ADVANCES IN MATHEMATICS **132**, 1–23 (1997)
ARTICLE NO. AI971676

Coupling Surfaces and Weak Bernoulli in One and Higher Dimensions

Robert M. Burton*

Department of Mathematics, Oregon State University, Corvallis, Oregon 97331

and

Jeffrey E. Steif†

Department of Mathematics, Chalmers University of Technology,

We propose a notion of weak Bernoulli in all dimensions which generalizes the usual definition in dimension 1. The key idea is the concept of a *coupling surface*. We relate this notion to previously studied properties and discuss a number of possible variants in dimension 1. We also show that the Ising model, at low temperature, is weak Bernoulli with an explicit description of the coupling surface.

1. INTRODUCTION

After the isomorphism problem for Bernoulli shifts in ergodic theory was solved by Ornstein [16], conditions were sought which would guarantee that a stationary process is isomorphic to an i.i.d. process. A stationary process is isomorphic to another if there is a shift-invariant measure-preserving transformation between the corresponding measures on sequence space. A process which is isomorphic to an i.i.d. process is called a Bernoulli shift. (These definitions extend immediately to stationary random fields). In [11], the notion of weak Bernoulli (which is a certain mixing condition) was introduced and used to show that an irreducible aperiodic Markov chain on a finite state space is a Bernoulli shift.

The point of this paper is to make a closer examination of the definition of weak Bernoulli in dimension 1 and to propose a natural definition in

* Research supported by AFOSR Grant 91-0215 CT and NSF Grant DMS-9103738.
† Research supported by a grant from the Swedish National Science Foundation.

0001-8708/97 $25.00

dimensions larger than 1 which agrees with the usual definition in dimension 1. The results in [5] tell us that the natural naive extension of the definition of weak Bernoulli to higher dimensions is uninteresting (this point is described in detail later) and that something more subtle is required. The key idea in this extension is the notion of a *coupling surface*.

All definitions of mixing are based on some type of approximate independence. The notion of independence, which will play a central role here, is given in the next lemma where three equivalent forms are given.

LEMMA 1.1. *Let $\{X_i\}_{i \in A \cup B}$ be random variables defined on the same probability space taking values in a finite set F, where A and B are disjoint and countable (possibly finite). For $U \subseteq A \cup B$, let P_U denote the measure on F^U induced by the random variables $\{X_i\}_{i \in U}$. For a configuration $\eta \in F^A$, let P_B^η denote the measure on F^B which is the conditional distribution of the variables $\{X_i\}_{i \in B}$ given that the variables $\{X_i\}_{i \in A}$ equal η. Finally let $\| \ \|$ denote the total variation norm of a finite signed measure. Then for all $\varepsilon > 0$, there exists $\delta > 0$ so that*

(1) *$\|P_{A \cup B} - P_A \times P_B\| \leqslant \delta$ implies that for ε most of the configurations η on A (with respect to P_A), $\|P_B - P_B^\eta\| \leqslant \varepsilon$.*

(2) *If for δ most η on A (with respect to P_A), we have that $\|P_B - P_B^\eta\| \leqslant \delta$, then there exists a coupling $\{(X'_i, X''_i)\}_{i \in A \cup B}$ of $\{X_i\}_{i \in A \cup B}$ with itself such that $\{X'_i\}_{i \in A \cup B}$ and $\{X''_i\}_{i \in A}$ are independent and such that*

$$P(X'_b = X''_b \text{ for all } b \in B) \geqslant 1 - \varepsilon.$$

(3) *If there exists a coupling $\{(X'_i, X''_i)\}_{i \in A \cup B}$ of $\{X_i\}_{i \in A \cup B}$ with itself such that $\{X'_i\}_{i \in A \cup B}$ and $\{X''_i\}_{i \in A}$ are independent and such that*

$$P(X'_b = X''_b \text{ for all } b \in B) \geqslant 1 - \delta$$

then $\|P_{A \cup B} - P_A \times P_B\| \leqslant \varepsilon$.

Finally, δ can be taken to be $\varepsilon^2/100$ in (1), (2), and (3) above.

We leave the proof of the above lemma to the reader as a fairly simple but instructive exercise. The last part of the lemma (which says that δ can be taken to be $\varepsilon^2/100$) will only be used in the proof of Theorem 1.14.

We begin with three definitions in dimension 1 which are equivalent. It is the first of these definitions which is usually taken as the definition of weak Bernoulli. The third one, which looks quite complicated, will be the one which generalizes to higher dimensions. All processes in this paper will be assumed to take on only a finite number of values.

DEFINITION 1.2. A stationary process $\{X_n\}_{n\in\mathbf{Z}}$ is called *one-sided weak Bernoulli* if

$$\lim_{n\to\infty} \|P_{(-\infty,0]\cup[n,\infty)} - P_{(-\infty,0]} \times P_{[n,\infty)}\| = 0.$$

We mention that this condition was introduced by Kolmogorov (see [20]) under the name absolutely regular and was discovered independently in [11].

DEFINITION 1.3. A stationary process $\{X_n\}_{n\in\mathbf{Z}}$ is called *two-sided weak Bernoulli* if

$$\lim_{n\to\infty} \sup_{k\geqslant 0} \|P_{(-\infty,-n]\cup[0,k]\cup[n+k,\infty)} - P_{(-\infty,-n]\cup[n+k,\infty)} \times P_{[0,k]}\| = 0.$$

DEFINITION 1.4. A stationary process $\{X_n\}_{n\in\mathbf{Z}}$ is called *weak Bernoulli* (WB) if there exists a nonnegative random variable C so that for all n, there exists a coupling $(\sigma_1, \sigma_2, \tilde{C}^1, \tilde{C}^2)$ of two copies of the distribution of $\{X_n\}_{n\in\mathbf{Z}}$ and two copies of the distribution of C (where we suppress the dependence on n in the notation) so that

(1) σ_1 and $\sigma_2|_{[-n,n]^c}$ are independent (where $\sigma_2|_U$ means the restriction of σ to U) and

(2) $A_1 \cap A_2 \subseteq \{x: \sigma_1(x) = \sigma_2(x)\}$ where

$$A_1 = \{x \in [-n, n]: \tilde{C}^1 \leqslant x + n\}$$

and

$$A_2 = \{x \in [-n, n]: \tilde{C}^2 \leqslant n - x\}.$$

For each n, we call the above coupling which depends on n the nth *coupling*.

Motivation for this definition will come after Definition 1.6 but we mention that there is some analogy between this definition and the one that arises in Theorem 4.4.7 in [1].

THEOREM 1.5. *Let $\{X_n\}_{n\in\mathbf{Z}}$ be a stationary process. Then the following are equivalent*:

(i) $\{X_n\}_{n\in\mathbf{Z}}$ *is one-sided weak Bernoulli.*

(ii) $\{X_n\}_{n\in\mathbf{Z}}$ *is two-sided weak Bernoulli.*

(iii) $\{X_n\}_{n\in\mathbf{Z}}$ *is weak Bernoulli.*

The notion of weak Bernoulli has been studied quite extensively and, since it is natural to also study random fields (and their mixing properties), it is natural as well to try to extend the notion of WB to higher dimensions. The definition we give is the following.

DEFINITION 1.6. A translation invariant measure μ on $A^{\mathbf{Z}^d}$ 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 $(\sigma_1, \sigma_2, \{\tilde{C}^1_m\}_{m\in\mathbf{Z}^{d-1}}, ..., \{\tilde{C}^{2d}_m\}_{m\in\mathbf{Z}^{d-1}})$ of two 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 (where $B_n=[-n, n]^d$ and $\sigma_2|_U$ means the restriction of σ to U) and

(2) $\bigcap_{i=1}^{2d} A_i \subseteq \{x: \sigma_1(x)=\sigma_2(x)\}$, where for $i=1, 2, ..., d$

$$A_i=\{x\in B_n: \tilde{C}^i_{\hat{x}_i}\leqslant x_i+n\}$$

and for $i=d+1, d+2, ..., 2d$

$$A_i=\{x\in B_n: \tilde{C}^i_{\hat{x}_{i-d}}\leqslant n-x_{i-d}\}$$

where $\hat{x}_j=\{x_1, ..., x_{j-1}, x_{j+1}, ..., x_d\}$.

For each n, we call the above coupling which depends on n the *n th coupling*.

The reader may naturally wonder why our definition of WB is based upon a generalization of Definition 1.4 rather than generalizations (which the reader can easily come up with herself) of the much simpler Definitions 1.2 or 1.3. The reason is that the methods in [5] show that any process which satisfies these natural generalizations are in fact finitely dependent, also called m-dependent. This means that there is some fixed number k so that the random variables associated to two index sets which are separated by hyperplanes more than k apart are completely independent. (This phenomenon is discussed further in [8].) This condition is then obviously too strong. This is one of the motivations behind our definition of weak Bernoulli in higher dimensions.

The idea behind the definition of WB is the following. We have two copies of our process which are independent outside B_n and we want to couple them inside the box. As we come inwards from the boundary of the box, we have $2d$ coupling surfaces corresponding to the $2d$ sides of B_n. The two processes are then coupled perfectly further inside B_n than all of the coupling surfaces. The lower $(d-1)$-dimensional process corresponds to how far these coupling surfaces are from their respective sides of B_n.

We mention that it is quite straightforward to show that a random field, which is WB according to the above definition, is a Bernoulli shift. The way to do this is to prove that it is very weak Bernoulli and to use the equivalence of this with being a Bernoulli shift (see [18]). Coupling techniques have become an essential tool in recent years in studying phase transitions in random fields (see [2–4]), as well as in studying mixing properties (see [8]). We feel that the definition of WB given above ties together the essence of these coupling techniques while extending the usual definition in dimension 1 which has been studied extensively [5].

We now proceed to give two other variants of the above definitions in dimension 1. We will only use the second definition when we generalize these to higher dimensions. We feel it worthwhile to give the first and to show the difference between these definitions because it illustrates the fact that "past–future" definitions, such as Definition 1.2, do not always agree with the corresponding "inside–outside" definition which is Definition 1.3 in this case.

DEFINITION 1.7. A stationary process $\{X_n\}_{n \in \mathbf{Z}}$ is called *one-sided quite weak Bernoulli* if for all $\varepsilon > 0$,

$$\lim_{n \to \infty} \|P_{(-\infty, 0] \cup [\varepsilon n+1, \varepsilon n+n]} - P_{(-\infty, 0]} \times P_{[\varepsilon n+1, \varepsilon n+n]}\| = 0.$$

DEFINITION 1.8. A stationary process $\{X_n\}_{n \in \mathbf{Z}}$ is called *two-sided quite weak Bernoulli* if for all $\varepsilon > 0$,

$$\lim_{n \to \infty} \|P_{(-\infty, -\varepsilon n] \cup [0, n] \cup [n+\varepsilon n, \infty)} - P_{(-\infty, -\varepsilon n] \cup [n+\varepsilon n, \infty)} \times P_{[0, n]}\| = 0.$$

Remark. We leave it to the reader to check that two-sided weak Bernoulli implies two-sided quite weak Bernoulli and that two-sided quite weak Bernoulli implies one-sided quite weak Bernoulli. Theorems 1.9 and 1.10 below tell us that the reverse implications fail.

THEOREM 1.9 *There exists a finite state stationary process $\{X_n\}_{n \in \mathbf{Z}}$ which is one-sided quite weak Bernoulli but not two-sided quite weak Bernoulli.*

THEOREM 1.10. *There exists a finite state stationary process $\{X_n\}_{n \in \mathbf{Z}}$ which is two-sided quite weak Bernoulli but not two-sided weak Bernoulli.*

We now relate the definition of weak Bernoulli to other related conditions. The notion of two-sided quite weak Bernoulli generalizes easily to higher dimensions and is studied in [8]. The notion of quite weak Bernoulli with exponential rate, defined later, also comes from [8].

Definition 1.11. A translation invariant measure μ on $F^{\mathbf{Z}^d}$ is called *quite weak Bernoulli* (*QWB*) if for all $\varepsilon > 0$,

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

where Λ_n denotes $[-n, n]^d$.

The next theorem tells us that (under certain moment conditions on the lower dimensional process) weak Bernoulli implies QWB.

Theorem 1.12. *Let μ be a d-dimensional weak Bernoulli process with associated $(d-1)$-dimensional process $\{C_m\}_{m \in \mathbf{Z}^{d-1}}$. If $E[|C_0|^{d-1}] < \infty$, then μ is QWB.*

Remark. We remarked earlier that in dimension 1, WB implies two-sided quite weak Bernoulli. This also follows from this last theorem since the moment condition becomes vacuous. We would believe that in $d \geqslant 2$ WB does not, in general, imply QWB and that examples of the type given in Theorem 1.10 could be constructed which would show that QWB does not imply WB. In addition, similar to the WB property, it can easily be shown that the QWB property also implies that the field is a Bernoulli shift.

A somewhat stronger condition than QWB is the following.

Definition 1.13. A translation invariant 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

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

for all n.

Various random fields satisfy the QWBE property. One of these is the plus state for the Ising model at a variety of parameter values. Some of the parameter values for which this property has been proved include $d > 2$, zero external field, and sufficiently high or sufficiently low temperature (see [13] or combine the methods in [3] and [8]). See [15] for other parameter values. Standard methods, together with a percolation result in [9] show that for $d = 2$ and zero external field, temperatures less than the critical temperature belong to this class. Finally, it is proved in [17] that for $d = 2$, we are in this class if there is a zero external field and the temperature is larger than the critical temperature or if there is a nonzero external field and the temperature is arbitrary. The last two results show that in two dimensions we always have the QWBE property, except if we are at the critical point where the external field is zero and the temperature is the critical temperature.

Another collection of measures for which the QWBE property has been established are certain measures of maximal entropy for subshifts of finite type (see [8]). Subshifts of finite type (see [6] for a discussion) are certain dynamical systems which arise in ergodic theory and which have connections to statistical mechanics. Such objects have natural measures associated to them, the so-called measures of maximal entropy. In [8], it is proved that for a certain family of subshifts of finite type, all the ergodic measures of maximal entropy satisfy the QWBE property. Within this class one can find, for any d and k, a subshift of finite type in d dimensions with exactly k ergodic (and, hence, QWBE by the above) measures of maximal entropy (see [7])

The QWBE property also implies a central limit theorem (under a nonsingularity assumption) (see [8]). In [19] it is shown that, under the QWBE property, one can use empirical distributions to consistently estimate the joint distribution of the random field assuming a necessary entropy constraint.

Our next theorem relates WB to QWBE under an exponential moment assumption.

THEOREM 1.14. *Let μ be a d-dimensional weak Bernoulli process with associated $(d-1)$-dimensional process $\{C_m\}_{m \in \mathbf{Z}^{d-1}}$. If there exists $\delta > 0$ such that $E[e^{\delta C_0}] < \infty$, then μ is QWBE.*

Remark. We have not discussed the notion of QWBE in dimension 1. We believe that there is no implication between WB and QWBE in dimension 1, although we have not pursued this relationship. Note that, while the moment condition in Theorem 1.12 is vacuous in dimension 1, this is not the case in Theorem 1.14.

Our next theorem gives a concrete nontrivial example (the Ising model with large interaction J and 0 external field) of a process which is weak Bernoulli with an explicit construction of the lower dimensional process which has a finite exponential moment. By the previous result this implies that the process is QWBE. The fact that this process is QWBE was proved in [8] and the methods used there will allow us to easily show that this process is weak Bernoulli with the explicit construction of the lower dimensional process. This will be the only part of the paper which is not self-contained. We will not describe the Ising model here, but refer the reader to [8] for a complete discussion. In any case, the methods and results in [8] will be needed for the proof of Theorem 1.15.

THEOREM 1.15. *Let* $\{Y_n\}_{n\in\mathbf{Z}^d}$ *be the plus state for the ferromagnetic Ising model with large interaction* J *and* 0 *external field. Then* $\{Y_n\}$ *is weak Bernoulli with a lower dimensional process whose one-dimensional marginal has a finite exponential moment.*

Remark. Note that Theorem 1.15, together with Theorem 1.14, implies that the plus state for the ferromagnetic Ising model with large interaction J and 0 external field is QWBE.

Before concluding this introduction, we return to our earlier discussion concerning how the natural extension of WB implies that the process is finitely dependent. For illustrative purposes we will only do this in the case where we are considering the "mixing" in the horizontal direction in two dimensions. We now give a quick proof of one of the results in [5]; one can certainly obtain stronger results. Since [5] assumes mixing, our result is actually stronger. Note that $\|P_{((-\infty,0]\cup[n,\infty))\times\mathbf{Z}}-P_{(-\infty,0]\times\mathbf{Z}}\times P_{[n,\infty)\times\mathbf{Z}}\|$ should be small for large n under a natural extension of the WB definition, and that this quantity being zero means that events whose index sets are more than n apart (horizontally) are independent. Therefore this last result says exactly that (this natural definition of) WB implies that the process is finitely dependent.

THEOREM 1.16. *Let* $\{X_n\}_{n\in\mathbf{Z}^2}$ *be a stationary process such that two independent copies of it are ergodic under vertical translation* (*so-called weak mixing*). *Then*

$$\|P_{((-\infty,0]\cup[n,\infty))\times\mathbf{Z}}-P_{(-\infty,0]\times\mathbf{Z}}\times P_{[n,\infty)\times\mathbf{Z}}\|$$

is either 0 *or* 2.

Proof. $P_{((-\infty,0]\cup[n,\infty))\times\mathbf{Z}}$ and $P_{(-\infty,0]\times\mathbf{Z}}\times P_{[n,\infty)\times\mathbf{Z}}$ are both measures on $F^{((-\infty,0]\cup[n,\infty))\times\mathbf{Z}}$ (where F is the state space of the process) which are ergodic under vertical translation. However, two ergodic measures are either the same or mutually singular. ∎

The rest of the paper is devoted to proofs. In Section 2 we will prove Theorem 1.5. In Section 3 we will prove Theorems 1.9 and 1.10. Finally in Section 4 we prove Theorems 1.12, 1.14, and 1.15.

2. EQUIVALENCE OF DEFINITIONS

In this section we prove Theorem 1.5.

Proof of Theorem 1.5. The fact that two-sided weak Bernoulli implies one-sided weak Bernoulli is easy and left to the reader. We first prove that one-sided weak Bernoulli implies two-sided weak Bernoulli. Let $\varepsilon > 0$. Choose N such that, for all $n \geqslant N$,

$$\|P_{(-\infty, 0] \cup [n, \infty)} - P_{(-\infty, 0]} \times P_{[n, \infty)}\| \leqslant \frac{\varepsilon}{3}.$$

Now let $k \geqslant 1$ be arbitrary, and let $A = (-\infty, -n]$, $B = [0, k]$, and $C = [n+k, \infty)$. We then have that for all $n \geqslant N$,

$$\begin{aligned}\|P_{A \cup B \cup C} - P_{A \cup C} \times P_B\| \leqslant & \|P_{A \cup B \cup C} - P_{A \cup B} \times P_C\| \\ & + \|P_{A \cup B} \times P_C - P_A \times P_B \times P_C\| \\ & + \|P_A \times P_B \times P_C - P_{A \cup C} \times P_B\|.\end{aligned}$$

However, each of these three terms is at most $\varepsilon/3$ and, as k is arbitrary, this proves that the process is two-sided weak Bernoulli.

The fact that Definition 1.6 implies Definition 1.2 is trivial and left to the reader.

For the more difficult direction we proceed as follows. We mention that Theorem 4.4.7 in [1] is exactly a one-sided version of our Theorem 1.5 and we refer the reader to this result. We will adapt the proof of this result to our setting. The main idea will be to obtain a version of so-called maximal couplings for a finite setting which is appropriate for our situation. See [14] for an excellent treatment of the coupling method and where a main result on maximal couplings due to Goldstein ([12]) is discussed. The subprobability measure μ is said to dominate the subprobability measure v if $v(B) \leqslant \mu(B)$ for all measurable sets B. If ρ and ρ' are subprobability measures, we let $\rho \wedge \rho'$ denote the maximum subprobability measure which is dominated by both ρ and ρ' (see [14]).

We start off by showing how to couple two measures μ_1 and μ_2 on $X = A^{\{-n, ..., n\}}$ (A is a finite set) such that they are "maximally coupled from the inside outwards."

Denote elements of X by $\{\eta(x), |x| \leqslant n\}$ and for each $i = 0, ..., n$, let $\mathscr{F}_i$ be the σ-algebra generated by the coordinates $-i, ..., i$ and let $\mathscr{F}_{-1} = \{X, \varnothing\}$. We will now construct a coupling $\mu_1 \otimes \mu_2$ such that, for all $k = 0, 1, ..., n$,

$$1 - \mu_1 \otimes \mu_2((\eta_1, \eta_2) \colon \eta_1(x) = \eta_2(x)\ \forall |x| \leqslant k) = \frac{\|(\mu_1)|_{\mathscr{F}_k} - (\mu_2)|_{\mathscr{F}_k}\|}{2}.$$

It is easy to see [14] that for any coupling of μ_1 and μ_2, the last above $=$ must be a $\geqslant$. Moreover, it is easy to obtain a coupling where equality holds for a fixed k (see [14]), but the point is to construct a coupling, such that equality holds for all k. The idea is to modify the method in [12]. To do this, we first consider $n+1$ subprobability measures,

$$(\mu_1, \mu_2)(-1), (\mu_1, \mu_2)(0), ..., (\mu_1, \mu_2)(n)$$

(where $(\mu_1, \mu_2)(i)$ is defined on the measurable space $(X, \mathscr{F}_i)$), which are defined as follows.

Let $(\mu_1, \mu_2)(n) = \mu_1 \wedge \mu_2$ and for $i = -1, 0, ..., n-1$, let

$$(\mu_1, \mu_2)(i) = (\mu_1)|_{\mathscr{F}_i} \wedge (\mu_2)|_{\mathscr{F}_i} - ((\mu_1)|_{\mathscr{F}_{i+1}} \wedge (\mu_2)|_{\mathscr{F}_{i+1}})_{\mathscr{F}_i}.$$

We leave to the reader to check that this yields a subprobability measure, in particular, that it does not give any set negative measure. Note that $\sum_{i=-1}^{n} \|(\mu_1, \mu_2)(i)\| = 1$. For $i = 0, ..., n$, let v_i be the probability measure on $(X, \mathscr{F}_i)$ obtained by normalizing $(\mu_1, \mu_2)(i)$ to be a probability measure. As in the proof of the maximal coupling result (see [14]), one can find probability measures $\{v_i'\}_{i=-1}^{n}$ on $X \times X$ such that

(1) For $i = 0, 1, ..., n$, v_i' restricted to $\mathscr{F}_i$, has v_i as its two marginals and couples them perfectly (i.e., $v_i'((\eta_1, \eta_2)\colon \eta_1(x) = \eta_2(x)\ \forall |x| \leqslant i) = 1$), and

(2) $\sum_{i=-1}^{n} \|(\mu_1, \mu_2)(i)\|\ v_i'$ is a coupling of μ_1 and μ_2.

Once we have this, it follows easily that

$$1 - \mu_1 \otimes \mu_2((\eta_1, \eta_2)\colon \eta_1(x) = \eta_2(x)\ \forall |x| \leqslant k) = \frac{\|(\mu_1)|_{\mathscr{F}_k} - (\mu_2)|_{\mathscr{F}_k}\|}{2}$$

for all $k = 0, 1, ..., n$ where

$$\mu_1 \otimes \mu_2 = \sum_{i=-1}^{n} \|(\mu_1, \mu_2)(i)\|\ v_i'.$$

Returning to our stationary process, which we view as a (shift invariant) measure μ on $X = A^{\mathbf{Z}}$, we let μ^n be the restriction of μ to the coordinates between $-n$ and n and, if $\eta \in X$, we let μ_η^n be the conditional distribution of μ on the coordinates between $-n$ and n, given that η is the configuration at the other coordinates.

We next let m_n be the probability measure on the nonnegative integers $\mathbf{N}$ given by

$$m_n(i) = \int_X \|(\mu^n, \mu_\eta^n)(n-i)\|\ d\mu(\eta).$$

Note that m_n is concentrated on $\{0, ..., n+1\}$. Heuristically, $m_n(i)$ is the average probability that we need to go i steps toward the origin before coupling μ^n and μ^n_η perfectly. We leave to the reader to check that the two-sided weak Bernoulli assumption implies that the family of probability measures $\{m_n\}_{n \geqslant 1}$ is tight. It is easily seen that tightness implies in turn the existence of a probability measure m on $\mathbf{N}$ such that for all n and for all i, $m_n([i, \infty)) \leqslant m([i, \infty))$. We claim that m is the zero-dimensional process (i.e., a probability measure on $\mathbf{N}$) required by Definition 1.6 to show that our process is weak Bernoulli. To do this, one uses the measures $\mu^n \otimes \mu^n_\eta$, together with the method of proof of Theorem 4.4.7 in [1], to complete the proof. ∎

3. 1*D* COUNTEREXAMPLES

In this section we prove Theorems 1.9 and 1.10.

Proof of Theorem 1.9. Let $\{Y_n\}_{n \in \mathbf{Z}}$ be an i.i.d. process with $P(Y_n = 1) = \frac{1}{2} = P(Y_n = 0)$.

For $M \geqslant 1$, we define an *M-block* to be a block

$$0111 \cdots 1110$$

with 2^M ones between 2 zeros (which has length $2^M + 2$). We say that an M-block is *good* if there are no $(M+I)$-blocks, $I \geqslant 1$, between this M-block and the preceding M-block (if one exists).

We next let $\{V_n\}_{n \in \mathbf{Z}}$ be a factor of (a function of) the process $\{Y_n\}$, taking values $\{0, 1, \varnothing\}$, which is defined as follows. If $n \in \mathbf{Z}$ is one of the first M ones in a good M-block for some $M \geqslant 1$, we then let

$$V_n = \sum_{i=1}^{i_0} Y_{n - iM} \pmod 2,$$

where $i_0 = \sup \{ i \mid Y_{n-jM}$ does not occur in an M-block for $j = 1, ..., i\}$. (Since n can belong to at most one M-block, V_n is well defined.) If $n \in \mathbf{Z}$ is not one of the first M ones in a good M-block for any $M \geqslant 1$, we let $V_n = \varnothing$. We finally consider the stationary process $\{X_n\}_{n \in \mathbf{Z}}$ defined by $\{(Y_n, V_n)\}_{n \in \mathbf{Z}}$ (which is a factor of $\{Y_n\}$).

We say that a stationary process $\{Z_n\}$ is *bilaterally deterministic* (BD) if $\bigcap_{n \geqslant 1} \sigma(Z_i, |i| \geqslant n) = \sigma(Z_i, i \in \mathbf{Z})$, where the above equality is meant modulo the probability measure. Recall that two σ-algebras are equal modulo a measure means that, for each set in the first σ-algebra, there is a set in the second σ-algebra whose symmetric complement with the first set has probability 0 and conversely.

The first step in this proof is to show that $\{X_n\}$ is BD since then it will easily follow that $\{X_n\}$ is not two-sided quite weak Bernoulli, as it is trivial to show that a two-sided quite weak Bernoulli process cannot be BD unless the entire process consists of one atom.

To do this, we first leave it to the reader to check that to show BD, it suffices to show that for all $\ell \geqslant 1$, $\sigma(X_i, |i| \leqslant \ell) \subseteq \sigma(X_i, |i| > \ell)$, where again the containment is meant modulo the probability measure. We now fix such an ℓ.

We will show that with probability 1 for all values of M except for finitely many, there exists a good M-block entirely to the right of ℓ with the next M-block to the left being entirely to the left of $-\ell$. Denote this event by U. Assuming U has probability 1, we proceed as follows.

We need to determine the values of the Y_n for $|n| \leqslant \ell$ as a function of $\{X_n\}_{|n| > \ell}$. (Once we do this, we can determine the values of the X_n for $|n| \leqslant \ell$, since the V_n's are simply functions of the Y_n's.) This is done as follows. Let s be the first 0 to the right of ℓ in the $\{Y_n\}$ process, and t be the first 0 to the left of $-\ell$. As $P(U)=1$, there is a.s. an M satisfying $2^M > s-t$ and $M \geqslant 2\ell+1$, such that there is a good M-block entirely to the right of ℓ whose preceding M-block is entirely to the left of $-\ell$. We consider the smallest such M and we call the corresponding good M-block B.

Now, let n be such that $|n| \leqslant \ell$. Let r be the unique point among the first M ones in B with the property that $r-n$ is a multiple of M. By construction,

$$V_r = \sum_{i=1}^{i_0} Y_{r-iM} \pmod 2,$$

where $i_0 = \sup\{i \mid Y_{r-jM}$ does not occur in an M-block for $j=1, ..., i\}$. Since $r-n$ is a multiple of M, Y_n is one of the terms in the above sum and, in addition, V_r, i_0, and all the terms in the above sum, except for Y_n, are in $\sigma(X_i, |i| > \ell)$, since $M \geqslant 2\ell+1$. Therefore, using the above sum, we see that Y_n is also in $\sigma(X_i, |i| > \ell)$. (We point out that the reason we required that $2^M > s-t$ is to ensure that $i_0 \in \sigma(X_i, |i| > \ell)$.)

We now show that U has probability 1. Let

$$\sigma_M = \inf\{i \geqslant \ell+1 : i \text{ is the left point of an } M\text{-block}\}$$

and

$$b_M = \sup\{i \leqslant -\ell-1 : i \text{ is the right point of an } M\text{-block}\}$$

which are well defined a.s. Let Ω be the event of probability 1, such that a_M and b_M exist and are finite for all M. Let E_M be the event that there is an M'-block between b_M and a_M for some $M' > M$. The first step is to show that

$$\sum_M P(E_M) < \infty, \tag{2.1}$$

a computation left to the reader. Since only finitely many M-blocks can intersect $\{-\ell, ..., \ell\}$, it is clear that $\Omega \cap \{E_M \text{ i.o.}\}^c = U$ and so U has probability 1 by the Borel–Cantelli lemma.

The next step of the proof is to show that $\{X_n\}$ is one-sided quite weak Bernoulli. To show this, it suffices by Lemma 1.1 to show that, $\forall \varepsilon > 0$, there exists N such that $\forall n \geqslant N$, there exists a coupling $\{(\tilde{X}_k, \tilde{\tilde{X}}_k)\}_{k \in \mathbf{Z}}$ of $\{X_k\}_{k \in \mathbf{Z}}$ with itself, such that

$$\{\tilde{X}_k\}_{k \in \mathbf{Z}} \text{ and } \{\tilde{\tilde{X}}_k\}_{k=-\infty}^{0} \text{ are independent} \tag{$*$}$$

and

$$P(\tilde{X}_k = \tilde{\tilde{X}}_k \ \forall k \in \{\varepsilon n + 1, ..., \varepsilon n + n\}) \geqslant 1 - \varepsilon. \tag{$**$}$$

We denote by $a(n)$ the nearest integer to $\log_2 \log_2(n)$.

Let $\varepsilon > 0$. Choose N so that for all $n \geqslant N$

(1) If W_i are i.i.d. with $P(W_i = 1) = 1/4 = 1 - P(W_i = 0)$, then $P(\min\{i \colon W_i = 1\} \leqslant \varepsilon n/4) \geqslant 1 - \varepsilon/4$

(2) $(2^{2^{a(n)-1}})^{4/3} (2^{a(n)-1} + 2) \leqslant \varepsilon n/2$

(3) If we have $\lfloor (2^{2^{a(n)-1}})^{4/3} \rfloor$ independent events, each with probability at least $\frac{1}{2}^{(2^{a(n)-1}+2)}$ of occurring, then at least one of them will occur with probability $\geqslant 1 - \varepsilon/4$,

(4) $\sum_{j=1}^{\infty} \frac{1}{2}^{(2^{a(n)+j}+2)} (2^{a(n)+j} + 1 + n) < \varepsilon/4$ and

(5) The probability that the first $a(n)$-block after 0 is good is $\geqslant 1 - \varepsilon/8$.

(A straightforward computation left to the reader is required to see that such an N exists.) We now let $n \geqslant N$.

We now construct the desired coupling satisfying $(*)$ and $(**)$. We first let $M_n = a(n)$. Let $\{X'_k\}_{k \in \mathbf{Z}}$ and $\{X''_k\}_{k \in \mathbf{Z}}$ be independent, each having the same distribution as $\{X_k\}_{k \in \mathbf{Z}}$. $\{\tilde{X}_k\}_{k \in \mathbf{Z}}$ and $\{\tilde{\tilde{X}}_k\}_{k \in \mathbf{Z}}$ will now be defined as functions of $\{X'_k\}_{k \in \mathbf{Z}}$ and $\{X''_k\}_{k \in \mathbf{Z}}$. We let $\tilde{X}_k = X'_k$ for all k and $\tilde{\tilde{X}}_k = X''_k$ for all $k \leqslant 0$. This of course implies $(*)$ above immediately. If $n_0 = \min\{i \geqslant 1 \colon Y'_i = Y''_i = 0\} \leqslant \varepsilon n/4$, then let $\tilde{\tilde{Y}}_k = Y''_k$ for all $k \leqslant n_0$, let

$\tilde{\tilde{Y}}_k = \tilde{Y}_k$ for all $k > n_0 + M_n$ and, for $k \in \{n_0+1, ..., n_0+M_n\}$, let $\tilde{\tilde{Y}}_k$ be such that

$$\sum_{j=0}^{j_0} \tilde{Y}_{k-jM_n} = \sum_{j=0}^{j'_0} \tilde{\tilde{Y}}_{k-jM_n} \tag{2.2}$$

where $j_0 = \sup\{j\colon \tilde{Y}_{k-\ell M_n}$ is not in an M_n-block for $\ell = 1, ..., j\}$ and $j'_0 = \sup\{j\colon \tilde{\tilde{Y}}_{k-\ell M_n}$ is not in an M_n-block for $\ell = 1, ..., j\}$. If $n_0 > \varepsilon n/4$, then let $\tilde{\tilde{Y}}_k = Y''_k$ for all k. The V_k's are then simply defined as they are forced to be. (The point of using n_0 is that, after time n_0, the Y_k's are independent of $\{X_k\}_{k \leqslant n_0}$, although the V_k's are not.)

To complete the proof, we need to show that

(a) $\{\tilde{\tilde{X}}_k\}_{k \in \mathbf{Z}}$ and $\{X_k\}_{k \in \mathbf{Z}}$ are equal in distribution and

(b) $P(\tilde{X}_k = \tilde{\tilde{X}}_k \ \forall k \in \{\varepsilon n+1, \varepsilon n+n\}) \geqslant 1-\varepsilon$.

(a) is straightforward and left to the reader.

For (b) one proceeds as follows. Let $A_1 = \{n_0 \leqslant \varepsilon n/4\}$, A_2 be the event that there is an (M_n-1)-block completely contained in $\{\varepsilon n/2+1, ..., \varepsilon n\}$ for $\{X'_k\}$, A_3 be the event that there is no (M_n+j)-block $(j \geqslant 1)$ intersecting $\{\varepsilon n+1, \varepsilon n+n\}$ for $\{X'_k\}$ and A_4 be the event that, for both processes $\{\tilde{Y}_k\}$ and $\{\tilde{\tilde{Y}}_k\}$, the first $a(n)$-block after $\varepsilon n/2$ is good. We now show that $P(A_1 \cap A_2 \cap A_3 \cap A_4) \geqslant 1-\varepsilon$ and $A_1 \cap A_2 \cap A_3 \cap A_4 \subseteq \{\tilde{X}_k = \tilde{\tilde{X}}_k \ \forall k \in \{\varepsilon n+1, \varepsilon n+n\}\}$.

The inequality is obtained by noting that (1) implies that $P(A_1) \geqslant 1-\varepsilon/4$, (2) and (3), together with an easy computation, imply that $P(A_2) \geqslant 1-\varepsilon/4$, (4), together with an easy computation, implies that $P(A_3) \geqslant 1-\varepsilon/4$, and (5) implies that $P(A_4) \geqslant 1-\varepsilon/4$.

Next, to show the desired containment, one first notes that if $n_0 \leqslant \varepsilon n/4$, then $n_0 + M_n$ is $\leqslant \varepsilon n/2$ (as (2) certainly implies that $a(n) \leqslant \varepsilon n/4$); thus $\tilde{Y}_k = \tilde{\tilde{Y}}_k$ for all $k \geqslant \varepsilon n/2$. We need to show that $\tilde{V}_k = \tilde{\tilde{V}}_k$ for all $k \in \{\varepsilon n+1, \varepsilon n+n\}$. The fact that $\tilde{Y}_k = \tilde{\tilde{Y}}_k$ for all $k \geqslant \varepsilon n/2$ and the event A_2 together imply that any M-block for $M \leqslant M_n - 1$, intersecting $\{\varepsilon n+1, \varepsilon n+n\}$, is good for $\tilde{Y}$, if and only if it is good for $\tilde{\tilde{Y}}$ and in this case, one does not need to look at the Y_j values for $j \leqslant \varepsilon n/2$ to determine the V_r values. So the V_r values in this case will be the same for $\tilde{X}$ and $\tilde{\tilde{X}}$. The event A_3 implies that there are no $M_n + j$-blocks intersecting $\{\varepsilon n+1, \varepsilon n+n\}$ for $j \geqslant 1$ and we need not worry about these. Finally, we need to consider M_n-blocks intersecting $\{\varepsilon n+1, \varepsilon n+n\}$. We claim that for any M_n-block for $\tilde{X}$ (or equivalently for $\tilde{\tilde{X}}$) which is entirely to the right of $\varepsilon n/2$, the corresponding V_r's in $\tilde{X}$ and $\tilde{\tilde{X}}$ will be the same. (Note that any M_n-block intersecting $\{\varepsilon n+1, \varepsilon n+n\}$ is necessarily entirely to the right of $\varepsilon n/2$, since $2^{a(n)+2} < \varepsilon n/2$). Letting B be the first (not necessarily good)

M_n-block entirely to the right of $\varepsilon n/2$, it is clear that this is the case for all such M_n-blocks with the possible exception of B. Finally, (2.2) together with (5), ensures that this holds for the block B as well. ▮

Proof of Theorem 1.10. Let $\{Y_n\}_{n\in\mathbf{Z}}$ be an i.i.d. process with $P(Y_n=1)=\frac{1}{2}=P(Y_n=0)$. For $M\geqslant 17$, we define (as in the proof of Theorem 1.9) an *M-block* to be a block

$$0111\cdots 1110$$

with 2^M ones between 2 zeros (and, hence, has length 2^M+2). We say that two successive M-blocks are *pairwise good* if there are no $(M+I)$-blocks, $I\geqslant 1$, between them.

Given a realization of the $\{Y_n\}$ process, we next let $\{V_n\}_{n\in\mathbf{Z}}$ be a process taking values $\{0, 1, \varnothing\}$, which is defined as follows. If $r, s\in\mathbf{Z}$ are such that there exist M-blocks B_1 and B_2 (in the $\{Y_n\}$ process) which are pairwise good (with B_1 being to the left of B_2) and with r being the last 1 in B_1 and s being the first 1 in B_2, we then let

$$(V_r, V_s)=\begin{cases}(1,1) & \text{with probability } \dfrac{1}{4}+\dfrac{1}{\sqrt{M}}\\[2ex] (0,0) & \text{with probability } \dfrac{1}{4}+\dfrac{1}{\sqrt{M}}\\ (1,0) & \text{with probability } \dfrac{1}{4}-\dfrac{1}{\sqrt{M}}\\[2ex] (0,1) & \text{with probability } \dfrac{1}{4}-\dfrac{1}{\sqrt{M}}.\end{cases}$$

These pairs (V_r, V_s) are chosen independent for different pairs r, s as above, and $V_k=\varnothing$ for the remaining integers k. It is clear that $\{X_n\}_{n\in\mathbf{Z}}=\{(Y_n, V_n)_{n\in\mathbf{Z}}\}$ is a stationary process. We claim that $\{X_n\}_{n\in\mathbf{Z}}$ is two-sided quite weak Bernoulli, but not weak Bernoulli. We first show the former.

Let $A(n,\varepsilon)=(-\infty, -\varepsilon n]\cup[\varepsilon n+n, \infty)$ and $B(n,\varepsilon)=[0, n]$. To show that $\{X_n\}_{n\in\mathbf{Z}}$ is two-sided quite weak Bernoulli, it suffices by Lemma 1.1 to show that $\forall\varepsilon>0$ there exists N such that $\forall n\geqslant N$, there exists a coupling $\{(\tilde{X}_k, \tilde{\tilde{X}}_k)\}_{k\in\mathbf{Z}}$ of $\{X_k\}_{k\in\mathbf{Z}}$ with itself such that

$$\{\tilde{X}_k\}_{k\in\mathbf{Z}} \text{ and } \{\tilde{\tilde{X}}_k\}_{k\in A(n,\varepsilon)} \text{ are independent} \qquad (*)$$

and

$$P(\tilde{X}_k = \tilde{\tilde{X}}_k \ \forall k \in B(n, \varepsilon)) \geqslant 1 - \varepsilon. \qquad (**)$$

As before, we denote by $a(n)$ the nearest integer to $\log_2 \log_2(n)$. Let $\varepsilon > 0$. Choose N so that for all $n \geqslant N$,

(1) $(2^{2^{a(n)-1}})^{4/3}(2^{a(n)-1}+2) \leqslant \varepsilon n$;

(2) If we have $\lfloor (2^{2^{a(n)-1}})^{4/3} \rfloor$ independent events, each with probability of at least $\frac{1}{2}^{(2^{a(n)-1}+2)}$ of occurring, then at least one of them will occur with probability $\geqslant 1 - \varepsilon/8$;

(3) $\sum_{j=1}^{\infty} \frac{1}{2}^{(2^{a(n)+j}+2)}(2^{a(n)+j}+1+3n) < \varepsilon/4$;

(4) $P(G_n) \geqslant 1 - \varepsilon/4$, where G_n will be given later in the proof (at which point, we will need to check that this is possible), and

(5) $4/\sqrt{a(n)} < \varepsilon/4$.

(A computation is required to see that such an N exists.) We now let $n \geqslant N$. We now construct the desired coupling satisfying $(*)$ and $(**)$. We first let $M_n = a(n)$.

Let $\{\tilde{Y}_k\}_{k \in \mathbf{Z}}$, $\{\tilde{\tilde{Y}}_k\}_{k \in A(n, \varepsilon)}$ be i.i.d. taking values 0 and 1 with probability $\frac{1}{2}$ each. For $k \notin A(n, \varepsilon)$, let $\tilde{\tilde{Y}}(k) = \tilde{Y}(k)$. Let E be the event that there is an $(M_n - 1)$-block completely contained in $[-\varepsilon n + 1, 0]$ and an $(M_n - 1)$-block completely contained in $[n, n + \varepsilon n - 1]$ for the process $\{\tilde{Y}_k\}_{k \in \mathbf{Z}}$ (or equivalently for the process $\{\tilde{\tilde{Y}}_k\}_{k \in \mathbf{Z}}$). Let F be the event that no $(M_n + j)$-block (for $j \geqslant 1$) intersects $[-\varepsilon n, n + \varepsilon n]$. If $(E \cap F)^c$ occurs, we then choose all the V_j's for the two processes independently. However, if $(E \cap F)$ occurs, we proceed as follows. Let B_1 denote the $(M_n - 1)$-block furthest to the right which is completely contained in $[-\varepsilon n + 1, 0]$ and let B_2 denote the $(M_n - 1)$-block furthest to the left, which is completely contained in $[n, n + \varepsilon n - 1]$. All V_j's for good pairs of $(M_n + j)$-blocks $(j \geqslant 1)$ are chosen independently in the two processes $\{\tilde{X}_k\}_{k \in \mathbf{Z}}$ and $\{\tilde{\tilde{X}}_k\}_{k \in \mathbf{Z}}$. For pairs of $(M_n - j)$-blocks $(j \geqslant 1)$ which are good and which are between (or at) B_1 and B_2 in the two processes (recall that the Y_k's are the same in the two process on $A(n, \varepsilon)^c$, and so these pairs would be in the same location in the two processes), we use the same V_k's for the two processes. For all other pairs of $(M_n - j)$-blocks $(j \geqslant 1)$ which are good, we use independent V_k's for the two processes. (Note that by construction such pairs necessarily do not intersect $[0, n]$.)

To finish the construction, we need to describe what we do for good M_n-block pairs. If a good pair of M_n-blocks is between B_1 and B_2 in the two processes, we use the same V_k's for the two processes. For good pairs of

M_n-blocks where neither block is between B_1 and B_2, we use independent V_k's in the two processes. Finally, consider the leftmost M_n-block between B_1 and B_2 (if one exists) in the two processes. If at least one of them is not the right block in a good pair of M_n-blocks, then let the corresponding V_k's be independent (and $\varnothing$ for at least one of them). If they are both the right block in a good pair of M_n-blocks, then, if z denotes the location of the left most 1 in this M_n-block, w denotes the location of the rightmost 1 in the corresponding M_n-block for the $\tilde{Y}$ process and w' denotes the rightmost 1 in the corresponding M_n-block for the $\tilde{\tilde{Y}}$ process (of course, w need not equal w'), we let

$$(\tilde{V}(w), \tilde{V}(z), \tilde{\tilde{V}}(w'), \tilde{\tilde{V}}(z)) = \begin{cases} (1, 1, 1, 1) & \text{with probability } \frac{1}{4}\left(\frac{1}{2}+\frac{2}{\sqrt{M_n}}\right) \\ (1, 0, 1, 0) & \text{with probability } \frac{1}{4}\left(\frac{1}{2}-\frac{2}{\sqrt{M_n}}\right) \\ (1, 1, 0, 1) & \text{with probability } \frac{1}{4}\left(\frac{1}{2}-\frac{2}{\sqrt{M_n}}\right) \\ (1, 0, 0, 0) & \text{with probability } \frac{1}{4}\left(\frac{1}{2}-\frac{2}{\sqrt{M_n}}\right) \\ (1, 1, 0, 0) & \text{with probability } \frac{1}{\sqrt{M_n}} \\ (0, 1, 1, 1) & \text{with probability } \frac{1}{4}\left(\frac{1}{2}-\frac{2}{\sqrt{M_n}}\right) \\ (0, 0, 1, 0) & \text{with probability } \frac{1}{4}\left(\frac{1}{2}-\frac{2}{\sqrt{M_n}}\right) \\ (0, 0, 1, 1) & \text{with probability } \frac{1}{\sqrt{M_n}} \\ (0, 0, 0, 0) & \text{with probability } \frac{1}{4}\left(\frac{1}{2}+\frac{2}{\sqrt{M_n}}\right) \\ (0, 1, 0, 1) & \text{with probability } \frac{1}{4}\left(\frac{1}{2}-\frac{2}{\sqrt{M_n}}\right). \end{cases}$$

The point of the above coupling is that we must do it so that $(\tilde{V}(w), \tilde{V}(z))$ has the correct joint distribution, $(\tilde{\tilde{V}}(w'), \tilde{\tilde{V}}(z))$ has the correct joint distribution, $(\tilde{V}(w), \tilde{V}(z))$ and $\tilde{\tilde{V}}(w')$ are independent, and so that

$\tilde{V}(z)$ and $\tilde{\tilde{V}}(z)$ are coupled in a good way. Note that the above procedure couples $\tilde{V}(z)$ and $\tilde{\tilde{V}}(z)$ correctly with probability $1-2/\sqrt{M_n}$.

Finally, we need to proceed analogously with the rightmost M_n-block between B_1 and B_2, if one exists.

It easily follows from our construction that $\{(\tilde{X}_k, \tilde{\tilde{X}}_k)\}_{k\in\mathbf{Z}}$ is a coupling of $\{X_k\}_{k\in\mathbf{Z}}$ with itself satisfying $(*)$ above.

To verify $(**)$, let G_n be the event that there is no M_n-block between B_1 and B_2, or if there is at least one, then the leftmost one is the right block in a good pair of M_n-blocks for both processes and the rightmost one is the left block in a good pair of M_n-blocks for both processes. It is clear that $P(G_n)$ is $\geqslant 1-\varepsilon/4$ for large n and so (4) above holds.

It follows from the construction and the degree to which $\tilde{V}(z)$ and $\tilde{\tilde{V}}(z)$ are coupled that, conditioned on $E\cap F\cap G_n$ occurring, $\{\tilde{X}_k=\tilde{\tilde{X}}_k\ \forall k\in B(n,\varepsilon)\}$ has exactly probability $(1-2/\sqrt{M_n})^2$ (which is $\geqslant 1-\varepsilon/4$ by (5)). Next, (1) and (2), together with a computation give $P(E)\geqslant 1-\varepsilon/4$, and (3), together with a computation, implies $P(F)\geqslant 1-\varepsilon/4$. This proves $(**)$ and completes the proof that $\{X_n\}$ is two-sided quite weak Bernoulli.

We now proceed to show that $\{X_n\}$ is not weak Bernoulli. Let $\varepsilon=\frac{1}{2}$. If $\{X_n\}$ were weak Bernoulli, it would follow that there exists ℓ such that there exists a coupling $\{(\tilde{X}_k, \tilde{\tilde{X}}_k)\}_{k\in\mathbf{Z}}$ of $\{X_k\}_{k\in\mathbf{Z}}$ with itself such that

$$\{\tilde{X}_k\}_{k\in(-\infty,0]\cup[\ell,\infty)} \text{ and } \{\tilde{\tilde{X}}_k\}_{k\in(-\infty,0]} \text{ are independent} \tag{$*$}$$

and

$$P(\tilde{X}_k=\tilde{\tilde{X}}_k\ \forall k\in[\ell,\infty))\geqslant \tfrac{1}{2}. \tag{$**$}$$

Note that the definition of weak Bernoulli really only tells us that there is a coupling $\{(\tilde{X}_k, \tilde{\tilde{X}}_k)\}_{k\in(-\infty,0]\cup[\ell,\infty)}$ of $\{X_k\}_{k\in(-\infty,0]\cup[\ell,\infty)}$ with itself with properties $(*)$ and $(**)$. However, if one has the latter coupling, then it is always possible to construct the former coupling (see Extension Lemma 4.2.4 in [1])

We now show that for all ℓ no such coupling exists. Let ℓ be given and assume we are given such a coupling. Let S be the random set of integers s with the property that for both the processes $\{\tilde{X}_k\}_{k\in\mathbf{Z}}$ and $\{\tilde{\tilde{X}}\}_{k\in\mathbf{Z}}$ there is no s-block intersecting $[0,\ell]$, and that there is a good pair of s-blocks with the left one contained in $(-\infty,0)$ and the right one contained in (ℓ,∞). If $s\in S$, let F_s be the left s-block and let H_s be the right s-block in the good s-block pair with the above property in the first process $\{\tilde{X}\}_{k\in\mathbf{Z}}$ (so that $F_s\subseteq(-\infty,0)$ and $H_s\subseteq(\ell,\infty)$), and let F'_s be the left s-block and let H'_s be the right s-block in the good s-block pair with the above property in the second process $\{\tilde{\tilde{X}}_k\}_{k\in\mathbf{Z}}$. Of course, if $\tilde{X}_k=\tilde{\tilde{X}}_k\ \forall k\in[\ell,\infty)$, then $H_s=H'_s$ for all $s\in S$.

Next, a computation similar to that needed to prove (2.1) shows that, with probability 1, $|n\backslash S|<\infty$. Enumerate S as $n_1, n_2, \ldots$. Let Q_k be the V_i variable at the rightmost 1 of F_{n_k} in the $\{\tilde{X}\}$ process, Q'_k be the V_i variable at the rightmost 1 of F'_{n_k} in the $\{\tilde{\tilde{X}}\}$ process, R_k be the V_i variable at the leftmost 1 of H_{n_k} in the $\{\tilde{X}\}$ process, and R'_k be the V_i variable at the leftmost 1 of H'_{n_k} in the $\{\tilde{\tilde{X}}\}$ process. By the assumption of the coupling, we have that

$$P(R_k = R'_k \ \forall k \geqslant 1) \geqslant \tfrac{1}{2}. \tag{2.3}$$

We now have the i.i.d. sequence $\{(Q'_k, R'_k)\}_{k\geqslant 1}$, where each marginal has distribution

$$\left(\frac{1}{4}+\frac{1}{\sqrt{n_k}}\right)\delta_{(1,1)}+\left(\frac{1}{4}+\frac{1}{\sqrt{n_k}}\right)\delta_{(0,0)}+\left(\frac{1}{4}-\frac{1}{\sqrt{n_k}}\right)\delta_{(1,0)}+\left(\frac{1}{4}-\frac{1}{\sqrt{n_k}}\right)\delta_{(0,1)},$$

and the i.i.d. sequence $\{(Q'_k, R_k)\}_{k\geqslant 1}$, where each marginal has distribution

$$\tfrac{1}{4}\delta_{(1,1)}+\tfrac{1}{4}\delta_{(0,0)}+\tfrac{1}{4}\delta_{(1,0)}+\tfrac{1}{4}\delta_{(0,1)};$$

(2.3) now gives us that

$$P((Q'_k, R_k)=(Q'_k, R'_k)\ \forall k\geqslant 1)\geqslant \tfrac{1}{2}.$$

Finally we obtain a contradiction as follows. Since there are only countably many cofinite sequences of integers (n_ℓ) and S is cofinite a.s., there must be a cofinite sequence (n'_ℓ) such that

$$P(\{S=(n'_\ell)\}\cap\{(Q'_k, R_k)=(Q'_k, R'_k)\ \forall k\geqslant 1\})>0.$$

However, the Kakutani dichotomy for product measures (see [10]) and the fact that

$$\sum_{\ell\geqslant 1}\left(\frac{1}{\sqrt{n'_\ell}}\right)^2=\infty$$

imply that the conditional distributions of (Q'_k, R_k) and (Q'_k, R'_k) given $S=(n'_\ell)$ (which are each product measures) are mutually singular. ∎

4. SOME PROPERTIES OF WEAK BERNOULLI

In this section, we give the proofs of Theorems 1.12, 1.14, and 1.15.

Proof of Theorem 1.12. Let $B'_n = [-n, n]^{d-1}$ ($\{0\}$ if $d=1$). We first note that for any coupling $(\{\tilde{C}^1_j\}_{j \in B'_n}, ..., \{\tilde{C}^{2d}_j\}_{j \in B'_n})$ of $2d$ copies of $\{C_j\}_{j \in B'_n}$, we have that for any $\varepsilon > 0$,

$$P(\max_{(i,j) \in [1, 2d] \times B'_n} \tilde{C}^i_j \geqslant \varepsilon n) \leqslant (2n+1)^{d-1} \, 2d \max_{i \in [1, 2d]} P(C_0 \geqslant \varepsilon n)$$

and the latter goes to 0 as $n \to \infty$ by the finite $(d-1)$th moment assumption.

For each n, let G_n be the random subset of B_n, where σ_1 and σ_2 agree in the nth coupling given to us in the definition of weak Bernoulli. We then have (where P denotes the nth coupling),

$$P(\Lambda_{n(1-\varepsilon)} \not\subseteq G_n) \leqslant P(\max_{(i,j) \in [1, 2d] \times B'_n} \tilde{C}^i_j \geqslant \varepsilon n)$$

which goes to 0 as $n \to \infty$ by the above. In view of Lemma 1.1, this proves that μ is QWB. ∎

Proof of Theorem 1.14. Let $B'_n = [-n, n]^{d-1}$ ($\{0\}$ if $d=1$). We first note that for any coupling $(\{\tilde{C}^1_j\}_{j \in B'_n}, ..., \{\tilde{C}^{2d}_j\}_{j \in B'_n})$ of $2d$ copies of $\{C_j\}_{j \in B'_n}$, we have that for any $\varepsilon > 0$,

$$P(\max_{(i,j) \in [1, 2d] \times B'_n} \tilde{C}^i_j \geqslant \varepsilon n) \leqslant (2n+1)^{d-1} \, 2d \max_{i \in [1, 2d]} P(C_0 \geqslant \varepsilon n)$$

and the latter is at most $c_\varepsilon e^{-\gamma_\varepsilon n}$ for all n and for some positive constants c_ε and γ_ε by the exponential moment assumption.

For each n, let G_n be the random subset of B_n, where σ_1 and σ_2 agree in the nth coupling given to us in the definition of weak Bernoulli. We then have (where P denotes the nth coupling),

$$P(\Lambda_{n(1-\varepsilon)} \not\subseteq G_n) \leqslant P(\max_{(i,j) \in [1, 2d] \times B'_n} \tilde{C}^i_j \geqslant \varepsilon n)$$

which is at most $c_\varepsilon e^{-\gamma_\varepsilon n}$ for all n by the above.

We now use the last statement in Lemma 1.1 which says that δ can be taken to be $\varepsilon^2/100$. This implies the QWBE property with perhaps two new constants c_ε and γ_ε. ∎

Proof of Theorem 1.15. The lower $(d-1)$-dimensional stationary process $\{C_k\}_{k \in \mathbf{Z}^{d-1}}$ will be a function of $\{(Y'_n, Y''_n)\}$, where $\{Y'_n\}$ and $\{Y''_n\}$ are independent copies of $\{Y_n\}$. For $k \in \mathbf{Z}^{d-1}$, let C_k be 1 plus the supremum of the nonnegative integers ℓ such that there is a path from the $(d-1)$-dimensional lattice $\{(j_1, j_2, ..., j_{d-1}, 0) : (j_1, j_2, ..., j_{d-1}) \in \mathbf{Z}^{d-1}\}$ to (k, ℓ) such that at every point on this path, at least one of Y' and Y'' takes the values -1. If there is no such nonnegative integer ℓ with this property, C_k is taken to be 0. While it is clear that $\{C_k\}_{k \in \mathbf{Z}^{d-1}}$ is stationary, it is not obvious that the value ∞ is not obtained with positive probability. However, the fact that C_0 has an exponential moment (and therefore of course that ∞ is obtained with zero probability) is contained in Proposition 2.4 in [8].

We now show that $\{Y_n\}$ is weak Bernoulli with respect to $\{C_k\}$. Given $n \geqslant 1$, we now construct a coupling $\{\sigma_1, \sigma_2, \{\tilde{C}^1_k\}_{k \in \mathbf{Z}^{d-1}}, ..., \{\tilde{C}^{2d}_k\}_{k \in \mathbf{Z}^{d-1}})$ of two copies of $\{Y_n\}$ and $2d$ copies of $\{C_k\}$ such that

(1) σ_1 and $\sigma_2|_{B_n^c}$ are independent, and

(2) $\bigcap_{i=1}^{2d} A_i \subseteq \{x: \sigma_1(x) = \sigma_2(x)\}$, where for $i = 1, 2, ..., d$

$$A_i = \{x \in B_n: \tilde{C}^i_{\hat{x}_i} \leqslant x_i + n\}$$

and for $i = d+1, d+2, ..., 2d$

$$A_i = \{x \in B_n: \tilde{C}^i_{\hat{x}_{i-d}} \leqslant n - x_{i-d}\},$$

where $\hat{x}_j = \{x_1, ..., x_{j-1}, x_{j+1}, ..., x_d\}$.

Our n is fixed. Let σ' and σ'' be independent copies of $\{Y_n\}$. We will now define $\sigma_1, \sigma_2, \{\tilde{C}^i_n\}_{n \in \mathbf{Z}^{d-1}}$, $i = 1, ..., 2d$, simply as functions of σ' and σ'', which gives a coupling with (1) and (2) satisfied. We first take σ_1 simply to be σ'.

For $i = 1, 2, ..., d$, let $\tilde{C}^i_k$ be 1 plus the supremum of the nonnegative integers ℓ, such that there is a path from the $(d-1)$-dimensional lattice $\{(j_1, ..., j_{i-1}, -n, j_{i+1}, ..., j_d) : (j_1, ..., j_{i-1}, j_{i+1}, j_d) \in \mathbf{Z}^{d-1}\}$ to $(k, \ell - n)$, such that at every point on this path at least one of σ' and σ'' takes the value -1. If there is no such nonnegative integer ℓ with this property, $\tilde{C}^1_k$ is taken to be 0.

For $i = d+1, d+2, ..., 2d$, let $\tilde{C}^i_k$ be 1 plus the supremum of the nonnegative integers ℓ, such that there is a path from the $(d-1)$-dimensional lattice $\{(j_1, ..., j_{i-1}, n, j_{i+1}, ..., j_d) : (j_1, ..., j_{i-1}, j_{i+1}, j_d) \in \mathbf{Z}^{d-1}\}$ to $(k, n - \ell)$, such that at every point on this path at least one of σ' and σ'' takes the value -1. If there is no such nonnegative integer ℓ with this property, $\tilde{C}^1_k$ is taken to be 0.

The final step is to use the notion of separating sets (see [8]). Letting $V = \{x \in B_n: \sigma'(x) = \sigma''(x)\}$, it is easily seen that any point $z \in \bigcap_{i=1}^{2d} A_i$ (which is now defined since the lower dimensional processes are defined) has the property that any path from z to the boundary of B_n necessarily intersects V. Consider the set $\mathscr{S}$ of all subsets W of V which have the property that for any point $z \in \bigcap_{i=1}^{2d} A_i$, any path from z to the boundary of B_n necessarily intersects W. (Of course, V is in $\mathscr{S}$). One can show (see [8]) that there is a set $W' \in \mathscr{S}$ which is minimal (no proper subset of W' is in $\mathscr{S}$) and has the property that any other set $W'' \in \mathscr{S}$ is contained within the volume surrounded by W', in that if $b \notin W'$ is such that there is a path from b to the boundary of B_n not intersecting W' (in words, b is outside the volume surrounded by W') then there is a path from b to the boundary of B_n not intersecting W''.

We finally let σ_2 be σ' on W' and at all points with the property that any path from them to the boundary of B_n intersects W' (in words, those points which are surrounded by W'), and let σ_2 be σ'' at all other points. This gives us a coupling which clearly satisfies all the required properties, except perhaps that σ_2 has the correct distribution. However, this last property follows from the Markov property (see [8] for more details). ∎

ACKNOWLEDGMENT

We thank the referee for generously providing us with very many concrete suggestions in order to make the paper more readable.

REFERENCES

1. H. C. P. Berbee, "Random Walks with Stationary Increments and Renewal Theory," Ph.D. thesis, Amsterdam, 1979.
2. J. van den Berg, A uniqueness condition for Gibbs measures with application to the 2-dimensional Ising antiferromagnet, *Comm. Math. Phys.* **152** (1993), 161–166.
3. J. van den Berg and C. Maes, Disagreement percolation for Markov random fields, *Ann. Probab.* **22** (1994), 749–763.
4. J. van den Berg and J. Steif, Percolation and the hard-core lattice gas model, *Stochastic Process. Appl.* **49**, No. 2 (1994), 179–197.
5. R. Bradley, A caution on mixing conditions for random fields, *Statist. Probab. Lett.* **8** (1989), 489–491.
6. R. Burton and J. E. Steif, Nonuniqueness of measures of maximal entropy for subshifts of finite type, *Ergodic Theory and Dynamical Systems* **14** (1994), 213–235.
7. R. Burton and J. E. Steif, New results on measures of maximal entropy, *Israel J. Math.* **89** (1995), 275–300.
8. R. M. Burton and J. E. Steif, Quite weak Bernoulli with exponential rate and percolation for random fields, *Stochastic Process. Appl.* **58**, No. 1 (1995), 35–55.
9. J. Chayes, L. Chayes, and R. H. Schonmann, Exponential decay of connectivities in the two-dimensional Ising model, *J. Statist. Phys.* **49** (1987), 433–445.
10. R. Durrett, "Probability: Theory and Examples," Wadsworth, Pacific Grove, CA, 1990.
11. N. A. Friedman and D. Ornstein, On Isomorphism of weak Bernoulli Transformations, *Adv. Math.* **5** (1971), 365–394.

12. S. Goldstein, Maximal coupling, *Z. Wahrsch. Verw. Gebiete* **46** (1979), 193–204.
13. F. di Liberto, G. Gallavotti, and L. Russo, Markov processes, Bernoulli schemes, and Ising model, *Commun. Math. Phys.* **33** (1973), 259–282.
14. T. Lindvall, “Lectures on the Coupling Method,” Wiley, New York, 1992.
15. F. Martinelli and E. Olivieri, Approach to equilibrium of Glauber dynamics in the one phase region I. The attractive case, *Commun. Math. Phys.* **161** (1994), 447–486.
16. D. S. Ornstein, “Ergodic Theory, Randomness and Dynamical Systems,” Yale Univ. Press, New Haven, CT, 1974.
17. R. H. Schonmann and S. B. Shlosman, Completely analyticity for 2D Ising completed, *Commun. Math. Phys.* **170** (1995), 453–482.
18. J. E. Steif, Space-Time Bernoullicity of the lower and upper stationary processes for attractive spin systems, *Ann. Probab.* **19** (1991), 609–635.
19. J. E. Steif, Consistent estimation of joint distributions for sufficiently mixing random fields, *Ann. Statist.* **25** (1997), 293–304.
20. V. A. Volkonskii and Yu. A. Rozanov, Some limit theorems for random functions I, *Theory Probab. Appl.* **4** (1959), 178–197.