J Stat Phys (2010) 139: 367–374
DOI 10.1007/s10955-010-9945-4

Exponential Rates of Convergence in the Ergodic Theorem: A Constructive Approach

G.G. Bosco · F.P. Machado · Thomas Logan Ritchie

Received: 14 October 2008 / Accepted: 7 March 2010 / Published online: 19 March 2010

Abstract We prove that, once an algorithm of perfect simulation for a stationary and ergodic random field F taking values in $S^{\mathbb{Z}^d}$, S a bounded subset of R^n, is provided, the speed of convergence in the mean ergodic theorem occurs exponentially fast for F. Applications from (non-equilibrium) statistical mechanics and interacting particle systems are presented.

Keywords Exponential rates · Ergodic theorem · Random fields · Perfect simulation

1 Introduction

This paper is about sufficient conditions for the occurrence of exponential rates of convergence in the mean ergodic theorem for random fields taking values in $\mathbb{R}^{\mathbb{Z}^d}$, that is, an array of real random variables indexed by the points of $\mathbb{Z}^d$.

If $F \equiv (F_i)_{i\in\mathbb{Z}^d}$ is a stationary and ergodic random field such that $\mathbb{E}(|F_0|) < \infty$, the mean ergodic theorem states that

$$\frac{\sum_{i\in\Lambda_n} F_i}{|\Lambda_n|} \xrightarrow{n} \mathbb{E}(F_0) \quad \text{a.s. and in } L^1$$

G.G. Bosco is supported by a grant from Capes.
T.L. Ritchie is supported by grants from CNPq and Fapesp.

G.G. Bosco (✉)
Faculdade de Filosofia Ciências e Letras de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, Brazil
e-mail: gerald@ffclrp.usp.br

F.P. Machado
Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, Brazil
e-mail: fmachado@ime.usp.br

T.L. Ritchie
Universidade Federal do ABC, Santo André, Brazil
e-mail: thomas.ritchie@ufabc.edu.br

We say that the convergence in the mean ergodic theorem occurs exponentially fast for F, if for all $\epsilon > 0$

$$\mathbb{P}\left(\left\{\left|\frac{\sum_{i\in\Lambda_n} F_i}{|\Lambda_n|} - \mathbb{E}(F_0)\right| > \epsilon\right\}\right) \xrightarrow{n} 0 \quad \text{exp. fast}$$

Since we will pursue a *constructive* approach, we make this meaning sufficiently clear at once.[1] Roughly speaking, we say that an ergodic and stationary random field F is *constructable*, if we can determine (with probability 1) the value of F_0, the spin value at the origin,[2] from the knowledge of a finite (but random) number of i.i.d. uniform random variables (usually also indexed by points of the $\mathbb{Z}^d$ lattice). This notion corresponds to the idea of *perfect simulation* (also known as *exact sampling*) stemming from the Propp-Wilson algorithm for finite Markov chains [7] and in the past ten years has been extended to other types of random objects, as in [2–4, 8, 9], for example. An important practical consequence of constructing a prescribed random field is that we can sample F exactly over any finite subset of $\mathbb{Z}^d$ generating for that purpose only a finite (random) number of i.i.d. uniform random variables.

The central result of this paper (Theorem 1) states that, if F is constructable and the spin space bounded, then convergence in the mean ergodic theorem occurs exponentially fast. Previously, this property was known to occur for i.i.d. random fields (Cramér's theorem), Markov sequences ([1]), Gibb's measures ([6]) and stationary distributions of attractive interacting particle systems ([5]). In the present case, it is somewhat remarkable that large deviations estimates may occur under such mild assumptions[3] on F and the authors were originally stricken when it was discovered that it was possible to prove large deviation estimates for F (under the above construction assumptions) assuming "only", in addition, that the probability to use more than n uniform random variables to determine F_0 should decay faster than $1/n$. However, as Jeffrey Steif pointed out, even this "mild" additional hypothesis was unnecessary. The proof (of Theorem 1) we present in Sect. 4 is a slight modification of our original proof made to dispense with the decay hypothesis above.

This paper is organized as follows. In Sect. 2 a simple example of construction is presented, so that the reader can bear a concrete image in mind while reading the abstract definitions of Sect. 3 and the proof of Theorem 1 in Sect. 4.

2 Random Sequential Adsorption: A Concrete Example of Construction

In order to make the forthcoming sections more clear to the non-specialist, we found it interesting to discuss briefly and informally a simple (but non trivial) example (from non-equilibrium statistical mechanics) of construction of a thermodynamic limit random field. We refer the reader to [8] for details and formal definitions.

For $n \in \mathbb{N}$, consider the following (random) algorithm for the (finite volume) one-dimensional nearest-neighbor-exclusion Random Sequential Adsorption model:

[1] See Sect. 3 for details.

[2] And consequently F_i, for all $i \in \mathbb{Z}^d$, due to the stationarity of F.

[3] Under these sole assumptions, for example, one can only prove that pair correlations decay to zero (as distance increases to infinity) but no further estimate is possible without further assumptions on F.

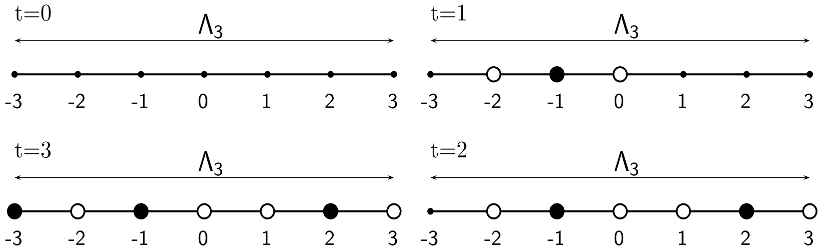

Fig. 1 At time $t = 0$, the box Λ_3 is empty. At time $t = 1$, a particle "chooses" site -1 and is fixed onto it forever. Sites -2 and 0 are declared *blocked*. At time $t = 2$, another particle chooses site 2 (uniformly amongst sites $-3, 1, 2, 3$). Sites 1 and 3 become blocked. At time $t = 3$, site -3 is chosen and the box Λ_3 gets jammed

> A site is chosen uniformly over the box $\Lambda_n = [-n, +n]$ and a particle is set onto it provided the site itself and both its nearest neighbors are unoccupied. Once occupied, a site remains so forever and both its neighbors are declared *blocked*. The process continues until all the sites of Λ_n are either blocked (spin 0/○) or occupied (spin 1/●).

Figure 1 illustrates a realization of the above algorithm for $n = 3$.

Call $X_n \in \{0, 1\}^{\Lambda_n}$ the terminal/jammed configuration and μ_n its corresponding probability distribution.

Our aim is to describe a limit parking algorithm over the whole $\mathbb{Z}$ lattice in such a way that $X \in \{0, 1\}^{\mathbb{Z}}$, the corresponding (thermodynamic) jammed configuration, is distributed according to $\mu \overset{\text{def}}{=} \lim_{n\to\infty} \mu_n$. This amounts to saying that the probability of occupying the origin of $\mathbb{Z}$ in the thermodynamic limit corresponds to the limit (as n increases) of the probabilities of occupying the center of Λ_n in the finite (volume) algorithms described above. And so forth, for every local event.[4]

We accomplish this task implementing all the finite algorithms together as follows

1. random numbers are (independently) generated throughout the $\mathbb{Z}$ lattice to produce a (random) profile $\omega \in [0, 1]^{\mathbb{Z}}$.
2. each entry of $\omega = (\omega_i)_{i\in\mathbb{Z}}$ is interpreted as the time a particle arrives at site i and
3. each box Λ_n is filled up to a jammed configuration according to the arrival times in ω and respecting the nearest-neighbor-exclusion rule.

Figure 2 illustrates this procedure for a given ω.

Call $l(\omega)$, the first local minimum of ω to the left of 0 and $r(\omega)$, the first local minimum of ω to the right of 0, so that $l(\omega) \le 0 \le r(\omega)$.

The key to the limit algorithm is to notice that, provided n is large enough so that $[l(\omega), r(\omega)] \subset \Lambda_n$, all finite algorithms will produce the same spin value at the origin in their corresponding jammed configurations. That is, given ω, all jammed configurations $X_n(\omega)$ will share the same spin value at the origin, $X_n(\omega)[0]$, for n sufficiently large. Therefore it is natural to define $X(\omega)[0]$, the spin value at the origin in the thermodynamic jammed configuration, as this constant.

In fact, it is not difficult to figure out that this constant is equal to 1 if, and only if, both $l(\omega)$ and $r(\omega)$ are even numbers and equal to 0, otherwise. In the above figure $l(\omega) = -3$

[4]The limit in the above definition of μ refers to the topology of weak convergence of probability measures.

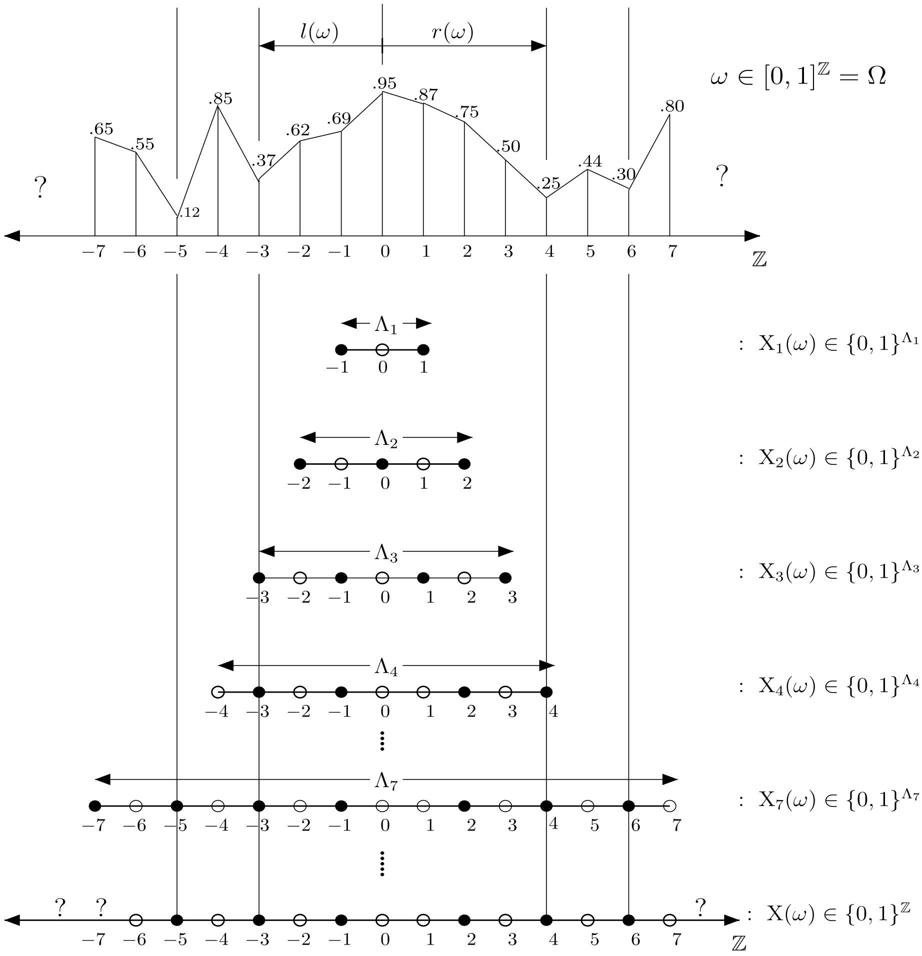

Fig. 2 Box Λ_1 as target: at time 0.69, the first particle occupies site -1, blocking site 0; at time 0.87 a second particle occupies site 1, site 0 is already blocked and the target region gets jammed, so that the third particle is lost at time 0.95; $(1, 0, 1)$ is assigned to $X_1(\omega)$. And so forth to the other boxes

and $r(\omega) = 4$, so that, provided $n \geq 4$, all X_n (and so X by definition) assign spin 0 to the origin.

Since the thermodynamic jamming limit X is translation invariant, the same procedure also applies to all sites of the $\mathbb{Z}$ lattice. Hence our task is completed.

In what follows, we call the rule of assigning a spin value to the origin from the knowledge of a finite (but random) number of uniform random variables the *construction function* for a prescribed (limit) random field.

3 Basic Definitions and Notation

Let $\Omega = [0, 1]^{\mathbb{Z}}$ be the set of doubly infinite sequences of real numbers in $[0, 1]$ endowed with the standard Borel (product) sigma-algebra $\mathcal{B}$ and Lebesgue product measure λ. Since $(\Omega, \mathcal{B}, \lambda)$ is a probability space, we will denote λ by $\mathbb{P}$ throughout this section.

Points in Ω will be denoted by ω, (ω_i), $(\omega_i)_{i\in\mathbb{Z}}$ or more explicitly by $(\ldots,\omega_{-1},\omega_0,\omega_1,\ldots)$. Moreover, for $n\in\mathbb{N}=\{0,1,2,\ldots\}$, we make $\Lambda_{i,n}\stackrel{\text{def}}{=}\{i-n,\ldots,i,\ldots,i+n\}$, the box with radius n centered at site $i\in\mathbb{Z}$ and $\omega|\Lambda_{i,n}\stackrel{\text{def}}{=}(\omega_{i-n},\ldots,\omega_i,\ldots,\omega_{i+n})$, the restriction of ω to $\Lambda_{i,n}$. For simplicity's sake, we denote $\Lambda_{0,n}$ by Λ_n.

Given a measurable function $f:\Omega\to\mathbb{R}$, we define on Ω the doubly infinite sequence of random variables $(Y_i)_{i\in\mathbb{Z}}$ by $Y_i=f\circ\theta_i$, where $[\theta_i(\omega)]_j\stackrel{\text{def}}{=}\omega_{j+i}$ is the usual shift operator in $[0,1]^{\mathbb{Z}}$. By its definition $Y\equiv(Y_i)$ is an ergodic and stationary random field (taking values) in $\mathbb{R}^{\mathbb{Z}}$.

In what follows, we will be interested in configurations $\omega\in\Omega$ which determine Y_i, the spin value at site i, only with the *information* inside the box $\Lambda_{i,n}$. For that purpose, for each $i\in\mathbb{Z}$, we define the random variable $R_i:\Omega\to\mathbb{N}$ by

$$R_i(\omega)\stackrel{\text{def}}{=}\inf\{n\in\mathbb{N}:(\omega'|\Lambda_{i,n}=\omega|\Lambda_{i,n})\Rightarrow(Y_i(\omega')=Y_i(\omega))\}$$

That is, on $\{R_i\le n\}$, Y_i, the spin at site i, is determined with the information inside the box $\Lambda_{i,n}$.

Note that, at first, the R_i's are allowed to assume the ∞ value but in what follows we will restrict this possibility to an event of probability zero.

Again, we denote R_0 by R for simplicity and since $R_i=R\circ\theta_i$, $(R_i)_{i\in\mathbb{Z}}$ is an ergodic and stationary random field. In particular $\mathbb{P}(\{R_i\le n\})=\mathbb{P}(\{R\le n\})$, for all $i\in\mathbb{Z}$.

Now suppose that, given μ, a probability measure in (the Borel sigma-algebra of) $\mathbb{R}^{\mathbb{Z}}$, we can identify a measurable map $f:\Omega\to\mathbb{R}$ such that the random field $(Y_i)_{i\in\mathbb{Z}}$, defined by $Y_i=f\circ\theta_i$ as above, is distributed exactly according to μ. If, besides, $\mathbb{P}(\{R>n\})$ converges to 0 (as n increases), we say that f is a *construction function* for μ. Correspondingly, in case μ is the probability distribution of a prescribed random field F (taking values) in $\mathbb{R}^{\mathbb{Z}}$, we call f, a *construction function* for F.[5]

The notion of construction in the above paragraph corresponds to the idea of *perfect simulation* (also known as *exact sampling*) in the sense that one can exactly simulate (in a computer, for example) a finite window of a prescribed (limit) distribution. In this sense, Y may be thought of as an explicit construction of F. Examples of perfect simulation of (limit) random fields can be found in [2–4, 8, 9].

4 Main Result

If $F\equiv(F_i)_{i\in\mathbb{Z}}$ is an ergodic and stationary random field in $\mathbb{R}^{\mathbb{Z}}$ such that $\mathbb{E}(|F_0|)<\infty$, the mean ergodic theorem states that

$$\frac{\sum_{i\in\Lambda_n}F_i}{|\Lambda_n|}\stackrel{n}{\longrightarrow}\mathbb{E}(F_0)\quad\text{a.s. and in } L^1 \tag{1}$$

We say that the convergence in the mean ergodic theorem occurs exponentially fast for F, if for all $\epsilon>0$

$$\mathbb{P}\left(\left\{\left|\frac{\sum_{i\in\Lambda_n}F_i}{|\Lambda_n|}-\mathbb{E}(F_0)\right|>\epsilon\right\}\right)\stackrel{n}{\longrightarrow}0\quad\text{exp. fast} \tag{2}$$

[5]Once identified, f will be well defined up to an event of probability zero.

in the sense that we can find positive constants C and α (depending on ϵ) such that the probabilities in (2) are bounded by $C\exp(-\alpha\cdot|\Lambda_n|)$. In this case we also say that exponential rates of convergence in the mean ergodic theorem apply to F.

Exponential rates of convergence in the mean ergodic theorem are known to occur for i.i.d. random fields (Cramér's theorem), Markov sequences ([1]), Gibb's measures ([6]) and stationary distributions of attractive interacting particle systems ([5]). In the sequel we prove that 2 also holds, provided we can identify a bounded construction function for F.[6]

Theorem 1 *If a bounded construction function f is identified for a random field F taking values in $\mathbb{R}^{\mathbb{Z}}$, then exponential rates of convergence in the mean ergodic theorem apply to F.*

Proof Suppose $\epsilon > 0$ is given.

Without loss of generality we assume that $f : \Omega \to [-1/2, 1/2]$ and $\mathbb{E}(F_0) = 0$. Moreover, we adhere to the notation of Sect. 3 and define on Ω the random field $Y \equiv (Y_i)_{i\in\mathbb{Z}}$ by $Y_i = f\circ\theta_i$. We recall that both Y and F share the same probability distribution μ in $\mathbb{R}^{\mathbb{Z}}$.

The key idea to the proof is to construct on an enlarged probability space a copy W of Y together with W', a field that approximates W locally and, at the same time, is block dependent, that is the local spins W'_i and W'_j are independent whenever the distance between sites i and j is sufficiently large. We proceed with this construction below.

Let $U \equiv (U_i)_{i\in\mathbb{Z}}$ and $V \equiv (V_{i,j})_{i\in\mathbb{Z},j\in\mathbb{Z}}$ be two independent i.i.d. families of uniform random variables in $[0, 1]$ defined on an abstract probability space $(X, \mathcal{X}, \mathbb{P})$.[7]

Since U is distributed (in Ω) according to λ, it is clear that the random field $W \equiv (W_i)_{i\in\mathbb{Z}} \stackrel{\text{def}}{=} Y(U)$, such that $W_i = f(\ldots, U_{i-n},\ldots, U_i,\ldots, U_{i+n},\ldots)$, is also distributed according to μ.

Choose n' large enough so that

$$\lambda(\{R > n'\}) < \epsilon/4 \tag{3}$$

and define the random field $W' \equiv (W'_i)_{i\in\mathbb{Z}}$ by

$$W'_i = f(\ldots, V_{i-n'-2,i}, V_{i-n'-1,i}, U_{i-n'},\ldots, U_i,\ldots, U_{i+n'}, V_{i+n'+1,i}, V_{i+n'+2,i},\ldots)$$

That is, at each site $i \in \mathbb{Z}$, W' shares the same information as W on the box $\Lambda_{i,n'}$, which implies that $\mathbb{P}(\{W'_i \neq W_i\}) < \epsilon/4$ in view of (3). Moreover W'_i and W'_j are independent, provided that $|i - j| > 2n'$ because under this condition W'_i and W'_j are measureable on independent sigma-algebras, viz $\sigma(\ldots, V_{i-n'-2,i}, V_{i-n'-1,i}, U_{i-n'},\ldots, U_i,\ldots, U_{i+n'}, V_{i+n'+1,i}, V_{i+n'+2,i},\ldots)$ and $\sigma(\ldots, V_{j-n'-2,j}, V_{j-n'-1,j}, U_{j-n'},\ldots, U_j,\ldots, U_{j+n'}, V_{j+n'+1,j}, V_{j+n'+2,j},\ldots)$ respectively. That is, W' is a block dependent stationary and ergodic random field. Therefore exponential rates of convergence in the mean ergodic theorem apply to W', so that

$$\mathbb{P}\left(\left\{\left|\frac{\sum_{i\in\Lambda_n} W'_i}{|\Lambda_n|}\right| > \epsilon/2\right\}\right) \xrightarrow{n} 0 \quad \text{exp. fast} \tag{4}$$

At last, define $I \equiv (I_i)_{i\in\mathbb{Z}}$ by

$$I_i = \begin{cases} 1 & \text{on } \{R_i(U) > n'\} \\ 0 & \text{on } \{R_i(U) \leq n'\} \end{cases}$$

[6] A straightforward application of the first Borel-Cantelli lemma shows that (2) is stronger than (1).

[7] To avoid ambiguity, henceforth we denote the probability measure of Ω (Sect. 3) only by λ.

In words, I_i indicates whether the information inside the box $\Lambda_{i,n'}$ is not sufficient to determine W_i. As before, the random variables I_i and I_j will independent whenever $|i - j| > 2n'$ because I_i and I_j are measureable on independent sigma-algebras, viz $\sigma(U_{i-n'}, \dots, U_i, \dots, U_{i+n'})$ and $\sigma(U_{j-n'}, \dots, U_j, \dots, U_{j+n'})$ respectively. Hence I is a block dependent, stationary and ergodic random field with $0 \le \mathbb{E}(I_0) = \mathbb{P}(\{R_i(U) > n'\}) < \epsilon/4$ and therefore

$$\mathbb{P}\left(\left\{\frac{\sum_{i\in\Lambda_n} I_i}{|\Lambda_n|} > \epsilon/2\right\}\right) \xrightarrow{n} 0 \quad \text{exp. fast} \tag{5}$$

Now observe that

$$\left|\frac{\sum_{i\in\Lambda_n} W_i - W_i'}{|\Lambda_n|}\right| \le \frac{\sum_{i\in\Lambda_n} |W_i - W_i'|}{|\Lambda_n|} \le \frac{\sum_{i\in\Lambda_n} I_i}{|\Lambda_n|} \tag{6}$$

since by our former assumption on the range of f, casual discrepancies between W_i and W_i' on $\{R_i(U) > n'\}$ are bounded by 1 and

$$\begin{aligned}\left|\frac{\sum_{i\in\Lambda_n} W_i}{|\Lambda_n|}\right| &= \left|\frac{\sum_{i\in\Lambda_n} (W_i - W_i') + W_i')}{|\Lambda_n|}\right| \\ &\le \left|\frac{\sum_{i\in\Lambda_n} W_i - W_i'}{|\Lambda_n|}\right| + \left|\frac{\sum_{i\in\Lambda_n} W_i'}{|\Lambda_n|}\right| \\ &\le \frac{\sum_{i\in\Lambda_n} I_i}{|\Lambda_n|} + \left|\frac{\sum_{i\in\Lambda_n} W_i'}{|\Lambda_n|}\right| \end{aligned} \tag{7}$$

So that

$$\left\{\left|\frac{\sum_{i\in\Lambda_n} W_i}{|\Lambda_n|}\right| > \epsilon\right\} \subset \left\{\frac{\sum_{i\in\Lambda_n} I_i}{|\Lambda_n|} > \epsilon/2\right\} \cup \left\{\left|\frac{\sum_{i\in\Lambda_n} W_i'}{|\Lambda_n|}\right| > \epsilon/2\right\} \tag{8}$$

and thus (5) and (4) together yield

$$\mathbb{P}\left(\left\{\left|\frac{\sum_{i\in\Lambda_n} W_i}{|\Lambda_n|}\right| > \epsilon\right\}\right) \xrightarrow{n} 0 \quad \text{exp. fast} \tag{9}$$

The result follows observing that W has the same distribution of F. □

4.1 Remarks and Open Problem about Theorem 1

Of course if the range of f is not $[-1/2, 1/2]$ or $\mathbb{E}(f) \ne 0$ the result follows applying a linear transformation to f. Moreover, a close observation to all steps in the above proof shows that it is dimension- independent, that is, it works out equally well in any dimension. At least, the range of f could be assumed to be any bounded subset of $\mathbb{R}^n$ so that Theorem 1 could be rephrased as follows

Theorem 2 *If a bounded construction function f is identified for a random field F taking values in $(\mathbb{R}^n)^{\mathbb{Z}^d}$, then exponential rates of convergence in the (multidimensional) mean ergodic theorem apply to F.*

A more subtle point is the following. In many applications (especially those coming from interacting particle systems) it is convenient to attach more than a single uniform random variable to each site of $\mathbb{Z}^d$. Again, a close examination of the above proof would indicate it also works out adequately in this slightly different context. The important fact being how far f takes information from and not how much data is attached to each site of $\mathbb{Z}^d$.

On the order hand, the boundedness assumption about the construction function f seems hard to weaken. It is reasonable to conjecture that Theorem 1 still remains valid (perhaps under further suitable assumptions on the tail of R), provided (the distribution of) f has exponential tails, since this is indeed the case for i.i.d. random fields (Cramér's theorem). Whether this is true or not, we don't know.

Acknowledgements The authors are thankful to Jeffrey Steif for his comments on the proof of Theorem 1 and to K. Marton and P. Shields from whom Steif's ideas probably stem.

References

1. Baum, L.E., Katz, M., Read, R.R.: Exponential convergence rates for the law of large numbers. Trans. Am. Math. Soc. **102**, 187–199 (1962)
2. Ferrari, P.A., Fernandez, R., Garcia, N.L.: Perfect simulation for interacting point processes, loss networks and Ising models. Stoch. Process. Appl. **102**, 63–88 (2002)
3. Galves, A., Garcia, N.L., Löcherbach, E.: Perfect simulation and finitary coding for multicolor systems with interactions of infinite range. Preprint
4. Häggström, O., Steif, J.E.: Propp-Wilson algorithms and finitary codings for high noise Markov random fields. Comb. Probab. Comput. **9**, 425–439 (2000)
5. Lebowitz, J.L., Schonmann, R.H.: Pseudo-free energies and large deviations for non Gibbsian FKG measures. Probab. Theory Relat. Fields **77**, 49–64 (1988)
6. Olla, S.: Large deviations for Gibbs random fields. Probab. Theory Relat. Fields **77**, 343–357 (1988)
7. Propp, J.G., Wilson, D.B.: Exact sampling with coupled Markov chains and applications to statistical mechanics. Random Struct. Algorithms **9**, 223–252 (1996)
8. Ritchie, T.L.: Construction of the thermodynamic jamming limit for the parking process and other exclusion schemes on $\mathbb{Z}^d$. J. Stat. Phys. **122**(3), 381–398 (2006)
9. van den Berg, J., Steif, J.E.: On the existence and non-existence of finitary codings for a class of random fields. Ann. Probab. **27**, 1501–1522 (1999)