DOUBLE JUMP PHASE TRANSITION IN A SOLITON CELLULAR AUTOMATON

LIONEL LEVINE, HANBAEK LYU, AND JOHN PIKE

ABSTRACT. In this paper, we consider the soliton cellular automaton introduced in [25] with a random initial configuration. We give multiple constructions of a Young diagram describing various statistics of the system in terms of familiar objects like birth-and-death chains and Galton-Watson forests. Using these ideas, we establish limit theorems showing that if the first n boxes are occupied independently with probability $p \in (0, 1)$, then the number of solitons is of order n for all p, and the length of the longest soliton is of order $\log n$ for $p < 1/2$, order $\sqrt{n}$ for $p = 1/2$, and order n for $p > 1/2$. Additionally, we uncover a condensation phenomenon in the supercritical regime: For each fixed $j \geq 1$, the top j soliton lengths have the same order as the longest for $p \leq 1/2$, whereas all but the longest have order $\log n$ for $p > 1/2$. As an application, we obtain scaling limits for the lengths of the kth longest increasing and decreasing subsequences in a random stack-sortable permutation of length n in terms of random walks and Brownian excursions.

1. INTRODUCTION

In 1990, Takahashi and Satsuma proposed a $1 + 1$ dimensional cellular automaton of filter type called the *soliton cellular automaton*, also known as the *box-ball system* [17, 25]. It is defined as a discrete-time dynamical system $(X_s)_{s \geq 0}$ whose states are binary sequences $X_s : \mathbb{N} \to \{0, 1\}$ with finitely many 1's. We may think of the states as configurations of balls in boxes where box k contains a ball at stage s if $X_s(k) = 1$ and is empty if $X_s(k) = 0$. The update rule $X_s \mapsto X_{s+1}$ is defined as follows: At the beginning of stage s, each ball has been moved a total of s times. To reach stage $s+1$, successively move the leftmost ball which has been moved a total of s times to the first empty box on its right, continuing until all balls have been moved. Alternatively, at each stage $s \geq 0$ a 'carrier' starts at the origin and sweeps rightward to infinity. Each time she encounters an occupied box, she *pushes* the ball to the top of her stack. Each time she encounters an empty box and her stack is nonempty, she *pops* any ball from her stack into the box. In keeping with this picture, we will refer to the stages of the box-ball system as *sweeps* henceforth.

As a concrete example, the system initially having balls in boxes $2, 3, 5, 6, 7, 11$ evolves through sweep $s = 3$ as

$$
\begin{array}{c|l}
s = 0 & 0\ 1\ 1\ 0\ 1\ 1\ 1\ 0\ 0\ 0\ 1\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ \ldots \\
1 & 0\ 0\ 0\ 1\ 0\ 0\ 0\ 1\ 1\ 1\ 0\ 1\ 1\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ \ldots \\
2 & 0\ 0\ 0\ 0\ 1\ 0\ 0\ 0\ 0\ 0\ 1\ 0\ 0\ 1\ 1\ 1\ 1\ 0\ 0\ 0\ 0\ 0\ \ldots \\
3 & 0\ 0\ 0\ 0\ 0\ 1\ 0\ 0\ 0\ 0\ 0\ 1\ 0\ 0\ 0\ 0\ 0\ 1\ 1\ 1\ 1\ 0\ \ldots
\end{array}
$$

In this model, a (non-interacting) *soliton of length* k is defined to be a string of k consecutive 1's followed by k consecutive 0's. During one sweep, such a soliton travels to the right at speed k. The physical interpretation is that of a traveling wave with velocity equal to its wavelength. If a k-soliton precedes a j-soliton with $j < k$, then the two will eventually collide, resulting in interference. The subsequent states of the system depend on the congruence class of their initial distance modulo their relative speed, $k-j$, but solitons are never created or destroyed in the course of these interactions. The case of three or more interacting solitons can be described similarly [25]. It is easy to see that since we have finitely many balls initially, after some finite time the system consists of non-interacting solitons whose lengths are nondecreasing from left to right. We will call such a configuration *stable*. This final macrostate of the system can be encoded in the Young diagram $\Lambda(X_0)$ having j^{th} column equal in length to the jth longest soliton.

In this paper, we start the soliton cellular automaton from a random initial configuration and study the limiting shape of the resulting Young diagram. We have two parameters, $n \in \mathbb{N}$ and $p \in (0,1)$. Let $X^{n,p}$ be a random coloring of $\mathbb{N}$ so that each site in $[1,n]$ is 1 with probability p and 0 with probability $1-p$, independently of all others, and all sites in (n,∞) are 0. Let $\Lambda^{n,p} = \Lambda(X^{n,p})$ be the corresponding random Young diagram and denote the lengths of its ith row and j^{th} column by $\rho_i(n)$ and $\lambda_j(n)$, respectively. (Thus $\lambda_j(n)$ gives the length of the jth longest soliton and $\rho_i(n)$ the number of solitons of length at least i.) We are going to show that each fixed row has order n for all values of p, but the column lengths vary drastically according to whether p is less than, equal to, or greater than 1/2. The asymptotics of the rows and columns of $\Lambda^{n,p}$ are summarized in the following table, for which Theorem 1 proves the ρ entries, and Theorem 2 proves the λ entries. For the precise meaning of the Landau notation employed, see Subsection 1.2.

$i \geq 1, j \geq 2$ fixed	$\rho_i(n)$	$\lambda_j(n)$	$\lambda_1(n)$
Subcritical phase ($p < 1/2$)	$\Theta(n)$	$\Theta(\log n)$	$\Theta(\log n)$
Critical phase ($p = 1/2$)	$\Theta(n)$	$\Theta(\sqrt{n})$	$\Theta(\sqrt{n})$
Supercritical phase ($p > 1/2$)	$\Theta(n)$	$\Theta(\log n)$	$\Theta(n)$

TABLE 1. Double jump phase transition for the order of the longest j solitons (j fixed as $n \to \infty$) in the random box-ball system. All entries are up to constant factors that do not depend on n. In the sub- and supercritical phases the λ_j are concentrated, and the constant factor depends only on p (and not on j). In the critical phase the λ_j are not concentrated, and the constant factor depends on j.

Erdős and Rényi coined the term *double jump* to describe the emergence of a giant component in the sparse random graph with n vertices, each pair independently joined by an edge with probability c/n, where $c > 0$ is a parameter. The analogy between random graph components and box-ball solitons becomes apparent if we take $c = p/(1-p)$. Then with high probability, all connected components of the Erdős-Rényi graph are of size $O(\log n)$ for $p < 1/2$; components of size $\Theta(n^{2/3})$ emerge at $p = 1/2$; and for $p > 1/2$, the largest component is of size $\Theta(n)$ while all the rest have size $O(\log n)$ [8]. Except for the exponent 2/3 (which becomes 1/2) this is exactly the behavior of the soliton lengths in the box-ball system as summarized in the last two columns of Table 1.

1.1. **Related work.** There have been some exciting recent developments involving the box-ball system with a bi-infinite random initial configuration. A central question is to understand the invariant measures on $\{0,1\}^{\mathbb{Z}}$ under the box-ball dynamics. Ferrari, Nguyen, Rolla, and Wang [9] showed that the Bernoulli product measure with density $p < 1/2$ is invariant and provided a recipe for constructing additional invariant measures based on a soliton decomposition of box-ball configurations. Croydon, Kato, Sasada, and Tsujimoto [4] found sufficient conditions for invariance using Pitman's $2M - X$ transformation and considered extending the box-ball system from $\mathbb{Z}$ to $\mathbb{R}$. See the references for more details.

1.2. **Notation.** We adopt the notation $\mathbb{R}^+ = [0,\infty)$, $\mathbb{N} = \{1,2,3,\ldots\}$, and $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$ throughout. We employ the Landau notation $O(\cdot), \Omega(\cdot), \Theta(\cdot)$ in the sense of stochastic boundedness. That is, given $\{a_n\}_{n=1}^{\infty} \subseteq \mathbb{R}^+$ and a sequence $\{W_n\}_{n=1}^{\infty}$ of nonnegative random variables, we say that $W_n = O(a_n)$ if for every $\varepsilon > 0$, there is a $C \in (0,\infty)$ such that $\mathbb{P}\{W_n > Ca_n\} < \varepsilon$ for all n. We say that $W_n = \Omega(a_n)$ if for every $\varepsilon > 0$, there is a $c \in (0,\infty)$ such that $\mathbb{P}\{W_n < ca_n\} < \varepsilon$ for all n, and we say $W_n = \Theta(a_n)$ if $W_n = O(a_n)$ and $W_n = \Omega(a_n)$. The constants c, C may depend on p and ε but not n.

1.3. **Main results.** Fix $p \in (0,1)$, and let $\xi_1, \xi_2, \ldots$ be a sequence of i.i.d. random variables with law $\mathbb{P}\{\xi_1 = 1\} = p$ and $\mathbb{P}\{\xi_1 = -1\} = 1 - p$. Define $X^p \in \{0,1\}^{\mathbb{N}}$ by

$$X^p(k) = \mathbf{1}\{\xi_k = 1\},$$

and for each $n \in \mathbb{N}$, set $X^{n,p} = X^p \mathbf{1}_{[1,n]}$. The interpretation is that $X^{n,p}$ corresponds to an arrangement of balls in boxes where boxes $1, \ldots, n$ are each occupied independently with probability p, and boxes $n+1, n+2, \ldots$ are empty.

For each fixed $n \geq 1$ and $p \in (0,1)$, we consider the box-ball system $(X_s)_{s \geq 0}$ with the random initial configuration $X_0 = X^{n,p}$. Recall that the soliton lengths are denoted by $\lambda_1(n) \geq \lambda_2(n) \geq \cdots$. This information can be summarized by the Young diagram $\Lambda^{n,p}$ whose j^{th} column has length $\lambda_j(n)$. The length of its ith row, $\rho_i(n)$, equals the number of solitons in the system having length at least i. In particular, $\rho_1(n)$ gives the total number of solitons.

Many properties of this Young diagram can be described in terms of the simple random walk $\{S_k\}_{k=0}^{\infty}$ defined by $S_0 = 0$ and $S_k = \xi_1 + \cdots + \xi_k$. Our first result shows that the i longest rows are of order n for any $p \in (0,1)$.

Theorem 1. *Let $X^{n,p}$ and S_k be as above. Then the following statements hold.*

(i)   *(SLLN for rows) Let $\varsigma = \inf\{k > 0 : S_k = 0\} \in \mathbb{N} \cup \{\infty\}$ be the first return time of S_k to 0. Then for any fixed $i \geq 1$,*

$$\frac{\rho_i(n)}{n} \to \mathbb{P}\left\{ \max_{0 \leq k \leq \varsigma} S_k = i \right\} > 0 \quad a.s. \ \ as \ n \to \infty.$$

(ii)   *(CLT for the first row)*

$$\frac{\rho_1(n) - np(1-p)}{\sqrt{np(1-p)[1 - 3p(1-p)]}} \Rightarrow Z$$

where $Z \sim \mathcal{N}(0,1)$, the standard normal distribution.

Denote by $C(\mathbb{R})$ the space of continuous functions $f : \mathbb{R} \to \mathbb{R}$ endowed with the topology of uniform convergence on compact sets, and let $C_0^+(\mathbb{R})$ be the subspace of $C(\mathbb{R})$ consisting of nonnegative compactly supported functions f such that $f \equiv 0$ on $(-\infty, 0]$. For any closed interval $I \subseteq \mathbb{R}$ containing 0,

denote by $C(I)$ and $C_0^+(I)$ the space of restrictions $f|_I$ where $f \in C(\mathbb{R})$ and $f \in C_0^+(\mathbb{R})$, respectively. For $b \in I$, define the operator $\mathcal{E}_b : C(I) \to C(I)$ by

$$\mathcal{E}_b(f)(t) = f(t) - \min_{b \wedge t \leq s \leq b \vee t} f(s),$$

where $y \wedge z = \min(y, z)$ and $y \vee z = \max(y, z)$. We call b the *pivot* of $\mathcal{E}_b$. Define $\mathfrak{m} : C_0^+(I) \to \mathbb{R}^+$ by $\mathfrak{m}(g) = \sup\{x \in I : g(x) = \max(g)\}$, the location of the rightmost global maximum of g. Finally, define the *excursion operator* $\mathcal{E}$ on $C_0^+(I)$ by $\mathcal{E}(g) = \mathcal{E}_{\mathfrak{m}(g)}(g)$. See Figure 4 for an illustration.

We now state the main result of the paper.

Theorem 2. *Let $X^{n,p}$ be as above and set $\theta = (1-p)/p$. Let $\lambda_j(n)$ denote the jth longest soliton length.*

(i) *(Subcritical phase) For $p < 1/2$, $\lambda_j(n)$ is concentrated around $\mu_n := \log_\theta\left(\frac{(1-2p)^2}{1-p}n\right)$ for each fixed $j \geq 1$ in the sense that for all $x \in \mathbb{R}$,*

$$\exp(-\theta^{-x}) \sum_{k=0}^{j-1} \frac{\theta^{-k(x+1)}}{k!} \leq \liminf_{n\to\infty} \mathbb{P}\left\{\lambda_j(n) \leq x + \mu_n\right\}$$

$$\leq \limsup_{n\to\infty} \mathbb{P}\left\{\lambda_j(n) \leq x + \mu_n\right\} \leq \exp(-\theta^{-(x+1)}) \sum_{k=0}^{j-1} \frac{\theta^{-kx}}{k!}.$$

In particular, $\lambda_j(n) = \Theta(\log n)$.

(ii) *(Critical phase) For $p = 1/2$, let $B = \{B_t\}_{0 \leq t \leq 1}$ be a standard Brownian motion on $[0,1]$. Then for each fixed $j \geq 1$,*

$$n^{-1/2}[\lambda_1(n), \lambda_2(n), \ldots, \lambda_j(n)] \Rightarrow [\max|B|, \max\mathcal{E}(|B|), \ldots, \max\mathcal{E}^{j-1}(|B|)],$$

In particular, $\lambda_j(n) = \Theta(\sqrt{n})$.
Furthermore, for any integers $j, k \geq 1$,

$$\lim_{n\to\infty} n^{-k/2}\mathbb{E}[(\lambda_j(n))^k] = \mathbb{E}\left[\left(\max\mathcal{E}^{j-1}(|B|)\right)^k\right].$$

(iii) *(Supercritical phase) For $p > 1/2$,*

$$\frac{\lambda_1(n) - (2p-1)n}{2\sqrt{p(1-p)n}} \Rightarrow Z \sim \mathcal{N}(0,1).$$

Furthermore, there exists a constant $c = c(p) > 0$ such that

$$\mathbb{P}\{|\lambda_1(n) - (2p-1)n| \geq x\} \leq c \exp(-x^2/(8n)),$$

and for all $j \geq 2$, $\lambda_j(n)$ is concentrated around $\hat{\mu}_n := \log_{\theta^{-1}}\left(\frac{(1-2p)^2}{p}n\right)$ in the sense that for all $x \in \mathbb{R}$,

$$\left(\exp(-\theta^{-\frac{x}{2}}) \sum_{k=0}^{j-1} \frac{\theta^{-k(\frac{x}{2}+1)}}{k!}\right) - c\theta^{\frac{x}{8}} \leq \liminf_{n\to\infty} \mathbb{P}\left\{\lambda_j(n) \leq x + \hat{\mu}_n\right\}$$

$$\leq \limsup_{n\to\infty} \mathbb{P}\left\{\lambda_j(n) \leq x + \hat{\mu}_n\right\} \leq \left(\exp(-\theta^{-(\frac{3x}{2}+1)}) \sum_{k=0}^{j-1} \frac{\theta^{-\frac{3kx}{2}}}{k!}\right) + c\theta^{\frac{x}{8}}.$$

In particular, $\lambda_1(n) = \Theta(n)$ and $\lambda_j(n) = \Theta(\log n)$ if $j \geq 2$.

We call the statement in Theorem 2 (iii) a *condensation phenomenon* because in the supercritical regime, a linear number of balls condense into the longest soliton while the next j longest solitons each have $\Theta(\log n)$ balls.

The methods that we develop in this paper to study the box-ball system yield several interesting results on lengths of monotone subsequences in random pattern avoiding permutations. The study of statistics involving longest increasing or decreasing subsequences in different types of random permutations has a long history and rich connections to many other fields [22]. In the context of the box-ball system, the class of 312-avoiding permutations arises naturally, and we are able to generalize some classical results on such permutations in multiple directions.

For each $n \in \mathbb{N}$, let $\mathfrak{S}_n$ be the set of all permutations on $\{1, 2, \ldots, n\}$. Given two permutations $\sigma \in \mathfrak{S}_n$ and $\tau \in \mathfrak{S}_k$ with $1 < k \leq n$, we say that σ is τ-*avoiding* if no subsequence of σ has the same relative order as τ. (For example, a permutation is 312-avoiding if there is no subsequence of the form z, x, y with $x < y < z$.) Denote by $\mathfrak{S}_n^\tau$ the set of all τ-avoiding permutations in $\mathfrak{S}_n$. Note that σ is τ-avoiding if and only if σ^{-1} is τ^{-1}-avoiding. (In particular, σ is 231-avoiding if and only if σ^{-1} is 312-avoiding.) Given a permutation $\sigma \in \mathfrak{S}_n$, define integers $\lambda_1, \ldots, \lambda_k$ (resp. $\rho_1, \ldots, \rho_k$) recursively so that $\lambda_1(\sigma) + \cdots + \lambda_k(\sigma)$ (resp. $\rho_1(\sigma) + \cdots + \rho_k(\sigma)$) equals the length of the longest subsequence in σ obtained by taking a disjoint union of k decreasing (resp. increasing) subsequences.

In a classic work [23], Rotem studied properties of 231-avoiding permutations chosen uniformly at random among all such permutations of a given length. He showed that if Σ^n is a permutation in $\mathfrak{S}_n^{231}$ chosen uniformly at random, then

$$\mathbb{E}[\rho_1(\Sigma^n)] = (n+1)/2, \quad \mathbb{E}[\lambda_1(\Sigma^n)] = \sqrt{\pi n} + O(1).$$

Our next theorem is an extension of the above result both to the higher moments and to 'subsequent' longest increasing and decreasing subsequences of Σ^n.

Theorem 3. *Let Σ^n be a uniformly chosen random 312- (or 231-) avoiding permutation of length n.*

(i) *Suppose that $T_1^n, T_2^n, \ldots, T_i^n$ is a sequence of rooted trees where T_1^n is chosen uniformly at random among all rooted plane trees on $n + 1$ nodes, and for $r \geq 1$, T_{r+1}^n is obtained from T_r^n by deleting all leaves. Then*

$$[\rho_1(\Sigma^n), \rho_2(\Sigma^n), \ldots, \rho_i(\Sigma^n)] =_d [\text{\# of leaves in } T_1^n, \text{\# of leaves in } T_2^n, \ldots, \text{\# of leaves in } T_i^n].$$

(ii) *Let $\{S_k\}_{k=0}^\infty$ be a simple symmetric random walk with $S_0 = 0$ and let $\varsigma = \inf\{k > 0 : S_k = 0\}$ be the time of its first return to 0. Then for any fixed $i \geq 1$,*

$$\frac{\rho_i(\Sigma^n)}{2n} \to \mathbb{P}\left\{\max_{0 \leq k \leq \varsigma} S_k = i\right\} > 0 \quad a.s. \quad as \ n \to \infty.$$

(iii) *Let $B^{ex} = (B_t^{ex})_{0 \leq t \leq 1}$ be a standard Brownian excursion on $[0, 1]$. Then for each fixed $j \geq 1$,*

$$n^{-1/2}[\lambda_1(\Sigma^n), \lambda_2(\Sigma^n), \ldots, \lambda_j(\Sigma^n)] \Rightarrow \sqrt{2}[\max B^{ex}, \max \mathscr{E}(B^{ex}), \ldots, \max \mathscr{E}^{j-1}(B^{ex})].$$

Furthermore, for any integers $j, k \geq 1$,

$$\lim_{n \to \infty} n^{-k/2} \mathbb{E}[(\lambda_j(\Sigma^n))^k] = 2^{k/2} \mathbb{E}\left[\left(\max \mathscr{E}^{j-1}(B^{ex})\right)^k\right].$$

We remark that given a 312-avoiding permuatation σ, we can actually interpret $\lambda_k(\sigma)$ as the length of the longest decreasing subsequence after successively deleting an arbitrary longest decreasing subsequence $k - 1$ times. For the rows, we can interpret ρ_k similarly but the longest increasing subsequence we delete at each step must be a special one; see Proposition 8.1. Note that such an interpretation is not valid for general permutations.

1.4. **Outline and organization.** Broadly speaking, we proceed by observing correspondences between various combinatorial objects related to box-ball configurations, such as Motzkin paths, rooted forests, and 312-avoiding permutations; see Figure 1. We can then interpret the rows and columns of the Young diagram associated with a box-ball configuration in terms of these objects (Table 2). This allows us to reformulate the original soliton problem in other languages and vice versa.
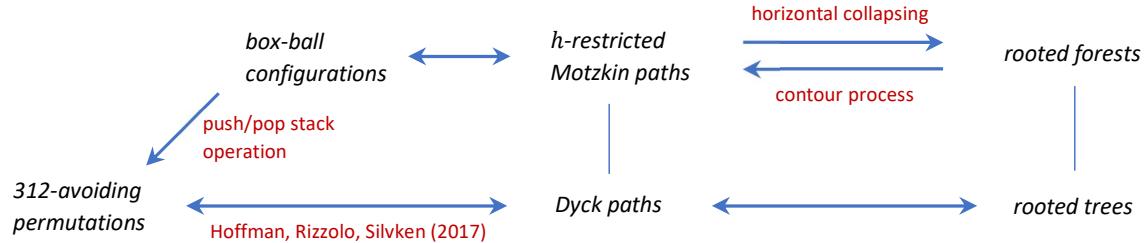

FIGURE 1. Correspondences and inclusions between six combinatorial objects. Objects in the same row are in bijective correspondence, while the vertical lines indicate that the bottom objects are special cases of the top.

For us, Motzkin paths provide the most useful framework, especially in the random setting. This is because the random box-ball configuration $X^{n,p}$ can be viewed as the increment sequence of the first n steps of a simple random walk driven by the Bernoulli(p) measure. The corresponding random (h-restricted) Motzkin path is the same simple random walk except that downstrokes at height 0 are censored. The problem then essentially boils down to studying properties of the excursions of such censored random walks. The results for random Motzkin paths can then be translated back to solitons or permutations.

	Box-ball configurations	Motzkin paths	Rooted forests	312-avoiding permutations
i^{th} row length of Young diagram	Number of solitons of length $\geq i$	Number of subexcursions of height i	Number of leaves after trimming them $i-1$ times	Length of ith longest increasing subsequence
j^{th} column length of Young diagram	Length of jth longest soliton	Maximum height after applying excursion operator $j-1$ times	Maximum height after contracting longest path $j-1$ times	Length of jth longest decreasing subsequence

TABLE 2. Interpretation of rows and columns of the Young diagram associated with four combinatorial objects.

This paper is organized as follows: In Section 2, we describe relations between box-ball configurations, Motzkin paths, and rooted forests, and show how to construct the Young diagram from these objects. In Section 3, we discuss a correspondence between random box-ball configurations, a birth-and-death chain, and a Galton-Watson forest. We prove Theorem 1 in Section 4, and the proof of Theorem 2 is given in Sections 5, 6, and 7. In Section 8, we discuss a connection between box-ball configurations and pattern-avoiding permutations and prove Theorem 3. Finally, in Appendix A, we prove the three lemmas stated in Subsection 2.2 along with some results concerning 312-avoiding permutations.

2. Constructing the time-invariant Young Diagram

In this section, we establish some important statements about the Young diagram which will be used crucially in later sections.

2.1. **Motzkin paths.** We begin with a bijection between box-ball states and a class of lattice paths we call *h-restricted Motzkin*, a minor variant of the bijection with Dyck paths in [26]. A function $f : \mathbb{R}^+ \to \mathbb{R}$ is a *lattice path* if f is the linear interpolation of some function $\gamma : \mathbb{N}_0 \to \mathbb{Z}$. A lattice path f is called *Motzkin* if it is nonnegative, compactly supported, and consists only of $(1,1)$, $(1,-1)$, and $(1,0)$ steps (which we refer to as 'upstrokes,' 'downstrokes,' and '*h*-strokes,' respectively). We say that a Motzkin path is *h-restricted* if its *h*-strokes occur only on the *x*-axis. Finally, if Γ is a Motzkin path, we write Γ_k for $\Gamma(k)$, $k \geq 0$.

The aforementioned bijection maps a (compactly supported) configuration $X : \mathbb{N} \to \{0,1\}$ to the h-restricted Motzkin path $\Gamma(X)$ defined by linear interpolation of its values on $\mathbb{N}_0^2$, which are given recursively by $\Gamma(X)_0 = 0$ and

$$\Gamma(X)_{k+1} - \Gamma(X)_k = \begin{cases} +1 & \text{if } X(k+1) = 1 \\ -1 & \text{if } X(k+1) = 0 \text{ and } \Gamma(X)_k \geq 1 \\ 0 & \text{if } X(k+1) = 0 \text{ and } \Gamma(X)_k = 0 \end{cases}$$

for all $k \geq 0$. The inverse map from paths to configurations proceeds by writing a 0 for each downstroke or h-stroke and a 1 for each upstroke. See Figure 2 for an illustration.
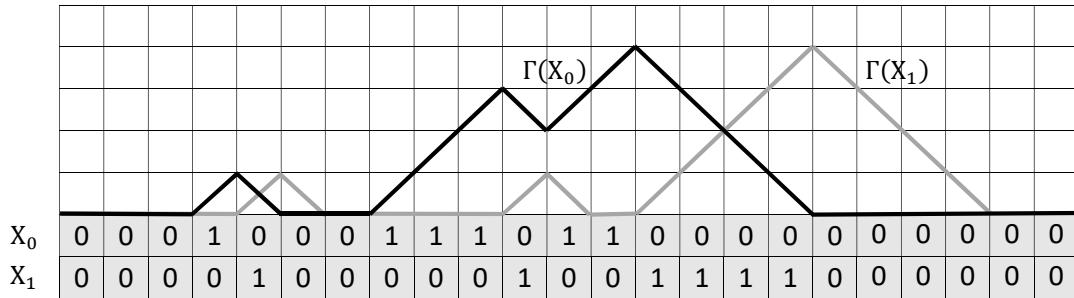
FIGURE 2. The top shaded row shows an initial box-ball configuration X_0, and the black path is $\Gamma(X_0)$. To update, balls are placed at downstrokes of $\Gamma(X_0)$, resulting in the configuration X_1 and the grey path $\Gamma(X_1)$.

The shape of this path tells us how to evolve the system by a single sweep: A ball is picked up at each upstroke and deposited at each downstroke. Specifically, label the balls $1,\ldots,m$ from left to right. (This labeling applies only to states, not the system as a whole. In subsequent sweeps, the label of a particular ball may change.) Then the jth upstroke occurs at the site where the carrier picks up the ball labeled j. The site at which she deposits ball j is determined by drawing a horizontal line from the center of the jth upstroke to the first downstroke on its right. From this description, we see that the height of the path at any site equals the number of balls in the carrier's stack after she visits that site. When the sweep is completed, the new state of the system corresponds to the unique path formed by converting each downstroke to an upstroke and then adding h-strokes and downstrokes so that it is h-restricted Motzkin.

Formally, the box-ball state X_{s+1} is given in terms of the Motzkin path $\Gamma(X_s)$ by

$$X_{s+1}(k+1) = \mathbf{1}\big\{\Gamma(X_s)_{k+1} - \Gamma(X_s)_k = -1\big\}$$

where $\mathbf{1}$ is the indicator function.

2.2. **Hill-flattening and excursion operators.** We now describe two methods of constructing a Young diagram $\Lambda(\Gamma)$ associated with a (not necessarily h-restricted) Motzkin path Γ. As usual, we denote the ith row and j^{th} column by $\rho_i(\Gamma)$ and $\lambda_j(\Gamma)$.

First we give the row-wise construction using the *hill-flattening operator* $\mathscr{H}$ defined on the set of all Motzkin paths. To begin, we say that an interval $[a,b]$ with $a, b \in \mathbb{N}_0$ and $a \le b$ is a *hill interval* of the Motzkin path Γ if for every $c \in [a,b]$, $\Gamma_{a-1} = \Gamma_c - 1 = \Gamma_{b+1}$. We write $\mathscr{I}(\Gamma)$ for the collection of all hill intervals of Γ, and denote the number of hill intervals by $\rho(\Gamma) = |\mathscr{I}(\Gamma)|$. The hill-flattening operator $\mathscr{H}$ is then defined by

$$\mathscr{H}(\Gamma)_k = \begin{cases} \Gamma_k - 1 & \text{if } k \text{ is contained a hill interval of } \Gamma \\ \Gamma_k & \text{otherwise} \end{cases}$$

for $k \in \mathbb{N}_0$.

A *hill* of Γ is the graph of Γ over $[a-1, b+1]$ with $[a,b]$ a hill interval. Thus hills consist of a single upstroke, followed by zero or more h-strokes, followed by a single downstroke. Call a hill with no h-strokes a *peak*. Then the hill-flattening operator $\mathscr{H}$, when applied to Γ, *flattens* each hill of Γ by replacing the upstroke and downstroke with h-strokes and then lowering any intermediate h-strokes so that the path remains connected.

Note that each application of the hill-flattening operator decreases the maximum height of the Motzkin path by 1 and never increases the number of hills, so

$$\rho(\Gamma) \ge \rho(\mathscr{H}(\Gamma)) \ge \rho(\mathscr{H}^2(\Gamma)) \ge \cdots \ge \rho(\mathscr{H}^{\max\Gamma}(\Gamma)) = 0.$$

We define the Young diagram $\Lambda(\Gamma)$ associated to the Motzkin path Γ as having i^{th} row of length $\rho_i(\Gamma) = \rho(\mathscr{H}^{i-1}(\Gamma))$ for $1 \le i \le \max\Gamma$. Here repeated applications of $\mathscr{H}$ are denoted by $\mathscr{H}^{j+1}(f) = \mathscr{H}\big(\mathscr{H}^j(f)\big)$ with $\mathscr{H}^0$ the identity operator. In particular, given a box-ball configuration $X : \mathbb{N}_0 \to \{0,1\}$ of finite support, we can construct the Young diagram $\Lambda(\Gamma(X))$. See Figure 3 for an illustration.
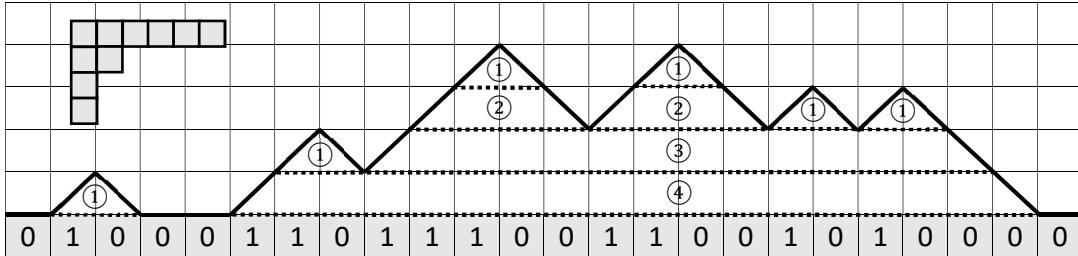

FIGURE 3. Construction of Young diagram via hill-flattening procedure applied to the Motzkin path associated with box-ball configuration X. The bottom row is the configuration X and the black path is $\Gamma = \Gamma(X)$. Trapezoidal regions with label i are the hills of $\mathscr{H}^{i-1}(\Gamma)$, each of which becomes a distinct cell in the ith row of $\Lambda(\Gamma)$. The resulting Young diagram $\Lambda(\Gamma)$ is depicted in the upper left corner.

Now consider a box-ball system $(X_s)_{s \geq 0}$ started from a configuration $X_0 : \mathbb{N}_0 \to \{0, 1\}$. The following lemma says that for each $s \geq 0$, the corresponding Young diagram $\Lambda(\Gamma(X_s))$ is independent of s and its column lengths correspond to the lengths of the solitons.

Lemma 2.1. $\Lambda(\Gamma(X_s)) = \Lambda(\Gamma(X_{s+1}))$ *for all* $s \geq 0$. *Moreover,* $\Lambda(\Gamma(X_0)) = \Lambda(X_0)$.

Next, we give the column-wise construction of $\Lambda(\Gamma)$. The key observation is that the jth longest column length, which we denote by λ_j, is obtained by successively applying the excursion operator to Γ $j - 1$ times and then taking a maximum.

Lemma 2.2. *Let* Γ *be a Motzkin path and let* $\lambda_j(\Gamma)$ *denote the length of the* j^{th} *column of* $\Lambda(\Gamma)$. *Then*

$$\lambda_j(\Gamma) = \max \mathscr{E}^{j-1}(\Gamma), \quad 1 \leq j \leq \rho(\Gamma).$$

In particular, if $(X_s)_{s \in \mathbb{N}_0}$ *is a finitely supported box-ball system with initial configuration* $X_0 : \mathbb{N} \to \{0, 1\}$, *then*

$$\lambda_j(X_0) = \max \mathscr{E}^{j-1}(\Gamma(X_0)).$$

We relegate the proofs of these lemmas, along with that of Lemma 2.3 below, to Appendix A in order to maintain the flow of the paper.

Lemma 2.2 gives the following column-wise construction of $\Lambda(\Gamma)$. Let $\mathtt{m} = \mathtt{m}(\Gamma)$ be the location of the rightmost global maximum of Γ, and set $\lambda_1(\Gamma) = \Gamma_{\mathtt{m}}$, the maximum height of Γ. To find $\lambda_2(\Gamma)$, one first computes $\mathscr{E}(\Gamma)$ by traversing Γ to the left and right of $\mathtt{m}$ as follows: Starting with height 0 at $\mathtt{m}$, move to the left, remaining at height 0 until the first local minimum, and then record the sequence of strokes until the original lattice path returns to the height of this minimum. Then repeat the process, staying at height 0 until encountering a local minimum and then recording the path of the second such excursion. Continue to the beginning of the path and then repeat the procedure moving to the right from $\mathtt{m}$. The resulting path precisely records all 'subexcursions' which are not subsumed by the maximum $(\mathtt{m}, \Gamma_{\mathtt{m}})$. $\lambda_2(\Gamma)$, the length of the second column of $\Lambda(\Gamma)$, is equal to the maximum of $\mathscr{E}(\Gamma)$. Continuing in this fashion gives $\lambda_j(\Gamma) = \max \mathscr{E}^{j-1}(\Gamma)$ for all $j \geq 1$.
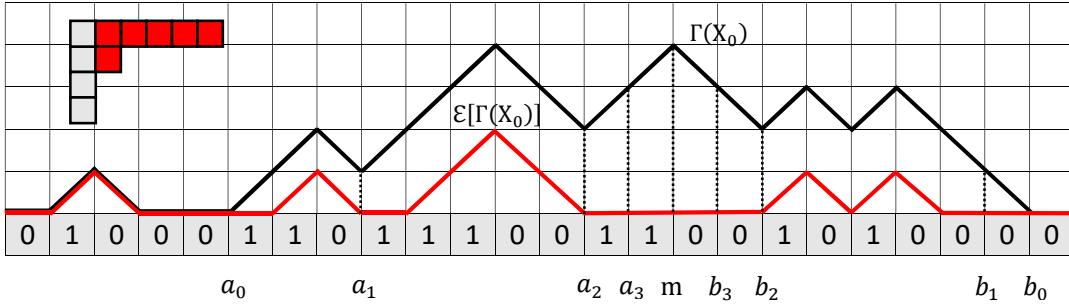

FIGURE 4. The black path is $\Gamma(X_0)$, the red path is $\mathscr{E}(\Gamma(X_0))$, and $\mathtt{m}$ is the location of the rightmost maximum of $\Gamma(X_0)$. The kth box in the bottom row is $X_0(k)$, and the red Young diagram is constructed from the red path via hill flattening. The a and b terms are defined in the proof of Proposition A.2.

In light of Lemma 2.2, it is natural to call $\max \mathscr{E}^{j-1}$ the jth *column length functional*. A crucial advantage of extracting the column length λ_j from the functional $\max \mathscr{E}^{j-1}$ is that this operation is continuous with respect to the topology of $C_0^+(\mathbb{R}^+)$ as stated in the lemma below. This enables us to take various scaling limits of the system.

Lemma 2.3. *For any interval $I \subseteq \mathbb{R}^+$, functions $f, g \in C_0^+(I)$, and $j \geq 1$,*

$$\left| \max \mathscr{E}^{j-1}(f) - \max \mathscr{E}^{j-1}(g) \right| \leq 2\|f - g\|_\infty.$$

Remark 2.4 (Depth process with drains). In private communication with Jim Pitman, we learned that an operator equivalent to $\mathscr{E}_b$ was used in studying Brownian paths and continuum random trees. In our context, given a Motzkin path Γ, flip it upside down and consider it as a bucket filled to the top with water. Given $b \in \mathbb{R}^+$, poke a hole at point $(b, -\Gamma(b))$. This will drain some of the water, and $-\mathscr{E}_b(\Gamma)(x)$ gives the water level at each $x \in \mathbb{R}^+$. For instance, the red path in Figure 4 can be obtained from the black one in this way with drain at $b = \mathfrak{m}(\Gamma)$. A similar procedure can be defined with multiple drains. This operation was applied to Brownian paths to study, for example, the line-breaking construction of the continuum random tree in a Brownian excursion [1]; sampling bridges, meanders, and excursions at independent uniform times [18]; and developments in the tree setting with different metaphors such as "forest growth" and "bead crushing" [19, 20].

2.3. **Rooted forests.** In this subsection, we develop an alternative perspective for constructing the Young diagram from an associated rooted forest. The idea is to collapse a Motzkin path to a rooted forest by horizontal identification. Intuitively, one paints the underside of the graph of each excursion with glue and then compresses it horizontally to obtain a tree. Then the original Motzkin path can be viewed as the contour process (or Harris walk in the random setting) of the rooted forest so constructed. This point of view will be especially useful for thinking about arguments in Section 7.

To begin, recall that a *rooted forest* is a sequence of vertex-disjoint plane trees $\{T_i\}_{i \geq 1}$ such that each T_i is rooted at a vertex $\mathtt{r}_i \in V(T_i)$. The *level* of a vertex $v \in T_i$ is defined as $\ell(v) = d(v, \mathtt{r}_i)$ where d is the graph distance. Given a Motzkin path Γ, we define a rooted forest $\mathfrak{F}(\Gamma)$ as follows: Let $G(\Gamma) = (V, E)$ be the graph with vertex set $V = \{(k, \Gamma_k)\}_{k \in \mathbb{N}_0} \subset \mathbb{N}_0^2$ and adjacency relation

$$(a, \Gamma_a) \overset{\text{adj}}{\sim} (b, \Gamma_b) \iff |a - b| = 1 \text{ and } \Gamma_a, \Gamma_b \text{ not both 0}.$$

In words, $G(\Gamma)$ is obtained from Γ by removing the h-strokes at 0 but retaining all vertices. Clearly each component of $G(\Gamma)$ is isomorphic to a path beginning and ending at height 0, and there are only finitely many such paths since Γ has finite support. Arranging the components from left to right so that their vertex labels are increasing, let P_i denote the ith component from the left. Define an equivalence relation $\sim$ on the vertex set of $G(\Gamma)$ by

$$(a, \Gamma_a) \sim (b, \Gamma_b) \iff 0 < \Gamma_a = \Gamma_b \leq \Gamma_x \text{ for all } x \in [a, b],$$

and write $T_i = P_i / \sim$ for the resulting rooted tree; see Figure 5. The rooted forest associated with Γ is $\mathfrak{F}(\Gamma) = \{T_i\}_{i \geq 1}$.

We can recover Γ from $\mathfrak{F}(\Gamma)$ by keeping track of the levels of the vertices explored in depth-first search. This exploration process begins at the root of T_1 and visits nodes from bottom to top and from left to right in such a way that it backtracks to the parent of the current node only if there is no child left to visit. After exhausting all nodes in T_1, the explorer moves to the second tree T_2, and so on.

More concretely, let $\iota : \mathbb{N}_0 \to V(\mathfrak{F})$ be the function which maps k to the location of the depth-first search at step k so that $\iota(0) = \mathtt{r}_1$, $\iota(k+1)$ is the leftmost unvisited child of $\iota(k)$ if such a child exists, and $\iota(k+1)$ is the parent of $\iota(k)$ if its children have all been visited. (Here the parent of $\mathtt{r}_i$ is taken to be $\mathtt{r}_{i+1}$.) The depth-first-search ordering of the vertices of $\mathfrak{F}$ is given by $u \prec v$ if $\min\{k : \iota(k) = u\} < \min\{k : \iota(k) = v\}$. Finally, the *contour process* on $\mathfrak{F}$ is the function $H(\mathfrak{F}) : \mathbb{N}_0 \to \mathbb{N}_0$ which maps k to the level of $\iota(k)$ in $\mathfrak{F}$. By construction, $H(\mathfrak{F})(k) = \Gamma_k$ for every $k \in \mathbb{N}_0$.
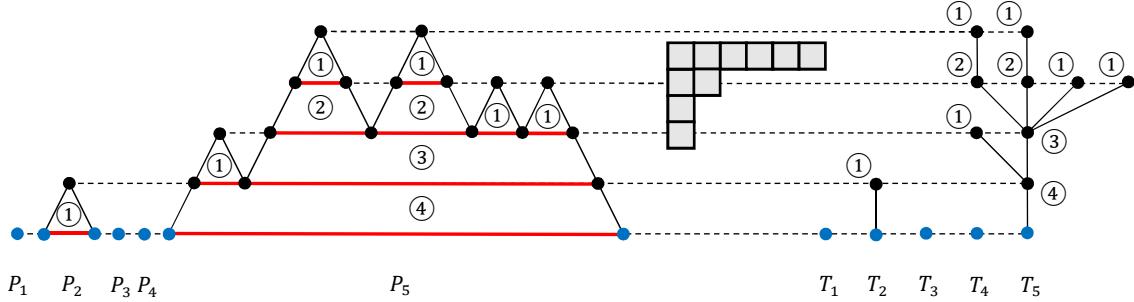
FIGURE 5. Rooted forest $\mathfrak{F}(X_0)$ corresponding to the box-ball configuration X_0 given in Figure 3. Each connected component of $G(\Gamma(X_0))$ (left) becomes a tree rooted at a blue node (right) by identifying vertices connected by the red horizontal lines. Flattening hills of $\Gamma(X_0)$ corresponds to trimming leaves from $\mathfrak{F}(X_0)$.

Now we discuss how to compute the Young diagram $\Lambda(\Gamma)$ from the corresponding rooted forest $\mathfrak{F}(\Gamma)$. In the previous subsection, we constructed the diagram from the Motzkin path via successive applications of the hill-flattening and excursion operators. In terms of rooted forests, these operators can be interpreted in terms of 'trimming' and 'lopping.' Namely, let Υ_0 be the collection of all rooted forests with finitely many vertices and consider the *trimming operator* $\mathcal{T} : \Upsilon_0 \to \Upsilon_0$ which deletes all leaves of the input forest; see Figure 5.

Next, the *lopping operator* $\mathcal{L} : \Upsilon_0 \to \Upsilon_0$ is defined as follows: Given a rooted forest $\mathfrak{F} = \{T_i\} \in \Upsilon_0$, find the rightmost node of maximal level, say $v_{\mathsf{m}} \in V(T_k)$. Set $q = \iota^{-1}(v_{\mathsf{m}})$ and let γ be the unique path from $\mathbf{r}_k$ to v_{m}. Now let $\mathfrak{F}_1$ and $\mathfrak{F}_2$ be the rooted forests induced from $\mathfrak{F}$ such that $V(\mathfrak{F}_1) = \iota([1, q])$ and $V(\mathfrak{F}_2) = \iota([q, \infty))$. Then $\mathcal{L}(\mathfrak{F})$ is obtained by first deleting all edges contained in the copies of γ from $\mathfrak{F}_1$ and $\mathfrak{F}_2$, and then taking the union of the resulting rooted forests with components ordered according to the depth-first search; see Figure 6.
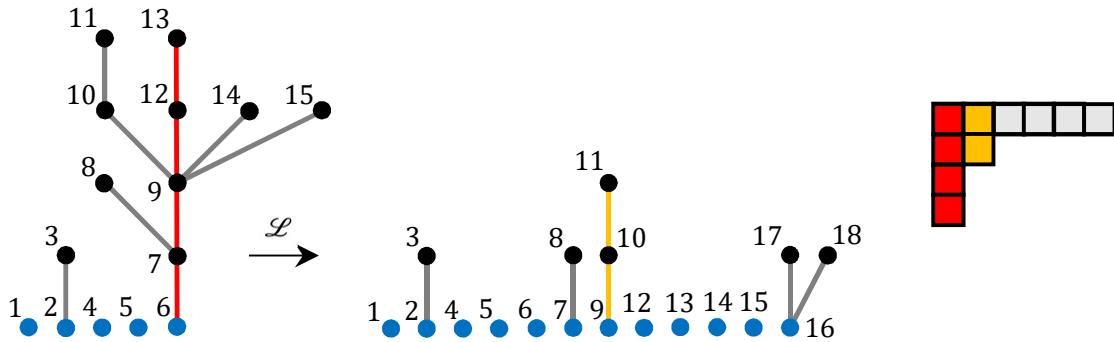

FIGURE 6. The rooted forest $\mathfrak{F} = \mathfrak{F}(\Gamma(X_0))$ on the left appeared in Figure 5, and the one on the right is $\mathcal{L}(\mathfrak{F})$. Numbers next to nodes indicate depth-first-search ordering and $q = 13$. Note that the maximum height of $\mathfrak{F}$ and $\mathcal{L}(\mathfrak{F})$ correspond to the first and second columns of $\Lambda(X_0)$, respectively.

The following proposition shows that these operators are compatible with each other and gives a way to construct the Young diagram $\Lambda(\Gamma)$ from $\mathfrak{F}(\Gamma)$.

Proposition 2.5. *For each Motzkin path* Γ*, we have the following:*

(i) $\mathfrak{F}\big(\mathcal{H}(\Gamma)\big) = \mathcal{T}\big(\mathfrak{F}(\Gamma)\big)$.

(ii) $\mathfrak{F}\big(\mathcal{E}(\Gamma)\big) = \mathcal{L}\big(\mathfrak{F}(\Gamma)\big)$.

(iii) *For each* $1 \le i \le \max\Gamma$, $\rho_i = $ *# of leaves in* $\mathcal{T}^{i-1}(\mathfrak{F}(\Gamma))$.

(iv) *For each* $1 \le j \le \rho(\Gamma)$, $\lambda_j = $ *maximal level of nodes in* $\mathcal{L}^{j-1}(\mathfrak{F}(\Gamma))$.

Proof. For (i), note that leaves in the forest correspond to hills in the path, so applying $\mathcal{H}$ to Γ results in the forest obtained by applying $\mathcal{T}$ to $\mathfrak{F}(\Gamma)$. For (ii), observe that $\mathcal{E}$ only affects the rightmost excursion of maximal height in Γ, $\mathcal{L}$ only affects the rightmost tree of maximal height in $\mathfrak{F}(\Gamma)$, and the 'bushes' growing off of the 'trunk' of this tree correspond precisely to the subexcursions in the corresponding path component which are not subsumed by the maximum.

Now assertion (i) shows that $\mathfrak{F}(\mathcal{H}^{i-1}(\Gamma)) = \mathcal{T}^{i-1}(\mathfrak{F}(\Gamma))$ for all $1 \le i \le \max\Gamma$, and ρ_i is the number of hill intervals of $\mathcal{H}^{i-1}(\Gamma)$, which equals the number of leaves in $\mathfrak{F}(\mathcal{H}^{i-1}(\Gamma)) = \mathcal{T}^{i-1}(\mathfrak{F}(\Gamma))$, and (iii) follows. Finally, given a rooted forest $\mathfrak{F}$, denote by $\|\mathfrak{F}\|$ the maximal level of nodes in $\mathfrak{F}$. Then $\|\mathfrak{F}(\Gamma)\| = \max\Gamma$, so (ii) implies

$$\left\| \mathcal{L}^{i-1}\big(\mathfrak{F}(\Gamma)\big) \right\| = \left\| \mathfrak{F}\big(\mathcal{E}^{i-1}(\Gamma)\big) \right\| = \max\mathcal{E}^{i-1}(\Gamma) = \lambda_i(\Gamma). \qquad \square$$

We remark that Proposition 2.5 (iv) holds if we replace the lopping operator $\mathcal{L}$ by the much simpler one which simply contracts the rightmost longest path into a single root. However, for this contraction operator Proposition 2.5 (ii) no longer holds.

3. RANDOM BOX-BALL SYSTEM AND HARRIS WALK

In this section, we describe stochastic objects corresponding to the random box-ball system introduced in Subsection 1.3.

3.1. **Harris walks.** Fix $p \in (0,1)$, and let $\xi_1, \xi_2, \dots$ be i.i.d. with $\mathbb{P}\{\xi_1 = 1\} = p$ and $\mathbb{P}\{\xi_1 = -1\} = 1 - p$. Let $X^p, X^{n,p} \in \{0,1\}^{\mathbb{N}}$ be as in Subsection 1.3, and let $\{S_k\}_{k=0}^{\infty}$ be the associated random walk, where $S_0 = 0$ and $S_k = \xi_1 + \cdots + \xi_k$. The *Harris walk* $\{H_k\}_{k=0}^{\infty}$ associated with X^p is defined by $H_0 = 0$ and $H_k = (H_{k-1} + \xi_k) \vee 0$ for $k \ge 1$. In other words, H_k is a simple random walk with increments ξ_j, except that downsteps at 0 are censored.

This defines an irreducible and aperiodic birth-and-death chain on $\mathbb{N}_0$ with transition probabilities $P(x, x+1) = p$, and $P(x, (x-1) \vee 0) = 1 - p$. One readily verifies that the chain is reversible with respect to the measure $\mu(x) = \theta^{-x}$ where $\theta = (1-p)/p$. Note that the sum $\sum_{k \ge 1} \theta^k$ converges if and only if $p > 1/2$, so the chain is transient for these values of p and recurrent for $p \le 1/2$. It is null recurrent when $p = 1/2$ since then $\sum_{k \ge 1} \theta^{-k} = \infty$, and it is positive recurrent for $p < 1/2$ as the latter sum converges in this case. (See [12] for background on recurrence criteria for birth-and-death chains.) In the ergodic regime, $p < 1/2$, we can normalize μ to obtain the stationary distribution $\pi(x) = [(1-2p)/(1-p)]\theta^{-x}$.

Now the random Motzkin path $\Gamma(X^{n,p})$ is given by the trajectory of the Harris walk up to time n, completed by appending downstrokes at the end until the height reaches 0 and appending h-strokes thereafter. More precisely, if we define $H : \mathbb{R}^+ \to \mathbb{R}^+$ to be the linear interpolation of the Harris walk, $H(t) = H_{\lfloor t \rfloor} + (t - \lfloor t \rfloor)(H_{\lfloor t+1 \rfloor} - H_{\lfloor t \rfloor})$, then we have

$$\Gamma(X^{n,p})(x) = H(x)\mathbf{1}_{[0,n]}(x) + \max\big(0, H(n) - x + n\big)\mathbf{1}_{[n,\infty)}(x).$$

Moreover, an easy induction argument shows that for all $k \in \mathbb{N}_0$,

$$H_k = S_k - \min_{0 \le r \le k} S_r.$$

Thus if $S : \mathbb{R}^+ \to \mathbb{R}$ is the linear interpolation of the random walk $\{S_k\}_{k=0}^\infty$, then $H = \mathscr{E}_0(S)$. This observation also shows that, marginally, $H_k =_d \max_{0 \le r \le k} S_r$.

3.2. **Galton-Watson forests.**

Following the procedure outlined in Subsection 2.3, one can construct a random rooted forest $\mathfrak{F}(X^{n,p}) = \mathfrak{F}\big(\Gamma(X^{n,p})\big)$ from the trajectory of the truncated Harris walk $\Gamma(X^{n,p})$, and it turns out that $\mathfrak{F}(X^{n,p})$ has the same law as the sub-forest of a Galton-Watson forest with mean offspring number $p/(1-p)$ consisting of the first n nodes revealed by depth-first search.

To be precise, let $\{\zeta_j^k\}_{j,k\ge1}$ be an array of i.i.d. $\mathbb{N}_0$-valued random variables, and define the sequence $\{Z_k\}_{k\ge0}$ by $Z_0 = 1$ and

$$ Z_{k+1} = \begin{cases} \zeta_1^{k+1} + \cdots + \zeta_{Z_k}^{k+1} & \text{if } Z_k > 0 \\ 0 & \text{if } Z_k = 0. \end{cases} $$

The interpretation is that Z_k is the population size in the kth generation of a species in which individuals survive for a single generation and produce an i.i.d. number of offspring before dying. ζ_j^{k+1} is the number of offspring of the jth individual in generation k, and the common law of the ζ's is called the *offspring distribution*. The family tree T for this population is known as a *Galton-Watson tree*. We will be interested in Galton-Watson trees with geometric offspring distribution

$$ \mathbb{P}\{\zeta_j^k = x\} = p^x(1-p), \quad x \in \mathbb{N}_0, $$

which is the number of independent Bern(p) trials preceding the first failure. Observe that $\mathbb{E}[\zeta_j^k] = p/(1-p)$, so T is subcritical if $0 < p < 1/2$, critical if $p = 1/2$, and supercritical if $1/2 < p < 1$. The law of a Galton-Watson tree with Geom$(1-p)$ offspring distribution will be denoted by GWT(p).

We call a sequence of i.i.d. Galton Watson trees $\mathfrak{F}_{GW} = \{T_i\}_{i\ge1}$ a *Galton-Watson forest*, and write GWF(p) for the law of a forest of i.i.d. GWT(p) trees. It is well known that for $0 < p \le 1/2$, each component T_i is finite with full probability [7, Ch. 5.3.4], so the depth-first-search visits all nodes in the forest. However, for $p > 1/2$, each component has a positive probability of being infinite, so almost surely there exists an index $I < \infty$ such that $|T_i| < \infty$ for all $i < I$ and $|T_I| = \infty$. Thus for $p > 1/2$, the depth-first-search cannot pass beyond the leftmost infinite branch in T_I; see Figure 8.

Now let $\mathfrak{F}_p \sim$ GWF(p), write $\mathfrak{F}_{n,p}$ for the vertex-induced subforest of $\mathfrak{F}_p$ on the nodes $\iota([1,n]) \subseteq V(\mathfrak{F}_p)$ which are visited by the depth-first-search in the first n steps, and write GWF(n,p) for the law of $\mathscr{F}_{n,p}$.

Proposition 3.1. $\mathfrak{F}(X^{n,p}) \sim$ GWF(n,p).

Proof. Let $\Gamma = \Gamma(X^p)$ and $\mathfrak{F} = \mathfrak{F}(\Gamma)$. Denote by Z_v the number of children of node $v \in V(\mathfrak{F})$. We will show that the Z_v's are i.i.d. and have the law of the number of independent Bern(p) trials before the first failure. This will imply that the Harris walk $\{H_k\}_{k=0}^\infty$ is distributed as the contour process of $\mathfrak{F}_p$. Then the relation between $\Gamma(X^{n,p})$ and H from the previous subsection yields the assertion.

Let $\mathfrak{F}(X^p) = \{T_i\}_{i\ge1}$ and fix a node $v \in V(T_i)$ for some $i \ge 1$. Let P_i be the path component in $G(\Gamma)$ which is collapsed to T_i via the equivalence relation $\sim$. Note that the number of nodes in $P_i \setminus \{v\}$ which are identified with v equals the number of children of v. Let $x = (a_0, \Gamma_{a_0})$ be such a vertex of P_i with a_0 minimal. If $\Gamma_{a_0+1} - \Gamma_{a_0} = \xi_{a_0+1}$ is 1, then the depth-first search finds the first child of v; otherwise, v is childless and the search moves to its parent or to the root of next tree T_{i+1} depending on whether $\Gamma_{a_0} \ge 1$ or $\Gamma_{a_0} = 0$. If $\xi_{a_0+1} = 1$, then let $a_1 = \min\{k \ge a_0 : \Gamma_k = \Gamma_{a_0}\}$ be the first return time to level Γ_{a_0} after a_0. (a_1 may be infinite if $p > 1/2$.) As before, the depth-first search finds the second child of v if and only if $\xi_{a_1+1} = 1$. Continuing thusly, we see that Z_v has a Geom$(1-p)$ distribution, and the proof is complete. $\square$

Proposition 3.1 allows us to describe the joint distribution of the first i rows or the first j columns in the random box-ball system started at $X^{n,p}$ in terms of Galton-Watson Forests.

Corollary 3.2. *Suppose that $\mathfrak{F} \sim \mathrm{GWF}(n, p)$. For each $i \geq 1$, let $\mathfrak{l}_i$ and $\mathfrak{h}_i$ be the number of leaves in $\mathcal{T}^{i-1}(\mathfrak{F})$ and the maximum height of $\mathcal{L}^{i-1}(\mathfrak{F})$, respectively. Then for any $1 \leq i \leq \max(\Gamma(X^{n,p}))$ and $1 \leq j \leq \rho(\Gamma(X^{n,p}))$, we have*

$$[\rho_1(n), \rho_2(n), \ldots, \rho_i(n)] =_d [\mathfrak{l}_1, \mathfrak{l}_2, \ldots, \mathfrak{l}_i]$$

and

$$[\lambda_1(n), \lambda_2(n), \ldots, \lambda_j(n)] =_d [\mathfrak{h}_1, \mathfrak{h}_2, \ldots, \mathfrak{h}_j].$$

4. ASYMPTOTICS FOR THE ROWS

In this section, we prove our first main result, Theorem 1. From the construction described in Subsection 2.2, we have that $\rho_1(n)$, the length of the first row of $\Lambda^{n,p}$, equals the number of peaks in $\Gamma(X^{n,p})$, which equals the number of 10 patterns in $X^{n,p}$. In general, $\rho_i(n)$ is the number of subexcursions of height i in the Harris walk $\{H_k\}_{k=0}^n$, and these can also be understood in terms of certain binary patterns in the initial configuration.

We begin with a proof of the $i = 1$ case of Theorem 1 using arguments from renewal theory. Strong laws for the other rows can be deduced similarly by considering analogous (delayed) renewal processes, but we will find it more convenient to pursue an alternative approach that will be of use in Section 8.

Proof of Theorem 1 for $i = 1$. First observe that the number of solitons in $X^{n,p}$ is equal to the number of 10 patterns, so $\rho_1(n) = \mathbf{1}\{\xi_n = 1\} + N_{10}(n)$ where $N_{10}(n)$ is the number of 10 patterns in the first n terms. Because of the scaling, it suffices to prove that $N_{10}(n) = \sum_{k=1}^{n-1} \mathbf{1}\{\xi_k = 1, \xi_{k+1} = -1\}$ satisfies the asserted limit theorems.

Now $N_{10}(n)$ counts occurrences of 'head, tail' patterns in a sequence of independent coin flips, which we view as a renewal process. Let T_{10} be distributed as the inter-event times in this process. Then the elementary renewal theorem gives $\mathbb{E}[N_{10}(n)]/n \to 1/\mathbb{E}[T_{10}]$. Since $\mathbb{E}[N_{10}(n)] = (n-1)p(1-p)$, it follows from the strong law for renewal processes that

$$\frac{N_{10}(n)}{n} \to \frac{1}{\mathbb{E}[T_{10}]} = p(1-p) \text{ a.s.}$$

Renewal theory also shows that $N_{10}(n)$ converges weakly to a standard normal random variable when appropriately normalized [3]. To compute the variance, we write $W_k = \mathbf{1}\{\xi_k = 1, \xi_{k+1} = -1\}$ and observe that $\mathbb{E}[W_k] = p(1-p)$, $\mathbb{E}[W_k W_{k+1}] = 0$, and $\mathbb{E}[W_k W_\ell] = p^2(1-p)^2$ when $|k - \ell| > 1$, hence

$$\mathbb{E}[N_{10}(n)^2] = \sum_{k=1}^{n-1} \mathbb{E}[W_k^2] + \sum_{|k-\ell|>1} \mathbb{E}[W_k W_\ell] = (n-1)p(1-p) + (n-2)(n-3)p^2(1-p)^2,$$

so

$$\mathrm{Var}(N_{10}(n)) = \mathbb{E}[N_{10}(n)^2] - \mathbb{E}[N_{10}(n)]^2 = (n-1)p(1-p) - (3n-5)p^2(1-p)^2.$$

The second part of the theorem follows upon invoking Slutsky's theorem to simplify the expression $(N_{10}(n) - \mathbb{E}[N_{10}(n)])/\mathrm{Var}(N_{10}(n))^{1/2}$. $\square$

Remark 4.1. The normal convergence of $\rho_1(n)$ can also be established using Stein's method for sums of locally dependent random variables (see [2, Ch. 9]). Though this approach is more involved, it has the upshot of supplying a Berry-Esseen rate of order $O(n^{-1/2})$. One can show that a central limit theorem also holds for the other row lengths by a similar renewal theory argument, but the corresponding variance computations are not as straightforward.

To treat the $i > 1$ case, we need to establish some more terminology and a useful lemma. Let $\gamma : \mathbb{R}^+ \to \mathbb{R}$ be any nearest neighbor lattice path (so $|\gamma_{k+1} - \gamma_k| \in \{-1, 0, 1\}$ for all $k \in \mathbb{N}_0$). We say that γ has a *subexcursion of height h* on the interval $[r, t]$ if $\gamma_r = \gamma_t < \gamma_s$ for all $s \in (r, t)$ and $\max_{r < s < t} \gamma_s - \gamma_r = h$. Such a subexcursion is said to begin at r and end at t.

Let $\{S_k\}_{k=0}^{\infty}$ be the simple random walk with increment distribution $\mathbb{P}\{S_{k+1} - S_k = 1\} = 1 - \mathbb{P}\{S_{k+1} - S_k = -1\} = p$. For each $i \geq 1$ and $\ell \geq 0$, define the indicator random variable

$$J_\ell^i = \mathbf{1}\{S \text{ has subexcursion of height } i \text{ beginning at time } \ell\}$$

and let τ_ℓ^i be the length of the subexcursion of S_k beginning at $k = \ell$, conditional on $J_\ell^i = 1$. Note that the distribution of τ_ℓ^i does not depend on ℓ by the Markov property of S_k, so we may drop the subscript when notationally convenient. Moreover, due to the negative drift of S_k for $0 < p < 1/2$, it is not hard to see that τ^i has an exponential tail.

The following lemma establishes a polynomial tail bound for the sum of centered indicators and thereby a strong law for the number of subexcursions of fixed height in the interval $[0, n]$. This bound (with $m = 4$ and $\varepsilon = 1/\log n$) will also be used in in the proof of Theorem 3 in Section 8.

Lemma 4.2. *Let $\varsigma = \inf\{k > 0 : S_k = 0\}$ be the first return time of S_k to zero. Fix $i \geq 1$ and $\varepsilon > 0$. Set $\mu_i = \mathbb{P}\{\max_{0 \leq k \leq \varsigma} S_k = i\}$. Then for each fixed $m \geq 2$, there exists a constant $C = C(m, i, p) > 0$ such that for each $n \geq 1$,*

$$\mathbb{P}\left\{ \left| \sum_{\ell=0}^{n} (J_\ell^i - \mu_i) \right| > \varepsilon n \right\} \leq \frac{C}{\varepsilon^{2m} n^{m-1}}.$$

With the above lemma (proved at the end of this section), it is easy to deduce Theorem 1.

Proof of Theorem 1 for $i \geq 1$. The hill-flattening procedure produces a unique column of length at least i for each such subexcursion, so $\rho_i(n)$ is the number of height i subexcursions of H on $[0, n]$. Since the Harris walk H_k and the associated simple random walk S_k over $[0, n]$ share subexcursions of positive height, we may regard $\rho_i(n)$ as the number of subexcursions of S_k occuring on $[0, n]$. Furthermore, we can approximate $\rho_i(n)$ by $N_i(n) := \sum_{\ell=0}^{n} J_\ell^i$ since the two only differ when H has a subexcursion of height at least i beginning at or after $n - i$, hence $|N_i(n) - \rho_i(n)| \leq 1$. Therefore, the assertion follows from Lemma 4.2 with $m = 3$, $\varepsilon = 1/\log n$ and the first Borel-Cantelli lemma. $\square$

Our proof of Lemma 4.2 is based on joint moment estimates of the random variables $J_\ell^i - \mu_i$. Before undertaking this task, we give some preliminary calculations and remarks to set the stage. Fix integers $i \geq 0$ and $0 \leq \ell_1 < \ell_2 < \ell_3$. Clearly $\mathbb{E}[J_{\ell_1}^i - \mu_i] = 0$, and we compute

$$\begin{aligned}
\mathbb{E}\big[(J_{\ell_1}^i - \mu_i)(J_{\ell_2}^i - \mu_i)\big] &= \mu_i \mathbb{E}[J_{\ell_2}^i - \mu_i \,|\, J_{\ell_1}^i = 1] - \mu_i \mathbb{E}[J_{\ell_2}^i - \mu_i] \\
&= \mu_i \mathbb{E}[J_{\ell_2}^i - \mu_i] \mathbb{P}\{\tau_{\ell_1}^i \leq \ell_2 - \ell_1\} - \mu_i^2 \mathbb{P}\{\tau_{\ell_1}^i > \ell_2 - \ell_1\} \\
&= -\mu_i^2 \mathbb{P}\{\tau^i > \ell_2 - \ell_1\},
\end{aligned}$$

where we used the fact that $J_{\ell_2}^i$ is independent of $\{J_{\ell_1}^i = 1\}$ if the excursion starting at ℓ_1 ends at or before ℓ_2, and $J_{\ell_2}^i = 0$ otherwise. Arguing analogously, we find that

$$\begin{aligned}
&\mathbb{E}\big[(J_{\ell_1}^i - \mu_i)(J_{\ell_2}^i - \mu_i)(J_{\ell_3}^i - \mu_i)\big] \\
&\quad = \mu_i \mathbb{E}\big[(J_{\ell_2}^i - \mu_i)(J_{\ell_3}^i - \mu_i) \,|\, J_{\ell_1}^i = 1\big] - \mu_i \mathbb{E}\big[(J_{\ell_2}^i - \mu_i)(J_{\ell_3}^i - \mu_i)\big] \\
&\quad = \mu_i \mathbb{E}\big[(J_{\ell_2}^i - \mu_i)(J_{\ell_3}^i - \mu_i)\big] \mathbb{P}\{\tau_{\ell_1}^i \leq \ell_2 - \ell_1\} - \mu_i^2 \mathbb{E}[J_{\ell_3}^i - \mu_i] \mathbb{P}\{\tau_{\ell_1}^i \in (\ell_2 - \ell_1, \ell_3 - \ell_1]\} \\
&\qquad + \mu_i^3 \mathbb{P}\{\tau_{\ell_1}^i > \ell_3 - \ell_1\} - \mu_i \mathbb{E}\big[(J_{\ell_2}^i - \mu_i)(J_{\ell_3}^i - \mu_i)\big]
\end{aligned}$$

$$= -\mu_i \mathbb{E}\big[(J^i_{\ell_2} - \mu_i)(J^i_{\ell_3} - \mu_i)\big]\mathbb{P}\{\tau^i_{\ell_1} > \ell_2 - \ell_1\} + \mu_i^3 \mathbb{P}\{\tau^i > \ell_3 - \ell_1\}$$

$$= \mu_i^3 \mathbb{P}\{\tau^i > \ell_3 - \ell_2\}\mathbb{P}\{\tau^i > \ell_2 - \ell_1\} + \mu_i^3 \mathbb{P}\{\tau^i > \ell_3 - \ell_1\}.$$

This shows that $J^i_1, J^i_2, \ldots, J^i_n$ are *not* negatively associated for $n \geq 3$, so an immediate Chernoff-Hoeffding type bound is not applicable in our case.

Now in order to prove Lemma 4.2, we need to estimate the joint central moments of the random variables J^i_ℓ. For the sake of readability, this is split up into two propositions. Here and henceforth, the empty product is understood to equal one.

Proposition 4.3. *Fix integers $r \geq 2$, $0 \leq \ell_1 < \ell_2 < \cdots < \ell_r$, and $\alpha_1, \ldots, \alpha_r > 0$. Then*

$$\left| \mathbb{E}\Big[\prod_{k=1}^{r} (J^i_{\ell_k} - \mu_i)^{\alpha_k} \Big] \right| \leq \sum_{s=2}^{r} \left| \mathbb{E}\Big[\prod_{k=s+1}^{r} (J^i_{\ell_k} - \mu_i)^{\alpha_k} \Big] \right| \mathbb{P}\{\tau^i > \ell_s - \ell_1\}$$

$$+ \left(\left| \mathbb{E}\big[(J^i_{\ell_1} - \mu_i)^{\alpha_1}\big] \right| + \mathbb{P}\{\tau^i > \ell_2 - \ell_1\} \right) \left| \mathbb{E}\Big[\prod_{k=2}^{r} (J^i_{\ell_k} - \mu_i)^{\alpha_k} \Big] \right|.$$

Proof. Write $\beta_s = \prod_{k=2}^{s-1}(-\mu_i)^{\alpha_k}$. Casing out according to whether $J^i_{\ell_1}$ is 0 or 1, we see that $(J^i_{\ell_1} - \mu_i)^{\alpha_1} = c_1 J^i_{\ell_1} + d_1$ where $d_1 = (-\mu_i)^{\alpha_1}$ and $c_1 = (1-\mu_i)^{\alpha_1} - (-\mu_i)^{\alpha_1}$. Since $\mu_i \in [0,1]$, a little calculus shows that $|c_1|, |d_1| \leq 1$. Now the strong Markov property for S_k implies that for any $s \geq 2$, $J^i_{\ell_s}$ is independent of $J^i_{\ell_1}$ if the excursion starting at ℓ_1 ends at a site less than or equal to ℓ_s; otherwise $J^i_{\ell_s} = 0$. By partitioning according to the length $\tau^i_{\ell_1}$ of the first excursion we compute

$$\mathbb{E}\Big[\prod_{k=1}^{r} (J^i_{\ell_k} - \mu_i)^{\alpha_k} \Big] = d_1(1-\mu_i)\mathbb{E}\Big[\prod_{k=2}^{r} (J^i_{\ell_k} - \mu_i)^{\alpha_k} \Big| J^i_{\ell_1} = 0 \Big] + (c_1 + d_1)\mu_i \mathbb{E}\Big[\prod_{k=2}^{r} (J^i_{\ell_k} - \mu_i)^{\alpha_k} \Big| J^i_{\ell_1} = 1 \Big]$$

$$= c_1 \mu_i \mathbb{E}\Big[\prod_{k=2}^{r} (J^i_{\ell_k} - \mu_i)^{\alpha_k} \Big| J^i_{\ell_1} = 1 \Big] + d_1 \mathbb{E}\Big[\prod_{k=2}^{r} (J^i_{\ell_k} - \mu_i)^{\alpha_k} \Big]$$

$$= c_1 \mu_i \sum_{s=2}^{r} \mathbb{E}\Big[\prod_{k=2}^{r} (J^i_{\ell_k} - \mu_i)^{\alpha_k} \Big| J^i_{\ell_1} = 1, \tau^i_{\ell_1} \in (\ell_{s-1} - \ell_1, \ell_s - \ell_1] \Big]\mathbb{P}\{\tau^i_{\ell_1} \in (\ell_{s-1} - \ell_1, \ell_s - \ell_1]\}$$

$$+ c_1 \mu_i \mathbb{E}\Big[\prod_{k=2}^{r} (J^i_{\ell_k} - \mu_i)^{\alpha_k} \Big| J^i_{\ell_1} = 1, \tau^i_{\ell_1} > \ell_r - \ell_1 \Big]\mathbb{P}\{\tau^i_{\ell_1} > \ell_r - \ell_1\} + d_1 \mathbb{E}\Big[\prod_{k=2}^{r} (J^i_{\ell_k} - \mu_i)^{\alpha_k} \Big]$$

$$= c_1 \mu_i \sum_{s=2}^{r} \beta_s \mathbb{E}\Big[\prod_{k=s}^{r} (J^i_{\ell_k} - \mu_i)^{\alpha_k} \Big]\mathbb{P}\{\tau^i \in (\ell_{s-1} - \ell_1, \ell_s - \ell_1]\}$$

$$+ c_1 \mu_i \beta_{r+1}\mathbb{P}\{\tau^i > \ell_r - \ell_1\} + d_1 \mathbb{E}\Big[\prod_{k=2}^{r} (J^i_{\ell_k} - \mu_i)^{\alpha_k} \Big]$$

$$= c_1 \mu_i \sum_{s=3}^{r} \beta_s \mathbb{E}\Big[\prod_{k=s}^{r} (J^i_{\ell_k} - \mu_i)^{\alpha_k} \Big]\mathbb{P}\{\tau^i \in (\ell_{s-1} - \ell_1, \ell_s - \ell_1]\} + (c_1 \mu_i + d_1)\mathbb{E}\Big[\prod_{k=2}^{r} (J^i_{\ell_k} - \mu_i)^{\alpha_k} \Big]$$

$$- c_1 \mu_i \mathbb{E}\Big[\prod_{k=2}^{r} (J^i_{\ell_k} - \mu_i)^{\alpha_k} \Big]\mathbb{P}\{\tau^i > \ell_2 - \ell_1\} + c_1 \mu_i \beta_{r+1}\mathbb{P}\{\tau^i > \ell_r - \ell_1\}.$$

Since $|c_1|, |\mu_i|, |\beta_s| \leq 1$, $\mathbb{P}\{\tau^i \in (\ell_{s-1} - \ell_1, \ell_s - \ell_1]\} \leq \mathbb{P}\{\tau^i > \ell_{s-1} - \ell_1\}$, and $c_1 \mu_i + d_1 = \mathbb{E}[(J^i_{\ell_1} - \mu_i)^{\alpha_1}]$, the triangle inequality yields

$$\left| \mathbb{E}\left[\prod_{k=1}^r (J^i_{\ell_k} - \mu_i)^{\alpha_k} \right] \right| \leq \sum_{s=3}^r \left| \mathbb{E}\left[\prod_{k=s}^r (J^i_{\ell_k} - \mu_i)^{\alpha_k} \right] \right| \mathbb{P}\{\tau^i > \ell_{s-1} - \ell_1\} + \mathbb{P}\{\tau^i > \ell_r - \ell_1\}$$

$$+ \left| \mathbb{E}[(J^i_{\ell_1} - \mu_i)^{\alpha_1}] \right| \left| \mathbb{E}\left[\prod_{k=2}^r (J^i_{\ell_k} - \mu_i)^{\alpha_k} \right] \right| + \left| \mathbb{E}\left[\prod_{k=2}^r (J^i_{\ell_k} - \mu_i)^{\alpha_k} \right] \right| \mathbb{P}\{\tau^i > \ell_2 - \ell_1\}$$

$$= \sum_{s=2}^r \left| \mathbb{E}\left[\prod_{k=s+1}^r (J^i_{\ell_k} - \mu_i)^{\alpha_k} \right] \right| \mathbb{P}\{\tau^i > \ell_s - \ell_1\}$$

$$+ \left(\left| \mathbb{E}[(J^i_{\ell_1} - \mu_i)^{\alpha_1}] \right| + \mathbb{P}\{\tau^i > \ell_2 - \ell_1\} \right) \left| \mathbb{E}\left[\prod_{k=2}^r (J^i_{\ell_k} - \mu_i)^{\alpha_k} \right] \right|. \qquad \square$$

The key intuition for the next step is that each linear factor $(J^i_{\ell_k} - \mu_k)$ effectively decreases the 'degrees of freedom' by at least a half. This idea is codified in the following proposition.

Proposition 4.4. *Fix integers $r \geq 2$, $0 \leq \ell_1 < \ell_2 < \cdots < \ell_r$, and $\alpha_1, \ldots, \alpha_r \geq 1$. If $I := \{1 \leq k < r : \alpha_k = 1\}$ is nonempty, then there exist constants $C_r, D > 0$ such that*

$$\left| \mathbb{E}\left[\prod_{k=1}^r (J^i_{\ell_k} - \mu_i)^{\alpha_k} \right] \right| \leq C_r \sum_{\substack{I_0 \subseteq [1,r): \\ 2|I_0| \geq |I|}} \exp\left(-D \sum_{j \in I_0} (\ell_{j+1} - \ell_j) \right).$$

Proof. Since the length τ^i of a subexcursion of height i in S_k has exponential tail, we may choose constants $D, D_0 > 0$ such that

$$\mathbb{P}(\tau^i > t) \leq D_0 \exp(-Dt) \tag{1}$$

for all $t \geq 0$.

Also, the exponential is nonnegative, so it's enough to establish the inequality when the outer sum on the right-hand side is taken over a subset of those $I_0 \subseteq [1, r)$ with cardinality at least half that of I. Thus, for instance, we may dispense with the $I = \{1\}$ case by showing that the expectation on the left is bounded by a constant multiple of $\exp(-D(\ell_2 - \ell_1))$. This is an immediate consequence of Proposition 4.3 and Equation (1) since $\mathbb{E}[J^i_{\ell_1} - \mu_i] = 0$, $\left| \mathbb{E}\left[\prod_{k=s}^r (J_{\ell_k} - \mu_i)^{\alpha_k} \right] \right| \leq 1$, and $\mathbb{P}\{\tau^i > \ell_s - \ell_1\} \leq \mathbb{P}\{\tau^i > \ell_2 - \ell_1\}$ for $s \geq 2$.

We now proceed by induction on r. The base case follows from the previous observation as the assumption that $I \neq \emptyset$ implies $I = \{1\}$ when $r = 2$. For the induction step, let $r \geq 3$. Denote by B_1 and B_2 the first and second term in the right-hand side of the displayed inequality in Proposition 4.3, and let K denote the sum over I_0 in the right-hand side of the displayed inequality in Proposition 4.4. By Proposition 4.3, it suffices to show that both B_1 and B_2 can be bounded by some constant times K.

For the bound on B_2, note that the induction hypothesis gives

$$\left| \mathbb{E}\left[\prod_{k=2}^r (J^i_{\ell_k} - \mu_i)^{\alpha_k} \right] \right| \leq C_{r-1} \sum_{\substack{I'_0 \subseteq [2,r): \\ 2|I'_0| \geq |I \cap [2,r)|}} \exp\left(-D \sum_{j \in I'_0} (\ell_{j+1} - \ell_j) \right).$$

If $\alpha_1 \geq 2$, then $I \subseteq [2, r)$, so we have $B_2 \leq 2C_{r-1}K$. Otherwise $\alpha_1 = 1$ and we are assuming $I \neq \{1\}$, so the induction hypothesis and Equation (1) imply

$$\left| \mathbb{E}\left[\prod_{k=2}^{r} (J_{\ell_k}^i - \mu_i)^{\alpha_k} \right] \right| \mathbb{P}\{\tau^i > \ell_2 - \ell_1\} \leq C_{r-1} \sum_{\substack{I_0' \subseteq [2,r): \\ 2|I_0'| \geq |I \cap [2,r)|}} \exp\left(-D\left(\ell_2 - \ell_1 + \sum_{j \in I_0'} (\ell_{j+1} - \ell_j) \right) \right).$$

If we set $I_0 := \{1\} \cup I_0'$ for each $I_0' \subseteq [2, r)$ in the above summation, then $2|I_0| \geq 2 + |I \cap [2, r)| = 2 + |I| - 1 > |I|$. Moreover, the exponential terms can be written as $\exp(-D(\sum_{j \in I_0} \ell_{j+1} - \ell_j))$. Accordingly, we have that $B_2 \leq C_{r-1}K$.

Next, we show that B_1 can be bounded by some constant times K. Writing $m_1 = \max(I)$, we see that $|I \cap [s+1, r)| \geq 1$ for $s < m_1$, so it follows from the inductive hypothesis, Equation (1), and the fact that all central moments are bounded in absolute value by one that

$$B_1 \leq \sum_{s=2}^{m_1-1} C_{r-s} D_0 \sum_{\substack{I_0' \subseteq [s+1,r): \\ 2|I_0'| \geq |I \cap [s+1,r)|}} \exp\left(-D\left(\ell_s - \ell_1 + \sum_{j \in I_0'} (\ell_{j+1} - \ell_j) \right) \right)$$

$$+ \sum_{s=m_1}^{r} D_0 \exp\left(-D(\ell_s - \ell_1) \right),$$

with the convention that the empty sum is zero. For the first term, we view its inner sum as ranging over all $I_0 \subseteq [0, r)$ with $I_0 := [1, s) \cup I_0'$. Note that $2|I_0| = 2(s-1) + 2|I_0'| \geq 2(s-1) + |I \cap [s+1, r)| \geq |I|$ since $s \geq 2$. Furthermore, the sum of the $\ell_{j+1} - \ell_j$ terms over $j \in [1, s)$ is exactly $\ell_s - \ell_1$. Thus the first term above is at most some constant times K. Finally, taking $m_2 = \min(m_1, r - |I| + 1) \geq 2$, we see that the second term is bounded by $D_0 \sum_{s=m_2}^{r} \exp\left(-D(\ell_s - \ell_{s-1}) \right)$, which is a single summand in K and so less than K. This completes the inductive step and the proof. $\qquad\square$

We are now ready to prove Lemma 4.2.

Proof of Lemma 4.2. Fix $m \in \mathbb{N}$, and use Chebyshev's inequality and the linearity of expectation to write

$$\mathbb{P}\left\{ \left| \sum_{\ell=0}^{n} (J_\ell^i - \mu_i) \right| \geq t \right\} \leq t^{-2m} \sum_{0 \leq \ell_1 \leq \cdots \leq \ell_{2m} \leq n} \mathbb{E}\left[(J_{\ell_1}^i - \mu_i) \cdots (J_{\ell_{2m}}^i - \mu_i) \right]$$

Our goal is to show that the right-hand side of the above inequality is $O(t^{-2m} n^{m+1})$. Then letting $t = \varepsilon n$ gives the assertion. (The Landau notation is in terms of $n \to \infty$ throughout this proof.) We first observe that it suffices to bound the contribution from expectations involving at least $m + 1$ distinct ℓ_k's as there are $O(n^m)$ summands involving fewer and each is $O(1)$.

Fix $m < r \leq 2m$ and let $\alpha_1, \ldots, \alpha_r$ be positive integers such that $\sum_{k=1}^{r} \alpha_k = 2m$. Write $r = u + w$ where $w = \sum_{k=1}^{r} \mathbf{1}\{\alpha_k = 1\}$, and let $I = \{1 \leq k < r : \alpha_k = 1\}$ as in the preceding proposition. Since there are $O(1)$ choices for the r's and α_k's, we need only to demonstrate the existence of a constant $C_1 = C_1(r, i, p) > 0$ such that

$$\left| \sum_{0 \leq \ell_1 < \cdots < \ell_r \leq n} \mathbb{E}\left[(J_{\ell_1}^i - \mu_i)^{\alpha_1} \cdots (J_{\ell_r}^i - \mu_i)^{\alpha_r} \right] \right| \leq C_1 n^{m+1}$$

for all $n \geq 1$.

Note that $w \geq 2$ so that $|I| \geq 1$ and Proposition 4.4 applies. Thus we will be done upon showing that for each subset $I_0 \subseteq [1, r]$ such that $2|I_0| \geq |I|$, there exists a constant $C_2 = C_2(i, p) > 0$ such that

$$\left| \sum_{0 \leq \ell_1 < \cdots < \ell_r \leq n} \exp\left(-D \sum_{j \in I_0} (\ell_{j+1} - \ell_j) \right) \right| \leq C_2 n^{m+1} \tag{2}$$

for all $n \geq 1$. (There are $O(1)$ subsets I_0 in the sum from Proposition 4.4.)

To verify Equation (2), first observe that if $\ell_{j+1} - \ell_j > n^{1/2m}$ for some $j \in I_0$, then the corresponding summand is of order $O(\exp(-Dn^{1/2m}))$. As there are $O(n^{2m})$ choices, the contribution from such terms is of order $O(1)$. Accordingly, it suffices to show that there are $O(n^{m+1})$ sequences $0 \leq \ell_1 < \cdots < \ell_r \leq n$ not verifying this condition. To this end, let L be the set of maps $\ell : [r] \to [n] \cup \{0\}$ such that $\ell(j+1) - \ell(j) \leq n^{1/2m}$ for all $j \in I_0$, and let $G = ([r], E)$ be the graph with vertex set $[r]$ and edge set $E = \{\{j, j+1\} : j \in I_0\}$. Then G contains at most $r - |E| = r - |I_0|$ connected components, say $P_1, \ldots, P_N$ where P_i is a path consisting of vertices $\{j_i, j_i + 1, \ldots, j_i + s_i - 1\}$, $s_i = |P_i|$. Now for any $\ell \in L$ and $1 \leq i \leq N$, there at most $n^{1+s_i/2m}$ possible choices for $\ell(P_i)$—n for $\ell(j_i)$ and $n^{1/2m}$ for each of the $s_i - 1$ successive vertices. Since $N \leq r - |I_0|$ and $\sum_{i=1}^N s_i = r \leq 2m$, this gives

$$|L| \leq \prod_{i=1}^N n^{1+s_i/2m} = n^{N+r/2m} \leq n^{r-|I_0|+1}.$$

The assertion then follows since

$$2r - 2|I_0| + 2 \leq 2r - w + 2 \leq 2u + w + 2 \leq 2m + 2,$$

where we have used the fact that $2|I_0| \geq |I|$ and $2u + w \leq \sum_{k=1}^r \alpha_i = 2m$. $\qquad \square$

5. TOP SOLITON LENGTHS IN THE SUBCRITICAL REGIME

In this section, we prove Theorem 2 (i). Fix $p \in (0, 1/2)$ and let $\{H_k\}_{k=0}^\infty$ denote the Harris walk associated with the random box-ball configuration X^p. The main insight is that the jth longest soliton length, $\lambda_j(n)$, is asymptotically equal to the jth largest excursion height of H_k over the interval $[0, n]$, which we denote by $h_j(n)$ (Lemma 5.2). This allows us to obtain limit theorems for the $\lambda_j(n)$ in terms of the $h_j(n)$ (Lemma 5.1).

Before getting into the details, we discuss the main issue in comparing soliton lengths with excursion heights. Clearly $\lambda_j(n) \geq h_j(n)$ due to the hill-flattening construction of the invariant Young diagram (Lemma 2.1). For $j = 1$, we also have $\lambda_1(n) = h_1(n)$ since $\lambda_1(n)$ equals the maximum height of the Harris walk over $[0, n]$ by Lemma 2.2. However, this identity does not hold for $j \geq 2$. Indeed, Lemma 2.2 shows that $\lambda_2(n) = \max \mathscr{E}(\Gamma(X^{n,p}))$, the maximum excursion height of the modified Motzkin path $\mathscr{E}(\Gamma(X^{n,p}))$. While all but the highest excursion of $\Gamma(X^{n,p})$ are preserved after applying the excursion operator $\mathscr{E}$, it might be the case that there is a large subexcursion within the highest excursion which dominates the contribution from the second highest excursion of $\Gamma(X^{n,p})$. In Subsection 5.3, we show that this is not the case asymptotically.

5.1. **Overview and main results.** We begin by stating the main results of this section and using them to prove Theorem 2 (i). Our first step is to obtain limit theorems for the $h_j(n)$ (which will be defined more carefully in the following subsection).

Lemma 5.1. *Set $\theta = (1-p)/p$, $\sigma = (1-2p)/(1-p)$, and $\mu_n = \log_\theta\big((1-2p)\sigma n\big)$. Let $h_j(n)$ be the jth largest excursion height of the associate Harris walk over $[0, n]$. Then for any nondecreasing real sequence $\{x_n\}_{n\geq 1}$,*

$$\liminf_{n\to\infty} \exp\big(\theta^{-x_n}\big)\left(\sum_{k=0}^{j-1} \frac{\theta^{-k(x_n+1)}}{k!}\right)^{-1} \mathbb{P}\left\{h_j(n) \leq x_n + \mu_n\right\} \geq 1,$$

and

$$\limsup_{n\to\infty} \exp\big(\theta^{-(x_n+1)}\big)\left(\sum_{k=0}^{j-1} \frac{\theta^{-kx_n}}{k!}\right)^{-1} \mathbb{P}\left\{h_j(n) \leq x_n + \mu_n\right\} \leq 1.$$

Next, we show that the soliton lengths and excursion heights are essentially the same objects.

Lemma 5.2. *Fix $p \in (0, 1/2)$. Then for each $j \geq 1$,*

$$\lim_{n\to\infty} \mathbb{P}\left\{\lambda_j(n) \neq h_j(n)\right\} = 0.$$

It is then straightforward to derive the main result for soliton lengths in the subcritical regime.

***Proof of Theorem** 2* (i). Fix $j \geq 1$, $x \in \mathbb{R}$, and let $\mu_n = \log_\theta\big((1-2p)\sigma n\big)$. Since $\lambda_j(n) \geq h_j(n)$, we have

$$\mathbb{P}\left\{\lambda_j(n) \leq x + \mu_n\right\} \leq \mathbb{P}\left\{h_j(n) \leq x + \mu_n\right\}.$$

Hence Lemma 5.1 shows

$$\limsup_{n\to\infty} \mathbb{P}\left\{\lambda_j(n) \leq x + \mu_n\right\} \leq \exp(-\theta^{-(x+1)}) \sum_{k=0}^{j-1} \frac{\theta^{-kx}}{k!}.$$

For the other inequality, we have

$$\mathbb{P}\left\{h_j(n) \leq x + \mu_n\right\} \leq \mathbb{P}\left\{\lambda_j(n) \leq x + \mu_n\right\} + \mathbb{P}\left\{\lambda_j(n) \neq h_j(n)\right\},$$

so Lemmas 5.1 and 5.2 show that

$$\liminf_{n\to\infty} \mathbb{P}\left\{\lambda_j(n) \leq x + \mu_n\right\} \geq \exp(-\theta^{-x}) \sum_{k=0}^{j-1} \frac{\theta^{-k(x+1)}}{k!}. \qquad \square$$

5.2. **Excursion heights.** This subsection is devoted to proving Lemma 5.1. Roughly speaking, we proceed by showing that the Harris walk has $\Theta(n)$ excursions by time n. By relating the excursion heights to a gambler's ruin problem, we argue that their distribution has an exponential tail. Taking the maximum over the $\Theta(n)$ excursions shows that the law of $h_1(n)$ is approximated by a Gumbel distribution after scaling appropriately. The other order statistics are handled similarly.

To begin, set $\tau_1 = 0$ and for $k > 1$, define $\tau_k = \inf\{j > \tau_{k-1} : H_j = 0\}$ to be the time of the kth visit to 0. Thus τ_k is the beginning of the kth excursion above the x-axis, and τ_{k+1} is the end of the kth such excursion. (In this section, if the random walk stays at 0, this counts as an excursion of height 0.) Let

$$h_k = \sup\{H_t : \tau_k < t \leq \tau_{k+1}\} = \sup\left\{\sum_{i=\tau_k+1}^{t} \xi_i : \tau_k < t \leq \tau_{k+1}\right\} \vee 0$$

be the maximum height of the kth excursion. The strong Markov property ensures that $h_1, h_2, \ldots$ are i.i.d. $\mathbb{N}_0$-valued random variables. To compute their distribution function, $F(x) = \mathbb{P}\{h_1 \leq x\}$, we observe that $\mathbb{P}\{h_1 = 0\} = 1 - p$ and $\mathbb{P}\{h_1 \leq x\} = \mathbb{P}\{1 \leq h_1 \leq x\} + \mathbb{P}\{h_1 = 0\}$ for $x \geq 1$. In order for the event $\{1 \leq h_1 \leq x\}$ to occur, the random walk must begin with an upstep and then visit zero before visiting $x+1$. The latter occurs with the 'gambler's ruin' probability that a simple random walker, started at the origin and moving right with probability p, hits -1 before hitting x, which is given by $(\theta^x-1)/(\theta^x-\theta^{-1})$

[7, Ch. 5.7]. Putting all of this together shows that $F(x) = (1-p) + p(\theta^x - 1)/(\theta^x - \theta^{-1})$ for all $x \in \mathbb{N}_0$. After a bit of rearranging, we get

$$F(x) = \left(1 - \frac{1-2p}{\theta^{\lfloor x \rfloor + 1} - 1}\right) \mathbf{1}_{[0,\infty)}(x). \tag{3}$$

Now let $h_{1:m}, \ldots, h_{m:m}$ denote the (reversed) order statistics of $h_1, \ldots, h_m$ so that $h_{1:m} \geq \cdots \geq h_{m:m}$ and $\{h_{1:m}, \ldots, h_{m:m}\} = \{h_1, \ldots, h_m\}$ as multisets. Then

$$F_{j:m}(x) := \mathbb{P}\{h_{j:m} \leq x\} = \sum_{k=0}^{j-1} \binom{m}{k} F(x)^{m-k}(1-F(x))^k, \quad j = 1, \ldots, m.$$

In particular, the maximum $h_{1:m}$ has distribution function

$$F_{1:m}(x) = \left(1 - \frac{1-2p}{\theta^{\lfloor x \rfloor + 1} - 1}\right)^m \mathbf{1}_{[0,\infty)}(x).$$

Write $M_n = \sup\{k : \tau_{k+1} \leq n\}$ for the number of excursions completed by time n and let $r_n = \max\{\sum_{i=\tau_{M_n+1}}^r \xi_i : \tau_{M_n+1} \leq r \leq n\}$ be the maximum height attained after the last complete excursion. The *excursion heights* $h_1(n) \geq h_2(n) \geq \cdots \geq h_{M_n+1}(n)$ are the (reversed) order statistics for $h_1, \ldots, h_{M_n}, r_n$. We begin by showing that M_n is sharply concentrated around its mean so that we can essentially treat it as a deterministic sequence.

Proposition 5.3. *If M_n is the number of excursions of H completed by time n, then*

$$\frac{M_n}{n} \to \frac{1-2p}{1-p} \quad a.s. \quad as \ n \to \infty.$$

Proof. We may write $M_n = \sum_{k=1}^n \mathbf{1}\{H_k = 0\}$, the number of visits to 0 in $[1, n]$. Since the Harris walk is ergodic with stationary distribution $\pi(x) = [(1-2p)/(1-p)]\theta^{-x}$ for $p < 1/2$, we can apply the Markov chain ergodic theorem to obtain

$$\frac{M_n}{n} \to \pi(0) = \frac{1-2p}{1-p} \quad a.s. \qquad \square$$

The next ingredient in our argument is a simple stochastic monotonicity result.

Proposition 5.4. *Set $\sigma = (1-2p)/(1-p)$, $p \in (0, 1/2)$. For any real sequence $\{x_n\}_{n\geq 1}$ and any positive integer j, we have that for all $\varepsilon > 0$,*

$$\limsup_{n\to\infty} \mathbb{P}\{h_j(n) \leq x_n\} \leq \limsup_{n\to\infty} \mathbb{P}\{h_{j:\lfloor(\sigma-\varepsilon)n\rfloor} \leq x_n\}$$

and

$$\liminf_{n\to\infty} \mathbb{P}\{h_j(n) \leq x_n\} \geq \liminf_{n\to\infty} \mathbb{P}\{h_{j:\lceil(\sigma+\varepsilon)n\rceil} \leq x_n\}.$$

Proof. Define

$$N^-(n, \varepsilon) = \sup\{t : M_t \leq (\sigma - \varepsilon)n\}$$

and

$$N^+(n, \varepsilon) = \inf\{t : M_t \geq (\sigma + \varepsilon)n\}.$$

It follows from Proposition 5.3 that there is an a.s. finite N such that

$$\{h_1, \ldots, h_{M_{N^-(n,\varepsilon)}}\} \subseteq \{h_1, \ldots, h_{M_n+1}\} \subseteq \{h_1, \ldots, h_{M_{N^+(n,\varepsilon)}}\}$$

with probability one for all $n \geq N$. Because $r_n \leq h_{M_n+1}$ and the probability that h_{M_n+1} is among the j largest of $h_1, \ldots, h_{M_n+1}$ goes to zero as $n \to \infty$, we see that for any $\varepsilon > 0$,

$$\mathbb{P}\{h_{j:M_{N^-(n,\varepsilon)}} \leq h_j(n) \leq h_{j:M_{N^+(n,\varepsilon)}}\} > 1 - \varepsilon$$

when n is sufficiently large, hence

$$\limsup_{n\to\infty} \mathbb{P}\{h_j(n) \le x_n\} \le \limsup_{n\to\infty} \mathbb{P}\{h_{j:M_{N^-(n,\varepsilon)}} \le x_n\}$$

and

$$\liminf_{n\to\infty} \mathbb{P}\{h_j(n) \le x_n\} \ge \liminf_{n\to\infty} \mathbb{P}\{h_{j:M_{N^+(n,\varepsilon)}} \le x_n\}.$$

The desired assertion follows by noting that $M_{N^-(n,\varepsilon)} = \lfloor(\sigma - \varepsilon)n\rfloor$ and $M_{N^+(n,\varepsilon)} = \lceil(\sigma + \varepsilon)n\rceil$ a.s. since 0 is a recurrent state of $\{H_k\}$.   □

We are now in a position to prove the main result of this subsection.

Proof of Lemma 5.1. First, we claim that for any sequence $\{b_n\}_{n\ge1}$ with $\lim_{n\to\infty} b_n/n = c > 0$ and any nondecreasing sequence $\{y_n\}_{n\ge1}$, we have

$$\lim_{n\to\infty} \exp\left((c/\sigma)\theta^{-y_n}\right)\left(1 - \frac{\theta^{-y_n}}{\sigma n - (1-2p)^{-1}\theta^{-y_n}}\right)^{b_n} = 1. \tag{4}$$

Indeed,

$$\log\left[\left(1 - \frac{\theta^{-y_n}}{\sigma n - (1-2p)^{-1}\theta^{-y_n}}\right)^{b_n}\exp\left((c/\sigma)\theta^{-y_n}\right)\right]$$

$$= \theta^{-y_n}(b_n/\theta^{-y_n})\log\left(1 - \frac{\theta^{-y_n}}{\sigma n - (1-2p)^{-1}\theta^{-y_n}}\right) + (c/\sigma)\theta^{-y_n}$$

$$= \theta^{-y_n}\left[\frac{c}{\sigma} + \frac{\log\left(1 - \frac{\theta^{-y_n}}{\sigma n - (1-2p)^{-1}\theta^{-y_n}}\right)}{\theta^{-y_n}/b_n}\right].$$

Since $\theta > 1$ and $\{y_n\}_{n\ge1}$ is nondecreasing, θ^{-y_n} is bounded. The claim follows since a Taylor expansion of the log term shows that

$$\lim_{n\to\infty} \frac{\log\left(1 - \frac{\theta^{-y_n}}{\sigma n - (1-2p)^{-1}\theta^{-y_n}}\right)}{\theta^{-y_n}/b_n} = -\frac{c}{\sigma}.$$

Now fix $\varepsilon > 0$ and a nondecreasing sequence $\{x_n\}_{n\ge1}$. Recall that for any deterministic sequence of integers $\{b_n\}_{n\ge1}$,

$$\mathbb{P}\{h_{j:b_n} \le x_n + \mu_n\} = \sum_{k=0}^{j-1}\binom{b_n}{k}\left(1 - \frac{1-2p}{\theta^{\lfloor x_n+\mu_n\rfloor+1} - 1}\right)^{b_n-k}\left(\frac{1-2p}{\theta^{\lfloor x_n+\mu_n\rfloor+1} - 1}\right)^k$$

$$= \left(1 - \frac{1-2p}{\theta^{\lfloor x_n+\mu_n\rfloor+1} - 1}\right)^{b_n}\sum_{k=0}^{j-1}b_n^{-k}\binom{b_n}{k}\left(1 - \frac{1-2p}{\theta^{\lfloor x_n+\mu_n\rfloor+1} - 1}\right)^{-k}\left(\frac{(1-2p)b_n}{\theta^{\lfloor x_n+\mu_n\rfloor+1} - 1}\right)^k$$

when $x_n + \mu_n \ge 0$. (Since $\{x_n\}_{n\ge1}$ is nondecreasing and $\mu_n = \log_\theta\left((1-2p)\sigma n\right) \nearrow \infty$, this restriction is satisfied for all large n.)

Writing $\nu_n = (x_n + \mu_n) - \lfloor x_n + \mu_n\rfloor$, we have

$$\frac{1-2p}{\theta^{\lfloor x_n+\mu_n\rfloor+1} - 1} = \frac{1-2p}{\theta^{x_n+\mu_n}\theta^{1-\nu_n} - 1} = \frac{\theta^{-x_n}}{\sigma n\theta^{1-\nu_n} - (1-2p)^{-1}\theta^{-x_n}},$$

so, since $\theta > 1$ and $1 - \nu_n \in (0,1]$, we see that

$$\frac{\theta^{-(x_n+1)}}{\sigma n - (1-2p)^{-1}\theta^{-(x_n+1)}} \le \frac{1-2p}{\theta^{\lfloor x_n+\mu_n\rfloor+1} - 1} \le \frac{\theta^{-x_n}}{\sigma n - (1-2p)^{-1}\theta^{-x_n}}. \tag{5}$$

Set $b_n = \lfloor(\sigma - \varepsilon)n\rfloor$ and note that $\lim_{n\to\infty} b_n^{-k}\binom{b_n}{k} = \frac{1}{k!}$. Then the above estimates and show that for all sufficiently large n,

$$\mathbb{P}\{h_{j:b_n} \le x_n + \mu_n\} \le \left(1 - \frac{\theta^{-(x_n+1)}}{\sigma n - (1-2p)^{-1}\theta^{-(x_n+1)}}\right)^{b_n} \sum_{k=0}^{j-1} \frac{1+\varepsilon}{k!}\left(\frac{(\sigma+\varepsilon)n\theta^{-x_n}}{\sigma n - (1-2p)^{-1}\theta^{-x_n}}\right)^k.$$

Thus Equation (4) with $y_n = x_n + 1$ and $b_n = \lfloor(\sigma - \varepsilon)n\rfloor$ gives

$$\limsup_{n\to\infty} \exp\left(\left(1 - \frac{\varepsilon}{\sigma}\right)\theta^{-(x_n+1)}\right)\left(\sum_{k=0}^{j-1} \frac{\theta^{-kx_n}}{k!}\right)^{-1} \mathbb{P}\{h_{j:b_n} \le x_n + \mu_n\} \le (1+\varepsilon)\left(1 + \frac{\varepsilon}{\sigma}\right)^k.$$

By taking $y_n = x_n$ and $b'_n = \lceil(\sigma + \varepsilon)n\rceil$, a similar argument shows that

$$\liminf_{n\to\infty} \exp\left(\left(1 + \frac{\varepsilon}{\sigma}\right)\theta^{-x_n}\right)\left(\sum_{k=0}^{j-1} \frac{\theta^{-k(x+1)}}{k!}\right)^{-1} \mathbb{P}\{h_1(n) \le x_n + \mu_n\} \ge (1-\varepsilon)\left(1 - \frac{\varepsilon}{\sigma}\right)^k.$$

Letting $\varepsilon \searrow 0$ and applying Proposition 5.3 completes the proof. $\qquad\square$

Remark 5.5. Because we are taking the maximum of a random number of excursions, the sequence $\{h_j(n) - \mu_n\}_{n\ge 1}$ does not have a weak limit (and thus neither do the normalized subcritical soliton lengths). To see this, we first recall that

$$\mathbb{P}\{h_{1:b_n} \le x + \mu_n\} = \left(1 - \frac{1-2p}{\theta^{\lfloor x+\mu_n\rfloor+1} - 1}\right)\mathbf{1}\{x + \mu_n \ge 0\}$$

for any real sequence $\{b_n\}_{n\ge 1}$ and any $x \in \mathbb{R}$. Now fix $x > \mu_1$, write $v_n = (x + \mu_n) - \lfloor x + \mu_n\rfloor$, and choose subsequences $\{v_{n_k}\}_{k\ge 1}$ and $\{v_{n_\ell}\}_{\ell\ge 1}$ such that $v_{n_k} \le \frac{1}{3}$ and $v_{n_\ell} \ge \frac{2}{3}$ for all $k, \ell \in \mathbb{N}$. This is possible since $\mu_n = \log_\theta\big((1-2p)\sigma\big) + \log_\theta(n)$ and the fractional part of $\log_\theta(n)$ is dense in $[0,1]$.

Since

$$\frac{1-2p}{\theta^{\lfloor x+\mu_n\rfloor+1} - 1} = \frac{\theta^{-x}}{\sigma n\theta^{1-v_n} - (1-2p)^{-1}\theta^{-x}},$$

$\theta > 1$, $1 - v_\ell \in [0, 1/3]$, and $1 - v_k \in [2/3, 1]$, we have the following analogues of Equation (5):

$$\frac{\theta^{-(x+1/3)}}{\sigma n_\ell - (1-2p)^{-1}\theta^{-(x+1/3)}} \le \frac{1-2p}{\theta^{\lfloor x+\mu_{n_\ell}\rfloor+1} - 1} \le \frac{\theta^{-x}}{\sigma n_\ell - (1-2p)^{-1}\theta^{-x}}$$

and

$$\frac{\theta^{-(x+1)}}{\sigma n_k - (1-2p)^{-1}\theta^{-(x+1)}} \le \frac{1-2p}{\theta^{\lfloor x+\mu_{n_k}\rfloor+1} - 1} \le \frac{\theta^{-(x+2/3)}}{\sigma n_k - (1-2p)^{-1}\theta^{-(x+2/3)}}.$$

Repeating the last part of the proof of Lemma 5.1 (and restricting attention to $h_1(n) = \lambda_1(n)$ to simplify notation) shows that

$$e^x \le \liminf_{\ell\to\infty} \mathbb{P}\{\lambda_1(n_\ell) \le x + \mu_{n_\ell}\} \le \limsup_{\ell\to\infty} \mathbb{P}\{\lambda_1(n_\ell) \le x + \mu_{n_\ell}\} \le e^{x+1/3}$$

and

$$e^{x+2/3} \le \liminf_{k\to\infty} \mathbb{P}\{\lambda_1(n_k) \le x + \mu_{n_k}\} \le \limsup_{k\to\infty} \mathbb{P}\{\lambda_1(n_k) \le x + \mu_{n_k}\} \le e^{x+1}.$$

In particular,

$$\limsup_{\ell\to\infty} \mathbb{P}\{\lambda_1(n_\ell) \le x + \mu_{n_\ell}\} < \liminf_{k\to\infty} \mathbb{P}\{\lambda_1(n_k) \le x + \mu_{n_k}\}. \tag{6}$$

Since the sequence $\{\lambda_1(n) - \mu_n\}$ is tight by Lemma 5.1, both $\{\lambda_1(n_k) - \mu_{n_k}\}$ and $\{\lambda_1(n_\ell) - \mu_{n_\ell}\}$ have subsequential weak limits. As Inequality (6) implies that the limiting distribution functions disagree at x, it follows that $\{\lambda_1(n) - \mu_n\}$ does not converge weakly.

5.3. **Subexcursions within an excursion.** Given an excursion γ of H with length ς and rightmost global maximum at (m^*, h), define $a_\ell = \max\{t \le m^* : \gamma(t) = \ell\}$ and $b_\ell = \min\{t \ge m^* : \gamma(t) = \ell\}$ for $\ell = 0, \ldots, h$. Write $\gamma_{a,\ell} = \gamma|_{[a_{\ell-1}, a_\ell]} - \ell$ and $\gamma_{b,\ell} = \gamma|_{[b_\ell, b_{\ell-1}]} - \ell$. These paths correspond to the portions of γ which, moving away from m^*, begin at the point where γ first descends to height ℓ and end where γ first descends to height $\ell - 1$, except that they are shifted down by ℓ; see Figure 7.

Set $\widetilde{\gamma}_{a,\ell} = \gamma_{a,\ell} \vee 0$, $\widetilde{\gamma}_{b,\ell} = \gamma_{b,\ell} \vee 0$ (which has the effect of changing the downstroke furthest from m^* to an h-stroke) and define

$$\vartheta(k) = \begin{cases} \max \widetilde{\gamma}_{a,k}, & k \le h \\ \max \widetilde{\gamma}_{b,k-h}, & h < k \le 2h \end{cases}.$$

Then $\mathscr{E}(\gamma)$ is the concatenation of $\widetilde{\gamma}_{a,1}, \ldots, \widetilde{\gamma}_{a,h}, \widetilde{\gamma}_{b,h}, \ldots, \widetilde{\gamma}_{b,1}$, so $\max \mathscr{E}(\gamma) = \max_{1 \le k \le 2h} \vartheta(k)$.
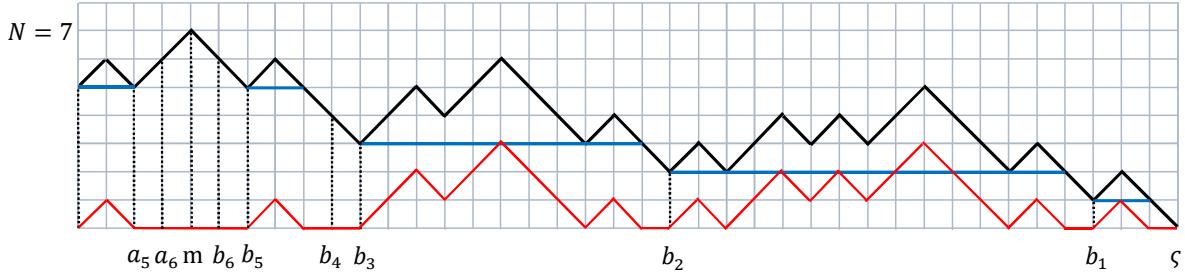
FIGURE 7. The black path is γ and the red path is $\mathscr{E}(\gamma)$. a_ℓ and b_ℓ are the first locations to the left and right of $\mathtt{m}$ that γ is at height ℓ and the blue lines indicate paths between successive values. In these regions, γ and $\mathscr{E}(\gamma)$ differ only by vertical translation.

Now let m_* denote the leftmost global maximum of γ and set $c_\ell = \min\{t \ge m_* : \gamma(t) = \ell\}$, $\gamma_{c,\ell} = \gamma|_{[c_\ell, c_{\ell-1}]} - \ell$, $\widetilde{\gamma}_{c,\ell} = \gamma_{c,\ell} \vee 0$. Define $\omega(k) = \max \widetilde{\gamma}_{c,k} = \max \gamma_{c,k}$ for $k = 1, \ldots, h$. We first observe that $\max\{\omega(1), \ldots, \omega(h)\} \ge \max\{\vartheta(h+1), \ldots, \vartheta(2h)\}$. To see that this is so, let $j = \min_{m_* \le t \le m^*} \gamma(t)$. Then $b_k = c_k$ for all $k < j$, hence $\omega(k) = \max_{c_k \le t \le c_{k-1}} \gamma(t) - k \ge \max_{b_k \le t \le b_{k-1}} \gamma(t) - k = \vartheta(h+k)$ for all $k \le j$ because $c_\ell \le b_\ell$ for all ℓ. On the other hand, since $c_j \le m^* < c_{j-1}$, $\omega(j) = h - j > h - k \ge \vartheta(h+k)$ for all $k > j$. It follows that $\max \mathscr{E}(\gamma) = \max\{\vartheta(1), \ldots, \vartheta(2h)\} \le \max\{\vartheta(1), \ldots, \vartheta(h), \omega(1), \ldots, \omega(h)\}$.

Next we observe that γ is symmetric about $\varsigma/2$ in distribution. This is because, conditional on the excursion length, the law of γ depends only on the number of up and down steps. Accordingly, $\mathscr{E}\gamma|_{[0,m^*]}$ and $\mathscr{E}\gamma|_{[m_*,\varsigma]}$ have the same distribution, so $\max\{\vartheta(1), \ldots, \vartheta(h)\} =_d \max\{\omega(1), \ldots, \omega(h)\}$, and thus

$$\mathbb{P}\{\max \mathscr{E}(\gamma) > x\} \le \mathbb{P}\{\max\{\vartheta(1), \ldots, \vartheta(h), \omega(1), \ldots, \omega(h)\} > x\}$$

$$\le \mathbb{P}\{\max\{\vartheta(1), \ldots, \vartheta(h)\} > x\} + \mathbb{P}\{\max\{\omega(1), \ldots, \omega(h)\} > x\}$$

$$= 2\mathbb{P}\{\max\{\omega(1), \ldots, \omega(h)\} > x\}.$$

To treat the latter probability, note that given $\{h = r\}$, $m_*, c_{r-1}, \ldots, c_0$ are stopping times with respect to the natural filtration, so $\omega(1), \ldots, \omega(r)$ are independent by the strong Markov property. Also, each $\omega(k)$ is stochastically dominated by the random variable Y which gives the maximum value taken by a simple random walker started at 0 and moving right with probability p before hitting -1 (as the path $\gamma_{c,k}$ is constrained to be at height at most $h - k$ whereas the random walker's path has no such restriction). We conclude that on the event $\{h \le r\}$,

$$\mathbb{P}\{\max \mathscr{E}(\gamma) > x\} \le 2\mathbb{P}\{\max\{\omega(1), \ldots, \omega(h)\} > x\} \le 2\left(1 - G(x)^r\right) \tag{7}$$

where $G(x) = \mathbb{P}\{Y \le x\}$ is the gambler's ruin probability [7, Ch. 5.7]

$$G(x) = \frac{\theta^{\lfloor x \rfloor + 1} - 1}{\theta^{\lfloor x \rfloor + 1} - \theta^{-1}} \mathbf{1}_{[0,\infty)}(x).$$

Proof of Lemma 5.2. Fix $j \ge 1$ and let $\varepsilon > 0$ be given. Lemma 5.1 implies that there exist $\delta > 0$, $N \in \mathbb{N}$ such that for each $n \ge N$, the event

$$E_{j,\delta,n} = \{h_1(n) \le n^\delta, h_j(n) \ge 2\delta \log_\theta(n)\}$$

has probability at least $1 - \varepsilon$. Write $H^{(k,n)}$ for the kth highest excursion of $H|_{[0,n]}$, so that $h_k(n) = \max H^{(k,n)}$. As each application of the excursion operator affects only one excursion, $\lambda_1(n), \dots, \lambda_j(n)$ are the j largest values among $\mathscr{E}^{i-1}(H^{(k,n)})$ as i and k range over $\{1, \dots, j\}$. On $E_{j,\delta,n}$, these coincide with $h_1(n), \dots, h_j(n)$ when $\mathscr{E}(H^{(k,n)}) \le 2\delta \log_\theta(n)$ for $k = 1, \dots, j$. Since

$$G(x) = \frac{\theta^{\lfloor x \rfloor + 1} - 1}{\theta^{\lfloor x \rfloor + 1} - \theta^{-1}} \ge 1 - \frac{\theta}{\theta^x - 1}$$

for $x > 0$, Equation (7) implies

$$\mathbb{P}\{\mathscr{E}(H^{(k,n)}) > 2\delta \log_\theta(n); E_{j,\delta,n}\} \le 2\left[1 - G\left(2\delta \log_\theta(n)\right)^{n^\delta}\right] \le 2\left[1 - \left(1 - \frac{\theta}{n^{2\delta} - 1}\right)^{n^\delta}\right]$$

$$\le 2\left[1 - \left(1 - n^\delta \frac{\theta}{n^{2\delta} - 1}\right)\right] \le \frac{4\theta}{n^\delta}.$$

Consequently,

$$\mathbb{P}\{\lambda_k(n) = h_k(n) \text{ for } k = 1, \dots, j\} \le \varepsilon + \frac{4j\theta}{n^\delta},$$

and the claim follows since ε is arbitrary. $\square$

6. TOP SOLITON LENGTHS AT CRITICALITY

In this section we observe that when $p = 1/2$, the (suitably scaled) Harris walk converges weakly to a reflected Brownian motion at the process level. In fact, this weak convergence can be strengthened to "polynomial convergence" by appealing to a result from Drmota [6]. This enables us to deduce scaling limits for the top soliton lengths.

Recall that $C([0,1])$ denotes the space of continuous functions $f : [0,1] \to \mathbb{R}$ equipped with the supremum norm. We say a continuous functional $F : C([0,1]) \to \mathbb{R}$ is of *polynomial growth* if there exists $r \ge 1$ such that $|F(\gamma)| \le \|\gamma\|_\infty^r$ for all $\gamma \in C([0,1])$.

Theorem 6.1 (Theorem 9 of [6]). *Suppose that a sequence of stochastic processes $x_n(t)$ defined on $C([0,1])$ converges weakly to $x(t)$. Furthermore suppose that there exists $s_0 \in [0,1]$ such that for all $r \ge 0$,*

$$\sup_{n \ge 0} \mathbb{E}\left[|x_n(s_0)|^r\right] < \infty,$$

and that for every $\alpha > 1$, there exists $\beta > 0$ and $C > 0$ with

$$\mathbb{E}\left[|x_n(t) - x_n(s)|^\beta\right] < C|t - s|^\alpha \text{ for all } s, t \in [0,1].$$

If $F : C([0,1]) \to \mathbb{R}$ is any continuous functional of polynomial growth, then

$$\lim_{n \to \infty} \mathbb{E}[F(x_n)] = \mathbb{E}[F(x)].$$

We show the following polynomial convergence of Harris walk to the reflected Brownian motion.

Theorem 6.2. *Let $\{B(t) : 0 \le t \le 1\}$ be a standard Brownian motion and define $H^n(t) = H(nt)/\sqrt{n}$ for $0 \le t \le 1$. Then for $p = 1/2$,*

$$\{H^n(t) : 0 \le t \le 1\} \Rightarrow \{|B(t)| : 0 \le t \le 1\} \text{ in } C([0,1]).$$

Furthermore, if $F : C([0,1]) \to \mathbb{R}$ is any continuous functional of polynomial growth, then

$$\lim_{n\to\infty} \mathbb{E}[F(H^n)] = \mathbb{E}(F(|B|)).$$

Proof. Since the rescaled Harris walk $H^n(t)$ is uniformly bounded by $n^{-\frac{1}{2}}|S^n(t)|$, which has moments of all orders and satisfies the Hölder criterion in Theorem 6.1, we only need to show that H^n converges weakly to $|B|$. To this end, recall from Subsection 1.3 that the linear interpolation of the $p = 1/2$ Harris walk is given by $H(t) = \mathcal{E}_0(S)(t) = S(t) - \min_{0 \le r \le t} S(r)$, where S is the linear interpolation of symmetric simple random walk.

Donsker's Theorem shows that after scaling diffusively, $S(t)$ converges weakly to a standard Brownian motion in the space $C([0,1])$. That is, writing $S^n(t) = S(nt)/\sqrt{n}$, we have

$$\lim_{n\to\infty} \mathbb{E}[F(S^n)] = \mathbb{E}[F(B)]$$

for every bounded and continuous functional $F : C([0,1]) \to \mathbb{R}$.

A direct computation shows that for any fixed $b \in [0,1]$, $\mathcal{E}_b$ is (2-Lipschitz) continuous and satisfies $\mathcal{E}_b(cf) = c\mathcal{E}_b(f)$ for all $b, c \ge 0$ (see Proposition A.6 (i) in Subsection A.3), so for every bounded and continuous $G : C([0,1]) \to \mathbb{R}$,

$$\lim_{n\to\infty} \mathbb{E}[G(H^n)] = \lim_{n\to\infty} \mathbb{E}\big[G\big(\mathcal{E}_0(S^n)\big)\big] = \mathbb{E}\big[G\big(\mathcal{E}_0(B)\big)\big],$$

hence H^n converges weakly to $\mathcal{E}_0(B)$. As

$$\mathcal{E}_0(B)(t) = B(t) - \min_{0 \le s \le t} B(s) =_d -B(t) - \min_{0 \le s \le t}\big(-B(s)\big) = \max_{0 \le s \le t} B(s) - B(t),$$

Lévy's $M - B$ theorem (see [16, Ch. 2.3]) implies $\mathcal{E}_0(B) =_d |B|$ and the proof is complete. $\qquad\square$

Now we can use the Lipschitz continuity of column length functionals $\max \mathcal{E}^{j-1}$ to obtain Theorem 2 (ii).

Proof of Theorem 2 **(ii).** First recall that the Motzkin path $\Gamma = \Gamma(X^{n,1/2})$ agrees with the Harris walk H on $[0,n]$, and has only downstrokes until it reaches height 0 on $[n, \infty)$, hence all of its peaks are contained in $[0,n]$. Recall also that the excursion operator deletes the peak at the rightmost maximum and preserves all the other peaks. Thus by Lemma 2.2, we have

$$n^{-1/2}\lambda_j(n) = n^{-1/2}\max \mathcal{E}^{j-1}(\Gamma) = n^{-1/2}\max_{[0,n]}\mathcal{E}^{j-1}(H|_{[0,n]}) = \max_{0 \le t \le 1}\mathcal{E}^{j-1}(H^n).$$

Lemma 2.3 in Section 2 shows that the column length functionals $\max \mathcal{E}^{j-1} : C([0,1]) \to \mathbb{R}$ are Lipschitz, so taking powers gives continuous functionals of polynomial growth, and the claimed convergence follows from Theorem 6.2. A stronger version of the second part of the assertion (concerning orders of column lengths) is shown in Theorem 6.4 below. $\qquad\square$

To establish the order of the other top soliton lengths, we appeal to known results about the marginal densities of the ranked maxima of $|B|$ over all excursions. To state our conclusions precisely, note that the continuity of B ensures that the random subset $\{t : B(t) \ne 0\}$ of $[0,1]$ is a countable union of

maximal disjoint intervals, called the *excursion intervals* of B. We call an excursion interval (a, b) *complete* if $B(a) = B(b) = 0$, and *incomplete* otherwise. All of the excursion intervals are complete except possibly the last one $(g(t), 1]$, where $g(t) = \sup\{0 \le t \le 1 : B(t) = 0\}$ is the last zero of B.

Let $h_1 \ge h_2 \ge \cdots > 0$ be the ranked sequence of values $\sup_{t \in (a,b)} |B_t|$ as (a, b) ranges over all excursion intervals of B. The marginal distributions of the ranked heights over excursions in the reflected Brownian bridge were first obtained by Pitman and Yor [21]. Lagnoux, Mercier, and Vallois [15] pointed out that the probability that the maximum of reflected Brownian motion is obtained during the last incomplete excursion is approximately 0.3069. Csaki and Hu [5] obtained the following explicit expressions for the marginal densities of ranked maxima of reflected Brownian motion over all excursions, including the final meander:

Theorem 6.3. *For each $j \ge 1$ and $y > 0$,*

$$\mathbb{P}\{h_j \ge y\} = 2^{j+1} \sum_{k=0}^{\infty} (-1)^k \binom{k+j-1}{k} \left(1 - \Phi\left((2k+2j-1)y\right)\right)$$

where $\Phi(\cdot)$ is the standard normal distribution function.

Accordingly, Theorem 6.2 and Lemma 2.2 imply

Theorem 6.4. *At criticality, we have that for each $x > 0$*

$$\lim_{n\to\infty} \mathbb{P}\left\{\lambda_1(n) \le x\sqrt{n}\right\} = 1 - 4\sum_{k=0}^{\infty} (-1)^k (1 - \Phi[(2k+1)x]).$$

Furthermore,

$$\limsup_{n\to\infty} \mathbb{P}\left\{\lambda_j(n) \le x\sqrt{n}\right\} \le 1 - 2^{j+1} \sum_{k=0}^{\infty} (-1)^k \binom{k+j-1}{k} (1 - \Phi[(2k+2j-1)x]). \tag{8}$$

In particular, for any $j \ge 1$, $\lambda_j(n) = \Theta(\sqrt{n})$.

Remark 6.5. One might wonder whether the top soliton lengths agree with the top excursion heights as in the subcritical phase. This would imply that the right-hand side of (8) gives the limiting distribution of $\lambda_j(n)/\sqrt{n}$ for all $j \ge 1$. However, we conjecture that this is *not* the case for $p = 1/2$. This is because the random variable Y appearing in the proof of Lemma 5.2 would then have distribution function $G(x) = 1 - 1/(x+2)$ [7, Ch. 4.1], and one cannot find $x_n \in O(\sqrt{n})$, $r_n \in \Omega(\sqrt{n})$ with $1 - \left(1 - \frac{1}{x_n+2}\right)^{r_n} \to 0$.

7. Top soliton lengths in the supercritical regime

In this section, we fix $p \in (1/2, 1)$ and prove Theorem 2 (iii). The intuition is the following. According to Proposition 3.1, the top soliton lengths are encoded in the first n nodes of a Galton-Watson forest $\mathfrak{F} = (T_i)_{i\ge 1} \sim \mathtt{GWF}(p)$. Since the offspring distribution has mean $p/(1-p) > 1$ in the supercritical regime, the random index $I = \min\{i : |T_i| = \infty\}$ is almost surely finite. For n large, about np nodes of the infinite component T_I will be exposed by the Harris walk, which climbs up along the 'leftmost' infinite branch in T_I. Hence $\lambda_1(n)$ should behave like the maximum of a random walk with positive drift, and λ_2 will be the maximum height of the first few finite components $T_1, \ldots, T_{I-1}$ together with the 'bushes' attached to the infinite branch in T_I. We prove the $\lambda_1(n)$ assertion by approximating $\{H_k\}$ by $\{S_k\}$. For subsequent soliton lengths, we appeal to a duality argument: A backward Harris walk started at the highest node will encounter a subcritical Galton-Watson forest, so λ_2 for density p should behave as λ_1 for density $1 - p$; see Figure 8.
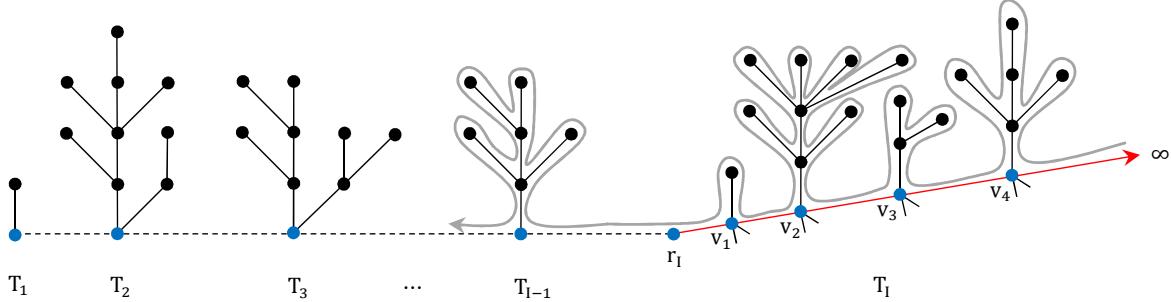
FIGURE 8. Supercritical Galton-Watson forest. T_I is the first infinite component and the red ray is the leftmost infinite branch in T_I on which the usual Harris walk climbs up. The grey contour is the backward Harris walk starting from the last vertex of level N, which encounters a subcritical Galton-Watson forest.

7.1. **Duality and proof of Theorem 2 (iii).** To make the above sketch rigorous, we introduce the notion of flip and dual configurations, which will be used to provide a coupling between the random box-ball configurations $X^{n,p}$ and $X^{n,1-p}$.

Given a random box-ball configuration $X^{n,p} = X^p \mathbf{1}_{[1,n]}$, define the associated box-ball configurations $\widetilde{X}^{n,p}, \widehat{X}^{n,p} : \mathbb{N} \to \{0,1\}$ (which we call the *flip* and *dual*) by

$$\text{(flip)} \qquad \widetilde{X}^{n,p}(k) = 1 - X^{n,p}(k),$$
$$\text{(dual)} \qquad \widehat{X}^{n,p}(k) = \left(1 - X^{n,p}(n-k+1)\right) \mathbf{1}\{1 \le k \le n\}.$$

For each $j \ge 1$, denote $\lambda_j(n) = \lambda_j(X^{n,p})$ and $\widehat{\lambda}_j(n) = \lambda_j(\widehat{X}^{n,p})$.
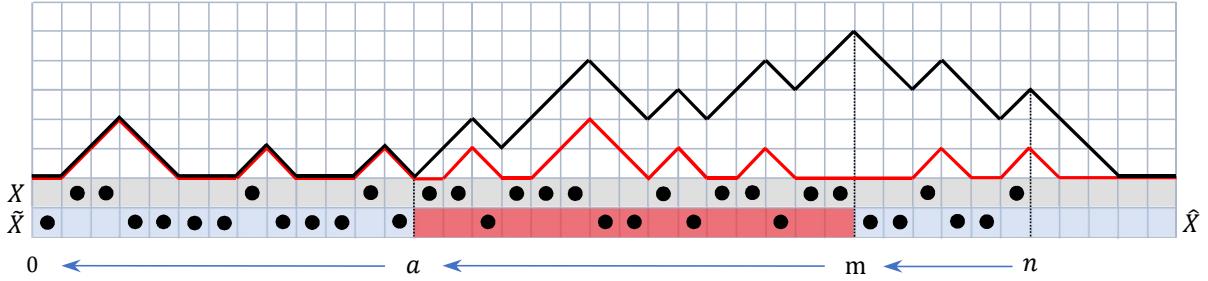
FIGURE 9. Box-ball configuration X, its flip $\widetilde{X}$, and its dual $\widehat{X}$, the latter being read from right to left. The black path is $\Gamma(X)$ and the red is $\mathscr{E}(\Gamma(X))$. Note that if one reads the Motzkin path $\mathscr{E}(\Gamma(X))$ from the righmost maximum $\mathtt{m}$ to the beginning of the associated excursion, $\mathtt{a}$, the corresponding box-ball configuration coincides with the dual $\bar{X}$.

For $1/2 < p < 1$, it is easy to see from the postive drift that $\lambda_1(n) = \max_{1 \le k \le n} H_k \approx \max_{1 \le k \le n} S_k \approx S_n$, where $\{S_k\}_{k \ge 0}$ and $\{H_k\}_{k \ge 0}$ denote the random walk and Harris walk associated with X^p. For the subsequent soliton lengths, we establish a duality with corresponding soliton lengths in an appropriate subcritical configuration.

Lemma 7.1. *Fix $\varepsilon > 0$, $j \in \mathbb{N}$, and $\theta = (1-p)/p < 1$. Then there exists a constant $c = c(p) > 1$ such that for each $n, x \geq 1$,*

$$\mathbb{P}\{|\lambda_1(n) - S_n| \geq x\} \leq c\theta^{x/2}$$

and

$$\mathbb{P}\{|\lambda_{j+1}(n) - \widehat{\lambda}_j(n)| \geq x\} \leq c\theta^{x/4}.$$

It is straightforward to deduce Theorem 2 from the above lemma.

Proof of Theorem 2 (iii). First, we may write

$$\frac{\lambda_1(n) - (2p-1)n}{2\sqrt{p(1-p)n}} = \frac{\lambda_1(n) - S_n}{2\sqrt{p(1-p)n}} + \frac{S_n - (2p-1)n}{2\sqrt{p(1-p)n}}.$$

The first term on the right-hand side converges in probability to zero by Lemma 7.1, and the second term converges in distribution to a standard normal by the usual central limit theorem, so the first part of the assertion follows from Slutsky's theorem.

The concentration inequality for $\lambda_1(n)$ is a consequence of Lemma 7.1 and Hoeffding's inequality applied to the associated random walk S_n:

$$\begin{aligned}
\mathbb{P}\{|\lambda_1(n) - (2p-1)n| \geq x\} &= \mathbb{P}\{|\lambda_1(n) - (2p-1)n| \geq x, |\lambda_1(n) - S_n| \geq x/2\} \\
&\quad + \mathbb{P}\{|\lambda_1(n) - (2p-1)n| \geq x, |\lambda_1(n) - S_n| < x/2\} \\
&\leq \mathbb{P}\{|\lambda_1(n) - S_n| \geq x/2\} + \mathbb{P}\{|S_n - (2p-1)n| \geq x/2\} \\
&\leq c\theta^{x/4} + 2e^{-\frac{x^2}{8n}} \leq Ce^{-\frac{x^2}{8n}}
\end{aligned}$$

for a suitable constant C.

Now let $\hat{\mu}_n = \log_{\theta^{-1}}\left(\frac{(1-2p)^2}{p}n\right)$. (This is the μ_n term from Section 5 but with p and $1-p$ switched since we are now working in the supercritical regime.) Then for $j \geq 1$ fixed, Lemma 7.1 implies

$$\begin{aligned}
\mathbb{P}\{\lambda_{j+1}(n) \leq x + \hat{\mu}_n\} &= \mathbb{P}\{\lambda_{j+1}(n) \leq x + \hat{\mu}_n, |\lambda_{j+1}(n) - \hat{\lambda}_j(n)| \leq x/2\} \\
&\quad + \mathbb{P}\{\lambda_{j+1}(n) \leq x + \hat{\mu}_n, |\lambda_{j+1}(n) - \hat{\lambda}_j(n)| > x/2\} \\
&\leq \mathbb{P}\{\hat{\lambda}_j(n) \leq 3x/2 + \hat{\mu}_n\} + c\theta^{x/8}.
\end{aligned}$$

The lower bound is established similarly:

$$\begin{aligned}
\mathbb{P}\{\hat{\lambda}_j(n) \leq (x/2) + \hat{\mu}_n\} &= \mathbb{P}\{\lambda_{j+1}(n) \leq (x/2) + \hat{\mu}_n, |\lambda_{j+1}(n) - \hat{\lambda}_j(n)| \leq x/2\} \\
&\quad + \mathbb{P}\{\hat{\lambda}_j(n) \leq x + \hat{\mu}_n, |\lambda_{j+1}(n) - \hat{\lambda}_j(n)| > x/2\} \\
&\leq \mathbb{P}\{\lambda_{j+1}(n) \leq x + \hat{\mu}_n\} + c\theta^{x/8},
\end{aligned}$$

and the assertion then follows from Theorem 2 (i). $\qquad\square$

7.2. **Proof of Lemma 7.1.** We now prove Lemma 7.1, establishing a duality principle between the super- and sub-critical box ball systems. Positive drift ensures that S and H are not too different, so the first claim seems reasonable since S should attain its maximum over $[0, n]$ near n. To explain why the second claim is true, let $\widehat{H} \in C_0^+(\mathbb{R}^+)$ be the Harris walk for the dual configuration so that $\widehat{\lambda}_1(n) = \max \widehat{H}$. Now H and $\widehat{H}$ are coupled in such a way that the latter is a time-reversal of $\mathscr{E}_n(S)$, which is approximated by $\mathscr{E}_n(H)$. Thus it all boils down to showing that the path $\mathscr{E}_n(H)$ pivoted at n is close to $\mathscr{E}(H) = \mathscr{E}_{\mathtt{m}}(H)$, pivoted at the actual location $\mathtt{m} = \mathtt{m}(H)$ of the rightmost maximum of H. But

again positive drift ensures that H attains its maximum near the end. Continuity of the column length functionals can then be used to show that the two paths must be close to each other in an appropriate sense.

We begin by introducing the following random variable:

$$R = \sup_{k \in \mathbb{N}} \left| \min_{1 \leq i \leq k} S_i \right| = -\inf_{k \in \mathbb{N}} S_k. \tag{9}$$

Also, let $\widetilde{S}_k$ and $\widetilde{H}_k$ be the random walk and Harris walk associated with the flip $\widetilde{X}^{n,p}$. Observe that $\widetilde{X}^{n,p}$ has the same law as $X^{n,1-p}$, and for each $1 \leq k \leq n$, we have $\widetilde{S}_k = (-\xi_1) + \cdots + (-\xi_k) = -S_k$ and

$$\widetilde{H}_k = \widetilde{S}_k - \min_{1 \leq i \leq k} \widetilde{S}_i = \max_{1 \leq i \leq k} S_i - S_k. \tag{10}$$

In the following proposition, we show that the maximum of the Harris walk $\{H_k\}_{0 \leq k \leq n}$ on the interval $[0, n]$ is exponentially concentrated around its last value H_n.

Proposition 7.2. *Fix* $1/2 < p < 1$ *and let* $\widehat{\theta} = p/(1 - p)$. *Then for any* $n, x \geq 1$,

$$\mathbb{P}\Big\{ \max_{0 \leq k \leq n} H_k - H_n \geq x \Big\} = \mathbb{P}\big\{ \widetilde{H}_n \geq x \big\} \leq \frac{2p - 1}{\widehat{\theta}^{\lfloor x \rfloor} - 1}.$$

Proof. To show the first inequality, let $\mathtt{a} = \mathtt{a}(X^{n,p})$ be the location of the leftmost global minimum of the random walk $\{S_k\}_{0 \leq k \leq n}$. Then for any $k \geq \mathtt{a}$,

$$H_k = S_k - \min_{0 \leq j \leq k} S_j = S_k - S_{\mathtt{a}}.$$

It follows that

$$\max_{0 \leq k \leq n} H_k - H_n = \max_{0 \leq k \leq n} (S_k - S_{\mathtt{a}}) - (S_n - S_{\mathtt{a}}) = \max_{0 \leq k \leq n} S_k - S_n = \widetilde{H}_n.$$

Now $\widetilde{H}_k$ gives the height of the subcritical Harris walk which moves up with probability $1 - p$, so writing $\mathtt{b}$ for the beginning of the excursion interval containing n, Equation (3) shows that

$$\mathbb{P}\Big\{ \max_{0 \leq k \leq n} H_k - H_n \geq x \Big\} = \mathbb{P}\big\{ \widetilde{H}_n \geq x \big\} \leq \mathbb{P}\Big\{ \max_{\mathtt{b} \leq k \leq n} \widetilde{H}_k > x - 1 \Big\}$$

$$\leq 1 - \left(1 - \frac{1 - 2(1 - p)}{\widehat{\theta}^{\lfloor x - 1 \rfloor + 1} - 1} \right) \leq \frac{2p - 1}{\widehat{\theta}^{\lfloor x \rfloor} - 1}. \qquad \square$$

Proposition 7.3. *Fix* $1/2 < p < 1$ *and* $\widehat{\theta} = p/(1 - p) > 1$. *Let* R *and* $\widetilde{H}_n$ *be as defined at* (9) *and* (10). *Then there exists a constant* $c = c(p) > 0$ *such that for all* $n, x \geq 1$,

$$\mathbb{P}\big\{ R + 2\widetilde{H}_n \geq x \big\} \leq c\theta^{x/2}.$$

Proof. Casing out according to the value of $\xi_1 = \mathbf{1}\big\{ X^p(1) = 1 \big\} - \mathbf{1}\big\{ X^p(1) = 0 \big\}$ shows that for any integer $k \geq 1$, $\mathbb{P}\{R \leq k\} = p\mathbb{P}\{R \leq k + 1\} + (1 - p)\mathbb{P}\{R \leq k - 1\}$, hence

$$\mathbb{P}\{R = k\} = \mathbb{P}\{R \leq k\} - \mathbb{P}\{R \leq k - 1\} = p\big(\mathbb{P}\{R \leq k + 1\} - \mathbb{P}\{R \leq k - 1\}\big) = p\big(\mathbb{P}\{R = k\} + \mathbb{P}\{R = k + 1\}\big),$$

so $\mathbb{P}\{R = k + 1\} = \widehat{\theta}^{-1}\mathbb{P}\{R = k\}$. It follows that

$$\mathbb{P}\{R \geq x\} = \sum_{k=x}^{\infty} \mathbb{P}\{R = 1\}\widehat{\theta}^{1-k} = \frac{\widehat{\theta}^{1-x}}{1 - \widehat{\theta}^{-1}} \mathbb{P}\{R = 1\}$$

for each $x \in \mathbb{N}$. Thus Proposition 7.2 implies that there is a $c = c(p) > 0$ such that

$$\mathbb{P}\big\{ R + 2\widetilde{H}_n \geq x \big\} \leq \mathbb{P}\{R \geq x/2\} + \mathbb{P}\{H \geq x/2\} \leq \frac{\widehat{\theta}^{1-\lfloor x/2 \rfloor}}{1 - \widehat{\theta}^{-1}} + \frac{2p - 1}{\widehat{\theta}^{\lfloor x/2 \rfloor} - 1} \leq c\widehat{\theta}^{-x/2}$$

for all $x \geq 2$. $\square$

We are now ready to prove Lemma 7.1.

Proof of Lemma *7.1.* Fix $n \geq j$ and let R and $\widetilde{H}_n$ be as defined at (9) and (10), respectively. According to the exponential bound in Proposition 7.3, it suffices to show the following inequalities:

$$\left| \lambda_1(n) - S_n \right| \leq R + 2\widetilde{H}_n \quad \text{and} \quad \left| \lambda_{j+1}(n) - \widehat{\lambda}_j(n) \right| \leq 2R + 4\widetilde{H}_n. \tag{11}$$

Note that the first inequality in (11) follows from Lemma 2.2 and the triangle inequality upon observing that

$$\left| \max_{1 \leq k \leq n} H_k - \max_{1 \leq k \leq n} S_n \right| \leq \max_{1 \leq k \leq n} \left| H_k - S_k \right| = \max_{1 \leq k \leq n} \left| \min_{1 \leq i \leq k} S_i \right| \leq \sup_{k \in \mathbb{N}} \left| \min_{1 \leq i \leq k} S_i \right| = R.$$

To establish the second inequality, let $n^* := \mathtt{m}(S\mathbf{1}_{[0,n]})$ denote the rightmost maximum of S on $[0, n]$, and define the sequence of random variables $\{\check{S}_k\}_{0 \leq k \leq n}$ by $\check{S}_k = S_k$ for all $k \neq n$ and $\check{S}_n = S_{n^*}$. As usual, let $\check{S}$ denote the linear interpolation of $\{\check{S}_k\}$. By construction, $\|\check{S} - S\|_\infty = \widetilde{H}_n$. Also, observe that $\mathscr{E}_n(S)(n) = 0 = \mathscr{E}(\check{S})(n)$, and for $0 \leq j < n$, writing $m_j = \min(S_j, \ldots, S_{n-1})$, we have $\mathscr{E}_n(S)(j) = S_j - \min(m_j, S_n)$ and $\mathscr{E}(\check{S})(j) = S_j - \min(m_j, S_{n^*}) = S_j - m_j$. If $\min(m_j, S_n) = S_n$, then $m_j = S_n + \mathbf{1}\{S_n < m_j\}$. It follows that

$$\mathscr{E}_n(S)(j) = \mathscr{E}(\check{S})(j) + m_j - \min(m_j, S_n) = \mathscr{E}(\check{S})(j) + \mathbf{1}\{S_n < m_j\}.$$

Writing $\widehat{S}_k = -(S_n - S_{n-k})$ for the random walk associated with the dual configuration, we see that the Harris walk $\widehat{H}_k$ can be written as

$$\widehat{H}_k = (S_{n-k} - S_n) - \min_{0 \leq j \leq k}(S_{n-j} - S_n) = S_{n-k} - \min_{n-k \leq i \leq n} S_i = \mathscr{E}(\check{S})(n - k) + \mathbf{1}\{S_n < m_{n-k}\}$$

for all $0 \leq k \leq n$. As $S_n < m_{n-k}$ implies $\widetilde{H}_n = \|\check{S} - S\|_\infty \geq 1$, we have

$$\left| \widehat{H}_k - \mathscr{E}(\check{S})(n - k) \right| \leq \widetilde{H}_n.$$

for all $k \geq 1$. Since the functional $\max \mathscr{E}^{j-1}$ is invariant under time reversal, the above observation together with Lemmas 2.2 and 2.3 yields

$$\left| \widehat{\lambda}_j(n) - \max \mathscr{E}^j(\check{S}) \right| = \left| \max \mathscr{E}^{j-1}(\widehat{H}) - \max \mathscr{E}^j(\check{S}) \right| \leq 2\widetilde{H}_n.$$

Finally, the triangle inequality, Lemma 2.2, and Lemma 2.3 give

$$\left| \lambda_{j+1}(n) - \widehat{\lambda}_j(n) \right| \leq \left| \max \mathscr{E}^j(H) - \max \mathscr{E}^j(S) \right| + \left| \max \mathscr{E}^j(S) - \max \mathscr{E}^j(\check{S}) \right| + 2\widetilde{H}_n$$

$$\leq 2\|H - S\|_\infty + 2\|S - \check{S}\|_\infty + 2\widetilde{H}_n \leq 2R + 4\widetilde{H}_n. \qquad \square$$

8. RANDOM 312-AVOIDING PERMUTATIONS

In this section, we discuss some relations between box-ball systems and 312-avoiding permutations and prove Theorem 3.

Recall that for a given permutation σ, one can use the Robinson-Schensted algorithm (see [24, Ch. 3.1]) to obtain a pair of standard Young tableaux with common shape $\mathtt{RS}(\sigma)$. Greene's theorem [10] relates the sum of the lengths of the first k rows (resp. columns) of the Young diagram $\mathtt{RS}(\sigma)$ to the length of a longest subsequence in σ which can be obtained by taking the union of k increasing (resp. decreasing) subsequences. In Proposition 8.1, we show that if σ is 312-avoiding, then a 'naive' version of Greene's theorem holds: We can subsequently delete longest increasing/decreasing subsequences to obtain subsequent row/column lengths of $\mathtt{RS}(\sigma)$. Hence, roughly speaking, Theorem 3 gives the asymptotics of the 'k^{th} longest' increasing/decreasing subsequences of a random 312- (or 231-) avoiding permutation.

For a precise statement, we introduce some notation. Given two finite sequences α, β of positive integers, denote by $\alpha \setminus \beta$ the sequence obtained by deleting all elements in β from α. Denote by α_+ (resp. α_-) an arbitrary longest increasing (resp. decreasing) subsequence of α. Furthermore, let α_+^* (resp. α_-^*) be the unique longest increasing (resp. decreasing) subsequence in α such that the sum of all numbers used in $(\alpha_+^*)^{-1}$ (resp. $(\alpha_-^*)^{-1}$) is as small (resp. large) as possible. This ensures that σ_+^* (resp. σ_-^*) is the 'leftmost' (resp. 'rightmost') longest increasing (resp. decreasing) subsequence in σ. For instance, if $\sigma = 146532$, then both 146 and 145 are longest increasing subsequences, where the former is σ_+^*. The following proposition is proved in Appendix A.4.

Proposition 8.1. *Let τ be a 312-avoiding permutation and fix arbitrary τ_-. Then $\mathrm{RS}(\tau \setminus \tau_-)$ is obtained from $\mathrm{RS}(\tau)$ by deleting its first column. Moreover, $\mathrm{RS}(\tau \setminus \tau_+^*)$ is obtained from $\mathrm{RS}(\tau)$ by deleting its first row.*

In order to prove Theorem 3, we begin by explaining (an equivalent version of) the construction of the time-invariant Young diagram introduced in [26], which was built upon a connection between box-ball configurations and 312-avoiding permutations. The first step is to map a box-ball configuration X_0 of m balls to a 312-avoiding permutation $\sigma = \sigma(X_0) \in \mathfrak{S}_m^{312}$ using the pushing and popping stack operations from [13, Ch. 2.2.1]. To do so, label the balls $1, \ldots, m$ from left to right so that the ith ball gets label i. Then the one-line notation for σ gives the left to right labels of the balls after a single update $X_0 \mapsto X_1$. That is, we push the symbol 1 onto an empty stack at the first ball and then, advancing to the right, pop the top of the stack off for storage at each empty box and push k onto the stack upon encountering the kth ball. See Figure 10 for an illustration.
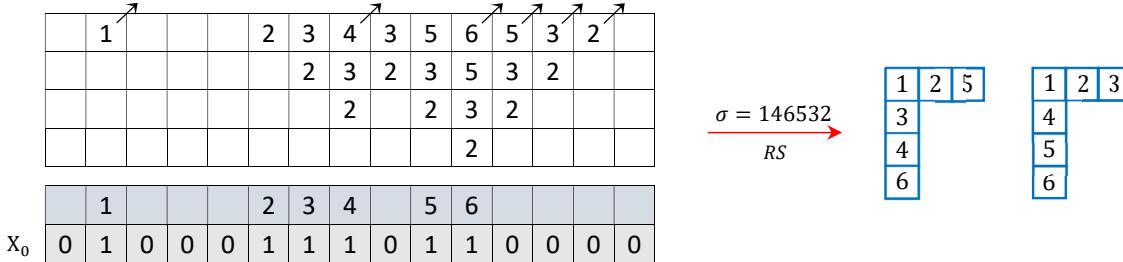

FIGURE 10. Construction of the 312-avoiding permutation corresponding to the box-ball environment in the bottom row via push-pop operations. The second row from the bottom indicates the labels of the balls, and the columns in the upper table give the contents of the right-sweeping stack. The resulting permutation is $\sigma = 146532$.

To get a Young diagram from this stack-representable permutation $\sigma(X_0)$, one applies the Robinson-Schensted algorithm to obtain a pair of standard Young tableaux, and records their common shape as $\mathrm{RS}(\sigma(X_0))$. It was shown in [26] that $\mathrm{RS}(\sigma(X_s))$ is invariant in $s \geq 0$ and its jth column length is the jth longest soliton length in the system. Thus, by Lemma 2.1, this construction gives the same Young diagram which was obtained by hill-flattening operations applied to the Motzkin path.

Proposition 8.2. *Let $X_0 : \mathbb{N}_0 \to \{0, 1\}$ be a finitely supported box-ball configuration. Then*

$$\mathrm{RS}(\sigma(X_0)) = \Lambda(X_0) = \Lambda(\Gamma(X_0)).$$

The following proposition (proved in Appendix A.4) shows that there is a bijection between 312-avoiding permutations of length n and Dyck paths of length $2n$ which 'factors through' box-ball configurations in a natural way. Let $\mathfrak{S}_n^{312}$ be the set of all 312-avoiding permutations of length n and let

Dyck_{2n} be the set of all Dyck paths of length $2n$—that is, lattice paths from $(0,0)$ to $(2n,0)$ consisting only of upstrokes and downstrokes and never dipping below the horizontal axis.

Proposition 8.3.

(i) *There exists a bijection $\varphi : \mathrm{Dyck}_{2n} \to \mathfrak{S}_n^{312}$.*

(ii) *For each $\tau \in \mathfrak{S}_n^{312}$ and $F \in \mathrm{Dyck}_{2n}$ such that $\varphi(F) = \tau$, there is a box-ball configuration X_0 such that $\tau = \sigma(X_0)$ and $F = \Gamma(X_0)$.*

We now prove Theorem 3 using similar ideas from the proof of Theorem 1 together with some known results on random Dyck paths and random walk excursions.

Proof of Theorem 3. Recall that $\sigma \in \mathfrak{S}_n^{312}$ if and only if $\sigma^{-1} \in \mathfrak{S}_n^{231}$, so the map $\sigma \mapsto \sigma^{-1}$ preserves the uniform distribution on the sets $\mathfrak{S}_n^{312}$ and $\mathfrak{S}_n^{231}$. Moreover, $\mathrm{RS}(\sigma) = \mathrm{RS}(\sigma^{-1})$. Hence it suffices to prove the assertion only for the 312-avoiding permutations.

Let F be a Dyck path of length $2n$ and let $\tau = \varphi(F)$ be the corresponding 312-avoiding permutation. Proposition 8.3 enables us to choose a box-ball configuration X_0 such that $\tau = \sigma(X_0)$ and $F = \Gamma(X_0)$, and Proposition 8.2 implies that $\mathrm{RS}(\tau) = \Lambda(F)$. If we denote by Σ^n and F^n uniformly random elements of $\mathfrak{S}_n^{312}$ and Dyck_{2n}, this yields

$$\mathrm{RS}(\Sigma^n) =_d \Lambda(F^n). \tag{12}$$

Now the contour process described in Subsection 2.3 gives a bijection between Dyck paths of length $2n$ and rooted plane trees with $n+1$ nodes, so part (i) of Theorem 3 follows from (12) and Proposition 2.5.

Part (iii) also follows easily from known results. Indeed, it is well known that under diffusive scaling the random walk excursion converges weakly to a standard Brownian excursion [1]. Moreover, by Theorem 6.1, the convergence is polynomial. Thus (iii) follows from (12) and Lemmas 2.2 and 2.3.

Lastly, we establish the strong law for $\rho_i(F^n)$ stated in part (ii). To begin, fix $i \geq 1$, and let $\{S_k\}_{k \geq 0}$ be a simple symmetric random walk with $S_0 = 0$. We may view the uniformly random Dyck path F^n of length $2n$ as the trajectory of S_k over the interval $[0,2n]$ conditioned to stay non-negative and satisfy $S_{2n} = 0$. By (12) and the hill-flattening procedure, $\rho_i(F^n)$ equals the number of subexcursions of F^n of height i. Recall the definitions of μ_i and J_ℓ^i given in Lemma 4.2 and above the same lemma, respectively. Let $N_i(n) = \sum_{\ell=0}^n J_\ell^i$. Then $N_i(2n) = \rho_i(F^n)$, so for all $n \geq 1$ and $\varepsilon < 1/2n$,

$$
\mathbb{P}\left\{\left|\frac{\rho_i(F^n)}{2n} - \mu_i\right| > \varepsilon\right\} \leq \mathbb{P}\left\{\left|\frac{N_i(2n)}{2n} - \mu_i\right| > \varepsilon \,\middle|\, S_k \text{ is a Dyck path over } [0,2n]\right\}
$$

$$
\leq \frac{\mathbb{P}\{|N_i(2n)/2n - \mu_i| > \varepsilon\}}{\mathbb{P}\{S_k \text{ is a Dyck path over } [0,2n]\}}.
$$

It is well known that the number of Dyck paths of length $2n$ is the nth Catalan number $\frac{1}{n+1}\binom{2n}{n}$, so by Stirling's approximation, $\mathbb{P}\{S_k \text{ is a Dyck path over } [0,2n]\} \sim n^{-3/2}/\sqrt{\pi}$. Now by Lemma 4.2 with $m = 4$ and $\varepsilon = \varepsilon(n) = 1/\log n \searrow 0$, we get

$$
\mathbb{P}\left\{\left|\frac{\rho_i(F^n)}{2n} - \mu_i\right| > \varepsilon(n)\right\} = O((\log n)^8 n^{-3/2}).
$$

In particular, these probabilities are summable, so the first Borel-Cantelli lemma implies $\rho_i(F^n)/2n \to \mu_i$ a.s. as $n \to \infty$. $\qquad \square$

APPENDIX A. PROOFS OF COMBINATORIAL LEMMAS

In this appendix, we provide proofs of Lemmas 2.1, 2.2, and 2.3, and Propositions 8.1 and 8.3.

A.1. **Time invariance of the Young diagram.** Our proof of Lemma 2.1 is similar to the argument from [26], which is formulated in terms of Dyck words intead of Motzkin paths. The argument is simplified by Proposition A.1.

To begin, recall that given a box-ball configuration X_s of finite support, the associated lattice path $\Gamma(X_s)$ is constructed by reading X_s from left to right: Starting at height 0, increase by 1 every time a 1 is encountered, decrease by 1 whenever a 0 is encountered at positive height, and remain at height 0 otherwise. A simple but useful observation is that reading X_s from right to left produces the lattice path $\Gamma(X_{s-1})$. More precisely, let $(X_s)_{s\geq 0}$ be a box-ball system started from a finitely supported configuration X_0. For each $s \geq 0$, let $r_s = \max\{k \geq 0 : X_s(k) = 1\}$ be the location of the rightmost 1 at time s. Construct a (backward) lattice path $\overleftarrow{\Gamma}(X_s) : \mathbb{N}_0 \to \mathbb{N}_0$ by $\overleftarrow{\Gamma}(X_s)_k = 0$ for $k \geq r_s$ and

$$\overleftarrow{\Gamma}(X_s)_k = \begin{cases} \overleftarrow{\Gamma}(X_s)_{k+1} + 1 & \text{if } X_s(k+1) = 1 \\ \overleftarrow{\Gamma}(X_s)_{k+1} - 1 & \text{if } X_s(k+1) = 0 \text{ and } \overleftarrow{\Gamma}(X_s)_{k+1} \geq 1 \\ 0 & \text{if } \overleftarrow{\Gamma}(X_s)_{k+1} = X_s(k+1) = 0 \end{cases}$$

for $0 \leq k < r_s$. See Figure A.1 for an illustration. In this appendix, we denote the ordinary lattice path Γ by $\overrightarrow{\Gamma}$ to emphasize the reading direction.

| 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 0 |
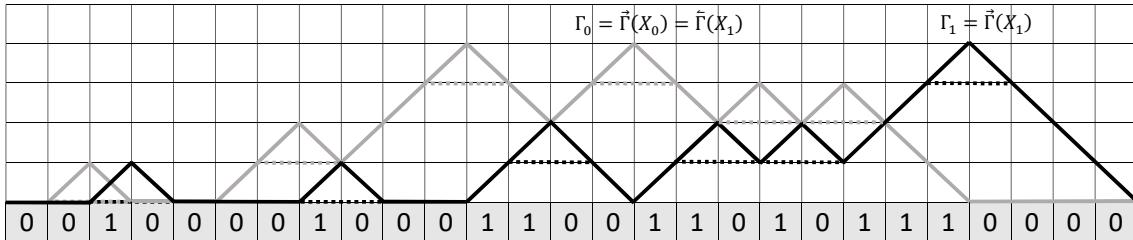
FIGURE A.1. The environment is X_1 where X_0 is the environment given in Figure 3. The black path is $\overrightarrow{\Gamma}(X_1)$ and the grey path is $\overleftarrow{\Gamma}(X_1)$. Notice that the latter coincides with the black path in Figure 3.

Proposition A.1. $\overleftarrow{\Gamma}(X_{s+1}) = \overrightarrow{\Gamma}(X_s)$ for all $s \geq 0$.

Proof. Fix $s \geq 0$, and observe that both paths are 0 on $[r_{s+1}, \infty)$, so the assertion holds on this interval. Now suppose the paths agree on $[k+1, \infty)$ for some $k < r_{s+1}$. We must show that $\overleftarrow{\Gamma}(X_{s+1})_k = \overrightarrow{\Gamma}(X_s)_k$.

The definition of the box-ball dynamics shows that $X_{s+1}(k+1) = 1$ if and only if $\overrightarrow{\Gamma}(X_s)_k - 1 = \overrightarrow{\Gamma}(X_s)_{k+1}$, hence

$$\overrightarrow{\Gamma}(X_s)_k - \overrightarrow{\Gamma}(X_s)_{k+1} = 1 \iff X_{s+1}(k+1) = 1 \iff \overleftarrow{\Gamma}(X_{s+1})_k - \overleftarrow{\Gamma}(X_{s+1})_{k+1} = 1.$$

The inductive hypothesis implies

$$\overrightarrow{\Gamma}(X_s)_k = \overrightarrow{\Gamma}(X_s)_{k+1} \iff X_{s+1}(k+1) = 0 \text{ and } \overrightarrow{\Gamma}(X_s)_{k+1} = 0$$

$$\iff X_{s+1}(k+1) = 0 \text{ and } \overleftarrow{\Gamma}(X_{s+1})_{k+1} = 0$$

$$\iff \overleftarrow{\Gamma}(X_{s+1})_k = \overleftarrow{\Gamma}(X_{s+1})_{k+1}$$

and

$$\vec{\Gamma}(X_s)_k - \vec{\Gamma}(X_s)_{k+1} = -1 \iff X_{s+1}(k+1) = 0 \text{ and } \vec{\Gamma}(X_s)_{k+1} \geq 1$$
$$\iff X_{s+1}(k+1) = 0 \text{ and } \bar{\Gamma}(X_{s+1})_{k+1} \geq 1$$
$$\iff \bar{\Gamma}(X_{s+1})_k - \bar{\Gamma}(X_{s+1})_{k+1} = -1. \qquad \square$$

To facilitate the proof of Lemma 2.1, it is convenient to reformulate the procedure for building Young diagrams row by row: Rather than flatten hills, we *contract* peaks by deleting the upstroke-downstroke pair and then identifying the endpoints so that the path remains connected. The number of hills after flattening is the same as the number of peaks after contracting, so everything is exactly same as before. The advantage here is that if one begins with an h-restricted Motzkin path, then the hills are always peaks and the Motzkin paths are always h-restricted. Moreover, the contraction operation can be understood in terms of the environment as deleting 10 patterns.

Proof of Lemma 2.1. The second part of the assertion clearly holds for all stable box-ball configurations $X_0 : \mathbb{N} \to \{0, 1\}$ of finite support. Since the system always stabilizes, the second part follows from the time invariance as stated in the first part.

Now let $(X_s)_{s \geq 0}$ be as before. To show the time invariance of $\Lambda(X_s)$, recall that the construction of $\Lambda(X_s)$ begins by counting the number of peaks in the path corresponding $X_s = X_s^{(0)}$. This is equal to the number of 10 patterns, which is equal to the number of 1-strings, which is equal to the number of 01 patterns. The length of the first row of $\Lambda(X_s)$ is given by this number. The peaks are then contracted by deleting the 10 patterns from X_s to obtain $X_s^{(1)}$ and the process is repeated with $\Gamma(X_s^{(1)})$. At each step, the 1-strings are counted, the diagram is updated, and the 10 patterns are deleted, continuing until the path consists only of h-strokes.

The key insights are that the number of 1 strings is the same regardless of whether the environment is read from left to right or conversely, and that the number of 1-strings after 10 patterns are deleted is the same as the number of 1 strings after 01 patterns are deleted. In the first case, each 1 string either decreases in length by 1 (possibly disappearing), or it merges with the string on its right. In the second, each string either decreases in length by 1 or merges with the string on its left.
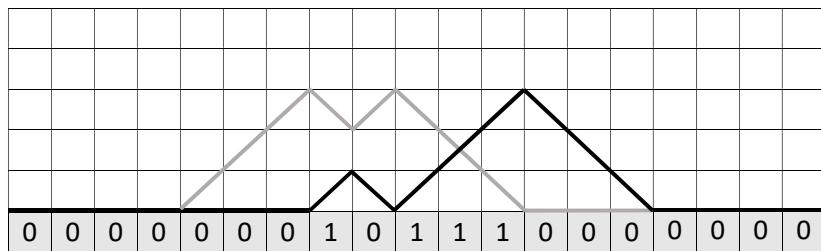

FIGURE A.2. The environment is formed by deleting either 10 patterns or 01 patterns from the environment in Figure A.1. The corresponding left-right (black) and right-left (grey) lattice paths have the same number of hills as the flattened paths in Figure A.1.

Now for any fixed $s \geq 0$, $\vec{\Gamma}(X_s)$ and $\vec{\Gamma}(X_{s+1})$ can be read off from X_{s+1} by proceeding from right to left and from left to right, respectively. The update rule for the former is to count 1-strings and then delete 10 patterns, and the update rule for the latter is to count 1-strings and then delete 01 patterns. By the previous observations, both result in the same final Young diagram.

At this point, it remains only to show that soliton lengths are given by the column lengths of the Young diagram $\Lambda(X_0)$. To see that this is so, observe that the path $\Gamma(X_\tau)$, which corresponds to the first stable configuration, consists of a series of single peaks of nondecreasing height, each as tall as the length of the associated soliton. As each flattening step reduces the height of the peaks by 1, we see that the number of rows of $\Lambda(X_\tau)$ having length at least ℓ corresponds to the number of solitons of length at least ℓ. Therefore, the columns of $\Lambda(X_\tau)$ encode the soliton lengths, so the same is true of $\Lambda(X_0)$ by invariance. $\qquad\square$

A.2. **Extracting column lengths with excursion operators.** In this subsection, we prove Lemma 2.2. The key observation is that the hill-flattening and excursion operators commute on the space of Motzkin paths.

To begin, we need to establish a couple of technical results. First, for any interval $I \subseteq \mathbb{R}$, recall that $C_0^+(I)$ denotes the space of continuous functions $I \to [0,\infty)$ with compact support. For any $f \in C_0^+(I)$, we denote by $\mathrm{supp}^+(f)$ the open set $\{x \in I : f(x) > 0\}$, which is a finite disjoint union of open intervals. Accordingly, we may write $\mathrm{supp}^+(f) = \bigsqcup_{i=1}^n (c_i, d_i)$, where $d_i < c_j$ if $i < j$. We call $J_i := (c_i, d_i)$ the ith *excursion interval* of f. Recall that $\mathscr{I}(\Gamma)$ denotes the set of hill intervals of Γ (see the beginning of Subsection 2.2).

Proposition A.2. *Fix a Motzkin path Γ and let $x \in \mathbb{N}$ be contained in a hill interval I_x of Γ. Denote* $\mathrm{supp}^+(\mathscr{E}_x(\Gamma)) = \bigsqcup_{i=1}^n J_i$ *as above. Then $\Gamma - \mathscr{E}_x(\Gamma)$ is constant on each J_i. In addition, $\mathscr{I}(\mathscr{E}_x(\Gamma)) = \mathscr{I}(\Gamma) \setminus \{I_x\}$ and $\max \mathscr{E}^{j-1}(\Gamma) \geq 1$ for all $1 \leq j \leq \rho(\Gamma)$.*

Proof. To establish the first part, write $M = \Gamma_x \geq 0$, and define integers $a_0 < a_1 < \cdots < a_{M-1} < a_M = x = b_M < b_{M-1} < \cdots < b_1 < b_0$ by

$$a_i = \max\{k \leq x : \Gamma_k = i\}, \quad b_i = \min\{k \geq x : \Gamma_k = i\}$$

for each $0 \leq i \leq M$. In words, they are the first locations where Γ has height i when moving to the left and right from x; see Figure 4. To simplify notation, we set $a_{-1} = 0$ and $b_{-1} = \infty$. Now $\Gamma_y - \mathscr{E}_x(\Gamma)_y = \min\{\Gamma_z : x \wedge y \leq z \leq x \vee y\}$, so on $\mathbb{N}_0$

$$\Gamma - \mathscr{E}_x(\Gamma) = \sum_{i=0}^{M-1} i\left(\mathbf{1}_{(a_{i-1}, a_i]} + \mathbf{1}_{[b_i, b_{i-1})}\right) + M\mathbf{1}_{(a_{M-1}, b_{M-1})}.$$

It follows that $\mathscr{E}_x(\Gamma)$ vanishes at the a_i's and b_i's, and differs from Γ by a constant on (a_{M-1}, b_{M-1}) and each interval of the form $(a_{k-1}, a_k]$ or $[b_k, b_{k-1})$, $0 \leq k \leq M-1$. J_i is the ith such interval (from left to right) where $\mathscr{E}_x(\Gamma)$ is not constant. This shows the first part of the assertion.

The preceding argument also implies that $\mathscr{I}(\mathscr{E}_x(\Gamma)) \subseteq \mathscr{I}(\Gamma)$. In addition, $\mathscr{E}_x(\Gamma) = 0$ on $[a_{M-1}, b_{M-1}]$ and $I_x \subseteq (a_{M-1}, b_{M-1})$, so I_x is not a hill interval of $\mathscr{E}_x(\Gamma)$. Finally, the definition of the a and b terms ensures that if $J \in \mathscr{I}(\Gamma) \setminus \{I_x\}$, then either $J \subseteq (a_{i-1}, a_i]$ or $J \subseteq [b_i, b_{i-1})$ for some $0 \leq i \leq M-1$. Since $\mathscr{E}_x(\Gamma)$ is a vertical translate of Γ on these intervals, J must be a hill interval of $\mathscr{E}_x(\Gamma)$. This shows $\mathscr{I}(\mathscr{E}_x(\Gamma)) = \mathscr{I}(\Gamma) \setminus \{I_x\}$.

Lastly, taking $x = \mathfrak{m}$ in the first part gives $\mathscr{I}(\mathscr{E}(\Gamma)) = \mathscr{I}(\Gamma) \setminus \{I_\mathfrak{m}\}$, and the second part of the second assertion follows from the first since each application of $\mathscr{E}$ removes a single hill interval and the height of a Motzkin path is at least one while hill intervals remain. $\qquad\square$

Proposition A.3. *For any interval $I \subseteq \mathbb{R}^+$, $f \in C_0^+(I)$, $x, y \in I$, if f is constant on the interval $[x, y] \subseteq I$, then $\mathscr{E}_x(f) = \mathscr{E}_y(f)$.*

Proof. Casing out according to whether $t < x$, $x \leq t \leq y$, or $t > y$ shows that

$$\min_{t \wedge x \leq s \leq t \vee x} f(s) = \min_{t \wedge y \leq s \leq t \vee y} f(s). \qquad\square$$

Proposition A.4. *For any Motzkin path Γ and any $x \in \mathbb{N}$ contained in a hill interval of Γ, $\mathscr{E}_x \circ \mathscr{H}(\Gamma) = \mathscr{H} \circ \mathscr{E}_x(\Gamma)$. In particular, $\mathscr{E} \circ \mathscr{H}(\Gamma) = \mathscr{H} \circ \mathscr{E}(\Gamma)$.*

Proof. Let $\mathtt{m} = \mathtt{m}(\Gamma)$ and $\mathtt{m}^* = \mathtt{m}(\mathscr{H}(\Gamma))$. Note that $\mathtt{m} < \mathtt{m}^*$ and that $\mathscr{H}(\Gamma)$ is constant on $[\mathtt{m}, \mathtt{m}^*]$. This holds for any Motzkin path Γ. Thus by Proposition A.3 with $I = \mathbb{R}^+$, it suffices to prove the first part. To this end, we first note that for any $k \in \mathbb{N}_0$,

$$\min_{k \wedge x \leq y \leq k \vee x} \Gamma_y - \min_{k \wedge x \leq y \leq k \vee x} \mathscr{H}(\Gamma)_y = \mathbf{1}\{k \in I_x\},$$

where I_x denotes the hill interval of Γ containing x. Indeed, $\mathscr{H}(\Gamma) = \Gamma - 1$ on I_x, so the left-hand side is 1 for all $k \in I_x$. Now fix $k \notin I_x$, and let x_* be the location of the leftmost minimum of Γ over the interval $[k \wedge x, k \vee x]$. Then x_* is an integer which is not contained in any hill interval of Γ, so $\mathscr{H}(\Gamma)_{x_*} = \Gamma_{x_*}$. Moreover, x_* minimizes $\mathscr{H}(\Gamma)$ on $[k \wedge x, k \vee x]$ since the only integer points with $\mathscr{H}(\Gamma)_y < \Gamma_y$ are those contained in a hill interval of Γ, in which case $\Gamma_y \geq \Gamma_{x_*} + 1$. This shows that the left-hand side is 0 for $k \notin I_x$ as desired.

In conjunction with Proposition A.2, we have

$$
\begin{aligned}
\mathscr{E}_x(\mathscr{H}(\Gamma))_k &= \mathscr{H}(\Gamma)_k - \min_{k \wedge x \leq y \leq k \vee x} \mathscr{H}(\Gamma)_y \\
&= \mathscr{H}(\Gamma)_k - \min_{k \wedge x \leq y \leq k \vee x} \Gamma_y + \mathbf{1}\{k \in I_x\} \\
&= \begin{cases} \mathscr{E}_x(\Gamma)_k - 1 & \text{if } k \in \bigcup \mathscr{I}(\Gamma) \setminus \{I_x\} \\ \mathscr{E}_x(\Gamma)_k & \text{otherwise} \end{cases} \\
&= \mathscr{H}(\mathscr{E}_x(\Gamma))_k.
\end{aligned}
$$

$\square$

Now we prove Lemma 2.2.

Proof of Lemma 2.2. Let Γ be a Motzkin path and write λ_j for the length of the jth column of $\Lambda(\Gamma)$ for each $1 \leq j \leq \rho(\Gamma)$. We show

$$\lambda_j = \max \mathscr{E}^{j-1}(\Gamma)$$

by induction on $\max \Gamma$. If the maximum is zero, then the assertion is trivial, so we may assume that it holds for all Motzkin paths with maximum less than $M \in \mathbb{N}$. Now fix a path Γ with $\max \Gamma = M$. The inductive hypothesis implies that the assertion holds for $\mathscr{H}(\Gamma)$ since it has maximum $M - 1 \geq 0$. Moreover, $\Lambda(\mathscr{H}(\Gamma))$ is obtained by deleting the first row of $\Lambda(\Gamma)$. Thus by Proposition A.4, we have

$$
\begin{aligned}
\lambda_j - 1 &= \max \mathscr{E}^{j-1}(\mathscr{H}(\Gamma)) \\
&= \max \mathscr{H}(\mathscr{E}^{j-1}(\Gamma)) \\
&= \max \mathscr{E}^{j-1}(\Gamma) - 1,
\end{aligned}
$$

where the final equality used the second part of Proposition A.2 to ensure $\max \mathscr{E}^{j-1}(\Gamma) \geq 1$ for any $1 \leq j \leq \rho(\Gamma)$. $\square$

Remark A.5. An easy modification of Proposition A.4 and applying the same proof of Lemma 2.2 shows that the excursion operator $\mathscr{E} = \mathscr{E}_{\mathtt{m}}$ in the statement of Lemma 2.2 could be replaced by $\mathscr{E}_{\mathtt{m}^*}$, where the pivot $\mathtt{m}^* = \mathtt{m}^*(\Gamma)$ is chosen to be an arbitrary element in the set $\{x \geq 0 : \Gamma(x) = \max \Gamma\}$ where the Motzkin path Γ achieves its maximum.

A.3. **Regularity of the column length functionals.** In this subsection we prove Lemma 2.3, establishing Lipschitz continuity of the 'column length functionals' $\max \mathcal{E}^{j-1}(\cdot)$. The general strategy is to show that the column length functionals satisfy a Lipschitz condition on Motzkin paths and then extend the result to arbitrary functions in $C_0^+(\mathbb{R}^+)$ by an approximation argument. We begin by establishing some preparatory results.

Proposition A.6.

(i) *Fix an interval $I \subseteq \mathbb{R}^+$, a point $b \in I$, and functions $f, g \in C_0^+(I)$. Then*

$$\|\mathcal{E}_b(f) - \mathcal{E}_b(g)\|_\infty \leq 2\|f - g\|_\infty.$$

(ii) *For any Motzkin paths $f, g \in C_0^+(\mathbb{R}^+)$,*

$$\|\mathcal{H}(f) - \mathcal{H}(g)\|_\infty \leq \|f - g\|_\infty.$$

Proof. For (i), the triangle inequality gives

$$\|\mathcal{E}_b(f) - \mathcal{E}_b(g)\|_\infty \leq \|f - g\|_\infty + \sup_{t \in I} \left| \min_{[t \wedge b, t \vee b]} f - \min_{[t \wedge b, t \vee b]} g \right| \leq 2\|f - g\|_\infty$$

since the minima of two functions over a given interval can differ by no more than their maximum difference over the interval.

For (ii), observe that the maximum distance between Motzkin paths is necessarily $\mathbb{N}_0$-valued and the claim is clearly true if $f = g$, so we may assume that $\|\mathcal{H}(f) - \mathcal{H}(g)\|_\infty \geq 1$. Let

$$x^* = \max \left\{ x \in \mathbb{N} : |\mathcal{H}(f)_x - \mathcal{H}(g)_x| = \|\mathcal{H}(f) - \mathcal{H}(g)\|_\infty \right\},$$

and assume without loss of generality that $\mathcal{H}(f)_{x^*} > \mathcal{H}(g)_{x^*}$. If x^* is not in a hill interval of g, then $g(x^*) = \mathcal{H}(g)_{x^*} < \mathcal{H}(f)_{x^*} \leq f(x^*)$, so

$$\left\| \mathcal{H}(f) - \mathcal{H}(g) \right\|_\infty = \left| \mathcal{H}(f)_{x^*} - \mathcal{H}(g)_{x^*} \right| \leq \left| f(x^*) - g(x^*) \right| \leq \left\| f - g \right\|_\infty.$$

If x^* is in a hill interval of both f and g, then

$$\left\| \mathcal{H}(f) - \mathcal{H}(g) \right\|_\infty = \left| \mathcal{H}(f)_{x^*} - \mathcal{H}(g)_{x^*} \right| = \left| \left(f(x^*) - 1 \right) - \left(g(x^*) - 1 \right) \right| \leq \left\| f - g \right\|_\infty.$$

Finally, suppose that x^* is in a hill interval $[a, b]$ of g but is not in any hill interval of f. Then g is constant on $[a, b]$, so our choice of x^* implies that $f(x^*) \geq f(y)$ for all $y \in [a, b]$. By considering whether or not $x^* < b$, we see that we must have $f(x^* + 1) = f(x^*) - 1$. A similar consideration of whether $f(x^*) = f(y)$ for all $a \leq y \leq x^*$ leads to the contradiction that x^* is in a hill interval of f. $\qquad\square$

To state our next result, we say that a function $\varphi : \mathbb{R} \to \mathbb{R}$ is an *affine scaling* if $\varphi(x) = ax + b$ for some $a > 0$, $b \in \mathbb{R}$. The set of all affine scalings forms a group under composition. Given $f \in C_0^+(\mathbb{R})$ and an affine scaling φ, we write $\varphi^*(f)$ for the function $f \circ \varphi^{-1}$. A function $\Gamma : \mathbb{R} \to \mathbb{R}^+$ is an *extended Motzkin path* if $\Gamma(n) = 0$ for all $n \leq 0$ and $\Gamma|_{[0,\infty)}$ is a Motzkin path.

Proposition A.7. *For any $f_1, f_2 \in C_0^+(\mathbb{R})$ which are not identically zero and any $\varepsilon > 0$, there exist affine scalings φ, ψ and extended Motzkin paths Γ_1, Γ_2 such that $\psi(0) = 0$ and for $i = 1, 2$, the function $\bar{f}_i = \psi \circ \varphi^*(\Gamma_i) \in C_0^+(\mathbb{R})$ satisfies*

$$\|f_i - \bar{f}_i\|_\infty < \varepsilon \quad and \quad \mathtt{m}(\bar{f}_i) = \mathtt{m}(f_i).$$

Proof. By hypothesis, $\mathtt{m}(f_1), \mathtt{m}(f_2) \in (0, \infty)$. Also, the f_i's are uniformly continuous, so there is some $\delta > 0$ such that $|x - y| < \delta$ implies $|f_1(x) - f_1(y)|, |f_2(x) - f_2(y)| < \varepsilon/4$. Set $\mathtt{s} = |\mathtt{m}(f_1) - \mathtt{m}(f_2)| + \mathbf{1}\{\mathtt{m}(f_1) = \mathtt{m}(f_2)\}$ and choose N large enough that $\Delta := \mathtt{s}/2^N < \delta$. Define the lattice

$$\mathfrak{L} = \{\mathtt{m}(f_1) + k\Delta\}_{k \in \mathbb{Z}}.$$

Note that $\mathrm{m}(f_1), \mathrm{m}(f_2) \in \mathfrak{L}$. Set $a = 2\Delta/\varepsilon$, $\mathfrak{L}^+ = \mathfrak{L} \cap [0, \infty)$, and let ℓ_0 denote the smallest element of $\mathfrak{L}^+$. Observe that $0 \le \ell_0 < \Delta$ by construction.

For $i = 1, 2$, define the function $\gamma_i : \mathfrak{L} \to \mathfrak{L}^+$ by

$$\gamma_i(\ell) = \begin{cases} \ell_0 & \text{if } \ell \le \ell_0 \\ \Delta \lceil (a f_i(\ell))/\Delta \rceil + \ell_0 & \text{otherwise.} \end{cases}$$

Note that $a f_i$ changes by no more than $\Delta/2$ when the argument changes by no more than Δ. In conjunction with the fact that $f_i \equiv 0$ on $(-\infty, 0]$, $f_i \ge 0$, and $\ell_0 \in [0, \Delta)$, this implies that γ_i is an extended Motzkin path on $\mathfrak{L}$. That is, $\gamma_i(\ell) = \ell_0$ for all $\ell \in \mathfrak{L} \cap (-\infty, \ell_0]$ and for each $\ell, \ell' \in \mathfrak{L}$ with $|\ell - \ell'| = \Delta$, we have $\gamma_i(\ell) \ge \ell_0$ and $|\gamma_i(\ell) - \gamma_i(\ell')| \in \{0, \Delta\}$.

Let $\varphi(x) = \Delta \cdot x + \ell_0$. Then φ is an affine scaling which maps $\mathbb{Z}$ bijectively to $\mathfrak{L}$. Also define the affine scaling $\sigma(x) = (x - \ell_0)/a$. By a slight abuse of notation, we will henceforth let γ_i denote its extension to $\mathbb{R}$ by linear interpolation. Let $\Gamma_i \in C_0^+(\mathbb{R})$ be the extended Motzkin path defined by $\Gamma_i = \varphi^{-1} \circ \gamma_i \circ \varphi$. Now define

$$\bar{f}_i = (\gamma_i - \ell_0)/a = \sigma \circ \varphi \circ \Gamma_i \circ \varphi^{-1} = \psi \circ \varphi^*(\Gamma_i)$$

where $\psi(x) = \sigma \circ \varphi(x) = \frac{\varepsilon}{2} x$. Then $\psi(0) = \sigma(\ell_0) = 0$ and $\mathrm{m}(\bar{f}_i) = \mathrm{m}(\gamma_i) = \mathrm{m}(f_i)$. For $x \in \mathfrak{L}$, a direct computation gives $|f_i(x) - \bar{f}_i(x)| < \varepsilon/2$. For $x \notin \mathfrak{L}$, writing ℓ_x for the nearest lattice point to x gives

$$|f_i(x) - \bar{f}_i(x)| \le |f_i(x) - f_i(\ell_x)| + |f_i(\ell_x) - \bar{f}_i(\ell_x)| + |\bar{f}_i(\ell_x) - \bar{f}_i(x)| < \frac{\varepsilon}{4} + \frac{\varepsilon}{2} + \frac{\Delta}{2a} = \varepsilon,$$

hence $\|f_i - \bar{f}_i\|_\infty < \varepsilon$ as desired. $\qquad\square$

We are now ready to prove Lemma 2.3.

Proof of Lemma 2.3. Fix $j \ge 1$. To begin, we observe that it is enough to show the assertion for $I = \mathbb{R}$. Indeed, for any $I \subseteq \mathbb{R}$ and any $h \in C_0^+(I)$, we can define a function $\tilde{h} \in C_0^+(\mathbb{R})$ which equals h on I and drops linearly to zero on $[b, b+1]$ where b is the rightmost boundary point of I. This construction ensures that $\max \mathscr{E}^{j-1}(h) = \max \mathscr{E}^{j-1}(\tilde{h})$ and $\|h_1 - h_2\|_\infty = \|\tilde{h}_1 - \tilde{h}_2\|_\infty$.

Next we show that the result holds if the graphs of f and g are (extended) Motzkin paths by induction on $m = \max \mathscr{E}^{j-1}(f) + \max \mathscr{E}^{j-1}(g)$. The assertion is trivial when $j = 1$ or $m = 0$. If $\max \mathscr{E}^{j-1}(f) \ge 1$ and $\max \mathscr{E}^{j-1}(g) = 0$, write $\mathrm{m}_j := \mathrm{m}(\mathscr{E}^{j-1}(f))$. Let $J = [a, b]$ be the excursion interval of $\mathscr{E}^{j-1}(\Gamma)$ which contains m_j. By Proposition A.2, $\Gamma - \mathscr{E}^{j-1}(\Gamma)$ is constant on the excursion intervals of $\mathscr{E}^{j-1}(\Gamma)$. Hence we get

$$\max \mathscr{E}^{j-1}(f) = \mathscr{E}^{j-1}(f)(\mathrm{m}_j) - \mathscr{E}^{j-1}(f)(a) = f(\mathrm{m}_j) - f(a).$$

As $f, g \ge 0$, consideration of whether or not $g(\mathrm{m}_j) \ge g(a)$ shows that

$$\max \mathscr{E}^{j-1}(f) = f(\mathrm{m}_j) - f(a) \le f(\mathrm{m}_j) - f(a) + |g(\mathrm{m}_j) - g(a)|$$
$$\le |f(\mathrm{m}_j) - g(\mathrm{m}_j)| + |f(a) - g(a)| \le 2\|f - g\|_\infty,$$

By symmetry, the result also holds when $m \ge 1$ and $\max \mathscr{E}^{j-1}(f) = 0$, so we may assume that both $\max \mathscr{E}^{j-1}(f)$ and $\max \mathscr{E}^{j-1}(g)$ are at least 1. As the maxima are necessarily attained on hill intervals, Proposition A.4, the inductive hypothesis, and part (ii) of Proposition A.6 imply

$$\left| \max \mathscr{E}^{j-1}(f) - \max \mathscr{E}^{j-1}(g) \right| = \left| \max \mathscr{H} \circ \mathscr{E}^{j-1}(f) - \max \mathscr{H} \circ \mathscr{E}^{j-1}(g) \right|$$
$$= \left| \max \mathscr{E}^{j-1} \circ \mathscr{H}(f) - \max \mathscr{E}^{j-1} \circ \mathscr{H}(g) \right|$$
$$\le 2 \left\| \mathscr{H}(f) - \mathscr{H}(g) \right\|_\infty \le 2 \left\| f - g \right\|_\infty.$$

This completes the proof for Motzkin paths.

Now we show the assertion for $f, g \in C_0^+(\mathbb{R})$ by induction on $j \geq 1$. The base case is tautological. For the inductive step, choose $\psi, \varphi, \Gamma_1, \Gamma_2, \bar{f}, \bar{g}$ as in Proposition A.7 with $f_1 = f, f_2 = g$. Then by the choice of $\bar{f}$, Proposition A.4, the inductive hypothesis, and Proposition A.6 (i), we have

$$\left|\max \mathscr{E}^k(f) - \max \mathscr{E}^k(\bar{f})\right| = \left|\max \mathscr{E}^{k-1} \circ \mathscr{E}_{\mathfrak{m}(f)}(f) - \max \mathscr{E}^{k-1} \circ \mathscr{E}_{\mathfrak{m}(f)}(\bar{f})\right|$$
$$\leq \|\mathscr{E}_{\mathfrak{m}(f)}(f) - \mathscr{E}_{\mathfrak{m}(f)}(\bar{f})\|_\infty$$
$$\leq 2\|f - \bar{f}\|_\infty < 2\varepsilon,$$

and similarly for g. Also, since $\psi(0) = 0$, the triangle inequality gives

$$\psi(\|\Gamma_1 - \Gamma_2\|_\infty) = \psi(\|\varphi^* \Gamma_1 - \varphi^* \Gamma_2\|_\infty)$$
$$= \|\psi \circ \varphi^* \Gamma_1 - \psi \circ \varphi^* \Gamma_2\|_\infty$$
$$= \|\bar{f} - \bar{g}\|_\infty < 4\varepsilon + \|f - g\|_\infty.$$

Lastly, observe that the functional $\max \mathscr{E}^k$ satisfies

$$\max \mathscr{E}^k(\bar{f}_i) = \psi \circ \max \mathscr{E}^k(\Gamma_i).$$

Thus in conjunction with the assertion for the Motzkin paths, we obtain

$$\left|\max \mathscr{E}^k(f) - \max \mathscr{E}^k(g)\right| < 4\varepsilon + \left|\max \mathscr{E}^k(\bar{f}) - \max \mathscr{E}^k(\bar{g})\right|$$
$$\leq 4\varepsilon + \psi\left(\left|\max \mathscr{E}^k(\Gamma_1) - \max \mathscr{E}^k(\Gamma_2)\right|\right)$$
$$\leq 4\varepsilon + \psi(\|\Gamma_1 - \Gamma_2\|_\infty) < 8\varepsilon + \|f - g\|_\infty.$$

Letting $\varepsilon \searrow 0$ completes the inductive step and the proof. $\qquad\square$

A.4. **Statistics of 312-avoiding permutations.** In this subsection, we provide proofs of Propositions 8.3 and 8.1.

Recall that for each $n \geq 1$ and permutation $\tau \in \mathfrak{S}_3$ of length 3, we denote by $\mathfrak{S}_n^\tau$ the set of all τ-avoiding permutations of length n. Also recall that Dyck_{2n} denotes the set of all Dyck paths of length $2n$. Note that a permutation σ is 312-avoiding iff its inverse σ^{-1} is 231-avoiding. Also, if we denote by $\bar{\sigma}$ the *reversal* of σ obtained by reading σ from right to left, then σ is 231-avoiding iff its reversal $\bar{\sigma}$ is 132-avoiding.

There are a number of bijections between τ-avoiding permutations and Dyck paths in the literature. For instance, Krattenthaler [14] obtained a bijection $\mathfrak{S}_n^{132} \to \mathrm{Dyck}_{2n}$, and later Hoffman, Rizzolo, and Silvken [11] used a bijection $\mathrm{Dyck}_{2n} \to \mathfrak{S}_n^{231}$ to study random 231-avoiding permutations in terms of random walks and Brownian excursions. In fact, the inverse of the latter bijection is the conjugation of the former by reversals of permutations and Dyck paths, where the reversal of a Dyck path is its left-right mirror image. In the forthcoming proof of Proposition 8.3, we will make use of the bijection $\mathrm{Dyck}_{2n} \to \mathfrak{S}_n^{231}$ mentioned above, which we give below in a slightly more general version.

For a given h-restricted Motzkin path Γ, we define a permutation $\sigma(\Gamma)$ as follows: Let v_k be the location of the kth upstroke of Γ. (Thus if $\Gamma = \Gamma(X_0)$, then v_k is the location of the kth ball in X_0). Then we define a 231-avoiding permutation $\sigma(\Gamma)$ by

$$\sigma(\Gamma)(k) = k + \frac{1}{2} \inf\{r \geq 0 : \Gamma_{v_k+r} = \Gamma_{v_k-1}\} - \Gamma_{v_k}$$
$$= k + \frac{1}{2} \sup\{r \geq 0 : \Gamma_{v_k+r} \geq \Gamma_{v_k}\} + 1 - \Gamma_{v_k}. \tag{A.1}$$

When restricted to Dyck paths, this map $F \mapsto \sigma(F)$ is shown to be a bijection between Dyck_{2n} and $\mathfrak{S}_n^{231}$ in [11, Thm. 4.3].
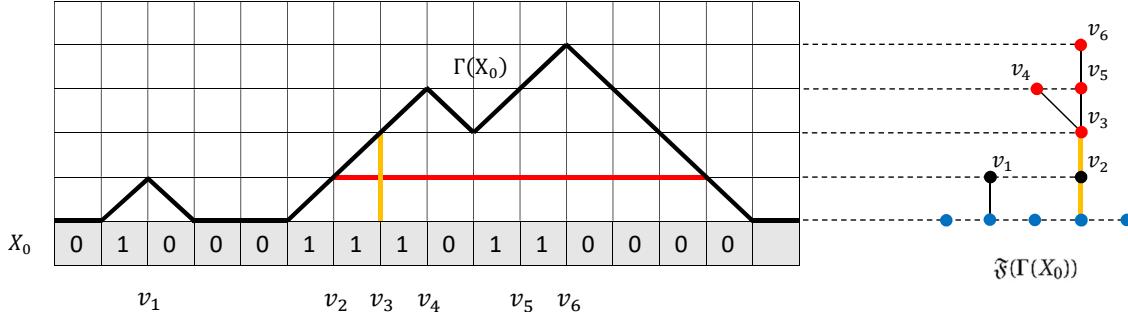
FIGURE A.3. Construction of the inverse 165243 of the 312-avoiding permutation $\sigma(X_0) = 146532$ directly from the corresponding Motzkin path $\Gamma(X_0)$ and rooted forest $\mathfrak{F}(\Gamma(X_0))$. On the left, the lengths of the red and orange paths correspond to the supremum and Γ_{v_k} terms in (A.1) for $k = 3$. On the right, the subtree rooted at v_3 consists of the four red nodes and the level of v_3 is the number of edges in the orange path. Thus (A.1) and (A.2) each show that $\sigma^{-1}(3) = 5$.

Remark A.8.   For a given rooted forest $\mathfrak{F}$, a permutation $\sigma(\mathfrak{F})$ can be defined similarly: Let v_k be the kth non-root node in $\mathfrak{F}$ according to the depth-first order and define

$$\sigma(\mathfrak{F})(k) := k + \#\{\text{nodes in the subtree of } \mathfrak{F} \text{ rooted at } v_k\} - (\text{level of } v_k \text{ in } \mathfrak{F}). \tag{A.2}$$

Note that the maps (A.2) and (A.1) yield the same permutation for corresponding rooted forest and its contour process. Namely, let Γ be the h-restricted Motzkin path which is a contour process of $\mathfrak{F}$. Then $\sigma(\Gamma) = \sigma(\mathfrak{F})$; see Figure A.3 for an illustration.

Proof of Proposition 8.3. We define a map $\varphi : \mathrm{Dyck}_{2n} \to \mathfrak{S}_n^{312}$ by

$$F \mapsto \sigma(F) \mapsto \sigma(F)^{-1},$$

where the first map is given by (A.1). As a composition of two bijections, φ is a bijection from Dyck_{2n} to $\mathfrak{S}_n^{312}$. This shows (i).

To show (ii), fix $F \in \mathrm{Dyck}_{2n}$ and let X_0 be the box-ball configuration obtained from F by

$$X_0(i) = \mathbf{1}\{F(i+1) - F(i) = 1\}$$

for all $i \geq 0$. It then suffices to show that

$$\sigma(X_0) = \sigma(F)^{-1}.$$

To this end, label the balls $1, \ldots, n$ from left to right, and recall the push-pop stack construction $X_0 \mapsto \sigma(X_0)$ described in Section 8. Fix a label $1 \leq k \leq n$. We are going to track the trajectory of ball k during the push-pop stack construction. Using the notation from Equation (A.1), let ball k be at site v_k. Note that F_{v_k} equals the number of balls in the stack after ball k is pushed onto it. Hence the number of balls which have been popped off in previous steps equals $k - F_{v_k}$. Next, while the stack sweeps sites to the right of v_k, balls with larger labels will be pushed on and popped off until ball k is finally deposited. This happens precisely when F first hits height $F_{v_k} - 1$ after location v_k. Accordingly, the

number of balls that are deposited during the period when ball k is in the stack equals the height of the subexcursion of F started at ν_k, which equals to half of the duration of this excursion. Thus

$$\text{\# balls popped out before ball } k = k - F_{\nu_k} + \frac{1}{2} \sup \left\{ r \geq 0 : F_{\nu_k + r} \geq F_{\nu_k} \right\}.$$

Therefore, $\sigma(F)(k)$, which is one more than the above quantity, is the position of k in $\sigma(X_0)$ as desired.
□

Proof of Proposition 8.1. Before we begin, recall the definition of the longest 'leftmost' increasing and 'rightmost' decreasing subsequences τ_+^* and τ_-^* given above the statement of Proposition 8.1.

We first show the assertion for τ_+^*. By induction on the length of the permutation, we suppose that the assertion holds for all 312-avoiding permutations of length less than n for some $n \geq 3$, and fix a 312-avoiding permutation τ of length n. (The result is true by inspection when $n = 3$.) Using Proposition 8.3, choose a box-ball configuration X_0 and a Dyck path F such that $\tau = \sigma(X_0)$ and $F = \Gamma(X_0)$.

By Greene's theorem ([10]), we know that the length of the first row of $\text{RS}(\tau)$ equals the length of any longest increasing subsequence in τ. Since $\text{RS}(\tau) = \Lambda(F)$, we see that the length of the longest increasing subsequence of τ equals the number of peaks in F.

Let X_0' be the box-ball configuration obtained from X_0 by deleting all 10 patterns from X_0, as in the proof of Lemma 2.1, and let $\Gamma' = \Gamma(X_0')$ and $\tau' = \sigma(X_0')$ be the h-restricted Motzkin path and 312-avoiding permutation constructed from X_0' (see the commutative diagram (A.3)). It is easy to see that Γ' can be directly obtained from Γ by first applying the hill-flattening operator $\mathcal{H}$ and then contracting new h-strokes which are not at height 0.

On the other hand, let L be the number of 10 patterns in X_0, which is the same as the number of peaks in Γ. When reading X_0 from left to right, let ℓ_i be the label of the ball that corresponds to the '1' in the ith 10 pattern. Then $\tilde{\tau} := \ell_1 \ell_2 \cdots \ell_L$ is an increasing subsequence in σ satisfying $\tau' = \tau \setminus \tilde{\tau}$. Moreover, Greene's theorem shows that this is a longest increasing subsequence. By an easy induction argument, one sees that ℓ_{i+1} is the first number to the right in σ that exceeds ℓ_i. Thus by definition, $\tilde{\tau} = \tau_+^*$, is the 'leftmost' longest increasing subsequence in σ. From the construction it is clear that $\tau' = \tau \setminus \tau_+^*$.

$$\begin{array}{ccc}
\tau & \longleftarrow X_0 \longrightarrow & \Gamma \\
\vdots & \downarrow & \vdots \\
\tau' & \longleftarrow X_0' \longrightarrow & \Gamma'
\end{array} \tag{A.3}$$

To complete the argument, recall that $\Lambda(\Gamma')$ is obtained from $\Lambda(\Gamma)$ by deleting the first row. Since $\text{RS}(\tau) = \Lambda(\Gamma)$ and $\text{RS}(\tau') = \Lambda(\Gamma')$ by Proposition 8.2, we have that $\text{RS}(\tau')$ is obtained from $\text{RS}(\tau)$ by deleting its first row. Since τ' can be obtained from τ by deleting a longest increasing subsequence, the inductive hypothesis applied to τ' completes the proof.

Next, we show the assertion for the columns. Let τ, X_0, Γ be as before. To begin, observe that in the stack construction of τ from X_0, every decreasing subsequence in τ is generated by the balls that occupy the stack at the same time. (For instance, in Figure 10, the decreasing subsequence 432 in $\sigma = 14632$ is generated by the balls in the stack on top of the ball of label 4.) Thus, every longest decreasing subsequence in τ is generated by the balls in the stack where Γ achieves its maximum.

Let $\mathtt{m}^*$ be any location where Γ attains its global maximum. During the stack operation to construct τ from X_0, let $\bar{\tau} = \ell_1 \ell_2 \cdots \ell_M$ be the decreasing sequence consisting of the numbers in the stack after pushing all the balls over the interval $[1, \mathtt{m}^*]$. This is a longest decreasing subsequence in τ. Denote

$\tau^\dagger = \tau \setminus \tau_-^*$. Let $X_0^\dagger$ be the box-ball configuration that is obtained by converting 1's that correspond to balls with labels in $\bar{\tau}$ to 0's. Then observe that $\sigma(X_0^\dagger) = \tau^\dagger$ and $\Gamma^\dagger = \mathscr{E}_{\mathtt{m}^*}(\Gamma)$, where $\mathscr{E}_{\mathtt{m}^*}$ is the excursion operator pivoted at location $\mathtt{m}^*$ instead of the rightmost one $\mathtt{m}$. According to Lemma 2.2 and the following remark, $\Lambda(\mathscr{E}_{\mathtt{m}^*}(\Gamma))$ is obtained by deleting the first column of $\Lambda(\Gamma)$. Since $\Lambda(\mathscr{E}_{\mathtt{m}^*}(\Gamma)) = \mathrm{RS}(\tau)$ and $\Lambda(\Gamma) = \mathrm{RS}(\tau^\dagger)$, the assertion follows. $\qquad\square$

ACKNOWLEDGMENTS

REFERENCES

[1] David Aldous. The continuum random tree. III. *Ann. Probab.*, 21(1):248–289, 1993.

[2] Louis H. Y. Chen, Larry Goldstein, and Qi-Man Shao. *Normal approximation by Stein's method.* Probability and its Applications (New York). Springer, Heidelberg, 2011.

[3] David Roxbee Cox. *Renewal theory.* Methuen & Co. Ltd., London; John Wiley & Sons, Inc., New York, 1962.

[4] David A Croydon, Tsuyoshi Kato, Makiko Sasada, and Satoshi Tsujimoto. Dynamics of the box-ball system with random initial conditions via pitman's transformation. *arXiv preprint arXiv:1806.02147*, 2018.

[5] Endre Csáki and Yueyun Hu. Lengths and heights of random walk excursions. In *Discrete random walks (Paris, 2003)*, Discrete Math. Theor. Comput. Sci. Proc., AC, pages 45–52. Assoc. Discrete Math. Theor. Comput. Sci., Nancy, 2003.

[6] Michael Drmota. Stochastic analysis of tree-like data structures. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, 460(2041):271–307, 2004. Stochastic analysis with applications to mathematical finance.

[7] Rick Durrett. *Probability: theory and examples.* Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, fourth edition, 2010.

[8] Paul Erdős and Alfred Rényi. On the evolution of random graphs. *Magyar Tud. Akad. Mat. Kutató Int. Közl.*, 5:17–61, 1960.

[9] Pablo A Ferrari, Chi Nguyen, Leonardo Rolla, and Minmin Wang. Soliton decomposition of the box-ball system. *arXiv preprint arXiv:1806.02798*, 2018.

[10] Curtis Greene. An extension of schensted's theorem. In *Young Tableaux in Combinatorics, Invariant Theory, and Algebra*, pages 39–50. Elsevier, 1982.

[11] Christopher Hoffman, Douglas Rizzolo, and Erik Slivken. Pattern-avoiding permutations and Brownian excursion, part II: fixed points. *Probab. Theory Related Fields*, 169(1-2):377–424, 2017.

[12] Samuel Karlin and James McGregor. The classification of birth and death processes. *Trans. Amer. Math. Soc.*, 86:366–400, 1957.

[13] Donald E. Knuth. *The art of computer programming. Vol. 1.* Addison-Wesley, Reading, MA, third edition, 1997. Fundamental algorithms.

[14] Christian Krattenthaler. Permutations with restricted patterns and dyck paths. *Advances in Applied Mathematics*, 27(2-3):510–530, 2001.

[15] Agnès Lagnoux, Sabine Mercier, and Pierre Vallois. Probability that the maximum of the reflected Brownian motion over a finite interval $[0, t]$ is achieved by its last zero before t. *Electron. Commun. Probab.*, 20:no. 62, 9, 2015.

[16] Peter Mörters and Yuval Peres. *Brownian motion*, volume 30 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, 2010. With an appendix by Oded Schramm and Wendelin Werner.

[17] Atsushi Nagai, Daisuke Takahashi, and Tetsuji Tokihiro. Soliton cellular automaton, Toda molecule equation and sorting algorithm. *Phys. Lett. A*, 255(4-6):265–271, 1999.

[18] Jim Pitman. Brownian motion, bridge, excursion, and meander characterized by sampling at independent uniform times. *Electron. J. Probab.*, 4:no. 11, 33, 1999.

[19] Jim Pitman and Matthias Winkel. Growth of the brownian forest. *Ann. of Probab.*, 33(6):2188–2211, 2005.

[20] Jim Pitman and Matthias Winkel. Regenerative tree growth: Markovian embedding of fragmenters, bifurcators, and bead splitting processes. *Ann. of Probab.*, 43(5):2611–2646, 2015.

[21] Jim Pitman and Marc Yor. On the distribution of ranked heights of excursions of a Brownian bridge. *Ann. Probab.*, 29(1):361–384, 2001.

[22] Dan Romik. *The surprising mathematics of longest increasing subsequences*, volume 4. Cambridge University Press, 2015.

[23] Doron Rotem. Stack sortable permutations. *Discrete Math.*, 33(2):185–196, 1981.

[24] Bruce E. Sagan. *The symmetric group*, volume 203 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 2001. Representations, combinatorial algorithms, and symmetric functions.

[25] Daisuke Takahashi and Junkichi Satsuma. A soliton cellular automaton. *J. Phys. Soc. Japan*, 59(10):3514–3519, 1990.

[26] Makoto Torii, Daisuke Takahashi, and Junkichi Satsuma. Combinatorial representation of invariants of a soliton cellular automaton. *Phys. D*, 92(3-4):209–220, 1996.

LIONEL LEVINE, DEPARTMENT OF MATHEMATICS, CORNELL UNIVERSITY, ITHACA, NY 14853.
Email address: `levine@math.cornell.edu`

HANBAEK LYU, DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA, LOS ANGELES, CA 90095.
Email address: `colourgraph@gmail.com`

JOHN PIKE, DEPARTMENT OF MATHEMATICS, BRIDGEWATER STATE UNIVERSITY, BRIDGEWATER, MA 02324.
Email address: `john.pike@bridgew.edu`