \cap [L_t = r]\right\} \\
&\quad \leq \frac{2^{\ell'+1}k^2\ell'}{(1-\epsilon)^2 n} + \frac{2k\epsilon(\ell')^2}{(1-\epsilon)} + 2(\epsilon n + 1)(\ell')^2 \exp\left\{-\frac{(1-\epsilon)n}{2^{\ell'+1}} \times f(\epsilon)\right\} \\
&\qquad + 4\ell' \exp\left\{-\frac{n}{2^{\ell'+1}} \times f(\epsilon)\right\}.
\end{aligned}
$$

□

Lemma 4.5.5. *We have*

$$
P\left\{[T'_n = 0] \cup \left[\cup_{t=1}^{T'_n - 1} R_t \neq G_t\right]\right\} \to 0,
$$

as $n \to \infty$.

Proof. Recall that $T'_n \leq \ell' + 1$. It follows from this inequality, Lemma 4.5.4 and Lemma 4.5.1 that

$$
\begin{aligned}
&P\left\{[T'_n = 0] \cup \left[\cup_{t=1}^{T'_n - 1} R_t \neq G_t\right]\right\} \\
&\quad \leq P\left\{\cup_{t=1}^{T'_n - 1} [R_t \neq G_t]\right\} + P\{T'_n = 0\} \\
&\quad \leq P\left\{\cup_{t=1}^{\ell'} \{[t < T'_n] \cap R_t \neq G_t\}\right\} + P\{T'_n = 0\} \\
&\quad \leq \sum_{t=1}^{\ell'} P\{[t < T'_n] \cap R_t \neq G_t\} + P\{T'_n = 0\} \\
&\quad \leq \ell' \times \max_{t \in [1,\ell']} [P\{[t < T'_n] \cap R_t \neq G_t\}] + P\{T'_n = 0\} \\
&\quad \leq \min_{\epsilon \in (0,1)} \left[\frac{2^{\ell'+1} k^2 (\ell')^2}{(1-\epsilon)^2 n} + \frac{2k\epsilon(\ell')^3}{(1-\epsilon)} + 2(\epsilon n + 1)(\ell')^3 \exp\left\{-\frac{(1-\epsilon)n}{2^{\ell'+1}} \times f(\epsilon)\right\}\right. \\
&\qquad \left. + 4(\ell')^2 \exp\left\{-\frac{n}{2^{\ell'+1}} \times f(\epsilon)\right\} + 2\exp\left\{-\frac{n}{2^{\ell'+1}} \times f(\epsilon)\right\}\right] \\
&\quad \stackrel{\text{def}}{=} \min_{\epsilon \in (0,1)} Z(\epsilon).
\end{aligned}
$$

Now choose $\epsilon = 2^{-m}$ where $m = (\log_2 n)^{1/4}$. Recall that

$$\ell' = \left\lceil \log_2 n - \sqrt{\log_2 n} \right\rceil - 1 < \log_2 n - \sqrt{\log_2 n}.$$

Therefore, we have

$$\ell' < (m^4 - m^2) < m^4,$$

and also

$$\frac{2^{\ell'}}{n} < 2^{-\sqrt{\log_2 n}} = 2^{-m^2}.$$

Thus we have

$$Z(\epsilon) \leq O(I_1 + I_2 + I_3 + I_4),$$

with

$$I_1 = \frac{2^{\ell'}(\ell')^2}{n}, \qquad I_2 = \epsilon(\ell')^3,$$
$$I_3 = \epsilon n(\ell')^3 \exp\left\{-\frac{n(1-\epsilon)f(\epsilon)}{2^{\ell'+1}}\right\}, \qquad I_4 = (\ell')^2 \exp\left\{-\frac{nf(\epsilon)}{2^{\ell'+1}}\right\}.$$

We verify that $Z(\epsilon)$ indeed goes to zero as $n \to \infty$. For I_1, I_2, we have

$$I_1 = \frac{2^{\ell'}(\ell')^2}{n} < m^8 2^{-m^2} = o(1),$$

and

$$I_2 = \epsilon\,(\ell')^3 < 2^{-m} m^{12} = o(1).$$

For I_3, I_4, we take a logarithm:

$$\begin{aligned}
\log_2(I_3) &= \log_2\left\{\epsilon n(\ell')^3 \exp\left\{-\frac{f(\epsilon)n}{2^{\ell'+1}}\right\}\right\} \\
&= \log_2\epsilon + \log_2 n + 3\log_2\ell' - \frac{n}{2^{\ell'+1}} \times \frac{\epsilon^2(1-\epsilon)}{2(1+\epsilon/3)} \\
&< -m + m^4 + 12\log_2 m - \frac{2^{m^2-2m}(1-2^{-m})}{4(1+2^{-m}/3)},
\end{aligned}$$

which goes to $-\infty$ as $n \to \infty$. Finally, we have

$$\begin{aligned}
\log_2(I_4) &= \log_2\left\{(\ell')^2 \exp\left\{-\frac{n}{2^{\ell'+1}} \times f(\epsilon)\right\}\right\} \\
&= 2\log_2\ell' - \frac{n}{2^{\ell'+1}} \times \frac{\epsilon^2}{2(1+\epsilon/3)} \\
&< 8\log_2 m - \frac{2^{m^2-2m}}{4(1+2^{-m}/3)},
\end{aligned}$$

which goes to $-\infty$ as $n \to \infty$. □

Lemma 4.5.6. *For all $t > 0$ and $k > 0$, if $r \in [0, \ell')$, we have*

$$P\{C \mid E(r)\} \le \min_{\epsilon\in(0,1)}\left[\frac{k^2 2^{\ell'}}{n(1-\epsilon)} + k\epsilon\right],$$

and if $r = \ell'$*, we have*

$$P\{C \mid E(r)\} = 0,$$

where

$$C = \left[\sum_{t=1}^{T'_n} G_t < \log_2 n - \sqrt{\log_2 n}\right],$$

and $E(r)$ *is defined by*

$$E(r) = [L_{T'_n} = r] \cap A_r \cap B_r \cap [T'_n > 0] \cap [\cap_{t=1}^{T'_n - 1} R_t = G_t].$$

Proof. Recall that

$$\ell' = \left\lceil \log_2 n - \sqrt{\log_2 n} \right\rceil - 1.$$

Given $T'_n > 0$, we have

$$\sum_{t=1}^{T'_n} R_t \geq \log_2 n - \sqrt{\log_2 n}.$$

Since $\sum_{t=1}^{T'_n - 1} R_t = L_{T'_n} = r$, the last inequality implies that

$$\begin{aligned} R_{T'_n} &\geq \log_2 n - \sqrt{\log_2 n} - \sum_{t=1}^{T'_n - 1} R_t \\ &\geq \log_2 n - \sqrt{\log_2 n} - r. \end{aligned}$$

Thus, since $R_{T'_n}$ is an integer, we have

$$\begin{aligned} R_{T'_n} &\geq \left\lceil \log_2 n - \sqrt{\log_2 n} \right\rceil - r \\ &= \ell' - r + 1. \end{aligned}$$

Given $\cap_{t=1}^{T'_n - 1} [G_t = R_t]$, we have

$$\sum_{t=1}^{T'_n - 1} G_t = \sum_{t=1}^{T'_n - 1} R_t.$$

Thus, event C implies that

$$\sum_{t=1}^{T'_n} G_t = G_{T'_n} + \sum_{t=1}^{T'_n-1} R_t = G_{T'_n} + r < \log_2 n - \sqrt{\log_2 n}.$$

In other words, we have

$$\begin{aligned} G_{T'_n} &\leq \log_2 n - \sqrt{\log_2 n} - 1 - r \\ &\leq \left\lceil \log_2 n - \sqrt{\log_2 n} \right\rceil - 1 - r \\ &= \ell' - r. \end{aligned}$$

Note that, when $r = \ell'$, the last inequality implies that $G_{T'_n} = 0$, which is impossible. Therefore, we have $P\{C \mid E(\ell')\} = 0$.

When $0 \leq r < \ell'$, there are two reasons that we may have $G_{T'_n} \leq \ell' - r$ but $R_{T'_n} > \ell' - r$. First, it might be because of $K_{T'_n} > k$, i.e., at step T'_n, two points among $(B_i)_{0 \leq i \leq k}$ fall in the same unit interval and thus $K_{T'_n} > k$. The probability that this happens is at most $1/M_{L_{T'_n}}$, as all the points are selected uniformly from the interval $[0, M_{L_{T'_n}}]$. Given B_r and $L_{T'_n} = r$, we have

$$M_{L_{T'_n}} = M_r \geq (1-\epsilon)\frac{n}{2^{r+1}}.$$

Therefore, we have

$$\begin{aligned} P\{K_t > k \mid E(r)\} &= P\left\{ \cup_{\substack{i,j \in [1,k] \\ i \neq j}} \left[\overline{B}_{t,i} = \overline{B}_{t,j} \right] \;\middle|\; E(r) \right\} \\ &\leq k^2 P\left\{ \left[\overline{B}_{t,1} = \overline{B}_{t,2} \right] \;\middle|\; E(r) \right\} \\ &\leq k^2 \frac{1}{M_r} \\ &\leq \frac{k^2 2^{r+1}}{n(1-\epsilon)} \\ &\leq \frac{k^2 2^{\ell'}}{n(1-\epsilon)}. \end{aligned}$$

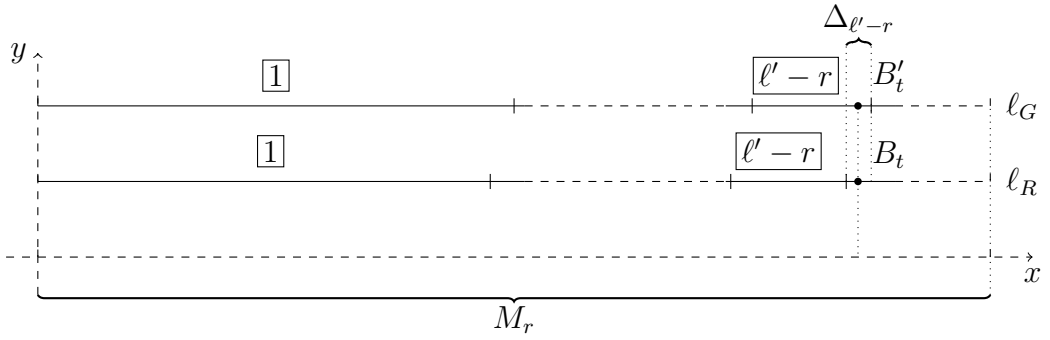

Figure 4–10: An example of the second possible reason that might cause event C. $\Delta_{\ell'-r}$ is the interval between the right endpoint of the $(\ell' - r)$-th segment of ℓ_R and that of the $(\ell' - r)$-th segment of ℓ_G. In this example, we have $B_t = B'_t \in \Delta_{\ell'-r}$, which implies that $G_{T'_n} \leq \ell' - r \leq$ and $R_{T'_n} > \ell' - r$.

Another reason why we may have $G_{T'_n} \leq \ell' - r$ but $R_{T'_n} > \ell' - r$ is that $B_{T'_n}$ falls into the interval between the right endpoint of the $(\ell' - r)$-th segment of ℓ_R and that of the $(\ell' - r)$-th segment of ℓ_G, as shown in Figure 4–10. We denote this interval by $\Delta_{\ell'-r}$. Given A_r, we can bound the width of $\Delta_{\ell'-r}$, denoted by $|\Delta_{\ell'-r}|$ as follows:

$$\begin{aligned}
|\Delta_{\ell'-r}| &= \left| \sum_{j=1}^{\ell'-r} \frac{M_r}{2^j} - \sum_{j=1}^{\ell'-r} N_{r+j} \right| \\
&\leq \sum_{j=1}^{\ell'-r} \left| \frac{M_r}{2^j} - N_{r+j} \right| \\
&\leq \sum_{j=1}^{\ell'-r} \frac{\epsilon M_r}{2^j} \\
&\leq \epsilon M_r.
\end{aligned}$$

Therefore, we have

$$\begin{aligned} P\left\{B'_t \in \Delta_{\ell'-r} \mid E(r)\right\} &= P\left\{\cup_{i\in[1,k]}\left[B_{t,i} \in \Delta_{\ell'-r}\right] \;\middle|\; E(r)\right\} \\ &\leq kP\left\{\left[\overline{B}_{t,1} \in \Delta_{\ell'-r}\right] \;\middle|\; E(r)\right\} \\ &= k\frac{\epsilon M_r}{M_r} \\ &= k\epsilon. \end{aligned}$$

In summary, for $r \in [0, \ell')$, we have

$$\begin{aligned} &P\left\{\sum_{t=1}^{T'_n} G_t < \log_2 n - \sqrt{\log_2 n} \;\middle|\; E(r)\right\} \\ &\leq P\left\{[B'_t \in \Delta_{\ell'-r}] \cup [K_t > k] \mid E(r)\right\} \\ &\leq P\left\{B'_t \in \Delta_{\ell'-r} \mid E(r)\right\} + P\left\{K_t > k \mid E(r)\right\} \\ &\leq \frac{k^2 2^{\ell'}}{n(1-\epsilon)} + k\epsilon. \end{aligned}$$

□

Lemma 4.5.7. *We have*

$$P\left\{\sum_{t=1}^{T'_n} G_t \geq \log_2 n - \sqrt{\log_2 n}\right\} \to 1,$$

as $n \to \infty$.

Proof. Let $r \in [0, \ell']$ be integer. Write the events

$$C = \left[\sum_{t=1}^{T'_n} G_t < \log_2 n - \sqrt{\log_2 n}\right],$$

$$D = [T'_n > 0] \cap \left[\cap_{t=1}^{T'_n - 1} R_t = G_t\right],$$

and

$$E(r) = [L_{T'_n} = r] \cap A_r \cap B_r \cap D.$$

For $r < \ell'$, it follows from Lemma 4.4.1, Lemma 4.5.2, and Lemma 4.5.6 that

$$\begin{aligned}
&P\left\{C \cap D \cap [L_{T'_n} = r]\right\} \\
&\quad \le P\left\{C \cap D \cap [L_{T'_n} = r] \cap A_r \cap B_r\right\} + P\left\{A_r^c \cap B_r\right\} + 2P\left\{B_r^c\right\} \\
&\quad \le P\left\{C \mid E(r)\right\} + P\left\{A_r^c \cap B_r\right\} + 2P\left\{B_r^c\right\} \\
&\quad \le \min_{\epsilon \in (0,1)} \left[\frac{k^2 2^{\ell'}}{n(1-\epsilon)} + k\epsilon + 2(\epsilon n + 1)\ell' \exp\left\{ -\frac{(1-\epsilon)n}{2^{\ell'+1}} \times f(\epsilon) \right\} \right. \\
&\qquad\qquad \left. + 4\exp\left\{ -\frac{f(\epsilon)n}{2^{\ell'+1}} \right\} \right].
\end{aligned}$$

For $r = \ell'$, it follows from Lemma 4.5.6 that

$$P\left\{C \cap D \cap [L_{T'_n} = r]\right\} = 0.$$

Recall that $L_{T'_n} \le \ell'$. Summing over all possible r and applying Lemma 4.5.1, we have

$$\begin{aligned}
&P\left\{C \cap D\right\} \\
&\quad = P\left\{ C \cap D \cap \left[\cup_{r=0}^{\ell'} [L_{T'_n} = r] \right] \right\} \\
&\quad = \sum_{r=0}^{\ell'} P\left\{C \cap D \cap [L_{T'_n} = r]\right\} \\
&\quad = \sum_{r=0}^{\ell'-1} P\left\{C \cap D \cap [L_{T'_n} = r]\right\} \\
&\quad \le \min_{\epsilon \in (0,1)} \left[\frac{k^2 \ell' 2^{\ell'}}{n(1-\epsilon)} + k\epsilon\ell' \right. \\
&\qquad + 2(\epsilon n + 1)(\ell')^2 \exp\left\{ -\frac{(1-\epsilon)n}{2^{\ell'+1}} \times f(\epsilon) \right\} \\
&\qquad \left. + 4\ell' \exp\left\{ -\frac{f(\epsilon)n}{2^{\ell'+1}} \right\} \right] \\
&\quad \stackrel{\texttt{def}}{=} \min_{\epsilon \in (0,1)} Z(\epsilon).
\end{aligned}$$

Now choose $\epsilon = 2^{-m}$ where $m = (\log_2 n)^{1/4}$. Recall that

$$\ell' = \left\lceil \log_2 n - \sqrt{\log_2 n} \right\rceil - 1 < \log_2 n - \sqrt{\log_2 n}.$$

Therefore, we have

$$\ell' < (m^4 - m^2) < m^4,$$

and also

$$\frac{2^{\ell'}}{n} < 2^{-\sqrt{\log_2 n}} = 2^{-m^2}.$$

Thus we have

$$Z(\epsilon) \leq O(I_1 + I_2 + I_3),$$

with

$$I_1 = \frac{\ell' 2^{\ell'}}{n}, \quad I_2 = \epsilon \ell', \quad I_3 = \epsilon n (\ell')^2 \exp\left\{ -\frac{(1-\epsilon)n}{2^{\ell'+1}} \times f(\epsilon) \right\}.$$

It is very easy to verify that I_1, I_2, I_3 all go to zero as n goes to infinity. For I_1 and I_2, we have

$$I_1 < m^4 2^{-m^2} = o(1),$$

and

$$I_2 < m^4 2^{-m} = o(1).$$

For I_3, we take a logarithm:

$$\begin{aligned} \log_2 I_3 &= \log_2 \epsilon + \log_2 n - 2 \log_2 \ell' - \frac{(1-\epsilon)n}{2^{\ell'+1}} \times f(\epsilon) \\ &< -m + m^4 - 8 \log_2 m - \frac{(1 - 2^{-m}) 2^{m^2 - 2m}}{4(1 + 2^{-m}/3)} \\ &\to -\infty, \end{aligned}$$

as $n \to \infty$. Therefore, we have $P\{C \cap D\} \to 0$. As Lemma 4.5.5 shows that $P\{D^c\} \to 0$, we have

$$P\{C\} \leq P\{C \cap D\} + P\{D^c\} \to 0,$$

as $n \to \infty$. □

Lemma 4.5.8. *For all constant $\zeta > 0$, We have*

$$P\left\{M_{L_{T'_n}} > (1+\zeta)2^{\sqrt{\log_2 n}}\right\} \to 0,$$

as $n \to \infty$.

Proof. M_{L_t+1} is the number of candidate leaves for the hop $z_{T'_n+1}$ on $\hat{\rho}$, that is the one right after step T'_n. Given that $T'_n > 0$, we have

$$L_{T'_n+1} \geq \log_2 n - \sqrt{\log_2 n}.$$

Recalling that

$$\ell' = \left\lceil \log_2 n - \sqrt{\log_2 n} \right\rceil - 1,$$

we have $\ell' < L_{T'_n+1}$. Since $(M_i)_{i\geq 0}$ is a decreasing sequence, it follows from Lemma 4.4.1 that

$$\begin{aligned}
&P\left\{M_{L_{T'_n+1}} > (1+\zeta)2^{\sqrt{\log_2 n}} \cap [T'_n > 0]\right\} \\
&\quad \leq P\left\{M_{\ell'} > (1+\zeta)2^{\sqrt{\log_2 n}}\right\} \\
&\quad \leq P\left\{M_{\ell'} > (1+\zeta)\frac{n}{2^{\ell'+1}}\right\} \\
&\quad \leq 2\exp\left\{-\frac{nf(\zeta)}{2^{\ell'+1}}\right\}.
\end{aligned}$$

This inequality together with Lemma 4.5.1 gives us

$$\begin{aligned}
&P\left\{M_{L_{T'_n+1}} > (1+\zeta)2^{\sqrt{\log_2 n}}\right\} \\
&\leq P\left\{M_{L_{T'_n+1}} > (1+\zeta)2^{\sqrt{\log_2 n}} \cap [T'_n > 0]\right\} + P\left\{T'_n = 0\right\} \\
&\leq 2\exp\left\{-\frac{nf(\zeta)}{2^{\ell'+1}}\right\} + 2\exp\left\{-\frac{nf(\zeta)}{2^{\ell'+1}}\right\} \\
&\leq 4\exp\left\{-2^{\sqrt{\log_2 n}} f(\zeta)\right\} \\
&\to 0,
\end{aligned}$$

as $n \to \infty$. □

The following lemma establish the relationship between T'_n and $\widehat{T}$.

Lemma 4.5.9. *We have*

$$\frac{\widehat{T} - T'_n}{\log n} \xrightarrow{p} 0,$$

as $n \to \infty$.

Proof. Note that we can think of the part of $\hat{\rho}$ after the T'_n-th hop as a routing process with $M_{L_{T'_n+1}}$ nodes as input, $z_{T'_n}$ as starting node, and $\widehat{T} - T'_n$ as the routing time of this process. Let $\zeta > 0$ be a constant. Write the event

$$M = \left[M_{L_{T'_n+1}} \leq (1+\zeta)2^{\sqrt{\log_2 n}}\right].$$

Given M, it follows from theorem 3.0.8 that, whatever IDs these $M_{L_{T'_n+1}}$ nodes have, whatever the starting node $z_{T'_n}$ is and the destination ID is, we always have

$$\mathbb{E}\left[\widehat{T} - T'_n \mid M\right] \leq (1+o(1))\frac{\log\left[(1+\zeta)2^{\sqrt{\log_2 n}}\right]}{H_k} = O(\sqrt{\log_2 n}).$$

It follows from Lemma 4.5.8 that, for all $\epsilon > 0$, we have

$$\begin{aligned} P\left\{\left|\frac{\widehat{T}-T_n'}{\log n}\right| \geq \epsilon\right\} &\leq P\left\{\left|\frac{\widehat{T}-T_n'}{\log n}\right| \geq \epsilon \;\middle|\; M\right\} + P\left\{M^c\right\} \\ &\leq \mathbb{E}\left[\left|\frac{\widehat{T}-T_n'}{\log n}\right| \;\middle|\; M\right] \Big/ \epsilon + o(1) \\ &= \frac{O(\sqrt{\log_2 n})}{\log(n)\epsilon} + o(1) \\ &\to 0, \end{aligned}$$

as $n \to \infty$. □

Theorem 4.5.10. *We have*

$$\frac{\widehat{T}}{\log n} \xrightarrow{p} \frac{1}{g(k)},$$

as $n \to \infty$, where

$$g(k) = H_k + o(1)$$

is a function of k.

Proof. We first extend $(G_t)_{t \leq \widehat{T}}$ to an infinite sequence of random variables. Let $(G_t)_{t > \widehat{T}}$ be i.i.d. random variables with the distribution

$$P\left\{G_t = i\right\} = (1 - 1/2^i)^k - (1 - 1/2^{i-1})^k, \quad t > \widehat{T}.$$

Therefore, we can define a random variable T_n by

$$T_n = \min\left\{t : \log_2 n \leq \sum_{i=1}^{t} G_i\right\}.$$

Write the events

$$E_1 = \left[\sum_{t=1}^{T_n'} G_t \geq \log_2 n - \sqrt{\log_2 n}\right],$$

and

$$E_2 = \left[\sum_{t=1}^{T'_n - 1} G_t = \sum_{t=1}^{T'_n - 1} R_t < \log_2 n - \sqrt{\log_2 n} \right] \cap \left[T'_n > 0 \right].$$

Given E_1 and E_2, we have

$$T_n - \sqrt{\log_2 n} \leq T'_n \leq T_n,$$

which implies that

$$\left| \frac{T_n - T'_n}{\log n} \right| = o(1).$$

Thus for all constant $\epsilon > 0$, we have

$$P\left\{ \left| \frac{T_n - T'_n}{\log n} \right| > \epsilon \;\middle|\; E_1 \cap E_2 \right\} \to 0,$$

as $n \to \infty$. Together with Lemma 4.5.7 and Lemma 4.5.5, we have for all constant $\epsilon > 0$

$$\begin{aligned} P\left\{ \left| \frac{T_n - T'_n}{\log n} \right| > \epsilon \right\} &\leq P\left\{ \left| \frac{T_n - T'_n}{\log n} \right| > \epsilon \;\middle|\; E_1 \cap E_2 \right\} + P\left\{E_1^c\right\} + P\left\{E_2^c\right\} \\ &\to 0, \end{aligned}$$

as $n \to \infty$. In other words, we have

$$\frac{T'_n}{\log n} \xrightarrow{p} \frac{T_n}{\log n}.$$

It follows from Corollary 4.3.3 that

$$\frac{T'_n}{\log n} \xrightarrow{p} \frac{1}{g(k)}.$$

Together with Lemma 4.5.9, we have

$$\frac{\widehat{T}}{\log n} \xrightarrow{p} \frac{1}{g(k)}. \qquad \square$$

Let $Y_1, Y_2, \ldots$ be random variables on some probability space, we say $Y_n \to Y$ in r-th mean [10, chap. 7], written $Y_n \xrightarrow{r} Y$, if $\mathbb{E}[Y_n^r] < \infty$ for all n and

$$\mathbb{E}[|Y_n - Y|^r] \to 0,$$

as $n \to \infty$. To prove our last result, we need the following lemma about convergence in probability and convergence in the mean [10, chap. 7.2].

Lemma 4.5.11. *If $Y_n \xrightarrow{p} Y$ and $P\{|Y_n| \le C\} = 1$ for all n and some C, then we have $Y_n \xrightarrow{r} Y$ for all $r \ge 1$.*

Theorem 4.5.12. *If $d/\log n \le C$ for some constant C, then we have*

$$\frac{\widehat{T}}{\log n} \xrightarrow{1} \frac{1}{g(k)},$$

as $n \to \infty$, where

$$g(k) = H_k + o(1)$$

is a function of k.

Proof. Since $\widehat{T} \le \log_2 n \le d$, we have

$$\frac{\widehat{T}}{\log n} \le \frac{d}{\log n} \le C.$$

Thus it follows from Theorem 4.5.10 and Lemma 4.5.11 that

$$\frac{\widehat{T}}{\log n} \xrightarrow{1} \frac{1}{g(k)}.$$

□

Recall that

$$\mathbb{E}[G_1] = \sum_{i=0}^{\infty} \left[1 - \left(1 - \frac{1}{2^i}\right)^k\right],$$

and $g(k) = \log(2)\mathbb{E}[G_1]$. Since (Lemma 4.3.2)

$$\frac{H_k}{\log 2} \le \mathbb{E}[G_1] \le \frac{H_k}{\log 2} + 1,$$

we have

$$H_k \leq g(k) \leq H_k + \log 2.$$

Because it is not difficult to verify that $\mathbb{E}[G_1] \sim \Phi(k)$ as $k \to \infty$, where

$$\Phi(k) = \sum_{i=0}^{\infty} \left[1 - e^{-i/2^k}\right],$$

we can approximate $g(k)$ with $\log(2)\Phi(k)$. Flajolet and Sedgewick [8, p. 311] showed that

$$\Phi(k) = \log k + \frac{\gamma}{\log 2} + \frac{1}{2} + P(k) + O(k^{-1}),$$

where γ is the Euler's constant and P is an oscillating function with tiny fluctuations of the order of 10^{-6}. Figure 4–11 shows the numeric values of the first 100 terms of $(H_k)_{k\geq 1}$, $(H_k + \log 2)_{k\geq 1}$ and $(\log(2)\Phi(k))_{k\geq 1}$.

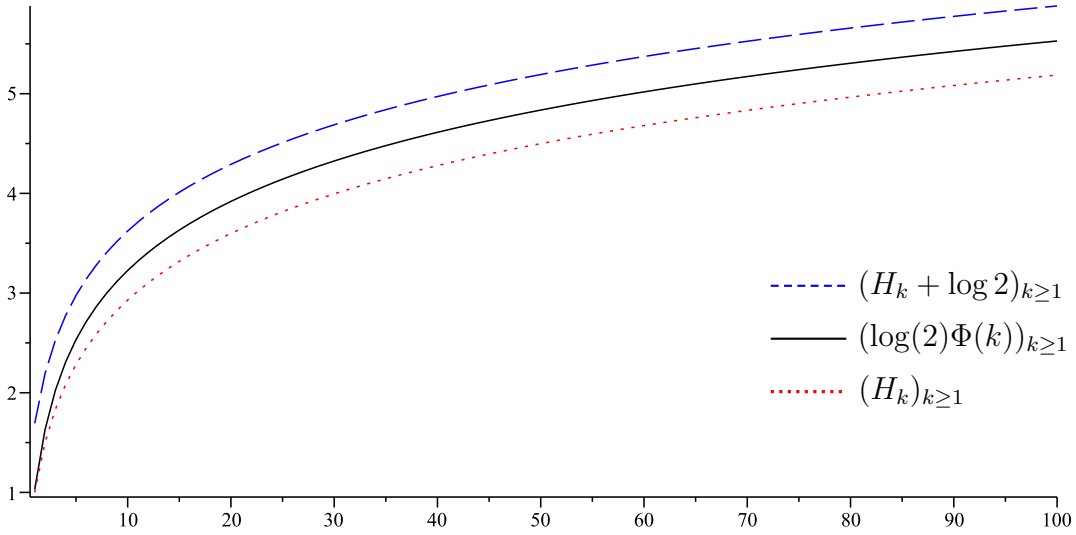

Figure 4–11: The first 100 terms of $(H_k)_{k\geq 1}$, $(H_k+\log 2)_{k\geq 1}$ and $(\log(2)\Phi(k))_{k\geq 1}$.

CHAPTER 5
Open Questions

Recall that T_{MAX}, the maximal routing time, is defined by

$$T_{\text{MAX}} = \sup_{x_1,\ldots,x_n} \sup_{y\in\{0,1\}^d} \mathbb{E}\left[T_{x_1}(y)\right],$$

and Theorem 3.0.8 shows that

$$T_{\text{MAX}} \leq (1+o(1))\frac{\log n}{H_k}.$$

We conjecture that this upper bound is tight in the case that $d/n \to \infty$.

Given $\mathcal{V}(\mathcal{K})$, whether it is deterministically or randomly decided, we can define random variables

$$\max_{y\in\{0,1\}^d} T_{X_1}(y), \qquad \max_{i\in[1,n]} T_{X_i}(\overline{1}), \qquad \max_{i\in[1,n]} \max_{y\in\{0,1\}^d} T_{X_i}(y),$$

and study their distributions and expectations.

In Kademlia's protocol [17], when a node tries to locate an ID, it is allowed to send querying messages at the same time to α neighbors, where α is a system-wide parameter. To approximate the routing process when this kind of parallel searching is permitted, we can define $\rho_x^\alpha(y) = (z_t)_{t\geq 1}$, where $(z_t)_{t\geq 1}$ is a sequence of vertex sets. Let $z_0 = \{x\}$. Given z_t and $t \geq 1$, we choose up to α vertices that are closest to y among all the neighbors of all the vertices in z_t. In the case $\alpha = 1$, we have $\rho_x^1(y) = \rho_x(y)$, which has been studied in this work. For $\alpha > 1$, we can ask similar questions like: what is the maximal routing time?

Given a vertex x and a positive integer i, let $\mathcal{N}_i(x)$ be the number of vertices that can reach x within i hops. When all the IDs are selected uniformly

at random from $\{0,1\}^d$ without replacement, what is the expectation of $\mathcal{N}_i(x)$? This question is of particularly interest to researchers studying Kademlia-based botnets.

References

[1] S. Androutsellis-Theotokis and D. Spinellis. A survey of peer-to-peer content distribution technologies. *ACM Computing Surveys*, 36(4):335 – 371, 2004. doi: 10.1145/1041680.1041681. URL http://doi.acm.org/10.1145/1041680.1041681.

[2] H. Balakrishnan, M. F. Kaashoek, D. Karger, R. Morris, and I. Stoica. Looking up data in P2P systems. *Communications of the ACM*, 46(2):43 – 48, 2003. doi: 10.1145/606272.606299. URL http://doi.acm.org/10.1145/606272.606299.

[3] B. Cohen. Incentives build robustness in BitTorrent. San Francisco, CA, USA, 2003.

[4] S. A. Crosby and D. S. Wallach. An analysis of BitTorrent's two Kademlia-based DHTs. Rice University, Houston, TX, USA, 2007.

[5] H. David and H. Nagaraja. *Order Statistics*. Wiley Series in Probability and Mathematical Statistics. Probability and Mathematical Statistics. John Wiley & Sons, Hoboken, NJ, USA, 2003. URL http://books.google.ca/books?id=bdhzFXg6xFkC.

[6] C. Davis, J. Fernandez, S. Neville, and J. McHugh. Sybil attacks as a mitigation strategy against the storm botnet. In *Proceedings of the 3rd International Conference on Malicious and Unwanted Software, Malware '08*, pages 32 – 40, 2008. doi: 10.1109/Malware.2008.4690855.

[7] J. Falkner, M. Piatek, J. P. John, A. Krishnamurthy, and T. Anderson. Profiling a million user DHT. In *Proceedings of the 7th ACM SIGCOMM*

conference on Internet measurement, IMC '07, pages 129 – 134, New York, NY, USA, 2007. ACM. doi: 10.1145/1298306.1298325. URL http://doi.acm.org/10.1145/1298306.1298325.

[8] P. Flajolet and R. Sedgewick. *Analytic Combinatorics*. Cambridge University Press, Cambridge, UK, 2009. URL http://books.google.ca/books?id=0h-4QcA1c1QC.

[9] E. Fredkin. Trie memory. *Communications of the ACM*, 3(9):490 – 499, 1960. doi: 10.1145/367390.367400. URL http://doi.acm.org/10.1145/367390.367400.

[10] G. Grimmett and D. Stirzaker. *Probability and Random Processes*. Oxford University Press, New York, NY, USA, 2001. URL http://books.google.ca/books?id=G3ig-0M4wSIC.

[11] J. B. Grizzard, V. Sharma, C. Nunnery, B. B. Kang, and D. Dagon. Peer-to-peer botnets: overview and case study. In *Proceedings of the 1st Conference on 1st Workshop on Hot Topics in Understanding Botnets*, HotBots '07, pages 1 – 1, Berkeley, CA, USA, 2007. USENIX Association. URL http://dl.acm.org/citation.cfm?id=1323128.1323129.

[12] W. Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58(301):13 – 30, 1963. URL http://www.jstor.org/stable/2282952.

[13] N. Johnson, S. Kotz, and N. Balakrishnan. *Continuous Univariate Distributions*, volume 2 of *Wiley Series in Probability and Mathematical Statistics. Applied Probability and Statistics*. John Wiley & Sons, Hoboken, NJ, USA, 1995. URL http://books.google.ca/books?id=0QzvAAAAMAAJ.

[14] N. Johnson, A. Kemp, and S. Kotz. *Univariate Discrete Distributions*. Wiley Series in Probability and Statistics. John Wiley & Sons, Hoboken,

NJ, USA, 2005. URL http://books.google.ca/books?id=7S6Qhs0ZxWoC.

[15] Y. Liu, Y. Guo, and C. Liang. A survey on peer-to-peer video streaming systems. *Peer-to-Peer Networking and Applications*, 1:18 – 28, 2008. URL http://dx.doi.org/10.1007/s12083-007-0006-y. 10.1007/s12083-007-0006-y.

[16] E. K. Lua, J. Crowcroft, M. Pias, R. Sharma, and S. Lim. A survey and comparison of peer-to-peer overlay network schemes. *Communications Surveys Tutorials, IEEE*, 7(2):72 – 93, 2005. doi: 10.1109/COMST.2005.1610546.

[17] P. Maymounkov and D. Mazières. Kademlia: A peer-to-peer information system based on the XOR metric. In *Peer-to-Peer Systems*, volume 2429 of *Lecture Notes in Computer Science*, pages 53 – 65. Springer, Berlin / Heidelberg, Germany, 2002.

[18] J. Menn. *All the Rave: the Rise and Fall of Shawn Fanning's Napster*. Crown Business, New York, NY, USA, 2003. URL http://books.google.ca/books?id=vz4UAQAAIAAJ.

[19] D. Milojicic, V. Kalogeraki, R. Lukose, K. Nagaraja, J. Pruyne, B. Richard, S. Rollins, and Z. Xu. Peer-to-peer computing. Technical Report HPL-2002-57R1, HP Laboratories, Palo Alto, CA, USA, 2003.

[20] P. V. Mockapetris. RFC 882: Domain names: concepts and facilities, 1983. URL http://tools.ietf.org/html/rfc882.

[21] P. V. Mockapetris. RFC 883: Domain names: implementation specification, 1983. URL http://tools.ietf.org/html/rfc883.

[22] A. Oram. *Peer-to-Peer: Harnessing the Benefits of a Disruptive Technology*. O'Reilly Series. O'Reilly, Sebastopol, CA, USA, 2001. URL http://books.google.ca/books?id=Mb8kQf0SNv0C.

[23] S. Ratnasamy, P. Francis, M. Handley, R. Karp, and S. Shenker. A scalable content-addressable network. *SIGCOMM Computer Communication Review*, 31(4):161 – 172, 2001. doi: 10.1145/964723.383072. URL http://doi.acm.org/10.1145/964723.383072.

[24] RFC-Gnutella. The annotated gnutella protocol specification v0.4. URL http://rfc-gnutella.sourceforge.net/developer/stable/.

[25] J. Risson and T. Moors. Survey of research towards robust peer-to-peer networks: Search methods. *Computer Networks*, 50(17):3485 – 3521, 2006. doi: 10.1016/j.comnet.2006.02.001.

[26] A. Rowstron and P. Druschel. Pastry: Scalable, decentralized object location, and routing for large-scale peer-to-peer systems. In *Middleware 2001*, volume 2218 of *Lecture Notes in Computer Science*, pages 329 – 350. Springer, Berlin / Heidelberg, Germany, 2001. URL http://dx.doi.org/10.1007/3-540-45518-3_18.

[27] R. Schollmeier. A definition of peer-to-peer networking for the classification of peer-to-peer architectures and applications. In *Proceedings of 1st International Conference on Peer-to-Peer Computing*, pages 101 – 102, 2001. doi: 10.1109/P2P.2001.990434.

[28] M. Shaked and J. Shanthikumar. *Stochastic Orders*. Springer Series in Statistics. Springer, New York, NY, USA, 2007. URL http://books.google.ca/books?id=rPiToBK2rwwC.

[29] M. Steiner, T. En-Najjary, and E. W. Biersack. A global view of Kad. In *Proceedings of the 7th ACM SIGCOMM Conference on Internet Measurement*, IMC '07, pages 117 – 122, New York, NY, USA, 2007. ACM. doi: 10.1145/1298306.1298323. URL http://doi.acm.org/10.1145/1298306.1298323.

[30] M. Steiner, T. En-Najjary, and E. W. Biersack. Exploiting Kad: possible uses and misuses. *SIGCOMM Computer Communication Review*, 37(5):65 – 70, 2007. doi: 10.1145/1290168.1290176. URL http://doi.acm.org/10.1145/1290168.1290176.

[31] R. Steinmetz and K. Wehrle. *Peer-to-Peer Systems And Applications*. Lecture Notes in Computer Science. Springer, Berlin / Heidelberg, Germany, 2005. URL http://books.google.ca/books?id=A8CLZ1FB4qoC.

[32] I. Stoica, R. Morris, D. Karger, M. F. Kaashoek, and H. Balakrishnan. Chord: A scalable peer-to-peer lookup service for internet applications. In *Proceedings of the 2001 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications*, SIGCOMM '01, pages 149 – 160, New York, NY, USA, 2001. ACM. doi: 10.1145/383059.383071. URL http://doi.acm.org/10.1145/383059.383071.

[33] W. Szpankowski. *Average Case Analysis of Algorithms on Sequences*. Wiley Series in Discrete Mathematics and Optimization. John Wiley & Sons, Hoboken, NJ, USA, 2011. URL http://books.google.ca/books?id=WfQg-1nsDCkC.

[34] B. Y. Zhao, L. Huang, J. Stribling, S. C. Rhea, A. D. Joseph, and J. D. Kubiatowicz. Tapestry: A resilient global-scale overlay for service deployment. *IEEE Journal on Selected Areas in Communications*, 22:41 – 53, 2004.