BBS invariant measures with independent soliton components

Pablo A. Ferrari, Davide Gabrielli

December 4, 2018

Abstract

The Box-Ball System (BBS) is a one-dimensional cellular automaton in $\{0,1\}^{\mathbb{Z}}$ introduced by Takahashi and Satsuma, who also identified conserved sequences called *solitons*. Integers are called boxes and a ball configuration indicates the sites occupied by balls. For each integer $k \geq 1$, a k-soliton consists of k boxes occupied by balls and k empty boxes. Ferrari, Nguyen, Rolla and Wang define the k-slots of a configuration as the places where k-solitons can be appended. Labeling the k-slots with integer numbers, they define the k-component of the configuration as the element of $\mathbb{Z}_{\geq 0}^{\mathbb{Z}}$ giving the number of k-solitons appended to each k-slot. They also show that shift-invariant distribution with independent soliton components are invariant for the automaton. We show that for each $\lambda \in [0, 1/2)$ the product measure on ball configurations with parameter λ has independent soliton components and that its k-component is a product measure of geometric random variables with parameter $1 - q_k(\lambda)$, an explicit function of λ. The construction is used to describe a large family of invariant measures with independent components, including Ising like measures.

Keywords: Box-Ball System, soliton components, conservative cellular automata

AMS 2010 Subject Classification: 37B15, 37K40, 60C05

1 Introduction

Takahashi and Satsuma [5], referred by 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 $\eta(i) = 1$ indicates that there is a ball at box $i \in \mathbb{Z}$. A carrier visits successively boxes from left to right picking balls from occupied boxes and deposing one ball, if carried, at the current visited box, if empty. We denote $T\eta$ the configuration obtained when the carrier has visited all boxes and $T^t\eta$ the configuration obtained after iterating this procedure t times, for positive integer t.

TS show the existence of conserved quantities in the BBS that they called *basic sequences* and Levine, Lyu and Pike [3] call *solitons*. An isolated k-soliton consists of

k-successive occupied boxes followed by k successive empty boxes. Not being other balls in the system, a 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 [5]. A k-soliton consists always of k occupied boxes and k empty boxes. Different solitons occupy disjoint sets of boxes and the trajectory of each tagged soliton can be identified along time [2] (referred by FNRW in the sequel).

A configuration can be mapped to a walk that jumps one unit up at occupied boxes and one unit down at empty boxes [1] [2]. The *excursions* of the walk are the pieces of configuration between two consecutive down *records*. Walks coming from configurations with density of balls less than $\frac{1}{2}$ have positive density of records, so that they contain only finite excursions. FNRW introduce a soliton decomposition of each ball configuration. An m-soliton contains $2(m-k)$ boxes for each $k<m$, where any finite number of k-solitons may be inserted without destroying the m-soliton; those boxes are called *k-slots*. Records are k-slots for all k and any k-slot is also a $(k-1)$-slot. Enumerating the k-slots, a k-soliton γ is *appended* to k-slot number i if the boxes occupied by γ are contained in the interval between k-slots numbered i and $i+1$. The *k-component* of a ball configuration η is denoted $D_k\eta \in (\mathbb{Z}_{\geq 0})^{\mathbb{Z}}$, where $D_k\eta(i)$ is the number of k-solitons appended to the i-th k-slot of η, $i \in \mathbb{Z}$. FNRW also proved that the k-component of the configuration $T\eta$ is a shift of the k-component of η, the amount shifted depends on the m-components of η for $m>k$.

Let μ be a shift-invariant measure on the set of ball configurations with density less than $1/2$ and call $\widehat{\mu}$ the record Palm measure of μ, defined as the measure μ conditioned to have a record at the origin. FNRW show that if μ is shift-invariant and $\widehat{\mu}$ has independent k-components, then μ is invariant for the dynamics. FNRW also study the asymptotic speed of solitons when the initial distribution of particles is translation invariant and ergodic.

Let $\lambda \in [0,1)$ and call μ_λ the product measure of Bernoulli(λ) random variables on the space $\{0,1\}^{\mathbb{Z}}$. Let $\widehat{\mu}_\lambda$ be its record Palm-measure, that is, the measure conditioned to have a record at the origin. In this paper we show that for $\lambda \in [0,\frac{1}{2})$, if η is distributed according to $\widehat{\mu}_\lambda$, then the components $(D_k\eta)_{k\geq 1}$ are independent and each component $(D_k\eta(i))_{i\in\mathbb{Z}}$ consists of i.i.d. Geometric random variables with parameter $1-q_k(\lambda)$, computed later in Corollary 3. We construct many other invariant measures with independent components, being each component i.i.d. Geometric random variables. A particular case is the distribution μ_Q of a stationary Markov chain with state space $\{0,1\}$ and transitions $Q(1,0)>Q(0,1)$, to guarantee that density of 1's is less than $\frac{1}{2}$; these are also nearest neighbor Ising-like measures with a negative external field.

The independence of components combined with Theorem 1 below (that is proved in [2]) imply that μ_λ and the Ising-like measures are invariant for BBS. These facts were proven directly by Croydon, Kato, Sasada and Tsujimoto [1], using space reversibility of the so called carrier process of BBS; see also [2].

To prove the results just described we introduce a family of probability measures ν_α on the set of finite excursions indexed by $\alpha=(\alpha_k)_{k\geq 1}$, a collection of parameters 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. Theorem 2, that is the main result of this paper, shows that conditioned on the components $m > k$, the distribution of the k-component is a product of s_k geometric distributions with a parameter $1-q_k(\alpha)$, where $s_k := 1 + \sum_{m>k} 2(m-k)n_m$ is the number of k-slots induced by the m-components, $m > k$ and $q(\alpha) = (q_k(\alpha))_{k\geq 1}$ is an explicit function of α, see (20), (21). 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 configuration by putting a record at the origin and concatenating the excursions separated by records. The resulting configuration is record-shift-invariant, that is, its distribution is the same as seen from any record and the mean distance between successive records is finite. Using the inverse-Palm transformation, we obtain a shift-invariant and T-invariant measure.

We call a *slot diagram* a finite portion of the array of the components that codify the combinatorial arrangement of solitons of one single excursion of the ball configuration. We show that the law of the slot diagram of the excursion between Records 0 and Record 1 of a configuration with law $\widehat{\mu}_\lambda$ coincides with the law of the slot diagram of any excursion of the configuration $\eta = D^{-1}\zeta$ where $\zeta = (\zeta_k)_{k\in\mathbb{Z}}$ are independent and $\zeta_k := (\zeta_k(i))_{i\in\mathbb{Z}}$ are i.i.d. Geometric$(1-q_k(\lambda))$. See Theorem 2 and Corollary 3.

2 Soliton decomposition

In this section we describe a variant of the Takahashi-Satsuma Algorithm to identify solitons and then the soliton decomposition proposed by FNRW.

A configuration of balls is an element $\eta \in \{0,1\}^{\mathbb{Z}}$, where for $y \in \mathbb{Z}$, $\eta(y) = 1$ means that there is a ball at box y, otherwise box y is empty. For each $\lambda \in [0,1]$ denote the set of configurations with density λ by

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

the set of configurations with density less than $\frac{1}{2}$.

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

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

We fix the constant by choosing $\xi(0) = 0$. Let $\mathcal{W}_\lambda := \{W\eta : \eta \in \mathcal{X}_\lambda\}$ and $\mathcal{W} := \{W\eta : \eta \in \mathcal{X}\}$ denote the corresponding sets in the space of walks. 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 $\xi(z) - \xi(z') = \sum_{y=z'+1}^{z}(2\eta(y) - 1)$. 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 $R\eta := \{z \in \mathbb{Z} : z \text{ is a record of } W\eta\}$ and $r(\eta, i) := \min\{z \in \mathbb{Z} : W\eta(z) = -i\}$ the position of *Record* i.

Excursion deconcatenation of configurations. Given $\eta \in \mathcal{X}$ and two successive records $r(\eta, i)$ and $r(\eta, i+1)$, the configuration ε_i defined by

$$\varepsilon_i(z) = \eta(r(\eta, i) + z)\mathbf{1}\{0 < z < r(\eta, i+1) - r(\eta, i)\} \tag{3}$$

is called *Excursion* i of η. If $r(\eta, i+1) = r(\eta, i) + 1$, we say that Excursion i is *empty*. Denote $\underline{\varepsilon} = (\varepsilon_i)_{i\in\mathbb{Z}}$. To make explicit the dependence on η, we may write $\varepsilon_i[\eta]$ and $\underline{\varepsilon}[\eta]$.

We call $\mathcal{E}$ the set of *Excursions*, that are configurations of balls ε with finitely many balls and such that $r(\varepsilon, i) - r(\varepsilon, i-1) = 1$ for any $i \neq 1$ and $r(\varepsilon, 0) = 0$. Observe that an excursion is completely identified by the ball configuration between Record 0 and Record 1. The corresponding finite walk is what is usually called excursion in the literature. We identify the two notions.

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

$$\widehat{\mathcal{X}} := \{\eta \in \mathcal{X} : 0 \in R\eta\}. \tag{4}$$

The map $\eta \mapsto \underline{\varepsilon}[\eta]$ is a bijection between $\widehat{\mathcal{X}}$ and $\mathcal{E}^{\mathbb{Z}}$. 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 satisfying

$$r(\eta, i+1) - r(\eta, i) = r(\varepsilon_i, 1),$$

and then setting the configuration between Records i and $i+1$ to be a translation by $r(\eta, i)$ of ε_i:

$$\eta(r(\eta, i) + z) = \varepsilon_i(z), \quad 0 \le z < r(\varepsilon_i, 1). \tag{5}$$

Identification of solitons. We describe a variant of the Takahashi-Satsuma algorithm to identify the solitons of a non-empty configuration $\eta \in \mathcal{X}$. Empty configurations have no solitons. A *run* of a configuration η is any segment $[x, y]$ with $x \le y$ such that $\eta(x) = \eta(z)$, for $z \in [x, y]$, $\eta(x-1) \neq \eta(x)$ if $x > -\infty$ and $\eta(y) \neq \eta(y+1)$ if $y < \infty$. An excursion has two semi-infinite runs and a finite number of finite runs. The algorithm applied to a configuration with finitely many balls 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. Go to the leftmost run of size k. The restriction of η to the k boxes of this run and the first k boxes of the successive run is called k-soliton.

2. Ignore the boxes belonging to already identified solitons, update the runs of the remaining configuration and go to 1.

Assume that η has infinitely many records to the left and to the right of the origin and identify all solitons of η by repeating the procedure above with each (finite) excursion of η.

The *support* $\{\gamma\}$ of a k-soliton γ consists on 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 $\Gamma_k \eta$ the set of k-solitons of η. Every box in $\mathbb{Z}$ is either a record or belongs to $\{\gamma\}$ for some k-soliton γ, for some $k \geq 1$.

Slots. A box z is a *k-slot* if either $z \in R\eta$ 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 $S_{k+1}\eta \subseteq S_k \eta$.

Enumerate the k-slots setting $s_k(\eta, 0) := r(\eta, 0)$, that is, k-slot 0 is at Record 0 for all k, and

$$s_k(\eta, j) := \text{position of the } j\text{-th } k\text{-slot, counting from } k\text{-slot } 0, \text{ for } j \in \mathbb{Z}. \tag{6}$$

Soliton decomposition of ball configurations [2]. 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 $j+1$:

$$\{\gamma\} \subset \big(s_k(\eta, j), s_k(\eta, j+1)\big).$$

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

$$\zeta_k(j) := \#\{\gamma \in \Gamma_k \eta : \gamma \text{ is appended to } k\text{-slot } j\}. \tag{7}$$

Denote $D : \widehat{\mathcal{X}} \to \big((\mathbb{Z}_{\geq 0})^{\mathbb{Z}}\big)^{\mathbb{N}}$ the transformation given by

$$D\eta \mapsto \zeta = (\zeta_k)_{k \geq 0}. \tag{8}$$

In fact §3.2 in [2] shows that D is a bijection between the sets

$$\big\{\eta \in \{0,1\}^{\mathbb{Z}} : 0 \in R\eta \text{ and all excursions of } \eta \text{ are finite}\big\} \text{ and} \tag{9}$$

$$\big\{\zeta \in \big((\mathbb{Z}_{\geq 0})^{\mathbb{Z}}\big)^{\mathbb{N}} : \sup\{k : \zeta_k(j) > 0\} < \infty, \ \text{ for all } j \in \mathbb{Z}\big\}. \tag{10}$$

We give a construction of D^{-1} in §5.2. The variables $\zeta = D\eta$ are called the *soliton components* (or just components) of the configuration η.

If $\zeta \in \big((\mathbb{Z}_{\geq 0})^{\mathbb{Z}}\big)^{\mathbb{N}}$ is a collection of random variables satisfying

$$\sum_k k E\zeta_k(j) < \infty \qquad \forall j \in \mathbb{Z}\,, \tag{11}$$

then ζ belongs to the set (10) almost surely and $D^{-1}\zeta$ is well defined a.s.. Let $\zeta = (\zeta_k)_{k \geq 1}$ be a random family of independent elements satisfying (11) with shift-invariant distribution. Denote $\widehat{\mu}$ the law of $\eta := D^{-1}\zeta$. Then $\widehat{\mu}(\widehat{\mathcal{X}}) = 1$ and $\widehat{\mu}$ is *record-shift-invariant* (see Theorem 4.1 in [2]), that is, for test functions f we have

$$\int f(\tau^{r(\eta,i)}\eta)\, \widehat{\mu}(\mathrm{d}\eta) = \int f(\eta)\, \widehat{\mu}(\mathrm{d}\eta), \quad \text{for all Record } i \text{ of } \eta\,. \tag{12}$$

Furthermore we have that the mean distance between records under $\widehat{\mu}$ is finite:

$$\int \big(r(\eta, i+1) - r(\eta, i)\big)\, \widehat{\mu}(\mathrm{d}\eta) = \int r(\eta, 1)\, \widehat{\mu}(\mathrm{d}\eta) < \infty, \quad \text{for all Record } i \text{ of } \eta. \tag{13}$$

Palm measures [6]. Given a shift-invariant measure μ on $\mathcal{X}$ with ball density $\lambda \in [0, \frac{1}{2})$, hence with record density $1 - 2\lambda$, define a measure $\mathrm{Palm}(\mu)$ on $\widehat{\mathcal{X}}$ which acts on test functions by

$$\int f(\eta)\,\mathrm{Palm}(\mu)(\mathrm{d}\eta) = \frac{1}{1-2\lambda}\int \mathbf{1}\{0 \in R\eta\} f(\eta)\mu(\mathrm{d}\eta). \tag{14}$$

This is the measure μ conditioned to have a record at the origin.

Reciprocally, for a record-shift-invariant measure $\widehat{\mu}$ on $\widehat{\mathcal{X}}$ such that $\int r(\eta,1)\widehat{\mu}(\mathrm{d}\eta) = (1-2\lambda)^{-1}$ for some $\lambda \in [0, \frac{1}{2})$ define $\mathrm{Palm}^{-1}(\widehat{\mu})$ as the measure acting on test functions f as

$$\int f(\eta)\ \mathrm{Palm}^{-1}(\widehat{\mu})(\mathrm{d}\eta) = (1-2\lambda)\int \sum_{z=1}^{r(\eta,1)} f(\tau^z\eta)\,\widehat{\mu}(\mathrm{d}\eta) \tag{15}$$

The measure $\mu := \mathrm{Palm}^{-1}(\widehat{\mu})$ is shift-invariant and has ball density $\int \eta(z)\mu(\mathrm{d}\eta) = \lambda$ for all $z \in \mathbb{Z}$.

Dynamics of the BBS. The present paper studies properties of measures invariant for the BBS, without discussing the dynamics. However, to motivate our result, we state an important dynamic result of FNRW [2].

The BBS dynamics is defined by the transformation $T : \mathcal{X} \to \{0,1\}^{\mathbb{Z}}$ given by

$$T\eta(z) := (1 - \eta(z))\mathbf{1}\{z \notin R\eta\} \tag{16}$$

This operator coincides with the TS carrier transformation described in Introduction when applied to finite configurations. The configuration $T\eta$ coincides with η at the records of η and invert the contents of the other boxes. This is because at each iteration of T the balls in each excursion go to the empty boxes of the same excursion; 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. In turn, this implies that density is conserved by T: $T\mathcal{X}_\lambda = \mathcal{X}_\lambda$ for any $\lambda \in [0, 1/2)$.

We say that μ is *T-invariant* if $\mu \circ T^{-1} = \mu$. FNRW have established conditions under which shift-invariant measures with independent soliton components are T-invariant:

Theorem 1 (FNRW [2])**.** *Let $\zeta = (\zeta_k)_{k\geq 1}$ be a family of independent random elements satisfying* (11) *with shift-invariant distribution. Let $\widehat{\mu}$ be the law of $D^{-1}\zeta$. Then $\widehat{\mu}$ concentrates on $\widehat{\mathcal{X}}$ and it is record-shift-invariant. The measure $\mu := \mathrm{Palm}^{-1}(\widehat{\mu})$ concentrates on $\mathcal{X}$, it is shift-invariant and T-invariant.*

3 Results

We introduce a probability measure ν_α on the set of excursions satisfying that under ν_α, the probability of an excursion ε is proportional to $\prod_k \alpha_k^{n_k(\varepsilon)}$, where α_k are parameters and $n_k(\varepsilon)$ is the number of k-solitons in ε. Under a suitable choice of α_k, the excursion has finite expected length. Our main result, Theorem 2, states that the measure obtained by concatenating i.i.d. copies of those excursions has independent components. Applying then Theorem 1 we conclude that this measure is the Palm measure of a T-invariant measure. As particular cases, we deduce in Corollaries 3 and 4 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 [2] and [1] using classical arguments and reversibility properties of queues.

3.1 Independent soliton measures

Random finite excursions. Recall that $\mathcal{E}$ is the set of finite excursions between records 0 and 1 and for each $\varepsilon \in \mathcal{E}$ denote

$$n_k(\varepsilon) := \text{number of } k\text{-solitons in } \varepsilon. \tag{17}$$

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

$$Z_\alpha := \textstyle\sum_{\varepsilon\in\mathcal{E}} \prod_{k\geq 1} \alpha_k^{n_k(\varepsilon)} < \infty \tag{18}$$

and define the measure ν_α on $\mathcal{E}$ by

$$\nu_\alpha(\varepsilon) := \frac{1}{Z_\alpha} \textstyle\prod_{k\geq 1} \alpha_k^{n_k(\varepsilon)}\,. \tag{19}$$

Define $q = q(\alpha) = (q_k)_{k\geq 1}$ by

$$q_1 := \alpha_1 \quad \text{and iteratively,} \tag{20}$$

$$q_k := \frac{\alpha_k}{\prod_{j=1}^{k-1}(1-q_j)^{2(k-j)}}\,, \qquad k \geq 2. \tag{21}$$

We are interested in α such that the mean excursion size under ν_α is finite:

$$\sum_{\varepsilon\in\mathcal{E}} \left[\Big(\sum_{j\geq 1} 2j\, n_j(\varepsilon) \Big) \prod_{k\geq 1} \alpha_k^{n_k(\varepsilon)} \right] < \infty. \tag{22}$$

Clearly (22) implies $Z_\alpha < \infty$ and hence that ν_α is well defined. Denoting the mean number of k-solitons per excursion by

$$\rho_k(\alpha) := \sum_{\varepsilon\in\mathcal{E}} n_k(\varepsilon)\nu_\alpha(\varepsilon), \tag{23}$$

inequality (22) is equivalent to:

$$\rho(\alpha) := \sum_{k\geq 1} 2k\rho_k(\alpha) < \infty. \tag{24}$$

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

$$P(Y=j) = (1-p)p^j,\ j\geq 0; \quad EY = \frac{p}{1-p}. \tag{25}$$

with the convention $0^0 = 1$. Our main result is the following.

Theorem 2. *Let α satisfy (22) and $\underline{\varepsilon} = (\varepsilon_i)_{i\in\mathbb{Z}}$ be a sequence of i.i.d. random excursions with distribution ν_α given by (19). Let $\eta = \eta[\underline{\varepsilon}]$ be the ball configuration with Record 0 at the origin and excursions $(\varepsilon_i)_{i\in\mathbb{Z}}$, defined in (5). Denote $\widehat{\mu}_\alpha$ the distribution of η. Denote $\zeta := D\eta$, the soliton decomposition of η, defined in (8). Then $\zeta = (\zeta_k)_{k\geq 1}$ is a family of independent components and for each k, $(\zeta_k(j))_{j\in\mathbb{Z}}$ are i.i.d. Geometric$(1-q_k)$ random variables, where $q = (q_k)_{k\geq 1}$ is given by (20)-(21).*

Furthermore let $\rho = \rho(\alpha)$ defined in (24) and $\lambda := \frac{\rho}{2(\rho+1)} < \frac{1}{2}$. Then $\widehat{\mu}_\alpha$ concentrates on $\widehat{\mathcal{X}}_\lambda$ and it is record shift-invariant and $\mu_\alpha := \mathrm{Palm}^{-1}(\widehat{\mu}_\alpha)$ is shift-invariant, concentrates on $\mathcal{X}_\lambda$ and is T-invariant.

When the mean excursion size under ν_α is infinite, the mean inter-record distance under μ_α is also infinite. The independence of components is still valid in this case but $\mathrm{Palm}^{-1}(\widehat{\mu})$ cannot be defined unless the mean distance between records is finite, that is unless (24) holds [6]. The expected values of $\zeta_k(i)$ under $\widehat{\mu}_\alpha$ are given by $q_k/(1-q_k)$.

A basic ingredient to prove Theorem 2 is the computation of the components of an excursion ε with law ν_α, using hierarchic partitions functions. This is done in Section 4. The converse construction is simpler. We consider a family $\zeta = (\zeta_k)_{k\geq 1}$ of independent vectors each consisting on i.i.d. random variables Geometric$(1-q_k)$ and show that the law of the excursions of $D^{-1}\zeta$ coincides with ν_α when α depend on q by the inverse of relations (20) (21) that is given later in (75). We present this construction in §5. Note that while given the α's it is in general not possible to obtain a closed form to the corresponding q's, the expression of the α's in terms of the q's is always explicit (75).

Product measures on $\mathcal{X}$ and stationary trajectories of Markov chains on $\{0,1\}$ with ball density less than $\frac{1}{2}$ can be constructed with the recipe of Theorem 2, by choosing α conveniently. Hence these measures have independent components with i.i.d. Geometric-distributed entries, as stated in the next corollaries.

Corollary 3 (Product measures)**.** *Let $\lambda \in [0, \frac{1}{2})$ and μ_λ be the product measure on $\mathcal{X}$ with density λ. Let $\widehat{\mu}_\lambda := \mathrm{Palm}(\mu_\lambda)$ and η be distributed with $\widehat{\mu}_\lambda$. Define*

$$\alpha_k := (\lambda(1-\lambda))^k . \tag{26}$$

Then $\alpha = (\alpha_k)_{k\geq 1}$ satisfies (24) and the excursions $\big(\varepsilon_i[\eta]\big)_{i\in\mathbb{Z}}$ 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 $(D_k\eta(i))_{i\in\mathbb{Z}}$ is a sequence of i.i.d. Geometric$(1-q_k)$ where $q_k = q_k(\lambda)$ are given by (20), (21). By Theorem 2 μ_λ is therefore T invariant.

Proof. Let $\mathcal{E}_n$ be the set of excursions occupying $2n$ boxes and let $\varepsilon \in \mathcal{E}_n$. The probability of ε under the product measure with density λ is $(\lambda(1-\lambda))^n(1-\lambda)$, where the last $(1-\lambda)$ is the probability to go to the record to the right of the excursion. This is the same as $\nu_\alpha(\varepsilon)$ with $\alpha_k = (\lambda(1-\lambda))^k$ and $Z_\alpha = (1-\lambda)^{-1}$.

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

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

where the Catalan number $\frac{1}{n+1}\binom{2n}{n}$ is the number of excursions of length $2n$. The last expression is the generating function of the Catalan numbers. Hence, $Z_\alpha = \frac{2}{1+\sqrt{1-4\beta}} = \frac{1}{1-\lambda}$, when $\beta = \lambda(1-\lambda)$. $\square$

Corollary 3 is a special case of the next corollary for Markov chains.

Corollary 4 (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 $\pi = (\pi_0, \pi_1)$ of Q satisfies $\pi_1 \in (0, \frac{1}{2})$. Let μ_Q be the distribution of a double infinite stationary trajectory of the Markov process. Denote $\widehat{\mu}_Q := \mathrm{Palm}(\mu_Q)$ and η be a configuration with law $\widehat{\mu}_Q$. Define $\alpha = (\alpha_k)_{k\geq 1}$ by*

$$\alpha_k := ab^k, \quad k \geq 1, \tag{27}$$

where

$$\begin{cases} a = Q(0,1)Q(1,0)\left[Q(1,1)Q(0,0)\right]^{-1}, \\ b = Q(1,1)Q(0,0). \end{cases} \tag{28}$$

Then α satisfies (22) and the 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 $(D_k\eta(i))_{i\in\mathbb{Z}}$ is a sequence of i.i.d. Geometric$(1-q_k)$ random variables, where $q_k = q_k(\alpha)$ is given by (20), (21). By Theorem 2 μ_Q is therefore T invariant.

Proof. In this case we get 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 Record 1 at the end of the excursion. So that

$$\nu_\alpha(\varepsilon) = Q(0,0)\prod_k \left(ab^k\right)^{n_k(\varepsilon)} \tag{29}$$

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 [4]. The partition function of our Lemma is given by

$$Z_\alpha = 1 + \sum_{n=1}^{+\infty}\sum_{k=1}^{n} N(n,k)a^k b^n = 1 + F(b,a). \tag{30}$$

where F is the generating function of the Narayana numbers and it is known to be

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

Inserting (28) in (30) 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. $\square$

4 Excursions and Slot diagrams

Slot diagrams A slot diagram is just a decomposition of an excursion; it facilitates the description of measures on the set of excursions.

Let $x = (x_k)_{k\geq 1}$ be a family of vectors $x_k = (x_k(0), \dots, x_k(s_k - 1))$ with $s_k \in \mathbb{N}$ and $x_k(j) \in \mathbb{Z}_{\geq 0}$. Denote $|x_k| := x_k(0) + \cdots + x_k(s_k - 1)$. We say that x is a *slot diagram* if it satisfies the following:

$$M(x) := \max\{k : x_k(0) > 0\} < \infty$$

$$s_\ell = 1, \text{ for } \ell \geq M(x) \text{ and } x_\ell(0) = 0 \text{ for } \ell > M(x) \tag{31}$$

$$s_k = 1 + \sum_{\ell > k} 2(\ell - k)|x_\ell|, \tag{32}$$

Denote $\mathcal{S}$ the set of slot diagrams. We show that $\mathcal{E}$ is in bijection with $\mathcal{S}$. We now construct the map $\varepsilon \mapsto x[\varepsilon]$. The map $x \mapsto \varepsilon[x]$ is given later in §5.1.

Construction of $x[\varepsilon]$. If $r(\varepsilon, 1) = r(\varepsilon, 0) + 1$, the excursion is empty and the slot diagram is defined as $s_k \equiv 1$ and $x_k(0) \equiv 0$. If ε is not empty, let $M = M(x)$ 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 (32) 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 $k = M - 1, \dots, 1$.

The proof of Theorem 2 is based on the following expression of ν_α in terms of slot diagrams.

Proposition 5. *Let α satisfy* (18), *ν_α be given by* (19) *and $x = x[\varepsilon]$ be the slot diagram of a finite excursion $\varepsilon \in \mathcal{E}$. Then,*

$$\nu_\alpha(\varepsilon) = \prod_{k\geq 1} q_k^{|x_k|}(1 - q_k)^{s_k(x)}. \tag{33}$$

where $(q_k)_{k\geq 1} = (q_k(\alpha))_{k\geq 1}$ are given in (20)-(21).

The proposition is proven at the end of this section after some auxiliary lemmas.

Denoting $x_k^\infty = (x_k, x_{k+1}, \dots)$, formula (33) 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 associated to a random excursion with law ν_α.

$$\nu_\alpha\left(M(x) = m\right) = q_m \prod_{\ell > m} \big(1 - q_\ell\big), \quad m \geq 0, \tag{34}$$

$$\nu_\alpha\big(x_m(0) \big| M(x) = m\big) = q_m^{|x_m(0)|-1}(1 - q_m), \tag{35}$$

$$\nu_\alpha\big(x_k \big| x_{k+1}^\infty\big) = q_k^{|x_k|}(1 - q_k)^{s_k(x)}, \tag{36}$$

where we abuse notation writing x_m as “the set of excursions ε whose m-component in $x[\varepsilon]$ is x_m”, and so on. Then, to construct a slot diagram of an excursion with law ν_α, first choose a maximal soliton-size m with probability (34) and use (35) to determine the number of maximal solitons $x_m(0)$ (a Geometric$(1-q_m)$ random variable conditioned to be strictly positive). Then we use (36) to construct iteratively the lower components. Under the measure ν_α and conditioned on x_{k+1}^∞, the variables $(x_k(0), \dots x_k(s_k-1))$ are i.i.d. Geometric$(1-q_k)$.

Partition functions In the next three lemmas we compute the partition functions needed to show Proposition 5.

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

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

We have that τx is also a slot diagram. For $\alpha = (\alpha_k)_{k\in\mathbb{N}} \in [0,1]^{\mathbb{N}}$ we define another "shift" operator θ by

$$\big(\theta\alpha\big)_k := \frac{\alpha_{k+1}}{(1-\alpha_1)^{2k}} \qquad k = 1, 2, \dots . \tag{37}$$

so that we can write (20)-(21) as

$$q_k = \big(\theta^{k-1}\alpha\big)_1, \quad k \geq 1, \tag{38}$$

with the convention $\theta^0\alpha = \alpha$. Define the partition functions

$$Z_\alpha(x_k^\infty) := \prod_{n\geq k} \alpha_n^{|x_n|} \sum_{\{y: y_k^\infty = x_k^\infty\}} \prod_{\ell=1}^{k-1} \alpha_\ell^{|y_\ell|}, \qquad Z_\alpha(x) = \prod_{n\geq 1} \alpha_n^{|x_n|}, \tag{39}$$

that is, the sum of weights of the slot diagrams y that are compatible with x_k^∞. These partition functions satisfy a useful recurrence:

Lemma 6. *We have*

$$Z_\alpha(x_k^\infty) = \frac{Z_{\theta\alpha}((\tau x)_{k-1}^\infty)}{(1-\alpha_1)}, \qquad k > 1 . \tag{40}$$

Proof. From (39) we have

$$Z_\alpha(x_k^\infty) = \prod_{i=k}^{\infty} \alpha_i^{|x_i|} \sum_{\left\{y_2^\infty : y_k^\infty = x_k^\infty\right\}} \prod_{j=2}^{k-1} \alpha_j^{|y_j|} \sum_{y_1\in\mathbb{N}^{s_1}} \alpha_1^{|y_1|}, \tag{41}$$

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

$$s_1 = s_1\left(y_2^{k-1} x_k^{+\infty}\right) = 1 + 2\sum_{i=2}^{k-1}(i-1)|y_i| + 2\sum_{i=k}^{+\infty}(i-1)|x_i| . \tag{42}$$

Substituting this in (41) we get

$$
\begin{aligned}
Z_\alpha(x_k^\infty) &= \frac{1}{(1-\alpha_1)} \prod_{i=k}^{\infty} \left[\frac{\alpha_i}{(1-\alpha_1)^{2(i-1)}}\right]^{|x_i|} \sum_{\left\{y_2^\infty : y_k^{+\infty} = x_k^{+\infty}\right\}} \prod_{j=2}^{k-1} \left[\frac{\alpha_j}{(1-\alpha_1)^{2(j-1)}}\right]^{|y_j|} \\
&= \frac{1}{(1-\alpha_1)} \prod_{i=k-1}^{+\infty} (\theta\alpha)_i^{|(\tau x)_i|} \sum_{\left\{y_1^\infty : y_{k-1}^\infty = (\tau x)_{k-1}^\infty\right\}} \prod_{j=1}^{k-2} (\theta\alpha)_j^{|y_j|} \,, \qquad (43)
\end{aligned}
$$

which gives (40). $\square$

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

Lemma 7. *For any fixed $k \geq 2$ and x_k^∞ we have*

$$
Z_\alpha(x_k^\infty) = \left[\prod_{i=0}^{k-2} \left(\frac{1}{(1-(\theta^i\alpha)_1)}\right)\right] \left[\prod_{j=k}^{+\infty} (\theta^{k-1}\alpha)_{j-k+1}^{|x_j|}\right] . \qquad (44)
$$

Proof. Iterating $k-1$ times the recursion (40) 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. $\square$

We now compute the partition function Z_α. The weight of the slot diagrams that have m as maximum soliton size is denoted

$$
Z_\alpha^m := \sum_{x : M(x) = m} Z_\alpha(x) \,, \qquad m \geq 0 \,. \qquad (45)
$$

Lemma 8. *The partition function Z_α is finite if and only if*

$$
\sum_{m=0}^{+\infty} (\theta^m\alpha)_1 < \infty. \qquad (46)
$$

In this case we have for $m \geq 1$

$$
Z_\alpha^m = (\theta^{m-1}\alpha)_1 \prod_{j=0}^{m-1} \left(\frac{1}{1-(\theta^j\alpha)_1}\right) , \qquad m \geq 1 \qquad (47)
$$

and

$$
Z_\alpha = 1 + \sum_{m=1}^{+\infty} (\theta^{m-1}\alpha)_1 \prod_{j=0}^{m-1} \left(\frac{1}{1-(\theta^j\alpha)_1}\right) . \qquad (48)
$$

Proof. Since the weight of the empty excursion is 1, we have $Z^0_\alpha = 1$ and (48) is obtained from (47) from the relation $Z_\alpha = \sum_{m=0}^{+\infty} Z^m_\alpha$. To show (47) write

$$Z^m_\alpha = \sum_{x_m(0)=1}^{+\infty} \alpha_m^{x_m(0)} \sum_{\{x_{m-1}\in\mathbb{N}^{s_{m-1}}\}} \alpha_{m-1}^{|x_{m-1}|} \cdots \sum_{\{x_1\in\mathbb{N}^{s_1}\}} \alpha_1^{|x_1|}\,, \tag{49}$$

where $s_k = s_k(x_{k+1}^{+\infty})$ given by (31) (32) 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 $+\infty$ so that 0 is an element of $\mathbb{N}$ according to our convention. Sum on x_1, use (31) (32) and change name to the summed variables to obtain

$$Z^m_\alpha = \frac{1}{1-\alpha_1} \sum_{x_{m-1}(0)=1}^{+\infty} \Big(\frac{\alpha_m}{(1-\alpha_1)^{2(m-1)}}\Big)^{x_{m-1}(0)} \cdots \sum_{\{x_1\in\mathbb{N}^{s_1}\}} \Big(\frac{\alpha_2}{(1-\alpha_1)^2}\Big)^{|x_1|} \tag{50}$$

$$= \frac{Z^{m-1}_{\theta\alpha}}{1-\alpha_1} = \cdots = Z^1_{\theta^{m-1}\alpha} \prod_{l=0}^{m-2} \left(\frac{1}{1-(\theta^l\alpha)_1}\right)\,. \tag{51}$$

Hence (47) follows from

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

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 (46) is satisfied the generic term in (48) is the product of a term of a converging series times a term converging to a finite value and therefore the series in (48) is converging. While instead when condition (46) is violated the generic term in the series in (48) is the product of a term of a diverging series times a diverging term and therefore the series in (48) is diverging. □

Remark 9. *Note that all the results of Lemma 8 can be written naturally in terms of the parameters q using* (38)*. In particular the convergence condition reads*

$$\sum_{k=1}^{+\infty} q_k < +\infty\,. \tag{52}$$

We have in addition

$$Z^m_\alpha = q_m \prod_{j=1}^{m} \left(\frac{1}{1-q_j}\right)\,, \qquad m \geq 1 \tag{53}$$

and

$$Z_\alpha = 1 + \sum_{m=1}^{+\infty} q_m \prod_{j=1}^{m} \left(\frac{1}{1-q_j}\right)\,. \tag{54}$$

Proof of Proposition 5. Under condition (46) the measure $\big(q_m \prod_{\ell>m}(1-q_\ell)\big)_{m\geq 0}$ is a probability in $\mathbb{N}$, Multiplying (48) by $\prod_{k\geq 1}(1-q_k)$ we have

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

We deduce therefore that we have the alternative (with respect to (54)) useful representation

$$Z_\alpha = \prod_{k\geq 0}(1-(\theta^k\alpha)_1)^{-1} = \prod_{k\geq 1}(1-q_k)^{-1}\,. \tag{55}$$

We prove (34)-(36) which are equivalent to (33). By definition

$$\nu_\alpha(M(x)=m) = \frac{Z_\alpha^m}{Z_\alpha} \tag{56}$$

Using (47) and the representation (55) of the partition function we get (34).

By definition we have

$$\nu_\alpha\big(x_k\big|x_{k+1}^{+\infty}\big) = \frac{Z_\alpha(x_k^\infty)}{Z_\alpha(x_{k+1}^\infty)}\,. \tag{57}$$

Using (44) and observing that

$$\frac{\big(\theta^{k-1}\alpha\big)_{i+1}}{(\theta^k\alpha)_i} = \big(1-(\theta^{k-1}\alpha)_1\big)^{2i} \tag{58}$$

we obtain directly (35), (36). □

5 Soliton components and excursions

5.1 From slot diagrams to excursions

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

$$I_{k,j}\eta(z) = \begin{cases} \eta(z) & \text{if } z\leq u\\ 1-\eta(u) & \text{if } u<z\leq u+k\\ \eta(u) & \text{if } u+k<z\leq u+2k\\ \eta(z-2k) & \text{if } u+2k<z. \end{cases} \tag{59}$$

Denote $I_{k,j}^n$ the operator: *iterate n times the operator $I_{k,j}$*, which corresponds to insert n k-solitons one after the other on the same slot. When $n=0$ we just have the identity.

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

$$\begin{aligned} \eta_\ell &\equiv 0 \quad \text{for } \ell>M, \quad \text{and iteratively,}\\ \eta_k &:= I_{k,0}^{x_k(0)}\dots I_{k,s_k-1}^{x_k(s_k-1)}\eta_{k+1}, \quad \text{for } k=M,\dots,1.\\ \varepsilon[x] &:= \eta_1. \end{aligned} \tag{60}$$

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

$$n_k(\varepsilon[x]) = \sum_{j=0}^{s_k-1} x_k(j) = |x_k|. \tag{61}$$

Example. Consider the following slot diagram x:

$$\begin{aligned} x_\ell &= (0), \quad \text{for } \ell > 3 \\ x_3 &= (2) \\ x_2 &= (0,0,1,0,0) \\ x_1 &= (3,0,4,1,0,0,0,0,2,0,1) \end{aligned} \tag{62}$$

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-soliton to 3-slot 0) I^2_{3,0}
0111000110011100 0              (attach 1 2-soliton to 2-slot 2) I^1_{2,2}
01010101110001100111000         (attach 3 1-soliton to 1-slot 0) I^3_{1,0}
0101010111010101010001100111000 (attach 4 1-soliton to 1-slot 2) I^4_{1,2}
010101011101010101001001100111000 (attach 1 1-soliton to 1-slot 3) I^1_{1,3}
01010101110101010100100110011101 01000 (attach 2 1-solitons to 1-slot 8) I^2_{1,8}
0101010111010101010010011001110101000010 (attach 1 1-soliton to 1-slot 10) I^1_{1,10}
```

The resulting excursion is given by

$$\begin{aligned} \varepsilon[x] &= I^1_{1,10} I^2_{1,8} I^1_{1,3} I^4_{1,2} I^3_{1,0} I^1_{2,2} I^2_{3,0} \eta_4 \\ &= \ldots 1010101110101010101000110011101010000 10 \ldots \end{aligned}$$

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 $M = 3$.

5.2 From components to configurations

We construct the map $\zeta \mapsto D^{-1}\zeta = \eta$ by first constructing a sequence of slot diagrams $x^i[\zeta]$, then a sequence of excursions $\varepsilon^i = \varepsilon[x^i[\zeta]]$; finally we construct η concatenating the excursions.

Let $\zeta = \big((\zeta_k(j))_{j\in\mathbb{Z}}\big)_{k\geq 1}$ belong to the set (10). We construct a slot-diagram $x = x[\zeta]$ as follows. Set

$$M(x) := \sup\{k \geq 0 : \zeta_k(0) > 0\} < +\infty\,, \tag{63}$$

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

$$\begin{aligned} s_k &= 1,\ \text{for } k \geq m, \\ x_k(0) &= 0,\ \text{for } k > m \\ x_m(0) &= \zeta_m(0) \end{aligned} \tag{64}$$

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

$$\begin{aligned} |x_\ell| &= \sum_{j=0}^{s_\ell-1} x_\ell(j) \\ s_k &= 1+2\sum_{\ell>k}(\ell-k)|x_\ell| \\ x_k(j) &= \zeta_k(j), \quad j=0,\dots,s_k-1. \end{aligned} \tag{65}$$

We have constructed a slot diagram

$$x := \big(x_k(0),\dots,x_k(s_k-1)\big)_{k\ge 1}. \tag{66}$$

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

$$\phi\zeta = (\tau^{s_k(\zeta)}\zeta_k)_{k\ge 1}\,. \tag{67}$$

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 (67) acts on the index labeling the slots, more precisely

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

Since $s_k=1$ for all $k\ge m$, we have $(\phi\zeta)_k(i)=\zeta_k(i+1)$ for all $k\ge m$. Hence, since ζ belongs to the set (10), so does $\phi\zeta$ and we can define iteratively

$$x^i := x[\phi^i\zeta], \quad i\ge 0. \tag{68}$$

For negative i let ζ' be the reflection of ζ with respect to the origin $\zeta'(i):=\zeta(-i)$ and define

$$x^i := (x[\phi^{-i-1}\zeta'])', \quad i<0\,, \tag{69}$$

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 (69) 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

$$\varepsilon^i := \varepsilon[x^i], \quad \underline{\varepsilon} = (\varepsilon^i)_{i\in\mathbb{Z}}. \tag{70}$$

Lemma 10. *The configuration $\eta=\eta[\underline{\varepsilon}]$ satisfies $D\eta=\zeta$.*

See [2] for a proof of this Lemma. This implies that D is a bijection between (9) and (10) and we can write $\eta=D^{-1}\zeta$.

5.3 From slot diagrams to components

We now construct a vector of components starting with a sequence of slot diagrams.

Let $\underline{x} = (x^i)_{i\in\mathbb{Z}}$ be a sequence of slot diagrams with s_k^i = number of k-slots in x^i. Define

$$S_k^0 = 0; \quad S_k^{i+1} - S_k^i = s_k^i \tag{71}$$

Let $\zeta = \zeta[\underline{x}]$ be defined by

$$\zeta(S^i + j) = x^i(j), \quad j = 0, \dots, s_k^i - 1. \tag{72}$$

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:

$$\zeta = D\big[\eta\big[(\varepsilon[x^i])_{i\in\mathbb{Z}}\big]\big] \tag{73}$$

5.4 Independent components with i.i.d. entries

We present the inverse statement of Proposition 5. Let $q = (q_k)_{k\geq 1}$, $q_k \in (0,1]$ satisfy

$$\sum_{k\geq 1} q_k < \infty \tag{74}$$

and define

$$\alpha_k := q_k \prod_{\ell=1}^{k-1} (1-q_\ell)^{2(k-\ell)}, \quad k = 1, 2, \dots \tag{75}$$

Let $(\zeta_k)_{k\geq 1}$ be independent elements with $\zeta_k = (\zeta_k(j))_{j\in\mathbb{Z}}$ i.i.d. random variables with distribution Geometric$(1-q_k)$. Since $q_k = P(\zeta_k(0) > 0)$, ζ satisfies (63) almost surely, which in turn implies that we can construct $D^{-1}\zeta$ using §5.2. Recall $x[\zeta]$ is the slot diagram of Excursion 0 in $D^{-1}\zeta$. We have the following Proposition.

Proposition 11. *Let $(q_k)_{k\geq 1}$ satisfy (74) and $\zeta = (\zeta_k)_{k\geq 1}$ be independent random vectors with $(\zeta_k(j))_{j\in\mathbb{Z}}$ Geometric$(1-q_k)$ i.i.d. random variables. Then the excursion $\varepsilon[x[\zeta]]$ has distribution ν_α defined in (19) with $(\alpha_k)_{k\geq 1}$ given by (75).*

Proof. By definition of ζ, for fixed s and x_k we have

$$P\big(\zeta_k(0) = x_k(0), \dots, \zeta_k(s-1) = x_k(s-1)\big) = \prod_{j=0}^{s-1} q_k^{x_k(j)}(1-q_k) = q_k^{|x_k|}(1-q_k)^s, \tag{76}$$

and using that $s_k(\zeta)$ is a function of $(\zeta_\ell)_{\ell>k}$ which is independent of ζ_k, we have that for any slot diagram x,

$$P(x[\zeta] = x) = \Big(\prod_{\ell>M}(1-q_\ell)\Big) \prod_{k=1}^{M} q_k^{|x_k|}(1-q_k)^{s_k} \tag{77}$$

$$= \Big(\prod_{\ell\geq 1}(1-q_\ell)\Big) \prod_{k=1}^{M} q_k^{|x_k|}(1-q_k)^{s_k-1} \tag{78}$$

where $M = \min\{k \geq 0 : x_{k'}(0) = 0 \text{ for all } k' > k\}$ and using the notation of §5.2. Since $s_\ell = 1 + 2\sum_{k>\ell}(k-\ell)|x_k|$ we can write (78) as

$$P(x[\zeta] = x) = \Big(\prod_{n\geq 1}(1-q_n)\Big) \prod_{k=1}^{M} \Big[q_k \prod_{\ell=1}^{k-1}(1-q_\ell)^{2(k-\ell)}\Big]^{|x_k|}. \tag{79}$$

Comparing this expression with (19), taking account of (61) and recalling (55) we see that Excursion 0 of $D^{-1}\zeta$ has distribution ν_α, with $(\alpha_k)_{k\geq 1}$ given by (75). □

Remark 12. *The excursions ε^i of $D^{-1}\zeta$ are independently and identically distributed all having distribution ν_α. In view of* (68) *this is equivalent to show that $x[\phi^i\zeta]$ are i.i.d.*

Remark 13. *Proposition 11 together with Proposition 5 imply that the transformations* (20),(21) *and* (75) *define a bijection between the $(\alpha_k)_{k\geq 1}$ satisfying* (18) *and the $(q_k)_{k\geq 1}$ satisfying* (52)*.*

6 Proof of Theorem 2

We now prove Theorem 2. Take independent random excursions $(\varepsilon^i)_{i\in\mathbb{Z}}$ satisfying the hypothesis of the Theorem and let $x^i = x[\varepsilon^i]$ be the corresponding slots diagrams. By Proposition 5 x^i satisfies (35) and (36), that is, given the number of k-slots $(s^i)_k$, the variables $((x^i)_k)_0^{(s^i)_k-1}$ are i.i.d. Geometric$(1-q_k)$. Let $\mathcal{F}_k$ be the sigma field generated by $((x^i)_k)_{i\in\mathbb{Z}}$ and $\mathcal{F}_{>k}$ the sigma field generated by $\big((x^i)_{k+1}^\infty : i \in \mathbb{Z}\big)$. Condition on $\mathcal{F}_{>k}$ and construct ζ_k using (72), 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$(1-q_k)$. This implies that $\mathcal{F}_k$ and $\mathcal{F}_{>k}$ are independent concluding the proof of the first part of the Theorem.

The measure $\widehat{\mu}_\alpha$ is Record shift-invariant because the excursions ε_i are i.i.d.. This and the law of large numbers give the $\widehat{\mu}_\alpha$-a.s. limit

$$\lim_{n\to+\infty} \frac{\#\{\text{balls in } \eta \text{ in } [0,n]\}}{n} = \lim_{k\to+\infty} \frac{\frac{\#\big\{\text{balls in } (\varepsilon_i)_{i=0}^{k-1}\big\}}{k}}{\frac{2\#\big\{\text{balls in } (\varepsilon_i)_{i=0}^{k-1}\big\}+k}{k}} = \frac{\frac{\rho}{2}}{\rho+1} = \lambda\,.$$

The measure μ_α is shift-invariant by definition of Palm measures and its inverse [6]. The fact that μ_α is T-invariant will follow by Theorem 1 once we show that (11) is satisfied. But, (22) implies (24) which together with excursion independence imply that for any fixed k-slot j the following limit exists $\widehat{\mu}_\alpha$-a.s. and

$$E\zeta_k(j) = \lim_{n\to\infty} \frac{\#\{k\text{-solitons in } \eta \text{ in } [0,n]\}}{\#\{k\text{-slots in } \eta \text{ in } [0,n]\}} = \frac{\rho_k}{1+\sum_{m>k} 2(m-k)\rho_m}. \tag{80}$$

The last expression is the mean number of k-solitons per excursion divided by the mean number of k-slots per excursion, see also [2] §3.2. Hence $\sum_k kE\zeta_k(j) \leq \sum_k k\rho_k < \infty$, by (24). This is (11) and concludes the proof of Theorem 2.

Note that using (25) the condition (11) is equivalent to

$$\sum_{k=1}^{+\infty} kq_k < +\infty\,. \tag{81}$$

Hence, the condition that the excursion is finite almost surely (18) is equivalent to (74) which says $M < \infty$ when written in terms of the q's. Condition (22) (bounded expected excursion length under ν_α) is equivalent to (81).

Acknowledgments

This project started when the first author was visiting Gran Sasso Scientific 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.

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] Ferrari, P. A., Nguyen, C., Rolla, L. T., and Wang, M. Soliton decomposition of the box-ball-system in Z. Preprint. arXiv:1806.02798, 2018.

[3] Levine, L., Lyu, H., and Pike, J. Phase transition in a random soliton cellular automaton. Preprint. arXiv:1706.05621, 2017.

[4] 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.

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

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