university of groningen

faculty of science and engineering

mathematics and applied mathematics

# Collisions of Independent Random Walks in Infinite Graphs

Bachelor's Project Mathematics

July 2020

Student: W. Koops

First supervisor: prof. dr. D. Rodrigues Valesin

Second assessor: dr. C.P. Hirsch

## Abstract

A recurrent graph has the (in)finite collision property if two independent random walks started from the same point collide (in)finitely often almost surely. Krishnapur and Peres [7] show that the graph $\mathrm{Comb}(\mathbb{Z})$, which is obtained from $\mathbb{Z}^2$ by removing all horizontal edges not on the $x$-axis, is a recurrent graph with the finite collision property. Barlow, Peres and Sousi [3] further study the collision properties of power-law combs, subgraphs of $\mathrm{Comb}(\mathbb{Z})$ where all vertices $(x, y)$ that do not satisfy $0 \leq y \leq f(x)$ and the corresponding edges are removed. In this thesis, these results are explained in detail. Finally, the case where the heights $f(n)$ of the comb graph are i.i.d. random variables with law $\mu$ is considered. In particular a condition on $\mu$ is given, which implies that the resulting comb graph has the finite collision property.

## Acknowledgments

I would like to thank my first supervisor, prof. dr. Daniel Valesin, for proposing this exciting topic and for guiding and supervising me throughout the whole process. I thank him for the pleasant collaboration and his patience and his encouragement and confidence in me.

I would like to thank dr. Hirsch for agreeing to be the second assessor for this project, for his enthusiasm and encouragement and for his comments on the first version of this thesis.

# Contents

**1 Introduction** **5**

**2 Problem definition** **7**

**3 Preliminary results** **8**

3.1 Results on random walks . . . 8

3.2 Electric network theory . . . 20

3.3 Results from probability theory . . . 29

3.4 Results from martingale theory . . . 34

**4 Exposition of known results** **37**

4.1 Results on collisions of random walks . . . 37

4.2 The finite collision property of Comb($\mathbb{Z}$) . . . 41

4.3 The Green function criterion and applications . . . 47

4.4 The finite collision property of Comb($\mathbb{Z}, \alpha$) for $\alpha > 1$ . . . 51

**5 Comb graphs with random heights** **60**

**6 Conclusion and discussion** **70**

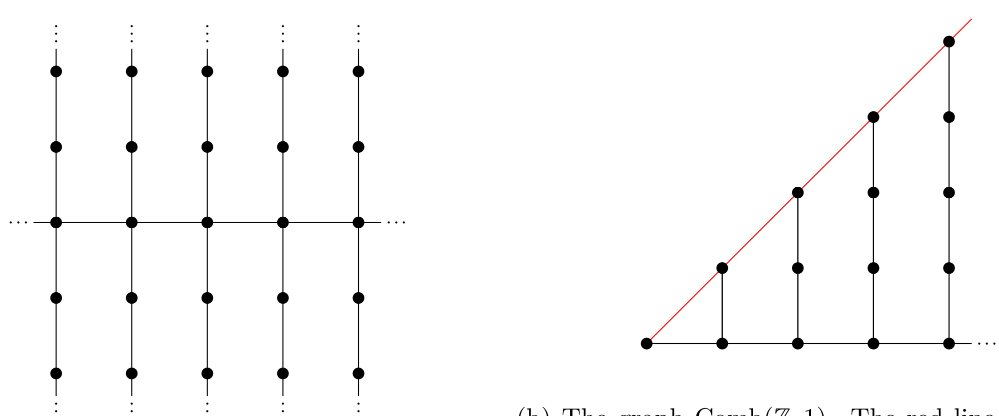
# 1 Introduction

In this thesis, collisions of independent random walks on infinite graphs are studied. A graph consists of a vertex set $V$ and an edge set $E$. Each edge connects two distinct vertices $x, y \in V$. Two vertices are called neighbours if they are connected by an edge. A graph is locally finite if every vertex has only finitely many neighbours. An important example of a graph is the $d$-dimensional grid, denoted by $\mathbb{Z}^d$, whose vertex set is the set of all points in $d$ dimensions with integer coordinates and where two vertices are neighbours if and only if they have distance 1.

A random walk on a locally finite graph $G$ is a sequence of random variables $(X_n)_{n\geq 0}$ taking values in $V$ such that for all $n \in \mathbb{N}$, $X_{n+1}$ is a neighbour of $X_n$ chosen uniformly at random. Given $X_n$, the distribution of $X_{n+1}$ does then not depend on $X_0, \ldots, X_{n-1}$ or the value of $n$. For more details, see Woess [13] or Lyons and Peres [10].

A random walk on a graph is called recurrent if it visits its starting position infinitely often. Instead of considering just one random walk, one can also consider two random walks and ask whether they meet infinitely often. The question arises whether there exists a graph in which any simple random walk is recurrent, but where two independent random walks collide only finitely often. This was first shown to be possible by Krishnapur and Peres [7] using the graph $\mathrm{Comb}(\mathbb{Z})$, which is obtained from $\mathbb{Z}^2$ by removing all horizontal edges not on the $x$-axis.

If two independent random walks collide (in)finitely often almost surely, we say that the graph has the (in)finite collision property. Barlow, Peres and Sousi [3] study further graphs with the (in)finite collision property, in particular the graph $\mathrm{Comb}(\mathbb{Z}, f)$ for some function $f$, which is defined as the subgraph of $\mathrm{Comb}(\mathbb{Z})$ by only including vertices $(x, y) \in \mathbb{Z}^2$ satisfying $0 \leq y \leq f(x)$ and the edges between these vertices. Of special interest is the case $f(x) = x^\alpha$ for $x > 0$, whose graph will be denoted by $\mathrm{Comb}(\mathbb{Z}, \alpha)$. Barlow, Peres and Sousi [3] show that in this case the graph has the infinite collision property if $\alpha \leq 1$ and the finite collision property if $\alpha > 1$.



(a) The graph $\mathrm{Comb}(\mathbb{Z})$

(b) The graph $\mathrm{Comb}(\mathbb{Z}, 1)$. The red line shows the bounding function $f(x) = x$. Note that there are no vertices with $x < 0$, since $f(x) < 0$ for $x < 0$.

Figure 1: Visual representations of the graphs $\mathrm{Comb}(\mathbb{Z})$ and $\mathrm{Comb}(\mathbb{Z}, 1)$.

This topic is mostly of theoretical interest, but Bertacchi, Lanchier and Zucca [4] give an application of the (in)finite collision property (of the infinite percolation cluster) to voter processes. While this specific class of graphs is not studied in this thesis, it still shows that this topic does have applications. Furthermore, random walks and stochastic processes in general, have a wide range of applications in statistical physics, economics, finance and population dynamics.

This thesis has two goals. The first goal is to explain the proofs in Krishnapur and Peres [7] and Barlow, Peres and Sousi [3] related to subgraphs of $\mathrm{Comb}(\mathbb{Z})$, and to prove the lemmas used in this proof as well. The second goal is to investigate comb graphs with random heights. Let $\{f(n)\}_{n\in\mathbb{Z}}$ be i.i.d. random variables with cumulative distribution function $F_X$ supported on $[1,\infty)$. For which $F_X$ does $\mathrm{Comb}(\mathbb{Z}, f)$ have the finite collision property? It is known that the height must have infinite expectation. In this thesis, an example of such cumulative distribution function $F_X$ is provided. To the knowledge of the author, this is the first such example. This thesis therefore provides an answer to Question 6 from Section 6 in Barlow, Peres and Sousi [3], which was first raised in Chen, Wei and Zhang [5].

In Chapter 2 of this thesis, the problem definition section, the notions introduced in the introduction are defined more formally and some additional graph theoretic notions are defined.

In Chapter 3, the necessary background theory to understand the papers by Krishnapur and Peres [7] and Barlow, Peres and Sousi [3], as well as the lemmas necessary to completely verify the proofs, is given. Some proofs of standard facts are omitted, but they can be found easily in the literature. In case of more specific lemmas, a detailed proof is always provided. In this chapter, knowledge of probability and analysis at bachelor level is assumed. An introduction to measure theory is useful, but most of the material does not depend on it.

In Chapter 4, an exposition of the known results on the (in)finite collision property of subgraphs of $\mathrm{Comb}(\mathbb{Z})$ is given. The first section shows some small results on the (in)finite collision property. The second section exposits the proof by Krishnapur and Peres [7] that $\mathrm{Comb}(\mathbb{Z})$ has the finite collision property. This exposition follows the original proof closely. The third section gives the criterion for the infinite collision property from Barlow, Peres and Sousi [3] and uses this criterion to show that $\mathrm{Comb}(\mathbb{Z}, \alpha)$ has the infinite collision property if $\alpha \le 1$. The final section shows the converse, namely that $\mathrm{Comb}(\mathbb{Z}, \alpha)$ has the finite collision property if $\alpha > 1$. Here the proof deviates in some cases from the original to make it more accessible.

In Chapter 5, the graph $\mathrm{Comb}(\mathbb{Z}, f)$ is considered for the case where the heights $f(n)$ itself are i.i.d. random variables. In particular, the following theorem is proven:

**Theorem.** Let $F_X : \mathbb{N} \to [0,1]$ be a cumulative distribution function such that

- $F_X(n) \le 1 - n^{-1/\alpha}$ for all $n \in \mathbb{N}$, for some constant $\alpha > 3$, and
- $F_X(n) \ge 1 - n^{-E}$ for all $n \in \mathbb{N}$, for some constant $E > 0$.

Let $\{f(n)\}_{n\in\mathbb{Z}}$ be i.i.d. random variables with cdf $F_X$.
Then $G = \mathrm{Comb}(\mathbb{Z}, f)$ has the finite collision property almost surely.

## 2 Problem definition

In this section, some of the topics introduced in the introduction are defined more formally. A graph $G$ is a pair $(V, E)$, where $V$ is a set and $E$ is a set of sets of the form $\{x, y\}$, where $x$ and $y$ are distinct elements of $V$. The sets $V$ and $E$ are called the vertex set and edge set of $G$, respectively. Two vertices $x, y \in V$ are said to be neighbours if $\{x, y\} \in E$. The number of neighbours of a vertex $v \in V$ is called the degree of $v$ and denoted by $\deg(v)$. A graph $G$ is locally finite if the degree of every vertex is finite. A path is a finite sequence of vertices $\mathcal{P} = (p_0, p_1, \ldots, p_n)$ such that $\{p_i, p_{i+1}\} \in E$ for all $0 \leq i \leq n-1$. A graph is called connected if for any two vertices $x, y \in V$ there exists a path $\mathcal{P}$ with $p_0 = x$ and $p_n = y$. Finally, a graph is called a tree if for any two vertices $x, y \in V$ there exists exactly one path $\mathcal{P}$ with $p_0 = x$ and $p_n = y$ such that all vertices in the path are different.

The main goal of the thesis is to determine whether certain classes of subgraphs of $\text{Comb}(\mathbb{Z})$ have the finite collision property or the infinite collision property. We now first define these graphs formally. $\text{Comb}(\mathbb{Z})$ is the graph with vertex set $\mathbb{Z}^2 = \{(x, y) : x, y \in \mathbb{Z}\}$ and edge set

$$\{\{(x,0),(x+1,0)\} : x \in \mathbb{Z}\} \cup \{\{(x,y),(x,y+1)\} : x, y \in \mathbb{Z}\}.$$

$\text{Comb}(\mathbb{Z}, f)$ is defined as the subgraph of $\text{Comb}(\mathbb{Z})$ by only including vertices $(x, y) \in \mathbb{Z}^2$ that satisfy $0 \leq y \leq f(x)$ and the edges between these vertices. It follows that $\text{Comb}(\mathbb{Z}, f)$ is the graph with vertex set $\{(x, y) : x, y \in \mathbb{Z}, 0 \leq y \leq f(x)\}$ and edge set

$$\begin{aligned} &\{\{(x,0),(x+1,0)\} : x \in \mathbb{Z}, f(x) \geq 0, f(x+1) \geq 0\} \\ &\cup \{\{(x,y),(x,y+1)\} : x, y \in \mathbb{Z},\ f(x) \geq y+1,\ y \geq 0\}. \end{aligned}$$

Note that $f$ should be nonnegative on a connected subset of $\mathbb{Z}$, otherwise $\text{Comb}(\mathbb{Z}, f)$ is not connected. Note that all connected subsets of $\mathbb{Z}$ are intervals. Of special interest is the case where $f(x) = x^\alpha$ for $x \geq 0$ and $f(x) < 0$ for $x < 0$. The latter condition implies that there are no vertices with $x < 0$. This graph is denoted by $\text{Comb}(\mathbb{Z}, \alpha)$.

#### Conventions and notation

In this thesis, the following conventions are used:

- $\mathbb{N} = \{0, 1, 2, \ldots\}$.
- $\text{Geo}\,(p)$: the geometric distribution with success probability $p$, counting the number of failures, i.e. if $G \sim \text{Geo}(p)$, then $\mathbb{P}(G_i = k) = (1-p)^k p$ for $k \geq 0$.
- $x \wedge y$: the minimum of $x$ and $y$, i.e. $\min\{x, y\}$.

# 3 Preliminary results

## 3.1 Results on random walks

This section presents a number of properties of random walks, based on the lecture notes by Verbitskiy and Valesin [12]. We first give some fundamental definitions from random walk theory and Markov chain theory. After that, some results that are used in later proofs are given. In this thesis, only discrete-time random walks are considered.

Let $S$ a countable set, the state space. A **transition function** on $S$ is a function $p : S \times S \to \mathbb{R}_{\geq 0}$ such that $\sum_{y \in S} p(x, y) = 1$ for all $x \in S$. In this setting, we interpret $p(x, y)$ as the probability of transiting to state $y$, given that the current state is $x$.
If $p$ is a one-step transition function, we might also be interested in an $n$-step transition function, which gives the probability of going from state $x$ to $y$ in $n$ steps. To do this, let us first define the product of two transition functions. For two transition functions $p$ and $q$, we can make a new transition function $p \circ q$, which gives the transition probabilities when first making a step according to $p$ and then a step according to $q$. By conditioning on the state $z$ reached after making a step according to $p$ and using the law of total probability, we see that defining

$$(p \circ q)(x, y) = \sum_{z \in S} p(x, z) q(z, y)$$

gives what we want. Of course, it needs to be proven that $p \circ q$ actually is a transition function according to the definition just given. To do this, we will interpret the transition function $p$ as a matrix $P = [p(i, j)]_{i,j \in S}$. Such a matrix is called a **stochastic matrix**, and in this section we will always denote the matrix corresponding to the transition function $p$ by $P$. If $P$ and $Q$ represent $p$ and $q$ respectively, then $PQ$ represents $p \circ q$. Note that the condition $\sum_{y \in S} p(x, y) = 1$ for all $x \in S$ is equivalent to $P \cdot \mathbf{1} = \mathbf{1}$, where $\mathbf{1} \in \mathbb{R}^S$ is the vector containing ones only. If $P$ and $Q$ are stochastic matrices, it hence follows that $(P \cdot Q)\mathbf{1} = P(Q\mathbf{1}) = P\mathbf{1} = \mathbf{1}$, so then $PQ$ is also a stochastic matrix and hence $p \circ q$ is a transition function.
Now that we have rigorously discussed the product of two transition functions, it is easy to define the $n$-step transition function $p^{(n)}$ recursively by $p^{(1)} = p$ and $p^{(n+1)} = p^{(n)} \circ p$.

Consider a measure $\mu$ on the power set of $S$. Since $S$ is countable, $\mu$ is completely determined by the values on the singleton sets. We write $\mu(x) = \mu(\{x\})$. The measure $\mu$ can be interpreted as a distribution, describing how likely it is to be in a given state. For a given measure $\mu$, we may be interested in the distribution among the states after doing one step according to the transition function $p$. This is a function of the initial state $y$, and it is given by

$$(\mu P)(y) = \sum_{x \in S} \mu(x) p(x, y).$$

A measure $\mu$ is called a **stationary measure** if $\mu P = \mu$.

Given a function $f : S \to \mathbb{R}$, we may be interested in the average value of $f$ after doing one step according to the transition function $p$. This is a function of the initial state $x$, and it is given by

$$(Pf)(x) = \sum_{y \in S} p(x, y) f(y).$$

The function $f$ is called **harmonic** if $f = Pf$.
Note that, in terms of matrices, we can interpret $f$ as a column vector and $\mu$ as a row vector. Then the notations $Pf$ and $\mu P$ indeed correspond to the definition of matrix multiplication.

A stochastic process is a sequence of random variables $(X_n)_{n \geq 0}$ defined on the same probability space. A Markov chain is a stochastic process where a transition function is used to go from the state $X_n$ to a new state, $X_{n+1}$. Given $X_n$, the distribution of $X_{n+1}$ does then not depend on $X_0, \ldots, X_{n-1}$ or the value of $n$. The formal definition is as follows:

**Definition 3.1.** (**Markov chain**)
A Markov chain with state space $S$, transition function $p$ and initial distribution $\mu$ is a sequence of random variables $(X_n)_{n \geq 0}$ defined in some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ satisfying $X_0 \sim \mu$ and

$$\mathbb{P}(X_{n+1} = x_{n+1} \mid X_0 = x_0, \ldots, X_n = x_n) = p(x_n, x_{n+1})$$

for all integers $n \geq 0$ and all $x_0, \ldots, x_{n+1} \in S$.

It is common to indicate initial distribution on the probability measure $\mathbb{P}$ and the corresponding expectation operator $\mathbb{E}$. We write $\mathbb{P}_\mu$ to indicate that we consider a Markov chain starting from $X_0 \sim \mu$. We write $\mathbb{P}_x$ to indicate that we consider a Markov chain starting from $X_0 = x$. When we have two Markov chains starting from $X_0 = x$ and $Y_0 = y$, we write $\mathbb{P}_{x,y}$.

In this thesis, we will study a specific type of Markov chains, namely a simple random walk on a graph $G = (V, E)$. For two vertices $x, y \in V$, we write $x \sim y$ if $x$ and $y$ are connected. The number of neighbours of a vertex $x \in V$ is called the degree of $x$; formally $\deg(x) = \#\{y : x \sim y\}$. We assume that $G$ is locally finite, which means that the degree of every vertex is finite. A simple random walk is a Markov chain where in every step we choose a neighbour of $x$ uniformly at random. Formally, we have the following definition:

**Definition 3.2.** (**Simple random walk, SRW**)
A simple random walk on $G = (V, E)$ is a Markov chain with state space $V$ and transition function

$$p(x, y) = \frac{1}{\deg(x)} \cdot \mathbb{1}_{\{x \sim y\}}, \quad x, y \in V.$$

Note that there is no condition on the initial distribution for a simple random walk. We will usually consider random walks starting from a fixed vertex.

### 3.1.1 Irreducibility, recurrence and transience

A Markov chain $(X_n)$ is called **irreducible** if for any two states $x, y \in S$ there exists an $n \in \mathbb{N}$ such that $p^{(n)}(x, y) > 0$. This means that any two states can be reached from each other with positive probability. A graph is called connected if there is a path of edges connecting any two vertices. The simple random walk on a connected graph is irreducible.

Let $\tau_x^+ = \inf\{n \geq 1 : X_n = x\}$ be the first return time. Note that this is a random variable. A state $x \in S$ is called **recurrent** if $\mathbb{P}_x(\tau_x^+ < \infty) = 1$, that is: when starting in state $x$, we return to state $x$ almost surely. If a state is not recurrent, it is **transient**.

Information about the recurrence and transience of a state (and, as we will see later, about the collisions of random walks), is given by the Green function, which is the expected number of visits to a state $y$ from a Markov chain starting in state $x$:

**Definition 3.3.** (**Green function**)
The Green function $G : S \times S \to [0, \infty]$ of a Markov chain is given by

$$G(x, y) = \mathbb{E}_x[\#\{n \geq 0 : X_n = y\}], \quad x, y \in S.$$

Note that we can rewrite $G(x, y)$ as follows:

$$G(x, y) = \mathbb{E}_x[\#\{n \geq 0 : X_n = y\}] = \mathbb{E}_x\left[\sum_{n=0}^{\infty} \mathbb{1}_{\{X_n = y\}}\right] = \sum_{n=0}^{\infty} \mathbb{P}_x(X_n = y).$$

Assume that there exists a recurrent state $x$ in an irreducible Markov chain. Then $x$ is visited infinitely often almost surely. By irreducibility it follows that for any other state $y \in S$, there exists an $n \in \mathbb{N}$ such that $p^{(n)}(x, y) > 0$, so we also visit $y$ infinitely often $n$ states after visiting $x$ almost surely. In particular, $y$ is also recurrent. It follows that in an irreducible chain either all states are recurrent or all states are transient. If all states of an irreducible Markov chain are recurrent, we call the chain **recurrent**, otherwise we call the chain **transient**. The following proposition links recurrence and transience to the Green function:

**Proposition 3.1.** In an irreducible Markov chain, either all states are recurrent or all states are transient. In the recurrent case the Green function satisfies $G(x, y) = \infty$ for all $x, y$. In the transient case, $G(x, y) < \infty$ for all $x, y$.

A fundamental theorem in random walk theory by Pólya deals with the question for what dimension $d$ the simple random walk on $\mathbb{Z}^d$ is recurrent. The theorem can be proven using electric network theory. The cases $d \in \{1, 2\}$ can also be shown more directly by calculating the transition probabilities $p^{(2n)}(0, 0)$. For the latter approach, see Woess [13], Chapter 1A.

**Theorem 3.1.** (**Pólya's theorem**)
The simple random walk on $\mathbb{Z}^d$ is recurrent if $d = 1$ or $d = 2$ and transient if $d \geq 3$.

### 3.1.2 The Markov property and the strong Markov property

Given a sequence of random variables $(X_n)_{n\geq 0}$, the natural filtration $(\mathcal{F}_n)_{n\geq 0}$ is the sequence of $\sigma$-algebras such that $\mathcal{F}_n$ is generated by the random variables $X_0, \ldots, X_n$. We can write

$$\mathcal{F}_n = \{\{\omega \in \Omega : (X_0(\omega), \ldots, X_n(\omega)) \in B\} : B \subseteq S^{n+1}\}, \qquad n \geq 0.$$

Note that $\Omega$ could be a general set here, but one possibility is letting $\Omega = S^{\mathbb{N}}$, so then $\omega$ is a sequence of states and the natural interpretation of the random variables is then $X_i(\omega) = \omega_i$. The $\sigma$-algebra $\mathcal{F}_n$ is the collection of measurable sets, i.e. the collection of events, that we can assign a probability to. In this case, $\mathcal{F}_n$ is exactly the collection of events that can be defined using the information up to time $n$, i.e. $(X_0, \ldots, X_n)$.

In a Markov chain, all information about the path from time $n$ onward known at time $n$ is already contained in $X_n$, so also knowing $X_0, \ldots, X_{n-1}$ does not help to predict the future of the path. If we have a function that only depends on the path from time $n$ onward, we can therefore discard the information from before time $n$. This is the essence of the Markov property:

**Proposition 3.2.** (**Markov property**)
Let $(X_n)$ be a Markov chain with state space $S$ and $X_0 \sim \mu$. Let $(\mathcal{F}_n)_{n\geq 0}$ be the corresponding natural filtration. Let $f : S^{\mathbb{N}} \to \mathbb{R}$ be a measurable and bounded function. Then, for any $n$:

$$\mathbb{E}_\mu[f(X_n, X_{n+1}, \ldots) \mid \mathcal{F}_n] = \mathbb{E}_{X_n}[f(X_0, X_1, \ldots)] \qquad \mathbb{P}_\mu\text{-almost surely.}$$

Note that on the left hand side the random walk starts from a random state drawn from the distribution $\mu$. On the right hand side, we start from a random state which is distributed in the same way as $X_n$ is distributed when starting from $X_0 \sim \mu$.

The Markov property has a generalization, where $n$ does not need to be deterministic, but it must be a stopping time. A random variable $\tau : \Omega \to \mathbb{N} \cup \{\infty\}$ is called a **stopping time** for the chain if $\{\tau \leq n\} \in \mathcal{F}_n$ for all $n \in \mathbb{N}$. This roughly means that a decision to stop at time $n$, i.e. to have $\tau = n$, can only depend on the information that we have at time $n$. An important example of the stopping time is the first return time $\tau_x^+ = \inf\{n \geq 1 : X_n = x\}$. The corresponding Markov property for stopping times is called the strong Markov property:

**Proposition 3.3.** (**Strong Markov property**)
Let $(X_n)$ be a Markov chain with state space $S$ and $X_0 \sim \mu$. Let $(\mathcal{F}_n)_{n\geq 0}$ be the corresponding natural filtration. Let $f : S^{\mathbb{N}} \to \mathbb{R}$ be a measurable and bounded function. Let $\tau$ be a stopping time. Then we have:

$$\mathbb{E}_\mu[f(X_\tau, X_{\tau+1}, \ldots) \mid \mathcal{F}_\tau] \cdot \mathbb{1}_{\{\tau<\infty\}} = \mathbb{E}_{X_\tau}[f(X_0, X_1, \ldots)] \cdot \mathbb{1}_{\{\tau<\infty\}} \qquad \mathbb{P}_\mu\text{-almost surely.}$$

Note that the property only makes sense if $\tau < \infty$. Multiplying by $\mathbb{1}_{\{\tau<\infty\}}$ ensures this.
For the proof of the Markov property and the strong Markov property we refer to the lecture notes by Verbitskiy and Valesin [12].

### 3.1.3 Inequalities on transition probabilities of random walks

In this subsection, we present two inequalities on transition properties of random walks. The content of this subsection is based on inequalities used in Krishnapur and Peres [7] and are hence directly related to our investigations.

Recall that a measure $\mu$ is called a stationary measure if $\mu P = \mu$. If $P$ represents a simple random walk on a locally finite graph $G = (V, E)$, then a stationary measure is given by $\mu(v) = c \deg(v)$ for some positive constant $c$. We then have

$$(\mu P)(y) = \sum_{x\sim y} \frac{1}{\deg x}\mu(x) = \sum_{x\sim y} \frac{c \deg x}{\deg x} = c\#\{x : x \sim y\} = c \deg(y) = \mu(y),$$

so $\mu P = \mu$, and hence this is indeed a stationary measure. We shall take $c = 1$ in what follows. It can be proven that in an irreducible and recurrent Markov chain, a stationary measure exists and is unique up to a multiplicative constant. Note that for a graph with bounded degrees, the stationary measure is bounded.

As $n$ increases, we expect that a simple random walk $(X_n)_{n\geq 0}$ on an infinite graph gets more 'spread out'. For example, if $X_n$ is a random walk on $\mathbb{Z}$, then $\mathrm{Var}[X_n] = n$. In particular, we expect that it becomes less likely (asymptotically) to find the random walk in a given vertex. Corollary 14.6 of Woess [13] immediately implies the following for simple random walks:

**Proposition 3.4.** Consider a simple random walk on a locally finite, infinite graph $G = (V, E)$. Let $p^{(n)}$ be the $n$-step transition function and let $\mu$ be a stationary measure with $\inf\limits_{v\in V} \mu(v) > 0$. Then there exists a constant $C > 0$ such that

$$\sup_{x,y\in V} \frac{p^{(n)}(x, y)}{\mu(y)} \leq \frac{C}{\sqrt{n}} \quad \text{for all } n > 0.$$

If $G$ has bounded degrees, then there exists a constant $C' > 0$ such that $p^{(n)}(x, y) \leq \frac{C'}{\sqrt{n}}$ for all vertices $x, y \in V$ and all integers $n > 0$.

If we consider a sufficiently small subset of the vertices of a given graph, the probability that we spend more than a fixed proportion of the time in this subset decreases exponentially in $n$. The difficulty in proving this fact lies in the dependence of the $X_i$.

**Proposition 3.5.** Let $G = (V, E)$ be an infinite, locally finite graph and let $(X_n)_{n\geq 0}$ be a simple random walk on $G$. Let $\mathcal{H} \subset V$ and $\alpha \in (0, 1)$ be given.
Assume that there exists an $n_0 \in \mathbb{N}$ such that

$$\mathbb{P}_v\left(\#\{i \leq n_0 : X_i \in \mathcal{H}\} \geq (\alpha/8)n_0\right) < \alpha/16 \qquad \text{for all } v \in V.$$

Let $c = \alpha/(8n_0)$. Then the following inequality holds:

$$\mathbb{P}_v\left(\#\{i \leq n : X_i \in \mathcal{H}\} \geq \alpha n\right) \leq 2\exp\{-cn\} \qquad \text{for all } v \in V,\ n \in \mathbb{N}.$$

*Proof.* We will divide time into blocks of size $n_0$, so that we can use the given inequality. Define

$$\xi_k = \mathbb{1}\{X_{kn_0} + \cdots + X_{(k+1)n_0-1} \geq (\alpha/8)n_0\}, \quad k \in \mathbb{N}.$$

The given inequality implies that $\mathbb{P}_v(\xi_k = 1) < \alpha/16$ for all $v$ and $k$. By conditioning on the starting vertex $v$, it follows that in fact $\mathbb{P}_\nu(\xi_k = 1) < \alpha/16$ for any initial distribution $\nu$. Note that the $\xi_k$'s are not independent. To deal with this, we prove the following claim:

*Claim 3.5.1.* Let $X_0 \sim \mu$. Then the following inequality holds:

$$\mathbb{E}_\mu \left[\exp\{\xi_0 + \cdots + \xi_m\}\right] \leq \exp\left\{\frac{\alpha(m+1)}{8}\right\} \qquad \text{for all } m \in \mathbb{N}.$$

*Proof.* Let $q = \mathbb{P}_\nu(\xi_0 = 1)$ for an arbitrary initial distribution $\nu$. Then we have

$$\mathbb{E}_\nu \left[\exp\{\xi_0\}\right] = q \cdot e + (1-q) = 1 + q \cdot (e-1) \leq 1 + 2q < 1 + 2\frac{\alpha}{16} = 1 + \frac{\alpha}{8} \leq \exp\{\alpha/8\},$$

by the inequality $\mathbb{P}_\nu(\xi_0 = 1) < \alpha/16$ and the fact that $1 + x \leq e^x$.

We prove the claim by induction on $m$. The base case has just been given, so we proceed with the induction step. Let $(\mathcal{F}_n)_{n\geq 0}$ be the natural filtration corresponding to $(X_n)_{n\geq 0}$. We bound

$$\begin{aligned}
\mathbb{E}_\mu \left[\exp\{\xi_0 + \cdots + \xi_{m+1}\}\right] &= \mathbb{E}_\mu \left[\mathbb{E}_\mu \left[\exp\{\xi_0 + \cdots + \xi_{m+1}\} \mid \mathcal{F}_{mn_0}\right]\right] \\
&= \mathbb{E}_\mu \left[\exp\{\xi_0 + \cdots + \xi_m\}\mathbb{E}_\mu \left[\exp\{\xi_{m+1}\} \mid \mathcal{F}_{mn_0}\right]\right] \\
&= \mathbb{E}_\mu \left[\exp\{\xi_0 + \cdots + \xi_m\}\mathbb{E}_{X_{mn_0}} \left[\exp\{\xi_0\}\right]\right] \\
&\leq \mathbb{E}_\mu \left[\exp\{\xi_0 + \cdots + \xi_m\} \exp\left\{\alpha/8\right\}\right] \\
&= \exp\left\{\alpha/8\right\} \mathbb{E}_\mu \left[\exp\{\xi_0 + \cdots + \xi_m\}\right],
\end{aligned}$$

by the law of iterated expectations, the fact that $\xi_0, \ldots, \xi_m$ are determined by $X_0, \ldots, X_{mn_0}$ and hence by $\mathcal{F}_{mn_0}$, the Markov property and the fact that $\mathbb{E}_\nu \left[\exp\{\xi_0\}\right] \leq \exp\{\alpha/8\}$ for any distribution $\nu$. This completes the induction step and hence the proof of the claim. □

We are now ready to prove the proposition. Let $N = 5n_0/\alpha$ and fix $n \geq N$. Let $K = \lfloor n/n_0 \rfloor - 1$. Then $K > n/n_0 - 2 \geq 5/\alpha - 2 \geq 3/\alpha$ and $n/n_0 \geq 5/\alpha > 5$ and hence $K > n/n_0 - 2 > n/(2n_0)$. Let $I_n$ be the number of indices $m \in \{0, \ldots, K\}$ such that $\xi_m = 1$. We observe that

$$\begin{aligned}
\#\{i \leq n : X_i \in \mathcal{H}\} &\leq I_n \cdot n_0 + (K - I_n) \cdot \tfrac{\alpha}{8} n_0 + n_0 \\
&\leq I_n \cdot n_0 + K \cdot \tfrac{\alpha}{8} n_0 + \tfrac{\alpha}{3} K n_0 < I_n \cdot n_0 + \tfrac{\alpha}{2} K n_0
\end{aligned}$$

since $\frac{\alpha}{3}K > 1$. In particular,

$$\text{if } \; I_n \leq \frac{\alpha}{2}K, \qquad \text{then } \; \frac{\#\{i \leq n : X_i \in \mathcal{H}\}}{n} \leq \frac{\frac{\alpha}{2}Kn_0 + \frac{\alpha}{2}Kn_0}{n} = \alpha\frac{Kn_0}{n} < \alpha.$$

This implies that the probability that $\#\{i \le n : X_i \in \mathcal{H}\} < \alpha n$ is larger than the probability that $I_n \le \frac{\alpha}{2}K$. By the claim and Markov's inequality, this yields:

$$\begin{aligned}
\mathbb{P}_v\left(\#\{i \le n : X_i \in \mathcal{H}\} \ge \alpha n\right) &\le \mathbb{P}_v\left(\xi_0 + \cdots + \xi_K \ge \tfrac{\alpha}{2}K\right) \\
&\le \mathbb{P}_v\left(\xi_0 + \cdots + \xi_K \ge \alpha n/(4n_0)\right) \\
&= \mathbb{P}_v\left(\exp\{\xi_0 + \cdots + \xi_K\} \ge \exp\{\alpha n/(4n_0)\}\right) \\
&\le \exp\{-\alpha n/(4n_0)\}\mathbb{E}_v\left[\exp\{\xi_0 + \cdots + \xi_K\}\right] \\
&\le \exp\{-\alpha n/(4n_0)\}\exp\left\{\tfrac{1}{8}\alpha(K+1)\right\} \le \exp\{-\alpha/(8n_0)n\}.
\end{aligned}$$

Since $c = \alpha/(8n_0)$, this proves the required inequality for $n \ge N$. For $n < N$ we have $2\exp\{-cn\} > 2\exp\{-cN\} = 2\exp\{-5/8\} > 1$, so then the inequality holds trivially. □

To apply the proposition, we need to check the condition that there exists an $n_0 \in \mathbb{N}$ such that

$$\mathbb{P}_v\left(\#\{i \le n_0 : X_i \in \mathcal{H}\} \ge (\alpha/8)n_0\right) < \alpha/16 \qquad \text{for all } v \in V.$$

A way to prove this condition is to show that $\mathbb{E}_v\left(\#\{i \le n_0 : X_i \in \mathcal{H}\}\right)$ grows sublinearly in $n_0$. Using this, we formulate a corollary of this result.

**Corollary 3.1.** Let $G = (V, E)$ be an infinite, locally finite graph and let $(X_n)_{n \ge 0}$ be a simple random walk on $G$. Let $\mathcal{H} \subset V$ be a subset of the vertex set such that

$$\lim_{n \to \infty} \frac{\mathbb{E}_v\left[\#\{i \le n : X_i \in \mathcal{H}\}\right]}{n} = 0.$$

for all $v \in V$ uniformly. Then for any $\alpha \in (0, 1)$, there exists a constant $c$ such that

$$\mathbb{P}_v\left(\#\{i \le n : X_i \in \mathcal{H}\} \ge \alpha n\right) \le 2\exp\{-cn\} \qquad \text{for all } v \in V,\ n \in \mathbb{N}.$$

If $n_0$ satisfies $\mathbb{E}_v\left(\#\{i \le n_0 : X_i \in \mathcal{H}\}\right) \le (\alpha/8)n_0 \cdot \alpha/16$ for all $v \in V$, we can take $c = \frac{\alpha}{8n_0}$.

*Proof.* If $n_0$ satisfies $\dfrac{\mathbb{E}_v\left[\#\{i \le n_0 : X_i \in \mathcal{H}\}\right]}{n_0} \le \dfrac{\alpha^2}{8 \cdot 16}$, then by Markov's inequality we have

$$\mathbb{P}_v\left(\#\{i \le n_0 : X_i \in \mathcal{H}\} \ge (\alpha/8)n_0\right) < \frac{\mathbb{E}_v\left[\#\{i \le n_0 : X_i \in \mathcal{H}\}\right]}{(\alpha/8)n_0} \le \frac{\alpha}{16},$$

and hence the result follows from Proposition 3.5.
The existence of such $n_0$ is guaranteed by the given uniform limit. □

Note that we have been very explicit about the constant $c$, which is somewhat unusual. The reason for this is that we will apply Corollary 3.1 in a setting where $n_0$ is not constant.

### 3.1.4 Combinatorics of the simple random walk on $\mathbb{Z}$

Let $(S_n)_{n\geq 0}$ be a simple random walk on $\mathbb{Z}$ with $S_0 = 0$. Let $H_n$ be the number of meetings of $S_n$ with zero up to time $n$ after the initial meeting, i.e. $H_n = \sum_{i=1}^n \mathbf{1}(S_i = 0)$. In Kirshnapur and Peres [7] it is claimed, but not proven, that there exists a constant $C > 0$ such that $P(H_n = k) \leq \frac{C}{\sqrt{n}}$, for all integers $n \geq 1$ and $k \geq 0$. Note that $S_i$ is odd for odd $i$, so $H_{2n}$ and $H_{2n-1}$ are identically distributed for any integer $n \geq 1$. We will therefore focus on $H_{2n}$.

We first recall the following lemma, which is easily proven by induction:

**Lemma 3.1. (Hockeystick Lemma)**
Let $n, k \geq 0$ be integers. Then

$$\sum_{i=k}^{n} \binom{i}{k} = \binom{n+1}{k+1}.$$

We also recall the Catalan numbers [11]. The $n$th Catalan number is defined by $C_n = \frac{1}{n+1}\binom{2n}{n}$. A standard interpretation of the $n$th Catalan number is the number of paths of $2n$ incrementing or decrementing steps that remain nonnegative and start and end at 0. By adding one additional incrementing step in the beginning and one additional decrementing step, we see that $C_n$ is also the number of paths of $2n+2$ incrementing or decrementing steps that start and end at 0 and remain positive in between. This yields the following lemma:

**Lemma 3.2. (Catalan numbers)**
The number of paths of $2n+2$ incrementing or decrementing steps that start and end at 0 and remain positive in between, is equal to the $n$th Catalan number $C_n = \frac{1}{n+1}\binom{2n}{n}$.

The following theorem gives the distribution of $H_{2n}$:

**Theorem 3.2.** Let $(S_n)_{n\geq 0}$ be a SRW on $\mathbb{Z}$ with $S_0 = 0$. Let $H_n = \sum_{i=1}^{n} \mathbf{1}(S_i = 0)$. Then

$$\mathbb{P}(H_{2n} = k) = \frac{2^k}{2^{2n}} \binom{2n-k}{n}, \quad \text{for } 0 \leq k \leq n.$$

*Proof.* Note that the total number of paths of length $2n$ is $2^{2n}$, so we can equivalently show that the number of paths of length $2n$ that meets zero exactly $k$ times is equal to $2^k\binom{2n-k}{n}$. Since the path has $k$ meetings with zero excluding the meeting at $n = 0$, we have $k$ paths in total from zero to zero. Such a path should not cross zero except at the endpoints, so it either only visits positive numbers or only negative numbers. Note that negating such a path again gives a path from zero to zero satisfying the requirements. Let us therefore count the number of paths of length $2n$ that meets zero exactly $k$ times and that are nonnegative until the last meeting with zero. The total number of paths is then $2^k$ times as large, since each of the $k$ paths between zero and zero can be negated independently.

Let us call paths of length $2n$ that meets zero exactly $k$ times and that are nonnegative until the last meeting with zero, $(n,k)$-special paths. So we need to show that the number of $(n,k)$-special paths is equal to $\binom{2n-k}{n}$ for all $0 \leq k \leq n$. This can be shown by induction on $n \geq 1$.

*Base case:* Let $n = 1$. Then the relevant values for $k$ are 0 and 1. For $k = 0$, there are two paths: namely the path consisting of two incrementing steps and the path consisting of two decrementing steps. For $k = 1$, we only have the path consisting of first an incrementing step and then a decrementing step. This matches the formula, as $\binom{2}{1} = 2$ and $\binom{1}{1} = 1$.

*Induction step*: Let $r \geq 1$ be given and assume that the number of $(r,k)$-special paths is equal to $\binom{2r-k}{r}$ for all $0 \leq k \leq r$ (Induction Hypothesis).
We will now first prove that the claim also holds for $n = r+1$ and $k = 0$. Consider any path of length $2n$ that does not meet zero again after $n = 0$. If we add two steps, then there are two possiblities: either this yields a path of length $2n+2$ that also does not meet zero again, or it yields a path that meets zero again for the first time at step $2n+2$. Note that we have $\binom{2n}{n}$ paths of length $2n$ that do not meet zero (by the Induction Hypothesis), and 4 ways to add two additional steps, so $4\binom{2n}{n}$ paths in total.
By Lemma 3.2, the number of paths that meets zero again for the first time at step $2n+2$ is equal to $2C_n = \frac{2}{n+1}\binom{2n}{n}$, where the factor 2 comes from the fact that the path can also be negative in between. Hence, the total number of $(r+1,0)$-special paths is equal to

$$\left(4 - \frac{2}{n+1}\right)\binom{2n}{n} = \frac{(2n+2)(2n+1)}{(n+1)^2}\binom{2n}{n} = \binom{2n+2}{n+1},$$

which shows that the claim also holds for $n = r+1$ and $k = 0$.

We will now show that the claim also holds for $n = r+1$ and $k \geq 1$. We will give a bijection between the set of $(n,\ell)$-special paths with $\ell \geq k-1$ and the set of $(n+1,k)$-special paths. For a $(n,\ell)$-special path, construct a path of length $2n+2$ by adding an incrementing in front of it and adding the decrementing step after the $(\ell-k+1)$th meeting in the original path (with the understanding that if $\ell = k-1$, then we add the decrementing step immediately after the incrementing step). Because we consider nonnegative paths, the path after the additional incrementing step remains positive until the additional decrementing step is added. Hence, $\ell-k+1$ meetings with zero are removed, but one is added back through the added decrementing step, so the total number of meetings equals $k$. Hence, this yields a $(n+1,k)$-special path.

We now describe the inverse operation. Consider a $(n+1,k)$-special path. Since this path is nonnegative until the last meeting with 0, the first step must be incrementing. Now find the first decrementing step which decrements to 0. Then the path must be positive in between (otherwise it would not be the first step which decrements to 0), so when removing the incrementing and this decrementing step, we get a nonnegative path. Moreover, at least $k-1$ meetings with zero remain, since the path after the first meeting remains unchanged. We conclude that this results in a $(n,\ell)$-special path with $\ell \geq k-1$.

Hence, there is a bijection between the set of $(n,\ell)$-special paths with $\ell \geq k-1$ and the set of $(n+1,k)$-special paths. Using the Induction Hypothesis and the Hockeystick Lemma (Lemma 3.1), we see that the number of $(n+1,k)$-special paths is therefore equal to

$$\sum_{\ell=k-1}^{n} \binom{2n-\ell}{n} = \sum_{i=n}^{2n-k+1} \binom{i}{n} = \binom{(2n-k+1)+1}{n+1} = \binom{2(n+1)-k}{n+1}.$$

This shows that the claim also holds for $n = r+1$ and $k \geq 1$. This completes the induction. Hence, the number of $(n,k)$-special paths is equal to $\binom{2n-k}{n}$ for all $0 \leq k \leq n$. By the preceding discussion, this completes the proof of the theorem. $\square$

We have now shown that

$$\mathbb{P}(H_{2n} = k) = \frac{2^k}{2^{2n}} \binom{2n-k}{n}, \quad \text{for } 0 \leq k \leq n.$$

We now show that $\mathbb{P}(H_{2n} = k)$ is decreasing in $k$ and hence find an upper bound for $\mathbb{P}(H_{2n} = k)$. This implies the required result, which we formulate as a corollary.

**Corollary 3.2.** Let $(S_n)_{n\geq 0}$ be a SRW on $\mathbb{Z}$ with $S_0 = 0$. Let $H_n = \sum_{i=1}^{n} \mathbf{1}(S_i = 0)$. Then there exists a constant $C > 0$ such that

$$\mathbb{P}(H_n = k) \leq \frac{C}{\sqrt{n}},$$

for all integers $n \geq 1$ and $k \geq 0$.

*Proof.* Since

$$\binom{2n-k}{n} = \frac{2n-k}{n-k} \binom{2n-(k+1)}{n} \geq \frac{2n-2k}{n-k} \binom{2n-(k+1)}{n} = 2\binom{2n-(k+1)}{n},$$

it follows that $\frac{2^k}{2^{2n}}\binom{2n-k}{n} \geq \frac{2^{k+1}}{2^{2n}}\binom{2n-(k+1)}{n}$ and hence $\mathbb{P}(H_{2n} = k)$ is decreasing in $k$.
In particular, we have

$$\mathbb{P}(H_{2n} = k) \leq \mathbb{P}(H_{2n} = 0) = \frac{1}{2^{2n}} \binom{2n}{n} \leq \frac{C'}{\sqrt{n}},$$

for some constant $C'$ and all integers $n \geq 1$ and $k \geq 0$.
Since $\mathbb{P}(H_{2n-1} = k) = \mathbb{P}(H_{2n} = k)$, we conclude that

$$\mathbb{P}(H_n = k) \leq \frac{C}{\sqrt{n}},$$

for some constant $C$ and all integers $n \geq 1$ and $k \geq 0$. $\square$

Even though Theorem 3.2, is formulated in terms of probabilities, the proof is mostly of a combinatorial nature. Apart from this result, there are two more results on the combinatorics of the simple random walk on $\mathbb{Z}$ that will be needed. The exposition here is based on the notes by Alm [1]. Let $N_n(a,b)$ denote the number of paths from $a$ to $b$ in $n$ steps, where each step either increments or decrements the position by 1. Let $N_n^{\neq 0}(a,b)$ denote the number of paths from $a$ to $b$ in $n$ steps that do not (re)visit 0. If $a=0$ then the visit at the starting point does not count. Note that $N_n(a,b)=0$ if $n$ and $a-b$ do not have the same parity. If $a-b$ and $n$ do have the same parity, then define $h=\frac{1}{2}(n-(a-b))$. It then follows that $N_n(a,b)=\binom{n}{h}$. The number of paths that do not visit zero, $N_n^{\neq 0}(a,b)$, can be counted using a mirroring argument:

**Lemma 3.3.** Let $a,b>0$ be integers. Then

$$N_n^{\neq 0}(a,b)=N_n(a,b)-N_n(-a,b).$$

*Proof.* If $n$ and $a-b$ do not have the same parity, then all terms in this equation are zero, so assume from now on that $a-b$ and $n$ do have the same parity and define $h=\frac{1}{2}(n-(a-b))$. Consider a path from $a$ to $b$ that does visit 0. By mirroring the path up to the first visit to 0, we obtain a path from $-a$ to $b$. Note that mirroring a path $\mathcal{P}=(p_0,\ldots,p_{k-1},0,p_{k+1},\ldots,p_n)$ up to a point $k$ such that $p_k=0$, results in the path $\mathcal{P}'=(-p_0,\ldots,-p_{k-1},0,p_{k+1},\ldots,p_n)$. Conversely, any path from $-a$ to $b$ must visit 0 since $-a<0<b$. So by mirroring the path up to the first visit to 0, we obtain a path from $a$ to $b$ that does visit 0. Note that the mirroring operation is its own inverse. Hence, this is a bijection between the paths that from $a$ to $b$ that visit 0 and the paths from $-a$ to $b$. The number of paths that visit 0 can be computed by subtracting the number of paths that do not visit 0 from the total number of paths. Hence, $N_n(a,b)-N_n^{\neq 0}(a,b)=N_n(-a,b)$, and rewriting this equality proves the lemma. □

For $a=0$ the situation is different, since in the proof the random walk first makes a step to avoid counting the initial visit to 0. This problem is known as the ballot problem, because paths that increment or decrement by 1 each step can be interpreted as the net number of votes for a candidate $A$ against candidate $B$, if the votes are made known one by one. A vote in favor of $A$ increments the net amount by 1, and a vote for $B$ decrements the net amount by 1. Assume that candidate $A$ wins with $b>0$ votes more than $B$. If the votes were made known in a random order, what is the probability that candidate $A$ was strictly ahead of candidate $B$ the whole time? This question is answered by the ballot theorem:

**Theorem 3.3.** (**Ballot theorem**)
Let $b>0$ be an integer. Then

$$N_n^{\neq 0}(0,b)=\frac{b}{n}N_n(0,b).$$

*Proof.* If $n$ and $b$ do not have the same parity, then both sides of the equation are zero, so assume that $b$ and $n$ do have the same parity and define $h=\frac{1}{2}(n+b)$. Then $N_n(0,b)=\binom{n}{h}$.

If the first step is decrementing, then the remainder of the path must go from $-1$ to $b > 0$, so it revisits zero at least once. So a path which does not revisit zero must start with an incrementing step and the remaining path must be a path of $n-1$ steps from 1 to $b$ which does not visit 0. By Lemma 3.3, it follows that

$$\begin{aligned} N_n^{\neq 0}(0,b) &= N_{n-1}^{\neq 0}(1,b) = N_{n-1}(1,b) - N_{n-1}(-1,b) \\ &= \binom{n-1}{h-1} - \binom{n-1}{h} = \binom{n}{h} \cdot \frac{h}{n} - \binom{n}{h} \cdot \frac{n-h}{n} = \binom{n}{h} \cdot \frac{2h-n}{n} = \frac{b}{n}\binom{n}{h}, \end{aligned}$$

and since $N_n(0,b) = \binom{n}{h}$, this proves the required equality. $\square$

In terms of random walks, the ballot theorem can be formulated as follows: Let $(S_n)_{n\geq 0}$ be a SRW on $\mathbb{Z}$ and let $b > 0$. Then $\mathbb{P}\left(S_i > 0 \text{ for all } 1 \leq i \leq n \mid S_n = b\right) = \frac{b}{n}$.

A mirroring argument can also be used to prove the following theorem:

**Theorem 3.4.** (**Maximum of a random walk**)
Let $(S_n)_{n\geq 0}$ be a SRW on $\mathbb{Z}$ and let $M_n = \max\left(S_0, S_1, \ldots, S_n\right)$. Then for $r > 0$ it holds that

$$\mathbb{P}(M_n \geq r) = \mathbb{P}(S_n = r) + 2\mathbb{P}(S_n > r).$$

*Proof.* Since $r > 0$, we have $M_n \geq r$ if and only if $S$ visits $r$ at some time $1 \leq k \leq n$. Let us first show that

$$\mathbb{P}(M_n \geq r, S_n < r) = \mathbb{P}(M_n \geq r, S_n > r).$$

By multiplying by $2^n$, this equality can be interpreted combinatorially as showing that the number of paths that hit $r$ that also satisfy $S_n < r$, is equal to the number of paths that hit $r$ that also satisfy $S_n > r$. Let $k$ be the time of the first visit to $r$.
We construct a bijection between these sets of paths by mirroring the paths around $k$ from time $k$ onwards. In this case, mirroring a path $\mathcal{P} = (p_0, \ldots, p_{k-1}, 0, p_{k+1}, \ldots, p_n)$ around $r$ from a time $k$ such that $p_k = r$ onwards, results in the path $\mathcal{P}' = (p_0, \ldots, p_{k-1}, r, r - p_{k+1}, \ldots, r - p_n)$. Note that the mirroring operation is its own inverse. Moreover, it maps a path with $p_n > r$ to a path with $p_n' = r - p_n < r$. Hence, this is a bijection between the paths that hit $r$ with $S_n < r$ and the paths that hit $r$ with $S_n < r$. So $\mathbb{P}(M_n \geq r, S_n < r) = \mathbb{P}(M_n \geq r, S_n > r)$.

It now follows that

$$\begin{aligned} \mathbb{P}(M_n \geq r) &= \mathbb{P}(M_n \geq r, S_n < r) + \mathbb{P}(M_n \geq r, S_n = r) + \mathbb{P}(M_n \geq r, S_n > r) \\ &= \mathbb{P}(M_n \geq r, S_n = r) + 2\mathbb{P}(M_n \geq r, S_n > r) = \mathbb{P}(S_n = r) + 2\mathbb{P}(S_n > r), \end{aligned}$$

where the last inequality holds since $S_n = r$ and $S_n > r$ already imply that $M_n \geq r$. $\square$

## 3.2 Electric network theory

In this section, we consider graphs as electric networks, where vertices are nodes and edges are have a given resistance. Since we are dealing with simple random walks only, we assume that every edge has unit resistance. The content of this section is based on Section 3.1 of Barlow, Peres and Sousi [3], augmented with the lecture notes by Valesin and Verbitskiy [12].

Let $d(x)$ denote the degree of a vertex $x$ of a graph $G$. Let $X$ be a simple random walk on $G$. The **transition density** $q_t(x, y)$ is defined by $q_t(x, y) = \frac{\mathbb{P}_x[X_t=y]}{d(y)}$. The reason for dividing by the degree of the vertex is to make $q_t$ a symmetric function, as the following lemma shows:

**Lemma 3.4.** Let $G = (V, E)$ be a locally finite graph and let $q_t(x, y)$ be the transition density corresponding to a simple random walk on $G$. Then $q_t(x, y) = q_t(y, x)$ for all $x, y \in V$.

*Proof.* Let $\mathcal{P} = (p_0, \dots, p_t)$ be a path in $G$ starting from $p_0 = x$ with endpoint $p_t = y$. Let $\mathcal{P}' = (p_t, \dots, p_0)$ be the reverse of $\mathcal{P}$. Let $X$ be a simple random walk on $G$. Then:

$$\mathbb{P}_x[\{X_k\}_{k\le t} = \mathcal{P}] = \prod_{k=0}^{t-1} \frac{1}{d(p_k)} = \frac{d(p_t)}{d(p_0)} \prod_{k=0}^{t-1} \frac{1}{d(p_{t-k})} = \frac{d(y)}{d(x)} \mathbb{P}_y[\{X_k\}_{k\le t} = \mathcal{P}'].$$

Note that there exists a bijection between the paths $\mathcal{P}$ from $x$ to $y$ of length $t$ and the paths $\mathcal{P}'$ of length $t$ from $y$ to $x$, namely reversing the path. This is a bijection since it is equal to its own inverse. By summing over all possible paths $\mathcal{P}$, we hence find $\mathbb{P}_x[X_t = y] = \frac{d(y)}{d(x)} \mathbb{P}_y[X_t = x]$, so

$$q_t(x, y) = \frac{\mathbb{P}_x[X_t = y]}{d(y)} = \frac{\mathbb{P}_y[X_t = x]}{d(x)} = q_t(y, x),$$

which is the required equality. $\square$

Let $G = (V, E)$ be a graph. For functions $f, g : V \to \mathbb{R}$, we define the quadratic form

$$\mathcal{E}(f, g) = \frac{1}{2} \sum_{x\sim y} (f(x) - f(y))(g(x) - g(y)),$$

where we sum over all vertices $x, y \in V$ that are connected by an edge. Since we consider an undirected graph, we have $x \sim y$ if and only if $y \sim x$, so every pair of vertices that is connected by an edge is actually counted twice in this sum. In terms of electric networks, the functions $f$ will typically be the electric potential of the vertices $x$ and $y$. Then $f(x) - f(y)$ is the potential difference of $x$ and $y$. Since $x$ and $y$ are connected by an edge of unit resistance, this is then also the current flowing from $x$ to $y$. In this context, $\mathcal{E}(f, f)$ is the energy flow corresponding to the potential $f$. Note that $\mathcal{E}(f, f)$ is a sum of squares, so it is nonnegative.

Let $A$ and $B$ be two subsets of $V$. The **effective resistance** between $A$ and $B$ is defined by

$$R_{\text{eff}}(A, B)^{-1} = \inf\{\mathcal{E}(f, f) : \mathcal{E}(f, f) < \infty, f|_A = 1, f|_B = 0\}.$$

We now interpret the effective resistance physically, in terms of electric networks. Suppose that all vertices in $A$ are connected to a single additional vertex $a$ and that all vertices in $B$ are connected to a single additional vertex $b$. When applying a potential of $V$ to $a$ and connecting $b$ to ground, a current $I$ flows from $a$ to $b$. Since resistors are linear elements, the ratio $\frac{V}{I}$ is constant, which is equal to the effective resistance $R_{\text{eff}}(A, B)$.

The effective resistance has several useful properties. For stating these properties, a few graph theoretic notions are needed. A graph $G' = (V', E')$ is a subgraph of $G = (V, E)$ if $G'$ can be obtained from $G$ by deleting vertices and edges. Formally, the required conditions are $V' \subseteq V$ and $E' \subseteq \{\{x, y\} \in E : x, y \in V'\}$, although the last condition can be simplified to $E' \subseteq E$ when the condition that $G'$ is a graph is already included.
The subgraph of $G$ induced by a vertex set $V' \subseteq V$ is the subgraph of $G$ with vertex set $V'$ and edge set $E' = \{\{x, y\} \in E : x, y \in V'\}$, i.e. no more edges are removed than necessary.

In a connected graph, the graph distance $d(x, y)$ between two vertices $x, y \in V$ is defined by the length of the shortest path between $x$ and $y$. For example, $d(x, y) = 1$ if and only if $x \sim y$.

**Lemma 3.5.** Let $G = (V, E)$ be a connected locally finite graph. Let $\mathcal{E}_G(f, f)$ be the quadratic form representing the energy flow corresponding to a potential $f : V \to \mathbb{R}$. Then:

- If $G' = (V', E')$ is a connected subgraph of $G$, then $\mathcal{E}_{G'}(f|_{V'}, f|_{V'}) \leq \mathcal{E}_G(f, f)$.
- For any connected graph, $R_{\text{eff}}(x, y) \leq d(x, y)$. If $G$ is a tree, then $R_{\text{eff}}(x, y) = d(x, y)$.

*Proof.* Note that the quadratic form can be seen as a sum over all edges, where every edge $\{x, y\} \in E$ is counted twice, but this is compensated for by the factor $\frac{1}{2}$. Since $E' \subseteq E$, the inequality $\mathcal{E}_{G'}(f|_{V'}, f|_{V'}) \leq \mathcal{E}_G(f, f)$ directly follows from the fact that the sum defining $\mathcal{E}_{G'}(f|_{V'}, f|_{V'})$ included all terms that the sum defining $\mathcal{E}_G(f, f)$ does, but $\mathcal{E}_G(f, f)$ possibly includes some additional nonnegative terms.

If $G$ is a tree, then for any two vertices $x, y$ there exists exactly one path $(x_0, \dots, x_d)$ such that $x_0 = x$, $x_d = y$, $x_i \sim x_{i+1}$ for all $0 \leq i \leq d - 1$ and $d = d(x, y)$. Let $f$ be a function satisfying $f(x) = 1$ and $f(y) = 0$. It now follows that

$$\begin{aligned} \mathcal{E}_G(f, f) &= \frac{1}{2} \sum_{u \sim v} (f(u) - f(v))^2 \geq \sum_{i=0}^{d-1} (f(x_i) - f(x_{i+1}))^2 \\ &\geq d \left( \frac{1}{d} \sum_{i=0}^{d-1} |f(x_i) - f(x_{i+1})| \right)^2 \geq d \left( \frac{1}{d} \left| \sum_{i=0}^{d-1} f(x_i) - f(x_{i+1}) \right| \right)^2 = d^{-1}, \end{aligned}$$

by leaving out all edges not on the path, the inequality of quadratic and arithmetic means, the triangle inequality and the fact that the sum defining the arithmetic mean telescopes to $f(x) - f(y) = 1$. Moreover, equality can be achieved by choosing $f(x_i) = \frac{d-i}{d}$. This means that all differences are the same and also all positive, which yields equality in the final two inequalities. We can now inductively define $f$ for vertices off the path by stating that for all vertices $v, w \in V$ such that $v \sim w$ but such that $v, w$ are not both on the path, we have $f(v) = f(w)$. Suppose that, when following this procedure, there exists a vertex $v$ such that $f(v)$ is assigned two different values. Then there exist paths from $v$ to $x_i$ and from $v$ to $x_j$, which do not make use of any of the edges on the path $(x_0, \dots, x_d)$, for two different $0 \leq i, j \leq d$. By following the path $(x_0, \dots, x_d)$ from $x_j$ to $x_i$, this yields two different paths from $v$ to $x_i$. But this contradicts that $G$ is a tree. So function values can be assigned using this procedure, and this gives equality in the first inequality. Hence, $\mathcal{E}_G(f, f) \geq d^{-1}$ for all $f$ satisfying $f(x) = 1$ and $f(y) = 0$ and equality can be achieved, so $R_{\text{eff}}(x, y)^{-1} = d^{-1}$, so $R_{\text{eff}}(x, y) = d(x, y)$ as required.

Finally, if $G$ is a connected graph, then the above inequality still holds, but it is not always possible to achieve equality. So then $R_{\text{eff}}(x, y)^{-1} \geq d(x, y)^{-1}$, and hence $R_{\text{eff}}(x, y) \leq d(x, y)$. $\square$

### 3.2.1 Harmonic functions

To further analyze the energy flow, some elementary properties of harmonic functions are needed: in particular, the existence and uniqueness of harmonic functions subject to some conditions. The next three propositions are from the lecture notes by Verbitskiy and Valesin [12].

**Proposition 3.6.** (**Existence principle**)
Let $G = (V, E)$ be a connected locally finite graph and let $W$ be a proper subset of $V$. Let $f_0 : V \setminus W \to \mathbb{R}$ be a bounded function. Then there exists a function $f : V \to \mathbb{R}$ whose restriction to $V \setminus W$ is equal to $f_0$ and which is harmonic on $W$.

*Proof.* Let $(X_n)_{n \geq 0}$ be a simple random walk on $G$ and consider the stopping time

$$\tau := \inf\{n \geq 0 : X_n \in V \setminus W\}.$$

Define the random variable $Y$ by $Y = f_0(X_\tau)$ if $\tau < \infty$ and $Y = 0$ otherwise. Define the function $f : V \to \mathbb{R}$ by $f(x) = \mathbb{E}[Y \mid X_0 = x]$. If $x \in V \setminus W$, then $\tau = 0$ and hence $f(x) = f_0(x)$. It remains to show that $f$ is harmonic on $W$. Let $x \in W$ be given. Then

$$\begin{aligned} f(x) = \mathbb{E}[Y((X_n)_{n \geq 0}) \mid X_0 = x] &= \sum_{x \in V} \mathbb{E}[Y((X_n)_{n \geq 0}) \mid X_1 = y, X_0 = x] \mathbb{P}[X_1 = y \mid X_0 = x] \\ &= \sum_{x \in V} \mathbb{E}[Y((X_n)_{n \geq 1}) \mid X_1 = y] p(x, y) = \sum_{x \in V} f(y) p(x, y), \end{aligned}$$

by the law of total expectation, the Markov property and the definition of a Markov chain, and the fact that $x \in W$ implies that $\tau \geq 1$ which makes it possible to forget $X_0$ and shift all indices by 1. Hence, the function $f$ has the requested properties. This completes the proof. $\square$

**Proposition 3.7.** (**Maximum principle**)
Let $G = (V, E)$ be a locally finite graph and let $W$ be a finite subset of $V$. Assume that the subgraph of $G$ induced by $W$ is connected. Let $f : V \to \mathbb{R}$ be harmonic on $W$, such that

$$\max_{x \in W} f(x) = \sup_{x \in V} f(x).$$

Then $f$ is constant on the set $\overline{W} := \{y \in V : \exists x \in W : x \sim y\}$.

*Proof.* Assume that $w \in W$ satisfies $f(w) = \max_{x \in W} f(x)$. We prove by induction on $d(w, x)$ that $f(x) = f(w)$ if $x \in \overline{W}$. Here $d$ denotes the graph distance where only edges from the subgraph of $G$ induced by $W$ can be used, except for possibly the last edge which can be from an element of $W$ to an element outside $W$, but inside $\overline{W}$. Since the subgraph of $G$ induced by $W$ is connected, the result then follows.
If $d(w, x) = 0$, then $w = x$ and the result trivially holds. Now let $k \in \mathbb{N}$ and assume that $f(x) = f(w)$ for all $x \in \overline{W}$ with $d(w, x) = k$. Let $y \in \overline{W}$ with $d(w, y) = k + 1$ be given. Then there exists a $w' \in W$ with $d(w, w') = k$ and $w' \sim y$. Since $w' \in W$, it follows that

$$f(w') = \frac{1}{\deg(w')} \sum_{x \sim w'} f(x) \leq \frac{1}{\deg(w')} \sum_{w' \sim x} \sup_{x \in V} f(x) = \frac{1}{\deg(w')} \cdot \deg(w') f(w) = f(w'),$$

so the inequality is in fact an equality, and hence $f(x) = \sup_{x \in V} f(x) = f(w)$ for all $x$ with $w' \sim x$. In particular, $f(y) = f(w)$. Since $y$ was arbitrary, we conclude that $f(y) = f(w)$ for all $y \in \overline{W}$ with $d(w, y) = k + 1$. This completes the induction and hence the proof. □

If the set $W$ is finite, then the harmonic function, whose existence is asserted by the existence principle, can be shown to be unique.

**Proposition 3.8.** (**Uniqueness principle**)
Let $G = (V, E)$ be a connected locally finite graph and let $W$ be a finite proper subset of $V$. If $f, g : V \to \mathbb{R}$ are harmonic on $W$ and equal outside $W$, then they are equal everywhere.

*Proof.* Let $h = f - g$. Then $h$ is identically 0 outside $W$ and harmonic on $W$. Since $W$ is finite, it follows that $h$ attains only finitely many positive values. Suppose that $h$ attains a positive value somewhere on $W$. So $h$ attains its supremum on a connected component $W'$ of $W$ so by the maximum principle, $h$ is constant on the set $\overline{W'}$. Since $W$ is a proper subset of $V$, and since $G$ is connected, $\overline{W'} \setminus W$ is nonempty. If $v \in \overline{W'} \setminus W$, then $h(v) = 0$. Hence, $h(w) = h(v) = 0$ for all $w \in \overline{W'}$. But $h$ attains its supremum on $W'$, so then the supremum of $h$ is 0, which contradicts the assumption that $h$ attains a positive value. So $f - g$ does not attain positive values. By switching $f$ and $g$, it follows that $g - f$ does not attain positive values either. So $f - g$ must be identically zero, and hence $f$ and $g$ are equal everywhere. □

We now have the tools to continue with the investigation of the energy flow. If $A^c \cap B^c$ is finite, the infimum $\inf\{\mathcal{E}(f,f) : \mathcal{E}(f,f) < \infty, f|_A = 1, f|_B = 0\}$ is attained by a function $f$, which is harmonic on the set $A^c \cap B^c$.

**Lemma 3.6.** Let $G = (V, E)$ be a locally finite graph and let $A, B \subset V$ be two nonempty disjoint subsets of the vertex set such that $A^c \cap B^c$ is finite. There exists a function $h : V \to \mathbb{R}$ such that $h|_A = 1$, $h|_B = 0$ and such that $R_{\text{eff}}(A, B)^{-1} = \mathcal{E}(h, h)$. This function $h$ is harmonic on the set $A^c \cap B^c$.

*Proof.* Since $A^c \cap B^c$ is finite and the function $f$ is already determined on $A \cup B$, this is a finite-dimensional optimization problem. Assume that $v \in A^c \cap B^c$. The first order condition is:

$$\frac{\mathrm{d}\,\mathcal{E}(f,f)}{\mathrm{d}[f(v)]} = \frac{1}{2}\sum_{v\sim y} 2(f(v) - f(y)) - \frac{1}{2}\sum_{x\sim v} 2(f(x) - f(v)) = 2\sum_{v\sim y}(f(v) - f(y)) = 0,$$

which implies $f(v) = \frac{1}{\deg(v)} \sum\limits_{v\sim y} f(y)$. Since this holds for all $v \in A^c \cap B^c$, it follows that any function $f$ which minimizes $\mathcal{E}(f,f)$ must be harmonic on $A^c \cap B^c$. Note that $A \cup B$ is nonempty and that a function $h : V \to \mathbb{R}$ such that $h|_A = 1$, $h|_B = 0$ is bounded on $A \cup B$, so the existence principle applies. Hence, there actually exists a function $h : V \to \mathbb{R}$ such that $h|_A = 1$, $h|_B = 0$ which is harmonic on the set $A^c \cap B^c$. Moreover, it is unique by the uniqueness principle, which applies since $A \cup B$ is finite, it follows that $h$ is uniquely determined.

It remains to prove that $h$ actually minimizes $\mathcal{E}(f,f)$. To do this, let us show that $\mathcal{E}(f,f)$ is convex as function of the variables $f(v)$ for $v \in A^c \cap B^c$ by calculating the Hessian. We have

$$\frac{\mathrm{d}^2\,\mathcal{E}(f,f)}{\mathrm{d}[f(v)]^2} = 2\deg(v) \qquad \text{and} \qquad \frac{\mathrm{d}^2\,\mathcal{E}(f,f)}{\mathrm{d}[f(v)]\mathrm{d}[f(w)]} = \begin{cases} -2 & \text{if } v \sim w \\ 0 & \text{otherwise.} \end{cases}$$

Hence, the Hessian is two times the Laplacian matrix of the graph, which is positive semidefinite. To prove this, write $A^c \cap B^c = \{v_1, \ldots, v_n\}$ and order the rows and columns in the Hessian accordingly. Denote the Hessian by $\mathcal{L} = [\ell_{ij} : 1 \le i, j \le n]$. Then

$$|\ell_{ii}| = 2\deg(v) = \sum_{i\neq j} |\ell_{ij}|,$$

so $\mathcal{L}$ is diagonal dominant. So the Hessian is symmetric and diagonal dominant, which is a sufficient condition for positive semidefiniteness.

Since the Hessian is a constant positive semidefinite matrix, the function is convex, so it has a global interior minimum. Hence, it follows that $h$ minimizes $\mathcal{E}(f,f)$. Hence,

$$R_{\text{eff}}(A, B)^{-1} = \inf\{\mathcal{E}(f,f) : \mathcal{E}(f,f) < \infty, f|_A = 1, f|_B = 0\} = \mathcal{E}(h,h),$$

so it follows that $R_{\text{eff}}(A, B)^{-1} = \mathcal{E}(h, h)$, as required. $\square$

### 3.2.2 Properties of Green's function

Let $G = (V, E)$ be a graph and let $(X_t)_{t\geq 0}$ be a simple random walk on $G$. Recall that Green's function is defined by $G(x, y) = \sum_{j=0}^{\infty} \mathbb{P}_x(X_j = y)$. It is however more convenient to work with the Green kernel, which is Green's function divided by $\deg y$. Put differently, this means working with the transition density $q_j(x, y)$ instead of the probability $\mathbb{P}_x(X_j = y)$.
Let $B$ be a finite subset of the vertex set $V$ Let $q_j^B(x, y)$ be the $j$-step transition density for a random walk which is killed when it exits $B$. Formally, this means that we consider the Markov chain where for every state in $B^c$ becomes an absorbing state. In addition, set $q_j^B(x, y) = 0$ if $x \in B^c$ or $y \in B^c$. There is also a Green kernel corresponding to the transition density $q_j^B$. Finally, we also need Green kernel up to time $t$:

**Definition 3.4.** (**Green kernels**)

- The Green kernel is defined by $g(x, y) = \sum_{j=0}^{\infty} q_j(x, y)$.
- The Green kernel of the Markov chain killed after exiting $B$ is $g_B(x, y) = \sum_{j=0}^{\infty} q_j^B(x, y)$.
- The Green kernel up to time $t$ is defined by $g_t(y, x) = \sum_{j=0}^{t} q_j(x, y)$.

By applying Lemma 3.6 to $\{x\}$ and $B^c$, it follows that there exists a function $h : V \to \mathbb{R}$ such that $h(x) = 1$, $h|_{B^c} = 0$, which is harmonic on $B \setminus \{x\}$ and such that $R_{\text{eff}}(x, B^c)^{-1} = \mathcal{E}(h, h)$. In the next lemma, it is shown that $h$ can be expressed in terms of Green kernels.

**Lemma 3.7.** Let $x \in B$ be given and define the function $\widetilde{g} : V \to \mathbb{R}$ by $\widetilde{g}(y) = g_B(x, y)$. Define the function $h : V \to \mathbb{R}$ by $h(y) = \frac{g_B(x,y)}{g_B(x,x)}$. Then:

- $\widetilde{g}$ is harmonic on $B \setminus \{x\}$.
- $\widetilde{g}$ satisfies the reproducing property $\mathcal{E}(\widetilde{g}, f) = f(x)$ for any function $f$ with $f|_{B^c} = 0$.
- $h$ is harmonic on $B \setminus \{x\}$, and satisfies $h(x) = 1$, $h|_{B^c} = 0$.
- $R_{\text{eff}}(x, B^c) = g_B(x, x)$.

*Proof.* Using the law of total probability, it can be proven that the transition density $q_j^B(x, y)$ satisfies the following property for any $x \in B$ and any $j \geq 1$:

$$\begin{aligned} q_j^B(x, y) = \frac{\mathbb{P}_x[X_j = y]}{\deg y} &= \frac{1}{\deg y} \sum_{\substack{w \in B \\ w \sim y}} \mathbb{P}_x[X_j = y \mid X_{j-1} = w] \mathbb{P}[X_{j-1} = w] \\ &= \frac{1}{\deg y} \sum_{\substack{w \in B \\ w \sim y}} \frac{\mathbb{P}[X_{j-1} = w]}{\deg w} = \frac{1}{\deg y} \sum_{\substack{w \in V \\ w \sim y}} q_{j-1}^B(x, w). \end{aligned}$$

In the final step, the summation can be changed from $w \in B$ to $w \in V$ since $q_{j-1}^B(x, w) = 0$ for all $w \in V \setminus V$. Note that $j \geq 1$ is used when applying the law of total probability.

This result can be applied to show that $\widetilde{g}$ is harmonic on $B \setminus \{x\}$. Let $y \in B \setminus \{x\}$. Then:

$$\begin{aligned}
\widetilde{g}(y) = g_B(x,y) &= \sum_{j=0}^{\infty} q_j^B(x,y) = \sum_{j=1}^{\infty} q_j^B(x,y) = \sum_{j=1}^{\infty}\left[\frac{1}{\deg y}\sum_{\substack{w\in V\\ w\sim y}} q_{j-1}^B(x,w)\right] \\
&= \frac{1}{\deg y}\sum_{\substack{w\in V\\ w\sim y}}\left[\sum_{j=1}^{\infty} q_{j-1}^B(x,w)\right] = \frac{1}{\deg y}\sum_{\substack{w\in V\\ w\sim y}}\left[\sum_{j=0}^{\infty} q_j^B(x,w)\right] = \frac{1}{\deg y}\sum_{\substack{w\in V\\ w\sim y}} \widetilde{g}(w),
\end{aligned}$$

where the lower limit of the summation can be changed from $j = 0$ to $j = 1$ since $q_0^B(x,y) = 0$ since $x \neq y$, and where we may interchange the summations since all quantities are positive. This shows that $\widetilde{g}$ is harmonic on $B \setminus \{x\}$, proving the first part of the lemma. For $\widetilde{g}(x)$, we have by the same computation that

$$\widetilde{g}(x) = g_B(x,x) = \sum_{j=0}^{\infty} q_j^B(x,x) = \frac{1}{\deg x} + \sum_{j=1}^{\infty} q_j^B(x,x) = \frac{1}{\deg x} + \frac{1}{\deg x}\sum_{\substack{w\in V\\ w\sim x}} \widetilde{g}(w),$$

since $q_0^B(x,x) = \dfrac{1}{\deg x}$. From this it follows that

$$\widetilde{g}(y) - \frac{1}{\deg y}\sum_{y\sim z}\widetilde{g}(z) = \frac{1}{\deg x}\mathbb{1}_{\{x=y\}}$$

for all $y \in B$. This can be used to show the reproducing property:

$$\begin{aligned}
\mathcal{E}(\widetilde{g}, f) &= \frac{1}{2}\sum_{y\sim z}(f(y)-f(z))(\widetilde{g}(y)-\widetilde{g}(z)) \\
&= \sum_{y\in V}\deg(y)f(y)\widetilde{g}(y) - \sum_{y\in V}\left[f(y)\sum_{y\sim z}\widetilde{g}(z)\right] = \sum_{y\in V}\deg(y)f(y)\left[\widetilde{g}(y) - \frac{1}{\deg y}\sum_{y\sim z}\widetilde{g}(z)\right] \\
&= \sum_{y\in B}\left[\deg(y)f(y)\frac{\mathbb{1}_{\{x=y\}}}{\deg(x)}\right] = \deg(x)f(x)\frac{1}{\deg(x)} = f(x).
\end{aligned}$$

In the first step, the sum is expanded and rewritten. There are $\deg(y)$ vertices $z$ such that $y \sim z$, which contribute a term $f(y)\widetilde{g}(y)$ and similarly $\deg(y)$ vertices $z$ such that $z \sim y$, which also contribute a term $f(y)\widetilde{g}(y)$, so in total we have a contribution of $\deg(y)f(y)\widetilde{g}(y)$, since the factor 2 cancels against the $\frac{1}{2}$ in front. For every pair $(y,z)$ with $y \sim z$, there are two terms of the form $f(y)\widetilde{g}(z)$. This yields the second term. This sum can now be written so that the previously shown equality can be used and since $f|_{B^c} = 0$, we can sum over $y \in V$ instead of $y \in B$. This proves the reproduction property.

The next part of the lemma considers the function $h(y) = \frac{g_B(x,y)}{g_B(x,x)}$, which is just $\widetilde{g}(y)$ divided by a constant. So it follows that $h$ is harmonic on $B \setminus \{x\}$ from the first part of the lemma. Moreover, $h(x) = \frac{g_B(x,x)}{g_B(x,x)} = 1$ and $h(y) = \frac{g_B(x,y)}{g_B(x,x)} = 0$ for $y \in B^c$ since $g_B(x,y) = 0$ for $y \in B^c$.

To show the final part of the lemma, first note that $h$ satisfies $h(x) = 1$, $h|_{B^c} = 0$ and that harmonic on $B \setminus \{x\}$. Since $B \setminus \{x\}$ is finite, it follows by the uniqueness principle that this is the only such function $h$. By applying Lemma 3.6 to $\{x\}$ and $B^c$, it follows that $h$ must therefore satisfy $R_{\text{eff}}(x, B^c)^{-1} = \mathcal{E}(h, h)$. To compute $\mathcal{E}(h, h)$, the reproducing property can be used. From this, it follows that $\mathcal{E}(\widetilde{g}, h) = h(x) = 1$. Hence, $\mathcal{E}(h, h) = (g_B(x, x))^{-1}$, since $h = (g_B(x, x))^{-1}\widetilde{g}$ and constants can just be pulled out of the energy quadratic form. It follows that $R_{\text{eff}}(x, B^c)^{-1} = \mathcal{E}(h, h) = (g_B(x, x))^{-1}$ and hence that $R_{\text{eff}}(x, B^c) = g_B(x, x)$. $\square$

The inequality $R_{\text{eff}}(x, B^c) = g_B(x, x)$ is very useful and is used in many places in Barlow, Peres and Sousi [3]. There is one more lemma on Green functions that is needed in [3].

**Lemma 3.8.** Let $g_t(y, x)$ be the Green kernel until time $t$. Then $g_t(y, x) \leq g_t(x, x)$.

*Proof.* To show this inequality, we condition on the first hitting time $\tau_x$ of $x$:

$$\begin{aligned} g_t(y, x) &= \sum_{j=0}^{t} q_j(y, x) = \sum_{j=0}^{t} \sum_{k=0}^{j} \mathbb{P}_y(\tau_x = k) q_{j-k}(x, x) = \sum_{\ell=0}^{t} \sum_{m=0}^{t-\ell} \mathbb{P}_y(\tau_x = m) q_\ell(x, x) \\ &= \sum_{\ell=0}^{t} \mathbb{P}_y(\tau_x \leq t - \ell) q_\ell(x, x) \leq \sum_{\ell=0}^{t} q_\ell(x, x) = g_t(x, x). \end{aligned}$$

In the second equality, the strong Markov property is used. The summations can be interchanged since both are finite. The inequality follows from $\mathbb{P}_y(\tau_x \leq t - \ell) \leq 1$ and $q_\ell(x, x) \geq 0$. $\square$

### 3.2.3 Spectral theory

As in the previous subsection, let $B$ be a finite subset of $V$. We may consider the restriction of the transition function $p$ to $B \times B$, which we simply denote by $p_B$. Note that $p_B(x, y) \geq 0$ for all $x, y \in V$ and that $\sum_{y \in B} p_B(x, y) \leq 1$ for all $x \in B$.

**Proposition 3.9.** (**Spectral decomposition**)
If $B$ is a finite subset of $V$, then $q_t^B$ can be decomposed as

$$q_t^B(x, y) = \sum_{i=1}^{|B|} \lambda_i^t \varphi_i(x) \varphi_i(y).$$

Moreover, the eigenvalues $\lambda_i$ are real and satisfy $|\lambda_i| \leq 1$ for all $1 \leq i \leq |B|$, and the eigenfunctions $\varphi_i(x)$ and $\varphi_i(y)$ are real-valued for all $1 \leq i \leq |B|$ and all $x, y \in B$.

This proposition is proven in Levin and Peres [8], Chapter 12, Lemma 12.1 and 12.2.

The spectral decomposition can be used to prove a number of useful inequalities.

**Lemma 3.9.** Assume that $G$ is locally finite and let $B$ be a (not necessarily finite) subset of $V$. Then the transition density $q_t^B$ satisfies the following inequalities:

- $q_{2t+1}^B(x,x) \leq q_{2t}^B(x,x)$ and $q_{2t+2}^B(x,x) \leq q_{2t}^B(x,x)$.
- $g_B(x,x) \geq \sum_{t=0}^{\infty} q_{2t}^B(x,x) \geq \frac{1}{2} g_B(x,x)$.

*Proof.* First consider the case where $B$ is finite. By the spectral decomposition, we can write

$$q_t^B(x,x) = \sum_{i=1}^{|B|} \lambda_i^t (\varphi_i(x))^2.$$

Since the eigenvalues $\lambda_i$ are real, it follows that $|\lambda_i|^2 = \lambda_i^2$, so $\lambda_i^{2t} = |\lambda_i|^{2t} \geq |\lambda_i|^{2t+1} \geq \lambda_i^{2t+1}$ and $\lambda_i^{2t} \geq \lambda_i^{2t+2}$. Since $(\varphi_i(x))^2$ is also nonnegative, this can be used to prove the following two inequalities:

$$q_{2t}^B(x,x) = \sum_{i=1}^{|B|} \lambda_i^{2t} (\varphi_i(x))^2 \geq \sum_{i=1}^{|B|} \lambda_i^{2t+1} (\varphi_i(x))^2 = q_{2t+1}^B(x,x),$$
$$q_{2t}^B(x,x) = \sum_{i=1}^{|B|} \lambda_i^{2t} (\varphi_i(x))^2 \geq \sum_{i=1}^{|B|} \lambda_i^{2t+2} (\varphi_i(x))^2 = q_{2t+2}^B(x,x).$$

This proves the inequalities in the case where $B$ is finite. Now let $B$ be arbitrary. Define $B(x,r) = \{y \in V : d(x,y) \leq r\}$. Since $G$ is locally finite, it follows that $B(x,r)$ is finite for all $x \in V$ and all finite $r \geq 0$. For a given $t$, define $B'(t) = B(x,t) \cap B$. Then we have $q_j^{B'(t)}(x,x) = q_j^B(x,x)$ for all $j \leq t$, since a random walk starting from $X_0 = x$ can never reach a vertex at graph distance larger than $t$. So by applying the inequalities to the finite set $B'(2t+2)$, the general case follows. This proves the first part of the lemma.

Applying the inequality $q_{2t+1}^B(x,x) \leq q_{2t}^B(x,x)$ now yields

$$g_B(x,x) = \sum_{t=0}^{\infty} q_t^B(x,x) \geq \sum_{t=0}^{\infty} q_{2t}^B(x,x) \geq \frac{1}{2} \sum_{t=0}^{\infty} \left( q_{2t}^B(x,x) + q_{2t+1}^B(x,x) \right) \geq \frac{1}{2} g_B(x,x),$$

so $g_B(x,x) \geq \sum_{t=0}^{\infty} q_{2t}^B(x,x) \geq \frac{1}{2} g_B(x,x)$ as required. □

### 3.3 Results from probability theory

In this section, a variety of results from probability theory is covered. The first proposition is very simple, but shows an ingenious way to show that the probability of some event is small. This proposition is used in Krishnapur and Peres [7].

**Proposition 3.10.** Let $X$ be a nonnegative random variable and let $A$ be an event such that $\mathbb{P}[A]$ and $\mathbb{E}[X|A]$ are positive. Then the following inequality holds:

$$\mathbb{P}[A] \leq \frac{\mathbb{E}[X]}{\mathbb{E}[X|A]}.$$

*Proof.* By the law of total expectation we have

$$\mathbb{E}[X] = \mathbb{P}[A]\mathbb{E}[X|A] + \mathbb{P}[A^c]\mathbb{E}[X|A^c] \geq \mathbb{P}[A]\mathbb{E}[X|A],$$

where the inequality holds since $X$ is nonnegative. Dividing by $\mathbb{E}[X|A]$ yields the result. $\square$

We now recall the geometric and the negative binomial distribution. Geo $(p)$ denotes the geometric distribution with success probability $p$, counting the number of failures. The probability mass function of the geometric distribution is given by $\mathbb{P}(G = k) = (1-p)^k p$ for $k \geq 0$. Its mean and variance are given by $\mathbb{E}[G] = \frac{1-p}{p}$ and $\mathrm{Var}[G] = \frac{1-p}{p^2}$.

The sum of $k$ independent geometrically distributed random variables with the same success probability $p$ has the negative binomial distribution with success probability $p$ and $k$ successes required. It follows immediately from the independence that if $\widetilde{G}$ has the negative binomial distribution, then its mean and variance are given by $\mathbb{E}[\widetilde{G}] = \frac{1-p}{p} \cdot k$ and $\mathrm{Var}[\widetilde{G}] = \frac{1-p}{p^2} \cdot k$.

If $(S_n)_{n\in\mathbb{N}}$ is a random walk on $\mathbb{Z}$ starting from $S_0 = 0$, then we can write $S_n = \sum_{j=1}^{n} B_j$ where the $B_j$ are i.i.d. random variables with $\mathbb{P}(B_j = 1) = \mathbb{P}(B_j = -1) = \frac{1}{2}$. The Central Limit Theorem yields that $\frac{S_n}{\sqrt{n}}$ converges in distribution to a normal distribution as $n \to \infty$. In particular, $\frac{S_n}{\sqrt{n}}$ can attain arbitrarily large numbers with positive probability as $n \to \infty$. On the other hand, the Law of Large Numbers states that $\frac{S_n}{n}$ converges to 0 almost surely. A question is whether there is a scaling between $\sqrt{n}$ and $n$, such that the resulting random variable does not converge to 0, but such that it also does not attain arbitrarily large numbers with positive probability. The law of the iterated logarithm provides this scaling:

**Proposition 3.11.** (**Law of the iterated logarithm**)
Let $(B_n)_{n\in\mathbb{N}}$ be a sequence of i.i.d. random variables with mean zero and variance 1.
Let $S_n = \sum_{j=1}^{n} B_j$. Then the following holds:

$$\limsup_{n\to\infty} \frac{\pm S_n}{\sqrt{2n \log\log n}} = 1 \quad \text{almost surely}$$

Here, the $\pm$ denotes that the law holds for both signs.

### 3.3.1 Inequalities for the random walk on $\mathbb{Z}$

The Chernoff bound is an inequality based on the moment-generating function which usually gives sharper bounds than Markov's inequality and Chebyshev's inequality. It can be proven using Markov's inequality.

**Proposition 3.12. (Chernoff bounds)**
Let $X$ be a random variable and let $t$ be a real number such that $\mathbb{E}\left[e^{tX}\right]$ exists. If $t > 0$, then $\mathbb{P}(X \geq a) \leq e^{-ta}\mathbb{E}\left[e^{tX}\right]$ for all $a \in \mathbb{R}$. If $t < 0$, then $\mathbb{P}(X \leq a) \leq e^{-ta}\mathbb{E}\left[e^{tX}\right]$ for all $a \in \mathbb{R}$.

In this subsection, some additional inequalities for the random walk on $\mathbb{Z}$ are proven. Using the Chernoff bound, we can now prove the following lemma:

**Lemma 3.10.** If $(S_n)_{n\in\mathbb{N}}$ is a random walk on $\mathbb{Z}$ starting from $S_0 = 0$, then $\mathbb{P}(S_n \geq \delta n) \leq e^{-\frac{1}{6}n\delta^2}$, $\mathbb{P}(S_n \leq -\delta n) \leq e^{-\frac{1}{6}n\delta^2}$ and $\mathbb{P}(|S_n| \geq \delta n) \leq 2e^{-\frac{1}{6}n\delta^2}$, for all $\delta > 0$.

*Proof.* Since $S_0 = 0$, we can write $S_n = \sum_{j=1}^n B_j$ where the $B_j$ are independent random variables with $\mathbb{P}(B_j = 1) = \mathbb{P}(B_j = -1) = \frac{1}{2}$. The moment-generating function of $B_j$ is

$$\mathbb{E}[e^{tB_j}] = e^t\mathbb{P}(B_j = 1) + e^{-t}\mathbb{P}(B_j = -1) = \tfrac{1}{2}\left(e^t + e^{-t}\right) = e^{-t}\left(1 + \tfrac{1}{2}\left(e^{2t} - 1\right)\right) \leq e^{-t}e^{\frac{1}{2}\left(e^{2t}-1\right)},$$

by the inequality $1 + x \leq e^x$. Hence, $\mathbb{E}[e^{tB_j}] \leq e^{\frac{1}{2}\left(e^{2t}-1\right)-t}$ for $1 \leq j \leq n$. Since $B_1, \ldots, B_n$ are independent, the same holds for $e^{tB_1}, \ldots, e^{tB_n}$. Hence,

$$\mathbb{E}\left[e^{tS_n}\right] = \mathbb{E}\left[\prod_{j=1}^n e^{tB_j}\right] = \prod_{j=1}^n \mathbb{E}\left[e^{tB_j}\right] \leq e^{\frac{1}{2}n\left(e^{2t}-1\right)-tn}.$$

By Chernoff's bound, it follows that $\mathbb{P}(S_n \geq \delta n) \leq e^{-tn\delta}\mathbb{E}\left[e^{tS_n}\right] \leq e^{\frac{1}{2}n\left(e^{2t}-1\right)-tn(1+\delta)}$ for $t > 0$. We choose $t = \frac{1}{2}\log(1+\delta)$, since this minimizes the right hand expression. Hence,

$$\mathbb{P}(S_n \geq \delta n) \leq e^{\frac{1}{2}n(\delta-(1+\delta)\log(1+\delta))}.$$

Let $g(x) = x\ln x$, then $g'(x) = 1 + \ln x$, so $g''(x) = \frac{1}{x}$ and hence $g'''(x) = -\frac{1}{x^2}$.
By the Taylor Remainder Theorem it hence follows that

$$g(1+\delta) = g(1) + \delta g'(1) + \tfrac{1}{2}\delta^2 g''(1) + \tfrac{1}{6}\delta^3 g'''(\xi), \qquad \text{where } \xi \in (1, 1+\delta).$$

Note that $\frac{1}{6}\delta^3 g'''(\xi) \geq \frac{1}{6}\delta^2 g'''(\xi) = \frac{1}{6}\delta^2 \cdot \left(-\frac{1}{\xi^2}\right) \geq -\frac{1}{6}\delta^2$, where the first inequality holds since $\delta^3 \leq \delta^2$ for $\delta \leq 1$ and the sign changes since $g'''(\xi)$ is negative. Hence,

$$g(1+\delta) \geq \delta + \tfrac{1}{2}\delta^2 - \tfrac{1}{6}\delta^2 = \delta + \tfrac{1}{3}\delta^2,$$

so $\delta - (1+\delta)\log(1+\delta) \leq -\frac{1}{3}\delta^2$. We conclude that $\mathbb{P}(S_n \geq \delta n) \leq e^{-\frac{1}{6}n\delta^2}$.

By symmetry considerations, we also have $\mathbb{P}(S_n \leq -\delta n) \leq e^{-\frac{1}{6}n\delta^2}$ for all $\delta > 0$.
It hence follows that $\mathbb{P}(|S_n| \geq \delta n) \leq 2e^{-\frac{1}{6}n\delta^2}$ for all $\delta > 0$. $\square$

If we substitute $\delta = n^{-1/2+\varepsilon}$, we obtain $\mathbb{P}(|S_n| \geq n^{1/2+\varepsilon}) \leq 2e^{-\frac{1}{6}n^{2\varepsilon}}$. So the probability that $|S_n| \geq n^{1/2+\varepsilon}$ decreases faster than any polynomial in $n$ for any $\varepsilon > 0$. In particular, it only happens finitely often, which is in accordance with the law of the iterated logarithm.

Lemma 3.10 can be used to prove several other bounds, especially in combination with Theorem 3.4, which states that $\mathbb{P}(\max\{Y_i : 0 \leq i \leq t\} \geq k) = \mathbb{P}(Y_t = k) + 2\mathbb{P}(Y_t > k) \leq 2\mathbb{P}(Y_t \geq k)$.

**Lemma 3.11.** Let $(Y_t)_{t\geq 0}$ be a random walk on $\mathbb{Z}$ starting from $Y_0 = 0$ and let $\tau$ be the first hitting time of $\{-k, k\}$. Then $\mathbb{P}(\tau \leq t) \leq 4e^{-\frac{1}{6}k^2/t}$.

*Proof.* If $\tau \leq t$, then there must exist an $1 \leq i \leq t$ such that $Y_i \leq -k$ or $Y_i \geq k$. By Theorem 3.4 and the analogous variant for the minimum of a random walk, it follows that

$$\begin{aligned}\mathbb{P}(\tau \leq t) &\leq \mathbb{P}(\min\{Y_i : 0 \leq i \leq t\} \leq -k) + \mathbb{P}(\max\{Y_i : 0 \leq i \leq t\} \geq k) \\ &\leq 2\mathbb{P}(Y_t \leq -k) + 2\mathbb{P}(Y_t \geq k) = 2\mathbb{P}(|Y_t| \geq k) \leq 4e^{-\frac{1}{6}k^2/t},\end{aligned}$$

where for the last inequality, Lemma 3.10 is used. $\square$

This bound can in turn be applied to get yet another bound, on the hitting times of a random walk which is restricted to an interval. This bound will be used in a slightly different setting.

**Lemma 3.12.** Let $(X_t)_{t\geq 0}$ be a random walk on $\mathbb{Z}\cap[0, n]$ starting from $X_0 = k$ for some $k \leq n$, and let $\tau$ be the first hitting time of 0. Then there exist constants $c_1, c_2 > 0$ not depending on $n$ or $k$ such that $\mathbb{P}(\tau \geq c_1k^2) \geq c_2$.

*Proof.* The simple random walk $X$ on $\mathbb{Z} \cap [0, n]$ can be generated from a simple random walk on $\mathbb{Z} \cap [0, 2n]$ by identifying $x$ and $2n - x$ with each other. We then consider the first hitting time of $\{0, 2n\}$ instead of the first hitting time of 0. Note that this hitting time stochastically dominates the first hitting time $\tau'$ of $\{0, 2k\}$, since $k \leq n$, so $\mathbb{P}(\tau \leq t) \leq \mathbb{P}(\tau' \leq t)$ for all $t \geq 0$. Before the first hitting time of $\{0, 2k\}$, the simple random walk on $\mathbb{Z} \cap [0, 2n]$ coincides with a simple random walk $(Y_t)_{t\geq 0}$ on $\mathbb{Z}$, starting from $Y_0 = k$. After shifting the random walk $k$ units to the left so that it starts from 0, Lemma 3.11 can be used. Hence,

$$\mathbb{P}(\tau \geq t) \geq P(\tau' \geq t) \geq 1 - 4e^{-\frac{1}{6}k^2/t}.$$

Choosing $t = c_1k^2$ for some constant $c_1 > 0$ yields the inequality $\mathbb{P}(\tau \geq c_1k^2) \geq 1 - 4e^{-1/(6c_1)}$. In particular, for every $c_1 < \frac{1}{6\ln 4}$ such a constant $c_2 = 1 - 4e^{-1/(6c_1)} > 0$ exists. $\square$

Finally, a stronger bound on the probability that $X_t = k$ is needed, especially in the cases where $k$ is larger than $\sqrt{t}$. This is a combination of the bound on $\mathbb{P}(\tau_k \leq t)$ from Lemma 3.11 and a bound on the probability that $X_t = k$ given by Proposition 3.4.

**Lemma 3.13.** Let $(X_t)_{t\geq 0}$ be a random walk on $\mathbb{Z}$ starting from $X_0 = 0$. Then there exists a constant $C > 0$ such that $\mathbb{P}(X_t = k) \leq \frac{C}{\sqrt{t}} e^{-\frac{1}{12}k^2/t}$.

*Proof.* If $k \leq \sqrt{t}$, then $e^{-\frac{1}{12}k^2/t} \geq e^{-\frac{1}{12}}$, so then the bound immediately follows from Proposition 3.4 after adjusting the constant $C$. Moreover, if $t/k \leq 12$ then the bound follows immediately from Lemma 3.10 by using the inequalities $e^{\frac{1}{12}k^2/t} \geq \frac{1}{12}k^2/t \geq \frac{1}{12^3}t$ and hence

$$\mathbb{P}(X_t = k) \leq \mathbb{P}(X_t \geq k) \leq 2e^{-\frac{1}{6}k^2/t} \leq C/t \cdot e^{-\frac{1}{12}k^2/t} \leq \frac{C}{\sqrt{t}} e^{-\frac{1}{12}k^2/t},$$

by choosing $C \geq 2 \cdot 12^3$. We now write

$$\mathbb{P}(X_t = k) = \frac{\mathbb{P}(X_t = k)}{\mathbb{P}(X_t \geq k)} \mathbb{P}(X_t \geq k).$$

Note that $\mathbb{P}(X_t \geq k) \leq 2e^{-\frac{1}{6}k^2/t}$ by Lemma 3.10.
By the inequalities $e^{\frac{1}{12}k^2/t} \geq 1 + \frac{1}{12}\frac{k^2}{t} \geq 2\sqrt{\frac{1}{12}\frac{k^2}{t}} \geq \frac{1}{2}\frac{k}{\sqrt{t}}$, it follows that $\mathbb{P}(X_t \geq k) \leq \frac{\sqrt{t}}{k} e^{-\frac{1}{12}k^2/t}$.

We now bound $\dfrac{\mathbb{P}(X_t = k)}{\mathbb{P}(X_t \geq k)}$. First note that for $a \geq 0$ we have

$$\mathbb{P}(X_t = k + 2a) = \frac{1}{2^t}\binom{t}{\frac{t+k+2a}{2}} = \frac{1}{2^t}\binom{t}{\frac{t+k}{2}} \prod_{j=1}^{a} \frac{\frac{t-k}{2} - j}{\frac{t+k}{2} + j} \geq \left(\frac{\frac{t-k}{2} - a}{\frac{t+k}{2} + a}\right)^a \mathbb{P}(X_t = k).$$

Since we can assume $t/k \geq 12$ it follows that $\left(1 - \frac{6k}{t+3k}\right)^{t/k} \geq \left(1 - \frac{6}{t/k}\right)^{t/k} \geq \frac{1}{4^6}$.

Summing this for $1 \leq a \leq t/k$ and using that $\left(\frac{\frac{t-k}{2} - a}{\frac{t+k}{2} + a}\right)^a = \left(\frac{t-k-2a}{t+k+2a}\right)^a$ is decreasing in $a$ and that $\left(\frac{t-k-2t/k}{t+k+2t/k}\right)^{t/k} \geq \left(\frac{t-3k}{t+3k}\right)^{t/k} = \left(1 - \frac{6}{t/k}\right)^{t/k}$ since $t/k \leq k$ since $\sqrt{t} \leq k$, it follows that

$$\mathbb{P}(X_t \geq k) \geq \sum_{a=1}^{t/k} \left(\frac{\frac{t-k}{2} - a}{\frac{t+k}{2} + a}\right)^a \mathbb{P}(X_t = k) \geq \frac{t}{k}\left(1 - \frac{6k}{t+3k}\right)^{t/k} \mathbb{P}(X_t = k) \geq \frac{1}{4^6}\frac{t}{k}\mathbb{P}(X_t = k).$$

Combining this yields $\mathbb{P}(X_t = k) \leq \frac{4^6}{\sqrt{t}} e^{-\frac{1}{12}k^2/t}$, for $k \geq \sqrt{t}$ and $t/k \geq 12$, so by choosing $C \geq 4^6$, the bound follows. Hence, the bound holds in all cases, which completes the proof. $\square$

### 3.3.2 Zero-one laws

When considering infinitely many events, one can be interested whether only finitely many or infinitely many occur. Often, it can be proven that the event occurs infinitely often (i.o.) with probability 1, and finitely often with probability 0, or vice versa. The simplest such result is the Borel-Cantelli lemma, which can be used to show that almost surely finitely many events occur:

**Proposition 3.13. (Borel-Cantelli lemma)**
Let $(A_n)_{n\in\mathbb{N}}$ be a sequence of events. If $\sum_{k=0}^{\infty}\mathbb{P}(A_k)$ is finite, then $\mathbb{P}(A_k \text{ occurs i.o.}) = 0$.

For more advanced 0-1 laws, some definitions from measure theory are needed. The following theorems are from Durrett [6]. A $\sigma$-algebra $\mathcal{T}$ is called **trivial** if it only contains events that happen with probability 0 or 1, i.e. $\mathbb{P}(A) \in \{0,1\}$ for all $A \in \mathcal{T}$. Now consider a probability space where $\Omega = S^{\mathbb{N}}$. Then elements of $\Omega$ are sequences of random variables taking values in the state space. A finite permutation of $\mathbb{N}$ is a bijection $\pi : \mathbb{N} \to \mathbb{N}$ such that $\pi(i) \neq i$ for only finitely many $i$. For a finite permutation $\pi$, define $\pi\omega$ by $(\pi\omega_i) = \omega_{\pi(i)}$. Under the interpretation that $X_i(\omega) = \omega_i$, this is the same as rearranging the random variables. An event $A$ is called permutable if $\pi^{-1}A = A$ for all finite permutations $\pi$, where $\pi^{-1}A = \{\omega : \pi\omega \in A\}$. The collection of permutable events is called the **exchangeable** $\sigma$-field $\mathcal{T}_E$. The notion of a permutable event allows us to formulate the following 0-1 law:

**Proposition 3.14. (Hewitt-Savage 0-1 law)**
Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of i.i.d. random variables. If the event $A$ is in the exchangeable $\sigma$-field of the sequence $(X_n)_{n\in\mathbb{N}}$, then $\mathbb{P}(A) \in \{0,1\}$.

The Hewitt-Savage 0-1 law can be used to prove a 0-1 law for Markov chains. For a Markov chain, define $\mathcal{F}'_n = \sigma(X_{n+1}, X_{n+2}, \ldots)$ and $\mathcal{T} = \bigcap_{n=0}^{\infty}\mathcal{F}'_n$. An event in $\mathcal{T}$ is called a **tail event**: it only depends on the tail of the sequence $(X_n)_{n\in\mathbb{N}}$, not on any number of initial terms. The difficulty in proving this lays in the fact that the random variables $(X_n)_{n\in\mathbb{N}}$ are not independent. So we break up the Markov chain in vectors that are i.i.d., so that the Hewitt-Savage 0-1 law can be applied.

**Theorem 3.5. (Orey)**
Let $(X_n)_{n\geq 0}$ be a recurrent Markov chain with $\mathbb{P}(X_0 = x) = 1$. Then $\mathcal{T}$ is trivial.

*Proof.* Define the $m$th hitting time of $x$ by $T_0 = 0$ and $T_m = \inf\{n > T_{m-1} : X_n = x\}$ for $m \geq 1$. Since the Markov chain is recurrent, $T_m$ is almost surely finite for all $m$. Consider the vector $V_n = (X(T_{n-1}), \ldots, X(T_n - 1))$. Note that these are vectors of variable length. By the strong Markov property, these vectors are independent.

Since finite permutations of the $V_i$ only affect a finite number of initial terms of the sequence $(X_n)_{n\in\mathbb{N}}$, it follows that any event contained in the tail field $\mathcal{T}$ is contained in the exchangeable $\sigma$-field of the sequence $(V_n)_{n\in\mathbb{N}}$. By the Hewitt-Savage 0-1 law, the exchangeable $\sigma$-field of the sequence $(V_n)_{n\in\mathbb{N}}$ is trivial, so the same holds for the tail field $\mathcal{T}$. $\square$

## 3.4 Results from martingale theory

Recall from Section 3.1.2 that for a sequence of random variables $(X_n)_{n\geq 0}$, the natural filtration $(\mathcal{F}_n)_{n\geq 0}$ is defined as the $\sigma$-algebra generated by the random variables $X_0, \dots, X_n$. Note that $\mathcal{F}_n$ is the collection of events that can be distinguished up to time $n$. In the context of martingale theory, $\mathcal{F}_n$ is also called the information set. The following exposition is based on lecture notes by Balázs [2]. We only consider martingales with respect to the natural filtration.

**Definition 3.5.** (**Martingale**)
A stochastic process $(X_n)_{n\geq 0}$ in a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is a martingale with respect to the natural filtration $(\mathcal{F}_n)_{n\geq 0}$ if the following two properties hold:

- $\mathbb{E}[|X_n|] < \infty$ for all $n \geq 0$.
- $\mathbb{E}[X_{n+1}|\mathcal{F}_n] = X_n$ almost surely, for all $n \geq 0$.

From the law of iterated expectations it follows that $\mathbb{E}[X_{n+1}] = \mathbb{E}[\mathbb{E}[X_{n+1}|\mathcal{F}_n]] = \mathbb{E}[X_n]$ for all $n \geq 0$. By induction, it hence follows that $\mathbb{E}[X_n] = \mathbb{E}[X_0]$ for all $n \geq 0$.

As stated, this only holds for fixed $n$, but under some conditions this can be extended to the case where $n$ is a stopping time $\tau$, which means that $\{\tau \leq n\} \in \mathcal{F}_n$ for all $n \geq 0$. If $(\mathcal{F}_n)_{n\geq 0}$ is the natural filtration, then this means that whether the event $\{\tau \leq n\}$ occurs can be decided by just knowing the values $X_0, \dots, X_n$. This yields Doob's optional stopping theorem:

**Theorem 3.6.** (**Doob's optional stopping theorem**)
Let $(X_n)_{n\geq 0}$ be a martingale and let $\tau$ be a stopping time. If $X$ is of bounded increments (i.e. there exists a constant $C$ such that $|X_{n+1} - X_n| \leq C$ for all $n \geq 1$) and $\mathbb{E}[\tau] < \infty$, then $\mathbb{E}[X_\tau] = \mathbb{E}[X_0]$.

It may be interesting to note the analogy to Section 3.1.2: Doob's optional stopping theorem is to the property that $\mathbb{E}[X_n] = \mathbb{E}[X_0]$ as the strong Markov property is to the Markov property.

We now consider a segment of $\mathbb{Z}$ from 0 to $B$ for some $B > 0$ and a simple random walk $(X_n)_{n\geq 0}$ starting from $X_0 = 1$. Let $\tau$ be the first time that either 0 or $B$ is reached. Note that $\{\tau \leq n\} \in \mathcal{F}_n$ since we can decide whether $\{\tau \leq n\}$ occurs by simply checking whether one of the realizations of $X_0, \dots, X_n$ is equal to 0 or $B$.

We now calculate the expectation of $\tau$ using two martingales.

**Lemma 3.14.** Consider a segment of $\mathbb{Z}$ from 0 to $B$ for some $B > 0$ and a simple random walk $(X_n)_{n\geq 0}$ starting from $X_0 = 1$. Let $\tau$ be the first time that 0 or $B$ is reached. Then $\mathbb{E}[\tau] = B - 1$.

*Proof.* Let us first prove that $(X_n)_{n\geq 0}$ is a martingale. Note that $|X_n| \leq n+1$ for all $n \in \mathbb{N}$ and hence $\mathbb{E}[|X_n|] < \infty$ for all $n \geq 0$. Also, we have

$$\mathbb{E}[X_{n+1}|\mathcal{F}_n] = \mathbb{E}[X_{n+1}|X_n] = \tfrac{1}{2}(X_n+1) + \tfrac{1}{2}(X_n-1) = X_n,$$

since $(X_n)_{n\geq 0}$ is a Markov chain. So $(X_n)_{n\geq 0}$ is a martingale. Note that $(X_n)$ is of bounded increments, in fact $|X_{n+1} - X_n| = 1$ for all $n \geq 1$.

Note that from any given point between 0 and $B$, doing $B$ steps in the same direction means that we cross 0 or $B$. We can see such a successive block of $B$ steps as a trial and doing $B$ steps in the same direction as success. From this, we see that $\tau$ is upper bounded by $B$ times a geometric random variable with success probability $1/2^{B-1}$, so we have $\mathbb{E}[\tau] \leq B2^{B-1} < \infty$.

Since $(X_n)_{n\geq 0}$ is a martingale of bounded increments and $\mathbb{E}[\tau] < \infty$, Doob's optional stopping theorem applies. Hence, $\mathbb{E}[X_\tau] = \mathbb{E}[X_0] = 1$. By definition, $X_\tau$ can only be 0 or $B$, so this implies $B \cdot \mathbb{P}(X_\tau = B) = \mathbb{E}[X_\tau] = 1$, so $\mathbb{P}(X_\tau = B) = \frac{1}{B}$ and $\mathbb{P}(X_\tau = 0) = \frac{B-1}{B}$.

Let $Y_n = X_n^2 - n$. We prove that $(Y_n)_{n\geq 0}$ is a martingale. Note that $|Y_n| \leq |X_n|^2 + n \leq (n+1)^2 + n$ for all $n \in \mathbb{N}$ and hence $\mathbb{E}[|Y_n|] < \infty$ for all $n \geq 0$. Also, we have

$$\mathbb{E}[Y_{n+1}|\mathcal{F}_n] = \mathbb{E}[Y_{n+1}|X_n] = \tfrac{1}{2}(X_n+1)^2 + \tfrac{1}{2}(X_n-1)^2 - (n+1) = X_n^2 + 1 - (n+1) = Y_n,$$

since $(X_n)_{n\geq 0}$ is a Markov chain. So $(Y_n)_{n\geq 0}$ is a martingale. Note that before stopping we have $0 \leq X_n \leq B$, so $|Y_{n+1} - Y_n| \leq 2B$ for all $n \geq 1$.

Since $(Y_n)_{n\geq 0}$ is a martingale of bounded increments and $\mathbb{E}[\tau] < \infty$, Doob's optional stopping theorem applies. Hence, $E[Y_\tau] = \mathbb{E}[Y_0] = 1$. Note that $\mathbb{E}[X_\tau^2] = B^2 \cdot \mathbb{P}(X_\tau = B) = B$, so

$$\mathbb{E}[\tau] = \mathbb{E}[X_\tau^2 - Y_\tau] = E[X_\tau^2] - \mathbb{E}[Y_\tau] = B - 1,$$

which is the required result. □

By symmetry considerations, this lemma also holds when $X_0 = B - 1$. From this we also see that the first return time to 0 or $B$, when starting from 0, has expectation $B$.

We now consider the situation where the random walk is confined to a two-sided segment $[-A, B]$.

**Lemma 3.15.** Consider a segment of $\mathbb{Z}$ from $-A$ to $B$ for some $A, B > 0$ and a simple random walk $(X_n)_{n\geq 0}$ starting from $X_0 = 0$. Then:

1. Let $\tau$ be the first return time to $-A, 0$ or $B$.
   Then $\mathbb{P}(X_\tau = -A) = \frac{1}{2A}$, $\mathbb{P}(X_\tau = B) = \frac{1}{2B}$ and $\mathbb{P}(X_\tau = 0) = \frac{1}{2}\left(\frac{A-1}{A} + \frac{B-1}{B}\right)$.
2. Let $\tau'$ be the hitting time of $\{-A, B\}$. Then $\mathbb{P}(X_{\tau'} = -A) = \frac{B}{A+B}$ and $\mathbb{P}(X_{\tau'} = B) = \frac{A}{A+B}$. Moreover, the number of visits to 0 excluding the first before $\tau'$ is geometrically distributed with success probability $\frac{1}{2}\left(\frac{1}{A} + \frac{1}{B}\right)$.

*Proof.* If $X_1 = 1$, then the random walk can either return to 0 or $B$, but not to $-A$. As shown in the proof of Lemma 3.14, the random walk then returns to 0 with probability $\frac{B-1}{B}$ and to $B$ with probability $\frac{1}{B}$. If $X_1 = -1$, then the random walk can either return to $-A$ or 0, but not to $B$. By symmetry considerations, the proof of Lemma 3.14 in this case implies that the random walk then returns to 0 with probability $\frac{A-1}{A}$ and to $-A$ with probability $\frac{1}{A}$. By the law of total probability it hence follows that

$$\begin{aligned}
\mathbb{P}(X_\tau = -A) &= \mathbb{P}(X_\tau = -A \mid X_1 = -1)\mathbb{P}(X_1 = -1) + \mathbb{P}(X_\tau = -A \mid X_1 = 1)\mathbb{P}(X_1 = 1) = \tfrac{1}{2A} \\
\mathbb{P}(X_\tau = 0) &= \mathbb{P}(X_\tau = 0 \mid X_1 = -1)\mathbb{P}(X_1 = -1) + \mathbb{P}(X_\tau = 0 \mid X_1 = 1)\mathbb{P}(X_1 = 1) = \tfrac{1}{2}\left(\tfrac{A-1}{A} + \tfrac{B-1}{B}\right) \\
\mathbb{P}(X_\tau = B) &= \mathbb{P}(X_\tau = B \mid X_1 = -1)\mathbb{P}(X_1 = -1) + \mathbb{P}(X_\tau = B \mid X_1 = 1)\mathbb{P}(X_1 = 1) = \tfrac{1}{2B},
\end{aligned}$$

which proves the first part of the lemma.

We now turn to the second part. Let us do experiments where a success is defined when $-A$ or $B$ is reached before 0 and a failure is defined as reaching 0 before $-A$ or $B$. If 0 is reached, then a new experiment is started. Since $\tau'$ is a stopping time, it follows by the strong Markov property that the experiments are independent. Moreover, by the first part of the lemma it follows that the success probability is $\frac{1}{2}\left(\frac{1}{A} + \frac{1}{B}\right)$. Hence the number of visits to 0 excluding the first before $\tau'$ is geometrically distributed with success probability $\frac{1}{2}\left(\frac{1}{A} + \frac{1}{B}\right)$.

To calculate the probabilities of reaching $-A$ and $B$, we again use the fact that $(X_n)_{n\geq 0}$ is a martingale of bounded increments, as shown in the proof of Lemma 3.14. Moreover, it follows from a similar argument as in this lemma that $\mathbb{E}[\tau'] < \infty$. Hence, Doob's optional stopping theorem applies, so $\mathbb{E}[X_{\tau'}] = \mathbb{E}[X_0] = 0$. By definition, $X_{\tau'}$ can only be $-A$ or $B$, so this implies $B \cdot \mathbb{P}(X_{\tau'} = B) - A \cdot \mathbb{P}(X_{\tau'} = -A) = \mathbb{E}[X_{\tau'}] = 0$ and $\mathbb{P}(X_{\tau'} = -A) + \mathbb{P}(X_{\tau'} = B) = 1$. Solving the system of equations implies $\mathbb{P}(X_{\tau'} = -A) = \frac{B}{A+B}$ and $\mathbb{P}(X_{\tau'} = B) = \frac{A}{A+B}$. $\square$

# 4 Exposition of known results

## 4.1 Results on collisions of random walks

In this section, some preliminary results on the collisions of random walks are given. The following proposition is proven in Krishnapur and Peres [7].

**Proposition 4.1.** Let $G = (V, E)$ be an infinite graph of bounded degree.
Then the expected number of collisions is finite if and only if $G$ is transient.

*Proof.* Let $X$ and $Y$ be two independent simple random walks on $G$ starting from a vertex $v$. The expected number of meetings of $X$ and $Y$ is given by $\sum_{n=0}^{\infty} \sum_{v \in V} (p^{(n)}(v, w))^2$. By Lemma 3.4, it holds that $q_n(v, w) = q_n(w, v)$ and hence $p^{(n)}(w, v) = p^{(n)}(v, w) \frac{\deg v}{\deg w}$. Hence, we have

$$\sum_{n=0}^{\infty} p^{(2n)}(v, v) = \sum_{n=0}^{\infty} \sum_{w \in V} p^{(n)}(v, w) p^{(n)}(w, v) = \sum_{n=0}^{\infty} \sum_{w \in V} (p^{(n)}(v, w))^2 \frac{\deg v}{\deg w}.$$

Since $G$ is of bounded degree, $\frac{\deg v}{\deg w}$ is bounded, so it follows that $\sum_{n=0}^{\infty} p^{(2n)}(v, v)$ is finite if and only if $\sum_{n=0}^{\infty} \sum_{v \in V} (p^{(n)}(v, w))^2$ is finite. By the second part of Lemma 3.9, it follows that $\sum_{n=0}^{\infty} p^{(2n)}(v, v)$ is finite if and only if $\sum_{n=0}^{\infty} p^{(n)}(v, v)$ is finite.

So $\sum_{n=0}^{\infty} \sum_{v \in V} (p^{(n)}(v, w))^2$ is finite if and only if $\sum_{n=0}^{\infty} p^{(n)}(v, v)$ is finite. The latter holds if and only if $v$ is transient and by Proposition 3.1 this holds if and only if $G$ is transient. Hence, the expected number of collisions is finite if and only if $G$ is transient. $\square$

This proposition can also be formulated as: the expected number of collisions is infinite if and only if $G$ is recurrent. However, the fact the expectation of a random variable is infinite, does not mean that it can be infinite with positive probability. Of course, the converse does hold: if a nonnegative random variable is infinite with positive probability, then its expectation is infinite.

The aim in Krishnapur and Peres [7] is to give a recurrent graph which has the finite collision property. Therefore, we first prove that $\mathrm{Comb}(\mathbb{Z})$ is recurrent.

**Proposition 4.2.** $\mathrm{Comb}(\mathbb{Z})$ is recurrent.

*Proof.* Consider a random walk $X$ on $\mathrm{Comb}(\mathbb{Z})$ starting from $X_0 = (0, 0)$. Let $U$ be the random walk corresponding to the horizontal steps of $X$, and let $V$ be the random walk corresponding to the vertical steps of $X$. Then $U$ and $V$ are both simple random walks on $\mathbb{Z}$. Since $\mathbb{Z}$ is recurrent, it follows that $U$ visits 0 infinitely often almost surely if $X$ makes infinitely many horizontal steps. Since $X$ can only make a horizontal step if the vertical position is 0, it then follows that $X$ visits $(0, 0)$ infinitely often. Now assume for the sake of contradiction that $X$

makes only finitely many horizontal steps with positive probability. Then $X$ does make infinitely many vertical steps. Since $V$ is a simple random walks on $\mathbb{Z}$, it follows that $V$ visits 0 infinitely often almost surely. At each such visit, $X$ makes a horizontal transition with probability $\frac{1}{2}$, so then $X$ does make infinitely many horizontal steps almost surely. Contradiction. Hence, $(0,0)$ is recurrent. By Proposition 3.1, it follows that $\mathrm{Comb}(\mathbb{Z})$ is recurrent. □

The following proposition is a generalization of the result used in Krishnapur and Peres [7] to bound the probability that two simple random walks collide, at two (possibly different) given points in time. The argument is very similar to the proof of Lemma 3.4.

**Proposition 4.3.** Let $G = (V, E)$ be an infinite graph of bounded degree. Consider two independent simple random walks $U$ and $U'$ on $G$, starting from the same vertex: $U_0 = U'_0 = v$. Then there exists a constant $C > 0$ such that

$$\mathbb{P}_{v,v}[U_i = U'_j] \leq \frac{C}{\sqrt{i+j}}$$

for all $i$ and $j$ such that $i + j > 0$.

*Proof.* Consider two paths $\mathcal{P} = (p_0, \dots, p_i)$ and $\mathcal{P}' = (p'_0, \dots, p'_j)$ of lengths $i$ and $j$ in $G$ starting from $p_0 = p'_0 = v$ and having the same endpoint $p_i = p'_j = w$. Let $\mathcal{P}'' = (p_0, \dots, p_i, p'_{j-1}, \dots, p'_0)$ be the path obtained by traversing $\mathcal{P}$ first and then returning to $v$ via $\mathcal{P}'$ in reverse. Then:

$$\begin{aligned}\mathbb{P}_v[\{U_k\}_{k\leq i+j} = \mathcal{P}''] &= \prod_{k=0}^{i} \frac{1}{\deg(p_k)} \prod_{k=1}^{j-1} \frac{1}{\deg(p'_k)} = \frac{\deg(p'_0)}{\deg(p_i)} \prod_{k=0}^{i-1} \frac{1}{\deg(p_i)} \prod_{k=0}^{j-1} \frac{1}{\deg(p'_{j-1})} \\ &= \frac{\deg(v)}{\deg(w)} \mathbb{P}_v[\{U_k\}_{k\leq i} = \mathcal{P}] \mathbb{P}_v[\{U'_k\}_{k\leq j} = \mathcal{P}'].\end{aligned}$$

Since $G$ is by assumption of bounded degree, we have $\frac{\deg(v)}{\deg(w)} \geq \frac{1}{b}$ for some $b > 0$. Note that given a path $\mathcal{P}''$ of length $i + j$ starting and ending at $v$, we can also make paths $\mathcal{P}$ and $\mathcal{P}'$ of length $i$ and $j$ by taking $\mathcal{P}$ to be the first $i + 1$ vertices of $\mathcal{P}''$ and $\mathcal{P}'$ to be the reverse of the last $j + 1$ vertices. Hence, there exists a bijection between the paths $\mathcal{P}''$ from $v$ to $v$ of length $i + j$ and the paths $\mathcal{P}, \mathcal{P}'$ of length $i$ and $j$ from $v$ to $w$.

By summing over all $w$ and all possible paths $\mathcal{P}, \mathcal{P}'$, we hence find that

$$\mathbb{P}_v[U_{i+j} = v] \geq \frac{1}{b} \mathbb{P}_{v,v}[U_i = U'_j].$$

By Proposition 3.4, there exists a $C' > 0$ such that $\mathbb{P}_v[U_{i+j} = v] = p^{(i+j)}(v, v) \leq \frac{C'}{\sqrt{i+j}}$. Multiplying by $b$ and letting $C = C'b$ yields

$$\mathbb{P}_{v,v}[U_i = U'_j] \leq \frac{C}{\sqrt{i+j}}$$

for all $i$ and $j$ such that $i + j > 0$. □

For two simple random walks on $\mathbb{Z}$, the reverse inequality also holds (with a different constant):

**Proposition 4.4.** Consider two independent simple random walks $U$ and $U'$ on $\mathbb{Z}$, starting from $U_0 = U'_0 = 0$. Then there exists a constant $C' > 0$ such that

$$\mathbb{P}_{0,0}[U_i = U'_j] \geq \frac{C'}{\sqrt{i+j}}$$

for all $i$ and $j$ such that $i + j$ is even and larger than 0.

*Proof.* Note that all vertices of $\mathbb{Z}$ have degree 2, so $\frac{\deg(v)}{\deg(w)} = 1$ for all vertices $v$ and $w$ of $\mathbb{Z}$. Using the same argument as in Proposition 4.3, it now follows that $\mathbb{P}_0[U_{i+j} = 0] = \mathbb{P}_{0,0}[U_i = U'_j]$.

Since $U$ is a random walk on $\mathbb{Z}$, we can write $U_n = \sum_{k=1}^{n} B_k$ where the $B_k$ are i.i.d. random variables with $\mathbb{P}(B_k = 1) = \mathbb{P}(B_k = -1) = \frac{1}{2}$. Hence, we have $U_n = 0$ if and only if $\mathbb{P}(B_k = 1)$ for exactly $\frac{n}{2}$ values of $k$ and hence also $\mathbb{P}(B_k = -1)$ for exactly $\frac{n}{2}$ values of $k$.

Let $h = \frac{1}{2}(i + j)$, which is an integer by assumption. Then there exist constants $C, C'$ such that

$$\mathbb{P}_{0,0}[U_i = U'_j] = \mathbb{P}_0[U_{i+j} = 0] = \binom{2h}{h}\frac{1}{2^{2h}} \geq \frac{C}{\sqrt{h}} = \frac{C'}{\sqrt{i+j}},$$

which is the required inequality. □

Finally, we use Orey's theorem (Theorem 3.5), to prove a 0-1 law for the finite collision property. This proposition can be found in Barlow, Peres and Sousi [3].

**Proposition 4.5. (0-1 law)**
Let $G$ be a connected recurrent graph and let $X$ and $Y$ be independent random walks on $G$. Let $Z$ be the number of collisions of $X$ and $Y$. Then for all $(a, b) \in G \times G$ we have

$$\mathbb{P}_{a,b}(Z = \infty) \in \{0, 1\}.$$

If there exist $a_0, b_0$ such that $\mathbb{P}_{a_0,b_0}(Z = \infty) > 0$ then $\mathbb{P}_{a,b}(Z = \infty) = 1$ for all $a, b$ such that there exists an $m$ with $\mathbb{P}_{a,b}(X_m = a_0, Y_m = b_0) > 0$. In particular, either $\mathbb{P}_{a,a}(Z = \infty) = 0$ for all $a$ or else $\mathbb{P}_{a,a}(Z = \infty) = 1$ for all $a$.

*Proof.* Let $\mathcal{T}_n^X = \sigma(X_{n+1}, X_{n+2}, \ldots)$ and $\mathcal{T}^X = \bigcap_{n=0}^{\infty} \mathcal{T}_n^X$ and similarly $\mathcal{T}_n^Y = \sigma(Y_{n+1}, Y_{n+2}, \ldots)$ and $\mathcal{T}^Y = \bigcap_{n=0}^{\infty} \mathcal{T}_n^Y$. By Orey's theorem (Theorem 3.5), it follows that $\mathcal{T}^X$ and $\mathcal{T}^Y$ are trivial. Define $\mathcal{T} = \bigcap_{n=0}^{\infty} \sigma\left(\mathcal{T}_n^X, \mathcal{T}_n^Y\right)$. By Lemma 2 of Lindvall and Rogers [9], it now follows that $\mathcal{T} = \sigma\left(\mathcal{T}^X, \mathcal{T}^Y\right)$ since $X$ and $Y$ are independent and hence that $\mathcal{T}$ is trivial since $\mathcal{T}^X$ and $\mathcal{T}^Y$ are trivial. Since the event $\{X_t = Y_t \text{ i.o.}\}$ is $\sigma\left(\mathcal{T}_n^X, \mathcal{T}_n^Y\right)$-measurable for all $n \in \mathbb{N}$, it follows that it is $\mathcal{T}$-measurable. Hence, the event $\{X_t = Y_t \text{ i.o.}\}$ has probability 0 or 1.

Now assume that $\mathbb{P}_{a_0,b_0}(Z=\infty)>0$. By the 0-1 law, it then follows that $\mathbb{P}_{a_0,b_0}(Z=\infty)=1$. Let $a,b$ and $m$ such that $\mathbb{P}_{a,b}(X_m=a_0,Y_m=b_0)\geq 0$ be given. Then

$$\mathbb{P}_{a_0,b_0}(Z=\infty)\geq \mathbb{P}(Z=\infty \mid X_m=a_0,Y_m=b_0)\mathbb{P}_{a,b}(X_m=a_0,Y_m=b_0)>0,$$

and by the 0-1 law it hence follows that $\mathbb{P}_{a,b}(Z=\infty)=1$. Since the graph is connected, there exists an $m$ such that $\mathbb{P}_{a',a'}(X_m=a,Y_m=a)=\mathbb{P}_{a'}(X_m=a)^2>0$. Hence, applying this with $a_0=b_0$ and $a=b$ yields that $\mathbb{P}_{a,a}(Z=\infty)=0$ for all $a$ or else $\mathbb{P}_{a,a}(Z=\infty)=1$ for all $a$. □

From the 0-1 law, it follows that for a connected recurrent graph there are only two possibilities: either $\mathbb{P}_{a,a}(Z=\infty)=0$ for all $a$ or else $\mathbb{P}_{a,a}(Z=\infty)=1$ for all $a$. So when both random walks start from the same vertex, they collide either infinitely often almost surely or finitely often almost surely. We define these two properties as the infinite collision property and the finite collision property, respectively.

**Definition 4.1.** (**The finite and infinite collision property**)
If $\mathbb{P}_{a,a}(Z=\infty)=1$ for all $a\in G$, then $G$ has the **infinite collision property**.
If $\mathbb{P}_{a,a}(Z=\infty)=0$ for all $a\in G$, then $G$ has the **finite collision property**.

We now give a corollary of the 0-1 law.

**Corollary 4.1.** Let $(B_n)_{n\in\mathbb{N}}$ be a collection of finite subsets of $G$ and let

$$Z(B_n)=\sum_{t=0}^{\infty}\mathbb{1}(X_t=Y_t\in B_n)$$

be the number of collisions in $B_n$.
Let $A_n$ be the event that $Z(B_n)$ is positive, i.e. $A_n=\{Z(A_n)>0\}$. Then

1. If $G=\bigcup_{n\in\mathbb{N}}B_n$ and $\mathbb{P}(A_n \text{ occurs i.o.})=0$, then $G$ has the finite collision property.
2. If the $A_n$ are disjoint and $\mathbb{P}(A_n)>c>0$ for all $n$, then $G$ has the infinite collision property.

*Proof.* 1. Note that we can view the two random walks $X$ and $Y$ as one random walk on the product graph $G\times G$. If $G\times G$ is recurrent, then the random walk on the product graph visits every vertex infinitely often and hence $X$ and $Y$ collide infinitely often at every vertex of $G$, which implies $\mathbb{P}(A_n \text{ occurs i.o.})=1$. So $G\times G$ is transient, which implies that the number of collisions in the set $B_n$ is finite almost surely for all $n\in\mathbb{N}$. Since there are only finitely many sets $B_n$ with a positive amount of collisions, it follows that there are only finitely many collisions in total and hence that $G$ has the finite collision property.

2. Note that $\mathbb{P}(A_n)>c$ implies that $\mathbb{P}(A_n \text{ occurs i.o.})>c$. Since the $A_n$ are disjoint, we have $Z\geq \mathbb{1}(A_n)$, so $\mathbb{P}(Z=\infty)>c$ and hence by the 0-1 law it follows that $\mathbb{P}(Z=\infty)=1$. □

## 4.2 The finite collision property of $\mathrm{Comb}(\mathbb{Z})$

This section is devoted to the exposition of the proof of the following theorem from Krishnapur and Peres [7]:

**Theorem 4.1.** $\mathrm{Comb}(\mathbb{Z})$ has the finite collision property.

In Krishnapur and Peres [7] a more general statement is proven: $\mathrm{Comb}(G)$ has the finite collision property for any recurrent infinite graph with constant vertex degree. Since in the entire thesis we focus on $\mathrm{Comb}(\mathbb{Z})$ and its subgraphs, we only give the proof for $\mathrm{Comb}(\mathbb{Z})$. However, only minor modifications are needed to make the proof work for graphs with constant degree.

Let $(X_n)_{n\geq 0}$ and $(Y_n)_{n\geq 0}$ be simple random walks on $\mathrm{Comb}(\mathbb{Z})$ both starting from $(0,0)$. In the proof we will use the following random variables and events:

$$\begin{aligned}
Z_{n,\ell} &= |\{(N,L) : n \leq N < 2n, \ell \leq L < 2\ell \text{ and } X_N = Y_N = (v,L) \text{ for some } v \in \mathbb{Z}\}| \\
A_{n,\ell} &= \{Z_{n,\ell} > 0\}. \\
W_{n,\ell} &= \sum_{k \in \{\frac{\ell}{2}, \ell, 2\ell\}} Z_{n,k} + \sum_{k \in \{\frac{\ell}{2}, \ell, 2\ell\}} Z_{2n,k}
\end{aligned}$$

The goal is to show that only finitely many of the events $A_{n,\ell}$ occur. To do this, we find an upper bound for $\mathbb{E}[W_{n,\ell}]$ and a lower bound for $\mathbb{E}[W_{n,\ell}|A_{n,\ell}]$ and then use Proposition 3.10. We use several constants $C, C', C_1, C_2$, etc., which can have different values at each appearance.

**Lemma 4.1.** There exists a constant $C$ such that $\mathbb{E}[Z_{n,\ell}] \leq C\ell n^{-1/4}$ for all $n, \ell \geq 1$.

*Proof.* We decompose $X$ and $Y$ into a horizontal random walk and a vertical random walk. Let $U$ and $U'$ be independent simple random walks on $\mathbb{Z}$ starting from 0. Let $V$ and $V'$ be independent simple random walks on $\mathbb{Z}$ starting from 0 with a self-loop probability of $\frac{1}{2}$ at 0. Let $K_n$ and $K'_n$ be the number of transitions of $V$ and $V'$ from 0 to 0 respectively up to time $n$. Note that $K_n$ and $K'_n$ are also the number of transitions of $X$ and $Y$ on the horizontal copy of $\mathbb{Z}$, so they are the number of transitions of $U$ and $U'$ up to time $n$ respectively. Then we have

$$X = (U_{K_n}, V_n) \text{ and } Y = (U'_{K'_n}, V'_n).$$

Let us now fix $L \in \mathbb{Z}$. By the law of total probability, we have

$$\begin{aligned}
\mathbb{P}[X_n = Y_n = (x,L) \text{ for some } x \in \mathbb{Z}] &= \sum_{k=0}^{\infty} \sum_{k'=0}^{\infty} \mathbb{P}[V_n = V'_n = L, K_n = k, K'_n = k', U_k = U'_{k'}] \\
&= \sum_{k,k'} \mathbb{P}[V_n = V'_n = L, K_n = k, K'_n = k']\mathbb{P}[U_k = U'_{k'}],
\end{aligned}$$

since $U$ and $U'$ are independent of $V$ and $V'$ and hence also of $K_n$ and $K'_n$.

By Proposition 4.3, there exists a constant $C$ such that $\mathbb{P}[U_k = U'_{k'}] \leq \frac{C}{\sqrt{k+k'}}$. By taking the term with $k = k' = 0$ seperately and then applying this result, it follows that the probability $\mathbb{P}[X_n = Y_n = (x, L)$ for some $x \in \mathbb{Z}]$ is upper bounded by

$$\mathbb{P}\left[V_n = V'_n = L; K_n = K'_n = 0\right] + C \cdot \mathbb{E}\left[\frac{\mathbb{1}(V_n = V'_n = L; K_n + K'_n > 0)}{\sqrt{K_n + K'_n}}\right]. \tag{1}$$

To bound the second term we further decompose $V_n$ to get rid of the self-loop. Let $(S_k)_{k\geq 0}$ be a simple random walk on $\mathbb{Z}$ starting from 0 and let $(G_k)_{k\geq 0}$ be a sequence of independent geometrically distributed random variables with success probability $\frac{1}{2}$, i.e. $G_k \sim \mathrm{Geo}\left(\frac{1}{2}\right)$. Then $V$ is generated by following the path of $S$, except that it makes $G_i$ transitions from 0 to 0 on the $i$th visit to the origin by $S$. Note that $S$ starts at $S_0 = 0$. This counts as the 0th visit to the origin. Similarly, $V'$ is generated using $S'$ and $(G'_k)_{k\geq 0}$.

We now give a lower bound for $K_n$ based on the geometric random variables $(G_k)_{k\geq 0}$. Assume that $K_n < \frac{n}{2}$. Let $H_n = \sum_{i=1}^{\lfloor n/2 \rfloor} \mathbb{1}(S_i = 0)$. Note that in the first $n$ steps, we make $n - K_n$ steps according to $S$ and $K_n$ transitions from 0 to 0. These $K_n$ transitions at least include all transitions from 0 to 0 following the first $n - K_n - 1$ steps of $S$. Let $\widetilde{H}_n = \sum_{i=1}^{n-K_n-1} \mathbb{1}(S_i = 0)$ be the number of times $S$ visits 0 in the first $n - K_n - 1$ steps, then $K_n$ is at least the number of transitions from 0 to 0 following the first $\widetilde{H}_n$ visits of $S$ to 0. Also, note[1] that $K_n < \frac{n}{2}$ implies $n - K_n - 1 \geq \left\lfloor \frac{n}{2} \right\rfloor$ and hence $\widetilde{H}_n \geq H_n$. Therefore, the following inequalities hold:

$$K_n \geq \sum_{i=0}^{\widetilde{H}_n} G_i \geq \sum_{i=0}^{H_n} G_i \geq \sum_{i=1}^{H_n} G_i.$$

Let $R_n = \sum_{i=1}^{H_n} G_i$. Then $K_n \geq R_n$ or else $K_n \geq \frac{n}{2}$, so $K_n \geq R_n \wedge \frac{n}{2}$. Note that the summations for $H_n$ and $\widetilde{H}_n$ start from $i = 1$ to avoid counting the 0th visit to origin.

Since $K_n \geq R_n \wedge \frac{n}{2}$ and similarly $K'_n \geq R'_n \wedge \frac{n}{2}$, it follows that $\mathbb{E}\left[\frac{\mathbb{1}(V_n = V'_n = L; K_n + K'_n > 0)}{\sqrt{K_n + K'_n}}\right]$ is bounded by

$$\mathbb{E}\left[\frac{\mathbb{1}(V_n = V'_n = L; R_n + R'_n > 0)}{\sqrt{(R_n \wedge \frac{n}{2}) + (R'_n \wedge \frac{n}{2})}}\right] + \mathbb{P}\left[V_n = V'_n = L; R_n = R'_n = 0\right]. \tag{2}$$

To compute this expectation, condition on $\{S_i : i \leq n/2\}$ and $\{G_i : i \leq H_n\}$ and their primed versions: this completely determines the path of $V$ and $V'$ up to time $\frac{n}{2} + R_n$ and $\frac{n}{2} + R'_n$ respectively. Moreover, it determines $R_n$ and $R'_n$. First assume that $\max\{R_n, R'_n\} < \frac{n}{4}$, then $V$ and $V'$ have at least $\frac{n}{4}$ more steps to go. Since the self-loop probability of $V$ at 0 is $\frac{1}{2}$, $V$ is actually a simple random walk on $\mathbb{Z}$ with a self-loop added at 0. It hence follows from Proposition 3.4 that $p^{(n/4)}(x, L) \leq \frac{C}{\sqrt{n}}$ for all $x \in \mathbb{Z}$ and hence that $\mathbb{P}[V_n = L] \leq \frac{C}{\sqrt{n}}$.

[1] Proof: Since $K_n$ is an integer, $K_n < \frac{n}{2}$ implies $K_n \leq \frac{n-1}{2}$ if $n$ is odd and $K_n \leq \frac{n-2}{2}$ if $n$ is even. Hence, $n - K_n - 1 \geq n - \left(\frac{n-1}{2}\right) - 1 = \frac{n-1}{2} = \left\lfloor \frac{n}{2} \right\rfloor$ if $n$ is odd and $n - K_n - 1 \geq n - \left(\frac{n-2}{2}\right) - 1 = \frac{n}{2} = \left\lfloor \frac{n}{2} \right\rfloor$ if $n$ is even.

Similarly, $\mathbb{P}[V_n' = L] \leq \frac{C}{\sqrt{n}}$ and by independence it follows that $\mathbb{P}[V_n = V_n' = L] \leq \frac{C^2}{n}$ if $\max\{R_n, R_n'\} < \frac{n}{4}$. Hence, the expectation in (2) can be upper bounded by

$$C'\mathbb{E}\left[\frac{\mathbb{1}(R_n + R_n' > 0)}{\sqrt{R_n + R_n'}}\frac{1}{n} + \frac{\mathbb{1}(\max\{R_n, R_n'\} \geq \frac{n}{4})}{\sqrt{n}}\right]. \tag{3}$$

We now focus on the first term in this expectation. First note that we can bound

$$\begin{aligned}\mathbb{E}\left[\frac{\mathbb{1}(R_n + R_n' > 0)}{\sqrt{R_n + R_n'}}\frac{1}{n}\right] &\leq \mathbb{E}\left[\frac{\mathbb{1}(R_n > 0) + \mathbb{1}(R_n' > 0)}{\sqrt{R_n + R_n'}}\frac{1}{n}\right]\\ &\leq \mathbb{E}\left[\frac{\mathbb{1}(R_n > 0)}{\sqrt{R_n}}\frac{1}{n}\right] + \mathbb{E}\left[\frac{\mathbb{1}(R_n' > 0)}{\sqrt{R_n'}}\frac{1}{n}\right] = 2\mathbb{E}\left[\frac{\mathbb{1}(R_n > 0)}{\sqrt{R_n}}\right]\cdot\frac{1}{n}.\end{aligned}$$

Recall from Corollary 3.2 that $\mathbb{P}[H_n = k] \leq \frac{C}{\sqrt{n}}$. Although $H_n$ is defined slightly different here, this only changes the value of the constant. Using this inequality, we bound

$$\mathbb{E}\left[\frac{\mathbb{1}(H_n > 0)}{\sqrt{H_n}}\right] \leq \frac{\mathbb{P}[H_n \geq n^{1/2}]}{\sqrt{n^{1/2}}} + \sum_{k=1}^{n^{1/2}}\frac{\mathbb{P}[H_n = k]}{\sqrt{k}} \leq \frac{1}{n^{1/4}} + \frac{C}{\sqrt{n}}\sum_{k=1}^{n^{1/2}}\frac{1}{\sqrt{k}} \leq C_1 n^{-1/4}.$$

Since $(G_k)_{k\geq 0}$ are independent and geometric, it follows that $\widetilde{G}_r = \sum_{i=1}^r G_i$ has the negative binomial distribution with success probability $\frac{1}{2}$ and $r$ successes required. Hence, $\mathbb{E}[\widetilde{G}_r] = r$ and $\mathrm{Var}[\widetilde{G}_r] = 2r$. It follows by Chebyshev's inequality[2] that

$$\mathbb{P}\left(\widetilde{G}_r \leq \frac{r}{2}\right) \leq \mathbb{P}\left(\left|\widetilde{G}_r - r\right| \geq \frac{r}{2}\right) \leq \frac{8}{r},$$

and hence

$$\mathbb{E}\left[\frac{\mathbb{1}\left(\sum_{i=1}^r G_i \neq 0\right)}{\sqrt{\sum_{i=1}^r G_i}}\right] = \mathbb{E}\left[\frac{\mathbb{1}(\widetilde{G}_r \neq 0)}{\sqrt{\widetilde{G}_r}}\right] \leq \mathbb{P}\left(\widetilde{G}_r \leq \frac{r}{2}\right) + \sqrt{\frac{2}{r}} \leq \frac{8}{r} + \sqrt{\frac{2}{r}} \leq \frac{C_2}{\sqrt{r}}.$$

This inequality holds for fixed $r$, but we can use it for $H_n$ after conditioning on it using the law of iterated expectations. Combining these inequalities therefore yields

$$\begin{aligned}\mathbb{E}\left[\frac{\mathbb{1}(R_n + R_n' > 0)}{\sqrt{R_n + R_n'}}\frac{1}{n}\right] &\leq 2\mathbb{E}\left[\frac{\mathbb{1}(R_n > 0)}{\sqrt{R_n}}\right]\cdot\frac{1}{n} = 2\mathbb{E}\left[\frac{\mathbb{1}\left(\sum_{i=1}^{H_n} G_i \neq 0\right)}{\sqrt{\sum_{i=1}^{H_n} G_i}}\right]\cdot\frac{1}{n}\\ &= 2\mathbb{E}\left[\mathbb{E}\left[\frac{\mathbb{1}\left(\sum_{i=1}^{H_n} G_i \neq 0\right)}{\sqrt{\sum_{i=1}^{H_n} G_i}}\middle| H_n\right]\right]\cdot\frac{1}{n} \leq 2\mathbb{E}\left[\mathbb{E}\left[\frac{C_2\cdot\mathbb{1}(H_n > 0)}{\sqrt{H_n}}\middle| H_n\right]\right]\cdot\frac{1}{n}\\ &= 2C_2\cdot\mathbb{E}\left[\frac{\mathbb{1}(H_n > 0)}{\sqrt{H_n}}\right]\cdot\frac{1}{n} \leq \frac{2C_2}{n}C_1 n^{-1/4} = \frac{C_3}{n^{5/4}},\end{aligned}$$

which provides a bound for the first term in equation (3).

[2] Krishnapur and Peres [7] use Cramér's theorem to bound this probability.

We now bound the second term in equation (3). Note that $R_n \leq |\{i \leq n : V_i = 0\}|$, so

$$\mathbb{P}\left[\max\{R_n, R'_n\} \geq \tfrac{n}{4}\right] \leq \mathbb{P}\left[R_n \geq \tfrac{n}{4}\right] + \mathbb{P}\left[R'_n \geq \tfrac{n}{4}\right] = 2\mathbb{P}\left[R_n \geq \tfrac{n}{4}\right] \leq 2\mathbb{P}\left[|\{i \leq n : V_i = 0\}| \geq \tfrac{n}{4}\right].$$

Note that $\mathbb{P}[V_k = 0] \leq \frac{C}{\sqrt{k}}$ for $1 \leq k \leq n$, so summing this yields $\mathbb{E}[|\{i \leq n : V_i = 0\}|] \leq C'\sqrt{n}$. By Corollary 3.1, it hence follows that $\mathbb{P}\left[|\{i \leq n : V_i = 0\}| \geq \frac{n}{4}\right]$ and hence $\mathbb{P}\left[R_n \geq \frac{n}{4}\right]$ are upper bounded by $2\exp\{-cn\}$ for some constant $c$.

Finally, we bound $\mathbb{P}[V_n = V'_n = L; K_n = K'_n = 0]$ and $\mathbb{P}[V_n = V'_n = L; R_n = R'_n = 0]$. First note that $K_n = K'_n = 0$ implies $R_n = R'_n = 0$, so the latter probability is larger. It hence suffices to find an upper bound for $\mathbb{P}[V_n = V'_n = L; R_n = R'_n = 0]$. Note that $R_n = 0$ if and only if $G_i = 0$ for all $1 \leq i \leq H_n$. Since the $G_i$ are independent and $\mathbb{P}[G_i > 0] = \frac{1}{2}$ for all $i$, we therefore have

$$\mathbb{P}[R_n = 0] = \sum_{k=0}^{n} \frac{\mathbb{P}[H_n = k]}{2^k} \leq \frac{C}{\sqrt{n}} \sum_{k=0}^{n} \frac{1}{2^k} \leq \frac{2C}{\sqrt{n}},$$

since $\mathbb{P}[H_n = k] \leq \frac{C}{\sqrt{n}}$ by Corollary 3.2. Since $R$ and $R'$ are independent, it follows that $\mathbb{P}[R_n = R'_n = 0] \leq \frac{4C^2}{n}$. Hence, it follows that

$$\begin{aligned}\mathbb{P}\left[V_n = V'_n = L; R_n = R'_n = 0\right] &= \mathbb{P}\left[V_n = V'_n = L \mid R_n = R'_n = 0\right] \mathbb{P}\left[R_n = R'_n = 0\right] \\ &\leq \tfrac{4C^2}{n} \mathbb{P}\left[V_n = V'_n = L \mid R_n = R'_n = 0\right].\end{aligned}$$

To bound the latter probability, again condition on $\{S_i : i \leq n/2\}$ and $\{G_i : i \leq H_n\}$ and their primed versions. Note that this determines $R_n$ and $R'_n$. By a similar argument as before it now follows that, $\mathbb{P}[V_n = V'_n = L \mid R_n = R'_n = 0] \leq \frac{C'^2}{n}$.
This together yields that there exists a constant $C$ such that

$$\mathbb{P}\left[V_n = V'_n = L; R_n = R'_n = 0\right] \leq \frac{C}{n^2}.$$

We now have all results needed to prove the lemma. From combining equations (1), (2) and (3), we see that there exist constants $C_4$ and $C_6$ such that

$$\begin{aligned}&\mathbb{P}[X_n = Y_n = (x, L) \text{ for some } x \in \mathbb{Z}] \\ &\quad \leq C_4 \mathbb{E}\left[\frac{\mathbb{1}(R_n + R'_n > 0)}{\sqrt{R_n + R'_n}} \frac{1}{n} + \frac{\mathbb{1}(\max\{R_n, R'_n\} > \frac{n}{4})}{\sqrt{n}}\right] + C_6 \mathbb{P}\left[V_n = V'_n = L; R_n = R'_n = 0\right].\end{aligned}$$

Of these three terms, the first one dominates. Hence, there exists a constant $C$ such that

$$\mathbb{P}[X_n = Y_n = (x, L) \text{ for some } x \in \mathbb{Z}] \leq \frac{C}{n^{5/4}} \quad \text{for all } L \in \mathbb{Z}, n \geq 1 \tag{4}$$

To bound $\mathbb{E}[Z_{n,\ell}]$, it remains to sum this bound over $n$ and $\ell$:

$$\mathbb{E}[Z_{n,\ell}] \leq \sum_{N=n}^{2n-1} \sum_{L=\ell}^{2\ell-1} \mathbb{P}[X_N = Y_N = (x, L) \text{ for some } x \in \mathbb{Z}] \leq C\ell n^{-1/4},$$

so we conclude that there exists a constant $C$ such that $\mathbb{E}[Z_{n,\ell}] \leq C\ell n^{-1/4}$ for all $n, \ell \geq 1$. □

Recall that $Z_{n,\ell}$ is the number of collisions bounded in space and time by $\ell \leq L < 2\ell$ and $n \leq N < 2n$, and $A_{n,\ell}$ is the event that a positive number of these collisions occur. If such a collision occurs, then we expect many more collisions to occur soon after this and also nearby in space. Recall that $W_{n,\ell}$ is the number of collisions bounded in space and time by $\frac{\ell}{2} \leq L < 4\ell$ and $n \leq N < 4n$. This can be expressed as the sum of 6 random variables of the form $Z_{n,\ell}$. From Lemma 4.1 we immediately conclude that there exists a constant $C_1$ such that $\mathbb{E}[W_{n,\ell}] \leq C_1 \ell n^{-1/4}$ for all $n, \ell \geq 1$. We now find a lower bound for $\mathbb{E}[W_{n,\ell}|A_{n,\ell}]$:

**Lemma 4.2.** Let $0 < \alpha < 1$ be given. There exists a constant $C > 0$, depending on $\alpha$ but not on $\ell$ or $n$, such that for all $n, \ell$ with $1 \leq \ell < 2(2n)^{1/(2\alpha)}$, we have $\mathbb{E}[W_{n,\ell}|A_{n,\ell}] \geq C\ell^\alpha$.

*Proof.* Assume that $A_{n,\ell}$ occurs. Then the two random walks collide at a time $N_c$ with $n \leq N_c < 2n$ at some vertex $(v, L_c)$ with $\ell \leq L_c < 2\ell$. Note that $W_{n,\ell}$ is at least the number of collisions bounded in space and time by $\frac{\ell}{2} \leq L < 4\ell$ and $n \leq N < 4n$, and in particular it is at least number of collisions bounded in space and time by $L_c - \frac{\ell}{2} \leq L \leq L_c + \frac{\ell}{2}$ and $N_c \leq N \leq N_c + 2n$. This is in turn at least the number of collisions that occur before one of the random walks hits one of the vertices $(v, L_c \pm \frac{\ell}{2})$ or $2n$ steps are done, whichever occurs earlier. Note that during this time interval, both random walks are confined to a segment of $\mathbb{Z}$, so we may assume that they in fact occur on a segment of $\mathbb{Z}$. We center the segment around 0 instead of around $L_c$.
This can be formalized as follows. Let $U$ and $V$ be two independent random walks on $\mathbb{Z}$ starting from 0. Let $T_U$ be the first time $U$ hits $\pm\frac{\ell}{2}$ and similarly define $T_V$. Let

$$Y_{n,\ell} = \sum_{k=0}^{2n \wedge T_U \wedge T_V} \mathbb{1}[U_k = V_k]$$

be the number of collisions that occur before one of the random walks hits one of the vertices $(v, L_c \pm \frac{\ell}{2})$ or $2n$ steps are done. Given that $A_{n,\ell}$ occurs, $W_{n,\ell}$ stochastically dominates $Y_{n,\ell}$, as explained before. Note that $\ell < 2(2n)^{1/(2\alpha)}$ implies $2n \geq (\ell/2)^{2\alpha}$. Hence,

$$\begin{aligned}
\mathbb{E}[W_{n,\ell}|A_{n,\ell}] \geq \mathbb{E}[Y_{n,\ell}] \geq &\sum_{k=0}^{(\ell/2)^{2\alpha} \wedge T_U \wedge T_V} \mathbb{P}[U_k = V_k] \\
&= \sum_{k=0}^{(\ell/2)^{2\alpha}} \mathbb{P}[U_k = V_k; T_U \wedge T_V > k] \geq \sum_{k=0}^{(\ell/2)^{2\alpha}} (\mathbb{P}[U_k = V_k] - \mathbb{P}[T_U \wedge T_V \leq k]) \\
&\geq \left( \sum_{k=0}^{(\ell/2)^{2\alpha}} \mathbb{P}[U_k = V_k] \right) - (\ell/2)^{2\alpha} \mathbb{P}[T_U \wedge T_V \leq (\ell/2)^{2\alpha}] \\
&\geq \left( \sum_{k=1}^{(\ell/2)^{2\alpha}} \frac{C'}{\sqrt{k}} \right) - 2(\ell/2)^{2\alpha} \mathbb{P}[T_U \leq (\ell/2)^{2\alpha}]
\end{aligned}$$

where the last inequality holds by Proposition 4.4. We have $\sum_{k=0}^{(\ell/2)^{2\alpha}} \frac{C'}{\sqrt{k}} \geq C'' \ell^\alpha$.

Lemma 3.11 provides an upper bound for $\mathbb{P}[T_U \le (\ell/2)^{2\alpha}]$. We have

$$\mathbb{P}[T_U \le (\ell/2)^{2\alpha}] \le 4\exp\{-\tfrac{1}{6}(\ell/2)^{2-2\alpha}\}.$$

Hence, $2(\ell/2)^{2\alpha}\mathbb{P}[T_U \le (\ell/2)^{2\alpha}] \le 8(\ell/2)^{2\alpha}\exp\{-\frac{1}{6}(\ell/2)^{2(1-\alpha)}\}$. Since $1-\alpha > 0$, goes faster to 0 than any polynomial in $\ell$. In particular, there exists an $L$ such that for $\ell \ge L$ we have $2(\ell/2)^{2\alpha}\mathbb{P}[T_U \le (\ell/2)^{2\alpha}] \le \frac{1}{2}C''\ell^\alpha$. Hence, $\mathbb{E}[W_{n,\ell}|A_{n,\ell}] \ge \frac{1}{2}C''\ell^\alpha$ for $\ell \ge L$.

Note that $\mathbb{E}[W_{n,\ell}|A_{n,\ell}] \ge 1$ since $W_{n,\ell} \ge Z_{n,\ell} > 0$ by assumption. Take $0 < C \le \min\left\{\frac{1}{2}C'', \frac{1}{L^\alpha}\right\}$. Then it follows that $\mathbb{E}[W_{n,\ell}|A_{n,\ell}] \ge C\ell^\alpha$ for all $n, \ell$ with $1 \le \ell < 2(2n)^{1/(2\alpha)}$. □

We have now shown that for a given $0 < \alpha < 1$ there exist constants $C_1, C_2$ such that

$$\begin{aligned} \mathbb{E}[W_{n,\ell}] &\le C_1 \ell n^{-1/4} && \text{for all } n, \ell \ge 1, \\ \mathbb{E}[W_{n,\ell}|A_{n,\ell}] &\ge C_2 \ell^\alpha && \text{for all } \ell \le 2(2n)^{1/(2\alpha)}. \end{aligned}$$

Hence, it follows by Proposition 3.10 that there exists a constant $C$ such that

$$\mathbb{P}[A_{n,\ell}] \le \frac{\mathbb{E}[W_{n,\ell}]}{\mathbb{E}[W_{n,\ell}|A_{n,\ell}]} \le C\frac{\ell^{1-\alpha}}{n^{1/4}} \qquad \text{for all } \ell \le 2(2n)^{1/(2\alpha)}.$$

We are now ready for the main proof. For any $(N, L)$ with $N, L \ge 1$ there exist integers $r, k \ge 0$ such that for $n = 2^r$ and $\ell = 2^k$ we have $n \le N < 2n$ and $\ell \le L < 2\ell$. If $\ell \le 2(2n)^{1/(2\alpha)}$, then

$$k = \log_2 \ell \le 1 + \tfrac{\log_2(2n)}{2\alpha} = 1 + \tfrac{r+1}{2\alpha}.$$

Conversely, if we count up to $k = 1 + \frac{r+1}{2\alpha}$, then we include at least all $L \le 2(2n)^{1/(2\alpha)}$. Let us now choose $\alpha > 2/3$. Then $1/(2\alpha) - 3/4 < 0$, so $2^{(1/(2\alpha)-3/4)} < 1$. Hence, we have

$$\sum_{r=0}^{\infty}\sum_{k=0}^{1+\frac{r+1}{2\alpha}} \mathbb{P}[A_{2^r,2^k}] \le \sum_{r=0}^{\infty}\sum_{k=0}^{1+\frac{r+1}{2\alpha}} C\frac{2^{k(1-\alpha)}}{2^{r/4}} \le \sum_{r=0}^{\infty} C'\frac{2^{r(1-\alpha)/(2\alpha)}}{2^{r/4}} = \sum_{r=0}^{\infty} C'2^{r(1/(2\alpha)-3/4)} < \infty.$$

By the Borel-Cantelli lemma (Proposition 3.13), it follows that only finitely many of the events $\mathbb{P}[A_{2^r,2^k}]$ with $k \le 1 + \frac{r+1}{2\alpha}$ occur. As explained above, this counts all collisions with time $N$ and vertical displacement $L$ satisfying $1 \le L \le 2(2n)^{1/(2\alpha)}$. By symmetry, the same holds for negative displacements satisfying $1 \le |L| \le 2(2n)^{1/(2\alpha)}$. Hence, for $2/3 < \alpha < 1$ the set

$$\{n : X_n = Y_n = (x, L) \text{ for some } x \in \mathbb{Z} \text{ and } L \in \mathbb{Z} \text{ with } 1 \le |L| \le 2(2n)^{1/(2\alpha)}\}$$

is finite almost surely. The number of collisions with $L = 0$ is finite almost surely by equation (4). Finally, $\{n : |V_n| > 2(2n)^{1/(2\alpha)} \text{ or } |V_n'| > 2(2n)^{1/(2\alpha)}\}$ is finite almost surely by the law of the iterated logarithm (Proposition 3.11) and in particular the number of collisions with $|L| > 2(2n)^{1/(2\alpha)}$ is finite almost surely. Hence, two random walks $X$ and $Y$ on Comb($\mathbb{Z}$) collide only finitely often almost surely. This concludes the proof of Theorem 4.1. □

### 4.3 The Green function criterion and applications

In this section, the Green function criterion for infinitely many collisions from Barlow, Peres and Sousi [3] is studied. Recall from Section 3.2 that $q_t^B(x,y)$ is the $t$-step transition density divided by $d(y)$ for a random walk which is killed when it exits $B$ and let $g_B(x,y) = \sum_{t=0}^{\infty} q_t^B(x,y)$ be the corresponding Green kernel. Throughout this section, let $X$ and $Y$ be two independent random walks on a graph $G$ and let $X^B$ and $Y^B$ be the corresponding walks which are killed after exiting $B$. Let $\widetilde{Z}_B$ be the number of edges crossed at the same time by $X^B$ and $Y^B$, i.e.

$$\widetilde{Z}_B = \sum_{t=0}^{\infty} \mathbb{1}\{X_t^B = Y_t^B, X_{t+1} = Y_{t+1}\}.$$

To prove the criterion, first a bound for $\mathbb{E}_{o,o}[\widetilde{Z}_B]$ is given. This bound will be needed in the proof of the criterion, but will also be used on a few other occasions.

**Lemma 4.3.** Let $G = (V,E)$ be a graph with distinguished vertex $o$. Let $B$ be a subset of the vertex set $V$. Let $\widetilde{Z}_B$ be the number of edges crossed at the same time by $X^B$ and $Y^B$. Then

$$\tfrac{1}{2} g_B(o,o) \le \mathbb{E}_{o,o}[\widetilde{Z}_B] \le g_B(o,o).$$

*Proof.* We start by proving that $\mathbb{E}_{o,o}[\widetilde{Z}_B] = \sum_{t=0}^{\infty} q_{2t}^B(o,o)$:

$$\begin{aligned}
\mathbb{E}_{o,o}[\widetilde{Z}_B] &= \sum_{t=0}^{\infty} \sum_{x\in B} \sum_{y\sim x} \mathbb{P}_{o,o}(X_t^B = Y_t^B = x, X_{t+1} = Y_{t+1} = y) \\
&= \sum_{t=0}^{\infty} \sum_{x\in B} \sum_{y\sim x} \mathbb{P}_o(X_t^B = x)\mathbb{P}(X_{t+1} = y | X_t^B = x)\mathbb{P}_o(Y_t^B = x)\mathbb{P}(Y_{t+1} = y | Y_t^B = y) \\
&= \sum_{t=0}^{\infty} \sum_{x\in B} \left[\mathbb{P}_o(X_t^B = x)\mathbb{P}_o(Y_t^B = x) \sum_{y\sim x} \frac{1}{d(x)^2}\right] = \sum_{t=0}^{\infty} \sum_{x\in B} \left[\mathbb{P}_o(X_t^B = x)^2 \frac{1}{d(x)}\right] \\
&= \sum_{t=0}^{\infty} \sum_{x\in B} \left[\frac{\mathbb{P}_o(X_t^B = x)\mathbb{P}_x(X_t^B = o)}{d(o)}\right] = \sum_{t=0}^{\infty} \frac{\mathbb{P}_o(X_{2t}^B = o)}{d(o)} = \sum_{t=0}^{\infty} q_{2t}^B(o,o).
\end{aligned}$$

In the first equality, we sum the collision probabilities over all possible times and all pairs of vertices connected by an edge. In the second equality, we use the fact that the random walks $X$ and $Y$ are independent and also the fact that by the Markov property the probability that $X_{t+1} = y$ given that $X_t = x$ does not depend on the path the walk has taken before time $t$. In the third equality, we take out terms that do not depend on $y$ and use the fact that $X$ and $Y$ are simple random walks. In the fourth equality, we use the fact that $X$ and $Y$ are identically distributed. In the fifth equality we use the property that $\mathbb{P}_x(X_t^B = o) = \frac{d(o)}{d(x)}\mathbb{P}_o(X_t^B = x)$. The sixth equality holds by conditioning on $X_t^B$ and using the Markov property. The final equality follows from the definition of the transition density $q$.

By Lemma 3.9 we have $g_B(o,o) \ge \sum_{t=0}^{\infty} q_{2t}^B(o,o) \ge \frac{1}{2} g_B(o,o)$. Hence, the lemma follows. $\square$

**Theorem 4.2.** (**Green function criterion**)
Let $G = (V, E)$ be a recurrent graph with distinguished vertex $o$. Let $(B_r)_r$ be an increasing sequence of vertex sets such that $B_r \neq V$ for all $r$ and $\bigcup_r B_r = V$. Suppose that there exists a constant $C < \infty$ such that for all $r$ we have

$$g_{B_r}(x, x) \leq C g_{B_r}(o, o) \quad \text{for all } x \in B_r.$$

Then $G$ has the infinite collision property.

*Proof.* Let $r$ be given and write $B = B_r$. By Markov's inequality, it follows that

$$\mathbb{P}\left(\widetilde{Z}_B \geq \tfrac{1}{2}\mathbb{E}_{o,o}[\widetilde{Z}_B]\right) = \mathbb{P}\left(\widetilde{Z}_B^2 \geq \tfrac{1}{4}\mathbb{E}_{o,o}[\widetilde{Z}_B]^2\right) \geq \frac{(\mathbb{E}_{o,o}[\widetilde{Z}_B])^2}{4\mathbb{E}_{o,o}[\widetilde{Z}_B^2]}.$$

To be able to use this bound, the second moment of $\widetilde{Z}_B$ needs to be computed. To do this, write

$$\mathbb{E}_{o,o}[\widetilde{Z}_B^2] = \mathbb{E}_{o,o}[\widetilde{Z}_B] + 2\mathbb{E}_{o,o}\left[\tbinom{\widetilde{Z}_B}{2}\right],$$

and note that $\binom{\widetilde{Z}_B}{2}$ can be interpreted as the number of pairs of edges crossed at the same time. The probability that this occurs for a given pair of edges and at given times $s, t$ with $s \geq t+1$, can be written as follows:

$$\begin{aligned}
&\mathbb{P}_{o,o}(X_t^B = Y_t^B = x, X_{t+1} = Y_{t+1} = y, X_s^B = Y_s^B = z, X_{s+1} = Y_{s+1} = w) \\
&\quad = \mathbb{P}_o(X_t^B = x)\mathbb{P}(X_{t+1} = y | X_t^B = x)\mathbb{P}_o(Y_t^B = x)\mathbb{P}(Y_{t+1} = y | Y_t^B = y) \\
&\quad\quad \times \mathbb{P}_y(X_{s-t-1}^B = z)\mathbb{P}(X_{s+1} = w | X_s^B = w)\mathbb{P}_o(Y_{s-t-1}^B = z)\mathbb{P}(Y_{s+1} = w | Y_s^B = w) \\
&\quad = \frac{\mathbb{P}_o(X_t^B = x)^2}{d(x)^2}\frac{\mathbb{P}_y(X_{s-t-1}^B = z)^2}{d(z)^2} = q_t^B(o,x)^2 q_{s-t-1}^B(y,z)^2.
\end{aligned}$$

Denote this probability by $p(s, t, x, y, z, w)$. By summing the collision probabilities over all possible pairs of times $s, t$ with $s \geq t+1$ and all pairs of pairs of vertices connected by an edge, we find the following result:

$$\begin{aligned}
\mathbb{E}_{o,o}\left[\binom{\widetilde{Z}_B}{2}\right] &= \sum_{t=0}^{\infty}\sum_{s=t+1}^{\infty}\sum_{x\in B}\sum_{y\sim x}\sum_{z\in B}\sum_{w\sim z} p(s,t,x,y,z,w) \\
&= \sum_{t=0}^{\infty}\sum_{s=t+1}^{\infty}\sum_{x\in B}\sum_{y\sim x}\sum_{z\in B}\sum_{w\sim z} q_t^B(o,x)^2 q_{s-t-1}^B(y,z)^2 \\
&= \sum_{t=0}^{\infty}\sum_{s=0}^{\infty}\sum_{x\in B}\sum_{y\sim x}\sum_{z\in B} q_t^B(o,x)^2 q_s^B(y,z)^2 d(z)
\end{aligned}$$

Note that $\sum_{z\in B} q_s^B(y,z)^2 d(z) = \sum_{z\in B} q_s^B(y,z) q_s^B(z,y) d(y) = q_{2s}^B(y,y)$. After summing over $s$, this yields a factor $\sum_{s=0}^{\infty} q_{2s}^B(y,y)$, which is upper bounded by $g_B(y,y)$.

Using this result, we can write

$$\begin{aligned}
\mathbb{E}_{o,o}\left[\binom{\widetilde{Z}_B}{2}\right] &\leq \sum_{t=0}^{\infty}\sum_{x\in B}\sum_{y\sim x} q_t^B(o,x)^2 g_B(y,y) \leq \max_{y\in B} g_B(y,y)\sum_{t=0}^{\infty}\sum_{x\in B}\sum_{y\sim x} q_t^B(o,x)^2 \\
&= \max_{y\in B} g_B(y,y)\sum_{t=0}^{\infty}\sum_{x\in B} q_t^B(o,x)^2 d(x) = \max_{y\in B} g_B(y,y)\sum_{t=0}^{\infty} q_{2t}^B(o,o) \\
&\leq g_B(o,o)\max_{y\in B} g_B(y,y),
\end{aligned}$$

where in the second inequality, we can write $g_B(y,y) \leq \max_{y\in B} g_B(y,y)$ because $g_B(y,y) = 0$ for all $y \notin B$, and the other steps follow in the same way as in Lemma 4.3.
Recall that $g_B(o,o) \geq \mathbb{E}_{o,o}[\widetilde{Z}_B] \geq \frac{1}{2}g_B(o,o)$ by Lemma 4.3. Since $G$ is recurrent and $B_r \neq V$, it follows that $g_B(o,o) < \infty$. By the given inequality, we have $\max_{y\in B} g_B(y,y) \leq Cg_B(o,o)$. Hence, $\mathbb{E}_{o,o}\left[\binom{\widetilde{Z}_B}{2}\right] \leq Cg_B(o,o)^2$ and hence $\mathbb{E}_{o,o}[\widetilde{Z}_B^2] = \mathbb{E}_{o,o}[\widetilde{Z}_B] + 2\mathbb{E}_{o,o}\left[\binom{\widetilde{Z}_B}{2}\right] \leq g_B(o,o) + Cg_B(o,o)^2$. It follows that

$$\mathbb{P}\left(\widetilde{Z}_B \geq \tfrac{1}{4}g_B(o,o)\right) \geq \mathbb{P}\left(\widetilde{Z}_B \geq \tfrac{1}{2}\mathbb{E}_{o,o}[\widetilde{Z}_B]\right) \geq \frac{(\mathbb{E}_{o,o}[\widetilde{Z}_B])^2}{4\mathbb{E}_{o,o}[\widetilde{Z}_B^2]} \geq \frac{g_B(o,o)}{16(1+2Cg_B(o,o))}.$$

Since $g_B(o,o) \geq d(o)^{-1}$ and since the right hand side of the inequality is decreasing in $r$, it follows that $\mathbb{P}\left(\widetilde{Z}_{B_r} \geq \frac{1}{4}g_B(o,o)\right) \geq c$ for all $r > 0$. As $r \to \infty$, we have $\widetilde{Z}_{B_r} \uparrow \widetilde{Z}$, where $\widetilde{Z}$ is the number of edges crossed by $X$ and $Y$ at the same time. Letting $r \to \infty$, it follows that $\mathbb{P}_{o,o}(\widetilde{Z} = \infty) \geq c$ and since $Z \geq \widetilde{Z}$, it then also follows that $\mathbb{P}_{o,o}(Z = \infty) \geq c$. So by the 0-1 law it follows that $\mathbb{P}_{o,o}(Z = \infty) = 1$. Hence, $G$ has the infinite collision property. □

Before applying this criterion, it is useful to give the following corollary of Lemma 4.3:

**Corollary 4.2.** Let $G = (V,E)$ be a graph with distinguished vertex $o$. Let $B$ be a subset of the vertex set $V$ and assume that $d(x) \leq D$ for all $x \in B$. Let $Z_B$ be the number of collisions of the random walks $X^B$ and $Y^B$ that are killed when exiting $B$. Then

$$\tfrac{1}{2}g_B(o,o) \leq \mathbb{E}_{o,o}[Z_B] \leq Dg_B(o,o).$$

*Proof.* Let $\widetilde{Z}_B$ be the number of edges crossed at the same time by $X^B$ and $Y^B$. Then $Z_B \geq \widetilde{Z}_B$, from which it immediately follows that $\mathbb{E}_{o,o}[Z_B] \geq \mathbb{E}_{o,o}[\widetilde{Z}_B] \geq \frac{1}{2}g_B(o,o)$.
For the other inequality, observe that every collision of $X^B$ and $Y^B$ yields a probability of at least $\frac{1}{D}$ that $X^B$ and $Y^B$ cross the same edge. Hence, $\mathbb{E}_{o,o}[Z_B] \leq D\mathbb{E}_{o,o}[\widetilde{Z}_B] \leq Dg_B(o,o)$. □

The Green function criterion has a wide range of applications. In particular, the Green function criterion can be used to prove that $\mathrm{Comb}(\mathbb{Z},\alpha)$ has the infinite collision property for $\alpha \leq 1$.

**Theorem 4.3.** Comb$(\mathbb{Z}, \alpha)$ has the infinite collision property for $\alpha \leq 1$.

*Proof.* Let $\alpha \leq 1$ be given. Let $V$ denote the vertex set of Comb$(\mathbb{Z}, \alpha)$. Let $B_r$ denote the set of vertices to the right of the origin and on horizontal distance at most $r$, i.e.

$$B_r := \{(x, y) \in V : 0 \leq x \leq r\}.$$

Set $o = (0, 0)$. Note that $o \in B_r$ and $(B_r)_r$ be an increasing sequence of vertex sets such that $B_r \neq V$ for all $r$ and $\bigcup_r B_r = V$. So to apply the Green function criterion, it just needs to be proven that $g_{B_r}(x, x) \leq C g_{B_r}(o, o)$ for all $r$ and all $x \in B_r$. In fact, this holds with $C = 1$.

By Lemma 3.7, it follows that $g_{B_r}(x, x) = R_{\text{eff}}(x, B_r^c)$. Note that

$$\begin{aligned} R_{\text{eff}}(x, B_r^c)^{-1} &= \inf\{\mathcal{E}(f, f) : \mathcal{E}(f, f) < \infty, f(x) = 1, f|_{B_r^c} = 0\} \\ &\geq \inf\{\mathcal{E}(f, f) : \mathcal{E}(f, f) < \infty, f(x) = 1, f((r+1, 0)) = 0\} = R_{\text{eff}}(x, (r+1, 0))^{-1}, \end{aligned}$$

since for the second set there are less constraints and hence the first set is a subset of the second set, so the infimum of the second set cannot be larger. However, note that the only edge between a vertex of $B_r$ and a vertex of $B_r^c$ is the edge between $(r, 0)$ and $(r+1, 0)$. So given that $f((r+1, 0)) = 0$, it is optimal for $f$ to be constant 0 on $B_r^c$ when minimizing $\mathcal{E}(f, f)$, and hence the two infima coincide. So $R_{\text{eff}}(x, B_r^c) = R_{\text{eff}}(x, (r+1, 0))$. Since a comb graph is a tree, we have $R_{\text{eff}}(x, (r+1, 0)) = d(x, (r+1, 0))$ by Lemma 3.5, so

$$g_{B_r}(x, x) = R_{\text{eff}}(x, B_r^c) = R_{\text{eff}}(x, (r+1, 0)) = d(x, (r+1, 0))$$

for all $x \in B$. In particular, $g_{B_r}(o, o) = d((0, 0), (r+1, 0)) = r + 1$.

Since $\alpha \leq 1$, we can write $x = (x_1, x_2)$ with $x_2 \leq x_1^\alpha \leq x_1$ for all $x \in B$. Hence,

$$g_{B_r}(x, x) = d((x_1, x_2), (r+1, 0)) = r + 1 - x_1 + x_2 \leq r + 1$$

for all $x \in B$. So $g_{B_r}(x, x) \leq r + 1 = g_{B_r}(o, o)$ for all $r$ and all $x \in B_r$. By the Green function criterion, we conclude that Comb$(\mathbb{Z}, \alpha)$ has the infinite collision property for $\alpha \leq 1$. $\square$

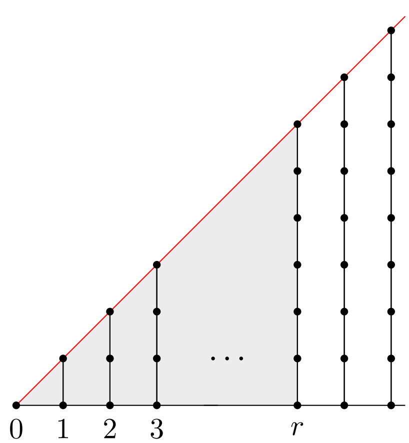



Figure 2: The graph Comb$(\mathbb{Z}, 1)$. The grey area shows the set $B_r$ used in the proof.

### 4.4 The finite collision property of $\mathrm{Comb}(\mathbb{Z}, \alpha)$ for $\alpha > 1$

The goal of this section is to show that $\mathrm{Comb}(\mathbb{Z}, \alpha)$ has the finite collision property for $\alpha > 1$. This means that if the teeth on the comb graph at horizontal coordinate $x$ have height $f(x) = x^\alpha$, then two random walks on this comb graph collide only finitely often almost surely. There is no general criterion known for the finite collision property of graphs. To prove this result, sufficiently accurate estimates of the transition density $q_t(x, y)$ are needed.

As in Barlow, Peres and Sousi [3], write $\alpha' = \alpha \wedge 2$ and $\beta' = \frac{1+\alpha'}{2+\alpha'}$. Note that $1 < \alpha' \leq 2$ and $\frac{2}{3} < \beta' \leq \frac{3}{4}$. Write $Q_{k,h} = \{(k, y) : 0 \leq y \leq h\}$. Let $Z_{k,h}$ be the number of collisions of the two random walks in $Q_{k,h}$ and let $\widetilde{Z}_{k,h} = Z_{k,2h/3} - Z_{k,h/3}$. After proving a bound for the transition density $q_t(x, y)$, the proof can be finished in the same way as the proof of the finite collision property of $\mathrm{Comb}(\mathbb{Z})$. This means that we find an upper bound for $\mathbb{E}[Z_{k,h}]$ and a lower bound for $\mathbb{E}[Z_{k,h} \mid \widetilde{Z}_{k,h}]$ and then apply Proposition 3.10.

Although we will present the same lemmas as in Barlow, Peres and Sousi [3], in some cases a slightly different proof is given. In some cases this is done to avoid the use of lemmas which are difficult to prove, or to make the lemmas easier to generalize.

The first lemma is very general and provides an upper bound for the transition density.

**Lemma 4.4.** Let $G = (V, E)$ be a graph and let $B$ be a subset of the vertex set $V$.
Then the following inequality holds:

$$q_t(x, x) \leq \frac{2 g_B(x, x)}{t \mathbb{P}_x(\tau_B \geq t)}.$$

*Proof.* Using the inequalities presented in Lemma 3.9, it follows that

$$\frac{2}{t} g_t(x, x) = \frac{2}{t} \sum_{j=0}^{t} q_j(x, x) \geq \frac{2}{t} \sum_{j=0}^{t/2} q_{2j}(x, x) \geq \frac{2}{t} \frac{t}{2} q_t(x, x) = q_t(x, x),$$

$$\frac{2}{t} g_t(x, x) = \frac{2}{t} \sum_{j=0}^{t} q_j(x, x) \geq \frac{2}{t} \sum_{j=0}^{\lfloor t/2 \rfloor} q_{2j}(x, x) \geq \frac{2}{t} \frac{t}{2} q_{2\lfloor t/2 \rfloor}(x, x) \geq q_t(x, x),$$

for even and odd $t$ respectively, since both sums contain $\lfloor t/2 \rfloor + 1 \geq t/2$ terms. This proves the inequality $\frac{2}{t} g_t(x, x) \geq q_t(x, x)$.

Let $\overline{B} = \{y \in V : \exists x \in B : x \sim y\}$. Since $B$ is finite and since the graph $G$ is locally finite, $\overline{B}$ is also finite. Write $\partial B = \overline{B} \setminus B$. By conditioning on $\tau_B$ and using the strong Markov property, it follows that the transition density $q_j$ satisfies:

$$q_j(x, x) = q_j^B(x, x) + \sum_{k=0}^{j} \sum_{y \in \partial B} \left[ \mathbb{P}(\tau_B = k, X_k = y) q_{j-k}(y, x) \right].$$

This equality can now be used to bound $g_t(x,x)$. We have:

$$\begin{aligned}
g_t(x,x) &= \sum_{j=0}^{t} q_j(x,x) = \sum_{j=0}^{t}\left[q_j^B(x,x) + \sum_{k=0}^{j}\sum_{y\in\partial B}\left[\mathbb{P}(\tau_B = k, X_{\tau_B} = y)q_{j-k}(y,x)\right]\right] \\
&= \sum_{j=0}^{t} q_j^B(x,x) + \sum_{j=0}^{t}\sum_{k=0}^{j}\sum_{y\in\partial B}\left[\mathbb{P}(\tau_B = k, X_{\tau_B} = y)q_{j-k}(y,x)\right] \\
&\leq \sum_{j=0}^{\infty} q_j^B(x,x) + \sum_{y\in\partial B}\sum_{j=0}^{t}\sum_{k=0}^{j}\left[\mathbb{P}(\tau_B = k, X_{\tau_B} = y)q_{j-k}(y,x)\right] \\
&= g_B(x,x) + \sum_{y\in\partial B}\sum_{\ell=0}^{t}\sum_{m=0}^{t-\ell}\left[\mathbb{P}(\tau_B = m, X_{\tau_B} = y)q_{\ell}(y,x)\right] \\
&\leq g_B(x,x) + \sum_{y\in\partial B}\sum_{\ell=0}^{t}\sum_{m=0}^{t-1}\left[\mathbb{P}(\tau_B = m, X_{\tau_B} = y)q_{\ell}(y,x)\right] \\
&= g_B(x,x) + \sum_{y\in\partial B}\sum_{\ell=0}^{t}\left[\mathbb{P}(\tau_B < t, X_{\tau_B} = y)q_{\ell}(y,x)\right] \\
&= g_B(x,x) + \sum_{y\in\partial B}\mathbb{P}(\tau_B < t, X_{\tau_B} = y)g_t(y,x) \\
&\leq g_B(x,x) + \sum_{y\in\partial B}\mathbb{P}(\tau_B < t, X_{\tau_B} = y)g_t(x,x) \\
&= g_B(x,x) + g_t(x,x)\sum_{y\in\partial B}\mathbb{P}(\tau_B < t, X_{\tau_B} = y) \\
&= g_B(x,x) + g_t(x,x)\mathbb{P}(\tau_B < t).
\end{aligned}$$

In the first line, we apply the definition and the equality for $q_j(x,x)$ just shown. In the second step, the summations are split. In the third line, we upper bound the first term by changing the summation limit to $\infty$ and we interchange the summations in the second term, which is allowed since all terms are positive. In the fourth line, the definition of $g_B(x,x)$ is applied and the summation variables are changed. In the fifth line, we change the summation limit from $t-\ell$ to $t-1$, which yields an upper bound unless $\ell = 0$. Since $x \in B$ and $y \notin B$, we have $x \neq y$ and hence $q_\ell(y,x) = 0$, and hence this step is also valid for $\ell = 0$. In the sixth line, we sum over $m$. In the seventh line, we use that $\mathbb{P}(\tau_B < t, X_{\tau_B} = y)$ can be taken out of the summation since it does not depend on $\ell$, and then use the definition of $g_t(y,x)$. In the eighth line, we apply the inequality $g_t(y,x) \leq g_t(x,x)$, which is Lemma 3.8. In the ninth line, $g_t(x,x)$ is taken out of the summation since it does not depend on $y$, and finally we sum over $y$ in the tenth line.

This proves the inequality $g_t(x,x) \leq g_B(x,x) + g_t(x,x)\mathbb{P}(\tau_B < t)$, which can be rearranged to $\mathbb{P}(\tau_B \geq t)g_t(x,x) \leq g_B(x,x)$. Combining this with $q_t(x,x) \leq \frac{2}{t}g_t(x,x)$ yields,

$$\tfrac{t}{2}q_t(x,x)\mathbb{P}(\tau_B \geq t) \leq \mathbb{P}(\tau_B \geq t)g_t(x,x) \leq g_B(x,x),$$

which can be rearranged to the required inequality. $\square$

Let $H(a,b) := \{(x,y) \in V : a \leq x \leq b\}$. This is the set of vertices with a horizontal coordinate between $a$ and $b$. For notational convenience, write $\mathbb{P}_k$ for $\mathbb{P}_{(k,0)}$.

**Lemma 4.5.** Let $k \geq 0$, $k_1 \geq 1$ and let $T = \tau_{B(k-k_1,k+k_1)}$ be the first exit of $X$ from the set $B(k-k_1, k+k_1)$. Then there exist constants $c, c'$ such that

$$\mathbb{P}_k(T \leq t) \leq c \exp\left(-c'\left(k_1^{2+\alpha'}/t\right)^{1/3}\right).$$

*Proof.* We first prove this inequality in the case $k_1 \leq k$.
For $t < k_1$ the inequality is trivial since then the left hand side is 0. We now prove that the inequality holds for $t < \widetilde{c}_1 k_1^{2-\alpha'/2}$ by only considering the horizontal steps made by $X$, for some constant $\widetilde{c}_1$ to be chosen later. Let $U$ be the random walk on $\mathbb{Z} \cap [k-k_1, k+k_1]$, corresponding to the horizontal steps made by $X$. By Lemma 3.11, it follows that

$$\mathbb{P}_k\left(T \leq t\right) \leq 4e^{-\frac{1}{6}k_1^2/t} \leq c \exp\left(-c'\left(k_1^{2+\alpha'}/t\right)^{1/3}\right),$$

since $\left(k_1^{2+\alpha'}/t\right)^{1/3} = k_1^{2/3+\alpha'/3}/t^{1/3} \leq k_1^{2/3+\alpha'/3}\left(\widetilde{c}_1 k_1^{2-\alpha'/2}\right)^{2/3}/t = \widetilde{c}_1^{2/3} k_1^2/t$, so by choosing the constant $c'$ to be at most $\frac{1}{6}\widetilde{c}_1^{-2/3}$ the second inequality holds.
If $t \geq \widetilde{c}_2 k_1^{2+\alpha'}$ for some constant $\widetilde{c}_2$ to be chosen later, then the inequality holds by choosing $c \geq 2$ and $c' \leq \frac{1}{2}\widetilde{c}_2^{-1/3}$. From now on, we can assume that the time $t$ satisfies $\widetilde{c}_1 k_1^{2-\alpha'/2} \leq t \leq \widetilde{c}_2 k_1^{2+\alpha'}$.

Let $L$ be the number of horizontal steps that the random walk makes until it leaves the set $H(k-k_1/2, k+k_1/2)$. Let $\lambda > 0$ and $\theta \leq \frac{1}{4}$ be constants, which will be chosen later. Then

$$\mathbb{P}_k\left(T \leq t\right) \leq \mathbb{P}_k\left(L \leq k_1^2/\lambda\right) + \mathbb{P}_k\left(T \leq t, L \geq k_1^2/\lambda\right)$$

By Lemma 3.10, we have $\mathbb{P}_k\left(L \leq k_1^2/\lambda\right) \leq 4e^{-\frac{1}{6}(k_1/2)^2/(k_1^2/\lambda)} = 4e^{-\lambda/24}$.

Let $n = \theta k_1^{\alpha'}$, and assume that $\theta$ is chosen such that $n \geq 1$. At each meeting of $X$ with a vertex $(k,0)$ we perform an independent experiment, where we succeed if we hit $n$ on the tooth and then spend at least $n^2$ steps on this tooth. Since $k - k_1/2 \geq k_1/2$, it follows that each tooth has length at least $\left(\frac{k_1}{2}\right)^{\alpha'} \geq \frac{1}{4}k_1^{\alpha'} \geq n$, so there is enough room in every tooth in $H(k-k_1/2, k+k_1/2)$. The independence of the experiments follows from the strong Markov property. Note that the probability of starting a walk on the tooth is $\frac{1}{3}$. If the random walk makes the transition from $(k,0)$ to $(k,1)$, then by Lemma 3.14, there is a probability of $\lceil n \rceil^{-1} \leq \frac{1}{2}n^{-1}$ of reaching height $n$ before returning to 0. By Lemma 3.12 is a probability lower bounded by a constant $c_1$ that $X$ takes at least $c_2 n^2 = c_2\theta^2 k_1^{2\alpha'}$ steps on the tooth. Let $c_1 = \frac{1}{6}c_1'$. Combining this gives a success probability of at least $c_1 n^{-1}$, since by the strong Markov property these three events are independent so the probabilities can be multiplied.
Hence, the number of successes is binomially distributed with at least $k_1^2/\lambda$ trials and success probability at least $c_1 n^{-1}$. Denote such random variable by $\mathrm{Bin}\left(k_1^2/\lambda,\, c_1 n^{-1}\right)$.

Let $\gamma = t/k_1^{2+\alpha'}$ and take $\lambda = \gamma^{-1/3}$ and $\theta = \frac{2}{c_1c_2}\gamma^{2/3}$. Note that this choice of $\lambda$ implies that

$$\mathbb{P}_k\left(L \le k_1^2/\lambda\right) \le 4e^{-\lambda/24} \le 4e^{-c'\gamma^{-1/3}}.$$

Note that

$$n = \frac{2}{c_1c_2}\gamma^{2/3}k_1^{\alpha'} = \frac{2}{c_1c_2}t^{2/3}k_1^{\alpha'/3-4/3} \ge \frac{2}{c_1c_2}\widetilde{c}_1^{2/3}$$

since $t \ge \widetilde{c}_1 k_1^{2-\alpha'/2}$, so by taking $\widetilde{c}_1 \ge (c_1c_2)^{3/2}$ this ensures that $n \ge 1$. Moreover, $\gamma \le \widetilde{c}_2$ since $t \le \widetilde{c}_2 k_1^{2+\alpha'}$, so by taking $\widetilde{c}_2 \le (c_1c_2/8)^{3/2}$ this ensures that $\theta \le \frac{1}{4}$.

Write $N = k_1^2/\lambda$, $p = c_1n^{-1}$ and $s = \frac{t}{c_2n^2}$. We compute:

$$\frac{s}{Np} = \frac{tc_2^{-1}n^{-2}}{k_1^2/\lambda c_1 n^{-1}} = \frac{t\lambda}{k_1^2c_1c_2\theta k_1^{\alpha'}} = \frac{t\lambda}{2\gamma^{2/3}k_1^{2+\alpha'}} = \frac{1}{2},$$
$$Np = k_1^2/\lambda c_1n^{-1} = k_1^2\gamma^{1/3}c_1\theta^{-1}k_1^{-\alpha'} = \tfrac{1}{2}c_1^2c_2\gamma^{-1/3}k_1^{2-\alpha'}.$$

Hence, $s = \frac{1}{2}Np$ and $Np = \frac{c_1^2c_2}{2}\gamma^{-1/3}k_1^{2-\alpha'}$.

For $-1 \le \mu \le 0$ we have $e^\mu - 1 \le \left(1 - \frac{1}{e}\right)\mu$. The moment generating function of a binomially distributed random variable with $N$ trials and success probability $p$ is

$$M(\mu) = (1 + p(e^\mu - 1))^N \le \left(1 + (1 - \tfrac{1}{e})\mu p\right)^N \le \exp\left(\left(1 - \tfrac{1}{e}\right)\mu Np\right).$$

Note that $T$ is at least $c_2n^2$ times the number of successes, since with each success at least $c_2n^2$ steps are done on a tooth and all successes occur before time $T$. If $L \ge N$, it therefore follows that $T$ stochastically dominates $c_2n^2\,\mathrm{Bin}\,(N,p)$. In particular, the event $T \le t$ is less likely than the event $\mathrm{Bin}\,(N,p) \le \frac{t}{c_2n^2} = s$. By Chernoff's bound with $\mu = -1$, we now find that

$$\begin{aligned}\mathbb{P}\left(T \le t,\ L \ge N\right) &\le \mathbb{P}\left(\mathrm{Bin}\,(N,\,p) \le s\right) \le e^s\exp\left(\left(\tfrac{1}{e} - 1\right)Np\right) = \exp\left(\left(\tfrac{1}{e} - \tfrac{1}{2}\right)Np\right)\\ &= \exp\left(-\mu'\tfrac{1}{2}c_1^2c_2\gamma^{-1/3}k_1^{2-\alpha'}\right) = \exp\left(-c'\gamma^{-1/3}k_1^{2-\alpha'}\right) \le \exp\left(-c'\gamma^{-1/3}\right),\end{aligned}$$

where $\mu' = \frac{1}{2} - \frac{1}{e} > 0$ and where it is used that $k_1^{2-\alpha'} \ge 1$. Hence, it follows that both terms in $\mathbb{P}_k\left(T \le t\right) \le \mathbb{P}_k\left(L \le k_1^2/\lambda\right) + \mathbb{P}_k\left(T \le t, L \ge k_1^2/\lambda\right)$ are bounded by a term of the form

$$c\exp\left(-c'\gamma^{-1/3}\right) = c\exp\left(-c'\left(k_1^{2+\alpha'}/t\right)^{1/3}\right),$$

and hence the result follows in the case $k_1 \le k$.

If $k_1 > k$, then $B(k - k_1, k + k_1) = B(0, k + k_1)$ and in that case the random walk can only escape at $k + k_1$. Since $\frac{k+k_1}{2} > k$, the random walk must then first visit $\frac{k+k_1}{2}$ before it can visit $k + k_1$. Let $k' = \lceil\frac{k+k_1}{2}\rceil$. Then $k' \ge k_1$, so

$$\mathbb{P}_k(T \le t) \le \mathbb{P}_{k'}(T \le t) \le c\exp\left(-c'\left(\left(k'\right)^{2+\alpha'}/t\right)^{1/3}\right) \le c\exp\left(-c''\left(k_1^{2+\alpha'}/t\right)^{1/3}\right),$$

so by modifying the constants the bound still holds. □

The previous two lemmas enable us to provide the first bound on the transition density.

**Lemma 4.6.** Let $u = (k, 0)$ be a vertex on the horizontal axis and let $t \geq 1$. Then $q_t(u, u) \leq \frac{c}{t^{\beta'}}$.

*Proof.* Let $k_1 = bt^{1/(\alpha'+2)}$, where $b \geq 1$ is a constant that will be chosen later.

We use Lemma 4.4 with $B = H(k - k_1, k + k_1)$. To upper bound $\mathbb{P}(\tau_B \leq t)$ we use Lemma 4.5 with this choice for $k$ and $k_1$. From this Lemma 4.5 it follows that

$$\mathbb{P}_u(\tau_B \leq t) \leq c \exp\left(-c'\left(k_1^{2+\alpha'}/t\right)^{1/3}\right) = c \exp\left(-c' b^{(2+\alpha')/3}\right).$$

By taking $b$ sufficiently large, it can be ensured that $\mathbb{P}_k(\tau_B \leq t) \leq \frac{1}{2}$.

By Lemma 3.7 and Lemma 3.5, it follows that

$$g_B(u, u) \leq R_{\text{eff}}(u, B^c) \leq R_{\text{eff}}(u, (k + k_1, 0)) = d((k, 0), (k + k_1, 0)) = k_1.$$

By Lemma 4.4 it now follows that $q_t(u, u) \leq \dfrac{2g_B(u, u)}{t\mathbb{P}_u(\tau_B \geq t)} \leq 4k_1 t^{-1} = 4bt^{-\frac{\alpha'+1}{\alpha'+2}} \leq ct^{-\beta'}$. □

For the following two lemmas, the condition is each time implied by the previous lemma. However, formulating them more generally allows directly reusing these lemmas in the next chapter. Write 0 for the vertex $(0, 0)$.

**Lemma 4.7.** Assume that $q_t(u, u) \leq \frac{c}{t^{\beta'}}$ for all $u$ on the horizontal axis and all $t \geq 1$. Then $q_t(0, u) \leq \frac{c'}{t^{\beta'}}$ for all $t \geq 1$ and all points $u = (k, 0)$ on the horizontal axis.

*Proof.* Let $t$ be given and let $B$ be the union of the balls of radius $t$ around 0 and $u$. Then a random walk starting from 0 or starting from $u$ remains in $B$ for the first $t$ steps. Hence, $q_t(0, u) = q_t^B(0, u)$ and $q_t(0, 0) = q_t^B(0, 0)$ and $q_t(u, u) = q_t^B(u, u)$.

If $t = 2s$ is even, then it follows by the Cauchy-Schwarz inequality for sequences and the spectral decomposition (Lemma 3.9) that

$$\begin{aligned}
q_t(0, u) = q_{2s}^B(0, u) &= \sum_{i=1}^{|B|} \lambda_i^{2s} \varphi_i(0)\varphi_i(u) = \sum_{i=1}^{|B|} \left(\lambda_i^s \varphi_i(0)\right)\left(\lambda_i^s \varphi_i(u)\right) \\
&\leq \sqrt{\sum_{i=1}^{|B|} \lambda_i^{2s} \varphi_i(0)^2} \sqrt{\sum_{i=1}^{|B|} \lambda_i^{2s} \varphi_i(u)^2} \leq \sqrt{\sum_{i=1}^{|B|} \lambda_i^{t} \varphi_i(0)^2} \sqrt{\sum_{i=1}^{|B|} \lambda_i^{t} \varphi_i(u)^2} \\
&= \sqrt{q_t^B(0, 0)}\sqrt{q_t^B(u, u)} = \sqrt{q_t(0, 0)}\sqrt{q_t(u, u)} \leq \frac{c}{t^{\beta'}}.
\end{aligned}$$

Hence, the result holds for even $t$.

If $t = 1$, then this inequality holds by choosing $c' \geq 1$.
If $t = 2s + 1$ is odd and $t > 1$, then it similarly follows that

$$\begin{aligned} q_t(0,u) &= q^B_{2s+1}(0,u) = \sum_{i=1}^{|B|} \lambda_i^{2s+1}\varphi_i(0)\varphi_i(u) = \sum_{i=1}^{|B|} \left(\lambda_i^s \varphi_i(0)\right)\left(\lambda_i^{s+1}\varphi_i(u)\right) \\ &\leq \sqrt{\sum_{i=1}^{|B|} \lambda_i^{2s}\varphi_i(0)^2}\sqrt{\sum_{i=1}^{|B|} \lambda_i^{2s+2}\varphi_i(u)^2} \leq \sqrt{\sum_{i=1}^{|B|} \lambda_i^{t-1}\varphi_i(0)^2}\sqrt{\sum_{i=1}^{|B|} \lambda_i^{t+1}\varphi_i(u)^2} \\ &= \sqrt{q^B_{t-1}(0,0)}\sqrt{q^B_{t+1}(u,u)} = \sqrt{q_{t-1}(0,0)}\sqrt{q_{t+1}(u,u)} \leq \frac{c}{(t^2-1)^{\beta'/2}} \leq \frac{c'}{t^{\beta'}}. \end{aligned}$$

Hence, the lemma holds for all $t \geq 1$. □

This lemma is now extended to vertices that are not on the horizontal axis.

**Lemma 4.8.** Assume that $q_t(0,(k,0)) \leq ct^{-\beta}$ for all times $t \geq 1$ and all points $(k,0)$ on the horizontal axis. Then $q_t(0,(k,h)) \leq c'' t^{-\beta} e^{-h^2/(c't)}$ for all times $t \geq 1$ and all points $(k,h)$.

*Proof.* Let $T_A$ be the first hitting time of the set $A$ for a simple random walk $(S_n)_{n\in\mathbb{N}}$ on $\mathbb{Z}$ starting from $S_0 = 0$. By the ballot theorem (Theorem 3.3) and Lemma 3.13, it follows that

$$\mathbb{P}_h(T_0 = s) = \frac{h}{s}\mathbb{P}_h(S_s = 0) \leq c\frac{h}{s}\frac{1}{\sqrt{s}}e^{-h^2/(c's)},$$

where the first inequality follows from the ballot theorem by considering the reverse problem: the probability that a random walk starting from 0 revisits 0 when taking $s$ steps to reach $h$ is equal to $\frac{h}{s}$. Reversing the paths only gives a factor $\frac{\deg((k,0))}{\deg((k,h))}$ to the probabilities of each of the paths, so probability that the reverse path, which starts from $s$, visits 0 only once as the last vertex in the path, is also $\frac{h}{s}$, so this proves the first equality.

Note that $e^{x^2/(2c't)} \geq 1 + \frac{x^2}{2c't} \geq 2\sqrt{\frac{x^2}{2c't}} = c''\frac{x}{\sqrt{t}}$. Hence, there exists a constant $c''$ such that $\sqrt{t}e^{-x^2/(2c't)} \geq c'' x e^{-x^2/(c't)}$. This gives $\mathbb{P}_h(T_0 = s) \leq c/s \cdot e^{-h^2/(c's)}$, for different constants $c, c'$.

Let $T^b_a$ be the first hitting time of $a$ of a simple random walk restricted to the interval $[a,b]$. A simple random walk restricted to the interval $[a,b]$ can be generated from a simple random walk restricted to the interval $[a, 2b-a]$ by identifying $b-x$ and $b+x$ for all $1 \leq x \leq b-a$. In that case, the hitting time of $a$ becomes the hitting time of the set $\{a, 2b-a\}$. Hence, we have:

$$\begin{aligned} \mathbb{P}_h(T^m_0 = s) &= \mathbb{P}_{h-m}(T^0_{-m} = s) = \mathbb{P}_{h-m}(T_{\{-m,m\}} = s) \\ &\leq \mathbb{P}_{h-m}(T_{-m} = s) + \mathbb{P}_{h-m}(T_m = s) = \mathbb{P}_h(T_0 = s) + \mathbb{P}_{2m-h}(T_0 = s) \end{aligned}$$

by translating or reflecting the intervals a number of times. Note that $h$ belongs to the interval $[0,m]$, so $h \leq m \leq 2m-h$. Hence, $c/s \cdot e^{-(2m-h)^2/(c's)} \leq c/s \cdot e^{-h^2/(c's)}$. So we can bound $\mathbb{P}_h(T^m_0 = s) \leq c/s \cdot e^{-h^2/(c's)}$ for yet another constant $c$.

Let $m$ denote the height of the tooth at position $k$. By Lemma 3.4, it holds that $q_t(0,(k,h)) = q_t((k,h),0)$. Since $q_t(0,(k,0)) \le ct^{-\beta}$, it now follows that

$$q_t((k,h),0) = \sum_{s=1}^{t-1} \mathbb{P}_h(T_0^m = s) q_{t-s}((k,0),0) \le c \sum_{s=1}^{t-1} 1/s \cdot e^{-h^2/(c's)} (t-s)^{-\beta}$$

for again another constant $c$. Hence,

$$\begin{aligned} q_t((k,h),0) &\le c e^{-h^2/(c't)} \sum_{s=1}^{t-1} 1/s \cdot (t-s)^{-\beta} \le c e^{-h^2/(c't)} \int_0^t s^{-1}(t-s)^{-\beta} \mathrm{d}s \\ &= ct^{-\beta} e^{-h^2/(c't)} \int_0^1 u^{-1}(1-u)^{-\beta} \mathrm{d}u = c'' t^{-\beta} e^{-h^2/(c't)}. \end{aligned}$$

The sum can be bounded by the integral since $1/s \cdot (t-s)^{-\beta}$ is decreasing for $1 < s \le t/(\beta+1)$ and increasing for $s \ge t/(\beta+1)$ and hence the sum from 1 to $\lfloor t/(\beta+1) \rfloor$ can be upper bounded by the integral from 0 to $\lfloor t/(\beta+1) \rfloor$ whereas the sum from $\lceil t/(\beta+1) \rceil$ to $t-1$ can be upper bounded by the integral from $\lceil t/(\beta+1) \rceil$ to $t$. In the integral, we make the change of variables $u = s/t$. Finally, observing that the resulting integral is a constant implies the lemma. $\square$

Finally, a lemma is needed to get a better bound for small $t$.

**Lemma 4.9.** Let $x = (k,0)$. Then if $t < k^{2+\alpha'}$, we have

$$q_t(0,x) \le ck^{-(\alpha'+1)} = c\left(k^{2+\alpha'}\right)^{-\beta}.$$

*Proof.* Let $m$ be an integer within distance 1 of $k/2$. Let $T$ be the first hitting time of $(m,0)$ for a simple random walk $X$ on the comb graph starting from $X_0 = 0$. Then

$$\mathbb{P}_0(X_t = x) = \mathbb{P}_0(X_t = x, T_m \le t/2) + \mathbb{P}_0(X_t = x, T_m \ge t/2).$$

Consider a path $(p_0, \ldots, p_t)$ such that $p_0 = 0$, $p_t = x$ and $p_i \ne (m,0)$ for all $i \le t/2$. Since $0 \le m \le k$, the path must visit $(m,0)$, so $p_i = (m,0)$ for some $i > t/2$. Then the reverse of this path $(p'_0, \ldots, p'_t) = (p_t, \ldots, p_0)$ satisfies $p_0 = x$, $p_t = 0$ and $p_i = (m,0)$ for some $i < t/2$, so in particular for this path the hitting time of $m$ is smaller than $t/2$.
Reversing the paths only gives a factor $\frac{\deg((0,0))}{\deg((k,0))}$ to the probabilities of each of the paths, which is bounded by a constant $c$. Hence, by summing this over all possible paths satisfying $p_0 = 0$, $p_t = x$ and $p_i \ne (m,0)$ for all $i \le t/2$, we obtain

$$\mathbb{P}_0(X_t = x, T_m > t/2) \le c\mathbb{P}_x(X_t = 0, T_m \le t/2).$$

Since $T_m$ is a stopping time, it follows from the strong Markov property that

$$\mathbb{P}_0(X_t = x, T_m \le t/2) \le \mathbb{P}_0(T_m \le t/2) \max_{0 \le s \le t/2} \mathbb{P}_m(X_{t-s} = x).$$

Since $t-s \geq t/2$, it follows that the second factor is bounded by $ct^{-\beta'}$. The first term is bounded by Lemma 4.5. Write $\eta = k^{2+\alpha'}/t$. This yields

$$
\begin{aligned}
\mathbb{P}_0(X_t = x, T_m \leq t/2) &\leq ct^{-\beta'} \exp\left(-c'\left(k_1^{2+\alpha'}/t\right)^{1/3}\right) \\
&\leq ck^{-(\alpha'+1)}\eta^{-\beta'} \exp\left(-c'\eta^{1/3}\right) \\
&\leq ck^{-(\alpha'+1)} \sup_{\eta>0}\left(\eta^{-\beta'} \exp\left(-c'\eta^{1/3}\right)\right) \leq c''k^{-(\alpha'+1)}.
\end{aligned}
$$

Since $m$ is an integer within distance 1 of $k/2$, the term $\mathbb{P}_x(X_t = 0, T_m \leq t/2)$ can be bounded in exactly the same way as the term as $\mathbb{P}_0(X_t = x, T_m \leq t/2)$. So both terms are bounded by terms of the form $ck^{-(\alpha'+1)}$, and hence $q_t(0,x) = c'\mathbb{P}_0(X_t = x)$ is bounded by $ck^{-(\alpha'+1)}$ for some constant $c$. This completes the proof of the lemma. $\square$

These lemmas can now be combined to prove the following lemma on the transition density:

**Lemma 4.10.**
Let $x = (k,h) \in V$. The transition density $q$ satisfies

$$
q_t(0,x) \leq \begin{cases} ct^{-\beta} & \text{if } t \geq k^{2+\alpha'}, \\ c\left(k^{2+\alpha'}\right)^{-\beta} & \text{if } t \leq k^{2+\alpha'}. \end{cases}
$$

*Proof.* The case $t \geq k^{2+\alpha'}$ follows directly from combining Lemma 4.6, Lemma 4.7 and Lemma 4.8. If $t \leq k^{2+\alpha'}$, then by conditioning on the first hit of $(k,0)$ we find

$$
q_t(0,x) \leq \mathbb{P}_0(X_t = x) \leq \max_{0\leq s\leq t} \mathbb{P}_0(X_s = (k,0)) \leq c\left(k^{2+\alpha'}\right)^{-\beta}
$$

by Lemma 4.9. This completes the proof. $\square$

Recall that $Q_{k,h} = \{(k,y) : 0 \leq y \leq h\}$ and that $Z_{k,h}$ is the number of collisions of the two random walks in $Q_{k,h}$ and that $\widetilde{Z}_{k,h} = Z_{k,2h/3} - Z_{k,h/3}$. These notions allow us to formulate and prove the final lemma needed for proving the finite collision property of $\mathrm{Comb}(\mathbb{Z},\alpha)$ for $\alpha > 1$.

**Lemma 4.11.**

(a) $\mathbb{E}[Z_{k,h}] \leq chk^{-\alpha'}$.

(b) $\mathbb{E}[Z_{k,h} \mid \widetilde{Z}_{k,h} > 0] \geq ch$.

*Proof.* (a) By Lemma 4.10, we have

$$\mathbb{E}[Z_{k,h}] = \sum_{t=0}^{\infty} \sum_{x \in Q_{k,h}} q_t(0,x)^2 \leq \sum_{t=0}^{k^{2+\alpha'}} \frac{c'h}{k^{2(1+\alpha')}} + \sum_{t=k^{2+\alpha'}}^{\infty} \frac{c''h}{t^{2\beta'}} \leq \frac{c'hk^{2+\alpha'}}{k^{2+2\alpha'}} + \frac{c''h}{(k^{2+\alpha'})^{2\beta'-1}} \leq \frac{ch}{k^{\alpha'}},$$

since $(2+\alpha')(2\beta'-1) = (2+\alpha')\frac{\alpha'}{\alpha'+2} = \alpha'$. So $\mathbb{E}[Z_{k,h}] \leq chk^{-\alpha'}$.

(b) Since we consider the expectation conditioned on the event that $\{\widetilde{Z}_{k,h} > 0\}$, there is a collision at a point $x = (k,y)$ for some $h/3 \leq y \leq 2h/3$. Consider the two random walks from this point onwards. The total number of collisions that happen in the set $Q_{k,h}$ is at least the number of collisions that happen before one of the walks first exits $Q_{k,h}$ after the collision at point $x = (k,y)$. This can be modeled by considering the walks that are killed when exiting $Q_{k,h}$. By Corollary 4.2 and Lemma 3.7 it now follows that

$$\mathbb{E}[Z_{k,h} \mid \widetilde{Z}_{k,h} > 0] \geq \tfrac{1}{2} g_{Q_{k,h}}(x,x) = \tfrac{1}{2} R_{\text{eff}}\left(x, Q_{k,h}^c\right) \geq \tfrac{1}{2} \cdot \tfrac{2}{9} h \geq ch.$$

This proves the lemma. □

These lemmas together allow to prove the main result of this section:

**Theorem 4.4.** $\mathrm{Comb}(\mathbb{Z},\alpha)$ has the finite collision property for $\alpha > 1$.

*Proof.* Let $\alpha > 1$ be given. By Proposition 3.10 and Lemma 4.11, it follows that

$$\mathbb{P}(\widetilde{Z}_{k,h} > 0) \leq \frac{\mathbb{E}[Z_{k,h}]}{\mathbb{E}[Z_{k,h} \mid \widetilde{Z}_{k,h} > 0]} \leq \frac{c'hk^{-\alpha'}}{c''h} = ck^{-\alpha'},$$

so $\mathbb{P}(\widetilde{Z}_{k,h} > 0) \leq ck^{-\alpha'}$. Summing over all $k$ and all $j$ satisfying $2^j \leq k^\alpha$ yields

$$\sum_{k=0}^{\infty} \sum_{j=0}^{\alpha \log_2 k} \mathbb{P}(\widetilde{Z}_{k,2^j} > 0) \leq \sum_{k=0}^{\infty} \alpha \log_2(k) ck^{-\alpha'} < \infty$$

since $\alpha' > 1$. By the Borel-Cantelli lemma, it follows that $\mathbb{P}(\widetilde{Z}_{k,2^j} > 0 \text{ occurs i.o.}) = 0$. By Corollary 4.1, it follows that $\mathrm{Comb}(\mathbb{Z},\alpha)$ has the finite collision property. □

# 5 Comb graphs with random heights

In this chapter we consider comb graphs with finite random heights. Let $\alpha > 3$ be given and let $F_X : \mathbb{N} \to [0,1]$ be a cumulative distribution function satisfying $F_X(n) \leq 1 - n^{-1/\alpha}$ for all $n \in \mathbb{N}$. Moreover, assume that there exists a constant $E > 0$ such that $F_X(n) \geq 1 - n^{-E}$ for all $n \in \mathbb{N}$. Let $\{f(n)\}_{n\in\mathbb{Z}}$ be i.i.d. random variables with cdf $F_X$.
We now prove that $G = \mathrm{Comb}(\mathbb{Z}, f)$ has the finite collision property almost surely.

If we have a random walk $S$ on $\mathbb{Z} \cap [-k,k]$, then we expect that $S$ takes about $k^2$ steps before hitting $\pm k$. Lemma 3.10 implies that the probability that $S$ takes less than $k^{2-\varepsilon}$ steps before hitting $\pm k$ decays superpolynomially in $k$ for all $\varepsilon > 0$. If we have a typical subset of $[-k,k]$ of size $\ell$, then we similarly expect that $S$ visits this subset about $k\ell$ times before hitting $\pm k$ and that the probability that $S$ takes less than $\ell k^{1-\varepsilon}$ steps before hitting $\pm k$ decays superpolynomially in $k$. Here 'typical' means that the elements of the set are spread out over the entire interval $[-k,k]$. In particular, the elements are not all (or mostly) close to each other or close to $\pm k$.

It is beyond the scope of this thesis to prove this in full generality. We first prove it for centric sets, which are sets containing 0 such that the gap between successive elements increases with the distance to 0. The result can then be generalized to unions of centric sets that only have 0 in common. We first give the definition of a centric set.

**Definition 5.1. (Centric set)**
A set $D = \{d_{-\ell}, \ldots, d_{-1}, d_0, d_1, \ldots, d_\ell\} \subset [-k,k]$ is called centric if $d_0 = 0$, $d_i < d_j$ for $i < j$ and $d_i - d_{i-1} \geq d_{i+1} - d_i$ for $-\ell < i < -1$ and $d_i - d_{i-1} \leq d_{i+1} - d_i$ for $0 < i < \ell$. For a centric set, we call $m = \min\{d_1 - d_0, d_0 - d_{-1}\}$ the minimal gap and $\ell$ the length.

When considering the Markov chain $T$ on a centric set resulting from a random walk $S$ on $\mathbb{Z} \cap [-k,k]$, then it follows from the property that the gaps are increasing with the distance to 0 that $T$ is biased towards 0, which makes it easier to bound the probability of reaching $\pm k$ early. This is done in the following lemma.

**Lemma 5.1. (Centric sets are visited often with high probability)**
Consider a random walk $(S_t)_{t\geq 0}$ on $\mathbb{Z} \cap [-k,k]$ starting from $S_0 = 0$ and let $\tau_k$ be the first hitting time of $\pm k$. Let $D$ be a centric set with minimal gap at least $m \geq 3$ and length at least $\ell$. Let $0 < \varepsilon < 1$. Then

$$\mathbb{P}(\#\{t \leq \tau_k : S_t \in D \setminus \{0\}\} \leq m\ell^{2-\varepsilon}) \leq c \exp\left(-c'\ell^{\varepsilon}\right).$$

*Proof.* Write $D = \{d_{-\ell}, \ldots, d_{-1}, d_0, d_1, \ldots, d_\ell\}$ as in the definition of a centric set.
Let $D^+ = \{d_0, d_1, \ldots, d_\ell\}$ be the set of nonnegative elements of $D$, and identify the elements of $D^+$ with elements of $\mathbb{Z} \cap [0,\ell]$ by mapping $d_i$ to $i$ for $0 \leq i \leq \ell$.

Let $(T_r)_{r\geq 0}$ be the random walk on $\mathbb{Z}\cap[0,\ell]$ corresponding to the random walk on $S$. More precisely, define $t(0)=0$ and $t(r+1)=\inf\{s\geq t(r): S_s\in D^+\setminus\{S_{t(r)}\}\}$. Note that this infimum is finite almost surely since $S$ is recurrent. Then $S_{t(r)}=d_i$ for some $0\leq i\leq \ell$. If $S_{t(r)}=d_i$, then we define $T_r=i$. If $T_r=i$ for some $0<i<\ell$, then $T_{r+1}=i\pm 1$, since by definition revisits to $d_i$ are not counted, and the random walk $S$ visits either $d_{i-1}$ or $d_{i+1}$ before visiting any of the other $d_i$. So $T$ is a random walk on $\mathbb{Z}\cap[0,\ell]$, but not a simple random walk.

Let $e_i=d_i-d_{i-1}$. Then $e_{i+1}\geq e_i$ for $0<i<\ell$. Assume that $T_r=i$, then the random walk is at $d_i$. Translate the segment such that $d_i$ falls on 0. Then $d_{i-1}$ falls on $-e_i$, while $d_{i+1}$ falls on $e_{i+1}$. By Lemma 3.15.2, the random walk hits $d_{i-1}$ first with probability $\frac{e_{i+1}}{e_{i+1}+e_i}\geq\frac{1}{2}$ and it hits $d_{i+1}$ first with probability $\frac{e_i}{e_{i+1}+e_i}\leq\frac{1}{2}$.

We now generate a simple random walk $(\widetilde{T}_r)_{r\geq 0}$ on $\mathbb{Z}\cap[0,\ell]$ such that $\widetilde{T}_r\geq T_r$ for all $r\geq 0$ before the first time $\widetilde{T}_r$ hits $\ell$. Let $\widetilde{T}_0=0$. If $T_r\neq 0$ and $T_r$ increments by 1, which happens with probability at most $\frac{1}{2}$, then $\widetilde{T}_r$ also increments by 1. If $T_r\neq 0$ and $T_r$ decrements by 1, then $\widetilde{T}_r$ increments with a certain probability such that the overall probability that $\widetilde{T}_r$ increments is $\frac{1}{2}$, and decrement otherwise. If $T_r=0$ and $\widetilde{T}_r\neq 0$, then $\widetilde{T}_r$ simply chooses at random between incrementing and decrementing. If $T_r=\widetilde{T}_r=0$, then $\widetilde{T}_r$ moves to 1 with probability 1.

We now prove by induction on $r$ that $\widetilde{T}_r\geq T_r$ for all $r\geq 0$ before the first time $\widetilde{T}_r$ hits $\ell$, from which it also follows that this definition gives all possibilities for $\widetilde{T}_r$ before the first time $\widetilde{T}_r$ hits $\ell$. The base case is $T_0=\widetilde{T}_0=0$. Let now $s\in\mathbb{N}$ be given and assume that $\widetilde{T}_s\geq T_s$ and $\widetilde{T}_s\neq\ell$, from which it follows that $T_s\neq\ell$. If $T_s=0$, then $T_{s+1}=1$. Since $T_{s+1}\equiv\widetilde{T}_{s+1}\mod 2$, it follows that $\widetilde{T}_{s+1}$ cannot be 0 and hence $\widetilde{T}_{s+1}\geq 1=T_{s+1}$. Otherwise, we have $T_s\neq 0$ and hence also $\widetilde{T}_s\neq 0$. If $T_s$ increments, then $\widetilde{T}_s$ also increments, so $\widetilde{T}_{s+1}=\widetilde{T}_s+1\geq T_s+1=T_{s+1}$. Otherwise $T_s$ decrements, and then $\widetilde{T}_{s+1}\geq\widetilde{T}_s-1\geq T_s-1=T_{s+1}$. This completes the induction.

Since $\widetilde{T}_r\geq T_r$ for all $r\geq 0$ before the first time $\widetilde{T}_r$ hits $\ell$, it follows that the first hitting time $\tau_{T,+}$ of $\ell$ of $T_r$ stochastically dominates the hitting time of $\ell$ of $\widetilde{T}_r$. To provide a bound for the hitting time of $\ell$ of $\widetilde{T}_r$, we do one more transformation. The simple random walk on $\mathbb{Z}\cap[0,\ell]$ can be generated from a simple random walk on $\mathbb{Z}\cap[-\ell,\ell]$ by identifying $x$ and $-x$ with each other. Moreover, before the first hitting time of $\pm\ell$, the simple random walk on $\mathbb{Z}\cap[-\ell,\ell]$ coincides with a simple random walk on $\mathbb{Z}$. So the first hitting time of $\ell$ of $\widetilde{T}_r$ is identically distributed as the first hitting time $\tau_U$ of $\pm\ell$ of a simple random walk $(U_r)_{r\geq 0}$ on $\mathbb{Z}$. By Lemma 3.11, we have

$$\mathbb{P}(\tau_U\leq 4\ell^{2-\varepsilon})\leq 4e^{-\frac{1}{24}\ell^{\varepsilon}}\leq ce^{-c'\ell^{\varepsilon}}$$

for some constants $c,c'>0$. By the above discussion, it hence follows that the first hitting time $\tau_{T,+}$ of $\ell$ of $T_r$ satisfies $\mathbb{P}(\tau_{T,+}\leq 4\ell^{2-\varepsilon})\leq ce^{-c'\ell^{\varepsilon}}$.
So with a probability larger than $1-ce^{-c'\ell^{\varepsilon}}$, $S$ visits the set $D^+$ at least $4\ell^{2-\varepsilon}$ times, not counting consecutive revisits to the same element of $D^+$, before hitting $d_\ell$.

Using the same steps for the nonpositive elements of $D$, it follows that the first hitting time $\tau_{T,-}$ of $\ell$ of $T_r$ satisfies $\mathbb{P}(\tau_{T,-} \le 4\ell^{2-\varepsilon}) \le ce^{-c'\ell^\varepsilon}$ for possibly different constants $c, c' > 0$. So with a probability larger than $1 - ce^{-c'\ell^\varepsilon}$, $S$ visits the set $D^- = \{d_{-\ell}, \ldots, d_{-1}, d_0\}$ at least $2\ell^{2-\varepsilon}$ times, not counting consecutive revisits to the same element of $D^-$, before hitting $d_{-\ell}$.
In particular, with a probability larger than $1 - ce^{-c'\ell^\varepsilon}$ (with again different constants and using the union bound), $S$ visits the set $D$ at least $4\ell^{2-\varepsilon}$ times, not counting consecutive revisits to the same element of $D$, before hitting either $d_{-\ell}$ or $d_\ell$. Since $[d_{-\ell}, d_\ell] \subseteq [-k, k]$, it follows that with probability larger than $1 - ce^{-c'\ell^\varepsilon}$ (with again different constants and using the union bound), $S$ visits the set $D$ at least $4\ell^{2-\varepsilon}$ times, not counting consecutive revisits to the same element of $D$, before $\tau_k$, the hitting time of $\pm k$. Since at most half of the visits are to 0, $S$ visits the set $D \setminus \{0\}$ at least $2\ell^{2-\varepsilon}$ times, not counting consecutive revisits to the same element of $D$, before $\tau_k$, the hitting time of $\pm k$.

We now have to take into account that an element of $D$, say $d_i$, will be revisited a number of times before $S$ visits $d_{i-1}$ or $d_{i+1}$. By Lemma 3.15.2, the number of revisits is geometrically distributed with success probability $\frac{1}{2}\left(\frac{1}{e_i} + \frac{1}{e_{i+1}}\right)$. Note that $e_i \ge e_1$ for $i \ge 1$ and $e_i \ge e_0$ for $i \le 0$, so $e_i \ge \min\{e_0, e_1\} = m$ for all $-\ell < i < \ell$. Hence, the number of revisits stochastically dominates a geometrically distributed random variable with success probability $\frac{1}{m}$. Moreover, by the strong Markov property these random variables are independent.

Hence, it follows that with probability larger than $1 - ce^{-c'\ell^\varepsilon}$ the number of visits of $S$ to $D \setminus \{0\}$, counting revisits, stochastically dominates a sum $\widetilde{G}$ of $2\ell^{2-\varepsilon}$ independent geometrically distributed random variable with success probability $\frac{1}{m}$. The moment generating function of a geometric random variable with success probability $\frac{1}{m}$ is $(m - (m-1)e^t)^{-1}$, so the moment generating function of $\widetilde{G}$ is $M_{\widetilde{G}}(t) = \left(m - (m-1)e^t\right)^{-2\ell^{2-\varepsilon}}$.
By Chernoff's bound it follows that for $t < 0$ we have

$$\mathbb{P}(\widetilde{G} \le m\ell^{2-\varepsilon}) \le e^{-tm\ell^{2-\varepsilon}} \left(m - (m-1)e^t\right)^{-2\ell^{2-\varepsilon}} = \left(e^{tm}\left(m - (m-1)e^t\right)^2\right)^{-\ell^{2-\varepsilon}}.$$

For $m = 3$ and $m = 4$ choosing $t = \log(0.9) < 0$ yields $e^{tm}\left(m - (m-1)e^t\right)^2 > 1$.
For $m \ge 5$, let $t = \log\left(\frac{m}{m+1}\right)$. Then $e^{tm} = \left(1 - \frac{1}{m+1}\right)^m \ge \frac{1}{e}$ and $m - (m-1)e^t = \frac{2m}{m+1} \ge \frac{5}{3}$. Then $e^{tm}\left(m - (m-1)e^t\right)^2 > \frac{25}{9e} > 1$. So there exists a constant $c$ such that

$$\mathbb{P}(\widetilde{G} \le m\ell^{2-\varepsilon}) \le e^{-c\ell^{2-\varepsilon}}.$$

Since $\varepsilon < 1$, this term is dominated by a term of the form $ce^{-c'\ell^\varepsilon}$. So with probability larger than $1 - ce^{-c'\ell^\varepsilon}$ the number of visits of $S$ to $D \setminus \{0\}$ stochastically dominates $\widetilde{G}$ and $\mathbb{P}(\widetilde{G} \ge m\ell^{2-\varepsilon})$. Hence, we conclude that $\mathbb{P}(\#\{t \le \tau_k : S_t \in D \setminus \{0\}\} \le m\ell^{2-\varepsilon}) \le ce^{-c'\ell^\varepsilon}$. $\square$

**Corollary 5.1. (Collections of centric sets are visited often with high probability)** Consider a random walk $(S_t)_{t\geq 0}$ on $\mathbb{Z}\cap[-k,k]$ starting from $S_0 = 0$ and let $\tau_k$ be the first hitting time of $\pm k$. Let $\mathcal{D} = \{D_1,\dots,D_n\}$ be a collection of $n$ centric sets with minimal gap at least $m \geq 3$ and length at least $\ell$ such that $D_i \cap D_j = \{0\}$ for all $1 \leq i,j \leq n$ with $i \neq j$. Let $0 < \varepsilon < 1$. Then

$$\mathbb{P}(\#\{t \leq \tau_k : S_t \in \bigcup_{D\in\mathcal{D}} D \setminus \{0\}\} \leq nm\ell^{2-\varepsilon}) \leq cn\exp\left(-c'\ell^{\varepsilon}\right).$$

*Proof.* By Lemma 5.1, we have

$$\mathbb{P}(\#\{t \leq \tau_k : S_t \in D_i \setminus \{0\}\} \leq m\ell^{2-\varepsilon}) \leq c\exp\left(-c'\ell^{\varepsilon}\right).$$

Since $D_1 \setminus \{0\},\dots,D_n\setminus\{0\}$ are mutually disjoint, it follows that

$$\#\{t \leq \tau_k : S_t \in \bigcup_{D\in\mathcal{D}} D \setminus \{0\}\} = \sum_{i=1}^{n} \#\{t \leq \tau_k : S_t \in D_i \setminus \{0\}\}.$$

Hence, $\#\{t \leq \tau_k : S_t \in \bigcup_{D\in\mathcal{D}} D\setminus\{0\}\} \leq nm\ell^{2-\varepsilon}$ implies that $\#\{t \leq \tau_k : S_t \in D_i\setminus\{0\}\} \leq m\ell^{2-\varepsilon}$ for at least one $1 \leq i \leq n$. By the union bound, it follows that

$$\mathbb{P}(\#\{t \leq \tau_k : S_t \in \bigcup_{D\in\mathcal{D}} D \setminus \{0\}\} \leq \sum_{i=1}^{n} \mathbb{P}(\#\{t \leq \tau_k : S_t \in D_i \setminus \{0\}\} \leq cn\exp\left(-c'\ell^{\varepsilon}\right),$$

which is the required result. □

We now attach independent random variables to the elements of the set $\mathbb{Z}\cap[-k,k]$. These will be interpreted as the heights of the teeth of a comb graph. The following lemma shows that with high probability we can construct a collection of centric sets such that at all elements in these centric sets, the realization of the random variable is large.

**Lemma 5.2.**
Let $p,q \in [0,1]$ and let $k,\ell$ and $m$ be positive integers. Assume that $m \geq \sqrt{k/(pq)}$ and $4m\ell \leq k$. Let $(H_i)_{i\in\mathbb{Z}\cap[-k,k]}$ be independent random variables with cdf $F_i$. Then with probability exceeding $1-2\ell^2\exp\left(-c'/q\right)$ there exists a collection $\mathcal{D} = \{D_1,\dots,D_\ell\}$ of $\ell$ centric sets with minimal gap at least $m$ and length at least $\ell$ such that $D_i\cap D_j = \{0\}$ for all $1 \leq i,j \leq \ell$ with $i \neq j$ and such that $H_d \geq (F_d)^{-1}(1-p)$ for all $d \in \bigcup_{D\in\mathcal{D}} D\setminus\{0\}$.

*Proof.* For a centric set $D_i$, let $D_i^+$ denote the set of nonnegative elements of $D_i$. We first construct the nonnegative parts $D_i^+$ of the centric sets. Let $d_{i,j}$ denote the $j$th element of the $i$th centric set that has to be constructed, for $1 \leq i,j \leq \ell$. By definition, $d_{i,0} = 0$ for all $1 \leq i \leq \ell$. Let $n = \left\lceil \frac{1}{pq} \right\rceil$. We choose the $d_{i,j}$ such that

$$jm + (j-1)n\ell + (j-1)^2 n + (i-1)n \leq d_{i,j} < jm + (j-1)n\ell + (j-1)^2 n + in.$$

For constant $j$, we have $d_{i,j} < d_{i+1,j}$. In particular, for a fixed $j$, all $d_{i,j}$ are different. Moreover,

$$d_{1,j+1} \geq (j+1)m + jn\ell + j^2 n > jm + (j-1)n\ell + n\ell + (j-1)^2 n > d_{\ell,j},$$

which implies that all $d_{i,j}$ are different for $j > 0$. We now show that $d_{i,j} - d_{i,j-1} \leq d_{i,j+1} - d_{i,j}$ by using these upper and lower bounds for the $d_{i,j}$. We have

$$\begin{aligned} d_{i,j} - d_{i,j-1} &\leq jm + (j-1)n\ell + (j-1)^2 n + in - ((j-1)m + (j-2)n\ell + (j-2)^2 n + (i-1)n) \\ &= m + n\ell + (2j-3)n + n = m + n\ell + (2j-1)n - n \\ &= (j+1)m + jn\ell + j^2 n + (i-1)n - (jm + (j-1)n\ell + (j-1)^2 n + in) \\ &\leq d_{i,j+1} - d_{i,j}, \end{aligned}$$

as required. Hence, the $D_i^+$ can be the nonnegative parts of a centric set. We now show that the minimal gap of the nonnegative part is indeed at least $m$. This follows from

$$m \leq m + (i-1)n \leq d_{i,1} - d_{i,0}.$$

Finally, we have to show that the largest element constructed in this way is not larger than $k$. Note that $n = \left\lceil \frac{1}{pq} \right\rceil \leq \frac{2}{pq}$ since $pq \leq 1$. Hence,

$$n\ell \leq \frac{2}{pq}\frac{k}{4m} \leq \frac{k}{2pq\sqrt{k/(pq)}} = \frac{1}{2}\sqrt{\frac{k}{pq}} \leq \frac{1}{2}m.$$

The largest value among the $d_{i,j}$ is $d_{\ell,\ell}$, which is upper bounded by

$$\ell m + (\ell-1)n\ell + (\ell-1)^2 n + \ell n \leq \ell m + 2\ell^2 n \leq 2m\ell < k.$$

Hence, the nonnegative parts of the centric set have minimal gap at least $m$ and length $\ell$. Hence, the collection of centric sets chosen in this way has the required properties.

Now choose the $d_{i,j}$ such that the inequality is satisfied and such that $F_{d_i}(H_{d_i})$ is maximal among all $n$ possible values of $d_{i,j}$. Note that $F_k(H_k)$ has the uniform distribution on $[0,1]$. Since $H_d \geq (F_d)^{-1}(1-p)$ is equivalent to $F_d(H_d) \geq 1-p$, it follows that the probability that $H_d \geq (F_d)^{-1}(1-p)$ is $p$. In particular, the probability that there is no $k$ among the $n \geq 1/(pq)$ possible choices satisfying $H_d \geq (F_d)^{-1}(1-p)$, is equal to $(1-p)^n \leq (1-p)^{1/(pq)} \leq \exp(-c/q)$. By the union bound, the probability that there exists a pair $(i,j)$ with $1 \leq i,j \leq \ell$ where it is not possible to choose $d_i$ such that $H_d \geq (F_d)^{-1}(1-p)$ is at most $\ell^2 \exp(-c/q)$.

In particular, with probability exceeding $1 - \ell^2 \exp(-c/q)$ we can construct such a collection of sets $D^+$ with the required property and similarly with probability exceeding $1 - \ell^2 \exp(-c/q)$ (with possibly a different constant) we can construct such a set $D^-$ satisfying the constraints. By the union bound it follows that with probability exceeding $1 - 2\ell^2 \exp(-c/q)$ we can construct a collection of centric sets with all required properties. □

We now have proven all lemmas on centric sets that we need. For proving that $\mathrm{Comb}(\mathbb{Z}, f)$ has the finite collision property almost surely, we use an approach similar to Barlow, Peres and Sousi [3]. Note that Lemma 4.4 can directly be reused. As in [3], write $Q_{k,h} = \{(k, y) : 0 \le y \le h\}$. Let $Z_{k,h}$ be the number of collisions of the two random walks in $Q_{k,h}$ and let $\widetilde{Z}_{k,h} = Z_{k,2h/3} - Z_{k,h/3}$. Write $B(a, b) := \{(x, y) \in G : a \le x \le b\}$. We now prove a lemma analogous to Lemma 4.5. Write $\mathbb{P}_k$ for $\mathbb{P}_{(k,0)}$. Let $X$ be a random walk on the comb graph.

**Lemma 5.3.**
Let $k \in \mathbb{Z}$, $k_1 \ge 1$ and let $T = \tau_{B(k-k_1,k+k_1)}$ be the first exit of $X$ from $B(k - k_1, k + k_1)$. Then there exist constants $\alpha' > 1$ and $\alpha'' > 0$ (which can depend on $\alpha$) such that

$$\mathbb{P}_k(T \le t) \le c \exp\left(-c' \left(k_1^{2+\alpha'}/t\right)^{\alpha''}\right).$$

*Proof.* Let $\delta = \frac{\alpha-3}{2\alpha+2} > 0$. Note that $\delta < \frac{\alpha+1}{2\alpha+2} \le \frac{1}{2}$. Let $\alpha' = 1 + \delta/2$ and $\alpha'' = \delta/32$. By choosing the constants $c, c'$ suitably, the inequality holds for all sufficiently small $k_1$. This will be indicated in a few places in the proof. Similarly, the constants can be chosen such that the inequality holds for all $t \ge k_1^{2+\alpha'}$, in both cases by making the right hand side at least 1. We now prove that the inequality holds for $t < k_1^{7/4}$ by only considering the horizontal steps made by $X$. Let $U$ be the random walk on $\mathbb{Z} \cap [k - k_1, k + k_1]$, corresponding to the horizontal steps made by $X$. Note that until the time where $k \pm k_1$ is hit, $U$ behaves as a random walk on $\mathbb{Z}$. By Lemma 3.11, it follows that

$$\mathbb{P}\left(T \le k_1^{7/4}\right) \le 4e^{-\frac{1}{6}k_1^{1/4}} \le ce^{-c'k_1^{1/4}} \le c \exp\left(-c' \left(k_1^{2+\alpha'}/t\right)^{\alpha''}\right),$$

where the final inequality holds since $\alpha' \le 2$ and $\alpha'' \le 1/16$ so $\left(k_1^{2+\alpha'}/t\right)^{\alpha''} \le k_1^{4\alpha''} \le k_1^{1/4}$. From now on, we can assume that the time $t$ satisfies $k_1^{7/4} \le t \le k_1^{2+\alpha'}$.

We translate the interval from $[k - k_1, k + k_1]$ to $[-k_1, k_1]$ and forget $k$. To each integer $i$ in the interval $[-k_1, k_1]$ a random variable $H_i$ with cdf $F_X$ is attached, which is the height of the tooth at that point. Note that we assume that $F_X(n) \le 1 - n^{-1/\alpha}$ for all $n \in \mathbb{N}$ for some $\alpha > 3$. We apply Lemma 5.2 with

$$p = k_1^{-1/2+\delta}, \quad q = k_1^{-\delta/2}, \quad k = k_1, \quad m = \left\lceil k_1^{3/4-\delta/4} \right\rceil \quad \text{and} \quad \ell = c_\ell k_1^{1/4+\delta/4}$$

for some $\frac{1}{16} \le c_\ell \le \frac{1}{8}$ such that $\ell$ is an integer. Note that the inequalities $m \ge \sqrt{k_1/(pq)}$ and $4m\ell \le k_1$ hold. This yields that with probability exceeding

$$1 - 2k_1^{1/2+\delta/2} \exp\left(-c'' k_1^{\delta/2}\right) \ge 1 - c \exp\left(-c' k_1^{\delta/2}\right)$$

there exists a collection $\mathcal{D} = \{D_1, \dots, D_\ell\}$ of $\ell$ centric sets with minimal gap at least $m$ and length at least $\ell$ such that $D_i \cap D_j = \{0\}$ for all $1 \le i, j \le \ell$ with $i \ne j$ and such that $H_d \ge (F_X)^{-1}(1 - p)$ for all $d \in \bigcup_{D \in \mathcal{D}} D \setminus \{0\}$. Let $\widetilde{D} = \bigcup_{D \in \mathcal{D}} D \setminus \{0\}$.

Let $\varepsilon = \frac{1}{2}\delta$. Note that $\delta \le \frac{1}{2}$, so certainly $0 < \varepsilon < 1$. Moreover,

$$m\ell^{3-\varepsilon} \ge k_1^{3/4-\delta/4}\left(c_\ell k_1^{1/4+\delta/4}\right)^{3-\varepsilon} \ge \lambda k_1^{3/2+\delta/2-1/2\varepsilon} \ge \lambda k_1^{3/2+\delta/4},$$

for some constant $\lambda$, e.g. $\lambda = \left(\frac{1}{16}\right)^3$. We now apply Corollary 5.1, to the collection of centric sets constructed above. Let $S$ be a random walk on $\mathbb{Z} \cap [-k_1, k_1]$ starting from $S_0 = 0$. Then

$$\mathbb{P}(\#\{t \le \tau_k : S_t \in \widetilde{D}\} \le m\ell^{3-\varepsilon}) \le \tilde{c}\ell \exp\left(-\tilde{c}'\ell^{\varepsilon}\right) \le c' \exp\left(-c''\ell^{\varepsilon}\right)$$

so in this case, $\mathbb{P}\left(\#\{t \le \tau_k : S_t \in \widetilde{D}\} \le \lambda k_1^{3/2+\delta/4}\right) \le c' \exp\left(-c'' k_1^{\delta/8}\right)$ for again other constants $c', c''$. Note that this term dominates a term of the form $c\exp\left(-c' k_1^{\delta/2}\right)$.

Note that the condition $F_X(n) \le 1 - n^{-1/\alpha}$ for all $n \in \mathbb{N}$ also extends to $x \in \mathbb{R}_{>0}$ since $X$ is integer-valued and hence $F_X(x) = F_X(\lfloor x \rfloor) \le 1 - \lfloor x \rfloor^{-1/\alpha} \le 1 - x^{-1/\alpha}$ for all $x \in \mathbb{R}_{>0}$. By taking $x = y^{-\alpha}$ and applying the inverse of the cdf, this yields $(F_X)^{-1}(1-y) \ge y^{-\alpha}$. Hence,

$$H_d \ge (F_X)^{-1}(1-p) = (F_X)^{-1}\left(1 - k_1^{-1/2+\delta}\right) \ge k_1^{\alpha(1/2-\delta)} = k_1^{3/2+\delta}$$

since for all $d \in \widetilde{D}$ since $\delta = \frac{\alpha-3}{2\alpha+2}$ rewrites to $\alpha = \frac{\delta+3/2}{1/2-\delta}$.

Let $n = \theta k_1^{3/2+\delta}$, where $\theta \le 1$ will be chosen later such that $n \ge 1$. Let $L$ be the number of meetings of $X$ with a vertex $(k, 0)$ with a tooth of length at least $n$ before time $T$. From the discussion above it follows that with probability exceeding $1 - c\exp\left(-c' k_1^{\delta/8}\right)$, there exists a set $\widetilde{D}$ such that all teeth have length at least $n$ and such that $\#\{t \le \tau_k : S_t \in \widetilde{D}\} \ge \lambda k_1^{3/2+\delta/4}$. This implies that $L \ge \lambda k_1^{3/2+\delta/4}$, so $\mathbb{P}\left(L < \lambda k_1^{3/2+\delta/4}\right) \le c\exp\left(-c' k_1^{\delta/8}\right)$. Hence,

$$\begin{aligned}\mathbb{P}_k(T \le t) &\le \mathbb{P}\left(L < \theta k_1^{3/2+\delta/4}\right) + \mathbb{P}\left(T \le t,\ L \ge \lambda k_1^{3/2+\delta/4}\right) \\ &\le c\exp\left(-c' k_1^{\delta/8}\right) + \mathbb{P}\left(T \le t,\ L \ge \lambda k_1^{3/2+\delta/4}\right). \end{aligned} \tag{5}$$

At each meeting of $X$ with a vertex $(k, 0)$ with a tooth of length at least $n$, we perform an independent experiment, where we succeed if we hit $n$ on the tooth and then spend at least $n^2$ steps on this tooth. The independence of the experiments follows from the strong Markov property. Note that the probability of starting a walk on the tooth is $\frac{1}{3}$. If the random walk makes the transition from $(k, 0)$ to $(k, 1)$, then by Lemma 3.14, there is a probability of $\lceil n \rceil^{-1} \le \frac{1}{2}n^{-1}$ of reaching height $n$ before returning to 0. There now is a probability lower bounded by a constant $c_1'$ that $X$ takes at least $n^2 = \theta^2 k_1^{3+\delta/4}$ steps on the tooth. Let $c_1 = \frac{1}{6}c_1'$. Combining this gives a success probability of at least $c_1 n^{-1}$, since by the strong Markov property these three events are independent so the probabilities can be multiplied.
Hence, the number of successes is binomially distributed with at least $\lambda k_1^{3/2+\delta/4}$ trials and success probability at least $c_1 n^{-1}$. Denote such random variable by $\mathrm{Bin}\left(\lambda k_1^{3/2+\delta/4},\ c_1 n^{-1}\right)$.

Since $c_1$ is a constant not depending on $k_1$ and $\lambda = \left(\frac{1}{16}\right)^3$ is a constant, we have for sufficiently large $k_1$ that $\frac{1}{2}k_1^{3/4\cdot\delta}c_1\lambda \geq 1$. Let

$$\theta = \frac{2t}{c_1\lambda k_1^{3+5/4\cdot\delta}} \leq \frac{t}{k_1^{3+\delta/2}} \leq 1,$$

where the first inequality holds for $k_1$ sufficiently large and the last inequality holds for $t \leq k_1^{2+\alpha'}$. Recall that we can choose the constants $c, c'$ such that the desired inequality holds if one of these conditions does not hold, so we may assume that these conditions hold. We also need to check that $n \geq 1$. Recall that we may assume that $t \geq k_1^{7/4}$. Since $c_1\lambda \leq 1$, it then follows that

$$\theta \geq \frac{2k_1^{7/4}}{k_1^{3+5/4\cdot\delta}} > \frac{1}{k_1^{5/4+5/4\cdot\delta}} > \frac{1}{k_1^{3/2+\delta}},$$

so $n = \theta k_1^{3/2+\delta} \geq 1$ as required.

Write $N = \lambda k_1^{3/2+\delta/4}$, $p = c_1 n^{-1}$ and $s = tn^{-2}$. We compute:

$$\frac{s}{Np} = \frac{tn^{-2}}{\lambda k_1^{3/2+\delta/4}c_1 n^{-1}} = \frac{t}{\lambda k_1^{3/2+\delta/4}c_1\theta k_1^{3/2+\delta}} = \frac{tc_1\lambda k_1^{3+5/4\cdot\delta}}{2tc_1\lambda k_1^{3/2+\delta/4}k_1^{3/2+\delta}} = \frac{1}{2}$$

$$Np = \lambda k_1^{3/2+\delta/4}c_1\theta^{-1}k_1^{-3/2-\delta} = \lambda c_1 k_1^{-3/4\cdot\delta}\frac{c_1\lambda k_1^{3+5/4\cdot\delta}}{2t} = \tfrac{1}{2}(c_1\lambda)^2\cdot k_1^{3+1/2\cdot\delta}/t = \widetilde{c}_1 k_1^{2+\alpha'}/t.$$

Hence, $s = \frac{1}{2}Np$ and $Np = \widetilde{c}_1 k_1^{2+\alpha'}/t$.

For $-1 \leq \mu \leq 0$ we have $e^\mu - 1 \leq \left(1 - \frac{1}{e}\right)\mu$. The moment generating function of a binomially distributed random variable with $N$ trials and success probability $p$ is

$$M(\mu) = (1 + p(e^\mu - 1))^N \leq \left(1 + (1 - \tfrac{1}{e})\mu p\right)^N \leq \exp\left(\left(1 - \tfrac{1}{e}\right)\mu Np\right).$$

Note that $T$ is at least $n^2$ times the number of successes, since with each success at least $n^2$ steps are done on a tooth and all successes occur before time $T$. If $L \geq N$, it therefore follows that $T$ stochastically dominates $n^2\,\mathrm{Bin}\,(N,p)$. In particular, the event $T \leq t$ is less likely than the event $\mathrm{Bin}\,(N,p) \leq tn^{-2} = s$. By Chernoff's bound with $\mu = -1$, we now find that

$$\begin{aligned}\mathbb{P}\left(T \leq t,\ L \geq N\right) &\leq \mathbb{P}\left(\mathrm{Bin}\,(N,\,p) \leq s\right) \leq e^s \exp\left(\left(\tfrac{1}{e} - 1\right)Np\right) = \exp\left(\left(\tfrac{1}{e} - \tfrac{1}{2}\right)Np\right)\\ &= \exp\left(-\mu'\widetilde{c}_1 k_1^{2+\alpha'}/t\right) = \exp\left(-c'k_1^{2+\alpha'}/t\right) \leq \exp\left(-c'\left(k_1^{2+\alpha'}/t\right)^{\alpha''}\right)\end{aligned}$$

where $\mu' = \frac{1}{2} - \frac{1}{e} > 0$, and where the last inequality holds for $t \leq k_1^{2+\alpha'}$ since $\alpha'' = \delta/32 < 1$. Finally, note that $\alpha' < 2$, so $2 + \alpha' < 4$ and hence

$$c\exp\left(-c'k_1^{\delta/8}\right) = c\exp\left(-c'\left(k_1^4\right)^{\alpha''}\right) \leq c\exp\left(-c'\left(k_1^{2+\alpha'}/t\right)^{\alpha''}\right).$$

By equation (5), this proves the lemma. $\square$

From now on, let $\alpha'$ be a constant such that Lemma 5.3 holds.
Let $\beta' = \frac{\alpha'+1}{\alpha'+2}$. We now state a lemma analogous to Lemma 4.6.

**Lemma 5.4.** $q_t(u,u) \leq \frac{c}{t^{\beta'}}$ for any $u = (k,0)$ on the horizontal axis and any $t \geq 1$.

*Proof.* The proof is identical to the proof of Lemma 4.6, except that Lemma 5.3 is used instead of Lemma 4.5 and hence the 1/3 is replaced by $\alpha''$. The upper bound for $\mathbb{P}_u(\tau_B \leq t)$ now becomes

$$\mathbb{P}_u(\tau_B \leq t) \leq c \exp\left(-c'\left(k_1^{2+\alpha'}/t\right)^{\alpha''}\right) = c \exp\left(-c' b^{(2+\alpha')\alpha''}\right).$$

By taking $b$ sufficiently large, it can still be ensured that $\mathbb{P}_k(\tau_B \leq t) \leq \frac{1}{2}$. □

We also need a lemma analogous to Lemma 4.9.

**Lemma 5.5.** Let $x = (k,0)$. Then if $t < |k|^{2+\alpha'}$, we have

$$q_t(0,x) \leq c|k|^{-(\alpha'+1)} = c\left(|k|^{2+\alpha'}\right)^{-\beta}.$$

*Proof.* For positive $k$, the proof is identical to the proof of Lemma 4.9, except that Lemma 5.3 is used instead of Lemma 4.5 and hence the 1/3 is replaced by $\alpha''$.
That means that we observe that $\sup_{\eta>0}\left(\eta^{\beta'} e^{-c\eta^{\alpha''}}\right)$ is finite and hence constant.

For negative $k$, the proof is similar to the proof for positive $k$. □

These lemmas can now be combined to prove the following lemma on the transition density:

**Lemma 5.6.**
Let $x = (k,h) \in V$. The transition density $q$ satisfies

$$q_t(0,x) \leq \begin{cases} ct^{-\beta} & \text{if } t \geq |k|^{2+\alpha'}, \\ c\left(|k|^{2+\alpha'}\right)^{-\beta} & \text{if } t \leq |k|^{2+\alpha'}. \end{cases}$$

*Proof.* The proof is identical to the proof of Lemma 4.10, except that we refer to Lemma 5.4 and 5.5 instead of Lemma 4.6 and 4.9. (Note that Lemma 4.7 and 4.8 can be reused directly). □

Recall that $Q_{k,h} = \{(k,y) : 0 \leq y \leq h\}$ and that $Z_{k,h}$ is the number of collisions of the two random walks in $Q_{k,h}$ and that $\widetilde{Z}_{k,h} = Z_{k,2h/3} - Z_{k,h/3}$. These notions allow us to formulate and prove the final lemma needed for proving the finite collision property of $\text{Comb}(\mathbb{Z},\alpha)$ for $\alpha > 1$.

**Lemma 5.7.**

(a) $\mathbb{E}[Z_{k,h}] \leq ch|k|^{-\alpha'}$.

(b) $\mathbb{E}[Z_{k,h} \mid \widetilde{Z}_{k,h} > 0] \geq ch$.

*Proof.* The proof is identical to the proof of Lemma 4.11. $\square$

These lemmas together allow to prove the main result of this chapter:

**Theorem 5.1.** Let $\alpha > 3$ be given and let $F_X : \mathbb{N} \to [0,1]$ be a cumulative distribution function satisfying $F_X(n) \leq 1 - n^{-1/\alpha}$ for all $n \in \mathbb{N}$. Moreover, assume that there exists a constant $E > 0$ such that $F_X(n) \geq 1 - n^{-E}$ for all $n \in \mathbb{N}$. Let $\{f(n)\}_{n\in\mathbb{Z}}$ be i.i.d. random variables with cdf $F_X$. Then $G = \mathrm{Comb}(\mathbb{Z}, f)$ has the finite collision property almost surely.

*Proof.* Let $\alpha > 1$ be given. By Proposition 3.10 and Lemma 5.7, it follows that

$$\mathbb{P}(\widetilde{Z}_{k,h} > 0) \leq \frac{\mathbb{E}[Z_{k,h}]}{\mathbb{E}[Z_{k,h} \mid \widetilde{Z}_{k,h} > 0]} \leq \frac{c' h k^{-\alpha'}}{c'' h} = ck^{-\alpha'},$$

so $\mathbb{P}(\widetilde{Z}_{k,h} > 0) \leq ck^{-\alpha'}$. Since there exists a constant $E > 0$ such that $F_X(n) \geq 1 - n^{-E}$ for all $n \in \mathbb{N}$, it follows that $\mathbb{P}(f(n) \geq |n|^{2/E}) \leq n^{-2}$ for all $n \in \mathbb{Z}$. In particular, this almost surely happens only finitely often by the Borel-Cantelli Lemma.

Summing over all $k$ and all $j$ satisfying $2^j \leq f(k)$ yields

$$\sum_{k=-\infty}^{\infty} \sum_{j=0}^{\log_2 f(k)} \mathbb{P}(\widetilde{Z}_{k,2^j} > 0) \leq C + \sum_{k=-\infty}^{\infty} \sum_{j=0}^{\log_2 |k|^{2/E}} \mathbb{P}(\widetilde{Z}_{k,2^j} > 0) \leq C + \sum_{k=\infty}^{\infty} \log_2 |k|^{2/E} ck^{-\alpha'} < \infty$$

since $\alpha' > 1$. The first inequality holds since there are only finitely many $k$ for which the inequality $f(k) \geq |k|^{2/E}$ holds and each of these $k$s gives only a finite number of additional terms. By the Borel-Cantelli lemma, it follows that $\mathbb{P}(\widetilde{Z}_{k,2^j} > 0 \text{ occurs i.o.}) = 0$.
By Corollary 4.1, it follows that only finitely many collisions occur in the graph $\mathrm{Comb}(\mathbb{Z}, f)$ almost surely. Hence, $\mathrm{Comb}(\mathbb{Z}, f)$ has the finite collision property almost surely. $\square$

# 6 Conclusion and discussion

To conclude this thesis, let us briefly summarize the main results. In the Chapter 3, the groundwork is laid for Chapter 4 and 5. Chapter 4 reviews the main results known on comb graphs from the literature: the finite collision property of $\text{Comb}(\mathbb{Z})$ (from Kirshnapur and Peres [7]), the finite collision property of $\text{Comb}(\mathbb{Z}, \alpha)$ for $\alpha \leq 1$ and the infinite collision property of $\text{Comb}(\mathbb{Z}, \alpha)$ for $\alpha > 1$ (from Barlow, Peres and Sousi [3]).

In the final chapter, it is proven that if $F_X : \mathbb{N} \to [0, 1]$ is a cumulative distribution function satisfying $F_X(n) \leq 1 - n^{-1/\alpha}$ for some $\alpha > 3$, and a technical condition (namely, that there exists a constant $E > 0$ such that $F_X(n) \geq 1 - n^{-E}$ for all $n \in \mathbb{N}$), then the comb graph with i.i.d. heights with cdf $F_X$ has the finite collision property almost surely.

The author is very satisfied with this result. Two important questions however still remain. The first is whether the 3 is optimal in this theorem, or that it could be lowered. The second question is whether the technical condition can be removed. It seems that the latter would require different sets, possibly also involving the time, i.e. an approach more similar to Kirshnapur and Peres [7], at least when it comes to the sets in space and time the collisions are divided in.

Many related problems are still open, but can possibly be solved using the centric set approach. One problem of particular interest to the author is the case of comb graphs with gaps between vertical copies of $\mathbb{Z}$. This means that some vertical copies of $\mathbb{Z}$ are removed from $\text{Comb}(\mathbb{Z})$. A bit more formally, let $\text{Comb}_D(\mathbb{Z})$ be the comb graph with vertical copies of $\mathbb{Z}$ at locations given by a set $D \subset \mathbb{Z}$. The distance between two successive elements of $D$ is called a gap, and we say that $\text{Comb}_D(\mathbb{Z})$ has bounded gaps, if all gaps are bounded by a constant and $D$ moreover contains infinitely many positive and infinitely many negative elements. This situation leads to the following conjecture:

**Conjecture 6.1.**
If $\text{Comb}_D(\mathbb{Z})$ has bounded gaps, then $\text{Comb}_D(\mathbb{Z})$ has the finite collision property.

While this case is very similar to the collision property of $\text{Comb}(\mathbb{Z})$, the proof does not carry over easily. An important step in the proof of Krishnapur and Peres [7] is the independence of the horizontal walk $(U)$ and the vertical walk $(V)$. This step no longer works in the case with gaps, and it seems hard to prove that the covariance is sufficiently small as well. However, this conjecture may be solvable using the centric set approach.

# References


[1] S.E. Alm. *Simple random walk*. 2006.

[2] M. Balázs. *Martingale Theory Lecture Notes*. 2018.

[3] M. T. Barlow, Y. Peres, and P. Sousi. Collisions of random walks. *Annales de l'IHP Probabilités et statistiques*, 48(4):922–946, 2012.

[4] D. Bertacchi, N. Lanchier, and F. Zucca. Contact and voter processes on the infinite percolation cluster as models of host-symbiont interactions. *The Annals of Applied Probability*, 21(4):1215–1252, 2011.

[5] D. Chen, B. Wei, and F. Zhang. A note on the finite collision property of random walks. *Statistics & Probability Letters*, 78(13):1742–1747, 2008.

[6] R. Durrett. *Probability: theory and examples*, volume 49. Cambridge University Press, 2019.

[7] M. Krishnapur and Y. Peres. Recurrent graphs where two independent random walks collide finitely often. *Electronic Communications in Probability*, 9:72–81, 2004.

[8] D. Levin and Y. Peres. *Markov chains and mixing times*, volume 107. American Mathematical Soc., 2017.

[9] T. Lindvall and L. C. G. Rogers. Coupling of multidimensional diffusions by reflection. *The Annals of Probability*, 14(3):860–872, 1986.

[10] R. Lyons and Y. Peres. *Probability on Trees and Networks*, volume 42 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, New York, 2016. Available at `http://pages.iu.edu/~rdlyons/`.

[11] R. Stanley and E. W. Weisstein. Catalan numbers. From MathWorld—A Wolfram Web Resource. Visited March 30 at `mathworld.wolfram.com/CatalanNumber.html`.

[12] D. Valesin and E. Verbitskiy. *Lecture Notes Random Walks*. 2019.

[13] W. Woess. *Random Walks on Infinite Graphs and Groups*, volume 138. Cambridge University Press, 2000.