Electron. J. Probab. **25** (2020), article no. 78, 1–26.
ISSN: 1083-6489 https://doi.org/10.1214/20-EJP475

BBS invariant measures with independent soliton components

Pablo A. Ferrari* Davide Gabrielli†

Abstract

The Box-Ball System (BBS) is a one-dimensional cellular automaton in the configuration space $\{0,1\}^{\mathbb{Z}}$ introduced by Takahashi and Satsuma [8], who identified conserved quantities called *solitons*. Ferrari, Nguyen, Rolla and Wang [4] map a configuration to a family of *soliton components*, indexed by the soliton sizes $k \geq 1$. Building over this decomposition, we give an explicit construction of a large family of invariant measures for the BBS that are also shift invariant, including Ising-like Markov and Bernoulli product measures. The construction is based on the concatenation of iid excursions of the associated walk trajectory. Each excursion has the property that the law of its k component given the larger components is product of a finite number of geometric distributions with a parameter depending on k. As a consequence, the law of each component of the resulting ball configuration is product of identically distributed geometric random variables, and the components are independent. This last property implies invariance for BBS, as shown by [4].

Keywords: Box-Ball System; soliton components; conservative cellular automata.
AMS MSC 2010: 37B15; 37K40; 60C05.
Submitted to EJP on December 6, 2018, final version accepted on June 8, 2020.
Supersedes arXiv:1812.02437.

1 Introduction

Takahashi and Satsuma [8], referred to as TS in the sequel, introduced the *Box-Ball System* (BBS), a cellular automaton describing the deterministic evolution of a finite number of balls on the infinite lattice $\mathbb{Z}$. A ball configuration η is an element of $\{0,1\}^{\mathbb{Z}}$, where $\eta(z) = 1$ indicates that there is a ball at box $z \in \mathbb{Z}$. A carrier visits successively boxes from left to right picking balls from occupied boxes and depositing one ball, if carried, at the current visited box, if empty. We denote by $T\eta$ the configuration obtained after the carrier has visited all boxes and $T^t\eta$ the configuration obtained after iterating this procedure t times, for t a positive integer.

*University of Buenos Aires, Argentina.
E-mail: pferrari@dm.uba.ar, http://mate.dm.uba.ar/~pferrari/
†University of L'Aquila, Italy.
E-mail: davide.gabrielli@univaq.it, http://people.disim.univaq.it/~gabriell/

The evolution of the Box-Ball dynamics is shown by the following example:

$$\begin{array}{rl} \eta & 01101011010001111010000 \\ \text{Carrier Load} & 01212123232101234343210 \\ T\eta & 00010100101110000101111 \end{array} \tag{1.1}$$

The configurations η and $T\eta$ are identically 0 outside the finite window shown. In the second line we write the number of balls which are transported by the carrier; we assume that the carrier is always empty outside of the window shown in the picture.

TS show the existence of conserved quantities in the BBS called *solitons*, see also Levine, Lyu and Pike [6]. In the absence of other solitons, a k-soliton consists of k successive occupied boxes followed by k successive empty boxes. In this case, the k-soliton travels at speed k, because the carrier picks the k balls and deposits them in the k empty boxes of the soliton. Solitons with different speeds "collide" but still can be identified at collisions, see §2. More generally, a k-soliton consists always of k occupied boxes and k empty boxes which are however not necessarily consecutive; different solitons occupy disjoint sets of boxes.

A configuration of balls can be mapped to a walk trajectory that jumps one unit up at occupied boxes and one unit down at empty boxes [1] [4]. The down *records* of the path are the new minimal values and the *excursions* are the pieces of configuration between two consecutive records. Walks coming from configurations with density of balls less than half have positive density of records, hence any box is either a record or belongs to a finite excursion.

The TS identification of solitons can be performed in configurations with infinitely many balls and finite excursions by applying it to each excursion. Ferrari, Nguyen, Rolla and Wang, referred to as FNRW in the sequel, introduce a soliton decomposition of any ball configuration η based on the concept of k-slots. The *k-slots* are boxes determined by η with the property that any k-soliton is strictly contained between two successive k-slots; k-slots have integer labels with record 0 being k-slot 0 for all k. We say that each k-soliton is *appended* to the nearest k-slot to its left. The decomposition of η consists of an infinite array $(\zeta_k(j))$ with k natural and j integer, where $\zeta_k(j)$ represents the number of k-solitons appended to the k-slot number j.

Since the soliton decomposition is performed independently inside each excursion, it is convenient to introduce the finite array of components associated to one single excursion. This combinatorial object is called a *slot diagram*. The components of an infinite configuration of balls with a record at the origin are obtained by suitably joining the slot diagrams of its excursions. We use Palm theory to associate measures conditioned to have a record at the origin with shift stationary measures.

We show that the soliton components of the product of Bernoulli measures of parameter $\lambda < 1/2$ as seen from a record have the following properties. There is a function $q_k(\lambda)$ such that the variables $\zeta_k(i)$ form a mutually independent array, such that its marginal $\zeta_k(i)$ has geometric distribution with parameter $1 - q_k(\lambda)$ for all k-slot label i. We construct many other measures with the same property, including stationary Markov chains with state space $\{0, 1\}$ and ball density less than half.

The independence of components combined with a result of FNRW imply that the measures whose decomposition is described in the previous paragraph are invariant for BBS. In the case of Bernoulli and Markov, these facts were proven directly by Croydon, Kato, Sasada and Tsujimoto [1], using reversibility of the carrier process illustrated in (1.1); see also [4].

To prove the results just described we introduce two families of probability measures on the set of finite excursions. The first family denoted ν_α is indexed by parameters

$\alpha_k \in [0,1)$ satisfying a summability condition. Under ν_α each excursion has weight $\prod_{k\geq 1} \alpha_k^{n_k}$, where n_k is the number of k-solitons in the excursion. The second family, called φ_q is indexed by parameters $q_k \in [0,1)$, $k \geq 1$, also satisfying some summability condition. Conditioning on the components $m > k$ of the slot diagram of the random excursion with law φ_q, the distribution of the k-component is a product of s_k geometric distributions with mean $q_k/(1-q_k)$, where s_k is the number of k-slots in the excursion; s_k is determined by the m-components, for m bigger than k. Theorem 3.1 shows a bijection between those two families with an explicit relation between α and q. Under suitable assumptions, the resulting random excursion has finite mean length.

We then consider a sequence of i.i.d. excursions with law ν_α and finite expected excursion length and construct a ball configuration η by putting a record at the origin and concatenating the excursions separated by records. We show that the components of η are independent and that the k-component consists of i.i.d. Geometric random variables with mean $q_k/(1-q_k)$, where q is a function of α. Using the inverse-Palm transformation, we obtain a shift invariant and T-invariant measure. The T-invariance is deduced from the independence of the components, as explained before. We show that product of Bernoulli and Ising-like measures conditioned to have a record at the origin have i.i.d. excursions with distribution ν_α for suitable α, which in turn implies that they have independent components and are T-invariant.

The paper is organized as follows.

In Section 2 we introduce notation, illustrate the soliton decomposition, define the slot diagrams and show that they are in bijection with excursions.

In Section 3 we introduce the two aforementioned families of probability measures on the set of excursions parameterized by an infinite collection of parameters and show in Theorem 3.1 that these are two different parametrization of the same family of probability measures with a non trivial relationship between the two families of parameters.

In Section 4 we obtain a T-invariant measure μ_α via the concatenation of i.i.d. random excursions with distribution ν_α. This is obtained in Theorem 4.5 by the combination of Theorem 4.3 (independence of components of μ_α) and Theorem 4.4 (independence of components implies T-invariance), a result of [4]. These are the remaining main results of the paper.

2 Excursions, solitons and slot diagrams

In this section we define *excursions*, describe a variant of the Takahashi-Satsuma Algorithm of [8] to identify solitons in excursions and call a *slot diagram* the FNRW soliton decomposition of an excursion.

A configuration of balls is an element $\eta \in \{0,1\}^{\mathbb{Z}}$, where for each *box* $y \in \mathbb{Z}$, $\eta(y) = 1$ means that there is a *ball* at box y, otherwise $\eta(y) = 0$ means y is empty. In this section we consider configurations with a finite number of balls.

Map a ball configuration η to a walk $\xi = W\eta \in \mathbb{Z}^{\mathbb{Z}}$ defined up to a global additive constant by

$$\xi(z) - \xi(z-1) = 2\eta(z) - 1. \tag{2.1}$$

We fix the constant by choosing $\xi(0) = 0$. The configuration of balls is completely determined by the walk and if $\xi = W\eta$ we write also $\eta = W^{-1}\xi$.

We call $z \in \mathbb{Z}$ a *record* for ξ if $\xi(z) < \xi(z')$ for any $z' < z$. This depends just on η as $\xi(z) - \xi(z') = \sum_{y=z'+1}^{z}(2\eta(y)-1)$ and we can therefore say equivalently that z is a record for the configuration η.

Excursions We introduce the set $\mathcal{E}$ of *finite excursions* (sometimes called *soft excursions* in the literature). An element $\varepsilon \in \mathcal{E}$ is a finite walk which starts and ends at zero, it is always non-negative and it has length $2n(\varepsilon)$. More precisely $\varepsilon = \big(\varepsilon(0), \dots, \varepsilon(2n(\varepsilon))\big)$ with the constraints $|\varepsilon(z) - \varepsilon(z-1)| = 1$, $\varepsilon(z) \geq 0$ for $0 \leq z \leq 2n(\varepsilon)$ and $\varepsilon(0) = \varepsilon(2n(\varepsilon)) = 0$. The empty excursion $\emptyset$ is also an element of $\mathcal{E}$ with $n(\emptyset) = 0$. We call $\mathcal{E}_n$ the set of finite excursions of length $2n$, hence $\mathcal{E} = \cup_{n=0}^{+\infty} \mathcal{E}_n$. It is well known [7] that the number of excursions of length $2n$ is given by

$$|\mathcal{E}_n| = \frac{1}{n+1}\binom{2n}{n}; \tag{2.2}$$

the right hand side is the Catalan number C_n.

The underlying configuration of balls of an excursion ε is called $W^{-1}\varepsilon$ and is defined by

$$W^{-1}\varepsilon(z) := \frac{\varepsilon(z) - \varepsilon(z-1) + 1}{2}, \qquad z = 1, \dots, 2n(\varepsilon).$$

This is a configuration of balls restricted to the interval $[1, 2n(\varepsilon)]$ but we can naturally extend it to a configuration on the whole axis $\mathbb{Z}$ just considering empty all the remaining boxes. This corresponds to extending the excursion to an infinite walk by adding to the left and to the right downward oriented steps.

We use the same notation both for configurations of balls/walks restricted to a finite interval and for configurations of balls/walks on the whole $\mathbb{Z}$ axis. The exact meaning will be clear from the context. We call an excursion both the walk ε and the corresponding configuration of balls $W^{-1}\varepsilon$, since they are bijectively related.

Takahashi-Satsuma Identification of solitons We describe a variant of the Takahashi-Satsuma algorithm [8] to identify the solitons of a finite ball configuration η. The empty configuration $\eta(z) \equiv 0$ has no solitons. Assume η is nonempty. A *run* of η is any segment $[z, y]$ with $-\infty \leq z \leq y \leq \infty$ such that $\eta(z) = \eta(z')$, for $z' \in [z, y]$, $\eta(z-1) \neq \eta(z)$ if $z > -\infty$ and $\eta(y) \neq \eta(y+1)$ if $y < \infty$. The ball configuration underlying an excursion (considered on the whole lattice) has two semi-infinite runs and a finite number of finite runs. The algorithm is the following:

If there are finite runs in the configuration, do:

1. Let k be the size of the smallest run in the configuration. Select the leftmost run of size k. Set the restriction of η to the k boxes of this run and the first k boxes of the successive run as a k-soliton.
2. Ignore the boxes belonging to already identified solitons, update the runs of the remaining configuration and go to 1.

For a k-soliton γ we call *support* of γ, denoted by $\{\gamma\} \subset \mathbb{Z}$, the union of two sets of boxes: the *head* $\{h_0(\gamma), \dots, h_{k-1}(\gamma)\}$ and the *tail* $\{t_0(\gamma), \dots, t_{k-1}(\gamma)\}$, satisfying $\eta(h_i) = 1$ and $\eta(t_i) = 0$ and $h_i(\gamma) < h_{i+1}(\gamma)$, $t_i(\gamma) < t_{i+1}(\gamma)$ for $i = 0, \dots, k-2$. Either $h_i(\gamma) < t_j(\gamma)$ for all i, j or $t_j(\gamma) < h_i(\gamma)$ for all i, j. We denote by $\Gamma_k\eta$ the set of k-solitons of η. When η has infinitely many records to the right and left of the origin, every box in $\mathbb{Z}$ is either a record or belongs to $\{\gamma\}$ for some k-soliton γ, for some $k \geq 1$.

An example of the application of this algorithm is illustrated in Fig. 1. The final decomposition into solitons is illustrated in Fig. 2

Slots Given an excursion ε, a box z is a *k-slot* if either z is Record 0, i.e. the record assuming the value 0 that is $z = 0$, or $z \in \{h_i(\gamma), t_i(\gamma)\}$ for some $i \geq k$, some $\gamma \in \Gamma_m\eta$ for some $m > k$. Let $S_k\eta$ be the set of k-slots of η. We have $S_{k+1}\eta \subseteq S_k\eta$.

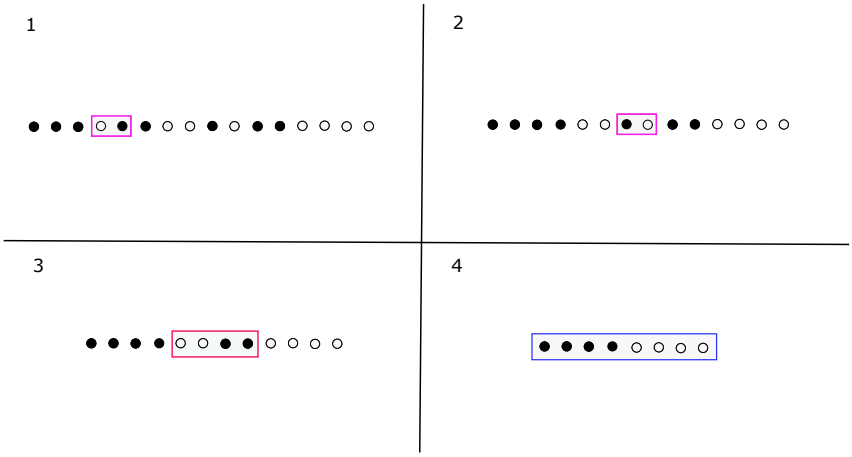

Figure 1: The Takahashi-Satsuma algorithm applied to the finite configuration in quadrant 1 which is indeed an excursion. Identified solitons are surrounded by rectangles of different colors (violet for 1-solitons, red for 2-solitons, blue for 4-solitons). The algorithm stops after 4 iterations.

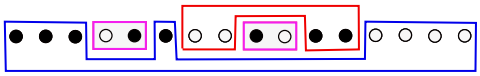

Figure 2: The final decomposition into solitons of the configuration of Fig. 1. Balls and boxes belonging to the same soliton are surrounded by colored lines. The lines are violet for 1-solitons, red for 2-solitons, blue for 4-solitons.

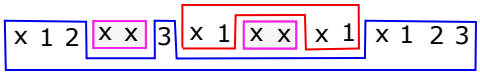

Figure 3: Slots associated to our sample configuration of Fig. 1. To each box we associate the number of the maximal slot. The symbol $\times$ means that the box is not a slot for any $k \geq 1$. A box with number m is a k-slot for each $k \leq m$.

Enumerate the k-slots setting $\mathtt{s}_k(\eta, 0) := 0$, that is, k-slot 0 is at record 0 for all k, and

$$\mathtt{s}_k(\eta, j) := \text{ position of the } j\text{-th } k\text{-slot, counting from } k\text{-slot } 0. \tag{2.3}$$

We show in Figures 3 and 4 an example of identifications of the slots using the sample configuration η of Fig. 1.

We get all slots together in Fig. 5.

Soliton decomposition of ball configurations [4] We say that a k-soliton γ is *appended* to k-slot j of η if its support is strictly included in the open integer interval with extremes in the k-slots j and $j+1$:

$$\{\gamma\} \subset \big(\mathtt{s}_k(\eta, j), \mathtt{s}_k(\eta, j+1)\big). \tag{2.4}$$

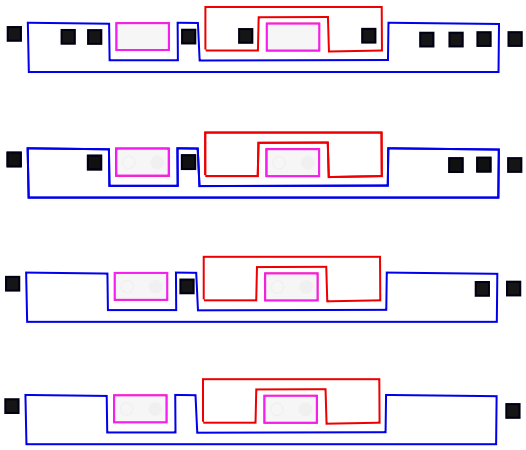

Figure 4: From the top to the bottom we represents respectively the sets $S_1\eta$, $S_2\eta$, $S_3\eta$ and $S_4\eta$. The configuration η is our sample configuration of Fig. 1. Boxes belonging to the sets are marked by a ■. We marked also the slots associated to the records on the left and on the right of the finite configuration (which is indeed an excursion). The origin is the leftmost black square. Numbers to the slots are given starting counting from this slot.

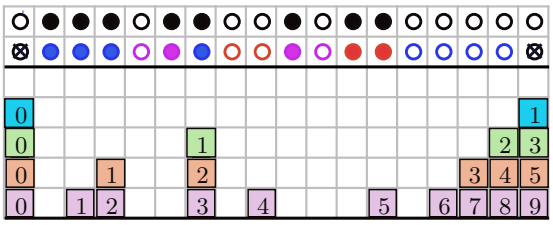

Figure 5: Slot enumeration. The configuration η is our sample configuration of Fig. 1, an excursion between two records. In the second line the solitons have been identified and colored: a blue 4-soliton, a red 2-soliton and two purple 1-solitons. Below, colored square boxes identify k-slots for $k = 1, 2, 3, 4$. Records are k-slots for all k but we have only depicted until $k = 4$ as there are no solitons bigger than 4 and hence no slots bigger than 4 besides the records. There is a 4-soliton appended at 4-slot 0, a 2-soliton appended at 2-slot 2 and two 1-solitons appended at 1-slots 2 and 4, respectively. The slots associated to the record located to the right of the excursion do not strictly belong to the excursion; they are still depicted because they are needed to identify the solitons appended to the slots to the left of it.

Any finite number of k-solitons may be appended to a single k-slot. Define

$$\zeta_k(j) := \#\{\gamma \in \Gamma_k\eta : \gamma \text{ is appended to } k\text{-slot } j\}. \tag{2.5}$$

Consider the example of Fig. 4. Starting from the bottom we have that the blue 4-soliton is between $\mathtt{s}_4(\eta, 0)$ and $\mathtt{s}_4(\eta, 1)$ so that it is appended to the 4-slot number 0 and $\zeta_4(0) = 1$; the red 2-soliton is between $\mathtt{s}_2(\eta, 2)$ and $\mathtt{s}_2(\eta, 3)$ so that it is appended to the 2-slot number 2 and $\zeta_2(2) = 1$; the violet 2-solitons are respectively between $\mathtt{s}_1(\eta, 2)$ and $\mathtt{s}_1(\eta, 3)$ and $\mathtt{s}_1(\eta, 4)$ and $\mathtt{s}_1(\eta, 5)$ so that the leftmost 1-soliton is appended to the 1-slot number 2 while the rightmost 1-soliton is appended to the 1-slot number 4 and therefore

we have $\zeta_1(2) = 1$ and $\zeta_1(4) = 1$. All the remaining ζ's are identically zero. See also Fig. 5.

Slot diagrams A slot diagram is a combinatorial object which encodes the components of a single excursion.

We start giving a formal definition. A *Slot Diagram* is a family $x = (x_k)_{k\geq 1}$ of vectors $x_k = (x_k(0), \dots, x_k(s_k - 1))$ with $s_k \in \mathbb{N}$ and $x_k(j) \in \mathbb{Z}_{\geq 0}$, satisfying the following conditions: denoting by $|x_k| := x_k(0) + \cdots + x_k(s_k - 1)$, we have

$$1) \qquad M(x) := \max\{k : x_k(0) > 0\} < \infty\,, \tag{2.6}$$

$$2) \qquad s_\ell = 1, \text{ for } \ell \geq M(x) \text{ and } x_\ell(0) = 0 \text{ for } \ell > M(x)\,, \tag{2.7}$$

$$3) \qquad s_k = 1 + \sum_{\ell > k} 2(\ell - k)|x_\ell|\,. \tag{2.8}$$

The complete structure of a slot diagram is determined by the finite collection of vectors $(x_k)_{1\leq k\leq M}$ but for notational convenience we consider also the indices $k > M = M(x)$. An example of a slot diagram is the following

$$\begin{array}{ccc} k & \mapsto & x_k \\ 4 & \mapsto & (1) \\ 3 & \mapsto & (0,0,0) \\ 2 & \mapsto & (0,0,1,0,0) \\ 1 & \mapsto & (0,0,1,0,1,0,0,0,0) \end{array} \tag{2.9}$$

In this case we have $M = 4$, $s_4 = 1$, $s_3 = 3$, $s_2 = 5$ and $s_1 = 9$. For any $k > 4$ we have $s_k = 1$ and $x_k(0) = 0$ and therefore the slot diagram is completely determined by the finite diagram (2.9).

Let $\mathcal{S}$ be the set of slot diagrams. We have that $\mathcal{E}$ is in bijection with $\mathcal{S}$ so that a slot diagram completely codifies an excursion. We now construct the map $\varepsilon \mapsto x[\varepsilon]$ and its inverse $x \mapsto \varepsilon[x]$ (see [4] and [3] for more details).

Construction of $x[\varepsilon]$ Consider an excursion ε. If the excursion is empty then the slot diagram is defined as $s_k \equiv 1$ and $x_k(0) \equiv 0$. If ε is not empty, then let M be the maximal soliton size in ε and define $s_\ell = 1$ for $\ell \geq M$, $x_\ell(0) = 0$ for $\ell > M$ and set $x_M(0) =$ number of M-solitons in the excursion. Assume we have set $x_{k+1}, \dots, x_M$. Use (2.8) to define the number of k-slots s_k and set $x_k(j) =$ number of k-solitons appended to k-slot j in the excursion. Iterate for $k = M - 1, \dots, 1$.

In short, considering the excursion ε as an infinite walk we have that $s_k - 1$ is the number of k-slots of the excursion which are not records and $x_k(j)$ is the number of k-solitons appended to the k-slot number j. For example (2.9) is the slot diagram associated to the excursion corresponding to the ball configuration in Fig. 1 and Fig. 5.

Construction of $\varepsilon[x]$ Given a configuration η with no ℓ-solitons for $\ell < k$, define $I_{k,j}$ the operator that insert a k-soliton at k-slot j of η, as follows. Denote by $u = \mathtt{s}_k(\eta, j)$ the position of k-slot j in η and

$$I_{k,j}\eta(z) = \begin{cases} \eta(z) & \text{if } z \leq u \\ 1 - \eta(u) & \text{if } u < z \leq u + k \\ \eta(u) & \text{if } u + k < z \leq u + 2k \\ \eta(z - 2k) & \text{if } u + 2k < z. \end{cases} \tag{2.10}$$

Denote by $I_{k,j}^n$ the n-th iteration of $I_{k,j}$, which corresponds to insert n k-solitons one after the other on the same slot j. When $n=0$ we just have the identity, meaning that no k-soliton is inserted at slot j.

Denoting $M := M(x)$, define

$$\begin{aligned} &\eta_\ell \equiv 0 \quad \text{for } \ell > M, \quad \text{and iteratively,} \\ &\eta_k := I_{k,0}^{x_k(0)} \dots I_{k,s_k-1}^{x_k(s_k-1)} \eta_{k+1}, \quad \text{for } k = M, \dots, 1. \\ &\varepsilon[x] := W\eta_1. \end{aligned} \tag{2.11}$$

Observe that the number n_k of k-solitons in the excursion $\varepsilon[x]$ coincides with the sum over j of $x_k(j)$:

$$n_k(\varepsilon[x]) = \sum_{j=0}^{s_k-1} x_k(j) = |x_k|. \tag{2.12}$$

Example. Consider the following slot diagram x:

$$\begin{aligned} x_\ell &= (0), \quad \text{for } \ell > 3 \\ x_3 &= (2) \\ x_2 &= (0,0,1,0,0) \\ x_1 &= (3,0,4,1,0,0,0,0,2,0,1) \end{aligned} \tag{2.13}$$

that is, $M=3$, $s_k = 1$ for $k \geq 3$, $s_2 = 5$ and $s_1 = 11$.

In this example the algorithm works as follows. Active k-slots are red and k-solitons being appended at each step are blue.

0 (record 0 = k-slot 0 for all k)
0111000111000 (attach 2 3-solitons to 3-slot 0) $I_{3,0}^2$
0111000110011100 0 (attach 1 2-soliton to 2-slot 2) $I_{2,2}^1$
0101010111000110011100 0 (attach 3 1-solitons to 1-slot 0) $I_{1,0}^3$
0101010111010101010001100111000 (attach 4 1-solitons to 1-slot 2) $I_{1,2}^4$
0101010111010101010010011001110 00 (attach 1 1-soliton to 1-slot 3) $I_{1,3}^1$
0101010111010101010010011001110101000 (attach 2 1-solitons to 1-slot 8) $I_{1,8}^2$
0101010111010101010010011001110101000 10 (attach 1 1-soliton to 1-slot 10) $I_{1,10}^1$

The resulting excursion is given by

$$\begin{aligned} \varepsilon[x] &= W\left(I_{1,10}^1 I_{1,8}^2 I_{1,3}^1 I_{1,2}^4 I_{1,0}^3 I_{2,2}^1 I_{3,0}^2 \eta_4\right) \\ &= W\left(\dots 10101011101010101010001100111010100010 \dots\right) \end{aligned}$$

where the dots represent records and we have painted blue, green and red the 1-, 2- and 3-solitons, respectively. Record 0 is the dot preceding the leftmost 1 and record 1 is the dot following the rightmost 0. Here we start with the empty excursion η_4 because $M=3$.

3 Random excursions

We introduce two natural families of probability measures on the set of excursions $\mathcal{E}$ depending on two collections of parameters α and q. The main result of this section is that the two families coincide with a non trivial relationship between the parameters.

For $p \in (0,1]$ we say that a random variable Y is Geometric(p) when

$$P(Y=j) = p(1-p)^j, \ j \geq 0; \quad EY = \frac{1-p}{p}. \tag{3.1}$$

with the convention $0^0 = 1$.

3.1 Probability measures on excursions

First family For each excursion $\varepsilon \in \mathcal{E}$ define

$$n_k(\varepsilon) := \text{number of } k\text{-solitons in } \varepsilon, \tag{3.2}$$

where this number is given by the Takahashi-Satsuma algorithm in §2 applied to ε.

Let $\alpha = (\alpha_k)_{k\geq 1}$ be a family of parameters with $\alpha_k \in [0,1)$, define

$$Z_\alpha := \textstyle\sum_{\varepsilon\in\mathcal{E}} \prod_{k\geq 1} \alpha_k^{n_k(\varepsilon)} \tag{3.3}$$

and call

$$\mathcal{A} := \{\alpha \,:\, Z_\alpha < +\infty\}\,. \tag{3.4}$$

For $\alpha \in \mathcal{A}$ define the measure ν_α on $\mathcal{E}$ by

$$\nu_\alpha(\varepsilon) := \frac{1}{Z_\alpha} \prod_{k\geq 1} \alpha_k^{n_k(\varepsilon)}\,, \tag{3.5}$$

here again we use the convention $0^0 = 1$ so that if $\alpha_k = 0$ then the measure ν_α gives full measure to excursions without k-solitons. Note that by (2.12) we can write (3.5) in terms of the slot diagram of ε by

$$\nu_\alpha(\varepsilon) = \frac{1}{Z_\alpha} \prod_{k\geq 1} \alpha_k^{|x_k[\varepsilon]|}\,. \tag{3.6}$$

We denote the mean number of k-solitons per excursion by

$$\rho_k(\alpha) := \sum_{\varepsilon\in\mathcal{E}} n_k(\varepsilon)\, \nu_\alpha(\varepsilon), \tag{3.7}$$

and therefore the mean excursion size is

$$\sum_{k\geq 1} 2k\, \rho_k(\alpha) = \frac{1}{Z_\alpha} \sum_{\varepsilon\in\mathcal{E}} \left[\Big(\sum_{j\geq 1} 2j\, n_j(\varepsilon) \Big) \prod_{k\geq 1} \alpha_k^{n_k(\varepsilon)} \right]\,. \tag{3.8}$$

We call $\mathcal{A}^+$ the set of α such that the mean excursion size under ν_α is finite:

$$\mathcal{A}^+ := \Big\{ \alpha \,:\, \textstyle\sum_{k\geq 1} 2k\rho_k(\alpha) < +\infty \Big\}\,. \tag{3.9}$$

By definition we have $\mathcal{A}^+ \subseteq \mathcal{A}$.

Second family Let $q = (q_k)_{k\geq 1}$ be a family of parameters with $q_k \in [0,1)$ and introduce the sets

$$\mathcal{Q} := \Big\{ q \,:\, \textstyle\sum_{k\geq 1} q_k < +\infty \Big\}\,, \tag{3.10}$$

$$\mathcal{Q}^+ := \Big\{ q \,:\, \textstyle\sum_{k\geq 1} k q_k < +\infty \Big\}\,. \tag{3.11}$$

For $q \in \mathcal{Q}$ consider the probability measure φ_q on $\mathcal{E}$ defined by

$$\varphi_q(\varepsilon) := \textstyle\prod_{k\geq 1} q_k^{|x_k[\varepsilon]|} (1-q_k)^{s_k(x[\varepsilon])}. \tag{3.12}$$

The fact that (3.12) is a probability measure on $\mathcal{E}$ when $q \in \mathcal{Q}$ is a consequence of the following argument. Writing $x = x[\varepsilon]$ and denoting $x_k^\infty = (x_k, x_{k+1}, \dots)$, formula (3.12) is equivalent to the following three formulas (with the convention $q_0 = 1$ to take care of

the empty excursion), which give a recipe to construct/simulate the random slot diagram of an excursion with distribution (3.12),

$$\varphi_q\left(M(x)=m\right)=q_m\prod_{\ell>m}\big(1-q_\ell\big),\quad m\geq 0, \tag{3.13}$$

$$\varphi_q\big(x_m(0)\big|M(x)=m\big)=q_m^{|x_m(0)|-1}(1-q_m), \tag{3.14}$$

$$\varphi_q\big(x_k\big|x_{k+1}^\infty\big)=q_k^{|x_k|}(1-q_k)^{s_k(x)}, \tag{3.15}$$

where we abuse notation writing x_m as "the set of slot diagrams y such that $y_m=x_m$", and so on. Then, to construct a slot diagram with law φ_q, first choose a maximal soliton-size m with probability (3.13). This is a probability on $\mathbb{Z}_{\geq 0}$ since $q\in\mathcal{Q}$. Then use (3.14) to determine the number of maximal solitons $x_m(0)$ (a Geometric$(1-q_m)$ random variable conditioned to be strictly positive). Finally we use (3.15) to construct iteratively the lower components. Under the measure φ_q and conditioned on x_{k+1}^∞, the variables $(x_k(0),\dots,x_k(s_k-1))$ are i.i.d. Geometric$(1-q_k)$.

3.2 Equivalence of measures

Given the parameters α and q we define the transformation $q=q(\alpha)$ by

$$q_1:=\alpha_1\quad\text{ and }\quad q_k:=\frac{\alpha_k}{\prod_{j=1}^{k-1}(1-q_j)^{2(k-j)}}\,,\ \text{for } k\geq 2, \tag{3.16}$$

and $\alpha=\alpha(q)$ by

$$\alpha_k:=q_k\prod_{\ell=1}^{k-1}(1-q_\ell)^{2(k-\ell)}\,,\quad \text{for } k\geq 1. \tag{3.17}$$

Theorem 3.1 (Equivalence of measures)**.** *The maps $\alpha\to q(\alpha)$ and $q\to\alpha(q)$ defined by equations* (3.16) *and* (3.17) *are the inverse of each other. Moreover we have*

$$\alpha\in\mathcal{A}\text{ if and only if }q\in\mathcal{Q}, \tag{3.18}$$

$$\alpha\in\mathcal{A}^+\text{ if and only if }q\in\mathcal{Q}^+. \tag{3.19}$$

In particular, the transformations (3.16)–(3.17) *map bijectively $\mathcal{A}\leftrightarrow\mathcal{Q}$ and $\mathcal{A}^+\leftrightarrow\mathcal{Q}^+$. Furthermore, if $\alpha\in\mathcal{A}$, we have*

$$\nu_\alpha=\varphi_q, \tag{3.20}$$

defined in (3.5) *and* (3.12)*.*

The remaining of this subsection is devoted to the proof of Theorem 3.1. We start with some notation and preliminary results. In the next three lemmas we compute the partition function Z_α.

Given a slot diagram x we define the *translation* τ by

$$\big(\tau x\big)_k=x_{k+1},\qquad k=1,2,\dots.$$

We have that τx is again a slot diagram. For $\alpha=(\alpha_k)_{k\in\mathbb{N}}\in\mathcal{A}$ we define another "translation" operator θ by

$$\big(\theta\alpha\big)_k:=\frac{\alpha_{k+1}}{(1-\alpha_1)^{2k}}\,,\qquad k=1,2,\dots\,, \tag{3.21}$$

so that we can write (3.16) as

$$q_k=\big(\theta^{k-1}\alpha\big)_1\,,\quad k\geq 1, \tag{3.22}$$

with the convention $\theta^0\alpha = \alpha$. We define and compute some restricted partition functions. We call $Z_\alpha(x_k^\infty)$ the sum of the weights $\prod_{k\geq 1}\alpha_k^{n_k}$ over all the excursions ε such that $x_k^\infty[\varepsilon] = x_k^\infty$. We have

$$Z_\alpha(x_k^\infty) := \prod_{n\geq k}\alpha_n^{|x_n|}\sum_{\{y:y_k^\infty = x_k^\infty\}}\prod_{\ell=1}^{k-1}\alpha_\ell^{|y_\ell|}, \qquad Z_\alpha(x) = \prod_{n\geq 1}\alpha_n^{|x_n|}\,, \tag{3.23}$$

where we sum the weights of the slot diagrams y which are compatible with x_k^∞. These partition functions satisfy a useful recurrence:

Lemma 3.2 (Iterating tail partition functions)**.** *We have*

$$Z_\alpha(x_k^\infty) = \frac{Z_{\theta\alpha}((\tau x)_{k-1}^\infty)}{(1-\alpha_1)}\,, \qquad k>1\,. \tag{3.24}$$

Proof. From (3.23) we have

$$Z_\alpha(x_k^\infty) = \prod_{i=k}^{\infty}\alpha_i^{|x_i|}\sum_{\left\{y_2^\infty : y_k^\infty = x_k^\infty\right\}}\prod_{j=2}^{k-1}\alpha_j^{|y_j|}\sum_{y_1\in\mathbb{Z}_{\geq 0}^{s_1}}\alpha_1^{|y_1|}\,, \tag{3.25}$$

where $y_k^\ell = (y_k, y_{k+1}, \dots, y_\ell)$. Note that the last sum gives $(1-\alpha_1)^{-s_1}$. If for $k<\ell<m$ we write $y_k^\ell y_{\ell+1}^m = y_k^m$, then

$$s_1 = s_1\left(y_2^{k-1}x_k^{+\infty}\right) = 1 + 2\sum_{i=2}^{k-1}(i-1)|y_i| + 2\sum_{i=k}^{+\infty}(i-1)|x_i|\,. \tag{3.26}$$

Substituting this in (3.25) we get

$$\begin{aligned} Z_\alpha(x_k^\infty) &= \frac{1}{(1-\alpha_1)}\prod_{i=k}^{\infty}\left[\frac{\alpha_i}{(1-\alpha_1)^{2(i-1)}}\right]^{|x_i|}\sum_{\left\{y_2^\infty : y_k^{+\infty} = x_k^{+\infty}\right\}}\prod_{j=2}^{k-1}\left[\frac{\alpha_j}{(1-\alpha_1)^{2(j-1)}}\right]^{|y_j|} \\ &= \frac{1}{(1-\alpha_1)}\prod_{i=k-1}^{+\infty}(\theta\alpha)_i^{|(\tau x)_i|}\sum_{\left\{y_1^\infty : y_{k-1}^\infty = (\tau x)_{k-1}^\infty\right\}}\prod_{j=1}^{k-2}(\theta\alpha)_j^{|y_j|}\,, \end{aligned} \tag{3.27}$$

which gives (3.24). □

We now compute $Z_\alpha(x_k^\infty)$.

Lemma 3.3 (Tail partition function)**.** *For any fixed $k\geq 2$ and x_k^∞ we have*

$$Z_\alpha(x_k^\infty) = \left[\prod_{i=0}^{k-2}\left(\frac{1}{(1-(\theta^i\alpha)_1)}\right)\right]\left[\prod_{j=k}^{+\infty}\left(\theta^{k-1}\alpha\right)_{j-k+1}^{|x_j|}\right]. \tag{3.28}$$

Proof. Iterating $k-1$ times the recursion (3.24) we have

$$Z_\alpha(x_k^\infty) = \left[\prod_{i=0}^{k-2}\left(\frac{1}{(1-(\theta^i\alpha)_1)}\right)\right]Z_{\theta^{k-1}\alpha}\left(\left(\tau^{k-1}x\right)_1^{+\infty}\right).$$

The statement is now obtained observing that for any x we have

$$Z_\alpha(x_1^{+\infty}) = \prod_{i=1}^{\infty}\alpha_i^{|x_i|}\,,$$

because the complete slot diagram is fixed so that there are no sums to be done. □

We now compute the partition function Z_α. Denoting by Z_α^m the weight of the excursions having m as maximum soliton size, we have

$$Z_\alpha^m := \sum_{x:M(x)=m} Z_\alpha(x)\,, \qquad m \geq 0\,. \tag{3.29}$$

Lemma 3.4 (Finiteness of the partition function)**.** *The partition function Z_α is finite if and only if*

$$\sum_{m=0}^{+\infty} (\theta^m \alpha)_1 < \infty\,. \tag{3.30}$$

Furthermore,

$$Z_\alpha^m = \left(\theta^{m-1}\alpha\right)_1 \prod_{j=0}^{m-1} \left(\frac{1}{1-(\theta^j\alpha)_1}\right), \qquad m \geq 1, \tag{3.31}$$

and

$$Z_\alpha = 1 + \sum_{m=1}^{+\infty} \left(\theta^{m-1}\alpha\right)_1 \prod_{j=0}^{m-1} \left(\frac{1}{1-(\theta^j\alpha)_1}\right). \tag{3.32}$$

Proof. Since the weight of the empty excursion is 1, we have $Z_\alpha^0 = 1$ and (3.32) is obtained from (3.31) from the relation $Z_\alpha = \sum_{m=0}^{+\infty} Z_\alpha^m$. To show (3.31) we sum over all possible slot diagrams

$$Z_\alpha^m = \sum_{x_m(0)=1}^{+\infty} \alpha_m^{x_m(0)} \sum_{\left\{x_{m-1}\in\mathbb{Z}_{\geq 0}^{s_{m-1}}\right\}} \alpha_{m-1}^{|x_{m-1}|} \cdots \sum_{\left\{x_1\in\mathbb{Z}_{\geq 0}^{s_1}\right\}} \alpha_1^{|x_1|}\,, \tag{3.33}$$

where $s_k = s_k(x_{k+1}^{+\infty})$ given by (2.7)–(2.8) with $x_k(0) = 0$ for any $k > m$. Note that $x_m(0)$ has to be summed from 1 up to $+\infty$ since at level m there must be at least one soliton. All the other variables are summed from 0 to $+\infty$. Sum on x_1, use (2.7)–(2.8), change name to the summed variables and iterate to obtain

$$Z_\alpha^m = \frac{1}{1-\alpha_1} \sum_{x_{m-1}(0)=1}^{+\infty} \left(\frac{\alpha_m}{(1-\alpha_1)^{2(m-1)}}\right)^{x_{m-1}(0)} \cdots \sum_{\{x_1\in\mathbb{N}^{s_1}\}} \left(\frac{\alpha_2}{(1-\alpha_1)^2}\right)^{|x_1|} \tag{3.34}$$

$$= \frac{Z_{\theta\alpha}^{m-1}}{1-\alpha_1} = \cdots = Z_{\theta^{m-1}\alpha}^1 \prod_{l=0}^{m-2} \left(\frac{1}{1-(\theta^l\alpha)_1}\right). \tag{3.35}$$

Hence (3.31) follows from

$$Z_\alpha^1 = \sum_{x_1(0)=1}^{+\infty} \alpha_1^{x_1(0)} = \frac{\alpha_1}{1-\alpha_1}.$$

It remains to discuss the convergence. We use that if $0 < \beta_m < 1$ then $\sum_m \beta_m < +\infty$ if and only if $\prod_m (1-\beta_m) > 0$. When (3.30) is satisfied the generic term in (3.32) is the product of a term of a converging series times a term converging to a finite value and therefore the series in (3.32) is converging. While instead when condition (3.30) is violated the generic term in the series in (3.32) is the product of a term of a diverging series times a diverging term and therefore the series in (3.32) is diverging. □

Proof of Theorem 3.1. The fact that the maps (3.16) and (3.17) are the inverse of each other can be checked by a direct verification. Consider $\alpha \in \mathcal{A}$ and $q = q(\alpha)$. By (3.22) we have

$$\sum_{k\geq 1} q_k = \sum_{k\geq 1} (\theta^{k-1}\alpha)_1 < \infty, \tag{3.36}$$

by (3.30). This proves $q \in \mathcal{Q}$.

Substituting (3.22) into (3.31) and (3.32), we get

$$Z_\alpha^m = q_m \prod_{j=1}^{m} \left(\frac{1}{1-q_j} \right), \qquad m \geq 1 \tag{3.37}$$

and

$$Z_\alpha = 1 + \sum_{m=1}^{+\infty} q_m \prod_{j=1}^{m} \left(\frac{1}{1-q_j} \right). \tag{3.38}$$

Under condition (3.36) the measure $\big(q_m \prod_{\ell>m}(1-q_\ell)\big)_{m\geq 0}$ (recall that $q_0 = 1$) is a probability in $\mathbb{Z}_{\geq 0}$; multiplying therefore (3.38) by $\prod_{k\geq 1}(1-q_k)$ we have

$$Z_\alpha \prod_{k\geq 1}(1-q_k) = \prod_{k\geq 1}(1-q_k) + \sum_{m\geq 1} q_m \prod_{k>m}(1-q_k) = 1.$$

This gives the alternative useful representation

$$Z_\alpha = \prod_{k\geq 1}(1-q_k)^{-1}. \tag{3.39}$$

We will show that ν_α satisfies the following identities.

$$\nu_\alpha\left(M(x) = m\right) = q_m(\alpha) \prod_{\ell>m} \big(1 - q_\ell(\alpha)\big), \quad m \geq 0, \tag{3.40}$$

$$\nu_\alpha\big(x_m(0)\big|M(x) = m\big) = q_m^{|x_m(0)|-1}(\alpha)(1-q_m(\alpha)), \tag{3.41}$$

$$\nu_\alpha\big(x_k\big|x_{k+1}^\infty\big) = q_k^{|x_k|}(\alpha)(1-q_k(\alpha))^{s_k(x)}. \tag{3.42}$$

Since these are the identities (3.13)–(3.15) characterizing φ_q, (3.40)–(3.42) imply $\nu_\alpha = \varphi_q$.

By definition we have

$$\nu_\alpha\big(M(x) = m\big) = \frac{Z_\alpha^m}{Z_\alpha}. \tag{3.43}$$

Using (3.37) and (3.39) we get (3.40). Again by definition we have

$$\nu_\alpha\big(x_k\big|x_{k+1}^{+\infty}\big) = \frac{Z_\alpha(x_k^\infty)}{Z_\alpha(x_{k+1}^\infty)}. \tag{3.44}$$

Using (3.28) and observing that

$$\frac{\big(\theta^{k-1}\alpha\big)_{i+1}}{(\theta^k\alpha)_i} = \frac{\big(\theta^{k-1}\alpha\big)_{i+1}}{(\theta^{k-1}\alpha)_{i+1}\big(1-(\theta^{k-1}\alpha)_1\big)^{-2i}} = \big(1-(\theta^{k-1}\alpha)_1\big)^{2i} \tag{3.45}$$

we obtain directly (3.41), (3.42). This proves $\nu_\alpha = \varphi_{q(\alpha)}$.

Conversely, assume $q \in \mathcal{Q}$. Then $\prod_{\ell\geq 1}(1-q_\ell) > 0$ and we have

$$\varphi_q(x) = \Big(\prod_{\ell>M}(1-q_\ell)\Big) \prod_{k=1}^{M} q_k^{|x_k|}(1-q_k)^{s_k} \tag{3.46}$$

$$= \Big(\prod_{\ell\geq 1}(1-q_\ell)\Big) \prod_{k=1}^{M} q_k^{|x_k|}(1-q_k)^{s_k-1}. \tag{3.47}$$

$$= \Big(\prod_{\ell\geq 1}(1-q_\ell)\Big) \prod_{k=1}^{M} \Big[q_k \prod_{\ell=1}^{k-1}(1-q_\ell)^{2(k-\ell)}\Big]^{|x_k|}. \tag{3.48}$$

where we used $s_\ell = 1 + 2\sum_{k>\ell}(k-\ell)|x_k|$. Comparing this expression with (3.6), we get $Z_{\alpha(q)} = (\prod_{\ell\geq 1}(1-q_\ell))^{-1} < \infty$ and $\varphi_q = \nu_{\alpha(q)}$, using (3.17).

It remains to prove (3.19). It suffices to show that $q \in \mathcal{Q}^+$ if and only if the mean excursion length under φ_q has finite expectation.

Since by (2.8) the value of s_k depends just on $|x_\ell|$ with $\ell > k$ and by the property (3.15) of the measure φ_q we can apply Wald Theorem (see for example [2] section 4.1) getting

$$E_{\varphi_q}(|x_k|) = E_{\varphi_q}\left(\sum_{j=1}^{s_k} x_k(j)\right) = m_k\beta_k, \tag{3.49}$$

where

$$\begin{aligned} m_k &:= \frac{q_k}{1-q_k} \quad \text{is the mean of a Geometric}(1-q_k) \text{ random variable (3.1)} \\ \beta_k &:= E_{\varphi_q}(s_k)\,,\ k \geq 1. \end{aligned}$$

By definition we have that $\beta_k \geq 1$.

The mean excursion size under φ_q is therefore given by

$$E_{\varphi_q}\left(\sum_{k=1}^{\infty} 2k|x_k|\right) = \sum_{k=1}^{\infty} 2km_k\beta_k\,. \tag{3.50}$$

Consider relationship (2.8) and take expected value with respect to the measure φ_q on both sides. Using (3.49) we get the recurrence

$$\beta_k = 1 + \sum_{\ell>k} 2(\ell-k)m_\ell\beta_\ell\,, \qquad k = 0, 1, 2, \dots \tag{3.51}$$

where we observe that when the system has a unique solution then β_0 is 1 plus the mean excursion size under φ_q. Section 3.3 of [4] shows that if $\sum_k km_k < +\infty$, then the recursion (3.51) has a unique finite solution $(\beta_k)_{k\geq 0}$. Since $\sum_k km_k < +\infty$ is equivalent to $\sum_k kq_k < \infty$, we have proven that if $q \in \mathcal{Q}^+$ then the mean excursion size under φ_q is finite, which in turn implies $\alpha(q) \in \mathcal{A}^+$.

Conversely, if the mean excursion size under φ_q is finite, then the series on the right hand side of (3.50) is convergent. Since $\beta_k \geq 1$ this implies that $\sum_k kq_k < \infty$ holds. □

Remark 3.5 ($\alpha \in \mathcal{A}$ is a local property)**.** We point out that while the sets $\mathcal{Q}$ and $\mathcal{Q}^+$ are identified just by asymptotic properties of the parameters q (i.e. changing the values of a finite number of them does not change the belonging or not to these sets), this is not the case for the sets $\mathcal{A}$ and $\mathcal{A}^+$. For example, consider $\alpha = (\alpha_1, \alpha_2, 0, \dots)$ such that $\alpha_k = 0$ for any $k > 2$. Then also $q_k = 0$ for any $k > 2$ and the partition function can be explicitly computed. Using (3.39) and (3.16) we obtain

$$Z_\alpha = \frac{1-\alpha_1}{(1-\alpha_1)^2 - \alpha_2}\,. \tag{3.52}$$

We have that (3.52) is finite and positive if and only if $0 \leq \alpha_1 < 1$ and $0 \leq \alpha_2 \leq (1-\alpha_1)^2$. A similar but more involved computation can be done for any finite numbers of α's different from zero. Notice that all the constraints on the parameters α are also important in order that $0 \leq (\theta\alpha)_k < 1$ in definition (3.21).

3.3 Random walks and Markov chains

We apply Theorem 3.1 to Bernoulli product measures and Markov chains to show that those measures as seen from a record have independent components and, as a corollary, that they are T-invariant.

Lemma 3.6 (Random walks)**.** *Consider the law of an excursion of a random walk which moves upwards with probability $\lambda \in [0, 1/2)$ and downwards with probability $1-\lambda$. This measure corresponds to ν_α with $\alpha \in \mathcal{A}^+$ and given by*

$$\alpha_k = (\lambda(1-\lambda))^k\,, \qquad k \geq 1, \tag{3.53}$$

$$Z_\alpha = (1-\lambda)^{-1}. \tag{3.54}$$

Proof. Let $\mathcal{E}_n$ be the set of excursions of length $2n$. The distribution of an excursion $\varepsilon \in \mathcal{E}_n$ starting at record 0 for the random walk is $(\lambda(1-\lambda))^n(1-\lambda)$, where the last $(1-\lambda)$ is the probability to go to -1 at step $2n+1$. Since $\sum_{k=1}^\infty k n_k[\varepsilon] = n$ for $\varepsilon \in \mathcal{E}_n$, we have that the random walk excursion has law ν_α with parameters (3.53)–(3.54).

Notice that Z_α can also be computed when $\alpha_k = \beta^k$ for some β as follows

$$Z_\alpha = \sum_{\varepsilon\in\mathcal{E}} \prod_{k\geq 1} \alpha_k^{n_k} = \sum_{n=0}^{+\infty} \sum_{\varepsilon\in\mathcal{E}_n} \beta^n = \sum_{n=0}^{+\infty} \frac{1}{n+1}\binom{2n}{n}\beta^n, \tag{3.55}$$

where we used (2.2). The last expression is the generating function of the Catalan numbers [7]. Hence, $Z_\alpha = \frac{2}{1+\sqrt{1-4\beta}} = \frac{1}{1-\lambda}$, when $\beta = \lambda(1-\lambda)$. The fact that $\alpha \in \mathcal{A}^+$ can be verified computing

$$\sum_{k\geq 1} k\rho_k(\alpha) = \beta \frac{\partial}{\partial\beta} \log Z_\alpha \Big|_{\beta=\lambda(1-\lambda)} = \frac{\lambda}{1-2\lambda}. \tag{3.56}$$

The corresponding parameters $q(\alpha)$ can be computed by (3.16) but it seems that there is not a simple analytical expression valid for each k. □

The following is a generalization of the previous Lemma. Let $Q = (Q(i,j))_{i,j\in\{0,1\}}$ be the transition matrix of a Markov chain on $\{0,1\}$ and assume that the stationary probability measure $p = (p_0, p_1)$ of Q satisfies $p_1 \in (0, \frac{1}{2})$, that is, $Q(0,1) < Q(1,0)$. Let μ_Q be the distribution of a double infinite stationary trajectory of the Markov chain.

Lemma 3.7 (Markov chains)**.** *Consider the law of an excursion of $W\eta$ when the configuration of balls η is distributed as μ_Q. This law corresponds to ν_α with $\alpha \in \mathcal{A}^+$ given by*

$$\alpha_k = ab^k, \quad k \geq 1\,, \tag{3.57}$$

where

$$\begin{cases} a = Q(0,1)Q(1,0)\big[Q(1,1)Q(0,0)\big]^{-1}\,, \\ b = Q(1,1)Q(0,0)\,. \end{cases} \tag{3.58}$$

We have moreover that $Z_\alpha = 1/Q(0,0)$.

Proof. The probability of an excursion of the chain has a factor $(Q(0,0)Q(1,1))^{k-1}$ for each k-soliton, a factor $Q(0,1)Q(1,0)$ for each soliton and a global factor $Q(0,0)$ coming from the probability to go to -1 at the end of the excursion. That is, the probability of an excursion ε is given by

$$Q(0,0)\prod_k \big(ab^k\big)^{n_k(\varepsilon)} \tag{3.59}$$

that is $\nu_\alpha(\varepsilon)$ with α given by (3.57) and $Z_\alpha = 1/Q(0,0)$.

We can also obtain Z_α by summing the weights. A classic result says that the number of excursions of length $2n$ and having exactly k local maxima is given by the Narayana numbers

$$N(n,k) = \frac{1}{n}\binom{n}{k}\binom{n}{k-1},$$

see for example exercise 6.36 of [7]. Since $\sum_{k=1}^{\infty} n_k$ coincides with the number of local maxima and $n = \sum_{k=1}^{\infty} k n_k$, the partition function of our Lemma is given by

$$Z_\alpha = 1 + \sum_{n=1}^{+\infty}\sum_{k=1}^{n} N(n,k) a^k b^n = 1 + F(b,a)\,. \tag{3.60}$$

where F is the generating function of the Narayana numbers and it is known ([7] exercise 6.36) to be

$$F(b,a) = \frac{1 - b(1+a) - \sqrt{(1-b-ba)^2 - 4b^2 a}}{2b}\,.$$

Inserting (3.58) in (3.60) and using $Q(0,0) > Q(1,1)$ (which holds because the density is below $1/2$) we get $Z_\alpha = 1/Q(0,0)$ after some elementary steps. The fact that $\alpha \in \mathcal{A}^+$ can be obtained for example using (3.60) since the mean excursion size is given by

$$2b\frac{\partial}{\partial b}\log\left(1 + F(b,a)\right)\,.$$

Here too the corresponding parameters $q(\alpha)$ can be computed by (3.16) but it seems that there is not a simple analytical expression valid for each k. □

4 Infinitely many balls

In this section we consider the space of configurations with infinitely many balls, discuss the BBS dynamics, define measures on this space concatenating excursions, discuss the soliton decomposition of these measures and show that random configurations obtained by concatenating independent excursions with law ν_α have independent components. As a consequence we describe a set of invariant measures with independent components. To make these statements precise we need to introduce Palm measures.

For each $\lambda \in [0,1]$ denote the set of configurations with density λ by

$$\mathcal{X}_\lambda := \Big\{\eta \in \{0,1\}^{\mathbb{Z}} : \lim_{y\to\infty}\frac{1}{y}\sum_{z=-y}^{0}\eta(z) = \lim_{y\to\infty}\frac{1}{y}\sum_{z=0}^{y}\eta(z) = \lambda\Big\}, \quad \text{and}$$
$$\mathcal{X} := \cup_{0\le\lambda<\frac{1}{2}}\mathcal{X}_\lambda, \tag{4.1}$$

the set of configurations with density strictly less than $\frac{1}{2}$. As we see below this space is conserved by the dynamics.

Consider a walk ξ. Recall that $z \in \mathbb{Z}$ is a *record* for ξ if $\xi(z) < \xi(z')$ for any $z' < z$. Notice that if $\eta \in \mathcal{X}_\lambda$ and $\lambda < \frac{1}{2}$, then the records have density $1 - 2\lambda$, as the number of empty boxes equals the number of balls between records. Denote the set of records of η by

$$R\eta := \{z \in \mathbb{Z} \,:\, z \text{ is a record of } W\eta\}$$

and

$$r(\eta, i) := \min\{z \in \mathbb{Z} : W\eta(z) = -i\}$$

the position of *record* i of the walk $W\eta$. When $\eta \in \mathcal{X}$ the position $r(\eta,i)$ is well defined and belongs to $\mathbb{Z}$ for each $i \in \mathbb{Z}$. We use the notation $r(\xi,i) := r(\eta,i)$ when $\xi = W\eta$.

4.1 Concatenating excursions

Given $\eta \in \mathcal{X}$ and $i \in \mathbb{Z}$, call $\eta^{(i)}$ the configuration between records i and $i+1$ translated to the origin:

$$\eta^{(i)}(z) := \eta(r(\eta,i)+z)\,\mathbf{1}\{0 < z < r(\eta,i+1) - r(\eta,i)\} \text{ and} \tag{4.2}$$

$$\varepsilon^{(i)} := W\eta^{(i)} \tag{4.3}$$

be the corresponding walk which is indeed an excursion. The walk $\varepsilon^{(i)}$ is called excursion i of η. If $r(\eta,i+1) = r(\eta,i)+1$, we say that excursion i is empty. Again $\eta^{(i)}$ and $\varepsilon^{(i)}$ can be considered either on a finite interval or on the whole $\mathbb{Z}$, since all the boxes are empty outside of a finite window. We denote by $\underline{\varepsilon} = (\varepsilon^{(i)})_{i\in\mathbb{Z}}$ the collection of excursions of $\eta \in \mathcal{X}$. To make the dependence on η explicit, we may write $\varepsilon^{(i)}[\eta]$ and $\underline{\varepsilon}[\eta]$.

The set of configurations in $\mathcal{X}$ with a record at the origin is denoted by

$$\widehat{\mathcal{X}} := \{\eta \in \mathcal{X} : 0 \in R\eta\}. \tag{4.4}$$

Since $\mathcal{X}$ has ball density less than $1/2$, if $\eta \in \widehat{\mathcal{X}}$ then η has infinitely many records to the right and left of the origin, and hence, all its excursions are finite. As a consequence, the map $\eta \mapsto \underline{\varepsilon}[\eta]$ is a bijection between $\widehat{\mathcal{X}}$ and a suitable subset of $\mathcal{E}^{\mathbb{Z}}$. The reverse map $\underline{\varepsilon} \mapsto \eta = \eta[\underline{\varepsilon}]$ puts record 0 of η at the origin: $r(\eta,0) = 0$ and recursively the other records using the iteration

$$r(\eta,i+1) - r(\eta,i) = r(\varepsilon^{(i)},1) = 2n(\varepsilon^{(i)}) + 1, \tag{4.5}$$

and inserting excursion i between records i and $i+1$:

$$\eta(r(\eta,i)+z) = W^{-1}\varepsilon^{(i)}(z), \quad 0 \le z < r(\varepsilon^{(i)},1). \tag{4.6}$$

Given a configuration $\eta = \eta[\underline{\varepsilon}] \in \widehat{\mathcal{X}}$, define the slots of η by a translation of the slots of the excursions:

$$\text{If } z \text{ is a } k\text{-slot for } \varepsilon^{(i)}, \text{ then } r(\eta,i)+z \text{ is a } k\text{-slot for } \eta[\underline{\varepsilon}]. \tag{4.7}$$

Since 0 is always a k-slot for $\varepsilon^{(i)}$, we have that the records of η are also k-slots for all k.

Give label 0 to the k-slot at record 0: $\mathtt{s}_k(\eta,0) := r(\eta,0)$, and enumerate the other k-slots by

$$\mathtt{s}_k(\eta,j) := \text{ position of the } j\text{-th } k\text{-slot, counting from } k\text{-slot 0, for } j \in \mathbb{Z}. \tag{4.8}$$

See Fig. 6.

FNRW Soliton decomposition of ball configurations Recall the definition (2.4) and the notation (2.5) where $\zeta_k(j)$ is the number of k-solitons appended to k-slot j. Define $D : \widehat{\mathcal{X}} \to \left(\left(\mathbb{Z}_{\ge 0}\right)^{\mathbb{Z}}\right)^{\mathbb{N}}$ the transformation given by

$$\eta \mapsto D\eta = \zeta = \Big(\big(\zeta_k(j)\big)_{j\in\mathbb{Z}}\Big)_{k\in\mathbb{N}}; \tag{4.9}$$

here $\zeta_k(j) \in \mathbb{Z}_{\ge 0}$. In fact §3.2 in [4] shows that D is a bijection between $\widehat{\mathcal{X}}$ and

$$\mathcal{Z} := \left\{\zeta \in \left(\left(\mathbb{Z}_{\ge 0}\right)^{\mathbb{Z}}\right)^{\mathbb{N}} : \sup\{k : \zeta_k(j) > 0\} < \infty, \text{ for all } j \in \mathbb{Z}\right\}. \tag{4.10}$$

We give a construction of D^{-1} in §4.1.2. The array $\zeta = D\eta$ is called the *soliton decomposition* of the configuration η. The *k-component* of the configuration η is $\zeta_k = (\zeta_k(j))_{j\in\mathbb{Z}}$, the k-th row of the array ζ; we also use the notation $\zeta_k = D_k\eta$.

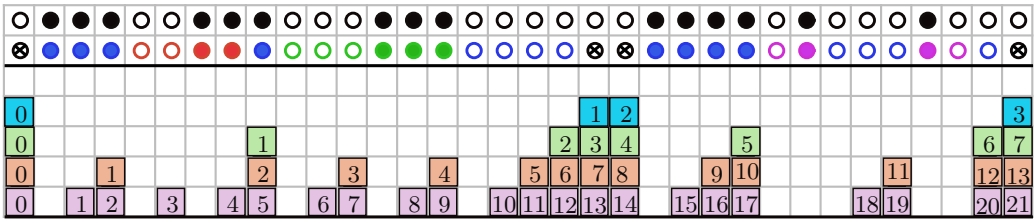

Figure 6: Enumeration of slots. The upper line is a ball configuration η. The second line identifies records and solitons: records are crossed circles, 4-solitons are blue, 3-solitons green, 2-solitons red and 1-solitons purple; there are 4 records and 3 excursions, the second one being empty. In the lines below slots are represented by colored squares piled up below the corresponding box: pink, orange, green and blue light colors correspond to 1,2,3 and 4-slots, respectively. Each k-slot located at the leftmost record has label 0, for each $k \geq 1$; successive k-slots at each line are then enumerated. The k-slots labels for $k \geq 4$ coincide with the labels of the records, as there are no k-solitons for $k > 4$ in this example.

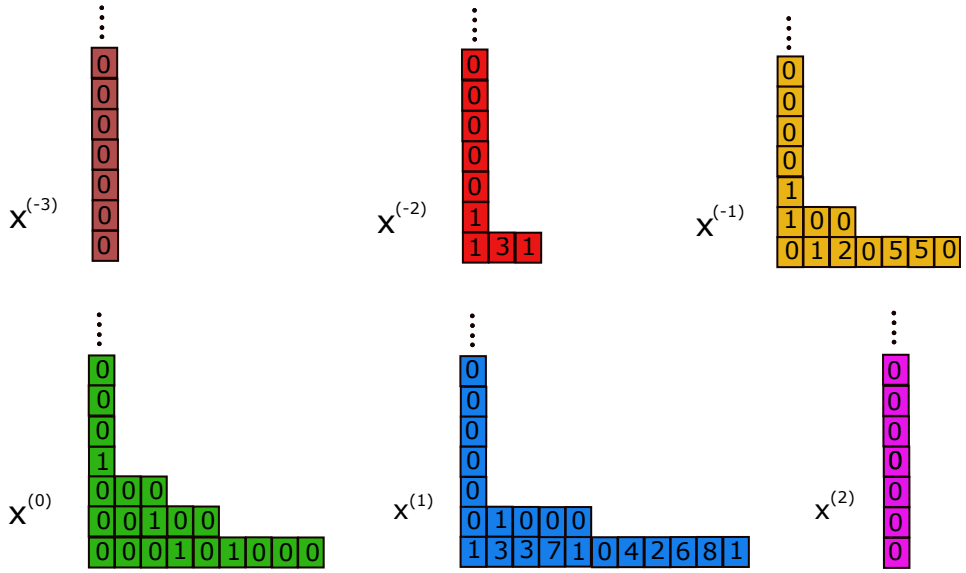

Figure 7: Six different slot diagrams. Line k from bottom on each diagram represents the values of the vector $x_k^{(i)}$ $i = -3, \dots 2$. Observe that $x^{(-3)} = x^{(2)} = \emptyset$.

4.1.1 Concatenation of slot diagrams

Applying the Takahashi-Satsuma algorithm to each excursion $\varepsilon^{(i)}$ we get the corresponding slot diagram $x^{(i)} = x[\varepsilon^{(i)}]$. We can concatenate the slot diagrams to obtain the components $\zeta = D\eta$ of the configuration $\eta \in \widehat{\mathcal{X}}$.

The concatenation of the slot diagrams in Fig. 7 is illustrated in Fig. 8. In Fig. 7 we represent the values of the vectors of the slot diagrams inside boxes left justified. The values on line k from the bottom on each diagram are the values of the vector $x_k^{(i)}$. The values on the column j counting from left (and calling column 0 the leftmost) represent the values $x_k^{(i)}(j)\,, k = 1, 2, \dots$.

The concatenation procedure is the following. The slot diagram $x^{(0)}$ maintains its shape and the column $\left(x_k^{(0)}(0)\right)_{k\geq 1}$ coincides with $\left(\zeta_k(0)\right)_{k\geq 1}$. The remaining slot

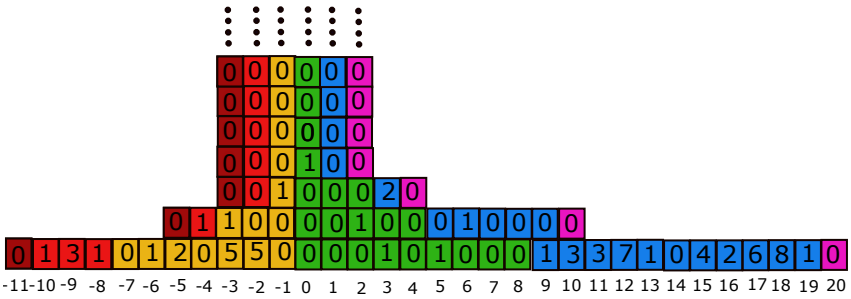

Figure 8: The result of the concatenation procedure of 6 slot diagrams. We have that $x^{(i)}, i = -3, -1, 0, 1, 2$ are the ones illustrated in Fig. 7 where $x^{(-3)} = x^{(2)} = \emptyset$. The integer labels below the picture represents the coordinates. On the column with label j it is possible to read the values of $\big(\zeta_k(j)\big)_{k\geq 1}$.

diagrams are glued joining the rows of the same height in an unique row respecting the order of the labels. Boxes of the row k of the slot diagram $x^{(i)}$ are to the right of the boxes of the row k of the slot diagram $x^{(i-1)}$ and to the left of the boxes of the row k of the slot diagram $x^{(i+1)}$. Recall that each slot diagram has an infinite column containing just zeros above the column number 0.

In Fig. 8 we represent the concatenation of 6 slot diagrams (we do not draw the infinite columns of zeros which should be drawn on the columns $-3, -2, -1, 0, 1, 2$). Concatenating all the slot diagrams $\underline{\varepsilon}[\eta]$ we obtain an infinite array such that on the column j we read the values $\big(\zeta_k(j)\big)_{k\geq 1}$ of the components of the configuration η. A formal description is given in the following paragraph.

More formally, denoting $s_k^{(i)} :=$ number of k-slots in $x^{(i)}$. Define

$$S_k^{(0)} = 0; \quad S_k^{(i+1)} - S_k^{(i)} = s_k^{(i)}\,. \tag{4.11}$$

Consider $\underline{x} := \big(x^{(i)}\big)_{i\in\mathbb{Z}}$ and let $\zeta = \zeta[\underline{x}]$ be defined by

$$\zeta_k(S_k^i + j) = x_k^{(i)}(j), \quad j = 0, \dots, s_k^{(i)} - 1\,;\ k \geq 1\,;\ i \in \mathbb{Z}\,. \tag{4.12}$$

It is not hard to see that the ζ so constructed is the decomposition of the configuration η whose excursions have slot diagrams $x^{(i)}$:

$$\zeta\Big(\underline{x}\big[\underline{\varepsilon}[\eta]\big]\Big) = D\eta, \tag{4.13}$$

where $\underline{x}\big[\underline{\varepsilon}[\eta]\big]$ denotes the slot diagrams $\big(x\big[\varepsilon^{(i)}[\eta]\big]\big)_{i\in\mathbb{Z}}$.

4.1.2 From components to slot diagrams

We explain now how to construct a family of slot diagrams starting from an array $\zeta \in \mathcal{Z}$, that is, with the property $\sup\{k \geq 0 : \zeta_k(j)\} < \infty$ for all $j \in \mathbb{Z}$. In Fig. 9 we show a portion of the infinite array ζ and discuss how to generate the slot diagrams $x^{(i)}$, $i \geq 0$. In the first step (top picture) we search for the maximal row in column 0 such that the corresponding value is strictly positive. We color by red the square, add it to the slot diagram $x^{(0)}$ and set $M(x^{(0)}) = 4$. Then we compute $s_3^{(0)}$ using (2.8), color by yellow a corresponding number of squares in the row 3 and add them to the slot diagram $x^{(0)}$. Now we compute $s_2^{(0)}$ again using (2.8), color a corresponding number of squares in the

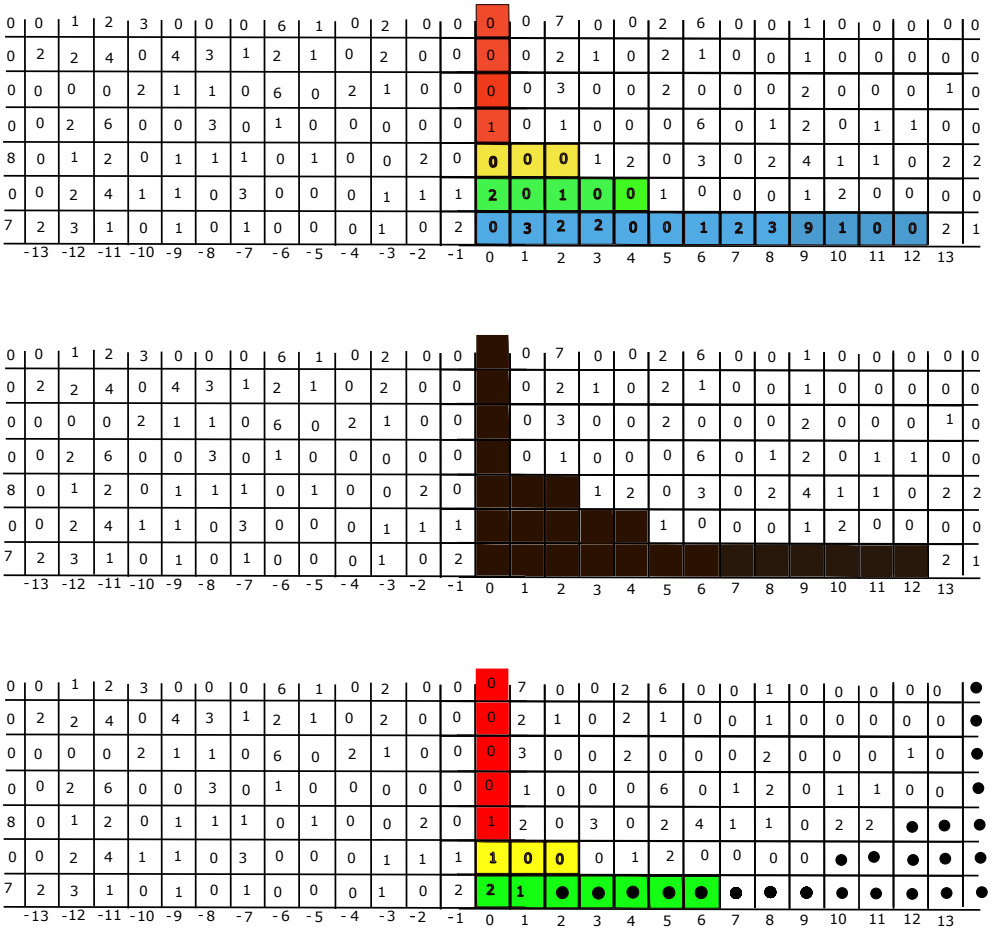

Figure 9: Construction of the slot diagrams $x^{(i)}$, $i = 0, 1$, from an infinite array ζ. In the first picture from above the slot diagram $x^{(0)}$ consists of the colored region. The squares has been added to the diagram in this order: red, yellow, green and blue. In the middle picture the squares corresponding to $x^{(0)}$ are black colored and then removed. In the bottom figure the lines have been shifted to the left to fill the empty spaces. Squares previously outside of the picture are drawn with a • inside. The slot diagram $x^{(1)}$ corresponds now to the colored region with squares added following the same order as before.

row 2 by green and add them to the slot diagram $x^{(0)}$. Finally compute $s_1^{(0)}$, color by blue a corresponding number of squares on the first line and add them to the slot diagram $x^{(0)}$. The final slot diagram number zero $x^{(0)}$ consists of all the colored region.

To construct slot diagram number 1, erase all colored boxes and shift the non erased region of each positive row to the left, until we have again an array. This is illustrated in the middle and bottom picture of Fig. 9. In the middle picture we colored black those squares to be deleted, while in the bottom picture we shifted the lines to the left. Each line has been shifted by the corresponding number of • appearing on the right. Apply now the algorithm we have used above to identify slot diagram zero and call the result slot-diagram 1. This is illustrated again in the bottom picture of Fig. 9 using the same order of the colors. Repeat the procedure to construct the slot diagrams with nonnegative label.

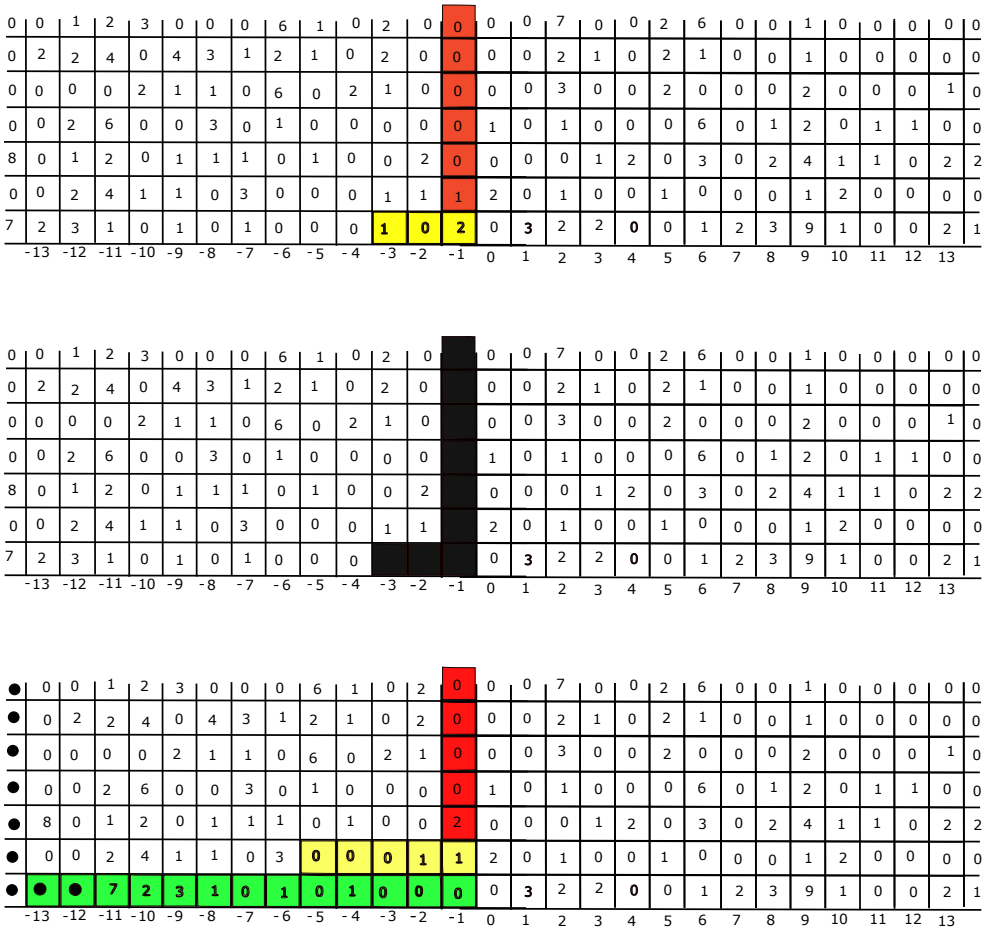

Figure 10: Construction of the slot diagrams $x^{(i)}$, $i = -1, -2$ from the infinite array ζ. The slot diagrams consist of the colored region added according the same rules as before. On the top picture we have $x^{(-1)}$, on the middle picture we color black the squares to be removed, on the bottom picture we shift to the right the rows and construct $x^{(-2)}$.

To construct the slot diagrams with negative label, use the same algorithm as for label zero but working from right to left, and the procedure illustrated in Fig. 10.

Finally the slot diagrams produced by the above iterations of the algorithm are the following

This construction is formally described as follows. Let $\zeta = \big((\zeta_k(j))_{j\in\mathbb{Z}}\big)_{k\geq 1}$ belong to $\mathcal{Z}$. We construct a slot-diagram $x = x[\zeta]$ as follows. Set

$$M(x) := \sup\{k \geq 0 : \zeta_k(0) > 0\} < +\infty\,, \tag{4.14}$$

a bounded nonnegative integer. Call $m = M(x)$ and set

$$\begin{aligned} s_k &= 1, \text{ for } k \geq m, \\ x_k(0) &= 0, \text{ for } k > m, \\ x_m(0) &= \zeta_m(0). \end{aligned} \tag{4.15}$$

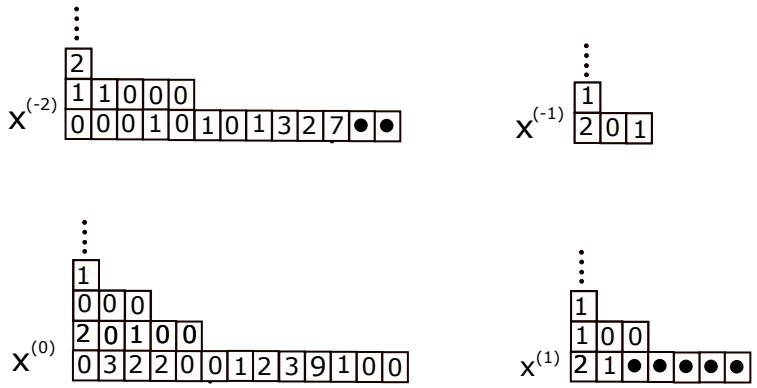

Figure 11: The four slot diagrams produced by the iteration of the algorithm in Figures 9 and 10.

Assume $\big(x_\ell(0),\dots,x_\ell(s_\ell-1)\big)$ is known for $\ell>k$ and iteratively define

$$\begin{aligned} |x_\ell| &= \sum_{j=0}^{s_\ell-1} x_\ell(j),\\ s_k &= 1+2\sum_{\ell>k}(\ell-k)|x_\ell|,\\ x_k(j) &= \zeta_k(j),\quad j=0,\dots,s_k-1. \end{aligned} \tag{4.16}$$

We have constructed a slot diagram

$$x := \big(x_k(0),\dots,x_k(s_k-1)\big)_{k\ge1}. \tag{4.17}$$

Write $x[\zeta]$ and $s_k[\zeta]$ to stress that x and s_k are functions of ζ and define the hierarchical translation

$$\phi\zeta = (\tau^{s_k[\zeta]}\zeta_k)_{k\ge1}\,. \tag{4.18}$$

The coordinate $s_k[\zeta]$ is the leftmost positive coordinate of ζ_k not used in the construction of $x[\zeta]$. We stress that the translation $\tau^{s_k[\zeta]}$ in (4.18) acts on the index labeling the slots, more precisely

$$(\phi\zeta)_k(j) = \tau^{s_k[\zeta]}\zeta_k(j) = \zeta_k(j+s_k[\zeta])\,.$$

Since $s_k=1$ for all $k\ge m$, we have $(\phi\zeta)_k(j)=\zeta_k(j+1)$ for all $k\ge m$. Hence, since ζ belongs to the set (4.10), so does $\phi\zeta$ and we can define iteratively

$$x^{(i)} := x[\phi^i\zeta],\quad i\ge0. \tag{4.19}$$

For negative i let ζ' be the reflection of ζ with respect to the origin translated by -1: $\zeta'(j):=\zeta(-j-1)$ for $j\in\mathbb{Z}$ and define

$$x^{(i)} := (x[\phi^{-i-1}\zeta'])',\quad i<0\,, \tag{4.20}$$

that is, construct the slots diagrams for ζ', reflect the obtained slot diagrams, assign the reflected slot diagram of 0 to -1 and so on. In (4.20) for a slot diagram x we defined the reflected one x' by $x'_k(j)=x_k(s_k-j-1)$. The corresponding excursions are then given by

$$\varepsilon^{(i)} := \varepsilon[x^{(i)}],\quad \underline{\varepsilon} = (\varepsilon^{(i)})_{i\in\mathbb{Z}}. \tag{4.21}$$

Lemma 4.1 (FNRW)**.** *The configuration $\eta = \eta[\underline{\varepsilon}]$ satisfies $D\eta = \zeta$.*

See §2.3, "Reconstructing the configuration from the components" in [4] for a proof of this lemma. This implies that D is a bijection between $\cup_{\lambda<1/2}\widehat{\mathcal{X}}_\lambda$ and $\mathcal{Z}$ and we can write $\eta = D^{-1}\zeta$.

4.2 Measures on ball configurations and soliton components

We here define distributions on arrays of components and ball configurations starting with independent families of iid excursions and vice-versa.

Palm measures We consider configurations with all records and the underlying point process of the records. We start by recalling the definition of Palm and anti-Palm measures, see Chapter 8 of Thorisson [9] for background and proofs of the following facts, which are stated with respect to the point process of the records.

Let μ be a translation invariant measure on $\mathcal{X}$ and define $\lambda = \lambda(\mu) := \int \eta(0)\,\mu(\mathrm{d}\eta)$ its mean density; the density of records is then $1-2\lambda$. Define the measure $\mathrm{Palm}(\mu)$ on $\widehat{\mathcal{X}}$ by its action on test functions f, which is given by

$$\int f(\eta)\,\mathrm{Palm}(\mu)(\mathrm{d}\eta) = \frac{1}{1-2\lambda}\int \mathbf{1}\{0 \in R\eta\} f(\eta)\mu(\mathrm{d}\eta). \tag{4.22}$$

This is the measure μ conditioned to have a record at the origin. $\mathrm{Palm}(\mu)$ is record-translation invariant:

$$\int f(\eta)\,\mathrm{Palm}(\mu)(\mathrm{d}\eta) = \int f(\tau^{r(\eta,i)}\eta)\,\mathrm{Palm}(\mu)(\mathrm{d}\eta), \quad \text{for all record } i. \tag{4.23}$$

Reciprocally, for a record-translation invariant measure $\widehat{\mu}$ on $\widehat{\mathcal{X}}$ with finite average inter-record distance

$$\kappa(\widehat{\mu}) := \int r(\eta,1)\widehat{\mu}(\mathrm{d}\eta) \in [1,\infty), \tag{4.24}$$

define the *anti-Palm* measure $\mathrm{Palm}^{-1}(\widehat{\mu})$ acting on test functions f as

$$\mathrm{Palm}^{-1}(\widehat{\mu})f := \frac{1}{\kappa(\widehat{\mu})}\int \sum_{z=1}^{r(\eta,1)} f(\tau^z\eta)\,\widehat{\mu}(\mathrm{d}\eta). \tag{4.25}$$

The measure $\mu := \mathrm{Palm}^{-1}(\widehat{\mu})$ is translation invariant and has mean ball density

$$\lambda(\mu) = \frac{\kappa(\widehat{\mu})-1}{2\kappa(\widehat{\mu})} \in [0, \tfrac{1}{2}), \tag{4.26}$$

indeed $\frac{1}{2}(\kappa(\widehat{\mu})-1)$ is the mean number of balls per excursion, that is, between two successive records and $\kappa(\widehat{\mu})$ is the mean distance between successive records. There are record-translation invariant measures $\widehat{\mu}$ with infinite average inter-record distance, but concentrating on the set of configurations with all records finite. The anti-Palm transformation of those measures is not defined.

The next proposition proven by FNRW says that random arrays in $\mathcal{Z}$ with translation invariant distribution and independent components produce record-translation invariant distributions on the space of ball configurations.

Proposition 4.2 (FNRW, Independent components and Palm measures)**.** *Let ζ be a random array with translation invariant distribution concentrating on $\mathcal{Z}$ and satisfying $(\zeta_k)_{k\geq 1}$ independent. Then the law of $D^{-1}\zeta$, denoted by $\widehat{\mu}$, is record-translation invariant. Furthermore, if $\sum_k kE\left[\zeta_k(0)\right] < \infty$, then the inter-record distance under $\widehat{\mu}$ is finite and the measure $\mathrm{Palm}^{-1}(\widehat{\mu})$ is translation invariant and concentrates on $\mathcal{X}$.*

We have the following result.

Theorem 4.3 (Soliton weights and independent geometric components)**.** *a) Let $\alpha \in \mathcal{A}$ and $\underline{\varepsilon} = (\varepsilon^{(i)})_{i\in\mathbb{Z}}$ be a sequence of i.i.d. random excursions with distribution ν_α given by* (3.5)*. Let $\widehat{\mu}_\alpha$ be the distribution of $\eta = \eta[\underline{\varepsilon}]$, the random ball configuration with Record 0 at the origin and excursions $(\varepsilon^{(i)})_{i\in\mathbb{Z}}$, defined in* (4.6)*. Define $\zeta := D\eta$, the soliton decomposition of η, defined in* (4.9)*. Then $\zeta_k(j)$ are independent Geometric$(1-q_k(\alpha))$ random variables, for all $j \in \mathbb{Z}, k \geq 1$.*

b) Reciprocally, let $q \in \mathcal{Q}$ and $\zeta = \big((\zeta_k(j))_{j\in\mathbb{Z}}\big)_{k\geq 1}$ be an array of independent random variables with $\zeta_k(j)$ distributed according to Geometric$(1-q_k)$, for all $j \in \mathbb{Z}$ for all $k \geq 1$. Then $\zeta \in \mathcal{Z}$ with probability 1 and denoting $\eta := D^{-1}\zeta$, we have that $\varepsilon^{(i)}[\eta]$ are i.i.d. excursions with law $\nu_{\alpha(q)}$, so that η has law $\widehat{\mu}_{\alpha(q)}$, a record-translation invariant measure.

Proof. a) Let $x^{(i)} = x[\varepsilon^{(i)}]$ be the slot diagram of the excursion $\varepsilon^{(i)}$. By Theorem 3.1 $x^{(i)}$ satisfies (3.14) and (3.15), that is, given the number of k-slots $s_k^{(i)}$, the variables $x_k^{(i)}(0), \dots, x_k^{(i)}(s_k^{(i)} - 1)$ are i.i.d. Geometric$(1-q_k(\alpha))$ random variables. Let $\mathcal{F}_k$ be the sigma field generated by the k-th row $(x_k^{(i)})_{i\in\mathbb{Z}}$ and denote by $\mathcal{F}_{>k}$ the sigma field generated by $\big((x_{k+1}^{(i)})_{i\in\mathbb{Z}}, (x_{k+2}^{(i)})_{i\in\mathbb{Z}}, \dots\big)$, the rows bigger than k. Condition on $\mathcal{F}_{>k}$ and construct ζ_k using (4.12), that is juxtaposing the k-component of each slot diagram one after the other. Since the excursions are independent, the resulting component $\zeta_k \in (\mathbb{Z}_{\geq 0})^{\mathbb{Z}}$ consists of i.i.d. Geometric$(1-q_k(\alpha))$ random variables independently of the conditioning. This implies that $\zeta_k(j)$ are independent Geometric$(1-q_k(\alpha))$ random variables, for all $j \in \mathbb{Z}, k \geq 1$, concluding the proof of item a.

b) It suffices to show that the excursions generated by the slot diagrams $(x^{(i)}[\zeta])_{i\in\mathbb{Z}}$ have marginal law $\nu_{\alpha(q)}$ and are independent. It is immediate from the construction illustrated in Fig. 9 (top picture) that $x^{(0)}[\zeta]$ satisfies (3.13)–(3.15). Let the array $\zeta^{(1)}$ obtained by erasing the entries used by $x^{(0)}[\zeta]$ and sliding the remaining entries to the left (Fig. 9). Since the set of erased entries does not depend on de contents of the non-erased entries and the entries in ζ are independent, $\zeta^{(1)}$ has the same law as ζ and it is independent of $x^{(0)}[\zeta]$. Then $x^{(1)}[\zeta] = x^{(0)}[\zeta^{(1)}]$, which is independent of the previous slots diagrams. The same argument applies to the construction of the slot diagrams of ζ with negative label (see Fig. 10). □

4.3 Invariant measures for the BBS

Theorem 4.4 below, proven by FNRW, states that a translation invariant measure whose Palm transform has independent components is invariant for the BBS dynamics. As a consequence of Theorems 4.4 and 4.3, we will conclude that the measure $\mu_\alpha = \mathrm{Palm}^{-1}(\hat{\mu}_\alpha)$ introduced in Theorem 4.3 is invariant for the dynamics.

The BBS dynamics can be described by the operator T acting on configurations $\eta \in \mathcal{X}$ by

$$T\eta(z) := (1 - \eta(z))\, \mathbf{1}\{z \notin R\eta\}. \tag{4.27}$$

The configuration $T\eta$ coincides with η at the records of η and the contents of the other boxes are inverted. Indeed, at each iteration of T the balls in each excursion go to the empty boxes of the same excursion and the record boxes remain empty. In particular, the number of balls and empty boxes of η and $T\eta$ between two successive records of η are the same. Since the records have a positive density, this implies that density is conserved by T: $T\mathcal{X}_\lambda = \mathcal{X}_\lambda$ for any $\lambda \in [0, 1/2)$ and that $T : \mathcal{X} \to \mathcal{X}$ indeed. When η has finitely many balls $T\eta$ coincides with the configuration obtained after the carrier has visited all boxes of the configuration η, as described in the introduction.

We say that a measure μ is T-*invariant* if $\mu \circ T^{-1} = \mu$. The next theorem of FNRW establishes conditions under which translation invariant measures with independent soliton components are T-invariant.

Theorem 4.4 (FNRW. Independent components and T-invariance)**.** *Let $\zeta = (\zeta_k)_{k\geq 1}$ be a random array with translation invariant distribution and independent rows satisfying $\sum_k kE\zeta_k(0) < \infty$. Let $\widehat{\mu}$ be the law of $D^{-1}\zeta$. Then $\mu := \mathrm{Palm}^{-1}(\widehat{\mu})$ is T-invariant.*

We have proven in Theorem 4.3 that for $\alpha \in \mathcal{A}$ the measure obtained by concatenating i.i.d. copies of excursions with law ν_α has independent components. Applying then Theorem 4.4 we conclude in Theorem 4.5 below that if $\alpha \in \mathcal{A}^+$ this measure is the Palm measure of a T-invariant measure. As particular cases, we deduce in Corollaries 4.6 and 4.8 that product measures and stationary Markov chains in $\{0,1\}$ with density of balls less than $\frac{1}{2}$ are T-invariant, a fact proven in [4] and [1] using classical arguments and reversibility properties of queues.

We now show that if $\alpha \in \mathcal{A}^+$, then μ_α is T-invariant and that if $q \in \mathcal{Q}^+$, then $\mu_{\alpha(q)}$ is T-invariant. When $\alpha \in \mathcal{A}^+$ we have

$$\kappa(\alpha) := \kappa(\widehat{\mu}_\alpha) = 1 + 2\sum_k k\rho_k(\alpha) < \infty\,, \tag{4.28}$$

where $\kappa(\widehat{\mu})$ is defined in (4.24). We define also $\lambda(\alpha) := \lambda(\mu_\alpha)$ where $\lambda(\mu)$ is defined in (4.26).

Theorem 4.5 (μ_α is T-invariant)**.** *a) Assume the conditions of Theorem 4.3. If $\alpha \in \mathcal{A}^+$, then $\widehat{\mu}_\alpha$ concentrates on $\widehat{\mathcal{X}}_{\lambda(\alpha)}$ and it is record-translation invariant and the measure $\mu_\alpha := \mathrm{Palm}^{-1}(\widehat{\mu}_\alpha)$ concentrates on $\mathcal{X}_{\lambda(\alpha)}$, is translation invariant and T-invariant.*

b) If $q \in \mathcal{Q}^+$, then $\widehat{\mu}_{\alpha(q)}$ concentrates on $\widehat{\mathcal{X}}_{\lambda(\alpha(q))}$ and $\mu_{\alpha(q)} := \mathrm{Palm}^{-1}(\widehat{\mu}_{\alpha(q)})$ is translation invariant, concentrates on $\mathcal{X}_{\lambda(\alpha(q))}$ and is T-invariant.

The next corollaries prove that product measures on $\mathcal{X}$ and stationary trajectories on $\mathcal{X}$ of Markov chains on $\{0,1\}$ may be expressed as μ_α of Theorem 4.5, by choosing the appropriate α and/or q. In particular those measures are T-invariant, a fact already proven by using reversibility of those trajectories by [1] and [4].

Corollary 4.6 (Product measures)**.** *Let $\lambda \in [0, \frac{1}{2})$ and π_λ be the product measure on $\mathcal{X}$ with density λ. Let $\widehat{\pi}_\lambda := \mathrm{Palm}(\pi_\lambda)$ and η be distributed with $\widehat{\pi}_\lambda$. Define*

$$\alpha_k := (\lambda(1-\lambda))^k\,. \tag{4.29}$$

Then $\alpha \in \mathcal{A}^+$ and the random excursions $\left(\varepsilon^{(i)}[\eta]\right)_{i\in\mathbb{Z}}$ are i.i.d. with distribution ν_α, the soliton components $(D_k\eta)_{k\geq 0}$ are mutually independent and the k-soliton component $(D_k\eta(j))_{j\in\mathbb{Z}}$ is a sequence of i.i.d. Geometric$(1-q_k(\alpha))$ random variables. As a consequence, the measure π_λ is T-invariant.

Remark 4.7 (Mean number of solitons per box)**.** Denote by δ_k the mean number of k-solitons per box under the product measure π_λ. It is given by $\delta_k = \rho_k\,(1-2\lambda)$, where ρ_k is the mean number of k-solitons per excursion, that is between successive records and $(1-2\lambda)^{-1}$ is the mean distance between successive records under π_λ. Kuniba and Lyu [5] have computed an explicit expression for δ_k in terms of λ.

Corollary 4.6 is a special case of the next corollary for Markov chains.

Corollary 4.8 (Markov chains and Ising models)**.** *Let $Q = (Q(i,j))_{i,j\in\{0,1\}}$ be the transition matrix of a Markov chain in $\{0,1\}$ and assume that the stationary probability measure (p_0, p_1) of Q satisfies $p_1 \in (0, \frac{1}{2})$. Let π_Q be the distribution of a double infinite stationary trajectory of the chain. Define $\widehat{\pi}_Q := \mathrm{Palm}(\pi_Q)$ and η be a configuration with law $\widehat{\pi}_Q$. Define $\alpha = (\alpha_k)_{k\geq 1}$ by $\alpha_k := ab^k$ for a, b defined in function of Q by* (3.58)*. Then $\alpha \in \mathcal{A}^+$*

and the random excursions $\left(\varepsilon^{(i)}[\eta]\right)_{i\in\mathbb{Z}}$ *are i.i.d. with distribution* ν_α*, the soliton components* $(D_k\eta)_{k\geq 0}$ *are mutually independent and the* k*-soliton component* $(D_k\eta(j))_{j\in\mathbb{Z}}$ *is a sequence of i.i.d. Geometric*$(1-q_k(\alpha))$ *random variables. As a consequence,* μ_Q *is* T*-invariant.*

Remark 4.9 (Infinite expected excursion size)**.** When $\alpha \in \mathcal{A}\setminus\mathcal{A}^+$, the mean excursion size under ν_α is infinite and $\widehat{\mu}_\alpha$ defined in Theorem 4.3 has infinite mean inter-record distance. The independence of components is still valid in this case but $\mathrm{Palm}^{-1}(\widehat{\mu}_\alpha)$ cannot be defined [9]. In particular, when α is given by (4.29) with $\lambda=\frac{1}{2}$, ν_α is the law of an excursion of the symmetric random walk, $\widehat{\mu}_\alpha$ is well defined but its inverse-Palm measure is not, as the density of records is 0.

Proof of Theorem 4.5. a) Since $\alpha\in\mathcal{A}^+$ the mean inter-record distance is finite under $\widehat{\mu}_\alpha$ (4.28) and therefore the measure μ_α is well defined and translation invariant, as we saw in §4.2. The fact that μ_α is T-invariant will follow by Theorem 4.4 once we show that $\sum_k kE\zeta_k(j)<\infty$. Since $\zeta_k(j)$ is Geometric$(1-q_k(\alpha))$, the condition $\sum_k kE\zeta_k(j)<\infty$ is equivalent to $q(\alpha)\in\mathcal{Q}^+$ that follows by Theorem 3.1.

b) Since $q\in\mathcal{Q}^+$ and $\zeta_k(j)$ is Geometric$(1-q_k)$ we have $\sum_k kE\zeta_k(j)<\infty$ and we can apply Theorem 4.4. □

References

[1] Croydon, D. A., Kato, T., Sasada, M., and Tsujimoto, S. *Dynamics of the box-ball system with random initial conditions via Pitman's transformation*. Preprint arXiv:1806.02147, 2018.

[2] Durrett, R. *Probability: theory and examples*. Fourth edition. Cambridge Series in Statistical and Probabilistic Mathematics, **31**. Cambridge University Press, Cambridge, 2010. MR-2722836

[3] Ferrari, P. A., Gabrielli, D. *Box-ball system: soliton and tree decomposition of excursions*. arXiv:1906.06405

[4] Ferrari, P. A., Nguyen, C., Rolla, L. T., and Wang, M. *Soliton decomposition of the Box-Ball-System*. Preprint arXiv:1806.02798, 2018.

[5] Kuniba, A., Lyu, H. *Large deviations and one-sided scaling limit of randomized multicolor box-ball system*. arXiv:1808.08074 MR-4056651

[6] Levine, L., Lyu, H., and Pike, J. *Double jump phase transition in a soliton cellular automaton*. Preprint. arXiv:1706.05621, 2017.

[7] Stanley, R. P. *Enumerative combinatorics. Vol. 2*, vol. **62** of Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge, 1999. With a foreword by Gian-Carlo Rota and appendix 1 by Sergey Fomin. MR-1676282

[8] Takahashi, D., and Satsuma, J. *A soliton cellular automaton*. Journal of The Physical Society of Japan **59**, 10 (1990), 3514–3519. MR-1082435

[9] Thorisson, H. *Coupling, stationarity, and regeneration*. Probability and its Applications (New York). Springer-Verlag, New York, 2000. MR-1741181

Acknowledgments. We thank the referees of the Electronic Journal of Probability for their careful reading and for many helpful comments for the presentation of the results. PAF thanks many illuminating discussions with Leo Rolla.

This project started when the first author was visiting Gran Sasso Science Institute in L'Aquila in 2016 and then developed during the stay of the authors at the Institut Henri Poincare, Centre Emile Borel, during the trimester *Stochastic Dynamics Out of Equilibrium* in 2017. We thank both institutions for warm hospitality and support.