

# On the equivalence of certain ergodic properties for Gibbs states

**Citation for published version (APA):**

Hollander, den, W. T. F., & Steif, J. E. (1997). *On the equivalence of certain ergodic properties for Gibbs states*. (Katholieke Universiteit Nijmegen. Mathematisch Instituut : report; Vol. 9724). Radboud Universiteit Nijmegen.

**Document status and date:**

Published: 01/01/1997

**Document Version:**

Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

**Please check the document version of this publication:**

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

**General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

[www.tue.nl/taverne](http://www.tue.nl/taverne)

**Take down policy**

If you believe that this document breaches copyright please contact us at:

[openaccess@tue.nl](mailto:openaccess@tue.nl)

providing details and we will investigate your claim.

# On the equivalence of certain ergodic properties for Gibbs states

Frank den Hollander, University of Nijmegen <sup>\*</sup>  
Jeffrey E. Steif, Chalmers University of Technology <sup>†</sup>

## Abstract

We extend our previous work by proving that for translation invariant Gibbs states on  $\mathbb{Z}^d$  with a translation invariant interaction potential  $\Psi = (\Psi_A)$  satisfying  $\sum_{A \ni 0} |A|^{-1} [\text{diam}(A)]^d \|\Psi_A\| < \infty$  the following hold:  
(1) the Kolmogorov-property implies a Trivial Full Tail;  
(2) the Bernoulli-property implies Følner Independence.  
The existence of bilaterally deterministic Bernoulli Shifts tells us that neither (1) nor (2) is true for random fields in general without some further assumption (even when  $d = 1$ ).

## 1 Introduction

The purpose of this paper is to extend some results for Markov random fields, that were proved in [HS], to a large class of (possibly infinite range) Gibbs states. In §1 we give some notations and definitions. In §2 we formulate our theorems. In §3 and §4 we give proofs.

**Notations and definitions.** Throughout this paper we consider stationary stochastic processes  $X = \{X_x\}_{x \in \mathbb{Z}^d}$  taking values in a *finite* set  $F$ . We also view  $X$  as a probability measure  $\mu$  on  $\Omega = F^{\mathbb{Z}^d}$  that is invariant under the natural  $\mathbb{Z}^d$ -action.

We write  $B_n = [-n, n]^d \cap \mathbb{Z}^d$  to denote the  $n$ -box in  $\mathbb{Z}^d$ . If  $\mu$  is a probability measure on  $F^{\mathbb{Z}^d}$  and  $A \subseteq \mathbb{Z}^d$ , then we let  $\mu_A$  denote the probability measure on  $F^A$  obtained by projecting  $\mu$  onto  $A$ . We also let  $X_A$  denote the process restricted to  $A$ , so that  $\mu_A$  is just the distribution of  $X_A$ .

---

<sup>\*</sup>Research partially carried out while visiting the Department of Mathematics, Chalmers University of Technology, Sweden in January 1996 and October 1997.

<sup>†</sup>Research supported by grants from the Swedish Natural Science Research Council and from the Royal Swedish Academy of Sciences.

In order to save space, rather than repeating verbatim a number of definitions we will frequently refer to [HS]. In particular, the reader can find there the definitions of the  $\underline{d}$ -distance between two probability measures  $\mu_A$  and  $\nu_A$  with finite  $A$ , entropy, ergodicity, K-AUTOMORPHISM (K), TRIVIAL FULL TAIL (TFT), BERNOULLI (B), VERY WEAK BERNOULLI (VWB), and FØLNER INDEPENDENCE (FI).

For translation invariant ergodic random fields the following orderings hold (see [HS], §1 and Theorem 2.4 and references):

$$\begin{aligned} FI &\subsetneq VWB; \quad TFT \subsetneq K \\ FI &\subsetneq TFT; \quad VWB \subsetneq K \\ B &= VWB. \end{aligned}$$

A Gibbs state is defined as follows (see [G], Chapter 2). An *interaction potential* is a family  $\Psi = (\Psi_A)$  of maps  $\Psi_A: F^A \rightarrow \mathbb{R}$  satisfying

$$\sum_{A: A \cap \Lambda \neq \emptyset} \|\Psi_A\| < \infty \text{ for all } \Lambda \subseteq \mathbb{Z}^d \text{ non-empty and finite,}$$

where  $\|\Psi_A\| = \sup_{\eta \in F^A} |\Psi_A(\eta)|$  and where  $A$  runs over the non-empty finite subsets of  $\mathbb{Z}^d$ . For a given  $\Psi$ , a *Gibbs state for  $\Psi$*  is any random field  $\mu$  whose conditional probabilities on  $\Lambda$  given  $\sigma$  on  $\Lambda^c$  are of the form

$$\mu(\cdot | \sigma) = \frac{1}{Z_{\Lambda, \sigma}} \exp[-H_{\Lambda}(\cdot | \sigma)] \text{ for all } \Lambda \subseteq \mathbb{Z}^d \text{ non-empty and finite and } \sigma \in F^{\Lambda^c},$$

where  $Z_{\Lambda, \sigma}$  is the normalizing constant (or partition sum),

$$H_{\Lambda}(\eta | \sigma) = - \sum_{A: A \cap \Lambda \neq \emptyset} \Psi_A([\eta \vee \sigma]_A) \quad (\eta \in F^{\Lambda})$$

is the Hamiltonian on  $\Lambda$  given  $\sigma$  on  $\Lambda^c$ , and  $[\eta \vee \sigma]_A$  is the configuration  $\eta \vee \sigma$  restricted to  $A$ .

The class of interaction potentials that we allow in this paper are the ones satisfying

$$(*) \quad \begin{cases} \Psi_A = \Psi_{A+z} \text{ for all } A \text{ and all } z \in \mathbb{Z}^d \\ \sum_{A \ni 0} \frac{1}{|A|} [\text{diam}(A)]^d \|\Psi_A\| < \infty, \end{cases}$$

where  $\text{diam}(A) = \sup_{x, y \in A} |x - y|_1$ . The second of these conditions means that for large sets the total interaction across the boundary of the set is of the order of the surface of the set.

Despite the fact that the interaction potential is assumed to be translation invariant, there may – and in general will – be Gibbs states that are not translation invariant. In this paper, however, we only consider translation invariant Gibbs states.

## 2 Main theorems

The goal of this paper is to show that the converses of ‘FI implies VWB’ and ‘TFT implies K’, though not true in general (see [HS] for a discussion), are true for all  $\mathbb{Z}^d$ -invariant Gibbs states for interactions satisfying (\*). That is, we prove the following two theorems.

**Theorem 2.1** *If  $\mu$  is a  $\mathbb{Z}^d$ -invariant Gibbs state for an interaction satisfying (\*) and is VWB, then  $\mu$  is FI.*

**Theorem 2.2** *If  $\mu$  is a  $\mathbb{Z}^d$ -invariant Gibbs state for an interaction satisfying (\*) and is K, then  $\mu$  is TFT.*

The proofs of these theorems are given in §3 and §4. Thus, for the class (\*) we obtain the following ordering:

$$(**) \quad FI = VWB \subseteq TFT = K.$$

### Remarks:

(1) For  $d = 1$ , (\*) precisely coincides with the well-known sufficient condition for uniqueness of the Gibbs state ([G], p. 166). Being the unique Gibbs state, the measure is necessarily TFT ([G], Theorem 7.7(a)). So Theorem 2.2 is of no interest for this case. In fact, for  $d = 1$ , (\*) is known to imply that the unique Gibbs state is Weak Bernoulli ([G], p. 461), which is stronger than FI. Therefore Theorem 2.1 is also of no interest in this case.

(2) Theorem 2.2 is trivial, for any  $d \geq 1$ , if all (!) Gibbs states for the given interaction are  $\mathbb{Z}^d$ -invariant. In fact, then ergodicity is already enough to imply TFT. The reason for this is that any such ergodic Gibbs state cannot be decomposed as a convex combination of two Gibbs states for the same interaction, since these would necessarily be  $\mathbb{Z}^d$ -invariant and by ergodicity would be identical. Hence, any such ergodic Gibbs state is extremal within the class of all Gibbs states, and therefore must be TFT (again by [G], Theorem 7.7(a)).

(3) In [OW1] it is proved that for the Ising model with ferromagnetic nearest-neighbor interaction both the ‘+ state’ and the ‘− state’ are B. So for this case all four properties in (\*\*) hold. The proof shows that the same is true for all interactions satisfying the FKG lattice condition ([G], p. 445), the technical reason being that then the conditional measure in a finite set is stochastically increasing as a function of the configuration outside the set.

(4) As will become clear from the proofs, both Theorem 2.1 and Theorem 2.2 are statements of the type: if a certain property holds ‘one-sided’ then it also holds ‘two-sided’ (i.e., if the property holds with respect to the lexicographic past of a large box, then it holds with respect to the entire outside of the large box). In

the theory of Gibbs states similar types of statements occur, for instance, for the notions of Markov property ([G], Section 10.1) and entropy [E].

(5) An open question is whether  $\text{TFT} = \text{VWB}$  for the class (\*). In [H] an example is constructed of a Markov random field on  $\mathbb{Z}^2$  that is K but not VWB. Since [HS] shows that  $\text{K} = \text{TFT}$  for Markov random fields in general, this example violates  $\text{TFT} = \text{VWB}$ . However, it is not Gibbsian (because it is not strictly positive on all cylinder sets). Perhaps a Gibbsian counterexample can be found in the class of nearest-neighbor ‘clock models’ [FS], where Gibbs states are known to exist that are unique and yet have arbitrarily slow decay of correlations.

(6) Another open question is whether (\*\*) also holds for the larger class of interactions where the second condition in (\*) is weakened to  $\sum_{A \ni 0} \|\Psi_A\| < \infty$ , i.e., the usual summability condition.

### 3 Proof of Theorem 2.1

**3.1 Key lemma.** We will need the following property of a Gibbs state for an interaction satisfying (\*), which plays an important role in the proofs of both Theorem 2.1 and Theorem 2.2.

**Lemma 3.1** *Fix an interaction satisfying (\*) and let  $\mu$  be a  $\mathbb{Z}^d$ -invariant Gibbs state for this interaction. Then, given  $\ell, m \in \mathbb{N}$  and  $\delta > 0$ , there exists a  $C(\ell, m, \delta)$ , satisfying*

$$\lim_{\ell \rightarrow \infty} C(\ell, m, \delta) = 1 \text{ for fixed } m \text{ and } \delta,$$

*such that for any  $k \in [\ell, m\ell] \cap \mathbb{N}$ , any  $\sigma, \sigma' \in F^{B_k^c}$  that agree on  $B_{k+[\delta\ell]} \setminus B_k$ , and any  $\eta \in F^{B_k}$ , the following bounds hold a.s.:*

$$\frac{1}{C(\ell, m, \delta)} \leq \frac{\mu_{B_k}(\eta|\sigma)}{\mu_{B_k}(\eta|\sigma')} \leq C(\ell, m, \delta).$$

PROOF: Fix  $m \in \mathbb{N}$  and  $\delta > 0$ . For  $k, \ell \in \mathbb{N}$ , let  $\mathcal{A}_{k, \ell, \delta}$  denote the collection of finite sets  $A$  satisfying  $A \cap B_k \neq \emptyset$  and  $A \cap B_{k+[\delta\ell]}^c \neq \emptyset$ . Given any finite set  $A$ , let  $T_A(k, \ell, \delta)$  denote the number of translates of  $A$  that are contained in  $\mathcal{A}_{k, \ell, \delta}$ . Some elementary combinatorial geometry (left to the reader) shows that there exists a  $C_1(m, \delta)$  such that

$$\sup_A \sup_{\ell \in \mathbb{N}} \sup_{k \in [\ell, m\ell] \cap \mathbb{N}} \frac{T_A(k, \ell, \delta)}{[\text{diam}(A)]^d} \leq C_1(m, \delta).$$

Next, for any  $l \in \mathbb{N}$ , any  $k \in [\ell, m\ell] \cap \mathbb{N}$ , any  $\sigma, \sigma' \in F^{B_k^c}$  that agree on  $B_{k+[\delta\ell]} \setminus B_k$ , and any  $\eta \in F^{B_k}$ , we have

$$\begin{aligned} |H_{B_k}(\eta|\sigma) - H_{B_k}(\eta|\sigma')| &\leq \sum_{A \in \mathcal{A}_{k,\ell,\delta}} \|\Psi_A\| \\ &= \sum_{A \ni 0} \frac{T_A(k,\ell,\delta)}{|A|} \|\Psi_A\| \\ &\leq C_1(m, \delta) \sum_{A \ni 0, \text{diam}(A) \geq [\delta\ell]} \frac{1}{|A|} [\text{diam}(A)]^d \|\Psi_A\|. \end{aligned}$$

By assumption (\*), the sum in the right-hand side tends to zero as  $\ell \rightarrow \infty$ . Hence there exists a  $C_2(\ell, m, \delta)$ , satisfying  $\lim_{\ell \rightarrow \infty} C_2(\ell, m, \delta) = 1$  for fixed  $m$  and  $\delta$ , such that

$$\frac{1}{C_2(\ell, m, \delta)} \leq \frac{e^{-H_{B_k}(\eta|\sigma)}}{e^{-H_{B_k}(\eta|\sigma')}} \leq C_2(\ell, m, \delta)$$

for any  $l, k, \sigma, \sigma', \eta$  as above. These inequalities being true for all  $\eta$ , the ratio of the corresponding partition functions also satisfies the exact same inequalities. This proves the claim with  $C(\ell, m, \delta) = C_2(\ell, m, \delta)^2$ .  $\square$

**3.2 Proof of Theorem 2.1.** If a process is VWB, then it is B (see §1). The latter is in turn equivalent to the following condition, called extremality (see [HS], §3 and references).

**Definition 3.2** A  $\mathbb{Z}^d$ -invariant probability measure  $\nu$  is called EXTREMAL if for all  $\epsilon > 0$  there exist an  $N \in \mathbb{N}$  and a  $\delta > 0$  such that: for all  $n \geq N$  and for all decompositions of  $\nu_{B_n}$  of the form

$$\nu_{B_n} = \sum_{i=1}^M p_i \nu_i$$

with  $(p_1, \dots, p_M)$  a probability vector and  $M \leq 2^{\delta|B_n|}$ , most of the  $\nu_i$ 's are  $\bar{d}$ -close to  $\nu_{B_n}$  in the sense that

$$\sum_{i: \bar{d}(\nu_{B_n}, \nu_i) < \epsilon} p_i > 1 - \epsilon.$$

In words, any 'not too large' decomposition of the measure on large blocks must have almost all components close to the original measure.

We must contrast extremality with the definition of FI, which reads:

**Definition 3.3** A  $\mathbb{Z}^d$ -invariant probability measure  $\nu$  is called FÖLNER INDEPENDENT if for all  $\epsilon > 0$  there exists an  $N \in \mathbb{N}$  such that: if  $n \geq N$  and  $S \subseteq B_n^c$  with  $S$  finite, then

$$\bar{d}(\mu_{B_n}, \mu_{B_n}/\sigma) < \epsilon$$

for all  $\sigma \in F^S$  except for an  $\epsilon$ -portion as measured by  $\mu$  (where  $\mu_{B_n}/\sigma$  denotes  $\mu_{B_n}$  conditioned on  $\sigma$ ).

In words, for large  $n$  and for most configurations on  $B_n^c$  the conditional distribution on  $B_n$  is  $\bar{d}$ -close to the unconditional distribution.

To show that ‘ $\mu$  is B’ implies ‘ $\mu$  is FI’, let  $\epsilon > 0$  and pick  $N_1, \delta$  from Definition 3.2. Next, choose  $\gamma > 0$  sufficiently small and pick  $N_2$  such that  $|F|^{|B_{n+[\gamma n]} \setminus B_n|} \leq 2^{\delta|B_n|}$  for all  $n \geq N_2$ . Next, pick  $N_3$  from Lemma 3.1 such that  $C(n, 1, \gamma) \leq 1 + \epsilon$  for all  $n \geq N_3$ . For such  $n$ , it follows readily from the bounds in Lemma 3.1 that, for any  $\sigma, \sigma' \in F^{B_n^c}$  that agree on  $B_{n+[\gamma n]} \setminus B_n$ , the measures  $\mu_{B_n}(\cdot|\sigma)$  and  $\mu_{B_n}(\cdot|\sigma')$  are within  $\epsilon$  in total variation distance.

By Lemma 3.2 in [HS], to verify the FI-condition in Definition 3.3 it suffices to consider  $n \geq \max\{N_1, N_2, N_3\}$  and finite sets  $S \subseteq B_n^c$  that contain  $B_{n+[\gamma n]} \setminus B_n$ . Since  $|F|^{|B_{n+[\gamma n]} \setminus B_n|} \leq 2^{\delta|B_n|}$ , extremality yields that there exist configurations  $\eta_1, \dots, \eta_M$  on  $B_{n+[\gamma n]} \setminus B_n$ , with  $M \leq |F|^{|B_{n+[\gamma n]} \setminus B_n|}$ , such that their total measure is at least  $1 - \epsilon$  and such that also  $\bar{d}(\mu_{B_n}, \mu_{B_n}/\eta_i) < \epsilon$  for each  $\eta_i$ .

Now consider all configurations  $\sigma$  on  $S$  such that the restriction of  $\sigma$  to  $B_{n+[\gamma n]} \setminus B_n$  is  $\eta_i$  for some  $i \in \{1, \dots, M\}$ . Clearly, these configurations have total measure at least  $1 - \epsilon$ , and so we need only show that for each such  $\sigma$ ,

$$\bar{d}(\mu_{B_n}, \mu_{B_n}/\sigma) < 2\epsilon.$$

For this it suffices to show that

$$\bar{d}(\mu_{B_n}/\eta, \mu_{B_n}/\sigma) < \epsilon$$

whenever  $\sigma$  is a configuration on  $S$  whose restriction to  $B_{n+[\gamma n]} \setminus B_n$  is  $\eta$ . However,  $\mu_{B_n}/\eta$  and  $\mu_{B_n}/\sigma$  are each averages of measures that, as we saw earlier, are all within  $\epsilon$  in total variation distance of each other. Hence  $\mu_{B_n}/\eta$  and  $\mu_{B_n}/\sigma$  are within  $\epsilon$  in total variation distance, and therefore also within  $\epsilon$  in  $\bar{d}$ -distance.  $\square$

## 4 Proof of Theorem 2.2

We will prove the result only for  $d = 2$ , the extension to higher dimensions being straightforward. The proof is a variation on the proof of the analogous statement for Markov random fields given in [HS]. The main point is to implement Lemma 3.1, which requires some estimates.

TFT means that the  $\sigma$ -algebra  $T$  defined by

$$\begin{aligned} T &= \bigcap_{m \geq 1} T_m \\ T_m &= \sigma(X_x, x \in B_m^c) \end{aligned}$$

is trivial. On the other hand,  $K$  is equivalent to the smaller  $\sigma$ -algebra  $T'$  defined by

$$\begin{aligned} T' &= \sigma(\bigcup_{m \geq 1} T'_m) \\ T'_m &= \bigcap_{n \geq 1} T'_{m,n} \\ T'_{m,n} &= \sigma(X_x, x \in \{(x_1, x_2) : x_2 \leq -n \text{ or } (x_1 \leq -n \text{ and } x_2 \leq m)\}) \end{aligned}$$

( $T'_{m,n}$  is the lexicographic past of the rectangle  $[-n, n] \times [-n, m]$  in  $\mathbb{Z}^2$ ) being trivial (see [HS], §1 and references). We will show that  $T = T'$  a.s, which more than implies the claim that  $K = \text{TFT}$ .

In order to do so, we appeal to Lemma 2.10 in [BH] (which is stated there only for  $d = 1$  but whose proof for higher dimensions is identical). According to this lemma, since  $T' \subseteq T$  it suffices to show that

$$h(X_{B_n}|T') = h(X_{B_n}|T) \quad \text{for all } n \geq 0,$$

where  $h(\cdot|\cdot)$  denotes conditional entropy.

Fix  $n \geq 0$ . Since  $T'_n \subseteq T' \subseteq T$ , it suffices to show that

$$h(X_{B_n}|T'_n) \leq h(X_{B_n}|T). \quad (1)$$

To achieve this, we will show that there exists a function  $\Delta(k, \ell, \delta) \geq 0$ , defined for  $k, \ell \in \mathbb{N}$  with  $k > 2n$  and for  $\delta > 0$ , satisfying

$$\lim_{\ell \rightarrow \infty} \frac{\Delta(k, \ell, \delta)}{(2\ell + 1)^2} = 0 \quad \text{for fixed } k \text{ and } \delta, \quad (2)$$

such that

$$h(X_{B_n}|T'_{n,k-n}) \leq h(X_{B_n}|T_{k(2\ell+1)-n}) + \alpha_{k,\ell,\delta} h(X_0) + \frac{\Delta(k, \ell, \delta)}{(2\ell + 1)^2}, \quad (3)$$

where  $h(\cdot)$  denotes entropy and

$$\alpha_{k,\ell,\delta} = \frac{\lfloor \delta \ell \rfloor (6r - 1) + \lfloor \delta \ell \rfloor (\lfloor \delta \ell \rfloor + 1)}{(2\ell + 1)^2} \quad \text{with } r = k(\ell + 1) - n.$$

Assuming the latter, we can let  $\ell \rightarrow \infty$ ,  $\delta \rightarrow 0$ ,  $k \rightarrow \infty$  (in this order) in (3) and use (2) to obtain (1). Note that  $\alpha_{k,\ell,\delta}$  vanishes in this limit and that, by the backwards martingale convergence theorem, the two entropies in (3) converge to the two entropies in (1).

To construct  $\Delta(k, \ell, \delta)$ , we define

$$C_{k,\ell} = \cup_{x,y: |x| \leq \ell, |y| \leq \ell} \{B_n + (kx, ky)\}$$

and note that the  $(2\ell + 1)^2$  translates of  $B_n$  comprising  $C_{k,\ell}$  are disjoint and have distance at least  $k - 2n$  between them. Let  $r = k(\ell + 1) - n$  as above and define

$$\begin{aligned} E_r &= \{(i, j) : j < -r\} \\ D_{r,\delta} &= B_{r+\lfloor \delta \ell \rfloor} \setminus (B_r \cup E_r). \end{aligned}$$

In words,  $E_r$  is the lower half plane adjacent to the bottom segment of the boundary of  $B_r$ , while  $D_{r,\delta}$  consists of  $\lfloor \delta \ell \rfloor$  layers adjacent to the left, right and top



segments of the boundary of  $B_r$ . Note that the boundary of  $B_r$  encloses  $C_{k,\ell}$  and is a distance  $k - 2n$  away from it.

We next order the  $(2\ell + 1)^2$  translates of  $B_n$  in  $C_{k,\ell}$  lexicographically. Namely, we say that  $B_n + (x, y)$  precedes  $B_n + (x', y')$  if  $y < y'$  or  $(y = y'$  and  $x < x')$ . In this way, we get an ordering of the translates of  $B_n$ , which we enumerate as  $B^1, B^2, \dots, B^{(2\ell+1)^2}$ . The idea of the proof is to compute the conditional entropy

$$(\dagger) = h(X_{D_{r,\delta}} \vee X_{C_{k,\ell}} | X_{E_r})$$

in *two different ways*, to derive an upper, respectively, lower bound for the two resulting expressions, and in this way obtain an inequality between these bounds. This inequality will then be exploited to complete the proof.

For the lower bound, we estimate

$$\begin{aligned} (\dagger) &\geq h(X_{C_{k,\ell}} | X_{E_r}) = h(\vee_{i=1}^{(2\ell+1)^2} X_{B^i} | X_{E_r}) \\ &= \sum_{i=1}^{(2\ell+1)^2} h(X_{B^i} | X_{E_r} \vee X_{B^1 \cup \dots \cup B^{i-1}}). \end{aligned}$$

Clearly, each of the terms in the sum is bounded below by  $h(X_{B_n} | T'_{n,k-n})$ , because the distance between the translates  $B^i$  is  $k - 2n$  and so is the distance between  $\cup_i B^i$  and  $E_r$ . Hence

$$(\dagger) \geq (2\ell + 1)^2 h(X_{B_n} | T'_{n,k-n}). \quad (4)$$

For the upper bound, we write

$$(\dagger) = h(X_{D_{r,\delta}} | X_{E_r}) + h(X_{C_{k,\ell}} | X_{E_r} \vee X_{D_{r,\delta}}).$$

The first term is at most  $|D_{r,\delta}|h(X_0)$ , where  $|D_{r,\delta}| = \sum_{i=1}^{\lfloor \delta \ell \rfloor} (6r - 1 + 2i) = \lfloor \delta \ell \rfloor (6r - 1) + \lfloor \delta \ell \rfloor (\lfloor \delta \ell \rfloor + 1)$ . We express the second term as  $h(X_{C_{k,\ell}} | X_{B_r^c}) + \Delta(k, \ell, \delta)$  with

$$\Delta(k, \ell, \delta) = h(X_{C_{k,\ell}} | X_{E_r} \vee X_{D_{r,\delta}}) - h(X_{C_{k,\ell}} | X_{B_r^c}) \geq 0$$

(the inequality coming from  $E_r \cup D_{r,\delta} \subseteq B_r^c$ ). We develop  $h(X_{C_{k,\ell}} | X_{B_r^c})$  as

$$\begin{aligned} h(X_{C_{k,\ell}} | X_{B_r^c}) &= h(\vee_{i=1}^{(2\ell+1)^2} X_{B^i} | X_{B_r^c}) \\ &\leq \sum_{i=1}^{(2\ell+1)^2} h(X_{B^i} | X_{B_r^c}) \\ &\leq (2\ell + 1)^2 h(X_{B_n} | T_{2r-(k-n)}), \end{aligned}$$

using the fact that the largest distance between the boundary of  $B_r$  and the center of a translate  $B^i$  is  $2r - (k - n)$ . Thus

$$(\dagger) \leq (2\ell + 1)^2 h(X_{B_n} | T_{2r-(k-n)}) + |D_{r,\delta}|h(X_0) + \Delta(k, \ell, \delta). \quad (5)$$

Comparing (4) and (5), noting that  $2r - (k - n) = k(2\ell + 1) - n$  and dividing by  $(2\ell + 1)^2$ , we obtain (3). Hence we need only verify (2) with the above definition of  $\Delta(k, \ell, \delta)$ .

To achieve the latter, we need the following trivial lemma.

**Lemma 4.1** *Let  $p = \{p_i\}_{i \in I}$  and  $q = \{q_i\}_{i \in I}$  be two finite probability vectors satisfying*

$$\frac{1}{C} \leq \frac{p_i}{q_i} \leq C \text{ for all } i \in I.$$

*Then  $h(q) \geq -\log C + \frac{1}{C}h(p)$ , where  $h(\cdot)$  denotes entropy.*

PROOF: Write

$$h(p) = \sum_i p_i \log \left( \frac{1}{p_i} \right) \leq \sum_i C q_i \log \left( \frac{C}{q_i} \right) = C \log C + Ch(q).$$

□

We want to apply Lemma 4.1 when  $p$  is the conditional law of  $X_{C_{k,\ell}}$  given  $X_{E_r} \vee X_{D_{r,\delta}}$  and  $q$  is the conditional law of  $X_{C_{k,\ell}}$  given  $X_{B_r^c}$ . Fix  $k$  and  $\delta$ . Applying Lemma 3.1 and averaging over the configuration in  $B_r \setminus C_{k,\ell}$ , we find that there exists a  $C(\ell)$  (namely,  $C(\ell) = C(\ell, 2k, \delta)$  in the notation of Lemma 3.1 because  $kl \leq r \leq 2kl$ ), satisfying

$$\lim_{\ell \rightarrow \infty} C(\ell) = 1,$$

such that for any  $\ell \in \mathbb{N}$ , any  $\eta \in F^{C_{k,\ell}}$ , any  $\sigma \in F^{E_r \cup D_{r,\delta}}$  and any  $\sigma' \in F^{B_r^c}$  whose restriction to  $E_r \cup D_{r,\delta}$  is  $\sigma$ , the following bounds hold a.s.:

$$\frac{1}{C(\ell)} \leq \frac{\mu(X_{C_{k,\ell}} = \eta | X_{E_r} \vee X_{D_{r,\delta}} = \sigma)}{\mu(X_{C_{k,\ell}} = \eta | X_{B_r^c} = \sigma')} \leq C(\ell)$$

(use that  $B_{r+[\delta\ell]} \setminus B_r \subseteq E_r \cup D_{r,\delta} \subseteq B_r^c$ ). Using Lemma 4.1, we now obtain (integrate over  $\eta, \sigma, \sigma'$ )

$$h(X_{C_{k,\ell}} | X_{B_r^c}) \geq -\log C(\ell) + \frac{1}{C(\ell)} h(X_{C_{k,\ell}} | X_{E_r} \vee X_{D_{r,\delta}})$$

and so

$$\begin{aligned} 0 \leq \Delta(k, \ell, \delta) &= h(X_{C_{k,\ell}} | X_{E_r} \vee X_{D_{r,\delta}}) - h(X_{C_{k,\ell}} | X_{B_r^c}) \\ &\leq \log C(\ell) + \left(1 - \frac{1}{C(\ell)}\right) h(X_{C_{k,\ell}} | X_{E_r} \vee X_{D_{r,\delta}}). \end{aligned}$$

But  $h(X_{C_{k,\ell}} | X_{E_r} \vee X_{D_{r,\delta}})$  can be bounded above by  $(2\ell + 1)^2 h(X_{B_n})$ . Hence (2) follows because  $\lim_{\ell \rightarrow \infty} C(\ell) = 1$ . □

**Acknowledgment** The authors thank A. van Enter for critical remarks while the paper was in progress.

## References

- [BH] Berbee, H.C.P. and den Hollander, W.Th.F., Tail triviality for sums of stationary random variables, *Ann. Probab.* **17** (1989) 1635–1645.
- [E] van Enter, A., On a question of Bratteli and Robinson, *Lett. Math. Phys.* **6** (1982) 289–291.
- [FS] Fröhlich, J. and Spencer, T., The Kosterlitz-Thouless transition in two-dimensional Abelian spin systems and the Coulomb gas, *Commun. Math. Phys.* **81** (1981) 527–602.
- [G] Georgii, H.-O., *Gibbs Measures and Phase Transitions*, de Gruyter 1988, New York.
- [H] Hoffman, C., A Markov random field which is  $K$  but not Bernoulli, *Israel J. Math.* (1998), to appear.
- [HS] den Hollander, F. and Steif, J.E., On  $K$ -automorphisms, Bernoulli shifts and Markov random fields, *Ergod. Th. Dynam. Sys.* **17** (1997) 405–415.
- [OW1] Ornstein, D.S. and Weiss, B.,  $\mathbb{Z}^d$ -actions and the Ising model, unpublished manuscript (1973).
- [OW2] Ornstein, D.S. and Weiss, B., Every transformation is bilaterally deterministic, *Israel J. Math.* **24** (1975) 154–158.

Frank den Hollander  
Department of Mathematics  
University of Nijmegen  
Toernooiveld 1  
6525 ED Nijmegen  
The Netherlands  
*denholla@sci.kun.nl*

Jeffrey E. Steif  
Department of Mathematics  
Chalmers University of Technology  
S-41296 Gothenburg  
Sweden  
*steif@math.chalmers.se*