RESEARCH ARTICLE

Scaling limit of soliton lengths in a multicolor box-ball system

Joel Lewis[1], Hanbaek Lyu [2], Pavlo Pylyavskyy [3] and Arnab Sen[4]

[1]Department of Mathematics, The George Washington University, 801 22nd St. NW, Washington, DC, 20052, United States;
E-mail: jblewis@gwu.edu.
[2]Department of Mathematics, University of Wisconsin, 480 Lincoln Dr, Madison, Wisconsin, 53706, United States;
E-mail: hlyu@math.wisc.edu (corresponding author).
[3]Department of Mathematics, University of Minnesota, 127 Vincent Hall 206 Church St. SE Minneapolis, 55455, United States;
E-mail: pylyavskyy@gmail.com.
[4]Department of Mathematics, University of Minnesota, 127 Vincent Hall 206 Church St. SE Minneapolis, 55455, United States;
E-mail: arnab@umn.edu.

Abstract
The box-ball systems are integrable cellular automata whose long-time behavior is characterized by soliton solutions, with rich connections to other integrable systems such as the Korteweg-de Vries equation. In this paper, we consider a multicolor box-ball system with two types of random initial configurations and obtain sharp scaling limits of the soliton lengths as the system size tends to infinity. We obtain a sharp scaling limit of soliton lengths that turns out to be more delicate than that in the single color case established in [LLP20]. A large part of our analysis is devoted to studying the associated carrier process, which is a multidimensional Markov chain on the orthant, whose excursions and running maxima are closely related to soliton lengths. We establish the sharp scaling of its ruin probabilities, Skorokhod decomposition, strong law of large numbers and weak diffusive scaling limit to a semimartingale reflecting Brownian motion with explicit parameters. We also establish and utilize complementary descriptions of the soliton lengths and numbers in terms of modified Greene-Kleitman invariants for the box-ball systems and associated circular exclusion processes.

Contents

1. Introduction

1.1. The κ-color BBS

The box-ball systems (BBS) are integrable cellular automata in 1+1 dimension whose longtime behavior is characterized by soliton solutions. The *κ-color BBS* is a cellular automaton on the half-integer lattice $\mathbb{N}$, which we think of as an array of boxes that can fit at most one ball of any of the κ colors. At each discrete time $t \geq 0$, the system configuration is given by a coloring $\xi^{(t)} : \mathbb{N} \to \mathbb{Z}_{\kappa+1} := \mathbb{Z}/(\kappa+1)\mathbb{Z} = \{0, 1, \cdots, \kappa\}$ with finite support – that is, such that $\xi_x^{(t)} = 0$ for all but finitely many sites x. When $\xi_x^{(t)} = i$, 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 subset $(\mathbb{Z}_{\kappa+1})^{\mathbb{N}}$ of all $(\kappa + 1)$-colorings on $\mathbb{N}$ with finite support defined as follows:

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

2. (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 κ-color BBS is given by

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

A typical 5-color BBS trajectory is shown below.

$$
\begin{array}{ll}
t = 0 : & 3210000513004112520000000000000000000000000000000000 \cdots \\
t = 1 : & 0003210001530001415220000000000000000000000000000000 \cdots \\
t = 2 : & 0000003210105300104105220000000000000000000000000000 \cdots \\
t = 3 : & 0000000003021153010041005220000000000000000000000000 \cdots \\
t = 4 : & 0000000000030002150311041000522000000000000000000000 \cdots \\
t = 5 : & 0000000000003000025100310411000522000000000000000000 \cdots \\
t = 6 : & 0000000000000300002051003100411000522000000000000000 \cdots \\
t = 7 : & 0000000000000030000200510031000411000522000000000000 \cdots
\end{array}
$$

The grounding observation in the κ-color BBS with finitely many balls of positive colors is that the system eventually decomposes into *solitons*, which are sequences of consecutive balls of positive and nonincreasing colors, whose length and content are preserved by the BBS dynamics in all future steps. For instance, all of the nonincreasing consecutive sequences of balls in $\xi^{(6)}$ in the example (specifically, 3, 2, 51, 31, 411, 522) above are solitons, and they are preserved in $\xi^{(7)}$ up to their location and will be so in all future configurations. Note that a soliton of length k travels to the right with speed k. Therefore, the lengths of solitons in a soliton decomposition must be nondecreasing from left to right. In the early dynamics, longer solitons can collide into shorter solitons (e.g., 321 during $t = 0, 1, 2$) and undergo a nonlinear interaction.

The soliton decomposition of the BBS trajectory initialized at $\xi^{(0)}$ can be encoded in a Young diagram $\Lambda = \Lambda(\xi^{(0)})$ having j^{th} column equal in length to the jth-longest soliton. For instance, the Young diagram corresponding to the soliton decomposition of the instance of the 5-color BBS given before is

$$\Lambda(\xi^{(0)}) = \quad \tag{2}$$

Note that the ith row of the Young diagram $\Lambda(\xi^{(0)})$ is precisely the number of solitons of length at least i.

1.2. Overview of main results

We consider the κ-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'. For both models, we denote the kth row and column lengths of the Young diagram encoding the soliton decomposition by $\rho_k(n)$ and $\lambda_k(n)$, respectively,

In the permutation model, the BBS is initialized by a uniformly chosen random permutation Σ^n of colors $\{1, 2, \cdots, n\}$. 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] tells us that the row and column lengths of the random Young diagram constructed from Σ^n via the RS correspondence scale as $\sqrt{n}$. In Theorem 2.1, we show that for the random Young diagram constructed via BBS, the columns scale as $\sqrt{n}$ but the rows scale as n. Namely,

$$\rho_k(n) \sim \frac{n}{k(k+1)}, \qquad \lambda_k(n) \sim \frac{2\sqrt{n}}{\sqrt{k-1}+\sqrt{k}}. \tag{3}$$

While the row lengths in RS-constructed Young diagram are related to the longest increasing subsequences, we show that the row lengths in the BBS-constructed Young diagram are related to the number of *ascents* (Lemma 3.5). This will show that the majority of solitons have a length of order $O(1)$. Hence, the row and column scalings in (3) are consistent.

In the independence model, which we denote $\xi^{n,\mathbf{p}}$, the colors of the sites in the interval $[1, n]$ are independently drawn from a fixed distribution $\mathbf{p} = (p_0, p_1, \cdots, p_\kappa)$ on $\mathbb{Z}_{\kappa+1}$. Recently, Lyu and Kuniba obtained sharp asymptotics for the row lengths as well as their large deviations principle in this independence model [KL20]. In Theorems 2.4–2.7, we establish a sharp scaling limit for the column lengths for the independence model, as summarized in Table 1 and as bullet points below. Let $p^* := \max(p_1, \ldots, p_\kappa)$ denote the density of the maximum positive color and let r denote the multiplicity of p^* (i.e., number of p_i's such that $p_i = p^*$ for $i = 1, \ldots, \kappa$).

○ In the subcritical regime ($p_0 > p^*$), top soliton lengths have sharp scaling $\log_\theta n + (r-1)\log_\theta \log n + O(1)$, where $\theta = p^*/p_0$.
○ In the critical regime ($p_0 = p^*$), $n^{-1/2}\lambda_1(n)$ converges weakly to the maximum L_1-norm of a κ-dimensional semimartingale reflecting Brownian motion (SRBM).

Table 1. *Asymptotic scaling of the ith row length ρ_i and the jth column length λ_j for the independence model with ball density* $\mathbf{p} = (p_0, p_1, \cdots, p_\kappa)$ *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 [LLP20], but the scaling inside the subcritical and supercritical regimes depends on the multiplicity of the maximum positive color p^*. Sharp asymptotics for the row lengths have been obtained in [KL20]. c_i's are constants depending on $\mathbf{p}$ and i; Constnts c, c' do not depend on j; D is a nonnegative and nondegenerate random variable.*

$i \geq 1, j \geq 2$ fixed		$\rho_i(n)$	$\lambda_1(n)$	$\lambda_j(n)$
Subcritical phase $(p^* < p_0)$	Simple ($p^* = p_\ell$ for unique ℓ)	$c_i\, n + \Theta(\sqrt{n})$	$c \log n + \Theta(1)$	
	Non-simple ($p^* = p_\ell$ for multiple ℓ)		$c \log n + c' \log\log n + \Theta(1)$	
Critical phase ($p^* = p_0$)		$c_i\, n + \Theta(\sqrt{n})$	$D\sqrt{n} + o(\sqrt{n})$	$\Theta(\sqrt{n})$
Supercritical phase ($p^* > p_0$)	Simple ($p^* = p_\ell$ for unique ℓ)	$c_i\, n + \Theta(\sqrt{n})$	$c\, n + \Theta(\sqrt{n})$	$\Theta(\log n)$
	Non-simple ($p^* = p_\ell$ for multiple ℓ)			$\Theta(\sqrt{n})$

○ In the supercritical regime ($p_0 < p^*$), $\lambda_1(n) = (p^* - p_0)n + \Theta(\sqrt{n})$. If $r = 1$, then all subsequent top solitons are of order $\log n$; if $r \geq 2$, they are of order $\sqrt{n}$.

○ The fluctuation of $\lambda_1(n)$ depends explicitly on a κ-dimensional SRBM, which arises as the diffusive scaling limit of the associated carrier process.

We establish a similar 'double-jump' phase transition for the $\kappa = 1$ case established by Levine, Lyu and Pike [LLP20]. We find that in the multicolor ($\kappa \geq 2$) case, the maximum positive ball density $p^* = \max(p_1, \cdots, p_\kappa)$ compared to the zero density p_0 dictates general phase transition structure. However, we find that the scaling inside the subcritical and supercritical regimes depends on the multiplicity r of the maximum positive color p^*. Furthermore, the fluctuation of the top soliton length $\lambda_1(n)$ about its mean behavior is described by a κ-dimensional semimartingale reflecting Brownian motion (SRBM) lurking behind, whose covariance matrix depends on $\mathbf{p}$ explicitly. Such SRBM arises as the diffusive scaling limit of the associated carrier process.

A large part of our analysis is devoted to studying the associated carrier process, which is a Markov chain on the κ-dimensional nonnegative integer orthant, whose excursions and running maxima are closely related to soliton lengths (see Lemmas 3.1-3.2). We establish its sharp scaling of ruin probabilities, strong law of large numbers and weak diffusive scaling limit to an SRBM with explicit parameters (Theorems 2.3–2.5). We also establish and utilize alternative descriptions of the soliton lengths and numbers in terms of the modified Greene-Kleitman invariants for the box-ball systems (Lemma 3.5) and associated circular exclusion processes.

1.3. Background and related works

The κ-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 $\{0, 1, \cdots, \kappa\}$, and the time evolution is defined by successive application of the combinatorial R (cf. [FYO00, HHI+01, KOS+06, IKT12]). For a friendly introduction to the combinatorial R, see [KL20, Sec. 3]. The κ-color BBS treated in this paper corresponds to the case where the tableau shape is a single box, which was called the *basic κ-color BBS* in [KL20, Kon20]. The 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 [LLP20, CKST18, KL20, FG18, KL20,

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 [LLP20, FNRW18, FG18, CKST18] as well as for the multicolor case [KL20, KLO18, Kon20]. More recently, invariant measures of the discrete KdV and Toda-type systems have been investigated [CS20].

Three important works are strongly related to this paper. In [LLP20], 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 a 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 [KLO18] and Kuniba and Lyu [KL20]. The central observation is that, when the initial configuration is given by a product measure, 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 [KLO18], 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 [KL20], for the (basic) κ-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 κ-color BBS. We develop two main tools for our analysis, which are a modified version of Greene-Kleitman invariants for BBS (Section 3.3) and the carrier process (see Definition 2.2). For the independence model, we obtain the scaling limit of the carrier process as an SRBM [Wil95], and it plays a central role in our analysis. For the permutation model, the carrier process 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, for example, [F+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 in the clockwise direction. It turns out that this process is crucial in analyzing the permutation model (Section 4.2), 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 (Section 3.1). Interestingly, a cylindric version of Schur functions has been used to study rigged configurations and BBS [LPS14].

1.4. Organization

In Section 2, we define the carrier process, state the permutation and the independence model for the κ-color BBS, and state our main results. We also provide numerical simulation to validate our results empirically. In Section 3, we introduce infinite and finite capacity carrier processes for the κ-color BBS

and state the three key combinatorial lemmas (Lemmas 3.1, 3.3, 3.5). In Section 4, we prove our main result for the permutation model (Theorem 2.1) by using the modified GK invariants for BBS (Lemma 3.5) and analyzing the associated circular exclusion process. In Section 5, we prove Theorem 2.3 **(i)** about the stationary behavior of the subcritical carrier process. Next, in Section 6, we introduce the 'decoupled carrier process' and develop the 'Skorokhod decomposition' of the carrier process. These will play critical roles in the analysis in the following sections. In Section 7, we analyze the decoupled carrier process over the i.i.d. ball configuration. In Section 8, we prove Theorem 2.3 **(ii)** and Theorem 2.4. In Sections 9 and 10, we establish a linear and diffusive scaling limit of the carrier process, which is stated in Theorem 2.5. Background on SRBM and an invariance principle for SRBM are also provided in Section 10. In Section 11, we prove Theorems 2.6 and 2.7. Lastly, in Section 12, we provide postponed proofs for the combinatorial lemmas stated in Section 3.

1.5. *Notation*

We use the convention that summation and product over the empty index set equal zero and one, respectively. For any probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and any event $A \in \mathcal{F}$, we let $\mathbf{1}(A)$ denote the indicator variable of A. Let $C^d(0, \infty)$ denote the space of continuous functions $f : [0, \infty) \to \mathbb{R}^d$ endowed with the topology of uniform convergence on compact intervals. We let $\text{tridiagonal}_d(a, b, c)$ denote the $d \times d$ matrix which has a on its subdiagonal, b on its diagonal and c on its superdiagonal entries, and zeros elsewhere.

We adopt the notations $\mathbb{R}^+ = [0, \infty)$, $\mathbb{N} = \{1, 2, 3, \ldots\}$ and $\mathbb{Z}_{\geq 0} = \mathbb{N} \cup \{0\}$ throughout. For a sequence of events $(A_n)_{n \geq 1}$, we say A_n occurs *with high probability* if $\mathbb{P}(A_n) \to 1$ as $n \to \infty$. We employ the Landau notations $O(\cdot)$, $\Omega(\cdot)$, $\Theta(\cdot)$ 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 $W_n = O(a_n)$ with high probability if for each $\varepsilon > 0$, there is a constant $C \in (0, \infty)$ such that $\mathbb{P}(W_n < Ca_n) \geq 1 - \varepsilon$ for all sufficiently large n. We say that $W_n = \Omega(a_n)$ if for each $\varepsilon > 0$, there is a $c \in (0, \infty)$ such that $\mathbb{P}(W_n > ca_n) \geq 1 - \varepsilon$ for all sufficiently large n, and we say $W_n = \Theta(a_n)$ with high probability if $W_n = O(a_n)$ and $W_n = \Omega(a_n)$ both with high probability. In all of these Landau notations, we require that the constants c, C do not depend on n.

2. Statement of results

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

2.1. *The permutation model*

For the *permutation model*, let $(U_x)_{x \geq 1}$ be a sequence of i.i.d. 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 Σ^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

$$\xi_x^n := \Sigma^n(x) \cdot \mathbf{1}(1 \leq x \leq n). \tag{4}$$

We now state our main result 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 2.1 (The permutation model). *Let ξ^n be the permutation model as above. For each $k \geq 1$, denote $\rho_k(n) = \rho_k(\xi^n)$ and $\lambda_k(n) = \lambda_k(\xi^n)$. Then for each fixed $k \geq 1$, almost surely,*

$$\lim_{n\to\infty} n^{-1}\rho_k(n) = \frac{1}{k(k+1)}, \qquad \lim_{n\to\infty} n^{-1/2}\lambda_k(n) = \frac{2}{\sqrt{k-1}+\sqrt{k}}. \tag{5}$$

Our proof of Theorem 2.1 proceeds as follows. We first establish a combinatorial lemma (Lemma 3.5) that associates the soliton lengths and numbers with a modified version of Greene-Kleitman invariants for BBS. We then utilize the tail bounds on longest increasing subsequences in uniformly random permutations in Baik, Deift and Johansson [BDJ99] for establishing the scaling limit for the lengths of the columns. For the row lengths, we use the characterization of soliton numbers as an additive functional of finite-capacity carrier processes [KL20]. Such a process becomes an exclusion process on the unit circle for the permutation model.

2.2. The independence model

To define the *independence model*, fix integers $n, \kappa \geq 1$. Let $\mathbf{p} = (p_0, p_1, \cdots, p_\kappa)$ be a probability distribution on $\{0, 1, \cdots, \kappa\}$. Let $\xi = \xi^{\mathbf{p}}$ be the sequence $(\xi_x)_{x\in\mathbb{N}}$ of i.i.d. random variables ξ_x where

$$\mathbb{P}(\xi_x = i) = p_i \qquad \text{for } i = 0, 1, \ldots, \kappa. \tag{6}$$

For each integer $n \geq 1$, define κ-color BBS configuration $\xi^{n,\mathbf{p}}$ of size n by

$$\xi_x^{n,\mathbf{p}} = \xi_x^{\mathbf{p}} \cdot \mathbf{1}(1 \leq x \leq n). \tag{7}$$

We may further assume, without loss of generality, that $p_i > 0$ for all $1 \leq i \leq \kappa$. 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 $\{i, \cdots, \kappa - 1\}$.

Through various combinatorial lemmas (see Section 3), we will establish that the soliton lengths $\lambda_j(n)$ of for the i.i.d. model are closely related to the extreme behavior of a Markov chain $(W_x)_{x\in\mathbb{N}}$ defined on the nonnegative integer orthant $\mathbb{Z}_{\geq 0}^\kappa$, which we call the '$\kappa$-color carrier process'. Denote $\mathbf{e}_i \in \mathbb{Z}^\kappa$ whose coordinates are all zero except the ith coordinate being 1.

Definition 2.2 (κ-color carrier process). Let $\xi := (\xi_x)_{x\in\mathbb{N}}$ be κ-color ball configuration. The *(κ-color) carrier process* over ξ is a process $(W_x)_{x\in\mathbb{N}}$ on the state space $\Omega := \mathbb{Z}_{\geq 0}^\kappa$ defined by the following evolution rule: Denoting $i := \xi_{x+1}$ if $\xi_{x+1} \in \{1, \ldots, \kappa\}$ and $i := \kappa + 1$ if $\xi_{x+1} = 0$,

$$W_{x+1} - W_x = \begin{cases} \mathbf{e}_i - \mathbf{1}(i_* \neq 0)\,\mathbf{e}_{i_*} & \text{if } 1 \leq i \leq \kappa \\ -\mathbf{1}(i_* \neq 0)\,\mathbf{e}_{i_*} & \text{if } i = \kappa + 1, \end{cases} \tag{8}$$

where $i_* := \sup\{1 \leq j < i : W_x(j) \geq 1\}$ with the convention $\sup \emptyset = 0$. Unless otherwise mentioned, we take $W_0 = \mathbf{0}$ and $\xi = \xi^{\mathbf{p}}$ with density $\mathbf{p} = (p_0, \ldots, p_\kappa)$.

In words, at location x, the carrier holds $W_x(i)$ balls of color i for $i = 1, \ldots, \kappa$. When a new ball of color $1 \leq \xi_{x+1} \leq \kappa$ is inserted into the carrier W_x, then a ball of the largest available color that is smaller than ξ_x is excluded from W_x; if there is no such ball in W_x, then no ball is excluded. If $\xi_{x+1} = 0$, then no new ball is inserted, and a ball of the largest available color that is smaller than ξ_x is excluded from W_x. The resulting state of the carrier is W_{x+1}. We call the transition rule (8) as the 'circular exclusion' (since a ball in the carrier's possession is excluded from the carrier upon the insertion of a new ball according to the circular ordering). One can also view the carrier process as a multitype queuing system, where W_x denotes the state of the queue and $W_x(i)$ is the number of jobs of 'cyclic hierarchy' i to be processed.

A large portion of this paper will be devoted to analyzing scaling limits of the carrier process W_x over the i.i.d. configuration $\xi^{\mathbf{p}}$. In this case, W_x is a Markov chain on the state space of the nonnegative integer orthant Ω. See Figure 1 for an illustration.
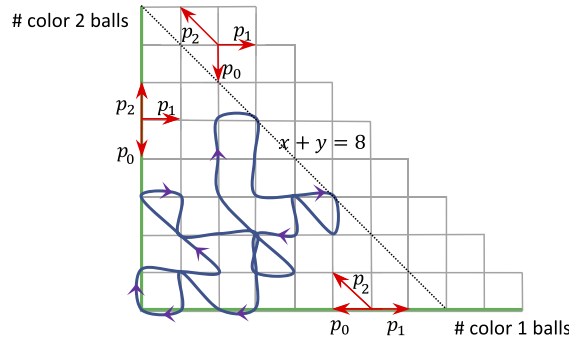
Figure 1. *State space diagram for the carrier process W_x for $\kappa = 2$. Red arrows illustrate the transition kernel at the 'interior' (gray) and 'boundary' (green) points in the state space. A single excursion (starting and ending at the origin) of 'height' 8 is shown in a blue path with arrows.*

Theorem 2.3 states the behavior of the carrier process in the subcritical regime $p_0 > \max(p_1, \cdots, p_\kappa)$. Define a function $\pi : \Omega \to \mathbb{R}$ by

$$\pi(n_1, n_2, \cdots, n_\kappa) = \prod_{i=1}^{\kappa} \left(1 - \frac{p_i}{p_0}\right)\left(\frac{p_i}{p_0}\right)^{n_i}. \tag{9}$$

This is a valid probability distribution on Ω when $p_0 > \max(p_1, \cdots, p_\kappa)$ since

$$\sum_{n_1=0}^{\infty} \cdots \sum_{n_\kappa=0}^{\infty} \prod_{i=1}^{\kappa} \left(\frac{p_i}{p_0}\right)^{n} = \prod_{i=1}^{\kappa}\left(1 - \frac{p_i}{p_0}\right)^{-1} \in (0, \infty). \tag{10}$$

Note that π is the the product of geometric distributions of means $p_i/(p_0 - p_i) > 0$ for $i = 1, \ldots, \kappa$.

Theorem 2.3 (The carrier process at the subcritical regime). *Let $p^* := \max(p_1, \cdots, p_\kappa)$ and suppose $p_0 > p^*$. Let r denote the multiplicity of p^* (i.e., number of i's in $\{1, \ldots, \kappa\}$ s.t. $p_i = p^*$).*

(i) *(Convergence) The carrier process W_x is an irreducible, aperiodic and positive recurrent Markov chain on $\mathbb{Z}_{\geq 0}^{\kappa}$ with π in (9) as its unique stationary distribution. Thus, writing d_{TV} for the total variation distance and denoting the distribution of W_x by π_x, then*

$$\lim_{x \to \infty} d_{TV}(\pi_x, \pi) = 0. \tag{11}$$

(ii) *(Multidimensional Gambler's ruin) Let T_1 denote the first return time of W_x to the origin and let $h_1 := \max_{0 \leq x \leq T_1} \|W_x\|_1$. Then for all $N \geq 1$, there exists a constant $\delta > 0$ such that*

$$\delta \binom{N + r - 1}{r - 1}\left(\frac{p^*}{p_0}\right)^N \leq \mathbb{P}(h_1 \geq N) \leq C \binom{N + r - 1}{r - 1}\left(\frac{p^*}{p_0}\right)^N, \tag{12}$$

where $C = 1$ if $r = \kappa$ and $C = \left(\frac{p^*}{p^* - p^{(2)}}\right)^{\kappa - r}$ if $r < \kappa$ with $p^{(2)}$ being the second largest value among $p_1, \ldots, p_\kappa$.

By using Theorem 2.3, we establish sharp scaling limit of soliton lengths for the independence model in the subcritical regime, which is stated in Theorem 2.4 below. (See Section 1.5 for a precise definition of Landau notations.)

Theorem 2.4 (The independence model – Subcritical regime). *Fix $\kappa \geq 1$ and let $\xi^{n,\mathbf{p}}$ be as the i.i.d. model above. Denote $\lambda_j(n) = \lambda_j(\xi^{n,\mathbf{p}})$, $p^* := \max_{1 \leq i \leq \kappa} p_i$, and $r := |\{1 \leq i \leq \kappa : p_i = p^*\}|$. Suppose*

$p_0 > p^*$ *and denote* $\theta := p^*/p_0$. *Then for each fixed* $j \geq 1$,

$$\lambda_j(n) = \log_\theta n + (r-1) \log_\theta \log n + \Theta(1). \tag{13}$$

Furthermore, denote $\nu_n := (1 + \delta_n) \log_\theta(\sigma n/(r-1)!)$, *where* $\sigma := \prod_{i=1}^\kappa \left(1 - \frac{p_i}{p_0}\right)$ *and* $\delta_n :=$ $\frac{(r-1) \log \log_\theta(\sigma n/(r-1)!) + \log(r-1)!}{\log \sigma n/(r-1)!}$. *Then for all* $x \in \mathbb{R}$,

$$\exp(-\delta\theta^{-x}) \leq \liminf_{n\to\infty} \mathbb{P}(\lambda_j(n) \leq x + \nu_n) \tag{14}$$

$$\leq \limsup_{n\to\infty} \mathbb{P}(\lambda_j(n) \leq x + \nu_n) \leq \exp\left(-\frac{C}{(r-1)!}\theta^{-(x-1)}\right) \sum_{k=0}^{j-1} \frac{\theta^{-k(x-1)}}{k!(r-1)!}, \tag{15}$$

where $\delta > 0$, $C \geq 1$ *are constants in Theorem* 2.3.

Next, we turn our attention to the critical and the supercritical regime, where $p_0 \leq \max(p_1, \cdots, p_\kappa)$. In this regime, the carrier process does not have a stationary distribution, and we are interested in identifying the limit of the carrier process in the linear and diffusive scales. A natural candidate for the diffusive scaling limit (if it exists) would be the semimartingale reflecting Brownian motion (SRBM) [Wil95], whose definition we recall in Section 10. Roughly speaking, an SRBM on a domain $S \subseteq \mathbb{R}^\kappa$ is a stochastic process $\mathcal{W}$ that admits a Skorokhod-type decomposition

$$\mathcal{W} = X + RY, \tag{16}$$

where X is a κ-dimensional Brownian motion with drift θ, covariance matrix Σ and initial distribution ν. The 'interior process' X gives the behavior of $\mathcal{W}$ in the interior of S. When it is at the boundary of S, it is pushed instantaneously toward the interior of S along the direction specified by the 'reflection matrix' R and an associated 'pushing process' Y. We say such $\mathcal{W}$ is a SRBM associated with $(S, \theta, \Sigma, R, \nu)$. If $R = I - Q$ for some nonnegative matrix Q with spectral radius less than one, then such $\mathcal{W}$ is unique (pathwise) for possibly degenerate Σ when $S = \mathbb{R}^\kappa_{\geq 0}$ [HR81]. If Σ is nondegenerate and S is a polyhedron, a necessary and sufficient condition for the existence and uniqueness of such SRBM is that R is 'completely-$\mathcal{S}$' (see Definition 10.2) [Wil95, KW07].

A crucial observation for analyzing the carrier process in the critical and supercritical regimes is the following. Of all the κ coordinates of W_x, some have a negative drift and some others do not. We call an integer $1 \leq i \leq \kappa$ an *unstable color* if $p_i \geq \max(p_{i+1}, \cdots, p_\kappa, p_0)$ and a *stable color* otherwise. Since balls of color i can only be excluded by balls of colors in $\{i+1, \ldots, \kappa, 0\}$, then the coordinate $W_x(i)$ is likely to diminish if the color i is stable but not if i is unstable. Denote the set of all unstable colors by $\mathcal{C}^\mathbf{p}_u = \{\alpha_1, \cdots, \alpha_r\}$ with $\alpha_1 < \cdots < \alpha_r$ and let $\mathcal{C}^\mathbf{p}_s := \{0, 1, \cdots, \kappa\} \setminus \mathcal{C}^\mathbf{p}_u$ denote the set of stable colors. (See Figure 8 for illustration.) By definition, we have

$$p_{\alpha_1} \geq p_{\alpha_2} \geq \cdots \geq p_{\alpha_r} \geq p_{\alpha_{r+1}} := p_0. \tag{17}$$

Now, we will construct a new process X_x, which we call the 'decoupled carrier process' (see Section 6.1), that mimics the behavior of W_x, but the values of X_x on the unstable colors are unconstrained and thus can be negative. Since W_x is confined in the nonnegative orthant $\mathbb{Z}^\kappa_{\geq 0}$ but X_x is not, we need to add some correction process to X_x that 'pushes' it toward the orthant $\mathbb{Z}^\kappa_{\geq 0}$ whenever X_x has some of its coordinates going to negative. More precisely, in Lemma 6.3, we identify a 'reflection matrix' $R \in \mathbb{R}^{\kappa \times \kappa}$ and a 'pushing process' Y_x on $\mathbb{Z}^\kappa$ such that

$$W_x = X_x + RY_x \quad \text{for } x \geq 0, \tag{18}$$
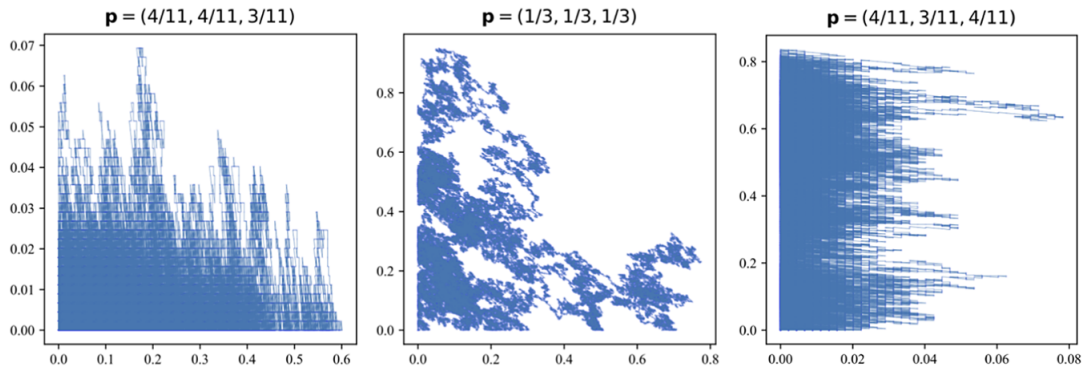
Figure 2. *Simulation of the carrier process W_x in diffusive scaling for $\kappa = 2$, $n = 2 \times 10^5$, at three critical ball densities (left) $\mathbf{p} = (4/11, 4/11, 3/11)$, (middle) $\mathbf{p} = (1/3, 1/3, 1/3)$ and (right) $\mathbf{p} = (4/11, 3/11, 4/11)$. In all cases, the process converges weakly to a semimartingale reflecting Brownian motion on $\mathbb{R}^2_{\geq 0}$ whose covariance matrix is nondegenerate in the middle and degenerate in the other two cases.*

where $Y_0 = \mathbf{0}$, and for each $i \in \{1, \ldots, \kappa\}$, the ith coordinate of Y_x is nondecreasing in x and can only increase when $W_x(i) = 0$. We call the above as a *Skorokhod decomposition of the carrier process* (Our definition is motivated by the Skorokhod problem; see Definition 10.3.) This and the classical invariance principle for SRBM [RW88] are the keys to establishing the following result on the scaling limit of the carrier process.

Theorem 2.5 (Linear and diffusive scaling limit of the carrier process). *Suppose $p_0 \leq \max(p_1, \cdots, p_\kappa)$. Let $\alpha_1 < \cdots < \alpha_r$ as before and define*

$$\boldsymbol{\mu} = (\mu_1, \ldots, \mu_\kappa) := \sum_{j=1}^{r} \mathbf{e}_{\alpha_j}(p_{\alpha_j} - p_{\alpha_{j+1}}), \tag{19}$$

where we let $p_{\alpha_{r+1}} = p_0$.

(i) *(Linear scaling) Almost surely,*

$$\lim_{x \to \infty} x^{-1} W_x = \lim_{x \to \infty} x^{-1}\left(\max_{0 \leq t \leq x} W_t(i) \, ; \, i = 1, \ldots, \kappa\right) = \boldsymbol{\mu}. \tag{20}$$

(ii) *(Diffusive scaling) Let $(\overline{W}_t)_{t \in \mathbb{R}_{\geq 0}}$ denote the linear interpolation of $(W_x - x\boldsymbol{\mu})_{x \in \mathbb{N}}$. Then as $n \to \infty$,*

$$(x^{-1/2}\overline{W}_{xt} \, ; \, 0 \leq t \leq 1) \Longrightarrow \mathcal{W} \text{ in } C([0,1]), \tag{21}$$

where $\mathcal{W}$ is an SRBM associated with data $(S, \mathbf{0}, \Sigma, R, \delta_0)$ (see Definition 10.1) with $S := \{(x_1, \ldots, x_\kappa) \in \mathbb{R}^\kappa : x_i \geq 0 \text{ if } \mu_i = 0\}$, Σ the limiting covariance matrix (possibly degenerate) in (177), $R := \mathrm{tridiag}_\kappa(0, 1, -1)$, and δ_0 the point mass at $\mathbf{0}$.

In Figures 2 and 3, we provide simulations of the carrier process $W_x = (W_x(1), W_x(2))$ for $\kappa = 2$ in various regimes, numerically verifying Theorem 2.5. In Figure 2, we show the carrier process in diffusive scaling ($n^{-1/2}$) at three different critical ball densities $\mathbf{p}$. The carrier process in diffusive scaling converges weakly to an SRBM in $\mathbb{R}^2_{\geq 0}$, whose covariance matrix depends on $\mathbf{p}$ and can be degenerate. For instance, at $\mathbf{p} = (4/11, 4/11, 3/11)$, $W_x(2)$ is subcritical (since $p_2 = 3/11 < 4/11 = p_0$), and $W_x(1)$ is critical, so the SRBM degenerates in the second axes.

In Figure 3, we show the carrier process in diffusive scaling at three different supercritical ball densities $\mathbf{p}$. The carrier process has a nonzero drift $\boldsymbol{\mu} = (\mu_1, \mu_2) \in \mathbb{R}^2_{\geq 0}$. If $\mu_1, \mu_2 > 0$, then the centered
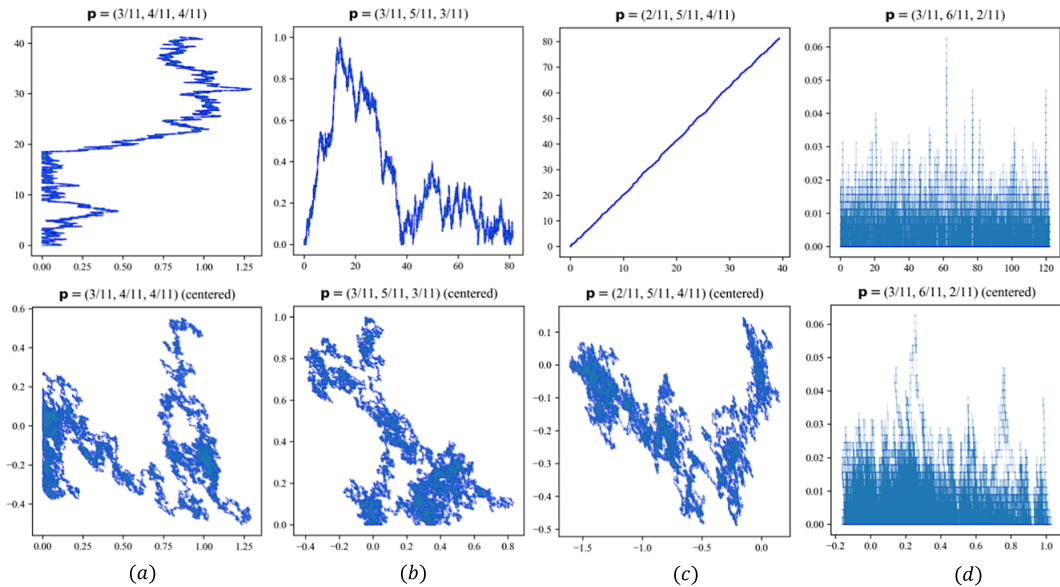
Figure 3. *Simulation of the carrier process W_x in diffusive scaling for $\kappa = 2$, $n = 2 \times 10^5$, at four supercritical ball densities (a) $\mathbf{p} = (3/11, 4/11, 4/11)$, (b) $\mathbf{p} = (3/11, 5/11, 3/11)$, (c) $\mathbf{p} = (2/11, 5/11, 4/11)$ and (d) $\mathbf{p} = (3/11, 6/11, 2/11)$. The processes grow linearly at least in one dimension (the top row shows uncentered processes in diffusive scaling). As shown in the second row, after centering by the mean drift $\boldsymbol{\mu}$, the processes converge weakly to semimartingale reflecting Brownian motion on domains (a) $\mathbb{R}_{\geq 0} \times \mathbb{R}$, (b) $\mathbb{R} \times \mathbb{R}_{\geq 0}$, (c) $\mathbb{R}^2$ (no reflection) and (d) $\mathbb{R} \times \mathbb{R}_{\geq 0}$ (with a degenerate covariance matrix).*

carrier process $W_x - x\boldsymbol{\mu}$ converges weakly to a 2-dimensional Brownian motion in diffusive scaling. If either μ_1 or μ_2 equals zero, then the diffusive scaling limit is an SRBM on $\mathbb{R}_{\geq 0} \times \mathbb{R}$ or $\mathbb{R} \times \mathbb{R}_{\geq 0}$, which is the domain S in the statement of Theorem 2.5 **(ii)**. For instance, for $\mathbf{p} = (3/11, 6/11, 2/11)$ as in Figure 3 (d), the SRBM is on domain $S = \mathbb{R} \times \mathbb{R}_{\geq 0}$ and has a degenerate covariance matrix, since $W_x(2)$ is subcritical and vanishes in the diffusive scale.

Using the linear and the diffusive scaling limit of the carrier process in Theorem 2.5, we obtain a sharp scaling limit of soliton lengths for the independence model in the critical and subcritical regimes. These results are stated in Theorems 2.6 and 2.7 below.

Theorem 2.6 (The independence model – Critical regime). *Suppose $p^* = p_0$. Then for each fixed $j \geq 1$, $\lambda_j(n) = \Theta(\sqrt{n})$. Furthermore, let Σ be a $\kappa \times \kappa$ covariance matrix defined explicitly in (177) and $R = \text{tridiag}_{\kappa \times \kappa}(0, 1, -1)$. Let $\mathcal{W}$ be a semimartingale reflecting Brownian motion associated with data $(\mathbb{R}^\kappa_{\geq 0}, \mathbf{0}, \Sigma, R, \delta_{\mathbf{0}})$ (see Definition 10.1). Then as $n \to \infty$,*

$$n^{-1/2}\lambda_1(n) \Longrightarrow \sup \|W\|_1, \tag{22}$$

where $\Longrightarrow$ denotes weak convergence.

Theorem 2.7 (The independence model – Supercritical regime). *Suppose $p^* > p_0$.*

(i) *(Top soliton length in the supercritical regime) It holds that*

$$\lim_{n \to \infty} n^{-1}\lambda_1(n) \overset{a.s.}{=} p^* - p_0 \quad \text{and} \quad \lambda_1(n) = (p^* - p_0)n + \Theta(\sqrt{n}). \tag{23}$$

More precisely, let $\alpha_1 < \cdots < \alpha_r$ denote the unstable colors and let $\alpha_{r+1} := 0$. Let $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_\kappa)$ be as in (19) and $J := \{i \ : \ \mu_i > 0\}$. Let $\mathcal{W} = (\mathcal{W}^1, \ldots, \mathcal{W}^\kappa)$ denote the SRBM in Theorem 2.5 (ii). Then

$$\sum_{i=1}^{\kappa} \mathcal{W}^i(1) \preceq \liminf_{n \to \infty} \frac{\lambda_1(n) - (p_* - p_0)n}{\sqrt{n}} \tag{24}$$

$$\preceq \limsup_{n \to \infty} \frac{\lambda_1(n) - (p_* - p_0)n}{\sqrt{n}} \tag{25}$$

$$\preceq \sum_{j \in J} B^j(1) + \sup_{0 \leq v \leq 1} \sum_{j \in \{1, \ldots, \kappa\} \setminus J}^{\kappa} \mathcal{W}^i(v), \tag{26}$$

where $\preceq$ denotes stochastic dominance and $B = (B^1, \ldots, B^\kappa)$ is a Brownian motion in $\mathbb{R}^\kappa$ with zero drift and the same covariance matrix with $\mathcal{W}$.

(ii) *(Subsequent soliton lengths in the simple supercritical regime) Suppose $r = 1$. Then for any fixed $j \geq 2$, $\lambda_j(n) = \Theta(\log n)$ with high probability.*

(iii) *(Subsequent soliton lengths in the nonsimple supercritical regime) Suppose $r \geq 2$. Then for any fixed $j \geq 2$, $\lambda_j(n) = \Theta(\sqrt{n})$ with high probability; that is, for each $\varepsilon > 0$, there exists constants $c_1, c_2 > 0$ such that $\liminf_{n \to \infty} \mathbb{P}(\lambda_j(n)/\sqrt{n} \in [c_1, c_2]) \geq 1 - \varepsilon$.*

Multiple remarks on Theorems 2.4–2.7 are in order. These results extend the 'double-jump' phase transition on soliton lengths for the $\kappa = 1$ case established by Levine, Lyu and Pike [LLP20] to the multicolor case. As in the $\kappa = 1$ case, we find that there exist three regimes – subcritical ($\lambda_1(n) = \Theta(\log n)$), critical ($\lambda_1(n) = \Theta(\sqrt{n})$) and supercritical ($\lambda_1(n) = \Theta(n)$) – depending on whether the maximum ball density $p^* = \max(p_1, \ldots, p_\kappa)$ exceeds the empty box density p_0. However, we find that the scaling behavior of the soliton lengths inside each regime is significantly more nuanced in the multicolor case than in the single-color case.

In the subcritical regime $p^* < p_0$, we find all top soliton lengths $\lambda_j(n)$ for $j \geq 1$ are concentrated around $\log_\theta n + (r - 1) \log_\theta \log n$, where $\theta = p^*/p_0$ and r denotes the multiplicity of the maximum positive color p^*, and the tail of $\lambda_n(n)$ has a Gumbel-type tail distribution. While this scaling coincides with that in the $\kappa = 1$ case for $r = 1$, if $r \geq 2$, then the top solitons are an asymptotically 'a tad' longer by $(r - 1) \log_\theta \log n$, which is caused by the competition between multiple maximal colors.

In the critical regime $p^* = p_0$, we find that $\lambda_1(n)/\sqrt{n} \Rightarrow D$, where the distribution of the nondegenerate random variable D depends on a SRBM on the orthant $\mathbb{R}_{\geq 0}^\kappa$ with zero drift and an explicit covariance matrix Σ. This is the same SRBM to which the entire carrier process converges weakly in diffusive scaling as in Theorem 2.5. For instance, if p^* is uniquely achieved, then the SRBM $\mathcal{W}$ is degenerate in all but one dimension. In particular, for $\kappa = 1$, our result recovers the corresponding result in [LLP20]. In general, Σ can depend on the entire $\mathbf{p}$, capturing the intertwined interaction between balls of all colors in the multicolor case.

In the supercritical regime $p^* > p_0$, Theorem 2.7 shows that $\lambda_1(n)/n \to p^* - p_0$ almost surely and the fluctuation of $\lambda_1(n)$ about its mean is of order $\sqrt{n}$. While a central limit theorem (CLT) for $\lambda_1(n)$ in the supercritical regime was shown in [LLP20] for the $\kappa = 1$ case, we find in the multicolor case that the distribution of the fluctuation of $\lambda_1(n)$ does not always satisfy CLT. More precisely, the following corollary shows that CLT holds for $\lambda_1(n)$ if and only if the ball density is strictly decreasing on the unstable colors. (Recall (17).)

Corollary 2.8. *(Fluctuation of λ_1 in the supercritical regime) Keep the same setting as in Theorem 2.4. Suppose supercritical regime $p^* > p_0$. Let $\alpha_1 < \cdots < \alpha_r$ denote the unstable colors.*

(i) *Further assume $p_{\alpha_1} > \cdots > p_{\alpha_r}$, Then $\lambda_1(n)$ satisfies the following central limit theorem:*

$$\frac{\lambda_1(n) - (p_* - p_0)n}{\sqrt{n}} \Longrightarrow N(0, \|\Sigma\|_1), \tag{27}$$

where the limiting distribution is the normal distribution with mean zero and variance $\|\Sigma\|_1$ for Σ the covariance matrix in Theorem 2.5.

(ii) *If $p_{\alpha_j} = p_{\alpha_{j+1}}$ for some $1 \leq j \leq r - 1$, then*

$$\mathbb{E}\left[\liminf_{n \to \infty} \frac{\lambda_1(n) - (p_* - p_0)n}{\sqrt{n}} \right] > 0. \tag{28}$$

In particular, $\lambda_1(n)$ does not satisfy the central limit theorem.

Indeed, suppose $p_{\alpha_1} > \cdots > p_{\alpha_r}$ as in Corollary 2.8 **(i)**. Then Theorem 2.5 states that $x^{-1/2}(W_x - \mu x)$ converges weakly to the (nonreflecting) Brownian motion in $\mathbb{R}^\kappa$ with covariance matrix Σ. Hence, in this case, Theorem 2.7 **(i)** immediately implies that

$$\frac{\lambda_1(n) - (p_* - p_0)n}{\sqrt{n}} \implies \sum_{i=1}^{\kappa} B^i(1), \tag{29}$$

where $B = (B^1, \ldots, B^\kappa)$ is a Brownian motion in $\mathbb{R}^\kappa$ with zero drift and covariance matrix Σ in Theorem 2.5. Since $B(1)$ is a standard normal vector with mean zero and covariance matrix Σ, the result in Corollary 2.8 **(i)** follows.

If we are in the situation as in Corollary 2.8 **(ii)**, then some of the consecutive unstable colors have the same ball density (i.e., $p_{\alpha_j} = p_{\alpha_{j+1}}$). For every such α_j, the corresponding coordinate has to remain nonnegative in the limiting SRBM. So in this case, the fluctuation of λ_1 about its mean in the diffusive scaling has a positive expectation. As an example, consider the case $\mathbf{p} = (p_0, p_1, p_2)$ with $p_1 > p_2 = p_0$ (see Figure 3 (b)). In this case, the limiting SRBM $W = (W^1, W^2)$ is on the domain $\mathbb{R} \times \mathbb{R}_{\geq 0}$, so the lower bound $W^1(1) + W^2(1)$ on the fluctuation in (24) has a positive expectation. This can be understood for the following reasons. Since $p_1 > \max(p_0, p_2)$, the number of color 1 balls in the carrier grows linearly and makes the dominant contribution (of order n) to $\lambda_1(n)$. However, the number of color 2 balls in the carrier still contributes to $\lambda_1(n)$ by order $\sqrt{n}$ since $p_2 = p_0$. While the fluctuation of the number of color 1 balls around its mean $(p_1 - p_0)n$ has mean zero, the contribution of color 2 balls of order $\sqrt{n}$ is only visible in the diffusive scaling, and it is almost always of a positive amount.

Another interesting behavior of the multicolor BBS is the order of subsequent soliton lengths, $\lambda_j(n)$ for $j \geq 2$, in the supercritical regime, which depends drastically on the multiplicity r of the maximal ball density p^*. That is, $\lambda_j(n)$ for all $j \geq 2$ is of order $\log n$ if $r = 1$, but they are of order $\sqrt{n}$ if $r \geq 2$. The former case agrees with the results for the $\kappa = 1$ case in [LLP20]. There, it was shown that $\lambda_2(n)$ comes from the subexcursions of the carrier process below its running maximum. The height of such subexcursions has exponential tails, so we have order $\log(n)$ as the order of the maximum of n subexponential random variables. However, if $r \geq 2$ in the multicolor case, the discrepancy between the number of balls of two maximal colors is of order $\sqrt{n}$ and contributes to $\lambda_2(n)$ (see the proof of Theorem 2.7 **(iii)**). We remark that a duality between the subcritical and the supercritical regimes for $\kappa = 1$ was established in [LLP20], in the sense that λ_{j+1} in the supercritical regime corresponds to λ_j in the subcritical regime for $j \geq 1$. Our results confirm a similar correspondence still holds asymptotically for the simple ($r = 1$) supercritical regime; but λ_{j+1} in the nonsimple ($r \geq 2$) supercritical regime corresponds to λ_j in the critical regime.

3. Key combinatorial lemmas

3.1. Infinite capacity carrier process and soliton lengths

The definition of κ-color BBS dynamics we gave in the introduction involves the nonlocal movement of balls. It can instead be defined using a 'carrier', which gives a localized characterization of the process and reveals a number of important invariants that fully determine the resulting solitons. For the simplest case $\kappa = 1$, imagine a carrier of infinite capacity sweeps through the time-t configuration $\xi^{(t)}$ from the

Γ_x

0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	**2**	**2**	**2**	**1**	**2**	**2**	**5**	**5**	**5**	0	0	0	0	0	0	0	0	0
0	0	0	0	0	**3**	**3**	**6**	**6**	**2**	**2**	**5**	**5**	**5**	**6**	**5**	0	0	**2**	**3**	0	0	0	0
0	0	**2**	**5**	**7**	**7**	**7**	**7**	**7**	**6**	**6**	**6**	**6**	**7**	**7**	**6**	**5**	**6**	**6**	**6**	**3**	0	0	0

ξ_x

0	2	5	7	3	2	6	1	0	2	5	5	7	6	0	0	6	2	3	0	0	0	0

ξ'_x

0	0	2	5	0	0	3	0	7	1	2	2	6	5	7	6	5	0	2	6	3	0	0	0

Figure 4. *Time evolution of the infinite capacity carrier process $(\Gamma_x)_{x \geq 0}$ over the 7-color initial configuration ξ, producing new configuration ξ' consisting of existing ball colors. For instance, $\xi_2 = 2$, $\Gamma_2 = [2, 0, 0, \cdots]$, and $\xi'_4 = 5$. Notice that ξ' can also be obtained by the time evolution of the 7-color BBS applied to ξ.*

left, picking up each ball it encounters and depositing a ball into each empty box whenever it can. We will see that after we run this carrier over $\xi^{(t)}$, the resulting configuration is in fact $\xi^{(t+1)}$. Moreover, the maximum number of balls in the carrier during the sweep is in fact the first soliton length λ_1.

Now we introduce the *infinite-capacity carrier process* and the carrier version of the κ-color BBS dynamic. Denote

$$\mathcal{B}_\infty := \left\{ \mathbf{x} \in \{0, 1, \cdots, \kappa\}^{\mathbb{N}} \mid \mathbf{x} \text{ is nonincreasing and has finite support} \right\}, \tag{30}$$

which is the set of 'reversed' semi-standard Young tableaux of shape $1 \times \infty$ and letters from $\{0, \ldots, \kappa\}$. Namely, an element in this set is an infinite string of letters consisting of finitely many nonincreasing nonzero letters followed by an infinite string of zeros. An element $\mathbf{x}$ in $\mathcal{B}_\infty$ describes the state of the infinite-capacity carrier. If the carrier at state $\mathbf{x}$ encounters a new ball of color y, it produces a new carrier state $\mathbf{x}'$ and a new ball color y' according to the 'circular exclusion rule': *Inserting y into $\mathbf{x}$, y' is the largest letter in $\mathbf{x}$ with $y' < y$, and $\mathbf{x}'$ is obtained by replacing the leftmost letter y' in $\mathbf{x}$ with y.* More precisely, define a map $\Psi : \mathcal{B}_\infty \times \{0, 1, \cdots, \kappa\} \to \{0, 1, \cdots, \kappa\} \times \mathcal{B}_\infty$, $(\mathbf{x}, y) \mapsto (y', \mathbf{x}')$ by

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

$$\mathbf{x}'(i) = \mathbf{x}(i)\mathbf{1}(i \neq i^*) + y\mathbf{1}(i = i^*) \qquad \forall i \geq 1. \tag{31}$$

(ii) Suppose $y = 0$. Then $y' = \mathbf{x}(1) = \max(\mathbf{x})$ and

$$\mathbf{x}'(i) = \mathbf{x}(i + 1) \qquad \forall i \geq 1. \tag{32}$$

Fix a κ-color BBS configuration $\xi : \mathbb{N} \to \{0, 1, \cdots, \kappa\}$. Fix $\Gamma_0 \in \mathcal{B}_\infty$, and recursively define a new κ-color BBS configuration ξ' and a sequence $(\Gamma_x)_{x \geq 0}$ of elements of $\mathcal{B}_\infty$ by

$$(\xi'_{x+1}, \Gamma_{x+1}) = \Psi(\Gamma_x, \xi_{x+1}) \qquad \forall x \in \mathbb{N}. \tag{33}$$

We call the sequence $(\Gamma_x)_{x \geq 0}$ the *infinite capacity carrier process over ξ*. The carrier state Γ_x is determined by the balls in the interval $[1, x]$ (see Figure 4 for an illustration). Unless otherwise mentioned, we will assume $\Gamma_0 = \mathbf{0} = [0, 0, 0, \cdots] \in \mathcal{B}_\infty$. The induced update map $\xi \mapsto \xi'$ turns out to coincide with the κ-color BBS evolution (1). See Remark 3.4 for more details.

It is important to note that the carrier process $(W_x)_{n \in \mathbb{N}}$ we introduced in (8) can be derived from the infinite-capacity carrier process $(\Gamma_x)_{x \in \mathbb{N}}$ above by simply recording the number of balls of each color

$i = 1, \ldots, \kappa$. That is,

$$W_x = (m_1(\Gamma_x), \ldots, m_\kappa(\Gamma_x)) \quad \text{for all } x \geq 0, \tag{34}$$

where $m_i(\Gamma_x)$ denotes the number of balls of color (letter) i in Γ_x for $i = 1, \ldots, \kappa$.

Lemma 3.1 below states that the first soliton length λ_1 equals the maximum number of balls of positive colors in the associated carrier process.

Lemma 3.1. *Suppose the initial κ-color BBS configuration ξ has finite support. Let $(W_x)_{x \geq 0}$ and $(\Gamma_x)_{x \geq 0}$ be as before. Then*

$$\lambda_1(\xi) = \max_{x \geq 0} \|W_x\|_1 = \max_{x \geq 0}(\# \text{ of positive letters in } \Gamma_x). \tag{35}$$

For $\kappa = 1$, it is possible to precisely characterize all subsequent soliton lengths $\lambda_2, \lambda_3, \ldots$ by applying the 'excursion operator' to the carrier process multiple times and taking maximum [LLP20]. Roughly speaking, given the 1-dimensional carrier process $W = (W_x)_{x \geq 0}$ for $\kappa = 1$, which starts at 0 and takes value 0 for all large x, let $\mathcal{E}(W)$ denote the new lattice path that describes the excursion heights above the record minimum of W away from the rightmost global maximizer of W. Then $\lambda_2 = \max(\mathcal{E}(W))$, and $\lambda_3 = \max(\mathcal{E}^2(W))$, and so on. We currently do not have a similar κ-dimensional excursion operator for exactly describing the subsequent soliton lengths for the general multicolor case. However, we provide a lower bound on λ_j in terms of the jth largest 'excursion height' of the carrier process, which is enough to obtain sharp asymptotics for λ_j in the subcritical regime.

We introduce some notation. Let $\mathbf{0} = (0, 0, \cdots, 0) \in (\mathbb{Z}_{\geq 0})^\kappa$ denote the origin, and write

$$M_n := \sum_{x=1}^{n} \mathbf{1}(W_x = \mathbf{0}) \tag{36}$$

for the number of visits of W_x to $\mathbf{0}$ during $[1, n]$. For each $k \geq 1$, let T_k denote the time of the kth visit of W_x to $\mathbf{0}$ and set $T_0 = 0$. We say that the trajectories of W_x restricted to the time intervals $[T_{k-1}, T_k]$ between consecutive visits to $\mathbf{0}$ are its *excursions*. Also note that M_n defined at (36) equals the number of complete excursions of the carrier process during $[1, n]$. We will define the *height* of the carrier at site x by

$$\|W_x\|_1 = W_x(1) + \cdots + W_x(\kappa), \tag{37}$$

which equals the number of balls of positive color that the carrier possesses at site x. Define the *kth excursion height h_k* and *height of the final meander r_n* by

$$h_k = \max_{T_{k-1} \leq t \leq T_k} \|W_x\|_1, \qquad r_n = \max_{T_{M_n} \leq t \leq n} \|W_x\|_1. \tag{38}$$

Denote by $\mathbf{h}_1(n) \geq \mathbf{h}_2(n) \geq \cdots \geq \mathbf{h}_{M_n}(n)$ the order statistics of the excursion heights $h_1, \cdots, h_{M_n}$. We then have the following lemma.

Lemma 3.2. *Soliton decomposition of ξ is obtained as the union of the soliton decomposition of the support of each excursion of the carrier process over ξ. In particular, for $j, n \geq 1$, $\lambda_j(n) \geq \mathbf{h}_j(n)$.*

Proofs of Lemmas 3.1 and 3.2 are relagated to Section 12.

3.2. Finite capacity carrier processes and soliton numbers

In [KL20], it is shown that the row lengths of the invariant Young diagram of any κ-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 combinatorial lemmas in the present paper.

0	0	0	0	0	0	2	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	3	3	6	2	1	2	2	5	5	6	0	0	0	2	3	0	0	0	0
0	0	2	5	7	7	7	7	6	2	2	5	5	7	7	6	0	6	6	6	3	0	0	0

Γ_x (above)

ξ_x: | 0 | 2 | 5 | 7 | 3 | 2 | 6 | 1 | 0 | 2 | 5 | 5 | 7 | 6 | 0 | 0 | 6 | 2 | 3 | 0 | 0 | 0 | 0 |

ξ'_x: | 0 | 0 | 2 | 5 | 0 | 0 | 3 | 7 | 6 | 1 | 2 | 2 | 5 | 5 | 7 | 6 | 0 | 0 | 2 | 6 | 3 | 0 | 0 | 0 |

Figure 5. *Time evolution of the capacity-3 carrier process $(\Gamma_x)_{x\geq 0}$ over the 7-color initial configuration ξ, with new configuration ξ' consisting of existing ball colors. For instance, $\xi_2 = 2$, $\Gamma_2 = [2,0,0]$, and $\xi'_4 = 5$. Notice that while ξ is the same as in the example in Figure 4, the new 7-color BBS configuration ξ' is different. In this case, the map $\xi \mapsto \xi'$ does not agree with the 7-color BBS time evolution.*

First, fix an integer parameter $c \geq 1$ that we call *capacity*. Denote

$$\mathcal{B}_c = \{[x_1, \cdots, x_c] \in \{0, 1, \cdots, \kappa\}^c \mid x_1 \geq \cdots \geq x_c\}, \tag{39}$$

which can also be identified as the set of all $(1 \times c)$ semistandard tableaux with letters from $\{0, 1, \cdots, \kappa\}$. Define a map $\Psi_c : \mathcal{B}_c \times \{0, 1, \cdots, \kappa\} \to \{0, 1, \cdots, \kappa\} \times \mathcal{B}_c$, $([x_1, \cdots, x_c], y) \mapsto (y', [x'_1, \cdots, x'_c])$ 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

$$[x'_1, \cdots, x'_c] = [x_1, \cdots, x_{i^*-1}, y, x_{i^*+1}, \cdots, x_c]. \tag{40}$$

(ii) Suppose $x_c \geq y$. Then $y' = x_1$ and

$$[x'_1, x'_2, \cdots, x'_c] = [x_2, \cdots, x_c, y]. \tag{41}$$

Fix a κ-color BBS configuration $\xi : \mathbb{N} \to \{0, 1, \cdots, \kappa\}$. Let $\Gamma_0 = [0, \cdots, 0] \in \mathcal{B}_c$, and recursively define a new κ-color BBS configuration ξ' and a sequence $(\Gamma_x)_{x\geq 0}$ of elements of $\mathcal{B}_c$ by

$$(\xi'_{x+1}, \Gamma_{x+1}) = \Psi_c(\Gamma_x, \xi_{x+1}) \qquad \forall x \in \mathbb{N}. \tag{42}$$

We call the sequence $(\Gamma_x)_{x\geq 0}$ the *capacity-c carrier process over ξ*. See Figure 5 for an illustration.

The following lemma, which is proven in [KL20], gives a closed-form expression of the row sums of the invariant Young diagram:

Lemma 3.3. *Let $(\xi^{(t)})_{t\geq 0}$ be a κ-color BBS trajectory such that $\xi^{(0)}$ has finite support. For each $c \geq 1$, let $(\Gamma_{x;c})_{x\geq 0}$ denote the capacity-c carrier process over $\xi^{(t)}$. Then for all $k \geq 1$ and $t \geq 0$, we have*

$$\rho_1(\xi^{(0)}) + \cdots + \rho_k(\xi^{(0)}) \equiv \sum_{x=1}^{\infty} \mathbf{1}(\xi_x^{(t)} > \min \Gamma_{x-1;k}), \tag{43}$$

where $\min \Gamma_{x-1;k}$ denotes the smallest letter in $\Gamma_{x-1;k}$.

Proof. See equation (13) and Proposition 4.5 in [KL20]. We also provide a self-contained proof in Section 12.2. □

Remark 3.4. It is well-known that, if the capacity $c \geq 1$ is large enough compared to the number of balls of color ≥ 1 in the system, then the induced update map $\xi \mapsto \xi'$ agrees with the κ-color BBS time evolution (see, for example, [HKT01]). Also, once the capacity c is large enough, the capacity-c carrier process is equivalent to the infinite capacity carrier process in the sense that they always contain the same number of each positive letter. Hence, it follows that the map $\xi \mapsto \xi'$ defined in (33) coincides with the κ-color BBS time evolution defined in the introduction. In other words, the κ-color BBS dynamic

can be equivalently defined by repeatedly applying the infinite-capacity carrier process to the current ball configuration, analogously as in the $\kappa = 1$ case in [LLP20].

3.3. Modified Greene-Kleitman invariants for BBS

One 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 tableaux of the same shape. For each permutation σ, record the common shape of the Young tableaux as $\Lambda_{\mathrm{RS}}(\sigma)$. Let $\rho_i^{\mathrm{RS}}(\sigma)$ and $\lambda_j^{\mathrm{RS}}(\sigma)$ denote its ith row length and its jth column lengths, respectively. 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 the longest subsequence in σ that can be obtained by taking the union of k decreasing (resp. increasing) subsequences. That is, for each $k \geq 1$,

$$\rho_1^{\mathrm{RS}}(\sigma)) + \cdots + \rho_k^{\mathrm{RS}}(\sigma)) = \max\left(\left|\bigsqcup k \text{ increasing subsequences of } \sigma\right|\right), \tag{44}$$

$$\lambda_1^{\mathrm{RS}}(\sigma)) + \cdots + \lambda_k^{\mathrm{RS}}(\sigma)) = \max\left(\left|\bigsqcup k \text{ decreasing subsequences of } \sigma\right|\right). \tag{45}$$

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

If we consider the κ-color BBS trajectory started at $\xi^{(0)} = \sigma \mathbf{1}([1, n])$, then we obtain another Young diagram $\Lambda(\sigma) := \Lambda(\xi^{(0)})$, whose j^{th} column length equals the jth longest soliton length. Then a natural question arises: Do the sums of the first k rows and columns of $\Lambda(\sigma)$ 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. However, for the columns, it turns out that we just need to impose that the k decreasing subsequences be *noninterlacing*. In fact, in Lemma 3.5, we establish these modified Greene-Kleitman invariants for BBS in the more general setting when σ is an arbitrary κ-color BBS configuration with finite support, where having 0's and repetitions are both allowed.

Let $\xi : \mathbb{N} \to \{0, 1, \cdots, \kappa\}$ be a κ-color BBS configuration with finite support. For subsets $A, B \subseteq \mathbb{N}$, denote $A \prec B$ if $\max(A) < \min(B)$. We say A, B are *noninterlacing* if $A \prec B$ or $B \prec A$. We say ξ is *nonincreasing* on $A \subseteq \mathbb{N}$ if $\xi_{a_1} \geq \xi_{a_2}$ 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 ξ in A* by

$$\mathrm{NA}(A, \xi) := 1 + \sum_{i=2}^{|A|} \mathbf{1}(\xi_{a_{i-1}} < \xi_{a_i}). \tag{46}$$

Moreover, define the *penalized length of A with respect to ξ* by

$$\mathrm{L}(A, \xi) := \left[|A| - \sum_{i=\min A}^{\max A} \mathbf{1}(\xi_i = 0)\right] \mathbf{1}(\xi \text{ is non-increasing on } A). \tag{47}$$

Note that the summation in (47) is over the interval $[\min A, \max A] \cap \mathbb{Z}$, which may contain A properly.

Lemma 3.5. *Let $(\xi^{(t)})_{t \geq 0}$ be a κ-color BBS trajectory such that $\xi^{(0)}$ has finite support. Then for each $k, t \geq 0$, we have*

$$\rho_1(\xi^{(0)}) + \cdots + \rho_k(\xi^{(0)}) \equiv \max_{A_1 \sqcup \cdots \sqcup A_k = \mathbb{N}} \sum_{i=1}^{k} \mathrm{NA}(A_i, \xi^{(t)}), \tag{48}$$

$$\lambda_1(\xi^{(0)}) + \cdots + \lambda_k(\xi^{(0)}) \equiv \max_{A_1 \prec \cdots \prec A_k \subseteq \mathbb{N}} \sum_{i=1}^{k} \mathrm{L}(A_i, \xi^{(t)}). \tag{49}$$

The proof of Lemma 3.5 may be found in Section 12.2.

4. Proof of Theorem 2.1

In this subsection, we prove our first main result, Theorem 2.1. Let Σ^n be a uniformly chosen random permutation of the set $\{1, 2, \cdots, n\}$, and let $\xi^n = \Sigma^n \mathbf{1}([1, n])$ be the random n-color BBS configuration induced from Σ^n. Let $\lambda_k(n) = \lambda_k(\xi^n)$ denote the length of the kth longest soliton in ξ^n.

4.1. Proof of Theorem 2.1 for the columns

Our proof of Theorem 2.1 for the columns relies on Lemma 3.5 and the sharp asymptotic of longest decreasing subsequence of a uniform random permutation due to Baik, Deift and Johansson [BDJ99].

Proof of Theorem 2.1 for the columns. Fix an integer $k \geq 1$. It suffices to show that, almost surely,

$$\lim_{n \to \infty} n^{-1/2} \sum_{i=1}^{k} \lambda_i(n) = 2\sqrt{k}. \tag{50}$$

For each integer $k \geq 1$, let $L(k)$ denote the length of the longest increasing subsequence in a uniformly random permutation of k letters. By Lemma 3.5, recall that

$$\lambda_1(n) + \cdots + \lambda_k(n) = \max \left\{ \sum_{i=1}^{k} L(A_i, \xi^n) \,\Big|\, A_1 \prec \cdots \prec A_k \subseteq [1, n] \right\}. \tag{51}$$

We view a random permutation as a ranking among n i.i.d. Uniform$([0, 1])$ random variables $U_1, \cdots, U_n$. If $A \subseteq \{1, \cdots, n\}$, then the ranking of U_i for $i \in A$ gives a uniform random permutation of A, which we call a random permutation of $[n]$ restricted on A. Moreover, one can also see that if we restrict a random permutation on multiple disjoint subsets, then these smaller permutations are independent. Hence, if $A_1 \prec \cdots \prec A_k$ are noninterlacing subsets of $[0, n]$, then the permutations restricted on these subsets are independent. Moreover, since the random permutation model ξ^n does not assign color 0 on any site in $[0, n]$, for any increasing subsequence $A \subseteq [0, n]$ and its supporting interval $I = [\min A, \max A]$,

$$\mathrm{L}(A, \xi^n) = |A| \leq |I| = \mathrm{L}(I, \xi^n) \overset{d}{=} L(|I|). \tag{52}$$

It follows that

$$\sum_{i=1}^{k} \lambda_i(n) \overset{d}{=} \max \left\{ \sum_{i=1}^{k} L(n_i) \,\Big|\, \sum_{i=1}^{k} n_i = n, \; L(n_1), \ldots, L(n_k) \text{ are indepenent} \right\}. \tag{53}$$

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 $m \geq 1$,

(Lower tail): $\mathbb{P}\left(m^{-1/6}(L(m) - 2\sqrt{m}) \leq -t\right) \leq C \exp(-ct^3)$ for all $t \in [M, 2m^{1/3}]$; (54)

(Upper tail): $\mathbb{P}\left(m^{-1/6}(L(m) - 2\sqrt{m}) \geq t\right) \leq C \exp(-ct^{3/5})$ for all $t \in [M, m^{5/6} - 2m^{1/3}]$. (55)

Taking $t = (\log m)^2$, we obtain

$$\mathbb{P}\left(|L(m) - 2\sqrt{m}| \geq (\log m)^2 m^{1/6}\right) \leq 2C \exp(-c(\log m)^{6/5}). \tag{56}$$

Fix $\varepsilon > 0$. Note that if $m \geq \varepsilon\sqrt{n}$, then for any fixed $d > 0$,

$$\mathbb{P}\left(|L(m) - 2\sqrt{m}| \geq (\log m)^2 m^{1/6}\right) = O(n^{-d}). \tag{57}$$

Now, denote the random variable in the right-hand side of (53) by X. We write $X = \max(Y, Z)$, where

$$Y = \max\{L(n_1) + \cdots + L(n_k) : n_1 + \cdots + n_k = n, n_i \geq \varepsilon\sqrt{n} \text{ for all } i\}, \tag{58}$$

$$Z = \max\{L(n_1) + \cdots + L(n_k) : n_1 + \cdots + n_k = n, n_i < \varepsilon\sqrt{n} \text{ for at least one } i\}. \tag{59}$$

Denote $\mathcal{A} := \{(n_1, \ldots, n_k) : n_1 + \cdots + n_k = n, n_i \geq \varepsilon\sqrt{n} \text{ for all } i\}$. For each $\eta = (n_1, \ldots, n_k) \in \mathcal{A}$, denote $Y_\eta := L(n_1) + \cdots + L(n_k)$ and $M_\eta := 2(\sqrt{n_1} + \cdots + \sqrt{n_k})$. Then by a union bound and (57),

$$\mathbb{P}(|Y_\eta - M_\eta| \geq k(\log m)^2 m^{1/6}) = O(n^{-d}). \tag{60}$$

Note that $Y = \max_{\eta \in \mathcal{A}} Y_\eta$, and since there are at most n^k partitions of $[n]$ into k intervals, $|\mathcal{A}| \leq n^k$. So by a union bound, we have

$$\mathbb{P}\left(|Y - \max_{\eta \in \mathcal{A}} M_\eta|\right) \leq \sum_{\eta \in \mathcal{A}} \mathbb{P}\left(|Y_\eta - M_\eta| \geq k(\log m)^2 m^{1/6}\right) = O(n^{-d}) \tag{61}$$

for any fixed $d > 0$. The deterministic optimization problem

$$\max_{\eta \in \mathcal{A}} M_\eta = \max\{2\sqrt{n_1} + \cdots + 2\sqrt{n_k} : n_1 + \cdots + n_k = n, n_i \geq \varepsilon\sqrt{n} \forall \ i\} \tag{62}$$

achieves its maximum when $\sum_{i=1}^k |n_i - (n/k)|$ is minimized, in which case we have $|n_i - (n/k)| \leq 1$ for all $1 \leq i \leq k$. Denoting the maximizer as $n_1, \cdots, n_k$, it follows that, for all $1 \leq i \leq k$,

$$|\sqrt{n_i} - \sqrt{n/k}| \leq \frac{1}{\sqrt{n_i} + \sqrt{n/k}} \leq \frac{1}{2\sqrt{(n/k) - 1}}. \tag{63}$$

So this yields, for all sufficiently large $n \geq 1$,

$$\mathbb{P}\left(|Y - 2\sqrt{kn}| > 2k(\log n)^2 n^{1/6}\right) \tag{64}$$

$$\leq \mathbb{P}\left(|Y - 2\sqrt{kn}| > k(\log n)^2 n^{1/6} + \frac{k}{\sqrt{(n/k) - 1}}\right) = O(n^{-d})$$

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 $n_i > \varepsilon\sqrt{n}$, we continue to use the tail bound for $|L(n_i) - 2\sqrt{n_i}|$ in (57). Hence,

$$\mathbb{P}\left(Z > 2\sqrt{(k-1)n} + 2k(\log n)^2 n^{1/6} + k\varepsilon\sqrt{n}\right) = O(n^{-d}), \tag{65}$$

where the first term bounds the contribution from at most $k - 1$ intervals of size $\geq \varepsilon\sqrt{n}$, the second term is given by the BDJ tail bound in (57) and the last term gives a trivial bound for intervals of size $< \varepsilon\sqrt{n}$. Hence, if we choose $\varepsilon < 2/k(\sqrt{k-1} + \sqrt{k})$, then (64) and (65) give us
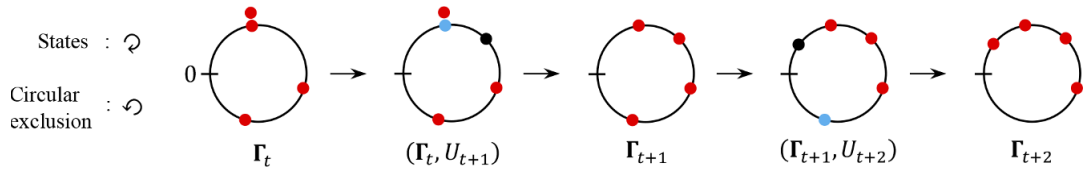
Figure 6. *Evolution of a 4-point circular exclusion process. The states in the unit circle are ordered clockwise. Each newly inserted point (black dot) annihilates the closest preexisting point in the counterclockwise direction (light blue dot).*

$$\mathbb{P}(Z > Y) \le \mathbb{P}\left(Y < 2\sqrt{kn} + 2k(\log n)^2 n^{1/6}\right) \tag{66}$$

$$+ \mathbb{P}\left(Z > 2\sqrt{(k-1)n} + 2k(\log n)^2 n^{1/6} + \frac{2\sqrt{n}}{\sqrt{k-1} + \sqrt{k}}\right)$$

$$= O(n^{-d})$$

for each fixed $d > 0$. Now note that, for each $t > 0$,

$$\mathbb{P}\left(\left|\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{k}\lambda_i(n)\right) - 2\sqrt{k}\right| > t\right) = \mathbb{P}\left(|\max(Y,Z) - 2\sqrt{kn}| > t\sqrt{n}\right) \tag{67}$$

$$\le \mathbb{P}\left(|Y - 2\sqrt{kn}| > t\sqrt{n}\right) + \mathbb{P}(Z > Y).$$

Hence, by choosing $t = 1/\log n$, for any fixed $d > 0$, (64) and (66) yield

$$\mathbb{P}\left(\left|\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{k}\lambda_i(n)\right) - 2\sqrt{k}\right| > \frac{1}{\log n}\right) = O(n^{-d}). \tag{68}$$

Then the assertion follows from the Borel-Cantelli lemma. □

4.2. Circular exclusion process and the row lengths

In this subsection, we prove Theorem 2.1 for the rows. By Lemma 3.3, this can be done by analyzing the carrier process over the uniform random permutation ξ^n. Let $\mathbf{X} := (U_x)_{x \ge 1}$ be a sequence of i.i.d. Uniform$[0,1]$ random variables. For each capacity $k \ge 1$, we may define the carrier process $(\mathbf{\Gamma}_x)_{x \ge 0}$ over $\mathbf{X}$ using the same 'circular exclusion rule' we used to define the map Ψ in Section 3.2. More precisely, denote $\mathcal{C}_k = \{(x_1, \cdots, x_k) \in [0,1]^k \mid x_1 \ge \cdots \ge x_k\}$. Define a map $\phi : \mathcal{C}_k \times [0,1] \to \mathcal{C}_k$, $[x_1, \cdots, x_k, y] \mapsto [x_1', \cdots, x_k']$ by

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

$$[x_1', \cdots, x_k'] = [x_1, \cdots, x_{i^*-1}, y, x_{i^*+1}, \cdots, x_k]. \tag{69}$$

(ii) If $x_k \ge y$, then $[x_1', \cdots, x_k'] = [x_2, \cdots, x_k, y]$.

Then the *k-point circular exclusion process* $(\mathbf{\Gamma}_x)_{x \ge 0}$ *over* $\mathbf{X}$ is defined recursively by

$$\mathbf{\Gamma}_{x+1} = \phi(\mathbf{\Gamma}_x, U_{x+1}). \tag{70}$$

See Figure 6 for an illustration. Note that $(\mathbf{\Gamma}_x)_{x \ge 0}$ forms a Markov chain on state space $\mathcal{C}_k$. When $\mathbf{\Gamma}_0 = [0, 0, \cdots, 0]$, we call $(\mathbf{\Gamma}_x)_{x \ge 0}$ the *carrier process over* $\mathbf{X}$ *with capacity* k.

In the following lemma, which will be proved in Section 4.3, we show that the k-point circular exclusion process converges to its unique stationary measure π, which is the distribution of the order statistics from k i.i.d. Uniform([0, 1]) variables.

Lemma 4.1. *Fix an integer $k \geq 1$ and let $(\Gamma_x)_{x \geq 0}$ denote the k-point circular exclusion process with an arbitrary initial configuration.*

 (i) *Let π denote the distribution of the order statistics from k i.i.d. uniform random variables on $[0, 1]$. Then π is the unique stationary distribution for the Markov chain $(\Gamma_x)_{x \geq 0}$.*
(ii) *For each $x \geq 0$, let π_x denote the distribution of Γ_x. Then π_x converges to π in total variation distance. More precisely,*

$$d_{TV}(\pi_x, \pi) := \sup_{A \subseteq [0,1]^k} |\pi_x(A) - \pi(A)| \leq \left(1 - \frac{1}{(2k)^{k-1} k!} \right)^{\lfloor x/k \rfloor}, \tag{71}$$

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

Now we derive Theorem 2.1 for the row asymptotics.

Proof of Theorem 2.1 for the rows. Let $\mathbf{X} = (U_x)_{x \geq 1}$ denote an infinite sequence of i.i.d. Uniform([0, 1]) random variables, Σ^n be the random permutation on $[n]$ induced by $U_1, \cdots, U_n$, and $\xi^n = \Sigma^n \mathbf{1}([1, n])$ be the random n-color BBS configuration as defined at (4). Fix an integer $k \geq 1$ and let $(\Gamma_x)_{x \geq 0}$ be the k-point circulr exclusion process over $\mathbf{X}$. Also, let $(\Gamma_x)_{x \geq 0}$ be the capacity-k carrier process over ξ^n as defined in Section 3.2. By construction, for each $1 \leq x \leq n$, we have

$$\mathbf{1}(\xi^n(x) > \min \Gamma_{x-1}) = \mathbf{1}(U_x > \min \Gamma_{x-1}). \tag{72}$$

Thus, according to Lemma 3.3, almost surely,

$$n^{-1}(\rho_1(\xi^n) + \cdots + \rho_k(\xi^n)) = n^{-1} \sum_{x=1}^{n} \mathbf{1}(U_x > \min \Gamma_{x-1}). \tag{73}$$

By Lemma 4.1 and Markov chain ergodic theorem, almost surely,

$$\lim_{n \to \infty} n^{-1}(\rho_1(\xi^n) + \cdots + \rho_k(\xi^n)) = \mathbb{P}(U_{k+1} > \min(U_1, \cdots, U_k)) = \frac{k}{k+1}. \tag{74}$$

Then the assertion follows. □

4.3. Stationarity and convergence of the circular exclusion process

We prove Lemma 4.1 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 section.

Proposition 4.2. *Fix an integer $k \geq 1$ and let π denote the distribution of the order statistics from k i.i.d. uniform random variables on $[0, 1]$. Then π is a stationary distribution of the k-point circular exclusion process.*

Proof of Lemma 4.1. For convergence, we use a standard coupling argument. Namely, fix arbitrary distributions π_0 and $\bar{\pi}_0$ on $\mathcal{C}_k$ and let $\mathbf{X} = (U_x)_{x \geq 1}$ denote a sequence of i.i.d. Uniform([0, 1]) variables. Let $(\Gamma_x)_{x \geq 0}$ be k-point circular exclusion processes over $\mathbf{X}$ with initial distribution π_0 and let $(\bar{\Gamma}_x)_{x \geq 0}$ be k-point circular exclusion processes over $\mathbf{X}$ with initial distribution $\pi_{\bar{0}}$. These two processes are naturally coupled since they evolve simultaneously over the same environment $\mathbf{X}$. Let $\tau = \inf\{x \geq 0 \mid \Gamma_x = \bar{\Gamma}_x\}$ denote the first meeting time of the two chains (see Figure 7). By the coupling, $\Gamma_s = \bar{\Gamma}_s$ and $s \leq x$ imply
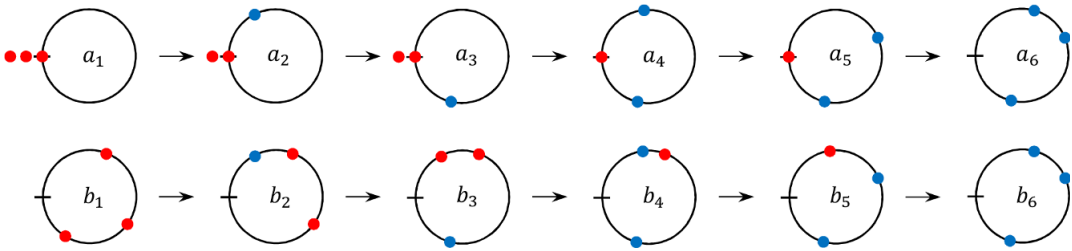
Figure 7. *Joint evolution of two 3-point circular exclusion processes. The states in the unit circle are ordered clockwise. A newly inserted point annihilates one of the closest preexisting points in the counterclockwise direction. Blue (resp., red) dots represent points that are shared (resp., not shared) in both processes. The two chains meet after the fifth transition.*

$\mathbf{\Gamma}_x = \bar{\mathbf{\Gamma}}_x$. A standard argument shows

$$d_{TV}(\pi_x, \bar{\pi}_x) \leq \mathbb{P}(\mathbf{\Gamma}_x \neq \bar{\mathbf{\Gamma}}_x) = \mathbb{P}(\tau > x), \tag{75}$$

where π_x and $\bar{\pi}_x$ denote the distributions of $\mathbf{\Gamma}_x$ and $\bar{\mathbf{\Gamma}}_x$. We claim that

$$\mathbb{P}(\tau > t) \leq \mathbb{P}(\mathbf{\Gamma}_0 \neq \bar{\mathbf{\Gamma}}_0)\left(1 - \frac{1}{(2k)^{k-1}k!}\right)^{\lfloor t/k \rfloor}. \tag{76}$$

According to Proposition 4.2, this will imply Lemma 4.1 by choosing $\bar{\pi}_0 = \pi$.

To bound the tail probability of meeting time τ, we will show that two circular exclusion processes 'synchronize' after k steps with probability at least $1/k!$, in the sense that

$$\mathbb{P}(\mathbf{\Gamma}_{x+k} = \bar{\mathbf{\Gamma}}_{x+k} \mid \mathbf{\Gamma}_x \neq \bar{\mathbf{\Gamma}}_x) \geq \frac{1}{(2k)^{k-1}k!} \qquad \text{for all } x \geq 0. \tag{77}$$

Then the claim (76) follows since

$$\mathbb{P}(\tau > Nk) = \mathbb{P}(\mathbf{\Gamma}_{Nk} \neq \bar{\mathbf{\Gamma}}_{Nk} \mid \mathbf{\Gamma}_0 \neq \bar{\mathbf{\Gamma}}_0)\mathbb{P}(\mathbf{\Gamma}_0 \neq \bar{\mathbf{\Gamma}}_0) \tag{78}$$

$$\leq \mathbb{P}(\mathbf{\Gamma}_0 \neq \bar{\mathbf{\Gamma}}_0)\prod_{i=1}^{N}\mathbb{P}(\mathbf{\Gamma}_{ik} \neq \bar{\mathbf{\Gamma}}_{ik} \mid \mathbf{\Gamma}_{(i-1)k} \neq \bar{\mathbf{\Gamma}}_{(i-1)k}) \tag{79}$$

$$\leq \mathbb{P}(\mathbf{\Gamma}_0 \neq \bar{\mathbf{\Gamma}}_0)\left(1 - \frac{1}{(2k)^{k-1}k!}\right)^{N}. \tag{80}$$

We begin with the following simple observation for a sufficient condition of meeting. Let $\mathbf{X} = (U_t)_{t \geq 1}$ be a sequence of i.i.d. Uniform$([0,1])$ variables. Fix $t \geq 1$ and let $\mathbf{\Gamma}_x = [x_1, \cdots, x_k]$ and $\bar{\mathbf{\Gamma}}_x = [\bar{x}_1, \cdots, \bar{x}_k]$ 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

$$\mathbf{\Gamma}_{x+k} = [U_{t+1}, U_{t+2}, \cdots, U_{t+k}] = \bar{\mathbf{\Gamma}}_{x+k}, \tag{81}$$

as all of the k points in $\mathbf{\Gamma}_x$ and $\mathbf{\Gamma}_x$ will be successively annihilated from the largest to the smallest by inserting $U_{t+1}, \cdots, U_{t+k}$.

For the general case, regard each U_s as a uniformly chosen point from the unit circle S^1. Then the $2k$ points $y_1, \cdots, y_{2k}$ will divide S^1 into disjoint arcs of lengths, say, $\ell_1, \cdots, \ell_m$, for some $2 \leq m \leq 2k$. If the points $U_{t+1}, \cdots, U_{t+k}$ are strictly decreasing in the counterclockwise order within one of the m

arcs, then by circular symmetry and a similar observation, we will have $\Gamma_{x+k} = \bar{\Gamma}_{x+k}$. Noting that

$$
\mathbb{P}\left(\begin{array}{c} U_{t+1}, \cdots, U_{t+k} \text{ are strictly decreasing in the} \\ \text{counterclockwise order within an arc of length } \ell \end{array} \right) = \frac{\ell^k}{k!} \tag{82}
$$

and $\ell_1 + \cdots + \ell_m = 1$, Hölder's inequality yields

$$
\mathbb{P}\left(\Gamma_{x+k} = \bar{\Gamma}_{x+k} \mid \Gamma_x = [x_1, \cdots, x_k], \bar{\Gamma}_x = [\bar{x}_1, \cdots, \bar{x}_k] \right) \tag{83}
$$

$$
\geq \sum_{i=1}^m \frac{\ell_i^k}{k!} \geq \frac{1}{k!} \frac{(\ell_1 + \cdots + \ell_m)^k}{m^{k-1}} = \frac{1}{m^{k-1}k!} \geq \frac{1}{(2k)^{k-1}k!}.
$$

This shows the assertion. □

Lastly in this section, we prove Proposition 4.2.

Proof of Proposition 4.2. We show π 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

$$
X_{(1)} < \cdots < X_{(I-1)} < Y < X_{(I+1)} < \cdots < X_{(k)}, \tag{84}
$$

where $I \in \{1, 2, \cdots, k\}$ is the random index such that $Y \in (X_{(I)}, X_{(I+1)})$. For $I = k$, the interval $(X_{(k)}, X_{(k+1)})$ denotes the union of $(0, X_{(1)})$ and $(X_{(k)}, 1)$. In this case, the point $X_{(k)}$ is deleted and Y is added as the smallest or largest point depending on which sub-intervals it falls.

We claim that (84) 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

$$
\mathbb{E}\left[f(X_{(1)}, \cdots, X_{(I-1)}, Y, X_{(I+1)}, \cdots, X_{(k)}) \right] \tag{85}
$$

$$
= \sum_{i=1}^k \mathbb{E}[f(X_{(1)}, \cdots, X_{(i-1)}, Y, X_{(i+1)}, \cdots, X_{(k)}) \mathbf{1}_{Y \in (X_{(i)}, X_{(i+1)})}] \tag{86}
$$

$$
= \sum_{i=1}^{k-1} \frac{1}{k!} \int_{z_1 < \cdots < z_i < y < z_{i+1} < \cdots < z_k} f(z_1, \cdots, z_{i-1}, y, z_{i+1}, \cdots, z_k) \, dz_1 \cdots dz_k \, dy \tag{87}
$$

$$
+ \frac{1}{k!} \int_{z_1 < \cdots < z_k < y} f(z_1, \cdots, z_{k-1}, y) \, dz_1 \cdots dz_k \, dy \tag{88}
$$

$$
+ \frac{1}{k!} \int_{y < z_1 < \cdots < z_k} f(y, z_1, \cdots, z_{k-1}) \, dz_1 \cdots dz_k \, dy. \tag{89}
$$

Integrating out z_i and denoting $z_0 := 0$,

$$
= \sum_{i=1}^{k-1} \frac{1}{k!} \int_{z_1 < \cdots < z_{i-1} < y < z_{i+1} < \cdots < z_k} f(z_1, \cdots, z_{i-1}, y, z_{i+1}, \cdots, z_k)(y - z_{i-1}) \tag{90}
$$

$$
dz_1 \cdots z_{i-1} z_{i+1} \cdots dz_k \, dy \tag{91}
$$

$$
+ \frac{1}{k!} \int_{z_1 < \cdots < z_{k-1} < y} f(z_1, \cdots, z_{k-1}, y)(y - z_{k-1}) \, dz_1 \cdots dz_{k-1} \, dy \tag{92}
$$

$$
+ \frac{1}{k!} \int_{y < z_1 < \cdots < z_{k-1}} f(y, z_1, \cdots, z_{k-1})(1 - z_{k-1}) \, dz_1 \cdots dz_{k-1} \, dy. \tag{93}
$$

We then rename y as z_i for the first integral above and as z_k for the second integral above. For the last integral, we rename y as z_1 and z_i as z_{i+1} for $i = 1, \ldots, k-1$. This gives

$$= \frac{1}{k!} \int_{z_1 < \cdots < z_k} f(z_1, \cdots, z_k) \left[(1 - z_k) + \left(\sum_{i=1}^{k-1} z_i - z_{i-1} \right) + (z_k - z_{k-1}) \right] dz_1 \cdots dz_k \qquad (94)$$

$$= \mathbb{E}\left[f(X_{(1)}, \cdots, X_{(I-1)}, X_{(I)}, X_{(I+1)}, \cdots, X_{(k)}) \right]. \qquad (95)$$

This shows the assertion. □

5. Proof of Theorem 2.3(i)

We prove Theorem 2.3 **(i)** in this section. Recall the probability distribution π in (9). We assume $p_0 > p^* := \max(p_1, \ldots, p_\kappa)$ in the following proof.

Proof of Theorem 2.3 (i). We first show the irreducibility and aperiodicity of the chain W_x. For its irreducibility, fix $\mathbf{x}, \mathbf{y} \in \mathcal{B}_\infty$ and write $\mathbf{y} = [y_1, y_2, \cdots]$. Since all elements of $\mathcal{B}_\infty$ have finite support, there exists an integer $m \geq 1$ such that $\mathbf{x}(i) \equiv 0$ and $\mathbf{y}(i) \equiv 0$ for all $i \geq m$. Then note that

$$\mathbb{P}(\Gamma_{x+2m} = \mathbf{y} \mid \Gamma_x = \mathbf{x}) \qquad (96)$$

$$\geq \mathbb{P}\left(\xi^{\mathbf{P}}(x+1) = 0, \cdots, \xi^{\mathbf{P}}(x+m) = 0, \xi^{\mathbf{P}}(x+m+1) = y_1, \cdots, \xi^{\mathbf{P}}(x+2m) = y_m \right) \qquad (97)$$

$$= p_0^m p_{y_1} \cdots p_{y_m} > 0. \qquad (98)$$

Since $\mathbf{x}, \mathbf{y} \in \mathcal{B}_\infty$ were arbitrary, this shows the Markov chain W_x is also irreducible. Then for its aperiodicity, it is enough to observe that

$$\mathbb{P}(\Gamma_{x+1} = [0, 0, \cdots] \mid \Gamma_x = [0, 0, \cdots]) = p_0 > 0. \qquad (99)$$

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

$$\exp(\mathrm{wt}(\mathbf{x})) = \prod_{i=1}^{\kappa} \left(\frac{p_i}{p_0} \right)^{m_i(\mathbf{x})}, \quad \exp(\mathrm{wt}(i)) = p_i. \qquad (100)$$

Recall the definition of the map $\Psi : \mathcal{B}_\infty \times \{0, 1, \cdots, \kappa\} \to \{0, 1, \cdots, \kappa\} \times \mathcal{B}_\infty$ given in Section 3.1. Note that for each pair $(\mathbf{x}, y) \in \mathcal{B}_\infty \times \{0, 1, \cdots, \kappa\}$ and $(y', \mathbf{x}') \in \{0, 1, \cdots, \kappa\} \times \mathcal{B}_\infty$ such that $\Psi(\mathbf{x}, y) = (y', \mathbf{x}')$, $y' = y'(\mathbf{x}, y)$, we have

$$\exp(\mathrm{wt}(\mathbf{x})) \exp(\mathrm{wt}(y)) = p_y p_0^{-\|\mathbf{x}\|_1} \prod_{i=1}^{\kappa} p_i^{m_i(\mathbf{x})} \qquad (101)$$

$$= p_{y'} p_0^{-\|\mathbf{x}'\|_1} \prod_{i=1}^{\kappa} p_i^{m_i(\mathbf{x}')} = \exp(\mathrm{wt}(y')) \exp(\mathrm{wt}(\mathbf{x}')). \qquad (102)$$

Indeed, the total number of each letter $1 \leq i \leq \kappa$ in both pairs $(\mathbf{x}, y)$ and $(y', \mathbf{x}')$ is the same. So if $y' \geq 1$, then some ball of positive color in $\mathbf{x}$ is replaced by a ball of positive color y', so $\|\mathbf{x}\|_1 = \|\mathbf{x}'\|_1$ and the above identity holds. If $y' = 0$ and $y \geq 1$, then $\mathbf{x}'$ has one more ball of color y than $\mathbf{x}$ does, so the above identity holds; if $y' = y = 0$, then both $\mathbf{x}'$ and $\mathbf{x}$ do not contain any ball of positive color, so the above identity holds.

Now, observe that for each fixed $\mathbf{x}' \in \mathcal{B}_\infty$, Ψ gives a bijection between $\{0, 1, \cdots, \kappa\} \times \{\mathbf{x}'\}$ and its inverse image under Ψ. If we denote the second coordinate of Ψ by Ψ_2, then this yields

$$\sum_{\substack{(\mathbf{x},y)\in\mathcal{B}_\infty\times\{0,1,\cdots,\kappa\} \\ \Psi_2(\mathbf{x},y)=\mathbf{x}'}} \exp(\mathrm{wt}(\mathbf{x}))\exp(\mathrm{wt}(y)) = \sum_{\substack{(\mathbf{x},y)\in\mathcal{B}_\infty\times\{0,1,\cdots,\kappa\} \\ \Psi_2(\mathbf{x},y)=\mathbf{x}'}} \exp(\mathrm{wt}(y'(\mathbf{x},y)))\exp(\mathrm{wt}(\mathbf{x}')) \quad (103)$$

$$= \exp(\mathrm{wt}(\mathbf{x}')) \sum_{y'\in\{0,1,\cdots,\kappa\}} \exp(\mathrm{wt}(y')) \quad (104)$$

$$= \exp(\mathrm{wt}(\mathbf{x}')). \quad (105)$$

Dividing both sides by

$$\sum_{\mathbf{x}\in\mathcal{B}_\infty} \exp(\mathrm{wt}(\mathbf{x})) = \sum_{n_1=0}^{\infty} \cdots \sum_{n_\kappa=0}^{\infty} \prod_{i=1}^{\kappa} \left(\frac{p_i}{p_0}\right)^{n_i} = \prod_{i=1}^{\kappa} \left(1 - \frac{p_i}{p_0}\right)^{-1} > 0, \quad (106)$$

we get

$$\sum_{\substack{(\mathbf{x},i)\in\mathcal{B}_\infty\times\{0,1,\cdots,\kappa\} \\ \Psi_2(\mathbf{x},i)=\mathbf{x}'}} \pi(m_1(\mathbf{x}), \cdots, m_\kappa(\mathbf{x}))p_i = \pi(m_1(\mathbf{x}'), \cdots, m_\kappa(\mathbf{x}')). \quad (107)$$

This shows that π is a stationary distribution of the Markov chain $(W_x)_{x\geq 0}$, as desired.

Lastly, positive recurrence follows from the irreducibility and the existence of stationary distribution [LP17, Thm. 21.13]. Convergence of the distribution of W_x to the stationary distribution in total variation distance then follows from the irreducibility, aperiodicity and positive recurrence (see [LP17, Thm. 21.16]). □

Remark 5.1. The statement and the proof of Theorem 2.3 **(i)** are reminiscent of [KL20, Thm. 1], where the authors show that for all $\mathbf{p} = (p_0, \cdots, p_\kappa)$, the (finite) capacity-$c$ carrier process over $\xi^{\mathbf{p}}$ is irreducible with unique stationary distribution

$$\pi_c(\mathbf{x}) = \frac{1}{Z_c} \prod_{i=0}^{\kappa} p_i^{m_i(\mathbf{x})}, \quad \mathbf{x} \in \mathcal{B}_\infty, \quad (108)$$

where Z_c denotes the partition function. In fact, their result applies to more general finite-capacity carriers whose state space is the set $B_c^{(a)}(\kappa)$ of all semistandard tableaux of rectangular shape $(c \times a)$ with letters from $\{0, 1, \cdots, \kappa\}$. In this general case, the partition function $Z_c = Z_c^{(a)}(\kappa, \mathbf{p})$ is identified with the Schur polynomial associated with the $(a \times c)$ Young tableau with constant entries c and parameters $p_0, p_1, \cdots, p_\kappa$.

6. The Skorohkod decomposition of the carrier process

In this section, we develop the Skorohkod decomposition of the carrier process, which we briefly mentioned in the introduction. The idea is to write the carrier process, which is confined in the nonnegative integer orthant $\mathbb{Z}_{\geq 0}^\kappa$, as the sum of a less confined process and a boundary correction. Namely, let $(W_x)_{x\geq 0}$ be the carrier process over an arbitrary ball configuration ξ as in (8). We seek for the following decomposition

$$W_x = X_x + RY_x \quad \text{for } x \geq 0, \quad (109)$$

where
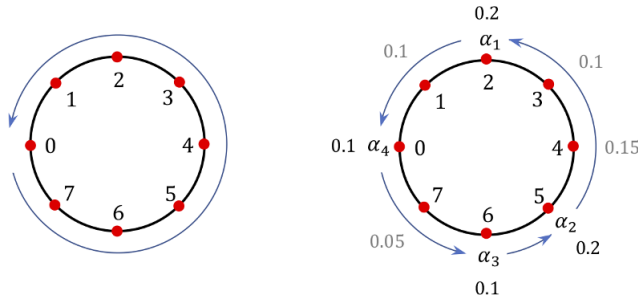
Figure 8. *Illustration of the original circular exclusion rule (left) and its decoupled version (right) for* $\kappa = 7$ *and ball density* $\mathbf{p} = (.1, .1, .25, .05, .15, .2, .1, .05)$. *We take the set of exceptional colors* $\mathcal{C}_e$ *to be the set of unstable colors* $\mathcal{C}_u^{\mathbf{P}} = \{2, 5, 6\}$. *For instance, in the decoupled carrier process, inserting new balls of color* 5 *into the carrier only excludes existing balls of colors* 2, 3 *and* 4.

1. $(X_x)_{x \geq 0}$ is the 'decoupled carrier process', which is a version of the carrier process that allows the number of balls of certain 'exceptional colors' to be negative;
2. $R = \mathrm{tridiag}_\kappa(0, 1, -1)$ is the $\kappa \times \kappa$ 'reflection matrix' (see (118));
3. $(Y_x)_{x \geq 0}$ is the 'pushing process': $Y_0 = \mathbf{0}$ and for each $i \in \{1, \ldots, \kappa\}$, the ith coordinate of Y_x is nondecreasing in x and can only increase when $W_x(i) = 0$.

We will first introduce the decoupled carrier process $(X_x)_{x \geq 0}$ in Section 6.1 and establish its basic properties in Proposition 6.2. In Section 6.2, we will introduce the reflection matrix R and the pushing process $(Y_x)_{x \geq 0}$ and verify the Skorohkod decomposition (109) in Lemma 6.3. All results in this section are for a deterministic ball configuration ξ.

6.1. Definition of the decoupled carrier process

In this section, we introduce a 'decoupled version' of the carrier process W_x in (8), which will be critical in proving Theorem 2.3 **(ii)** as well as Theorems 2.6–2.7.

To illustrate the idea, consider the carrier process W_x with $\kappa = 2$ as in Figure 1. While the transition kernel for this Markov chain depends on whether it is in the interior or at the boundary of the state space $\mathbb{Z}_{\geq 0}^2$, we may consider a similar Markov chain on the entire integer lattice $\mathbb{Z}^2$ that only uses the kernel in the interior, by allowing the counts of color 1 and 2 balls in W_x to be negative. In the general construction of decoupled carrier processes, we will allow the freedom to choose positive colors $\alpha_1 < \cdots < \alpha_r$ in $\{1, \ldots, \kappa\}$ whose count can be negative. Recall that inserting a ball of color i to the carrier W_n will exclude the largest color i_* in W_n that is less than i. In the decoupled carrier process, the color wheel $\mathbb{Z}_{\kappa+1}$ is divided into intervals $[0, \alpha_1], [\alpha_1, \alpha_2], \ldots, [\alpha_r, \kappa]$, and inserting a color i in $(\alpha_j, \alpha_{j+1}]$ can only exclude a color in the interval $[\alpha_j, \alpha_{j+1}]$. Hence, the interaction between colors in distinct intervals is 'decoupled'. See Figure 8 for an illustration.

Definition 6.1 (Decoupled carrier process). Let $\xi := (\xi_x)_{x \in \mathbb{N}}$ be κ-color ball configuration and fix a set $\mathcal{C}_e \subseteq \{1, \ldots, \kappa\}$ of 'exceptional colors'. Let

$$\Omega := \{(x_1, \ldots, x_\kappa) \in \mathbb{Z}^\kappa \ : \ x_i \geq 0 \text{ if } i \notin \mathcal{C}_e\}. \tag{110}$$

The *decoupled carrier process* over ξ associated with $\mathcal{C}_e$ is a process $(X_x)_{x \in \mathbb{N}}$ on the state space Ω defined as follows. If $\mathcal{C}_e = \emptyset$, then we take $X_x \equiv W_x$, where W_x is the carrier process in (8). Suppose $\mathcal{C}_e = \{\alpha_1, \ldots, \alpha_r\}$ for some $r \geq 1$ with $\alpha_1 < \cdots < \alpha_r$. Denote $\alpha_{r+1} := \kappa + 1$. Having defined $X_1, \ldots, X_x$,

denote $i := \xi_{x+1}$ if $\xi_{x+1} \in \{1, \ldots, \kappa\}$ and $i := \kappa + 1$ if $\xi_{x+1} = 0$. Then

$$X_{x+1} - X_x := \begin{cases} \mathbf{e}_i - \mathbf{1}(i_* \neq 0)\,\mathbf{e}_{i_*} & \text{if } 1 \leq i \leq \alpha_1 \\ \mathbf{e}_i - \mathbf{e}_{i'} & \text{if } \alpha_1 < i \leq \kappa \\ -\mathbf{e}_{i'} & \text{if } i = \kappa + 1, \end{cases} \tag{111}$$

where $i_* := \sup\{j : 1 \leq j < i, \ X_x(j) \geq 1\}$ (with the convention $\sup \emptyset = 0$) and

$$i' := \begin{cases} \alpha_j & \text{if } \alpha_j < i \leq \alpha_{j+1} \text{ and } X_x(\alpha_j) = \cdots = X_x(i-1) \leq 0 \\ q & \text{if } \alpha_j < q < i \leq \alpha_{j+1} \text{ and } X_x(q) \geq 1, X_x(q+1) = \cdots = X_x(i-1) = 0. \end{cases} \tag{112}$$

Unless otherwise mentioned, we take $X_0 = \mathbf{0}$ and $\xi = \xi^{\mathbf{p}}$ with density $\mathbf{p} = (p_0, \ldots, p_\kappa)$.

It is helpful to compare the recursion (111) for the decoupled carrier process to that of the carrier process in (8). Notice that in (8), inserting i into W_x can decrease by one at coordinate i_* only when $W_x(i_*) \geq 1$. Hence, W_x is confined in the nonnegative orthant $\mathbb{Z}_{\geq 0}^\kappa$. In comparison, when a ball of color i is inserted to the decoupled carrier X_x, it decreases by one at coordinate, say $\ell \in \{i', i_*\}$. If $\ell \notin \mathcal{C}_e$, then the above construction ensures that $X_x(\ell) \geq 1$. From this, one can observe that $X_x(j) \geq 0$ for all $x \geq 0$ whenever $j \notin \mathcal{C}_e$. In contrast, if $\ell \in \mathcal{C}_e$, then $X_{x+1}(\ell) = X_x(\ell) - 1$ regardless of whether $X_x(\ell) \geq 1$. Hence, X_x can take negative values on the exceptional colors. We call the recursion in (111) as the 'decoupled circular exclusion'.

In the proposition below, we establish a basic coupling result between the carrier and the decouple carrier processes. For its proof, we will introduce the following notation. Define the following function $f_W : \mathbb{Z}_{\geq 0}^\kappa \times \{0, \ldots, \kappa\} \to \{0, \ldots, \kappa\}$ as

$$f_W(\mathbf{w}, y) := \begin{cases} 0 & \text{if } [W_0 = \mathbf{w} \text{ and } \xi_1 = y \implies W_1 - W_0 = \mathbf{e}_y] \\ j & \text{if } [W_0 = \mathbf{w} \text{ and } \xi_1 = y \implies W_1 - W_0 = \mathbf{e}_y - \mathbf{e}_j \text{ or } -\mathbf{e}_j]. \end{cases} \tag{113}$$

Roughly speaking, if $f_W(\mathbf{w}, y) = j$, then j is the color of the ball that is excluded when a ball of color y is inserted into the carrier of state $\mathbf{w}$. The circular exclusion rule says $f_W(\mathbf{w}, y) = \sup\{i : 1 \leq i < y, \mathbf{w}(i) \geq 1\}$ with the convention $\kappa + 1 \equiv 0$ and $\sup \emptyset = 0$. Similarly, define a function $f_X : \Omega \times \{0, \ldots, \kappa\} \to \{0, \ldots, \kappa\}$ as

$$f_X(\mathbf{w}, y) := \begin{cases} 0 & \text{if } [X_0 = \mathbf{w} \text{ and } \xi_1 = y \implies X_1 - X_0 = \mathbf{e}_y] \\ j & \text{if } [X_0 = \mathbf{w} \text{ and } \xi_1 = y \implies X_1 - X_0 = \mathbf{e}_y - \mathbf{e}_j \text{ or } -\mathbf{e}_j]. \end{cases} \tag{114}$$

Intuitively, if $f_X(\mathbf{w}, y) = j$, then j is the color of the ball that is excluded when a ball of color y is inserted into the decoupled carrier of state $\mathbf{w}$.

For each $x \in \mathbb{N}$, define $\hat{X}_x \in \mathbb{Z}_{\geq 0}^\kappa$ by

$$\hat{X}_x(i) := X_x(i) - \min_{0 \leq s \leq x} X_s(i) \quad \text{for all } i = 1, \ldots, \kappa. \tag{115}$$

Note that $\hat{X}_x(i) \geq \max(0, X_x(i))$ for all i by definition and $X_0 = \mathbf{0}$. Also, $\hat{X}_x(i) \equiv X_x(i)$ for all $i \notin \mathcal{C}_e$ since $X_x(i) \geq 0$ for all $x \in \mathbb{N}$ and all $i \notin \mathcal{C}_e$.

Proposition 6.2 (Basic coupling between the carrier and the decoupled carrier processes). *Let $(W_x)_{x \geq 0}$ be the carrier process in (8) and let $(X_x)_{x \geq 0}$ be the decoupled carrier process in (111) associated with $\mathcal{C}_e = \{\alpha_1, \ldots, \alpha_r\}$ for some $r \geq 1$. Suppose these two processes evolve over the same ball configuration ξ and $W_0 = X_0 = \mathbf{0} \in \mathbb{Z}_{\geq 0}^\kappa$. Then the following hold.*

(i) $W_x(i) = X_x(i)$ for all $\alpha_r < i \le \kappa$ and $x \ge 0$. Furthermore,

$$W_x(\alpha_r) = X_x(\alpha_r) \quad \text{if } X_1(\alpha_r), \ldots, X_{x-1}(\alpha_r) \ge 1. \tag{116}$$

(ii) $W_x(i) \le \hat{X}_x(i)$ for all $1 \le i \le \kappa$ and $x \ge 0$. Furthermore, for each $x \ge 0$, denoting $y := \xi_{x+1}$ if $\xi_{x+1} \in \{1, \ldots, \kappa\}$ and $y := \kappa + 1$ if $\xi_{x+1} = 0$,

$$f_W(W_x, \xi_{x+1}) \le f_X(X_x, \xi_{x+1}) < y. \tag{117}$$

Proof. In this proof, we denote $y_X := f_X(X_x, \xi_{x+1})$ and $y_W := f_W(W_x, y)$. Note that $y_W, y_X \in [0, \xi_{x+1})$ (recall that $\sup \emptyset = 0$).

The second part of **(i)** follows from the first part of **(i)** and definition. Now we show the first part of **(i)** by induction on $x \ge 0$. For $x = 0$, we have $W_0 = X_0 = \mathbf{0}$. Denote $\ell := \alpha_r$ and suppose $W_x(i) = X_x(i)$ for all $\ell < i \le \kappa$ for some $x \ge 0$. If $y \le \ell$, then inserting a ball of color y into the carrier W_x and the decoupled carrier X_x does not affect their state for colors strictly larger than ℓ. Hence, $W_{x+1}(i) = W_x(i) = X_x(i) = X_{x+1}(i)$ for all $\ell < i \le \kappa$. So suppose $y > \ell$. In this case, $y_W = \sup\{1 \le j < y : W_x(j) \ge 1\}$ and $y_X = \max\{\ell, \sup\{1 \le j < y : X_x(j) \ge 1\}\}$. Note that W_{x+1} is obtained from W_x by increasing its value on color y by one and decreasing its value on color y_W by one. If $y_W > \ell$, then by the induction hypthesis, $y_W = y_X$, so X_{x+1} is obtained from $X_x|_{(\ell,\kappa]} = W_x|_{(\ell,\kappa]}$ by the same way, so $X_{x+1}|_{(\ell,\kappa]} = W_{x+1}|_{(\ell,\kappa]}$. Otherwise, suppose $y_W \le \ell$. Then X_{x+1} is obtained from $X_x|_{(\ell,\kappa]} = W_x|_{(\ell,\kappa]}$ by increasing its value on color y by one and decreasing its value on color ℓ by one. Hence, $W_{x+1}|_{(\ell,\kappa]} = X_{x+1}|_{(\ell,\kappa]}$, as desired.

Now we prove **(ii)** by an induction on $x \ge 0$. The base step when $x = 0$ follows by definition ($W_0 = \hat{X}_0 = \hat{X} = \mathbf{0}$ and $0 = y_W \le y_X < y$). For the induction step, suppose $W_x \le \hat{X}_x$ coordinatewise for some $x \ge 0$. We first show that $y_W \le y_X < y$. That $y_X < y$ follows from the definition (111). To show $y_W \le y_X$, we assume $y_W \ge 1$ since otherwise the claim holds trivially. Since a ball of color $y_W \ge 1$ is excluded from the carrier W_x, we have $W_x(y_W) \ge 1$. If $y_W \notin C_e$, then by the induction hypothesis, $1 \le W_x(y_W) \le X_x(y_W)$, so it follows that $y_W \le y_X$. Otherwise, suppose $y_W \in C_e$. Then since y_X is at least the largest exceptional color that is $< y$, it follows that $y_W \le y_X$, as desired.

It remains to show $W_{x+1} \le \hat{X}_{x+1}$ coordinatewise. First suppose $y_W = 0$. Then $W_x(1) = \cdots = W_x(y-1) = 0$, so $W_{x+1}(1) = \cdots = W_{x+1}(y-1) = 0$ and $W_{x+1} - W_x = \mathbf{e}_y$. Hence, $W_{x+1}(i) = 0 \le \hat{X}_{x+1}(y)$ for all $1 \le i < y$. Noting that $X_{x+1}(y) = X_x(y) + 1$, by definition, we have $\hat{X}_{x+1}(y) = \hat{X}_x(y) + 1$. Then by the induction hypothesis, we have $W_{x+1}(y) = W_x(y) + 1 \le \hat{X}_x(y) + 1 = \hat{X}_{x+1}(y)$. Furthermore, $W_{x+1}(i) = W_x(i) \le \hat{X}_x(i) = \hat{X}_{x+1}(i)$ for all $y < i \le \kappa$, where the middle inequality is from the induction hypothesis and the equalities are from the definition. Thus, we have shown that $W_{x+1} \le \hat{X}_{x+1}$ coordinatewise.

Lastly, we suppose $y_W \ge 1$ and show $W_{x+1} \le \hat{X}_{x+1}$ coordinatewise. Then $1 \le y_W \le y_X < y$, $W_{x+1} - W_x = \mathbf{e}_y - \mathbf{e}_{y_W}$ and $X_{x+1} - X_x = \mathbf{e}_y - \mathbf{e}_{y_X}$. By the induction hypothesis and the definition, we only need to verify $W_{x+1}(y_X) \le \hat{X}_{x+1}(y_X)$. This holds when $y_W = y_X$ since then $W_{x+1}(y_X) = W_x(y_X) - 1 \le \hat{X}_x(y_X) - 1 \le \hat{X}_{x+1}(y_X)$. So we may assume $y_W < y_X$. By definition of y_W, we have $W_x(y_W + 1) = \cdots = W_x(y-1) = 0$, and so $W_{x+1}(y_W + 1) = \cdots = W_{x+1}(y-1) = 0$. Then by definition, $W_{x+1}(y_X) = 0 \le \hat{X}_{x+1}(y_X)$. This completes the induction. $\qquad\square$

6.2. *Proof of the Skorokhod decomposition of the carrier process*

Now we give an explicit construction of the Skorokhod decomposition of $(W_x)_{x \ge 0}$. First, let R be the $\kappa \times \kappa$ tridiagonal matrix with 0 on the subdiagonal, 1 on the main diagonal and -1 on the superdiagonal

entries:

$$R := \text{tridiag}_\kappa(0, 1, -1) = \begin{bmatrix} 1 & -1 & 0 & & & \\ 0 & 1 & -1 & 0 & & \\ \vdots & & \ddots & & & \\ 0 & \cdots & 0 & 1 & -1 \\ 0 & & \cdots & 0 & 1 \end{bmatrix} = I - Q, \tag{118}$$

where I is the $\kappa \times \kappa$ identity matrix and $Q = I - R$. Notice that the spectral radius of Q is zero for all $\kappa \geq 2$ being an upper triangular matrix with zero diagonal entries. The above reflection matrix also has the property of being 'completely-$\mathcal{S}$'; see Definition 10.2 and the proof of Theorem 2.5 for justification.

Next, we define the pushing process $(Y_x)_{x \geq 0}$ on $\mathbb{Z}^\kappa_{\geq 0}$ recursively as follows: Set $Y_0 = \mathbf{0}$. Having defined Y_x, denoting $y_W := f_W(W_x, \xi_{x+1})$ (see (113)) and $y_X := f_X(X_x, \xi_{x+1})$ (see (114)), define

$$Y_{x+1} - Y_x := \begin{cases} \mathbf{0} & \text{if } y_W = y_X \\ \mathbf{e}_{y_W+1} + \cdots + \mathbf{e}_{y_X} & \text{if } y_W < y_X. \end{cases} \tag{119}$$

Note that (119) covers all cases since $y_W \leq y_X$ due to Proposition 6.2. From the definition, it is clear that every coordinate of Y_x is nondecreasing. Also, clearly, Y_x is determined by the first x ball colors $\xi_1, \ldots, \xi_x$.

Lemma 6.3 (Skorokhod decomposition of the carrier process). *Let* W_x, X_x, R *and* Y_x *as before. Then*

(i) $W_x = X_x + RY_x$ *for all* $x \geq 0$;

(ii) $Y_0 = \mathbf{0}$ *and for each* $i \in \{1, \ldots, \kappa\}$, *the ith coordinate of* Y_x *is nondecreasing in* x *and can only increase when* $W_x(i) = 0$; *that is,* $\sum_{x \geq 0} \mathbf{1}(W_x(i) \geq 1)(Y_{x+1}(i) - Y_x(i)) = 0$.

Proof. Let $y := \xi_{x+1}$ if $\xi_{x+1} \neq 0$ and $y := \kappa + 1$ if $\xi_{x+1} = 0$. Also let $y_W := f_W(W_x, \xi_{x+1})$ and $y_X := f_X(X_x, \xi_{x+1})$ (see (113) and (114)). We first show (ii). According to (117) in Proposition 6.2, we have $y_W \leq y_X < y$. Also, by the definition of y_W, we have $W_x(y_W + 1) = \cdots = W_x(y - 1) = 0$. Hence, if $Y_{x+1}(i) - Y_x(i) > 0$, then $i \in \{y_W + 1, \ldots, y - 1\}$, and hence, $W_x(i) = 0$. This shows (ii).

Next, we show (i) by induction on $x \geq 0$. It holds trivially when $x = 0$, so suppose for the induction step that it holds for some $x \geq 0$. We wish to show that

$$W_{x+1} = X_{x+1} + RY_{x+1}. \tag{120}$$

From (113)–(114), note that

$$(W_{x+1} - W_x) - (X_{x+1} - X_x) = \begin{cases} \mathbf{0} & \text{if } y_W = y_X \\ \mathbf{e}_{y_X} - \mathbf{e}_{y_W} & \text{if } 1 \leq y_W < y_X \\ \mathbf{e}_{y_X} & \text{if } 0 = y_W < y_X. \end{cases} \tag{121}$$

If $y_W = y_X$, then $R(Y_{x+1} - Y_x) = \mathbf{0}$, so (120) holds by the induction hypothesis. Next, suppose $1 \leq y_W < y_X$. Note that

$$R(Y_{x+1} - Y_x) = R(\mathbf{e}_{y_W+1} + \cdots + \mathbf{e}_{y_X}) \tag{122}$$
$$= (\mathbf{e}_{y_W+1} - \mathbf{e}_{y_W}) + (\mathbf{e}_{y_W+2} - \mathbf{e}_{y_W+1}) + \cdots + (\mathbf{e}_{y_X} - \mathbf{e}_{y_X-1}) \tag{123}$$
$$= \mathbf{e}_{y_X} - \mathbf{e}_{y_W}. \tag{124}$$

Lastly, suppose $0 = y_W < y_X$. Then

$$R(Y_{x+1} - Y_x) = R(\mathbf{e}_1 + \cdots + \mathbf{e}_{y_X}) \tag{125}$$
$$= \mathbf{e}_1 + (\mathbf{e}_2 - \mathbf{e}_1) + (\mathbf{e}_3 - \mathbf{e}_2) + \cdots + (\mathbf{e}_{y_X} - \mathbf{e}_{y_X-1}) = \mathbf{e}_{y_X}. \tag{126}$$

Hence, in all cases, the induction step holds by the induction hypothesis, (121) and (119). $\qquad \square$

7. Probabilistic analysis of the decoupled carrier process

In the previous section, we defined the decoupled carrier process $(X_x)_{x \geq 0}$ associated with an arbitrary set $\mathcal{C}_e = \{\alpha_1, \ldots, \alpha_r\} \subseteq \{1, \ldots, \kappa\}$ of exceptional colors over a deterministic ball configuration ξ. In this section, we establish various important probabilistic results for the decoupled carrier process $(X_x)_{x \geq 0}$ over the i.i.d. ball configuration $\xi^{\mathbf{p}}$ with a particular choice of the associated set $\mathcal{C}_e$ of exceptional colors.

7.1. Decomposition of the decoupled carrier process

Let $\mathbf{p} = (p_0, \ldots, p_\kappa)$ be the ball density at each site. We choose the set of exceptional colors $\mathcal{C}_e$ so that it satisfies the following 'stability condition':

$$\text{For all } 1 \leq j \leq r, \quad \max\{p_i : \alpha_j < i < \alpha_{j+1}\} < p_{\alpha_{j+1}}, \tag{127}$$

where we set $\alpha_0 = 0 = \alpha_{r+1}$. Since balls of a nonexceptional color i in (α_j, α_{j+1}) can be excluded by balls of color α_{j+1} in the decoupled carrier, the above condition ensures that $(X_x(i))_{x \geq 0}$ do not blow up. A canonical choice of such $\mathcal{C}_e$ is the set of unstable colors $\mathcal{C}_u^{\mathbf{p}}$ that we defined above the statement of Theorem 2.5.

Define the following processes:

$$\begin{cases} X_x := \text{The decoupled carrier process over } \xi = \xi^p \text{ associated with } \mathcal{C}_e \text{ satisfying (127)} \\ X_x^s := (\mathbf{1}(i \notin \mathcal{C}_e) \, X_x(i) \, ; \, i = 1, \ldots, \kappa) \qquad (\triangleright \text{ The 'stable part' of } X_x) \\ X_x^u := (\mathbf{1}(i \in \mathcal{C}_e) \, X_x(i) \, ; \, i = 1, \ldots, \kappa) \qquad (\triangleright \text{ The 'unstable part' of } X_x). \end{cases} \tag{128}$$

Namely, X_x^s (resp., X_x^u) agrees with X_x on the nonexceptional (resp., exceptional) colors but its coordinates on exceptional (resp., nonexceptional) colors are zero. Clearly, we have the following decomposition:

$$X_x = X_x^s + X_x^u \qquad \text{for all } x \geq 0. \tag{129}$$

In Lemma 7.1, we will show that $(X_x^s)_{x \geq 0}$ defines an irreducible Markov chain whose empirical distribution converges to its unique stationary distribution π^s defined as

$$\pi^s(n_1, \ldots, n_\kappa) = \prod_{j \in \mathcal{C}_u^{\mathbf{p}}} \mathbf{1}(n_j = 0) \prod_{j=0}^{r} \left[\prod_{\alpha_j < i < \alpha_{j+1}} \left(1 - \frac{p_i}{p_{\alpha_{j+1}}}\right)\left(\frac{p_i}{p_{\alpha_{j+1}}}\right)^{n_i} \right], \tag{130}$$

where we set $\alpha_0 = 0 = \alpha_{r+1}$. Hence, the expression in the bracket above is a nondegenerate geometric distribution. Thus, the above is the product of $\kappa - r$ geometric distributions, so it is indeed a probability distribution on Ω^s. Comparing (130) with (9), we see that the exceptional color α_{j+1} plays the role of color 0 for the nonexceptional colors in the interval $(\alpha_j, \ldots, \alpha_{j+1})$.

Lemma 7.1. *Let $(X_x^s)_{x \geq 0}$ be the process defined in (128). Then it is an aperiodic Markov chain on the state space $\mathbb{Z}_{\geq 0}^\kappa$ and has a unique communicating class with unique stationary distribution π^s defined in (130). Furthermore, if we denote the distribution of X_x^s by π_x^s, then*

$$\lim_{x \to \infty} d_{TV}(\pi_x^s, \pi^s) = 0. \tag{131}$$

Proof. First we show $(X_x^s)_{x \geq 0}$ defines a Markov chain. Clearly, the full decoupled carrier process $(X_x)_{x \geq 0}$ over $\xi = \xi^{\mathbf{p}}$ defines a Markov chain on $\mathbb{Z}^\kappa$. Hence, it is enough to show that X_{x+1}^s is determined from X_x^s and ξ_{x+1} for each $x \geq 0$. Fix $x \geq 0$ and denote $y := \xi_{x+1}$. Fix a nonexceptional color i. Let j be such that $\alpha_j < i < \alpha_{j+1}$. If $y \notin [i, \alpha_{j+1}]$, then $X_{x+1}^s(i) = X_x^s(i)$. If $y = i$, then $X_{x+1}^s(i) = X_x^s(i) + 1$. If $y \in (i, \alpha_{j+1}]$, then $X_{x+1}^s(i) - X_x^s(i) = -1$ if $X_x^s(i) \geq 1$ and $X_x^s(i+1) = \cdots = X_x^s(\alpha_{j+1} - 1) = 0$;

otherwise, $X_{x+1}^s(i) - X_x^s(i) = 0$. In all cases, $X_{x+1}^s(i)$ is determined by X_x^s and y. Since i was an arbitrary nonexceptional color, this verifies that $(X_x^s)_{x\geq0}$ is a Markov chain.

Next, let Ω^s denote the subset of $\mathbb{Z}_{\geq0}^\kappa$ consisting of all points whose coordinates on exceptional colors are zeroed out. Clearly, $(X_x^s)_{x\geq0}$ lives in Ω^s. We show the irreducibility of the chain $(X_x^s)_{x\geq0}$ on Ω^s. Aperiodicity will follow from irreducibility by noting that $\mathbf{0} \in \Omega^s$ is aperiodic. Observe that X_x^s visits every state eventually in Ω^s with positive probability starting from the initial state $\mathbf{0}$. Hence, it suffices to show the converse transition. Fix $\mathbf{x} = (x_1, \ldots, x_\kappa) \in \Omega^s$. Denote $n_1 = x_1 + \cdots + x_{\alpha_1-1}$, which is the number of balls of color in $[1, \alpha_1)$. Observe that inserting n_1 balls of color α_1 into the decoupled carrier X_x removes all balls of colors in $[1, \alpha_1)$ and leaves with $x_{\alpha_1} + n_1$ balls of color α_1. Next, we insert $x_{\alpha_1} + n_1 + n_2$ balls of color α_2 into the decoupled carrier, where $n_2 = x_{\alpha_1+1} + \cdots + x_{\alpha_2-1}$. This will remove all remaining balls of colors in $[1, \alpha_2)$ and leave $x_{\alpha_2} + (x_{\alpha_1} + n_1 + n_2)$ balls of color α_2. Repeating this process, we can remove all balls of stable colors in the decoupled carrier, so X_x^s visits $\mathbf{0}$ with a positive probability.

Next, we can verify that π^s is a stationary distribution of $(X_x^s)_{x\geq0}$ by using a similar argument as in the proof of Theorem 2.3 **(i)**. The key idea is the following: The evolution of balls of colors in (α_j, α_{j+1}) in the decoupled carrier X_x depends only on balls of colors in $(\alpha_j, \alpha_{j+1}]$ and inserting balls of color α_{j+1} can exclude any color in that interval. Moreover, the 'stable component' X_x^s of X_x does not count the number of balls of color α_{j+1} and recall the 'stability condition' (127). So one can treat α_{j+1} as color 0 in the subcritical carrier. We omit the details.

Lastly, the convergence of the empirical distribution in (131) follows from the same soft argument given at the end of the proof of Theorem 2.3 **(i)**. □

Next, we introduce a representation of the decoupled carrier process as a (truncated) partial sums process. By Lemma 7.1, $(X_x^s, \xi_{x+1})_{x\geq0}$ defines an aperiodic Markov chain on $\mathbb{Z}_{\geq0}^\kappa \times \{0, \ldots, \kappa\}$ with unique stationary distribution $\pi^s \otimes \mathbf{p}$. For each $\ell \in \{1, \ldots, \kappa\}$, define a functional $g^\ell : \mathbb{Z}^\kappa \times \{0, \ldots, \kappa\} \to \mathbb{Z}$ by

$$g^\ell(\mathbf{w}, i) := \begin{cases} 1 & \text{if } i = \ell \\ -1 & \text{if } \alpha_j \leq \ell < i \leq \alpha_{j+1} \text{ for some } j \in \{0, \ldots, r-1\} \\ & \qquad \text{and } \mathbf{w}(\ell+1) = \cdots = \mathbf{w}(i-1) = 0 \\ -1 & \text{if } \alpha_r \leq \ell, i = 0, \text{ and } \mathbf{w}(\ell+1) = \cdots = \mathbf{w}(i-1) = 0 \\ 0 & \text{otherwise,} \end{cases} \tag{132}$$

where we denoted $\alpha_0 := 0$. It is easy to verify that, for each $\ell \in \{1, \ldots, \kappa\}$ and $x \geq 0$,

$$X_{x+1}(\ell) = \begin{cases} X_x(\ell) + g^\ell(X_x^s, \xi_{x+1}) & \text{if } \ell \in \mathcal{C}_e \\ \max(0, X_x(\ell) + g^\ell(X_x^s, \xi_{x+1})) & \text{if } \ell \notin \mathcal{C}_e. \end{cases} \tag{133}$$

In words, the random variable $g^\ell(X_x^s, \xi_{x+1})$ gives the increment of $X_{x+1}(\ell)$ for exceptional ℓ; for nonexceptional ℓ, the same holds but with additional truncation at 0 to ensure the value of $X_x(\ell)$ stays nonnegative. In particular, we can view $X_x(\ell)$ for nonexceptional ℓ as a Lindley process in queuing theory.

Another consequence of the observation in (133) is that the decoupled carrier process X_x^u on the exceptional colors (the unstable component of X_x) can be written as an additive function of the Markov chain $(X_x^s, \xi_{x+1})_{x\geq0}$:

$$X_x^u = \sum_{z=1}^x \sum_{\ell \in \{\alpha_1, \ldots, \alpha_r\}} g^\ell(X_z^s, \xi_{z+1}) \mathbf{e}_\ell. \tag{134}$$

This representation will be used critically in Sections 7, 9 and 10.

In the following proposition, we compute the stationary expectation of the increments $g^\ell(X_x^s, \xi_{x+1})$ in (133).

Proposition 7.2 (Bias of the decoupled carrier). *Let g^ℓ be the function in (132). Then*

$$\mathbb{E}_{\pi^s \otimes \mathbf{p}}[g^\ell(X_x^s, \xi_{x+1})] = p_\ell - p_{\ell^+}, \tag{135}$$

where ℓ^+ is the smallest exceptional color strictly larger than ℓ. (If $\ell > \alpha_r$, then take $\ell^+ = 0$.)

Proof. Fix $j \in \{0, \ldots, r\}$ and $\alpha_j \le \ell < \alpha_{j+1}$. Denote $\ell^+ := \alpha_{j+1}$, where we take $\alpha_0 = 0$ and $\alpha_{r+1} = \kappa + 1 \equiv 0 \pmod{\kappa + 1}$. Denote $\zeta_x := g^\ell(X_x^s, \xi_{x+1})$. It is clear from the definition that

$$\mathbb{P}_{\pi^s \otimes \mathbf{p}}(\zeta_x = 1) = p_\ell. \tag{136}$$

It remains to show

$$\mathbb{P}_{\pi^s \otimes \mathbf{p}}(\zeta_x = -1) = p_{\ell^+}. \tag{137}$$

To this end, observe that

$$\mathbb{P}_{\pi^s \otimes \mathbf{p}}(\zeta_x = -1) = p_{\ell+1} + \sum_{i=\ell+2}^{\ell^+} \mathbb{P}_{\pi^s}(X_x^s(\ell+1) = \cdots = X_x^s(i-1) = 0)\, p_i. \tag{138}$$

Since X_x^s is distributed as the stationary distribution π^s for all $x \ge 0$,

$$\mathbb{E}_{\pi^s}\left[\sum_{\ell < i < \ell^+} X_{x+1}^s(i) - \sum_{\ell < i < \ell^+} X_x^s(i) \right] = 0. \tag{139}$$

Let T denote the random variable in the expectation above. Then

$$\mathbb{P}_{\pi^s \otimes \mathbf{p}}(T = -1) = \left(1 - \mathbb{P}_{\pi^s}\left(\sum_{\ell < i < \ell^+} X_x^s(i) = 0 \right) \right) p_{\ell^+}, \tag{140}$$

$$\mathbb{P}_{\pi^s \otimes \mathbf{p}}(T = 1) = p_{\ell+1} + \sum_{i=\ell+2}^{\ell^+-1} \mathbb{P}_{\pi^s}(X_x^s(\ell+1) = \cdots = X_x^s(i-1) = 0)\, p_i. \tag{141}$$

Since $T \in \{-1, 0, 1\}$ and (139) holds, this yields

$$p_{\ell^+} = p_{\ell+1} + \sum_{i=\ell+2}^{\ell^+} \mathbb{P}_{\pi^s}(X_x^s(\ell+1) = \cdots = X_x^s(i-1) = 0)\, p_i. \tag{142}$$

Note that the right-hand side equals $\mathbb{P}_{\bar\pi \otimes \mathbf{p}}(\zeta_x = -1)$ in (138), as desired. This shows the assertion. $\qquad\square$

7.2. Finite moments of return times of the decoupled carrier process

The main goal of this section is to prove Theorem 7.3 below, which shows that the first return time to the origin of the stable part of the decoupled carrier process $(X_x^s)_{x \ge 0}$ has finite moments of all orders. In fact, we prove this result in a more general setting that includes the excursions of $X_x(i)$ under the past maximum for exceptional colors i with a positive drift. (Handling such a general setting will be useful

in the proof of Proposition 9.2.) Define a new process $(\widetilde{X}_x)_{x \geq 0}$ on $\mathbb{Z}_{\geq 0}^\kappa$ by

$$\widetilde{X}_x(i) := \begin{cases} X_x(i) & \text{if } i \notin \mathcal{C}_e \\ \max_{0 \leq t \leq x} X_t^u(i) - X_x^u(i) & \text{if } i = \alpha_j \in \mathcal{C}_e \text{ for some } j \text{ and } p_{\alpha_j} > p_{\alpha_{j+1}} \\ 0 & \text{otherwise.} \end{cases} \tag{143}$$

Notice that $(\widetilde{X}_x)_{x \geq 0}$ defines a Markov chain on the nonnegative orthant $\mathbb{Z}_{\geq 0}^\kappa$.

Theorem 7.3. *Let $(\widetilde{X}_x)_{x \geq 0}$ be the Markov chain on $\mathbb{Z}_{\geq 0}^\kappa$ in (143). Assume (128) holds. Let τ denote its first return time to the origin. Then τ has finite moments of all orders. Furthermore, $(\widetilde{X}_x)_{x \geq 0}$ is irreducible, aperiodic and positive recurrent, and has a unique stationary distribution.*

We recall the following geometric ergodic theorem for Markov chains on a countable state space. It is an important tool for showing finite exponential moments of return times.

Theorem 7.4 (Geometric Ergodic theorem; Special case of Theorem 15.0.1 in [MT12]). *Let $(\mathbf{X}_t)_{t \geq 0}$ be a Markov chain on a countable state space Ω with transition kernel P, which is irreducible and aperiodic. Then the following conditions are equivalent:*

  (i) *There exists a state $x \in \Omega$ such that the return time of the chain to x has a finite exponential moment;*
  (ii) *The chain is geometrically ergodic; that is, there exists a function $V : \Omega \to [1, \infty)$, constant $\varepsilon \in (0, 1)$ and a finite set $\mathcal{C}$ such that*

$$PV(x) \leq (1 - \varepsilon)V(x) \quad \text{for all } x \in \Omega \setminus \mathcal{C}. \tag{144}$$

In order to prove Theorem 7.3, we will establish a general lemma on the first return time of Markov chains defined on the nonnegative integer orthant that abstracts important structure of the subcritical carrier process W_x. Its proof is relegated to the end of this section.

Lemma 7.5. *Let $Z_x = (Z_x(1), \ldots, Z_x(d))$ be an aperiodic and irreducible Markov chain on $\mathbb{Z}_{\geq 0}^d$. Suppose $Z_0 = \mathbf{0}$ and assume the following three properties:*

  (A1) *(Geometric ergodicity of top coordinate) The return time of $Z_x(d)$ to zero has a finite exponential moment.*
  (A2) *(Hierarchical dependence) There is a sequence of i.i.d. random variables $(\xi_x)_{x \in \mathbb{N}}$ with distribution $\mathbf{p}$ and functions $f_i : \mathbb{Z}^{d-i-1} \times \mathbb{R} \to \{-1, 0, 1\}$ such that*

$$Z_{x+1}(i) = \max(0, Z_x(i) + f_i(Z_x^{>i}, \xi_{x+1})) \quad \text{for all } x \in \mathbb{N} \text{ and } i \in \{0, \ldots, d-1\}, \tag{145}$$

  where $Z_x^{>i} := (Z_x(i+1), \ldots, Z_x(d))$. Furthermore, $Z_x^{>i}$ has a unique stationary distribution, say $\lambda^{>i}$.
  (A3) *(Coordinatewise negative drift) For all $i = 0, \ldots, d-1$,*

$$\mathbb{E}_{\lambda^{>i} \otimes \mathbf{p}}\left[f_i(Z_x^{>i}, \xi_{x+1}) \right] < 0. \tag{146}$$

Now fix $i \in \{0, \ldots, d-1\}$. For each $j \geq 1$, let τ_j be the jth return time of $(Z_x^{>i})_{x \geq 0}$ to the origin. Then τ_1 has finite moments of all orders. Furthermore, denote $R_j := Z_{\tau_j}(i)$ for $j \geq 0$. Then $(R_j)_{j \geq 1}$ is a Markov chain on $\mathbb{Z}_{\geq 0}$ such that there exists constants $c, K > 0$ for which

$$\sup_{m \geq K} \mathbb{E}[R_1 - R_0 \mid R_0 = m] \leq -c. \tag{147}$$

In addition, $(R_j)_{j \geq 1}$ is geometrically ergodic (see Theorem 7.4).

We now deduce Theorem 7.3 assuming Lemma 7.5.

Proof of Theorem 7.3. Let $\mathcal{C}_u^{\mathbf{p}}$ denote the set of unstable colors, which is empty in the subcritical regime $p_0 > \max(p_1, \ldots, p_\kappa)$ and nonempty in the critical and the supercritical regimes $p_0 \leq \max(p_1, \ldots, p_\kappa)$. In the latter case, we let $\alpha_1 < \cdots < \alpha_r$ denote the unstable colors. For each $x \geq 0$, we write $\widetilde{X}_x = (Y_x(0), Y_x(1), \ldots, Y_x(r))$, where

$$Y_x(0) := \big(X_x(1), \cdots, X_x(\alpha_1 - 1)\big), \tag{148}$$

and for each $j \in \{1, \ldots, r\}$ (setting $p_{\alpha_{r+1}} = p_0$),

$$Y_x(j) := \begin{cases} \big(X_x(\alpha_j), X_x(\alpha_j + 1), \cdots, X_x(\alpha_{j+1} - 1)\big) & \text{if } p_{\alpha_j} = p_{\alpha_{j+1}} \\ \big(\max_{1 \leq s \leq x} X_s(\alpha_j) - X_x(\alpha_j), X_x(\alpha_j + 1), \cdots, X_x(\alpha_{j+1} - 1)\big) & \text{if } p_{\alpha_j} > p_{\alpha_{j+1}}. \end{cases} \tag{149}$$

We will show that for each j, the return time to the origin of $(Y_x(j))_{x \geq 0}$ has finite moments of all orders. Then by an inductive argument (see the proof of Lemma 7.5), it follows that the return time of $(\widetilde{X}_x)_{x \geq 0}$ also has finite moments of all orders.

Denote $R_x := Y_x(j)$. Note that R_x is a Markov chain on $\mathbb{Z}_{\geq 0}^{\ell^+ - \ell}$ with $\ell = \alpha_j$. We wish to show that the return time to the origin of R_x has finite moments of all orders. We will only show this for the case of $p_{\alpha_j} > p_{\alpha_{j+1}}$, as a similar and simpler argument will show the desired statement for the case $p_{\alpha_j} = p_{\alpha_{j+1}}$.

First, consider a partial sums process $S_n = \sum_{k=1}^n \eta_k$, $S_0 = 0$, where the increments η_k take values from $\{-1, 0, 1\}$ and they are not necessarily i.i.d. Consider the new process $\overline{S}_n := \max_{1 \leq k \leq n} S_k - S_n$, which measures the height of the excursion of $(S_k)_{1 \leq k \leq n}$ below the running maximum. Note that $\overline{S}_n$ satisfies the following recursion:

$$\overline{S}_n - \overline{S}_{n-1} = \begin{cases} -\eta_n & \text{if } \eta_n = -1 \text{ or } \overline{S}_{n-1} \geq 1, \\ 0 & \text{if } \overline{S}_{n-1} = 0 \text{ and } \eta_n \in \{0, 1\}. \end{cases} \tag{150}$$

Equivalently, we have

$$\overline{S}_n = \max(0, \overline{S}_{n-1} - \eta_n). \tag{151}$$

Now suppose $\ell = \ell^+ - 1$ so that $R_x := \max_{1 \leq s \leq x} X_s(\ell) - X_x(\ell)$. In this case, $X_x(\ell)$ a simple random walk on $\mathbb{Z}$ with positive drift $p_\ell - p_{\ell^+} > 0$, so R_x is a birth-deach chain on $\mathbb{Z}_{\geq 0}$ with negative drift $p_{\ell^+} - p_\ell < 0$. In this case, the claim follows immediately. Hence, we may assume $\ell < \ell^+ - 1$. Notice that $(X_x(\ell^+ - 1))_{x \in \mathbb{N}}$ is a birth-deach chain on $\mathbb{Z}_{\geq 0}$ which moves to the right with probability p_{ℓ^+-1} and to the left with probability p_{ℓ^+}. Since $\ell < \ell^+ - 1$, by the choice of ℓ and ℓ^+, we have $p_{\ell^+-1} < p_{\ell^+}$. Hence, $X_x(\ell^+ - 1)$ has negative drift $p_{\ell^+-1} - p_{\ell^+} < 0$ on $\mathbb{Z}_{>0}$. Thus, the return time to the origin of $X_x(\ell^+ - 1)$ has a finite exponential moment. This verifies the hypothesis **(A1)** in Lemma 7.5; **(A2)** follows from the observation in the previous paragraph and (133); **(A3)** follows from Proposition 7.2. Therefore, by Lemma 7.5, we deduce that the return time to the origin of R_x has finite moments of all orders.

One can easily check the irreducibility of $\overline{X}_x$ by using a similar argument as in the proof of Lemma 7.1. Aperiodicity is clear, as one can stay at the origin in one step when a color 0 is encountered. We have established that the return time to the origin of $\overline{X}_x$ has finite moments of all orders. This implies that the chain is positive recurrent. Hence, the chain has a stationary distribution [LLP20, Thm. 21.13], and it is unique from the irreducibility and Kac's theorem [LLP20, Lem. 21.12]. □

We now prove Lemma 7.5. The argument is soft and inductive in nature.

Proof of Lemma 7.5. We first claim the following:

$$\begin{array}{c} \text{For each } i \in \{0, \ldots, d-1\}, \text{ the first return time of } (Z_x^{>i})_{x \geq 0} \\ \text{to some state } \mathbf{x} \text{ has a finite exponential moment.} \end{array} \tag{152}$$

We show the (152) by induction on $i = d - 1, \ldots, 0$. Fix $i \in \{0, \ldots, d-1\}$. The base step for $i = d - 1$ is given by the hypothesis (**A1**). For the induction step, suppose the first return time of $(\mathbb{Z}_x^{>i})_{x \geq 0}$ to some state $\mathbf{x}'$ has a finite exponential moment. Let τ_j denote the jth return time of $(Z_x^{>i})_{x \geq 0}$ to $\mathbf{x}'$. Consider a new process

$$(Q_j, \mathbf{x}') := (Z_{\tau_j}(i), Z_{\tau_j}(i+1), \ldots, Z_{\tau_j}(d)). \tag{153}$$

By the strong Markov property, this defines a Markov chain $(Q_j)_{j \geq 1}$ on $\mathbb{Z}_{\geq 0}$.

Step 1. (147) holds for $(Q_j)_{j \geq 1}$. We would like to show

$$\sup_{m \geq K} \mathbb{E}[Q_1 - Q_0 \mid Q_0 = m] \leq -c \tag{154}$$

for some constants $c, K > 0$. Instead of $Z_x(i)$, we consider its 'untruncated version'

$$\overline{Z}_x(i) := \sum_{\ell=1}^{j} f_i(Z_x^{>i}, \xi_{x+1}) \tag{155}$$

with $\overline{Z}_0(i) = 0$. (Note that $\overline{Z}_x(0) = Z_x(0)$ by the hypothesis.) Since $(Z_x^{>i})_{x \geq 0}$ is a Markov chain by the hypothesis (**A2**), by the strong Markov property, excursions from $\mathbf{x}$ for the recurrent chain $Z_x^{>i}$ are i.i.d. Hence, $\overline{Q}_j := \overline{Z}_{\tau_j}(i)$ for $j \geq 1$ forms a random walk, whose increments are i.i.d. and have the same distribution as $\overline{Q}_1$. We claim that this random walk has a negative drift:

$$\mathbb{E}[\overline{Q}_1] < 0. \tag{156}$$

To see this, first, note that

$$\lim_{x \to \infty} \frac{1}{x} \overline{Z}_x(i) = \mathbb{E}_{\lambda^{>i} \otimes \mathbf{p}}[f_i(Z_0^{>i}, \xi_1)] =: \alpha < 0 \tag{157}$$

by the hypothesis (**A3**). Since $\tau_1, \tau_2 - \tau_1, \tau_3 - \tau_2, \ldots$ are i.i.d. by the strong Markov property and since τ_1 has a finite exponential moment by the induction hypothesis, $\tau_j \to \infty$ almost surely. So $\overline{Z}_{\tau_j}/\tau_j \to \alpha < 0$ almost surely. Also, to the strong law of large numbers and the previous results,

$$\mathbb{E}[\overline{Q}_1] = \lim_{j \to \infty} \frac{\overline{Q}_j}{j} = \lim_{j \to \infty} \frac{\overline{Z}_{\tau_j}}{\tau_j} \frac{\tau_j}{j} = \alpha \mathbb{E}[\tau_1] < 0. \tag{158}$$

This shows the claim.

Now note that

$$\mathbb{E}[Q_1 - Q_0 \mid W_0 = m] = \mathbb{E}[Z_{\tau_1}(i) - Z_0(i) \mid Z_0(i) = m] \tag{159}$$

$$= \mathbb{E}[(Z_{\tau_1}(i) - Z_0(i))\mathbf{1}_{\tau_1 \leq m} \mid Z_0(i) = m] \tag{160}$$

$$+ \mathbb{E}[(Z_{\tau_1}(i) - Z_0(i))\mathbf{1}_{\tau_1 > m} \mid Z_0(i) = m] \tag{161}$$

$$= \mathbb{E}[\overline{Z}_{\tau_1}\mathbf{1}_{\tau_1 \leq m}] + \mathbb{E}[(Z_{\tau_1}(i) - Z_0(i))\mathbf{1}_{\tau_1 > m} \mid Z_0(i) = m] \tag{162}$$

$$= \mathbb{E}[\overline{Z}_{\tau_1}] - \mathbb{E}[\overline{Z}_{\tau_1}\mathbf{1}_{\tau_1 > m}] + \mathbb{E}[(Z_{\tau_1}(i) - Z_0(i))\mathbf{1}_{\tau_1 > m} \mid Z_0(i) = m]. \tag{163}$$

For the third equality, we have used the fact that $\tau_1 \leq m$ and $Z_0(i) = m$ in conjunction with the hypothesis imply $Z_x(i) \geq 0$ for all $0 \leq x \leq \tau_1$. Note that $|\overline{Z}_{\tau_1}| \leq \tau_1$ and τ_1 has a finite expectation by the induction hypothesis, so $\mathbb{E}[\overline{Z}_{\tau_1}\mathbf{1}_{\tau_1 > m}] \to 0$ as $m \to \infty$ by the dominated convergence theorem. Also,

$$\mathbb{E}[(Z_{\tau_1}(i) - Z_0(i))\mathbf{1}_{\tau_1 > m} \mid Z_0(i) = m] \leq \mathbb{E}[(\tau_1 - m)^+ \mathbf{1}_{\tau_1 > m}], \tag{164}$$

so again by the dominated convergence theorem, the above tends to zero as $m \to \infty$. Since $\mathbb{E}[\overline{Z}_{\tau_1}] < 0$ by (156), we have shown (154).

Step 2. $(Q_j)_{j \geq 1}$ **is geometrically ergodic.** Next, we show that the Markov chain $(Q_j)_{j \geq 0}$ on $\mathbb{Z}_{\geq 0}$ is geometrically ergodic. To this end, first note that $|Q_{j+1} - Q_j| \leq \tau_{j+1} - \tau_j$, so it has finite exponential moment by the hypothesis. By the dominated convergence theorem,

$$\lim_{\beta \searrow 0} \mathbb{E}\left[\frac{\exp(\beta(Q_{j+1} - Q_j)) - 1}{\beta} \,\middle|\, Q_j = m \right] = \mathbb{E}[Q_{j+1} - Q_j \,|\, Q_j = m]. \tag{165}$$

Let $c, K > 0$ be the constants in (154). Then by choosing sufficiently small $\beta > 0$, we can find $\varepsilon > 0$ such that

$$\mathbb{E}[\exp(\beta(Q_{j+1} - Q_j)) \,|\, Q_j = m] \leq 1 - \varepsilon \qquad \forall m \geq K. \tag{166}$$

So, by taking $V(x) = \exp(\beta x)$, we have $PV(x) \leq (1-\varepsilon)V(x)$ for all x outside the finite set $\{0, 1, \cdots, K\}$, verifying the geometric ergodicity condition for the chain Q_j.

Step 3. Completing the induction step. By the geometric ergodic theorem (Theorem 7.4), the first return time σ of the geometrically ergodic chain $(Q_j)_{j \geq 1}$ to some sate $x' \in \mathbb{Z}_{\geq 0}$ has a finite exponential moment. Denote $\mathbf{x} = (x', \mathbf{x}') \in \mathbb{Z}_{\geq 0}^{d-i}$. We now show that the first return time S of the chain $(Z_x^{\geq i})_{x \geq 0}$ to the state $\mathbf{x}$ has a finite exponential moment. Note that $S = \tau_\sigma$. Since σ has a finite exponential moment, there exists a constant $c > 0$ such that $\mathbb{P}(\sigma = l) \leq e^{-cl}$ for all $\ell \geq 1$. Also, by the induction hypothesis, τ_1 has a finite exponential moment. Hence, there exists $A > 1$ such that $\mathbb{E}[A^{\tau_1}] < \infty$. By choosing A sufficiently close to 1, and applying dominated convergence, we can assume $\mathbb{E}[A^{2\tau_1}] \leq e^{c/2}$. Now by Cauchy-Schwarz,

$$\mathbb{E}[A^S] = \mathbb{E}[A^{\tau_\sigma}] = \sum_{l=1}^{\infty} \mathbb{E}[A^{\tau_l} \mathbf{1}_{\sigma = l}] \leq \sum_{l=1}^{\infty} \sqrt{\mathbb{E}[A^{2\tau_l}]} \sqrt{\mathbb{P}(\sigma = l)} \tag{167}$$

$$= \sum_{l=1}^{\infty} \sqrt{\mathbb{E}[A^{2\tau_1}]^l} \sqrt{\mathbb{P}(\sigma = l)} \leq \sum_{l=1}^{\infty} e^{cl/4} e^{-cl/2} = \sum_{l=1}^{\infty} e^{-cl/4} < \infty. \tag{168}$$

This shows that S has a finite exponential moment, as desired. Thus far, we have shown (152).

Step 4. Concluding for the return time to the origin. Fix $i \in \{0, 1, \ldots, d-1\}$. By (152), there exists a state $\mathbf{x} \in \mathbb{Z}_{\geq 0}^{d-i}$ such that the first return time τ_1 of $(\mathbb{Z}_x^{>i})_{x \geq 0}$ to $\mathbf{x}$ has a finite exponential moment. Thus, τ_1 has finite moments of all orders. It is well-known that, for any recurrent and irreducible Markov chain on a countable state space, if for any state i the first moment of the first return time is finite, then this also applies to any other state. This generalizes to moments of all orders of the first return time [HJR53]. Therefore, we can conclude that the first return time of $(\mathbb{Z}_x^{>i})_{x \geq 0}$ to the origin has finite moments of all orders.

Lastly, let σ_j denote the jth return time of $(Z_x^{>i})_{x \geq 0}$ to the origin and denote $R_j := Z_{\tau_j}(0)$ for $j \geq 1$. We know that σ_1 has finite moments of all orders. We can repeat Steps 1–2 above for the chain $(R_j)_{j \geq 1}$ to conclude (147) and its geometric ergodicity. This completes the proof. $\qquad \square$

Remark 7.6. In [ADOS11], Aurzada, Döring, Ortgiese and Scheutzow show that having a finite exponential moment for first return times is actually *not* a class property. Hence, in the proof of Lemma 7.5, knowing that the first return time to some state $\mathbf{x}$ has a finite exponential moment does not necessarily imply that the first return time to the origin also has a finite exponential moment.

7.3. Linear and diffusive scaling limit of the decoupled carrier process

In this section, we establish linear and diffusive scaling limits of the decoupled carrier process. We start with an illustrating example.

Example 7.7. Suppose $\mathcal{C}_e = \{1, \ldots, \kappa\}$ so that all positive colors are exceptional. Denote $\eta_x := X_x - X_{x-1}$ for $x \geq 1$. Then $(\eta_k)_{k \geq 1}$ are i.i.d. random vectors in $\mathbb{Z}^\kappa$ with the following distribution:

$$\mathbb{P}(\eta_i = -\mathbf{e}_\kappa) = p_0, \quad \mathbb{P}(\eta_i = \mathbf{e}_1) = p_1, \quad \mathbb{P}(\eta_i = \mathbf{e}_j - \mathbf{e}_{j-1}) = p_j \text{ for } j = 2, \ldots, \kappa. \tag{169}$$

Then note that

$$\boldsymbol{\mu} := \mathbb{E}[\eta_i] = -p_0 \mathbf{e}_\kappa + p_1 \mathbf{e}_1 + \sum_{j=2}^{\kappa} (\mathbf{e}_j - \mathbf{e}_{j-1}) p_j \tag{170}$$

$$= \mathbf{e}_1 (p_1 - p_2) + \mathbf{e}_2 (p_2 - p_3) + \cdots + \mathbf{e}_\kappa (p_\kappa - p_0), \tag{171}$$

$$\Sigma := \mathbb{E}[\eta_i \eta_i^T] = \mathbf{e}_\kappa \mathbf{e}_\kappa^T p_0 + \mathbf{e}_1 \mathbf{e}_1^T p_1 + \sum_{j=2}^{\kappa} (\mathbf{e}_j - \mathbf{e}_{j-1})(\mathbf{e}_j - \mathbf{e}_{j-1})^T p_j \tag{172}$$

$$= \begin{bmatrix} p_1 + p_2 & -p_2 & 0 & & \\ -p_2 & p_2 + p_3 & -p_3 & 0 & \\ 0 & -p_3 & p_3 + p_4 & & \\ & 0 & & \ddots & -p_\kappa \\ & & & -p_\kappa & p_\kappa + p_0 \end{bmatrix}. \tag{173}$$

In this case, the decoupled carrier process $(X_x)_{x \geq 0}$ is a Markov chain on $\mathbb{Z}^\kappa$ with the mean and the covariance matrix of the increments η_x given as above. Then the linear interpolation of the linear interpolation of the d-dimensional process $(\frac{1}{\sqrt{n}}(X_n - n\boldsymbol{\mu})_{n \in \mathbb{N}}$ converges weakly to the d-dimensional Brownian motion with covariance matrix Σ (see, for example, [DMR94, Thm. 1] and the following remark). Note that $\boldsymbol{\mu} = \mathbf{0}$ if $p_0 = p_1 = \cdots = p_\kappa = 1/(\kappa + 1)$, which is a special case of the critical regime for the multicolor BBS (i.e., $p_0 = \max(p_1, \ldots, p_\kappa)$). See the simulation in Figure 2 for $\kappa = 2$ and uniform ball density. ▲.

Next, we compute the mean and the variance of the increments of the unstable part of the decoupled carrier process.

Proposition 7.8 (Mean and limiting covariance matrix). *Let $(X_x)_{x \geq 0}$ be the decoupled carrier process in (128). Denote $\zeta_x := X_x^u - X_{x-1}^u$ for $x \geq 1$. Then the following hold:*

(i) *We have*

$$\boldsymbol{\mu} := \mathbb{E}_{\pi^s \otimes \mathbf{p}}[\zeta_1] \equiv \mathbf{e}_{\alpha_1}(p_{\alpha_1} - p_{\alpha_2}) + \mathbf{e}_{\alpha_2}(p_{\alpha_2} - p_{\alpha_3}) + \cdots + \mathbf{e}_{\alpha_r}(p_{\alpha_r} - p_{\alpha_0}), \tag{174}$$

$$\mathbb{E}_{\pi^s \otimes \mathbf{p}}[\zeta_1 \zeta_1^T] = \sum_{\ell \in \{\alpha_1, \ldots, \alpha_r\}} \mathbf{e}_\ell \mathbf{e}_\ell^T \left(p_\ell + \sum_{\ell < q \leq \ell^+} p_q \prod_{\ell < j < q} \left(1 - \frac{p_j}{p_{\ell^+}}\right) \right) \tag{175}$$

$$- \sum_{\ell \in \{\alpha_1, \ldots, \alpha_{r-1}\}} (\mathbf{e}_\ell \mathbf{e}_{\ell^+}^T + \mathbf{e}_{\ell^+} \mathbf{e}_\ell^T) p_{\ell^+} \prod_{\ell < j < \ell^+} \left(1 - \frac{p_j}{p_{\ell^+}}\right). \tag{176}$$

(ii) *Define the 'limiting covariance matrix' $\Sigma \in \mathbb{R}^{\kappa \times \kappa}$ as*

$$\Sigma := \lim_{n \to \infty} n^{-1} \mathbb{E}_{\pi^s \otimes \mathbf{p}}[(X_n^u - n\boldsymbol{\mu})(X_n^u - n\boldsymbol{\mu})^T]. \tag{177}$$

Then Σ is well-defined, nonzero, symmetric and positive semidefinite.

Proof. We first show **(i)**. The stationary expectation of ζ_1 can be easily verified from Proposition 7.2. Denote $m_{(a,b)}(X_x^s) := \sum_{a < i < b} X_x^s(i)$, which is set to zero if $b \le a + 1$. From (134), we can write

$$\zeta_x = \sum_{\ell \in \{\alpha_1, \dots, \alpha_r\}} \mathbf{e}_\ell \left(\mathbf{1}(\xi_x = \ell) - \sum_{\ell + 1 \le q \le \ell^+} \mathbf{1}(\xi_x = q)\mathbf{1}(m_{(\ell,q)}(X_{x-1}^s) = 0) \right). \quad (178)$$

Then it is straightforward to compute

$$\zeta_x \zeta_x^T = \sum_{\ell \in \{\alpha_1, \dots, \alpha_r\}} \mathbf{e}_\ell \mathbf{e}_\ell^T \left[\mathbf{1}(\xi_x = \ell) + \sum_{\ell + 1 \le q \le \ell^+} \mathbf{1}(\xi_x = q)\mathbf{1}(m_{(\ell,q)}(X_{x-1}^s) = 0) \right] \quad (179)$$

$$- \sum_{\ell \in \{\alpha_1, \dots, \alpha_{r-1}\}} (\mathbf{e}_\ell \mathbf{e}_{\ell^+}^T + \mathbf{e}_{\ell^+} \mathbf{e}_\ell^T)\mathbf{1}(\xi_x = \ell^+)\mathbf{1}(m_{(\ell,\ell^+)}(X_{x-1}^s) = 0). \quad (180)$$

Thus, by taking the stationary expectation of $\zeta_x \zeta_x^T$ in conjunction with (130), we obtain the second identity in **(i)**.

Lastly, we show **(ii)**. Assuming Σ is well-defined and that it is symmetric and positive semidefinite is clear from the definition. Next, we argue that Σ is well-defined. Let $\bar{\zeta}_x := \zeta_x - \mathbb{E}[\zeta_x]$. For $i \ge 0$, let σ_i denote the number of steps that the Markov chain Z_x takes until it returns to the origin for the $i+1$st time, By strong Markov property, σ_i's are i.i.d. Furthermore, the excursions of Z_x from the origin (that is, Z_x restricted on the time intervals $[0, \sigma_0], [\sigma_0, \sigma_1], \dots$) are i.i.d. Furthermore, by Theorem 7.3 and the fact that ξ_x's are i.i.d. with distribution $\mathbf{p} = (p_0, \dots, p_\kappa)$, $p_0 > 0$ (in fact, we assume $\min(p_1, \dots, p_\kappa) > 0$), it follows that σ_0 has finite moments of all orders. Hence, there exists some $\lambda > 0$ such that $\mathbb{E}[\lambda^{\sigma_0}] < \infty$. Moreover, by Kac's theorem [LLP20, Lem. 21.12],

$$\lim_{n \to \infty} \frac{1}{\mathbb{E}[\sigma_1]} = \pi^s \otimes \mathbf{p}(\mathbf{0}, 0) = \pi^s(\mathbf{0})\, p_0, \quad (181)$$

where the stationary distribution π^s is explicitly given in (130).

Now consider decomposing the trajectory of Z_x into excursions from the origin. Write $s_i := \sum_{k=\sigma_i}^{\sigma_{i+1}-1} \bar{\zeta}_k \bar{\zeta}_k^T$. By the strong Markov property, $s_1, s_2, \dots$ are i.i.d. and also note that $\mathbb{E}[s_i] = \mathbf{0}$. Denote $\Sigma_n := \mathbb{E}[(X_n^u - n\boldsymbol{\mu})(X_n^u - n\boldsymbol{\mu})^T]$. Observe that

$$\Sigma_{\sigma_n} = \mathbb{E}[(X_{\sigma_n}^u - \sigma_n\boldsymbol{\mu})(X_{\sigma_n}^u - \sigma_n\boldsymbol{\mu})^T] = \mathbb{E}[(s_1 + \dots + s_n)(s_1 + \dots + s_n)^T] = n\mathbb{E}[s_1 s_1^T]. \quad (182)$$

So by the elementary renewal theorem, almost surely,

$$\lim_{n \to \infty} \frac{1}{\sigma_n}\Sigma_{\sigma_n} = \lim_{n \to \infty} \frac{n}{\sigma_n}\mathbb{E}[s_1 s_1^T] = \frac{\mathbb{E}[s_1 s_1^T]}{\mathbb{E}[\sigma_1]} = p_0 \pi^s(\mathbf{0})\, \mathbb{E}[s_1 s_1^T]. \quad (183)$$

To show the convergence holds along the whole sequence, let $T(n)$ denote the total number of visits of Z_x to the origin in the first n steps. Denote $r_n := \sum_{k=\sigma_{T(n)}}^{n} \bar{\zeta}_k \bar{\zeta}_k^T$. Then since $s_1, \dots, s_{T(n)}, r_n$ are independent and $\mathbb{E}[s_i] = 0$,

$$\Sigma_n = \mathbb{E}[(s_1 + \dots + s_{T(n)} + r_n)(s_1 + \dots + s_{T(n)} + r_n)^T] = \Sigma_{\sigma_{T(n)}} + \mathbb{E}[r_n r_n^T]. \quad (184)$$

Denote $\Lambda_n := \mathbb{E}[\sum_{x=1}^{n} \|\bar{\zeta}_x \bar{\zeta}_x^T\|]$, which is nondecreasing in n. Then similar argument as before shows that $\frac{1}{\sigma_n}\Lambda_{\sigma_n}$ converges a.s., and by the monotonicity of Λ_n, an elementary renewal theory argument shows that $n^{-1}\Lambda_n$ converges as $n \to \infty$. Now by Jensen's inequality,

$$\|\mathbb{E}[r_n r_n^T]\| \le \Lambda_n - \Lambda_{\sigma_{T(n)}}. \quad (185)$$

Since $\sigma_{T(n)} \leq n \leq \sigma_{T(n)+1}$ and $\mathbb{E}[\sigma_1] < \infty$, it follows that $\sigma_{T(n)}/n \to 1$ a.s. as $n \to \infty$. Hence, dividing both sides of (185) by n and letting $n \to \infty$ shows that $n^{-1}\|\mathbb{E}[r_n r_n^T]\| \to 0$ as $n \to \infty$. Then from (184), we deduce

$$\lim_{n \to \infty} n^{-1} \Sigma_n = \lim_{n \to \infty} \frac{\sigma_{T(n)}}{n} \frac{1}{\sigma_{T(n)}} \Sigma_{\sigma_{T(n)}} + \lim_{n \to \infty} \mathbb{E}[r_n r_n^T] \tag{186}$$

$$= \lim_{n \to \infty} \sigma_n^{-1} \Sigma_{\sigma_n} \tag{187}$$

$$= p_0 \pi^s(\mathbf{0}) \, \mathbb{E}[s_1 s_1^T] \tag{188}$$

$$= \pi^s(\mathbf{0}) \, p_0 \, \mathbb{E}\left[(\bar{\zeta}_1 + \cdots + \bar{\zeta}_{\sigma_1})(\bar{\zeta}_1 + \cdots + \bar{\zeta}_{\sigma_1})^T\right]. \tag{189}$$

Finally, since $\bar{\zeta}_x$'s are uniformly bounded and σ_1 has a finite expectation, the last expression is a matrix with finite entries by Wald's identity. From this formula, it is also easy to verify that Σ is nonzero. □

Now we establish linear and diffusive scaling limits of the decoupled carrier process on unstable colors. This is the main outcome of this section.

Proposition 7.9 (Limit theorems for the decoupled carrier process on unstable colors). *Let* $(X_x)_{x \geq 0}$ *be the decoupled carrier process in* (128). *Denote* $\zeta_x := X_x^u - X_{x-1}^u$ *for* $x \geq 1$. *Then the following hold.*

(i) *(SLLN) Almost surely,*

$$\lim_{n \to \infty} n^{-1} X_n = \mathbf{e}_{\alpha_1}(p_{\alpha_1} - p_{\alpha_2}) + \mathbf{e}_{\alpha_2}(p_{\alpha_2} - p_{\alpha_3}) + \cdots + \mathbf{e}_{\alpha_r}(p_{\alpha_r} - p_{\alpha_0}) := \boldsymbol{\mu}. \tag{190}$$

(ii) *(FCLT) Let* $(\overline{X}_v)_{v \in \mathbb{R}_{\geq 0}}$ *denote the linear interpolation of the lattice path* $(X_x - x\boldsymbol{\mu})_{x \in \mathbb{N}}$. *Let* $B = (B_t : 0 \leq t \leq 1)$ *denote the standard Brownian motion. Then as* $n \to \infty$,

$$(n^{-1/2}\overline{X}_{nt} \, ; \, 0 \leq v \leq 1) \Longrightarrow (B_t \, ; \, 0 \leq t \leq 1) \text{ in } C([0,1]), \tag{191}$$

where $B = (B_v : 0 \leq v \leq 1)$ *is the Brownian motion in* $\mathbb{R}^\kappa$ *with mean zero and covariance matrix* Σ *defined in* (177). *Here* $\Longrightarrow$ *denotes weak convergence in* $C([0,1])$.

Proof. Recall the decomposition $X_x = X_x^u + X_x^s$. From Lemma 7.1 and Theorem 7.3, we know that X_x^s is a geometrically mixing Markov chain on a subset of $\mathbb{Z}_{\geq 0}^\kappa$ with unique stationary distribution π^s in (130). Hence, $n^{-1}X_n^s$ converges to zero almost surely. Also, the linear interpolation of $(X_x^s)_{x \in \mathbb{N}}$ in diffusive scaling converges almost surely to zero in $C([0,1])$. Thus, it is enough to verify **(i)** and **(ii)** with X_n replaced by X_n^u.

Recall the Markov additive function representation (134) of X_x^u, where the underlying Markov chain $(X_x^s, \xi_x)_{x \geq 0}$ has the unique stationary distribution $\pi^u \otimes \mathbf{p}$ and is geometrically ergodic (see Theorem 7.3). Thus, **(i)** follows from the standard Markov chain ergodic theorem for positive Harris chains (see, for example, [MT12, Thm. 17.1.7]). Recall that the limiting covariance matrix Σ defined in (177) is well-defined and nontrivial by Proposition 7.8. Then **(ii)** follows from the functional CLT for multivariate strongly mixing processes (see, for example, [DMR94, Thm. 1] and the following remark). See also [RS10, Thm. 3.1]. For a functional central limit theorem for additive functionals (univariate) of a positive Harris chain, see [MT12, Thm. 17.4.4 and eq. (17.38)]. □

8. Proofs of Theorem 2.3(ii) and Theorem 2.4

We prove Theorem 2.3 **(ii)** and Theorem 2.4 in this section. Throughout this section, we fix a probability distribution $\mathbf{p} = (p_0, p_1, \cdots, p_\kappa)$ on $\{0, 1, \cdots, \kappa\}$, and let $(W_x)_{x \geq 0}$ be the carrier process in (8) over the i.i.d. configuration $\xi = \xi^\mathbf{p}$.

8.1. *Strong stability of the subcritical carrier process*

In order to prove Theorem 2.3 **(ii)**, we need stronger stability properties of the carrier process than what is stated in Theorem 2.3. More specifically, (1) if $W_0 = \mathbf{0}$, then its first return time to the origin has finite moments of all orders; and (2) if $W_0 \sim \pi$ and conditional on $\|W_0\|_1 = N$, it has a uniformly positive probability to visit the origin before it visits 'level' $N + 1$. These results are established in the following proposition. In the remainder of this section, we will denote $W_x^{\geq a} := (W_x(a), \ldots, W_x(\kappa))$ and $W_x^{<a} := (W_x(1), \ldots, W_x(a-1))$ and use similar notation for $X_x^{\geq a}$ and $X_x^{<a}$. This is the content of Proposition 8.1 below, and proving this result is the main goal of this section.

Proposition 8.1. *Suppose $p_0 > p^* := \max(p_1, \cdots, p_\kappa)$ and let $(W_x)_{x \geq 0}$ be the carrier process over $\xi^{\mathbf{p}}$. The following hold.*

(i) *The first return time of W_x to the origin has finite moments of all orders.*

(ii) *For each $m \in \mathbb{N}$, let $\tau_m = \inf\{x \geq 1 \,:\, \|W_x\|_1 = m\}$. There exists constants $L_0, N_0 \geq 1$ and $c_0 > 0$ such that*

$$\inf_{N \geq N_0} \mathbb{P}_\pi\left(\tau_0 < \min(\tau_N, c_0 N^2 + L_0) \,|\, \|W_0\|_1 = N\right) > 0. \tag{192}$$

We prove a series of lemmas in order to prepare for the proof of Proposition 8.1.

Lemma 8.2 (Birth-deach chain domination of excursions of the carrier). *Let $(W_x)_{x \geq 0}$ be the carrier process in (8) and suppose $p_0 > p^* := \max(p_1, \ldots, p_\kappa)$. Fix $a \in \{1, \ldots, \kappa\}$ and define a birth-deach chain $(S_x)_{x \geq 0}$ on $\mathbb{Z}_{\geq 0}$ by $S_0 := W_x(a)$ and*

$$S_{x+1} - S_x = \begin{cases} 1 & \text{if } \xi_{x+1} = a \\ -1 & \text{if } \xi_{x+1} = 0 \text{ and } S_x \geq 1 \\ 0 & \text{otherwise.} \end{cases} \tag{193}$$

Note that $(S_x)_{x \geq 0}$ is a birth-death chain on $\mathbb{Z}_{\geq 0}$ with negative drift $p_a - p_0 < 0$. For all $x \geq 0$,

$$\|W_x^{\geq a}\|_1 \leq S_x \quad \text{if } \min_{0 \leq t \leq x} W_t(a) \geq 1. \tag{194}$$

Proof. The proposition says that as long as the carrier has at least one ball of color a, then the total load $\|W_x^{\geq a}\|_1$ is dominated by S_x. This is easy to verify by induction. The inequality could be violated when $W_x(a) = 0$, since then the total load can increase by inserting balls of color $> a$ while S_x does not. $\quad\square$

In the statement and proofs below, we denote $\mathbb{P}_{\mathbf{x}}(\cdot) = \mathbb{P}(\cdot \,|\, W_0 = \mathbf{x})$.

Lemma 8.3 (Quadratic first hitting time of the origin of the subcritical carrier). *Let $(W_x)_{x \geq 0}$ be the carrier process in (8) and suppose $p_0 > p^* := \max(p_1, \ldots, p_\kappa)$. There exists a constant $c > 0$ such that*

$$\inf_{\mathbf{x} \in \mathbb{Z}_{\geq 0}^\kappa} \mathbb{P}_{\mathbf{x}}(\|W_x\|_1 = 0 \text{ for some } x \leq c\|\mathbf{x}\|_1^2) > 0. \tag{195}$$

Proof. We prove the assertion by induction on κ. If $\kappa = 1$, then the assertion follows easily since W_x then is a birth-deach chain on $\mathbb{Z}_{\geq 0}$ with negative bias $p_1 - p_0 < 0$ (e.g., see Lemma 8.5). For the induction step, note that $W_x^{\geq 2}$ behaves as the subcritical carrier process with ball colors $\{0, 2, 3, \ldots, \kappa\}$. That is, it evolves by the circular exclusion restricted on colors $\{0, 2, \ldots, \kappa\}$ while ignoring balls of color 1. Thus, $W_x^{\geq 2}$ is a lazy version of a carrier process with subcritical ball density as $\max(p_2, \ldots, p_\kappa) < p_0$. Let τ_i for $i = 1, 2, \cdots$ denote the ith time that $W_x^{\geq 2}$ returns to the origin. By the strong Markov property, $\tau_{i+1} - \tau_i$ for $i \geq 1$ are i.i.d., and they have finite moments of all orders by Lemma 7.5. Also, by the induction hypothesis, there exists a constant $c_1 > 0$ such that

$$\inf_{\mathbf{x} \in \mathbb{Z}_{\geq 0}^{\kappa-1}} \mathbb{P}_{\mathbf{x}}(\tau_1 \leq c_1\|\mathbf{x}\|_1^2) > 0. \tag{196}$$

Denote $Q_i := W_{\tau_i}(1)$ for $i \geq 1$. Then $(Q_i)_{i \geq 1}$ is a Markov chain on $\mathbb{Z}_{\geq 0}$. Denote $\sigma := \inf\{i \geq 1 : Q_i \leq L\}$, where $L \geq 1$ is a constant. Let $M := \|W_0\|_1$ and let $c_2 > 0$ be a constant to be determined. Introduce the following events:

$$E_1 := \{\tau_1 \leq c_1 M^2\}, \tag{197}$$

$$E_2 := \big\{ \max_{1 \leq k \leq \lfloor 2c_2 M \rfloor} |\tau_{i+1} - \tau_i| < M \big\}, \tag{198}$$

$$E_3 := \{W_{\tau_1}(1) \leq 2M\}, \tag{199}$$

$$E_4 := \{\sigma \leq c_2 W_{\tau_1}(1)\}. \tag{200}$$

These events depend on constants $M, L, c_2 > 0$ that we will subsequently choose below. Note that

$$\sigma_\tau = \tau_1 + \sum_{i=1}^{\sigma-1} (\tau_{i+1} - \tau_i) \leq \tau_1 + \sigma \max_{1 \leq i \leq \sigma} (\tau_{i+1} - \tau_i), \tag{201}$$

so $\sigma \leq 2c_2 M$ on $E_3 \cap E_4$. Hence, $\tau_\sigma \leq (c_1 + 2c_2)M^2$ on $E := \bigcap_{i=1}^4 E_i$. Also note that $\|W_{\tau_\sigma}\|_1 = W_{\tau_\sigma}(1) = Q_\sigma \leq L$. Hence, denoting $c := (c_1 + 2c_2) \vee 1$,

$$\big\{\|W_x\|_1 \leq L \text{ for some } x \leq cM^2\big\} \supseteq E. \tag{202}$$

Moreover,

$$\mathbb{P}_{\mathbf{x}}(\|W_x\|_1 = 0 \text{ for some } x \leq cM^2 + L) \geq p_0^L \, \mathbb{P}_{\mathbf{x}}(\|W_x\|_1 \leq L \text{ for some } x \leq cM^2). \tag{203}$$

Furthermore, since $c \geq 1$,

$$\inf_{\|\mathbf{x}\|_1 < M} \mathbb{P}_{\mathbf{x}}(\|W_x\|_1 = 0 \text{ for some } x \leq c\|\mathbf{x}\|_1^2 + L) \geq p_0^M. \tag{204}$$

Therefore, it suffices to show that for some constant $M_0 \geq 1$,

$$\inf_{M \geq M_0} \inf_{\|\mathbf{x}\|_1 = M} \mathbb{P}_{\mathbf{x}}(E) > 0. \tag{205}$$

Since E_1 has a uniformly positive probability by the induction hypothesis (196), it is enough to show that E_2, E_3, E_4 have high probaiblity to occur.

For E_2, since $\tau_{i+1} - \tau_i$ for $i \geq 1$ are i.i.d. and have finite moments of all orders, it follows that E_2 occurs with a high probability if M is sufficiently large. To see this, note that

$$\mathbb{P}(E_2) = (1 - \mathbb{P}(\tau_2 - \tau_1 \geq M))^{\lfloor 2c_2 M \rfloor} \geq \left(1 - \frac{\mathbb{E}[(\tau_2 - \tau_1)^2]}{M^2}\right)^{\lfloor 2c_2 M \rfloor} \to 1 \quad \text{as } M \to \infty. \tag{206}$$

For E_3, by Lemma 8.2, on the event that $W_{\tau_1}(1) > 2M$, a negatively biased birth-death chain $(S_x)_{x \geq 0}$ on $\mathbb{Z}_{\geq 0}$ makes an up-crossing of height at least M in $c_1 M^2$ steps, so

$$1 - \mathbb{P}_{\mathbf{x}}(E_3) \leq \mathbb{P}\left(\max_{0 \leq x \leq \lfloor c_1 M^2 \rfloor} S_x > M \,\Big|\, S_0 = 0\right). \tag{207}$$

Since S_x is a negatively biased simple random walk, the probability in the last expression is exponentially small in M.

For E_4, by Lemma 7.5, there are constants $K, c_3 > 0$ such that

$$\sup_{m \geq K} \mathbb{E}[Q_1 - Q_0 \,|\, Q_0 = m] \leq -c_3. \tag{208}$$

By Lemma 8.5, $\sigma \leq c_4 Q_1 = c_4 W_{\tau_1}(1)$ occurs with probability at least $1 - \frac{1}{c_3 c_4}$ for some constant $c_4 > 0$. Hence, by choosing $L \geq K$, $c_2 \geq c_4$, and letting c_4 sufficiently large, E_4 occurs with a high probability. This shows the assertion. $\qquad\square$

Lemma 8.4 (Growth of (sub-)critical carrier). *Let $(W_x)_{x \geq 0}$ be the carrier process in (8) with arbitrary initial state W_0 and suppose $p_0 \geq p^* := \max(p_1, \ldots, p_\kappa)$. Then for each $\varepsilon > 0$, almost surely,*

$$\limsup_{n \to \infty} n^{-1} \max_{0 \leq x \leq n} \|W_x\|_1 \leq \varepsilon. \tag{209}$$

Proof. Suppose $W_0 = (W_0(1), \ldots, W_0(\kappa))$ is arbitrary and write $M := \|W_0\|_1$. We may prepend to the ball configuration ξ the following sequence:

$$(\underbrace{\kappa, \ldots, \kappa}_{W_0(\kappa)}, \underbrace{\kappa - 1, \ldots, \kappa - 1}_{W_0(\kappa-1)}, \ldots, \underbrace{1, \ldots, 1}_{W_0(1)}) \tag{210}$$

and denote the extended configuration $\tilde\xi = (\tilde\xi_1, \ldots, \tilde\xi_M, \xi_1, \xi_2, \ldots)$. Let $\tilde W$ denote the carrier process with zero initial state run on $\tilde\xi$. Then after scanning the first M in the extended configuration, the new carrier $\tilde W$ attains exactly the same state W_0 (i.e., $\tilde W_M = W_0$), and thereafter, it undergoes the same dynamics as W (i.e., $\tilde W_{x+M} = W_x$ for all $x \geq 0$). Furthermore, $\max_{0 \leq x \leq n}\|W_x\|_1 \leq \max_{0 \leq x \leq n+M}\|\tilde W_x\|_1$, so it is enough to show the assertion for $\tilde W$. For simplicity, below we will denote $\tilde W$ and $\tilde\xi$ as W and ξ, respectively, and assume that the first M entries of ξ may be deterministic.

Fix $\varepsilon > 0$. By Lemmas 3.1 and 3.5,

$$\max_{0 \leq x \leq n} \|W_x\|_1 = \lambda_1(n) = \max_{A_1 \subseteq [0,n]} L(A_1, \xi), \tag{211}$$

where the right-hand side equals the penalized length of the longest nonincreasing subsequence in $\xi(n) := (\xi_0, \xi_1, \ldots, \xi_n)$. Let $D_i(x_1, x_2)$ denote the number of i's minus the number of 0's in $(\xi_{x_1}, \xi_{x_1+1}, \ldots, \xi_{x_2})$. If $\lambda_1(n) \geq \varepsilon n + M$, then $D_i(x_1, x_2) \geq \varepsilon n/\kappa$ for some i and $M < x_1 \leq x_2 \leq n$. Note that $D_i(x_1, x_2)$ is the sum of $x_2 - x_1$ i.i.d. Bernoulli variables with success probability $p_i - p_0 \leq 0$. Hence, by union bound and Hoeffding's inequality,

$$\mathbb{P}(\lambda_1(n) \geq \varepsilon n + M) \leq \sum_{i=1}^{\kappa} \sum_{M < x_1 \leq x_2 \leq n} \mathbb{P}(D_i(x_1, x_2) \geq \varepsilon n/\kappa) \tag{212}$$

$$\leq \kappa n^2 \exp(-cn) \tag{213}$$

for some constant $c > 0$. By Borel-Cantelli lemma, it follows that $\limsup_{n \to \infty} \lambda_1(n)/n \leq \varepsilon$ almost surely. Then the assertion follows. $\qquad\square$

We remark Theorem 2.5, which will be proved in Section 10, establishes the exact asymptotic $\max_{0 \leq x \leq n}\|W_x\|_1 \sim C\sqrt{n}$ for some constant $C > 0$.

Lemma 8.5 (Drift and bound on hitting time). *Let $(Y_t)_{t \geq 0}$ be a Markov chain on $\mathbb{Z}_{\geq 0}$ with transition kernel P. Suppose $\mathbb{E}_x[|Y_t|] < \infty$ for all $x, t \geq 0$ and there exists constants $c, L > 0$ such that*

$$\mathbb{E}_x[Y_1 - x] \leq -c \qquad \text{for all } x \geq L. \tag{214}$$

Let $\tau := \inf\{t \geq 0 : Y_t \leq L\}$. Then

$$\mathbb{P}_x(\tau \geq Cx) \leq \frac{1}{cC} \qquad \text{for all } x \geq 0 \text{ and } C > 0. \tag{215}$$

Proof. For any function $g : \mathbb{Z}_{\geq 0} \to \mathbb{R}$, denote $Pg(x) := \sum_y g(y) P(x, y)$ and $PY := P \operatorname{id}(Y)$. Note that the condition (214) reads

$$Px - x \leq -c \qquad \text{for all } x \notin [0, L]. \tag{216}$$

Define the compensator $(K_t)_{t \geq 0}$ of $(Y_t)_{t \geq 0}$ as $K_0 = 0$ and

$$K_n := \sum_{k=0}^{n-1} (PY_k - Y_k). \tag{217}$$

Then $Y_n - K_n$ is a martingale with respect to the natural filtration $(\mathcal{F}_t)_{t \geq 0}$, $\mathcal{F}_t := \sigma(Y_0, \dots, Y_t)$. Also note that by (216), $K_{n \wedge \tau} \leq -c(n \wedge \tau)$, for if $k < n \wedge \tau$, then $PY_k - Y_k \leq -c$. Now using the martingale condition,

$$x = \mathbb{E}_x[Y_0 - K_0] = \mathbb{E}_x[Y_{n \wedge \tau} - K_{n \wedge \tau}] \geq c\, \mathbb{E}_x[n \wedge \tau]. \tag{218}$$

Now if $n \geq Cx$, then $\{\tau \geq Cx\} = \{n \wedge \tau \geq Cx\}$. Hence, by Markov's inequality, by choosing $n \geq Cx$, we can conclude as

$$\mathbb{P}_x(\tau \geq Cx) = \mathbb{P}_x(n \wedge \tau \geq Cx) \leq \frac{\mathbb{E}_x[n \wedge \tau]}{Cx} \leq \frac{1}{cC}. \tag{219}$$

$\square$

We now prove Proposition 8.1.

Proof of Proposition 8.1. Part **(i)** follows immediately from Theorem 7.3 with $\mathcal{C}_e = \emptyset$. Such choice of the set $\mathcal{C}_e$ of the exceptional colors satisfies the stability condition (127) in the subcritical regime $p_0 > p^*$.

Next, we show **(ii)**. Suppose the maximum ball density p^* is achieved at positive colors $i_1 \leq i_2 \leq \cdots \leq i_r$. That is,

$$p_0 > p_{i_1} = \cdots = p_{i_r} > \max\{p_j \,:\, 1 \leq j \leq \kappa, \, , j \notin \{i_1, \dots, i_r\}\}. \tag{220}$$

Denote $\mathcal{C}^* := \{i_1, \dots, i_r\}$. Fix $\lambda \in (0, 1)$ and define a set

$$\mathcal{X}_{\lambda, M} := \left\{ \mathbf{x} = (x_1, \dots, x_\kappa) \in \mathbb{Z}_{\geq 0}^\kappa \,:\, \|\mathbf{x}\|_1 = M, \, x_{i_1} \geq \lambda M \right\}. \tag{221}$$

We will omit λ from the subscript of the above sets unless otherwise mentioned. By Proposition 8.6,

$$\mathbb{P}_\pi(\|W_0\|_1 = N) = \Theta\left(\binom{N + r - 1}{r - 1} \left(\frac{p^*}{p_0} \right)^N \right). \tag{222}$$

Noting that

$$\mathbb{P}_\pi(W_0 \in \mathcal{X}_N \mid \|W_0\|_1 = N) = \frac{\mathbb{P}_\pi(W_0 \in \mathcal{X}_N)}{\mathbb{P}_\pi(\|W_0\|_1 = N)} \tag{223}$$

$$\geq \frac{\mathbb{P}_\pi(\|W_0\|_1 = N - \lceil \lambda N \rceil)}{\mathbb{P}_\pi(\|W_0\|_1 = N)} \left(\frac{p^*}{p_0} \right)^{\lceil \lambda N \rceil}, \tag{224}$$

it follows that

$$\inf_{N \geq 1} \mathbb{P}_\pi(W_0 \in \mathcal{X}_N \mid \|W_0\|_1 = N) \geq c_* > 0 \tag{225}$$

for some constant $c_* = c_*(\lambda) > 0$.

For each $\mathbf{x} \in \mathbb{Z}_{\geq 0}^{\kappa}$, let $\mathbb{P}_{\mathbf{x}}$ denote the law of $(W_x)_{x \geq 0}$ with $W_0 = \mathbf{x}$. We claim that there exist constants $L_0, M_0 \geq 1$ and $\lambda, c_0 > 0$ such that

$$\inf_{M \geq M_0} \inf_{\mathbf{x} \in \mathcal{X}_M} \mathbb{P}_{\mathbf{x}} \Big(\tau_0 < \min(\tau_{M+L_0}, c_0 M^2) \Big) > 0. \tag{226}$$

Due to (225), this is enough to conclude (192). Indeed, since $W_0 \in \mathcal{X}_N$ implies $\|W_0\|_1 = N$, (225) implies

$$\mathbb{P}_{\pi}(\cdot \mid \|W_0\|_1 = N) \geq c_* \mathbb{P}_{\pi}(\cdot \mid W_0 \in \mathcal{X}_N) \geq c_* \inf_{\mathbf{x} \in \mathcal{X}_N} \mathbb{P}_{\pi}(\cdot). \tag{227}$$

Also note that, for any integer $L_0 \geq 1$,

$$\mathbb{P}_{\mathbf{x}} \Big(\tau_0 < \min(\tau_N, c_0 N^2 + L_0) \Big) \geq p_0^{L_0} \mathbb{P}_{\mathbf{y}} \Big(\tau_0 < \min(\tau_{N+L_0}, c_0 N^2) \Big), \tag{228}$$

where $\mathbf{y} \in \mathcal{X}_{N-L_0}$ is the carrier state obtained by inserting L_0 0's into the carrier with state $\mathbf{x}$. This yields

$$\inf_{N \geq N_0} \mathbb{P}_{\pi} \Big(\tau_0 < \min(\tau_N, c_0 N^2 + L_0) \mid \|W_0\|_1 = N \Big) \geq c_* \inf_{N \geq N_0} \inf_{\mathbf{x} \in \mathcal{X}_{N-L_0}} \mathbb{P}_{\mathbf{x}} \Big(\tau_0 < \min(\tau_{N+L_0}, c_0 N^2) \Big),$$

where the right-hand side is positive due to (226) by choosing $N_0 = M_0 + L_0$.

For the rest of the proof, we will show (226). Let $a := i_1$, $\rho := \inf\{x \geq 0 : W_x(i_1) = 0\}$ and $\tau_0 := \inf\{x \geq 0 : \|W_x\|_1 = 0\}$. According to Lemma 8.3, there exists a constant $c_0 > 0$ such that $\tau_0 \leq c_0 \|W_0\|_1^2$ with a positive probability. Denote $M := \|W_0\|_1$ and fix $\varepsilon, L > 0$. Define the following events:

$$A_1 := \{\tau_0 \leq c_0 M^2\}, \tag{229}$$

$$A_2 := \left\{ \|W_x^{\geq a}\|_1 \leq M + \frac{L}{2} - 2\varepsilon x \text{ for all } x \in [0, \rho] \right\}, \tag{230}$$

$$A_3 := \left\{ \|W_x^{<a}\|_1 \leq \frac{L}{2} + \varepsilon x \text{ for all } x \geq 0 \right\}, \tag{231}$$

$$A_4 := \{\|W_x\|_1 \leq M \text{ for all } x \in [\rho, \tau_0]\}. \tag{232}$$

Note that

$$\big\{ \|W_x\|_1 \text{ hits } 0 \text{ before it hits } M + L \text{ for some } x \leq c_0 M^2 \big\} \supseteq A := \bigcap_{i=1}^4 A_i. \tag{233}$$

Thus, it suffices to show that, for M_0, L sufficiently large and $\varepsilon > 0$ sufficiently small,

$$\inf_{M \geq M_0} \inf_{\mathbf{x} \in \mathcal{X}_M} \mathbb{P}_{\mathbf{x}}(A) > 0. \tag{234}$$

To this effect, first note that A_1 occurs with a uniformly positive probability by Lemma 8.3. Next, we observe that A_2 and A_3 occur with high probability. For A_2, according to Lemma 8.2, $\|W_x^{\geq a}\|_1 \leq S_x$ for all $x \in [0, \rho)$, where $(S_x)_{s \geq 0}$ is a biased random walk on $\mathbb{Z}$ with a negative drift $p_a - p_0 < 0$. Let ρ' denote the first time that $(S_x)_{s \geq 0}$ hits the origin. Then $\rho \leq \rho'$ by the coupling, so

$$\mathbb{P}_{\mathbf{x}}(A_2^c) \leq \mathbb{P}\left(S_x > S_0 + \frac{L}{2} - 2\varepsilon x \text{ for some } x \geq 0 \right). \tag{235}$$

The right-hand side above is the probability that a biased simple random walk on $\mathbb{Z}$ with mean increment $p_a - p_0 + 2\varepsilon$ starts at zero and ever reaches height $L/2$. We choose $\varepsilon > 0$ small so that $p_a - p_0 + 2\varepsilon < 0$.

Then by gambler's ruin for a negatively biased simple random walk on $\mathbb{Z}$, this probability is exponentially small in L. Thus, by choosing L large and $\varepsilon > 0$ small, we can make $\inf_{M \geq 1, \mathbf{x} \in \mathcal{X}_M} \mathbb{P}_\mathbf{x}(A_2)$ arbitrarily close to one.

For A_3, let X_x denote the decoupled carrier process with exceptional colors $\mathcal{C}_e = \{a\}$. Then by Proposition 6.2, $\|W_x^{<a}\|_1 \leq \|\hat{X}_x^{<a}\|_1 = \|X_x^{<a}\|_1$ for all $x \geq 0$. Note that $\|W_0^{<a}\|_1 \leq (1 - \lambda)M$ since $W_0 \in \mathcal{X}_M$. Moreover, note that $X_x^{<a}$ behaves exactly as the subcritical carrier process with ball colors in $\{1, \ldots, a\}$ and balls of color a acting as the empty box. That is, $X_x^{<a}$ evolves by the circular exclusion restricted on colors $\{1, \ldots, a\}$ while ignoring balls of colors in $\{a + 1, \ldots, \kappa, 0\}$. Thus, $X_x^{<a}$ is a lazy version of a carrier process with subcritical ball density as $\max(p_1, \ldots, p_{i_1-1}) < p_{i_1}$. Thus, by Lemma 8.4, $\limsup_{n \to \infty} \max_{0 \leq x \leq n} n^{-1}\|X_x^{<a}\|_1 \leq \varepsilon$ almost surely. Hence, A_3 occurs with high probability for any fixed $\varepsilon > 0$ if L is large enough.

Next, we show that $\bigcap_{i=1}^4 A_i$ occur with a uniformly positive probability. By definition, $\rho < \tau_0$. By the definition of the set $\mathcal{X}_M$ in (221), we get $W_0(a) \geq \lambda M$. Since $W_x(a)$ can decrease at most by one, it follows that $\rho \geq \lambda M$ almost surely. On $A_2 \cap A_3$, $\|W_\rho\|_1 \leq M(1 - \lambda\varepsilon) + L$. Thus,

$$A_1 \cap A_2 \cap A_3 \cap A_4^c \tag{236}$$

$$\subseteq \{\|W_x\|_1 \text{ makes an up-crossing from } M(1 - \lambda\varepsilon) + L \text{ to } M + L \text{ in } c_0 M^2 \text{ steps}\} \tag{237}$$

$$\subseteq \bigcup_{1 \leq i \leq \kappa} \{W_x(i) \text{ makes an up-crossing of length } M\lambda\varepsilon/\kappa \text{ in } c_0 M^2 \text{ steps}\}. \tag{238}$$

By the coupling (194) in Lemma 8.2, the last up-crossing probability is exponentially small in M. This shows

$$\mathbb{P}_\mathbf{x}\left(\bigcap_{i=1}^4 A_i\right) \geq \mathbb{P}_\mathbf{x}\left(\bigcap_{i=1}^3 A_i\right) - e^{-O(M)}. \tag{239}$$

Since A_1 has uniformly positive probability and $A_2 \cap A_3$ has a high probability, by union bound, the above is uniformly positive for M sufficiently large. This finishes the proof. $\qquad\square$

8.2. Order statistics of the excursion heights and multidimensional Gambler's ruin

According to Theorem 2.3 **(i)**, the carrier process $(W_x)_{x \geq 0}$ in the subcritical regime $p_0 > \max(p_1, \cdots, p_\kappa)$ will visit the origin $\mathbf{0} := (0, 0, \cdots, 0) \in (\mathbb{Z}_{\geq 0})^\kappa$ infinitely often with finite mean excursion time $\pi(\mathbf{0})^{-1}$. Namely, the number M_n of visits of W_x to $\mathbf{0}$ during $[1, n]$ (defined in (36)) satisfies

$$\frac{M_n}{n} \to \pi(\mathbf{0}) = \prod_{i=1}^\kappa \left(1 - \frac{p_i}{p_0}\right) \quad \text{a.s.} \quad \text{as } n \to \infty \tag{240}$$

by Theorem 2.3 **(i)** and the Markov chain ergodic theorem.

According to Lemma 3.1, 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, each excursion height is $O(1)$ with an exponential tail. Since there are $M_n \sim \pi(\mathbf{0})n$ i.i.d. excursions, their maximum height behaves as $O(\log n)$.

To make this estimate more precise, we analyze the order statistics of the excursion heights of the carrier process during $[1, n]$. For this, let $h_{1:m} \geq h_{2:m} \geq \cdots \geq h_{m:m}$ denote the order statistics of the first m excursion heights $h_1, \cdots, h_m$. The strong Markov property ensures that these excursion heights

are i.i.d., so we have

$$\mathbb{P}\{h_{j:m} \leq N\} = \sum_{\ell=0}^{j-1} \binom{m}{\ell} \mathbb{P}(h_1 \leq N)^{m-\ell} \, \mathbb{P}(h_1 > N)^\ell, \quad j = 1, \cdots, m. \tag{241}$$

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

$$\mathbb{P}(h_1 \leq N) = \left(1 - \frac{1-2p}{\theta^{N+1} - 1}\right) \mathbf{1}(N \geq 0), \tag{242}$$

where $\theta = p_0/p_1$ (see [LLP20, Sec. 4]). In order to obtain sharp asymptotics for top soliton lengths in the multicolor case, we need a similar result for a generalized gambler's ruin problem. That is, we need an asymptotic expression of the probability that the subcritical carrier process reaches 'height' N (see (37)) before coming back to the origin.

However, solving the 'carrier's ruin' problem asymptotically for $N \to \infty$ seems to be a nontrivial problem. 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 1. Assuming $p_0 > \max(p_1, p_2)$, the carrier process has a drift toward the origin in the interior and the right boundary of the state space $\mathbb{Z}_{\geq 0}^2$, but this is not necessarily true when there is no ball of color 1 (e.g., consider $\mathbf{p} = (0.4, 0.3, 0.3)$). 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 for this reason. Another standard approach is the one-step analysis, which is computationally challenging since it involves inverting a large matrix (with blocks of expanding sizes) at every N, and one needs to obtain an asymptotic expression of the solution of a $N^\kappa \times N^\kappa$ linear equation as $N \to \infty$.

Despite the technical difficulties we mentioned above, as stated in Theorem 2.3 (ii), we are able to obtain exact asymptotic expression on the probability that an excursion reaches height N as $N \to \infty$. Our analysis uses a novel idea of 'stationary balancing', which we believe to be useful for solving other multidimensional ruin problems. A major technical component we will use in the proof is Proposition 8.1(ii).

The following combinatorial observation will be used in the proof of Theorem 2.3 (ii) below. It states that if we have k independent geometric random variables of parameters $p_1/p_0, \ldots, p_\kappa/p_0$, and if we condition on their sum being N, then the total mass should be concentrated on the most probable colors. We note that in the statement, the $1 - \frac{p_i}{p_0}$ terms are omitted from the product since they are all between $1 - \frac{p^*}{p_0}$ and 1. The proof is given at the end of this section.

Proposition 8.6. *Let $p_0 > p^* = \max(p_1, \ldots, p_\kappa)$. Let r denote the number of i's in $\{1, \ldots, \kappa\}$ such that $p_i = p^*$. If $p_1 = \cdots = p_\kappa$, then*

$$\sum_{x_1 + \cdots + x_\kappa = N} \prod_{i=1}^\kappa \left(\frac{p_i}{p_0}\right)^{x_i} = \left(\frac{p^*}{p_0}\right)^N \binom{N + \kappa - 1}{\kappa - 1}. \tag{243}$$

Suppose $r < \kappa$ and let $p^{(2)}$ denote the second largest value among $p_1, \ldots, p_\kappa$. Then

$$\left(\frac{p^*}{p_0}\right)^N \binom{N + r - 1}{r - 1} \leq \sum_{x_1 + \cdots + x_\kappa = N} \prod_{i=1}^\kappa \left(\frac{p_i}{p_0}\right)^{x_i} \leq \left(\frac{p^*}{p_0}\right)^N \binom{N + r - 1}{r - 1} \left(\frac{p^*}{p^* - p^{(2)}}\right)^{\kappa - r}. \tag{244}$$

We are now ready to prove Theorem 2.3 (ii).

Proof of Theorem 2.3 (ii). Fix two disjoint subsets $A, B \subseteq \mathbb{Z}_{\geq 0}^\kappa$. Let τ_i for $i \geq 1$ denote the ith time that the Markov chain $(W_x)_{x \geq 0}$ hits the union $A \cup B$. Then by strong Markov property, the subsequential process $\widetilde{W}_i := W_{\tau_i}$ for $i \geq 1$ is a Markov chain on the state space $A \cup B$. Since $(W_x)_{x \geq 0}$ is irreducible and

aperiodic, so is the restricted chain $(\widetilde{W}_i)_{i \geq 1}$. So if the restricted chain has a stationary distribution, it has to be unique. Note that the following probability distribution $\pi_{A \cup B}$ on $A \cup B$ is a stationary distribution for $(\widetilde{W}_i)_{i \geq 1}$:

$$\pi_{A \cup B}(\mathbf{x}) = \pi(\mathbf{x}) / \pi(A \cup B) \qquad \text{for } \mathbf{x} \in A \cup B, \tag{245}$$

where for each subset $R \subseteq \mathbb{Z}_{\geq 0}^{\kappa}$, we denote $\pi(R) := \sum_{\mathbf{y} \in R} \pi(\mathbf{y})$. Here, π is the stationary distribution for the subcritical carrier process defined in (9). This can be justified by using the Markov chain ergodic theorem (see, for example, [AF02, Sec. 2.7.1]).

Let $(W'_x)_{x \geq 0}$ be a carrier process on the ball configuration $\xi^{\mathbf{p}}$ but initialized as $W'_0 \sim \pi_{A \cup B}$. If we restrict this chain at hitting times of $A \cup B$, then the restricted chain is stationary with distribution $\pi_{A \cup B}$. That is, if we denote the ith time that W'_t visits $A \cup B$ as τ'_i, then W'_0 and $W'_{\tau'_1}$ has the same distribution $\pi_{A \cup B}$. The key idea is to treat the restricted stationary process $(W'_{\tau'_i})_{i \geq 0}$ as if it is a two-state process on $\{A, B\}$ and then derive a 'balance equation' for the mass transport between A and B.

By using (245),

$$\mathbb{P}\left(W'_x \text{ visits } B \text{ before } A\right) = \mathbb{P}\left(W'_{\tau'_1} \in B\right) = \pi_{A \cup B}(B) = \frac{\pi(B)}{\pi(A \cup B)}. \tag{246}$$

This gives

$$\pi_{A \cup B}(B) = \mathbb{P}\left(W'_x \text{ visits } B \text{ before } A\right) \tag{247}$$

$$= \mathbb{P}\left(W'_x \text{ visits } B \text{ before } A, W'_0 \in A\right) + \mathbb{P}\left(W'_t \text{ visits } B \text{ before } A, W'_0 \in B\right) \tag{248}$$

$$= \mathbb{P}\left(W'_x \text{ visits } B \text{ before } A \,\middle|\, W'_0 \in A\right) \pi_{A \cup B}(A) \tag{249}$$

$$\qquad + \mathbb{P}\left(W'_x \text{ visits } B \text{ before } A \,\middle|\, W'_0 \in B\right) \pi_{A \cup B}(B). \tag{250}$$

Simplifying using (245), we obtain the following 'balance equation':

$$\mathbb{P}\left((W'_x)_{x \geq 1} \text{ visits } B \text{ before } A \,\middle|\, W'_0 \in A\right) \tag{251}$$

$$\qquad = \mathbb{P}\left((W'_x)_{x \geq 1} \text{ visits } A \text{ before } B \,\middle|\, W'_0 \in B\right) \frac{\pi(B)}{\pi(A)}. \tag{252}$$

Now we specialize in the above result. Take $A = \{\mathbf{0}\}$ and $B = \{\mathbf{x} \in \mathbb{Z}_{\geq 0}^{\kappa} : \|\mathbf{x}\|_1\|_1 = N\}$. Note that

$$\mathbb{P}\left((W'_x)_{x \geq 1} \text{ visits } B \text{ before } \mathbf{0} \,\middle|\, W'_0 = \mathbf{0}\right) = \mathbb{P}(h_1 \geq N). \tag{253}$$

Recalling the the formula for π in (9), it follows that

$$\mathbb{P}(h_1 \geq N) = \mathbb{P}\left(W'_x \text{ visits } \mathbf{0} \text{ before } B \,\middle|\, W'_0 \in B\right) \sum_{x_1 + \cdots + x_{\kappa} = N} \prod_{i=1}^{\kappa} \left(\frac{p_i}{p_0}\right)^{x_i}, \tag{254}$$

where the sum is over all integers $x_1, \ldots, x_{\kappa} \geq 0$ that sum to N. The above along with Proposition 8.6 is enough to deduce the upper bound in (12).

To obtain a lower bound of matching order, we need to show that the probability in the right-hand side of (254) is uniformly positive for all sufficiently large N. This requires a substantial analysis, which we have done in proving Proposition 8.1. By this result, there exists a constant $\delta > 0$ such that

$$\liminf_{N \geq 1} \mathbb{P}\left((W'_x)_{x \geq 1} \text{ visits } \mathbf{0} \text{ before } B \,\middle|\, W'_0 \in B\right) > \delta > 0. \tag{255}$$

Then the assertion follows from (254), (255) and Proposition 8.6. □

Proof of Proposition 8.6. Suppose we have real numbers $a_1 = a_2 = \cdots = a_r \geq a_{r+1} \geq \cdots \geq a_\kappa > 0$. Note that

$$\sum_{x_1+\cdots+x_\kappa=N} a_1^{x_1} \cdots a_\kappa^{x_\kappa} = a_1^N \sum_{x_1+\cdots+x_\kappa=N} \left(\frac{a_{r+1}}{a_1}\right)^{x_{r+1}} \cdots \left(\frac{a_\kappa}{a_1}\right)^{x_\kappa}$$

$$= a_1^N \sum_{q=0}^N \binom{q+r-1}{r-1} \sum_{x_{r+1}+\cdots+x_\kappa=N-q} \left(\frac{a_{r+1}}{a_1}\right)^{x_{r+1}} \cdots \left(\frac{a_\kappa}{a_1}\right)^{x_\kappa}. \tag{256}$$

If $a_1 = \cdots = a_\kappa$, then the above expression equals to $a_1^N \binom{N+\kappa-1}{\kappa-1}$. Hence, if $p_1 = \cdots = p_\kappa$, we get (243).

We now assume $a_1 = \cdots = a_r > a_{r+1} \geq \cdots \geq a_\kappa$ for some $r \in \{1, \ldots, \kappa\}$. Then the last expression in (256) is at most

$$a_1^N \sum_{q=0}^N \binom{q+r-1}{r-1} \sum_{x_{r+1}+\cdots+x_\kappa=N-q} \left(\frac{a_{r+1}}{a_1}\right)^{N-q} \tag{257}$$

$$= a_1^N \sum_{q=0}^N \binom{q+r-1}{r-1}\binom{N-q+\kappa-r-1}{\kappa-r-1} \left(\frac{a_{r+1}}{a_1}\right)^{N-q} \tag{258}$$

$$\leq a_1^N \binom{N+r-1}{r-1}\left[\sum_{n\geq 0} \binom{n}{\kappa-r-1} \left(\frac{a_{r+1}}{a_1}\right)^{n-(\kappa-r-1)}\right]. \tag{259}$$

Note that the sum in the bracket above equals

$$\left(\frac{a_1}{a_{r+1}}\right)^{(\kappa-r-1)} \sum_{n\geq 0} \binom{n}{\kappa-r-1} \left(\frac{a_{r+1}}{a_1}\right)^n = \left(\frac{a_1}{a_{r+1}}\right)^{(\kappa-r-1)} \frac{\left(\frac{a_{r+1}}{a_1}\right)^{(\kappa-r-1)}}{\left(1-\frac{a_{r+1}}{a_1}\right)^{\kappa-r}} = \left(\frac{a_1}{a_1-a_{r+1}}\right)^{\kappa-r}, \tag{260}$$

where we used the generating function $\sum_{n\geq 0} \binom{n}{k} y^n = \frac{y^k}{(1-y)^{k+1}}$ (with $\binom{n}{k} = 0$ for $n < k$). Hence, it follows that

$$\sum_{x_1+\cdots+x_\kappa=N} \prod_{i=1}^\kappa \left(\frac{p_i}{p_0}\right)^{x_i} \leq \left(\frac{p^*}{p_0}\right)^N \binom{N+r-1}{r-1}\left(\frac{p^*}{p^*-p^{(2)}}\right)^{\kappa-r}. \tag{261}$$

For the lower bound, note that the last expression in (256) is at least

$$a_1^N \sum_{q=0}^N \binom{q+r-1}{r-1} \sum_{x_{r+1}+\cdots+x_\kappa=N-q} \left(\frac{a_\kappa}{a_1}\right)^{N-q} \tag{262}$$

$$= a_1^N \sum_{q=0}^N \binom{q+r-1}{r-1}\binom{N-q+\kappa-r-1}{\kappa-r-1} \left(\frac{a_\kappa}{a_1}\right)^{N-q} \geq a_1^N \binom{N+r-1}{r-1}. \tag{263}$$

Hence, we get

$$\sum_{x_1+\cdots+x_\kappa=N} \prod_{i=1}^\kappa \left(\frac{p_i}{p_0}\right)^{x_i} \geq \left(\frac{p^*}{p_0}\right)^N \binom{N+r-1}{r-1}. \tag{264}$$

This shows the assertion. □

8.3. Proof of Theorem 2.4

Now that we have the asymptotic soliton to the 'carrier's ruin' problem (Theorem 2.3(**ii**)), we are ready to obtain sharp scaling limit for the top soliton lengths in the subcritical regime, as stated in Theorem 2.4. To do so, we first obtain the following scaling limit of $\mathbf{h}_j(n)$ using a similar argument to that developed in [LLP20]. For instance, the maximum excursion height $\mathbf{h}_1(n)$ of the subcritical carrier process during $[0, n]$ scales like $(1 + o(1)) \log n$, where its tail follows the Gumbel distribution up to a constant shift. The tail cannot have a tight scaling limit due to a rounding error even in the $\kappa = 1$ case; see [LLP20, Remark 5.5].

Proposition 8.7. *Suppose* $p_0 > p^* := \max(p_1, \cdots, p_\kappa)$. *Let* r *denote the multiplicity of* p^*, $\theta :=$ p_0/p^*, *and* $\sigma := \pi(\mathbf{0}) > 0$ *(see (240)). Let* $v_n := (1 + \delta_n) \log_\theta(\sigma n/(r-1)!)$, *where we set* $\delta_n :=$ $\frac{(r-1)\log\log_\theta(\sigma n/(r-1)!)+\log(r-1)!}{\log \sigma n/(r-1)!}$. *Fix* $j \geq 1$ *and* $x \in \mathbb{R}$. *Then*

$$\limsup_{n\to\infty} \left[\exp\left(-\frac{C}{(r-1)!}\theta^{-(x-1)} \right) \left(\sum_{\ell=0}^{j-1} \frac{\theta^{-\ell x}}{\ell!(r-1)!} \right) \right]^{-1} \left(\mathbb{P}\{\mathbf{h}_j(n) \leq x + v_n\} + o(1) \right) \leq 1, \quad (265)$$

$$\liminf_{n\to\infty} \left[\exp\left(-\frac{\delta}{(r-1)!}\theta^{-(x-1)} \right) \left(\sum_{\ell=0}^{j-1} \frac{\theta^{-\ell(x-1)}}{\ell!(r-1)!} \right) \right]^{-1} \left(\mathbb{P}\{\mathbf{h}_j(n) \leq x + v_n\} + o(1) \right) \geq 1, \quad (266)$$

where constants $\delta > 0$ *and* $C \geq 1$ *are as in the Theorem 2.3 (ii).*

Proof. Fix $\varepsilon \in (0, \sigma)$ and let $b_n = \lfloor(\sigma - \varepsilon)n\rfloor$. As $M_n/n \to \sigma$ a.s. (see (240)), we have that $M_n \geq b_n$ for all sufficiently large n almost surely. Hence, for each fixed $x \in \mathbb{R}$,

$$\mathbb{P}(\mathbf{h}_j(n) \leq x + v_n) \leq \mathbb{P}(h_{j:b_n} \leq x + v_n) + o(1). \quad (267)$$

Furthermore, according to Theorem 2.3 (**ii**) and (241),

$$\mathbb{P}(h_{j:b_n} \leq x + v_n) \quad (268)$$

$$\leq \sum_{\ell=0}^{j-1} \binom{b_n}{\ell} \left(1 - \frac{C\binom{\lfloor x+v_n\rfloor+r-1}{r-1}}{\theta^{\lfloor x+v_n\rfloor}} \right)^{b_n-\ell} \left(\frac{\binom{\lfloor x+v_n\rfloor+r-1}{r-1}}{\theta^{\lfloor x+v_n\rfloor}} \right)^\ell \quad (269)$$

$$= \left(1 - \frac{C\binom{\lfloor x+v_n\rfloor+r-1}{r-1}}{\theta^{\lfloor x+v_n\rfloor+1}} \right)^{b_n} \sum_{\ell=0}^{j-1} b_n^{-\ell} \binom{b_n}{\ell} \left(1 - \frac{C\binom{\lfloor x+v_n\rfloor+r-1}{r-1}}{\theta^{\lfloor x+v_n\rfloor}} \right)^{-\ell} \left(\frac{\binom{\lfloor x+v_n\rfloor+r-1}{r-1} b_n}{\theta^{\lfloor x+v_n\rfloor}} \right)^\ell. \quad (270)$$

Since $v_n = (1 + \delta_n) \log_\theta(\sigma n/(r-1)!)$, note that

$$\log\left(\frac{\sigma v_n^{r-1} n}{\theta^{v_n}} \right) = (r-1)\log(1+\delta_n) + (r-1)\log\log_\theta(\sigma n/(r-1)!) + \log(r-1)! \quad (271)$$

$$+ \delta_n(\log(r-1)! - \log\sigma n) \quad (272)$$

$$= (r-1)\log(1+\delta_n) \to 0 \qquad \text{as } n \to \infty, \quad (273)$$

where the second equality uses the definition of δ_n and the limit follows since $\delta_n = o(1)$. Using Stirling's approximation, $\binom{N+r-1}{r-1} = (1 + o(1))N^{r-1}/(r-1)!$ as $N \to \infty$, we get

$$\lim_{n\to\infty} \frac{\theta^{\lfloor x+v_n\rfloor}}{\theta^{v_n}} \log\left(1 - \frac{C\binom{\lfloor x+v_n\rfloor+r-1}{r-1}}{\theta^{\lfloor x+v_n\rfloor}} \right)^{b_n} = -\lim_{n\to\infty} \frac{b_n C\binom{\lfloor x+v_n\rfloor+r-1}{r-1}}{\theta^{v_n}} \quad (274)$$

$$= -\lim_{n\to\infty} \frac{1}{(r-1)!}\left(\frac{\lfloor x+v_n\rfloor}{v_n} \right)^{r-1} \frac{b_n C v_n^{r-1}}{\theta^{v_n}} \quad (275)$$

$$= -\frac{C(1-\frac{\varepsilon}{\sigma})}{(r-1)!}. \quad (276)$$

Similarly,

$$\lim_{n\to\infty} \frac{\theta^{\lfloor x+\nu_n\rfloor}}{\theta^{\nu_n}} \frac{\binom{\lfloor x+\nu_n\rfloor+r-1}{r-1} b_n}{\theta^{\lfloor x+\nu_n\rfloor}} = \frac{(1-\frac{\varepsilon}{\sigma})}{(r-1)!}. \tag{277}$$

Writing $\eta_n = (x+\nu_n) - \lfloor x+\nu_n\rfloor \in [0,1)$, since $\theta > 1$,

$$\theta^{x-1} \leq \frac{\theta^{\lfloor x+\nu_n\rfloor}}{\theta^{\nu_n}} = \theta^x \theta^{-\eta_n} \leq \theta^x. \tag{278}$$

Also note that $\lim_{n\to\infty} b_n^{-\ell} \binom{b_n}{\ell} = \frac{1}{\ell!}$. From the above computations, we deduce

$$\limsup_{n\to\infty} \exp\left(\frac{C}{(r-1)!}\left(1-\frac{\varepsilon}{\sigma}\right)\theta^{-x+1}\right)\left(1 - \frac{C\binom{\lfloor x+\nu_n\rfloor+r-1}{r-1}}{\theta^{\lfloor x+\nu_n\rfloor}}\right)^{b_n} \leq 1, \tag{279}$$

$$\limsup_{n\to\infty} \frac{(r-1)!}{1-\frac{\varepsilon}{\sigma}}\theta^{x-1}\frac{\binom{\lfloor x+\nu_n\rfloor+r-1}{r-1} b_n}{\theta^{\lfloor x+\nu_n\rfloor}} \leq 1. \tag{280}$$

Then we obtain

$$\limsup_{n\to\infty}\left[\exp\left(-\frac{C}{(r-1)!}\left(1-\frac{\varepsilon}{\sigma}\right)\theta^{-(x-1)}\right)\left(\sum_{\ell=0}^{j-1}(1+\varepsilon)\left(1-\frac{\varepsilon}{\sigma}\right)^\ell \frac{\theta^{-\ell(x-1)}}{\ell!(r-1)!}\right)\right]^{-1} \tag{281}$$

$$\times \mathbb{P}\big(h_{j:b_n} \leq x+\nu_n\big) \tag{282}$$

$$\leq 1. \tag{283}$$

Therefore, letting $\varepsilon \searrow 0$ and using (267) give the limsup in the statement. A similar argument using $b_n = \lceil(\sigma+\varepsilon)n\rceil$ shows the liminf in the statement. □

Now we are ready to establish sharp scaling for the top soliton lengths in the subcritical regime.

Proof of Theorem 2.4. Let ν_n be as in Theorem 2.4. Note that

$$\nu_n = \log_\theta n + (r-1)\log_\theta\log n + c + o(1) \tag{284}$$

for some constant c. Hence, the asymptotic (13) for $\lambda_j(n)$ follows from (14).

Now we derive (14). Fix $j \geq 1$ and $x \in \mathbb{R}$. Then by Proposition 8.7,

$$\liminf_{n\to\infty} \mathbb{P}(\mathbf{h}_1(n) \leq x+\nu_n) \geq \exp(-\delta\theta^{-x}), \tag{285}$$

$$\limsup_{n\to\infty} \mathbb{P}(\mathbf{h}_j(n) \leq x+\nu_n) \leq \exp\left(-\frac{C}{(r-1)!}\theta^{-(x-1)}\right)\sum_{k=0}^{j-1}\frac{\theta^{-k(x-1)}}{k!(r-1)!}. \tag{286}$$

Moreover, recall the quantities M_n and r_n in (36) and (38), respectively. By Lemma 3.1,

$$\mathbf{h}_1(n) = \max\{h_1, \cdots, h_{M_n}\} \leq \lambda_1(n) \leq \max\{h_1, \cdots, h_{M_n+1}\}. \tag{287}$$

Also, note that

$$0 \leq \mathbb{P}(\mathbf{h}_1(n) \leq x+\nu_n) - \mathbb{P}(\max\{h_1, \cdots, h_{M_n+1}\} \leq x+\nu_n) \leq \mathbb{P}(h_{M_n+1} > \mathbf{h}_1(n)) = o(1). \tag{288}$$

It follows that

$$\mathbb{P}(\lambda_1(n) \leq x+\nu_n) = \mathbb{P}(\mathbf{h}_1(n) \leq x+\nu_n) + o(1). \tag{289}$$

Moreover, since $\lambda_1(n) \geq \lambda_j(n) \geq \mathbf{h}_j(n)$ by Lemma 3.1,

$$\mathbb{P}(\lambda_1(n) \leq x + \nu_n) \leq \mathbb{P}(\lambda_j(n) \leq x + \nu_n) \leq \mathbb{P}(\mathbf{h}_j(n) \leq x + \nu_n). \tag{290}$$

Then (285)–(286) show (14), as desired. □

9. The linear scaling limit of the carrier process

In this section, we prove Theorem 2.5 **(i)**, concerning the linear scaling limit of the carrier process W_x in (8).

Throughout this section, we assume $p_0 \leq p^* = \max(p_1, \ldots, p_\kappa)$. In this case, the set $\mathcal{C}_s^{\mathbf{p}}$ of unstable colors (defined above the statement of Theorem 2.5) is nonempty. Let $\alpha_1 < \cdots < \alpha_r$ denote the unstable colors. Let $(X_x)_{x \geq 0}$ be the decoupled carrier process in (128) with $\mathcal{C}_e = \mathcal{C}_u^{\mathbf{p}}$. Recall the process $(\hat{X}_x)_{x \geq 0}$ in (115).

We first show that the coordinate $X_x(\ell)$ for ℓ an unstable color of supercritical density behaves like a random walk with a positive drift.

Proposition 9.1. *Fix $j \in \{1, \ldots, r\}$ and denote $\ell := \alpha_j$, $\ell^+ := \alpha_{j+1}$. If $p_\ell > p_{\ell^+}$, then $\bar{M} := -\inf_{k \in \mathbb{N}} X_k(\ell)$ has a finite exponential moment.*

Proof. Recall that $(X_x(\ell))_{x \geq 0}$ is a Markov additive functional with increments $g^\ell(X_x^s, \xi_{x+1})$ (see (133)). Under the hypothesis, it has a positive bias $\mathbb{E}_{\pi^s \otimes \mathbf{p}}[g^\ell(X_x^s, \xi_{x+1})] = \alpha := p_\ell - p_{\ell^+} > 0$ (see Proposition 7.2). Hence, one can expect that $X_x(\ell)$ will essentially behave as a simple random walk on $\mathbb{Z}$ with a positive bias. Since $\bar{M}$ measures the height of the excursion of $X_x(\ell)$ below the x-axis, it should have a finite exponential moment. Below we give a rigorous justification.

Consider the Markov chain

$$Y_x := \left(X_x(\ell + 1), \cdots, X_x(\ell^+ - 1)\right). \tag{291}$$

Let $\tau_j := j$ for $j \geq 0$ if $\ell + 1 = \ell^+$; Otherwise, let τ_j be the jth return time of $\left(X_x(\ell+1), \cdots, X_x(\ell^+ - 1)\right)$ to the origin. By strong Markov property, $\tau_1, \tau_2 - \tau_1, \tau_3 - \tau_2, \ldots$ are i.i.d., and they have finite moments of all orders by Lemma 7.5. Let $R_j := X_{\tau_j}(\ell)$ for $j \geq 1$. Then $(R_j)_{j \geq 1}$ is a random walk. Let $\eta_i := R_i - R_{i-1}$ denote the increments. It has a positive drift as

$$\mathbb{E}[\eta_1] = \lim_{j \to \infty} \frac{R_j}{j} = \lim_{j \to \infty} \frac{X_{\tau_j}(\ell)}{\tau_j} \frac{\tau_j}{j} = \alpha \, \mathbb{E}[\tau_1] > 0, \tag{292}$$

where the first two equalities use the strong law of large numbers and the Markov chain ergodic theorem.

Next, we claim that $X_x(\ell)$ returns to the origin only finitely many times almost surely. First note that by the strong law of large numbers, $n^{-1} R_n \to \alpha > 0$ almost surely. Hence, $n^{-1} R_n > \alpha/2$ infinitely often almost surely. Note that for each $j \geq 1$, since $\tau_{j+1} - \tau_j$ is independent from R_j and has the same distribution as τ_1, by Chebyshev's inequality,

$$\mathbb{P}(X_x(\ell) = 0 \text{ for some } x \in [\tau_j, \tau_{j+1})) \leq \mathbb{P}(R_j \leq \tau_{j+1} - \tau_j) \tag{293}$$

$$\leq \mathbb{E}[\mathbb{P}(\tau_{j+1} - \tau_j \geq R_j \mid R_j)] \tag{294}$$

$$\leq \mathbb{E}[R_j^{-2} \, \mathbb{E}[\tau_1^2]] \tag{295}$$

$$\leq \mathbb{E}[\tau_1^2]\left((\alpha j/2)^{-2} + c\mathbb{P}(R_j \leq (\alpha/2)j)\right), \tag{296}$$

where the last inequality follows by partitioning on two cases depending on $R_j \leq (\alpha/2)j$ or $R_j > (\alpha/2)j$. If we denote $\bar{\eta}_i := \mathbb{E}[\eta_i] - \eta_i$, then $\bar{\eta}_i$'s are mean zero i.i.d., so, noting that $\mathbb{E}[R_n] = \alpha n$,

$$\mathbb{P}(R_n \le (\alpha/2)n) \le \mathbb{P}(\mathbb{E}[R_n] - R_n \ge (\alpha/2)n) \tag{297}$$

$$\le \mathbb{P}\left(\sum_{i=1}^{n} -\overline{\eta}_i \ge (\alpha/2)n\right) \tag{298}$$

$$\le \mathbb{P}\left(\left(\sum_{i=1}^{n} -\overline{\eta}_i\right)^4 \ge (\alpha/2)^4 n^4\right) \tag{299}$$

$$\le \frac{C(\mathbb{E}[\overline{\eta}_1^2]^2 + \mathbb{E}[\overline{\eta}_1^4])}{n^2} \tag{300}$$

for some constant $C > 0$. Note that for the last inequality, we have used Chebyshev's inequality along with the fact that only the $O(n^2)$ terms of the form $\overline{\eta}_i^2 \overline{\eta}_j^2$ for $i \ne j$ and $\overline{\eta}_i^4$ have nonzero expectations. Sincet $|\eta_1| \le \tau_1$ has a finite moments of all orders, so does $\overline{\eta}_1$. Thus, (293) implies

$$\sum_{j \ge 1} \mathbb{P}(Z_x(0) = 0 \text{ for some } x \in [\tau_j, \tau_{j+1})) < \infty. \tag{301}$$

By the Borel-Cantelli lemma, it follows that $X_x(\ell)$ visits the origin only finitely many times almost surely. This shows the claim.

Now we conclude that $\overline{M}$ has a finite exponential moment. For this, we use the general result by Hansen [Han06] about the running maximum of a random walk with negative drift – that if the running maximum is uniformly bounded almost surely, then the supremum of the running maximum has a finite exponential moment. We apply this result to the random walk $(-R_j)_{j \ge 1}$. According to the claim, it follows that $\sup_{x \ge 0} -X_x(\ell) = -\inf_{x \ge 0} X_x(\ell)$ is almost surely finite. Hence, $\sup_{j \ge 1} -R_j$ is almost surely finite, so by [Han06, Thm. 2.1], $\sup_{j \ge 1} -R_j = -\inf_{j \ge 1} R_j$ has a finite exponential moment. Since the increments of R_j have finite exponential moments, we can conclude that $-\inf_{x \ge 0} X_x(\ell)$ also has a finite exponential moment. □

Proposition 9.2. *Let $j \in \{1, \dots, r\}$ be arbitrary with $\ell := \alpha_j$, $\ell^+ := \alpha_{j+1}$, and $p_\ell > p_{\ell^+}$. Then for each integer $d \ge 1$, there exists a constant $c > 0$ such that for all $n \ge 1$ and $s > 0$,*

$$\mathbb{P}\left(\max_{0 \le t \le n} X_t(\ell) - X_n(\ell) \ge s\right) \le \exp(-cs), \tag{302}$$

$$\mathbb{P}\left(\left|\max_{0 \le t \le n} \hat{X}_t(\ell) - X_n(\ell)\right| \ge \varepsilon\right) \le \exp(-cs). \tag{}$$

Proof. Consider the following Markov chain

$$Y_x := \left(\max_{1 \le s \le x} X_s(\ell) - X_x(\ell), X_x(\ell+1), \cdots, X_x(\ell^+ - 1)\right) \tag{303}$$

on $\mathbb{Z}_{\ge 0}^{\ell^+ - \ell}$. Note that $Y_0 = \mathbf{0}$. Let τ denote the first return time of Y_x to the origin. In Theorem 7.3, we have shown that τ has finite moments of all orders. Let $L_1, L_2, \dots$ denote the lengths of excursions of Y_x to the origin. Since $L_k \ge 1$ for all $k \ge 1$, $M_n \le n$. Let $h_1, h_2, \cdots$ denote its subsequent excursion heights of Y_x. Since $h_1 \le L_1 = \tau$, and using the elementary inequality,

$$1 - (1 - a)^n \le na \quad \text{for } a \in (0, 1), \tag{304}$$

for each $s > 0$,

$$\mathbb{P}\left(\left|\max_{1 \le t \le n} X_t(\ell) - X_n(\ell)\right| \ge s\right) \le \mathbb{P}(\|Y_n\|_1 \ge s) \tag{305}$$

$$\le \mathbb{P}(\max(h_1, \dots, h_{M_n}) \ge s) \tag{306}$$

$$\leq \mathbb{P}(\max(h_1, \ldots, h_n) \geq s) \tag{307}$$

$$\leq 1 - (1 - \mathbb{P}(h_1 \geq s))^n \tag{308}$$

$$\leq n\,\mathbb{P}(h_1 \geq s) \tag{309}$$

$$\leq n\,\mathbb{P}(\tau \geq s). \tag{310}$$

Note that $\mathbb{P}(\tau \geq s)$ is exponentially small in s. Hence, the first inequality in (302) follows.

Next, we show the second inequality in (302). By definition of $\hat{X}_x$, we have

$$\mathbb{P}\left(\left|\max_{1 \leq s \leq x} \hat{X}_s(\ell) - X_x(\ell)\right| \geq s\right) = \mathbb{P}\left(\left|\max_{0 \leq s \leq x}\left(X_s(\ell) - \min_{0 \leq t \leq s} X_t(\ell)\right) - X_x(\ell)\right| \geq s\right) \tag{311}$$

$$\leq \mathbb{P}\left(\max_{0 \leq s \leq x} X_s(\ell) - X_x(\ell) \geq s/2\right) \tag{312}$$

$$+ \mathbb{P}\left(\max_{0 \leq s \leq x}\left(- \min_{0 \leq t \leq s} X_t(\ell)\right) \geq s/2\right). \tag{313}$$

The second term in the last expression is exponentially small in s due to Proposition 9.1. Hence, the second inequality in (302) follows from the above and the first equality in (302). □

The following lemma shows half of Theorem 2.5 **(i)**.

Lemma 9.3. *Let* $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_\kappa) := \sum_{j=1}^r \mathbf{e}_{\alpha_j}(p_{\alpha_j} - p_{\alpha_{j+1}})$. *For* $i = 1, \ldots, \kappa$, *almost surely,*

$$\limsup_{n \to \infty} n^{-1}\left(\max_{0 \leq t \leq n} W_t(i)\right) \leq \mu_i. \tag{314}$$

Proof. By Proposition 6.2,

$$\max_{0 \leq t \leq x} W_x(i) \leq \max_{0 \leq t \leq x} \hat{X}_x(i) \quad i = 1, \ldots, \kappa, \tag{315}$$

where $\hat{X}_x(i) = X_x(i) - \min_{0 \leq s \leq x} X_s(i)$. Let $\tau_0 := 0$, let τ_j for $j \geq 1$ denote the jth return time of X_x^s to the origin, and let h_j denote the maximum value of $\|X_s^s\|_1$ during the interval $[\tau_{j-1}, \tau_j]$. By the strong Markov property, h_j's are i.i.d. By Lemma 7.1, $(X_x^s)_{x \geq 0}$ is a Markov chain on $\mathbb{Z}_{\geq 0}^\kappa$ with a unique stationary distribution, and its return time to the origin, say τ, has finite moments of all order by Theorem 7.3.

Now note that, for each $s > 0$,

$$\mathbb{P}\left(\max_{0 \leq t \leq n} \|X_t^s\|_1 \geq s\right) \leq \mathbb{P}(\max(h_1, \ldots, h_n) \geq s) = 1 - (1 - \mathbb{P}(h_1 \geq s))^n \tag{316}$$

$$\leq n\,\mathbb{P}(h_1 \geq s) \leq n\,\mathbb{P}(\tau \geq s). \tag{317}$$

Now choosing $s = n^{1/4}$, it follows that $\mathbb{P}\left(n^{-1/2} \max_{0 \leq t \leq n} \|X_t^s\|_1 \geq n^{-1/4}\right)$ is summable, so by the Borel-Cantelli lemma,

$$\lim_{n \to \infty} n^{-1/2} \max_{0 \leq t \leq n} \|X_t^s\|_1 = 0 \quad \text{a.s.} \tag{318}$$

Combining with (315) and recalling $\hat{X}_x(i) = X_x(i)$ for $i \in \mathcal{C}_s^{\mathbf{p}}$, we deduce (314) for all $i \in \mathcal{C}_s^{\mathbf{p}}$.

By the argument in the previous paragraph, we may assume the set $\mathcal{C}_u^{\mathbf{p}}$ of unstable colors is nonempty, and it remains to show the statement for unstable colors. Fix $j \in \{1, \ldots, r\}$ and let $\ell = \alpha_j$, $\ell^+ = \alpha_{j+1}$ (with $\alpha_{r+1} = 0$). Since ℓ is an unstable color, $p_\ell \geq p_{\ell^+}$. First, suppose $p_\ell > p_{\ell^+}$. Then Propositions 9.2

and 7.9 imply

$$\lim_{n\to\infty} n^{-1} \max_{1\le t \le n} \hat{X}_t(\ell) = \lim_{n\to\infty} n^{-1} X_n(\ell) = p_\ell - p_{\ell^+} \tag{319}$$

almost surely. Then the assertion follows from (315).

It remains to consider the case $p_\ell = p_{\ell^+}$. In this case, we wish to show

$$\limsup_{n\to\infty} n^{-1}\left(\max_{0\le t \le n} W_t(\ell)\right) = 0. \tag{320}$$

Rewrite (315) as

$$\max_{0\le t \le x} W_x(\ell) \le \max_{0\le t \le x}\left(X_t(\ell) - \min_{0\le k \le t} X_k(\ell)\right) \le \max_{0\le t \le x} X_t(\ell) + \max_{0\le t \le x}(-X_t(\ell)). \tag{321}$$

Hence, it suffices to show

$$\limsup_{n\to\infty} n^{-1}\left(\max_{0\le t \le n} X_t(\ell)\right) = \limsup_{n\to\infty} n^{-1}\left(\max_{0\le t \le n} -X_t(\ell)\right) = 0. \tag{322}$$

First assume $\ell + 1 = \ell^+$. In this case $X_x(\ell)$ is a lazy simple random walk on $\mathbb{Z}$. Hence, by the reflection principle,

$$\mathbb{P}\left(\max_{0\le t \le n}(-X_t(\ell)) \ge a\right) \le 2\mathbb{P}(-X_n(\ell) \ge a) \le \exp(-\frac{a^2}{n}). \tag{323}$$

The right-hand side is exponentially small in a by the bounded difference inequality. So taking $a = n^{2/3}$ and applying the Borel-Cantelli lemma show that $n^{-1} \max_{0\le t \le n}(-X_t(\ell))$ converges to zero almost surely. By a symmetric argument, the same conclusion holds for $n^{-1} \max_{0\le t \le n} X_t(\ell)$. Hence, this verifies the assertion.

Lastly, suppose $\ell + 1 < \ell^+$. In this case, $X_x(\ell)$ is not a random walk. Instead, from (133), we can write it as a Markov additive functional:

$$X_x(\ell) = \sum_{t=0}^{x} g^\ell(X_t^s, \xi_{t+1}). \tag{324}$$

Moreover, the increment $g^\ell(X_t^s, \xi_{t+1})$ does not depend on the whole X_t^s, but only on

$$Y_t := \big(X_t(\ell+1), \cdots, X_t(\ell^+ - 1)\big). \tag{325}$$

Let $\tau_0 = 0$ and τ_i for $i \ge 1$ denote the ith return time of Y_x to the origin. According to Theorem 7.3, τ_1 (and hence all τ_i's) have a finite moments of all orders.

Consider the process $R_i := -X_{\tau_i}(\ell)$. By the strong Markov property, the sequence R_i for $i \ge 1$ is a random walk. Denote its increment $\eta_i := R_i - R_{i-1}$. Then η_i has finite moments of all orders since each $\tau_i - \tau_{i-1}$ does so, and $X_x(\ell)$ changes at most by one in x. Moreover, by the strong law of large numbers and the Markov chain ergodic theorem,

$$\mathbb{E}[\eta_1] = \lim_{n\to\infty} \frac{R_n}{n} = \lim_{n\to\infty} \frac{\tau_n}{n} \frac{(-X_{\tau_n}(\ell))}{\tau_n} = \mathbb{E}[\tau_1]\, \mathbb{E}_{\pi^s \otimes \mathbf{p}}[g^\ell(X_x^s, \xi_{x+1})] = 0. \tag{326}$$

Hence R_i is a mean-zero random walk.

Denote $M := \max(\tau_1, \tau_2 - \tau_1, \ldots, \tau_n - \tau_{n-1})$. Since this is the maximum of i.i.d. random variables of finite moments of all orders, union bound and Chebyshev's inequality and (304) give

$$\mathbb{P}(M \geq a) = 1 - (1 - \mathbb{P}(\tau_1 \geq a))^n \leq n \, \mathbb{P}(\tau_1 \geq a) = O(n \, a^{-d}) \tag{327}$$

for any integer $d \geq 1$. Also, since the increments $X_{s+1}(\ell) - X_s(\ell)$ are bounded by 1,

$$\max_{0 \leq s \leq n} R_s \geq \max_{0 \leq s \leq \tau_n} -X_s(\ell) - M \geq \max_{0 \leq s \leq n} -X_s(\ell) - M. \tag{328}$$

Hence, combining the above inequalities and using Kolmogorov's maximal inequality, for any $b > 1/\sqrt{n}$,

$$\mathbb{P}\left(n^{-1} \max_{0 \leq s \leq n} -X_s(\ell) \geq b\right) \leq \mathbb{P}\left(\max_{0 \leq s \leq n} R_s(\ell) \geq nb - M\right) \tag{329}$$

$$\leq \mathbb{P}\left(\max_{0 \leq s \leq n} R_s(\ell) \geq nb - \sqrt{n}\right) + \mathbb{P}\left(M > \sqrt{n}\right) \tag{330}$$

$$= \frac{n \, \mathrm{Var}(\eta_1)}{(nb - \sqrt{n})^2} + O(n^{-2}). \tag{331}$$

Then taking $b = n^{-1/6}$ and denoting $T_n := \max_{0 \leq s \leq n} -X_s(\ell)$, we get

$$\mathbb{P}\left(n^{-1} T_n \geq n^{-1/3}\right) \leq \frac{c}{n^{2/3}} \tag{332}$$

for some constant $c > 0$. Notice that T_n is nondecreasing in n. By Borel-Cantelli Lemma and (332), we have that $n^{-2} T_{n^2} \to 0$ almost surely. Fix $k \geq 1$ and let $n = n(k)$ denote the largest inetger such that $n^2 \leq k < (n+1)^2$. Then using monotonicity,

$$\frac{n^2}{(n+1)^2} \frac{T_{n^2}}{n^2} \leq \frac{T_k}{k} \leq \frac{T_{(n+1)^2}}{(n+1)^2} \frac{(n+1)^2}{n^2}. \tag{333}$$

Taking $k \to \infty$, we deduce that $k^{-1} T_k \to 0$ almost surely as $k \to \infty$. Therefore, it follows that $n^{-1} \max_{0 \leq t \leq n}(-X_t(\ell))$ converges to zero almost surely. By a symmetric argument, the same conclusion holds for $n^{-1} \max_{0 \leq t \leq n} X_t(\ell)$. This completes the proof. □

Now we are ready to prove Theorem 2.5 **(i)**.

Proof of Theorem 2.5 (i). We wish to show that

$$\lim_{n \to \infty} n^{-1} W_n = \mu \quad \text{a.s.} \tag{334}$$

Note that by Lemma 9.3,

$$\limsup_{n \to \infty} n^{-1} W_n \leq \mu \quad \text{a.s.}, \tag{335}$$

where we interpret the inequality componentwise. Recall the Skorokhod decomposition $W_x = X_x + R Y_n$ in Lemma 6.3. We first consider the case when $\kappa \geq 3$. Then writing $R = I - Q$ with $Q = \mathrm{tridiag}_\kappa(0, 0, 1)$ and using the identity $(I - Q)(I + Q + Q^2 + \ldots) = I$, we see that R^{-1} is the following upper diagonal matrix whose nonzero entries equal to one:

$$R^{-1} = I + Q + \cdots + Q^{\kappa-1}. \tag{336}$$

Write

$$n^{-1} Y_n = R^{-1}(n^{-1} W_n - n^{-1} X_n). \tag{337}$$

Then by using (335) and the fact that $\lim_{n\to\infty} n^{-1} X_n = \mu$ a.s. (see Proposition 7.9),

$$\mathbf{a} := \limsup_{n\to\infty} (n^{-1} W_n - n^{-1} X_n) \leq \mathbf{0}, \tag{338}$$

where we applied limsup as well as inequality componentwise. It is crucial to note that R^{-1} has nonnegative entries. Hence,

$$\limsup_{n\to\infty} n^{-1} Y_n = R^{-1} \mathbf{a} \leq \mathbf{0}. \tag{339}$$

But since each Y_n is a nonnegative vector by definition, it follows that $\lim_{n\to\infty} n^{-1} Y_n = \mathbf{0}$ almost surely. Then using the Skorokhod decomposition once more, we get

$$\lim_{n\to\infty} n^{-1} W_n = \mu + R \lim_{n\to\infty} n^{-1} Y_n = \mu \tag{340}$$

almost surely, as desired.

It remains to verify (334) for the case when $\kappa = 1, 2$. Denote $\mathbf{y} := \limsup_{n\to\infty} n^{-1} Y_n$. Suppose $\kappa = 2$. Then the Skorokhod decomposition and (338) yield

$$\begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \mathbf{y} = \mathbf{a} \leq \mathbf{0}. \tag{341}$$

Note that $\mathbf{y} \geq \mathbf{0}$ since $Y_n \geq \mathbf{0}$ for all $n \geq 1$. Then it is easy to see that $\mathbf{y}$ must equal $\mathbf{0}$. The case for $\kappa = 1$ can be argued similarly. □

10. The diffusive scaling limit of the carrier process

In this section, we prove Theorem 2.5 **(ii)** on the diffusive scaling limit of the carrier process in the critical and the supercritical regime. The definition of SRBM below is adapted from [Wil98, Def. 3.1].

Definition 10.1 (Semimartingale reflecting Brownian motion). Fix an integer $\kappa \geq 1$ and a subset $J \subseteq \{1, \ldots, \kappa\}$. Let $S := \{(x_1, \ldots, x_\kappa) \in \mathbb{R}^\kappa : x_i \geq 0 \text{ for all } i \in J\}$ and let $\mathcal{B}$ denote the Borel σ-algebra on S, ν is a probability measure on $(S, \mathcal{B})$, θ is a constant vector in $\mathbb{R}^\kappa$, Σ is a $\kappa \times \kappa$ covariance matrix (symmetric and positive semidefinite[1]) and R is a $\kappa \times \kappa$ matrix. A *semimartingale reflecting Brownian motion* (SRBM) associated with the data $(S, \theta, \Sigma, R, \nu)$ is an $\{\mathcal{F}_t\}$-adapted, κ-dimensional process W defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and filtration $\{\mathcal{F}_t; t \geq 0\}$ (an increasing family of sub-σ-algebras of $\mathcal{F}$) such that

(i) $W = X + RY$, $\mathbb{P}$-a.s.;
(ii) $\mathbb{P}$-a.s., W has continuous paths and $W(t) \in S$ for all $t \geq 0$;
(iii) Under $\mathbb{P}$,
    (a) X is a κ-dimensional Brownian motion with drift vector θ, covariance matrix Σ and $X(0) \sim \nu$;
    (b) $\{X(t) - X(0) - \theta t, \mathcal{F}_t; t \geq 0\}$ is a martingale;
(iv) Y is an $\{\mathcal{F}_t\}$-adapted, κ-dimensional process such that $\mathbb{P}$-a.s. for $i = 1, \ldots, \kappa$,
    (a) $Y_i(0) = 0$;
    (b) Y_i is continuous and nondecreasing;
    (a) Y_i can increase only when W is on the face $F_i := \{x \in S : x_i = 0\}$, i.e., $\int_0^\infty \mathbf{1}(W_i(s) > 0) \, dY_i(s) = 0$.

Roughly speaking, an SRBM $W = X + RY$ behaves like the Brownian motion X in the interior of the domain S, and it is confined to the domain by the instantaneous 'reflection' (or 'pushing') at the boundary, where the direction of such 'reflection' on the ith face F_i is given by the ith column of the

---

[1] We allow the covariance matrix to be degenerate.

reflection matrix R. Note that in Definition 10.1, the domain S only requires coordinates in the set J be nonnegative, while it is standard to take S to be the nonnegative orthant $\mathbb{R}^{\kappa}_{\geq 0}$. We take this slightly more general domain to analyze the diffusive scaling limit of the centered carrier process $\overline{W}_t$ in Theorem 2.5, which can take negative values in coordinates corresponding to unstable colors.

A classical result of Reiman and Williams [RW88] (see also [Wil98, Thm. 3.1]) shows that an SRBM associated with $(S, \theta, \Sigma, R, \nu)$ with $S = \mathbb{R}^{\kappa}_{\geq 0}$ and Σ nondegenerate uniquely exists if and only if the reflection matrix R is *completely-$\mathcal{S}$* (see Definition 10.2). Roughly speaking, this condition means that at any boundary point of S, there exists a nonnegative linear combination of the reflection directions (i.e., columns of R) that points to the interior of S. When Σ is degenerate, then SRBM still exists but may not be unique.

Definition 10.2 (Completely-$\mathcal{S}$). A matrix $R \in \mathbb{R}^{d \times d}$ is *completely-$\mathcal{S}$* if for every principal submatrix R_0 of R, there is a nonnegative vector x_0 such that $R_0 x_0$ has strictly positive coordinates. Here, a principal submatrix of R is a matrix obtained by deleting all rows and columns of R with indices in some proper subset set $\mathcal{I} \subsetneq \{1, \ldots, d\}$ (possibly empty).

It is critical to notice that the reflection matrix R in (118) that gives a Skorokhod decomposition of the carrier process W_x as in Lemma 6.3 has the following property: For $\kappa \geq 3$, $R = I - Q$, where Q has a spectral radius less than one. In this case, we can say a lot about SRBM with a more direct argument. The first step is to recall that the problem that defines SRBM in Definition 10.1 is a particular instance of the classical Skorokhod problem stated below.

Definition 10.3 (Skorokod Problem). Fix a subset $J \subseteq \{1, \ldots, \kappa\}$ and let $S := \{(x_1, \ldots, x_{\kappa}) \in \mathbb{R}^{\kappa} : x_i \geq 0 \text{ for all } i \in J\}$. Let C_S denote the subspace of $C^{\kappa}(0, \infty)$ consisting of paths x with $x(0) \in S$. Fix matrix $R \in \mathbb{R}^{\kappa \times \kappa}$ and $x \in C_S$. A pair $(z, y) \in C^{\kappa}(0, \infty) \times C^{\kappa}(0, \infty)$ is a solution of the Skorohod problem for x w.r.t. R if the following conditions hold:

(i) $z(t) = x(t) + R\, y(t)$ for all $t \geq 0$.
(ii) $z(t) \in S$ for all $t \geq 0$.
(iii) For $i = 1, \ldots, \kappa$, $y_i(0) = 0$, $y_i(t)$ is nondecreasing, and $\int_0^{\infty} \mathbf{1}(i \in J)\mathbf{1}(z_i(t) \geq 0)\, dy_i(t) = 0$.

When the reflection matrix R can be written as $R = I - Q$ where Q is nonnegative and has a spectral radius less than one, then there is a unique solution (z, y) to the Skorokhod problem for each path x and the map $x \mapsto (z, y)$ (the Skorohod map) is continuous. This result is stated and proved in Theorem 10.4.

Theorem 10.4 (Harrison and Reiman '81). *Let $S = \mathbb{R}^d \times \mathbb{R}^{\kappa-d}_{\geq 0}$ and C_S be as in Definition 10.3. Suppose the reflection matrix R can be written as $R = I - Q$, where Q is nonnegative and has a spectral radius of less than one. Then for each path $x \in C_S$, there exists a unique pair of functions $(z, y) \in C^{\kappa}(0, \infty) \times C^{\kappa}(0, \infty)$ that solves the Skorokhod problem in Definition 10.3. Furthermore, denoting $z = \phi(x)$ and $y = \psi(x)$, both ϕ and ψ are continuous mappings $C_S \to C^{\kappa}(0, \infty)$.*

Proof. The original result [HR81, Thm. 1] is stated for $S = \mathbb{R}^{\kappa}_{\geq 0}$, where in our setting we allow S to be the intersection of axes-parallel half-spaces in $\mathbb{R}^{\kappa}$. A minor modification of the proof of [HR81, Thm. 1] will show the minor extension as stated above. We sketch the argument for completeness.

Without loss of generality, assume $S = \mathbb{R}^d \times \mathbb{R}^{\kappa-d}_{\geq 0}$ for some $d \in \{0, \ldots, \kappa\}$. Denote $C = C^{\kappa}(0, \infty)$ and fix $x \in C_S$. Let C_0 be the set of paths $y \in C$ such that $y(0) = \mathbf{0}$ and nondecreasing componentwise. Define a map $\pi = \pi_x : C_0 \to C_0$ such that

$$\pi(y)_i(t) = \begin{cases} 0 & \text{if } i = 1, \ldots, d \\ \sup_{0 \leq s \leq t}[y(s)Q - x(s)]^+ & \text{if } i = d+1, \ldots, \kappa. \end{cases} \tag{342}$$

Then one can check that (z, y) is a solution to the Skorokhod problem if and only if

$$y \in C_0, \qquad y = \pi(y), \qquad z = x + (I - Q)y. \tag{343}$$

One can then argue that there is a unique solution $y \in C_0$ such that $y = \pi(y)$.

To this end, for each square matrix Q, we let $\|Q\|_\infty$ denote its maximum absolute row sum. Since Q is nonnegative and has spectral radius < 1, there exists a positive diagonal matrix Λ such that $\tilde{Q} := \Lambda^{-1} Q \Lambda$ satisfies $\|\tilde{Q}\|_\infty < 1$ [Vei69, Lem. 3]. Observe that (z, y) satisfies (343) if and only if $(z\Lambda, y\Lambda)$ satisfies (343) with x and Q replaced by Λx and $\tilde{Q}$. Thus, without loss of generality, we may assume $\|Q\|_\infty < 1$.

Now fix $T \geq 0$ and define $C_0[0, T]$ and $C_S[0, T]$ in the obvious way. These are complete metric spaces endowed with the norm

$$\|y\| := \max_{1 \leq j \leq \kappa} \sup_{0 \leq t \leq T} |y_j(t)|. \tag{344}$$

Then one can show that the map π is a contraction on $C_0[0, T]$:

$$\|\pi(y) - \pi(y')\| \leq \|Q\|_\infty \|y - y'\|. \tag{345}$$

Since $\|Q\|_\infty < 1$, it follows that π is a contraction mapping, implying that there is a unique fixed point $y \in C_0$.

Now to show the continuity of the mapping $x \mapsto \phi(x)$, we observe that $\phi(x)$, being the unique fixed point of $y = \pi_x(y)$ of the contraction mapping π, can be explicitly constructed as the limit of $y^n(x) := \pi_x^n(y^0)$ with $y^0 \equiv 0$. Then note that for $x, x' \in C_0[0, T]$,

$$\|y^{n+1}(x) - y^{n+1}(x')\| \leq \|x - x'\| + \|Q\|_\infty \|y^n(x) - y^n(x')\|. \tag{346}$$

By an induction and taking $n \to \infty$, we get $\|\phi(x) - \phi(x')\| \leq \frac{1}{1 - \|Q\|_\infty} \|x - x'\|$. Thus ϕ is $\frac{1}{1 - \|Q\|_\infty}$-Lipschitz continuous on $C_0[0, T]$. Since T was arbitrary, this implies continuity of ϕ on $C_0(0, \infty)$. Thus, ϕ is continuous on $C^\kappa(0, \infty)$ in the topology of uniform convergence on compact intervals. The continuity of the mapping $x \mapsto \psi(x)$ is clear from the last identity in (343). $\qquad\square$

In the proof of Theorem 10.4, we have used the fact that if Q is a matrix of spectral radius less than one, then there exists a positive diagonal matrix Λ such that $\Lambda^{-1} Q \Lambda$ has maximum absolute row sum strictly less than one, appealing to [Vei69, Lem. 3]. In our case, $Q = \text{tridiag}_\kappa(0, 0, 1)$, and we can directly take Λ to have diagonal entries $\Lambda(i, i) = \kappa - i + 1$ for $i = 1, \ldots, \kappa$, in which case the maximum absolute row sum equals $\frac{\kappa - 1}{\kappa} < 1$.

Proof of Theorem 2.5 (ii). For this proof, we will appeal to the continuity of the Skorokod map $x \mapsto (y, z)$ we established in Theorem 10.4. Let $J = \{1, \ldots, \kappa\} \setminus \{\alpha_j ; j = 1, \ldots, r, p_{\alpha_j} > p_{\alpha_{j+1}}\}$ and $S := \{(x_1, \ldots, x_\kappa) \in \mathbb{R}^\kappa : x_i \geq 0 \text{ for all } i \in J\}$. Let $\mu = (\mu_1, \ldots, \mu_\kappa) := \sum_{j=1}^r \mathbf{e}_{\alpha_j}(p_{\alpha_j} - p_{\alpha_{j+1}})$. Then μ is nonzero in its jth coordinate if and only if $j \in J$. Recall the Skorokhod decomposition of the carrier process W_x in Lemma 6.3:

$$\overline{W}_x = \overline{X}_x + R Y_x \quad \text{for } x \in N, \tag{347}$$

where we denoted $\overline{W}_x = W_x - x\mu$ and $\overline{X}_x = X_x - x\mu$. Since $W_x \in \mathbb{R}_{\geq 0}^\kappa$, we have $\overline{W}_s \in S$ for all $s \in \mathbb{R}_{\geq 0}$. Note that (347) gives a Skorokhod decomposition of the centered carrier process $(\overline{W}_x)_{x \in \mathbb{N}}$. Namely, for each $i \in J$, $Y_x(i)$ can increase only if $\overline{W}_x(i) = 0$. This is because for $i \in J$, $\overline{W}_x(i) = W_x(i)$, and by Lemma 6.3, we know that $Y_x(i)$ increases only if $W_x(i) = 0$. From (347), we deduce

$$\widetilde{W}^n(t) = \widetilde{X}^n(t) + R\widetilde{Y}^n(t) \quad \text{for } t \in \mathbb{R}_{\geq 0}, \tag{348}$$

where $\widetilde{W}^n$, $\widetilde{X}^n$ and $\widetilde{Y}^n$ are the linear interpolations of $\frac{1}{\sqrt{n}}(W_x - x\mu)$, $\frac{1}{\sqrt{n}}(X_x - x\mu)$, and $\frac{1}{\sqrt{n}} Y_x$.

Since $R = \text{tridiag}_\kappa(0, 1, -1)$, we can write $R = I - Q$, where $Q = \text{tridiag}_\kappa(0, 0, 1)$, so Q has spectral radius zero for all $\kappa \geq 1$ since Q^κ is zero. Denoting $x \mapsto (\phi(x), \psi(x))$ by the Skorohod mapping as in

Theorem 10.4, according to (348), for each $n \geq 1$, we have

$$\phi(\widetilde{X}^n) = \widetilde{W}^n \quad \text{and} \quad \psi(\overline{X}^n) = \widetilde{Y}^n. \tag{349}$$

That is, the pair $(\widetilde{W}^n, \widetilde{Y}^n)$ is the unique solution of the Skorokhod problem for $\widetilde{X}^n$ with respect to the reflection matrix R. Recall that by Proposition 7.9, $\widetilde{X}^n$ converges weakly to the Brownian motion B in $\mathbb{R}^\kappa$ with zero drift and covariance matrix Σ. By continuity of the Skorohod mapping, it follows that

$$\widetilde{W}^n \implies \lim_{n \to \infty} \phi(\widetilde{X}^n) = \phi(B), \tag{350}$$

$$\widetilde{Y}^n \implies \lim_{n \to \infty} \psi(\widetilde{X}^n) = \psi(B). \tag{351}$$

In particular, $\widetilde{W}^n$ converges weakly to the SRBM associated with data $(S, \mathbf{0}, \Sigma, R, \delta_{\mathbf{0}})$, as desired. $\qquad\square$

11. Proofs of Theorems 2.7 and 2.7

In this section, we establish scaling limits of the top soliton lengths for the i.i.d. model in the critical and the supercritical regimes.

By now, it is easy to deduce Theorem 2.6.

Proof of Theorem 2.6. Suppose $p_0 = \max(p_1, \cdots, p_\kappa)$. Then $\mathcal{C}_u^{\mathbf{p}} = \{0 \leq i \leq \kappa : p_i = p_0\}$, and we may write $\mathcal{C}_u^{\mathbf{p}} = \{\alpha_0, \cdots, \alpha_r\}$ with $0 = \alpha_0 < \alpha_1 < \cdots < \alpha_r$. Then the weak convergence of the diffusively scaled first soliton length in (22) follows from Lemma 3.1, Theorem 2.5 and the continuous mapping theorem.

Next, we justify that $\lambda_j(n) = \Theta(n)$ with high probability for all $j \geq 1$. The upper bound follows since $\lambda_j(n) \leq \lambda_1(n) = O(\sqrt{n})$ with high probability. For the lower bound, we use the fact that the carrier process in the critical regime converges weakly to an SRBM as in Theorem 2.5. In particular, there are excursions of the carrier process of height (i.e., the L_1-norm) at least $c\sqrt{n}$ with high probability if $c > 0$ is small enough. Then the lower bound $\lambda_1(n) = \Omega(\sqrt{n})$ with high probability follows from Lemma 3.2. $\qquad\square$

In the rest of this section, we prove Theorem 2.7. Throughout, we will assume $p^* = \max(p_1, \cdots, p_\kappa) > p_0$. Let $\alpha_1 < \cdots < \alpha_r$ denote the unstable colors. Under the hypothesis, it holds that $p_{\alpha_1} = p^*$.

Proof of Theorem 2.7 (i). Let $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_\kappa)$ be as in Theorem 2.5. By Lemma 3.1 and Theorem 2.5, almost surely,

$$\lim_{n \to \infty} n^{-1} \lambda_1(n) = \lim_{n \to \infty} n^{-1} \|W_n\|_1 = \|\boldsymbol{\mu}\|_1 \tag{352}$$

$$= (p_{\alpha_1} - p_{\alpha_2}) + (p_{\alpha_2} - p_{\alpha_3}) + \cdots + (p_{\alpha_r} - p_0) = p^* - p_0. \tag{353}$$

Next, recall the Skorokhod decomposition $W_x = X_x + R Y_x$ in Lemma 6.3. Define $t(n) := \arg\max_{0 \leq t \leq n} \|W_t\|_1$. Let J denote the set of indices $i \in \{1, \ldots, \kappa\}$ such that $\mu_i > 0$. Then $\mu_i = 0$ if $i \notin J$, so

$$\lambda_1(n) - n\|\boldsymbol{\mu}\|_1 = \sum_{i=1}^\kappa W_{t(n)}(i) - n\mu_i \tag{354}$$

$$= \sum_{i \in J} W_{t(n)}(i) - n\mu_i + \max_{0 \leq t \leq n} \sum_{i \notin J} W_t(i). \tag{355}$$

By Proposition 6.2, it follows that

$$\sum_{i \in J} W_n(i) - n\mu_i + \sum_{i \notin J} W_n(i) \leq \lambda_1(n) - n\boldsymbol{\mu} \leq \sum_{i \in J} \hat{X}_{t(n)}(i) - n\mu_i + \max_{0 \leq t \leq n} \sum_{i \notin J} W_t(i). \qquad (356)$$

Recall that the linear interpolation of $n^{-1/2}(W_n - n\boldsymbol{\mu})$ converges weakly to the SRBM with specified data as in Theorem 2.5. Hence, the lower bound in Theorem 2.7 **(i)** follows from above. For the upper bound, we use Proposition 9.2 to note that, almost surely,

$$\lim_{n \to \infty} n^{-1/2} \left| \sum_{i \in J} \hat{X}_{t(n)}(i) - \sum_{i \in J} X_{t(n)}(i) \right| = 0. \qquad (357)$$

Hence, almost surely,

$$\limsup_{n \to \infty} n^{-1/2} \sum_{i \in J} (\hat{X}_{t(n)}(i) - n\mu_i) + n^{-1/2} \max_{0 \leq t \leq n} \sum_{i \notin J} W_t(i) \qquad (358)$$

$$= \limsup_{n \to \infty} n^{-1/2} \sum_{i \in J} (X_n(i) - n\mu_i) + n^{-1/2} \max_{0 \leq t \leq n} \sum_{i \notin J} W_t(i). \qquad (359)$$

Recall that by Proposition 7.9, the linear interpolation of $n^{-1/2}(X_n - n\boldsymbol{\mu})$ converges to a Brownian motion on $\mathbb{R}^\kappa$ with mean zero and an explicit covariance matrix Σ. Also, by Theorem 2.5 and the continuous mapping theorem,

$$n^{-1/2} \max_{0 \leq t \leq n} \sum_{i \notin J} W_t(i) \implies \sup_{0 \leq v \leq 1} \sum_{i \notin J} W^v(v), \qquad (360)$$

where $W = (W^1, \ldots, W^\kappa)$ is the SRBM in Theorem 2.5. Thus, the upper bound in (24) follows by the continuous mapping theorem.   □

Next, we complete the proof of Theorem 2.7 **(ii)**–**(iii)**. To this effect, it suffices to show the following statement.

Theorem 11.1. *Suppose $p^* > p_0$ and fix $j \geq 2$. Then the following hold.*

 (i) *Suppose $p_i = p^*$ for a unique $1 \leq i \leq \kappa$. Then $\lambda_j(n) = \Theta(\log n)$ with high probability.*
(ii) *Suppose $p_i = p^*$ at least two distinct colors $1 \leq i \leq \kappa$. Then $\lambda_j(n) = \Theta(\sqrt{n})$ with high probability.*

We begin with the following definition. For $0 \leq i, j \leq \kappa$ and a finite subset $H \subseteq \mathbb{N}$, define a random variable $D_{i,j}(H)$ by

$$D_{i,j}(H) = \sum_{x \in H} [\mathbf{1}(\xi^{\mathbf{p}}(x) = i) - \mathbf{1}(\xi^{\mathbf{p}}(x) = j)], \qquad (361)$$

which equals the difference of the number of color i and color j balls in H given by $\xi^{\mathbf{p}}$.

Proposition 11.2. *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 $C > 0$,*

$$\mathbb{P}(D_{j,i}(H) \geq 2C \log n) \leq \exp(-C(p_i - p_j) \log n) \qquad (362)$$

for all $n \geq 1$.

Proof. Let $\varepsilon = p_i - p_j > 0$ and denote $|H| = m$. Note that $\mathbb{E}[D_{j,i}(H)] = -\varepsilon m$. Since $D_{j,i}(H)$ is a sum of i.i.d. increments with absolute value at most one, by Hoeffding's inequality,

$$\mathbb{P}(D_{j,i}(H) - \mathbb{E}[D_{j,i}(H)] \geq t) \leq e^{-t^2/(2m)} \qquad (363)$$

for any $t > 0$. Let $t = \varepsilon m + 2C \log n$. Then $t/m \geq \varepsilon$, so

$$\mathbb{P}(D_{j,i}(H) \geq 2C \log n) = \mathbb{P}(D_{j,i}(H) - \mathbb{E}[D_{j,i}(H)] \geq t) \leq e^{-(\varepsilon/2)t} \leq e^{-\varepsilon C \log n}. \tag{364}$$

This shows the assertion. □

Proof of Theorem 11.1. Denote $\xi := \xi^{n,\mathbf{p}}$. Our argument is based on Lemma 3.5. In this proof, for integers $a < b$, an 'interval' $[a, b]$ will refer to the set $\{a, a+1, \ldots, b\}$. We say a subset $A \subseteq \mathbb{N}$ is a nonincreasing subsequence if ξ is nonincreasing 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 3.5:

$$\lambda_1(n) + \lambda_2(n) = \max_{A_1 < A_2 \subseteq [1,n]} \mathrm{L}(A_1, \xi) + \mathrm{L}(A_2, \xi). \tag{365}$$

Let $A_1 < A_2$ be an optimal choice of subsequences that achieves $\lambda_1(n) + \lambda_2(n)$ according to (365). Let $I = [a, b]$ and $J = [c, d]$ denote the supporting intervals of A_1 and A_2, respectively. We split A_1 into successive disjoint sub-subsequences $A'_\kappa, A'_{\kappa-1}, \cdots, A'_1$, where in each A'_ℓ, we only pick the balls of color ℓ in A_1. Let $I_j := [\min A'_j, \max A'_j]$. This gives a noninterlacing partition of $I = I_\kappa \sqcup \cdots \sqcup I_1$. We split A_2 similarly and obtain a noninterlacing partition $J = J_\kappa \sqcup \cdots \sqcup J_1$ similarly. This gives us a partition of the whole interval $[1, n]$ into the following collection of disjoint sub-intervals

$$\mathcal{H} = \{[1, a-1], I_\kappa, I_{\kappa-1}, \cdots, I_1, [b+1, c-1], J_\kappa, J_{\kappa-1}, \cdots, J_1, [d+1, n]\}, \tag{366}$$

ordered from left to right.

For $\lambda_1(n)$, we choose a sub-optimal nonincreasing subsequence $A^{(i)}$ by choosing all balls of color i in $[1, n]$. Then $\lambda_1(n) \geq \mathrm{L}(A^{(i)}, \xi)$ by Lemma 3.5, so (365) yields

$$\lambda_2(n) \leq \mathrm{L}(A_1, \xi) + \mathrm{L}(A_2, \xi) - \mathrm{L}(A^{(i)}, \xi). \tag{367}$$

Then breaking the right-hand side of (367) into sub-intervals given by the partition in (366), we may write

$$\mathrm{L}(A_1, \xi) + \mathrm{L}(A_2, \xi) - \mathrm{L}(A^{(i)}, \xi) = \sum_{H \in \mathcal{H}} f(H), \tag{368}$$

where if $H = I_j$ or J_j $(1 \leq j \leq k)$,

$$f(H) := (\text{number of balls of color } j \text{ in } H - \text{number of balls of color } 0 \text{ in } H) \tag{369}$$
$$- (\text{number of balls of color } i \text{ in } H - \text{ number of balls of color } 0 \text{ in } H) \tag{370}$$
$$= D_{j,i}(H), \tag{371}$$

else if $H = [1, a-1], [b+1, c-1]$ or $[d+1, n]$,

$$f(H) := (\text{number of balls of color } 0 \text{ in } H - \text{number of balls of color } i \text{ in } H) \tag{372}$$
$$= D_{0,i}(H). \tag{373}$$

Now suppose that p_i is the unique maximum among $p_1, \cdots, p_\kappa$ and assume $p_i > p_0$. Note that $\mathcal{H}$ contains $2\kappa + 3$ intervals. Noting that $D_{i,i}(H) = 0$, a union bound and Proposition 11.2 give

$$\mathbb{P}\left(\sum_{H \in \mathcal{H}} f(H) \geq 2(2\kappa + 3)C \log n\right) \leq \sum_{[s,t] \subseteq [1,n]} \sum_{\substack{0 \leq \ell \leq \kappa \\ \ell \neq i}} \mathbb{P}(D_{\ell,i}([s,t]) \geq 2C \log n) \tag{374}$$

$$\leq 3n^2 \sum_{\substack{0 \leq \ell \leq \kappa \\ \ell \neq i}} \exp(-C(p_i - p_\ell) \log n) \tag{375}$$

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

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

$$D_{j,i}(H) \le \max_{1 \le s \le t \le n} \left| D_{j,i}([s,t]) \right| \le 2 \max_{1 \le t \le n} \left| D_{j,i}([1,t]) \right|. \tag{376}$$

In this case, $D_{j,i}([1,t])$ is a symmetric random walk with t increments. Hence, for some large enough constant $C > 0$, the right-hand side of (376) is at most $C\sqrt{n}$ with probability at least $1 - \varepsilon$ by the functional central limit theorem. This shows that $\lambda_2(n) = O(\sqrt{n})$ with probability at least $1 - \varepsilon$.

Now we prove the lower bounds in **(i)** and **(ii)**. Fix $j \ge 2$. Let $A_1, \ldots, A_{j-1}$ denote an optimal choice of noninterlacing subsets of $[1, n]$ such that

$$\lambda_1(n) + \cdots + \lambda_{j-1}(n) = \sum_{i=1}^{j-1} \mathrm{L}(A_i, \xi). \tag{377}$$

Denote $I_i := [\min A_i, \max A_i]$ for $i = 1, \ldots, r - 1$, so that $I_1, \ldots, I_{j-1}$ are noninterlacing supporting intervals for $A_1, \ldots, A_{j-1}$. For each interval $J = [s, t]$, let $N_0(J)$ denote the maximum number of *consecutive* 0's in the sequence $\xi_s, \xi_{s+1}, \ldots, \xi_t$. For each integer $1 \le \ell \le \kappa$, let $M_\ell(J)$ denote the maximum number of ℓ's (not necessarily consecutive) in the sequence $\xi_s, \xi_{s+1}, \ldots, \xi_t$. We will use these notations for the rest of the proof.

Fix a constant $0 < c_1 < 1/(3 \log p_0^{-1})$. We first show that $\mathbb{P}(\lambda_j(n)/\log n \ge c_1) = 1 - o(1)$. To this end, we claim that

$$\mathbb{P}(N_0(I_i) \ge c_1 \log n \text{ for some } i = 1, \ldots, r - 1) = 1 - o(1). \tag{378}$$

Note that if $N_0(I_i) \ge c_1 \log n$, then we can split the nonincreasing subsequence A_i into two nonincreasing subsequences A_i' and A_i'' by removing the $c_1 \log n$ consecutive zeros in the supporting interval I_i. Then $A_1 \prec \ldots A_{i-1} \prec A_i' \prec A_i'' \prec \cdots \prec A_{j-1}$ is a noninterlacing collection of nonincreasing subsequences, whose total penalized length has now increased by at least $c_1 \log n$. Thus, by Lemma 3.5, $\lambda_j(n) \ge c_1 \log n$ with high probability if the claim (378) holds.

Now we show (378). Fix a constant $0 < c_2 < p^* - p_0$. Since $\mathrm{L}(A_i, \xi) \le |I_i|$,

$$\mathbb{P}\left(\sum_{i=1}^{j-1} |I_i| < c_2 n \right) \le \mathbb{P}\left(\lambda_1(n) + \cdots + \lambda_{j-1}(n) < c_2 n \right) \le \mathbb{P}(\lambda_1(n) < c_2 n). \tag{379}$$

Since $\lambda_1(n)/n \to p^* - p_0 > c_2$ a.s. by Theorem 2.7 **(i)**, the above probability is of order $o(1)$.

Next, by using a union bound,

$$\mathbb{P}\left(\sum_{i=1}^{j-1} |I_i| \ge c_2 n, \, N_0(I_i) < c_1 \log n \text{ for all } i = 1, \ldots, j - 1 \right) \tag{380}$$

$$\le \mathbb{P}\left(\bigcup_{J_1 \prec \cdots \prec J_{j-1} \subseteq [1,n]} \left\{ \sum_{i=1}^{j-1} |J_i| \ge c_2 n, \, N_0(J_i) < c_1 \log n \text{ for all } i = 1, \ldots, j - 1 \right\} \right) \tag{381}$$

$$\le \mathbb{P}\left(\bigcup_{J_1 \prec \cdots \prec J_{j-1} \subseteq [1,n]} \bigcup_{i=1}^{j-1} \left\{ |J_i| \ge \frac{c_2 n}{r - 1}, \, N_0(J_i) < c_1 \log n \right\} \right) \tag{382}$$

$$\leq (r-1)n^{2(r-2)} \sum_{J \subseteq [1,n], |J| \geq \frac{c_2 n}{r-1}} \mathbb{P}(N_0(J) < c_1 \log n) \tag{383}$$

$$\leq (r-1)n^{2(r-1)}\mathbb{P}(N_0([1,n]) < c_1 \log n), \tag{384}$$

where J_i and J above denote deterministic intervals. We can subdivide the interval $[1, n]$ into consecutive subintervals $K_1, K_2, \ldots$ of length $\lceil c_1 \log n \rceil$. There are at least $\lfloor \frac{n}{c_1 \log n} \rfloor$ such subintervals, and they can be fully occupied with balls of color 0 independently with probability $p_0^{\lceil c_1 \log n \rceil}$. Hence, recalling $0 < c_1 < 1/(3 \log p_0^{-1})$,

$$\mathbb{P}(N_0([1,n]) < c_1 \log n) \leq \left(1 - p_0^{c_1 \log n}\right)^{\lfloor \frac{n}{c_1 \log n} \rfloor} \tag{385}$$

$$\leq \exp\left(-p_0^{c_1 \log n} \lfloor \frac{n}{c_1 \log n} \rfloor\right) \leq \exp(-n^{1/3}). \tag{386}$$

Therefore, (384) is of order $o(1)$. Now (378) follows by a union bound. In particular, this completes the proof of **(i)**.

Finally, suppose $p_{\alpha_1} = p_{\alpha_2} = p^*$ for some $1 \leq \alpha_1 < \alpha_2 \leq \kappa$. Fix $\varepsilon > 0$. We will show that there exists a constant $c = c(\varepsilon, j) > 0$ such that

$$\liminf_{n \to \infty} \mathbb{P}\left(n^{-1/2}\lambda_j(n) \geq c\right) \geq 1 - \varepsilon. \tag{387}$$

To this end, we split each A_i into successive disjoint sub-subsequences $A_{i,\kappa}, \cdots, A_{i,2}, A_{i,1}$, where in each $A_{i,\ell}$, we only pick the balls of color ℓ in A_i. Denote $I_{i,\ell} := [\min A_{i,\ell}, \max A_{i,\ell}]$. By (379) and a union bound,

$$\mathbb{P}\left(|I_{i,\ell}| \geq \frac{c_2 n}{\kappa(j-1)} \text{ for some } 1 \leq i \leq j-1 \text{ and } 1 \leq \ell \leq \kappa\right) = 1 - o(1). \tag{388}$$

Fix $\delta > 0$. Partition $[0, n]$ into intervals $J_k := [k\delta n, (k+1)\delta n]$ of equal length $\lfloor \delta n \rfloor$. We can choose δ small enough so that any fixed interval of length $\frac{c_2 n}{\kappa(j-1)}$ in $[1, n]$ contains J_k for some $1 \leq k \leq \lfloor \delta^{-1} \rfloor$.

For each $1 \leq \ell \leq \kappa$, choose $\ell_* \in \{i_1, i_2\} \setminus \{\ell\}$. Fix a constant $\alpha > 0$ and define the following event:

$$E_{k,\ell} := \left\{\max_{t \leq \delta n} n^{-1/2}\left|D_{\ell,\ell_*}\left([(k-1)\lfloor \delta n \rfloor, k\lfloor \delta n \rfloor + t]\right)\right| \geq \alpha\right\}. \tag{389}$$

Since D_{i,i_*} on these disjoint intervals are i.i.d., by the functional central limit theorem, we have

$$\liminf_{n \to \infty} \mathbb{P}\left(\bigcap_{k=1}^{\lfloor 1/\delta \rfloor} \bigcap_{\ell=1}^{\kappa} E_{k,\ell}\right) \geq 1 - \frac{\varepsilon}{2} \tag{390}$$

as long as the constant $\alpha > 0$ is small enough. By a union bound, for all $n \geq 1$ sufficiently large,

$$\mathbb{P}\left(\{J_k \subseteq I_{i,\ell} \text{ for some } k, i, \ell\} \cap E_{k,\ell}\right) \geq 1 - \varepsilon. \tag{391}$$

We now claim that

$$\{J_k \subseteq I_{i,\ell} \text{ for some } k, i, \ell\} \cap E_{k,\ell} \subseteq \{\lambda_j(n) \geq \alpha\sqrt{n}\}, \tag{392}$$

which is enough to conclude the desired lower bound $\lambda_j(n) = \Omega(\sqrt{n})$. To show this claim, suppose the event on the left-hand side above holds. Denote $I_{i,\ell} = [e, f]$. The maximum of D_{ℓ,ℓ_*} in the event $E_{k,\ell}$ occurs at site m in J_k, so we may split the interval $[e, f]$ into $[e, m]$ and $[m+1, f]$. Suppose

$D_{\ell,\ell_*}([e,m]) \geq \alpha\sqrt{n}$. Let $A_{i,\ell}^-$ and $A_{i,\ell}^+$ denote the subsequences formed by picking up all ℓ's in $[e,m]$ and all ℓ_*'s in $[m+1,f]$, respectively. Now define two nonincreasing subsequences A_i', A_i'' by

$$\begin{cases} A_i' := [A_{i,\kappa}, \ldots, A_{i,\ell+1}, A_{i,\ell}^-, A_{i,\ell}^+], & A_i'' := [A_{i,\ell-1}, \ldots, A_{i,1}] & \text{if } \ell > \ell_* \\ A_i' := [A_{i,\kappa}, \ldots, A_{i,\ell+1}, A_{i,\ell}^-], & A_i'' := [A_{i,\ell}^+, A_{i,\ell-1}, \ldots, A_{i,1}] & \text{if } \ell < \ell_*. \end{cases} \tag{393}$$

Together with the other $j-2$ subsequences $A_1, \ldots, A_{i-1}, A_{i+1}, \ldots, A_{j-1}$, these j noninterlacing and nonincreasing subsequences achieve total penalized lengths at least $\lambda_1(n) + \cdots + \lambda_{j-1}(n) + \alpha\sqrt{n}$. By Lemma 3.5, this implies $\lambda_j(n) \geq \alpha\sqrt{n}$. If $D_{\ell,\ell_*}([e,m]) \leq -\alpha\sqrt{n}$, then let $A_{i,\ell}^-$ and $A_{i,\ell}^+$ denote the subsequences formed by picking up all ℓ_*'s in $[e,m]$ and all ℓ's in $[m+1,f]$, respectively, and define

$$\begin{cases} A_i' := [A_{i,\kappa}, \ldots, A_{i,\ell+1}, A_{i,\ell}^-], & A_i'' := [A_{i,\ell}^+, A_{i,\ell-1}, \ldots, A_{i,1}] & \text{if } \ell > \ell_*, \\ A_i' := [A_{i,\kappa}, \ldots, A_{i,\ell+1}, A_{i,\ell}^-, A_{i,\ell}^+], & A_i'' := [A_{i,\ell-1}, \ldots, A_{i,1}] & \text{if } \ell < \ell_*. \end{cases} \tag{394}$$

In this case, we can also conclude $\lambda_j(n) \geq \alpha\sqrt{n}$ similarly. This completes the proof. $\square$

12. Proofs of combinatorial lemmas

In this section, we establish various combinatorial statements about the κ-color BBS dynamics and the associated carrier processes. Our main goal is to show Lemmas 3.1, 3.2 and 3.5. We also provide an elementary and self-contained proof of Lemma 3.3, which has been proved in the more general form in [KL20, Prop. 4.5] using connections with combinatorial R.

12.1. *Proof of Lemmas 3.1 and 3.2*

In this subsection, we prove Lemmas 3.1 and 3.2. We rely on the finite-capacity carriers (see Section 3.2) and Lemma 3.3. 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 12.1. *Let* $\xi : \mathbb{N} \to \mathbb{Z}_{\kappa+1}$ *be any κ-color BBS configuration with finite support. Denote by* $(\Gamma_{x;c})_{x \geq 0}$ *and* $(\Gamma_{x;c+1})_{t \geq 0}$ *the carrier processes over ξ with finite capacities c and $c+1$, respectively. Then for any $t \geq 0$, $\Gamma_{x;c}$ viewed as a c-dimensional vector is obtained by omitting a single coordinate in $\Gamma_{x;c+1}$ viewed as a $c+1$-dimensional vector.*

Proof. Fix a κ-color BBS configuration $\xi : \mathbb{N} \to \mathbb{Z}_{\kappa+1}$. Let $(\Gamma_{x;c})_{x \geq 0}$ and $(\Gamma_{x;c+1})_{x \geq 1}$ denote the carrier processes over ξ with finite capacities c and $c+1$, respectively. We will show the assertion by induction on $x \geq 0$. For $x = 0$, 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 $x \geq 0$. Denote $S = \Gamma_{x;c}, T = \Gamma_{x+1;c} \in \mathcal{B}_c$ and $S' = \Gamma_{x;c+1}, T' = \Gamma_{x+1;c+1} \in \mathcal{B}_{c+1}$. Recall that the entries in carrier states are nonincreasing from left, which is the opposite of the convention for semistandard Young tableaux (as used in [KL20] and [KLO18]).

By the induction hypothesis, we may assume that S can be obtained from T by omitting its j_*th entry $T(j_*) = r$. Let B and A be the blocks to the left and right of the entry $T(j_*)$ of T. Hence, S is the concatenation of the blocks B and A (see Figure 9 left). Let $q := \xi_{x+1}$.

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 $r = T'(j_* - 1)$ 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 $r = T'(j_*)$ from T'. Otherwise, q replaces r in T to get T' (see in Figure 9 right). Then q must replace the largest

Figure 9. (Left) $S \in \mathcal{B}_c$ is obtained from $T \in \mathcal{B}_{c+1}$ 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.

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. □

Proof of Lemma 3.1. Fix a κ-color BBS configuration $\xi : \mathbb{N} \to \mathbb{Z}_{\kappa+1}$. For each integer $c \geq 1$, let $(\Gamma_{x;c})_{x \geq 0}$ denote the capacity-c carrier process over ξ. Let $(\Gamma_x)_{x \geq 0}$ denote the infinite capacity carrier process over ξ. We also write

$$M = \max_{s \geq 0}(\text{\# of nonzero entries in } \Gamma_s.) \tag{395}$$

Note that from Lemma 3.3, we can deduce that for any $1 \leq j \leq \rho_1(\xi)$,

$$\lambda_j(\xi) = |\{k \geq 1 : \rho_k(\xi) \geq j\}| = \max\left\{k \geq 1 \,\middle|\, E_k(\xi) \geq E_{k-1}(\xi) + j\right\}, \tag{396}$$

where $E_k(\xi)$ is defined in (405).

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

$$\tau_c := \inf\left\{x \geq 0 \mid \Gamma_{x;c} \text{ contains all positive entries and } 0 < \xi_{x+1} \leq \min \Gamma_{x;c}(x)\right\}. \tag{397}$$

We let $\tau_c = \infty$ if the set on the right-hand side is empty. Note that if we consider two carrier processes $\Gamma_{x;c}$ and $\Gamma_{x;c+1}$, then $\tau_c + 1$ is the first time that they contain distinct sets of nonzero entries. Moreover, $\Gamma_{\tau_c+1;c+1}$ 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

$$E_c = Const. \qquad \forall c \geq M. \tag{398}$$

Hence, $\lambda_1(\xi) \leq M$ by (396).

However, note that $x^* := \tau_{M-1} < \infty$ and ξ_{x^*+1} does not exceed the smallest entry in $\Gamma_{x^*;M-1}$ by definition of τ_{M-1}. So $\mathbf{1}(\xi_{x^*+1} > \min \Gamma_{x^*;M-1}) = 0$. Also, since $\Gamma_{x^*;M-1}$ and $\Gamma_{x^*;M}$ share the same positive entries, $\Gamma_{x^*;M}$ is obtained from $\Gamma_{x^*;M-1}$ by augmenting 0 to its right. Since $\xi_{x^*+1} > 0$ by definition of x^*, we have $\mathbf{1}(\xi_{x^*+1} > \min \Gamma_{x^*;M}) = 1$. Moreover, by Proposition 12.1,

$$\mathbf{1}(\xi_{x+1} > \min \Gamma_{x;c}) \geq \mathbf{1}(\xi_{x+1} > \min \Gamma_{x;c-1}) \tag{399}$$

for all $c \geq 1$ and $x \geq 0$. It follows that $E_M \geq E_{M-1} + 1$. Hence, by (396), we deduce $\lambda_1(\xi) \geq M$. This shows $\lambda_1(\xi) = M$, as desired. □

Proof of Lemma 3.2. Fix a κ-color BBS configuration ξ with finitely many balls of positive colors. Let $W := (W_x)_{x \geq 0}$ be the carrier process over ξ. Let $T_0 := 0$ and let T_k for $k \geq 1$ denote the kth site that the carrier returns to the origin. Define sub-configurations $\xi^{(1)} := (\xi_0, \xi_1, \ldots, \xi_{T_1-1})$, $\xi^{(2)} := (\xi_{T_1}, \xi_{T_1+1}, \ldots, \xi_{T_2-1})$, and so on. Let N denote the number of nontrivial excursions of the carrier process W. Then ξ is the concatenation of $\xi^{(1)}, \ldots, \xi^{(N)}$. We wish to show that the soliton decomposition

of ξ is the union of the soliton decomposition of $\xi^{(i)}$'s. Equivalently, we wish to show that

$$\rho_c(\xi) = \sum_{k=1}^{N} \rho_c(\xi^{(k)}) \quad \text{for all } c \geq 1. \tag{400}$$

To show the claim (400) above, let $(\Gamma_{x;c})_{x \geq 0}$ denote the capacity-c carrier process over ξ. By Proposition 12.1, we have $\Gamma_{T_k;c} = \mathbf{0}$ for all $k \geq 0$. In words, the capacity-c carrier resets to empty at each site T_k. Hence, if we let $(\Gamma_{x;c}^{(k)})_{T_{k-1} \leq x < T_k}$ denote the capacity-c carrier process over $\xi^{(k)}$, then

$$(\Gamma_{x;c}^{(k)})_{T_{k-1} \leq x < T_k} = (\Gamma_{x;c})_{T_{k-1} \leq x < T_k}. \tag{401}$$

It follows that

$$\sum_{x=1}^{N} \mathbf{1}(\xi_s > \min \Gamma_{x-1;c}) = \sum_{k=1}^{N} \sum_{T_{k-1} < x \leq T_k} \mathbf{1}(\xi_s^{(k)} > \min \Gamma_{x-1;c}^{(k)}). \tag{402}$$

By Lemma 3.3, the above yields

$$\rho_1(\xi) + \cdots + \rho_c(\xi) = \sum_{k=1}^{N} \rho_1(\xi^{(k)}) + \cdots + \rho_c(\xi^{(k)}). \tag{403}$$

The above holds for all $c \geq 1$. By using induction in c, one can then deduce (400).

The second part of the assertion that $\lambda_j(n) \geq \mathbf{h}_j(n)$ is immediate from the first part we have just shown above and in Lemma 3.1. □

12.2. *Proof of Lemmas 3.3 and 3.5*

Recall the notations introduced in Section 3.3. For any κ-color BBS configuration $X : \mathbb{N} \to \mathbb{Z}_{\kappa+1}$ with finite support and integer $k \geq 1$, we denote

$$R_k(\xi) := \max_{A_1 \sqcup \cdots \sqcup A_k} \sum_{i=1}^{k} \mathrm{NA}(A_i, \xi), \qquad L_k(\xi) := \max_{A_1 < \cdots < A_k \subseteq \mathbb{N}} \sum_{i=1}^{k} \mathrm{L}(A_i, \xi). \tag{404}$$

Lastly, we also denote

$$E_k(\xi) := \sum_{s=1}^{\infty} \mathbf{1}(\xi_s > \min \Gamma_{s-1;k}), \tag{405}$$

where $(\Gamma_{x;i})_{t \geq 0}$ is the capacity-i carrier process over ξ. We set $R_0(\xi) = L_0(\xi) = E_0(\xi) = 0$ for convenience. In this subsection, we will show with an elementary argument that the above quantities associated with a κ-color BBS configuration are invariant under time evolution. This will lead to the proof of Lemmas 3.5 and 3.3.

We remark that the invariants $E_k(\xi)$ 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 κ-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, KL20, KLO18].

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 κ-color BBS evolution $X \mapsto X'$ is defined by

$$\xi' := K_1 \circ K_2 \circ \cdots \circ K_\kappa(\xi). \tag{406}$$

We propose two ways to simplify the κ-color BBS dynamics. First, using the cyclic symmetry of the system, we can reformulate the update of a κ-color BBS configuration in terms of κ 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 κ jump according to the rule K_κ, and we relabel each of them with color 1 and increase the positive colors of all other balls by 1. Then we have

$$K_1 \circ K_2 \circ \cdots \circ K_\kappa(\xi) = (\mathcal{T}_\kappa)^\kappa(\xi). \tag{407}$$

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 κ-color BBS configuration $\xi : \mathbb{N} \to \mathbb{Z}_{\kappa+1}$ of finite support, we define its *standardization* to be the following map $\hat{\xi} : \mathbb{N} \to \mathbb{Z}_{\geq 0}$: For each $1 \leq i \leq \kappa$, let m_i denote the number of balls in X of color i. Then to produce $\hat{\xi}$, 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 $\hat{\xi}$ is an N-color BBS configuration with each color in $\{1, \cdots, N\}$ used for exactly one ball.

Proposition 12.2. *Let ξ and $\hat{\xi}$ denote a κ-color BBS configuration with finite support and its standardization, respectively. Then the following hold.*

(i) *Standardization preserves the number of ascents, noninterlacing nonincreasing sequences, and their penalized lengths. In particular, for each $k \geq 1$,*

$$R_k(\xi) = R_k(\hat{\xi}), \qquad L_k(\xi) = L_k(\hat{\xi}). \tag{408}$$

(ii) *ξ and $\hat{\xi}$ give the same soliton partition (i.e., $\Lambda(\xi) = \Lambda(\hat{\xi})$).*

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

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

$$\hat{\xi}' = (\hat{\xi})'. \tag{409}$$

To see this, observe that for the evolution $\xi \mapsto \xi'$, after all balls of color κ have jumped, they return to the same left-right order as before: if some ball of color κ, say in position x, jumped over some other ball of color κ, 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 κ in the previous step are triggered one by one from left, 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 $\hat{\xi} \mapsto \hat{\xi}'$. This shows (409), as desired. □

In the following proposition, we show the time-invariance of the three quantities associated with a given BBS configuration. This will show most of Lemma 3.5.

Proposition 12.3. *Let ξ be an arbitrary κ-color BBS configuration of finite support. Fix $j \geq 1$. The following hold.*

(i) $E_j(\xi) = E_j(\mathcal{T}_\kappa(\xi))$.

(ii) $R_j(\xi) = E_j(\xi)$.

(iii) $L_j(\xi) = L_j(\mathcal{T}_\kappa(\xi))$.

(iv) *If* $(\xi^{(t)})_{t \geq 0}$ *denotes the* κ-*color BBS trajectory with* $\xi = \xi_0$, *then for all* $t \geq 1$,

$$E_j(\xi^{(t)}) = R_j(\xi^{(t)}) \equiv E_j(\xi), \qquad L_j(\xi^{(t)}) \equiv L_j(\xi). \tag{410}$$

We first derive Lemmas 3.5 and 3.3 assuming Proposition 12.3.

Proof of Lemma 3.3 and 3.5. Let $(\xi^{(t)})_{t \geq 0}$ be a κ-color BBS trajectory such that ξ_0 has finite support. We take $T \geq 1$ large enough so that at time T, the system decomposes into noninteracting solitons whose lengths are nondecreasing from left. We can reformulate the condition that a κ-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 $\xi^{(T)} \mapsto \xi^{(T+1)}$, 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:

$$\begin{gathered} \text{The } i\text{th largest entry of the right soliton is strictly larger} \\ \text{than the } i + g\text{th largest entry of the left soliton.} \end{gathered} \tag{411}$$

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

$$\cdots 00433200431100 \cdots . \tag{412}$$

For each $k \geq 1$, let λ_k denote the length of the kth-longest soliton and let ρ_k denote the number of solitons of length $\geq k$. They both form the same Young diagram, whose kth column and row lengths are given by λ_k and ρ_k, respectively.

For each $j \geq 1$, let $(\Gamma_{s;j})_{s \geq 0}$ denote the capacity-j carrier process on $\xi^{(t)}$. As the carrier process over $\xi^{(t)}$ 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 soliton, this is clear, and otherwise, it is still true due to the separation condition (411). Hence, we have

$$E_j(\xi^{(T)}) = \sum_{k=1}^{\infty} \min(\lambda_k, j) = \sum_{k=1}^{j} \rho_k. \tag{413}$$

Then by Proposition 12.3, we deduce

$$R_j(\xi^{(t)}) = E_j(\xi^{(t)}) = E_j(\xi^{(T)}) = \sum_{k=1}^{j} \rho_k \tag{414}$$

for all $t \geq 0$, as desired. In the general case, the above equations hold due to the separation condition (411). This shows Lemma 3.3 as well as the first equation in Lemma 3.5.

Similarly, for the second equation in Lemma 3.5, it suffices to show $L_j(\xi^{(T)}) = \lambda_1 + \cdots + \lambda_j$. It is easy to see $L_j(\xi^{(T)}) \geq \lambda_1 + \cdots + \lambda_j$ by choosing the j longest nonincreasing sequences given by the top j solitons. It remains to show the converse inequality; choose a collection of noninterlacing nonincreasing subsequences on supports $A_1, A_2, \cdots, A_j$ that achieves $L_j(\xi^{(T)})$. We may assume that $|A_1| + \cdots + |A_j|$ is as small as possible, where $|\cdot|$ means (nonpenalized) 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 is achieved when A_i's are the support of the j longest nonincreasing sequences given by the solitons, which shows the assertion.

To show the claim, for each $i \geq 1$, let $\mathbf{u}_i$ denote the maximal nonincreasing subsequence of positive colors in the ith longest soliton in $\xi^{(T)}$. Schematically, we can write $\xi^{(T)}$ as

$$\xi^{(T)}: \quad \cdots \mathbf{u}_3 0 \cdots 0 \mathbf{u}_2 0 \cdots 0 \mathbf{u}_1 00 \cdots . \tag{415}$$

Let l_i denote the number of 0's between $\mathbf{u}_{i+1}$ and $\mathbf{u}_i$.

Suppose for contradiction that some A_k intersects with two $\mathbf{u}_i$'s. Let i be as small as possible so that A_k intersects with $\mathbf{u}_{i+1}$ and $\mathbf{u}_i$. We first suppose the case when the two solitons have a sufficient gap (i.e., $l_{i+1} \geq \lambda_{i+1}$). Let $A'_k = A_k \setminus \mathbf{u}_{i+1}$. Then $A_1, \cdots, A_{k-1}, A'_k, A_{k+1}, \cdots, A_j$ is a sequence of noninterlacing nonincreasing subsequences in $\xi^{(t)}$ with a strictly smaller total number of elements than the original sequence. Moreover, this new sequence achieves the optimum $L_j(\xi^{(T)})$ since

$$\mathrm{L}(A'_k, \xi^{(T)}) \geq \mathrm{L}(A_k, \xi^{(T)}) - \mathbf{u}_{i+1} + l_i \geq \mathrm{L}(A_k, \xi^{(T)}). \tag{416}$$

Namely, omitting all elements of $\mathbf{u}_{i+1}$ from A_k deletes at most $|\mathbf{u}_{i+1}|$ positive numbers but at least $l_i \geq |\mathbf{u}_{i+1}|$ zeros. This contradicts the minimality of the original sequence $A_1, \cdots, A_j$. This shows the claim. Lastly, when the gap between the solitons is small (i.e., $l_{i+1} < \lambda_{i+1}$), one can argue similarly by using the separation condition (411). This shows the claim, as desired. □

Lastly in this subsection, we prove Proposition 12.3.

***Proof of Proposition 12.3.*** **(iv)** immediately follows from **(i)**–**(iii)**. According to Proposition 12.2, the assertion is valid for arbitrary BBS if and only if it is true for the standardized system with initial configuration $\hat{\xi}$, where each positive color is used exactly once. Hence, without loss of generality, we may assume that each positive color in ξ is used exactly once. Furthermore, in proving (i)–(iii), we may assume that there is a ball of color κ in ξ, since otherwise, the cyclic update rule $\mathcal{T}_\kappa$ 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 $\xi' = \mathcal{T}_\kappa(\xi)$. For any string $\mathbf{u}$ of integers in $\{0, 1, \ldots, \kappa - 1\}$, we let $\mathbf{u}'$ denote the string obtained by incrementing the positive integers in $\mathbf{u}$ by one.

(i) Suppose $\xi_x = \kappa$ and the ball of color κ is in a contiguous block of balls whose labels are $\mathbf{u}\kappa\mathbf{v}0\mathbf{w}$ for some words $\mathbf{u}, \mathbf{v}$. Note that $\mathbf{u}$ and $\mathbf{w}$ consist of integers in $\{0, \ldots, \kappa - 1\}$, while $\mathbf{v}$ is either empty or only has positive integers $< \kappa$. After the update $\xi \mapsto \xi' := \mathcal{T}_\kappa(\xi)$, we reach an arrangement in which $\mathbf{u}, \mathbf{v}$ and $\mathbf{w}$ have had their labels incremented, the space between them is empty ($\xi'_x = 0$) and 1 follows $\mathbf{v}$. Let y be the site such that $\xi'_y = 1$. Here is a schematic:

configuration	arrangement		
ξ	$[\ \cdots \ \mathbf{u} \ \cdots \ ]$	κ	$[\ \cdots \ \mathbf{v} \ \cdots \ ] \ 0 \ \mathbf{w}$
$\xi' = \mathcal{T}_\kappa(\xi)$	$[\ \cdots \ \mathbf{u}' \ \cdots \ ]$	0	$[\ \cdots \ \mathbf{v}' \ \cdots \ ] \ 1 \ \mathbf{w}'$

Consider running the capacity-j carrier over ξ and $\mathcal{T}_\kappa(\xi)$ and computing their energies $E_j(\xi)$ and $E_j(\xi')$. Let the corresponding carrier processes be denoted by $\Gamma := (\Gamma_x)_{x \geq 0}$ and $\Gamma' := (\Gamma'_x)_{t \geq 0}$, respectively. Observe that up to 'time' $x-1$, the two carriers go through the equivalent environments $\mathbf{u}$ and $\mathbf{u}'$, so Γ'_{x-1} can be obtained from Γ_{x-1} by adding 1 to all positive colors in the latter carrier. It follows that the contributions to the energies of both carry up to this point are the same.

Next, after inserting $\xi_x = \kappa$ and $\xi'_x = 0$ into these carriers, we get carrier states $\Gamma_x = [\kappa, A, 0 \cdots 0]$ and $\Gamma'_x = [A', 0 \cdots 0]$ for some (possibly empty) positive decreasing sequence A (see Figure 10 left). This only adds 1 to the energy for the carrier Γ. Also note that, since κ is the unique largest color in the system, it sits in the carrier Γ and does not interact with any other incoming balls thereafter. We can think of this as the capacity of the carrier Γ being decreased
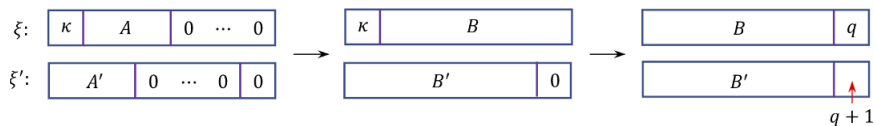
Figure 10. *Two capacity-j carriers over ξ and $\xi' = \mathcal{T}_\kappa(\xi)$. They end up with the same energy.*

to $j-1$ after time x. Then over the interval (x, ∞), the carriers go through the input $[\mathbf{v}0\mathbf{w}]$ and $[\mathbf{v}'1\mathbf{w}']$, respectively.

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

(ii) Let $(\Gamma_x)_{x \geq 0}$ denote the capacity-j carrier process over ξ. We will partition the sites that contain balls of positive colors into j disjoint sets $A_1, \ldots, A_j$ such that if $x \in A_i$ and the energy E_j increases when inserting the ball ξ_x into the carrier Γ_{x-1}, then either x is the rightmost (smallest) element of A_i or there exists a unique $y \in A_i$ such that (y, x) counts as an ascent in A_i. The existence of such subsets $A_1, \ldots, A_j$ implies that

$$R_j(\xi) \geq \sum_{i=1}^{j} \mathrm{NA}(A_i, \xi) \geq E_j(\xi). \tag{417}$$

For this proof, we will consider sites with color zero as having a ball of color zero. We will recursively construct sets $A_1(x), \ldots, A_j(x)$ for $x \geq 0$ as follows. Initially, make all j sets to be empty. Consider the ball at site x with color ξ_x (we may simply call it the 'ball ξ_x') is inserted into the carrier Γ_{x-1}. There are j positions in Γ_{x-1} at which ξ_x can be placed after the insersion, and let $r(x) \in \{1, \ldots, j\}$ denote that position. Note that $r(x) < j$ if and only if $\xi_x > \min \Gamma_x$ if and only if E_j increase by one. Now define $A_1(x), \ldots, A_j(x)$ as follows: For $i = 1, \ldots, j$,

$$\text{If } r(x) < j : \qquad A_i(x) = \begin{cases} A_i(x-1) \cup \{x\} & \text{if } r(x) = i \\ A_i(x-1) & \text{if } r(x) \neq i, \end{cases} \tag{418}$$

$$\text{If } r(x) = j : \qquad A_i(x) = \begin{cases} A_{i-1 \,(\mathrm{mod}\, j)}(x-1) \cup \{x\} & \text{if } i = j \\ A_{i-1 \,(\mathrm{mod}\, j)}(x-1) & \text{if } i \neq j. \end{cases} \tag{419}$$

That is, if the energy E_j increases by inserting the ball ξ_x into the carrier Γ_{x-1}, which occurs exactly when $r(x) < j$, we append x to the set $A_i(x-1)$ where the new ball ξ_x is placed at in Γ_{x-1}. Otherwise, the new ball ξ_x is inserted in position j, and all the other balls are shifted to the left by one, while the ball at position 1 is dropped out. In this case, we first shift the indices of all sets $A_1(x-1), \ldots, A_j(x-1)$ by -1 modulo j, and then append x to the set with index j (previously of index 1).

Then clearly, A_i's are disjoint and partitions $\mathbb{N}$. Moreover, we claim that it has the required properties. Indeed, suppose that the energy E_j increases when inserting the ball ξ_x into the carrier Γ_{x-1} (i.e., $\xi_x > \min \Gamma_{x-1}$). Then ξ_x replaces some ball ξ_y (possibly 0) in Γ_{x-1}. Then necessarily,

$\xi_y < \xi_x$. Moreover, if ξ_x is inserted in the ith position in Γ_{x-1}, then the ball ξ_y it is replacing should also be in the ith position in Γ_{x-1}. By construction, we have $y, x \in A_i$. So (y, x) is an ascent in A_i, as desired.

For the other direction, suppose that $R_j(\xi)$ is achieved by a collection of disjoint sets $A'_1, \cdots, A'_j$ that are 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

$$A''_\ell = \begin{cases} ([1,x] \cap A_i) \cup ((x,\infty) \cap A'_{i^*}) & \text{if } \ell = i \\ ([1,x] \cap A_{i^*}) \cup ((x,\infty) \cap A'_i) & \text{if } \ell = i^* \\ A'_\ell & \text{otherwise.} \end{cases} \tag{420}$$

Then by construction, this new collection of sets $A''_1, \cdots, A''_j$ has at least as many ascents as the A'-sequences do, and the point of disagreement with the A'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 $R_j(\xi) \leq E_j(\xi)$, as desired.

(iii) Let $L_j^{\text{new}} := L_j(\xi')$. 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 ξ, fix a set of k noninterlacing decreasing subsequences whose sum of penalized lengths is the maximum value L_j. We will produce a set of noninterlacing decreasing subsequences in ξ' that have the same sum of penalized lengths. We call the unique ball of color κ in ξ' by simply κ. Suppose κ 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 κ 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 κ 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 κ is in the interval spanned by a decreasing subsequence but does not belong to it, removing κ introduces a gap and so penalizes the length of that sequence by 1. If neither holds, removing κ 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, and so increases its penalized length by 1. If a sequence ends in I and no subsequence 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 holds, then inserting 1 does not change the penalized lengths of any of the subsequences. Then, 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, $L_j^{\text{new}} \geq L_j$, as claimed.

Finally, to show that actually $L_j^{\text{new}} = L_j$, we apply the 'reverse-complement' operation, reversing the order of $\mathbb{Z}$ and the order of the labels. This preserves decreasing subsequences, the noninterlacing 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, $L_j^{\text{new}} \leq L_j$. This shows $L_j = L_j^{\text{new}}$, as desired. □

13. Open questions and final remarks

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

Two-sided limiting shape of the Young diagrams. Many of the known results in scaling limits of invariant Young diagrams of randomized BBS ([LLP20, KL20, KLO18] 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 the proper two-sided limiting shape of the Young diagram as in the case of the Plancherel measure [KKR88] [IO02]? This question is not entirely obvious since the top rows (soliton numbers) obey the laws of large numbers, whereas the top columns (soliton lengths) obey extreme value statistics.

Column length scaling of higher order invariant Young diagrams. The κ-color BBS is known to have κ-tuple of invariant Young diagrams, where the 'higher order' Young diagrams describe the internal degrees of the freedom of the solitons [KL20]. It is our future work to extend the methods and results in the present paper for the first-order Young diagram of the κ-color BBS into higher-order Young diagrams.

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

$$u_t + 6uu_t + u_{xxx} = 0, \tag{421}$$

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 arises from KdV by discretizing space and time:

$$y_k^t + \frac{\delta}{y_{i+1}^t} = \frac{\delta}{y_k^{t+1}} + y_{k+1}^{t+1}. \tag{422}$$

A 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]:

$$U_n^{t+1} = \min\left(1 - U_n^t, \sum_{k=-\infty}^{n-1} (U_k^t - U_k^{t+1})\right), \tag{423}$$

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 [LLP20, KL20, KLO18], 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, suppose we initialize dKDV (422) so that the first n box states are independent Exp(1) random variables and evolve the system until solitons come out. What is 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 questions for dKdV because not everything decomposes into solitons: just like in the usual KdV, there is chaotic 'radiation' left behind.

Competing interest. The authors have no competing interest to declare.

References

[ADOS11] F. Aurzada, H. Döring, M. Ortgiese and M. Scheutzow, 'Moments of recurrence times for Markov chains'.

[AF02] D. Aldous and J. Fill, 'Reversible Markov chains and random walks on graphs', *monograph in preparation*, 2002.

[BDJ99] J. Baik, P. Deift and K. Johansson, 'On the distribution of the length of the longest increasing subsequence of random permutations', *J. Amer. Math. Soc.* **12**(4) (1999), 1119–1178.

[BFPS07] A. Borodin, P. L. Ferrari, M. Prähofer and T. Sasamoto, 'Fluctuation properties of the TASEP with periodic initial configuration', *J. Stat. Phys.* **129**(5–6) (2007), 1055–1080.

[BFS08] A. Borodin, P. L. Ferrari and T. Sasamoto, 'Transition between Airy1 and Airy2 processes and TASEP fluctuations', *Comm. Pure Appl. Math.* **61**(11) (2008), 1603–1629.

[CKST18] D. A. Croydon, T. Kato, M. Sasada and S. Tsujimoto, 'Dynamics of the box-ball system with random initial conditions via Pitman's transformation', Preprint, 2018, arXiv:1806.02147.

[CS19a] D. A. Croydon and M. Sasada, 'Duality between box-ball systems of finite box and/or carrier capacity' Preprint, 2019, arXiv:1905.00189.

[CS19b] D. A. Croydon and M. Sasada, 'Invariant measures for the box-ball system based on stationary Markov chains and periodic Gibbs measures', *J. Math. Phys.* **60**(8) (2019), 083301.

[CS20] D. A. Croydon and M. Sasada, 'Detailed balance and invariant measures for discrete kdv-and Toda-type systems', Preprint, 2020, arXiv:2007.06203.

[DMR94] P. Doukhan, P. Massart and E. Rio, 'The functional central limit theorem for strongly mixing processes', *Ann. I.S.U.P.* **30** (1994), 63–82.

[F⁺18] P. A. Ferrari et al., 'TASEP hydrodynamics using microscopic characteristics', *Probab. Surv.* **15** (2018), 1–27.

[FG18] P. A. Ferrari and D. Gabrielli, 'BBS invariant measures with independent soliton components', Preprint, 2018, arXiv:1812.02437.

[FNRW18] P. A. Ferrari, C. Nguyen, L. Rolla and M. Wang, 'Soliton decomposition of the box-ball system', Preprint, 2018, arXiv:1806.02798.

[FYO00] K. Fukuda, Y. Yamada and M. Okado, 'Energy functions in box ball systems', *Internat. J. Modern Phys. A* **15**(9) (2000), 1379–1392.

[Gre82] C. Greene, 'An extension of Schensted's theorem', in *Young Tableaux in Combinatorics, Invariant Theory, and Algebra* (Elsevier, 1982), 39–50.

[Han06] N. R. Hansen, 'The maximum of a random walk reflected at a general barrier', *Ann. Appl. Probab.* **16**(1) (2006), 15–29.

[HHI⁺01] G. Hatayama, K. Hikami, R. Inoue, A. Kuniba, T. Takagi and T. Tokihiro, 'The $A_M^{(1)}$ automata related to crystals of symmetric tensors', *J. Math. Phys.* **42**(1) (2001), 274–308.

[Hir81] R. Hirota, 'Discrete analogue of a generalized Toda equation', *J. Phys. Soc. Japan* **50**(11) (1981), 3785–3791.

[HJR53] J. L. Hodges Jr. and M. Rosenblatt, 'Recurrence-time moments in random walks', *Pacific J. Math.* **3** (1953), 127–136.

[HKT01] G. Hatayama, A. Kuniba and T. Takagi, 'Factorization of combinatorial R matrices and associated cellular automata', *J. Stat. Phys.* **102**(3–4) (2001), 843–863.

[HR81] J. M. Harrison and M. I. Reiman, 'Reflected Brownian motion on an orthant', *Ann. Probab.* **9**(2) (1981), 302–308.

[IKT12] R. Inoue, A. Kuniba and T. Takagi, 'Integrable structure of box–ball systems: crystal, Bethe ansatz, ultradiscretization and tropical geometry', *J. Phys. A* **45**(7) (2012), 073001.

[IO02] V. Ivanov and G. Olshanski, 'Kerov's central limit theorem for the plancherel measure on young diagrams', in *Symmetric Functions 2001: Surveys of Developments and Perspectives* (Springer, 2002), 93–151.

[KKR88] S. Kerov, A. Kirillov and N. Reshetikhin, 'Combinatorics, Bethe ansatz, and representations of the symmetric group', *J. Math. Sci.* **41**(2) (1988), 916–924.

[KL20] A. Kuniba and H. Lyu, 'Large deviations and one-sided scaling limit of randomized multicolor box-ball system', *J. Stat. Phys.* **178**(1) (2020), 38–74.

[KLO18] A. Kuniba, H. Lyu and M. Okado, 'Randomized box–ball systems, limit shape of rigged configurations and thermodynamic Bethe ansatz', *Nuclear Phys. B* **937** (2018), 240–271.

[Kon20] K. Kondo, 'Dynamics of the multicolor box-ball system with random initial conditions via pitman's transformation', Preprint, 2020, arXiv:2003.12974.

[KOS⁺06] A. Kuniba, M. Okado, R. Sakamoto, T. Takagi and Y. Yamada, 'Crystal interpretation of Kerov–Kirillov–Reshetikhin bijection', *Nuclear Phys. B* **740**(3) (2006), 299–327.

[KW07] W. Kang and R. J. Williams, 'An invariance principle for semimartingale reflecting brownian motions in domains with piecewise smooth boundaries'.

[LLP20] L. Levine, H. Lyu and J. Pike, 'Double jump phase transition in a soliton cellular automaton', *Int. Math. Res. Not.* **2022**(1) (2020), 665–727.

[LP17] D. A. Levin and Y. Peres, *Markov Chains and Mixing Times* vol. 107 (American Mathematical Soc., 2017).

[LPS14] T. Lam, P. Pylyavskyy and R. Sakamoto, 'Rigged configurations and cylindric loop Schur functions', Preprint, 2014, arXiv:1410.4455.

[MT12] S. P. Meyn and R. L. Tweedie, *Markov Chains and Stochastic Stability* (Springer Science & Business Media, 2012).

[Rom15] D. Romik, *The Surprising Mathematics of Longest Increasing Subsequences* vol. 4 (Cambridge University Press, 2015).

[RS10] A. Rohde and C. Strauch, 'Uniform central limit theorems for multidimensional diffusions', Preprint, 2010, arXiv:1010.3604.

[RW88] M. I. Reiman and R. J. Williams, 'A boundary property of semimartingale reflecting Brownian motions', *Probab. Theory Related Fields* **77**(1) (1988), 87–97.

[Sag01] B. E. Sagan, *The Symmetric Group (Graduate Texts in Mathematics)* vol. 203, second edn. (Springer-Verlag, New York, 2001).

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

[TS90] D. Takahashi and J. Satsuma, 'A soliton cellular automaton', *J. Phys. Soc. Japan* **59**(10) (1990), 3514–3519.

[Vei69] A. F. Veinott, 'Discrete dynamic programming with sensitive discount optimality criteria', *Ann. Math. Stat.* **40**(5) (1969), 1635–1660.

[Wil95] R. J. Williams, 'Semimartingale reflecting brownian motions in the orthant', *Stochastic networks*, 1995.

[Wil98] R. J. Williams, 'An invariance principle for semimartingale reflecting brownian motions in an orthant', *Queueing Systems* **30**(1) (1998), 5–25.