Occurence, repetition and matching of patterns in the low-temperature Ising model

J.R. Chazottes *
F. Redig †

December 2, 2003

Abstract

We continue our study of exponential law for occurrences and returns of patterns in the context of Gibbsian random fields. For the low temperature plus phase of the Ising model, we prove exponential laws with error bounds for occurrence, return, waiting and matching times. Moreover we obtain a Poisson law for the number of occurrences of large cylindrical events and a Gumbel law for the maximal overlap between two independent copies. As a by-product, we derive precise fluctuation results for the logarithm of waiting and return times. The main technical tool we use, in order to control mixing, is disagreement percolation.

Key-words: occurrence and repetition of patterns, low temperature Ising model, disagreement percolation, exponential law, Poisson law, Gumbel law, large deviations.

1 Introduction

The study of occurrence and return times for highly mixing random fields has been initiated by Wyner, see [17]. In the context of stationary *processes*, there is a vast literature on exponential laws with error bounds for α, φ, ψ-mixing processes, see e.g. [3] for a recent overview. In the last four years, very precise results were obtained by Abadi [2]. The advantage of his approach is that it gives sharp bounds on the error of the exponential approximation and it holds for *all* cylindrical events. Moreover, it

*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

can be generalized to a broad class of random fields, see [4] for the case of Gibbsian random fields in the Dobrushin uniqueness regime (high temperature).

Low temperature Gibbsian random fields do not share the mixing property of the Dobrushin uniqueness regime, i.e. they are not (non-uniformly) φ-mixing. So far, no results on exponential laws have been proved in this context. To study these questions for Gibbsian random fields at low temperature, the Ising model is a natural candidate to begin with. The typical picture of the low temperature plus phase of this model is a sea of plus spins with exponentially damped islands of minus spins. Therefore decay of correlations of local observables can be estimated using the technique of disagreement percolation as initiated in [5] and further exploited in [6].

In this paper we prove exponential law with error bounds for occurrences and returns of cylindrical events for the low temperature plus phase of the Ising model. As an application we also obtain the exponential law with error bounds for waiting and matching times. These results can then be further exploited to obtain a Poisson law for the number of occurrences of cylindrical events (the Poisson law for the number of large contour has been obtained in [10] in the limit of zero temperature). We also derive a 'Gumbel law' for the maximal overlap (in the spirit of [14]) between two independent copies of the low-temperature Ising model. Other applications are strong approximations and large deviation estimates of the logarithm of waiting and return times.

The paper is organized as follows. In Section 2 we introduce basic notations, define occurrence and return times, and collect the mixing results at low temperature based on disagreement percolation. In section 3 we state our results. Section 4 is devoted to proofs.

2 Notations, definitions

2.1 Configurations, Ising model

We consider the low temperature plus phase of the Ising model on $\mathbb{Z}^d$, $d \geq 2$. This is a probability measure $\mathbb{P}^+_\beta$ on lattice spin configurations $\sigma \in \Omega = \{+,-\}^{\mathbb{Z}^d}$, defined as the weak limit as $V \uparrow \mathbb{Z}^d$ of the following finite volume measures:

$$\mathbb{P}^+_{V,\beta}(\sigma_V) = \exp\left(-\beta \sum_{<xy>\in V} \sigma_x\sigma_y - \beta \sum_{<xy>, x\in\partial V, y\notin V} \sigma_x\right) \Big/ Z^+_{V,\beta} \tag{2.1}$$

where $Z^+_{V,\beta}$ is the partition function. In (2.1) $< xy >$ denotes nearest neighbor bonds and ∂V the inner boundary, i.e. the set of those $x \in V$ having at least one neighbor $y \notin V$. For the existence of the limit $V \uparrow \mathbb{Z}^d$ of $\mathbb{P}^+_{V,\beta}$, see e.g. [12].

For $\eta \in \Omega$, $V \subseteq \mathbb{Z}^d$ we denote by $\mathbb{P}^\eta_{V,\beta}$ the corresponding finite volume measure with boundary condition η:

$$\mathbb{P}^{\eta}_{V,\beta}(\sigma_V) = \exp\left(-\beta \sum_{<xy>\in V} \sigma_x\sigma_y - \beta \sum_{x\in\partial V, x\in\partial V, y\notin V} \sigma_x\eta_x\right) \Big/ Z^{\eta}_{V,\beta}\,. \tag{2.2}$$

Later on, we shall omit the indices $\beta, +$ (in $\mathbb{P}^{+}_{\beta}$) referring to the inverse temperature and plus boundary condition respectively. We will choose $\beta > \beta_0 > \beta_c$, i.e., temperature below the transition point, such that a certain mixing condition, defined in detail below, is satisfied.

Let $V_n \uparrow \mathbb{Z}^d_+$ be an increasing sequence of sets such that

$$\lim_{n\to\infty} \frac{|\partial V_n|}{|V_n|} = 0\,. \tag{2.3}$$

We need the following pressure function $q \mapsto P(q\beta)$, $q \in \mathbb{R}$:

$$P(q\beta) = \lim_{n\to\infty} \frac{1}{|V_n|} \log \sum_{\sigma_{V_n}\in\{-,+\}^{V_n}} \exp\left(-q\beta \sum_{<xy>\in V_n} \sigma_x\sigma_y\right)\,. \tag{2.4}$$

(See [12] for the existence of $P(q\beta)$.)

2.2 Patterns, occurrence and return times

A pattern supported on a set $V \subseteq \mathbb{Z}^d$ is a configuration $\sigma_V \in \{+,-\}^V$. Patterns will be denoted by A. We will identify A with its cylinder, i.e., with the set $\{\sigma \in \Omega : \sigma_V = A\}$, so that it makes sense to write e.g. $\sigma \in A$. For $x \in \mathbb{Z}^d$, θ_x denotes the shift over x. For a pattern A supported on V, $\theta_x A$ denotes the pattern supported on $V + x$ defined by $\theta_x A(y + x) = A(y)$, $y \in V$.

If A is a pattern supported on V, and $W \subseteq \mathbb{Z}^d$ then we denote by $(A \prec W)$ the event that there exists $x \in \mathbb{Z}^d$ such that $V + x \subseteq W$ and such that $\sigma_{V+x}(y) = \theta_x A$. In words this means that the pattern A appears in the set W.

Let $\mathbb{V} = (V_n)$ where $V_n \uparrow \mathbb{Z}^d_+$, is such that $\lim_{n\to\infty} \frac{|\partial V_n|}{|V_n|} = 0$, and A_n a pattern supported on V_n. We define

$$\mathbf{T}^{\mathbb{V}}_{A_n} = \min\{|V_k| : A_n \prec V_k\}\,. \tag{2.5}$$

In words, this is volume of the first set V_k in which we can see the pattern A_n.

We denote for $x \in \mathbb{Z}^d$: $C(x, n) = C_n + x$. For $x, y \in \mathbb{Z}^d$: $|x - y| = \max_{i=1}^d |x_i - y_i|$, and for subsets $A, B \subseteq \mathbb{Z}^d$: $d(A, B) = \min_{x\in A, y\in B} |x - y|$.

For $\sigma \in \Omega$, A a pattern supported on V, $W \supset V$, we define the number of occurrences of A in W:

$$N(A, W, \sigma) = \sum_{x\in W: V+x\subseteq W} I(\sigma_{V+x} = \theta_x A)\,. \tag{2.6}$$

For a sequence $V_n \uparrow \mathbb{Z}^d_+$, the return time is defined as follows:

$$\mathbf{R}_{\sigma_{V_n}}(\sigma) = \min\{|V_k| : N(\sigma_{V_n}, V_k, \sigma) \geq 2\}\,. \tag{2.7}$$

Finally, for $\mathbb{V} = V_n \uparrow \mathbb{Z}^d_+$, and σ, $\eta \in \Omega$, we define the waiting time:

$$\mathbf{W}(V_n, \eta, \sigma) = \mathbf{T}^{\mathbb{V}}_{\eta_{V_n}}(\sigma) \tag{2.8}$$

We are interested in this quantity for σ distributed according to $\mathbb{P}$ and η distributed according to another ergodic (sometimes Gibbsian) probability measure $\mathbb{Q}$.

Finally, we consider 'matching times', in view of studying maximal overlap between two independent samples of $\mathbb{P}$. For σ, $\eta \in \Omega$,

$$\mathbf{M}(V_n, \sigma, \eta) = \min\{|V_k| : \exists x :\ V_n + x \subseteq V_k,\ \sigma_{V_n+x} = \eta_{V_n+x}\}\,.$$

In words, this is the minimal volume of a set of type V_k such that inside V_k, σ and η match on a set of the form $V_n + x$.

In the sequel we will omit the reference to the sequence V_n, in order not to overburden notation. In fact, proofs will be done for $V_n = C_n = [0, n]^d \cap \mathbb{Z}^d$. The generalization to $\mathbb{V}$ is obvious provided that the following two (sufficient) conditions are fulfilled:

1. $\lim_{n\to\infty} \frac{|\partial V_n|}{|V_n|} = 0$;

2. There exists $c > 0$ such that, for all x with $|x| \geq 1$, $|(V_n + x)\Delta V_n| \geq cn$.

2.3 Mixing at low temperatures

In [4] we derived exponential laws for hitting and return times under a mixing condition of the type

$$\sup_{\sigma,\eta,\xi} |\mathbb{P}^\eta_V(\sigma_W) - \mathbb{P}^\xi_V(\sigma_W)| \leq |W| \exp(-cd(V^c, W)) \tag{2.9}$$

usually called 'non-uniform exponential φ-mixing'. This condition is of course not satisfied at low temperatures since boundary conditions continue to have influence. Take e.g. $W = \{0\}$, $\eta \equiv +$, $\xi \equiv -$, then for $\beta > \beta_c$:

$$\lim_{V\uparrow\mathbb{Z}^d} \mathbb{P}^\eta_V(\sigma_0 = +) - \mathbb{P}^\xi_V(\sigma_0 = +) = m^+_\beta > 0 \tag{2.10}$$

where $0 < m^+_\beta = \int \sigma_0 d\mathbb{P}(\sigma)$ is the magnetization. This clearly contradicts (2.9). However, for local functions f, g we do have an estimate like

$$\Big| \int f\ \theta_x g\ d\mathbb{P} - \int f d\mathbb{P} \int g d\mathbb{P} \Big| \leq C(f, g)\ e^{-c(\beta)|x|}\,. \tag{2.11}$$

The intuition here is that there can only be correlation between two functions if the clusters containing their dependence sets are finite (i.e. not contained in the sea of plusses) and intersect. Since finite clusters are exponentially small (in diameter), we have exponential decay of correlations of local functions.

This idea is formalized in the context of 'disagreement percolation'. To introduce this concept, we define a path $\gamma = \{x_1, \ldots, x_n\}$, i.e. a subset of $\mathbb{Z}^d$ such that x_i and x_{i-1} are neighbors for all $i = 1, \ldots, n$.

More formally, for $W \subseteq V$ and η and $\xi \in \Omega$, we have the following inequality:

$$|\mathbb{P}^\eta_V(\sigma_W) - \mathbb{P}^\xi_V(\sigma_W)| \leq |\partial W| \, |\mathbb{P}^\eta_V \times \mathbb{P}^V_\xi(W \leftrightarrow \partial V)| \,. \tag{2.12}$$

Here $(W \leftrightarrow \partial V)$ denotes the event of those couples $(\sigma_1, \sigma_2) \in \Omega_V \times \Omega_V$ where there is 'a path of disagreement' γ leading from W to the boundary of V such that $\sigma_1(x) \neq \sigma_2(x)$ for all $x \in \gamma$. Of course whether the probability of this event under the measure $\mathbb{P}^\eta_V \times \mathbb{P}^\xi_V$ will be small depends on the distance between V and W and on the chosen boundary conditions η, ξ. The estimate (2.12) as well as the ideas of disagreement percolation can be found in [6],[13].

On the top of inequality 2.12 we have the following estimate of [7], see [13]:

$$\mathbb{P} \times \mathbb{P}(W \leftrightarrow \partial V) \leq e^{-c(\beta) d(W, \partial V)} \tag{2.13}$$

as soon as $\beta > \beta_0$, and where $c(\beta) \to \infty$ as $\beta \to \infty$.

In the rest of the paper we always work with $\beta > \beta_0$, so that we can apply (2.12), (2.13).

3 Results

3.1 Exponential laws

Theorem 3.1. *There exist $0 < \Lambda_1 \leq \Lambda_2 < \infty$, $c, c' > 0$ such that for any pattern $A = A_n$ supported on C_n, there exist $\kappa > 0$ and $\lambda_A \in [\Lambda_1, \Lambda_2]$ such that for all n and $t < e^{\kappa n^d}$:*

$$\left| \mathbb{P}\left(\mathbf{T}_A \geq \frac{t}{\lambda_A \mathbb{P}(A)} \right) - e^{-t} \right| \leq e^{-ct} e^{-c' n^d}. \tag{3.2}$$

For return times we have to restrict to good patterns, i.e., patterns which are not 'badly self-repeating' in the following sense:

Definition 3.3. *A pattern A_n is called good if for any x with $|x| < n/2$, for the cylinders we have $A_n \cap \theta_x A_n = \varnothing$.*

Good patterns have a return time at least $(n/2+1)^d$ and as we will see later that this property guarantees that the return time is actually of the order e^{cn^d}.

The following lemma is proved in [4] for general Gibbsian random fields.

Lemma 3.4. *Let $\mathcal{G}_n$ be the set of all good patterns. There exists $c > 0$ such that*

$$\mathbb{P}(\mathcal{G}_n) \geq 1 - e^{-cn^d} \,. \tag{3.5}$$

We denote by $\mathbb{P}(\cdot|A)$ the measure $\mathbb{P}$ conditioned on the event $A \prec C_n$.

Theorem 3.6. *There exist $0 < \Lambda_1 \leq \Lambda_2 < \infty$, $c, c' > 0$ such that for any* good *pattern $A = A_n$ supported on C_n, there exist $\kappa > 0$ and $\lambda_A \in [\Lambda_1, \Lambda_2]$ such that for all n and $t < e^{\kappa n^d}$:*

$$\left|\mathbb{P}\left(\mathbf{R}_A \geq \frac{t}{\lambda_A \mathbb{P}(A)}\middle|A\right) - e^{-t}\right| \leq e^{-ct} e^{-c' n^d}. \tag{3.7}$$

We have the following analogue of Theorem 3.1 for matching times.

Theorem 3.8. *There exist $0 < \Lambda_1 \leq \Lambda_2 < \infty$, $c, c' > 0$ such that for any pattern $A = A_n$ supported on C_n, there exist $\kappa > 0$ and $\lambda_n \in [\Lambda_1, \Lambda_2]$ such that for all n and $t < e^{\kappa n^d}$:*

$$\left|\mathbb{P}\times\mathbb{P}\left((\sigma, \eta) : \mathbf{M}_n(\sigma, \eta) \geq \frac{t}{\lambda_n \mathbb{P}\times\mathbb{P}(\sigma_{C_n} = \eta_{C_n})}\right) - e^{-t}\right| \leq e^{-ct} e^{-c' n^d}. \tag{3.9}$$

3.2 Poisson law

Let $A = A_n$ be any pattern supported on C_n. Let $C(t/\mathbb{P}(A))$ be the maximal cube of the form $C_k = [0, k]^d \cap \mathbb{Z}^d$ such that $|C_k| \leq t/\mathbb{P}(A)$. Observe that

$$\frac{|C(t/\mathbb{P}(A))|}{t/\mathbb{P}(A)} \to 1$$

as $n \to \infty$. Define

$$N_t^n(\sigma) = N(A_n, C(t/\mathbb{P}(A)), \sigma)\,. \tag{3.10}$$

Then we have

Theorem 3.11. *If σ is distributed according to $\mathbb{P}$, and A_n is a sequence of good patterns, then the processes $\{N_t^n/\lambda_{A_n} : t \geq 0\}$ converge to a mean one Poisson process $\{N_t : t \geq 0\}$ weakly on path space, where λ_{A_n} is the parameter of Theorem 3.1.*

3.3 Gumbel law

To formulate the Gumbel law for certain extremes, we need simply connected subsets G_n, $n \geq 1$, such that $|G_n| = n$ and $G_{n^d} = C_n$. For instance, for $d = 2$, $G_1 = \{(0,0)\}$, $G_2 = \{(0,0),(1,0)\}$, $G_3 = \{(0,0),(1,0),(1,1)\}$, $G_4 = \{(0,0),(1,0),(1,1),(0,1)\}$, etc.

For $\eta \in \Omega$, define

$$\mathcal{M}_n(\eta, \sigma) = \max\{|G_k| : \exists x \in G_n \text{ with } G_n + x \subseteq G_k \text{ and } \eta_{G_n+x} = \sigma_{G_n+x}\} \tag{3.12}$$

In words this is the volume of the maximal subset of the type G_k on which η and σ agree. We have the following

Theorem 3.13. *For any $\eta \in \Omega$, there exists a sequence $u_n \uparrow \infty$, and constants $\lambda, \lambda', \nu, \nu' \in (0, \infty)$ such that for all $x \in \mathbb{Z}$*

$$\min\{e^{-\lambda' e^{-\nu' x}}, e^{-\lambda e^{-\nu x}}\} \le \liminf_{n\to\infty} \mathbb{P}\times\mathbb{P}\left((\eta,\sigma) : \mathcal{M}_n(\eta,\sigma) \le u_n + x\right) \le$$
$$\limsup_{n\to\infty} \mathbb{P}\times\mathbb{P}\left((\eta,\sigma) : M_n(\eta,\sigma) \le u_n + x\right) \le \max\{e^{-\lambda' e^{-\nu' x}}, e^{-\lambda e^{-\nu x}}\}\,. \quad (3.14)$$

Remark 3.15. *Notice that in theorem 3.13 we study the maximal matching between two configurations on a specific sequence of supports G_n. Since in the low temperature plus phase we have percolation of plusses, the same theorem would of course not hold for the cardinality of the maximal connected subset of C_n on which η and σ agree because the latter subset occupies a fraction of the volume of C_n.*

Remark 3.16. *The fact that in the Gumbel law we only have a lower and an upper bound is due to the discreteness of the $\mathbf{M}_n(\sigma,\eta)$. This situation can be compared to the study of the maximum of independent geometrically distributed random variables, see for instance [11].*

3.4 Fluctuations of waiting, return and matching times

We denote by $s(\mathbb{P})$ the entropy of $\mathbb{P}$ defined by

$$s(\mathbb{P}) = \lim_{n\to\infty} -\frac{1}{n^d} \sum_{A_n \in \{+,-\}^{C_n}} \mathbb{P}(A_n)\log\mathbb{P}(A_n)\,.$$

The next result (proved in subsection 4.7) shows how the repetition of typical patterns allows to compute the entropy using a single 'typical' configuration.

Theorem 3.17. *There exists $\epsilon_0 > 0$ such that for all $\epsilon > \epsilon_0$*

$$-\epsilon\log n \le \log\left[\mathbf{R}_{\sigma_{C_n}}(\sigma)\ \mathbb{P}(\sigma_{C_n})\right] \le \log\log n^{\epsilon} \quad \text{eventually } \mathbb{P}-\text{almost surely.} \quad (3.18)$$

In particular,

$$\lim_{n\to\infty} \frac{1}{n^d}\log \mathbf{R}_{\sigma_{C_n}}(\sigma) = s(\mathbb{P}) \quad \mathbb{P}-\text{almost surely}\,. \quad (3.19)$$

Note that (3.19) is a particular case of the result by Ornstein and Weiss in [16] where $\mathbb{P}$ is only assumed to be ergodic. Under our assumptions, we get the more precise result (3.18).

Remark 3.20. *It follows immediately from* (3.18) *that the sequence* $(\log \mathbf{R}_{\sigma_{C_n}}(\sigma))$ *satisfies the central limit theorem if and only if* $(-\log\mathbb{P}(\sigma_{C_n}))$ *does. However, in the low temperature regime, we are not able to prove the central limit theorem for* $(-\log\mathbb{P}(\sigma_{C_n}))$.

Suppose that η is a configuration randomly chosen according to an ergodic random field $\mathbb{Q}$ and, independently, σ is randomly chosen according to $\mathbb{P}$. We denote by $s(\mathbb{Q}|\mathbb{P})$ the relative entropy density of $\mathbb{Q}$ with respect to $\mathbb{P}$, where

$$s(\mathbb{Q}|\mathbb{P}) = \lim_{n\to\infty} \frac{1}{n^d} \sum_{A_n \in \{+,-\}^{C_n}} \mathbb{Q}(A_n) \log \frac{\mathbb{Q}(A_n)}{\mathbb{P}(A_n)} .$$

We have the following result (proved in Subsection 4.8):

Theorem 3.21. *Assume that $\mathbb{Q}$ is an ergodic random field. Then there exists $\epsilon_0 > 0$ such that for all $\epsilon > \epsilon_0$*

$$-\epsilon \log n \le \log \left(\mathbf{W}(C_n, \eta, \sigma)\right) \mathbb{P}(\eta_{C_n})) \le \log\log n^\epsilon \tag{3.22}$$

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

$$\lim_{n\to\infty} \frac{1}{n^d} \log \mathbf{W}(C_n, \eta, \sigma) = s(\mathbb{Q}) + s(\mathbb{Q}|\mathbb{P}) \quad \mathbb{Q}\times\mathbb{P} - \text{a.s}\,. \tag{3.23}$$

Remark 3.24. *If in* (3.23) *we choose $\mathbb{Q} = \mathbb{P}^-$, the low temperature minus phase, we conclude that the time to observe a pattern typical for the minus phase in the plus phase, is equal to the time to observe a pattern typical for the plus phase, at the logarithmic scale.*

The following theorem is proved in subsection 4.9.

Theorem 3.25. *For all $q \in \mathbb{R}$ the limit*

$$\mathcal{W}(q) = \lim_{n\to\infty} \frac{1}{n^d} \log \int \mathbf{W}(C_n, \eta, \sigma)^q \; d\mathbb{P}\times\mathbb{P} \tag{3.26}$$

exists and equals

$$\mathcal{W}(q) = \begin{cases} P((1-q)\beta) + (q-1)P(\beta), & \textit{for } q \ge -1, \\ P(2\beta) - 2P(\beta), & \textit{for } q < -1, \end{cases} \tag{3.27}$$

where P is the pressure defined in (2.4)*.*

From this result, it follows that the sequence $(\frac{1}{n^d} \log \mathbf{W}(C_n, \eta, \sigma))$ satisfies a generalized large deviation principle in the sense of theorem 4.5.20 in [8].

Remark 3.28. *A more general version of Theorem 3.25 can be easily derived: 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 3.27 has to be properly modified (see [4]).*

For the matching times, we have the following analogue of Theorem 3.21 (see subsection 4.10):

Theorem 3.29. *There exists $\epsilon_0 > 0$ such that for all $\epsilon > \epsilon_0$*

$$-\epsilon \log n \le \log \left(\mathbf{M}(C_n, \eta, \sigma) \; \mathbb{P}\times\mathbb{P}(\sigma_{C_n} = \eta_{C_n})\right) \le \log\log n^\epsilon \tag{3.30}$$

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

$$\lim_{n\to\infty} \frac{1}{n^d} \log \mathbf{M}(C_n, \eta, \sigma) = \mathcal{W}(-1) \quad \mathbb{P}\times\mathbb{P} - \text{a.s}\,. \tag{3.31}$$

4 Proofs

4.1 Positivity of the parameter

Lemma 4.1 (The parameter). *There exist strictly positive constants Λ_1, Λ_2 such that for any integer t with $t\mathbb{P}(A) \le 1/2$, one has*

$$\Lambda_1 \le \lambda_{A,t} := -\frac{\log \mathbb{P}(\mathbf{T}_A > t)}{t\mathbb{P}(A)} \le \Lambda_2 \,.$$

Proof. We proceed by a second moment estimate on the random variable $N(A, C_k, \sigma)$:

$$\mathbb{E}(N(A, C_k, \sigma))^2 = \sum_{x,y: x+C_n \subseteq C_k, y+C_n \subseteq C_k} \mathbb{P}(\theta_x A \cap \theta_y A). \tag{4.2}$$

We split the sum in three parts: $I_1 = \sum_{x=y}$, $I_2 = \sum_{x \ne y, |x-y| \le \Delta}$, $I_3 = \sum_{x \ne y, |x-y| > \Delta}$.

We now estimate I_1, I_2 and I_3. The quantities I_1 and I_2 are estimated as in [4]. For I_1 we have:

$$I_1 = (k+1)^d \mathbb{P}(A).$$

For I_2, using the Gibbs property and $d \ge 2$:

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

Only the third term involves the disagreement percolation estimate.

$$\begin{aligned} & I_3 - (k+1)^{2d} \mathbb{P}(A)^2 \\ \le & \sum_{x \ne y, |x-y| > \Delta} \mathbb{P}(A) \; |\mathbb{P}(\sigma_{C(x,n)} = A | \sigma_{C(y,n)} = A) - \mathbb{P}(A)|. \end{aligned} \tag{4.3}$$

Denote by $C'_{x,\Delta,n}$ the set of those sites which are at least at lattice distance $\Delta + 1$ away from $C(x, n)$, and $C^{\Delta}(x, n)$ the complement of that set. Then we have for $|x - y| > \Delta$:

$$\begin{aligned} & |\mathbb{P}(\sigma_{C(x,n)} = A | \sigma_{C(y,n)} = A) - \mathbb{P}(A)| \\ = & \left| \int \int \Big(\mathbb{P}(\sigma_{C(x,n)} = A | \eta_{C'_{x,\Delta,n}}) - \mathbb{P}(\sigma_{C(x,n)} = A | \xi_{C'_{x,\Delta,n}}) \Big) \, d\mathbb{P}(\eta | \sigma_{C(y,n)} = A) d\mathbb{P}(\xi) \right| \\ \le & \int \int \mathbb{P}^{\eta}_{C^{\Delta}(x,n)} \times \mathbb{P}^{\xi}_{C^{\Delta}(x,n)} \left(C(x,n) \nleftrightarrow \partial C^{\Delta}(x,n) \right) d\mathbb{P}(\eta | \sigma_{C(y,n)} = A) d\mathbb{P}(\xi) \\ \le & \frac{1}{\mathbb{P}(A)} \; \mathbb{P} \times \mathbb{P} \left(C(x,n) \nleftrightarrow \partial C^{\Delta}(x,n) \right) \\ \le & \frac{1}{\mathbb{P}(A)} \; |\partial C(x,n)| \; e^{-d(C(x,n), \partial C^{\Delta}(x,n))} \\ \le & e^{-cn^{d+1} + c' n^d} \le e^{-\tilde{c} n^{d+1}} \end{aligned} \tag{4.4}$$

where in the last step we used the choice $\Delta = \Delta_n = n^{d+1}$. Using the second moment estimate and proceeding as in [4], this gives the inequality:

$$\begin{aligned} \frac{\mathbb{P}(\mathbf{T}_A \le t)}{t\mathbb{P}(A)} & \ge \frac{1}{1 + e^{-\delta n} \Delta^d + t\mathbb{P}(A) + e^{-cn^{d+1}} t / \mathbb{P}(A)} \\ & \ge \frac{1}{1 + C_1 + 1/2 + C_2} \end{aligned} \tag{4.5}$$

where

$$C_1 = \sup_n n^{d(d+1)}\ e^{-\delta n} < \infty, \quad C_2 = \sup_A \sup_{t\leq 1/(2\mathbb{P}(A))} e^{-cn^{d+1}} t/\mathbb{P}(A) < \infty\,.$$

The upper bound is derived as in the high temperature case, see [4]. □

4.2 Iteration lemma and proof of Theorem 3.1

We consider k mutually disjoint cubes C_i such that $|C_i| = f_A = (\lfloor \mathbb{P}(A)^{-\theta/d}\rfloor + 1)^d$, where $0 < \theta < 1$ is fixed. The essential point is to make precise the approximation of $\mathbb{P}(A \not\prec \cup_{i=1}^k C_i)$ by $\mathbb{P}(A \not\prec C_1)^k$.

For a cube C_i we denote by $C_i^\Delta \subseteq C_i$ the largest cube inside C_i with the same midpoint as C_i and such that the boundary ∂C_i is at least at lattice distance Δ away from C_i^Δ

$$\mathbb{P}\left(A \not\prec \cup_{i=1}^k C_i\right) =$$
$$\mathbb{P}(A \not\prec C_1 | A \not\prec C_2 \cap A \not\prec C_3 \ldots A \not\prec C_k)\mathbb{P}(A \not\prec C_2 \cap A \not\prec C_3 \ldots A \not\prec C_k) =$$
$$\left(\mathbb{P}(A \not\prec C_1^\Delta | A \not\prec C_2 \cap A \not\prec C_3 \ldots A \not\prec C_k) + \epsilon_1\right)\mathbb{P}(A \not\prec C_2 \cap A \not\prec C_3 \ldots A \not\prec C_k) =$$
$$\left(\mathbb{P}(A \not\prec C_1^\Delta) + \epsilon_1 + \epsilon_2\right)\mathbb{P}(A \not\prec C_2 \cap A \not\prec C_3 \ldots A \not\prec C_k) =$$
$$\left(\mathbb{P}(A \not\prec C_1) + \epsilon_1 + \epsilon_2 + \epsilon_3\right)\mathbb{P}(A \not\prec C_2 \cap A \not\prec C_3 \ldots A \not\prec C_k)\,.$$

We now start to estimate the errors ϵ_i. For the first one:

$$\begin{aligned} |\epsilon_1| &\leq \mathbb{P}(A \not\prec C_1^\Delta \cap A \prec C_1 | A \not\prec C_2 \cap A \not\prec C_3 \ldots A \not\prec C_k) \\ &\leq \Delta f_A^{(d-1)/d}\mathbb{P}(A_n)\ e^{cn^{d-1}} \end{aligned} \tag{4.6}$$

In the last step, in order to obtain the factor $e^{cn^{d-1}}$ we used the following fact:

$$\sup_{\eta,\xi} \frac{\mathbb{P}(\sigma_{C_n} = A_n | \eta_{C_n^c})}{\mathbb{P}(\sigma_{C_n} = A_n | \xi_{C_n^c})} \leq e^{cn^{d-1}} \tag{4.7}$$

an inequality which is valid for any Gibbs measure. For ϵ_2 we use the disagreement percolation estimate, as in the proof of lemma 4.1:

$$|\epsilon_2| \leq \frac{\mathbb{P}\times\mathbb{P}(C_1^\Delta \leftrightarrow \partial C_1)}{\mathbb{P}(A \not\prec C_2 \cap A \not\prec C_3 \ldots A \not\prec C_k)}$$

Finally, as in the estimate of the first ϵ_1,

$$\epsilon_3 \leq \Delta f_A^{(d-1)/d}\mathbb{P}(A_n) \tag{4.8}$$

where now the boundary term $e^{cn^{d-1}}$ is not present since we do not have a conditioned measure. Put

$$\alpha_{k-p} = \mathbb{P}(A \prec \cup_{i=p+1}^k C_i). \tag{4.9}$$

We obtain the recursion inequality:

$$\alpha_k \leq (\alpha_1 + \epsilon)\alpha_{k-1} + \epsilon' \tag{4.10}$$

where $\epsilon = \epsilon_1 + \epsilon_3$, and $\epsilon' = \epsilon_2 \alpha_{k-1}$ This gives, combining the estimates for ϵ_i with the elementary inequality $a^n - b^n \leq (\max\{a,b\})^{n-1} n$:

$$\alpha_k - \alpha_1^k \leq$$
$$k\Big(2\Delta f_A^{(d-1)/d}\mathbb{P}(A_n)e^{cn^{d-1}}\Big)\Big(\mathbb{P}(A \not\prec C_1) + 2\Delta f_A^{(d-1)/d}\mathbb{P}(A_n)e^{cn^{d-1}}\Big)^{k-1} + ke^{-cn^{d+1}}$$

Now, fix $f_A = \mathbb{P}(A)^{-\theta}$, $\Delta = tn^{d+1}$ $k = \lfloor \frac{t}{\mathbb{P}(A) f_A} \rfloor$. Then we have

$$k\epsilon' \leq te^{-ctn^{d+1}}$$

and

$$k\epsilon \leq te^{-cn^d}.$$

Therefore as long as $t < e^{-\kappa n^d}$ with $\kappa < c$, we have

$$\alpha_k - \alpha_1^k \leq e^{-c' n^d} e^{-ct}. \tag{4.11}$$

The lower bound

$$\alpha_k - \alpha_1^k \geq e^{-c' n^d} e^{-ct} \tag{4.12}$$

is obtained analogously. At this stage, one can repeat the proof of [4] to obtain (3.2) in Theorem 3.1. □

4.3 Return time

For a pattern A_n and a configuration $\sigma \in \Omega$ such that $\sigma_{C_n} = A_n$ we write $A \prec^* C_k$ for the event that A appears at least twice C_k and $A \not\prec^* C_k$ is the event that A occurs in C_k only on C_n.

, i.e., the number of occurrences is equal to one.

In order to repeat the iteration lemma for pattern repetitions, we first prove the following lemma.

Lemma 4.13. *Let A_n be a good pattern, then there exists $c > 0$ such that for the cube C_k of volume $f_A = (\lfloor \mathbb{P}(A)^{-\theta/d} \rfloor + 1)^d$, we have*

$$|\mathbb{P}(A_n \not\prec^* C_k | A_n) - \mathbb{P}(A_n \not\prec C_k)| \leq e^{-cn^d} \tag{4.14}$$

Proof. Since A is good, A does not appear in any cube $\theta_x C_n$ for $|x| < n/2$. We will introduce a gap Δ' with a n-dependence to be chosen later on. Denote by $\mathcal{C}_n^{\Delta'}$ the minimal cube containing C_n such that its boundary is at distance at least Δ' from C_n. For simplicity, we write

$$\begin{aligned} |\mathbb{P}(A \not\prec^* \mathcal{C}^{\Delta'}|A) - \mathbb{P}(A \not\prec^* C \setminus \mathcal{C}_n^{\Delta'}|A)| &\leq \mathbb{P}(A \prec \mathcal{C}_n^{\Delta'} \setminus C_{n/2}|A) \\ &\leq {\Delta'}^d e^{-cn^d}. \end{aligned} \tag{4.15}$$

To get the last inequality, remark that

$$\mathbb{P}(A \prec \mathcal{C}_n^{\Delta'} \setminus C_{n/2}|A) \leq \sup_{V:|V|>(n/2)^d} \sup_{B\in\Omega_V} \sup_{\eta\in\Omega} \mathbb{P}(B|\eta_{V^c}) \tag{4.16}$$

since $|\theta_x C_n \setminus C_n| > (n/2)^d$ for $|x| \geq n/2$. The rhs of (4.16) is bounded by e^{-cn^d} by the Gibbs property. Now we can use the mixing property to obtain:

$$|\mathbb{P}(A \not\prec^* C \setminus \mathcal{C}_n^{\Delta'}|A) - \mathbb{P}(A \not\prec C \setminus \mathcal{C}_n^{\Delta'})| \leq e^{-c_1\Delta'} e^{c_2 n^d} f_A^{(d-1)/d} \tag{4.17}$$

and finally,

$$|\mathbb{P}(A \not\prec C) - \mathbb{P}(A \not\prec C \setminus \mathcal{C}_n^{\Delta'})| \leq \Delta' f_A^{(d-1)/d}\mathbb{P}(A) \tag{4.18}$$

which yields the statement of the lemma for the choice $f_A = (\lfloor \mathbb{P}(A)^{-\theta/d} \rfloor + 1)^d$ and $\Delta' = n^{d+1}$. □

We can now state the analogue of the iteration lemma for pattern repetitions:

Lemma 4.19. *Let $A = A_n \in \mathcal{G}_n$ be a good pattern. Let C_i be a collection of disjoint cubes of volume f_A^d. We have the following estimate:*

$$\begin{aligned} &\left(\mathbb{P}(A \not\prec^* \cup_{i=1}^k C_i|A) - [\mathbb{P}(A \not\prec C_1)]^k\right) \\ \leq\ & k\left(2\Delta f_A^{(d-1)/d}\mathbb{P}(A)e^{cn^{d-1}}\right)\left(\mathbb{P}(A \not\prec C_1) + 2\Delta f_A^{(d-1)/d}\mathbb{P}(A)e^{cn^{d-1}}\right)^{k-1} \\ +\ & ke^{-c\Delta} + e^{-cn^d}\mathbb{P}(A \not\prec C_1)^{k-1} \end{aligned} \tag{4.20}$$

Proof. Start with the following identity:

$$\mathbb{P}(A \not\prec^* \cup_{i=1}^k C_i|A) = \frac{\mathbb{P}(A \cap A \not\prec^* C_1 \cap A \not\prec C_2 \ldots A \not\prec C_k)}{\mathbb{P}(A)} \tag{4.21}$$

We can proceed now as in the proof of the iteration lemma to approximate the rhs of (4.21) by

$$\Pi_k = \frac{\mathbb{P}(A \cap A \not\prec^* C_1)}{\mathbb{P}(A)}\ \mathbb{P}(A \not\prec C_2)\ldots\mathbb{P}(A \not\prec C_k) \tag{4.22}$$

at the cost of an error ϵ which can be estimated by

$$\epsilon \leq k\left(2\Delta f_A^{(d-1)/d}\mathbb{P}(A)e^{cn^{d-1}}\left(\mathbb{P}(A \not\prec C_1) + 2\Delta f_A^{(d-1)/d}\mathbb{P}(A)e^{cn^{d-1}}\right)^{k-1} + ke^{-c\Delta} \tag{4.23}$$

Now, to replace Π_k by $\mathbb{P}(A \not\prec C_1)^k$, use lemma 4.13 to conclude that this replacement induces an extra error which is at most

$$e^{-cn^d}\mathbb{P}(A \not\prec C_1)^{k-1}\,. \tag{4.24}$$

□

4.4 Matching time

In order to prove the exponential law (3.8) for matching times, we first remark that for cylinders A_n defined on $\Omega \times \Omega = (\{+,-\} \times \{+,-\})^{\mathbb{Z}^d}$, we have the analogue of Theorem 3.1 under the measure $\mathbb{P} \times \mathbb{P}$ with the same proof. Indeed, a typical configuration drawn from $\mathbb{P} \times \mathbb{P}$ is a sea of $(+,+)$ with exponentially damped islands of non $(+,+)$. We now generalize the statement of Theorem 3.1 to the $\mathcal{F}_n$ measurable events that we need (which are not cylindrical).

Lemma 4.25. *Suppose $E_n = \{(\sigma,\eta) : \sigma_x = \eta_x,\ \forall x \in C_n\}$. Theorem 3.1 holds with A_n replaced by E_n and $\mathbb{P}$ replaced by $\mathbb{P} \times \mathbb{P}$.*

Proof. Clearly, the analogue of the iteration lemma does not pose any new problem. The main point is to prove the non-triviality of the parameter, i.e., the analogue of lemma 4.1. In order to obtain this, we have to estimate the second moment of

$$N^k_{E_n} = \sum_{x: C_n + x \subseteq C_k} I(\theta_x E_n)$$

under $\mathbb{P} \times \mathbb{P}$. As before we split

$$\mathbb{E} \times \mathbb{E}(N^k_{E_n})^2 \le I_1 + I_2 + I_3 \tag{4.26}$$

where $I_1 = \sum_{x=y} \mathbb{P} \times \mathbb{P}(E_n) \le (k+1)^d \mathbb{P}(E_n)$, $I_2 = \sum_{x \neq y, |x-y| \le \Delta} \mathbb{P} \times \mathbb{P}(\theta_x E_n \cap \theta_y E_n)$ and $I_3 = \sum_{x \neq y, |x-y| > \Delta} \mathbb{P} \times \mathbb{P}(\theta_x E_n \cap \theta_y E_n)$. The only problematic term here is I_2. As in the proof for cylindrical events, we will use the Gibbs property, and prove first the existence of $1 > \delta > 0$ such that

$$\delta \le \mathbb{P} \times \mathbb{P}(\sigma_x = \eta_x | (\sigma,\eta)_{\mathbb{Z}^d \setminus \{x\}}) \le 1 - \delta\,. \tag{4.27}$$

We now further estimate

$$\begin{aligned} \mathbb{P} \times \mathbb{P}(\sigma_x = \eta_x | (\sigma,\eta)_{\mathbb{Z}^d \setminus \{x\}}) &= \sum_{\epsilon = +,-} \mathbb{P}(\sigma_x = \epsilon | \sigma) \mathbb{P}(\eta_x = \epsilon | \eta) \\ &\le \sup_{\sigma,\eta} \left[\mathbb{P}(+|\sigma)\mathbb{P}(+|\eta) + (1 - \mathbb{P}(+|\sigma))(1 - \mathbb{P}(+|\eta)) \right]\,. \end{aligned} \tag{4.28}$$

Since by the Gibbs property $0 < \zeta < \mathbb{P}(+|\eta) < 1 - \zeta < 1$, we can bound (4.28) by

$$\max_{\zeta < x,y < 1-\zeta} (2uv - u - v - 1) < 1$$

where the last inequality follows from

$$2uv \le u^2 + v^2 < u + v$$

for $u, v < 1 - \zeta < 1$. From the inequality (4.27), we obtain using $d \ge 2$:

$$\begin{aligned} & \sum_{x \in C_k} \sum_{y \neq x, |y-x| \le \Delta} \mathbb{P} \times \mathbb{P}(\theta_y E_n | \theta_x E_n) \mathbb{P} \times \mathbb{P}(E_n) \\ \le\ & (k+1)^d (\Delta+1)^d \sup_{\sigma,\eta} \sup_{k \ge n} \sup_{x_1,\dots x_k \in \mathbb{Z}^d} \mathbb{P} \times \mathbb{P}(\sigma_{x_1} = \eta_{x_1}, \dots, \sigma_{x_k} = \eta_{x_k} | (\sigma,\eta)_{\mathbb{Z}^d \setminus \{x_1,\dots x_k\}}) \\ \le\ & (1-\delta)^n\,. \end{aligned} \tag{4.29}$$

Therefore, choosing $\Delta = n^{d+1}$, we obtain

$$\sum_{x\in C_k} \sum_{y\neq x, |y-x|\leq\Delta} \mathbb{P}\times\mathbb{P}(\theta_y E_n|\theta_x E_n)\mathbb{P}\times\mathbb{P}(E_n) \leq (k+1)^d C \tag{4.30}$$

where

$$C = \sup_n n^{d(d+1)}(1-\delta)^n < \infty\,.$$

The third term in the decomposition (4.26) is estimated as in the proof of lemma 4.1. At this point we can repeat the proof of lemma 4.1. □

4.5 Poisson law for occurrences

For a *good pattern* $A = A_n$ supported on C_n, we define the second occurrence time by the relation:

$$(T_A^2(\sigma) \leq k^d) = (N(A, V_k, \sigma) \geq 2) \tag{4.31}$$

and the restriction that T_A^2 can only take values $(k+1)^d$, $k \in \mathbb{N}$. Similarly we define the p-th occurrence time:

$$(T_A^p(\sigma) \leq k^d) = (N(A, V_k, \sigma) \geq p) \tag{4.32}$$

and the same restriction. The following proposition shows that in the limit $n \to \infty$, properly normalized increments of the process $\{T_{A_n}^k : k \in \mathbb{N}\}$ converge to a sequence of independent exponentials. This implies convergence of the finite dimensional distributions of the counting process to a Poisson process defined in (3.10).

Proposition 4.33. *Let A_n be a good pattern (in the sense of Definition 3.3). Define $\tau_{A_n}^p = T_{A_n}^p - T_{A_n}^{p-1}$, where $T_{A_n}^0 = 0$. For all $p \in \mathbb{N}$, $t_1, \ldots t_p \in [0,\infty)$,*

$$\lim_{n\to\infty} \mathbb{P}\left(\left[\tau_{A_n}^p \geq \sum_{i=1}^p t_i/\mathbb{P}(A_n)\right] \cap \left[\tau_{A_n}^{p-1} \leq \sum_{i=1}^{p-1} t_i/\mathbb{P}(A_n)\right] \cap \ldots \cap \left[\tau_{A_n}^1 \leq t_1/\mathbb{P}(A_n)\right]\right) =$$

$$e^{-(t_1+\ldots t_k)}(1-e^{-(t_1+\ldots t_{k-1})})\ldots(1-e^{-t_1})$$

Proof. We start with the case of two occurrence times T_1, T_2:

$$\begin{aligned}
&\mathbb{P}\left(T_1 \leq \frac{t}{\mathbb{P}(A)} \cap T_2 \geq \frac{s}{\mathbb{P}(A)} + T_1\right) \\
=& \sum_{k\leq\frac{t}{\mathbb{P}(A)}} \mathbb{P}\left(T_2 \geq \frac{s}{\mathbb{P}(A)} + k \mid T_1 = k\right)\,\mathbb{P}(T_1 = k)\,.
\end{aligned} \tag{4.34}$$

Let us denote by $\mathcal{C}_k$ the cube defined by the relation $(T_1 \leq k) = (A \prec \mathcal{C}_k)$, and by $A \prec^1 C_k$ the event that A appears for the first time in C_k (more precisely $A \prec^1 C_k$ abbreviates the event $(T_1 = k)$, i.e., $\cap_{l<k}(A \not\prec C_l) \cap (A \prec C_k))$.

Let us denote by $\mathcal{C}_k^\Delta$ the Δ-extension of $\mathcal{C}_k$, i.e., the minimal cube containing $\mathcal{C}_k$ such that $\partial\mathcal{C}_k^\Delta$ and ∂C_k are at least Δ apart. Recall that $C(t/\mathbb{P}(A))$ denotes the maximal cube of the form $C_k = [0,k]^d \cap \mathbb{Z}^d$ such that $|C_k| \le t/\mathbb{P}(A)$. Remember that

$$\frac{|C(t/\mathbb{P}(A))|}{t/\mathbb{P}(A)} \to 1$$

as $n \to \infty$.

Lemma 4.35. *If A is a good pattern, then we have the estimate*

$$\begin{aligned} &\mathbb{P}\left(T_2 \ge \frac{s}{\mathbb{P}(A)} + k \;\Big|\; A \prec^1 \mathcal{C}_k\right) - \mathbb{P}\Big(A \not\prec C\left(\frac{s}{\mathbb{P}(A)}\right) \setminus \mathcal{C}_k^\Delta \;\Big|\; A \prec^1 \mathcal{C}_k\Big) \\ \le\;\; & \Delta f_A^{(d-1)/d} e^{-cn^d}\,. \end{aligned} \tag{4.36}$$

Proof. The proof is identical to that of lemma 4.13. ∎

Now we want to replace

$$\mathbb{P}\Big(A \not\prec C\left(\frac{s}{\mathbb{P}(A)}\right) \setminus \mathcal{C}_k^\Delta \;\Big|\; A \prec^1 \mathcal{C}_k\Big) \tag{4.37}$$

by the unconditioned probability of the same event. We make the choice $\Delta = n^{d+1}$. By the disagreement percolation estimate, this gives an error which can be bounded by

$$\sum_{k\le t/\mathbb{P}(A)} \mathbb{P}(T_1 = k)\Big[\mathbb{P}\Big(A \not\prec C\left(\frac{s+t}{\mathbb{P}(A)}\right) \setminus \mathcal{C}_k^\Delta \Big| A \prec^1 \mathcal{C}_k\Big) - \mathbb{P}\Big(A \not\prec C\left(\frac{s+t}{\mathbb{P}(A)}\right) \setminus \mathcal{C}_k^\Delta\Big)\Big] \le$$

$$\sum_{k\le t/\mathbb{P}(A)} e^{-c\Delta} \le t^2 e^{cn^d} e^{-c' n^{d+1}}\,.$$

Finally,

$$\begin{aligned} &\sup_{k\le t/\mathbb{P}(A)} \Big[\mathbb{P}\Big(A \not\prec C\left(\frac{s+t}{\mathbb{P}(A)}\right) \setminus \mathcal{C}_k^\Delta\Big) - \mathbb{P}\Big(A \not\prec C\left(\frac{s+t}{\mathbb{P}(A)}\right) \setminus C\left(\frac{t}{\mathbb{P}(A)}\right)\Big)\Big] \\ \le\;\; & \Delta(t/\mathbb{P}(A))^{(d-1)/d}\mathbb{P}(A) = \Delta t^{(d-1)/d}\mathbb{P}(A)^{1/d}\,. \end{aligned} \tag{4.38}$$

By the exponential law, we have, using $|C((t+s)/\mathbb{P}(A)) \setminus C(t/\mathbb{P}(A))| = t/\mathbb{P}(A)$:

$$\mathbb{P}\Big(A \not\prec C\left(\frac{s+t}{\mathbb{P}(A)}\right) \setminus C\left(\frac{t}{\mathbb{P}(A)}\right)\Big) = \exp(-\lambda_A s) + \epsilon_n \tag{4.39}$$

where $\epsilon_n = \epsilon(n,t,s) \to 0$ as $n \to \infty$. Which gives:

$$\begin{aligned} &\lim_{n\to\infty} \Big(\mathbb{P}(\tau_2 \ge s/\mathbb{P}(A) \cap \tau_1 \le t/\mathbb{P}(A)) - \lim_n \mathbb{P}(\tau_1 \le t/\mathbb{P}(A)) e^{-\lambda_A s}\Big) \\ =\;\; & \lim_{n\to\infty} \Big(\mathbb{P}(\tau_2 \ge s/\mathbb{P}(A) \cap \tau_1 \le t/\mathbb{P}(A)) - (1 - e^{-\lambda_A t})e^{-\lambda_A s}\Big) = 0\,. \end{aligned} \tag{4.40}$$

This proves the statement of the proposition for $k = 2$, the general case is analogous and left to the reader. ∎

The following proposition follows immediately from proposition 4.33

Proposition 4.41. *Let $A_n \in \mathcal{G}_n$ be a good pattern supported on C_n. Then the finite dimensional marginals of the process $\{N^n_{t/\lambda_{A_n}} : t \geq 0\}$ converge to the finite dimensional marginals of a mean one Poisson process as n tends to infinity.*

In order to obtain convergence in the Skorokhod space, we have to prove tightness. This is an immediate consequence of the following simple lemma for general point processes, applied to

$$N^n_t = N(A_n, C(t/\mathbb{P}(A_n)), \sigma)\,.$$

Lemma 4.42. *Let $\{N^n_t : t \geq 0\}$ be a sequence of point processes with path space measures $\mathbb{P}^T_n$ on $D([0,T],\mathbb{N})$. If there exists $C > 0$ such that for all n and for all $t \leq T$ we have the estimate*

$$\mathbb{E}^T_n(N^n_t) \leq Ct \tag{4.43}$$

then the sequence $\mathbb{P}^T_n$ is tight.

Proof. From (4.43) we infer for all n, $t \leq T$

$$\mathbb{P}^T_n(N^n_t \geq K) \leq CT/K\,.$$

Hence

$$\lim_{K\uparrow\infty} \sup_{0\leq t\leq T} \sup_n \mathbb{P}^T_n(N^n_t \geq K) = 0 \tag{4.44}$$

For a trajectory $\omega \in D([0,T],\mathbb{N})$ one defines the modulus of continuity

$$w_\gamma(T,\omega) = \inf_{(t_i)_{i=1}^N} \sup_{i=1}^N |\omega_{t_i} - \omega_{t_{i-1}}| \tag{4.45}$$

where the infimum is taken over all partitions $t_0 = 0 < t_1 < \ldots < t_N = t$ such that $t_i - t_{i-1} \geq \gamma$. If for some $\epsilon > 0$ $w_\gamma(T,\omega) \geq \epsilon$, then the number of jumps of ω in $[0,T]$ is at least $[T/\gamma]$. Hence we obtain using (4.43):

$$\mathbb{P}^T_n(w_\gamma(T,\omega) \geq \epsilon) \leq \mathbb{P}^T_n(N^n_T \geq T/\gamma) \leq C\gamma\,. \tag{4.46}$$

This gives for all $\epsilon > 0$:

$$\lim_{\gamma\downarrow 0} \sup_n \mathbb{P}^T_n(w_\gamma(T,\omega) \geq \epsilon) = 0\,. \tag{4.47}$$

Combination of (4.44) and (4.47) with the tightness criterion [15] p. 152 yields the result. □

Remark 4.48. *With much more effort, one can obtain precise bounds for the difference*

$$\left|\mathbb{P}(N^n_t/\lambda_{A_n} = k) - \frac{t^k}{k!}e^{-t}\right|$$

which are well-behaved in n, t and k. In particular, from such bounds one can obtain convergence of all moments of N^n_t/λ_{A_n} to the corresponding Poisson moments. This is done in [1] in the context of mixing processes.

4.6 Gumbel law

For $\eta, \sigma \in \Omega$ denote

$$\mathcal{V}_0(\eta, \sigma) = \bigcup \{G_k : \sigma_{G_k} = \eta_{G_k}\} \,. \tag{4.49}$$

We start with the following simple lemma:

Lemma 4.50. *1. There exists $\delta > 0$ such that for all $\eta \in \Omega$:*

$$\inf_{k \in \mathbb{N}} \frac{\mathbb{P} \times \mathbb{P}(\mathcal{V}_0 \supset G_{k+1})}{\mathbb{P}(\mathcal{V}_0 \supset G_k)} \geq \delta \,. \tag{4.51}$$

2. There exists a non-decreasing sequence $u_n \uparrow \infty$ such that for all $n \in \mathbb{N}$:

$$1 \leq n\mathbb{P}(\mathcal{V}_0 \supset G_{u_n}) \leq \frac{1}{\delta} \,. \tag{4.52}$$

Proof. For item 1:

$$\begin{aligned} \frac{\mathbb{P} \times \mathbb{P}(\mathcal{V}_0 \supset G_{k+1})}{\mathbb{P} \times \mathbb{P}(\mathcal{V}_0 \supset G_k)} &= \mathbb{P} \times \mathbb{P}(\eta_{x_{n+1}} = \sigma_{x_{n+1}} | \sigma_{G_n} = \eta_{G_n}) \\ &\geq \inf_{\xi, \sigma} \mathbb{P} \times \mathbb{P}(\sigma_x = \eta_x | \sigma_{\mathbb{Z}^d \setminus \{x\}}, \xi_{\mathbb{Z}^d \setminus \{x\}}) \\ &= \delta > 0 \end{aligned} \tag{4.53}$$

where the last inequality follows from the fact that $\mathbb{P} \times \mathbb{P}$ is a Gibbs measures. For item 2, put

$$f(n) = \mathbb{P} \times \mathbb{P}(\mathcal{V}_0 \supset G_n)$$

and

$$\begin{aligned} u_n^+ &= \min\{k : f(k) \leq 1/n\} \\ u_n^- &= \max\{k : f(k) \geq 1/n\} \end{aligned} \tag{4.54}$$

Clearly,

$$u_n^- \leq u_n^+ \leq u_n^- + 1 \,.$$

Now choosing $u_n = u_n^-$ and using (4.53), we obtain

$$\begin{aligned} \frac{1}{n} &\leq \mathbb{P} \times \mathbb{P}(\mathcal{V}_0 \supset G_{u_n}) \\ &= \frac{\mathbb{P} \times \mathbb{P}(\mathcal{V}_0 \supset G_{u_n})}{\mathbb{P} \times \mathbb{P}(\mathcal{V}_0 \supset G_{u_n+1})} \mathbb{P} \times \mathbb{P}(\mathcal{V}_0 \supset G_{u_n+1}) \\ &\leq \frac{1}{\delta n} \,. \end{aligned} \tag{4.55}$$

□

We now adapt our definition of matching time to the sequence of sets G_n:

$$\tau_n^{\mathcal{G}}(\eta,\sigma) = \min\{k : \exists x : G_n + x \subseteq G_k \text{ such that } \sigma_{G_n+x} = \eta_{G_n+x}\}\,. \tag{4.56}$$

We have the relation

$$(\mathcal{M}_n(\eta,\sigma) \geq k) = (\tau_k^{\mathcal{G}}(\eta,\sigma) \leq n)\,. \tag{4.57}$$

In words: the maximal matching inside G_n is greater than or equal to k if and only if the first time that a matching on a set G_k happens is not larger than n. Now we choose $k = u_n + x$ ($x \in \mathbb{N}$) and use the exponential law for matching times:

$$\mathbb{P}\times\mathbb{P}(\tau_{u_n}^{\mathcal{G}}(\eta,\sigma) \leq n) = 1 - \exp(-\lambda_n \mathbb{P}\times\mathbb{P}(\sigma_{G_{u_n+x}} = \eta_{G_{u_n+x}})) + \epsilon_n$$

where ϵ_n goes to zero as n goes to infinity. By the choice of u_n,

$$\mathbb{P}\times\mathbb{P}(\sigma_{G_{u_n+x}} = \eta_{G_{u_n+x}}) = \mathbb{P}\times\mathbb{P}(\mathcal{V}_0(\eta,\sigma) \supset G_{u_n+x}) \in \left[\frac{A}{n}e^{-\nu x}, \frac{B}{n}e^{-\nu' x}\right] \tag{4.58}$$

where $A, B \in (0,\infty)$ and

$$0 < e^{-\nu} = \liminf_{n\to\infty} \frac{\mathbb{P}\times\mathbb{P}(\sigma_{G_{n+1}} = \eta_{G_{n+1}})}{\mathbb{P}\times\mathbb{P}(\sigma_{G_n} = \eta_{G_n})} < 1$$

and

$$0 < e^{-\nu'} = \limsup_{n\to\infty} \frac{\mathbb{P}\times\mathbb{P}(\sigma_{G_{n+1}} = \eta_{G_{n+1}})}{\mathbb{P}\times\mathbb{P}(\sigma_{G_n} = \eta_{G_n})} < 1\,.$$

Here the inequality for the $\liminf$ is an immediate consequence of lemma 4.50, and the inequality for the $\limsup$ is derived in a completely analogous way, using the Gibbs property. The theorem now follows immediately from (4.58).

4.7 Proof of Theorem 3.17

We start by showing the following summable upper-bound of

$$\mathbb{P}\{\sigma\ : \log(\mathbf{R}_{\sigma_{C_n}}(\sigma)\mathbb{P}(\sigma_{C_n})) \geq \log t\} \leq$$

$$\sum_{A_n\in\mathcal{G}_n} \mathbb{P}(A_n)\ \mathbb{P}\{\sigma\ : \log(\mathbf{R}_{A_n}(\sigma)\mathbb{P}(A_n)) \geq \log t \mid A_n\} + \sum_{A_n\in\mathcal{G}_n^c} \mathbb{P}(A_n)\,.$$

From Theorem 3.6 and Lemma 3.4 we get for all $0 < t < e^{\kappa n^d}$

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

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

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

An application of the Borel-Cantelli lemma leads to

$$\log\left(\mathbf{R}_{\sigma_{C_n}}(\sigma)\mathbb{P}(\sigma_{C_n})\right) \le \log\log(n^{\epsilon}) \quad \text{eventually a.s.}\,.$$

For the lower bound first observe that Theorem 3.6 gives, for all $0 < t < e^{cn^d}$

$$\mathbb{P}\{\sigma : \log(\mathbf{R}_{\sigma_{C_n}}(\sigma)\mathbb{P}(\sigma_{C_n})) \le \log t\} \le e^{-c'n^d} + 1 - \exp(-\Lambda_2 t) + e^{-cn^d}\,.$$

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

$$\log\left(\mathbf{R}_{\sigma_{C_n}}(\sigma)\mathbb{P}(\sigma_{C_n})\right) \ge -\epsilon \log n \quad \text{eventually a.s.}\,.$$

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

4.8 Proof of Theorem 3.21

We first show that the strong approximation formula (3.18) holds with $\mathbf{W}(C_n, \eta, \sigma)$ in place of $\mathbf{R}_{\sigma_{C_n}}(\sigma)$ with respect to the measure $\mathbb{Q} \times \mathbb{P}$. We have the following identity:

$$\int d\mathbb{Q}(\eta)\ \mathbb{P}\left\{\sigma : \mathbf{T}_{\eta_{C_n}}(\sigma) > \frac{t}{\mathbb{P}(\eta_{C_n})}\right\} =$$

$$(\mathbb{Q} \times \mathbb{P})\left\{(\eta, \sigma) : \mathbf{W}(C_n, \eta, \sigma) > \frac{t}{\mathbb{P}(\eta_{C_n})}\right\}.$$

This shows that Theorem 3.1 remains valid if we replace $\mathbf{T}_{\eta_{C_n}}(\sigma)$ with $\mathbf{W}(C_n, \eta, \sigma)$ and $\mathbb{P}$ with $\mathbb{Q} \times \mathbb{P}$, hence so is Theorem 3.17. Therefore for ϵ large enough, we obtain

$$-\epsilon \log n \le \log(\mathbf{W}(C_n, \eta, \sigma)\mathbb{P}(\eta_{C_n})) \le \log\log n^{\epsilon} \tag{4.59}$$

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

$$\log(\mathbf{W}(C_n, \eta, \sigma)\mathbb{P}(\sigma_{C_n})) = \log \mathbf{W}(C_n, \eta, \sigma) + \log \mathbb{Q}(\eta_{C_n}) - \log \frac{\mathbb{Q}(\eta_{C_n})}{\mathbb{P}(\eta_{C_n})}$$

and use (4.59). After division by n^d, we obtain (3.23) since $\lim_{n\to\infty} \frac{1}{n^d} \log \mathbb{Q}(\sigma_{C_n}) = -s(\mathbb{Q})$, $\mathbb{Q}$-a.s. by the Shannon-Mc Millan-Breiman theorem and $\lim_{n\to\infty} \frac{1}{n^d} \log \frac{\mathbb{Q}(\eta_{C_n})}{\mathbb{P}(\eta_{C_n})} = s(\mathbb{Q}|\mathbb{P})$, $\mathbb{Q}$-a.s. by the Gibbs variational principle (See e.g. [4] for a proof).

4.9 Proof of Theorem 3.25

We follow the line of proof of [4] to compute $\mathcal{W}(q)$. The only extra complication in our case is that the bound

$$\mathbb{P}\left(\mathbf{T}_{A_n} > \frac{t}{\mathbb{P}(A_n)}\right) \le e^{-ct}$$

for all $t > 0$ cannot be obtained directly from Theorem 3.1. Instead we will use the following lemma which shows that such a bound can be obtained by a rough version of the iteration lemma. Given this result, the proof of [4] can be repeated.

Lemma 4.60. *1. There exists $c > 0$ such that for all patterns $A_n \in \{+,-\}^{C_n}$*

$$\mathbb{P}\left(\mathbf{T}_{A_n} > \frac{t}{\mathbb{P}(A_n)}\right) \leq e^{-ct}.$$

2. There exists $\delta \in (0, \frac{1}{2})$ such that for all n and all pattern $A = A_n$

$$0 < \delta < \mathbb{P}(T_A > \frac{1}{2\mathbb{P}(A)}) < 1 - \delta < 1.$$

Proof. To prove the first inequality, we fill part of the cube $C(t/\mathbb{P}(A))$ with little cubes of size f_A (where f_A is defined in Lemma 4.2), with $k \geq t/(2\mathbb{P}(A)f_A)$. The gaps Δ separating the different cubes are taken equal to $\lceil tn^{d+1}\rceil$. We then have the following

$$\mathbb{P}(T_A > t/\mathbb{P}(A)) \leq \mathbb{P}(A \not\prec \cup_{i=1}^{K} C_i).$$

Notice that we do not have to estimate here the probability that the pattern is not in the gaps since we only need an upper bound. Now

$$\alpha_K = \mathbb{P}(A \not\prec \cup_{i=1}^{K} C_i) =$$

$$\mathbb{P}(A \not\prec C_1 | A \not\prec \cup_{i=2}^{K} C_i)\mathbb{P}(A \not\prec \cup_{i=2}^{K} C_i) := \mathbb{P}(A \not\prec C_1 | A \not\prec \cup_{i=2}^{K} C_i)\alpha_{K-1}.$$

Using the disagreement percolation estimate, we have

$$\mathbb{P}(A \not\prec C_1 | A \not\prec \cup_{i=2}^{K} C_i) - \alpha_1 \leq e^{-\Delta}.$$

Therefore

$$\alpha_K \leq \alpha_{K-1}\alpha_1 + \exp(-\Delta).$$

Iterating this inequality gives, using $\Delta = \lceil tn^{d+1}\rceil$,

$$\alpha_K \leq \alpha_1^K + e^{-tn^{d+1}} e^{cn^d} t$$

Now we use $K > t/2\mathbb{P}(A)f_A$, and Lemma 4.1 to obtain:

$$\alpha_K \leq (1 - \Lambda_1 f_A \mathbb{P}(A))^{t/(2f_A\mathbb{P}(A))} + e^{-ct}$$

which implies the first inequality of the lemma.

The second inequality follows directly from Lemma 4.1. □

4.10 Proof of Theorem 3.29

The proof of (3.30) is identical to the proof of (3.22) but using the exponential law for the matching time. Formula (3.31) follows from

$$\mathbb{P}\times\mathbb{P}(\sigma_{C_n} = \eta_{C_n}) = \sum_{\sigma_{C_n} \in \{+,-\}^{C_n}} \mathbb{P}(\sigma_{C_n})^2$$

and the definition of $\mathcal{W}$.

References

[1] M. Abadi, *Statistics and error terms of occurrence times in mixing processes*, preprint (2003).

[2] M. Abadi, *Sharp error terms and necessary conditions for exponential hitting times in mixing processes*, to appear in the Annals of Probab.

[3] 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.

[4] M. Abadi, J.R. Chazottes, F. Redig and E. Verbitskiy, *Exponential distribution for the occurence of rare patterns in Gibbsian random fields*, preprint 2003, to appear in Commun. Math. Phys.

[5] J. van den Berg, *A uniqueness condition for Gibbs measures, with application to the 2-dimensional Ising antiferromagnet*, Commun. Math. Phys. **152** (1993), no. 1, 161-166.

[6] J. van den Berg, C. Maes, *Disagreement percolation in the study of Markov fields*, Ann. Probab. **22** (1994), no. 2, 749–763.

[7] R. Burton and J. Steiff, *Quite weak Bernoulli with exponential rate and percolation for random fields*, Stoch. Proc. Appl. **58**, 35-55 (1995).

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

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

[10] P. Ferrari, P. Picco, *Poisson approximation for large-contours in low-temperature Ising models*, Physica A: Statistical Mechanics and its Applications **279**, Issues 1-4, 303–311 (2000).

[11] J. Galambos, *The asymptotic theory of extreme order statistics*. Second edition. Robert E. Krieger Publishing Co., Inc., Melbourne, FL, 1987.

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

[13] H.O. Georgii, O. Häggström and C. Maes, *The Random Geometry of Equilibrium Phases*, in Phase transitions and critical phenomena, Vol 18, Eds. C. Domb and J.L Lebowitz, p1-142, Academic Press London (2001)

[14] S. Karlin, A. Dembo, *Limit distributions of maximal segmental score among Markov-dependent partial sums*, Adv. in Appl. Probab. **24** (1992), no. 1, 113–140.

[15] C. Kipnis, C. Landim, *Scaling limits of interacting particle systems.* Grundlehren der Mathematischen Wissenschaften **320**, Springer-Verlag, Berlin, 1999.

[16] 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).

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