

# Erdős-Rényi Laws for Gibbs Measures

Francis M. Comets<sup>\*</sup>

Université Paris 7, UFR de Mathématiques, case 7012, 2, place Jussieu,  
F-75251 Paris Cédex 05, France, Email: comets@mathp7.jussieu.fr, fax: 33-1-44 27 69 35

Received: 23 December 1992/in revised form: 5 October 1993

**Abstract:** Can one detect a phase transition from a single, large sample of a Gibbs measure? What information does one get on the other Gibbs distributions with the same potential? We approach these questions via Erdős-Rényi laws. In particular we prove almost-sure limit theorems for sets of empirical distributions of sub-samples of the given one: for suitable sub-samples size this set converges to the set of stationary Gibbs measures. Moreover we formulate Erdős-Rényi laws for general families of random variables with suitable large deviation principles.

## I. Introduction

On a single realization of a random field on the lattice observed in a large box, one can see smaller windows where the sample shows a large deviation from its typical behavior. The smallest the size of the windows, the most unlikely the deviation. This is the underlying idea of Erdős-Rényi laws, which are well known from statisticians for independent identically distributed random variables.

In this paper, we prove Erdős-Rényi type laws for Gibbs distributions, with a particular emphasis on (first-order) phase transition. Let us illustrate our results in the case of a real valued, finite range interaction Gibbs random field  $P$ . Consider the

average spin  $M_\Lambda(\omega) = \frac{1}{|\Lambda|} \sum_{i \in \Lambda} \omega_i$  of the sample  $\omega = (\omega_i)_i$  on a cubic box  $\Lambda$  with cardinality  $|\Lambda|$ . It is well known that for large  $\Lambda$ ,  $P\{M_\Lambda \geq x\}$  behaves approximately like  $\exp\{-|\Lambda|\lambda(x)\}$ , with  $\lambda(x) > 0$  if  $x$  is larger than some number. Let us observe the sample on a box  $\Lambda$ ; for cubic windows  $\Lambda' \subset \Lambda$  the Erdős-Rényi statistics  $M_{\Lambda, \Lambda'}^+$  is the largest average spin  $M_{i+\Lambda'}$  among all the translates  $i + \Lambda'$  of  $\Lambda'$  which are included in  $\Lambda$ . Then, for all such  $x$  the Erdős-Rényi law proved in this paper states that

$$M_{\Lambda, \Lambda'}^+ \longrightarrow x \quad P\text{-a.s.} \quad (1.1)$$

---

<sup>\*</sup> URA CNRS 756 (Centre de Mathématiques Appliquées, Ecole Polytechnique) et 1321 (Statistique et Modèles Aléatoires, Université Paris 7)

if  $\Lambda$  and  $\Lambda'$  increase to the whole lattice in such a way that  $|\Lambda'| \sim \frac{1}{\lambda(x)} \text{Log} |\Lambda|$ . The reason is that the large number of windows balances the small probability for deviation. The limiting case of interest when  $|\Lambda'|/\text{Log} |\Lambda| \rightarrow \infty$ , may be investigated too: for example, if the interaction has finite range and if  $|\partial\Lambda'| \sim c \text{Log} |\Lambda|$  for  $c$  smaller than some  $c_0 > 0$ , then

$$M_{\Lambda, \Lambda'}^+ \longrightarrow m^+ \quad P\text{-a.s.}, \tag{1.2}$$

where  $m^+$  is the largest expected value of spin  $E^Q \omega_0$  for  $Q$  ranging over the set of ergodic Gibbs distributions with the same potential as  $P$ . Similarly, the smallest such expectation may be estimated with the smallest moving average  $M_{\Lambda, \Lambda'}^-$ .

Let us quote some consequences of (1.2). A single sample from  $P$  contains information – like  $m^-$  and  $m^+$  – on other Gibbs measures. If we assume that  $m^- \neq m^+$ , then with probability one the phase transition will be detected in the limit of an infinite sample size. In fact, we will give convergent estimators of the set of ergodic Gibbs measures itself. Here is another aspect of (1.2). Let us call “–state” any ergodic Gibbs distribution  $P$  such that  $E^P \omega_0 = m^-$ . Then the statement (1.2) gives a quantitative description of a “–state” realization as small islands typical from “+states” into a sea more typical of the “–state”; this is somewhat reminiscent of the classical picture of the Ising model in the two-phases region. Note that (1.2) yields a lower bound for the size of the larger island of “+state.”

In this paper, we consider rather general models. For instance, we assume translation invariance for the interaction potential only, but not for the underlying measure  $P$ . As a counterpart, the results involving deviation estimates like (1.2) inside the set of Gibbs measures will not provide tight constants (like the largest possible value of  $c_0$  in (1.2) and the exact size of islands). Getting the exact constants is more stringent. They should be derived both from tight estimates involving the surface tension like those from Dobrushin, Kotecky and Shlosman (1992) in the two dimensional Ising model, and from exponential mixing properties for the extremal states.

Knowing if multiplicity of phase occurs or not is important for statistical purposes, since it gives different asymptotics for estimators of parameters like temperature, external field, . . . Our result indicate possible hypothesis testing.

There is a vast literature on Erdős-Rényi laws for independent variables: we mention just a few below. Erdős and Rényi (1970) proved their “new strong law of large numbers” using the Cramér-Chernov large deviation estimates, in the more refined version given by Bahadur and Ranga-Rao. Later S. Csörgő (1979 a) observed that their result was a consequence of the original Cramér-Chernov estimates, and he developed applications to statistics via the empirical measure in the same paper as well as to the theory of efficiency in the sense of Bahadur (S. Csörgő, 1979 b). Erdős and Révész (1975) studied the size of the longest head-run in the Bernoulli model. The case of Markov chains can be traced in papers of Samarova (1981), Arratia and Waterman (1989); the latter contains applications to molecular biology. Related questions are the strong limit theorems for large exceedances of partial sums of independent or Markov-dependent variables (Dembo and Karlin 1991 a, b). Note that in the theory of small markovian perturbations of dynamical systems (Freidlin and Wentzell (1984)), the analysis of the long time behavior is somewhat analogous to Erdős-Rényi laws. A complete analysis of the fluctuations of Erdős-Rényi statistics was achieved in the independent case by Deheuvels, Devroye and Lynch (1986),

using refined large deviation estimates; such an estimate is not available in the Gibbs framework up to now.

A generalization of Erdős-Rényi laws is given in a paper by Book (1975) for maxima of moving, suitably normalized sums of independent variables. There, the moving window is larger than in (1.1)—including that of (1.2) for instance—the normalization is weaker than the window size, and the convergence results take into account only the moderately large deviations. Our result (1.2) shows that Gibbs random fields exhibit a completely different behavior.

The paper is organized as follows. We start with definitions of Gibbs measures and we recall their large deviation properties. In Sect. III we prove, for a general random field obeying a large deviation principle with uniformity in the lower bound, the Erdős-Rényi law at the level of the empirical field; the result also applies to hypermixing processes. Section IV is devoted to estimating the set of stationary Gibbs measures itself. In Sect. V, we concentrate on empirical averages of a given local function  $f$  of the sample: we prove a.s. convergence of the functional estimator of the corresponding large deviation rate function. This section was motivated by simulation experiments, where the functional estimator appeared to be far more informative than the (slowly converging) point estimators. Experimental results are reported in Sect. VI.

## II. Gibbs Measures

Let  $\Omega_0$  be a polish space (separable complete metric space) and  $D$  some positive integer. For all  $i \in \mathbb{Z}^D$ , we will denote by  $\theta_i$  the shift operator on  $\Omega = \Omega_0^{\mathbb{Z}^D}$  given by  $(\theta_i \omega)_j = \omega_{i+j}$  if  $\omega = (\omega_i)_{i \in \mathbb{Z}^D} \in \Omega$  and  $j \in \mathbb{Z}^D$ . A stationary random field on  $\Omega$  is a probability measure on the borel  $\sigma$ -field of  $\Omega$  which is invariant by the shift operators  $\theta_i, i \in \mathbb{Z}^D$ . The set  $\mathbb{P}_s(\Omega)$  of stationary random fields on  $\Omega$  endowed with the weak topology is itself a polish space, with the Prohorov metric  $d$ .

A stationary interaction potential on  $\Omega$  is a collection  $\Phi = \{\Phi_V; V \subset \mathbb{Z}^D, V \text{ finite}\}$  such that  $\Phi_V$  is a continuous and bounded function on  $\Omega$ , measurable with respect to the  $\sigma$ -field generated by the coordinates  $\omega_i, i \in V$ , and such that  $\Phi_V \circ \theta_i = \Phi_{i+V}$ .

In this paper, we are given such a potential  $\Phi$ , and we assume in addition

$$\|\Phi\| = \sum_{V; 0 \in V} \|\Phi_V\|_\infty < \infty. \tag{2.1}$$

The range  $r_\Phi \in [0, \infty]$  of the potential  $\Phi$  is the smallest  $r$  in  $[0, \infty]$  such that  $\Phi_V = 0$  for all  $V$  with diameter larger than  $r$ . Important examples of potentials are finite range ones.

For  $\omega, \omega' \in \Omega$  we will denote  $(\omega_i)_{i \in \Lambda}$  by  $\omega_\Lambda, (\omega_i)_{i \notin \Lambda}$  by  ${}_\Lambda \omega$ , and by  $\omega_\Lambda \vee \omega'$  the element of  $\Omega$  equal to  $\omega$  on  $\Lambda$  and to  $\omega'$  on  $\Lambda^c$ . Let  $\alpha$  be a given probability measure on  $\Omega_0$ . Gibbs measures are defined by the D.L.R. equations:

**Definition 2.1.** A probability measure  $P$  on the borel  $\sigma$ -field of  $\Omega$  is a Gibbs measure with potential  $\Phi$  if, for all finite subset  $\Lambda$  of  $\mathbb{Z}^D$ ,

$$P(d\omega_\Lambda / {}_\Lambda \omega = {}_\Lambda z) = Z_{\Lambda, z}^{-1} \exp \left\{ \sum_{V: V \cap \Lambda \neq \emptyset} \Phi_V(\omega_\Lambda \vee z) \right\} \prod_{i \in \Lambda} \alpha(d\omega_i)$$

for  $\mathbb{P}$ -a.e.  $z$  in  $\Omega$ , where  $Z_{\Lambda,z}$  is the normalizing constant

$$Z_{\Lambda,z} = E^{\alpha^{\otimes \mathbb{Z}^D}} \exp \left\{ \sum_{V:V \cap \Lambda \neq \emptyset} \Phi_V(\omega_\Lambda \vee z) \right\}.$$

We will denote by  $\mathbb{G}$  the set of Gibbs measures with potential  $\Phi$ .

Throughout this paper, we will consider a sequence of increasing finite rectangles  $\Lambda_n$  in  $\mathbb{Z}^D$  such that  $\Lambda_n \xrightarrow{n \rightarrow \infty} \mathbb{Z}^D$ . We will assume that the cardinality  $|\Lambda_n|$  of  $\Lambda_n$  is such that  $|\Lambda_n| \sim an^D$  for some  $a \in ]0, \infty[$ . For all  $\omega \in \Omega$  and all finite rectangle  $\Lambda$ , we define  $\tilde{\omega}(\omega, \Lambda) \in \Omega$  as the periodic extension of the restriction  $\omega_\Lambda$ . Then, given  $\omega$ , the empirical field based on the observation of  $\omega$  on  $i + \Lambda_n$  is the stationary random field

$$R_{i,n,\omega} = \frac{1}{|\Lambda_n|} \sum_{j \in \Lambda_n} \delta_{\theta_j \tilde{\omega}(\omega, i + \Lambda_n)} \tag{2.2}$$

with  $|\Lambda|$  the cardinality of  $\Lambda$ . When  $i = 0$ , we will write  $R_{n,\omega}$  for  $R_{0,n,\omega}$  since it coincides with the usual (periodized) empirical field.

A major ingredient needed in this paper is the large deviation principle for Gibbs measures obtained by Comets (1986), Föllmer and Orey (1988), Olla (1988). A quick inspection of the proof reveals that these estimates are uniform in the following sense:

**Theorem 2.2.** *There exists an affine function  $I : \mathbb{P}_s(\Omega) \rightarrow [0, +\infty]$  (called the rate function) such that*

- 1) *the level set  $\Gamma_t = \{R \in \mathbb{P}_s(\Omega); I(R) \leq t\}$  is compact for all  $t \in [0, +\infty[$ ; in particular,  $I$  is lower semi-continuous.*
- 2) *for all  $P \in \mathbb{G}$ , and all borel set  $A$  in  $\mathbb{P}_s(\Omega)$ ,*

$$\begin{aligned} - \inf \{I(R); R \in \mathring{A}\} &\leq \liminf_{n \rightarrow \infty} \inf_{z \in \Omega} \frac{1}{|\Lambda_n|} \text{Log } P\{R_{n,\omega} \in A / \Lambda_n \omega = \Lambda_n z\} \\ &\leq \limsup_{n \rightarrow \infty} \sup_{z \in \Omega} \frac{1}{|\Lambda_n|} \text{Log } P\{R_{n,\omega} \in A / \Lambda_n \omega = \Lambda_n z\} \\ &\leq - \inf \{I(R); R \in \bar{A}\} \end{aligned}$$

( $\mathring{A}$ , resp.  $\bar{A}$ , is the interior, resp. closure, of  $A$  in  $\mathbb{P}_s(\Omega)$  for the weak topology).

In addition, the function  $I$  depends only on  $\Phi$  and  $\alpha$ ,  $I(R)$  is equal to the specific entropy of  $R$  relative to any  $P \in \mathbb{G}$ ; the set  $\Gamma_0$  coincides with the set  $\mathbb{G}_s$  of stationary Gibbs measures.

The explicit form of the rate function  $I$  is given in the above references; it will not be needed here. Recently, Georgii (1992) proved the theorem for a stronger topology. Note that these estimates are uniform with respect to the configuration outside  $\Lambda$ ; uniformity in these estimates has been already mentioned in Deuschel and Stroock's book (1989), p. 289.

### III. Erdős-Rényi Type Laws

For any integer  $l \leq n$ , we define the set of empirical fields based on moving windows in  $\Lambda_n$  which are translates of  $\Lambda_l$ ,

$$\Delta_{n,l}(\omega) = \{R_{i,l,\omega}; i \in \mathbb{Z}^D, i + \Lambda_l \subset \Lambda_n\}. \tag{3.1}$$

In the next two sections, we are interested in the asymptotics of  $\Delta_{n,l(n)}(\omega)$  for  $P$ -a.e.  $\omega$  and for some sequences  $l(n) \rightarrow \infty$ . Recall the Hausdorff distance between closed sets  $\Delta$  and  $\Gamma$  in  $\mathbb{P}_s(\Omega)$ :

$$\text{dist}(\Delta, \Gamma) = \inf\{\delta > 0; \Delta \subset \Gamma^\delta \text{ and } \Gamma \subset \Delta^\delta\} \tag{3.2}$$

with the usual notation  $A^\delta$  for the open  $\delta$ -neighborhood of  $A$  in the metric  $d$ .

**Theorem 3.1.** *Let  $P \in \mathbb{G}$ ,  $t > 0$  and  $l(n)$  be a non-decreasing sequence of integers such that  $|\Lambda_{l(n)}| \sim t^{-1} \text{Log} |\Lambda_n|$ . Then*

$$\lim_{n \rightarrow \infty} \text{dist}(\Delta_{n,l(n)}(\omega); \Gamma_t) = 0, \quad P\text{-a.s.}$$

with  $\Gamma_t = \{R \in \mathbb{P}_s(\Omega); I(R) \leq t\}$ .

Let us point out that the theorem is rather general.

*Remark 1.* The proof of Theorem 3.1, given below, extends to other distributions  $P$  than Gibbs measures.

a) Uniform large deviations. In what follows it is enough to assume that for some rate function  $I$  with compact level sets, such that for all  $Q$  with  $0 < I(Q) < +\infty$  and all neighborhood  $N$  of  $Q$  there exists  $R \in N$  such that  $I(R) < I(Q)$ , and to assume

$$\begin{aligned} -\inf\{I(R); R \in \bar{A}\} &\leq \liminf_{n \rightarrow \infty} \inf_{z \in \Omega, i \in \mathbb{Z}^D} \frac{1}{|\Lambda_n|} \text{Log} P\{R_{i,n,\omega} \in A /_{i+\Lambda_n} \omega =_{i+\Lambda_n} z\} \\ &\leq \limsup_{n \rightarrow \infty} \sup_{i \in \mathbb{Z}^D} \frac{1}{|\Lambda_n|} \text{Log} P\{R_{i,n,\omega} \in A\} \\ &\leq -\inf\{I(R); R \in \bar{A}\} \end{aligned}$$

for all borel set  $A$  and some regular version of the conditional probabilities. Moreover, in dimension  $D = 1$ , it is clearly sufficient to assume this uniformity property on the conditional distributions “given the past”; therefore, the theorem applies in particular to Markov chains with the uniform ergodicity property denoted by  $(U)$  given in chapter IV of Deuschel and Stroock (1989).

b) Hypermixing processes (see Deuschel and Stroock (1989), §5.4). One can replace the condition of uniformity with respect to  $z$ , with the following mixing-type condition:  $\exists l \geq 1, \alpha > 1, \gamma > 0$  such that for all  $n \geq 1$  and all Borel sets  $A_1, \dots, A_n$  in  $\Omega$  which are  $l$ -measurably separated,

$$P\left(\bigcap_{k=1}^n A_k\right) \leq \gamma \left[\prod_{k=1}^n P(A_k)\right]^{1/\alpha}.$$

Minor changes are required in the proof of Lemma 3.3 only, and the details are left to the reader. In particular, Theorem 3.1 will apply to time-discrete, stationary hypermixing processes.

We start with two lemmas:

**Lemma 3.2.** *Let  $l(n)$  be a non-decreasing sequence of integers tending to infinity, and*

$$t_n = |A_{l(n)}|^{-1} \text{Log } |A_n|, \quad t = \limsup_{n \rightarrow \infty} t_n.$$

Then, for all  $\delta > 0$ ,

$$P \left[ \limsup_{n \rightarrow \infty} \left\{ \omega; \max_{R \in \Delta_{n,l(n)}(\omega)} d(R; \Gamma_t) \geq \delta \right\} \right] = 0.$$

Define

$$\begin{aligned} A_{n,\delta} &= \{ \omega; \max [d(R; \Gamma_t); R \in \Delta_{n,l(n)}(\omega)] \geq \delta \} \\ &= \bigcup_{i: i+A_{l(n)} \subset A_n} \{ \omega; d(R_{i,l(n),\omega}; \Gamma_t) \geq \delta \}. \end{aligned}$$

Since  $I$  is lower semi-continuous with compact level sets, we have  $\xi = \inf \{ I(R); d(R; \Gamma_t) \geq \delta \} - t > 0$ .  $P$  being a Gibbs measure, it has translation invariant conditional distributions; the upper bound in Theorem (2.2, 2) implies that there exists some  $\alpha_l$  with  $\lim_{l \rightarrow \infty} \alpha_l = 0$  and

$$\sup_i P \{ d(R_{i,l,\omega}; \Gamma_t) \geq \delta \} \leq \exp[-|A_l|(t + \xi - \alpha_l)].$$

Let  $m(l)$  be the largest integer  $m$  such that  $l(m) = l$ . We have for all  $l$

$$\begin{aligned} P(A_{m(l),\delta}) &\leq \sum_{i: i+A_l \subset A_{m(l)}} P \{ d(R_{i,l,\omega}; \Gamma_t) \geq \delta \} \\ &\leq |A_{m(l)}| \exp[-|A_l|(t + \xi - \alpha_l)] \\ &= \exp[-|A_l|(\xi - \alpha'_l)] \end{aligned}$$

with some  $\alpha'_l$  tending to zero. Here, we have used the definitions of  $t_{m(l)}$  and  $t$ . Therefore,  $\sum_l P(A_{m(l),\delta}) < \infty$  and  $P(\limsup_{l \rightarrow \infty} A_{m(l),\delta}) = 0$  from Borel-Cantelli lemma. Note that, in addition, the size of the moving window  $l(n)$  is constant and equal to  $l(m(l))$  for all  $n \in ]m(l-1), m(l)]$ . Hence  $A_{n,\delta} \subset A_{m(l),\delta}$  for such  $n$ 's, the set  $\limsup_{n \rightarrow \infty} A_{n,\delta}$  coincides with  $\limsup_{l \rightarrow \infty} A_{m(l),\delta}$  and has probability zero.  $\square$

**Lemma 3.3.** *Let  $a_n$  be a sequence with  $\lim_{n \rightarrow \infty} a_n |A_n|^{-\alpha} = \infty$  for some  $\alpha > 0$ , and let  $B$  be a borel subset of  $\mathbb{P}_s(\Omega)$  such that*

$$\liminf_{n \rightarrow \infty} a_n^{-1} \text{Log } \inf_{z \in \Omega} P \{ R_{n,\omega} \in B / A_n \omega = A_n z \} > -1.$$

Let  $l(n)$  be a sequence of integers such that  $a_{l(n)} \sim \text{Log } |A_n|$ , and  $A_n = \{ \omega; \exists i : i + A_{l(n)} \subset A_n, R_{i,l(n)\omega} \in B \}$ . Then,

$$P \left( \liminf_{n \rightarrow \infty} A_n \right) = 1.$$

Pick in  $\mathbb{Z}^D$  the coarser lattice such that the translates of the rectangle  $A_{l(n)}$  by the elements of the lattice form a partition of  $\mathbb{Z}^D$ . Denote by  $L_n$  the set of those elements  $i$  such that the translates  $i + A_{l(n)}$  are contained in  $\Lambda_n$ . Then,  $|L_n| \sim |\Lambda_n|/|A_{l(n)}|$  and the sets  $i + A_{l(n)}$ ,  $i \in L_n$ , are not intersecting. Let  $j \in L_n$ ; we have

$$\begin{aligned}
 A_n^c &= \bigcap_{i:i+A_{l(n)} \subset \Lambda_n} \{R_{i,l(n),\omega} \notin B\} \subset \bigcap_{i \in L_n} \{R_{i,l(n),\omega} \notin B\}, \\
 P(A_n^c) &\leq E^P \left[ \prod_{i \in L_n: i \neq j} 1_{R_{i,l(n),\omega} \notin B} \times P\{R_{j,l(n),\omega} \notin B /_{j+A_{l(n)}} \omega\} \right] \\
 &\leq \sup_{z \in \Omega} P\{R_{l(n),\omega} \notin B /_{\Lambda_{l(n)}} \omega = \Lambda_{l(n)} z\} \times P \left[ \bigcap_{i \in L_n: i \neq j} \{R_{i,l(n),\omega} \notin B\} \right] \\
 &\leq \left[ \sup_{z \in \Omega} P\{R_{l(n),\omega} \notin B /_{\Lambda_{l(n)}} \omega = \Lambda_{l(n)} z\} \right]^{|L_n|} \\
 &= \left[ 1 - \inf_{z \in \Omega} P\{R_{l(n),\omega} \in B /_{\Lambda_{l(n)}} \omega = \Lambda_{l(n)} z\} \right]^{|L_n|}, \tag{3.3}
 \end{aligned}$$

using the translation invariance of the conditional distributions under  $P$ . By assumption there exists some  $\xi > 0$  and  $\alpha_n$  tending to zero such that

$$\begin{aligned}
 \text{Log } P(A_n^c) &\leq |L_n| \log[1 - \exp\{-a_{l(n)}(1 - \xi - \alpha_n)\}] \\
 &\leq -|L_n| \exp -a_{l(n)}(1 - \xi - \alpha_n). \tag{3.4}
 \end{aligned}$$

The assumption on  $a_n$  implies that  $l(n) = o(\text{Log}^{1/\alpha} |A_n|)$ , and the bound in (3.4) is less than  $-|A_n|^{\xi/2}$  for large enough  $n$ . Therefore  $\sum_n P(A_n^c) < \infty$ , and the conclusion follows.  $\square$

*Remark 2.* Let  $\nu$  be a fixed positive integer, and let

$$A_{n,\nu} = \{\omega; |\Delta_{n,l(n)}(\omega) \cap B| \geq \nu\}.$$

Then, Lemma 3.3 holds with  $A_{n,\nu}$  instead of  $A_n$  (note that  $A_{n,1}$  is equal to the set  $A_n$  defined in the lemma).

Indeed, a straightforward computation yields the following bound for  $P(A_{n,\nu}^c)$  corresponding to (3.4):

$$P(A_{n,\nu}^c) \leq \sum_{k=0}^{\nu-1} \binom{|A_n|}{k} e^{(|L_n| - k) \exp -a_{l(n)}(1 - \xi - \alpha_n)},$$

which is once again the general term of a convergent serie.

We now prove Theorem 3.1: the statement is equivalent to

$$\forall \delta > 0, \limsup_{n \rightarrow \infty} \sup_{R \in \Delta_{n,l(n)}(\omega)} d(R; \Gamma_t) \leq \delta \quad P\text{-a.s.}, \tag{3.5 a}$$

and

$$\forall \delta > 0, \limsup_{n \rightarrow \infty} \sup_{Q \in \Gamma_t} d(Q; \Delta_{n,l(n)}(\omega)) \leq \delta \quad P\text{-a.s.} \tag{3.5 b}$$

The Lemma 3.2 is exactly (3.5 a). Let us prove (3.5 b).

The level set  $\Gamma_t$  being compact, we can choose a finite number  $q$  of points  $P_1, \dots, P_q$  in  $\Gamma_t$  such that the balls with center  $P_k$  and radius  $\delta/4$ ,  $k = 1, \dots, q$ , cover  $\Gamma_t$ . Since  $I$  is a convex function achieving the value 0 and since  $t$  is positive, there exists some  $Q_k$  in each such ball with  $I(Q_k) < t$ . Then, the balls with center  $Q_k$  and radius  $\delta/2$ ,  $k = 1, \dots, q$ , cover  $\Gamma_t$  and (3.5 b) will follow from

$$\forall k \leq q, \forall \delta > 0, \limsup_{n \rightarrow \infty} d(Q_k; \Delta_{n,l(n)}(\omega)) < \delta/2 \quad P\text{-a.s.} \quad (3.6)$$

But (3.6) is a consequence of Lemma (3.3): taking  $a_n = t|A_n|$  and  $B = \{R \in \mathbb{P}_s(\Omega); d(R, Q_k) < \delta/2\}$ , then  $A_n$  is the set under consideration in (3.6) and we see from the lower bound in Theorem 2.2, 2), that the lemma applies.  $\square$

*Remark 3.* Let  $\Lambda$  be a given finite subset of  $\mathbb{Z}^D$ , and

$$\Delta_{n,l(n),\Lambda}(\omega) = \{R_{i,l(n),\omega}; i \in \mathbb{Z}^D : i + \Lambda_{l(n)} + \Lambda \subset A_n\}.$$

Then, it is clear from the proofs that the Lemma 3.2 holds with  $\Delta_{n,l(n),\Lambda}$  instead of  $\Delta_{n,l(n)}$ , and that Lemma 3.3 holds with  $A'_n = \{\exists i : i + \Lambda_{l(n)} + \Lambda \subset A_n, R_{i,l(n),\omega} \in B\}$  instead of  $A_n$ . Therefore, under the assumptions of Theorem 3.1, we have also

$$\lim_{n \rightarrow \infty} \text{dist}(\Delta_{n,l(n),\Lambda}(\omega); \Gamma_t) = 0, \quad P\text{-a.s.}$$

Now, we formulate the Erdős-Rényi law in the classical fashion of (1.1).

Let  $\Lambda$  be a finite subset of  $\mathbb{Z}^D$ , and  $f : \Omega_0^\Lambda \rightarrow \mathbb{R}$  be a continuous and bounded function. Define

$$I_f(x) = \inf \{I(Q); E^Q f = x\}, \quad x \in \mathbb{R}. \quad (3.6 b)$$

Then, according to the contraction principle (Deuschel and Stroock, p. 37),  $I_f$  is the rate function governing the large deviations of  $\frac{1}{|A_n|} \sum_{i \in A_n} f \circ \theta_i(\omega)$  under  $P$ , and  $I_f$  is convex and lower semi-continuous. The domain of  $I_f$ ,  $\text{Dom } I_f = \{x \in \mathbb{R}; I_f(x) < \infty\}$ , is a bounded interval.

**Corollary 3.4.** *Let  $P \in \mathbb{G}$ ,  $t > 0$  and  $l(n)$  be as in Theorem 3.1. Let  $m \in I_f^{-1}\{0\}$ . Define the random variable*

$$X_{n,l(n)}^+(\omega) = \max \left\{ \frac{1}{|A_{l(n)}|} \sum_{j \in i + \Lambda_{l(n)}} f(\omega_{j+\Lambda}); i + \Lambda_{l(n)} + \Lambda \subset A_n \right\}.$$

Then,

$$X_{n,l(n)}^+ \xrightarrow[n \rightarrow \infty]{} x^+(t) \quad P\text{-a.s.},$$

with  $x^+(t) = \sup\{x; x \geq m \text{ and } I_f(x) \leq t\}$ . In particular,  $I_f(x^+(t)) \leq t$  and  $x^+(t)$  is the unique root of the equation  $I_f(x) = t$  in  $]m, +\infty[$  if (and only if)  $t$  is achieved by  $I_f$  on  $]m, +\infty[$ .

Similarly, the random variable

$$X_{n,l(n)}^-(\omega) = \min \left\{ \frac{1}{|A_{l(n)}|} \sum_{j \in i + \Lambda_{l(n)}} f(\omega_{j+\Lambda}); i + \Lambda_{l(n)} + \Lambda \subset A_n \right\}$$

converges  $P$ -a.s. to  $x^-(t) = \inf\{x; x \leq m, I_f(x) \leq t\}$ .  $I_f(x^-(t)) \leq t$ , and  $x^-(t)$  is the unique solution on  $] - \infty, m[$  of  $I_f(x) = t$  if there exists such a solution.

*Remark.*  $I_f(x^+(t)) = t$  will hold if  $\alpha^{\otimes \Lambda} \{ \omega_\Lambda; f(\omega_\Lambda) > x^+(t) \} > 0$ , as it follows from elementary computations on the Laplace transform of  $\sum_{\iota \in \Lambda_n} f(\omega_{\iota+\Lambda})$ .

Let us make first a general observation. If  $g$  is a continuous function from a metric space  $(E, d)$  to another space  $(E', d')$ , and if  $A_n \subset E$  are such that  $\lim_{n \rightarrow \infty} \text{dist}(A_n, A) = 0$  in the Hausdorff distance with some compact  $A \subset E$ , then  $\lim_{n \rightarrow \infty} \text{dist}'(g(A_n), g(A)) = 0$ .

We set  $g(R) = \mathbb{E}^R f$ ,  $A_n = \Delta_{n, l(n), \Lambda}$  and  $A = \Gamma_t$ . Combining this observation with Remark 3 about Theorem 3.1, we obtain  $P$ -a.s.,

$$\lim_{n \rightarrow \infty} \text{dist}' \left( \left\{ \int f dR_{i, l(n), \omega}; i + \Lambda_{l(n)+\Lambda} \subset A_n \right\}; \{ \mathbb{E}^Q f; Q \in \Gamma_t \} \right) = 0 \quad (3.7)$$

with  $\text{dist}'$  the Hausdorff distance on closed subsets of  $\mathbb{R}$ . On the other hand,  $\int f dR_{i, l(n), \omega} - \frac{1}{|\Lambda_{l(n)}|} \sum_{j \in i+\Lambda_{l(n)}} f(\omega_{j+\Lambda})$  converges to zero with  $n$ , uniformly in  $i \in \mathbb{Z}^D$  and  $\omega \in \Omega$ ; hence, taking suprema of the sets in (3.7), we have  $\lim_{n \rightarrow \infty} X_{n, l(n)}^+ = y^+(t)$  with  $y^+(t) = \max \{ E^Q f; Q \in \Gamma_t \}$ . Checking that  $y^+(t) = x^+(t)$  is standard matter and left to the reader. The inequality  $I_f(x^+(t)) \leq t$  follows from the lower semi-continuity of  $I_f$ . Since  $I_f$  is non-negative, convex, and  $I_f(m) = 0$ ,  $I_f(x) = t$  has at most one solution  $x \in ]m, +\infty[$  for all  $t > 0$ .

Hence, one half of the statements of the corollary are proved. The other ones follow when changing  $f$  into  $-f$ .  $\square$

*Remark 2'.* For all given integer  $\nu$ , the  $\nu^{\text{th}}$  maximum of the moving averages of  $f$  converges a.s. to  $x^+(t)$  too, according to Remark 2.

#### IV. Tracking the Set of Stationary Gibbs Measures

We will say that a *first order phase transition* occurs if the set  $\mathbb{G}_e$  of ergodic Gibbs measures contains more than one element. We are interest in this case below. Then the expected value of some statistics for the random field shows a discontinuity under small perturbations of the interaction potential.

Recall that  $\Gamma_0 = \mathbb{G}_s$ . In this section, we will obtain results like Theorem 3.1, but involving  $\Gamma_0$  instead of  $\Gamma_t$  with  $t > 0$ . We will consider ratios  $t_n = |\Lambda_{l(n)}|^{-1} \text{Log} |\Lambda_n|$  tending to 0, but slowly enough in view of Lemma 3.3. In order to choose the rate of convergence of  $t_n$ , we introduce first the following function:

$$\varepsilon_\Phi(\Lambda) = \left\| \frac{1}{|\Lambda|} \sum_{V \cap \Lambda \neq \emptyset, V \not\subset \Lambda} \Phi_V \right\|_\infty \quad (4.1)$$

for finite  $\Lambda \subset \mathbb{Z}^D$ . Here are some elementary properties:

**Lemma 4.1.** a) *There exists  $\varepsilon : \mathbb{R}^+ \rightarrow \mathbb{R}^+$  with  $\lim_{u \rightarrow \infty} \varepsilon(u) = 0$  such that*

$$\varepsilon_\Phi(\Lambda) \leq \varepsilon \left( \frac{|\partial \Lambda|}{|\Lambda|} \right), \quad \text{where } \partial \Lambda \text{ denotes the boundary of } \Lambda.$$

b) If the potential  $\Phi$  has finite range  $r_\Phi$ ,

$$\varepsilon_\Phi(\Lambda) \leq r_\Phi \|\Phi\| \frac{|\partial\Lambda|}{|\Lambda|}.$$

*Proof.* For  $r \geq 1$ , let  $\delta(r) = \sum \|\Phi_V\|_\infty$ , where the sum ranges over all  $V$  containing 0, with diameter larger than  $r$ . From (2.1),  $\delta(r)$  tends to 0 as  $r$  tends to infinity. Splitting the sum in (4.1), into  $V$ 's with diameter larger than  $r$  or not, we obtain for all  $r$ ,

$$\varepsilon_\Phi(\Lambda) \leq \delta(r) + r \frac{|\partial\Lambda|}{|\Lambda|} \|\Phi\|.$$

Optimizing the bound with respect to  $r$ , yields the desired function  $\varepsilon$ . The statement for finite range potential follows from the previous inequality with  $r = r_\Phi$ .  $\square$

In order to use Lemma 3.3, we give rough (but general) estimates for large deviations in the set of Gibbs measures. Recall that  $\mathbb{G}_e$  is the set of ergodic Gibbs measures.

**Lemma 4.2.** a)  $\forall P \in \mathbb{G}, \forall Q \in \mathbb{G}_e, \forall \delta > 0$ ,

$$\liminf_{n \rightarrow \infty} \frac{1}{2|\Lambda_n| \varepsilon_\Phi(\Lambda_n)} \text{Log} \inf_{z \in \Omega} P\{d(R_{n,\omega}; Q) < \delta / \Lambda_n \omega = \Lambda_n z\} \geq -1.$$

b) Assume that  $\varepsilon_\Phi(\Lambda) \leq a \left(\frac{|\partial\Lambda|}{|\Lambda|}\right)^\alpha$  for some  $\alpha \in ]0, 1]$  and  $a > 0$ . Then,  $\forall P \in \mathbb{G}, \forall \delta > 0, \forall Q \in \mathbb{G}_s$ ,

$$\liminf_{n \rightarrow \infty} \frac{1}{|\partial\Lambda_n|^\alpha |\Lambda_n|^{1-\alpha}} \text{Log} \inf_{z \in \Omega} P\{d(R_{n,\omega}; Q) < \delta / \Lambda_n \omega = \Lambda_n z\} > -\infty.$$

If, in addition,  $\mathbb{G}_e$  is finite, the previous limit is not less than  $-2a|\mathbb{G}_e|$ .

Let  $A_n = \{\omega; d(R_{n,\omega}; Q) < \delta\}$ .

$$\begin{aligned} \text{a)} \quad Q(A_n) &= \mathbb{E}^Q Q(A_n / \Lambda_n \omega) \\ &= \mathbb{E}^Q P(A_n / \Lambda_n \omega) \\ &\leq \inf_{z \in \Omega} P(A_n / \Lambda_n \omega = \Lambda_n z) \times \exp[2|\Lambda_n| \varepsilon_\Phi(\Lambda_n)]. \end{aligned} \tag{4.3}$$

If  $Q$  is ergodic,  $Q(A_n)$  goes to one by the ergodic theorem, and the statement is proved.

b) It is well known (Georgii (1988), Theorem 7.26) that  $Q$  is the mixture of ergodic Gibbs measures. Hence one can find a convex combination of finitely many ergodic Gibbs measures in any given neighborhood of  $Q$ , and therefore it is enough to prove the statement for  $Q = \sum_{k=1}^q \lambda_k Q_k$  with  $\lambda_k > 0, \sum_{k=1}^q \lambda_k = 1,$

$Q_k \in \mathbb{G}_e$ . Let  $(\Lambda_n^k)_{k \leq q}$  be a partition of  $\Lambda_n$  into parallel rectangular slices with  $|\Lambda_n^k| \sim \lambda_k |\Lambda_n|$  and  $|\partial\Lambda_n^k| \sim |\partial\Lambda_n| (D - 1 + \lambda_k) / D$ . Then,  $R_{n,\omega}$  is close to the convex combination of the empirical fields  $R_{n,\omega}^k$  based on  $\Lambda_n^k$ : more precisely, for large  $n, d\left(\sum_{k=1}^q \lambda_k R_{n,\omega}^k; R_{n,\omega}\right) < \delta/2$ .

Let us recall a convexity property of the Prohorov metric  $d : d(\lambda\mu + (1 - \lambda)\mu'; \lambda\nu + (1 - \lambda)\nu') \leq \max[d(\mu, \mu'), d(\nu, \nu')]$  for  $\lambda \in [0, 1]$  and all measures  $\mu, \mu', \nu, \nu'$ . Combining this with the previous inequality, we obtain for large  $n$

$$\bigcap_{k=1}^q \{d(R_{n,\omega}^k; Q_k) < \delta/2\} \subset \{d(R_{n,\omega}, Q) < \delta\}. \tag{4.4}$$

Repeating the argument used in (3.3), we have

$$\begin{aligned} &P \left[ \bigcap_{k=1}^q \{d(R_{n,\omega}^k; Q_k) < \delta/2\} / \Lambda_n \omega = \Lambda_n z \right] \\ &\geq \prod_{k=1}^q \inf_{y \in \Omega} P [d(R_{n,\omega}^k; Q_k) < \delta/2 / \Lambda_n^k \omega = \Lambda_n^k y] \\ &\geq \prod_{k=1}^q Q_k \{d(R_{n,\omega}^k; Q_k) < \delta/2\} \exp -2|\Lambda_n^k| \varepsilon_{\Phi}(\Lambda_n^k) \end{aligned}$$

using (4.3). Together with (4.4) and the ergodic theorem, this inequality implies that

$$\liminf_{n \rightarrow \infty} \left[ 2 \sum_{k=1}^q |\Lambda_n^k| \varepsilon_{\Phi}(\Lambda_n^k) \right]^{-1} \text{Log} \inf_{z \in \Omega} P \{d(R_{n,\omega}; Q) < \delta / \Lambda_n \omega = \Lambda_n z\} \geq -1.$$

If  $\varepsilon_{\Phi}(\Lambda) \leq a(|\partial\Lambda|/|\Lambda|)^{\alpha}$ , then

$$\begin{aligned} \sum_{k=1}^q |\Lambda_n^k| \varepsilon_{\Phi}(\Lambda_n^k) &\leq a|\Lambda_n|^{1-\alpha} |\partial\Lambda_n|^{\alpha} \sum_{k=1}^q \lambda_k^{1-\alpha} \left(1 - \frac{1 - \lambda_k}{D}\right)^{\alpha} \\ &\leq qa|\Lambda_n|^{1-\alpha} |\partial\Lambda_n|^{\alpha}, \end{aligned}$$

which ends the proof.  $\square$

In the two next theorems,  $P \in \mathbb{G}$ , and  $l(n)$  is a non-decreasing sequence of integers with

$$\lim_{n \rightarrow \infty} \frac{\text{Log} |\Lambda_n|}{|\Lambda_{l(n)}|} = 0. \tag{4.5}$$

**Theorem 4.3.** *In addition to (4.5), assume that  $\varepsilon_{\Phi}(\Lambda) \leq a \left(\frac{|\partial\Lambda|}{|\Lambda|}\right)^{\alpha}$ ,  $a > 0$ ,  $\alpha \in ]0, 1[$ , and that one of the two following conditions is fulfilled:*

- i)  $\lim_{n \rightarrow \infty} \frac{\text{Log} |\Lambda_n|}{|\partial\Lambda_{l(n)}|^{\alpha} |\Lambda_{l(n)}|^{1-\alpha}} = \infty$ ,
- ii)  $|\mathbb{G}_e| < \infty$  and  $\lim_{n \rightarrow \infty} \frac{\text{Log} |\Lambda_n|}{|\partial\Lambda_{l(n)}|^{\alpha} |\Lambda_{l(n)}|^{1-\alpha}} > 2a|\mathbb{G}_e|$ .

Then,  $\lim_{n \rightarrow \infty} \text{dist}(\Delta_{n,l(n)}(\omega); \mathbb{G}_s) = 0$   $P$ -a.s..

*Remark 4.* For finite range potential and cubic boxes  $\Lambda_n = [-n, n]^D$ , the above assumption i) is equivalent to  $l(n) = o(\text{Log } n)^{1/(D-1)}$ , though the next theorem only requires  $l(n) \sim b(\text{Log } n)^{1/(D-1)}$ , with  $b < (2^D D r_{\Phi} \|\Phi\|)^{-1/(D-1)}$ .

**Theorem 4.4.** *Let  $\xi_{n,l(n)}(\omega)$  be the convex hull of  $\Delta_{n,l(n)}(\omega)$  in  $\mathbb{P}_s(\Omega)$ . Assume (4.5), and*

$$\lim_{n \rightarrow \infty} \frac{\text{Log } |A_n|}{2|A_{l(n)}| \varepsilon_{\Phi}(A_{l(n)})} > 1. \tag{4.6}$$

*Then  $\lim_{n \rightarrow \infty} \text{dist}(\xi_{n,l(n)}(\omega); \mathbb{G}_s) = 0$  P-a.s..*

*Proof of Theorem 4.3.* We follow the lines of proof of Theorem 3.1, setting  $t = 0$ . Under the condition (4.5), we may use Lemma 3.2 and we obtain (3.5 a). We now prove (3.6), which in turn proves (3.5 b). Let us denote again by  $m(l)$  the largest integer  $m$  with  $l(m) = 1$ , and define  $a_l = \text{Log } |A_{m(l)}|$ . A straightforward computation using Lemma 4.2 b), shows that  $a_n$  satisfies the assumption in Lemma 3.3 with  $B = \{R; d(R; Q_k) < \delta/2\}$ , in both cases i) and ii). Hence (3.6) holds, and the theorem follows from the equality  $\Gamma_0 = \mathbb{G}_s$ .  $\square$

*Proof of Theorem 4.4.* Together with the convexity property of the Prohorov metric  $d$  and of the set  $\mathbb{G}_s$ , (4.5) implies that

$$\lim_{n \rightarrow \infty} \max \{d(R, \mathbb{G}_s); R \in \xi_{n,l(n)}(\omega)\} = 0 \text{ P-a.s..}$$

It remains to prove the analogue of (3.6),  $\forall Q \in \mathbb{G}_s, \forall \delta > 0$ ,

$$\limsup_{n \rightarrow \infty} d(Q; \xi_{n,l(n)}(\omega)) < \delta \text{ P-a.s..}$$

It is sufficient to prove the statement for a convex combination  $Q = \sum_{k=1}^q \lambda_k Q_k$ ,  $Q_k \in \mathbb{G}_e$ .

By definition of  $\xi$ , and from the convexity property of  $d$ , we check that it is still sufficient to prove

$$\forall Q \in \mathbb{G}_e, \forall \delta > 0, \limsup_{n \rightarrow \infty} d(Q; \Delta_{n,l(n)}(\omega)) < \delta \text{ P-a.s..}$$

But similar to the proof of Theorem 4.3, this follows from Lemmas 3.3 and 4.2 a).  $\square$

**Corollary 4.5.** *Let  $P \in \mathbb{G}$ , and  $f, X_{n,l(n)}^{\pm}$  be as in Corollary 3.4. Assume (4.5) and (4.6). Then,*

$$X_{n,l(n)}^+ \longrightarrow x^+(0)$$

and

$$X_{n,l(n)}^- \longrightarrow x^-(0) \text{ P-a.s.,}$$

with  $x^+(0) = \max \{E^Q f; Q \in \mathbb{G}_e\} = \max I_f^{-1}\{0\}$ , and  $x^-(0)$  the corresponding minimum.

The proof of the corollary is the same as for Corollary 3.4; indeed, taking the expectation of  $f$  and taking the convex hull are commuting operations on  $\mathbb{P}_s(\Omega)$ , and the convex hull of  $\mathbb{G}_e$  is  $\mathbb{G}_s$ .

Let us end this section with a remark. As mentioned in the introduction, (4.6) is a sufficient condition for the corollary, it yields only the right order of magnitude for  $l(n)$ , but not the right constants. Even if we would know, for ergodic  $P$ , the limit of

$$\frac{1}{|\partial A_n|} \text{Log } P \left\{ \frac{1}{|A_n|} \sum_{i \in A_n} f(\omega_i) \sim x \right\} \text{ for } x \in [x^-(0), x^+(0)],$$

we should need some extra property of exponential mixing for  $P$  in order to check that this limit yields the necessary and sufficient condition on  $l(n)$  for the corollary to hold.

### V. Estimating the Rate Function $I_f$

The convergence of the pointwise estimator  $X_{n,l(n)}^\pm$  to  $x^\pm(0)$  is very slow, as it can be observed on simulations. Analytical computations indicates that the difference is  $O(l(n)^{1/2})$  in general situations. Let us recall that we should take  $l(n) = O(\text{Log } n)^{1/(D-1)}$  for finite range interaction and cubic boxes (cf. Remark 4, §4) in order to investigate all Gibbs measures. Therefore we will make better use of a functional point of view, and estimate the rate function  $I_f$  itself to use more information from the data.  $I_f$  is equal to zero on the interval  $[x^-(0), x^+(0)]$ , and convex on  $\mathbb{R}$ . We will check on the simulation experiments of the next section, that the functional estimator is much more informative than the pointwise estimators.

In this section, we are given some function  $f$  as in Corollary (3.4), and we propose an estimator for  $I_f$ .

Let us extend  $X_{n,l}^+$  and  $X_{n,l}^-$  into functions on  $\mathbb{R}^+$ :

$$X_n^+(t) = \begin{cases} X_{n,l}^+ & \text{if } \frac{\text{Log } |A_n|}{|A_l|} \leq t < \frac{\text{Log } |A_n|}{|A_{l-1}|} \\ X_{n,n}^+ & \text{if } t < \frac{\text{Log } |A_n|}{|A_n|}. \end{cases} \tag{5.1}$$

$X_n^-$  is defined by the above expressions, changing  $+$ 's into  $-$ 's.

**Lemma 5.1.** *For all  $\delta \in ]0, 1]$ ,  $X_n^+(t)$  and  $X_n^-(t)$  converge uniformly on  $[\delta, 1/\delta]$  to  $x^+(t)$  and  $x^-(t)$  respectively,  $P$ -a.s.*

Let  $D$  be a dense, countable subset of  $[\delta, 1/\delta]$ . Corollary 3.4 implies that  $X_n^+$  converges pointwise on  $D$  to  $x^+$ ,  $P$ -a.s.. We will make use of the following sufficient Ascoli-type condition for relative compactness: if  $\phi_n : [\delta, 1/\delta] \rightarrow \mathbb{R}$  is a sequence of uniformly bounded, right-continuous and left limited functions and if

$$\forall \varepsilon > 0, \exists \eta > 0 \text{ and } n_0 : \sup_{n \geq n_0} \sup_{|t-s| < \eta} |\phi_n(t) - \phi_n(s)| \leq \varepsilon, \tag{5.2}$$

then the sequence is relatively compact in the uniform convergence topology and all limit points are continuous functions.

We check (5.2) below for  $X_n^+$ ; then pointwise convergence on  $D$ , continuity of the limit and density of  $D$  imply that  $x^+$  is the unique possible limit point, and so  $X_n^+$  converges uniformly to  $x^+$ ,  $P$ -a.s.

Let us check (5.2). Let  $\tau_{n,k} = |A_k|^{-1} \text{Log } |A_n|$ , and  $k(t)$  be such that  $\tau_{n,k(t)} \leq t < \tau_{n,k(t)-1}$ . we have

$$X_n^+(t) - X_n^+(s) = (\tau_{n,k(t)} - \tau_{n,k(s)}) \frac{X_{n,k(t)}^+ - X_{n,k(s)}^+}{\tau_{n,k(t)} - \tau_{n,k(s)}} \tag{5.3}$$

and  $\sup_{|t-s|<\eta} (\tau_{n,k(t)} - \tau_{n,k(s)}) \xrightarrow{n \rightarrow \infty} \eta$ , so it is enough to show that the ratio in (5.3)

is bounded; this ratio being a convex combination of ratios  $\frac{X_{n,l+1}^+ - X_{n,l}^+}{\tau_{n,l+1} - \tau_{n,l}}$  with coefficients  $\frac{\tau_{n,l+1} - \tau_{n,l}}{\tau_{n,k(t)} - \tau_{n,k(s)}}$ , we just need to prove that this last ratio is bounded.

Using the definition of  $X_{n,l}^+$  as a maximum, one obtains

$$\begin{aligned} |X_{n,l+1}^+ - X_{n,l}^+| &\leq \max_i \left| \frac{1}{|A_{l+1}|} \sum_{j \in i+A_{l+1}} f \circ \theta_j - \frac{1}{|A_l|} \sum_{j \in i+A_l} f \circ \theta_j \right| \\ &\leq \left[ \frac{1}{|A_l|} - \frac{1}{|A_{l+1}|} \right] |A_l| \|f\|_\infty + [|A_{l+1}| - |A_l|] \frac{1}{|A_{l+1}|} \|f\|_\infty \\ &= 2\|f\|_\infty (\tau_{n,l} - \tau_{n,l+1}) / \tau_{n,l} \\ &\leq (2\|f\|_\infty / \delta) (\tau_{n,l} - \tau_{n,l+1}) \end{aligned}$$

which implies boundedness of the ratios. Therefore (5.2) is proved, as well as the statement for  $X_n^+$ . Changing  $f$  into  $-f$ , one gets the one for  $X_n^-$ .  $\square$

Next, we consider some continuous non-decreasing piecewise linear upper-envelope of  $X_n^+$ : let  $\bar{X}_n^+ : \mathbb{R}^+ \rightarrow \mathbf{R}$  be the polygonal line with vertices at points  $t = \frac{\text{Log } |A_n|}{|A_k|}$ ,  $x = \max_{l \geq k} X_{n,l}^+$ ; ( $k = 1, \dots, n$ ), starting from  $(0, X_{n,n}^+)$  and constant for large  $t$ . We have  $\bar{X}_n^+(t) \geq X_n^+(t)$ , and we denote by  $T_n^+$  its inverse  $T_n^+(x) = \inf\{t > 0; \bar{X}_n^+(t) > x\}$ . Similarly, define  $\bar{X}_n^-$ , the continuous non-increasing piecewise linear lower-envelope of  $X_n^-$ , and  $T_n^-(x) = \sup\{t < 0; \bar{X}_n^-(x) < -t\}$ . Finally, remark that  $\bar{X}_n^+(0) = X_{n,n}^+ = X_{n,n}^- = \bar{X}_n^-(0)$ , and define the functional estimator

$$T_n(x) = \begin{cases} T_n^+(x), & \text{if } x > \bar{X}_n^+(0), \\ 0, & \text{if } x = \bar{X}_n^+(0), \\ T_n^-(x), & \text{if } x < \bar{X}_n^+(0). \end{cases}$$

**Theorem 5.2.** *For all  $a, b$  in the interior of the domain of  $I_f$ ,  $a < b$ ,  $T_n$  converges uniformly on  $[a, b]$  to  $I_f$ ,  $P$ -a.s.:*

Without loss of generality, we assume  $a < 0 < b$ . Let  $\varepsilon > 0$ . The function  $x^+$  is the inverse of  $I_f$  on  $[x^+(\varepsilon), b]$ , so  $x^+$  is a concave non-decreasing function with slopes bounded away from 0. Hence, choosing  $\eta > 0$  small enough, the  $\eta$ -neighborhood in the uniform norm of  $x^+$  on the interval  $[\varepsilon, (x^+)^{-1}(b)]$  has its symmetric (with respect to the bissector) contained in the  $\varepsilon$ -neighborhood of  $I_f$  on the interval  $[x^+(\varepsilon), b]$ .

In other respects, Lemma 5.1 shows that  $\bar{X}_n^+$  converges uniformly on  $[\delta, 1/\delta]$  to  $x^+$ , a.s.. Indeed,  $\sup_{[\delta, 1/\delta]} (X_n^+ - x^+) < \varepsilon$  implies that  $\sup_{[\delta, 1/\delta]} (\bar{X}_n^+ - x^+) < \varepsilon$  since

$x^+$  is concave, and we may take the absolute values inside the supremum since  $X_n^+ \leq \bar{X}_n^+$ . So we have  $P$ -a.s.,  $\sup_{x \in [x^+(\varepsilon), b]} |T_n^+(x) - I_f(x)| \leq \varepsilon$  for large  $n$ , and

similarly  $\sup_{x \in [a, x^-(\varepsilon)]} |T_n^-(x) - I_f(x)| \leq \varepsilon$  for large  $n$ .

Since  $T_n$  is non-decreasing (resp. non-increasing) on  $[\bar{X}_n^+(0), x^+(\varepsilon)]$  (resp.  $[x^-(\varepsilon), \bar{X}_n^-(0)]$ ), and since  $I_f(\cdot) \leq \varepsilon$  on  $[x^-(\varepsilon), x^+(\varepsilon)]$ , we have  $P$ -a.s.,  $\sup_{[a,b]} |T_n - I_f| \leq \varepsilon$  for large  $n$ , which is the claim.  $\square$

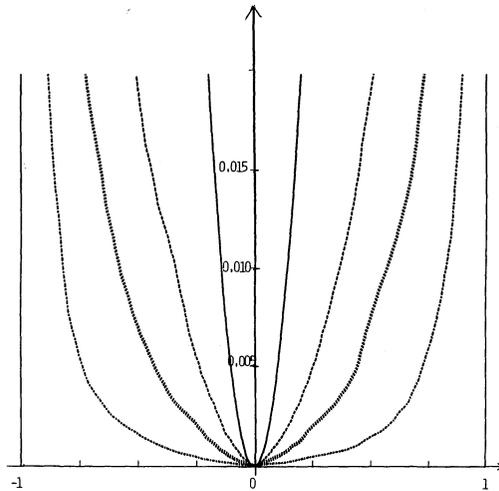
### VI. Numerical Results

We report here simulation experiments on the 2-dimensional nearest-neighbor Ising model on  $\Lambda_n = [-n, n]^2$ , whose hamiltonian is  $-H(\omega) = \beta \sum_{\|i-j\|_1=1} \omega_i \omega_j + h \sum \omega_i$ , with periodic boundary conditions.

Simulations were performed for different values of  $\beta, h$  and  $n$  (up to  $n = 800$ ), on Sun Sparc stations via the heat-bath algorithm with pre-computed transition probabilities. The function of interest was the magnetization,  $f(\omega) = \omega_0$ . The finite-size effects and periodic boundary condition made the “experimental critical temperature inverse” lower than the real value  $\beta_c = 0,4406868\dots$

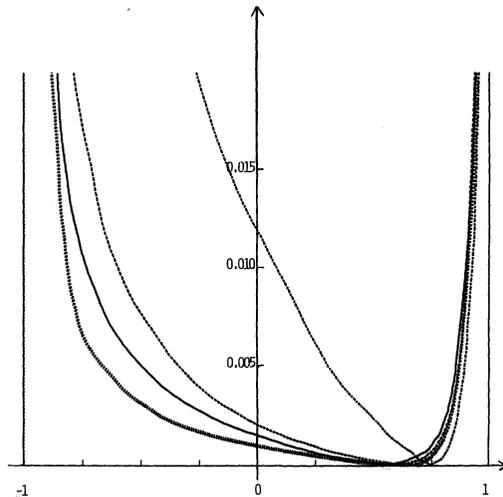
Trying an estimate  $x^\pm(0)$  using the point estimators  $X_{n,l}^\pm$  appeared to be unrealistic, due to slow convergence, to the existence of two different asymptotics according to Sect. III and IV, and to the choice of the scaling parameter in  $l(n)$ . On the other hand, the functional estimator gives encouraging descriptive results.

The empirical curves given in Fig. 1, 2 and 3 below were obtained by averaging the estimators  $T_n$  from ten samples in the same simulation but far apart in time one from another. Recall that they are estimates of the rate  $I_f$  of decay for large deviation probabilities.

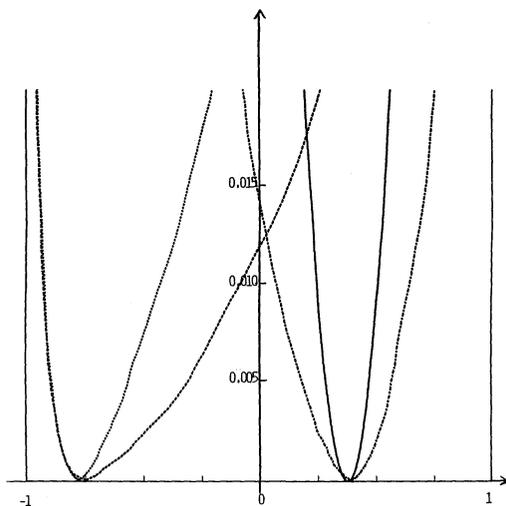


**Fig. 1.**  $h = 0$ , line curve:  $\beta = 0$ , exact function  $I_f$  (Cramer Transform); dashed curve:  $\beta = 0.35$ . dotted curve:  $\beta = 0.40$ ; dashed-dotted curve:  $\beta = 0.43$

For  $h = 0$ , the bottom of the graph of  $T_n$  becomes flatter as  $\beta$  increases up to  $\beta_c$ ; see Fig. 1. For larger  $\beta$ , spontaneous magnetization reflects in the total magnetization  $\bar{X}_n(0)$ , but  $T_n$  is composed of a “flat segment” from  $\bar{X}_n(0)$  to some value presumably close to the other spontaneous magnetization, and at each end of a much steeper,



**Fig. 2.**  $h = 0$ , line:  $\beta = 0.441$  and  $n = 200$ . dashed:  $\beta = 0.442$  and  $n = 200$ ; dotted:  $\beta = 0.442$  as before but  $n = 800$ ; dashed-dotted:  $\beta = 0.45$ ,  $n = 200$



**Fig. 3.** line:  $\beta = 0$ ,  $h = 0.4$ , exact function  $I_f$ ; dashed-dotted:  $\beta = 0.35$ ,  $h = 0.03$ ; dashed:  $\beta = 0.45$ ,  $h = 0$ ; dotted:  $\beta = 0.44$ ,  $h = 0.01$

strictly convex part; a drastic change in the curvature separates these three pieces, see Fig. 2.

Repeated experiments indicate that  $T_n$  is not much subject to fluctuations, for  $\beta$  smaller or slightly larger than the “experimental critical value”; for larger  $\beta$ , fluctuations from sample to sample become more important, but the shape itself is rather clearly the one described in the previous paragraph, independently of fluctuations. The estimator  $T_n$  improves slowly with  $n$  as predicted by the logarithmic rate of convergence (see Fig. 2).

The qualitative exchange of shape for  $T_n$  could be used for hypothesis testing, of hypothesis “ $x^-(0) = x^+(0)$ ”, comparing the minimal curvature of  $T_n$  close to

$\bar{X}_n(0)$ , to  $\frac{1}{2\hat{\sigma}_{k(n)}^2}$  with  $\hat{\sigma}_k^2$  some empirical estimator of  $\sigma_k^2 = \text{Var}\left(\frac{1}{\sqrt{|\Lambda_k|}} \sum_{i \in \Lambda_k} X_i\right)$  computed on the sample. Indeed, if  $\sigma^2 = \lim_{k \rightarrow \infty} \sigma_k^2$  is finite and  $x^-(0) = x^+(0)$ .  $I_f(x) = \frac{1}{2\sigma^2}(x - x^+(0))^2 + o(|x - x^+(0)|^2)$ , while if  $x^-(0) \neq x^+(0)$ ,  $I_f(x) = 0$  on  $[x^-(0), x^+(0)]$ . To determine the proper acceptance region for the test seems to be a difficult question. Note also that it is not clear how to discriminate from the last two experimental curves in Fig. 3 between the case of “large  $\beta$ ,  $h = 0$ ” and the case of “smaller  $\beta$ ,  $h \neq 0$ .”

## References

- Arratia, R., Waterman, M.S.: The Erdős-Rényi strong law for pattern matching with a given proportion of mismatches. *Ann. Probab.* **17**, 1152–1169 (1989)
- Book, S.A.: An extension of the Erdős-Rényi law of large numbers. *Proc. Am. Math. Soc.* **48**, 438–446 (1975)
- Comets, F.: Grandes déviations pour des champs de Gibbs sur  $Z^d$ . *C.R. Acad. Sc. Paris, Série I* **303**, 511–514 (1986)
- Csörgő, S.: Erdős-Rényi laws. *Ann. Statist.* **7**, 772–787 (1979 a)
- Csörgő, S.: Badahur efficiency and Erdős-Rényi maxima. *Sankhya Ser. A* **41**, 141–144 (1979 b)
- Deheuvels, P., Devroye, L., Lynch, J.: Exact convergence rate in the limit theorems of Erdős-Rényi an Shepp. *Ann. Probab.* **14**, 209–233 (1986)
- Dembo, A., Karlin, S.: Strong limit theorems of empirical functionals for large exceedance of partial sums of i.i.d. variables. *Ann. Probab.* **19**, 1737–1755 (1991 b)
- Dembo, A., Karlin, S.: Strong limit theorems of empirical distributions for large segmental exceedances of partial sums of Markov variables. *Ann. Probab.* **19**, 1756–1767 (1991 b)
- Deuschel, J.D., Stroock, D.W.: Large deviations. Boston: Academic Press (1989)
- Dobrushin, R.L., Kotecký, R., Shlosman, S.B.: Wulff construction: A global shape from local interaction. *Transl. of Math. Monographs* **104**. Providence: Am. Math. Soc. (1992)
- Erdős, P., Rényi, A.: On a new law of large numbers. *J. Analyse Math.* **23**, 103–111 (1970)
- Erdős, P., Révész, P.: On the length of the longest head-run. *Topics in Information theory Keszthely (Hungary) Coll. Math. Soc. János Bolyai* **16**, 219–228 (1975)
- Föllmer, H., Orey, S.: Large deviations for the empirical field of a Gibbs measure. *Ann. Probab.* **16**, 961–977 (1988)
- Freidlin, M.I., Wentzell, A.D.: Random perturbation of dynamical systems. Berlin, Heidelberg, New York: Springer (1984)
- Georgii, H.O.: Large deviations and maximum entropy principle for interacting random fields on  $Z^d$ . Preprint Univers. Munich (1992)
- Georgii, H.O.: Gibbs measures and phase transitions. Berlin: de Gruyter (1988)
- Olla, S.: Large deviations for Gibbs random fields. *Probab. Th. Rel. Fields* **77**, 343–357 (1988)
- Samarova, S.S.: On the length of the longest head-run for a Markov chain with two states. *Th. Prob. Appl.* **26**, 498–509 (1981)

Communicated by M. Aizenman

