The T, T^{-1}–process, finitary codings and weak Bernoulli

JEFFREY E. STEIF
Georgia Institute of Technology

June 4, 2001

Abstract

We give an elementary proof that the second coordinate (the scenery process) of the T, T^{-1}–process associated to any mean zero i.i.d. random walk on $\mathbf{Z}^d$ is not a finitary factor of an i.i.d. process. In particular, this yields an elementary proof that the basic T, T^{-1}–process is not finitarily isomorphic to a Bernoulli shift (the stronger fact that it is not Bernoulli was proved by Kalikow). This also provides (using past work of den Hollander and the author) an elementary example, namely the T, T^{-1}–process in 5 dimensions, of a process which is weak Bernoulli by not a finitary factor of an i.i.d. process. An example of such a process was given earlier by del Junco and Rahe. The above holds true for arbitrary stationary recurrent random walks as well. On the other hand, if the random walk is Bernoulli and transient, the T, T^{-1}–process associated to it is also Bernoulli. Finally, we show that finitary factors of i.i.d. processes with finite expected coding volume satisfy certain notions of weak Bernoulli in higher dimensions which have been previously introduced and studied in the literature. In particular, this yields (using past work of van den Berg and the author) the fact that the Ising model is weak Bernoulli throughout the subcritical regime.

1 Introduction

S. Kalikow [13] proved that the T, T^{-1}–process associated to simple symmetric random walk in 1 dimension (to be defined below) is not a Bernoulli shift, solving a problem that had been open for over 10 years. A relatively extensive study of the T, T^{-1}–processes associated to arbitrary (i.i.d.) random walks on $\mathbf{Z}^d$ was done in [11], studying how the properties of Bernoulli and weak Bernoulli are reflected in the behavior of the underlying random walk.

We first give the definition of the general T, T^{-1}–process. We will be slightly terse; the reader may refer to [11] for full details. For a fixed integer $d \geq 1$, let $\{X_i\}_{i \in \mathbf{Z}}$ be a stationary process taking values in $\mathbf{Z}^d$. Let $\{S_n\}_{n \in \mathbf{Z}}$ be the corresponding **random walk** on $\mathbf{Z}^d$ defined by

$$S_0 = 0, \; S_n = \sum_{i=1}^{n} X_i \; (n \geq 1), \; S_n = -\sum_{i=n+1}^{0} X_i \; (n \leq -1).$$

Next, let $\{C_z\}_{z \in \mathbf{Z}^d}$ be i.i.d. random variables taking values 1 and 0 each with probability 1/2.

Now consider the process

$$\{Z_i\}_{i \in \mathbf{Z}} \text{ where } Z_i = (X_i, C_{S_i}),$$

which we call the T, T^{-1}–**process associated with** $\{X_i\}_{i \in \mathbf{Z}}$. It is easy to see that $\{Z_i\}_{i \in \mathbf{Z}}$ is a stationary process; it is essentially a so-called skew-product. Note, importantly, that even if the random walk is living in $\mathbf{Z}^d$, the T, T^{-1}–process is always a process indexed by $\mathbf{Z}$.

Both in [13] and in [11], only the case where $\{X_i\}_{i \in \mathbf{Z}}$ is an i.i.d. process was considered while in this paper, we do not always make this assumption. We assume that the reader is familiar with the notions of (a) Bernoulli shifts, (b) weak Bernoulli (c) factor maps and (d) finitary factor maps (see e.g. [2] and [11] which together contain all of these definitions). By a *finite-valued process*, we mean a process which takes on only a finite number of values.

Our first result is the following.

Theorem 1.1 *Let $\{X_i\}_{i \in \mathbf{Z}}$ be an i.i.d. process taking values in $\mathbf{Z}^d$ and having mean 0. Then there is no finitary factor map from an (not necessarily finite-valued) i.i.d. process to $\{C_{S_i}\}_{i \in \mathbf{Z}}$, (i.e., to the second coordinate of the T, T^{-1}–process associated with $\{X_i\}_{i \in \mathbf{Z}}$).*

Remarks:

(a) In particular, this yields the fact that the basic T, T^{-1}–process (i.e., the one associated to simple symmetric random walk in 1 dimension) is not finitarily isomorphic to a Bernoulli shift, a weakening of Kalikow's theorem but one which is much more easily proved.

(b) One of the main results in [11] implies, as a special case, that the T, T^{-1}–process associated to simple symmetric random walk in $d \geq 5$ dimensions is weak Bernoulli. In view of Theorem 1.1, this yields an elementary example of a process which is weak Bernoulli but not a finitary factor of an i.i.d. process. The first example of such a process was given in [12] where a type of iterative construction is used.

(c) In order to continue the flow, we will give at the end of this introduction further comments concerning the above result.

Our second result is the following.

Theorem 1.2 *Let $\{X_i\}_{i \in \mathbf{Z}}$ be a stationary ergodic process taking values in $\mathbf{Z}^d$ and assume that $\{S_n\}_{n \in \mathbf{Z}}$ is recurrent; i.e., $P(S_n = 0$ for some $n > 0) = 1$. Then there is no finitary factor map from an (not necessarily finite-valued) i.i.d. process to $\{C_{S_i}\}_{i \in \mathbf{Z}}$.*

Remark: It is known (see [7], p. 347) that all mean 0 stationary ergodic processes taking values in $\mathbf{Z}$ are recurrent and so Theorem 1.2 is applicable in such cases. However, in Z^2, there exist stationary, ergodic, mean 0 random walks whose step distribution is finite but which are not recurrent (see [5]).

Theorems 1.1 and 1.2 will be proved in Section 2.

We mention another result which, while related to the above results in this paper, is a slight tangent. In [11], it was shown that the T, T^{-1}–process associated to any transient i.i.d. random walk is Bernoulli. The following is an extension of this.

Theorem 1.3 *Let $\{X_i\}_{i \in \mathbf{Z}}$ be a stationary process taking values in $\mathbf{Z}^d$ which is a Bernoulli shift and assume that $\{S_n\}_{n \in \mathbf{Z}}$ is transient; i.e., $P(S_n = 0$ for some $n > 0) < 1$. Then the T, T^{-1}–process associated to it is a Bernoulli shift.*

The proof given in [11] for the i.i.d. case (which verifies the so-called very weak Bernoulli condition) does not work here quite as nicely and we instead easily prove this by explicitly constructing a factor map from a Bernoulli shift to our process. In fact, in Section 3, we will construct a general class of processes, which includes all of the T, T^{-1}–processes as well as certain random fields indexed by $\mathbf{Z}^d$, and we will obtain a general result which yields Theorem 1.3 as a special case.

The second part of this paper deals with the relationship between finitary codings and weak Bernoulli in higher dimensions. We first need the following definition. Given a finitary mapping from a process $\{X_i\}_{i\in\mathbf{Z}}$ to a process $\{Y_i\}_{i\in\mathbf{Z}}$, we let T_j be defined (as usual) as the smallest integer such that the interval around j of radius T_j in the process $\{X_i\}_{i\in\mathbf{Z}}$ determines Y_j. If $E[T_0] < \infty$, we say that the finitary mapping has *finite expected coding length*. As explained in [12], it is elementary to show that if $\{Y_i\}_{i\in\mathbf{Z}}$ is a finitary factor of an i.i.d. process with finite expected coding length, then $\{Y_i\}_{i\in\mathbf{Z}}$ is weak Bernoulli. (A proof of this is given in this reference as well.) We want to extend such a result to higher dimensions. Finitary codings in higher dimensions are defined analogously to 1 dimension (see again [2]) and the definition of T_j is obtained by simply replacing the word *interval* by *box*. While it is not at first clear what should replace the condition $E[T_0] < \infty$, after some reflection, it is not hard to convince oneself that the correct replacement of this for $\mathbf{Z}^d$ is $E[T_0^d] < \infty$. However, what is much more subtle is how one generalizes the definition of weak Bernoulli to higher dimensions. It turns out that the first definitions one usually comes up with are inappropriate for various reasons; this is discussed at length in [4]. Given in this reference are the definitions of quite weak Bernoulli (QWB) and quite weak Bernoulli with exponential rate (QWBE) (which were originally introduced and studied in [3]) as well as a proposal for the correct definition of weak Bernoulli (WB) in higher dimensions. In Section 4, we will review these definitions and prove the following three results.

Theorem 1.4 *Let $\{Y_i\}_{i\in\mathbf{Z}^d}$ be a finitary factor of the (not necessarily finite-valued) i.i.d. process $\{X_i\}_{i\in\mathbf{Z}^d}$. If $E[T_0^d] < \infty$, then $\{Y_i\}_{i\in\mathbf{Z}^d}$ is QWB.*

Theorem 1.5 *Let $\{Y_i\}_{i\in\mathbf{Z}^d}$ be a finitary factor of the (not necessarily finite-valued) i.i.d. process $\{X_i\}_{i\in\mathbf{Z}^d}$. If $E[e^{\delta T_0}] < \infty$ for some $\delta > 0$, then $\{Y_i\}_{i\in\mathbf{Z}^d}$ is QWBE.*

Theorem 1.6 *Let $\{Y_i\}_{i\in\mathbf{Z}^d}$ be a finitary factor of the (not necessarily finite-valued) i.i.d. process $\{X_i\}_{i\in\mathbf{Z}^d}$. If $E[T_0^d]<\infty$, then $\{Y_i\}_{i\in\mathbf{Z}^d}$ is WB.*

Theorems 1.4, 1.5, and 1.6 will be proved in Section 4, and an application to the subcritical Ising model will also be given.

We make some final comments concerning Theorem 1.1. As mentioned in the first sentence of the paper, Kalikow proved that the T,T^{-1}–process associated to simple symmetric random walk in 1 dimension is not a Bernoulli shift. Hence, since factors of Bernoulli shifts are Bernoulli shifts (see [15]), there cannot be a finitary factor map from an i.i.d. process onto it. One cannot however immediately conclude that there also cannot be a finitary factor from an i.i.d. process onto $\{C_{S_i}\}_{i\in\mathbf{Z}}$ since it is not clear that the latter is not a Bernoulli shift. The fact that $\{Z_i\}_{i\in\mathbf{Z}}$ is not a Bernoulli shift and that the first coordinate is an i.i.d. process leads one to think that the second coordinate itself should then also not be a Bernoulli shift but the main result in [16] implies precisely that such reasoning is not correct. However, C. Hoffman [10] has in fact shown that $\{C_{S_i}\}_{i\in\mathbf{Z}}$ is not a Bernoulli shift. Hence Theorem 1.1 for simple symmetric random walk in 1 dimension does in fact follow from known results. Of course, Theorem 1.1 is most interesting in higher dimensions where the T,T^{-1}–process is Bernoulli or even weak Bernoulli.

2 Proofs of Theorems 1.1 and 1.2

In this section, we give simple proofs of Theorems 1.1 and 1.2. The first lemma which we need follows from the results in [14]. This is also explicitly discussed in [2] where it is stated that the result extends to higher dimensions and is used there to show that certain Markov random fields which are Bernoulli are not a finitary factor of an i.i.d. process. However, as we only need the result for 1 dimension, we state it only for that case. In words, it says that finitary factors of i.i.d. processes satisfy *standard large deviation behavior.*

Lemma 2.1 *Let $\{Y_i\}_{i\in\mathbf{Z}}$ be a real valued process taking on a finite number of values which is a finitary factor of an (not necessarily finite-valued) i.i.d. process $\{X_i\}_{i\in\mathbf{Z}}$. Then, for every*

$\varepsilon > 0$, there exist positive constants c_ε and γ_ε such that for all $n \geq 1$

$$P(|\frac{1}{n}\sum_{i=0}^{n-1} Y_i - E[Y_0]| \geq \varepsilon) \leq c_\varepsilon e^{-n\gamma_\varepsilon}.$$

The next lemma is a key ingredient.

Lemma 2.2 *Let $\{X_i\}_{i\in\mathbf{Z}}$ be an i.i.d. process taking values in $\mathbf{Z}^d$ and having mean 0, $\{S_n\}$ defined as usual, and $R_n := |\{S_0, S_1, \ldots, S_{n-1}\}|$ be the cardinality of the range of the random walk up until time $n-1$. Then for every $\varepsilon > 0$, it is the case that*

$$P(\frac{R_n}{n} \leq \varepsilon) \geq (\frac{1}{2})^{\varepsilon n}$$

holds for large n.

Proof: If the distribution of the X_i's is compact (or even satisfies much weaker assumptions), then this result follows from Lemma 5.1 in [6] which is a much stronger statement. See also [9] for related results. To extend to the general mean 0 case, we use a truncation argument. We need the following lemma which is proved afterwards.

Lemma 2.3 *Consider a probability distribution μ on $\mathbf{Z}^d$ whose mean is 0. Then given $\delta > 0$, there exists a probability distribution ν on $\mathbf{Z}^d$ which has compact support, mean 0, and such that*

$$\mu = (1-\delta)\nu + \delta\nu' \tag{1}$$

for some probability distribution ν' on $\mathbf{Z}^d$.

Given the result for the case of compactly supported distributions together with Lemma 2.3, we proceed with the proof of Lemma 2.2 as follows. Let $\varepsilon > 0$. Choose $\delta > 0$ such that $(1-\delta) \geq (1/2)^{\varepsilon/2}$. Denoting the distribution of X_i by μ, we can, by Lemma 2.3, choose a probability distribution ν having compact support, mean 0 and satisfying (1) for some probability distribution ν' on $\mathbf{Z}^d$. Letting $\{R_n^*\}$ denote the range process for the random walk whose steps are distributed according to ν, we can (as ν has compact support and we already know the result in this case) find N such that for all $n \geq N$,

$$P(\frac{R_n^*}{n} \leq \varepsilon) \geq (\frac{1}{2})^{\frac{\varepsilon n}{2}}.$$

It follows that for all $n \geq N$,

$$P(\frac{R_n}{n} \leq \varepsilon) \geq (1-\delta)^n P(\frac{R_n^*}{n} \leq \varepsilon) \geq (\frac{1}{2})^{\varepsilon n}$$

as desired. $\square$

Proof of Lemma 2.3: Fix a probability measure μ on $\mathbf{Z}^d$ with mean 0. We assume that μ does not live on a hyperplane intersected with $\mathbf{Z}^d$ for in that case, we could do the argument on a lower dimensional lattice. Fix $\delta > 0$. Given a map $p : \mathbf{Z}^d \to [0,1]$ with the property that there is a $y \in \mathbf{Z}^d$ with $p(y)\mu(y) > 0$, denote by μ_p the probability measure on $\mathbf{Z}^d$ given by

$$\mu_p(x) = \frac{p(x)\mu(x)}{\sum_{y \in \mathbf{Z}^d} p(y)\mu(y)}.$$

Note that if $p \equiv 1$, then $\mu_p = \mu$. Let $\mathcal{P}$ denote the set of p's as above which take on the value 0 for all but finitely many $x \in \mathbf{Z}^d$ and such that $\sum_{y \in \mathbf{Z}^d} p(y)\mu(y) = 1 - \delta$. Next, let $\mathcal{M}$ denote $\{\mu_p : p \in \mathcal{P}\}$. Clearly any $\nu \in \mathcal{M}$ has compact support and can be expressed as in (1) for some distribution ν' on $\mathbf{Z}^d$. We now show that some $\nu \in \mathcal{M}$ has mean 0, or equivalently that $\mathcal{M}' := \{E[\mu_p] : p \in \mathcal{P}\}$ contains the 0 vector, which will complete the proof. It is clear that $\mathcal{P}$ is a convex set and it is easy to check that for all $p, p' \in \mathcal{P}$, for all $\alpha \in [0,1]$,

$$\mu_{\alpha p + (1-\alpha)p'} = \alpha\mu_p + (1-\alpha)\mu_{p'}.$$

(This last equality is not necessarily true if $\sum_{y \in \mathbf{Z}^d} p(y)\mu(y) \neq \sum_{y \in \mathbf{Z}^d} p'(y)\mu(y)$.) Since expectation is linear, it follows that $\mathcal{M}'$ is a (not necessarily closed) convex set in $\mathbf{R}^d$. If $\mathcal{M}'$ does not contain 0, it would follow from the separating hyperplane theorem, that there exists a hyperplane H in $\mathbf{R}^d$ such that $\mathcal{M}'$ is on one side of (or on) H. We show however that $\mathcal{M}'$ contains points on both sides of H. Calling the two sides S_1 and S_2, choose $x \in S_2$ such that $\mu(x) > 0$ (the fact that μ does not live on a hyperplane and has mean 0 guarantees such an x). Let p be $1 - \delta/2$ on x and 1 elsewhere. Then $E[\mu_p]$ is in S_1. Now, clearly we can choose R sufficiently large so that if we modify p to be 0 for all points which are R or further away from the origin, we will obtain a (modified) p for which $E[\mu_p] \in S_1$, μ_p has compact support and $\sum_{y \in \mathbf{Z}^d} p(y)\mu(y) \geq 1 - \delta$. By multiplying this modified p by $(1-\delta)/\sum_{y \in \mathbf{Z}^d} p(y)\mu(y)$, we obtain a $p \in \mathcal{P}$ with $E[\mu_p] \in S_1$. By symmetry of the argument, $\mathcal{M}'$ contains points on both sides of H, a contradiction. $\square$

Proof of Theorem 1.1: Assume there exists a finitary coding from an i.i.d. process to $\{C_{S_i}\}_{i\in\mathbf{Z}}$. Lemma 2.1 now implies that there exists $\gamma > 0$ so that

$$P(|\frac{1}{n}\sum_{i=0}^{n-1} C_{S_i} - \frac{1}{2}| \geq \frac{1}{4}) \leq e^{-\gamma n}$$

for large n. Choose $\epsilon > 0$ so that

$$(\frac{1}{2})^{2\epsilon} > e^{-\gamma}.$$

If the colors at all the points of $\{S_0, S_1, \dots, S_{n-1}\}$ are 1, then certainly it follows that $|\frac{1}{n}\sum_{i=0}^{n-1} C_{S_i} - \frac{1}{2}| \geq \frac{1}{4}$. It follows that

$$P(|\frac{1}{n}\sum_{i=1}^{n} C_{S_i} - \frac{1}{2}| \geq \frac{1}{4}) \geq P(R_n \leq \epsilon n)(\frac{1}{2})^{\epsilon n}$$

which in turn is at least $(\frac{1}{2})^{2\epsilon n}$ for large n by Lemma 2.2. Hence $(\frac{1}{2})^{2\epsilon n} \leq e^{-\gamma n}$ for large n, contradicting the definition of ϵ and completing the proof. □

Remark: If we consider any symmetric d dimensional random walk with or without a finite mean, the argument given above easily yields the fact that the scenery process for the associated T, T^{-1}–process is not a finitary coding of an i.i.d. process. In particular, if in 1 dimension, we have a random walk where a step of size x has probability proportional to $1/|x|^\alpha$ for $x \neq 0$ with $\alpha \in (1, 3/2)$, then (using [11]) the associated T, T^{-1}–process is weak Bernoulli, giving us another example of a process which is weak Bernoulli but not a finitary coding of an i.i.d. process.

The next lemma can be found in [7], p. 346.

Lemma 2.4 *Let $\{X_i\}_{i\in\mathbf{Z}}$ be a stationary ergodic process taking values in $\mathbf{Z}^d$, $\{S_n\}$ defined as usual, and $R_n := |\{S_0, S_1, \dots, S_{n-1}\}|$ as above. If $P(S_n = 0$ for some $n > 0) = 1$ (i.e., if $\{S_n\}_{n\in\mathbf{Z}}$ is recurrent), then $\lim_{n\to\infty} R_n/n = 0$ a.s.*

Proof of Theorem 1.2: The proof of Theorem 1.1 essentially goes through here. The only change is that one notes that since $\lim_{n\to\infty} R_n/n = 0$ a.s. by Lemma 2.4, it immediately follows that for any $\varepsilon > 0$,

$$\lim_{n\to\infty} P(R_n \leq \epsilon n) = 1.$$

The rest of the proof is identical. □

3 A general class of stationary processes

We now introduce a general class of processes, indexed by Z^n, which has some resemblance to percolation theory when $n \geq 2$ but also contains all of the T, T^{-1}–processes as special cases in which case $n = 1$ (even though the random walk for the T, T^{-1}–process might be d–dimensional).

Let $\{X_k\}_{k \in \mathbf{Z}^n}$ be an arbitrary stationary process. Assume that each realization of this process defines for us an equivalence relation on $\mathbf{Z}^n$ in such a way that this rule is translation invariant. We call this a random equivalence relation. (Two examples below will clarify this definition.) Consider the following process. We first choose a realization from $\{X_k\}$. Then for each equivalence class C, assign all the elements of C the same value chosen from $\{1, \ldots, q\}$ (q is an integer here) each with probability $1/q$. Do this independently for different equivalence classes. The process we then consider is $\{(X_k, Y_k)\}_{k \in \mathbf{Z}^n}$ where Y_k is that value from $\{1, \ldots, q\}$ assigned to k in this two step procedure. It is straightforward to verify that this is a stationary process.

Examples:

(a) Let $n = 1$ and $\{X_k\}_{k \in \mathbf{Z}}$ be a stationary process taking values in $\mathbf{Z}^d$. For a realization from $\{X_k\}$, let, for $s > r$, $s \sim r$ if $\sum_{i=r+1}^{s} X_i = 0$. It is immediate that this is a shift invariant equivalence relation and moreover, a moment of thought shows that the process $\{(X_k, Y_k)\}_{k \in \mathbf{Z}^n}$ constructed above with $q = 2$ is precisely the T, T^{-1}–process associated to $\{X_k\}$.

(b) Let $n \geq 1$ and $\{X_k\}_{k \in \mathbf{Z}^n}$ be a stationary process taking values in a finite set. Given a realization from $\{X_k\}$, let $s \sim r$ if one can find a nearest neighbor path in $\mathbf{Z}^n$ from s to r on which $\{X_k\}$ is constant. It is immediate that this is a shift invariant equivalence relation; the study of this particular equivalence relation is exactly the study of percolation theory. Various aspects of the second coordinate of the resulting process $\{(X_k, Y_k)\}_{k \in \mathbf{Z}^n}$ for this equivalence relation have been studied in [8] (although this equivalence relation point of view was not taken there).

Theorem 3.1 *Let $\{X_k\}_{k \in \mathbf{Z}^n}$ be an arbitrary stationary process which is a Bernoulli shift. Assume that we have a random equivalence relation (defined exactly as above) such that a.s. all* *equivalence classes are finite. Then the corresponding process $\{(X_k, Y_k)\}_{k\in\mathbf{Z}^n}$ is a Bernoulli shift.*

Proof of Theorem 1.3: By example (a), our process is of the type described in Theorem 3.1 and moreover, it is immediate to check that the transience assumption implies (and is implied by) that all equivalence classes are finite. The result now follows from Theorem 3.1. $\square$

Proof of Theorem 3.1: Assume that the $\{X_k\}_{k\in\mathbf{Z}^n}$ process takes values in the set F. Consider an i.i.d. process $\{U_k\}_{k\in\mathbf{Z}^n}$ which takes values in $\{1,\ldots,q\}$ each with probability $1/q$. Taking these two processes to be independent, we obtain a process $\{(X_k, U_k)\}_{k\in\mathbf{Z}^n}$ which is a Bernoulli shift since a product of Bernoulli shifts is a Bernoulli shift. We now construct an explicit factor map from this latter process to the process $\{(X_k, Y_k)\}_{k\in\mathbf{Z}^n}$. The conclusion then follows from the fact that factors of Bernoulli shifts are Bernoulli shifts (see [15]).

Given a finite set $S \subseteq \mathbf{Z}^n$, let $\alpha(S)$ denote the smallest element of S in the lexocographic ordering on $\mathbf{Z}^n$. Also, given $x = \{x_k\}_{k\in\mathbf{Z}^n} \in F^{\mathbf{Z}^n}$ and $\ell \in \mathbf{Z}^n$, let $S_x(\ell)$ denote the equivalence class of ℓ under the realization x. Consider now the factor map

$$f : (F \times \{1,\ldots,q\})^{\mathbf{Z}^n} \to (F \times \{1,\ldots,q\})^{\mathbf{Z}^n}$$

given by

$$f((x_k, u_k)_{k\in\mathbf{Z}^n})_\ell := (x_\ell, u_{\alpha(S_x(\ell))})$$

where $x = \{x_k\}_{k\in\mathbf{Z}^n}$. It is immediate that f maps down to the process $\{(X_k, Y_k)\}_{k\in\mathbf{Z}^n}$, as desired. $\square$

The general class of processes introduced in this section are worth studying in their own right, in particular in investigating the ergodic properties of the process $\{(X_k, Y_k)\}_{k\in\mathbf{Z}^n}$ in terms of the given equivalence relation. This might be studied in a subsequent paper.

4 Proofs of Theorems 1.4, 1.5 and 1.6

We begin by giving the definitions of QWB, QWBE and WB which come from [3] and [4]. First we need some preliminaries.

Let $\{Y_i\}_{i\in I}$ be random variables defined on the same probability space taking values in a finite set F where I is countable (possibly finite). As usual, we will identify such a family of random variables with a probability measure μ on F^I (its distribution). Similarly, for $U \subseteq I$, we let μ_U denote the measure on F^U induced by the random variables $\{Y_i\}_{i\in U}$ and identify a stationary process $\{Y_i\}_{i\in \mathbf{Z}^d}$ taking values in the set F with a (shift invariant) probability measure μ on $F^{\mathbf{Z}^d}$. We also let $\mu_U(\cdot|A)$ denote μ_U conditioned on an event A, $\|\ \|$ denote the total variation norm of a finite signed measure, $B(n)$ denote $[-n,n]^d \cap \mathbf{Z}^d$, $B(x,n)$ denote $x+B(n)$ and for $x=(x_1,\ldots,x_d)\in \mathbf{Z}^d$, we let $|x| := \sum_i |x_i|$.

Definition 4.1 *A translation invariant probability measure μ on $F^{\mathbf{Z}^d}$ is called* **quite weak Bernoulli (QWB)** *if for all $\varepsilon > 0$,*

$$\lim_{n\to\infty} \|\mu|_{(\mathbf{Z}^d\setminus B(n))\cup B(n(1-\varepsilon))} - \mu|_{\mathbf{Z}^d\setminus B(n)} \times \mu|_{B(n(1-\varepsilon))}\| = 0.$$

Definition 4.2 *A translation invariant probability measure μ on $F^{\mathbf{Z}^d}$ is called* **quite weak Bernoulli with exponential rate (QWBE)** *if for all $\varepsilon > 0$, there exist constants $\gamma_\varepsilon > 0, c_\varepsilon > 1$ so that*

$$\|\mu|_{(\mathbf{Z}^d\setminus B(n))\cup B(n(1-\varepsilon))} - \mu|_{\mathbf{Z}^d\setminus B(n)} \times \mu|_{B(n(1-\varepsilon))}\| \le c_\varepsilon e^{-\gamma_\varepsilon n}$$

for all n.

Extensive discussions concerning the notion of WB in higher dimensions and why *natural* extensions of the 1 dimensional definition turn out to be uninteresting are given in [4]. Moreover, in [4], a definition of WB is proposed which is partially motivated by the work of Berbee (see [1]) who gives an equivalent definition of WB in 1 dimension in terms of couplings. The definition of WB given in [4] is the following; see also [4] for the idea behind this definition and the fact that in 1 dimension, this is equivalent to the usual definition.

Definition 4.3 *A translation invariant probability measure μ on $F^{\mathbf{Z}^d}$ is called* **weak Bernoulli** *(WB) if there exists a nonnegative integer valued stationary process indexed by $\mathbf{Z}^{d-1}$, $\{C_m\}_{m\in\mathbf{Z}^{d-1}}$ so that for all n, there exists a coupling $(\sigma_1, \sigma_2, \{\tilde{C}^1_m\}_{m\in\mathbf{Z}^{d-1}}, \ldots, \{\tilde{C}^{2d}_m\}_{m\in\mathbf{Z}^{d-1}})$ of 2 copies of the distribution of μ and $2d$ copies of the distribution of $\{C_m\}_{m\in\mathbf{Z}^{d-1}}$ (where we suppress the*

dependence on n in the notation) so that
(1) σ_1 and $\sigma_2|_{B(n)^c}$ are independent and
(2) $\cap_{i=1}^{2d} A_i \subseteq \{x : \sigma_1(x) = \sigma_2(x)\}$
where for $i = 1, 2, \ldots, d$

$$A_i = \{x \in B(n) : \tilde{C}^i_{\hat{x}_i} \leq x_i + n\}$$

and for $i = d+1, d+2, \ldots, 2d$

$$A_i = \{x \in B(n) : \tilde{C}^i_{\hat{x}_{i-d}} \leq n - x_{i-d}\}$$

where $\hat{x}_j = \{x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_d\}$.
*For each n, we call the above coupling which depends on n the n**th coupling**.*

Proof of Theorem 1.4: Let $E_{n,\varepsilon}$ be the event that

$$\Big(\cup_{x \in B(n)^c} B(x, T_x)\Big) \cap B(n - \frac{\varepsilon n}{2}) \neq \emptyset$$

and $F_{n,\varepsilon}$ be the event that

$$\Big(\cup_{x \in B(n-\varepsilon n)} B(x, T_x)\Big) \cap B(n - \frac{\varepsilon n}{2})^c \neq \emptyset.$$

We claim that for any $\varepsilon > 0$,

$$\lim_{n \to \infty} P(E_{n,\varepsilon} \cup F_{n,\varepsilon}) = 0. \tag{2}$$

To see this, we first note that there is a constant c_d so that for each $\ell \geq 1$, $|\{x : |x| = \ell\}| \leq c_d \ell^{d-1}$. Now

$$P(E_{n,\varepsilon}) \leq \sum_{\ell=n+1} c_d \ell^{d-1} P(T_0 \geq \ell - n + \frac{\varepsilon n}{2}) \leq \sum_{\ell=\lceil \frac{\varepsilon n}{2}+1 \rceil} c_d (\ell + n)^{d-1} P(T_0 \geq \ell).$$

Next for $\ell \geq 1 + \varepsilon n/2$, $\ell + n \leq (1 + 2/\varepsilon)\ell$ and so the above is at most

$$\sum_{\ell=\lceil \frac{\varepsilon n}{2}+1 \rceil} c_\varepsilon \ell^{d-1} P(T_0 \geq \ell)$$

which approaches 0 as $n \to \infty$ since $E[T_0^d] < \infty$. The exact same argument shows that $\lim_{n\to\infty} P(F_{n,\varepsilon}) = 0$, yielding (2).

Next, it is easy to check that

$$\mu|_{(\mathbf{Z}^d\setminus B(n))\cup B(n(1-\varepsilon))}(\cdot|(E_{n,\varepsilon}\cup F_{n,\varepsilon})^c) = \mu|_{(\mathbf{Z}^d\setminus B(n))}(\cdot|(E_{n,\varepsilon}\cup F_{n,\varepsilon})^c)\times\mu|_{B(n(1-\varepsilon))}(\cdot|(E_{n,\varepsilon}\cup F_{n,\varepsilon})^c),$$

and hence it follows that

$$\|\mu|_{(\mathbf{Z}^d\setminus B(n))\cup B(n(1-\varepsilon))} - \mu|_{\mathbf{Z}^d\setminus B(n)} \times \mu|_{B(n(1-\varepsilon))}\|$$

$$\leq \|\mu|_{(\mathbf{Z}^d\setminus B(n))\cup B(n(1-\varepsilon))} - \mu|_{(\mathbf{Z}^d\setminus B(n))\cup B(n(1-\varepsilon))}(\cdot|(E_{n,\varepsilon}\cup F_{n,\varepsilon})^c)\|$$

$$+\|\mu|_{(\mathbf{Z}^d\setminus B(n))}(\cdot|(E_{n,\varepsilon}\cup F_{n,\varepsilon})^c)\times\mu|_{B(n(1-\varepsilon))}(\cdot|(E_{n,\varepsilon}\cup F_{n,\varepsilon})^c) - \mu|_{\mathbf{Z}^d\setminus B(n)}\times\mu|_{B(n(1-\varepsilon))}\|.$$

Now, it is simple to check that if ν is a probability measure and $\nu(E)\geq 1-\delta$, then $\|\nu-\nu(\cdot|E)\|\leq 2\delta$ and that if $\max\{\|\nu_1-\mu_1\|,\|\nu_2-\mu_2\|\}\leq\delta$, then $\|\nu_1\times\nu_2-\mu_1\times\mu_2\|\leq 2\delta$. Combining the above with (2), the result immediately follows. □

Proof of Theorem 1.5: By the proof of Theorem 1.4, it suffices to show that for any $\varepsilon>0$, $P(E_{n,\varepsilon}\cup F_{n,\varepsilon})\leq c_\varepsilon e^{-\gamma_\varepsilon n}$ for some constants $\gamma_\varepsilon>0, c_\varepsilon>1$. To obtain such a bound on $P(E_{n,\varepsilon})$, the same computation as in the proof of Theorem 1.4 shows that it suffices to show that given positive constants c_1 and c_2, there are positive constants c_3 and c_4 such that

$$\sum_{\ell=c_1 n} c_2\ell^{d-1}P(T_0\geq\ell)\leq c_3e^{-c_4 n}.$$

However, if $E[e^{\delta T_0}]<\infty$ for some $\delta>0$, then by Markov's inequality, $P(T_0\geq\ell)$ decays exponentially in ℓ and the above easily follows. Similarly, one obtains an exponential bound for $P(F_{n,\varepsilon})$, completing the proof. □

Proof of Theorem 1.6: We first define $2d$ (possibly different) processes, each indexed by $\mathbf{Z}^{d-1}$. These will correspond to the $2d$ halfspaces H whose boundary, defined by $\partial H=\{x\in H:\exists y\in H^c \text{ with } \|x-y\|=1\}$, contains 0 and is perpendicular to one of the coordinate axes. Given such a halfspace H, we first let $\{C_x^H\}_{x\in\partial H}$ be defined by

$$C_x^H := \max\{k : B(x+kv_H, T_{x+kv_H})\cap\cup_{y\in H}B(y,T_y)\neq\emptyset\}$$

where v_H is that unit vector in $\mathbf{Z}^d$ pointing away from H. It is left to the reader to easily verify (in the same way that $\lim_{n\to\infty}P(E_{n,\varepsilon})=0$ was proved in the proof of Theorem 1.4) that for each halfspace H and each $x \in \partial H$, $C_x^H < \infty$ a.s. This yields for us $2d$ processes, each indexed by $\mathbf{Z}^{d-1}$ since each ∂H can be canonically identified with $\mathbf{Z}^{d-1}$. These processes certainly could be different due to, for example, a lack of isotropy. Now, order the $2d$ halfspaces $(H_1, \ldots, H_{2d})$ so that their corresponding v_H's are ordered as $(e_1, e_2, \ldots, e_d, -e_1, -e_2, \ldots, -e_d)$.

We now let $n \geq 1$. We now construct a coupling

$$(\sigma_1, \sigma_2, \{\tilde{C}_m^1\}_{m \in \mathbf{Z}^{d-1}}, \ldots, \{\tilde{C}_m^{2d}\}_{m \in \mathbf{Z}^{d-1}})$$

of 2 copies of the distribution of μ and the processes $C^{H_1}, \ldots, C^{H_{2d}}$ such that (1) and (2) in the definition of WB hold. At the end, we mention the small modification needed due to the fact that the processes corresponding to different C^H's might be different processes.

To construct the coupling, let $\{X_i'\}_{i \in \mathbf{Z}^d}$ and $\{X_i''\}_{i \in \mathbf{Z}^d}$ be two independent copies of the background i.i.d. process $\{X_i\}_{i \in \mathbf{Z}^d}$ of which our process $\{Y_i\}_{i \in \mathbf{Z}^d}$ is a finitary factor, the factor map being denoted by f. The final coupling will be defined in terms of only these processes and is as follows. For all $i \in \mathbf{Z}^d$, let $Y_i' = \Big(f(\{X_j'\}_{j \in \mathbf{Z}^d})\Big)_i$ and for all $i \notin B(n)$, let $Y_i'' = \Big(f(\{X_j''\}_{j \in \mathbf{Z}^d})\Big)_i$. Condition (1) is now satisfied. Consider the random set $G := \mathbf{Z}^d \setminus \cup_{x \in B(n)^c} B(x, T_x'')$. Clearly, given $\{T_i''\}_{i \in B(n)^c}$, (and hence the set G) and $\{X_i''\}_{i \in G^c}$ (and hence $\{Y_i''\}_{i \in B(n)^c}$), the random variables $\{X_i''\}_{i \in G}$ are i.i.d. with the original distribution. We now let $\{X_i'''\}_{i \in \mathbf{Z}^d}$ be defined by $X_i''' = X_i''$ for $i \notin G$ and $X_i''' = X_i'$ for $i \in G$. It follows from the above that the process $\{X_i'''\}_{i \in \mathbf{Z}^d}$ has the same distribution as $\{X_i\}_{i \in \mathbf{Z}^d}$.

We next let, for all $i \in \mathbf{Z}^d$, $Y_i'' = \Big(f(\{X_j'''\}_{j \in \mathbf{Z}^d})\Big)_i$. (One observes that this agrees with the definition for Y_i'' given earlier for $i \notin B(n)$.) It is clear that both the processes $\{Y_i'\}_{i \in \mathbf{Z}^d}$ and $\{Y_i''\}_{i \in \mathbf{Z}^d}$ are equal in distribution to the process $\{Y_i\}_{i \in \mathbf{Z}^d}$.

For property (2) of the coupling, it is clear that if $x \in G$ is such that $B(x, T_x''') \subseteq G$, then $Y_x' = Y_x''$. Hence if $Y_x' \neq Y_x''$, then there exists $y \in B(n)^c$ such that $B(y, T_y'') \cap B(x, T_x''') \neq \emptyset$ and hence such that $B(y, T_y''') \cap B(x, T_x''') \neq \emptyset$. Then there exists a halfspace H_i whose boundary contains one of the $2d$ faces of $B(n)$ such that $y \in H_i$. Letting $(\{\tilde{C}_m^1\}_{m \in \mathbf{Z}^{d-1}}, \ldots, \{\tilde{C}_m^{2d}\}_{m \in \mathbf{Z}^{d-1}})$ be the $2d$ different processes defined earlier but with respect to the $\{X_i'''\}_{i \in \mathbf{Z}^d}$ random variables and translated to the proper face of $B(n)$, it follows from the above that (2) also holds.

Finally, we take care of the fact that the processes $C^{H_1},\ldots,C^{H_{2d}}$ might be different. Consider the stationary process $\{C_m\}_{m\in\mathbf{Z}^{d-1}}$ given by $C_m=\sum_{i=1}^{2d}C_m^{H_i}$ where the $2d$ processes $C^{H_1},\ldots,C^{H_{2d}}$ are taken to be independent. It is easy to modify the above to show that $\{Y_i\}_{i\in\mathbf{Z}^d}$ is WB with respect to the process $\{C_m\}_{m\in\mathbf{Z}^{d-1}}$. $\square$

We end this section with an application to the Ising Model. To save space, we refer the reader to [2] for all concepts involving the Ising model.

Corollary 4.4 *The unique Gibbs state for the Ising model below the critical parameter in any dimension is WB.*

Proof: It follows from the proof of the main result in [2] that this unique Gibbs state is a finitary factor of an i.i.d. process with the factor map having finite expected coding volume. Hence Theorem 1.6 completes the proof. $\square$

Remark: It was shown in [4] that the plus state for the Ising model is WB in any dimension when the parameter is sufficiently large.

5 Further questions

In this final section, we list some questions.

Question 1: Is the T,T^{-1}–process associated to simple random walk with positive drift in 1 dimension a finitary factor of an i.i.d. process?

Remark: The proofs of the results in this paper do not allow us to conclude that this is not the case since it is easy to show that $\{R_n\}$ in this case satisfies standard large deviations behavior. Note, in addition, that the explicit factor map given in the proof of Theorem 1.3 is clearly not finitary.

Question 2: Is it the case that the T,T^{-1}–process associated to any mean 0 stationary random walk is not a finitary factor of an i.i.d. process?

Remark: A natural approach here would be to show that the nonstandard large deviation behavior for the range process in the i.i.d. case also holds in the general stationary case.

Question 3: Is the plus state for the Ising model WB for all parameter values?

Remark: The results in [2] imply that this plus state is not a finitary factor of an i.i.d. process with finite expected coding volume at or above the critical parameter and hence one cannot use Theorem 1.6 to obtain a positive answer to this question.

Acknowledgments. We are grateful to Harry Kesten for help concerning Lemmas 2.2 and 2.3 as well as for providing some references.

References

[1] Berbee, H. C. P., *Random Walks with Stationary Increments and Renewal Theory*, Ph.D. thesis, Amsterdam, 1979.

[2] Berg, J. van den and Steif, J. E., On the existence and non-existence of finitary codings for a class of random fields, *Ann. Probab.*, **27**, (1999), 1501-1522.

[3] Burton, R. M. and Steif, J. E., Quite weak Bernoulli with exponential rate and percolation for random fields, *Stochastic Process. Appl.*, **58**, (1995), 35–55.

[4] Burton, R. M. and Steif, J. E., Coupling surfaces and weak Bernoulli in one and higher dimensions, *Adv. Math.*, **132**, (1997), 1–23.

[5] Dekking, F.M., On transience and recurrence of generalized random walks, *Z. Wahrsch. Verw. Gebiete*, **61**, (1982), 459–465.

[6] Donsker, M. D. and Varadhan, S. R. S., On the number of distinct sites visited by a random walk, *Comm. Pure Appl. Math.* **32**, (1979), 721–747.

[7] Durrett, R., *Probability: Theory and Examples* (2nd edition), Wadsworth Publ. Co., Duxbury Press, Belmont (CA), 1996.

[8] Häggström, O., Coloring percolation clusters at random, in preparation.

[9] Hamana Y. and Kesten, H., A large deviation result for the range of random walk and for the Wiener sausage, preprint.

[10] Hoffman, C., The scenery of the T, T^{-1} is not Bernoulli, in preparation.

[11] den Hollander, F. and Steif, J. E., Mixing properties of the generalized T, T^{-1}–process, *Journal d'Analyse Mathématique*, **72**, (1997), 165–202.

[12] del Junco, A. and Rahe, M., Finitary codings and weak Bernoulli partitions, *Proc. Amer. Math. Soc.*, **75**, (1979), 259–264.

[13] Kalikow, S., T, T^{-1} transformation is not loosely Bernoulli, *Ann. Math.*, **115** (1982) 393–409.

[14] Marton, K. and Shields, P., The positive-divergence and blowing–up properties, *Isr. J. Math.*, **86**, (1994), 331–348.

[15] Ornstein, D.S., *Ergodic Theory, Randomness and Dynamical Systems*, Yale University Press New Haven, 1974.

[16] Smorodinsky, M. and Thouvenot, J-P., Bernoulli factors that span a transformation, *Isr. J. Math.*, **32**, (1979), 39–43.

Jeffrey E. Steif
School of Mathematics,
Georgia Institute of Technology, Atlanta.
`steif@math.gatech.edu`