BULLETIN BRAZILIAN MATHEMATICAL SOCIETY

*Bull Braz Math Soc, New Series 33(3), 427-446*


# Gumbel statistics for the longest interval of identical spins in a one-dimensional Gibbs measure

F. Redig

**Abstract.** We consider one-dimensional Gibbs measures on spin configurations $\sigma \in \{-1,+1\}^{\mathbb{Z}}$. For $N \in \mathbb{N}$ let $l_N$ denote the length of the longest interval of consecutive spins of the same kind in the interval $[0, N]$. We show that the distribution of a suitable continuous modification $l_c(N)$ of $l_N$ converges to the Gumbel distribution, i.e., for some $\alpha, \beta \in (0,\infty)$ and $\gamma \in \mathbb{R}$,

$$\lim_{N\to\infty} \mathbb{P}(l_c(N) \le \alpha \log N + \beta x + \gamma) = e^{-e^{-x}}.$$

**Keywords:** Gibbs measures, extreme values, Gumbel distribution.

**Mathematical subject classification:** Primary: 82C22; Secondary: 60K35.

## 1 Introduction

In the theory of extreme values, one is interested in the asymptotic distribution of the maximum of random variables. The typical question is the following: for a stationary $\mathbb{R}$-valued process $\{X_n : n \in \mathbb{N}\}$ find $u_n(x)$ such that the sequence of probabilities

$$\mathbb{P}\left(\max_{i=1}^{n} X_i \le u_n(x)\right) \tag{1.1}$$

has a non-trivial limit $G(x)$ as $n$ tends to infinity. For i.i.d. sequences, a complete picture is given, i.e., all possible limiting distributions $G(x)$ are known, and there is a very detailed description of the classes of distribution functions corresponding to the different possibilities of $G(x)$, see e.g. [6], [4]. For stationary dependent sequences, the first results were obtained in [7]. Generally, two conditions are needed in order to arrive at the usual extreme value distributions for independent maxima. The first one is a strong mixing condition, and the second

Received 2 September 2002.

one requires a control of certain conditional probabilities (see section 2 below for a precise formulation). These two conditions then imply that the statement

$$\lim_{n\to\infty} n\mathbb{P}\left(X_0 \geq u_n(x)\right) = x \tag{1.2}$$

is equivalent with the statement

$$\lim_{n\to\infty} \mathbb{P}\left(\max_{i=1}^{n} X_i \leq u_n(x)\right) = e^{-x}. \tag{1.3}$$

Note that the equivalence of (1.2) and (1.3) is trivial in the i.i.d. case. Once the equivalence between (1.2) and (1.3) is established, the problem of finding the extreme value distribution is reduced to finding the sequence $u_n(x)$ of (1.2). E.g. if $X_n$ are non-negative variables with

$$\lim_{x\to\infty} \frac{\mathbb{P}(X_0 \geq x)}{e^{-\alpha x}} = 1$$

then one can choose

$$u_n(x) = \frac{\log n}{\alpha} + \frac{x}{\alpha}$$

to obtain

$$\lim_{n\to\infty} \mathbb{P}\left(\max_{i=1}^{n} X_i \leq \frac{\log n}{\alpha} + \frac{x}{\alpha}\right) = e^{-e^{-x}}.$$

The distribution $G(x) = e^{-e^{-x}}$ is called the Gumbel distribution. If the limiting distribution is of the form $G(ax+b)$, then one says that it is of Gumbel-type. As we already mentioned, the conditions to obtain the equivalence between (1.2) and (1.3) are a mixing condition, and a condition involving conditional probabilities. The context of one-dimensional Gibbs measures seems therefore well-adapted to this kind of question because Gibbs measures have nice conditional probabilities. Surprisingly, no general results on extreme values for Gibbs measures are available in the literature. It is the aim of this paper to connect both notions. The context of one-dimensional Gibbs measures is then the first test-case. Although the one dimensional spin-systems are "trivial" from the point of view of critical phenomena, the question we address here is rather detailed, and becomes much more complicated in the two-dimensional situation.

A natural context in which Gibbs measures can be defined are discrete lattice spin systems. This means that a Gibbs measure is a probability measure $\mu$ on configurations $\sigma \in \{-1, +1\}^{\mathbb{Z}}$, specified by conditional probabilities of the form

$$\mu(\sigma_\Lambda | \eta_{\Lambda^c}) = \frac{\exp(-H_\Lambda^\eta(\sigma))}{Z}, \tag{1.4}$$

where the Hamiltonians $H_\Lambda^\eta(\sigma)$ are of the form

$$H_\Lambda^\eta(\sigma) = \sum_{A\cap\Lambda\neq\emptyset} U(A, \sigma_\Lambda\eta_{\Lambda^c}). \tag{1.5}$$

An extreme value question which can be asked in this context is the following. For $N \in \mathbb{N}$ large what is the asymptotic distribution $P_N$ of the length of the longest interval inside $[0, N]$ which contains only $+$ or $-$ spins? Unhappily, the question asked in this way is not well-defined: due to the discreteness, there is no such asymptotic distribution (not even in the case of independent $+, -$ spins, see theorem 2.4.5, and example 2.4.1 in [4]). Indeed, for a stationary sequence $X_j$ taking values in the set of non-negative integers, a necessary condition for the existence of a suitable normalization of the maximum (i.e., a sequence of numbers $a_n$ and $b_n$ such that $(max_{i=1}^n X_j - a_n)/b_n$ has a non-degenerate limiting distribution) is

$$\lim_{k\to\infty} \frac{P(X_1 = k)}{P(X_1 \geq k)} = 0$$

see [4], corollary 2.4.1. This condition cannot be satisfied if the tails of the distribution of $X_i$ are exponential, as e.g. in the case of the geometric distribution.

There are two ways to make the question well-defined. First, one considers all possible limit points of the distributions $P_N$ and shows that they lie in a well-defined neighborhood of the Gumbel distribution. Second, we can remove the discreteness of the random variables $l_N$ by putting around each spin an interval of exponentially distributed length, to which we give a colour (say black for $+$ spin, white for $-$ spin). We then ask for the length of the longest interval of one colour inside $[0, N]$ (denoted by $l_N^c$). In the second case, the distribution of $l_N^c$ does converge to a Gumbel-type distribution. Of course, the choice of the exponential interval around each spin seems quite ad hoc. We will see however that this choice is very well adapted to the question, because in the case of independent $+, -$ spins, the length of intervals of the same colour have an exponential distribution.

The paper is organized as follows: in section 2 we introduce the elements of extreme value theory and theory of Gibbs measures which we need, and formulate our results. In section 3 we consider the easy case of independent $+, -$ spins. Section 4 is devoted to the proofs in the general case.

## 2 Main Theorem

In this section we give some elementary background on Gibbs measures and extreme values and state our results.

## 2.1 Gibbs measures

We consider the configuration space $\Omega = \{-1, +1\}^{\mathbb{Z}}$. The symbol $S$ denotes the set of all finite subsets of $\mathbb{Z}$. The shift over $x \in \mathbb{Z}$ on configurations $\sigma$ is denoted by $\tau_x$: $\tau_x(\sigma)(y) = \sigma(y + x)$. For $A \subset \mathbb{Z}$ we denote $\mathrm{diam}(A) = \max(A) - \min(A)$. For $A \subset \mathbb{Z}$ we denote by $\mathcal{F}_A$ the $\sigma$-field generated by $\{\sigma(x) : x \in A\}$, and by $\Omega_A$ the configuration space $\{-1, 1\}^A$. For $\Lambda \in S$, and $\sigma, \eta \in \Omega$, we denote by $\sigma_\Lambda \eta_{\Lambda^c}$ the configuration coinciding with $\sigma$ on $\Lambda$ and with $\eta$ on $\Lambda^c$.

**Definition 2.1.** *An interaction is a map*

$$U : S \times \Omega \to \mathbb{R} \tag{2.2}$$

*satisfying*

1. *$U(A, \sigma)$ depends only on $\sigma(x), x \in A$.*
2. *Translation invariance:*
$$U(A + a, \sigma) = U(A, \tau_a \sigma) \tag{2.3}$$
3. *Strong Summability:*
$$\|U\| = \sum_{A \ni 0} \mathrm{diam}(A) \sup_\sigma |U(A, \sigma)| < \infty \tag{2.4}$$

**Remark 2.5.** We impose here a "strong summability" condition which is more than one needs to define Gibbs measures. This condition implies that we do not have phase transition (see e.g. [5] section 8.3).

Given an interaction we define, for $\Lambda \subset \mathbb{Z}^d$ a finite set, Hamiltonians in volume $\Lambda$ with boundary condition $\eta$:

$$H^\eta_\Lambda(\sigma) = \sum_{A \cap \Lambda \neq \emptyset} U(A, \sigma_\Lambda \eta_{\Lambda^c}), \tag{2.6}$$

and the finite volume Gibbs measures

$$\mu^{U,\eta}_\Lambda(\sigma) = \frac{\exp(-H^\eta_\Lambda(\sigma))}{Z^\eta_\Lambda}, \tag{2.7}$$

where the normalizing constant

$$Z_\Lambda^\eta = \sum_{\sigma\in\Omega_\Lambda} e^{-H_\Lambda^\eta(\sigma)} \tag{2.8}$$

is the finite-volume partition function with boundary condition $\eta$. $\mu_\Lambda^{U,\eta}$ is thus an $\eta$-dependent measure on $\{-1,+1\}^\Lambda$, and will serve as a candidate conditional probability distribution of the configuration inside $\Lambda$, given that the configuration outside $\Lambda$ is specified to be $\eta$. For a probability measure on $\Omega$, we denote $\mu_\Lambda^\eta$ to be the conditional probability distribution of the configuration inside $\Lambda$, given $\eta$ outside $\Lambda$ (defined for $\mu$-a.e. $\eta$). The following definition introduces the notion of Gibbs measure in the DLR-sense.

**Definition 2.9.** *A probability measure $\mu$ is a Gibbs measure with interaction $U$ if for $\mu$-a.e. $\eta$:*

$$\mu_\Lambda^\eta = \mu_\Lambda^{U,\eta}(\sigma) \tag{2.10}$$

The set of all Gibbs measure with interaction $U$ is denoted by $\mathcal{G}(U)$. In words, Definition 2.9 states that a Gibbs measure is a measure specified by a priori defined conditional probabilities. In our concrete one-dimensional context with interactions satisfying (2.4), $\mathcal{G}(U)$ is a singleton (i.e., no phase transition, see e.g. Theorem 8.93 in [5]) and hence we can use the symbol $\mu(U)$ to denote the unique Gibbs measure corresponding to $U$.

## 2.2 Extreme values

In this section we summarize the results on extreme value theory for stationary processes which we need in this paper. These results, and more background can be found in [6], [1] and [4]. We consider a stationary (two-sided) $\mathbb{R}$-valued process $\{X_n : n\in\mathbb{Z}\}$. For a finite or infinite $A\subset\mathbb{Z}$, we denote by $\mathcal{F}_A$ the $\sigma$-field generated by $X_i$, $i\in A$. For $A, B\subset\mathbb{Z}$, $d(A,B)$ denotes the distance between $A$ and $B$:

$$d(A,B) = \min\{|i-j| : i\in A,\ j\in B\}.$$

**Definition 2.11.** *The process $\{X_i : i\in\mathbb{Z}\}$ is called $\alpha$-mixing if there exists $\alpha:\mathbb{N}\to[0,\infty)$, with $\alpha(n)\to 0$ as $n\to\infty$ such that for any $A, B\subset\mathbb{Z}$, and any $F\in\mathcal{F}_A$, $G\in\mathcal{F}_B$:*

$$|\mathbb{P}(F\cap G) - \mathbb{P}(F)\mathbb{P}(G)| \le \alpha(d(A,B)). \tag{2.12}$$

In the following definition we write $[\cdot]$ for the greatest of integer function.

**Definition 2.13.** *Let $u_n$ be a sequence of real numbers. The stationary sequence is said to satisfy condition $D'(u_n)$ if*

$$\lim_{k\to\infty} \limsup_{n\to\infty} \sum_{j=2}^{[n/k]} \mathbb{P}(X_1 \geq u_n, X_j \geq u_n) = 0. \tag{2.14}$$

Let us denote

$$M_n = \max_{i=1}^{n} X_i.$$

The following theorem is proved in [6].

**Theorem 2.15.** *Let $u_n$ be a sequence of real numbers such that for an $\alpha$-mixing stationary process $\{X_i : i \in \mathbb{Z}\}$ the condition $D'(u_n)$ is satisfied. Then*

$$\lim_{n\to\infty} \mathbb{P}(M_n \leq u_n) = e^{-x} \tag{2.16}$$

*is equivalent with*

$$\lim_{n\to\infty} n\mathbb{P}(X_1 > u_n) = x \tag{2.17}$$

This theorem implies that as far as the behavior of the maxima of the stationary sequence is concerned, we can consider it as a sequence of i.i.d. copies of $X_1$. The condition 2.14 is usually called "short range condition" and ensures that large values do not occur in "clumps". E.g., a sequence with $X_{2n+1} = X_{2n}$ for all $n$, where large values would occur in pairs, is excluded by this condition.

## 2.3 The problem

Let $\mu$ be the Gibbs measure with interaction $U$ as in Definition 2.1. For $\sigma \in \Omega$ we define

$$\begin{aligned} X_0(\sigma) &= \min\{k \geq 0 : \sigma(k)\sigma(k+1) = -1\} \\ X_{n+1} &= \min\{k \geq X_n + 1 : \sigma(k)\sigma(k+1) = -1\} \text{ for } n \geq 0 \end{aligned} \tag{2.18}$$

We will suppose here that all $X_i < \infty$, excluding configurations $\sigma$ with a half-line of agreeing spins. This is not a restriction, since we will only need $\mu$-typical configurations. We then define the intervals

$$\begin{aligned} I_0(\sigma) &= [0, X_0(\sigma)] \\ I_k(\sigma) &= [X_{k-1}(\sigma) + 1, X_k(\sigma)], \text{ for } k \geq 1. \end{aligned} \tag{2.19}$$

Inside each interval $I_i$, the spins are of the same kind. By translation invariance of $\mu$, conditioned on the event $\sigma(-1) \neq \sigma(0)$, the lengths of the intervals with positive index

$$|I_k| = X_k - X_{k-1} + 1 \tag{2.20}$$

form a stationary sequence. We are interested in the distribution $P_N$ of the length of the longest interval of type $I_k$ inside $[0, N]$, as $N$ tends to infinity. As explained in the introduction, due to the discrete character of the variables $|I_k|$, the sequence $P_N$ will not converge (cf. example 1.7.15 in [6]). Therefore we first construct a natural continuous version of the variables $I_k$ and $l_k$ by an extra randomization as follows. Consider a sequence $\{\xi_k : k \in \mathbb{N}\}$ of i.i.d. exponential variables. We then define the random variables

$$Z_k = \sum_{i \in I_k(\sigma)} \xi_i. \tag{2.21}$$

and the corresponding random intervals

$$\begin{aligned} \mathcal{I}_0 &= [0, Z_0] \\ \mathcal{I}_k &= [\sum_{r=0}^{k-1} Z_r, \sum_{r=0}^{k} Z_r]. \end{aligned} \tag{2.22}$$

In words this means the following: to each $+$ or $-$ spin we associate a "colored" interval (say black for $+$, white for $-$) with exponential length and we glue these intervals together. The advantage of these extra randomized intervals is the fact that now the lengths $|\mathcal{I}_k|$ have a continuous distribution. Therefore these intervals can be considered from the point of view of extreme value theory. Later we will argue that the particular choice of randomization ($\xi_i$ exponentially distributed) is actually not important. We have chosen it here for the elegance of the presentation.

In what follows we will denote by $\mathbb{P}_{\mu_0}$ the joint distribution of the Gibbs measure $\mu$ (distribution of $\sigma$) conditioned on $\sigma(-1) \neq \sigma(0)$ and the independent i.i.d. sequence $\{\xi_k : k \in \mathbb{N}\}$. Under $\mathbb{P}_{\mu_0}$ the sequence $\{|\mathcal{I}_k| : k \in \mathbb{N}\}$ is stationary. The symbol $\mu_0$ will denote the measure $\mu$ conditioned on $\sigma(-1) \neq \sigma(0)$.

### 2.4 Results

**Theorem 2.23.** *Let $\mu$ be the Gibbs measure with interaction as in Definition 2.1. There exist $\alpha, \beta > 0$ and $\gamma \in \mathbb{R}$ such that for all $x \in \mathbb{R}$:*

$$\lim_N \mathbb{P}_{\mu_0}\left(\max_{k=1}^{N} |\mathcal{I}_k| \leq \alpha \log N + \beta x + \gamma\right) = e^{-e^{-x}} \tag{2.24}$$

For the following theorem, we need some more notation. For $N \in \mathbb{N}$ we define

$$\mathcal{M}'_N = \max\{|\mathcal{I}_k| : \mathcal{I}_k \subset [0, N]\}, \tag{2.25}$$

with the convention $\max(\emptyset) = \infty$. In words, $\mathcal{M}'_N$ is the length of the maximal interval of type $\mathcal{I}_k$ inside $[0, N]$. The variable $\mathcal{M}'_N$ is physically more relevant than the maximum $\mathcal{M}_N = \max_{i=1}^N |\mathcal{I}_k|$, where the index of the interval is varied. The following theorem shows that a similar extreme value theorem hold for $\mathcal{M}'_N$.

**Theorem 2.26.** *Let $\mu$ be a Gibbs measure with interaction as in Definition 2.1 and $\mathcal{M}'_N$ defined as in (2.25). Then there exists strictly positive $\alpha', \beta' \in (0, \infty)$ and $\gamma' \in \mathbb{R}$ such that*

$$\lim_N \mathbb{P}_{\mu_0}\left(\mathcal{M}'_N \leq \alpha' \log N + \beta' x + \gamma'\right) = e^{-e^{-x}}. \tag{2.27}$$

## 3 The independent case

As an easy test case, we consider the case of i.i.d. $\pm$ spins with $\mu(\sigma(x) = 1) = p \geq (1 - p)$. Under $\mu_0$, the interval lengths $|I_k|$ are i.i.d. with distribution

$$\mu(|I_k| = n) = \frac{1}{2} p^{n-1}(1-p) + \frac{1}{2}(1-p)^{n-1} p \tag{3.1}$$

The interval lengths $|\mathcal{I}_k|$ have distribution

$$\mathbb{P}_{\mu_0}(|\mathcal{I}_k| \geq x) = \frac{1}{2}\left(e^{-px} + e^{-(1-p)x}\right) \tag{3.2}$$

Therefore,

$$\lim_{n\to\infty} n\mathbb{P}_{\mu_0}\left(|\mathcal{I}_k| \geq \frac{\log n + \log\frac{1}{2} + x}{1-p}\right) = e^{-x} \tag{3.3}$$

and we conclude

$$\mathbb{P}_{\mu_0}\left(\max_{i=1}^{n} |\mathcal{I}_i| \leq \frac{\log n + \log\frac{1}{2} + x}{1-p}\right) = e^{-e^{-x}}. \tag{3.4}$$

In this special case we can compute everything explicitly, due to the exact distribution (3.1) of $|I_0|$. The following elementary lemma shows that it is enough to know the tail of the distribution of $|I_0|$ to decide about the tail of the distribution of $|\mathcal{I}_0|$.

**Lemma 3.5.** *If $\alpha_n$ and $\beta_n$ is a sequence of positive numbers such that $\sum_n \alpha_n < \infty$, $\sum \beta_n < \infty$ and*

$$\lim_{n\to\infty} \frac{\alpha_n}{\beta_n} = 1 \tag{3.6}$$

*then*

$$\lim_{x\to\infty} \frac{\sum_{n=1}^{\infty} \alpha_n \left(\int_x^{\infty} \frac{t^{n-1}}{(n-1)!} e^{-t} dt\right)}{\sum_{n=1}^{\infty} \beta_n \left(\int_x^{\infty} \frac{t^{n-1}}{(n-1)!} e^{-t} dt\right)} = 1 \tag{3.7}$$

**Proof:** Let us call $a$ the limit in the lhs of (3.7). Applying L'Hopital's rule $n+1$ times gives the identity

$$a = \lim_{x\to\infty} \frac{\sum_{r=0}^{\infty} \frac{x^r}{r!} \alpha_{r+n}}{\sum_{r=0}^{\infty} \frac{x^r}{r!} \beta_{r+n}}, \tag{3.8}$$

which gives

$$\inf_r \frac{\alpha_{r+n}}{\beta_{r+n}} \le a \le \sup_r \frac{\alpha_{r+n}}{\beta_{r+n}}. \tag{3.9}$$

Since this inequality holds for any $n \in \mathbb{N}$, we conclude $a = 1$. □

## 4 Proofs

The proof of Theorem 2.24 and Theorem 2.27 is divided in three steps:

1. First we prove that Theorem 2.24 implies Theorem 2.27.

2. The second step consists in verifying the condition $D'(u_n)$ for general one-dimensional Gibbs measures.

3. Finally, we have to study the tail of the distribution of $|\mathcal{I}_0|$ under $\mathbb{P}_{\mu_0}$. This amounts in proving that there exists $c, c' > 0$ such that

$$\lim_{n\to\infty} \frac{\mu(|I_0| \ge n)}{e^{-cn}} = c' \tag{4.1}$$

Note that this is not an immediate consequence of the large deviation property of Gibbs measures: what we need here is $\mu(|I_0| \geq n) \simeq e^{-cn}O(1)$ for $n \to \infty$. The fact that the correction to the large deviation behaviour is $O(1)$ is typical for one-dimensional Gibbs measures (in general corrections of the order of the boundary can occur).

## 4.1 Step 1

**Lemma 4.2.** *Let $\{X_n, n \in \mathbb{Z}\}$ a stationary $\alpha$-mixing sequence satisfying condition $D'(u_n)$. Suppose $T_n$ is a sequence of positive random variables on the same probability space such that*

$$\lim_{n\to\infty} \frac{T_n}{n} = \alpha. \tag{4.3}$$

*If*

$$\lim_{n\to\infty} n\mathbb{P}\left(X_1 \geq u_n(x)\right) = x \tag{4.4}$$

*then we have*

$$\lim_{n\to\infty} \mathbb{P}\left(\max_{i=1}^{T_n} \leq u_{[n\alpha]}(x)\right) = e^{-x} \tag{4.5}$$

**Proof:** We follow the lines of [6], Theorem 3.4.1. Abbreviate $F(x) = \mathbb{P}(X_1 \leq x)$, $M_n = \max_{i=1}^n X_i$.

$$\begin{aligned} &\mathbb{P}\left(\max_{i=1}^{T_n} X_i \leq u_{[n\alpha]}(x)\right) \leq \\ &\leq \mathbb{P}\left(\max_{i=1}^{[n\alpha]-[\epsilon n]} X_i \leq u_{[n\alpha]}(x)\right) + \mathbb{P}\left(T_n \leq [\alpha n] - [\epsilon n]\right). \end{aligned} \tag{4.6}$$

Fix $k \in \mathbb{N}$ and put $n' = [([n\alpha] - [n\epsilon])/k]$. Since

$$\{M_{n'} > u_{[n\alpha]}(x)\} = \cup_{j=1}^{n'}\{X_j > u_{[n\alpha]}(x)\} \tag{4.7}$$

we have the inequality

$$\begin{aligned} &\sum_{j=1}^{n'} \mathbb{P}(X_j > u_{[n\alpha]}(x)) - \sum_{1\leq i<j\leq n'} \mathbb{P}\left(X_i > u_{[n\alpha]}(x), X_j > u_{[n\alpha]}(x)\right) \leq \\ &\leq \mathbb{P}(M_{n'} > u_{[n\alpha]}(x)) \leq \sum_{j=1}^{n'} \mathbb{P}(X_j > u_{[n\alpha]}(x)), \end{aligned} \tag{4.8}$$

which gives, using stationarity,

$$\begin{aligned} 1 - n'(1 - F(u_{[n\alpha]}(x))) &\leq \mathbb{P}\left(M_{n'} \leq u_{[n\alpha]}(x)\right) \\ &\leq 1 - n'(1 - F(u_{[n\alpha]}(x))) + S_n, \end{aligned} \tag{4.9}$$

where

$$S_n = S_{n,k} = n' \sum_{j=2}^{n'} \mathbb{P}\left(X_1 > u_{[n\alpha]}(x),\, X_j > u_{[n\alpha]}(x)\right) \tag{4.10}$$

Condition $D'(u_n)$ implies that

$$\lim_{k\to\infty} \limsup_{n\to\infty} S_{n,k} = 0. \tag{4.11}$$

Next, condition (4.4) implies that

$$\begin{aligned} \lim_{n\to\infty} n'(1 - F(u_{[n\alpha]}(x)) &= \lim_{n\to\infty} \frac{[\alpha n] - [\epsilon n]}{[\alpha n]k}[\alpha n](1 - F(u_{[n\alpha]}(x)) \\ &= (1 - \frac{\epsilon}{\alpha})\frac{x}{k}. \end{aligned} \tag{4.12}$$

This gives

$$\begin{aligned} \left(1 - (1 - \frac{\epsilon}{\alpha}\frac{x}{k})\right) &\leq \liminf_{n\to\infty} \mathbb{P}(M_{n'} \leq u_{[n\alpha]}(x)) \\ &\leq \limsup_{n\to\infty} \mathbb{P}(M_{n'} \leq u_{[n\alpha]}(x)) \\ &\leq \left(1 - (1 - \frac{\epsilon}{\alpha})\frac{x}{k} + o(\frac{1}{k})\right). \end{aligned} \tag{4.13}$$

By the $\alpha$-mixing condition we obtain that for any $k$ fixed:

$$\lim_{n\to\infty} \left[\mathbb{P}\left(M_{[\alpha n]-[\epsilon n]} \leq u_{[n\alpha]}(x)\right) - \left(\mathbb{P}\left(M_{([[\alpha n]-[\epsilon n])/k]} \leq u_{[n\alpha]}(x)\right)\right)^k\right] = 0 \tag{4.14}$$

By taking $k$-th powers in (4.13) we then obtain

$$\begin{aligned} \left(1 - (1 - \frac{\epsilon}{\alpha}\frac{x}{k})\right)^k &\leq \liminf_{n\to\infty} \mathbb{P}\left(M_{[\alpha n]-[\epsilon n]} \leq u_{[n\alpha]}(x)\right) \\ &\leq \limsup_{n\to\infty} \mathbb{P}\left(M_{[\alpha n]-[\epsilon n]} \leq u_{[n\alpha]}(x)\right) \\ &\leq \left(1 - \left(1 - \frac{\epsilon}{\alpha}\frac{x}{k} + o(\frac{1}{k})\right)\right)^k. \end{aligned} \tag{4.15}$$

Letting $k$ tend to $\infty$ we obtain

$$\lim_{n\to\infty} \mathbb{P}\left(M_{[\alpha n]-[\epsilon n]} \leq u_{[n\alpha]}(x)\right) = e^{-(1-\frac{\epsilon}{\alpha})x}. \tag{4.16}$$

which gives, combined with (4.6):

$$\limsup_{n\to\infty} \mathbb{P}\left(\max_{i=1}^{T_n} X_i \leq u_{[n\alpha]}(x)\right) \leq e^{-(1-\frac{\epsilon}{\alpha})x}. \tag{4.17}$$

Since $\epsilon > 0$ is arbitrary, we conclude

$$\limsup_{n\to\infty} \mathbb{P}\left(\max_{i=1}^{T_n} X_i \leq u_{[n\alpha]}(x)\right) \leq e^{-x}. \tag{4.18}$$

To prove the opposite inequality, we start from

$$\begin{aligned} &\mathbb{P}(\max_{i=1}^{T_n} X_i \leq u_{[n\alpha]}(x)) \geq \\ &\geq \mathbb{P}(\max_{i=1}^{[\alpha n]+[\epsilon n]} X_i \leq u_{[n\alpha]}(x)) + \mathbb{P}\left(T_n \geq [\alpha n] + [\epsilon n]\right), \end{aligned} \tag{4.19}$$

and follow the same line of reasoning. □

In order to see that this lemma is enough to deduce Theorem 2.26 from Theorem 2.23, notice that (see (2.25))

$$\mathcal{M}'_N = \mathcal{M}_{T_N} \tag{4.20}$$

where

$$T_N = \max\{j \leq N : \mathcal{I}_j \subset [0, N]\}. \tag{4.21}$$

Under $\mathbb{P}_{\mu_0}$, the sequence $\{|\mathcal{I}_n|, n \in \mathbb{Z}\}$ is stationary and $\alpha$-mixing. This is a direct consequence of the renewal construction of chains of infinite order (of which our Gibbs measures are particular examples) of [3]. Therefore, $\mathbb{P}_{\mu_0}$ almost surely:

$$\lim_{N\to\infty} \frac{T_N}{N} = \alpha = (\mathbb{E}_\mu |I_0|)^{-1}. \tag{4.22}$$

### 4.2 Step 2: the condition $D'(u_n)$

Here we prove that the condition $D'(u_n)$ is satisfied for Gibbs measures with an interaction as in Definition 2.1.

**Lemma 4.23.** *Suppose $u_n(x)$ is chosen such that*

$$\lim_{n\to\infty} \mathbb{P}(X_1 \geq u_n(x)) = x. \tag{4.24}$$

*Then, if*

$$\sup_{j\neq 1, n\in\mathbb{N}} \left( \frac{\mathbb{P}(X_1 \geq u_n(x)|X_j \geq u_n(x))}{\mathbb{P}(X_1 \geq u_n(x))} < \infty, \right) \tag{4.25}$$

*condition $D'(u_n)$ is satisfied.*

**Proof:** First remark that, by stationarity

$$\begin{aligned} &\lim_{k\to\infty} \limsup_{n\to\infty} n \sum_{j=2}^{[n/k]} \mathbb{P}(X_j \geq u_n(x))\mathbb{P}(X_1 \geq u_n(x)) \\ &= \lim_{k\to\infty} \limsup_{n\to\infty} n[\frac{n}{k}](\mathbb{P}(X_1 \geq u_n(x))^2 \\ &= \lim_{k\to\infty} \frac{x^2}{k} = 0 \end{aligned} \tag{4.26}$$

Therefore,

$$\begin{aligned} &\lim_{k\to\infty} \limsup_{n\to\infty} n \sum_{j=2}^{[n/k]} \mathbb{P}(X_j \geq u_n(x), X_1 \geq u_n(x)) \\ &= \lim_{k\to\infty} \limsup_{n\to\infty} n \sum_{j=2}^{[n/k]} \left( \frac{\mathbb{P}(X_1 \geq u_n(x)|X_j \geq u_n(x))}{\mathbb{P}(X_1 \geq u_n(x))} - 1 \right) \mathbb{P}(X_1 \geq u_n(x))^2 \\ &\leq \sup_{j\neq 1, n\in\mathbb{N}} \left( \frac{\mathbb{P}(X_1 \geq u_n(x)|X_j \geq u_n(x))}{\mathbb{P}(X_1 \geq u_n(x))} \right) \cdot \\ &\quad \lim_{k\to\infty} \limsup_{n\to\infty} n[\frac{n}{k}](\mathbb{P}(X_1 \geq u_n(x))^2 = 0. \end{aligned} \tag{4.27}$$

□

**Lemma 4.28.** *If $U$ is an interaction satisfying the conditions of Definition 2.1, and $\mu = \mu(U)$ the unique Gibbs measure corresponding to $U$, then*

$$\sup_{a\in(0,\infty), n\in\mathbb{N}} \left( \frac{\mathbb{P}_{\mu_0}(|\mathcal{I}_1| \geq a \cap |\mathcal{I}_n| \geq a)}{\mathbb{P}_{\mu_0}(|\mathcal{I}_1| \geq a)^2} \right) < \infty \tag{4.29}$$

**Proof:** For $\xi_i$ i. i. d. exponential (mean one) random variables we abbreviate

$$\alpha_n(a) = \mathbb{P}(\sum_{i=1}^{n} \xi_i \geq a) = \int_a^\infty \frac{t^{n-1}}{(n-1)!} e^{-t} dt. \tag{4.30}$$

Then we write,

$$\begin{aligned} \frac{\mathbb{P}_{\mu_0}(|\mathcal{I}_1| \geq a \cap |\mathcal{I}_n| \geq a)}{\mathbb{P}_{\mu_0}(|\mathcal{I}_1| \geq a)^2} &= \frac{\sum_{n,m} \alpha_n(a)\alpha_m(a)\mu_0\left(|I_1| = n, |I_j| = m\right)}{\sum_{n,m} \alpha_n(a)\alpha_m(a)\mu_0(|I_1| = n)\mu_0(|I_j| = m)} \\ &\leq \sup_{n,m} \left( \frac{\mu_0\left(|I_1| = n, |I_j| = m\right)}{\mu_0(|I_1| = n)\mu_0(|I_j| = m)} \right). \end{aligned} \tag{4.31}$$

To prove that the supremum in the right hand side of (4.31) is finite, we can replace $\mu$ by $\mu_0$ and vice versa. This can be seen as follows. For $\sigma \in \Omega$ let us denote by $\sigma^x$ the configuration "flipped" at site $x$, i.e., $\sigma^x(y) = (-1)^{\delta_{x,y}} \sigma(y)$. The transformed measure $\mu^x$ defined via $\int f(\sigma^x)\mu(d\sigma) = \int f(\sigma)\mu^x(d\sigma)$ is absolutely continuous w.r.t. $\mu$ with Radon-Nikodym derivative

$$\frac{d\mu^x}{d\mu}(\sigma) = \exp\left( -\sum_{A \ni x} \left[ U(A, \sigma^x) - U(A, \sigma) \right] \right), \tag{4.32}$$

which clearly satisfies

$$e^{-2\|U\|} \leq \frac{d\mu^x}{d\mu}(\sigma) \leq e^{2\|U\|}, \tag{4.33}$$

where $\|U\|$ is the norm introduced in Definition 2.1. Note by $E$ the event $\{\sigma : \sigma(-1) \neq \sigma(0)\}$. For all $A \in \mathcal{F}_{[0,\infty)}$ we have

$$\frac{\int_{A \cap E} d\mu}{\int_{A \cap E^c} d\mu} = \frac{\int_{A \cap E^c} d\mu^{(-1)}}{\int_{A \cap E^c} d\mu}. \tag{4.34}$$

This gives, together with (4.33),

$$e^{-2\|U\|} \leq \frac{\mu_0(A)}{\mu(A)} \leq e^{2\|U\|}. \tag{4.35}$$

Denote

$$\begin{aligned} E_{kl} &= \{\sigma : \sigma(k) = \sigma(k+1) = \cdots = \sigma(k+l-1) \neq \sigma(k+l)\} \\ F_{kl} &= \{\sigma : \sigma(k) = \sigma(k+1) = \cdots = \sigma(k+l-1)\}. \end{aligned} \tag{4.36}$$

The bound (4.33) gives the existence of $0 < c_1 < c_2 < \infty$, such that for any $A \in \mathcal{F}_{[k+l+1,\infty)}$:

$$c_1 \leq \frac{\mu(E_{kl} \cap A)}{\mu(F_{kl} \cap A)} \leq c_2. \tag{4.37}$$

Now we can estimate

$$\begin{aligned}
&\mu\left(|I_1| = n \cap |I_j| = m\right) \\
&= \sum_{k_2,\dots,k_{j-1}} \mu\left(E_{0n} \cap |I_2(\sigma)| = k_2 \cap \dots \cap |I_j(\sigma)| = m\right) \\
&\leq C\mu(F_{0n}) \sum_{k_2,\dots,k_{j-1}} \mu\left[|I_1(\sigma)| = k_2 \cap \dots \cap |I_{j-1}(\sigma)| = m \mid \tau_{-n}(F_{0n})\right].
\end{aligned} \tag{4.38}$$

Here we used (4.37) and translation invariance of $\mu$ to arrive at the last equality. Now we can use the following property which is typical for one-dimensional Gibbs measures. For $A \in \mathcal{F}_{(-\infty,-1]}$ and $B \in \mathcal{F}_{[0,\infty)}$ we have the estimate

$$\begin{aligned}
\frac{\mu(B|A)}{\mu(B)} &\leq \sup_{\eta,\xi \in \Omega_{[0,\infty)}} \frac{\mu(B|\eta)}{\mu(B|\xi)} \\
&\leq \exp\left(2 \sup_{N\in\mathbb{N}} \sup_{K\in\mathbb{N}} \sup_{\eta,\xi,\sigma\in\Omega} |H^{\eta}_{[-K,N]}(\sigma) - H^{\xi}_{[-K,N]}(\sigma)|\right) \\
&\leq \exp\left(4 \sum_{A\subset\mathbb{Z},\min(A)\leq 0<\max(A)} \|U(A)\|_\infty\right) \\
&\leq \exp\left(4 \sum_{A\ni 0} \mathrm{diam}(A)\|U(A)\|_\infty\right) = \exp(4\|U\|) < \infty.
\end{aligned} \tag{4.39}$$

Using this estimate, we can proceed with (4.38):

$$\begin{aligned}
&\mu\left(|I_1| = n \cap |I_j| = m\right) \\
\leq\ & C_2\mu(F_{0n}) \sum_{k_2,\dots,k_{j-1}} \mu\left(|I_1(\sigma)| = k_2 \cap \dots \cap |I_{j-2}(\sigma)| = k_{j-1} \cap |I_{j-1}(\sigma)| = m\right) \\
=\ & C_2\mu(F_{0n})\mu\left(|I_{j-1}| = m\right) \\
\leq\ & C_3\mu(F_{0n})\mu_0\left(|I_j| = m\right) \\
\leq\ & C_4\mu(F_{0n})\mu\left(|I_j| = m\right).
\end{aligned} \tag{4.40}$$

Here in the last steps we used (4.35) and the stationarity of $\{|I_k| : k \in \mathbb{N}\}$ under $\mu_0$. From (4.40) we obtain, using (4.37)

$$\frac{\mu_0\left(|I_1| = n, |I_j| = m\right)}{\mu_0(|I_1| = n)\mu_0(|I_j| = m)} \leq C_4 \frac{\mu(F_{0n})}{\mu(E_{0n})} \leq C_5 < \infty. \tag{4.41}$$

$\square$

Since the estimate in the previous lemma is uniform in $a, n$, $D'(u_n(x))$ is satisfied for the stationary $\alpha$-mixing sequence $\{|\mathcal{I}_n|, n \in \mathbb{Z}\}$.

### 4.3 Step 3: tail probability

The last step in proving Theorem 2.24 consists in showing that the tail of the distribution of $|\mathcal{I}_0|$ under $\mathbb{P}_{\mu_0}$ is exponential. By Lemma 3.5 it suffices to prove the following.

**Lemma 4.42.** *If $U$ is an interaction as in Definition 2.1, and $\mu$ is the unique Gibbs measure corresponding to $U$, then there exist $c, c' \in (0, \infty)$ such that*

$$\lim_{n\to\infty} \frac{\mu(|I_0| = n)}{e^{-cn}} = c'. \tag{4.43}$$

**Proof:** We will use $c, c'$ for strictly positive constants, but their value can change from place to place. We have to prove that there exist $c, c' \in (0, \infty)$ such that

$$\lim_{n\to\infty} \frac{\mu\left(\sigma(0) = \sigma(1) = \cdots = \sigma(n) = 1 = -\sigma(n+1)\right)}{e^{-cn}} = c' \tag{4.44}$$

By the continuity of the conditional probabilities of $\mu$, the ratio

$$\frac{\mu\left(\sigma(0) = \sigma(1) = \cdots = \sigma(n) = 1 = -\sigma(n+1)\right)}{\mu\left(\sigma(0) = \sigma(1) = \cdots = \sigma(n) = 1 = \sigma(n+1)\right)} \tag{4.45}$$

converges, as $n$ goes to infinity, to

$$\mathbb{E}_\mu\left(\frac{d\mu^0}{d\mu}|\mathcal{F}_{(-\infty,0)}\right)(+), \tag{4.46}$$

where $+$ denotes the all-plus configuration, and $\frac{d\mu^0}{d\mu}$ the Radon Nikodym derivative of $\mu$ with respect to spin-flip at the origin. Therefore, it suffices to show that there exists constants $c, c' \in (0, \infty)$ such that

$$\lim_{n\to\infty} \frac{\mu(\sigma(0) = \sigma(1) = \cdots = \sigma(n) = 1)}{e^{-cn}} = c' \tag{4.47}$$

The probability in the denominator of the lhs of (4.47) can be rewritten as follows:

$$\mu(\sigma(0) = \sigma(1) = \cdots = \sigma(n) = 1) = \lim_{N\to\infty} \frac{Z^{+}_{[-N,1]\cup[n+1,N]} \exp(-H^{\emptyset}_{[0,n]}(+))}{Z^{+}_{[-N,N]}}, \tag{4.48}$$

where

$$Z^{+}_{\Lambda} = \sum_{\sigma\in\Omega_\Lambda} \exp(-H^{+}_{\Lambda}(\sigma)), \tag{4.49}$$

and

$$H^{\emptyset}_{[0,n]}(+) = \sum_{A\subset[0,n]} U(A,+) \tag{4.50}$$

Since we are in the one-dimensional situation, the partition function of a lattice interval $[a, b]$ satisfies

$$Z^{+}_{[a,b]} = \exp((b-a)P(U) + O(1)), \tag{4.51}$$

where the pressure $P(U)$ is defined as

$$P(U) = \lim_{N\to\infty} \frac{1}{2N+1} \log Z^{+}_{[-N,N]}, \tag{4.52}$$

and where $O(1)$ is bounded and converges to a constant when $b - a$ tends to infinity. Moreover,

$$\frac{Z^{+}_{[-N,1]\cup[n+1,N]}}{Z^{+}_{[-N,1]} Z^{+}_{[n+1,N]}} = \alpha_n, \tag{4.53}$$

where $\alpha_n$ converges to one, uniformly in $N$ as $n$ tends to infinity. Combination of (4.48), (4.51), (4.52) and (4.53) together with the following lemma finishes the proof. □

**Lemma 4.54.** *Let $U$ be an interaction as in Definition 2.1. Then the limit*

$$\lim_{n\to\infty} \Big( \sum_{A\subset[0,n]} U(A,+) - ne^{+}(U)\Big) = c, \tag{4.55}$$

*exists and defines a finite constant, where*

$$e^{+}(U) = \sum_{A\ni 0} \frac{U(A,+)}{|A|} \tag{4.56}$$

**Proof:** We rewrite, as in [5]

$$
\begin{aligned}
\sum_{A\subset[0,n]} U(A,+) &= \sum_{x\in[0,n]} \sum_{A\ni x, A\subset[0,n]} \frac{U(A,+)}{|A|} \\
&= \sum_{x\in[0,n]} \sum_{A\ni x} \frac{U(A,+)}{|A|} - \sum_{x\in[0,n]} \left( \sum_{A\ni x, A\cap[0,n]^c\neq\emptyset} \frac{U(A,+)}{|A|} \right) \\
&= ne^{+}(U) - \sum_{x\in[0,n]} \left( \sum_{A\ni x, A\cap[0,n]^c\neq\emptyset} \frac{U(A,+)}{|A|} \right). \qquad (4.57)
\end{aligned}
$$

Therefore, it is enough to show the existence of the limit

$$
\lim_{n\to\infty} \sum_{x\in[0,n]} \left( \sum_{A\ni x, A\cap[0,n]^c\neq\emptyset} \frac{U(A,+)}{|A|} \right). \qquad (4.58)
$$

As we will see later, cf. (4.63) the sum

$$
\sum_{A\ni 0} \left( \sum_{x\in[0,n], A\cap[-x,n-x]^c\neq\emptyset} \frac{|U(A,+)|}{|A|} \right) \qquad (4.59)
$$

is finite, uniformly in $n$. Therefore, we can interchange the sums in (4.58) and prove the existence of the limit

$$
\lim_{n\to\infty} \sum_{A\ni 0} \left( \sum_{x\in[0,n], A\cap[-x,n-x]^c\neq\emptyset} \frac{|U(A,+)|}{|A|} \right). \qquad (4.60)
$$

which can be rewritten as

$$
\lim_{n\to\infty} \sum_{A\ni 0} \frac{U(A,+)}{|A|} \sum_{0\leq x\leq n} I(\min(A) < -x \text{ or } (n-x) < \max(A)). \qquad (4.61)
$$

Clearly,

$$
|\{0 \leq x \leq n : \min(A) < -x \text{ or } (n-x) < \max(A)\}| \leq \mathrm{diam}(A). \qquad (4.62)
$$

so we obtain the uniform estimate

$$
\begin{aligned}
&\sum_{A\ni 0} \frac{|U(A,+)|}{|A|} \sum_{0\leq x\leq n} I(\min(A) < -x \text{ or } (n-x) < \max(A)) \leq \\
&\qquad \leq \sum_{A\ni 0} \mathrm{diam}(A)\|U(A)\|_\infty, \qquad (4.63)
\end{aligned}
$$

as we claimed before. Therefore, it is now enough to show that the limit

$$\lim_{n\to\infty} |\{0 \le x \le n : \min(A) < -x \text{ or } (n-x) < \max(A)\}| \tag{4.64}$$

exists for any finite subset $A \subset \mathbb{Z}$. This follows easily, since

$$|\{0 \le x \le n : \min(A) < -x \text{ or } (n-x) < \max(A)\}| = \operatorname{diam}(A) \tag{4.65}$$

as soon as $n > \operatorname{diam}(A)$. □

## 5 Additional Remarks

1. The choice of $\pm$ spins can be easily generalized to spins taking values in a finite alphabet. As long as we choose interactions satisfying the summability condition of Definition 2.1, Theorems 2.24 and 2.27 hold once one introduces the obvious modifications of the intervals $I_k$ and $\mathcal{I}_k$.

2. The extra randomization which we introduced in order to make the interval-lengths $|I_k|$ into continuous random variables $|\mathcal{I}_k|$ used the exponential distribution, i.e.,

$$\mathcal{I}_k = \sum_{j\in I_k} \xi_j. \tag{5.1}$$

   If for $\xi_j$ we choose i.i.d. random variables with distribution $F_\xi$ having an "exponential tail", then the same results hold. By "having an exponential tail" we mean here that the moment generating function of $\xi$ has to have a singularity of type $1/x$ when $x \simeq 1$, $x < 1$. In that case it is easily verified that for $X$ geometrically distributed with $P(X = n) = p^{n-1}(1-p)$ the moment generating function of

$$Z = \sum_{i=1}^{X} \xi_i \tag{5.2}$$

   has a singularity of type $1/x$ when $x \simeq 1-p$, $x < 1-p$.

**Acknowledgments.** I am grateful to Pablo Ferrari for useful discussions and warm hospitality at São Paulo University where this work was done. I also thank also C. Borgs for useful discussions on reference [2].

## References

[1] D. Aldous. *Probability Approximations via the Clumping Heuristic.* Springer, Berlin, (1987).

[2] M. Z. Bazant. *The largest cluster in subcritical percolation.* Preprint, available at the archive http://xxx.lanl.gov

[3] R. Fernandez, P.A. Ferrari and A. Galves. *Coupling, renewal and perfect simulation of chains of infinite order.* Notes for V Brazilian School of Probability, Ubatuba, August 2001. Available at http://www.ime.usp.br/∼pablo/publications.htm

[4] J. Galambos. *The Asymptotic Theory of Extreme Order Statistics.* Wiley, New-York, (1978).

[5] H.-O. Georgii. *Gibbs Measures and Phase Transitions*. Walter de Gruyter & Co., Berlin, (1988).

[6] M. R. Leadbetter, G. Lindgren and H. Rootzén. *Extremes and Related Properties of Random Sequences and Processes.* Springer-Verlag, New-York, Heidelberg, Berlin, (1982).

[7] G. L. O'Brien. *Limit theorems for the maximum term of a stationary process.* Ann. Prob. **2** (1974), 103–111.

**F. Redig**
Faculteit Wiskunde en Informatica
Technische Universiteit Eindhoven
Postbus 513, 5600 MB Eindhoven
THE NETHERLANDS