Exponential distribution for the occurrence of rare patterns in Gibbsian random fields

M. Abadi [*]
J.-R. Chazottes [†]
F. Redig [‡]
E. Verbitskiy [§]

October 21, 2003

Abstract: We study the distribution of the occurrence of rare patterns in sufficiently mixing Gibbs random fields on the lattice $\mathbb{Z}^d$, $d \geq 2$. A typical example is the high temperature Ising model. This distribution is shown to converge to an exponential law as the size of the pattern diverges. Our analysis not only provides this convergence but also establishes a precise estimate of the distance between the exponential law and the distribution of the occurrence of finite patterns. A similar result holds for the repetition of a rare pattern. We apply these results to the fluctuation properties of occurrence and repetition of patterns: We prove a central limit theorem and a large deviation principle.

Key-words: occurrence of patterns, repetition of patterns, exponential law, high temperature Gibbs random fields, non-uniform mixing, entropy, relative entropy, central limit theorem, large deviations.

[*]IME-USP, cp 66281, 05315-970, So Paulo, SP, Brasil, abadi@ime.usp.br.
[†]CPhT, CNRS-Ecole polytechnique, 91128 Palaiseau Cedex, France, jeanrene@cpht.polytechnique.fr
[‡]Faculteit Wiskunde en Informatica, Technische Universiteit Eindhoven, Postbus 513, 5600 MB Eindhoven, The Netherlands, f.h.j.redig@tue.nl
[§]Philips Research Laboratories, Prof. Holstlaan 4, 5656 AA Eindhoven, The Netherlands, evgeny.verbitskiy@philips.com

1 Introduction

In the last decade there has been an intensive study of exponential laws for rare events in the context of dynamical systems and stochastic processes, see e.g. the review paper [2]. In general, these laws are derived under the assumption of sufficiently strong mixing conditions, which basically ensures the possibility of writing the rare event as an intersection of almost independent events. The basic example of a rare event is the occurrence or return of a large cylindrical event. Other relevant examples are approximate cylindrical events (approximate matching in the sense of Hamming distance, see e.g. [8]), or large deviation events in certain interacting particle systems, see e.g. [3, 4].

The mixing conditions appearing in the context of dynamical systems or stochastic processes are typical for $\mathbb{Z}$-actions, e.g., the ψ-mixing condition is very naturally satisfied in the context of Bowen-Gibbs measures [5]. In turning to the context of random fields or $\mathbb{Z}^d$-actions, the ψ-mixing property is very restrictive and in many natural examples such as Gibbsian random fields, this property does not hold except (trivially) in the i.i.d. case and in non-interacting copies of one-dimensional Gibbs measures.

Gibbsian random fields have an obvious relevance to various applications, e.g., statistical physics, image processing, etc. Many interesting fluctuation properties such as large deviations principle, central limit theorems have been derived for them, and by now Gibbs measures constitute a well-established field of research, see e.g. [14], [16], [17].

The study of exponential laws for the occurrence or repetition of rare events in random fields has been initiated by A.J. Wyner [29] for the ψ-mixing case, using the Chen-Stein method. Because of the mixing condition, the results of that paper are not applicable to Gibbsian random fields like the Ising model in the high mixing regime (such as Dobrushin uniqueness, or analyticity regime).

As an example, consider the d-dimensional Ising model in the high-temperature regime and fix a pattern in a cubic box of size n: what is the size of the "observation window" in which we see this pattern for the first time ? This is clearly a rare event when the size of the pattern increases, and hence one expects in the "high mixing regime" that the size of this observation window is approximately exponentially distributed with parameter proportional to the probability of the pattern.

The main difficulty in making this intuition into a mathematical statement is caused by the typical non-uniform mixing of Gibbsian random fields: the influence of an event A on an event B is not only dependent on their distance but also on their size. More precisely, the difference between the conditional probabilities $\mathbb{P}(A|B)$ and $\mathbb{P}(A)$ can be estimated in the optimal situation of Dobrushin uniqueness regime as something of the form $|A|\exp(-\text{dist}(A,B))$. On a technical level, this "non-uniform mixing" implies that the rare event under consideration should be written as an intersection of events which at the same time are separated by a large distance and do not have an "excessive" size.

In this paper we concentrate on Gibbsian random fields in the Dobrushin uniqueness regime (e.g. high temperature case). This has to be considered as the first non-trivial test case for random fields, with a broad variety of examples. The regime of phase coexistence (such as in the low-temperature Ising model) poses an even larger non-uniformity in the mixing conditions, i.e., the difference between $\mathbb{P}(A|B)$ and $\mathbb{P}(A)$ will in that case also depend on which events B we are conditioning on. Recent techniques such as disagreement percolation constitute a powerful tool to tackle this situation. This is however not the subject of the present paper, where we want to deal with the basic non-uniformity in the mixing appearing in all non-trivial Gibbsian random fields.

Besides the mere derivation of exponential laws for the occurrence and repetition of rare events, we obtain a precise and uniform estimate of the error (i.e., the difference between the law and its exponential approximation). We show that obtaining this precise control of the error has many useful non-trivial applications in studying fluctuations of both "waiting times" and repetitions of rare patterns. The derivation of the exponential law is not via Chen-Stein method. Via a direct use of the (non-uniform) mixing we obtain more detailed information on the error term. The reason for that is that in the Chen-Stein method one gives an estimate of the variational distance between the "real counting process" and the Poisson process, whereas we only need one particular event. The precise estimation of the error turns to be crucial in the study of large deviations.

The problem of "waiting times" is to ask for the $\mathbb{P}$-typical size of the "observation window" in which a $\mathbb{Q}$-typical pattern occurs, where $\mathbb{P}$ is Gibbsian, and $\mathbb{Q}$ is any ergodic field. The logarithm of the size of this observation window properly normalized converges to the sum of the entropy of $\mathbb{Q}$ and the relative entropy density $s(\mathbb{Q}|\mathbb{P})$. To this "law of large numbers" we add precise large deviation estimates and a central limit theorem as a corollary of the exponential law with its precise error. The main point is that the exponential law provides an approximation of the logarithm of the waiting time by minus the logarithm of the probability of the corresponding pattern. For the cumulant generating function of the waiting times, we give an explicit expression in terms of the pressure. It coincides with the cumulant generating function of the probability of patterns in the interval $(-1, \infty)$ and is constant on $(-\infty, -1]$. A similar phenomenon was observed numerically for the cumulant generating function of the return times (that is in dimension one), see [18].

For repetition of patterns, we prove a similar exponential law with precise error bound. However, in that case we have to exclude "badly self-repeating" patterns, which have exponentially small probability for any Gibbs measure. As a corollary, we obtain a law of large numbers and a central limit theorem for repetitions. The large deviations are more subtle due to the presence of the bad patterns. We prove a full large deviation principle for the measure conditioned on good patterns, and a restricted large deviation principle for the full measure.

Our paper is organized as follows. In section 2 we give basic notations and definitions and state our main result and its corollaries. In section 3 we review basic properties of high-temperature Gibbs measures. Section 4 contains the proof of the exponential law for the occurrence of patterns, and section 5 is devoted to the derivation of its corollaries.

2 Definitions and results

We consider a random field $\{\sigma(\mathbf{x}) : \mathbf{x} \in \mathbb{Z}^d\}$ on the lattice $\mathbb{Z}^d$, $d \geq 2$, where $\sigma(\mathbf{x})$ takes values in a finite set $\mathcal{A}$. The joint distribution of $\{\sigma(\mathbf{x}) : \mathbf{x} \in \mathbb{Z}^d\}$ is denoted by $\mathbb{P}$. The configuration space $\Omega = \mathcal{A}^{\mathbb{Z}^d}$ is endowed with the product topology (making it into a compact metric space). The set of finite subsets of $\mathbb{Z}^d$ is denoted by $\mathcal{S}$. For $A, B \in \mathcal{S}$ we put $d(A, B) = \min\{|\mathbf{x} - \mathbf{y}| : \mathbf{x} \in A, \mathbf{y} \in B\}$, where $|\mathbf{x}| = \sum_{i=1}^d |x_i|$ ($\mathbf{x} = (x_1, x_2, ..., x_d)$). For $A \in \mathcal{S}$, $\mathcal{F}_A$ is the sigma-field generated by $\{\sigma(\mathbf{x}) : x \in A\}$. For $V \in \mathcal{S}$ we put $\Omega_V = \mathcal{A}^V$. For $\sigma \in \Omega$, and $V \in \mathcal{S}$, $\sigma_V \in \Omega_V$ denotes the restriction of σ to V. For $\mathbf{x} \in \mathbb{Z}^d$ and $\sigma \in \Omega$, $\tau_{\mathbf{x}}\sigma$ denotes the translation of σ by $\mathbf{x}$: $\tau_{\mathbf{x}}\sigma(\mathbf{y}) = \sigma(\mathbf{x} + \mathbf{y})$. For an event $E \subseteq \Omega$ the dependence set of E is the minimal $A \in \mathcal{S}$ such that E is $\mathcal{F}_A$ measurable. For any $n \in \mathbb{N}$ let $C_n = [0,n]^d \cap \mathbb{Z}^d$. An element $A_n \in \Omega_{C_n}$ is called a n-**pattern** or a pattern of size n.

Definition 2.1 (First occurrence of a pattern). *For every configuration $\sigma \in \Omega$ we define $\mathbf{t}_{A_n}(\sigma)$ to be the first occurrence of an n-pattern A_n in that configuration, that is the minimal $k \in \mathbb{N}$ such that there exists a non-negative vector $\mathbf{x} = (x_1, \ldots, x_d) \in \mathbb{Z}_+^d$ with $x_i \le k$, $i = 1, ..., d$, $|\mathbf{x}| > 0$, satisfying*

$$(\tau_{\mathbf{x}}\sigma)_{C_n} = A_n. \tag{2.2}$$

If such a vector x does not exist then we put $\mathbf{t}_{A_n}(\sigma) = \infty$.

We now come to the mixing hypothesis we make on our random fields. For $m > 0$ define

$$\varphi(m) = \sup \frac{1}{|A_1|} \mid \mathbb{P}\left(E_{A_1}|E_{A_2}\right) - \mathbb{P}\left(E_{A_1}\right) \mid, \tag{2.3}$$

where the supremum is taken over all finite subsets A_1, A_2 of $\mathbb{Z}^d$, with $d(A_1, A_2) \ge m$ and $E_{A_i} \in \mathcal{F}_{A_i}$, with $\mathbb{P}(E_{A_2}) > 0$. Note that this $\varphi(m)$ differs from the usual φ-mixing function since we divide by the size of the dependence set of the event E_{A_1}.

Definition 2.4. *A random field is* **non-uniformly exponentially φ-mixing** *if there exist constants $C_1, C_2 > 0$ such that*

$$\varphi(m) \le C_1 e^{-C_2 m} \quad \text{for all} \quad m > 0. \tag{2.5}$$

The examples that motivate this definition are Gibbsian random fields in the Dobrushin uniqueness regime (see Definition 3.8 below and examples thereafter). We leave their definition and properties till the next section. For a pattern $A_n \in \Omega_{C_n}$ we define the corresponding cylinder $\mathcal{C}(A_n)$ as

$$\mathcal{C}(A_n) = \{\sigma \in \Omega : \sigma_{C_n} = A_n\}.$$

Our main result reads:

Theorem 2.6. *For a translation-invariant Gibbs random field satisfying* (2.5)*, there exist strictly positive constants $C, c, \rho, \Lambda_1, \Lambda_2$, $\Lambda_1 \le \Lambda_2$, such that for any n and any n-pattern A_n, there exists $\lambda_{A_n} \in [\Lambda_1, \Lambda_2]$, such that*

$$\left| \; \mathbb{P}\left\{ \mathbf{t}_{A_n} > \Big(\frac{t}{\lambda_{A_n}\mathbb{P}\left(\mathcal{C}(A_n)\right)} \Big)^{1/d} \right\} - e^{-t} \; \right| \le \; C \; \mathbb{P}\left(\mathcal{C}(A_n)\right)^{\rho} \; e^{-ct} \tag{2.7}$$

for any $t > 0$.

Notice that $\mathbb{P}(\mathcal{C}(A_n))$ in the "error term" in (2.7) is bounded above by $\exp(-c' n^d)$, with $c' > 0$, by the Gibbs property, see (3.15).

The proof of this theorem is given in Section 4.

Remark 2.8. *The only results we are aware of in the context of random fields appeared in [29]. The results of that paper are valid under the assumption of a much stronger mixing condition than ours, namely ψ-mixing. Most Gibbs random fields (including the Ising model at high temperature) cannot satisfy such a property. As an examples of ψ-mixing Gibbsian* *random fields (in the sense of Wyner) on $\mathbb{Z}^2$, one can consider independent copies of a one-dimensional Markov chain, this gives a two-dimensional Gibbsian random field, but without interaction in the y-direction.*

From the technical point of view, Wyner uses the Chen-Stein method. This leads to an estimate which for fixed pattern size does not converge to zero as $t \to \infty$. Here we use a different approach allowing us to get a control in t in (2.7). This feature will turn to be fundamental when we prove large deviations for waiting times, see below.

From the proof of Theorem 2.6 it will be clear that we can generalize it to $(A_n)_n$'s that are finite patterns supported on a van Hove sequence of subsets of $\mathbb{Z}^d$.

We will show elsewhere how to prove an analog of Theorem 2.6 in order to obtain the same kind of result for the low temperature "plus phase" of the Ising model, where the mixing condition of Definition 2.4 is no longer satisfied.

We now state a number of corollaries of the previous theorem. We first consider the repetition of patterns.

Definition 2.9 (First repetition of the initial pattern). *For every configuration $\sigma \in \Omega$ and for all $n \in \mathbb{N}$, we define the first repetition, denoted by $\mathbf{r}_n(\sigma)$, as the minimal $k \in \mathbb{N}$ such that there exist a vector $\mathbf{x} = (x_1, \ldots, x_d) \in \mathbb{Z}_+^d$, with $0 \le x_i \le k$ and $|\mathbf{x}| > 0$, satisfying*

$$(\tau_{\mathbf{x}}\sigma)_{C_n} = \sigma_{C_n} \ . \tag{2.10}$$

To obtain a similar result for the repetition times we have to exclude certain patterns with "too quick repetitions". We will make this notion precise later. The following result is established in Subsection 5.1.

Theorem 2.11. *For a translation-invariant Gibbs random field satisfying (2.5), there exist*

(i) a set G_n, which is a union of cylinders;

(ii) strictly positive constants B, b, C, c, ρ

such that for any $n \ge 1$

$$\mathbb{P}(G_n^c) \le B e^{-bn^d}, \tag{2.12}$$

and for each A_n with $\mathcal{C}(A_n) \subseteq G_n$

$$\left| \ \mathbb{P}\left\{ \mathbf{r}_n > \Big(\frac{t}{\lambda_{A_n} \mathbb{P}(\mathcal{C}(A_n))} \Big)^{1/d} \Big| \ \mathcal{C}(A_n) \right\} - e^{-t} \ \right| \le \ C \ \mathbb{P}(\mathcal{C}(A_n))^{\rho} \ e^{-ct} \tag{2.13}$$

for all $t > 0$ and where λ_{A_n} is given in Theorem 2.6.

Notice that the constants appearing in the previous Theorems may be different. Nevertheless we used the same notations for the sake of simplicity.

We denote by $s(\mathbb{P})$ the entropy of $\mathbb{P}$ (see the next section for the definition). The next result (proved in subsection 5.2) shows how the repetition of typical patterns allows to compute the entropy using a single "typical" configuration.

Theorem 2.14. *For a translation-invariant Gibbs random field satisfying* (2.5)*, there exists* $\epsilon_0 > 0$ *such that for all* $\epsilon > \epsilon_0$

$$-\epsilon \log n \le \log \left[(\mathbf{r}_n(\sigma))^d \ \mathbb{P}(\mathcal{C}(\sigma_{C_n})) \right] \le \log \log n^{\epsilon} \quad \text{eventually } \mathbb{P}-\text{almost surely.} \tag{2.15}$$

In particular,

$$\lim_{n\to\infty} \frac{d}{n^d} \log \mathbf{r}_n(\sigma) = s(\mathbb{P}) \quad \mathbb{P} - \text{a.s.} \tag{2.16}$$

Note that (2.16) is a particular case of the result by Ornstein and Weiss in [23] where $\mathbb{P}$ is only assumed to be ergodic. Under our assumptions, we get the more precise result (2.15).

We now consider the occurrence of an n-pattern drawn from some ergodic random field in the configuration drawn from a possibly different Gibbsian random field. This is the natural d-dimensional analog of the waiting-time [26], [29].

Definition 2.17 ("Waiting time"). *For all configurations* $\xi, \sigma \in \Omega$ *and for all* $n \in \mathbb{N}$*, we define the "waiting time", denoted by* $\mathbf{w}_n(\xi,\sigma)$*, as the minimal* $k \in \mathbb{N}$ *such that there exist a non-negative vector* $\mathbf{x} = (x_1, \ldots, x_d) \in \mathbb{Z}_+^d$*, with* $0 \le x_i \le k$ *and* $|\mathbf{x}| > 0$*, satisfying*

$$(\tau_{\mathbf{x}}\sigma)_{C_n} = \xi_{C_n} \ . \tag{2.18}$$

Notice that $\mathbf{w}_n(\xi,\sigma) = \mathbf{t}_{\xi_{C_n}}(\sigma)$. We are going to consider the situation when ξ is "randomly chosen" according to an ergodic random field $\mathbb{Q}$ and σ is "randomly chosen" according to a non-uniformly exponentially φ-mixing Gibbs random field $\mathbb{P}$, i.e. (ξ,σ) is drawn with respect to the product measure $\mathbb{Q}\times\mathbb{P}$. We denote by $s(\mathbb{Q}|\mathbb{P})$ the relative entropy of $\mathbb{Q}$ with respect to $\mathbb{P}$; see section 3 for the definition and a more explicit form. We have the following result (proved in Subsection 5.3):

Theorem 2.19. *For a translation-invariant Gibbs random field* $\mathbb{P}$ *satisfying* (2.5)*, and an ergodic random field* $\mathbb{Q}$*, there exists* $\epsilon_0 > 0$ *such that for all* $\epsilon > \epsilon_0$

$$-\epsilon \log n \le \log \left[(\mathbf{w}_n(\xi,\sigma))^d \ \mathbb{P}(\mathcal{C}(\xi_{C_n})) \right] \le \log \log n^{\epsilon} \tag{2.20}$$

for $\mathbb{Q}\times\mathbb{P}$*-eventually almost every* (ξ,σ)*. In particular*

$$\lim_{n\to\infty} \frac{d}{n^d} \log \mathbf{w}_n(\xi,\sigma) = s(\mathbb{Q}) + s(\mathbb{Q}|\mathbb{P}) \quad \mathbb{Q}\times\mathbb{P} - \text{a.s.} \tag{2.21}$$

Statement (2.21) is the d-dimensional generalization of a result obtained in [7] in the case of Bowen-Gibbs measures. Using Theorem 2.14 we can rewrite (2.21), for a "typical" pair (ξ,σ), as follows:

$$\mathbf{w}_n(\xi,\sigma) \approx \mathbf{r}_n(\xi) \exp((n^d/d) s(\mathbb{Q}|\mathbb{P})) \, .$$

(The measure $\mathbb{Q}$ is supposed to be Gibbsian or only ergodic if we invoke the Ornstein-Weiss theorem alluded to above.) This gives an interpretation of relative entropy in terms of repetition and waiting times.

We now turn to the analysis of fluctuations of occurrence and repetitions of patterns. In the sequel, U is the interaction defining the Gibbs measure $\mathbb{P}$ (see Section 3 below). The following two theorems are proved in Subsection 5.4.

Theorem 2.22. *Let U be a finite range, translation-invariant interaction, and for β small enough let $\mathbb{P}_\beta$ be the unique Gibbs measure with interaction βU. There exists $\beta_0 > 0$ such that for all $\beta < \beta_0$ there exists $\theta = \theta_\beta > 0$ such that*

$$\frac{\log \mathbf{w}_n - \mathbb{E}(\log \mathbf{w}_n)}{n^{\frac{d}{2}}} \to \mathcal{N}(0,\theta^2) \ , \ \textit{as } n \to \infty, \ \textit{in } \mathbb{P}_\beta \times \mathbb{P}_\beta \ \textit{ distribution} \ . \tag{2.23}$$

where $\mathcal{N}(0,\theta^2)$ denotes the normal law with mean zero and variance θ^2, which is equal to

$$\frac{d^2}{dq^2} \left(P((1-q)\beta U) \right) \big|_{q=0} \, . \tag{2.24}$$

Theorem 2.25. *Let U be a finite range, translation-invariant interaction, and for β small enough let $\mathbb{P}_\beta$ be the unique Gibbs measure with interaction βU. There exists $\beta_0 > 0$ such that for all $\beta < \beta_0$ there exists $\theta = \theta_\beta > 0$ (the same as in the previous theorem) such that*

$$\frac{\log \mathbf{r}_n - \mathbb{E} \log \mathbf{r}_n}{n^{\frac{d}{2}}} \to \mathcal{N}(0,\theta^2) \ , \ \text{a.s. } n \to \infty, \ \textit{in } \mathbb{P}_\beta \ \textit{distribution}. \tag{2.26}$$

Remark 2.27. *From the proof of the previous theorem it follows that one can replace the measure $\mathbb{P}_\beta \times \mathbb{P}_\beta$ by the measure $\mathbb{Q} \times \mathbb{P}_\beta$, where $\mathbb{Q}$ is any ergodic random field, and $s(\mathbb{P}_\beta)$ by $s(\mathbb{Q}) + s(\mathbb{Q}|\mathbb{P}_\beta)$.*

Remark 2.28. *The β_0 of Theorems 2.25 and 2.22 determines the analyticity regime of the pressure. This is related to the regime where the high-temperature expansion is convergent. The restriction to finite range interactions is here for convenience only, and can be replaced by the requirement that the norm*

$$\|U\| = \sum_{A \ni 0} \|U(A,\cdot)\| \ \exp\left(\alpha(\mathrm{diam}(A))\right)$$

is finite for some $\alpha > 0$, see [27].

We end our corollaries with large deviation estimates. In the context of Gibbs measures, it is well-known that the sequence $\{-\frac{1}{n^d} \log \mathbb{P}(\mathcal{C}(\sigma_{C_n})) : n \in \mathbb{N}\}$ satisfies a large deviation principle see e.g., [10], [22]. Here we shall apply the more specific large deviation result of [24] that was already used in [9] to establish large deviations for $\log \mathbf{r}_n$ (in dimension one).

The following theorem is proved in subsection 5.5.

Theorem 2.29. *Let $\mathbb{P}$ be a translation-invariant Gibbs random field satisfying* (2.5)*. Then for all $q \in \mathbb{R}$ the limit*

$$\mathcal{W}(q) = \lim_{n\to\infty} \frac{1}{n^d} \log \int \mathbf{w}_n^{qd} \ d\mathbb{P}\times\mathbb{P} \tag{2.30}$$

exists. Moreover,

$$\mathcal{W}(q) = \begin{cases} P\left((1-q)U\right) + (q-1)P(U), & \textit{for } q \geq -1, \\ P(2U) - 2P(U), & \textit{for } q < -1, \end{cases} \tag{2.31}$$

where P is the pressure defined in (3.13) below.

The following theorem gives the precise consequence of Theorem 2.29 for the large deviations of $\log \mathbf{w}_n$ provided $P((1-q)U)$ is C^1 for all $q \geq -1$. For this we can apply the result of [24]. The pressure function is C^1 for example in the Ising model. In the case $P((1-q)U)$ is not differentiable everywhere on $[-1,\infty)$, the result of [24] will give us Large Deviations for u in some bounded interval.

Theorem 2.32. *Suppose U is a finite range translation-invariant interaction. Then there exists $\beta_1 > 0$ such that for $\beta \leq \beta_1$ there exists a unique Gibbs measure $\mathbb{P}_\beta$ with interaction βU, and for all $u \geq 0$ we have*

$$\lim_{n\to\infty} \frac{1}{n^d} \log\,(\mathbb{P}_\beta \times \mathbb{P}_\beta)\left(\frac{\log \mathbf{w}_n^d}{n^d} \geq s(\mathbb{P}_\beta)+u\right) = \inf_{q>-1} \{-(s(\mathbb{P}_\beta)+u)q + \mathcal{W}(q)\} \tag{2.33}$$

and for all $u \in (0,u_0)$, $u_0 = |\lim_{q\downarrow -1} \mathcal{W}'(q) - s(\mathbb{P})|$,

$$\lim_{n\to\infty} \frac{1}{n^d} \log\,(\mathbb{P}_\beta \times \mathbb{P}_\beta)\left(\frac{\log \mathbf{w}_n^d}{n^d} \leq s(\mathbb{P}_\beta)-u\right) = \inf_{q>-1} \{-(s(\mathbb{P}_\beta)-u)q + \mathcal{W}(q)\} \tag{2.34}$$

Remark 2.35. *A more general version of Theorem 2.29 can be easily deduced by following the same lines as its proof: The measure $\mathbb{P}\times\mathbb{P}$ can be replaced by the measure $\mathbb{Q}\times\mathbb{P}$ where $\mathbb{Q}$ is any Gibbsian random field (without any mixing assumption). Of course formula 2.31 has to be modified: Now $\mathcal{W}(q) = P(V-qU) - P(V) + qP(V)$ for $q \geq -1$, where V is the interaction of the Gibbs measure $\mathbb{Q}$. Accordingly, a version of Theorem 2.32 can be obtained under a differentiability condition on $\mathcal{W}$.*

Remark 2.36. *Under the assumption of Theorem 2.29, the sequence $\left\{\frac{d}{n^d}\log \mathbf{w}_n\right\}$ satisfies a Large Deviation Principle in the sense of [12] (Theorem 4.5.20 p. 157).*

The following theorem derives from Theorem 2.11. Since its derivation follows verbatim along the lines of [9], we omit the proof.

Theorem 2.37. *Suppose U is a finite range, translation-invariant interaction. There exists $\beta_1 > 0$ be such that for $\beta \leq \beta_1$ there exists a unique Gibbs measure $\mathbb{P}_\beta$ with interaction βU and there exists $\tilde{u} > 0$ such that for all $u \in [0,\tilde{u})$ we have*

$$\lim_{n\to\infty} -\frac{1}{n^d} \log \mathbb{P}_\beta\left(\frac{\log \mathbf{r}_n^d}{n^d} \geq s(\mathbb{P}_\beta)+u\right) = \mathcal{I}(s(\mathbb{P}_\beta)+u), \tag{2.38}$$

and

$$\lim_{n\to\infty} -\frac{1}{n^d} \log \mathbb{P}_\beta\left(\frac{\log \mathbf{r}_n^d}{n^d} \leq s(\mathbb{P}_\beta)-u\right) = \mathcal{I}(s(\mathbb{P}_\beta)-u), \tag{2.39}$$

where

$$\mathcal{I}(u) = \sup_{q\in\mathbb{R}} (uq - P((1-q)U) - (q-1)P(U))$$

Remark 2.40. *It follows from the proof of theorem 2.32 that we have the analogue theorem for repetition times, if we condition the measure $\mathbb{P}_\beta$ on good patterns, that is, patterns which are not "badly self-repeating", see Definition 5.6 below.*

Remark 2.41. *β_1 in Theorems 2.32 and 2.37 does not necessarily coincide with the critical inverse temperature β_c (below which there is a unique Gibbs measure), and is in general strictly larger than β_0 of Theorem 2.25 and 2.22, see [15].*

3 Gibbsian random fields and Dobrushin uniqueness

For the sake of convenience the present and next subsections are devoted to the notion of Gibbsian random fields and their mixing properties. More details on this subject can be found in [16], [17].

Definition 3.1. *A translation-invariant interaction is a function*

$$U : \mathcal{S} \times \Omega \to \mathbb{R}, \tag{3.2}$$

such that the following conditions are satisfied:

1. *$U(A, \sigma)$ depends on $\sigma(\mathbf{x})$, with $\mathbf{x} \in A$ only.*
2. *Translation invariance:*

$$U(A + \mathbf{x}, \tau_{-\mathbf{x}}\sigma) = U(A, \sigma) \qquad \forall A \in \mathcal{S}, \mathbf{x} \in \mathbb{Z}^d, \sigma \in \Omega. \tag{3.3}$$

3. *Uniform summability:*

$$\sum_{A \ni 0} \sup_{\sigma \in \Omega} |U(A, \sigma)| < \infty \ . \tag{3.4}$$

An interaction U is called *finite-range* if there exists an $R > 0$ such that $U(A, \sigma) = 0$ for all $A \in \mathcal{S}$ with $\mathrm{diam}(A) > R$.

The set of all such interactions is denoted by $\mathcal{U}$. Mostly we will give examples of Gibbs measures satisfying our mixing conditions with interactions $U \in \mathcal{U}$. This can be generalized easily to interactions such that

$$\|U\|_\alpha = \sum_{A \ni 0} \|U(A, \cdot)\| \ \exp\left(\alpha(\mathrm{diam}(A))\right)$$

is finite for some $\alpha > 0$.

For $U \in \mathcal{U}$, $\zeta \in \Omega$, $\Lambda \in \mathcal{S}$, we define the finite-volume Hamiltonian with boundary condition ζ as

$$H_\Lambda^\zeta(\sigma) = \sum_{A \cap \Lambda \neq \varnothing} U(A, \sigma_\Lambda \zeta_{\Lambda^c}) \, . \tag{3.5}$$

Corresponding to the Hamiltonian in (3.5) we have the finite-volume Gibbs measures $\mathbb{P}_\Lambda^{U,\zeta}$, $\Lambda \in \mathcal{S}$, defined on Ω by

$$\int f(\xi) \ d\mathbb{P}_\Lambda^{U,\zeta}(\xi) = \sum_{\sigma_\Lambda \in \Omega_\Lambda} f(\sigma_\Lambda \zeta_{\Lambda^c}) \ e^{-H_\Lambda^\zeta(\sigma)} / Z_\Lambda^\zeta \, , \tag{3.6}$$

where f is any continuous function and Z_Λ^ζ denotes the partition function normalizing $\mathbb{P}_\Lambda^{U,\zeta}$ to a probability measure. Because of the uniform summability condition, (3.4) the objects H_Λ^ζ and $\mathbb{P}_\Lambda^{U,\zeta}$ are continuous as functions of the boundary condition ζ.

For a probability measure $\mathbb{P}$ on Ω, we denote by $\mathbb{P}_\Lambda^\zeta$ the conditional probability distribution of $\sigma(\mathbf{x}), \mathbf{x} \in \Lambda$, given $\sigma_{\Lambda^c} = \zeta_{\Lambda^c}$. Of course, this object is only defined on a set of $\mathbb{P}$-measure one. For $\Lambda \in \mathcal{S}, \Gamma \in \mathcal{S}$ and $\Lambda \subseteq \Gamma$, we denote by $\mathbb{P}_\Gamma(\sigma_\Lambda | \zeta)$ the conditional probability to find σ_Λ inside Λ, given that ζ occurs in $\Gamma \setminus \Lambda$.

For $U \in \mathcal{U}$, we call $\mathbb{P}$ a Gibbs measure with interaction U if its conditional probabilities coincide with the ones prescribed by (3.6), i.e., if

$$\mathbb{P}_\Lambda^\zeta = \mathbb{P}_\Lambda^{U,\zeta} \qquad \mathbb{P}-a.s. \qquad \Lambda \in \mathcal{S}, \zeta \in \Omega. \tag{3.7}$$

We denote by $\mathcal{G}(U)$ the set of all translation invariant Gibbs measures with interaction U. For any $U \in \mathcal{U}$, $\mathcal{G}(U)$ is a non-empty compact convex set. In this paper we will in fact restrict ourselves to interactions with a unique Gibbs measure.

A basic example is the ferromagnetic Ising model, where $U(\{\mathbf{x},\mathbf{y}\},\sigma) = -\beta J\sigma(\mathbf{x})\sigma(\mathbf{y})$ if $|\mathbf{x}-\mathbf{y}| = 1$, $U(\{\mathbf{x}\},\sigma) = -h\beta\sigma(\mathbf{x})$. Here $\beta \in (0,\infty)$ represents the inverse temperature, $J > 0$ the coupling strength, and h the external magnetic field.

We turn to the mixing properties of Gibbs random fields. For an interaction $U \in \mathcal{U}$, the Dobrushin matrix is given by

$$\gamma_{\mathbf{xy}}(U) = \frac{1}{2}\sup\left\{|\ \mathbb{P}^{U,\zeta}_{\{\mathbf{x}\}}(\alpha) - \mathbb{P}^{U,\xi}_{\{\mathbf{x}\}}(\alpha)| :\ \zeta,\xi \in \Omega, \zeta_{\mathbb{Z}^d\setminus\{\mathbf{y}\}} = \xi_{\mathbb{Z}^d\setminus\{\mathbf{y}\}}\ ,\ \alpha \in \mathcal{A}\right\}\ .$$

The matrix γ measures the dependence of changing the spin at site $\mathbf{y}$ on the conditional probability at site $\mathbf{x}$.

Definition 3.8. *The interaction U is said to satisfy the Dobrushin uniqueness condition if*

$$\sup_{\mathbf{x}\in\mathbb{Z}^d}\sum_{\mathbf{y}\in\mathbb{Z}^d}\gamma_{\mathbf{xy}}(U) < 1\,. \tag{3.9}$$

The following result is proved in [16], see also [17], theorem 2.1.3, p. 52.

Theorem 3.10. *Let $U \in \mathcal{U}$ be a finite range interaction. Under the condition (3.9), there is a unique Gibbs measure $\mathbb{P} \in \mathcal{G}(U)$, and this $\mathbb{P}$ is non-uniformly exponentially φ-mixing, i.e., it satisfies the mixing property (2.5).*

Examples for which (3.9) is satisfied are:

1. The so-called high-temperature region where. $U \in \mathcal{U}$ is such that

$$\sup_{\mathbf{x}\in\mathbb{Z}^d}\sum_{A\ni\mathbf{x}}(|A|-1)\sup_{\sigma,\sigma'\in\Omega}|U(A,\sigma) - U(A,\sigma')| < 2. \tag{3.11}$$

   Inequality (3.11) implies the Dobrushin uniqueness condition (3.9) (see [16], p. 143, Proposition 8.8). In particular, it implies that $|\mathcal{G}(U)| = 1$ (i.e., no phase transition). Note that it is independent of the "single-site part" of the interaction, i.e., of the interactions $U(\{\mathbf{x}\},\sigma)$. For any finite range potential U there exists β_c such that βU satisfies (3.11) for all $\beta < \beta_c$. For the Ising model in $\mathbb{Z}^2$, much more is known: the mixing property (2.5) holds for *any* $\beta < \beta_c$ (see e.g. [13]).

2. Low temperature regime for an interaction with unique ground state, e.g., the Ising model in a homogeneous magnetic field and sufficiently large β. See [17] example (2.1.5)

3. Interactions in a large external field. See [17], example (2.1.4). For the Ising model in two dimensions this means that the field h should satisfy

$$|h| > 4\beta + \log(8\beta)\,.$$

Remark 3.12. *The Dobrushin uniqueness condition is not a necessary condition for the mixing property 2.4. More general versions, known as "Dobsruhin-Shlosman" conditions exist, see e.g., [21] for more details on general finite size conditions ensuring NUEM.*

We now recall some basic facts on entropy and relative entropy (or Kullback-Leibler information). We use the following shorthand to ease notation :

$$\sum_{\mathcal{C}_n} = \sum_{\mathcal{C}(A_n):A_n\in\Omega_{C_n}} .$$

The entropy $s(\mathbb{P})$ of $\mathbb{P}$ is defined as

$$s(\mathbb{P}) = \lim_{n\to\infty} -\frac{1}{n^d}\sum_{\mathcal{C}_n}\mathbb{P}(\mathcal{C}_n)\log\mathbb{P}(\mathcal{C}_n)\,.$$

The relative entropy $s(\mathbb{Q}|\mathbb{P})$ of a stationary random field $\mathbb{Q}$ with respect to a Gibbsian random field $\mathbb{P}$ is

$$s(\mathbb{Q}|\mathbb{P}) = \lim_{n\to\infty}\frac{1}{n^d}\sum_{\mathcal{C}_n}\mathbb{Q}(\mathcal{C}_n)\log\frac{\mathbb{Q}(\mathcal{C}_n)}{\mathbb{P}(\mathcal{C}_n)}$$

In terms of the interaction U of $\mathbb{P}$ the relative entropy is

$$s(\mathbb{Q}|\mathbb{P}) = P(U) + \int f_U\ d\mathbb{Q} - s(\mathbb{Q}),$$

where

$$f_U(\sigma) = \sum_{A\ni 0}\frac{U(A,\sigma)}{|A|}$$

and $P(U)$ is the pressure of U, which defined as follows

$$P(U) = \lim_{n\to\infty}\frac{1}{n^d}\log Z_{C_n}\,, \tag{3.13}$$

where

$$Z_{C_n} = \sum_{\sigma_{C_n}\in\Omega_{C_n}}\exp(-\sum_{A\subseteq C_n}U(A,\sigma))$$

is the partition function with the free boundary conditions.

Proposition 3.14. *Let $\mathbb{P}$ be a Gibbs random field and $\mathbb{Q}$ be an ergodic random field. Then*

$$\lim_{n\to\infty}\frac{1}{n^d}\log\frac{\mathbb{Q}(\mathcal{C}(\sigma_{C_n}))}{\mathbb{P}(\mathcal{C}(\sigma_{C_n}))} = s(\mathbb{Q}|\mathbb{P})$$

for $\mathbb{Q}$-almost every σ.

Proof. The proof is simple, but since we did not find it in the literature, we give it here for the sake of completeness. Write

$$g_n(\sigma)\sim h_n(\sigma)$$

if

$$\lim_{n\to\infty}\frac{1}{n^d}\sup_{\sigma}|g_n(\sigma)-h_n(\sigma)| = 0\,.$$

Let U be the potential of the Gibbsian field $\mathbb{P}$. Then we have

$$\log \mathbb{P}(\mathcal{C}(\sigma_{C_n})) \sim -\sum_{i\in C_n} \tau_i f_U(\sigma) - \log Z_{C_n} .$$

Therefore, by ergodicity of $\mathbb{Q}$

$$\frac{1}{n^d} \log \mathbb{P}(\mathcal{C}(\sigma_{C_n}))$$

converges $\mathbb{Q}$-a.s. to

$$-\int f_U d\mathbb{Q} - P(U) .$$

By the Shannon-Mc Millan-Breiman theorem [20, 28]

$$\frac{1}{n^d} \log \mathbb{Q}(\mathcal{C}(\sigma_{C_n}))$$

converges $\mathbb{Q}$-a.s. to $-s(\mathbb{Q})$. Hence the difference

$$\frac{1}{n^d}(\log \mathbb{Q}(\mathcal{C}(\sigma_{C_n})) - \log \mathbb{P}(\mathcal{C}(\sigma_{C_n})))$$

converges $\mathbb{Q}$-a.s. to

$$P(U) - \big(s(\mathbb{Q}) - \int f_U \; d\mathbb{Q}\big)$$

which is equal to $s(\mathbb{Q}|\mathbb{P})$ by the Gibbs variational principle, see[16]. □

A standard property of Gibbs measures which we will use often is the following: there exist positive constants C, c, C', c' such that

$$C' \; e^{-c'n^d} \leq \mathbb{P}(\mathcal{C}_n) \leq C \; e^{-cn^d} \tag{3.15}$$

for every cylinder $\mathcal{C}_n$ supported on C_n.

4 Proof of Theorem 2.6

To ease notation, we will write $\mathbb{P}(A)$ instead of $\mathbb{P}(\mathcal{C}(A))$ where $A = A_n$ is an n-pattern.

4.1 Preliminary results

In this section we prove Theorem 2.6. We follow the approach of [1].

For $V \in \mathcal{S}$, $\sigma \in \Omega$ and $A = A_n$ an n-pattern we say that "A is present in V", and write $A \prec V$, for the configuration σ if there exists $\mathbf{x} = (x_1, x_2, ..., x_d) \in \mathbb{Z}^d$ such that $W := \mathbf{x} + C_n \subseteq V$ and $(\tau_{\mathbf{x}}\sigma)_W = A$. By abusing notation, we will write $\mathbb{P}(A \prec V)$ for the probability of that event.

Lemma 4.1. *Let V be a finite subset of $\mathbb{Z}^d$, and let $A = A_n$ be a n-pattern. Then*

$$\mathbb{P}(A \prec V) \leq |V| \; \mathbb{P}(A).$$

Proof. $\mathbb{P}(A \prec V) \le \sum_{\mathbf{x}\in V} \mathbb{P}\big(\{\sigma :\, \sigma|_{\mathbf{x}+C_n} = A\}\big) = \sum_{\mathbf{x}\in V} \mathbb{P}(A) = |V|\; \mathbb{P}(A).$ □

For every $k \in \mathbb{N}$ define

$$N_k^A(\sigma) = \sum_{\substack{\mathbf{x}\in\mathbb{Z}^d \\ 0\le x_i\le k}} 1\!\!1\{\tau_{\mathbf{x}}(\sigma)_{C_n} = A\}.$$

Then the following events coincide:

$$\{\mathbf{t}_A \le k\} = \{N_k^A \ge 1\}. \tag{4.2}$$

Moreover,

$$\mathbb{E}N_k^A = (k+1)^d \mathbb{P}(A).$$

Lemma 4.3 (Second moment estimate). *Consider a non-uniformly exponentially φ-mixing Gibbsian random field. Then there exists $\delta > 0$ such that for every $n, k \in \mathbb{N}$, and every $\Delta > 2n$ one has*

$$\mathbb{E}(N_k^A)^2 \le (k+1)^d \mathbb{P}(A) \Big(1 + e^{-\delta n}\Delta^d + (k+1)^d \mathbb{P}(A) + (k+1)^d n^d \varphi(\Delta - 2n)\Big).$$

Proof. Define $C(\mathbf{x}, n) = \mathbf{x} + C_n$. We have to estimate the following expression

$$\mathbb{E}(N_k^A)^2 = \sum_{\substack{\mathbf{x}\in\mathbb{Z}^d \\ 0\le x_i\le k}} \sum_{\substack{\mathbf{y}\in\mathbb{Z}^d \\ 0\le y_i\le k}} \mathbb{P}(\sigma_{C(\mathbf{x},n)} = \sigma_{C(\mathbf{y},n)} = A). \tag{4.4}$$

We split the above double sum into the three following sums

$$I_1 = \sum_{\mathbf{x}=\mathbf{y}}, \quad I_2 = \sum_{\substack{\mathbf{x}\ne\mathbf{y} \\ |\mathbf{x}-\mathbf{y}|\le\Delta}}, \quad I_3 = \sum_{\substack{\mathbf{x}\ne\mathbf{y} \\ |\mathbf{x}-\mathbf{y}|>\Delta}}$$

Let us proceed with each of the sums separately. For I_1 one obviously has

$$I_1 = \sum_{\substack{\mathbf{x}\in\mathbb{Z}^d \\ 0\le x_i\le k}} \mathbb{P}(A) = (k+1)^d \mathbb{P}(A).$$

To estimate I_2 we use that for any Gibbsian random field there exists a constant $\delta > 0$ such that for any finite volume V, and any configuration σ and η, the conditional probability of observing σ on V, given η outside of V, can be estimated as follows [16]

$$\mathbb{P}(\sigma_V \,|\eta_{V^c}) \le \exp(-\delta |V|).$$

Therefore

$$\begin{aligned} I_2 = \sum_{\substack{\mathbf{x}\ne\mathbf{y} \\ |\mathbf{x}-\mathbf{y}|\le\Delta}} \mathbb{P}(\sigma_{C(\mathbf{x},n)} = \sigma_{C(\mathbf{y},n)} = A) &= \sum_{\substack{\mathbf{x}\ne\mathbf{y} \\ |\mathbf{x}-\mathbf{y}|\le\Delta}} \mathbb{P}\big(\sigma_{C(\mathbf{x},n)} = A \,\big|\, \sigma_{C(\mathbf{y},n)} = A\big)\; \mathbb{P}(A) \\ &\le \sum_{\substack{\mathbf{x}\ne\mathbf{y} \\ |\mathbf{x}-\mathbf{y}|\le\Delta}} \mathbb{P}(A) \exp\big(-\delta\; |C(\mathbf{x}, n) \setminus C(\mathbf{y}, n)|\big)\;. \end{aligned}$$

To complete the estimate, it is sufficient to observe that since $\mathbf{x} \neq \mathbf{y}$, the volume of the set $C(\mathbf{x}, n) \setminus C(\mathbf{y}, n)$ is at least n. Hence

$$I_2 \leq (k+1)^d \Delta^d \exp(-\delta n)\ \mathbb{P}(A).$$

Finally, using the mixing condition (2.5), for I_3 we obtain

$$\begin{aligned} I_3 &= \sum_{\substack{\mathbf{x}\neq\mathbf{y} \\ |\mathbf{x}-\mathbf{y}|>\Delta}} \mathbb{P}(\sigma_{C(\mathbf{x},n)} = \sigma_{C(\mathbf{y},n)} = A) \leq \sum_{\substack{\mathbf{x}\neq\mathbf{y} \\ |\mathbf{x}-\mathbf{y}|>\Delta}} \Big(\mathbb{P}(A) + n^d \varphi(\Delta - 2n)\Big)\ \mathbb{P}(A) \\ &\leq (k+1)^{2d}\ \mathbb{P}(A)\big(\mathbb{P}(A) + n^d\varphi(\Delta-2n)\big). \end{aligned}$$

Combining all the estimates together we obtain the statement of the lemma. □

Lemma 4.5 (The parameter). *There exist strictly positive constants Λ_1, Λ_2 such that for any integer t with $t\mathbb{P}(A) \leq 1/2$, one has*

$$\Lambda_1 \leq \lambda_{A,t} := -\frac{\log \mathbb{P}(\mathbf{t}_A > t^{1/d})}{t\mathbb{P}(A)} \leq \Lambda_2.$$

Proof. Taking into account (4.2) and the Cauchy-Schwartz inequality we obtain

$$\mathbb{P}(\mathbf{t}_A \leq k) \geq \frac{(\mathbb{E}N_k^A)^2}{\mathbb{E}(N_k^A)^2}. \tag{4.6}$$

We apply the basic inequalities

$$\frac{\kappa}{2} \leq 1 - e^{-\kappa} \leq \kappa\ , \tag{4.7}$$

where the left inequality is valid for all $\kappa \in [0,1]$, and the right inequality is true for $\kappa \geq 0$. Let now $\kappa = -\log \mathbb{P}(\mathbf{t}_A > t^{1/d})$. Then, using lemma 4.3 and (4.6), we conclude

$$\begin{aligned} \frac{-\log \mathbb{P}(\mathbf{t}_A > t^{1/d})}{t\mathbb{P}(A)} &\geq \frac{\mathbb{P}(\mathbf{t}_A \leq t^{1/d})}{t\mathbb{P}(A)} \\ &\geq \frac{1}{1 + e^{-\delta n}\Delta^d + (t+1)\mathbb{P}(A) + (t+1)n^d\varphi(\Delta-2n)} \\ &\geq \frac{1}{1 + c_1 + 3/2 + c_2} =: \Lambda_1, \end{aligned}$$

where we have chosen $\Delta = n^{d+1}$,

$$c_1 = \sum_{n\in\mathbb{N}} e^{-\delta n} n^{d(d+1)} < \infty, \text{ and } c_2 = \sup_{n\in\mathbb{N}} \Big\{(t+1)\, n^d \varphi(\Delta - 2n)\Big\}$$

We have to show that c_2 is finite. Indeed, since for a Gibbs random field $\mathbb{P}$ there exist $c', C' > 0$ such that

$$\mathbb{P}(A) \geq C' \exp(-c' n^d)$$

for every n-pattern A; t has been chosen such that $t\mathbb{P}(A) < 1/2$, we have

$$c_2 \leq \sup_{n\in\mathbb{N}} \left\{ \left[\frac{1}{2C'} \exp(c'n^d) + 1\right] n^d C_1 \exp\big(-C_2(n^{d+1} - 2n)\big) \right\} < \infty,$$

where we have used the mixing condition (2.5).

For the upper bound, we use (4.7) again, but first we have to check that

$$\kappa = -\log \mathbb{P}(\mathbf{t}_A > t^{1/d}) \in [0,1].$$

Indeed, since $t < (2\mathbb{P}(A))^{-1}$, by Lemma 4.1 we have

$$\mathbb{P}(\mathbf{t}_A > t^{1/d}) \geq \mathbb{P}\left(\mathbf{t}_A > \frac{1}{(2\mathbb{P}(A))^{1/d}}\right) = 1 - \mathbb{P}\left(\mathbf{t}_A \leq \frac{1}{(2\mathbb{P}(A))^{1/d}}\right) \geq 1 - \frac{\mathbb{P}(A)}{2\mathbb{P}(A)} = \frac{1}{2}.$$

Hence, $\kappa \leq \log(2) < 1$, therefore $\kappa \leq 2(1-e^{-\kappa})$, which means

$$-\log \mathbb{P}(\mathbf{t}_A > t^{1/d}) \leq 2\mathbb{P}(\mathbf{t}_A \leq t^{1/d}) \leq 2t\, \mathbb{P}(A) \leq 1$$

where we have used Lemma 4.1 for the second inequality. Hence, we can choose $\Lambda_2 = 2$. This finishes the proof. □

For positive numbers x_{A_n}, y_{A_n} depending on the n-pattern A_n we write $x_{A_n} \sim y_{A_n}$ if

$$\lim_{n\to\infty} \frac{x_{A_n}}{y_{A_n}} = 1.$$

For a positive integer t_A we set $C(t_A) = [0, t_A]^d \cap \mathbb{Z}^d$. For a subset $V \subseteq \mathbb{Z}^d$ let $A \not\prec V$ be the event that the n-pattern A cannot be found in V. (See above for the definition of $A \prec V$.)

The following lemma is crucial and gives the factorization property of the exponential distribution, i.e., the fact that asymptotically

$$\mathbb{P}(t_{A_n} > (t+s)/\mathbb{P}(A_n)) \sim \mathbb{P}(t_{A_n} > t/\mathbb{P}(A_n))\mathbb{P}(t_{A_n} > s/\mathbb{P}(A_n))$$

where the accuracy of the approximation marked $\sim$ is spelled out in detail. The idea is that the event of non-occurrence of the pattern in a cube of size $O(1/\mathbb{P}(A_n))$ can be viewed as the non-occurrence of the pattern in many sub-cubes of volume k_n, where $n^d << k_n << 1/\mathbb{P}(A_n)$. These sub-cubes will be separated by corridors of width Δ, where Δ is such that the pattern occurs with very small probability in the corridor, and on the other hand the mixing can be used to decouple the events of non-occurrence in different sub-cubes.

Our choice for the volume of the sub-cubes will be $k_n = O(\mathbb{P}(A_n)^{-\theta})$, with $\theta \in (0,1)$ and the corridors will have width $\Delta_n = O(n^k)$ with k big enough for the mixing to work well.

This explains the choices in the statement of the following lemma.

Lemma 4.8 (Iteration Lemma). *Let $A = A_n$ be a n-pattern and t_A be such that $t_A^d = [\mathbb{P}(A)^{-\vartheta}]$, where $[\cdot]$ denotes the integer part, and $\vartheta \in (0,1)$. For $i = 1, \dots k$, let $C_i(t_A)$ denote any collection of k disjoints cubes of the form $\mathbf{x}_i + C(t_A)$. Then, for n large enough, there exists $\delta \in (0,1)$, which depends only on the measure $\mathbb{P}$, such that the following inequality holds for all k:*

$$\left|\mathbb{P}\left(A \not\prec \bigcup_{i=1}^{k} C_i(t_A)\right) - \mathbb{P}\left(A \not\prec C(t_A)\right)^k\right| \leq k\mathbb{P}(A)^{\eta}\Big(\mathbb{P}\left(A \not\prec C(t_A)\right) + \mathbb{P}(A)^{\eta}\Big)^k, \tag{4.9}$$

where $\eta = (1 - \vartheta(d-1)/d)\,(1-\delta)$.

Proof. We will prove that

$$\mathbb{P}\left(A \not\prec \bigcup_{i=1}^{k} C_i(t_A)\right) \le \mathbb{P}\left(A \not\prec C(t_A)\right)^k + \mathbb{P}(A)^\eta \ k \ \left(\mathbb{P}\left(A \not\prec C(t_A)\right) + \mathbb{P}(A)^\eta\right)^k \tag{4.10}$$

The inequality

$$\mathbb{P}\left(A \not\prec \bigcup_{i=1}^{k} C_i(t_A)\right) \ge \mathbb{P}\left(A \not\prec C(t_A)\right)^k - \mathbb{P}(A)^\eta \ k \ \left(\mathbb{P}\left(A \not\prec C(t_A)\right) + \mathbb{P}(A)^\eta\right)^k$$

is derived analogously.

For any positive integer z, we write $\mathbf{z} = (z, \dots, z) \in \mathbb{Z}^d$. We denote by $C_i^z(t_A)$ the cube $\mathbf{x}_i + \mathbf{z} + C(t_A - 2z)$ and by $C^z(t_A)$ the cube $C(t_A - 2z)$. For any positive integer $\Delta < 2t_A$, we consider the difference

$$\begin{aligned}
&\left|\mathbb{P}\left(A \not\prec \bigcup_{i=1}^{k} C_i(t_A)\right) - \mathbb{P}\left(A \not\prec C_1^\Delta(t_A) \cup \bigcup_{i=2}^{k} C_i(t_A)\right)\right| \\
=\ & \mathbb{P}\left(\left(A \not\prec C_1^\Delta(t_A) \cup \bigcup_{i=2}^{k} C_i(t_A)\right) \ \cap \ \left(A \prec C_1(t_A) \backslash C_1^\Delta(t_A)\right)\right) \\
\le\ & \mathbb{P}\left(\left(A \not\prec \bigcup_{i=1}^{k} C_i^{2\Delta}(t_A)\right) \ \cap \ \left(A \prec C_1(t_A) \backslash C_1^\Delta(t_A)\right)\right).
\end{aligned}$$

Iterating the mixing property (2.3) and using Lemma 4.1, we bound the last term by

$$2d\Delta t_A^{d-1} \mathbb{P}(A) \ \left(\mathbb{P}\left(A \not\prec C^{2\Delta}(t_A)\right) + \varphi(\Delta) \ |C(t_A)| \ \right)^k .$$

On the other hand

$$\begin{aligned}
&\left|\mathbb{P}\left(A \not\prec C_1^\Delta(t_A) \cup \bigcup_{i=2}^{k} C_i(t_A)\right) - \mathbb{P}\left(A \not\prec C_1^\Delta(t_A)\right) \mathbb{P}\left(A \not\prec \bigcup_{i=2}^{k} C_i(t_A)\right)\right| \\
&\le \varphi(\Delta) \left|C_1^\Delta(t_A)\right| \ \mathbb{P}\left(A \not\prec \bigcup_{i=2}^{k} C_i(t_A)\right) \\
&\le \varphi(\Delta) t_A^d \left(\mathbb{P}\left(A \not\prec C^\Delta(t_A)\right) + \varphi(\Delta) \ |C(t_A)| \ \right)^{k-1} .
\end{aligned}$$

Put

$$\begin{aligned}
\epsilon_1 &= \epsilon_1(A, t_A, \Delta) = \varphi(\Delta) \ t_A^d, \\
\epsilon_2 &= \epsilon_2(A, t_A, \Delta) = 2d\Delta t_A^{d-1} \mathbb{P}(A), \\
\epsilon \ &= C(\epsilon_1 + \epsilon_2),
\end{aligned}$$

where C is a positive constant to be defined later on. Put also

$$\alpha_{k-j} = \mathbb{P}\left(A \not\prec \bigcup_{i=j+1}^{k} C_i(t_A)\right),$$

$$\alpha_{k-j}^z = \mathbb{P}\left(A \not\prec \bigcup_{i=j+1}^{k} C_i^z(t_A)\right).$$

We obtain the recursion

$$\alpha_k \le (\epsilon_1 + \epsilon_2)(\alpha_1^{2\Delta} + \epsilon_1)^{k-1} + \alpha_1^{\Delta}\alpha_{k-1} \ ,$$

which upon iteration leads to

$$\alpha_k \le (\epsilon_1 + \epsilon_2)\ k\ (\alpha_1^{2\Delta} + \epsilon_1)^{k-1} + (\alpha_1^{\Delta})^k.$$

We choose C such that, $\alpha_1^{2\Delta} + \epsilon_1 \le \alpha_1 + \epsilon$, so we have

$$\alpha_k \le \epsilon\ k\ (\alpha_1 + \epsilon)^{k-1} + (\alpha_1 + \epsilon)^k.$$

Now we use the following simple inequality: for $0 < x \le y < 1$ and N any positive integer

$$y^N - x^N = (y-x)(x^{N-1} + x^{N-2}y + \ldots + y^{N-1}) \le (y-x)Ny^{N-1}$$

to obtain

$$\alpha_k - \alpha_1^k \le 2\ \epsilon\ k\ (\alpha_1^{\Delta} + \epsilon_1)^{k-1} \le 4\ \epsilon\ k\ (\alpha_1 + \epsilon)^k\ . \tag{4.11}$$

Choose $\Delta = n^{d+1}$. Since $t_A^d \sim \mathbb{P}(A)^{-\vartheta}$ for some $\vartheta \in (0,1)$, and since we have exponential φ-mixing (2.5) we obtain for the "error terms" ϵ_1, ϵ_2:

$$\begin{aligned} \epsilon_1 &\sim e^{-c_1 n^{d+1}} e^{c_2 \vartheta n^d}, \\ \epsilon_2 &\sim 2dn^{d+1}\mathbb{P}(A)^{-\vartheta\frac{d-1}{d}}\mathbb{P}(A). \end{aligned}$$

This yields

$$\epsilon \le \mathbb{P}(A)^{\left(1-\vartheta\frac{d-1}{d}\right)(1-\delta)},$$

which together with (4.11) implies (4.10). □

4.2 Proof of Theorem 2.6

Let $t > 0$, and put $t = kf_A + r$, where $f_A = [1/(\mathbb{P}(A))^\gamma]$ ($\gamma \in (0,1)$, $[\cdot]$ denotes integer part), k is an integer and $r < f_A$. Put $t'_A = k[f_A]$, $t''_A = (k+1)[f_A]$, Without loss of generality we assume that the size n of a n-pattern A is sufficiently large, so $f_A\mathbb{P}(A) \sim \mathbb{P}(A)^{1-\gamma} < 1/2$. We remind that for n-patterns, Gibbs fields admit uniform estimates $\mathbb{P}(A) \le \exp(-cn^d)$ for some $c > 0$. Now, recall from Lemma 4.5 that

$$\lambda_A = -\frac{\log \mathbb{P}(\mathbf{t}_A > (f_A)^{1/d})}{f_A\mathbb{P}(A)} \in [\Lambda_1, \Lambda_2] \tag{4.12}$$

for some positive constants Λ_1, Λ_2. We also define

$$\tilde{\lambda}_A = -\frac{\log\left(\mathbb{P}(\mathbf{t}_A > (f_A)^{1/d}) + \mathbb{P}(A)\right)}{f_A\mathbb{P}(A)}.$$

It is not difficult to see that $\tilde{\lambda}_A \in [\Lambda_1/2, \Lambda_2]$, for n large enough.

Since $t'_A \le t \le t''_A$, one obviously has

$$\mathbb{P}(\mathbf{t}_A > t^{1/d}) - \exp(-\lambda_A\mathbb{P}(A)\ t) \ge \mathbb{P}(\mathbf{t}_A > (t''_A)^{1/d}) - \exp(-\lambda_A\mathbb{P}(A)\ t'_A),$$

and

$$\mathbb{P}(\mathbf{t}_A > t^{1/d}) - \exp(-\lambda_A\mathbb{P}(A)t) \le \mathbb{P}(\mathbf{t}_A > (t'_A)^{1/d}) - \exp(-\lambda_A\mathbb{P}(A)\ t''_A).$$

Now,

$$
\begin{aligned}
|\mathbb{P}(\mathbf{t}_A > (t'_A)^{1/d}) - \exp(-\lambda_A \mathbb{P}(A) t''_A)| \le\ & |\mathbb{P}(\mathbf{t}_A > (t'_A)^{1/d}\ ) - \mathbb{P}(\mathbf{t}_A > (f_A)^{1/d}\ )^k| \\
& + |\mathbb{P}(\mathbf{t}_A > (f_A)^{1/d}\ )^k - \exp(-\lambda_A \mathbb{P}(A)\ t'_A)| \\
& + |\exp(-\lambda_A \mathbb{P}(A)\ t'_A) - \exp(-\lambda_A \mathbb{P}(A)\ t''_A)|
\end{aligned}
$$

By Lemma 4.8,

$$
\begin{aligned}
|\mathbb{P}(\mathbf{t}_A > (t'_A)^{1/d}) - \mathbb{P}(\mathbf{t}_A > (f_A)^{1/d})^k| &\le \mathbb{P}(A)^{\gamma(1-\delta)/d}\ k\ \left(\mathbb{P}(\mathbf{t}_A > (f_A)^{1/d}) + \mathbb{P}(A)^{1-\gamma}\right)^k . \\
&= \mathbb{P}(A)^{\gamma(1-\delta)/d}\ t\ \mathbb{P}(A) \exp(-\tilde{\lambda}_A \mathbb{P}(A)\ t'_A) \\
&\le \mathbb{P}(A)^{\gamma(1-\delta)/d}\ t\ \mathbb{P}(A) \exp(-C_1 \mathbb{P}(A)\ t).
\end{aligned}
$$

By the choice of λ_A (4.12) , and since $t'_A = k f_A$,

$$
\mathbb{P}(\mathbf{t}_A > (f_A)^{1/d}\ )^k = \exp(-\lambda_A k f_A \mathbb{P}(A)) = \exp(-\lambda_A t'_A \mathbb{P}(A)).
$$

Finally,

$$
\begin{aligned}
|\exp(-\lambda_A \mathbb{P}(A)\ t'_A) - \exp(-\lambda_A \mathbb{P}(A)\ t''_A)| &\le \lambda_A\ \mathbb{P}(A)(t''_A - t'_A) \exp(-\lambda_A \mathbb{P}(A)\ t'_A) \\
&\le \Lambda_2\ \mathbb{P}(A) f_A \exp(-\Lambda_1 \mathbb{P}(A)\ t'_A) \\
&\le C_2\ \mathbb{P}(A)^{1-\gamma} \exp(-C_3 \mathbb{P}(A)\ t).
\end{aligned}
$$

The lower estimate is obtained in a similar way. This finishes the proof.

Remark 4.13. *Notice that in the iteration lemma it is not used that A_n is a pattern. Therefore, this lemma can be generalized to arbitrary measurable events $E_n \in \mathcal{F}_{C_{k_n}}$, where $k_n \lll 1/\mathbb{P}(E_n)$. The second moment estimate however uses that A_n is a pattern. Therefore Theorem 2.6 can be generalized as follows. Let $E_n \in \mathcal{F}_{C_{k_n}}$, where $|C_{k_n}| = O(n^\alpha)$, and $\mathbb{P}(E_n) = O(e^{-cn^d})$. Suppose furthermore that*

$$
\limsup_{n\to\infty} \sum_{0<|x|\le n^\alpha} \frac{\mathbb{P}(E_n \cap \theta_x E_n)}{\mathbb{P}(E_n)} < \infty \tag{4.14}
$$

then (2.7) holds for the occurrence time $\mathbf{t}_{E_n}$. Condition (4.14) takes care of the second moment estimate.

5 Proof of the other theorems

5.1 Proof of Theorem 2.11

We start with a lemma on "badly self-repeating" patterns.

Definition 5.1. *A pattern A_n is called badly self-repeating if there exists* $\mathbf{x}$, $0 < |\mathbf{x}| \le n/2$, *such that*

$$
\tau_{\mathbf{x}} \mathcal{C}(A_n) \cap \mathcal{C}(A_n) \neq \varnothing
$$

Correspondingly, a cylinder is called bad if it is of the form $\mathcal{C}(A_n)$ with A_n badly self-repeating. The union of bad n-cylinders is denoted by $\mathcal{B}_n$.

Lemma 5.2 (Conditioning on the initial pattern). *Let $A = A_n$ be a "good" pattern, that is, not a badly self-repeating pattern. Let t_A be such that $t_A^d \sim \mathbb{P}(A)^{-\vartheta}$, where $\vartheta \in (0,1)$. Then there exist positive constants b_1, b_2 such that for all integers $n \geq 1$, one has*

$$\left|\mathbb{P}\left(A \not\prec C(t_A)\backslash C_n \;\middle|\; A \prec C_n\right) - \mathbb{P}\left(A \not\prec C(t_A)\right)\right| \leq b_1\; e^{-b_2 n}\,.$$

Proof. We first observe that for any pattern A and any positive integer Δ such that $n+\Delta < t_A$, we have

$$\mathbb{P}\left(A \not\prec C(t_A)\backslash C_{n+\Delta})\right) - \mathbb{P}\left(A \not\prec C(t_A)\right) =$$
$$\mathbb{P}\left(A \prec C_{n+\Delta}\right) \leq (n+\Delta)^d\; \mathbb{P}(A)$$

where we used Lemma 4.1 to get the inequality. For the sake of convenience, "A is good" stands for $\forall \mathbf{x} \in \mathbb{Z}^d$ such that $0 < |\mathbf{x}| < n/2$, we have $(\tau_{\mathbf{x}}\sigma)_{C_n} \neq A$ for every $\sigma \in \mathcal{C}(A)$. Now we use that A is good to obtain

$$\mathbb{P}\left(A \not\prec C(t_A)\backslash C_{n+\Delta}, A\,\text{is good} \;\middle|\; A \prec C_n\right) - \mathbb{P}\left(A \not\prec C(t_A)\backslash C_n, A\,\text{is good} \;\middle|\; A \prec C_n\right) =$$
$$\mathbb{P}\left(\exists \mathbf{x}, n/2 < |\mathbf{x}| < n+\Delta \;:\; \sigma_{C_n + \mathbf{x}\backslash C_n} = P_n^{\mathbf{x}} \;\middle|\; A \prec C_n\right) \tag{5.3}$$

where $P_n^{\mathbf{x}}$ is a *fixed* pattern depending only on $A = A_n$ and $\mathbf{x}$. Using the Gibbs property we obtain

$$\begin{aligned}
(5.3) &\leq (n+\Delta)^d \sup_{|\mathbf{x}|>n/2} \sup_{\eta} \mathbb{P}\left(\sigma_{C_n+\mathbf{x}\backslash C_n} = P_n^{\mathbf{x}} \;\middle|\; \eta_{C_n} = A\right)\} \\
&\leq (n+\Delta)^d \sup_{|\mathbf{x}|>n/2} \exp(-c\;|C_n + \mathbf{x}\backslash C_n|) \\
&\leq (n+\Delta)^d\; \exp(-c' n^d)
\end{aligned}$$

where c, c' are positive constants. We now use the mixing property (2.3) to get, for any good pattern A :

$$\left|\mathbb{P}\left(A \not\prec C(t_A)\backslash C_{n+\Delta} \;\middle|\; A \prec C_n\right) - \mathbb{P}\left(A \not\prec C(t_A)\backslash C_{n+\Delta})\right)\right| \leq |C(t_A)|\; \varphi(\Delta)\,.$$

Putting together the above estimates, with the choice $\Delta = n^{d+1}$ and using (2.5), yields

$$\left|\mathbb{P}\left(A \not\prec C(t_A)\backslash C_n \;\middle|\; A \prec C_n\right) - \mathbb{P}\left(A \not\prec C(t_A)\right)\right| \leq$$
$$(n+n^{d+1})^d\; e^{-c'\;n^d} + C_1\; e^{c" n^d} e^{-C_2 n^{d+1}} + (n+n^{d+1})^d\; \mathbb{P}(A)\,.$$

This gives the desired result. □

We also need the following lemma.

Lemma 5.4 (Iteration Lemma for pattern repetitions). *Let t_A be such that $t_A^d \sim \mathbb{P}(A)^{-\vartheta}$, where $\vartheta \in (0,1)$. For $i = 2, \ldots k$, let $C_i(t_A)$ denote any collection of k disjoints cubes of the form $\mathbf{x}_i + C(t_A)$. Assume also that $C_1(t_A) = \mathbf{x}_1 + C(t_A)\backslash\{0\}$ is disjoint from $C_i(t_A)$, $i = 2, \ldots, k$. Then we have the following inequality for all k:*

$$\begin{aligned}
&\left|\mathbb{P}\left(A \not\prec \bigcup_{i=1}^{k} C_i(t_A) \mid A \prec C_n\right) - \mathbb{P}\left(A \not\prec C(t_A)\right)^k\right| \\
\leq\;& C_1 \exp\{-C_2 n\}\;\; \left(\mathbb{P}\left(A \not\prec C(t_A)\right) + C_1 \exp\{-C_2 n\}\right)^k\,.
\end{aligned}$$

Proof. Proceeding as in the proof of Lemma 4.8 we have:

$$\left|\mathbb{P}\left((A \prec C_n) \cap (A \not\prec \bigcup_{i=1}^{k} C_i(t_A))\right) - \mathbb{P}((A \prec C_n) \cap A \not\prec C(t_A)\backslash C_n)\ \mathbb{P}\left(A \not\prec C(t_A)\right)^{k-1}\right| \leq$$

$$\mathbb{P}(A)^{1-\vartheta}\ \left(\mathbb{P}\left(A \not\prec C(t_A)\right) + \mathbb{P}(A)^{1-\vartheta}\right)^k .$$

On the other hand, Lemma 5.2 tells us that

$$\left|\mathbb{P}(A \not\prec C(t_A)\backslash C_n | A \prec C_n) - \mathbb{P}(A \not\prec C(t_A))\right| \leq b_1\ e^{-b_2 n} .$$

□

The proof of (2.13) in Theorem 2.11 is now the same as that of Theorem 2.6. It remains to prove (2.12):

Lemma 5.5 (Probability of badly self-repeating patterns). *There exist* $c, C > 0$ *such that*

$$\mathbb{P}(\mathcal{B}_n) \leq Be^{-bn^d} \tag{5.6}$$

Proof. Put $C_n^+ = C_n \cap (C_n + \mathbf{x})$ and $C_n^+ = C_n \cap (C_n - \mathbf{x})$. By definition of $\mathcal{B}_n$, we have the inequality:

$$\mathbb{P}(\mathcal{B}_n) \leq \mathbb{P}\left(\exists \mathbf{x} : |\mathbf{x}| \leq n/2 : \sigma_{C_n^+(\mathbf{x})} = \sigma_{C_n^-(\mathbf{x})}\right) . \tag{5.7}$$

Define the event $E_{\mathbf{x}} = \{\sigma : \sigma_{C_n^+(\mathbf{x})} = \sigma_{C_n^-(\mathbf{x})}\}$. If $\sigma \in\ E_{\mathbf{x}}$, then there exists disjoint sets $S_n^+(\mathbf{x})$ and $S_n^-(\mathbf{x})$ such that $\sigma_{S_n^+(\mathbf{x})} = \sigma_{S_n^-(\mathbf{x})}$ and $|S_n^+(\mathbf{x})|, |S_n^-(\mathbf{x})| > \delta n^d$ for some positive δ. Therefore, we have

$$\begin{aligned}
\mathbb{P}(E_{\mathbf{x}}) &\leq \mathbb{P}\left(\sigma_{S_n^+(\mathbf{x})} = \sigma_{S_n^-(\mathbf{x})}\right) \\
&\leq \sup\left\{\mathbb{P}\left(\sigma_{S_n^+(\mathbf{x})} = \eta | \sigma_{(S_n^+(\mathbf{x}))^c} = \xi\right) : \eta \in \Omega_{S_n^+(\mathbf{x})}, \xi \in \Omega_{(S_n^+(\mathbf{x}))^c}\right\} \\
&\leq \exp(-c' n^d)
\end{aligned} \tag{5.8}$$

where in the last inequality we used the Gibbs property (3.15). Finally,

$$\mathbb{P}(\mathcal{B}_n) \leq \sum_{\mathbf{x}:|\mathbf{x}|<n/2} \mathbb{P}(E_{\mathbf{x}}) \leq Be^{-bn^d}. \tag{5.9}$$

□

5.2 Proof of Theorem 2.14

We start by showing the following summable upper-bound to

$$\mathbb{P}\{\sigma\ : \log(\mathbf{r}_n(\sigma)^d \mathbb{P}(\mathcal{C}(\sigma_{C_n}))) \geq \log t\} \leq$$

$$\sum_{\mathcal{C}_n \in \mathcal{B}_n^c} \mathbb{P}(\mathcal{C}_n)\ \mathbb{P}\{\sigma\ : \log(\mathbf{r}_n(\sigma)^d \mathbb{P}(\mathcal{C}_n)) \geq \log t \mid \mathcal{C}_n\} + \sum_{\mathcal{C}_n \in \mathcal{B}_n} \mathbb{P}(\mathcal{C}_n) .$$

From Theorem 2.11 and Lemma 5.5 we get for all $t > 0$

$$\mathbb{P}\{\sigma \ : \log(\mathbf{r}_n(\sigma)^d \mathbb{P}(\mathcal{C}(\sigma_{C_n}))) \geq \log t\} \leq (C' e^{-c' n^d} + e^{-\Lambda_1 t}) + C e^{-cn^d} \,.$$

Take $t = t_n = \log(n^\epsilon)$, $\epsilon > \Lambda_1^{-1}$, to get

$$\mathbb{P}\{\sigma \ : \log(\mathbf{r}_n(\sigma)^d \mathbb{P}(\mathcal{C}(\sigma_{C_n}))) \geq \log\log(n^\epsilon)\} \leq C' e^{-c' n^d} + \frac{1}{n^{\epsilon\Lambda_1}} + C e^{-cn^d} \,.$$

An application of the Borel-Cantelli lemma leads to

$$\log\Big[(\mathbf{r}_n(\sigma))^d \mathbb{P}(\mathcal{C}(\sigma_{C_n}))\Big] \leq \log\log(n^\epsilon) \quad \text{eventually a.s.}\,.$$

For the lower bound first observe that Theorem 2.11 gives, for all $t > 0$

$$\mathbb{P}\{\sigma \ : \log(\mathbf{r}_n(\sigma)^d \mathbb{P}(\mathcal{C}(\sigma_{C_n}))) \leq \log t\} \leq C' e^{-c' n^d} + (1 - \exp(-\Lambda_2 t)) + C e^{-cn^d} \,.$$

Choose $t = t_n = n^{-\epsilon}$, $\epsilon > 1$, to get, proceeding as before,

$$\log\Big[(\mathbf{r}_n(\sigma))^d \mathbb{P}(\mathcal{C}(\sigma_{C_n}))\Big] \geq -\epsilon \log n \quad \text{eventually a.s.}\,.$$

Finally, let $\epsilon_0 = \max(\Lambda_1^{-1}, 1)$.

5.3 Proof of Theorem 2.19

We first show that the strong approximation formula (2.15) holds with $\mathbf{w}_n$ in place of $\mathbf{r}_n$ with respect to the measure $\mathbb{Q} \times \mathbb{P}$. We have the following identity:

$$\int d\mathbb{Q}(\xi) \ \mathbb{P}\left\{\sigma : \mathbf{t}_{\xi_{C_n}}(\sigma) > \left(\frac{t}{\mathbb{P}(\mathcal{C}(\xi_{C_n}))}\right)^{1/d}\right\} =$$

$$(\mathbb{Q} \times \mathbb{P})\left\{(\xi, \sigma) : \mathbf{w}_n(\xi, \sigma) > \left(\frac{t}{\mathbb{P}(\mathcal{C}(\xi_{C_n}))}\right)^{1/d}\right\}$$

This shows immediately that Theorem 2.6 is valid with $\mathbf{w}_n(\xi, \sigma)$ in place of $\mathbf{t}_{\sigma_{C_n}}(\xi)$ and $\mathbb{Q} \times \mathbb{P}$ in place of $\mathbb{P}$, hence so is Theorem 2.14. Therefore for ϵ large enough, we obtain

$$-\epsilon \log n \leq \log\Big[(\mathbf{w}_n(\xi, \sigma))^d \mathbb{P}(\mathcal{C}(\xi_{C_n}))\Big] \leq \log\log n^\epsilon \tag{5.10}$$

for $\mathbb{Q} \times \mathbb{P}$-eventually almost every (ξ, σ). Write

$$\log\Big[(\mathbf{w}_n(\xi, \sigma))^d \mathbb{P}(\mathcal{C}(\sigma_{C_n}))\Big] = d \log \mathbf{w}_n(\xi, \sigma) + \log \mathbb{Q}(\mathcal{C}(\xi_{C_n})) - \log \frac{\mathbb{Q}(\mathcal{C}(\xi_{C_n}))}{\mathbb{P}(\mathcal{C}(\xi_{C_n}))}$$

and use (5.10). After division by n^d, we obtain (2.21) since $\lim_{n\to\infty} \frac{1}{n^d} \log \mathbb{Q}(\mathcal{C}(\sigma_{C_n})) = -s(\mathbb{Q})$, $\mathbb{Q}$-a.s. by the Shannon-Mc Millan-Breiman theorem and $\lim_{n\to\infty} \frac{1}{n^d} \log \frac{\mathbb{Q}(\mathcal{C}(\xi_{C_n}))}{\mathbb{P}(\mathcal{C}(\xi_{C_n}))} = s(\mathbb{Q}|\mathbb{P})$, $\mathbb{Q}$-a.s. (Proposition 3.14 in Section 3).

5.4 Proof of Theorem 2.25 and Theorem 2.22

We use the strong approximation formula (2.15) from Theorem 2.14 to get

$$\frac{d\log \mathbf{r}_n(\sigma)+\log \mathbb{P}_\beta(\mathcal{C}(\sigma_{C_n}))}{n^{\frac{d}{2}}} \to 0 \quad \text{when}\, n\to\infty, \text{ for } \mathbb{P}_\beta - \text{almost all}\, \sigma\,. \tag{5.11}$$

Therefore, it suffices to see that in the high-temperature regime we have a central limit theorem for $\{-\frac{1}{n^d}\log \mathbb{P}_\beta(\mathcal{C}(\sigma_{C_n}))\}$. By a standard argument presented below (5.15), one has

$$\lim_{n\to\infty}\frac{1}{n^d}\log\int \mathbb{P}_\beta(\mathcal{C}(\xi_{C_n}))^{-q}\, d\mathbb{P}(\xi) = P((1-q)\beta U)+(q-1)P(\beta U)\,, \tag{5.12}$$

for all $q\in[0,\infty)$. There exists $\beta_1>0$ such that for $|z|\le\beta_1$ the maps $z\mapsto P(zU)$ and

$$\Psi : z\mapsto \lim_{n\to\infty}\frac{1}{n^d}\log\int \mathbb{P}_\beta(\mathcal{C}(\xi_{C_n}))^{-z}\, d\mathbb{P}(\xi)$$

are analytic see e.g. [27], and [13]. Therefore, if $|(q-1)|\beta\le\beta_1$, the map $q\mapsto P((1-q)\beta U)+(q-1)P(\beta U)$ is analytic, and equality holds for all $q\in\mathbb{C}$.

By Bryc's theorem [6], this implies the CLT for $\{-\frac{1}{n^d}\log \mathbb{P}_\beta(\mathcal{C}(\sigma_{C_n}))\}$ with variance θ^2 given by

$$\theta^2 = \frac{d^2}{dq^2}\left(P((1-q)\beta U)\right)\big|_{q=0} \tag{5.13}$$

which is strictly positive by strict convexity of the pressure in the analyticity regime. The proof of Theorem 2.22 is the same once we observe that

$$\frac{d\log \mathbf{w}_n(\sigma)+\log \mathbb{P}_\beta(\mathcal{C}(\xi_{C_n}))}{n^{\frac{d}{2}}} \to 0 \quad \text{when}\, n\to\infty, \text{ for } \mathbb{P}_\beta\times\mathbb{P}_\beta - \text{almost all}\, (\xi,\sigma) \tag{5.14}$$

by using (5.10).

5.5 Proof of Theorem 2.29

Recall that for any Gibbs measure

$$-\log \mathbb{P}(\sigma_{C_n}) \sim \sum_{i\in C_n}\tau_i f_U(\sigma) + \log Z_{C_n}$$

and hence we have the identity

$$\lim_{n\to\infty}\frac{1}{n^d}\log\sum_{\mathcal{C}_n}\mathbb{P}(\mathcal{C}_n)^{1-q} = P((1-q)U)-(1-q)P(U). \tag{5.15}$$

In the sequel, we are going to show that

$$\int \mathbf{w}_n^{qd} d\mathbb{P}\times\mathbb{P} \approx \sum_{\mathcal{C}_n}\mathbb{P}(\mathcal{C}_n)^{1-q}, \tag{5.16}$$

for $q>-1$, and

$$\int \mathbf{w}_n^{qd} d\mathbb{P}\times\mathbb{P} \approx \sum_{\mathcal{C}_n}\mathbb{P}(\mathcal{C}_n)^{2}, \tag{5.17}$$

for $q \leq -1$. Here $a_n \approx b_n$ means that $\max\{a_n/b_n, b_n/a_n\}$ is bounded from above. Clearly (5.16) and (5.17) imply (2.31).

Let $q > 0$. Then

$$\int \mathbf{w}_n^{qd} d\mathbb{P} \times \mathbb{P} = \sum_{\mathcal{C}_n} \mathbb{P}(\mathcal{C}_n) \int \mathbf{t}_{\mathcal{C}_n}^{qd}(\sigma) d\mathbb{P}(\sigma) \tag{5.18}$$

$$= q \sum_{\mathcal{C}_n} \mathbb{P}(\mathcal{C}_n)^{1-q} \int_{\mathbb{P}(\mathcal{C}_n)}^{\infty} t^{q-1} \mathbb{P}\left\{\mathbf{t}_{\mathcal{C}_n}^d \geq \frac{t}{\mathbb{P}(\mathcal{C}_n)}\right\} dt. \tag{5.19}$$

By Theorem2.6, there exist positive constants A, B such that for any $t > 0$ one has

$$\mathbb{P}\left\{\mathbf{t}_{\mathcal{C}_n}^d \geq \frac{t}{\mathbb{P}(\mathcal{C}_n)}\right\} \leq Ae^{-Bt}.$$

Theorem2.6 also easily gives the lower bound :

$$\int_{\mathbb{P}(\mathcal{C}_n)}^{\infty} t^{q-1} \mathbb{P}\left\{\mathbf{t}_{\mathcal{C}_n}^d > \frac{t}{\mathbb{P}(\mathcal{C}_n}\right\} dt \geq K' - C\exp(-cn)\ K''$$

where $0 < K' := \int_1^\infty t^{q-1}\ e^{-\Lambda_2 t}\ dt < \infty$ and $0 < K'' := \int_0^\infty t^{q-1}\ e^{-\Lambda_1 t}\ dt < \infty$. For n large enough, $K' - C\exp(-cn)K''$ is strictly positive. Therefore we obtain

$$K_1 \sum_{\mathcal{C}_n} \mathbb{P}(\mathcal{C}_n)^{1-q} \leq \int \mathbf{w}_n^{qd} d\mathbb{P} \times \mathbb{P} \leq K_2 \sum_{\mathcal{C}_n} \mathbb{P}(\mathcal{C}_n)^{1-q},$$

where

$$K_1 := q\ (K' - C\exp(-cn_0)\ K''), \quad K_2 := qA \int_0^\infty t^{q-1} e^{-Bt} dt.$$

This establishes (5.16) for $q \geq 0$. The case $q = 0$ is trivial.

Let now $q \in (-1, 0)$.

$$\int \mathbf{w}_n^{-|q|d}\ d\mathbb{P} \times \mathbb{P} = \sum_{\mathcal{C}_n} \mathbb{P}(\mathcal{C}_n) \int \mathbf{t}_{\mathcal{C}_n}^{-|q|d}(\sigma)\ d\mathbb{P}(\sigma) \tag{5.20}$$

$$= \sum_{\mathcal{C}_n} \mathbb{P}(\mathcal{C}_n) \int_0^1 \mathbb{P}\left\{\mathbf{t}_{\mathcal{C}_n}^{-|q|d} \geq t\right\} dt \tag{5.21}$$

$$= |q| \sum_{\mathcal{C}_n} \mathbb{P}(\mathcal{C}_n)^{1+|q|} \int_{\mathbb{P}(\mathcal{C}_n)}^{\infty} t^{-|q|-1}\ \mathbb{P}\left\{\mathbf{t}_{\mathcal{C}_n}^d \leq \frac{t}{\mathbb{P}(\mathcal{C}_n)}\right\} dt. \tag{5.22}$$

The last integral is bounded from above by the integral where $\mathbb{P}(\mathcal{C}_n)$ replaced by 1 in the integration domain. From Theorem 2.6, we get the following lower bound, for every $t > 0$:

$$\mathbb{P}\left\{\mathbf{t}_{\mathcal{C}_n}^d \leq \frac{t}{\mathbb{P}(\mathcal{C}_n)}\right\} \geq 1 - e^{-\Lambda_1 t} - C'\mathbb{P}(\mathcal{C}_n)^\rho e^{-C'' t}$$

The number

$$K_1' := |q| \int_1^\infty t^{-|q|-1} \left(1 - e^{-\Lambda_1 t} - C'\mathbb{P}(\mathcal{C}_n)^\rho e^{-C'' t}\right)\ dt$$

is finite and strictly positive for n large enough.

Now, putting 0 instead of $\mathbb{P}(\mathcal{C}_n)$ gives an upper bound to the integral upon condideration. We use Lemma 4.5 to get immediately

$$\mathbb{P}\left\{\mathbf{t}^d_{\mathcal{C}_n} \le \frac{t}{\mathbb{P}(\mathcal{C}_n)}\right\} \le 1 - e^{-\Lambda_2 t} \ .$$

provided that $t \le \frac{1}{2}$. We have

$$\int_0^\infty t^{-|q|-1}\ \mathbb{P}\left\{\mathbf{t}^d_{\mathcal{C}_n} \le \frac{t}{\mathbb{P}(\mathcal{C}_n)}\right\} dt \le$$

$$\int_0^{\frac{1}{2}} t^{-|q|-1}\ \mathbb{P}\left\{\mathbf{t}^d_{\mathcal{C}_n} \le \frac{t}{\mathbb{P}(\mathcal{C}_n)}\right\} dt + \int_{\frac{1}{2}}^\infty t^{-|q|-1}\ dt \le$$

$$\frac{\Lambda_2\ 2^{1-|q|}}{1-|q|} + \frac{2^{-|q|}}{|q|} =: K_2' < \infty \ .$$

Hence, we conclude that for n large enough

$$K_1' \sum_{\mathcal{C}_n} \mathbb{P}(\mathcal{C}_n)^{1+|q|} \le \int \mathbf{w}_n^{-|q|d} d\mathbb{P}\times\mathbb{P} \le |q|\ K_2' \sum_{\mathcal{C}_n} \mathbb{P}(\mathcal{C}_n)^{1+|q|} \ .$$

Therefore we obtain (5.16) for $q \in (-1,0)$.

Finally, let us consider the remaining case $q \le -1$. Then for sufficiently large n (such that $\mathbb{P}(\mathcal{C}_n) < 1/2$) one has

$$\begin{aligned}
\int \mathbf{w}_n^{-|q|d} d\mathbb{P}\times\mathbb{P} &= |q| \sum_{\mathcal{C}_n} \mathbb{P}(\mathcal{C}_n)^{1+|q|} \int_{\mathbb{P}(\mathcal{C}_n)}^\infty t^{-|q|-1}\ \mathbb{P}\left\{\mathbf{t}^d_{\mathcal{C}_n} \le \frac{t}{\mathbb{P}(\mathcal{C}_n)}\right\} dt \\
&= |q| \sum_{\mathcal{C}_n} \mathbb{P}(\mathcal{C}_n)^{1+|q|} \left[\int_{\mathbb{P}(\mathcal{C}_n)}^{\frac{1}{2}} + \int_{\frac{1}{2}}^\infty\right] t^{-|q|-1}\mathbb{P}\left\{\mathbf{t}^d_{\mathcal{C}_n} \le \frac{t}{\mathbb{P}(\mathcal{C}_n)}\right\} dt \\
&= |q| \sum_{\mathcal{C}_n} \mathbb{P}(\mathcal{C}_n)^{1+|q|}\ [\ I_1(n,\mathcal{C}_n) + I_2(n,\mathcal{C}_n)\,] \,.
\end{aligned}$$

Clearly the second integral $I_2(n,\mathcal{C}_n)$ is uniformly bounded in n. Indeed,

$$I_2(n,\mathcal{C}_n) \le \int_{\frac{1}{2}}^\infty \frac{1}{t^{1+|q|}}\ dt < +\infty.$$

However, the first integral $I_1(n,\mathcal{C}_n)$ is diverging in the limit $n\to\infty$. Therefore the limiting behavior as $n\to\infty$ is determined by

$$|q| \sum_{\mathcal{C}_n} \mathbb{P}(\mathcal{C}_n)^{1+|q|} I_1(n,\mathcal{C}_n) = |q| \sum_{\mathcal{C}_n} \mathbb{P}(\mathcal{C}_n)^{1+|q|} \int_{\mathbb{P}(\mathcal{C}_n)}^{\frac{1}{2}} t^{-1-|q|}\mathbb{P}\left\{\mathbf{t}^d_{\mathcal{C}_n} \le \frac{t}{\mathbb{P}(\mathcal{C}_n)}\right\}\ dt.$$

We again use Lemma 4.5 to get

$$1 - e^{-\Lambda_1 t} \le \mathbb{P}\left\{\mathbf{t}^d_{\mathcal{C}_n} \le \frac{t}{\mathbb{P}(\mathcal{C}_n)}\right\} \le 1 - e^{-\Lambda_2 t} \ .$$

provided that $t \leq \frac{1}{2}$. Hence, using the Gibbs property (3.15), we have

$$I_1(n,\mathcal{C}_n) \leq \Lambda_2 \int_{\mathbb{P}(\mathcal{C}_n)}^{\frac{1}{2}} t^{-|q|}\ dt \leq \frac{\Lambda_2(1-2^{|q|-1}C'e^{-c'n^d})}{|q|-1}\ \mathbb{P}(\mathcal{C}_n)^{-|q|+1}$$

where we used the fact that for all $\kappa \in \mathbb{R}$, $1-e^{-\kappa} \leq \kappa$. Notice that for n large enough, the term between parentheses is strictly positive. Now, using the fact that $1-e^{-\kappa} \geq \kappa/2$ for any $\kappa \in [0,1]$, and remembering that $\Lambda_1/2 \leq 1$ ([1]), and using again the Gibbs property (3.15), we obtain

$$I_1(n,\mathcal{C}_n) \geq \frac{\Lambda_1(1-2^{|q|-1}Ce^{-cn^d})}{2(|q|-1)}\ \mathbb{P}(\mathcal{C}_n)^{-|q|+1}$$

where the term between parenthese is strictly positive provided that n is sufficiently large. Therefore, for n large enough, we end up with

$$\frac{|q|\Lambda_1(1-2^{|q|-1}Ce^{-cn^d})}{2(|q|-1)}\sum_{\mathcal{C}_n}\mathbb{P}(\mathcal{C}_n)^2 \leq \int \mathbf{w}_n^{-|q|d} d\mathbb{P}\times\mathbb{P} \leq \frac{2|q|\Lambda_2(1-2^{|q|-1}C'e^{-c'n^d})}{|q|-1}\sum_{\mathcal{C}_n}\mathbb{P}(\mathcal{C}_n)^2.$$

(Notice that L'Hôpital's rule shows that there is no problem at $q=-1$.) Thus, we obtain (5.17), which finishes the proof.

References

[1] M. Abadi, Exponential approximation for hitting times in mixing processes, Math. Phys. Electron. J. **7** (2001).

[2] M. Abadi, A. Galves, *Inequalities for the occurrence of rare events in mixing processes. The state of the art*, 'Inhomogeneous random systems' (Cergy-Pontoise, 2000), Markov Process. Related Fields 7 (2001), No. 1, 97–112.

[3] A. Asselah, P. Dai Pra, *Sharp estimates for the occurrence time of rare events for symmetric simple exclusion*, Stochastic Process. Appl. **71** (1997), No. 2, 259–273.

[4] A. Asselah, P. Dai Pra, *Occurrence of rare events in ergodic interacting spin systems*, Ann. Inst. H. Poincar Probab. Statist. **33** (1997), No. 6, 727–751.

[5] R. Bowen, *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*, Lecture Notes in Math. **470**, Springer, 1975.

[6] W. Bryc, *A remark on the connection between the large deviation principle and the central limit theorem*, Statist. & Probab. Lett. **18**, 253–256 (1993).

[7] J.-R. Chazottes, *Dimensions and waiting time for Gibbs measures*, J. Stat. Phys. **98** No. 3/4 , 305–320 (2000).

[8] Z. Chi, *The first-order asymptotic of waiting times with distortion between stationary processes*, IEEE Trans. Inform. Theory **47**, No. 1, 338–347 (2001).

[9] P. Collet, A. Galves, B. Schmitt, *Fluctuations of repetition times for gibbsian sources*, Nonlinearity **12**, 1225–1237 (1999).

[1] Indeed, $\Lambda_1 \leq \Lambda_2 = 2$, see the end of the proof of Lemma 4.5.

[10] F. Comets, *Grandes dviations pour des champs de Gibbs sur* $\mathbb{Z}^d$, CRAS, t. **303**, No. 11, 511-513 (1986).

[11] A. Dembo, I. Kontoyiannis, *Source coding, large deviations and approximate pattern matching*, IEEE Trans. Inf. Theory **48**, No. 6, 1590-1615 (2002).

[12] A. Dembo, O. Zeitouni, *Large Deviations Techniques & Applications*, Applic. Math. **38**, Springer, 1998.

[13] R.L. Dobrushin and S.B. Shlosman, *Completely analytical interactions: constructive description,* J. Stat. Phys. **46**, No. 5-6, 983–1014 (1987).

[14] R.S. Ellis, Entropy, large deviations, and statistical mechanics. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], 271. Springer-Verlag, New York, 1985.

[15] H. A. M. Danils, A. C. D. van Enter, *Differentiability properties of the pressure in lattice systems*, Comm. Math. Phys. **71**, No. 1, 65–76 (1980).

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

[17] X. Guyon. *Random Fields on a Network. Modeling, Statistics and Applications*, Springer Verlag, New York, Berlin, 1995.

[18] N. Haydn, J. Luvano, G. Mantica, S. Vaienti, *Multifractal Properties of Return Time Statistics*, Phys.Rev. Letters **88**, No.22 (2002).

[19] U. Krengel, *Ergodic theorems*, de Gruyter Studies in Mathematics **6**, Walter de Gruyter & Co., Berlin, 1985.

[20] H. Fllmer, *On entropy and information gain in random fields.* Z. Wahrsch. theorie Verw. Gebiete **26** (1973), 207–217.

[21] F. Martinelli, *An elementary approach to finite size conditions for the exponential decay of covariances in lattice spin models.* On Dobrushin's way. From probability theory to statistical physics, 169–181, Amer. Math. Soc. Transl. Ser. 2, 198, Amer. Math. Soc., Providence, RI, 2000.

[22] S. Olla, *Large deviations for Gibbs random fields*, Prob. Th. Rel. Fields, **77**, 343-357, (1988).

[23] D. Ornstein, B. Weiss, *Entropy and recurrence rates for stationary random fields*, Special issue on Shannon theory: perspective, trends, and applications. IEEE Trans. Inform. Theory **48**, No. 6, 1694–1697 (2002).

[24] D. Plachky, J.A. Steinebach, *A theorem about probabilities of large deviations with an application to queuing theory*, Periodica Math. Hungar. **6**, 343–345 (1975).

[25] D. Ruelle, Thermodynamic formalism. The mathematical structures of classical equilibrium statistical mechanics. Encyclopdia of Mathematics and its Applications **5**. Addison-Wesley Publishing Co., Reading, Mass., 1978

[26] P.C. Shields, *The ergodic theory of discrete sample paths.* AMS, Providence RI, 1996.

[27] B. Simon, The statistical mechanics of lattice gases. Vol. I. Princeton Series in Physics. Princeton University Press, Princeton, NJ, 1993.

[28] J.P. Thouvenot, *Convergence en moyenne de l'information pour l'action de* $\mathbb{Z}^2$. Z. Wahrsch. theorie Verw. Gebiete **24**, 135–137 (1972).

[29] A.J. Wyner, *More on recurrence and waiting times*, Ann. Appl. Probab. **9**, No. 3, 780–796 (1999).