ABSTRACT. The box-ball systems are integrable cellular automata whose long-time behavior is characterized by the soliton solutions, and have rich connections to other integrable systems such as Korteweg-de Vries equation. In this paper, we consider multicolor box-ball system with two types of random initial configuration and obtain the scaling limit of the soliton lengths as the system size tends to infinity. Our analysis is based on modified Greene-Kleitman invariants for the box-ball systems and associated circular exclusion processes.

1. Introduction

1.1. The $\kappa$-color BBS. The box-ball systems (BBS) are integrable cellular automata in 1+1 dimension whose long-time behavior is characterized by the soliton solutions. The $\kappa$-color BBS is a cellular automaton on the half-integer lattice $\mathbb{N}$, which we think of as an array of capacity-one boxes that can fit at most one ball of any of the $\kappa$ colors. At each discrete time $t \geq 0$, the system configuration is given by a coloring $X_t : \mathbb{N} \rightarrow \mathbb{Z}_{\kappa+1} := \{0, 1, \cdots, \kappa\}$ with finite support. When $X_t(x) = i$, we say the site $x$ is empty at time $t$ if $i = 0$ and occupied with a ball of color $i$ at time $t$ if $1 \leq i \leq \kappa$. To define the time evolution rule, for each $1 \leq a \leq \kappa$, let $K_a$ be the operator on the set $(\mathbb{Z}_{\kappa+1})^\mathbb{N}$ of all $(\kappa + 1)$-colorings on $\mathbb{N}$ defined as follows:

(i) Label the balls of color $a$ from left to right as $a_1, a_2, \cdots, a_m$.

(ii) Starting from $k = 1$ to $m$, successively move ball $a_k$ to the leftmost empty site to its right.

Then the time evolution $(X_t)_{t \geq 0}$ of the basic $\kappa$-color BBS is given by

$$X_{t+1} = K_1 \circ K_2 \circ \cdots \circ K_\kappa(X_t) \quad \forall t \geq 0.$$

A typical 5-color BBS trajectory is shown below.

$t = 0: 00312051300411252003211000000000000000000000000000$

$t = 1: 0001320153001415220032110000000000000000000000000$

$t = 2: 0000103021530010410522000321100000000000000000000$

$t = 3: 000000103002153010041005220000032110000000000000$

$t = 4: 000000010300021503100410005220000003211000000000$

$t = 5: 000000001030000251031004100052200000032110000000$

$t = 6: 000000000103000020510310041000522000000032110000$

Note that a sequence of $k$ balls of non-increasing colors travel to the right with speed $k$ until it interferes other balls in front. We call such as sequence a soliton of length $k$ if its length and content are preserved by the BBS dynamics in all future steps. For instance, all of the non-increasing consecutive sequences of balls in $X_5$ in the example above are solitons, since they are is preserved in $X_6$ up to their location changes and will be so in all future configurations. The grounding observation in the $\kappa$-color BBS is that any finite system eventually decomposes into solitons of non-decreasing lengths from left to right, which is called the soliton decomposition of the system $(X_t)_{t \geq 0}$. This final macrostate of the system
can be encoded in a Young diagram \( \Lambda = \Lambda(X_0) \) having \( j \)th column equal in length to the \( j \)th longest soliton. For instance, below is the Young diagram corresponding to the soliton decomposition of instance of the 5-color BBS given before:

\[
\Lambda(X_0) = 
\begin{array}{cccc}
\_ & \_ & \_ & \_ \\
\_ & \_ & \_ & \_ \\
\_ & \_ & \_ & \_ \\
\_ & \_ & \_ & \_ \\
\_ & \_ & \_ & \_ \\
\end{array}
\]

Note that the \( i \)th row of the Young diagram \( \Lambda(X_0) \) is precisely the number of solitons of length at least \( i \).

1.2. **Overview of main results.** We consider \( \kappa \)-color BBS initialized by a random BBS configuration of system size \( n \), and analyze the limiting shape of the random Young diagrams as \( n \) tends to infinity. We consider two models that we call the ‘permutation model’ and ‘independence model’.

In the permutation model, the BBS is initialized by a uniformly chosen random permutation \( \Sigma^n \) of colors \( \{1, 2, \cdots, \kappa\} \). A classical way of associating a Young diagram to a permutation is via the Robinson-Schensted correspondence (see [Sag01, Ch. 3.1]). A famous result of Baik, Deift, and Johansson [BDJ99] independently drawn from a fixed distribution \( p \) corresponds as \( \kappa = 2 \). We consider \( \kappa \)-color BBS given before: \( \rho_{k}(n) = \rho_{k}(n) \), respectively, then

\[
\rho_{k}(n) \sim \frac{n}{k(k+1)}, \quad \lambda_{k}(n) \sim \frac{2\sqrt{n}}{\sqrt{k}+\sqrt{k+1}}.
\]

The row and column scalings are consistent since the majority of solitons have length of order \( O(1) \). Moreover, the linear scaling for row lengths is not too surprising in the BBS literature, as they are given by an additive functional of some Markov chains (namely, the carrier processes) over the initial configuration (see [KL18, KLO18b]).

| \( i \geq 1, j \geq 2 \) fixed | \( \rho_i(n) \) | \( \lambda_1(n) \) | \( \lambda_j(n) \) |
|---|---|---|---|
| Subcritical phase \((p^* < p_0)\) | \( \Theta(n) \) | \( \Theta(\log n) \) | \( \Theta(\log n) \) |
| Critical phase \((p^* = p_0)\) | \( \Theta(n) \) | \( \Theta(\sqrt{n}) \) | \( \Theta(\sqrt{n}) \) |
| Supercritical phase \((p^* > p_0)\) | Simple \((p^* = p_\ell \) for unique \( \ell \)\) | \( \Theta(n) \) | \( \Theta(n) \) | \( \Theta(\log n) \) |
| Non-simple \((p^* = p_\ell \) for multiple \( \ell \)\) | | | | \( O(\sqrt{n}) \cap \Omega(\sqrt{n}/\log n) \) |

**Table 1.** Asymptotic scaling of column and row lengths for the independence model with ball density \( p = (p_0, p_1, \cdots, p_\kappa) \) and \( p^* = \max(p_1, \cdots, p_\kappa) \). The asymptotic soliton lengths undergo a similar ‘double-jump’ phase transition depending on \( p^* - p_0 \), as in the \( \kappa = 1 \) case established in [LLP17]. The existence of non-simple supercritical phase is unique to the multicolor \( (\kappa \geq 2) \) case, where subsequent soliton lengths scales as \( \sqrt{n} \) instead of \( \log n \). Sharp asymptotics for the row lengths has been obtained in [KL18]. The constant factors depend \( p, i, j \).

In the independence model, which we denote \( X^n_{\mathbf{p}} \), the color of each site in the interval \([1, n]\) are independently drawn from a fixed distribution \( \mathbf{p} = (p_0, p_1, \cdots, p_\kappa) \) on \( \mathbb{Z}_{k+1} \). Recently, Lyu and Kuniba obtained sharp asymptotics for the row lengths as well as their large deviations principle in this independence model [KL18]. Our main result in the present paper establishes the scaling limit for the column lengths for the independence model, as summarized in Table 1. We find a similar ‘double-jump’ phase
transition for the \( \kappa = 1 \) case established by Levine, Lyu, and Pike [LLP17]. In the multicolor \((\kappa \geq 2)\) case, the maximum positive ball density \( p^* = \max(p_1, \ldots, p_\kappa) \) compared to the zero density \( p_0 \) dictates general phase transition structure. Interestingly, we discover the ‘non-simple supercritical phase’, when the maximum ball density \( p^* \) is achieved by multiple colors \( i \). In this case, subsequent soliton lengths scale as \( \sqrt{n} \), instead of \( \log n \) as in the simple supercritical phase as well as the \( \kappa = 1 \) case in [LLP17].

1.3. Statement of results. Our main results concern the asymptotic behavior of top soliton lengths associated with the \( \kappa \)-color BBS trajectory for two models of random initial configuration \( X_0^{n, \kappa} = X_0 \): \((1)\) \( \kappa \) is fixed and \( X_0(x) = i \) independently with a fixed probability \( p_i, i \in \mathbb{Z}_{\kappa+1} \) for each \( x \in [1, n] \), and \((2)\) \( \kappa = n \) and \( X_0[1, n] \) is a random uniform permutation of length \( n \).

More precisely, for the permutation model, let \( X := (U_x)_{x \geq 1} \) be a sequence of i.i.d. \( \text{Uniform}([0,1]) \) random variables. For each integer \( n \geq 1 \), we denote by \( V_{1:n} < V_{2:n} < \cdots < V_{n:n} \) the order statistics of \( U_1, U_2, \cdots, U_n \). Then it is easy to see that the random permutation \( \Sigma^n \) on \([n]\) such that \( V_{i:n} = U_{\Sigma^n(i)} \) for all \( 1 \leq i \leq n \) is uniformly distributed among all permutations on \([n]\). Define \( X^n(x) = \Sigma^n(x) \cdot 1(1 \leq x \leq n) \).

To define the independence model, fix integers \( n, \kappa \geq 1 \). Let \( p = (p_0, p_1, \ldots, p_\kappa) \) be a probability distribution on \([0,1,\cdots,\kappa]\). Let \( X^p \) be a random map \( \mathbb{N} \to \{0,1,\cdots,\kappa\} \) such that \( \mathbb{P}(X^p(x) = i) = p_i \) independently for all \( x \in \mathbb{N} \) and \( 0 \leq i \leq \kappa \). Define \( \kappa \)-color and \( n \)-color BBS configurations \( X^{n,p} \) and \( X^n \) by

\[
X^{n,p}(x) = X^p(x) \cdot 1(1 \leq x \leq n).
\]

We may further assume, without loss of generality, that \( p_i > 0 \) for all \( 0 \leq i \leq \kappa \). Indeed, if \( p_i = 0 \) for some \( i \), then we can omit the color \( i \) entirely and consider the system as a \((\kappa - 1)\)-color BBS by shifting the colors \( \{i+1, \cdots, \kappa\} \) to \( \{i, \cdots, \kappa - 1\} \).

We now state our main results. For the permutation model, we obtain a precise first order asymptotic for the largest \( k \) rows and columns, as stated in the following theorem.

**Theorem 1.1.** Let \( X^n \) be as above. For each \( k \geq 1 \), denote \( \rho_k(n) = \rho_k(X^n) \) and \( \lambda_k(n) = \lambda_k(X^n) \). Then for each fixed \( k \geq 1 \), almost surely,

\[
\lim_{n \to \infty} n^{-1} \rho_k(n) = \frac{1}{k(k+1)}, \quad \lim_{n \to \infty} n^{-1/2} \lambda_k(n) = \frac{2}{\sqrt{k+\sqrt{k+1}}).
\]

For soliton lengths of the independence model, we establish the following double-jump phase transition behavior for top soliton lengths.

**Theorem 1.2.** Fix \( \kappa \geq 1 \) and let \( X^{n,p} \) be as above. Denote \( \lambda_j(n) = \lambda_j(X^{n,p}) \) and \( p^* = \max_{1 \leq i \leq \kappa} p_i \). Fix constant \( \varepsilon > 0 \).

(i) (Subcritical phase) Suppose \( p^* < p_0 \). Then there exists constants \( \theta_2 \geq \theta_1 \geq p_0/p_\kappa \) and \( C_1, C_2 > 0 \) such that for any non-decreasing real sequence \( \{x_n\}_{n \geq 1} \),

\[
\exp(-\theta_1^{-x_n}) \leq \liminf_{n \to \infty} \mathbb{P}(\lambda_1(n) \leq x_n + \mu_n^{(1)}) \quad \text{and} \quad \limsup_{n \to \infty} \mathbb{P}(\lambda_1(n) \leq x_n + \mu_n^{(2)}) \leq \exp(-\theta_2^{-x_n + 1}),
\]

where \( \mu_n^{(i)} = \log_{\theta_i}(C_i n) \) for \( i = 1,2 \). Furthermore, for any \( j \geq 2 \),

\[
\limsup_{n \to \infty} \mathbb{P}(\lambda_j(n) \leq x_n + \mu_n^{(2)}) \leq \exp(-\theta^{-x_n + 1}) \sum_{k=0}^{j-1} \theta^{-kx_n}.
\]
(ii) (Critical phase) Suppose \( p^* = p_0 \). Then \( \lambda_j(n) = \Theta(\sqrt{n}) \) for each fixed \( j \geq 1 \). Furthermore, let \( r = \left| \{ 1 \leq i \leq \kappa : p_i = p_0 \} \right| \). Then there exists (not necessarily independent) standard Brownian motions \( B^{(1)}, \ldots, B^{(r)} \) and constants \( \gamma_1, \ldots, \gamma_r \) such that

\[
y_{r, \max}|B^{(r)}| \leq \lim_{n \to \infty} \inf n^{-\frac{1}{2}} \lambda_1(n) \leq \lim_{n \to \infty} \sup n^{-\frac{1}{2}} \lambda_1(n) \leq \sum_{i=1}^{r} \gamma_i \max |B^{(i)}|, \tag{9}\]

where \( \leq \) denotes stochastic domination. In particular, if \( r = 1 \), then

\[
n^{-\frac{1}{2}} \lambda_1(n) \Rightarrow \gamma_1 \max |B^{(1)}|, \tag{10}\]

where \( \Rightarrow \) denotes weak convergence.

(iii) (Simple supercritical phase) Suppose \( p^* > p_0 \) and \( p_i = p_i^* \) for a unique \( 1 \leq i \leq \kappa \). Then almost surely,

\[
\lim_{n \to \infty} n^{-1} \lambda_1(n) = p^* - p_0 \quad \text{a.s.} \tag{11}\]

Moreover, for any fixed \( \varepsilon \in (0, 1) \) and \( j \geq 2 \), \( \lambda_j(n) = \Theta(\log n) \) with probability at least \( 1 - \varepsilon \). If we further assume that there is a single unstable color, then there exists a constant \( c > 0 \) such that

\[
\frac{\lambda_1(n) - (p^* - p_0)n}{\sqrt{cn}} \Rightarrow Z \sim \mathcal{N}(0, 1), \tag{12}\]

where \( \mathcal{N}(0, 1) \) denotes standard normal distribution and \( \Rightarrow \) denotes weak convergence.

(iv) (Non-simple supercritical phase) Suppose \( p^* > p_0 \) and \( p^* = p_i \) for more than one \( 1 \leq i \leq \kappa \). Then almost surely,

\[
\lim_{n \to \infty} n^{-1} \lambda_1(n) = p^* - p_0 \quad \text{a.s.} \tag{13}\]

Moreover, for any fixed \( \varepsilon \in (0, 1) \) and \( j \geq 2 \), \( \lambda_j(n) = O(\sqrt{n}) \) and \( \lambda_j = \Omega(\sqrt{n} \log n) \) with probability at least \( 1 - \varepsilon \).

1.4. Background and related works. The \( \kappa \)-color BBS was introduced in [Tak93], generalizing the original \( \kappa = 1 \) BBS first invented by Takahashi and Satsuma in 1990 [TS90]. In the most general form of the BBS, each site accommodates a semistandard tableau of rectangular shape with letters from \( \{ 0, 1, \ldots, \kappa \} \) and the time evolution is defined by successive application of the combinatorial \( R \) (cf. [FYOO0, HHI+01, KOS+06, IKT12]). The \( \kappa \)-color BBS treated in this paper corresponds to the case where the tableau shape is a single box, which was called the basic \( \kappa \)-color BBS in [KL18]. BBS is known to arise both from the quantum and classical integrable systems by the procedures called crystallization and ultradiscretization, respectively. This double origin of the integrability of BBS lies behind its deep connections to quantum groups, crystal base theory, solvable lattice models, the Bethe ansatz, soliton equations, ultradiscretization of the Korteweg-de Vries equation, tropical geometry and so forth; see for example the review [IKT12] and the references therein.

BBS with random initial configuration is an emerging topic in the probability literature, and has gained considerable attention with a number of recent works [LLP17, CKST18, KL18, FG18, KL18, CS19a, CS19b]. There are roughly two central questions that the researchers are aiming to answer: 1) If the random initial configuration is one-sided, what is the limiting shape of the invariant random Young diagram as the system size tends to infinity? 2) If one considers the two-sided BBS (where the initial configuration is a bi-directional array of balls), what are the two-sided random initial configurations that are invariant under the BBS dynamics? Some of these questions have been addressed for the basic 1-color BBS [LLP17, FNRW18, FG18, CKST18] as well as for the multicolor case [KL18, KLO18b].

There are two important works which are strongly related to this paper. In [LLP17], Levine, Lyu, and Pike studied various soliton statistics of the basic 1-color BBS when the system is initialized according to
a Bernoulli product measure with ball density $p$ on the first $n$ boxes. One of their main results is that 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, there is a condensation toward the longest soliton in the supercritical $p > 1/2$ regime in the sense that, 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 $p > 1/2$. Their analysis is based on geometric mappings from the associated simple random walks to the invariant Young diagrams, which enable robust analysis of the scaling limit of the invariant Young diagram. However, this connection is not apparent in the general $\kappa \geq 1$ case. In fact, one of the main difficulties in analyzing the soliton lengths in the multicolor BBS is that within a single regime, there is a mixture of behaviors that we see from different regimes in the single-color case.

The row lengths in the multicolor BBS are well-understood due to recent works by Kuniba, Lyu and Okado [KLO18b] and Kuniba and Lyu [KL18]. The central observation is that, when the initial configuration is given by a product measure, then the sum of row lengths can be computed via some additive functional (called ‘energy’) of carrier processes of various shapes, which are finite-state Markov chains whose time evolution is given by combinatorial $R$. In [KLO18b], the ‘stationary shape’ of the Young diagram for the most general type of BBS is identified by the logarithmic derivative of a deformed character of the KR modules (or Schur polynomials in the basic case). In [KL18], for the (basic) $\kappa$-color BBS that we consider in the present paper, it was shown that the row lengths satisfy a large deviations principle and hence the Young diagram converges to the stationary shape at an exponential rate, in the sense of row scaling.

The central subject of this paper is the column lengths of the Young diagram for the basic $\kappa$-color BBS. We develop two main tools for our analysis, which are a modified version of Greene-Kleitman invariants for BBS (Subsection 2.1) and infinite-capacity carrier process (Subsection 2.2). Especially, the latter gives rise to a ‘circular exclusion process’, which can be regarded as a circular version of the well-known Totally Asymmetric Simple Exclusion Process (TASEP) on a line (see, e.g., [F\textsuperscript{+}18, BFPS07, BFS08]). For its rough description, consider the following process on the unit circle $S^1$. Starting from some finite number of points, at each time, a new point is added to $S^1$ independently from a fixed distribution, which then deletes the nearest counterclockwise point already on the circle. Equivalently, one can think of each point in the circle trying to jump to the clockwise direction. It turns out that this process is crucial in analyzing the permutation model (Subsection 3.2), whereas for the independence model, the relevant circular exclusion process is defined on the integer ring $\mathbb{Z}_{\kappa+1}$ where points can stack up at the same location (Subsection 2.2). Interestingly, a cylindric version of Schur functions has been used to study rigged configurations and BBS [LPS14].

1.5. Organization. The rest of this paper is organized as follows. In Section 2, we introduce infinite and finite capacity carrier processes for the $\kappa$-color BBS and state the three key lemmas (Lemmas 2.1, 2.2, and 2.3). In Section 3, we prove our main result for the permutation model (Theorem 1.1) by using the modified GK invariants for BBS (Lemma 2.1) and analyzing the associated circular exclusion process. In Section 4, we begin our analysis on the infinite-capacity carrier process for the independence model. We decompose the carrier process into i.i.d. excursions, and show that the order statistics of the excursion heights give tight bounds for the first soliton length and lower bounds for the subsequent solitons (Lemma 4.1). In the following section, Section 5, we obtain the stationary distribution for the carrier process in the subcritical regime, and prove Theorem 1.2 (i), assuming the excursion lengths have finite exponential moment (Lemma 5.4). In Section 6, we introduce and analyze a ‘decoupled’ version of the infinite capacity carrier process in order to study the critical and supercritical regimes. This allows us to express the multiplicity of balls of ‘unstable’ colors in the carrier as an additive functional of a stationary
Markov chain that only depends on ‘stable’ colors. We prove Theorem 1.2 (ii)-(iv) in Section 7. Lastly, in Sections 8 and 9 we provide postponed proofs for the probabilistic and combinatorial lemmas.

1.6. Notation. We use the convention that summation and product over the empty index set equals zero and one, respectively. For any probability space $(\Omega, \mathcal{F}, P)$ and any event $A \in \mathcal{F}$, we let $1(A)$ the indicator variable of $A$.

We adopt the notations $\mathbb{R}^+ = [0, \infty)$, $\mathbb{N} = \{1, 2, 3, \ldots\}$, 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 $\{a_n\}_{n=1}^{\infty} \subset \mathbb{R}^+$ 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 $P\{W_n > Ca_n\} < \varepsilon$ 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 $P\{W_n < ca_n\} < \varepsilon$ 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 $\varepsilon$ but not $n$.

2. Key Lemmas

2.1. Modified Greene-Kleitman invariants for BBS. Perhaps one of the most natural way to associate a Young diagram with a given permutation is to use the celebrated Robinson-Schensted correspondence (see [Sag01, Ch. 3.1]), which gives a bijection between permutations and pairs of standard Young tableau of the same shape. For each permutation $\sigma$, record the common shape of the Young tableau as $\Lambda_{RS}(\sigma)$. According to Greene’s theorem [Gre82], the sum of the lengths of the first $k$ columns (resp. rows) of $\Lambda_{RS}(\sigma)$ is equal to the length of a longest subsequence in $\sigma$ that can be obtained by taking the union of $k$ (decreasing) subsequences. That is, for each $k \geq 1$,

$$\rho_1(\Lambda_{RS}(\sigma)) + \cdots + \rho_k(\Lambda_{RS}(\sigma)) = \max\{|\bigcup_{1 \leq i \leq k} \text{ increasing subsequences of } \sigma|\}, \quad (14)$$

$$\lambda_1(\Lambda_{RS}(\sigma)) + \cdots + \lambda_k(\Lambda_{RS}(\sigma)) = \max\{|\bigcup_{1 \leq i \leq k} \text{ decreasing subsequences of } \sigma|\} \quad (15)$$

The quantities on the right hand sides are call the Greene-Kleitman invariants.

On the other hand, if we consider the $n$-color BBS trajectory started at $X_0 = \sigma 1[1, n]$, then we obtain another Young diagram, which we denote by $\Lambda_{BBS}(\sigma) := \Lambda(X_0)$, whose $j$th column equals the $j$th longest soliton length. Then a natural question arises: Do the sums of the first $k$ rows and columns of $\Lambda_{BBS}(\sigma)$ relate to some type of Greene-Kleitman invariants? For the rows, we find that the correct modification is to localize the length of an increasing sequence into the number of ascents in a subsequence. On the other hand, for the columns, it turns out that we just need to impose that the $k$ decreasing subsequences be non-interlacing. In fact, in Lemma 2.1, we establish these modified Greene-Kleitman invariants for BBS in the more general setting when $\sigma$ is an arbitrary $\kappa$-color BBS configuration with finite support, where having 0’s and repetitions are both allowed.

Let $X : \mathbb{N} \to \{0, 1, \ldots, \kappa\}$ be a $\kappa$-color BBS configuration with finite support. For subsets $A, B \subseteq \mathbb{N}$, denote $A < B$ if $\max(A) < \min(B)$. We say $A, B$ are non-interlacing if $A < B$ or $B < A$. We say $X$ is non-increasing on $A \subseteq \mathbb{N}$ if $X(a_1) \geq X(a_2)$ for all $a_1, a_2 \in A$ such that $a_1 \leq a_2$. Denoting the elements of $A$ by $a_1 < a_2 < \cdots$, define the number of ascents of $X$ in $A$ by

$$\text{NA}(A, X) := 1 + \sum_{i=2}^{\infty} I(X(a_{i-1}) < X(a_i)). \quad (16)$$

Moreover, define the penalized length of $A$ with respect to $X$ by

$$L(A, X) := \left|A| - \sum_{\min A \leq i \leq \max A} I(X(i) = 0) \right| I(X \text{ is non-increasing on } A). \quad (17)$$
Lemma 2.1. Let \((X_t)_{t \geq 0}\) be a \(\kappa\)-color BBS trajectory such that \(X_0\) has finite support. Then for each \(k, t \geq 0\), we have

\[
\begin{align*}
\rho_1(\Lambda(X_0)) + \cdots + \rho_k(\Lambda(X_0)) &= \max_{A_1, \ldots, A_k = N} \sum_{i=1}^k N(A_i, X_t), \\
\lambda_1(\Lambda(X_0)) + \cdots + \lambda_k(\Lambda(X_0)) &= \max_{A_1, \ldots, A_k \subseteq \mathbb{N}} \sum_{i=1}^k L(A_i, X_t).
\end{align*}
\] (18) (19)

The proof of Lemma 2.1 may be found in Section 9.

2.2. Infinite capacity carrier process and first soliton length. The definition of \(\kappa\)-color BBS dynamics we gave in the introduction involves non-local movement of balls. It can instead be defined using a 'carrier', which gives a localized characterization of the process and reveals a number of invariants. For the simplest case of \(\kappa = 1\), imagine a carrier of infinite capacity sweeps through the time-\(t\) configuration \(X_t\) from left to right, picking up each ball it encounters and depositing a ball into each empty box whenever it can. After we run this carrier over \(X_t\), the resulting configuration is in fact \(X_{t+1}\). Moreover, the maximum number of balls in the carrier during the sweep is in fact the first soliton length \(\lambda_1\). For \(\kappa \geq 1\), we give a carrier version of the \(\kappa\)-color BBS dynamics, and we show that the maximum number of balls of positive color during the sweep also equals the first soliton length \(\lambda_1\). Furthermore, running finite capacity carriers will extract the row lengths of the invariant Young diagram.

We first define the infinite-capacity carrier process and the carrier version of the dynamics for general \(\kappa \geq 1\). Denote

\[
\mathcal{B}_\infty = \{x \in \{0,1,\cdots,\kappa\}^\mathbb{N} | x \text{ is non-increasing and has finite support}\}.
\] (20)

Define a map \(\Psi : \mathcal{B}_\infty \times \{0,1,\cdots,\kappa\} \rightarrow \{0,1,\cdots,\kappa\} \times \mathcal{B}_\infty, (x, y) \mapsto (y', x')\) by the following 'circular exclusion rule':

(i) Suppose \(y \geq 1\) and denote \(i^* = \min\{i \geq 1 \mid x(i) < y\}\). Then \(y' = x(i^*)\) and

\[
x'(i) = x(i)1(i \neq i^*) + y1(i = i^*) \quad \forall i \geq 1.
\] (21)

(ii) Suppose \(y = 0\). Then \(y' = x(1) = \max(x)\) and

\[
x'(i) = x(i+1) \quad \forall i \geq 1.
\] (22)

Fix a \(\kappa\)-color BBS configuration \(X : \mathbb{N} \rightarrow \{0,1,\cdots,\kappa\}\). Fix \(\Gamma_0 \in \mathcal{B}_\infty\), and recursively define a new \(\kappa\)-color BBS configuration \(X'\) and a sequence \((\Gamma_t)_{t \geq 0}\) of elements of \(\mathcal{B}_\infty\) by

\[
(X'(t+1), \Gamma_{t+1}) = \Psi(\Gamma_t, X(t+1)) \quad \forall t \geq 0.
\] (23)

We call the sequence \((\Gamma_t)_{t \geq 0}\) the infinite capacity carrier process over \(X\). Unless otherwise mentioned, we will assume \(\Gamma_0 = [0,0,0,\cdots] \in \mathcal{B}_\infty\). See Figure 1 for an illustration.

It turns out that the map \(X \mapsto X'\) defined in (23) coincides with the \(\kappa\)-color BBS time evolution defined in the introduction (see the following subsection for more discussion). Furthermore, in the following lemma, we show that first soliton length \(\lambda_1\) equals the maximum number of nonzero entries in the associated carrier process.

Lemma 2.2. Let \((X_t)_{t \geq 0}\) be a \(\kappa\)-color BBS trajectory such that \(X_0\) has finite support. Let \((\Gamma_t)_{t \geq 0}\) be the infinite capacity carrier process over \(X_0\). Then we have

\[
\lambda_1(X_0) = \max_{s \geq 0} \{\# \text{ of nonzero entries in } \Gamma_s\}.
\] (24)

The proof of Lemma 2.2 may be found in Section 9.
2.3. **Finite capacity carrier processes and soliton numbers.** In [KL18], it is shown that the row lengths of the invariant Young diagram of any \( \kappa \)-BBS trajectory can be extracted by running carrier processes of finite capacities, as we will summarize in this subsection. This will provide one of the key lemmas in the present paper.

First, fix an integer parameter \( c \geq 1 \) that we call *capacity*. Denote
\[
\mathcal{B}_c = \{ [x_1, \ldots, x_c] \in \{0, 1, \ldots, \kappa\}^c \mid x_1 \geq \cdots \geq x_c \},
\]
which can also be identified as the set of all \((1 \times c)\) semistandard tableaux with letters from \(\{0, 1, \ldots, \kappa\}\).

Define a map \( \Psi_c : \mathcal{B}_c \times \{0, 1, \ldots, \kappa\} \rightarrow \{0, 1, \ldots, \kappa\} \times \mathcal{B}_c, ([x_1, \ldots, x_c], y) \mapsto (y', [x'_1, \ldots, x'_c]) \) by the following ‘circular exclusion rule’:

(i) Suppose \( y > x_c \) and denote \( i^* = \min \{ i \geq 1 \mid x_i < y \} \). Then \( y' = x_{i^*} \)
and
\[
[x'_1, \ldots, x'_c] = [x_1, \ldots, x_{i^*-1}, y, x_{i^*+1}, \ldots, x_c].
\]

(ii) Suppose \( x_c \geq y \). Then \( y' = x_1 \)
and
\[
[x'_1, \ldots, x'_c] = [x_2, \ldots, x_c, y].
\]

Fix a \( \kappa \)-color BBS configuration \( X : \mathbb{N} \rightarrow \{0, 1, \ldots, \kappa\} \). Let \( \Gamma_0 = [0, \cdots, 0] \in \mathcal{B}_c \), and recursively define a new \( \kappa \)-color BBS configuration \( X' \) and a sequence \( (\Gamma_t)_{t \geq 0} \) of elements of \( \mathcal{B}_\infty \) by
\[
(X'(t + 1), \Gamma_{t+1}) = \Psi_c(\Gamma_t, X(t + 1)) \quad \forall t \geq 0.
\]

We call the sequence \( (\Gamma_t)_{t \geq 0} \) the *capacity-\( c \) carrier process over \( X \). See Figure 2 for an illustration.

It is well-known that, if the capacity \( c \geq 1 \) is large enough compared to the number of balls of color \( \geq 1 \) in the system, then the induced update map \( X \rightarrow X' \) agrees with the \( \kappa \)-color BBS time evolution (see, e.g., [HKT01]). Moreover, once \( c \) is large enough, the capacity-\( c \) carrier process is equivalent to the infinite capacity carrier process up to the number of 0’s in the carriers. Hence it follows that the infinite capacity carrier process induces the \( \kappa \)-color BBS time evolution, as claimed in the previous subsection. Furthermore, the following lemma, which is proven in [KL18], gives a carrier version of the BBS Greene-Kleitman invariant (18) for the row sums.

![Figure 1](image-url)
Lemma 2.3. Let \((X_t)_{t \geq 0}\) be a \(k\)-color BBS trajectory such that \(X_0\) has finite support. For each \(c \geq 1\), let \((\Gamma_{s,c})_{s \geq 0}\) denote the capacity-\(c\) carrier process over \(X_0\). Then for all \(k, t \geq 1\), we have
\[
\rho_1(\Lambda(X_0)) + \cdots + \rho_k(\Lambda(X_0)) = \sum_{s=1}^{\infty} \mathbf{1}(X_t(s) > \min \Gamma_{s-1;k}),
\]
where \(\min \Gamma_{s-1;k}\) denotes the smallest entry in \(\Gamma_{s-1;k}\).

Proof. See eq. (13) and Prop. 4.3 in [KL18].

3. PROOF OF THEOREM 1.1

In this subsection, we prove our first main result, Theorem 1.1. Let \(\Sigma^n\) be a uniformly chosen random permutation of the set \(\{1, 2, \cdots, n\}\), and let \(X^n = \Sigma^n(\{1, n\})\) be the random \(n\)-color BBS configuration induced from \(\Sigma^n\). Let \(\lambda_k(n) = \lambda_k(\Lambda(X^n))\) denote the length of the \(k\)th longest soliton in \(X^n\).

3.1. Proof of Theorem 1.1 for the columns. Denote by \(L(n)\) the length of a longest decreasing subsequence of \(\Sigma^n\). Our proof of Theorem 1.1 for the columns relies on Lemma 2.1 and the sharp asymptotic of longest decreasing subsequence of a uniform random permutation due to Baik, Deift, and Johansson [BDJ99].

Proof of Theorem 1.1 for the columns. Fix an integer \(k \geq 1\). It suffices to show that, almost surely,
\[
\lim_{n \to \infty} n^{-1/2} \sum_{i=1}^{k} \lambda_k(n) = 2\sqrt{k}.
\]
We start with a simple observation that a random permutation restricted to a subset (in particular, to an interval) is still a random permutation. Moreover, if we restrict a random permutation on multiple disjoint subsets, then these smaller permutations are independent (we don’t need the latter fact). These facts can be seen easily if we view a random permutation as a ranking among \(n\) i.i.d. \(\text{Uniform}(0,1)\) random variables. Hence by Lemma 2.1,
\[
\lambda_1(n) + \cdots + \lambda_k(n) \overset{d}{=} \max(L(n_1) + L(n_2) + \cdots + L(n_k) : n_1 + n_2 + \cdots + n_k = n).
\]
Baik, Deift, and Johansson [BDJ99] proved the following tail bounds for \(L_n\) (see also equations (1.7) and (1.8) in [BDJ99] or p. 149 in [Rom15]): There exist positive constants \(M, c, C\) such that for all \(m \geq 1\),
\[
\text{(Lower tail): } \quad \mathbb{P}(m^{-1/6}(L(m) - 2\sqrt{m}) \leq -t) \leq C \exp(-ct^3) \quad \text{for all } t \in [M, n^{5/6} - 2n^{1/3}];
\]
\[
\text{(Upper tail): } \quad \mathbb{P}(m^{-1/6}(L(m) - 2\sqrt{m}) \geq t) \leq C \exp(-ct^{3/5}) \quad \text{for all } t \in [M, n^{5/6} - 2n^{1/3}].
\]
Taking $t = (\log m)^2$, we obtain
\[ \mathbb{P}(|L(m) - 2\sqrt{m}| \geq (\log m)^2 m^{1/6}) \leq 2C \exp(-c(\log m)^{6/5}). \] (34)

Note that if $m \geq \varepsilon \sqrt{n}$, then for any fixed $d > 0$,
\[ \mathbb{P}(|L(m) - 2\sqrt{m}| \geq (\log m)^2 m^{1/6}) = O(n^{-d}). \] (35)

Now, denote the random variable in the right hand side of (31) by $X$. Fix $\varepsilon > 0$, and we write $X = \max(Y, Z)$, where
\begin{align*}
Y &= \max|L(n_1) + L(n_2) + \cdots + L(n_k) : n_1 + n_2 + \cdots + n_k = n, n_i \geq \varepsilon \sqrt{n} \text{ for all } i|,
Z &= \max|L(n_1) + L(n_2) + \cdots + L(n_k) : n_1 + n_2 + \cdots + n_k = n, n_i < \varepsilon \sqrt{n} \text{ for at least one } i|.
\end{align*}

(36) (37)

Since there are at most $n^k$ such partitions of $n$, by a union bound we have
\[ \mathbb{P}\left(|Y - 2\sqrt{kn}| > k(\log n)^2 n^{1/6}\right) = O(n^{-d}) \] (38)

for any fixed $d > 0$. The above deterministic optimization problem achieves its maximum when $n_1 = \cdots = n_k = n/k$, and so this yields
\[ \mathbb{P}\left(|Y - 2\sqrt{kn}| > k(\log n)^2 n^{1/6}\right) = O(n^{-d}) \] (39)

for any fixed $d > 0$.

Next, if $n_i < \varepsilon \sqrt{n}$, then we use the trivial upper bound $L(n_i) \leq n_i \leq \varepsilon \sqrt{n}$, otherwise if $n_i \geq \varepsilon \sqrt{n}$, we continue to use the tail bound for $|L(n_i) - 2\sqrt{m_i}|$ in (35). Hence
\[ \mathbb{P}\left(Z > 2\sqrt{(k-1)n} + k(\log n)^2 n^{1/6} + k\varepsilon \sqrt{n}\right) = O(n^{-d}), \] (40)

where the first term bounds the contribution from at most $k-1$ intervals of size $\geq \varepsilon \sqrt{n}$, second term is given by the BDJ tail bound in (35), and the last term gives a trivial bound for intervals of size $< \varepsilon \sqrt{n}$. Hence if we choose $\varepsilon < 2/(\sqrt{k} + \sqrt{k-1})$, then (39) and (40) give us
\[ \mathbb{P}(Z > Y) \leq \mathbb{P}\left(Y < 2\sqrt{kn} + k(\log n)^2 n^{1/6}\right) + \mathbb{P}\left(Z > 2\sqrt{(k-1)n} + k(\log n)^2 n^{1/6} + \frac{2\sqrt{n}}{\sqrt{k} + \sqrt{k-1}}\right) \] (41)
\[ = O(n^{-d}) \] (42)

for each fixed $d > 0$. Now note that, for each $t > 0$,
\[ \mathbb{P}\left(\left|\frac{1}{\sqrt{n}} \sum_{i=1}^{k} \lambda_i(n) - 2\sqrt{k}\right| > t\right) = \mathbb{P}\left(\max(Y, Z) - 2\sqrt{kn} > t\sqrt{n}\right) \leq \mathbb{P}\left(|Y - 2\sqrt{kn}| > t\sqrt{n}\right) + \mathbb{P}(Z > Y). \] (43) (44)

Hence by choosing $t = 1/\log n$, for any fixed $d > 0$, (39) and (41) yield
\[ \mathbb{P}\left(\left|\frac{1}{\sqrt{n}} \sum_{i=1}^{k} \lambda_i(n) - 2\sqrt{k}\right| > \frac{1}{\log n}\right) = O(n^{-d}). \] (45)

Then the assertion follows from Borel-Cantelli lemma. \[ \square \]
3.2. **Circular exclusion process and the row lengths.** In this subsection, we prove Theorem 1.1 for the rows. By Lemma 2.3, this can be done by analyzing the carrier process over the uniform random permutation $X^n$. Let $X := (U_t)_{t \geq 1}$ be a sequence of i.i.d. Uniform([0, 1]) random variables. For each capacity $k \geq 1$, we may define the carrier process $(\Gamma_t)_{t \geq 0}$ over $X$ using the same ‘circular exclusion rule’ we used to define the map $\Psi$ in Subsection 2.3. More precisely, denote $E_k = \{(x_1, \cdots, x_k) \in [0, 1]^k : x_1 \geq \cdots \geq x_k\}$. Define a map $\phi : \mathcal{E}_k \times [0, 1] \to \mathcal{E}_k$, $[x_1, \cdots, x_k, y] \mapsto [x'_1, \cdots, x'_k]$ by

(i) If $y > x_c$, then denote $i^* = \min\{i \geq 1 \mid x_i < y\}$ and let

$$[x'_1, \cdots, x'_k] = [x_1, \cdots, x_{i^*-1}, y, x_{i^*+1}, \cdots, x_c].$$

(ii) If $x_c \geq y$, then $[x'_1, \cdots, x'_k] = [x_2, \cdots, x_k, y]$. Then the $k$-point circular exclusion process $(\Gamma_t)_{t \geq 0}$ over $X$ is defined recursively by

$$\Gamma_{t+1} = \phi(\Gamma_t, U_{t+1}).$$

Note that $(\Gamma_t)_{t \geq 0}$ forms a Markov chain on state space $\mathcal{E}_k$. When $\Gamma_0 = [0, 0, \cdots, 0]$, we call $(\Gamma_t)_{t \geq 0}$ the carrier process over $X$ with capacity $k$.

![Figure 3. Evolution of a 4-point circular exclusion process. Each newly inserted point (black dot) annihilates the closest pre-existing point (red dot) in the counterclockwise direction.](image)

In the following lemma, which will be proved in Subsection 3.3, we show that the $k$-point circular exclusion process converges to its unique stationary measure $\pi$, which is the distribution of the order statistics from $k$ i.i.d. Uniform([0, 1]) variables.

**Lemma 3.1.** Fix an integer $k \geq 1$ and let $(\Gamma_t)_{t \geq 0}$ denote the $k$-point circular exclusion process with an arbitrary initial configuration.

(i) Let $\pi$ denote the distribution of the order statistics from $k$ i.i.d. uniform random variables on [0, 1]. Then $\pi$ is the unique stationary distribution for the Markov chain $(\Gamma_t)_{t \geq 0}$.

(ii) For each $t \geq 0$, let $\pi_t$ denote the distribution of $\Gamma_t$. Then $\pi_t$ converges to $\pi$ in total variation distance. More precisely,

$$d_{TV}(\pi_t, \pi) := \sup_{A \subset [0, 1]^k} |\pi_t(A) - \pi(A)| \leq \left(1 - \frac{1}{{k!}}\right)^{\lfloor t/k \rfloor},$$

where the supremum runs over all Lebesgue measurable subsets $A \subset [0, 1]^k$.

Now we derive Theorem 1.1 for the row asymptotics.

**Proof of Theorem 1.1 for the rows.** Let $X = (U_t)_{t \geq 1}$ be a sequence of i.i.d. Uniform([0, 1]) random variables, $\Sigma^n$ be the random permutation on $[n]$ induced by $U_1, \cdots, U_n$, and $X^n = \Sigma^n 1([1, n])$ be the random $n$-color BBS configuration as defined at (4). Fix an integer $k \geq 1$ and let $(\Gamma_t)_{t \geq 0}$ be the carrier process over $X$. Also, let $(\Gamma_t)_{t \geq 0}$ be the capacity-$k$ carrier process over $X^n$ as defined in Subsection 2.3. By construction, for each $1 \leq s \leq n$, we have

$$\mathbf{1}(X^n(s) > \min \Gamma_{s-1}) = \mathbf{1}(U_s > \min \Gamma_{s-1}).$$

(49)
Thus according to Lemma 2.3, almost surely,
\[ n^{-1} \left( \rho_1(\Lambda(X^n)) + \cdots + \rho_k(\Lambda(X^n)) \right) = n^{-1} \sum_{i=1}^{n} I(U_i > \min \Gamma_{s-1}). \] (50)

By Lemma 3.1 and Markov chain ergodic theorem, almost surely,
\[ \lim_{n \to \infty} n^{-1} \left( \rho_1(\Lambda(X_0)) + \cdots + \rho_k(\Lambda(X_0)) \right) = P(U_{k+1} > \min(U_1, \cdots, U_k)) = \frac{k}{k+1}. \] (51)

Then the assertion follows. \(\square\)

3.3. **Stationarity and convergence of the circular exclusion process.** We prove Lemma 3.1 in this subsection. We will assume the stationarity of the circular exclusion process as asserted in the following proposition, which will be proved at the end of this subsection.

** Proposition 3.2.** Fix an integer \(k \geq 1\) and let \(\pi\) denote the distribution of the order statistics from \(k\) i.i.d. uniform random variables on \([0, 1]\). Then \(\pi\) is a stationary distribution of the \(k\)-point circular exclusion process.

**Proof of Lemma 3.1.** For the convergence, we use a standard coupling argument. Namely, fix arbitrary distributions \(\pi_0\) and \(\tilde{\pi}_0\) on \(\mathcal{C}_k\) and let \((\Gamma_t)_{t \geq 0}\) and \((\tilde{\Gamma}_t)_{t \geq 0}\) be two \(k\)-point circular exclusion processes such that \(\Gamma_0\) and \(\tilde{\Gamma}_0\) are independently drawn from \(\pi_0\) and \(\tilde{\pi}_0\), respectively. We couple the two processes by using the same sequence of i.i.d. Uniform([0, 1]) variables \((U_t)_{t \geq 1}\) to evolve them simultaneously. Let \(\tau = \inf\{t \geq 0 \mid \Gamma_t = \tilde{\Gamma}_t\}\) denote the first meeting time of the two chains (see Figure 4). By the coupling, \(\Gamma_s = \tilde{\Gamma}_s\) and \(s < t\) imply \(\Gamma_t = \tilde{\Gamma}_t\). A standard argument shows
\[ d_{TV}(\pi_t, \tilde{\pi}_t) \leq P(\Gamma_t \neq \tilde{\Gamma}_t) = P(\tau > t), \] (52)

where \(\pi_t\) and \(\tilde{\pi}_t\) denote the distributions of \(\Gamma_t\) and \(\tilde{\Gamma}_t\). We claim that
\[ P(\tau > t) \leq P(\Gamma_0 \neq \tilde{\Gamma}_0) \left( 1 - \frac{1}{k!} \right)^{|t/k|}. \] (53)

According to Proposition 3.2, this will imply Lemma 3.1 by choosing \(\tilde{\pi}_0 = \pi\).

![Figure 4](image-url)

**Figure 4.** Joint evolution of two 3-point circular exclusion processes. Newly inserted point annihilates one of the closest pre-existing point in the counterclockwise direction. Blue (resp., red) dots represent points that are shared (resp., not shared) in both processes. The two chains meet after the fifth transition.

To bound the tail probability of meeting time \(\tau\), we will show that two circular exclusion processes 'synchronize' after \(k\) steps with probability at least \(1/k!\), in the sense that
\[ P(\Gamma_{t+k} = \tilde{\Gamma}_{t+k} \mid \Gamma_t \neq \tilde{\Gamma}_t) \geq \frac{1}{k!} \quad \text{for all } t \geq 0. \] (54)
Then the claim (53) follows since
\[ P(\tau > Nk) = P(\Gamma_{NK} \neq \hat{\Gamma}_{NK} | \Gamma_0 \neq \hat{\Gamma}_0) P(\Gamma_0 \neq \hat{\Gamma}_0) \]
\[ \leq P(\Gamma_0 \neq \hat{\Gamma}_0) \prod_{i=1}^{N} P(\Gamma_{ik} \neq \hat{\Gamma}_{ik} | \Gamma_{(i-1)k} \neq \hat{\Gamma}_{(i-1)k}) \]
\[ \leq P(\Gamma_0 \neq \hat{\Gamma}_0) \left( 1 - \frac{1}{k!} \right)^N. \]

We begin with following simple observation for a sufficient condition of meeting. Let \( X = (U_t)_{t \geq 1} \) be a sequence of i.i.d. Uniform\((0,1)\) variables. Fix \( t \geq 1 \) and let \( \Gamma_t = [x_1, \ldots, x_k] \) and \( \hat{\Gamma}_t = [\tilde{x}_1, \ldots, \tilde{x}_k] \) be arbitrary elements of \( \mathcal{C}_k \). Superpose the two \( k \)-point configurations into a one \( 2k \)-point configuration \( 0 \leq y_1 \leq y_2 \leq \cdots \leq y_{2k} \leq 1 \). For a special case, suppose \( y_{2k} < 1 \). Observe that on the event \( \{y_{2k} < U_{t+k} < \cdots < U_{t+1} \leq 1\} \), we have
\[ \Gamma_{t+k} = [U_{t+1}, U_{t+2}, \ldots, U_{t+k}] = \hat{\Gamma}_{t+k}, \]
as all of the \( k \) points in \( \Gamma_t \) and \( \Gamma_t \) will be successively annihilated from the largest to the smallest by inserting \( U_{t+1}, \ldots, U_{t+k} \).

For the general case, regard each \( U_i \) as a uniformly chosen point from the unit circle \( S^1 \). Then the \( 2k \) points \( y_1, \ldots, y_{2k} \) will divide \( S^1 \) into disjoint arcs of lengths, say, \( \ell_1, \ldots, \ell_m \), for some \( 2 \leq m \leq 2k \). If the points \( U_{t+1}, \ldots, U_{t+k} \) are strictly decreasing in the counterclockwise order within one of the \( m \) arcs, then by circular symmetry and a similar observation, we will have \( \Gamma_{t+k} = \hat{\Gamma}_{t+k} \). Noting that
\[ P\left( U_{t+1}, \ldots, U_{t+k} \text{ are strictly decreasing in the counterclockwise order within an arc of length } \ell \right) = \frac{\ell^k}{k!} \]
and \( \ell_1 + \cdots + \ell_m = 1 \), Jensen’s inequality yields
\[ P(\Gamma_{t+k} = \hat{\Gamma}_{t+k} | \Gamma_t = [x_1, \ldots, x_k], \hat{\Gamma}_t = [\tilde{x}_1, \ldots, \tilde{x}_k]) \geq \sum_{i=1}^{m} \frac{\ell_i^k}{k!} \geq \frac{1}{k!} (\ell_1 + \cdots + \ell_m)^k = \frac{1}{k!}. \]
This shows the assertion. \( \square \)

**Proof of Proposition 3.2.** We first show \( \pi \) is a stationary distribution for the Markov chain \( (\Gamma_s)_{s \geq 0} \). Let \( X_{(1)} < X_{(2)} < \cdots < X_{(k)} \) be the order statistics from \( k \) i.i.d. uniform RVs on \([0,1]\). Let \( Y \) be an independent Uniform\((0,1)\) random variable. After a new point \( Y \) is inserted to the preexisting list of \( k \) points \( X_{(1)} > X_{(2)} > \cdots > X_{(k)} \), the updated list of points will be
\[ X_{(1)} > \cdots > X_{(I-1)} > Y > X_{(I+1)} > \cdots > X_{(k)}, \]
where \( I \in \{1, 2, \ldots, k\} \) is the random index such that \( Y \in (X_{(I-1)}, X_{(I)}) \). For \( I = 1 \), the interval \( (X_{(1)}, X_{(I)}) \) denotes the union of \((0, X_{(1)}) \) and \((X_{(k)}, 1) \). In this case, the point \( X_{(1)} \) gets deleted and \( Y \) is added as the smallest or largest point depending on which sub-intervals it falls.

We claim that (61) is still the order statistics from \( k \) i.i.d. uniforms on \([0,1]\), which would prove that the distribution of \( k \) i.i.d. uniform points remains invariant under the transition rule. To show this, take a bounded test function \( f : [0,1]^k \to \mathbb{R} \). First we write
\[ E \left[ f(X_{(1)}, \ldots, X_{(I-1)}, Y, X_{(I+1)}, \ldots, X_{(k)}) \right] \]
\[ = \sum_{i=1}^{k} E \left[ f(X_{(1)}, \ldots, X_{(i-1)}, Y, X_{(i+1)}, \ldots, X_{(k)}) 1_{Y \in (X_{(i)}, X_{(i+1)})} \right] \]
\[ = \sum_{i=2}^{k} \frac{1}{k!} \int_{z_1 > \cdots > z_{i-1} > y > z_i > z_{i+1} > \cdots > z_k} f(z_1, \ldots, z_{i-1}, y, z_{i+1}, \ldots, z_k) \, dz_1 \cdots dz_k \, dy \]
\[ + \frac{1}{k!} \int_{z_1 > z_2 > \cdots > z_k > y \text{ or } z_1 > \cdots > z_k > y > z_1} f(z_2, \cdots, z_k, y) \, dz_1 \cdots dz_k \, dy. \]  

Integrated out \( z_i \), the integral becomes
\[ = \sum_{i=2}^{k} \frac{1}{k!} \int_{z_1 > \cdots > z_{i-1} > y > z_{i+1} > \cdots > z_k} f(z_1, \cdots, z_{i-1}, y, z_{i+1}, \cdots, z_k)(y - z_{i-1}) \, dz_1 \cdots z_{i-1} z_{i+1} \cdots dz_k \, dy \]

\[ + \frac{1}{k!} \int_{z_2 > \cdots > z_k > y} f(y_1, z_2, \cdots, z_k)(1 - z_k) \, dz_1 \cdots dz_k \, dy \]

\[ + \frac{1}{k!} \int_{z_2 > \cdots > z_k > y} f(y_1, z_2, \cdots, z_k)(y - 0) \, dz_1 \cdots dz_k \, dy. \]

We then rename \( y \) as \( z_i \) in the integral to obtain
\[ = \frac{1}{k!} \int_{z_1 > \cdots > z_k} f(z_1, \cdots, z_k) \left[ (z_1 - 0) + \left( \sum_{k=2}^{k} z_i - z_{i-1} \right) + (1 - z_k) \right] \, dz_1 \cdots dz_k \]

\[ = \mathbb{E} \left[ f(X_1, \cdots, X_{(I-1)}, X_I, X_{(I+1)}, \cdots, X_k) \right]. \]

This shows the assertion. \( \square \)

4. DECOMPOSING \( \kappa \)-COLOR CARRIER PROCESS INTO EXCURSIONS

Throughout this section, we fix a probability distribution \( \mathbf{p} = (p_0, p_1, \cdots, p_\kappa) \) on \( \{0, 1, \cdots, \kappa\} \), and let \( (\Gamma_t)_{t \geq 0} \) be the infinite capacity carrier process over the i.i.d. configuration \( X^\mathbf{p} \) as introduced in Subsection 2.2.

Unlike the circular exclusion process on continuum color space \([0, 1]\) we analyzed for the random permutation model in Subsection 3.2, it is important and will be conceptually more convenient to keep track of the multiplicity \( m_i(\Gamma_t) \) of the balls of color \( 1 \leq i \leq \kappa \) for the \( \kappa \)-color independence model. Hence we consider the Markov chain
\[ \Sigma_t := (m_1(\Gamma_t), \cdots, m_\kappa(\Gamma_t)) \in (\mathbb{Z}_{\geq 0})^\kappa, \]

which is determined by infinite capacity carrier process \( \Gamma_t \) (and vice versa). See Figure 5 for an illustration.

Let \( \mathbf{0} = (0, 0, \cdots, 0) \in (\mathbb{Z}_{\geq 0})^\kappa \) denote the origin, and write
\[ M_n = \sum_{i=1}^{n} 1(\Sigma_t = \mathbf{0}) \]

for the number of visits of \( \Sigma_t \) to \( \mathbf{0} \) during \([1, n]\). For each \( k \geq 1 \), let \( T_k \) denote the time of the \( k \)th visit of the chain \( \Sigma_t \) to \( \mathbf{0} \) and set \( T_0 = 0 \). We say that the trajectories of \( \Sigma_t \) restricted to the time intervals \([T_{k-1}, T_k]\) between consecutive visits to \( \mathbf{0} \) are its excursions. It is important to note that distinct excursions are independent due to the strong Markov property. Also note that \( M_n \) defined at (73) equals the number of complete excursions of the carrier process during \([1, n]\). We will define the height of the carrier at time \( t \) by
\[ \|\Sigma_t\|_1 = \sum_{i=1}^{\kappa} m_i(\Gamma_t), \]

(74)
which equals the number of balls of positive color that the carrier possesses at time $t$. Define the $k$th excursion height $h_k$ and height of the final meander $r_n$ by

$$h_k = \max_{T_{k-1} \leq t \leq T_k} \| \Sigma'_t \|_1, \quad r_n = \max_{T_{M_n} \leq t \leq n} \| \Sigma'_t \|_1.$$  \hspace{1cm} (75)

Following [LLP17], we are interested in the order statistics for $h_1, \ldots, h_{M_n}$, which we denote by $h_1(n) \geq h_2(n) \geq \cdots \geq h_{M_n}(n)$. Furthermore, let $h_{1:m} \geq h_{2:m} \geq \cdots \geq h_{m:m}$ denote the order statistics of the first $m$ excursion heights $h_1, \ldots, h_m$.

The main result in this section is the following lemma.

**Lemma 4.1.** Let $h_i(n)$, $h_j(n)$, and $M_n$ be as before. The following hold for each $n \geq 1$.

(i) $\max(h_1, \ldots, h_{M_n}) \leq h_1(n) \leq \max(h_1, \cdots, h_{M_n+1})$.

(ii) For each $j \geq 1$, $h_j(n) \geq h_j(n)$.

Now we prove Lemma 4.1. For the proof of (ii), we rely on the finite-capacity carriers (see Subsection 2.3) and Lemma 2.3. We need an additional combinatorial observation about the ‘coupling’ between the carrier processes of capacity $c$ and $c+1$ over the same BBS configuration, which is stated below.

**Proposition 4.2.** Let $X : \mathbb{N} \to \mathbb{Z}_{c+1}$ be any $\kappa$-color BBS configuration with finite support. Denote by $(\Gamma_{t,c})_{t \geq 0}$ and $(\Gamma_{t,c+1})_{t \geq 0}$ the carrier processes over $X$ with finite capacities $c$ and $c+1$, respectively. Then for any $t \geq 0$, $\Gamma_{t,c}$ is obtained by omitting a single entry of $\Gamma_{t,c+1}$.

**Proof.** See Section 9. \hspace{1cm} □

**Proof of Lemma 4.1.** Note that by Lemma 2.2, we have

$$\lambda_1(n) = \max_{0 \leq t \leq n} \| \Sigma'_t \|_1 = \max(h_1, h_2, \cdots, h_{M_n}, r_n).$$  \hspace{1cm} (76)

Moreover, since $r_n \leq h_{M_n+1}$, this immediately gives Lemma 4.1 (i).

For (ii) it is enough to show the assertion for a deterministic $\kappa$-color BBS configuration, which we will denote $X$. Let for each integer $k \geq 1$, let $p_k$ and $\lambda_k$ denote the $k$th row and column length of the corresponding Young diagram $\Lambda(X)$, respectively. Let $(\Gamma_{t})_{t \geq 0}$ denote the infinite capacity carrier process over $X$, and let the excursions $(SG_{t})_{t \geq 0}$ be denoted $h_1, h_2, \cdots, h_m$ from left to right. Also denote by...

\hspace{1cm} Figure 5. State space diagram for the Markov chain $SG_t = (m_1(\Gamma_t), m_2(\Gamma_t))$ for $\kappa = 2$. Red arrows illustrate the transition kernel at the ‘interior’ and ‘boundary’ points in the state space. A single excursion of height $h_1 = 8$ is shown in blue.
\(h_{1;m} \geq \cdots \geq h_{m;m}\) their order statistics. Moreover, it is easy to see that the capacity \(c := h_{1;m}\) carrier process over \(X\) reproduces the infinite capacity one \((\Gamma_t) \in \mathbb{R}_+\) in the sense that

\[m_i(\Gamma_t) = m_i(\Gamma_{t;c}) \quad \forall 1 \leq i \leq k \text{ and } t \geq 0. \tag{77}\]

Hence \(\Gamma_{t;c}\) has the same excursions as \(\Gamma_t\).

Fix \(1 \leq j \leq c\). We wish to show \(\lambda_j \geq h_{j;m}\). This will follow by showing

\[\rho h_{j;m} \geq j, \tag{78}\]

since this means there are at least \(j\) columns in the Young diagram \(\Lambda(X)\) that has length at least \(h_{j;m}\). To show this, for any \(1 \leq i \leq c\), we use Lemma 2.3 to write

\[\rho_i = \sum_{s=1}^{\infty} \left[ 1(X(s) > \min \Gamma_{s-1;i}) - 1(X(s) > \min \Gamma_{s-1;i-1}) \right] \tag{79}\]

\[= \sum_{k=0}^{\infty} \left[ \sum_{T_k \leq s \leq T_{k+1}} 1(X(s) > \min \Gamma_{s-1;i}) - 1(X(s) > \min \Gamma_{s-1;i-1}) \right], \tag{80}\]

where \(T_k\) denotes the \(k\)th return time to the origin of the carrier process \(\Sigma_{t;c}\) for each \(k \geq 1\). According to Proposition 4.2, the summands in (79) are nonnegative. Thus it is enough to show that there are at least \(j\) distinct \(k\)'s so that the corresponding summand in (80) equals 1.

First, we observe that during the \(k\)th excursion interval \([T_k, T_{k+1}]\), the three carriers \(\Gamma_{t;i-1}, \Gamma_{t;i}, \text{ and } \Gamma_{t;c}\) have the same trajectory (except the number of 0's) until the first time that the carrier \(\Gamma_{t;i}\) of capacity \(i\) becomes full of \(i\) balls of positive colors. Indeed, \(\Sigma_{T_k;i} = \Sigma_{T_k;i-1} = 0\) by Proposition 4.2 since \(\Sigma_{T_k;c} = 0\) by definition of \(T_k\). Moreover, all three carriers are evolved by the same input given by \(X\). Hence they must have the same trajectory until the first time that the lowest capacity carrier has to be overloaded.

Second, we observe that for each \(1 \leq i \leq c\), the following bound holds:

\[h_{k+1} \land i \leq \sum_{T_k \leq s \leq T_{k+1}} 1(X(s) > \min \Gamma_{s-1;i}) \leq i. \tag{81}\]

The upper bound comes from capacity constraint. For the lower bound, first note that the height \(\|\Sigma_{t;i}\|_1\) of the capacity- \(t\) carrier \(\Gamma_{t;i}\) must reach \(k_{k+1} \land i\) by the observation we made in the previous paragraph. Moreover, note that in order for the maximum capacity carrier \(\Gamma_{t-1;c}\) increase its height \(\|\Sigma_{t-1;c}\|_1 = m_1(\Gamma_{t-1;c}) + \cdots + m_k(\Gamma_{t-1;c})\) by one, it is necessary that \(\Gamma_{t;i}\) has at least one zero entry and the color of the new ball \(X(t+1)\) to be positive. This event is picked up by the indicator \(1(X(t) > \min \Gamma_{t-1;c})\). Hence, again by the observation in the previous paragraph, the indicator in (81) will be 1 every time the height of \(\Sigma_{t-1;i}\) increase as long as its height remains \(\leq h_{k+1} \lor i\). This shows the lower bound in (81).

Now we finish the proof. We choose \(i = h_{j;m}\) in (80). Then it suffices to show that the \(k\)th summand in (80) with \(i = h_{j;m}\) is at least 1 whenever \(h_k \geq h_{j;m}\). Suppose \(h_k \geq h_{j;m}\). First note that, by the second observation above, we may choose \(t^* \in (T_k, T_{k+1})\) such that \(\|\Sigma_{t^*};i\|_1 = i, \|\Sigma_{t^*};i\|_1 = i - 1, \text{ and } t^*\) is as small as possible. Then by the first observation above, \(\Sigma_{t^*};i-1 = \Sigma_{t^*};i-1\), and by the circular exclusion rule, we must have

\[X(t^*) \geq \text{smallest positive entry of } \Gamma_{t^*};i, \tag{82}\]

since otherwise the new ball \(X(t^*)\) will replace some existing ball in \(\Gamma_{t^*};i\) so that \(\|\Sigma_{t^*};i\|_1 = \|\Sigma_{t^*};i\|_1 = \|\Sigma_{t^*};i-1\|_1 = i\), which is a contradiction. Since \(\Sigma_{t^*};i-1 = \Sigma_{t^*};i\), the smallest positive entry of \(\Gamma_{t^*};i\) equals \(\min \Gamma_{t^*};i-1;1\). Thus

\[1(X(s) > \min \Gamma_{t^*};i) - 1(X(s) > \min \Gamma_{s-1;i-1}) = 1. \tag{83}\]

This shows the assertion. \(\square\)
5. PROOF OF THEOREM 1.2 IN THE SUBCRITICAL PHASE

We prove Theorem 1.2 (i) in this section. Throughout this section, we fix a probability distribution \( p = (p_0, p_1, \cdots, p_\kappa) \) on \([0, 1, \cdots, \kappa]\), and let \((\Gamma_t)_{t \geq 0}\) be the infinite capacity carrier process over \(X^p\) as introduced in Subsection 2.2.

5.1. Stationarity and convergence of the subcritical carrier process. Define a probability distribution \( \pi \) on \((\mathbb{Z}_{\geq 0})^\kappa\) by

\[
\pi(n_1, n_2, \cdots, n_\kappa) = \prod_{i=1}^{\kappa} \left( 1 - \frac{p_i}{p_0} \right) \left( \frac{p_i}{p_0} \right)^{n_i}.
\]

This is a valid probability distribution on \((\mathbb{Z}_{\geq 0})^\kappa\) for \(p_0 > \max(p_1, \cdots, p_\kappa)\) since

\[
\sum_{n_1=0}^{\infty} \cdots \sum_{n_\kappa=0}^{\infty} \prod_{i=1}^{\kappa} \left( \frac{p_i}{p_0} \right)^{n_i} = \prod_{i=1}^{\kappa} \left( 1 - \frac{p_i}{p_0} \right)^{-1} \in (0, \infty).
\]

The main result in this subsection is the following:

**Lemma 5.1.** Suppose \(p_0 > \max(p_1, \cdots, p_\kappa)\). Then \(\Sigma_\kappa = (m_1(\Gamma_t), \cdots, m_\kappa(\Gamma_t))\) is an irreducible and aperiodic Markov chain with \(\pi\) as its unique stationary distribution. Furthermore, if we denote its distribution at time \(t\) by \(\pi_t\), then

\[
\lim_{n \to \infty} d_{TV}(\pi_t, \pi) = 0.
\]

**Proof.** For its aperiodicity, it is enough to observe that

\[
P(\Gamma_{t+1} = [0, 0, \cdots] \mid \Gamma_t = [0, 0, \cdots]) = p_0 > 0.
\]

For its irreducibility, fix \(x, y \in \mathcal{B}_\infty\) and write \(y = [y_1, y_2, \cdots]\). Since all elements of \(\mathcal{B}_\infty\) have finite support, there exists an integer \(m \geq 1\) such that \(x(i) \equiv 0\) and \(y(i) \equiv 0\) for all \(i \geq m\). Then note that

\[
P(\Gamma_{t+2m} = y \mid \Gamma_t = x) \geq P(X^p(t+1) = 0, \cdots, X^p(t+m) = 0, X^p(t+m+1) = y_1, \cdots, X^p(t+2m) = y_m)
\]

\[
= p_0^m p_{y_1} \cdots p_{y_m} > 0.
\]

Since \(x, y \in \mathcal{B}_\infty\) were arbitrary, this shows the Markov chain \(\Sigma_\kappa\) is also irreducible.

Next, we show that \(\pi\) is a stationary distribution for \((\Sigma_\kappa)_{t \geq 0}\). The uniqueness of stationary distribution and total variation distance convergence will then follow from general results of countable state space Markov chain theory (see, e.g., [LP17, Thm. 21.13 and Thm. 21.16]). We work with the original carrier process \(\Gamma_t\). For each \(x \in \mathcal{B}_\infty\) and \(i \in \{0, 1, \cdots, \kappa\}\), denote

\[
\exp(\text{wt}(x)) = \prod_{i=1}^{\kappa} \left( \frac{p_i}{p_0} \right)^{m_i(x)}, \quad \exp(\text{wt}(i)) = \frac{p_i}{p_0}.
\]

Recall the definition of the map \(\Psi : \mathcal{B}_\infty \times \{0, 1, \cdots, \kappa\} \to \{0, 1, \cdots, \kappa\} \times \mathcal{B}_\infty\) defined after the statement of Lemma 5.1. Note that for each pair \((x, y) \in \mathcal{B}_\infty \times \{1, 2, \cdots, \kappa\}\) and \((y', x') \in \{1, 2, \cdots, \kappa\} \times \mathcal{B}_\infty\) such that \(\Psi(x, y) = (y', x')\), we have

\[
\exp(\text{wt}(x)) \exp(\text{wt}(y)) = \exp(\text{wt}(y')) \exp(\text{wt}(x')).
\]

Indeed, the total number of each letter \(1 \leq i \leq \kappa\) in both pairs \((x, y)\) and \((y', x')\) are the same so the equation holds.
Now, observe that for each fixed \( x' \in \mathcal{B}_\infty \), \( \Psi \) gives a bijection between \( \{0,1,\cdots,k\} \times \{x'\} \) and its inverse image under \( \Psi \). If we denote the second coordinate of \( \Psi \) by \( \Psi_2 \), then this yields
\[
\sum_{(x,y) \in \mathcal{B}_\infty \times \{0,1,\cdots,k\}} \exp(\text{wt}(x)) \exp(\text{wt}(y)) = \sum_{(x,y) \in \mathcal{B}_\infty \times \{0,1,\cdots,k\}} \exp(\text{wt}(y')) \exp(\text{wt}(x'))
\]
\[
= \exp(\text{wt}(x')) \sum_{y' \in \{0,1,\cdots,k\}} \exp(\text{wt}(y'))
\]
\[
= \exp(\text{wt}(x')).
\]
Dividing both sides by
\[
\sum_{x \in \mathcal{B}_\infty} \exp(\text{wt}(x)) = \sum_{n_1=0}^{\infty} \cdots \sum_{n_k=0}^{\infty} \prod_{i=1}^{k} \frac{p_i}{p_0} \prod_{i=1}^{k} (1 - \frac{p_i}{p_0})^{-1} > 0,
\]
we get
\[
\sum_{(x,i) \in \mathcal{B}_\infty \times \{0,1,\cdots,k\}} \pi(m_1(x),\cdots,m_k(x)) p_i = \pi(m_1(x'),\cdots,m_k(x')).
\]
This shows that \( \pi \) is a stationary distribution of the Markov chain \( (S\Gamma_t)_{t \geq 0} \), as desired. \( \square \)

**Remark 5.2.** The statement and proof of Lemma 5.1 are reminiscent of [KL18, Thm. 1], where the authors show that for all \( p = (p_0,\cdots,p_k) \), the (finite) capacity-\( c \) carrier process over \( X^p \) is irreducible with unique stationary distribution
\[
\pi_c(C) = \frac{1}{Z_c} \prod_{i=0}^{k} p_i^{m_i(C)},
\]
where \( Z_c \) denotes the partition function. In fact, their result applies to more general finite-capacity carriers whose state space is the set \( \mathcal{B}_c^{(a)}(\kappa) \) of all semistandard tableaux of rectangular shape \( (c \times a) \) with letters from \( \{0,1,\cdots,k\} \). In this general case, the partition function \( Z_c = Z_c^{(a)}(\kappa,p) \) is identified with the Schur polynomial associated with the \( (a \times c) \) Young tableau with constant entries \( c \) and parameters \( p_0, p_1, \cdots, p_k \).

**Remark 5.3.** Fix arbitrary \( p = (p_0,p_1,\cdots,p_k) \) and let \( (\Gamma_t)_{t \geq 0} \) be the infinite-capacity carrier process over \( X^p \). Define a functional \( \phi : \mathcal{B}_\infty \to \mathbb{R} \) by
\[
\phi(x) = \prod_{i=1}^{k} \left( \frac{p_0}{p_i} \right)^{m_i(x)}.
\]
Modifying the proof of Lemma 5.1, one can show that \( \phi(\Gamma_t) \) behaves as a martingale whenever \( m_i(\Gamma_t) \geq 1 \) for all \( 1 \leq i \leq k \), and as a sub- or super-martingale otherwise depending on the colors \( i \) such that \( \Gamma_t \) is missing.

More precisely, recall the definition of the map \( \Psi = (\Psi_1,\Psi_2) : \mathcal{B}_\infty \times \{0,1,\cdots,k\} \times \{0,1,\cdots,k\} \times \mathcal{B}_\infty \) defined in Subsection 2.2. Note that \( \Psi_1(x,i) \) and \( \Psi_2(x,i) \) respectively equal the color of the ball that pops out and the next carrier state when the carrier of state \( x \) encounters ball of color \( i \). Then we have
\[
\mathbb{E}[\phi(\Gamma_{t+1}) | \mathcal{F}_t] = \phi(\Gamma_t) \left( \sum_{i=0}^{k} p_{\Psi_1(\Gamma_t,i)} \right),
\]
where \( \mathcal{F}_t \) denotes the information up to time \( t \).
5.2. Order statistics of the excursion heights. According to Lemma 5.1, the Markov chain \((S_t^\Gamma)_{t \geq 0}\) in the subcritical phase \(p_0 > \max(p_1, \ldots, p_\kappa)\) will visit the origin \(0 := (0, 0, \ldots, 0) \in (\mathbb{Z}_{\geq 0})^\kappa\) infinitely often with finite mean excursion time \((\pi(0))^{-1}\). Namely, the number \(M_n\) of visits of \(S_t^\Gamma\) to \(0\) during \([1, n]\) (defined in (73)) satisfies
\[
\frac{M_n}{n} \to \pi(0) = \prod_{i=1}^\kappa \left(1 - \frac{p_i}{p_0}\right)^{-1} \quad \text{a.s. as } n \to \infty
\]
by Lemma 5.1 and the Markov chain ergodic theorem. According to Lemma 4.1 (i), the first soliton length \(\lambda_1(n)\) is essentially the same as the maximum of the first \(M_n\) excursion heights of the carrier process. Roughly speaking, there are \(M_n \sim \pi(0)n\) excursions of heights with exponential tail, so their maximum should behave as \(O(\log n)\).

To make this estimate more precise, following [LLP17], we analyze the order statistics of the excursion heights of the carrier process during \([1, n]\). For this, let \(h_{1:m} \leq h_{2:m} \leq \cdots \leq h_{m:m}\) denote the order statistics of the first \(m\) excursion heights \(h_1, \cdots, h_m\). The strong Markov property ensures that these excursion heights are i.i.d., so we have
\[
P(h_{j:m} \leq x) = \sum_{\ell=0}^{j-1} \binom{m}{\ell} P(h_1 \leq x)^{m-\ell} P(h_1 > x)^{\ell}, \quad j = 1, \cdots, m.
\]
In the simplest case \(\kappa = 1\), the distribution function of the excursion height \(h_1\) follows from the standard gambler’s ruin probability and is given by
\[
P(h_1 \leq x) = \left(1 - \frac{1 - 2p}{\theta|x| + 1 - 1}\right) \mathbf{1}_{[0, \infty)}(x),
\]
where \(\theta = p_0 / p_1\) (see [LLP17, Sec. 4]). For the general \(\kappa \geq 1\) case, computing the distribution function of \(h_1\) amounts to solving a high dimensional gambler’s ruin problem (see the illustration for \(\kappa = 2\) in Figure 5).

A standard martingale argument for the gambler’s ruin problem for \(\kappa = 1\) does not seem to readily apply for the general \(\kappa \geq 2\) dimensional case. The essential issue is that the subcritical carrier process for \(\kappa \geq 2\) may have a positive drift on a boundary of its state space. For instance, consider the \(\kappa = 2\) carrier process as in Figure 5. Assuming \(p_0 > p_1, p_2\), it might be plausible that there is a negative drift to the total sum
\[
m_1(\Gamma_t) + m_2(\Gamma_t) = \text{‘height’ of the chain at time } t.
\]
Indeed the subcritical condition ensures this in the interior and the right boundary of the state space, but this is not necessarily true along the left boundary where \(m_1(t) = 0\) (e.g., consider \(p = (0.4, 0.3, 0.3)\)).

In order to establish exponential tail bound on the excursion heights, we use an inductive argument that shows geometric ergodicity of the Markov chain \(S_t^\Gamma\) in the subcritical regime \(p_0 > \max(p_1, \cdots, p_\kappa)\). From this we deduce that the excursion length has finite exponential moment, as stated in the following lemma.

Lemma 5.4. Suppose \(p_0 > \max(p_1, \cdots, p_\kappa)\) and let \(\tau_0\) denote the first return time of the Markov chain \(S_t^\Gamma\) to the origin. Then there exists some constant \(\lambda > 1\) such that
\[
\mathbb{E}[\lambda^\tau_0] < \infty.
\]
In order to maintain the flow of the paper, we delay the proof of Lemma 5.4 until Section 8. Assuming this lemma, we can derive upper and lower exponential tail bound on the excursion heights.
Proposition 5.5. Suppose \( p_0 > \max(p_1, \cdots, p_\kappa) \). Then there exists some constants \( 0 < C_2 \leq C_1 \) and \( \theta_2 \geq \theta_1 \geq (p_0/p_\kappa) \) such that for any \( x \geq 0 \),
\[
\frac{C_1}{\theta_1^x} \leq \mathbb{P}(h_1 > x) \leq \frac{C_2}{\theta_2^x}.
\]

Proof. Let \( \tau_0 \) denote the excursion length of the Markov chain \( \Gamma \) and let \( \lambda > 1 \) be the constant as in Lemma 5.4. Then since the height of the chain \( \Gamma \) changes at most by 1 in each step, Markov’s inequality gives
\[
\mathbb{P}(h_1 > x) \leq \mathbb{P}(\tau_0 > x) \leq \mathbb{P}(\lambda^{\tau_0} > \lambda^x) \leq \lambda^{-x} \mathbb{E}[^{\tau_0}].
\]
This shows the upper bound in the assertion.

For the lower bound, we will show that
\[
\mathbb{P}(h_1 > x) \geq \frac{p_0 - p_\kappa}{p_0 + p_\kappa} \frac{1}{[x] + 1}
\]
for all \( x \geq 0 \). For each \( 1 \leq i \leq \kappa \), we let \( \tau^{(i)} \) denote the first time that \( m_\kappa(\Gamma_t) = 0 \) for \( t \geq 1 \). Also, let \( h^{(i)} = \max_{1 \leq j \leq \tau^{(i)}} m_\kappa(\Gamma_t) \) denote the maximum number of color \( \kappa \) balls in the carrier during \( [0, \tau^{(i)}] \).

Then since \( \tau \geq \tau^{(i)} \) and \( n \| \Gamma \| \geq m_\kappa(\Gamma_t) \), it follows that \( h_1 \geq h^{(i)} \) almost surely.

Observe that, conditional on \( m_\kappa(\Gamma_t) \geq 1 \), we have
\[
m_\kappa(\Gamma_{t+1}) - m_\kappa(\Gamma_t) = 1(X^\mathbb{P}(i) = \kappa) - 1(X^\mathbb{P}(i) = 0).
\]
If follows that during \( m_\kappa(\Gamma_t) \geq 1 \), \( (m_\kappa(\Gamma_t))_{t \geq 0} \) forms a simple lazy random walk of independent increments with the following transition probabilities
\[
\mathbb{P}(m_\kappa(\Gamma_{t+1}) = 1 | m_\kappa(\Gamma_t) = 1) = p_\kappa,
\]
\[
\mathbb{P}(m_\kappa(\Gamma_{t+1}) = -1 | m_\kappa(\Gamma_t) = 1) = p_0,
\]
\[
\mathbb{P}(m_\kappa(\Gamma_{t+1}) = 1 | m_\kappa(\Gamma_t) = 1) = 1 - p_0 - p_\kappa.
\]
Thus by Gambler’s ruin probability for simple random walks (see (103)), we have
\[
\mathbb{P}(h^{(i)} > x) = \frac{p_0 - p_\kappa}{p_0 + p_\kappa} \frac{1}{[x] + 1}
\]
for all \( x \geq 0 \). Then the lower bound (107) follows.

Assuming Proposition 5.5, we show the following scaling limit of \( h_j(n) \) using a similar argument developed in [LLP17].

Proposition 5.6. Suppose \( p_0 > \max(p_1, \cdots, p_\kappa) \). Let \( C_1, C_2, \theta_1, \theta_2 \) be as in Proposition 5.5. Denote \( \mu^{(i)}_n = \log_\theta_i (C_i \pi(0)n) \) for \( i = 1, 2 \). Let \( h_j(n) \) be the \( j \)-th largest excursion height of the carrier process over \( [0, n] \).

Then for any non-decreasing real sequence \( \{x_n\}_{n \geq 1} \),
\[
\exp(-\theta_1^{-x_n}) \leq \liminf_{n \to \infty} \mathbb{P} \left( h_1(n) \leq x_n + \mu^{(1)}_n \right), \quad \limsup_{n \to \infty} \mathbb{P} \left( h_1(n) \leq x_n + \mu^{(2)}_n \right) \leq \exp(-\theta_2^{-x_n+1})
\]
Furthermore, for any \( j \geq 2 \) and a real sequence \( \{x_n\}_{n \geq 1} \),
\[
\liminf_{n \to \infty} \left[ \exp \left( -\theta_2^{-x_n} 1(\theta_1 = \theta_2) \right) \sum_{\ell = 0}^{j-1} \theta_2^{-\ell x_n} \right]^{-1} \mathbb{P} \{ h_j(n) \leq x_n + \mu^{(2)}_n \} \geq 1,
\]
\[
\limsup_{n \to \infty} \left[ \exp \left( -\theta_1^{-x_n+1} 1(\theta_1 = \theta_2) \right) \sum_{\ell = 0}^{j-1} \theta_1^{-\ell x_n} \right]^{-1} \mathbb{P} \{ h_j(n) \leq x_n + \mu^{(1)}_n \} \leq 1.
\]
Proof. Fix a non-decreasing real sequence $\mu_n$. Denote $\sigma = \pi(0) > 0$ as in (101). Fix $\varepsilon > 0$ and let $b_n = \lceil (\sigma - \varepsilon) n \rceil$. As $M_n/l_n \to \sigma$ a.s., we have that $M_n \geq b_n$ for all sufficiently large $n$ almost surely (see [LLP17, Sec. 4] for more details). According to Proposition 5.5 and (102), we have

\[
P\{h_j(n) \leq x + \mu_n\} \leq P\{h_{1:b_n} \leq x + \mu_n\}
\]

(116)

\[
\leq \sum_{\ell=0}^{j-1} b_n\left(1 - \frac{C_2}{\theta_2^{x+\mu_n}}\right)^{b_n-\ell}\frac{C_1}{\theta_1^{x+\mu_n}} \ell
\]

(117)

\[
= \left(1 - \frac{C_2}{\theta_2^{x+\mu_n+1}}\right)^{b_n} \sum_{\ell=0}^{j-1} b_n^{\ell}\left(1 - \frac{C_2}{\theta_2^{x+\mu_n}}\right)^{-\ell}\frac{C_1 b_n}{\theta_1^{x+\mu_n}} \ell
\]

(118)

Note that $\lim_{n \to \infty} b_n^{-\ell} (b_n) = 1$. For $i = 1, 2$, also note that we have the following limits

\[
\lim_{n \to \infty} \exp\left(1 - \frac{\varepsilon}{\sigma}\theta_1^{x_n+1}\right)\left(1 - \frac{C_i}{\theta_i^{x_n+\mu_n^{(i)}}+1}\right) = 1, \quad \lim_{n \to \infty} \left(1 - \frac{\varepsilon}{\sigma}\right)\theta_i^{x_n+\mu_n^{(i)}} = 1
\]

(119)

Moreover, if $\theta_1 < \theta_2$, then we have

\[
\lim_{n \to \infty} \left(1 - \frac{C_2}{\theta_2^{x_n+\mu_n^{(i)}}+1}\right)^{b_n} = 1, \quad \lim_{n \to \infty} \frac{C_1 b_n}{\theta_1^{x_n+\mu_n^{(i)}}} = \infty.
\]

(120)

Thus for $\theta_1 = \theta_2$, we have

\[
\limsup_{n \to \infty} \left[\exp\left(-\left(1 - \frac{\varepsilon}{\sigma}\theta_1^{x_n+1}\right)\left(1 + \varepsilon\right)\theta_1^{-\ell x_n}\right)\right]^{-1} \leq 1,
\]

(121)

and for $\theta_1 < \theta_2$, we get

\[
\limsup_{n \to \infty} \left[\sum_{\ell=0}^{j-1} (1 + \varepsilon)\theta_1^{-\ell x_n}\right]^{-1} \leq 1.
\]

(122)

Therefore letting $\varepsilon \searrow 0$ gives the second part of the assertion. A similar argument shows the first assertion.

Now we are ready to prove our main result for the subcritical regime.

Proof of Theorem 1.2 (i). Fix $j \geq 1$, $x \in \mathbb{R}$, and let $\mu_n = \log_{\theta_j}(C_i n)$ for $i = 1, 2$ as in the assertion. The assertion for $\lambda_1(n)$ follows immediately from Lemma 4.1 (i) and the first part of Proposition 5.6. Since $\lambda_j(n) \leq \lambda_1(n)$, this also shows $\lambda_j(n) = O(\log n)$ for all $j \geq 1$. For the lower bound on the subsequent soliton lengths, fix $j \geq 2$ and use Lemma 4.1 (ii) to write

\[
P\{\lambda_j(n) \leq x_n + \mu_n^{(2)}\} \leq P\{h_j(n) \leq x_n + \mu_n^{(2)}\}
\]

for any real sequence $\{x_n\}_{n \geq 1}$. Then the assertion follows from the second part of Proposition 5.6. \(\square\)
6. Decoupling the Carrier Process for Critical and Supercritical Regimes

In this section, we introduce a ‘decoupled version’ of the infinite capacity process, which will be useful in getting upper bounds on the top soliton length in the critical and supercritical regimes. The key idea is to partition the color space \( \mathbb{Z}_{k+1} \) into intervals and modify the circular exclusion rule so that 1) the exclusion process among colors that belong to distinct intervals are decoupled, and 2) the exclusion process among colors from each interval behaves like the subcritical carrier process we analyzed in Section 5.

Let \( (\Gamma_t)_{t \geq 0} \) denote the infinite capacity carrier process over \( X^\mathbb{P} \) as introduced in Subsection 2.2 and let \( \lambda_j(n) = \lambda_j(X^n) \) for each \( n, j \geq 1 \). Throughout this section, we assume \( p^* := \max(p_1, \ldots, p_\kappa) \geq p_0 \).

6.1. The decoupled carrier process. In this subsection, we introduce the following ‘decoupled’ version of the infinite capacity carrier process. To begin, we call an integer \( 0 \leq i \leq \kappa \) unstable color if \( i = 0 \) or \( p_i \geq \max(p_{i+1}, \ldots, p_\kappa, p_0) \). Denote the set of all unstable colors by \( \mathcal{C}^u_\mathbb{P} = \{\alpha_0, \alpha_1, \ldots, \alpha_r\} \) where \( 0 = \alpha_0 < \alpha_1 < \cdots < \alpha_r \). We call the elements in the subset \( \mathcal{C}^s_\mathbb{P} = \{0, 1, \ldots, \kappa\} \setminus \mathcal{C}^u_\mathbb{P} \) stable colors. Write the elements of \( \mathcal{C}^u_\mathbb{P} \) as \( 0 = \alpha_0 < \alpha_1 < \cdots < \alpha_r \). This will partition the integer ring \( \mathbb{Z}_{k+1} \) into intervals \( [\alpha_j, \alpha_{j+1}] \), \( 0 \leq j \leq r \), where we take \( \alpha_{r+1} = \kappa + 1 \equiv 0 \mod (\kappa + 1) \). We then localize the circular exclusion rule within each such interval by letting each ball of color \( q \in [\alpha_j, \alpha_{j+1}] \setminus \{0\} \) in the carrier be replaced by new balls of colors only in the interval \( [\alpha_j, \alpha_{j+1}] \). More precisely, define a map \( \tilde{\Psi} : \mathcal{B}_\infty \times [0, 1, \ldots, \kappa] \to [0, 1, \ldots, \kappa] \times \mathcal{B}_\infty \), \( (x, y) \to (x', y') \) by the following localized circular exclusion rule:

(i) Suppose \( \alpha_j < y < \alpha_{j+1} \) for some \( 0 \leq j \leq r \). Denote \( i^* = \min\{i \geq 1 : x(i) < y\} \). If \( x(i^*) \in [\alpha_j, \alpha_{j+1}] \), then we let \( y' = x(i^*) \) and

\[
    x'(i) = x(i) \mathbf{1}(i \neq i^*) + y \mathbf{1}(i = i^*) \quad \forall i \geq 1.
\]

Otherwise, let \( y' = 0 \) and define \( x' \) to be the unique element in \( \mathcal{B}_\infty \) obtained from \( x \) by replacing a 0 with \( y \).

(ii) Suppose \( y = 0 \). Then apply (i) by regarding \( y = \kappa + 1 = \alpha_{r+1} \).

Now for the random \( \kappa \)-color BBS configuration \( X = X^\mathbb{P} \), we define a Markov chain \( (\tilde{\Gamma}_t)_{t \geq 0} \) on \( \mathcal{B}_\infty \) by \( \tilde{\Gamma}_0 = [0, 0, \ldots] \) and

\[
    \tilde{\Psi}(\tilde{\Gamma}_t, X(t + 1)) = (X'(t + 1), \tilde{\Gamma}_{t+1}) \quad \forall t \geq 0.
\]

We call this the decoupled carrier process over \( X^\mathbb{P} \). See Figures 6 and 7 for an illustration.

![Figure 6](image-url)

**Figure 6.** Illustration of the original circular exclusion rule (left) and its decoupled version (right) for \( \kappa = 7 \) and ball density \( p = (.1, .1, .25, .05, .15, .2, .1, .05) \). In this case \( \mathcal{C}^u_\mathbb{P} = \{0, 2, 5, 6\} \). For instance, in the localized rule, inserting new balls of color 5 into the carrier only excludes existing balls of colors 2, 3, 4 or 0 (empty spot).
Note that the decoupled carrier process \( \Gamma_t \) agrees with the usual carrier process \( \Gamma_t \) when \( \mathcal{C}^u = \emptyset \). In general, we show that \( \Gamma_t \) dominates \( \Gamma_t \), and also that they agree on colors between the largest unstable color \( \alpha_r \) and \( \kappa \).

**Proposition 6.1.** Let \( \Gamma_t \) and \( \Gamma_t \) denote the infinite capacity and localized carrier processes over \( X^P \), respectively. Write \( \mathcal{C}^u = \{ \alpha_0, \cdots, \alpha_\kappa \} \) for the set of unstable colors, where \( 0 = \alpha_0 < \alpha_1 < \cdots < \alpha_\kappa \leq \kappa \). Then the following hold.

(i) If \( p_0 > \max(p_1, \cdots, p_\kappa) \), then \( \Gamma_t = \Gamma_t \) for all \( t \geq 0 \).

(ii) \( m_i(\Gamma_t) = m_i(\Gamma_t) \) for all \( \alpha_r \leq i \leq \kappa \) and \( t \geq 0 \).

(iii) \( m_i(\Gamma_t) \leq m_i(\Gamma_t) \) for all \( 1 \leq i \leq \kappa \) and \( t \geq 0 \).

**Proof.** (i) follows easily since \( \mathcal{C}^u = \emptyset \) if \( p_0 > \max(p_1, \cdots, p_\kappa) \). In this case the two circular exclusion rules \( \Psi \) and \( \Psi \) that define the two carrier processes are the same. As \( \Gamma_0 = \Gamma_0 = \{0, 0, \cdots\} \), the assertion follows.

Next, we show (ii) by an induction on \( t \geq 0 \). For \( t = 0 \) we have \( \Gamma_0 = \Gamma_0 = \{0, 0, \cdots\} \). Suppose \( \Gamma_t = \Gamma_t \) for some \( t \geq 1 \). Denote \( \ell = \alpha_r \) and \( y = X^P(t+1) \). Let \( y', y'' \in \{0, 1, \cdots, \kappa\} \) so that

\[
\Psi(\Gamma_t, y) = (y', \Gamma_{t+1}), \quad \Psi(\Gamma_t, y) = (y'', \Gamma_{t+1}).
\]  

(125)

Then we wish to show that

\[
m_i(\Gamma_{t+1}) = m_i(\Gamma_{t+1}) \quad \forall \ell \leq i \leq \kappa.
\]  

(126)

First suppose \( 1 \leq y < \ell \), then \( 0 \leq y' < y \). Since the total number of each letter \( 1 \leq i \leq \kappa \) is the same in both pairs \( (\Gamma_t, y) \) and \( (y', \Gamma_{t+1}) \), it follows that \( \Gamma_t \) and \( \Gamma_{t+1} \) contain the same number of letter \( i \)'s for all \( \ell \leq i \leq \kappa \). Moreover, according to the definition of \( \Psi \), both \( \Gamma_t \) and \( \Gamma_{t+1} \) contain the same number of letter \( i \)'s for all \( \ell \leq i \leq \kappa \). Then by the induction hypothesis, for all \( \ell \leq i \leq \kappa \),

\[
m_i(\Gamma_t) = m_i(\Gamma_t) = m_i(\Gamma_{t+1}).
\]  

(127)

Second, suppose \( y \in \{0, \ell, \ell + 1, \cdots, \kappa\} \). If \( \ell \leq y'' < y \), then by the induction hypothesis, \( y' = y'' \) and (126) holds. Otherwise \( y'' = 0 \). If \( y \in \{\ell, \ell + 1, \cdots, \kappa\} \), then this means \( m_j(\Gamma_t) = 0 \) for all \( \ell \leq j < y \). By the induction hypothesis, this yields \( m_j(\Gamma_t) = 0 \) for all \( \ell \leq j < y \). Hence \( 0 \leq y' < \ell \), so (126) holds. A similar argument applies when \( y = 0 \). This completes the induction.
Finally, we show (iii) by an induction on \( t \). Since \( \Gamma_0 = \tilde{\Gamma}_0 = [0, 0 \cdots] \), (iii) holds for \( t = 0 \). For the induction step, suppose (iii) holds for \( t \geq 1 \). Suppose \( X^p(t+1) = y \) and let \( y', y'' \in [0, 1, \cdots, \kappa] \) be such that (125) holds. First note that since the number of color 0 balls in both \( \Gamma_t \) and \( \tilde{\Gamma}_t \) increase by 1 when inserting \( y \), we have

\[
m_y(\Gamma_{t+1}) = m_y(\Gamma_t) + 1 \leq m_y(\tilde{\Gamma}_t) + 1 = m_y(\tilde{\Gamma}_{t+1}).
\]

(128)

If there exists \( 0 \leq j \leq r \) such that \( y \in (\alpha_j, \alpha_{j+1}] \) and \( y' \in [\alpha_j, \alpha_{j+1}) \), then \( y' = y'' \) so (iii) holds. Otherwise, \( y'' = 0 \) so for each \( 1 \leq j \leq \kappa \) and \( j \neq y \),

\[
m_j(\Gamma_{t+1}) \leq m_j(\Gamma_t) \leq m_j(\tilde{\Gamma}_t) \leq m_j(\tilde{\Gamma}_{t+1}).
\]

(129)

This completes the induction.

Lastly in this subsection, we show a stability and convergence result for the decoupled carrier process, analogous to Lemma 5.1 for the subcritical carrier process. Namely, let \((\tilde{\Gamma}_t)_{t \geq 0}\) denote the decoupled carrier process over \( X^p \) and let \( \mathcal{C}_p = [\alpha_0, \alpha_1, \cdots, \alpha_r] \) denote the set of unstable colors. For each \( t \geq 0 \), define \( S_{\Gamma} \) to be the \( \kappa - r \) dimensional nonnegative integer vector whose coordinates are indexed by the stable colors and give the corresponding ball multiplicity:

\[
S_{\Gamma}: \mathcal{C}_p \to \mathbb{Z}_{\geq 0}, \quad S_{\Gamma}(i) = m_i(\tilde{\Gamma}_t).
\]

(130)

Note that the \( \kappa - r \) dimensional state space \((\mathbb{Z}_{\geq 0})^{|\mathcal{C}_p^0|}\) with \((\mathbb{Z}_{\geq 0})^{|\mathcal{C}_p|}\) in the canonical way.

It is important to note that, even though it misses the information on the \( r \) unstable colors, \((S_{\Gamma})_{t \geq 0}\) forms a Markov chain due to the decoupling. Moreover, in Lemma 6.2, we will show that this chain converges to its unique stationary distribution \( \tilde{\pi} \), which is defined by

\[
\tilde{\pi}(\kappa^{n}; i \in \mathcal{C}_p^0) = \prod_{j=0}^{r} \left( \prod_{\alpha_j < i < \alpha_{j+1}} \left( 1 - \frac{p_i}{p_{\alpha_{j+1}}} \right) \left( \frac{p_i}{p_{\alpha_{j+1}}} \right)^{n_{1}} \right),
\]

(131)

where we set \( \alpha_{r+1} = \kappa + 1 \).

**Lemma 6.2.** Let \((S_{\Gamma})_{t \geq 0}\) be the Markov chain defined above. Then it is an irreducible and aperiodic Markov chain with unique stationary distribution \( \tilde{\pi} \) on \( B^\infty \) defined in (131). Furthermore, if we denote the distribution of \( S_{\Gamma} \) by \( \tilde{\pi}_t \), then

\[
\lim_{n \to \infty} d_{TV}(\tilde{\pi}_t, \tilde{\pi}) = 0.
\]

(132)

**Proof.** We use a similar argument as in the proof of Lemma 5.1. We only show the irreducibility of the chain \((\tilde{\Gamma}_t)_{t \geq 0}\) and stationarity of the distribution \( \tilde{\pi} \). We first show that the decoupled carrier process \((\tilde{\Gamma}_t)_{t \geq 0}\) is irreducible on \( B^\infty \), which will show the irreducibility of the chain \((\tilde{\Gamma}_t)_{t \geq 0}\). To show this, first observe that \( \tilde{\Gamma}_t \) visits every state \( \mathbf{x} \in B^\infty \) with positive probability starting from the initial state \([0, 0, \cdots]\). Hence it suffices to show the converse transition.

Namely, fix \( \mathbf{x} \in B^\infty \) and suppose \( \tilde{\Gamma}_t = \mathbf{x} \). Denote \( n_1 = m_1(\mathbf{x}) + \cdots + m_{\alpha_1-1}(\mathbf{x}) \), which is the number of balls of color in \([1, \alpha_1]\). Observe that inserting \( n_1 \) balls of color \( \alpha_1 \) into the decoupled carrier \( \tilde{\Gamma}_t \) removes all balls of colors in \([1, \alpha_1]\) and leaves with \( m_{\alpha_1}(\mathbf{x}) + n_1 \) balls of color \( \alpha_1 \). Next, we insert \( m_{\alpha_1}(\mathbf{x}) + n_1 + n_2 \) balls of color \( \alpha_2 \) into the decoupled carrier, where \( n_2 = m_{\alpha_2}(\mathbf{x}) + \cdots + m_{\alpha_2-1}(\mathbf{x}) \). This will remove all remaining balls of colors in \([1, \alpha_2]\) and leave \( m_{\alpha_2}(\mathbf{x}) + (m_{\alpha_1}(\mathbf{x}) + n_1 + n_2) \) balls of color \( \alpha_2 \). Repeating this process, we can remove all balls of colors in \([1, \alpha_j]\) by inserting a finite string of balls. We can then insert balls of color 0 until the decoupled carrier does not contain any balls of nonzero colors. Since we can feed in any finite string of balls with a positive probability, this shows the desired irreducibility of the chain.
Next, we show that \( \pii \) is a stationary distribution of \( (\Gamma_t)_{t \geq 0} \). First of all, it defines a probability distribution on \((\mathbb{Z}_{\geq 0})^{\kappa^\ell} \) since \( p_i < p_{a_{j+1}} \) for all \( a_j < i < a_{j+1} \). For each \( x \in \mathcal{B}_\infty \) and \( i \in \{0, 1, \cdots, \kappa\} \), denote

\[
\exp(\text{wt}(x)) = \prod_{j=0}^{r} \left( \prod_{a_j < i < a_{j+1}} \left( \frac{p_i}{p_{a_{j+1}}} \right)^{m_i(x)} \right), \quad \exp(\text{wt}(i)) = p_i.
\]  

(133)

Note that for each pair \( (x, y) \in \mathcal{B}_\infty \times \{1, 2, \cdots, \kappa\} \) and \( (y', x') \in \{1, 2, \cdots, \kappa\} \times \mathcal{B}_\infty \) such that \( \Psi(x, y) = (y', x') \), we have

\[
\exp(\text{wt}(x)) \exp(\text{wt}(y)) = \exp(\text{wt}(y')) \exp(\text{wt}(x')).
\]  

(134)

Indeed, the total number of each letter \( \ell \leq i \leq \kappa \) in both pairs \( (x, y) \) and \( (y', x') \) are the same. Moreover, \( y \) and \( y' \) belong to the same interval \([a_j, a_{j+1})\) for some unique \( 0 \leq j \leq r \) according to the definition of the map \( \Psi \). The rest of the argument is the same as in the proof of Lemma 5.1. \( \square \)

6.2. Additive functional of the decoupled carrier process. The goal of this subsection is to reformulate the numbers of balls of unstable colors in the decoupled carrier process \( (\tilde{\Gamma}_t)_{t \geq 1} \) in such a form that we can analyze their scaling limit. We first give an outline of our approach below.

In the previous subsection, we have shown that the numbers of balls of stable colors in the decoupled carrier form a Markov chain with unique stationary distribution. More precisely, let \( 0 = a_0 < a_1 < \cdots < a_r \) denote the unstable colors, as before. In the decoupled carrier process \( (\tilde{\Gamma}_t)_{t \geq 1} \), the number of balls of unstable color \( a_j \) is a function of the numbers of balls of colors in the interval \([a_j, a_{j+1})\). Moreover, if we are only interested in one-step increment, we only need to store the numbers of balls of stable colors in \((a_j, a_{j+1})\) and see if the incoming ball has color \( a_j \) or \( a_{j+1} \). Hence it is possible to reformulate the increments of \( m_{a_j}(\tilde{\Gamma}_t) \) as a function of the ‘stable’ Markov chain \((m_{i}(\tilde{\Gamma}_t); a_j < i < a_{j+1})\) with the auxiliary information \( X^p(t+1) \). Consequently, this enables us to write \( m_{a_j}(\tilde{\Gamma}_t) \) as an ‘additive functional’ of a Markov chain with unique stationary distribution, for which various standard limit theorems are available from the general Markov chain theory.

Now we give precise formulation. For each \( 1 \leq \ell \leq \kappa \), we let \( \ell^+ \) denote the unique unstable color (modulo \( \kappa + 1 \)) that can replace balls of color \( \ell \) in the localized circular exclusion rule:

\[
\ell^+ = \begin{cases} 
\kappa + 1 & \text{if } \ell \geq \text{the largest unstable color,} \\
\min \{\ell \leq i \leq \kappa \mid i \in \mathcal{C}_{\ell}^p \} & \text{Otherwise.}
\end{cases}
\]  

(135)

For \( \ell^+ > \ell + 1 \), define a Markov chain

\[
S_{\ell}^\ell \tilde{\Gamma}_t := (m_{\ell+1}(\tilde{\Gamma}_t), \cdots, m_{\ell^+-1}(\tilde{\Gamma}_t)).
\]  

(136)

By Lemma 6.2 and the definition of the decoupled carrier process \( \tilde{\Gamma}_t \), it is easy to see that this chain is irreducible, aperiodic, and converges to the unique stationary distribution \( \pii^\ell \) defined by

\[
\pii^\ell(n_{\ell+1}, \cdots, n_{\ell^+-1}) = \prod_{i=\ell+1}^{\ell^+-1} \left( 1 - \frac{p_i}{p_{\ell^+}} \right) \left( \frac{p_i}{p_{\ell^+}} \right)^{n_i}.
\]  

(137)

Now define a Markov chain \( \tilde{Z}_{\ell}^\ell \) on the state space \( \mathbb{Z}_{\geq 0}^{\ell^+-\ell} \times \mathbb{Z}_{\kappa+1} \) by

\[
\tilde{Z}_{\ell}^\ell := \begin{cases} 
(S_{\ell}^\ell \tilde{\Gamma}_t, X^p(t+1)) & \text{if } \ell^+ > \ell + 1 \\
X^p(t+1) & \text{otherwise.}
\end{cases}
\]  

(138)

Then \( \tilde{Z}_{\ell}^\ell \) is irreducible, aperiodic, and converges to the unique stationary distribution \( \pii^\ell \otimes p \), which we understand as \( p \) when \( \ell^+ = \ell + 1 \).
Define a functional $g_\ell : (\mathbb{Z}_{\geq 0})^{\ell^* - \ell - 1} \times \{0, 1, \ldots, \kappa\} \rightarrow \{-1, 0, 1\}$ by

$$g_\ell(f, k) = \mathbf{1}(k = \ell) - \mathbf{1}(k = \ell + 1) - \sum_{q=\ell+2}^{\ell^*} \mathbf{1}(k = q)\mathbf{1}(f(q+1) = \cdots = f(1-q) = 0).$$

Define an additive functional $(\tilde{S}^\ell_t)_{t \geq 0}$ by

$$\tilde{S}^\ell_t = \sum_{k=1}^{t} g_\ell(\tilde{Z}^\ell_k), \quad \tilde{S}^\ell_0 = 0. \quad (139)$$

Note that $(\tilde{S}^\ell_t)_{t \geq 0}$ is a Markov additive functional built on top of the irreducible and stationary Markov chain $(\tilde{Z}^\ell_t)_{t \geq 0}$. The proposition below relates this to the number of balls of color $j$ in the decoupled carrier process.

**Proposition 6.3.** Let $(\tilde{\Gamma}_t)_{t \geq 0}$ be the decoupled carrier process over $X^\mathbb{P}$. Fix $1 \leq \ell \leq \kappa$ and let $(\tilde{S}^\ell_t)_{t \geq 0}$ be as defined at (139). Then for all $t \geq 0$,

$$m_\ell(\tilde{\Gamma}_t) = \tilde{S}^\ell_t - \min_{0 \leq s \leq t} \tilde{S}^\ell_s. \quad (140)$$

**Proof.** Observe that the functional $g_\ell$ gives the 1-step increment of $m_\ell(\tilde{\Gamma}_t)$ from the carrier process in the following sense:

$$m_\ell(\tilde{\Gamma}_{t+1}) - m_\ell(\tilde{\Gamma}_t) = \begin{cases} g_\ell(\tilde{Z}^\ell_t) & \text{if } m_\ell(\tilde{\Gamma}_t) > 0 \\ 1(g_\ell(\tilde{Z}^\ell_t) = 1) & \text{otherwise.} \end{cases}$$

In other words, $g(\tilde{Z}^\ell_t)$ is the honest increment of the ball count $m_\ell(\tilde{\Gamma}_t)$ if we allow it to take negative values. Then assertion follows by an easy inductive argument on $t$. \qed

Next, we give the distribution of the one-step increment of the Markov additive functional $\tilde{S}^\ell_t$ at stationarity. Clearly the number of balls of color $\ell$ increase by 1 if and only if the incoming ball has color $\ell$, which occurs with probability $p_\ell$. In Proposition 6.4 below, we show that the number decrease by 1 with probability $p_{\ell^*}$ under stationarity.

**Proposition 6.4.** Let $(\tilde{\Gamma}_t)_{t \geq 0}$ be the decoupled carrier process over $X^\mathbb{P}$. Fix $1 \leq \ell \leq \kappa$ and let $(\tilde{S}^\ell_t)_{t \geq 0}$ be as defined at (139). If $S^\ell \tilde{\Gamma}_0$ is distributed as its unique stationary distribution $\pi^\ell$, then for all $t \geq 0$,

$$\mathbb{P}_{\pi^\ell}(\tilde{S}^\ell_{t+1} - \tilde{S}^\ell_t = 1) = p_\ell, \quad \mathbb{P}_{\pi^\ell}(\tilde{S}^\ell_{t+1} - \tilde{S}^\ell_t = -1) = p_{\ell^*}, \quad (141)$$

where we set $p_{\kappa+1} = p_0$.

**Proof.** The first assertion is clear, since $\tilde{S}^\ell_t$ increase by 1 if and only if the new ball $X^\mathbb{P}(t+1)$ has color $\ell$. For the second equation, observe that

$$\mathbb{P}_{\pi^\ell}(\tilde{S}^\ell_{t+1} - \tilde{S}^\ell_t = -1) = p_{\ell+1} + \sum_{i=\ell+2}^{\ell^*} \mathbb{P}_{\pi^\ell}(m_{\ell+1}(\tilde{\Gamma}_t) = \cdots = m_{i-1}(\tilde{\Gamma}_t) = 0)p_i. \quad (142)$$

It remains to show that the right hand side above equals $p_{\ell^*}$.

Since $S^\ell \tilde{\Gamma}_t$ is distributed as the stationary distribution $\pi^\ell$ for all $t \geq 0$, if we denote

$$M_\ell(\tilde{\Gamma}_t) = m_{\ell+1}(\tilde{\Gamma}_t) + m_{\ell+2}(\tilde{\Gamma}_t) + \cdots + m_{\ell^*}(\tilde{\Gamma}_t), \quad (143)$$

then we have

$$\mathbb{E}_{\pi^\ell}[M_\ell(\tilde{\Gamma}_{t+1}) - M_\ell(\tilde{\Gamma}_t)] = 0. \quad (144)$$

Also note that

$$\mathbb{P}_{\pi^\ell}(M_\ell(\tilde{\Gamma}_{t+1}) - M_\ell(\tilde{\Gamma}_t) = -1) = (1 - \mathbb{P}_{\pi^\ell}(m_{\ell+1}(\tilde{\Gamma}_t) = \cdots = m_{\ell^*}(\tilde{\Gamma}_t) = 0))p_{\ell^*}. \quad (145)$$
\[ p_{\ell} = p_{\ell+1} + \sum_{i=\ell+2}^{\ell'} \mathbb{P}_{\overline{\pi}^{+}}(m_{\ell+1}(\overline{T}_t) = \cdots = m_{i-1}(\overline{T}_t) = 0)p_i \tag{146} \]

Since \( M_{\ell}(\overline{T}_{t+1}) \) and \( M_{\ell}(\overline{T}_t) \) can only differ by 1, (144), this yields

\[ p_{\ell'} = p_{\ell+1} + \left( \sum_{i=\ell+2}^{\ell'} \mathbb{P}_{\overline{\pi}^{+}}(m_{\ell+1}(\overline{T}_t) = \cdots = m_{i-1}(\overline{T}_t) = 0)p_i \right) + \mathbb{P}_{\overline{\pi}^{+}}(m_{\ell+1}(\overline{T}_t) = \cdots = m_{i-1}(\overline{T}_t) = 0)p_{\ell'}. \tag{147} \]

Note that the right hand side equals \( \mathbb{P}_{\overline{\pi}^{+}}(\tilde{S}_{t+1}^{\ell'} - \tilde{S}_{t}^{\ell'} = -1) \), as desired. \( \square \)

### 6.3. Limit theorems for the additive functionals

Next, we prove limit theorems for the additive functional \((\tilde{S}_{t}^{\ell})_{t \geq 0}\). We keep the same notations as we used in the previous subsection.

Since \((\tilde{Z}_{t}^{\ell})_{t \geq 0}\) is irreducible with a unique stationary distribution, all states are positive recurrent (see, e.g., [LP17, Lem. 21.13]). In Proposition 6.4, we have shown that

\[ \mathbb{E}_{\overline{\pi}^{+}}[g_{\ell}(\tilde{Z}_{t}^{\ell})] = p_{\ell} - p_{\ell'}. \tag{148} \]

Also, define the **limiting variance** \( \gamma_{\ell}^{2} \) of \((\tilde{S}_{t}^{\ell})_{t \geq 0}\) by

\[ \gamma_{\ell}^{2} := \text{Var}[g_{\ell}(\tilde{Z}_{0}^{\ell})] + 2 \sum_{k=1}^{\infty} \text{Cov}[g_{\ell}(\tilde{Z}_{0}^{\ell}), g_{\ell}(\tilde{Z}_{k}^{\ell})] \tag{149} \]

and denote \( \gamma_{\ell} = \sqrt{\gamma_{\ell}^{2}} \). In Proposition 6.6, we will obtain limit theorems for the additive functional \((\tilde{S}_{t}^{\ell})_{t \geq 0}\) by applying limit theorems for positive Harris chains with unique stationary measure. In doing so, a critical step is to show the following statement.

**Lemma 6.5.** The limiting variance \( \gamma_{\ell}^{2} \) of \((\tilde{S}_{t}^{\ell})_{t \geq 0}\) defined in (149) is positive and finite.

Our proof of Lemma 6.5 is based on showing that \( \tilde{Z}_{t} \) is geometrically ergodic, and hence its return time to the initial state has exponential tail. Recall that a similar statement (Lemma 5.4) was crucial for the subcritical regime. We postpone the proof of Lemma 6.5 to Section 8.

Now we state and derive limit theorems for the additive functional \((\tilde{Z}_{t}^{\ell})_{t \geq 0}\). Let \( \tilde{S}_{t}^{\ell}(\cdot) \) denote the linear interpolation of the points \((k, \tilde{Z}_{k}^{\ell} - (p_{\ell} - p_{\ell'}))k \in \mathbb{N} \times \mathbb{R} \) for all \( k \geq 0 \). Let \( C([0,1]) \) denote the space of continuous functions \( f: [0,1] \rightarrow \mathbb{R} \) equipped with the supremum norm.

**Proposition 6.6.** Let \((\tilde{Z}_{t}^{\ell})_{t \geq 0}\) be as before. Fix \( 1 \leq \ell \leq \kappa \) and let \((\tilde{S}_{t}^{\ell})_{t \geq 0}\) be as defined as in (139). Let \( \gamma_{\ell}^{2} \) be as in (149). Then the following hold.

**(i)** (SLLN) Almost surely,

\[ \lim_{n \to \infty} n^{-1}\tilde{S}_{n}^{\ell} = p_{\ell} - p_{\ell'}. \tag{150} \]

**(ii)** (CLT) Let \( Z \sim N(0,1) \) be a standard normal random variable. Then as \( n \to \infty \),

\[ n^{-1/2}(\tilde{S}_{n}^{\ell} - (p_{\ell} - p_{\ell'}))n) \Rightarrow Z. \tag{151} \]

**(iii)** (FCLT) Let \( B = (B_{u}: 0 \leq u \leq 1) \) denote the standard Brownian motion. Then as \( n \to \infty \),

\[ (t^{-1/2}\tilde{S}_{t}^{\ell}(tu) : 0 \leq u \leq 1) \Rightarrow (\gamma_{\ell}B_{u} : 0 \leq u \leq 1) \text{ in } C([0,1]), \tag{152} \]

where \( \Rightarrow \) denotes weak convergence in \( C([0,1]) \).
Proof. According to Lemma 6.2, the chain \((\tilde{Z}_t^\ell)_{t \geq 0}\) is ergodic, irreducible, and has a unique stationary distribution \(\tilde{\pi}^\ell \otimes p\), where \(\tilde{\pi}\) is defined in Lemma 6.2. Hence (i) follows from the law of large numbers for positive Harris chains (see, e.g., [MT12, Thm. 17.1.7]). Assuming Lemma 6.5, (ii) follows from the central limit theorem for positive Harris chains with an atom (see, e.g., [MT12, Thm. 17.2.2]). Also, (iii) follows from the functional central limit theorem for positive Harris chains (see, e.g., [MT12, Thm. 17.4.4 and eq. (17.38)]).

7. PROOF OF THEOREM 1.2 IN THE CRITICAL AND SUBCRITICAL REGIMES

7.1. Proof of Theorem 1.2 in critical regime. We prove Theorem 1.2 (ii) in this subsection. For each unstable color \(\ell \in C_u^p\), we let \(\tilde{h}^\ell(\cdot)\) denote the linear interpolation of the points \((k, m^\ell_1(\tilde{\Gamma}_k)) \in \mathbb{N}^2\). In case the corresponding additive process \((\tilde{S}_t^\ell)_{t \geq 0}\) is 'critical' i.e., \(p^\ell = p^{\ell,+}\), we show that \(\tilde{h}^\ell(\cdot)\) after rescaling converges weakly to the reflecting Brownian motion.

![Figure 8](image_url) Simulation of queue length paths for \(\kappa = 2\), \(n = 5 \times 10^6\), and a critical density \(p = (1/3,1/3,1/3)\). The red, green, and blue paths depict the paths \(m_1(\Gamma_t), m_2(\Gamma_t),\) and \(m_1(\Gamma_t) + m_2(\Gamma_t)\), respectively, under diffusive scaling.

Proposition 7.1. Fix an unstable color \(\ell \in C_u^p\) and suppose \(p^\ell = p^{\ell,+}\). Let \(\gamma^2_\ell\) be as defined in (149). Then the following hold.

(i) Let \(\tilde{h}^\ell(\cdot)\) be as before. Then as \(n \to \infty\),

\[
(t^{-1/2}\tilde{h}^\ell(tu) : 0 \leq u \leq 1) \Rightarrow (\gamma^\ell |B_u| : 0 \leq u \leq 1) \quad \text{in } C([0,1])
\]

where \(\Rightarrow\) denotes weak convergence in \(C([0,1])\) and \(B = (B_u : 0 \leq u \leq 1)\) denotes the standard Brownian motion.

(ii) As \(n \to \infty\),

\[
n^{-1/2} \max_{1 \leq s \leq n} m^\ell_1(\Gamma_s) \xrightarrow{d} \gamma^\ell \max |B|,
\]

where \(\xrightarrow{d}\) denotes convergence in distribution.

Proof. Let \(\gamma^2_\ell\) denote the limiting variance of the additive functional \((\tilde{S}_t^\ell)_{t \geq 0}\). Note that (ii) follows immediately from (i). To show (i), we follow a similar approach in [LLP17] developed for the \(\kappa = 1\) case at the critical phase. Define an operator \(\mathcal{E}_0 : C([0,1]) \to C([0,1])\) by

\[
\mathcal{E}_0(f)(u) = f(u) - \min_{0 \leq s \leq u} f(s).
\]

(155)
Let \( \tilde{S}_{f,t} \) and \( \tilde{H}_{f,t} \) denote the functions \((\gamma_t^{-1} r^{-1/2} \tilde{S}(tu) : 0 \leq u \leq 1)\) and \((\gamma_t^{-1} r^{-1/2} \tilde{H}(tu) : 0 \leq u \leq 1)\), respectively. According to Proposition 6.3, we have \( m_f(\Gamma_t) = \mathcal{E}_0(\tilde{S}_{f,t}) \). Also note that \( \mathcal{E}_0(cf) = c\mathcal{E}_0(f) \) for all \( c \geq 0 \) and \( f \in C([0,1]) \). Hence we have

\[
H_{f,t} = \mathcal{E}_0(\tilde{H}_{f,t}).
\]

By [LLP17, Prop. A.6], the operator \( \mathcal{E}_0 \) is (2-Lipschitz) continuous in \( C([0,1]) \). Hence if we fix a bounded and continuous functional \( G : C([0,1]) \to \mathbb{R} \), then \( G \circ \mathcal{E}_0 : C([0,1]) \to \mathbb{R} \) is also bounded and continuous. Thus by Proposition 6.6 (iii), we have

\[
\lim_{n \to \infty} \mathbb{E}[G(\tilde{H}_{f,t})] = \lim_{n \to \infty} \mathbb{E}[G(\mathcal{E}_0(\tilde{H}_{f,t}))] = \mathbb{E}[G(\mathcal{E}_0(B))].
\]

Since \( G : C([0,1]) \to \mathbb{R} \) was arbitrary, this shows that \( \tilde{H}_{f,t} \) converges weakly to \( \mathcal{E}_0(B) \). As

\[
\mathcal{E}_0(B)(t) = B(t) - \min_{0 \leq s \leq t} B(s) = B(t) - \max_{0 \leq s \leq t} (-B(s)) = \max B(s) - B(t),
\]

Lévy’s \( M-B \) theorem (see [MP10, Ch. 2.3]) implies \( \mathcal{E}_0(B) = \max |B| \). This shows the assertion. \( \square \)

**Proof of Theorem 1.2 (ii).** Suppose \( p_0 = \max(p_1, \ldots, p_k) \). Then \( \mathcal{C}_p^\mu = \{0 \leq i \leq k : p_i = p_0\} \) and we may write \( \mathcal{C}_p^\mu = \{a_0, \ldots, a_r\} \) with \( 0 = a_0 < a_1 < \cdots < a_r \). Note that for each \( 1 \leq \ell \leq k \), \( p_\ell = p_\ell \), if \( \ell \in \mathcal{C}_p^\mu \) and \( p_\ell < p_\ell+1 \) otherwise. Let \( (\tilde{\Gamma}_t)_{t \geq 0} \) denote the decoupled carrier process over \( X^\mathbb{P} \). By Lemma 2.2 and Proposition 6.1 (ii), we have

\[
\max_{1 \leq s \leq n} m_{a_s}(\tilde{\Gamma}_t) \leq \lambda_1(n) \leq \max_{1 \leq s \leq n} \left(m_1(\tilde{\Gamma}_t) + \cdots + m_k(\tilde{\Gamma}_t)\right) \leq \max_{1 \leq s \leq n} m_1(\tilde{\Gamma}_t) + \cdots + \max_{1 \leq s \leq n} m_k(\tilde{\Gamma}_t).
\]

Note that by Propositions 6.6 and 7.1, \( n^{-1/2} m_f(\tilde{\Gamma}_t) \) converges weakly to a constant multiple of a reflecting Brownian motion if \( p_\ell = p_0 \), and converges in probability to zero otherwise. Hence if we let \( \gamma_t^\ell \) denote the limiting variances defined in (149), then the above lower bound on \( \lambda_1(n) \) yield

\[
\gamma_a, \max_{0 \leq u \leq 1} |B_u| \leq \liminf_{n \to \infty} n^{-1/2} \lambda_1(n),
\]

whereas the upper bound and Slutsky’s theorem yield

\[
\limsup_{n \to \infty} n^{-1/2} \lambda_1(n) \leq \sum_{i=1}^r \gamma_{a_i} \max_{0 \leq u \leq 1} |B_u^{(i)}|,
\]

where \( B_u \) and \( B_u^{(i)} \)’s are (not necessarily independent) standard Brownian motions. In particular, if there is a single positive unstable color, i.e., \( \mathcal{C}_p^\mu = \{0, \alpha_1\} \), then we also have

\[
n^{-1/2} \lambda_1(n) \Rightarrow \gamma_\alpha, \max_{0 \leq u \leq 1} |B_u|.
\]

Lastly we argue that \( \lambda_j(n) = \Theta(\sqrt{n}) \). The upper bound on \( \lambda_1(n) \) shows \( \lambda_j(n) = O(\sqrt{n}) \), so we only need
to show that \( \lambda_j(n) \) grows at least in the order of \( \sqrt{n} \). For this, due to Lemma 4.1, we only need to show \( h_j(n) = \Theta(\sqrt{n}) \), where \( h_j(n) \) denotes the \( j \)th largest excursion height of the infinite capacity carrier \( \Gamma_t \) during \( [0,n] \). To this end, we note that

\[
\|S\Gamma_t\|_1 = m_1(\Gamma_t) + \cdots + m_k(\Gamma_t) = \max_{a \in \mathcal{C}_p^\mu} m_a(\tilde{\Gamma}_t) = \max_{0 \leq u \leq 1} |B_u|,
\]

where the last equality follows from Proposition 6.1 (ii). Due to the Brownian scaling limit of \( m_{a,c}(\tilde{\Gamma}_t) \), observe that all of the top finite number of excursions for both \( S\Gamma_t \) and \( m_{a,c}(\tilde{\Gamma}_t) \) have length of order \( \Theta(\sqrt{n}) \). It follows that each of the top \( j \) excursion of \( S\Gamma_t \) can contain at most a constant number (indep. of \( n \)) of top excursion of \( m_{a,c}(\tilde{\Gamma}_t) \). It follows that \( h_j(n) = \Theta(\sqrt{n}) \), as desired. This shows the assertion. \( \square \)
7.2. **Top soliton length in the supercritical regimes.** In this subsection, as stated in Theorem 7.2, we obtain the scaling limit of the top soliton length $\lambda_1(n)$ assuming $p^* := \max(p_1, \cdots, p_\kappa) > p_0$.

**Theorem 7.2.** Suppose $p^* > p_0$. Then almost surely,
\[
\lim_{n \to \infty} n^{-1} \lambda_1(n) = p^* - p_0.
\]
Further assume that there is a single unstable color. Then there exists a constant $c > 0$ such that
\[
\frac{\lambda_1(n) - (p_\kappa - p_0)n}{\sqrt{cn}} \Rightarrow Z \sim \mathcal{N}(0, 1),
\]
where $\mathcal{N}(0, 1)$ denotes standard normal distribution and $\Rightarrow$ denotes weak convergence.

Let $(\Gamma_t)_{t \geq 0}$ denote the infinite capacity carrier process over $X^p$ as introduced in Subsection 2.2 and let $\lambda_j(n) = \lambda_j(X^n, p)$ for each $n, j \geq 1$. Throughout this subsection we assume $p^* = \max(p_1, \cdots, p_\kappa) > p_0$ and write $\mathcal{C}_u^p = \{a_0, \cdots, a_\tau\}$, where $0 = a_0 < a_1 < \cdots < a_\tau$.

**Proposition 7.3.** Fix an unstable color $\ell \in \mathcal{C}_u^p$ such that $p_\ell > p_{\ell^*}$. Let $(\tilde{S}_t^\ell)_{t \geq 0}$ denote the additive functional defined in (139). Then
\[
P\left(\max_{1 \leq s \leq n} |\tilde{S}_s^\ell - \tilde{S}_n^\ell| \geq x\right) \leq e^{-cx}.
\]

**Proof.** The key idea here is to consider the following Markov chain
\[
\tilde{S}_t^\ell := (\max_{1 \leq s \leq t} \tilde{S}_s^\ell - \tilde{S}_n^\ell, m_{t+1}(\tilde{h}_t), \cdots, m_{t^*+1}(\tilde{h}_t))
\]
and its excursion heights. Namely, let $\tilde{M}_n$ be the number of its complete excursions during $[0, n]$, and let $\tilde{h}_1, \tilde{h}_2, \cdots$ denote its subsequent excursion heights. Then we can write
\[
P\left(\max_{1 \leq s \leq n} |\tilde{S}_s^\ell - \tilde{S}_n^\ell| \geq x\right) \leq \mathbb{P}(\|\tilde{\Gamma}_t\| \geq x) \leq \mathbb{P}(\tilde{h}_{\tilde{M}_{n+1}} \geq x) = \mathbb{P}(\tilde{h}_1 \geq x),
\]
where for the last inequality we have used the fact that excursion heights $h_1, h_2, \cdots$ are i.i.d. due to the strong Markov property of $\tilde{S}_t^\ell$. Hence it suffices to show that the first excursion height $\tilde{h}_1$ has exponential tail.

As in the proof Lemma 6.2, we can directly shown that the chain $\tilde{S}_t^\ell$ behaves as a subcritical carrier process with the following unique stationary distribution
\[
\mu(n_\ell, n_2, \cdots, n_\kappa) = \left(1 - \frac{p_0}{p_\ell}\right)^{n_\ell} \prod_{i=1}^{\kappa} \left(1 - \frac{p_i}{p_\ell}\right)^{n_i} \left(\frac{p_i}{p_0}\right)^{n_i}.
\]
(Note that the roles of $p_0$ and $p_\ell$ are exchanged as compared to (84).) Then using a similar argument as in the proof of Lemma 5.4, one can show that the excursion length of the chain $\tilde{S}_t^\ell$ has exponential moment. Since the increments are bounded, this implies that the excursion heights also have exponential moment, as desired. 

**Proposition 7.4.** Fix an unstable color $\ell \in \mathcal{C}_u^p$ such that $p_\ell > p_{\ell^*}$. Let $(\tilde{S}_t^\ell)_{t \geq 0}$ denote the additive functional defined in (139). Define a random variable
\[
R^\ell = \sup_{n \geq 1} \left(\min_{1 \leq s \leq n} \tilde{S}_s^\ell\right) = -\inf_{k \geq 1} \tilde{S}_k^\ell.
\]
Then there exists a constant $c > 0$ such that for each $x \geq 1$,
\[
P(R^\ell \geq x) \leq c(p_{\ell^*}/p_\ell)^x.
\]
Proof. Let \( S^f \tilde{\Gamma} \) and \( \tilde{\Gamma}_t = (S^f \tilde{\Gamma}_t, X^p(t + 1)) \) denote the Markov chains defined in (136) and (138), respectively. Let \( P_\mu \) denote the probability measure for the chain \( \tilde{\Gamma}_t \) when \( S^f \tilde{\Gamma}_0 \) is distributed as distribution \( \mu \). We denote \( P_\mu = P_\mathbf{x} \) when \( \mu \) is the Dirac mass at a particular state \( \mathbf{x} \).

We first show the assertion under the stationary distribution, that is, there exists a constant \( c' > 0 \) such that

\[
P_{\tilde{\Gamma}}(R^\ell \geq x) \leq c'(p_\ell/p_\ell^*)^x \tag{172}
\]

for all \( x \geq 1 \). Using Proposition 6.4, we first note that

\[
P_{\tilde{\Gamma}}(R^\ell \leq k) = p_\ell P_{\tilde{\Gamma}}(R^\ell \leq k + 1) + (1 - p_\ell - p_\ell^*) P_{\tilde{\Gamma}}(R^\ell = k) + p_\ell^* P_{\tilde{\Gamma}}(R^\ell = k - 1).
\]

This yields

\[
(p_\ell + p_\ell^*) P_{\tilde{\Gamma}}(R^\ell = k) = P_{\tilde{\Gamma}}(R^\ell = k) - P_{\tilde{\Gamma}}(R^\ell = k - 1)
\]

\[
= p_\ell [P_{\tilde{\Gamma}}(R^\ell \leq k + 1) - P_{\tilde{\Gamma}}(R^\ell = k - 1)]
\]

\[
= p_\ell [P_{\tilde{\Gamma}}(R^\ell = k + 1) + P_{\tilde{\Gamma}}(R^\ell = k)].
\]

Hence we get \( p_\ell P_{\tilde{\Gamma}}(R^\ell = k) = p_\ell^* P_{\tilde{\Gamma}}(R^\ell = k + 1) \), so we deduce

\[
P_{\tilde{\Gamma}}(R^\ell \geq x) = \sum_{k \geq x} P_{\tilde{\Gamma}}(R^\ell = 1)(p_\ell/p_\ell^*)^{k-1} = \frac{(p_\ell/p_\ell^*)^{1-[x]}}{1 - (p_\ell/p_\ell^*)} P_{\tilde{\Gamma}}(R^\ell = 1).
\]

This shows (172) as desired.

Now, conditioning on the initial state for \( S^f \tilde{\Gamma}_0 \), we have

\[
c'(p_\ell^*/p_\ell)^x \geq \sum_{y \in (Z)^{f^{\ell+1}}} P_{\tilde{\Gamma}^\ell}(R^\ell \geq x \mid S^f \tilde{\Gamma}_t = y) \tilde{\pi}^\ell(y)
\]

\[
= \sum_{y \in (Z)^{f^{\ell+1}}} P_{\tilde{\Gamma}}(R^\ell \geq x) \tilde{\pi}^\ell(y)
\]

\[
\geq P_0(R^\ell \geq x) \tilde{\pi}^\ell(0).
\]

The last inequality follows from the observation that if we run two carrier processes jointly over the same BBS configuration where one carrier has at least as many balls as the other one for each color, then this domination is maintained throughout the evolution. This can be shown easily by an induction. Then noting that by (137) and the definition of \( \ell^+ \), we have \( \tilde{\pi}^\ell(0) = \prod_{i=0}^{\ell^+ - 1} \left(1 - \frac{p_\ell}{p_\ell^*}\right) \eta > 0 \). This shows the assertion. \( \square \)

**Proposition 7.5.** Fix an unstable color \( \ell \in C^\mu_p \) such that \( p_\ell > p_\ell^* \). Let \( (\tilde{S}^\ell_i)^{t \geq 0} \) denote the additive functional defined in (139). Then there exists constants \( c > 0 \) such that for all \( n \geq 1 \) and \( x \geq 0 \),

\[
P\left(\max_{1 \leq i \leq n} m_\ell(\tilde{\Gamma}_t) - \tilde{S}^\ell_n \geq x\right) \leq e^{-cx}.
\]

**Proof.** According to Proposition 6.3, we can write \( m_\ell(\tilde{\Gamma}_t) - \tilde{S}^\ell_n = -\min_{0 \leq s \leq t} \tilde{S}^\ell_s \geq 0 \). Hence by union bound and triangle inequality, the probability in the assertion is bounded by

\[
P\left(\max_{1 \leq i \leq n} m_\ell(\tilde{\Gamma}_t) - \tilde{S}^\ell_n \geq x\right) \leq P\left(\max_{1 \leq i \leq n} m_\ell(\tilde{\Gamma}_t) - \tilde{S}^\ell_n \geq \frac{x}{x}\right).
\]

(181)
By Proposition 6.3, note that
\[
\mathbb{P}\left( \max_{1 \leq i \leq n} m_\ell(\tilde{T}_i) - \hat{S}_n^\ell \geq \frac{x}{k} \right) = \mathbb{P}\left( \max_{1 \leq i \leq n} \left( \hat{S}_i^\ell - \min_{0 \leq s \leq t} \hat{S}_s^\ell \right) - \hat{S}_n^\ell \geq \frac{x}{k} \right) \quad (183)
\]
\[
\leq \mathbb{P}\left( \max_{1 \leq i \leq n} \hat{S}_i^\ell - \hat{S}_n^\ell \geq \frac{x}{2k} \right) + \mathbb{P}\left( \max_{1 \leq i \leq n} - \min_{0 \leq s \leq t} \hat{S}_s^\ell \right) \geq \frac{x}{2k} \right) \quad (184)
\]
The last expression is exponentially small in \(x\) due to Propositions 7.3 and 7.4. Hence the assertion follows.

Now we are ready to prove Theorem 7.2.

**Proof of Theorem 7.2.** Suppose \(p^* = \max(p_1, \cdots, p_k) > p_0\). Let \(a_1 < \cdots < a_r\) denote the positive unstable colors. Under the assumption \(p_{a_i} = p^*\).

First we show that almost surely,
\[
\liminf_{n \to \infty} n^{-1} \lambda_1(n) \geq p^* - p_0 > 0. \quad (185)
\]
Indeed, according to Lemma 2.1, we have
\[
\lambda_1(n) = \max_{A \subseteq [1, n]} L(A, X^n) \quad (186)
\]
where \(L(A, X^n)\) denotes the penalized length of \(A\) w.r.t. \(X^n\) defined at (17). By choosing \(A \subseteq [1, n]\) to be the set of all locations of color \(i\) balls in \(X^n\), this yields
\[
n^{-1} \lambda_1(n) \geq n^{-1} \sum_{1 \leq i \leq n} \left[ I(X^n(x) = i) - I(X^n(x) = 0) \right]. \quad (187)
\]
By the strong law of large numbers, the right hand side converges almost surely to \(p_i - p_0\) as \(n \to \infty\). Since this holds for all \(1 \leq i \leq k\), the claimed lower bound (185) follows.

Next, we show that almost surely,
\[
\limsup_{n \to \infty} n^{-1} \lambda_1(n) \leq p^* - p_0. \quad (188)
\]
By Lemma 2.2 and Proposition 6.1, we have
\[
\lambda_1(n) \leq \max_{1 \leq i \leq n} m_1(\tilde{T}_i) + \cdots + \max_{1 \leq i \leq n} m_k(\tilde{T}_i). \quad (189)
\]
Note that by Propositions 6.6 and 7.1, \(n^{-1} \max_{1 \leq i \leq n} m_\ell(\tilde{T}_i)\) converges almost surely to zero if \(p_\ell \leq p_{\ell^*}\). Furthermore, suppose \(p_\ell > p_{\ell^*}\), and write
\[
\max_{1 \leq i \leq n} m_\ell(\tilde{T}_i) \leq \hat{S}_n^\ell + \left( \max_{1 \leq i \leq n} m_\ell(\tilde{T}_i) - \hat{S}_n^\ell \right). \quad (190)
\]
According to Proposition 7.5, the last two terms in the right hand side above have exponential tail, so by using Borel-Cantelli lemma, it converges to zero under scaling \(n^{-1}\). Moreover, \(n^{-1} \hat{S}_n^\ell \to p_\ell - p_0\) a.s. by Proposition 6.6. Hence dominated convergence theorem yields
\[
\limsup_{n \to \infty} n^{-1} \max_{1 \leq i \leq n} m_\ell(\tilde{T}_i) \leq p_\ell - p_0 \quad \text{a.s.} \quad (191)
\]
Combining this asymptotics, we deduce from (188) and another use of dominated convergence that, almost surely,
\[
\limsup_{n \to \infty} n^{-1} \lambda_1(n) \leq (p_{a_1} - p_{a_2}) + \cdots + (p_{a_{r-1}} - p_{a_r}) + (p_{a_r} - p_0) = p_{a_1} - p_0 = p^* - p_0, \quad (192)
\]
as desired.
Lastly, suppose that \( \ell = \alpha_1 \) is the unique unstable color. First write

\[
\frac{\lambda_1(n) - (p_\ell - p_0)n}{\sqrt{n}} = \frac{\lambda_1(n) - S_n^\ell + \tilde{S}_n^\ell - (p_\ell - p_0)n}{\sqrt{n}}.
\]

(193)

By Proposition 6.6, we know that the last term converges weakly to a constant multiple of standard normal random variable. Hence by Slutsky’s theorem, it suffices to show that the first term in the right hand side converges to zero in probability (or almost surely). To this end, we use Lemma 2.2 and Proposition 6.1 to write

\[
\max_{1 \leq i \leq n} m_\ell(\Gamma_i) \leq \lambda_1(n) \leq \max_{1 \leq i \leq n} m_1(\Gamma_i) + \cdots + \max_{1 \leq i \leq n} m_k(\Gamma_i),
\]

(194)

where \( \Gamma_i \) denotes the infinite capacity carrier process on \( X^p \). Since \( \ell = \alpha_1 = \alpha_r \) by assumption, Proposition 6.1 (ii) yields

\[
\lambda_1(n) \geq \max_{1 \leq i \leq n} m_\ell(\Gamma_i) = \max_{1 \leq i \leq n} m_\ell(\Gamma_i) \geq m_\ell(\Gamma_i) = S_n^\ell.
\]

(195)

Hence we get

\[
0 \leq \lambda_1(n) - \tilde{S}_n^\ell \leq \left( \max_{1 \leq i \leq n} m_\ell(\Gamma_i) - \tilde{S}_n^\ell \right) + \sum_{1 \leq i \leq n, i \neq \ell} \max_{1 \leq i \leq n} m_i(\Gamma_i).
\]

(196)

According to Proposition 7.5, the first term in the right hand side has exponential tail, so after dividing by \( \sqrt{n} \), it converges to zero almost surely by Borel-Cantelli lemma. Moreover, each summand in the second term is of order \( O(\log n) \) due to Proposition 6.6 and our results for the subcritical regime. Hence the second term converges to zero almost surely under the diffusive scaling \( \sqrt{n} \). It follows that \( n^{-1}(\lambda_1(n) - \tilde{S}_n^\ell) \to 0 \) a.s., as desired. \( \square \)

7.3. **Subsequent soliton lengths in the simple and non-simple supercritical regimes.** In this section, we complete the proof of Theorem 1.2 (iii) and (iv). Since we have shown the assertion for the top soliton length in Theorem 7.2 in the previous subsection, it suffices to show the following result.

**Theorem 7.6.** Suppose \( p^* > p_0 \). Fix \( \epsilon \in (0, 1) \) and \( j \geq 2 \). Then the following hold.

(i) Suppose \( p_i = p^* \) for a unique \( 1 \leq i \leq \kappa \). Then \( \lambda_j(n) = O(\log n) \) with probability at least \( 1 - \epsilon \).

(ii) Suppose \( p_i = p^* \) at least two distinct colors \( 1 \leq i \leq \kappa \). Then \( \lambda_j(n) = O(\sqrt{n}) \) and \( \lambda_j = \Omega(\sqrt{n} / \log n) \) with probability at least \( 1 - \epsilon \).

We begin with a simple observation. For \( 1 \leq i, j \leq \kappa \) and an interval \( H \), define a random variable \( D_{i,j}(H) \) by

\[
D_{i,j}(H) = \sum_{x \in H} \left[ 1(X^p(x) = i) - 1(X^p(x) = j) \right],
\]

(197)

which equals the difference of the number of color \( i \) and color \( j \) balls in \( H \) given by \( X^p \).

**Proposition 7.7.** Fix \( 1 \leq i, j \leq \kappa \) and suppose \( p_i > p_j \). Fix a finite subset \( H \subseteq \mathbb{N} \). Then for any constant \( C > 0 \),

\[
\mathbb{P}(D_{i,j}(H) \geq 2C \log n) \leq \exp(-C(p_i - p_j) \log n)
\]

(198)

for all \( n \geq 1 \).

**Proof.** Let \( \epsilon = p_i - p_j > 0 \) and denote \( |H| = m \). Note that \( \mathbb{E}[D_{i,j}(H)] = -\epsilon m \). Since \( D_{i,j}(H) \) is a sum of i.i.d. \( \pm 1 \) increments, by Hoeffding’s inequality,

\[
\mathbb{P}(D_{i,j}(H) - \mathbb{E}[D_{i,j}(H)] \geq t) \leq e^{-t^2/(2m)}
\]

(199)
for any \( t > 0 \). Let \( t = \epsilon m + 2C \log n \). Then \( t/m \geq \epsilon \), so
\[
\mathbb{P}(D_{j,i}(H) \geq 2C \log n) = \mathbb{P}(D_{j,i}(H) - \mathbb{E}[D_{j,i}(H)] \geq t) \leq e^{-t/2} \leq e^{-\epsilon C \log n}.
\] (200)
This shows the assertion. 

**Proof of Theorem 7.6.** Our argument is based on our earlier results and Lemma 2.1. In this proof, we say subset \( A \subseteq \mathbb{N} \) a non-increasing subsequence if \( X^nP \) is non-increasing on \( A \). The 'support' of \( A \) is the interval of integers \( [\min(A), \max(A)] \).

We first show the upper bounds in (i) and (ii). It suffices to obtain bounds on \( \lambda_2(n) \) in the corresponding regimes. Recall the formula for \( \lambda_1(n) + \lambda_2(n) \) given by Lemma 2.1:
\[
\lambda_1(n) + \lambda_2(n) = \max_{I \preceq J \subseteq \mathbb{N}} L(I, X^nP) + L(J, X^nP)
\] (201)

Find two non-increasing subsequences of non-interlacing support (say \( I = [a, b] \) and \( J = [c, d] \)) whose combined penalized length achieves \( \lambda_1(n) + \lambda_2(n) \). We split \( I \) into successive disjoint sub-intervals \( I_k, I_{k-1}, \ldots, I_1 \) where in each \( I_j \) we only pick the balls of color \( j \). Split \( J \) similarly. This gives us a partition of the whole interval \([1, n]\) into the following collection of disjoint sub-intervals
\[
\mathcal{H} = \{[1, a - 1], I_k, I_{k-1}, \ldots, I_1, [b + 1, c - 1], J_k, J_{k-1}, \ldots, J_1, [d + 1, n]\},
\] (202)
ordered from left to right.

For \( \lambda_1(n) \), we choose a sub-optimal non-increasing sequence by choosing all balls of color \( i \) in \([1, n]\). Denote its support by \( A^{(i)} \). Then Lemma 2.1 applied for \( \lambda_1(n) \) and (201) yield
\[
\lambda_2(n) \leq L(I, X^nP) + L(J, X^nP) - L(A^{(i)}, X^nP).
\] (203)

Then breaking the right hand side of (203) into sub-intervals given by the partition in (202), we may write
\[
L(I, X^nP) + L(J, X^nP) - L(A^{(i)}, X^nP) = \sum_{H \in \mathcal{H}} f(H),
\] (204)
where if \( H = I_j \) or \( J_j \) \((1 \leq j \leq k)\),
\[
f(H) = (\text{number of balls of color } j \text{ in } H - \text{number of balls of color } 0 \text{ in } H)
\] (205)
\[
- (\text{number of balls of color } i \text{ in } H - \text{number of balls of color } 0 \text{ in } H)
\] (206)
\[
= D_{j,i}(H),
\] (207)

else if \( H = [1, a - 1], [b + 1, c - 1] \) or \([d + 1, n]\),
\[
f(H) = - (\text{number of balls of color } i \text{ in } H - \text{number of balls of color } 0 \text{ in } H)
\] (208)
\[
= D_{0,i}(H).
\] (209)

Now suppose that \( p_i \) be the unique maximum among \( p_1, \ldots, p_\kappa \) and assume \( p_i > p_0 \). Note that \( \mathcal{H} \) contains \( 2\kappa + 3 \) intervals. Noting that \( D_{i,i}(H) = 0 \) a.s., union bound and Proposition 7.7 give
\[
\mathbb{P}\left( \sum_{H \in \mathcal{H}} f(H) \geq (2\kappa + 3)C \log n \right) \leq \sum_{H \in \mathcal{H}} \mathbb{P}(f(H) \geq C \log n) \leq 3 \sum_{0 \leq \ell \leq \kappa} \exp(-C(p_i - p_j)\log n)
\] (210)

for any fixed constant \( C > 0 \). Since the last expression tends to zero as \( n \to \infty \), this shows \( \lambda_2 = O(\log n) \) with high probability.

Next, suppose \( p_i = p^* \) at least two distinct colors \( 1 \leq i \leq \kappa \), and let \( p_i = p^* \) for some \( 1 \leq i \leq \kappa \). If we compare the number of balls of color \( j \) in \( H \) minus the number of balls of color \( i \) in \( H \). By using a
similar argument, it is $O(\log n)$ with high probability as long as $p_j < p^*$. If $p_j = p^*$, then by the triangle inequality,

$$D_{j,i}(H) \leq 2 \max_{1 \leq i \leq n} |D_{j,i}([1,t])|.$$  \hspace{1cm} (212)

In this case $D_{j,i}([1,t])$ is a simple symmetric random walk with $t$ increments. Hence for some large enough constant $C > 0$, the last quantity is at most $C\sqrt{n}$ with probability at least $1 - e$ by the central limit theorem. This shows $\lambda_2(n) = O(\log n)$ with high probability.

Now we prove the lower bounds in (i) and (ii). We first show that $\lambda_r(n) = \Omega(\log n)$ with high probability for any fixed $r \geq 2$. Consider an optimal choice of $j - 1$ non-interlacing non-increasing subsequences that achieves $\lambda_1(n) + \cdots + \lambda_{r-1}(n)$ in the sense of Lemma 2.1. The union of their support has total length at least $cn$ with high probability for some constant $c > 0$, so with high probability there is at least one subsequence of linear sized support. In such an interval, we can find consecutive $c \log n$ zeros whp. We can simply remove those zeros by considering two non-increasing subsequences instead of one. This gives rise to $r$ non-increasing disjoint subsequences whose total penalized length has now increased by at least $c \log n$. According to Lemma 2.1, this shows $\lambda_r(n) = \Omega(\log n)$ with high probability. In particular, this complete the proof of (i).

Lastly, we show for any fixed $r \geq 2$, there exists a constant $c > 0$ such that $\lambda_j(n) \geq c\sqrt{n}$ for all $n \geq 1$ with high probability. As before, we take an optimal choice of $(r-1)$ non-interlacing non-increasing subsequences that achieves $\lambda_1(n) + \cdots + \lambda_{r-1}(n)$. The total length of their support is at least linear in $n$ with high probability, so for some small constant $c_1 > 0$, at least one of them has to have length at least $c_1 n$ with high probability. Let $I$ be the support of such a subsequence. We split $I$ into successive disjoint sub-intervals $I_k, I_{k-1}, \cdots, I_1$ where in each $I_j$ we only pick the balls of color $j$ from $I$. Note that there should be come $I_j$ such that $p_j = p^*$ and $I_j$ has linear size, otherwise our choice of $r-1$ subsequences it will be non-optimal with high probability. Denote $I_1 = [e, f]$ where $f \geq e + c_1 n$.

Suppose $i'$ is another color such that $p_{i'} = p^*$. Suppose without loss of generality $i > i'$. We claim that there exists a constant $c > 0$ such that

$$\lim_{n \to \infty} P \left( \max_{t \in [1, c_1 n]} D_{i',i}([e, e + t]) \leq \frac{c \sqrt{n}}{\log n} \right) = 0.$$  \hspace{1cm} (213)

Suppose this claim holds. If we are allowed to increase the number of non-increasing subsequences by 1, then we can split the interval $I_1 = [e, f]$ into $[e, m]$ and $[m+1, f]$, where in the first subinterval we choose color $i$ balls and in the second one, we choose color $i'$ balls, while keeping everything else unchanged. According to the claim above, with high probability, we can choose $m$ to be the location where the maximum in (213) is achieved. Then we will gain at least $c \sqrt{n}/\log n$ while going from $\lambda_1(n) + \cdots + \lambda_{r-1}(n)$ to $\lambda_1(n) + \cdots + \lambda_r(n)$. This shows that the desired lower bound on $\lambda_r(n)$.

It suffices to show the claim above. Noting that $e \in [1, n]$ is random, we use a union bound to write

$$P \left( \max_{t \in [1, c_1 n]} D_{i',i}([e, e + t]) \leq \frac{c \sqrt{n}}{\log n} \right) \leq P \left( \max_{t \in [1, c_1 n]} D_{i',i}([x, t]) \leq \frac{c \sqrt{n}}{\log n} \text{ for some } x \in [1, n] \right)$$  \hspace{1cm} (214)

$$\leq \sum_{x \in [1, n]} P \left( \max_{t \in [1, c_1 n]} D_{i',i}([x, x + t]) \leq \frac{c \sqrt{n}}{\log n} \right)$$  \hspace{1cm} (215)

$$= n P \left( \max_{t \in [1, 1 + c_1 n]} D_{i',i}([1, t]) \leq \frac{c \sqrt{n}}{\log n} \right).$$  \hspace{1cm} (216)

where we have used translation invariance of $D_{i',i}([x, x + t])$ for the last equality. By the central limit theorem, the probability in the last expression is polynomially small in $n$, and we can make this $o(n)$ by
taking \( c > 0 \) small enough. Thus, there exists a constant \( c > 0 \) such that the probability in the left hand side tends to zero. This shows the assertion. 

\[ \square \]

8. Proof of probabilistic lemmas

8.1. Proof of Lemma 5.4. In this subsection, we show that the return time to the origin for the subcritical carrier process has finite exponential moment, which was claimed in Lemma 5.4.

Suppose \( p_0 > \max(p_1, \cdots, p_k) \) and let \( \Gamma_t \) denote the infinite capacity carrier process. Denote \( X_i^t = m_i(\Gamma_t) \) so that the carrier state at time \( t \) is given by \( \Sigma_t = (X_1^t, \cdots, X_k^t) \). Note that for each \( 1 \leq i \leq k \), the projection \( (X_i^t, X_{i+1}^t, \cdots, X_k^t)_{t \geq 0} \) is itself an irreducible and aperiodic Markov chain with unique stationary distribution given in (169). For each \( 1 \leq i \leq \kappa \), let \( \sigma_j; i \) denote the \( j \)th return time of \( (X_i^t, X_{i+1}^t, \cdots, X_k^t)_{t \geq 0} \) to the origin.

**Proposition 8.1.** Fix \( 1 \leq i \leq \kappa \), let \( Y_t = \tilde{S}_i^t \) denote the additive functional defined in (139). Then

\[
\mathbb{E}[Y_{\sigma_1; i+1}] < 0. \tag{217}
\]

**Proof:** Denote \( \tilde{S}_j = \sigma_j; i+1 \). We make a key observation that the excursions from 0 for the recurrent chain \( (X_{i+1}^t, X_{i+2}^t, \cdots, X_k^t)_{t \geq 0} \) are i.i.d. More precisely,

\[
(S_j - S_{j-1}, \text{color of the } t\text{-th ball for all } t \in [S_j, S_{j+1})) \text{ are i.i.d.} \tag{218}
\]

Hence \( (Y_{S_j})_{j \geq 0} \) is a random walk.

Now, let \( \pi \) denote the unique stationary distribution of \( (X_{i+1}^t, X_{i+2}^t, \cdots, X_k^t)_{t \geq 0} \) given in (169). Then by Markov chain ergodic theorem and Proposition 6.4,

\[
\lim_{t \to \infty} \frac{1}{t} \sum_{s=1}^{t} Y_s = \mathbb{E}_\pi[Y_{t+1} - Y_t] = p_i - p_0 < 0. \tag{219}
\]

So, \( Y_{S_j}/S_j \to p_i - p_0 \) almost surely. On the other hand, by law of large numbers, \( Y_{S_j}/j \to \mathbb{E}[Y_{S_j}] \) almost surely. Also note that \( \mathbb{E}[S_1] < \infty \) by Kac’s theorem (see, e.g., [LP17, Lem 21.13]). Hence

\[
\mathbb{E}[Y_{S_j}] = \lim_{j \to \infty} \frac{Y_{S_j}}{j} = \lim_{j \to \infty} \frac{Y_{S_j}}{S_j} \cdot \frac{S_j}{j} = (p_i - p_0)\mathbb{E}[S_1] < 0. \tag{220}
\]

\[ \square \]

**Proposition 8.2.** Fix \( 1 \leq i \leq \kappa \) and denote \( W_j = X_{j;i+1}^i \). Suppose \( \sigma_{j; i+1} \) has finite exponential moment.

(i) There exists constants \( c, K > 0 \) such that

\[
\mathbb{E}[W_j - W_{j-1} | W_{j-1} = m] \leq -c \quad \forall m \geq K. \tag{221}
\]

(ii) The chain \( (W_j)_{j \geq 1} \) is geometrically ergodic.

(iii) The first return time of \( W_j \) to the origin has finite exponential moment.

**Proof.** (iii) follows from (ii) from the geometric ergodic theorem [MT12, Thm. 15.0.1]. Hence it is enough to show (i) and (ii).

To show (i), we first write

\[
\mathbb{E}[W_1 - W_0 | W_0 = m] = \mathbb{E}[X_{S_1}^i - X_0^i \mid X_0^i = m] \tag{222}
\]

\[
= \mathbb{E}[(X_{S_1}^i - X_0^i)1_{S_1 \leq m} \mid X_0^i = m] + \mathbb{E}[(X_{S_1}^i - X_0^i)1_{S_1 > m} \mid X_0^i = m] \tag{223}
\]

\[
= \mathbb{E}[Y_{S_1}]1_{S_1 \leq m} + \mathbb{E}[(X_{S_1}^i - X_0^i)1_{S_1 > m} \mid X_0^i = m] \tag{224}
\]
\[
\gamma^2 = \prod_{j=0}^{r} \left( \prod_{i=\alpha_j+1}^{\alpha_{j+1}-1} \left( 1 - \frac{p_i}{p_{\alpha_{j+1}}} \right) \right) p_0 \mathbb{E} \left[ \left( \sum_{k=1}^{\tau_0} g_{\ell}(\tilde{Z}_k^\ell) - (p_{\ell} - p_\ell) \right)^2 \right],
\]

where the expectation assumes the all zero initial state. From the above expression, it is clear that \( \gamma^2_r > 0 \).
To show $\gamma^2 < \infty$, note that since $|g_\ell| \leq 1$, we have

$$
\mathbb{E}(0,0) \left( \left( \sum_{\ell=1}^{\tau_0} g_\ell(\tilde{Z}_k^\ell) - (p_\ell - p_\ell^\ell) \right)^2 \right) \leq 9 \mathbb{E}(0,0) [\tau_0^2].
$$

(232)

Hence it suffices to show $\mathbb{E}[\tau_0^2] < \infty$. In fact, $\tau_0$ has finite exponential moment. Indeed, let $\hat{S}^f \tilde{T}_t = (m_{\ell+1}(\tilde{T}_t), \ldots, m_{\ell-1}(\tilde{T}_t))$ denote the Markov chain defined in (136). Recall that this is an irreducible and aperiodic chain with unique stationary distribution $\pi^\ell$ defined in (137). Moreover, observe that this chain is a lazy version of the subcritical infinite capacity carrier process, where the roll of $p_0$ is played by $p_\ell^\ell$. Hence according to Lemma 5.4, we know that the return time to the origin for $\hat{S}^f \tilde{T}_t$ has finite exponential moment. This shows the assertion. \hfill \Box

9. Proof of combinatorial lemmas

In this section, we establish various combinatorial statements about $\kappa$-color BBS dynamics and the associated carrier processes. Our main goal is to show Lemma 2.1 and 2.2. We also provide an elementary and self-contained proof of Lemma 2.3, which has been proved in the more general form in [KL18, Prop. 4.3] using connections with combinatorial $R$.

9.1. Time invariants of the $\kappa$-color BBS. Recall the notations introduced in Subsection 2.1. For any $\kappa$-color BBS configuration $X : \mathbb{N} \rightarrow \mathbb{Z}_{\kappa+1}$ with finite support and integer $k \geq 1$, we denote

$$
R_k(X) = \max_{A_1 \cup \cdots \cup A_k} \sum_{i=1}^{\kappa} NA(A_i, X), \quad L_k(X) = \max_{A_1 < \cdots < A_k \subset \mathbb{N}} \sum_{i=1}^{\kappa} L(A_i, X).
$$

(233)

Lastly, we also denote

$$
E_k(X) = \sum_{s=1}^{\infty} 1(X(s) > \min_{\Gamma_{s-1} \cap \{k\}})
$$

(234)

where $(\Gamma_i \cap \{k\})_{i \in \mathbb{N}}$ is the capacity-$i$ carrier process over $X$. We set $R_0(X) = L_0(X) = E_0(X) = 0$ for convenience. In this subsection, we will show with an elementary argument that the above quantities associated with a $\kappa$-color BBS configuration are invariant under time evolution. This will lead to the proof of Lemmas 2.1 and 2.3.

We remark that the invariants $E_k(X)$ are called the energy. They were first introduced in [FYO00] for the $\kappa = 1$ BBS, and were recently used to define an energy matrix for the general $\kappa$-color BBS that characterizes the full set of invariants. Time invariance of the energy (and also the energy matrix) in the literature is usually shown by using the alternative characterization of the BBS dynamics in terms of combinatorial $R$ and connections to the Yang-Baxter equation [FYO00, IKT12, KL18, KLO18b].

Recall the BBS evolution rule defined in the introduction: For $i = \kappa, \kappa-1, \cdots, 1$, the balls of color $i$ each make one jump to the right, into the first available empty box (site with color 0), with balls that start to the left jumping before balls that start to their right. (This is the map $K_i$ defined in the introduction.) A single step of $\kappa$-color BBS evolution $X \rightarrow X'$ is defined by

$$
X' = K_1 \circ K_2 \circ \cdots \circ K_\kappa(X).
$$

(235)

We propose two ways to simplify the $\kappa$-color BBS dynamics. First, using the cyclic symmetry of the system, we can reformulate the update of a $\kappa$-color BBS configuration in terms of $\kappa$ applications of a single rule. Namely, let $\mathcal{T}_\kappa$ denote the following update rule for BBS configurations with finite support:
all the balls of color \( \kappa \) jump according to the rule \( K_\kappa \), and we relabel each of them with color 1 and increase the positive colors of all other balls by 1. Then we have

\[
K_1 \circ K_2 \circ \cdots \circ K_\kappa (X) = (\mathcal{F}_\kappa)^j (X).
\]

Second, we introduce “standardization” of BBS dynamics, which allows us to only consider BBS configurations with no repeated use of any positive color. Namely, given a \( \kappa \)-color BBS configuration \( X : \mathbb{N} \to \mathbb{Z}_{\geq 0} \) of finite support, we define its standardization to be the following map \( \hat{X} : \mathbb{N} \to \mathbb{Z}_{\geq 0} \): For each \( 1 \leq i \leq \kappa \), let \( m_i \) denote the number of balls in \( X \) of color \( i \). Then to produce \( \hat{X} \), we relabel first the color 1 balls from 1 to \( m_1 \) from right to left (so that the leftmost ball that was previously colored 1 is now colored \( m_1 \)), and then the original color 2 balls are relabeled with colors \( m_1 + 1 \) to \( m_1 + m_2 \) from right to left, and so on. Thus, if \( N = \sum_{i=1}^\kappa m_i \) is the total number of balls of positive color then \( \hat{X} \) is an \( N \)-color BBS configuration with each color in \( \{1, \ldots, N\} \) used for exactly one ball.

**Proposition 9.1.** Let \( X \) and \( \hat{X} \) denote a \( \kappa \)-color BBS configuration with finite support and its standardization, respectively. Then the following hold.

(i) Standardization preserves number of ascents, non-interlacing non-increasing sequences and their penalized lengths. In particular, for each \( k \geq 1 \),

\[
R_k (X) = R_k (\hat{X}), \quad L_k (X) = L_k (\hat{X}).
\]

(ii) \( X \) and \( \hat{X} \) give the same soliton partition, i.e., \( \Lambda (X) = \Lambda (\hat{X}) \).

**Proof.** By construction, standardization preserves ordering in the following sense: for \( y < z \), one has \( X(y) < X(z) \) if and only if \( \hat{X}(y) < \hat{X}(z) \). Thus, a given sequence of balls has an ascent in \( X \) if and only if it has an ascent in \( \hat{X} \), and likewise a given sequence of balls is non-increasing in \( X \) if and only if it is non-increasing in \( \hat{X} \). Part (i) follows immediately.

To show (ii), denote by \( X' \) and \( (\hat{X})' \) the BBS configurations obtained by applying one step of the BBS evolution rule to \( X \) and \( \hat{X} \), respectively. Since standardization does not change the location of balls, it suffices to show that standardization commutes with BBS time evolution rules, i.e.,

\[
\hat{X}' = (\hat{X})'.
\]

To see this, observe that for the evolution \( X \to X' \), after all balls of color \( \kappa \) have jumped, they return to the same left-right order as before: if some ball of color \( \kappa \), say in position \( x \), jumped over some other ball of color \( \kappa \), say in position \( y \), to land in position \( z \) (so \( x < y < z \)), it must be the case that sites between \( y \) and \( z \) were occupied. Therefore, when it is time for the ball in position \( y \) to jump, it jumps over all sites in \( (y, z] \). Hence in the first step, the balls of color \( \kappa \) in the previous step are triggered one by one from left to right, and since they restore the same left-right order, they will continue to be triggered in this order in all future steps. This exactly agrees with the time evolution \( \hat{X} \to \hat{X}' \). This shows (238), as desired.

In the following proposition, we show the time-invariance of the three quantities associated to a given BBS configuration. This will show most of Lemma 2.1.

**Proposition 9.2.** Let \( X \) be an arbitrary \( \kappa \)-color BBS configurations of finite support. Fix \( j \geq 1 \). The following hold.

(i) \( E_j (X) = E_j (\mathcal{F}_\kappa (X)) \).

(ii) \( R_j (X) = E_j (X) \).

(iii) \( L_j (X) = L_j (\mathcal{F}_\kappa (X)) \).

(iv) If \( (X_t)_{t \geq 0} \) denotes the \( \kappa \)-color BBS trajectory with \( X = X_0 \), then

\[
E_j (X_t) = R_j (X_t) = \text{Const.}, \quad L_j (X_t) = \text{Const.}
\]
We first derive Lemmas 2.1 and 2.3 assuming Proposition 9.2.

**Proof of Lemma 2.1 and 2.3.** Let \((X_t)_{t \geq 0}\) be a \(\kappa\)-color BBS trajectory such that \(X_0\) has finite support. We take \(T \geq 1\) large enough so that at time \(T\) the system decomposes into non-interacting solitons whose lengths are non-decreasing from left to right. We can reformulate the condition that a \(\kappa\)-color BBS configuration has reached its soliton decomposition as follows: Suppose two consecutive solitons are separated by \(g\) 0’s, where the left and right solitons have length \(l\) and \(r\), where ‘length’ of a soliton is its number of balls of positive colors. Suppose the gap is small, i.e., \(g < l\). In order for the left soliton to be preserved during the update \(X_T \rightarrow X_{T+1}\), all balls in the left soliton must be dropped by the carrier before any balls in the right soliton are dropped. It follows that for each \(i \geq 1\), the following ‘separation condition’ must hold at time \(T\):

\[
\text{The } i\text{th largest entry of the right soliton is strictly larger than the } i + g\text{th largest entry of the left soliton. (240)}
\]

When \(\kappa = 1\), this simply asserts that each soliton of length \(l\) must be followed by at least \(l\) empty sites. This is not the case for \(\kappa > 1\), as illustrated in the example

\[
\cdots00433200431100\cdots.
\]

For each \(k \geq 1\), let \(\lambda_k\) denote the length of the \(k\)th-longest soliton, and let \(\rho_k\) denote the number of solitons of length \(\geq k\). They both form the same Young diagram, whose \(k\)th column and row lengths are given by \(\lambda_k\) and \(\rho_k\), respectively.

For each \(j \geq 1\), let \((T_{s;j})_{s \geq 0}\) denote the capacity-\(j\) carrier process on \(X_T\). As the carrier process over \(X_T\) runs over a soliton of length \(k\), the carrier obtains \(\min(k, j)\) contribution to the energy. When the carrier was empty at the beginning of the solution, this is clear, and otherwise, it is still true due to the separation condition (240). Hence we have

\[
E_j(X_T) = \sum_{k=1}^{\infty} \min(\lambda_k, j) = \sum_{k=1}^{j} \rho_k.
\]

Then by Proposition 9.2, we deduce

\[
R_j(X_t) = E_j(X_t) = E_j(X_T) = \sum_{k=1}^{j} \rho_k
\]

for all \(t \geq 0\), as desired. In the general case, the above equations hold due to the separation condition (240). This shows Lemma 2.3 as well as the first equation in Lemma 2.1.

Similarly, for the second equation in Lemma 2.1, it suffices to show \(L_j(X_T) = \lambda_1 + \cdots + \lambda_j\). It is easy to see \(L_j(X_T) \geq \lambda_1 + \cdots + \lambda_j\) by choosing the \(j\) longest non-increasing sequences given by the top \(j\) solitons. It remains to show the converse inequality, choose a collection of non-interlacing non-increasing subsequences on supports \(A_1, A_2, \cdots, A_j\) that achieves \(L_j(X_T)\). We may assume that \(|A_1| + \cdots + |A_j|\) is as small as possible, where \(|\cdot|\) means (non-penalized) cardinality. We claim that every \(A_i\) is contained in the support of a single soliton (where it has positive colors). Then clearly the maximum sum of penalized lengths are achieved when \(A_i\)'s are the support of the \(j\) longest non-increasing sequences given by the solitons, which shows the assertion.

To show the claim, for each \(i \geq 1\), let \(u_i\) denote the maximal non-increasing subsequence of positive colors in the \(i\)th longest soliton in \(X_T\). Schematically, we can write \(X_T\) as

\[
X_T: \cdots u_30\cdots 0u_20\cdots 0u_100\cdots.
\]

Let \(l_i\) denote the number of 0’s between \(u_{i+1}\) and \(u_i\).
Suppose for contradiction that some $A_k$ intersects with two $u_i$'s. Let $i$ be as small as possible so that $A_k$ intersects with $u_{i+1}$ and $u_i$. We first suppose the case when the two solitons have sufficient gap, i.e., $\ell_{i+1} > \lambda_i + 1$. Let $A'_{k'} = A_k \setminus u_{i+1}$. Then $A_1, \ldots, A_{k-1}, A'_{k'}, A_{k+1}, \ldots, A_j$ is a sequence of non-interlacing non-increasing subsequences in $X_T$ with strictly smaller total number of elements than the original sequence. Moreover, this new sequence achieves the optimum $L_j(X_T)$ since

$$L(A_{k'}, X_T) \geq L(A_k, X_T) - u_{i+1} + \ell_i \geq L(A_k, X_T).$$

Namely, omitting all elements of $u_{i+1}$ from $A_k$ deletes at most $|u_{i+1}|$ positive numbers but at least $\ell_i \geq |u_{i+1}|$ zeros. This contradicts the minimality of the original sequence $A_1, \ldots, A_j$. This shows the claim. Lastly, when the gap between the solitons is small, i.e., $\ell_{i+1} < \lambda_k$, one can argue similarly by using the separation condition (240). This shows the claim, as desired.

Lastly in this subsection, we prove Proposition 9.2.

**Proof of Proposition 9.2.** (iv) immediately follows from (i)-(iii). According to Proposition 9.1, the assertion is valid for arbitrary BBS if and only if it is true for the standardized system with initial configuration $X$, where each positive color is used exactly once. Hence, without loss of generality, we may assume that each positive color in $X$ is used exactly once. Furthermore, in proving (i)-(iii), we may assume that there is a ball of color $\kappa$ in $X$, since otherwise the cyclic update rule $F_\kappa$ simply increases all positive colors by 1. Since all the invariants depend only on the relative ordering between ball colors, the assertion holds trivially. We will also denote $X' = F_\kappa(X)$.

(i) Suppose $X(x) = \kappa$ and the ball of color $\kappa$ is in a contiguous block of balls whose labels are $u v$ for some words $u, v$. After the update $X \xrightarrow{\kappa} X' := F_\kappa(X)$, we reach an arrangement in which $u$ and $v$ have had their labels incremented, the space between them is empty ($X'(x) = 0$), and 1 follows $v$. Let $y$ be the site such that $X'(y) = 1$. Here is a schematic:

| configuration | arrangement |
|---------------|-------------|
| $X$           | $\cdots \ 0 \ [\cdots \ u \ \cdots]$ | $\kappa$ | $\cdots \ v \ \cdots \ 0 \ \cdots$ |
| $X' = F_\kappa(X)$ | $\cdots \ 0 \ [\cdots \ u + \cdots]$ | $0$ | $\cdots \ v + \cdots \ 1 \ \cdots$ |

Consider running the capacity-$j$ carriers over $X$ and $F_\kappa(X)$ and computing their energies $E_j(X)$ and $E_j(X')$. Let the corresponding carrier processes be denoted by $(\Gamma_t)_{t \geq 0}$ and $(\Gamma'_t)_{t \geq 0}$, respectively. Observe that up to time $x - 1$, the two carriers go through the equivalent environments $[0 \cdots 0 u]$ and $[0 \cdots 0 (v + 1)]$, so $\Gamma'_t x - 1$ can be obtained from $\Gamma x - 1$ by adding 1 to all positive colors in the latter carrier. It follows that the contributions to the energies of both carriers up to this point are the same.

Next, after inserting $X(x) = \kappa$ and $X'(x) = 0$ into these carriers, we get carrier states $\Gamma x = [\kappa, A, 0 \cdots 0]$ and $\Gamma' x = [A + 1, 0 \cdots 0]$ for some positive decreasing sequence $A$ (see Figure 9 left). This only adds 1 to the energy for the carrier $\Gamma$. Also note that, since $\kappa$ is the unique largest color in the system, it sits in the carrier $\Gamma$ and does not interact with any other incoming balls thereafter. We can think of this as the capacity of the carrier $\Gamma$ being decreased to $j - 1$ after time $x$. Then over the interval $(x, \infty)$, the carriers go through the input $[v 0 0 \cdots]$ and $[(v + 1)0 0 \cdots]$, respectively.

Ignoring $\kappa$ in the carrier $\Gamma$ and shift by 1, they both have the same dynamics (and hence the same contribution to the energy) until the first time $x^*$ that $\Gamma x^*$ is full and a new ball of color $X(x^* + 1) = q > \min \Gamma x^*$. In this case, $q + 1$ replaces 0 in $\Gamma' x^*$ but it replaces $\kappa$ in $\Gamma x^*$. If such $x^*$ is not encountered up to the location $y$ of 1 in $X'$, then at site $y$, 0 replaces the maximum
entry in $\Gamma_j$, but 1 replaces 0 in $\Gamma'_j$, so this makes up the energy gap of 1 between the two carriers. Otherwise, suppose there exists such $x^*$ between $x$ and $y$. Then we can write the carrier states as $\Gamma_{x^*} = [x, B]$ and $\Gamma'_{x^*} = [B + 1, 0]$ for some positive decreasing sequence $B$ of length $j - 1$. Then since $X(x^* + 1) = q > \min \Gamma_{x^*}$, inserting $q$ (resp., $q + 1$) into $\Gamma_{x^*}$ (resp., $\Gamma'_{x^*}$) replaces $\kappa$ (resp., 0), only adding 1 to the energy for $\Gamma_j$. Then $\Gamma_{x^* + 1} = [B, q]$ and $\Gamma'_{x^* + 1} = [B + 1, q + 1]$ and all colors in $\Gamma'_j$ are at least 2, so inserting 0 and 1 at site $y$ do not increment energies of both carrier. Hence they end up with the same energy. This shows the assertion.

(ii) Let $(\Gamma_j)_{j\geq 0}$ denote the capacity-$j$ carrier process over $X$. For each $1 \leq i \leq j$, let $A_i$ denote the set of all sites $x$ such that $X(x) > \min \Gamma_{x-1}$ and inserting $X(x)$ into carrier $\Gamma_{x-1}$ replaces its $i$th entry. Then $A_1, \cdots, A_j$ are disjoint subsets, and since the energy $E_j$ gets increment 1 exactly when one of these subsets gets an ascent, this shows

$$R_j(X) \geq \sum_{i=1}^{j} NA(A_i, X) = E_j(X).$$

(246)

For the other direction, suppose that $R_j(X)$ is achieved by a collection of disjoint sets $A'_1, \cdots, A'_j$ that is different from the sets $A_1, \cdots, A_j$ computed by the carrier process. Find the first place that they differ, say that $x$ belongs to $A_i$ but to $A'_i$, for $i^* \neq i$. Then perform the following surgery: let

$$A''_\ell = \begin{cases} ([1, x] \cap A_i) \cup ((x, \infty) \cap A'_i) & \text{if } \ell = i \\ ([1, x] \cap A_i^*) \cup ((x, \infty) \cap A'_i) & \text{if } \ell = i^* \\ A'_\ell & \text{otherwise.} \end{cases}$$

(247)

Then by construction, this new collection of sets $A''_1, \cdots, A''_j$ has at least as many ascents as the $A'_i$-sequences do, and the point of disagreement with the $A'_i$s is moved later. Therefore repeating this process eventually produces the sets $A_1, \cdots, A_k$, and does not decrease the number of ascents. This shows $R_j(X) \leq E_j(X)$, as desired.

(iii) Let $L_j^{\text{new}} := L_j(X')$. We wish to show $L_j = L_j^{\text{new}}$. We begin by showing that $L_j \leq L_j^{\text{new}}$. In the original system $X$, fix a set of $k$ non-interlacing decreasing subsequences whose sum of penalized lengths is the maximum value $L_j$. We will produce a set of non-interlacing decreasing subsequences in $X'$ that have the same sum of penalized lengths. We call the unique ball of color $N$ in $X'$ by simply $N$. Suppose $N$ is in position $a$, and that positions $a + 1, a + 2, \ldots, b - 1$ have balls in them, but that position $b$ is empty; let $I = \{a, \cdots, b - 1\}$. There are cases, depending on two different questions: whether $N$ is part of a decreasing subsequence, or is in the interval spanned by a decreasing subsequence, or neither; and whether there is a decreasing subsequence whose interval spans $b$, or one that ends in $I$ with no other sequence that spans $b$, or neither.

If $N$ belongs to a decreasing subsequence, it is the largest entry. Therefore removing it decreases the length by 1 and does not add a penalty (because the gap created is not in the interior of any remaining sequence). If $N$ is in the interval spanned by a decreasing subsequence but doesn’t belong to it, removing $N$ introduces a gap and so penalizes the length of that sequence.

![Figure 9. Two capacity-$j$ carriers over X and X’ = F(X). They end up with the same energy.](image-url)
by 1. If neither hold, removing \(N\) does not change the penalized lengths of any subsequences. Adding 1 to every ball label does not change the penalized lengths of any subsequences. If a sequence spans \(b\) then inserting the new ball 1 removes a gap from that sequence, so increases its penalized length by 1. If a sequence ends in \(I\) and no subsequence sequence spans \(b\), then the 1 inserted in position \(b\) can be appended to this sequence; there are no gaps in \(I\), so this increases the penalized length by 1. And if neither hold, then inserting 1 does not change the penalized lengths of any of the subsequences. Finally, it is enough to observe that in either of the cases that result in a decrease of 1, it is necessarily the case that some sequence ends in \(I\) or spans \(b\). Thus, \(L_j \geq L_j',\) as claimed.

Finally, to show that actually \(L_j' \leq L_j\), we apply the “reverse-complement” operation, reversing the order of \(Z\) and the order of the labels. This preserves decreasing subsequences, the non-interlacing relation between them, and their penalized lengths; moreover, one time-step in the reverse-complement is exactly the reverse-complement of one inverse time-step in the original. Thus also \(L_j \geq L_j'\). This shows \(L_j = L_j'\), as desired.

9.2. Lemmas for finite capacity carrier processes. In this subsection, we prove Proposition 4.2 and Lemma 2.2.

**Proof of Proposition 4.2.** Fix a \(k\)-color BBS configuration \(X : \mathbb{N} \to Z_{k+1}\). Denote by \((\Gamma_{t,c})_{t \geq 0}\) and \((\Gamma_{t,c+1})_{t \geq 1}\) the carrier processes over \(X\) with finite capacities \(c\) and \(c + 1\), respectively. We will show the assertion by induction on \(t \geq 0\). For \(t = 0\), both carriers are filled with zeros so omitting any entry of \(\Gamma_{0,c+1}\) gives \(\Gamma_{0,c}\). For the induction step, suppose the assertion holds for some \(t \geq 1\). Denote \(T = \Gamma_{t,c}, S = \Gamma_{t+1,c} \in \mathcal{R}_c\) and \(T' = \Gamma_{t,c+1}, S' = \Gamma_{t+1,c+1} \in \mathcal{R}_{c+1}\). Recall that the entries in carrier states are non-decreasing from left to right, which is the opposite to the convention for semistandard Young tableau (as used in [KL18] and [KLO18b]).

By the induction hypothesis, we may assume that \(S\) can be obtained from \(T\) by omitting its \(j_\ast\)th entry entry \(T(j_\ast) = r\). Let \(B\) and \(A\) be the blocks to the left and right of the entry \(T(j_\ast)\) of \(T\). Hence \(S\) is the concatenation of the blocks \(B\) and \(A\) (see Figure 10 left). Let \(q := X(t + 1)\).

![Figure 10](image-url)

**Figure 10.** (Left) \(S \in \mathcal{R}_c\) is obtained from \(T \in \mathcal{R}_{c+1}\) by omitting an entry \(r\). (Right) After inserting \(q\) into \(T\) and \(S\) according to the circular exclusion rule, one can still omit a single entry from the larger tableau to get the smaller one.

First, suppose that \(q\) does not exceed the smallest entry of \(T\). In this case inserting \(q\) into \(T\) replaces the largest entry of \(T\), so \(T'\) is given by \(T'(j) = T(j + 1)\) for \(1 \leq j \leq c\) and \(T'(c + 1) = q\). We also have \(S'(j) = S(j + 1)\) for \(1 \leq j < c\) and \(S'(c) = q\). It follows that \(S'\) is obtained by omitting the same entry \(r = T'(j_\ast + 1)\) from \(T'\).

Second, suppose that \(q\) exceeds the smallest entry of \(T\). so that \(T'\) is computed from the pair \((T, q)\) using the reverse bumping. If \(q\) replaces some entry of \(A\) or \(B\) in \(T\) to get \(T'\), then the same replacement occurs to compute \(S'\) from the pair \((S, q)\). Hence in this case \(S'\) is obtained by omitting \(r = T'(j_\ast)\) from \(T'\). Otherwise, \(q\) replaces \(r\) in \(T\) to get \(T'\) (see in Figure 10 right). Then \(q\) must replace the largest entry of \(A\) in \(S\) to get \(S'\). Then \(S'\) is obtained from \(T'\) by deleting the largest entry in \(A\). This shows the assertion. \(\square\)
Proof of Lemma 2.2. Fix a \( k \)-color BBS configuration \( X : \mathbb{N} \to \mathbb{Z}_{k+1} \). For each integer \( c \geq 1 \), let \( (\Gamma_{t,c})_{t \geq 0} \) denote the capacity-\( c \) carrier process over \( X \). Let \( (\Gamma_t)_{t \geq 0} \) denote the infinite capacity carrier process over \( X \). We also write

\[
M = \max_{s \geq 0} (\# \text{ of nonzero entries in } \Gamma_s)
\]  
(248)

Note that from Lemma 2.3, we can deduce that for any \( 1 \leq j \leq \rho_1(\Lambda(X)) \),

\[
\lambda_j(\Lambda(X)) = \max \{ c \geq 1 \mid E_c(X) \geq E_{c-1}(X) - j \},
\]  
(249)

where \( E_k(X) \) is defined in (234).

Let \( \tau_c \) be the first time \( t \) that the carrier \( \Gamma_{t,c} \) is completely full with nonzero entries and \( X_0(t+1) > 0 \) does not exceed the smallest entry of \( \Gamma_{t,c} \). More precisely, let

\[
\tau_c := \inf \{ x \geq 0 \mid \Gamma_{t,c} \text{ contains all positive entries and } 0 < X_0(x+1) \leq \min \Gamma_{t,c}(x) \}.
\]  
(250)

We let \( \tau_c = \infty \) if the set in the right hand side is empty. Note that if we consider two carrier processes \( \Gamma_{t,c} \) and \( \Gamma_{t,c+1} \), then \( \tau_c + 1 \) is the first time that they contain distinct sets of nonzero entries. Moreover, \( \Gamma_{t,c+1,c+1} \) has \( c+1 \) nonzero entries. Hence if \( c \geq M \), then \( \tau_c = \infty \) and the two carrier processes have the same set of nonzero entries for all times. It follows that

\[
E_c = \text{Const.} \quad \forall c \geq M.
\]  
(251)

Hence \( \lambda_1(\Lambda(X)) \leq M \) by (249).

On the other hand, note that \( t^* := \tau_{M-1} < \infty \) and \( X(t+1) \) does not exceed the smallest entry in \( \Gamma_{t,M-1} \) by definition of \( \tau_{M-1} \). So \( 1(X(t^*+1) > \min \Gamma_{t,M-1}) = 0 \). Also, since \( \Gamma_{t,c} \) and \( \Gamma_{t,M} \) share the same positive entries, \( \Gamma_{t,c} \) is obtained from \( \Gamma_{t,M-1} \) by augmenting 0 to its right. Since \( X(t^*+1) > 0 \) by definition of \( t^* \), we have \( 1(X(t^*+1) > \min \Gamma_{t,M}) = 1 \). Moreover, by Proposition 4.2,

\[
1(X(t+1) > \min \Gamma_{t,c}) \geq 1(X(t+1) > \min \Gamma_{t,c-1})
\]  
(252)

for all \( c \geq 1 \) and \( t \geq 0 \). It follows that \( E_M \geq E_{M-1} + 1 \). Hence by (249), we deduce \( \lambda_1^{(1)} \geq M \). This shows \( \lambda_1^{(1)} = M \), as desired. \( \square \)

10. Open Questions and Final Remarks

In this section, we discuss some open problems and future directions.

10.1. Two-sided limiting shape of the Young diagrams. Many of the known result in scaling limits of invariant Young diagrams of randomized BBS ([LLP17, KL18, KLO18a] and the present paper) concern rescaling of the first finite rows or columns. Is it possible to jointly scale the rows and columns and obtain proper two-sided limiting shape of the Young diagram as in the case of the Plancheral measure [KKR88] [IO02]? This question is not entirely obvious since the top rows (soliton numbers) obey the law of large numbers, whereas the top columns (soliton lengths) obey extreme value statistics.

10.2. Obtaining sharper asymptotics. There are some rooms to improve our asymptotic results for the soliton lengths in independence model. First we only know \( \lambda_1(n) \sim (p^* - p_0)n \) in the supercritical regimes (both simple/non-simple), whereas sharp asymptotics for \( \lambda_1(n) \) is known for all regimes for \( \kappa = 1 \) case [LLP17]. In the subcritical regime, one may try to nail down the sharp constant in the asymptotic \( \lambda_1(n) = \Theta(\log n) \). This was done for the \( \kappa = 1 \) case in [LLP17] by solving the Gambler’s ruin problem, and for the multicolor case, this will amount to solve corresponding higher-dimensional Gambler’s ruin, which does not seem to be an obvious question. In this paper we obtained lower bound using longest nonincreasing subsequence and the upper bound by the existence of finite exponential moment of excursion heights. In the critical regime, we bounded \( \lambda_1(n) \) by sums of maximum of reflecting Brownian
motions, where only one appears for the lower bound but \( \kappa \) ones in the upper bound. There are some special case (even in the multicolor case, as stated in the main theorem) where these bounds coincide. But general sharp estimate does not seem immediate. Lastly, for the non-simple supercritical phase, we obtained lowerbound on \( \lambda_j(n) \) for \( j \geq 2 \) of order \( \sqrt{n} / \log n \). The log factor comes from using union bound over \( n \) different location of the random interval \([e, f]\) for linear size longest increasing subsequence. It is also interesting to improving this to \( \sqrt{n} \) lower bound.

10.3. **Column length scaling of higher order invariant Young diagrams.** The \( \kappa \)-color BBS is known to have \( \kappa \)-tuple of invariant Young diagrams, where the ‘higher order’ Young diagrams describe the internal degrees of the freedom of the solitons \[KL18\]. It is our future work to extend the methods and result in the present paper for the first order Young diagram of the \( \kappa \)-color BBS into higher order Young diagrams.

10.4. **Generalization to discrete KdV.** One of the most well-known integrable nonlinear partial differential equation is the Korteweg-de Veris (KdV) equation:

\[
    u_t + 6uu_x + u_{xxx} = 0, \tag{253}
\]

where \( u = u(x, t) \) is a function of two continuous parameters \( x \) and \( t \), and the lower indexes denote derivatives with respect to the specified variables. In 1981, Hirota \[Hir81\] introduced the following discrete KdV (dKdV) equation that arise from KdV by discretizing space and time:

\[
    y_i^{t} + \delta y_{i+1}^{t} = \delta y_{i+1}^{t+1} + y_{i+1}^{t+1}. \tag{254}
\]

If further discretization of the continuous box state in dKdV leads to the ultradiscrete KdV (udKdV) equation, which corresponds to the \( \kappa = 1 \) BBS by Takahashi-Satsuma \[TS90\]:

\[
    U_{n}^{t+1} = \min \left( 1 - U_{n}^{t}, \sum_{k=-\infty}^{n-1} (U_{k}^{t} - U_{k+1}^{t+1}) \right), \tag{255}
\]

where \( U_{k}^{t} \) denotes the number of balls at time \( t \) in box \( k \).

The scaling limit of soliton numbers and lengths of various BBS with random initial configuration has been studied extensively \[LLP17, KL18, KLO18a\], including the present paper. Hence a natural open question is to generalize the similar program to the case of discrete KdV (as opposed to ultradiscrete). For instance, if we initialize dKdV (254) so that the first \( n \) box states are independent \( \text{Exp}(1) \) random variables and evolve the system until solitons come out, what are the scaling limit of the soliton lengths and numbers as \( n \to \infty \)? Can we at least obtain estimates on their expectation? These are much harder question for dKdV because not everything decomposes into solitons: just like in the usual KdV, there so chaotic “radiation” left behind.

**ACKNOWLEDGMENTS**

JBL was supported in part by an ORAU Powe award and a grant from the Simons Foundation (634530).

**REFERENCES**

[BDJ99] Jinho Baik, Percy Deift, and Kurt Johansson, *On the distribution of the length of the longest increasing subsequence of random permutations*, Journal of the American Mathematical Society **12** (1999), no. 4, 1119–1178.

[BFPS07] Alexei Borodin, Patrik I Ferrari, Michael Prähofer, and Tomohiro Sasamoto, *Fluctuation properties of the TASEP with periodic initial configuration*, Journal of Statistical Physics **129** (2007), no. 5-6, 1055–1080.
Alexei Borodin, Patrik I. Ferrari, and Tomohiro Sasamoto, *Transition between Airy1 and Airy2 processes and TASEP fluctuations*, Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences 61 (2008), no. 11, 1603–1629.

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).

David A Croydon and Makiko Sasada, *Duality between box-ball systems of finite box and/or carrier capacity*, arXiv preprint arXiv:1905.00189 (2019).

David A Croydon and Makiko Sasada, *Invariant measures for the box-ball system based on stationary Markov chains and periodic Gibbs measures*, arXiv preprint arXiv:1905.00186 (2019).

Pablo A Ferrari et al., *TASEP hydrodynamics using microscopic characteristics*, Probability Surveys 15 (2018), 1–27.

Pablo A Ferrari and Davide Gabrielli, *BBS invariant measures with independent soliton components*, arXiv preprint arXiv:1812.02437 (2018).

Goro Hatayama, Kazuhiro Hikami, Rei Inoue, Atsuo Kuniba, Taichiro Takagi, and Tetsuji Tokihiro, *The $A_1^{(1)}$ automata related to crystals of symmetric tensors*, Journal of Mathematical Physics 42 (2001), no. 1, 274–308.

Ryogo Hirota, *Discrete analogue of a generalized toda equation*, Journal of the Physical Society of Japan 50 (1981), no. 11, 3785–3791.

Goro Hatayama, Atsuo Kuniba, and Taichiro Takagi, *Factorization of combinatorial $R$ matrices and associated cellular automata*, Journal of Statistical Physics 102 (2001), no. 3-4, 843–863.

Rei Inoue, Atsuo Kuniba, and Taichiro Takagi, *Integrable structure of box–ball systems: crystal, Bethe ansatz, ultradiscretization and tropical geometry*, Journal of Physics A: Mathematical and Theoretical 45 (2012), no. 7, 073001.

Vladimir Ivanov and Grigori Olshanski, *Kerov’s central limit theorem for the plancherel measure on young diagrams*, Symmetric functions 2001: surveys of developments and perspectives, Springer, 2002, pp. 93–151.

Sergei Kerov, Anatol Kirillov, and Nicolai Reshetikhin, *Combinatorics, bethe ansatz, and representations of the symmetric group*, Journal of Mathematical Sciences 41 (1988), no. 2, 916–924.

Atsuo Kuniba and Hanbaek Lyu, *Large deviations and one-sided scaling limit of multicolor box-ball system*, arXiv preprint arXiv:1808.08074 (2018).

Atsuo Kuniba, Hanbaek Lyu, and Masato Okado, *Randomized box–ball systems, limit shape of rigged configurations and thermodynamic Bethe ansatz*, Nuclear Physics B 937 (2018), 240–271.

Atsuo Kuniba, Masato Okado, Reiho Sakamoto, Taichiro Takagi, and Yasuhiko Yamada, *Crystal interpretation of Kerov–Kirillov–Reshetikhin bijection*, Nuclear Physics B 740 (2006), no. 3, 299–327.

Lionel Levine, Hanbaek Lyu, and John Pike, *Double jump phase transition in a random soliton cellular automaton*, arXiv preprint arXiv:1706.05621 (2017).

David A Levin and Yuval Peres, *Markov chains and mixing times*, vol. 107, American Mathematical Soc., 2017.

Thomas Lam, Pavlo Pylyavskyy, and Reiko Sakamoto, *Rigged configurations and cylindric loop schur functions*, arXiv preprint arXiv:1410.4455 (2014).

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

[MT12] Sean P Meyn and Richard L. Tweedie, Markov chains and stochastic stability, Springer Science & Business Media, 2012.

[Rom15] Dan Romik, The surprising mathematics of longest increasing subsequences, vol. 4, Cambridge University Press, 2015.

[Sag01] Bruce E. Sagan, The symmetric group, second ed., Graduate Texts in Mathematics, vol. 203, Springer-Verlag, New York, 2001, Representations, combinatorial algorithms, and symmetric functions.

[Tak93] D Takahashi, On some soliton systems defined by using boxes and balls, 1993 International Symposium on Nonlinear Theory and Its Applications,(Hawaii; 1993), 1993, pp. 555–558.

[TS90] Daisuke Takahashi and Junkichi Satsuma, A soliton cellular automaton, J. Phys. Soc. Japan 59 (1990), no. 10, 3514–3519.

JOEL LEWIS, DEPARTMENT OF MATHEMATICS, THE GEORGE WASHINGTON UNIVERSITY, WASHINGTON, DC 20052.
E-mail address: jblewis@gwu.edu

HANBAEK LYE, DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA, LOS ANGELES, CA 90095.
E-mail address: colourgraph@gmail.com

PAVLO PYLYAVSKYY, DEPARTMENT OF MATHEMATICS, UNIVERSITY OF MINNESOTA, MINNEAPOLIS, MN 55455.
E-mail address: ppylyavs@umn.edu

ARNAB SEN, DEPARTMENT OF MATHEMATICS, UNIVERSITY OF MINNESOTA, MINNEAPOLIS, MN 55455.
E-mail address: arnab@umn.edu